செங்காந்தள்

அறிஞர்களால் மதிப்பீடு செய்யப்படும் தமிழ் ஆய்விதழ்

தென் திருவிதாங்கூர் இந்துக் கல்லூரி

தமிழ்த்துறை உயராய்வு மையம்
நாகர்கோவில்

சங்க இலக்கியங்களில் வேளாண்மையும்
உணவு முறையும்

செங்காந்தள்

அறிஞர்களால் மதிப்பீடு செய்யப்படும் தமிழ் ஆய்விதழ்

தென் திருவிதாங்கூர் இந்துக் கல்லூரி

தமிழ்த்துறை உயராய்வு மையம்

நாகர்கோவில்

சங்க இலக்கியங்களில் வேளாண்மையும்

உணவு முறையும்

சங்க இலக்கியத்தில் சோற்றுணா

முனைவர் அ. சுஜாதா ஜாய்ஸ், தமிழ்த் துறை உதவிப்பேராசிரியர், நாஞ்சில் கத்தோலிக்க கலை மற்றும் அறிவியல் கல்லூரி, களியக்காவிளை, கன்னியாகுமரி மாவட்டம், தமிழ்நாடு,
ORCiD: https://orcid.org/0009-0004-6282-212X
DOI: 10.5281/zenodo.8397651

ஆய்வுச் சுருக்கம்

மனிதனின் அடிப்படைத் தேவைகளாக உணவு உடை உறைவிடம் காணப்படுகின்றது. சரியான உறைவிடமும் தரமான ஆடையுமின்றி மனிதன் வாழ்ந்தாலும் உணவின்றி உயிர் வாழ முடியாது என்பது மனுக்குலம் அறிந்த உண்மை. உணவு மனித உயிர்கள் மட்டுமின்றி பல தரப்பட்ட அஃறிணை உயிரினங்கள் மற்றும் மரம் செடி கொடிகளுக்குக் கூட தேவையாக உள்ளது. மனிதர்களாகிய நாம் கால ஓட்டத்தின் வளர்ச்சி நிலைக்கேற்ப உணவை பலவாறு சமைத்து ருசித்து வருகின்றோம். நம் முன்னோர்களாகிய பழந்தமிழர்கள் பலதரப்பட்ட உணவை பயன்படுத்தினாலும் சோற்றுணாவிற்கு அதிக முக்கியத்துவம் கொடுத்தனர். இதற்கு சங்க இலக்கியங்களில் காணப்படும் பலவிதமான சோற்றுணா வகைகள் சான்றாகிறது. சங்க காலத் தமிழர்கள் பலதரப்பட்ட அரிசி வகைகளையும் பலவாறு சமைத்து உண்டனர்.சில வகையிலான சோற்றுணா குறிப்பிட்ட சில நிகழ்வுகளில் மட்டுமே சமைத்து பயன்படுத்தப்பட்ட செய்தியும் அறியலாகிறது.

திறவுச்சொற்கள்: சங்க இலக்கியம், சோற்றுணா, பழந்தமிழர்கள்.

முன்னுரை

மனிதனின் அடிப்படைத் தேவையாக உணவும் உடையும், உறையுளும் அமைகின்றன. இதில் முதன்மை பெறுவது உணவேயாகும். உலக உயிர்கள் உணவின்றி வாழ்தலில்லை. உணவு என்பது உடல் வளர்ச்சியடைவதற்கான மூல ஆதாரமாக மட்டுமின்றி ஒரு சமுதாயத்தின் பண்பாட்டுக்குட்பட்டதாகவும் அமைகின்றது. உணவு முறை என்பது ஒவ்வொரு நாட்டின் சுற்றுப்புற சூழல் நில அமைப்பு மற்றும் அங்கு பயிரிடப்படும் விளபொருட்களின் அடிப்படையிலேயே அமைகின்றது. இதனை தொல்காப்பியரும் கருப்பொருள் பற்றிய விளக்கத்தின்போது 'தெய்வம் உணாவே மாமரம் புள் பறை' என்று தெய்வத்திற்கு அடுத்தப்படியாக உணவை கூறுவதிலிருந்து உணர்ந்து கொள்ளலாம். பழந்தமிழர் தம் உணவு முறைகள் நிலத்தின் அடிப்படையிலேயே அமைந்திருந்ததை காண முடிகிறது. பழந்தமிழர் சோற்றுணவுக்கு அதிக முக்கியத்துவம் கொடுத்ததை இக்கட்டுரை விளக்கியுரைக்கிறது.

உணவின் வகை

மனித குலம் உயிர் வாழ இன்றியமையாதது உணவு. உணவின் தன்மைக்கேற்ற அதன்

சுவையும் பயனும் அமையும். பொதுவாக உணவை திட உணவு, திரவ உணவு என இருபெரும் பிரிவாக பகுக்கலாம். திட உணவை தொல்காப்பியர் மரபியலில்,

> மெய்திரி வகையின் எண்வகை உணவின்

> செய்தியும் வரையார் (தொல்., பொருள்.,நூ. 623)

என்கிறார். இதற்கு உரைதரும் இளம்பூரணர் நெல்லு, காணம், வரகு, இறுங்கு, தினை, சாமை, புல்லு, கோதுமை என்கிறார். சங்கத் தமிழர்கள் சாமை, இறுங்கு, கோதுமை தவிர பிற உணவுகளை மிகுதியாக பயன்படுத்தினர். அதிலும் நெல்லரிசி சோற்றினை பலவாறு பக்குவப்படுத்தி உண்டமை சங்க இலக்கியம் முழுவதும் பரந்து விரிந்த செய்தியாகும். தண்ணீர், தேன், கள், பால், தயிர், மோர், நெய் போன்ற நீர்மத்தன்மைமிக்க உணவுகள் திரவ உணவாக விளங்குகின்றன.

அரிசி

சங்ககாலத் தமிழர்களின் முக்கிய உணவாக அரிசி விளங்கியது. காடைக்கண்ணியரிசி, தினையரிசி, வரகரிசி, மூங்கிலரிசி என பல்வேறு அரிசி வகைகளை பக்குவப்படுத்தி உணவு தயாரித்தனர். எனினும் நெல்லரிசியையே விரும்பி உண்டனர். அரிசியை பச்சையரிசி, புழுங்கலரிசி என இருவிதமாக பக்குவப்படுத்தி பயன்படுத்தினர். நெல்லை உரலிலிட்டு குற்றி பெறப்பட்ட அரிசி சாதாரணமாக பச்சரிசி எனப்படும். நெல்லை அவித்து காய வைத்து குற்றி தயாரிக்கப்படும் அரிசி புழுங்கலரிசி எனப்படும். நெல்லை குற்றி அரிசியாக்க பாறை உரல் மற்றும் நிலவுரலை பயன்படுத்தினர். மேலும் இரும்புப்பூணை கட்டி உலக்கை தயாரித்தனர். இதனை, *இருங்காழ் உலகை இரும்பு முகம் தேய்த்த / அமைப்புமாண் அரிசி அமலை வெண் சோறு* (சிறுபாண்., பா. 193 – 194) என்று *சிறுபாணாற்றுப்படை* எடுத்துரைக்கிறது. *...வயக்களிற்றுக் / கோடு லக்கையாக, நற்செம்பி னிலை சுளகா / ஆடுகழை நெல்லை யறை உரலுள் செய்து ருவாம்* (கு.க. 5: 1-3) இது பாறைகளில் குழி அமைத்து உரலாகப் பயன்படுத்தியதை விளக்கியுரைக்கிறது. மேலும் செம்பின் இலையை சுளகாகவும், யானைத் தந்தத்தையே உலக்கையாகவும் பயன்படுத்திய சேதி நம்மை வியக்கச் செய்கிறது. காவிரிப் பூம்பட்டினத்தில் சோறாக்கி வடித்த கஞ்சி அந்நகர் முழுவதும் சாலைகளில் பாய்ந்தோடியதாக *பட்டினப்பாலை*,

> சோறுவாக்கிய கொழுங்கஞ்சி

> யாறுபோலப் பரந்தொழுகி (பட்., பா: 44 – 45)

என்கிறது. எனவே தம்மை நாடி வருகின்ற ஏழை, எளியவர்களை பண்டைத் தமிழர்கள் இன்முகத்துடன் அழைத்து உணவு பரிமாறிய செய்தி புலனாகிறது.

சோற்றுணவு

பழந்தமிழர் ஐந்திணைக்கேற்ப தங்கள் வாழிடங்களை அமைத்துக் கொண்டனர். குறிஞ்சி நிலமக்கள் ஐவனம் என்னும் மலையரிசி, மூங்கிலரிசி போன்றவற்றையும் முல்லைநில

மக்கள் தினைச்சோறு, பால்சோறு போன்றவற்றையும் மருதநில மக்கள் செந்நெல் மற்றும் வெண்ணெல் சோற்றையும் அதிகமாக உட்கொண்டனர்.

அமிழ்தன மரபி ஞான்றுசை யடிசில் (புறம்., பா. 390: 17)

அயினியு மின்ற யின்றனனே (புறம்., பா. 77: 8)

அவிழ் வேண்டுநர்க் கிடையருள் (புறம்., பா. 366: 20)

கூழுடைக் கொழு மஞ்சிகை (பட்டினப்பாலை: 163)

இழித்தானாப் பல சொன்றி (மது.காஞ்சி., பா. 212)

பெருஞ்சேய் நெல்லின் கொக்குகிர் நிமிரல (புறம்., 395: 36)

விரலென நிமிர்ந்த நிரலமை புழுக்கல் (பொருந., பா. 114)

பாற்பெய் புன்கம் தோனரகு மயக்கி - (பாற்சோறு) (புறம்., பா. 34: 10)

குரூஉக் கணிறடிப் பொம்மல் (மலைபடு., பா. 169)

பாலு மிதவையும் பண்ணாது பெறுகுவீர் (மலைபடு., பா. 417)

இன்சுவை மூரற் பெறுகுவீர் (பெரும்பா., பா. 18)

சுடர்க் கடைப் பறவைப் பெயர்படு வத்தம் (பெரும்பா., பா. 305)

என்னும் சங்க இலக்கியப் பாடலடிகள் சங்ககாலத்தில் சோற்றுணவு பல பெயர்களில் வழங்கப்பட்டதையும், ஒவ்வொரு உணவும் தயாரிக்கும் முறைகளால் சிறு சிறு அளவில் சுவையில் மாறுபாட்டுடன் காணப்பட்டதையும் எடுத்துரைக்கிறது. அடிசில், அயினி, அவிழ், கூழ், சொன்றி, நிமிரல், புழுக்கல், புன்கம், பொம்மல், மிதவை, மூரல், வத்தம் என பல பெயர்களில் சோற்றுணவு வழங்கப்பட்டு வந்தது இதனால் அறியலாகிறது.

சோற்றின் வகைகள்

பழந்தமிழர்கள் பல வகையிலான சோற்றுணவை தயாரித்து உண்ட செய்தியை சங்க இலக்கியம் எடுத்துரைக்கிறது. சோற்றின் சுவைத் தன்மைக்கேற்ப பல பெயர்களால் சோற்றுணவு வழங்கப்பட்டது.

செஞ்சோறு

சங்க காலத்தில் செஞ்சோற்றுக்கு அதிக முக்கியத்துவம் கொடுத்தனர். இதனை, "செந்நெல் வல்சி " (பதி. பத்., பா. 75: 12) என்கிறது *பதிற்றுப் பத்து*. மருத நில மக்கள் செந்நெல், வெண்ணெல், போன்றவற்றிலிருந்து கிடைத்த சோற்றையே அதிகம் உண்டனர். எனினும் செந்நெல்லே உயர்ந்தது என்பதை இன்றுவரை சிரமேற்கொண்டு தமிழர்கள் 'சம்பா அரிசி' என்ற பெயரில் பயன்படுத்தி வருவது நாம் அறிந்த உண்மையாகும். அரிசியை நன்கு தீட்டாமல் பயன்படுத்தியதை சங்க இலக்கியம் எடுத்துரைக்கிறது.

வெண்சோறு

பழந்தமிழர்கள் வெண்ணெல் அரிசியைக் குற்றி பின்னர் சோறாக்கி உண்டனர். "அவைப்பு மாண் அரிசி அமலை வெண்சோறு" (சிறுபாண்., பா. 194) என்கிறது

சிறுபாணாற்றுப்படை நன்றாக வெண்மையாக தீட்டப்பட்ட அரிசியைக் கொண்டு உருவாக்கிய சோறு என்பதால் வெள்ளையரிசிச் சோறு எனப்பட்டது.

ஊன்சோறு

சோற்றோடு ஊன் (இறைச்சி) சேர்த்து ஆக்கும் சோறு 'ஊன்சோறு' எனப்பட்டது. இன்றைய 'பிரியாணியே' அன்றைய ஊன்சோறு என்றால் மிகையாகாது.

துடித்தோட்கை துடுப்பாக

ஆடுற்ற ஊன்சோறு

நெறியறிந்த கடிவாலுவன் (மதுரை.கா., பா. 34 – 36)

தனது பெரிய கைகளால் ஊன்சோற்றை விரவிய சமையலாள் தேவையறிந்து அதனை பிறருக்கு வழங்கியது தெரியவருகிறது.

நெய்ச்சோறு

நெய்யினை சோற்றோடு கலந்து ஆக்கப்படும் சோறு நெய்ச்சோறு எனப்பட்டது.

குய் கொள் கொழுந்துவை நெய்யுடை யடிசில்... (புறம்., பா: 190)

நெய்யுடை யடிசில் மெய்பட விதிர்த்தும் (புறம்., பா. 18: 4 – 5)

பைந்நிணம் ஒழுகிய நெய்ம்மரி அழகில்

வகையில் வான்றிணைப் புரையோர் கடும்பொடு

விருந்துண் டெஞ்சிய மிச்சில் (கு.பா., பா. 204 – 206)

என்னும் சங்க இலக்கியப் பாடலடிகள் வழி பழந்தமிழர்கள் நெய்கலந்த சோற்றினை மிகுதியாக பயன்படுத்திய செய்தி தெரிய வருகிறது.

உளுந்தஞ்சோறு

சோற்றுடன் உழுந்தைச் சேர்த்து ஆக்குவது உழுந்தஞ்சோறு ஆகும்.

நெய்யோடு மயக்கியவுழுந்து நூற்றன்ன (ஐங்குறு., பா. 211: 1)

உழுந்துதலைப் பெய்த கொழுங்களி மிதவை

பெருஞ்சோற் றமலை நிற்ப (அகம்., பா. 86: 1 – 2)

சோற்றுடன் நெய்யும் உழுந்தும் சேர்த்து சமைத்து உண்டது அறியலாகிறது.

பாற்சோறு

சங்க இலக்கியத்தில் பாற்சோறு பற்றிய செய்திகள் மிகுதியாகவே காணப்படுகின்றன. பல்யாட்டின நிரை எல்லினர் புகினே / பாலு மிதவை பண்ணாறு பெறுகுவீர் (மலைபடு கடாம், பா. 416 – 417) என்னும் *மலைபடுகடாம்* பாடிலடிகள் ஆநிரைகள் மிகுதியாக காணப்படும் இடையர்தம் வீட்டில் தங்கும் விருந்தினர்கள். பாலும், பாற்சோறும் பெற்றனர் என்றுரைக்கிறது.

புளிச்சோறு

சோற்றில் புளியை கலந்தாக்குவது புளிச்சோறு எனப்பட்டது. புளிக்கறி, புளிக்கூழ் மற்றும் புளியை ஊற்றி தயாரிக்கப்பட்ட புளிப்பு சுவையுடைய உணவுகளை பண்டைத்

தமிழர்கள் விரும்பி உண்டனர். "ஈயல் பெய்தட்ட இன்புளி வெஞ்சோறு" என்னும் பாடலடி ஈயலுடன் (ஈசல்) சேர்த்து சமைத்த புளிச்சோறு பற்றிக் கூறுகிறது.

இன்சோறு

இனிப்புப் பொருள்கள் சேர்த்து சோறு சமைத்து உண்பது பண்டைத் தமிழர்களின் வழக்கமாக இருந்தது. "இன்சோறு தருநர் பல்வயினுகர" (மதுரை. கா., பா. 535) விருந்தினர்களுக்கு இன்சோறு வழங்கி உபசரித்தை மதுரைக்காஞ்சி எடுத்துரைக்கிறது. 'விருந்து' என்ற சொல்லுக்கு 'புதுமை 'என்று பொருள் கூறுகின்றது தொல்காப்பியம். அவ்வாறானால் விருந்தினர் நமக்கு முன் பின் தெரியாதவர்களாகவும் இருக்கலாம் என்று பொருள். (Proceedings of the Ninth all India Conference of SRF, P. 11)

உப்பில்லாச் சோறு

வறுமையின் காரணமாக வாடிய சங்ககாலத் தமிழ்த் தாயொருத்தி உப்பில்லாச் சோற்றை உண்டு வாழ்ந்தாள் என புறநானூறு கூறுகின்றது. ...உப்பில்லா அவிப் புழுக்கல் / கைக்கொண்டு... (புறம்., பா. 363 : 12 – 13) பெருஞ்சித்திரனார் என்னும் புலவர் தனது வறுமை நிலையால் உப்பில்லாமல் கீரையை அவித்துத் தின்ற செய்தி இங்கு நினைவுகூரத் தக்கது. "இவ்வுலகில் வாழ்கின்றவர்களுக்கெல்லாம் உணவு கொடுத்தவர் உயிர் கொடுத்தவர் ஆவர் என்பதை மணிமேகலையில் பாத்திரம் பெற்ற காதையில் சாத்தனார் எடுத்துக்காட்டுகிறார்." (நவீனத் தமிழாய்வு, ப. 408) உணவின் அத்தியாவசிய தேவை உணர்த்தப்படும் இடம் இதுவாகும்.

மூங்கிலரிசிச் சோறு

தோரை நெல், முடந்தை நெல், வெதிர்நெல் என்றெல்லாம் அழைக்கப்படுவது மூங்கிலரிசியாகும். பழந்தமிழர்கள் மூங்கிலரிசி உணவை பயன்படுத்தியதை, புயற் புனிறு போகியயூமலி புறவி / னவற் பதங் கொண்டன வம்பொதிந் தோன (மலைபடு., பா. 120 – 121) என்னும் மலைபடுகடாம் பாடல் அடிகள் தெரிவிக்கின்றது. நம் முன்னோர் மூங்கிலரிசியை சோறாக்கி விரும்பி உண்ட செய்தியை இது எடுத்துரைக்கிறது.

புல்லரிசிச்சோறு

புல்லிலிருந்து கிடைக்கப்பெற்ற சிறிய அரிசியை பயன்படுத்தியும் உணவு சமைத்தனர். இன்றும் சில கிராமங்களில் சாமைப்புல், காடைக்கண்ணி புல் போன்றவற்றின் அரிசியை பயன்படுத்தியும் சோறு சமைக்கின்றனர். "பகைவர் புல்லார்க" என்னும் ஐங்குறு நூற்றுப் பாடலடி பகைவர்க்கு புல்லரிசிச் சோறு கொடுத்ததை எடுத்துரைக்கிறது. அவ்வாறெனில் பகைவர்களை இழிவுபடுத்தும் பொருட்டு புல்லரிசி சோறு வழங்கப்பட்டது தெளிவாகிறது.

அவல்

நெல்லை தண்ணீரில் ஊறவைத்து முளைகட்டி வரும் பருவத்தில் அதனை வறுத்து சூடு தணியும்முன் உரலிலிட்டு குற்றி இடித்து அதை முறத்தில் வைத்து புடைத்து உமி நீக்கி அவல் பெறப்பட்டது. மூங்கில் நெல்லிலிருந்தும் அவல் இடிக்கப்பட்டச் செய்தியை அகநானூறு,

பெருஞ் செய் நெல்லின் வரங்குகதிர் முறித்துப்

பாசவல் இடிக்கும் பெருங்காழ் உலக்கை (அகம்., பா. 121: 1 – 2)

என்கிறது. உரலில் உலக்கை கொண்டு நெல், அரிசி மற்றும் அவலை குற்றி பக்குவப்படுத்தியதை இதனால் அறியலாம்.

...கழனிக் கரும்பின்

விழைகழை பிழிந்த அந்தீஞ் சேற்றொடு

பால்பெய் செந்நெற் பாசவல் பகுக்கும்(அகம்., பா. 121: 1 – 2)

என்ற அகநானூற்றுப் பாடலடிகளும் இதற்கு சான்றாகிறது.

அவல் - நெற்பொரியிடியல் – (தமிழ்ப் பேரகராதி ப. 154)

என்று தமிழ்ப் பேரகராதி விளக்கம் தருகிறது.

பொரி

தமிழர்கள் அரிசியை வறுத்து அரிசிப்பொரி தயாரித்து உண்டனர். நெல்லை வறுத்து கிடைப்பது நெற்பொரி எனப்பட்டது.

புழுங்க நெல்லின் பொரி (ஐங்., பா. 4: 1 – 2)

செந்நெல் வான் பொரி (குறு., பா. 53 : 4)

...வெண்பொரி சிதறி (திருமுரு., பா. 231)

செம்பொற் புனைகலஞ் தம்பொரிக் கயந்த

பாலும் பலவென உண்ணர் (குறு., பா. 356: 6 – 7)

என்னும் பாடலடிகள் தமிழர்கள் பொரிக்கு கொடுத்த முக்கியத்துவத்தை பறைசாற்றுகின்றன.

தொகுப்புரை

சங்க இலக்கியத்தில் சோற்றுணா பற்றிய செய்திகள் மிகுந்து காணப்படுகின்றன. அடிசில், அயினி, அவிழ், கூழ், சொன்றி, நிமிரல், புழுக்கல், புன்கம், பொம்மல், மிதவை வத்தம், மூரல் என பல பெயர்களில் சங்க இலக்கியத்தில் சோறு வழங்கப்படுகிறது. சங்க காலத்தல் பல்வேறு தரப்பட்ட அரிசிகள் காணப்பட்டன. எனினும் செந்நெல்லரிசி, வெண்ணெல்லரிசி, காடைக்கண்ணியரிசி, தினையரிசி, வரகரிசி, மூங்கிலரிசி போன்றவை சிறப்பிடம் பெற்றன. சிறுசோறு மங்கள நிகழ்ச்சிகளிலும் பெருஞ்சோறு இறந்தவர்களின் நினைவாகவும், போரில் இறந்த வீர மறவர்களுக்கு எழுப்பப்பட்ட நடுகல் வழிபாட்டின்போதும் பிண்டம் வைத்தல் பொருட்டு தயாரிக்கப்பட்டன. பெருஞ்சோறு அல்லது பிண்டம் வைத்தல் என்பது காக்கைக்கு இன்றளவும் தம் முன்னோரை நினைத்து வைக்கின்ற 'பலி'ச் சோற்றைக் குறிக்கின்றது. பழந்தமிழர்கள் செஞ்சோறு, வெண்சோறு, உப்பில்லாச்சோறு, உழுந்தஞ்சோறு, ஊன்சோறு, நெய்ச்சோறு, புளிச்சோறு, பாற்சோறு, மூங்கிலரிச்சோறு, புல்லரிசிச்சோறு போன்ற பல்வேறுவிதமான சோற்று வகைகளை தயாரித்து உண்டு மகிழ்ந்தனர். சங்கத் தமிழர்கள் சோற்றுணவுக்கு அதிக முக்கியத்துவம் கொடுத்தனர். அவர்தம் கடின உழைப்பாகிய வேட்டையாடுதல், வேளாண்மை, ஆடு மாடு மேய்த்தல் போன்ற தொழில்களால் உண்ட

உணவு உரிய நேரத்தில் செரித்து உடலுக்கும், மனதுக்கும் புத்துணர்வு தந்தது. அதி நவீன யுகமாகிய இக்காலத்தில் காணப்படும் பல்வேறுவிதமான ஆயத்த உணவுகளில் அதிகமான கொழுப்பும், உடலுக்கு ஊறு விளைவிக்கக்கூடிய உணவுக் கலவைகளும் காணப்படுவதால் அவை உடலை வெகுவாக பாதித்து பல்வேறுபட்ட நோய்களுக்கு அடித்தளமிடுகிறது. 'சுவர் இருந்தால்தான் சித்திரம் வரைய முடியும்'. ஆனால் மனிதனின் கட்டுப்பாடற்ற வாழ்க்கை முறையினால் மனுக்குலம், முழுவதும் பாழாகி அழிந்து வருகிறது. இந்நிலை மாற வேண்டுமென்றால் ருசிக்காக இல்லாமல் பசிக்காக சாப்பிடும் நிலை வரவேண்டும். உணவே மருந்தாகும் நிலை வளர வேண்டும்.

துணைநூற்கள்

[1] அடைக்கலசாமி, எம்மார், *தமிழ் இலக்கிய வரலாறு.* ராசி பதிப்பகம், 90, கணேஷ் நகர், சென்னை – 73, 2007.

[2] ஆனந்தன், க. *தமிழ் இலக்கிய வரலாறு.* கண்மணி பதிப்பகம், திருச்சி – 2, 2002.

[3] இராமசுவாமி, க.அ. *குறுந்தொகை வசனம்.* ஒற்றுமை ஆபீஸ், வியாசராவ் தெரு, தியாகராய நகர், மதராஸ், 1941.

[4] இராமநாதன், இல. ப.கரு. *சங்ககாலத் தமிழர் வாழ்வு.* வெள்ளையன் பதிப்புக் கழகம், இராமநாதபுரம், 1958.

[5] இராமநாதன் செட்டியார், இல.ப.கரு. *சங்க காலத் தமிழர் வாழ்வு.* சென்னைப் பல்கலைக்கழகம், சென்னை – 5, 1964.

[6] கௌமாரீஸ்வரி, எஸ். (ப.அ) *தொல்காப்பியம் பொருளதிகாரம்.* இளம்பூரணர் உரை, சுரதா பதிப்பகம், திருவல்லிக்கேணி, சென்னை – 600 005, 2008.

[7] சுப்பிரமணியன். ச.வே. (உ.ஆ), *பரிபாடல்.* மணிவாசகர் பதிப்பகம், தொகுதி II பாரிமுனை, சென்னை – 108, 2010.

[8] சுப்பிரமணின், ச.வே. (ப.ஆ.), *தமிழ்ச் செவ்வியல் நூல்கள் மூலம் முழுவதும்.* மணிவாசகர் பதிப்பகம், பாரிமுனை, சென்னை, 2008.

[9] செயபால், இரா. *அகநானூறு – 1* (மூலமும் உரையும்) நியூ செஞ்சுரி புக் ஹவுஸ் (பி) லிட், சென்னை – 600 098, 2004.

[10] சோம சுந்தரனார், பொ.வே. (உ.ஆ), *ஐங்குறுநூறு.* திருநெல்வேலி தென்னிந்திய சைவ சித்தாந்த நூற்பதிப்புக் கழகம், சென்னை, 1966.

[11] தமிழண்ணல். *குறிஞ்சிப்பாட்டு.* மீனாட்சி புத்தக நிலையம், மதுரை, 1978.

[12] துரைசாமிப்பிள்ளை, சு. ஒளவை. *புறநானூறு.* தென்னிந்திய சைவ சித்தாந்த நூற்பதிப்புக் கழகம், 140 பிரகாசம் சாலை, சென்னை – 1, 1947.

[13] நச்சினார்க்கினியர் (உ.ஆ), *கலித்தொகை.* திருநெல்வேலி சைவசித்தாந்த நூற்பதிப்புக் கழகம் லிமிடெட், சென்னை – 18, 1943.

[14] நவீனத் தமிழாய்வு (பன்னாட்டுப் பன்முகத் தமிழ் காலாண்டு ஆய்விதழ்) மார்ச் 15, 2019 – சிறப்பிதழ்.

[15] நாகராசன், வி. (உ.ஆ), *குறுந்தொகை.* (மூ.உ), நியூ செஞ்சுரி புக் ஹவுஸ், பாவை பிரிண்டர்ஸ், சென்னை – 600 014, 2007.

[16] புலியூர்க் கேசிகன். *குறுந்தொகை* (மூ.உ, சாரதா பதிப்பகம், சென்னை – 14, 2010,.

[17] வேங்கடசாமி நாட்டார், ந.மு. (உ.ஆ), *அகநானூறு.* (மூ.உ), திருநெல்வேலி, தென்னிந்திய சைவ சித்தாந்த நூற்பதிப்புக் கழகம், சென்னை, 1957.

[18] வையாபுரிப்பிள்ளை, சு. *சங்க இலக்கியம்* (பாட்டும் தொகையும்). பாரி நிலையம், சென்னை, 1940.

தென் திருவிதாங்கூர் இந்துக் கல்லூரி

தமிழ்த்துறை உயராய்வு மையம், நாகர்கோவில்.

மற்றும்

செங்காந்தள்

அறிஞர்களால் மதிப்பீடு செய்யப்படும் தமிழ் ஆய்விதழ்

இணைந்து நடத்தும்

பன்னாட்டுக் கருத்தரங்கம்

சான்றிதழ்

திரு / திருமதி / செல்வி / முனைவர் __அ. சிதாரா ஜாய்ஸ்__

அவர்கள் 06.10.2023 அன்று தென் திருவிதாங்கூர் இந்துக் கல்லூரி மற்றும் *செங்காந்தள்* ஆய்விதழ் இணைந்து *சங்க இலக்கியங்களில்*

வேளாண்மையும் உணவு முறையும் என்னும் பொருண்மையில் நடத்திய பன்னாட்டுக் கருத்தரங்கத்தில் கலந்துக் கொண்டு

__சங்க இலக்கியத்தில் சோற்றுணா__

என்னும் தலைப்பில் வழங்கிய ஆய்வுக்கட்டுரை *செங்காந்தள்* மலர் -2 சிறப்பிதழ் - 3 புரட்டாசி 2054இல் மின்பிரசுரம் செய்யப்பெற்றது

என்று சான்றளிக்கப்படுகிறது.

முனைவர் பொ. இரவீந்திரன்
கருத்தரங்க ஒருங்கிணைப்பாளர்

முனைவர் பா. ம. ஜெயகலா
கருத்தரங்க ஒருங்கிணைப்பாளர்

முனைவர் க.கோ.அனிதாகுமாரி
தமிழ்த்துறைத் தலைவர்

முனைவர் தெ.ம. பத்மநாபன்
முதல்வர்

காண்க: Chenkaantal - http://pandianeducationaltrust.com/-chenkaantal.html MLA, DOAJ, Google Scholar, ROAD, Mir@bel Indexed

Experimental and theoretical investigation of structure activity relationship on L-Lysine Monohydrate for antioxidant efficacy

T. Brintha [a,e], J Jeni James [b,], M Amalanathan [c,e], P.J. Jegan Babu [d,e], M. Sony Michael Mary [d,e]

[a] Research Scholar, Reg No: 1921311213201, Nesamony Memorial Christian College Marthandam, Kanyakumari, Tamil Nadu, India
[b] Research Scholar, Reg No:21113102132001, Department of Physics & Research centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai 629153, Tamil Nadu, India
[c] Department of Physics & Research centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai 629153, Tamil Nadu, India
[d] Department of Physics, Nesamony Memorial Christian College Marthandam, Kanyakumari, Tamil Nadu, India
[e] Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli 627012, Tamil Nadu, India

ARTICLE INFO

ABSTRACT

Comprehensive spectroscopic research has been undertaken to investigate the structural behaviour of the L-lysine monohydrate molecule. The spectral properties of the L-lysine monohydrate molecule in solid phase were examined using Fourier Transform Infrared (FTIR) and Fourier Transform Raman methods. The B3LYP/6–311++G (d, p) computations were used to optimize the structure of the molecule. To provide complete vibrational spectral assignments, vibration energy distribution analysis (VEDA) was used. The Natural bond orbital (NBO) analysis explains the stability and distinct forms of hydrogen bonds with in the molecule. The chemical stability of the molecule is predicted by Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) analysis. Non-Covalent Interaction (NCI) analysis was done to identify weak interactions according to density of electron of the title compound. The fukui function identified the chemical reactivity sites. To predict its antioxidant efficacy, docking studies were done.

1. Introduction

In recent days degenerative illnesses like cancer and cardiovascular conditions are more likely to develop as a result of oxidative damage [1–3]. Antioxidants control the actions of reactive oxidants, reducing oxidative damage to biomolecules [4,5]. It is believed that antioxidants help our bodies fight free radicals, which enhances general health. In human nutrition, L-lysine is a necessary amino acid, which means that the body cannot generate it and must be acquired by diet or supplementation [6]. The most well-known use of amino acids is as the building blocks of proteins [7]. The amino acid lysine is found in relatively significant levels in nuclear histone [8]. Heinrich Drechsel, a German dentist, was the first to isolate lysine from casein (a milk phosphoprotein) in 1889 [9]. Lysine is necessary for normal growth and for the formation of carnitine, a nutrient that reduces cholesterol and transforms fatty acids into energy. This amino acid, which is one of the most important building blocks of muscle tissue, is often utilized by athletes to assist lean mass gain and overall muscle and bone health. All proteins in the body require L-lysine as a building block. It is necessary for the body to produce hormones, enzymes and antibodies as well as to absorb calcium, build muscle protein and to recover from incision [10]. Intermolecular charge transfer is made possible by the donor NH2 and acceptor COOH groups that are present in many amino acids [11]. It is used as a fine chemical, particularly as an ingredient in infusion solutions for pharmaceutical applications and as a precursor for industrial chemicals, in human medicine, cosmetics and the pharmaceutical industry [12,13].

In the ab initio discipline, the density functional theory (DFT) approach has become a prominent post-HF methodology for the calculation and optimization of structural properties, energies of the molecule, and frequencies in vibration, as well as the exact evaluation of a variety of molecular qualities. Researchers employ FT-IR, FT-Raman, and density functional theory (DFT) techniques to discover structural information, functional groups, and other quantum level properties [14–16]. In recent times, DFT calculations are essential for frequency and spectral intensity prediction. Over the past few years, significant
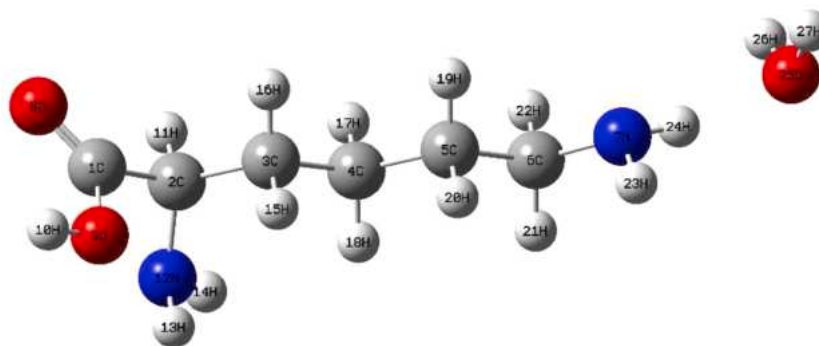
Fig. 1. Optimized molecular structure of ʟ-Lysine monohydrate.

Table 1
Fock matrix second order perturbation theory analysis in NBO basis at basic set.

Donor NBO (i)	ED	Acceptor NBO (j)	ED	E (2) Kcal/mol	E(j) – E(i) a.u	F(i, j) a.u
LP_1O_9	1.97703	$\sigma^*(C_1-O_8)$	0.02169	6.78	1.25	0.082
LP_1N_{12}	1.95292	$\sigma^*(C_2-C_3)$	0.03449	8.59	0.67	0.068
$\sigma^*(C_1-O_9)$	0.1033	$\sigma^*(C_2-N_{12})$	0.01816	6.94	0.01	0.032
LP_1N_7	1.9350	$\sigma^*(O_{25}-H_{27})$	0.03366	13.23	0.81	0.093

$E_{LUMO} = -0.63674$ eV

Gap = 6.2555 eV

$E_{HOMO} = -6.3674$ eV
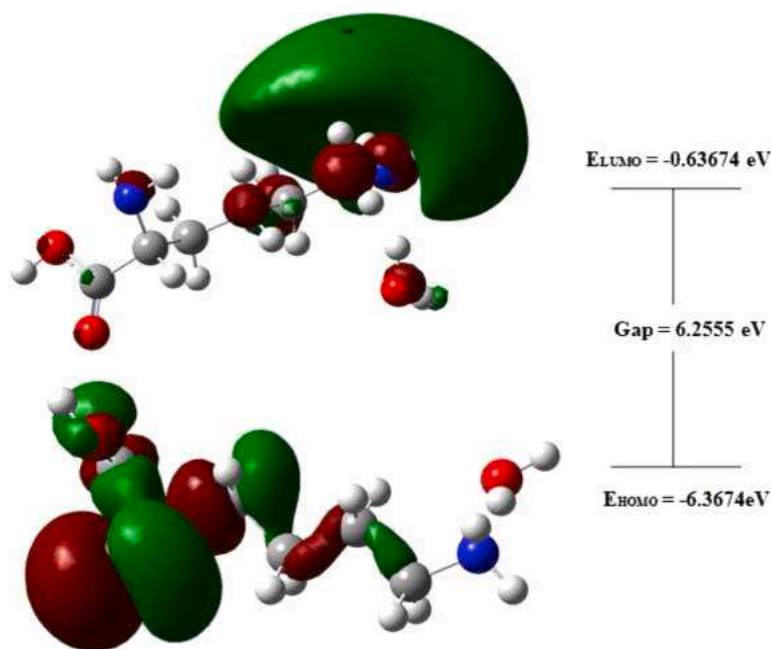
Fig. 2. ʟ-Lysine monohydrate HOMO-LUMO orbitals.

progress has been made to create novel DFT approaches that can be integrated with the available quantum chemical computing resources [17–19].

An analysis of bond orbitals (NBO) was done to predict the stability of the compound ʟ-lysine monohydrate. Fukui function investigation is used to determine the significant and qualitative characteristics of the reactive. The stability, or energy gap, is determined using frontier molecular orbital analysis. The fields of structural molecular biology, drug-gene testing, and computer-aided drug design all make use of molecular docking to some extent [20]. Molecular docking analysis of a molecule is a useful and effective computational method for forecasting a ligand's binding mechanism and affinity for proteins. The structure of the target protein structure can be determined using the protein data bank format. Molecular docking of protein-ligand can anticipate the ligand's preferred orientation in relation to the protein in order to create a stable complex and its activities. [21].

To the best of our knowledge, no DFT studies, vibrational assignments or molecular docking studies have been carried out for the title compound ʟ-Lysine Monohydrate, according to review of the literature. The main objective of this work focuses primarily on the complete vibrational and electronic analysis of ʟ-lysine monohydrate. FTIR and FT-Raman of the ʟ-lysine monohydrate is determined. The Mulliken and natural atomic charge distributions were also computed. The FMOs analysis was done to investigate the reactivity and stability of the

Fig. 3. (a) L-Lysine Monohydrate Reduced Density Gradient scatter plot. (b) L-Lysine Monohydrate Reduced Density Gradient structure.

three different proteins to identify the best activity of the title molecule.

2. Materials and methods

L-Lysine Monohydrate was acquired with 99 percent purity from Sigma-Aldrich and utilized as such without additional processing. By using a PERKIN ELMER FTIR spectrometer at 1.0 cm^{-1} resolution and in the range of 4000–450 cm^{-1} the spectra for FT-IR is obtained. For obtaining FT-IR spectra, KBr pellet technique in discharge mode is used. Bruker RFS 27 spectrometer in the range of 4000–100 cm^{-1} was used to acquire the FT-Raman spectrum of the Neodymium-doped Yttrium aluminum Garnet (Nd-YAG) laser with 200 mW powers.

3. Computation details

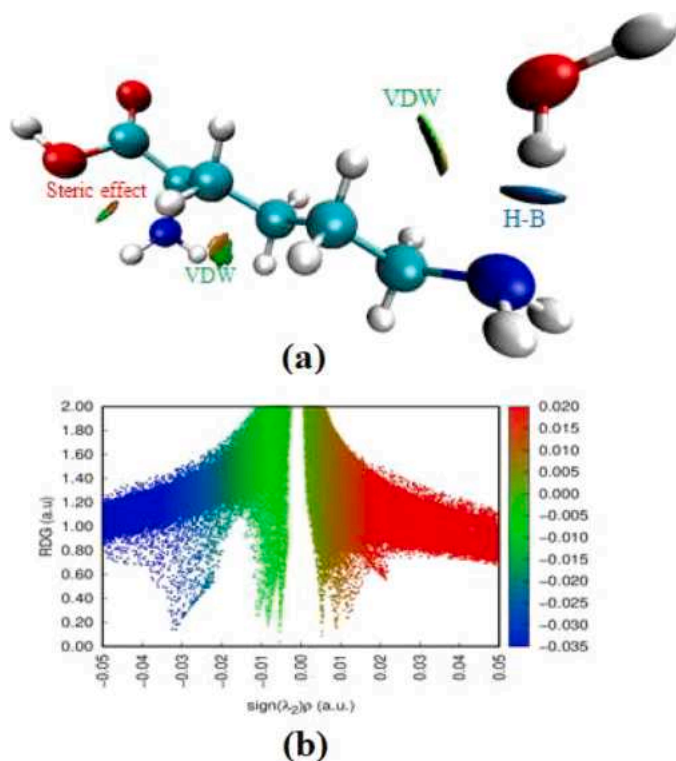The computations (DFT/B3LYP) at the 6–311+G (d, p) basic set level were performed using the programme package of Gaussian 09. The geometry of the molecule was optimized using Berny's optimization technique. HOMO and LUMO indicates the transfer of charge present in the molecule. In order to offer information about numerous interactions, NBO analyses were performed. The values for electronegativity, hardness, and softness were also calculated. Multiwfn software [22] was used to conduct AIM, Electron Localization Function (ELF), Localized Orbital Locator (LOL) and Non-Covalent Interaction (NCI) analysis. The NBO and AIM programs were used to determine charge in the atoms, molecular electrostatic potentials, bond ordering, and topological characteristics, while frontier orbitals and equations for certain common descriptors were used to determine reactivities and behaviors in different media [23,24]. The animation tool in the GAUSSVIEW programme facilitates in the assignment of determined wavenumbers by providing a visual depiction of the vibrational modes. Molecular docking investigations were conducted utilizing the Auto dock vina software to study the compound's biological activities. After that, the vibrational analysis was done utilizing contributions from the potential energy distribution (PED).

molecule. NBO analysis calculated the redistribution of electron density (ED) and E (2) energies in various antibonding and bonding orbitals using DFT calculations, providing clear insight into the stability resulting from hyperconjugation of various intramolecular interaction evidences. Electron Localization Function (ELF), Localized Orbital Locator (LOL) and Molecular Electrostatic Potential (MEP) were done to analyse the reactive sites of the molecule. Docking analysis were performed on
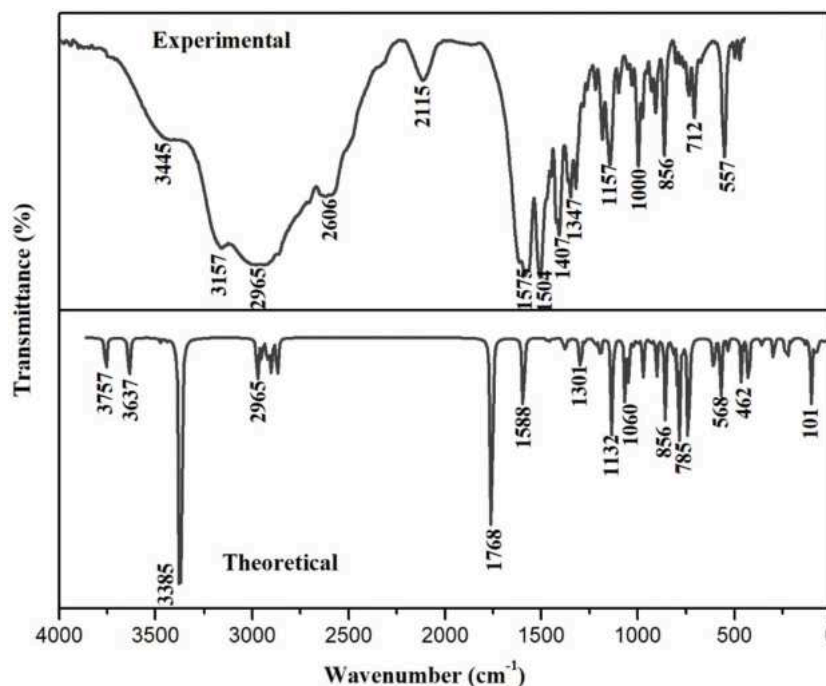

Fig. 4. Theoretical and experimental FT-IR spectra of the title molecule.
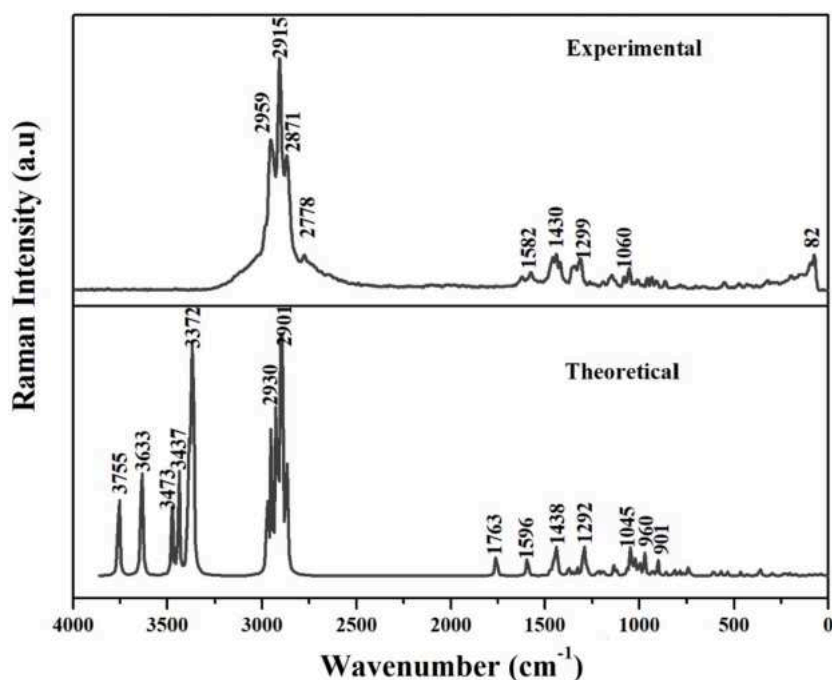
Fig. 5. Theoretical and experimental FT-Raman spectra of the title molecule.

4. Result and discussion

4.1. Molecular geometry

The arrangement of the atoms that make up a molecule in three dimensions is known as molecular geometry. It affects the phase of matter, polarity, colour, reactivity, biological activity and magnetism of a substance. The chemical bonds that connect each atom to its neighbour atoms define its position. The position of these atoms in space may be used to characterize the molecular geometry invoking two joined atoms which is bond length, three connected atoms which is the bond angles and three consecutive bonds which gives rise to dihedral angle [25]. The optimized structure of the title molecule, including enumeration of geometric properties of the atom including length of the bond, angle between the bond and dihedral angle, is produced using Gaussian09 software, as demonstrated in Fig. 1. (Supplementary Table 1) displays the bond length, bond angle, and dihedral angle parameters.

The Becke,3-parameter, Lee-Yang-Parr/6–31 G (d, p) method was used to calculate the optimal properties (length of the bond, angle between the bond and dihedral angle) of the title compound L-Lysine monohydrate. L-Lysine monohydrate has five Carbon-Carbon, two Carbon-Oxygen, nine Carbon-Hydrogen, two Carbon-Nitrogen, four Nitrogen-Hydrogen and three Oxygen-Hydrogen bonds. The bond lengths of C_1—C_2, C_2—C_3, C_3—C_4, C_4—C_5, and C_5—C_6 are deemed to be larger than the actual C—C bond length and this increase in the bond length occurs because of the transfer of electrons from nitrogen to the carbon atoms. The C_1—O_8 bond length is obtained as 1.208 Å and that is lower compared to the exact C—O bond and this predicts the existence of double bond. Other C_1—O_9 bond is higher with bond length 1.358 Å which is due to the presence of OH group and it shows the existence of single bond. C_2—H_{11}, C_3—H_{15}, C_3—H_{16}, C_4—H_{17}, C_4—H_{18}, C_5—H_{19}, C_5—H_{20}, C_6—H_{22} are 1.093 Å, 1.097 Å, 1.096 Å, 1.097 Å, 1.097 Å, 1.094 Å, 1.099 Å, 1.094 Å respectively, which are almost equal to the actual C—H bond lengths. The small disagreement with the theoretical value with the values from the experiment could be attributed to the conviction that the data that are obtained from experiment were done in solid state, whereas the computational values were obtained in the gaseous state. The bond length C_6—H_{21} is 1.1 Å which is higher compared to the

entire C—H bond and this is caused by the electron transfer within the atoms. The C—N bond lengths namely C_2—N_{12} and C_6—N_7 are higher and this is the result of charge transfer from C to N atom. N_7—H_{23}, N_7—H_{24}, N_{12}—H_{13}, and N_{12}—H_{14} have bond lengths of 1.017 Å, 1.016 Å, 1.015 Å, and 1.013 Å, respectively, which are comparable to the typical N—H bond length. The bond length of O_9—H_{10} and O_{25}—H_{26} are found to be 0.969 Å and 0.980 Å, which are also nearly equal to the actual O—H bond.

4.2. Natural bond orbital (NBO) analysis

Natural Bond Orbital (NBO) is an approach for studying intermolecular and intramolecular interactions [26]. The value of E (2) refers to the interactivity within the electron donor and the electron acceptor, i.e. the tendency of the electron donor to contribute to the electron acceptor. The higher the E (2) value, the more intense the connection between electron donors is and the more conjugation occurs across the system. Electron transfer between occupied NBO orbitals (bonding or lone pair) and occupied non-Lewis NBO orbitals determines the donor acceptor (antibonding or Rydberg) interaction. The second order fock matrix perturbation theory [27] completely describes the intermolecular delocalization in the molecule. The GAUSSIAN 09 programme is used to do the NBO analysis. The GAUSSIAN 09 programme is used to do the NBO analysis. The delocalization of interaction energy is calculated using the second order perturbation formula. Table 1 summarizes the results of NBO analysis on the title molecule.

The type of bond orbitals with their occupancy and energy as well as the percentage of NBOs in each hybridization along with percentage of s and p characters given by B3LYP/6–311++G (d, p) method is shown in supplementary Table 2. At significant natural atomic orbital energies, supplementary Table 2 displays the occupancy of electrons and the p character [28]. The title compound LP_2O_8 and LP_2O_9 possess low occupancy number of 0.9274 and 0.9059e with high energy values of −0.4948 and −0.4948 a.u with p-character which is approximately equal to 99.9 %. While LP_1O_9, LP_1N_{12} and LP_1N_7 are the orbitals with high occupancy and low energy values of −0.7886, −0.5573 and −0.5323 a.u, which possess the occupancies of 1.9770, 1.9529 and 1.9350e respectively.

Table 2

Vibrational analysis of L-lysine monohydrate using VEDA.

Wave numbers Experimental values		Scaled values	Assignments with PED (>10%)
FT-IR	FT-RAMAN		
-	-	3759	$\upsilon O_{25}H_{26}$ (98)
-	-	3637	υO_9H_{10} (100)
-	-	3476	$\upsilon N_{12}H_{13}$ (99)
3445	-	3443	υN_7H_{23} (97)
-	-	3392	$\upsilon N_{12}H_{14}$ (99)
-	-	3377	$\upsilon O_{25}H_{27}$ (94)
-	-	3368	υN_7H_{24} (96)
3157	-	2975	υC_2H_{11} (76)
2965	-	2971	υC_5H_{19} (69)
-	2959	2954	υC_3H_{15} (85)
-	-	2950	υC_2H_{11} (81)
-	-	2928	υC_4H_{17} (76)
-	2915	2916	υC_3H_{15} (85)
-	-	2902	υC_4H_{18} (80)
2606	-	2896	$\upsilon C_{17}H_4$ (82)
2115	2778	2870	υC_6H_{21} (95)
-	-	1759	υO_8C_1 (85)
-	-	1598	$\beta H_{24}N_7H_{23}$ (55) + $\tau H_{24}N_7H_{27}C_6$ (43)
1575	-	1595	$\beta H_{14}N_{12}H_{13}$ (76) + $\tau H_{13}N_{12}C_2C_3$ (21)
1504	1582	1591	$\beta H_{26}O_{25}H_{27}$ (87)
-	-	1471	$\beta H_{17}C_4H_{18}$ (90)
-	-	1455	$\beta H_{15}C_3H_{16}$ (14) + $\beta H_{17}C_4H_{18}$ (78)
-	-	1444	$\beta H_{17}C_4H_{18}$ (80)
1407	1430	1435	$\beta H_{19}C_5H_{20}$ (10)
-	-	1383	$\beta H_{21}C_6H_{27}$ (22) + $\tau H_{24}N_7H_{27}C_6$ (11) +
-	-	1374	$\tau H_{22}C_6H_{27}N_7$ (21)
1347	-	1351	$\beta H_{13}N_{12}C_2$ (17) + $\tau H_{11}C_2C_1O_9$ (26)
-	-	1328	$\tau H_{18}C_4C_3C_2$ (24)
-	-	1303	$\beta H_{15}C_3C_4$ (20) + $\tau H_{11}C_2C_1O_9$ (18)
-	1299	1299	$\beta H_{15}C_3C_4$ (14) + $\beta H_{19}C_5C_6$ (17)
-	-	1294	+ $\tau H_{15}C_3C_2N_{12}$ (11)
-	-	1281	υC_1C_2 (17) + $\beta H_{10}O_9C_1$ (21) + $\beta H_{17}C_4C_5$ (17)
-	-	1273	+ $\tau H_{21}C_6H_{27}N_7$ (10)
-	-	1225	$\beta H_{10}O_9C_1$ (10) + $\beta H_{17}C_4C_5$ (24) +
-	-	1212	$\tau H_{18}C_4C_3C_2$ (11) + $\tau H_{21}C_6H_{27}N_7$ (12)
-	-	1192	$\beta H_{11}C_2C_1$ (15) + $\tau H_{11}C_2C_1O_9$ (10) +
1157	-	1134	$\tau H_{20}C_5C_4C_3$ (20)
-	-	1121	$\beta H_{21}C_6C_{27}$ (23) + $\tau H_{16}C_3C_2N_{12}$ (23)
-	1060	1065	$\beta H_{11}C_2C_1$ (40)
-	-	1047	$\beta H_{13}N_{12}C_2$ (10) + $\beta H_{19}C_5C_6$ (24) + $\beta H_{17}C_4C_5$
-	-	1036	(23) + $\beta H_{21}C_6H_{27}$ (11) + $\tau H_{15}C_3C_2N_{12}$ (10)
-	-	1021	$\beta H_{10}O_9C_1$ (18) + $\beta H_{13}N_{12}C_2$ (11) +
1000	-	999	$\tau H_{18}C_4C_3C_2$ (10)
-	-	971	υC_1C_2 (20) + $\beta H_{10}O_9C_1$ (21)
-	-	930	$\tau H_{21}N_7H27C_6$ (16) + $\tau H_{22}C_6H_{27}N_7$ (12)
-	-	901	$\beta H_{13}N_7H_{22}$ (21)
856	-	858	υC_2C_3 (52)
-	-	816	υC_3C_4 (78)
-	-	789	υC_4C_5 (68)
-	-	767	υC_4C_5 (47)
-	-	744	υC_2C_3 (12) + υC_5C_6 (19) + $\tau H_{15}C_3C_2N_{12}$ (10)
-	-	738	υC_5C_6 (43)
712	-	722	$\upsilon N_{12}C_2$ (14) + $\tau H_{24}N_7H_{27}C_6$ (18)
-	-	609	υC_2C_3 (13) + $\beta H_{24}N_7H_{23}$ (12) + $\tau H_{24}N_7H_{27}C_6$
-	-	571	(17)
557	-	535	υO_9C_1 (52)
-	-	465	$\upsilon N_{12}C_2$ (22) + $\beta H_{14}N_{12}H_{13}$ (11)
-	-	426	$\beta H_{19}C_5C_6$ (12) + $\tau H_{15}C_3C_2N_{12}$ (21)
-	-	365	$\gamma O_8C_2O_9C_1$ (53)
-	-	357	$\beta H_{23}N_7H_{27}$ (13) + $\gamma O_{25}C_6N_7H_{27}$ (52)
-	-	299	$\beta H_{17}C_4C_5$ (14) + $\tau H_{21}C_6H_{27}N_7$ (51)
-	-	291	$\beta O_8C_1O_9$ (57)
-	-	235	$\tau H_{10}O_9C_1C_2$ (86)
-	-	222	$\beta N_{12}C_2C_3$ (12)
-	-	204	$\beta H_{27}C_6C_5$ (29)
-	-	174	$\beta H_{27}C_6C_5$ (53)
-	-	135	$\beta C_3C_4C_5$ (35)
-	-	132	$\beta N_{12}C_2C_3$ (28)
-	-	102	$\beta H_{23}N_7H_{27}$ (28) + $\beta C_2C_1O_9$ (35)
-	82	78	$\beta H_{23}N_7H_{27}$ (28) + $\gamma C_1C_3N_{12}C_2$ (53)

Table 2 (*continued*)

Wave numbers Experimental values		Scaled values	Assignments with PED (>10%)
FT-IR	FT-RAMAN		
-	-	72	υN_7H_{27} (12) + $\beta C_2C_3C_4$ (16) + $\tau H_{14}N_{12}C_2C_3$
-	-	62	(17)
-	-	42	υN_7H_{27} (10) + $\tau H_{14}N_{12}C_2C_3$ (45) +
-	-	28	$\gamma C_1C_3N_{12}C_2$ (17)
-	-	17	υN_7H_{27} (53)
			$\beta C_1C_2N_{12}$ (12)
			$\tau N_{12}C_2C_3C_4$ (14) + $\tau C_3C_4C_5C_6$ (33)
			$\tau c_2C_3C_4C_5$ (62)
			$\tau H_{26}O_{25}H_{27}C_6$ (74)
			$\beta C_1C_2N_{12}$ (16) + $\beta C_2C_3C_4$ (10) + $\tau C_5C_6H_{27}N_7$
			(14)
			$\beta H_{27}N_7C_6$ (14) + $\beta H_{27}C_6C_5$ (13) + $Bc_2C_3C_4$
			(12)
			$\tau O_9C_1C_2C_3$ (17) + $\tau C_3C_4C_5C_6$ (11) +
			$\tau C_2C_3C_4C_5$ (17)
			$\beta H_{27}N_7C_6$ (11) + $\tau O_9C_1C_2C_3$ (55)
			$\beta H_{23}N_7C_6$ (31) + $\tau C_2C_3C_4C_5$ (36)
			$\beta H_{27}C_6C_5$ (13) + $\tau C_5C_6H_{27}N_7$ (14) +
			$\tau N_{12}C_2C_3C_4$ (17) + $Tc_3C_4C_5C_6$ (20) +
			$\tau C_2C_3C_4C_5$ (17)

ν-stretching, β – in plane bending, γ -out of plane bending, τ -torsion.

In the present compound, between lone pair orbital and anti-bonding orbital intramolecular hyper conjugative interaction is produced. The hyper conjugative interactions between $LP(1)O_9 \rightarrow [C_1-O_8(\sigma^*)]$ and $LP(1)N_{12} \rightarrow [C_2-C_3(\sigma^*)]$ produce stabilization energies of 6.78 and 8.59 K cal mol-1, respectively, with electron density decreases of 0.0216e. $LP(1) N7 \rightarrow [C_{25}-H_{27}(\sigma^*)]$ is stabilized by an intramolecular delocalization of electron interaction within the lone pair of nitrogen and the molecule's C—H antibonding orbital. This interaction results in stabilization energies of 13.23 K cal mol^{-1}. This implies that intramolecular hydrogen bonding stabilizes the system.

4.3. Mulliken atomic charges

Mulliken population analysis is the simplest and fastest technique to compute Mulliken atomic charges theoretically. These atomic charges may be utilized to describe a molecule's electronic charge distribution. They are specifically handy for qualitatively estimating partial atomic charges [29]. To determine the net electrical charges density of the L-lysine molecule, Mulliken population analysis was performed. The DFT method was used to calculate the mulliken population, using B3LYP/6–311++ as the fundamental level. The variations in individual charge and charge due to electron cloud movement are identified using natural and Mulliken charge analysis. The title molecule's calculated Mulliken atomic charge values are shown in Supplementary Table 3. The analysis of the Mulliken atomic charges is represented graphically in Supplementary Figure 1.

The result of the calculation shows that all hydrogen atoms in the title compound is positively charged, while negative charge between carbon, nitrogen and oxygen atoms has been delocalized. The atoms in the L-lysine molecule with the smallest negative charges are those that form the hydrogen bonds. Out of entire hydrogen atoms in the title compound, H_{27} atom has the highest positive charge of value 0.3480 which occurs by the existence of nearby oxygen atom which is more electronegative and H_{18} having less positive charge. The result also shows that oxygen atom has negative charges which are donor atoms, with O_{25} (−0.9627) being the most negative of all oxygen atoms. The L-lysine molecule contains only two negatively charged N_7 (−0.8634) and N_{12} (−0.8390) nitrogen atoms. The theoretical Mulliken atomic charge of N_7 is found to be −0.8634 and the calculated Mulliken atomic charge of H_{27} is 0.3480 shows opposite negativity
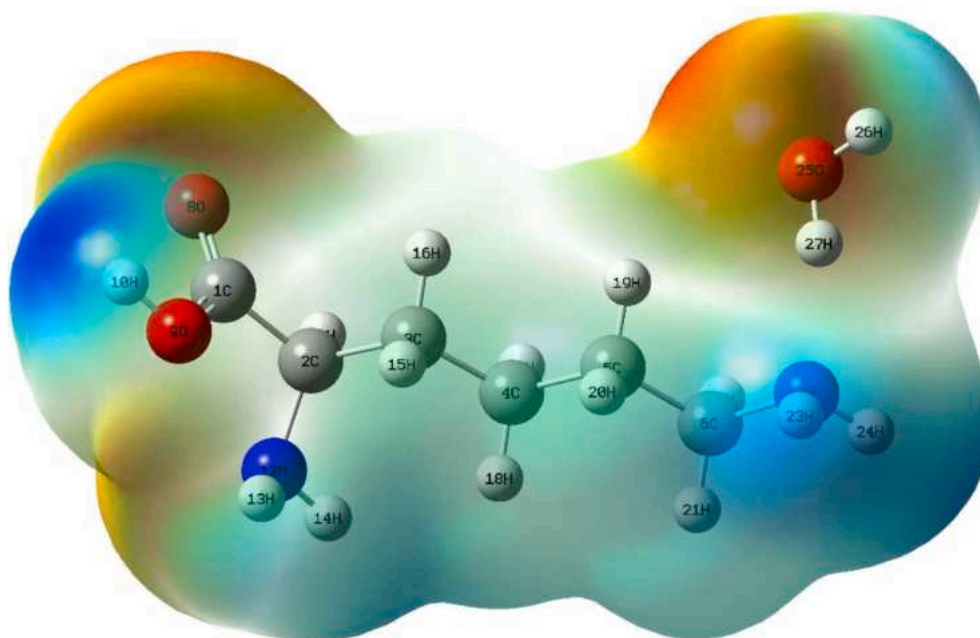
Fig. 6. MEP of ʟ-lysine monohydrate.

Table 3
The title compound's target proteins and its molecular docking activity.

Protein (PDB: ID)	Binding residues	Bond distances (Å)	Binding energy (kcal/mol)
6EUO	SER126	3.23	−5.4
	ASP125	3.05	
	THR129	2.99	
	THR129	2.92	
	THR129	3.16	
	ARG150	2.84	
	THR149	2.92	
7e58	SER171	2.88	−4.5
	ARG310	3.24	
	ARG310	2.80	
2QVD	CYS29	3.08	−4.9
	GLY26	2.84	
	GLY26	2.99	
	TYR120	3.19	
	TYR120	3.47	
	TYR120	3.03	
	PRO121	3.15	

4.4. HOMO – LUMO

The Molecular Orbital with the Highest Occupancy (HOMO) and low molecular orbital (LUMO) are the molecules' two most important orbitals. HOMO has the capability to give electrons, while LUMO has the capability to accept electrons [30]. The molecule's ionization potential and affinity of an electron of ʟ-lysine monohydrate are described by the notations HOMO and LUMO [31]. The HOMO-LUMO energies of ʟ-Lysine molecules were discovered to be −6.8923 eV and −0.63674 eV, respectively. Molecules with a large gap in energy are stable and don't react chemically, while those with a small gap do [32]. Fig. 2 depicts the graphical representations of HOMO and LUMO. The molecule's chemical stability is confirmed by the larger energy gap value (6.25556 eV). The B3LYP6–31 G (d, p) technique is applied to compute the ionization potential, affinity of electron, electronegativity, chemical potential, chemical hardness, chemical softness, and electrophilic index for the substance ʟ-Lysine. The properties listed above are in supplementary Table 4.

An affinity and the ability to undertake charge transfer operations is known as an electron acceptor. It takes electrons, as indicated by its computed electron affinity value of 0.63674 eV. Chemical hardness and electronegativity describe a molecule's response to a swap in the number of electrons at a constant potential and provide information about its reactivity, chemical binding, selectivity and stability [33,34].

The molecule's calculated electronegativity value is 3.76452 eV, which is high and indicates that it will be less nucleophilic. The chemical resistance of ʟ-lysine is very high (3.12778 eV), its large energy gap indicates lesser chemical reactivity and higher chemical stability. The electrophilicity index of 2.26544 eV calculated indicates that electrons are promoted from the nucleophile to the electrophile, increasing the molecule's chemical reactivity. The high electrophilicity index of the compound in comparison to its low chemical potential confirms its electrophilic nature (−3.76452 eV).

4.5. Topology analyses

4.5.1. AIM analysis

The Bader theory of "Atoms in Molecules" is one of the most efficient methods for evaluating the characteristics of an atom. (AIM) topological analysis is used to calculate the electron or molecule's charge density within the actual spatial boundaries of the atoms making up the system [35]. The basis for this strategy is determining critical points of the bond (BCP) [36]. BCP is characterized in terms of observable properties of topology including the charge density (ρ) and Laplacian field ($\Delta^2\rho(r)$) [37]. The ρ value of a bond describes its strength in which higher the ρ value, the stronger the related bond. The bonds are defined by the value $\Delta^2\rho(r)$. A covalent bond is identified when $\Delta^2\rho(r)$ is less than zero and a closed shell interaction is identified when $\Delta^2\rho(r)$ is greater than zero. The computed electron density and Laplacian values are shown in the below Table 6. By looking at the negative values of $\Delta^2\rho(r)$ it is clear that the interaction between the atoms of ʟ-lysine is covalent bond interaction. Supplementary Figure 2 illustrates the compound's molecular structure using an AIM analysis. The calculated ellipticity of the molecule is listed in supplementary Table 5.

As shown in the Supplementary figure 3, there are possibilities for hydrogen bond interaction between C—H…O and N—H…O in the molecule. From the previously reported data [38–40], a hydrogen bond is present when the electron density at the critical point of the bond is >0.002 a.u and the Laplacian of the electron density is >0.0004 a.u. As
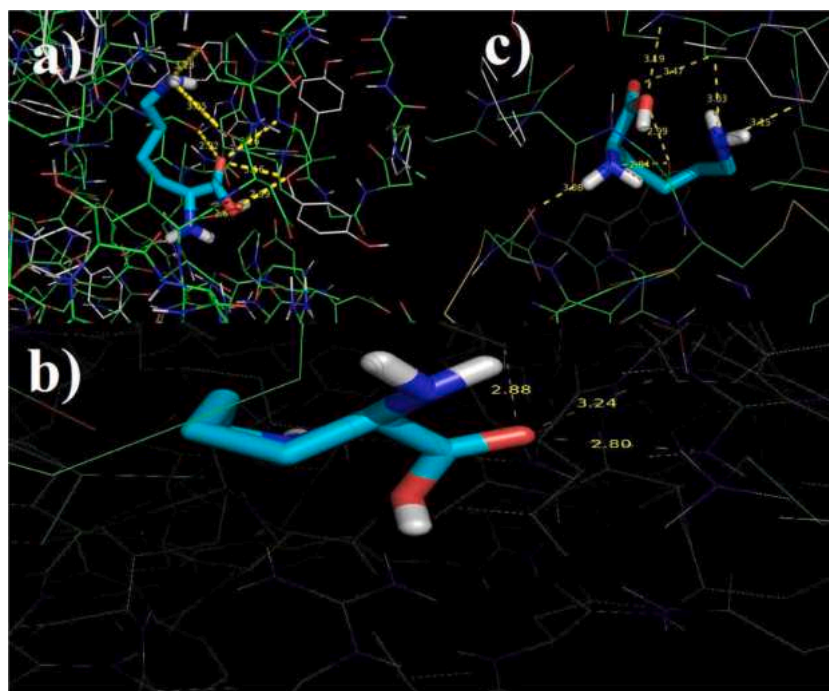
Fig. 7. (a) Docking of ʟ-Lysine Monohydrate with antioxidant protein, (b) Docking of ʟ-Lysine Monohydrate with antiviral protein, (c) Docking of ʟ-Lysine Monohydrate with Anti-inflammatory protein.

per the above data the electron densities of the critical points were found to be 0.0536 and 0.0322 a. u and the Laplacian electron densities were obtained as 0.0204 and 0.0885 a.u. The hydrogen bond energies are calculated using the relation proposed by Espinosa et al. [41], $E_{HB} = V$ $(r_b)/2$ between C—H…O is 4.4239 kcal/mol which shows there is weak interaction and N—H…O is 30.705 kcal/mol which shows the strong hydrogen bond interaction.

4.5.2. Electron localization function (ELF) and localized orbital locator (LOL)

The electron localization function and localized orbital locator which are widely known as ELF and LOL are used in covalent bond studies as they represent the region of molecular space where the most electron pairs can be found. According to the kinetic energy density, both the analysis has identical chemical compositions. ELF describes the electron pair density, on the other hand LOL describes the orbital gradient-induced maximum overlap of localized orbitals [42]. The ELF map is formed between 0.0 and 1.0; however, in sites where electron localization dominates electron density, the LOL reaches a significant value greater than 0.5, and the region below 0.5 reveals delocalized electronic regions [43].

The ELF and LOL colour shade maps are depicted in supplementary Figure. 3 (a, b) which clearly shows the presence of bonded and nonbonded electrons, respectively. The bonded and nonbonded localization of electrons in hydrogen atoms are shown in red, while the electron cloud delocalization is shown in blue colour around a few carbon and nitrogen atoms.

4.5.3. Non-covalent interaction (NCI) analysis

Reduced Density Gradient (RDG) is another name for the NCI interaction. Weak interactions such as Vander Waals, hydrogen bonds, and steric effects are detected using RDG, a topological technique. The interactions in the region and their illustration are obtained by the NCI analysis that is based upon the density of the electron and their derivatives [44]. The RDG function can be written as

$$RDG = \frac{1}{2(3\pi x^2)^{\frac{1}{3}}} \frac{\nabla\rho(r)}{\rho(r)^{\frac{4}{3}}}$$

Where $\rho(r)$ id the electron density [45]

There are three distinct interaction zones in our molecule, illustrated by blue, green, and red dots, all of which are low density and have small RDG gaps. The red colour indicates steric repulsion between atoms in a molecule, when $(\lambda 2) \rho > 0$. Green colour represents the van derwaals interaction in the region when $(\lambda 2) \rho = 0$. Strong electrostatic interactions such as hydrogen bonding can be seen in the blue spikes in the sign $(\lambda 2) \rho < 0$ area [46]. Large negative sign $(\lambda 2) \rho$ values in ʟ-lysine imply a more attractive and binding interaction. The red coloured interaction region indicates the repulsion as well as non-bonding interaction in the range (0.01–0.02) a.u. increased contact between hydrogen atoms. The Fig. 3(a, b) clearly shows that the ʟ-lysine molecule has greater hydrogen bonding interaction.

4.6. Vibrational analysis

Vibrational modes are identified using calculated vibrational wave numbers and atomic displacements corresponding to the various normal modes. Based on potential distribution analysis, the VEDA4 tool was used to assign vibrational modes [47]. In the solid state, in the region of 4000–400 cm^{-1} FTIR spectra and FT-Raman spectra in the region of 3500–250 cm^{-1} have been observed. The ʟ-lysine molecule contains 75 distinct typical vibrational modes and is made up of 27 atoms. Figs. 4 and 5 shows the computed and observed FT-IR and FT-Raman spectra. Table 2 depicts the distribution of various vibration modes.

4.6.1. N-H vibrations

In the FTIR spectrum, N—H bands have a lower population than Carbon-Hydrogen or Carbon-Carbon bands. Strong N—H bonds cause these bands to takes place at higher frequencies. The N—H stretching vibrations are most common in the 3500–3000 cm^{-1} range [48]. The infrared spectrum of the ʟ-lysine monohydrate molecule shows weak N—H stretching vibration at 3445 cm^{-1} and the stretching mode of N—H vibration is obtained at 3443 cm^{-1} with a theoretical PED

contribution of 97 %. Intramolecular charge transfers between the amino and carbonyl groups causes the minor difference between the theoretical and experimental value. The bending mode of N—H vibration is spotted at 1575 cm^{-1} (medium peak) in FTIR and 1595 cm^{-1} is the calculated value. A very weak peak is obtained at 1060 cm^{-1} in the Raman spectra with the comparable scaled value 1065 cm^{-1}. The C—N vibrations overlap with this vibration. The N—H in plane bending modes of L-lysine monohydrate are ascribed to the weak vibrational band which are found at 557 cm^{-1} in the IR spectrum and 82 in the IR-Raman spectrum; 535 cm^{-1} and 78 cm^{-1} are the calculated values.

4.6.2. C—H vibrations

Below 3000 cm^{-1}, there are noticeable peaks that are related with the aliphatic stretching mode vibrations of C—H. The aliphatic stretching vibrational mode for the C—H bond in the FTIR spectrum of the title compound, were determined to be between 2975 cm^{-1} and 2870 cm^{-1}, which corresponds to the values of experimental spectrum of 3157 cm^{-1}, 2965 cm^{-1}, 2606 cm^{-1} and 2115 cm^{-1}. The intramolecular C—H…N delocalization of electron interaction causes the experimental wave number to differ from the expected range. The in plane bending vibration modes of CH bonds are spotted in the present molecule at 1430 cm^{-1} and 1299 cm^{-1} in the Raman spectra and 1407 cm^{-1} and 1157 cm^{-1} in the IR spectrum with scaled DFT values of 1435 cm^{-1}, 1299 cm^{-1} and 1134 cm^{-1}. The strong band seen at 1347 cm^{-1} in the IR spectrum is designated to torsion mode. Due to the skeleton's highly electronegative nitro group, the wavenumbers obtained are blue-shifted from the anticipated value

4.6.3. C—C vibrations

The composition of the substituents has no effect on the C—C stretching vibrations, which are projected to be in the range of 1650 to 1100 cm^{-1} [49]. As a result, in the current analysis, the FT-Raman band at 1299 cm^{-1} and the FTIR band at 1157 cm^{-1} and 1000 cm^{-1} are allocated to the C—C stretching vibrations. In the FT-Raman spectra, the βCCH vibration is obtained at 712 cm^{-1}. The deviation in the wavenumbers may be due to the presence of nitrogen, which is an electronegative atom and hydrogen, which is an electropositive atom throughout the skeleton structure.

4.7. Molecular electrostatic potential (MEP)

MEP (Molecular Electrostatic Potential) has been widely applied in the investigations of biological identifications and hydrogen bonding interactions, as well as for predicting physiochemical properties and visualizing molecule size and shape [50]. The sites of relative reactivity in a species for nucleophilic and electrophilic attack are predicted using the Molecular electrostatic potential (MEP). Different hues represent different electrical potential values; the regions with red and green colour represent the lowest electric field potential, the region with blue colour represents the highest electric field potential [51]. The MEP map is constructed from the optimized geometry of the molecule using 6–311 ++ G (d, p) and is shown in Fig. 6. The compound's colour code ranges from −5.880e-2 to +5.880e-2. The colour red indicates that this compound surrounding the oxygen atoms are more negative. The nucleophilic area in blue represents the rest of the molecule.

4.8. Fukui functional analysis

Fukui functions are indices that tell you how likely a molecule is to donate or accept an electron, allowing you to forecast which atom in the molecule is most vulnerable to a nucleophilic or electrophilic attack. As a compound receives an electron, the Fukui function is $f+(r)$ that is also known as the index of nucleophilic assault. The Fukui function, commonly known as the index of electrophilic attack, is used when a molecule is prone to drop out an electron [52]. Condensed Fukui functions are used to quantify the Fukui function or to assign a numerical

value to each atom in a chemical system to indicate whether it can operate as a reactive site. For $f + r$ and f-r, the condensed Fukui function is written as:

$$f^+(r) = q_r (N + 1)\text{-}q_r(N)$$
$$f^-(r) = q_r (N\text{-}1)\text{-}q_r(N) \quad f^0(r) = (q_r (N + 1) \text{-}q_r (N\text{-}1))/2.$$

Morrel et al.[53] developed a new nucleophilicity and electrophilicity descriptor, f(r). It has the following definition:

$$\Delta f = f^+(r) - f^-(r)$$

The site is electrophilic if $\Delta f > 0$, and nucleophilic if $\Delta f < 0$. This is how the reactive site's nature may be determined using the dual descriptor. The reactive nature with its sites is indicated in supplementary figure 4. Supplementary Table 6 lists Fukui functions, local softness values, and local indices of electrophilicity and derivatives.

The order of susceptibility to nucleophilic attack for L-lysine monohydrate is H27>H13>H24>H20>H26>H15>H17>O25>H23>H16 >H10>H11>H18>O8>H14>H22>N12>H19>C4>N7, whereas susceptibility to electrophilic attack is C2>C5>O9>C3>C6>C1>H21.

4.9. Molecular docking

A useful technique to investigate the characteristics of how proteins interact with their ligands and to verify biological activity within any chemical structure is molecular docking study [54]. The ultimate aim of molecular docking simulation is to predict the optimal ligand-macromolecular partner binding arrangement. Within the protein binding site, it generates a variety of poses or ligand conformations. The ligands with the best posture were chosen based on which conformation has the lowest free binding energy [55]. The ability of a small molecule to determine its binding conformation to the exact target binding has made it one of the most widely used methods in structure-based drug discovery. The docking investigation was carried out using Auto-dock vina software. The target proteins PDB: 6EUO (Antioxidant), PDB: 7e58 (Antiviral) and PDB: 2QVD (Anti-inflammatory) were downloaded from RCSB protein data bank. For the docking purpose the downloaded protein was prepped by eliminating co-crystallized ligands and water molecules. The receptor was also prepared. Table 3 displays the docking parameters of L-Lysine monohydrate. The protein-Ligand interactions are represented in Fig. 7(a,b,c)

The protein is docked with the ligand L-lysine monohydrate. The molecule doped deeply with antioxidant protein creating seven bonds with amino acids SER126, ASP125, THR129, ARG150 and −5.4 kcal/mol is obtained as its binding affinity. The residues SER126 and ASP125 forms bond with values of 3.23 Å and 3.05 Å respectively from the ligand atom N$_7$. The residues THR129 and ARG150 with 2.99 Å and 2.84 Å are bonded with O$_9$. The oxygen atoms engaged in three bonds with residues THR129, THR129 and THR129 with bond distance 2.92 Å, 3.16 Å and 2.92 Å respectively. From the obtained result the L-lysine molecule shows good protein ligand interaction with binding score −5.4 kcal/mol and possess good antioxidant activity.

Secondly the molecule is docked with antiviral protein. The active sites of antiviral protein-PDB: 7e58 were determined as SER171, ARG310, ARG310. The binding affinity of the antiviral protein with ligand is −4.5 kcal/mol. The O$_8$ atom possess three bonds SER171, ARG310, ARG310 with bond distances 2.88 Å, 3.24 Å and 2.80 Å respectively. From the above result, the L-lysine molecule has a significant protein-ligand interaction with good binding score against 7e58 protein, indicating antiviral protein. Thirdly anti-inflammatory protein PDB: 2QVD was docked with the prepared ligand. It didn't form any hydrogen bonds in the active site, but it did for hydrophobic bonds with amino acids like CYS29, GLY26, GLY26, TYR120, TYR120, TYR120 and PRO121 as 3.08 Å, 2.84 Å, 2.99 Å, 3.19 Å, 3.47 Å, 3.03 Å and 3.15 Å respectively. The binding exponent appears to be −4.9 kcal/mol. It is clear from the figure that hydrogen atoms form hydrogen bonds with ligands and nitrogen or oxygen atoms in the protein. The MEP plot supports the hypothesis that positively charged hydrogen atoms attack a target protein's region that is rich in electrons.

The results of the molecular docking study indicated that all proteins displayed interactions with the ʟ-lysine monohydrate molecule that were sufficiently favourable. In these three proteins, none of the residues interacted with the monohydrate of the title molecule. The binding affinity for antioxidant protein is substantially higher (-5.4 kcal/mol) than for other activity proteins. This binding has a higher negative binding energy, resulting in greater stability, and does not significantly alter the active site's structure. The higher negative binding energy value suggests that it can host the target protein 6EUO more effectively than the proteins 7e58 and 2QVD. In summation these findings suggest that the title molecule makes a significant contribution to the management antioxidant activity.

5. Conclusion

The current study reports on the experimental and theoretical spectroscopic investigation of ʟ-lysine monohydrate using FT-IR and FT-Raman analyses. High level quantum chemistry computation was used to optimize the structure of ʟ-Lysine monohydrate. NBO analysis was done to find out how the title compound interacts with its donors and acceptors and how stable it is. The energy gap $E = 6.25556$ eV obtained by FMO analysis indicates that the title molecule is chemically stable. Using NPA atomic charges, the Fukui function was done to estimate the sites of chemical reactivity and topological analyses based on AIM, ELF and LOL were reported using the Multiwfn analyser. The NCI method revealed the title compound's strong hydrogen bonding, Vander Waals interaction and steric repulsion. ʟ-lysine molecule has considerable antioxidant activity against the PDB (6EUO), with greater binding affinity energy of -5.4 kcal/mol, indicating that the above compound may have antioxidant efficacy. These findings revealed that the title compound is the greatest and best for all features, which aids future study and innovative identifiers.

Funding

CRediT authorship contribution statement

T. Brintha: Conceptualization, Methodology, Writing – original draft. **J Jeni James:** Resources, Writing – review & editing. **M Amalanathan:** Software, Validation. **P.J. Jegan Babu:** Supervision, Investigation, Data curation. **M. Sony Michael Mary:** Formal analysis, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

Acknowledgement

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chphi.2023.100311.

References

[1] N.J. Temple, Antioxidants and disease: more questions than answers, Nutr, Res (2000) 449–459.
[2] J.R. Wagner, C. Hu, B.N. Ames, Endogenous oxidative damage of deoxycytidine in DNA, Proc. Natl. Acad. Sci, USA 89 (1992) 3380–3384.
[3] B.N. Ames, Identifying environmental chemicals causing mutations and cancer, Science 204 (1979) 587–593.
[4] C.G. Fraga, P.A. Motchnik, M.K. Shigenaga, H.J. Helbock, R.A. Jacob, B.N. Ames, Ascorbic acid protects against endogenous oxidative DNA damage in human sperm, Proc. Natl. Acad. Sci. USA 88 (1991) 11003–11006.
[5] S.J. Duthie, A. Ma, M.A. Ross, A.R. Collins, Antioxidant supplementation decreases oxidative DNA damage in human lymphocytes, Cancer Res 56 (1996) 1291–1295.
[6] S. Budavari, An Encyclopedia of Chemicals, Drugs, and Biologicals, 11th edition, The Merck Index, 1989.
[7] J.S. Ballantyne, Amino acid metabolism, Fish physiol. 20 (2001) 77–107.
[8] M. Aydın, Z. Kartal, S. Osmanoğlu, M.H. Başkan, R. Topkaya, EPR and FT-IR spectroscopic studies of ʟ-lysine monohydrochloride and ʟ-glutamic acid hydrochloride powders, J. Mol. Struct. 994 (2011) 150–154, https://doi.org/10.1016/j.molstruc.2011.03.010.
[9] L.C.R. Agnew, Dorland's Illustrated Medical Dictionary, 24th ed, WB Saunders, Company, 1967.
[10] M. Singh, D.M. Rao, S. Pande, S. Battu, K. Mahalakshmi, K.R. Dutt, M. Ramesh, Medicinal uses of ʟ-lysine: past and future, Int. J. Res. Pharm. Sci. 2 (2011) 637–642.
[11] N. Rani, N. Vijayan, M.A. Wahab, G. Bhagavannarayana, B. Riscob, K.K. Maurya, Growth and characterization analyses of pure and p-nitroaniline doped ʟ-lysine monohydrochloride single crystal for nonlinear optical applications, Optik 124 (2013) 1550–1554.
[12] O. Zelder, C. Klopprogge, H. Schöner, S. Häfner, B. Kröger, P. Kiefer, E. Heinzle, WO05059139A2 (2005).
[13] J.W. Oh, S.J. Kim, Y.J. Cho, N.H. Park, J.H. Lee, US5268293 (1993).
[14] Y. Zhang, Z. Guo, X.Z. You, Hydrolysis theory for cisplatin and its analogues based on density functional studies, J. Am. Chem. Soc. 123 (2001) 9378–9387, https://doi.org/10.1021/ja0023938.
[15] Y.B. Alpaslan, N. Süleymanoğlu, E. Öztekin, F. Erşahin, E. Ağar, S. Işık, Experimental and semi-empirical and DFT calculational studies on (E)-2- [(2, 4-Dichlorophenylimino) methyl]-p-cresol, J. Chem. Crystallogr. 40 (2010) 950–956, https://doi.org/10.1007/s10870-010-9769-8.
[16] R. Zhang, B. Du, G. Sun, Y. Sun, Experimental and theoretical studies on o-, m-and p-chlorobenzylideneaminoantipyrines, Spectrochimica. Acta. Part A, Mol. Biomol. Spectroscopy 75 (2010) 1115–1124, https://doi.org/10.1016/j.saa.2009.12.067.
[17] J. Devillers, Comparative QSAR, in:, CRC Press, 1998, pp. 200–201.
[18] Y.P. Rachelin, L.P. Nair, C. James, Electronic structure investigations and spectroscopic studies on the herbicidal molecule 4-Nitro phenyl-phenylether, J. Mol. Struct. 1036 (2013) 56–62.
[19] S.H. Haci Özis, ik Bayari, Semran, Vibrational analysis of Triclosan, American Inst. Phy. 1203 (2010) 1227.
[20] C.S. Abraham, S. Muthu, J.C. Prasana, S.J. Armaković, S. Armaković, B.G. AS, Spectroscopic profiling (FT-IR, FT-Raman, NMR and UV–Vis), autoxidation mechanism (H-BDE) and molecular docking investigation of 3-(4-chlorophenyl)-N, N-dimethyl-3-pyridin-2-ylpropan-1-amine by DFT/TD-DFT and molecular dynamics: a potential SSRI drug, Comput. Biol Chem 77 (2018) (2018) 131–145, https://doi.org/10.1016/j.compbiolchem.2018.08.010.
[21] S. Premkumar, T.N. Rekha, R.M. Asath, T. Mathavan, A.M.F. Benial, Vibrational spectroscopic, molecular docking and density functional theory studies on 2-acetylamino-5-bromo-6-methylpyridine, Eur. J. Pharm. Sci. 82 (2016) (2016) 115–125, https://doi.org/10.1016/j.ejps.2015.11.018.
[22] T. Lu, F. Chen, Multiwfn: A multifunctional wavefunction analyzer, J. Comput. Chem. 33 (2012) 580–592.
[23] E.D. Glendening, C.R. Landis, F. Weinhold, Natural bond orbital methods, Wiley interdisciplinary reviews, comput. mol. sci. 2 (2012) 1–42, https://doi.org/10.1002/wcms.51.
[24] K. Bamba, O.W. Patrice, N. Ziao, NBO population analysis and electronic calculation of four azopyridine ruthenium complexes by DFT method, Comput. Chem. 5 (2016) 51–64, https://doi.org/10.4236/cc.2017.51005.
[25] M.C. Murry, E. John, Organic Chemistry, 3rd ed, Wadsworth, Belmont.
[26] J.N. Liu, Z.R. Chen, S.F. Yuan, Study on the prediction of visible absorption maxima of azobenzene compounds, J. Zhejiang Univ. Sci. 6 (2005) 584–589.
[27] K. Singh, R. Kataria, Crystal structure, Hirshfeld surface and DFT based NBO, NLO, ECT and MEP of benzothiazole based hydrazone, Chem. Phys. 538 (2020) 110873, https://doi.org/10.1016/j.chemphys.2020.110873.
[28] S.S. Margreat, S. Ramalingam, S. Sebastian, S. Xavier, S. Periandy, J.C. Daniel, M. M. Julie, DFT, spectroscopic, DSC/TGA, electronic and molecular docking investigation of 2, 5-thiophenedicarboxylic acid: a promising anticancer agent, J. Mol. Struct. 1200 (2020) 127099.
[29] R. Jalilian, L.A. Jauregui, G. Lopez, J. Tian, C. Roecker, M.M. Yazdanpanah, Y. P. Chen, Scanning gate microscopy on graphene: charge inhomogeneity and extrinsic doping, Nanotechnology 22 (2011), 295705, https://doi.org/10.1088/0957-4484/22/29/295705.
[30] H.G. O.BeckerJan fleming. Frontier Orbitals and Organic Chemical Reactions, John wiley Sons LTD, london/New York/Syndney/Toronto, 1976.
[31] S. Renuga, M. Karthikesan, S. Muthu, Raman spectra, electronic spectra and normal coordinate analysis of N, N-dimethyl-3-phenyl-3-pyridin-2-yl-propan-1-amine by DFT method, Spectrochim. Acta Part A 127 (2014) 439–453, https://doi.org/10.1016/j.saa.2014.02.068.

[32] J. I.Aihara, Why are some polycyclic aromatic hydrocarbons extremely reactive, Phys. Chem. Chem. Phys. 1 (1999) 3193–3197, https://doi.org/10.1039/A902032B.

[33] P.K. Chattaraj, P. Fuentealba, P. Jaque, A. Toro-Labbé, Validity of the minimum polarizability principle in molecular vibrations and internal rotations: an ab initio SCF study, J. Phys. Chem. A 103 (1999) 9307–9312, https://doi.org/10.1021/jp9918656.

[34] L. Pogliani, Model with dual indices and complete graphs. The heterogeneous description of the dipole moments and polarizabilities, New J. Chem. 27 (2003) 919–927, https://doi.org/10.1039/B210474C.

[35] C. Anzline, P. Sivakumar, S. Israel, K. Sujatha, Comprehensive study on the topological properties of 5-Amino-2-Methyl Benzene Sulfonamide involving inter and intra molecular hydrogen bonds, J. Mol. Struct 1201 (2020) 127208, https://doi.org/10.1016/j.molstruc.2019.127208.

[36] L. Del Olmo, C. Morera-Boado, R. López, J. M.Garcia de la Vega, Electron density analysis of 1-butyl-3-methylimidazolium chloride ionic liquid, J. Mol. Model. 20 (2014) 1–10.

[37] R.F.W.BaderA Quantum Theory, Clarendon, Oxford, 1990.

[38] I. Majerz, A. Koll, Structural manifestations of proton transfer in complexes of 2, 6-dichlorophenols with pyridines, Acta Crystallogr. B: Struct. Sci. 60 (2004) 406–415.

[39] P.L.A. Popelier, R.F.W. Bader, Effect of twisting a polypeptide on its geometry and electron distribution,J, Phys. Chem. 98 (1994) 4473–4481.

[40] U. Koch, P.L.A. Popelier, Characterization of C-H-O hydrogen bonds on the basis of the charge density, J. Phys. Chem. 99 (1995) 9754–9797.

[41] E. Epsinosa, I. Alkorta, I. Rozas, J. Eiguero, R. Molins, About the evaluation of the local kinetic, potential and total energy densities in closed-shell interactions,Chem, Phys. Lett. 336 (2001) 457–461.

[42] H. Jacobsen, Localized-orbital locator (LOL) profiles of chemical bonding, Can. J. Chem 86 (2008) 695–702.

[43] B. Silvi, A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature 371 (1994) 683–686, https://doi.org/10.1038/371683a0.

[44] G. Saleh, C. Gatti, L.L. Presti, Non-covalent interaction via the reduced density gradient: independent atom model vs experimental multipolar electron densities, Comput. Theoretical Chem 998 (2012) 148–163, https://doi.org/10.1016/j.comptc.2012.07.014.

[45] E. Nemati-Kande, R. Karimian, V. Goodarzi, E. Ghazizadeh, Feasibility of pristine, Al-doped and Ga-doped Boron Nitride nanotubes for detecting SF4 gas: a DFT, NBO and QTAIM investigation, Appl. Surf. Sci 510 (2020) 145490, https://doi.org/10.1016/j.apsusc.2020.145490.

[46] S. Gatfaoui, N. Issaoui, T. Roisnel, H. Marouani, Synthesis, experimental and computational study of a non-centrosymmetric material 3-methylbenzylammonium trioxonitrate, J. Mol. Struct 1225 (2021) 129132, https://doi.org/10.1016/j.molstruc.2020.129132.

[47] M.H. Jamroz, Vibrational energy distribution analysis, VEDA 4 (2004).

[48] B. Smith. Infrared Spectral interpretation: a Systematic Approach, CRC press, 2018.

[49] N. Sundaraganesan, S. Ilakiamani, C. Meganathan, B.D. Joshua, Vibrational spectroscopy investigation using ab initio and density functional theory analysis on the structure of 3-aminobenzotrifluroide, Spectrochim. Acta. A 67 (2007) 214–224, https://doi.org/10.1016/j.saa.2006.07.004.

[50] J.S. Murray, K. Sen. Molecular Electrostatic Potentials, Concepts and Applications, Elsevier, Amsterdam, 1996.

[51] A. Sethi, R. Prakash, Novel synthetic ester of Brassicasterol, DFT investigation including NBO, NLO response, reactivity descriptor and its intramolecular interactions analyzed by AIM theory, J. Mol. Struct. 1083 (2015) 72–81, https://doi.org/10.1016/j.molstruc.2014.11.028.

[52] A.O. Zacharias, A. Varghese, K.B. Akshaya, M.S. Savitha, L. George, DFT, spectroscopic studies, NBO, NLO and Fukui functional analysis of 1-(1-(2, 4-difluorophenyl)-2-(1H-1, 2, 4-triazol-1-yl) ethylidene) thiosemicarbazide, J. Mol. Struct 1158 (2018) 1–13, https://doi.org/10.1016/j.molstruc.2018.01.002.

[53] C. Morell, A. Grand, A. Toro-Labbé, New dual descriptor for chemical reactivity, J. Phys. Chem 109 (2005) 205–212, https://doi.org/10.1021/jp046577a.

[54] F.B. Asif, F.L.A. Khan, S. Muthu, M. Raja, Computational evaluation on molecular structure (Monomer, Dimer), RDG, ELF, electronic (HOMO-LUMO, MEP) properties, and spectroscopic profiling of 8-Quinolinesulfonamide with molecular docking studies, Comput. Theor. Chem. 1198 (2021) 113169.

[55] M. Kandeel, Y. Kitade, Computational analysis of siRNA recognition by the Ago2 PAZ domain and identification of the determinants of RNA-induced gene silencing, PLoS One 8 (2013), 57140, https://doi.org/10.1371/journal.pone.0057140.

Full Length Article

Identification of interaction of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidine-carboxylic acid using DFT studies, molecular docking, biological activity and topology analysis for biological applications

M. Amalanathan [a,e,*], T. Brintha [b,e], S. Sijana [a,e], P.J. Jegan Babu [c,e], M. Sony Michael Mary [d]

[a] Department of Physics & Research Centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai, Tamil Nadu, 629 153, India
[b] Research Scholar, Reg No: 1921311213201, Nesamony Memorial Christian College Marthandam, Kanyakumari, Tamil Nadu, India
[c] Department of Physics & Research centre, Nesamony Memorial Christian College Marthandam, Kanyakumari, Tamil Nadu, India
[d] Department of Physics, Annai Velankanni College of Engineering, Azhagappapuram. District, Kanyakumari, Tamil Nadu 629 401, India
[e] Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu 627 012, India

ARTICLE INFO

ABSTRACT

1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acid and its salts are present in the cells and bodily fluids of many living things, and they play a significant role in biological systems as precursors to pyrimidine nucleosides. The work aims to determine the structural and biological properties of TDPCA by theoretical and experimental spectroscopic investigations. B3LYP/6–311++G(d,p) level computations were used for all theoretical calculations. Functional groups, vibrational modes, and the aromatic nature of TDPCA have all been predicted for the molecule using FT-IR and FT-Raman spectroscopy techniques. The NBO analysis was conducted to comprehend the likely charge transfer interaction that exists in the molecule. The delocalization can be determined by the analysis of HOMO and LUMO. Low energy gap between HOMO and LUMO is shown to be predictive of electron transport and results in bioactivity in the molecule. Analysis of the molecule's topology was done to determine its reactivity. Using molecular docking, the anticancer efficacy of the drug against IA PI 3-kinase inhibitor receptors for protein targets (2WXQ) was investigated.

1. Introduction

(TDPCA) (vitamin B13) and its salts are present in the cells and physiological fluids of many living species and which are important components of biological systems because they act as precursors of pyrimidine nucleosides [1–4]. Different metal complexes of were explored as biostimulators of ionic exchange processes in organisms, which are used in medicine [5,6]. TDPCA is also a popular topic of study in the fields of food safety and nutrition [7–9].Orotic acid's crystal and molecular structures were determined by Takusawaga and Shimada, [10] while Mutikainen [11] looked into the crystal structure of TDPCA metal complexes. A trustworthy assignment of the vibrational spectra of TDPCA and its metal complexes is a useful place to start when analyzing their interactions with other chemical species in the biological environment because of how crucial these compounds are to living systems.

The uracil or thymine molecule is joined to the TDPCA molecule. These kinds of molecules have been examined using vibrational spectroscopy, normal coordinate analysis (NCA), and ab initio calculations [12–14]. The results of these experiments may help to identify the spectra of TDPCA.

It has been established that metal ions, particularly Mg2+, are necessary for these processes, especially during the phosphoribosylation of TDPCA [15]. Metal orotates are frequently employed in medicine [16]. As cancer treatments, palladium, platinum and nickel orotate-complexes have been investigated [17]. Orotates have also been proposed to have a biological carrier role, allowing them to be used to treat syndromes caused by metal ion insufficiency in living creatures, such as magnesium, calcium, zinc, or iron [18–20]. In addition to their crucial biological function, the mono- and dianion of TDPCA (H2L and HL2) are possible polydentate ligands. Coordinated orotate anions can adopt a variety of interligand hydrogen bonding patterns depending on the metal coordination mode since their ligand surface has the ability to form double or triple hydrogen bonds. This enables the construction of lengthy, self-assembling structures [21].

DFT was used to investigate the compound's geometrical optimization, electronic, topological, and biological aspects. The charge transport inside the structure was determined using a HOMO/LUMO study.

The most common type of cancer is aberrant cell division and proliferation, which are brought on by a breakdown in the molecular signals that regulate these activities [22]. With 8.8 million deaths, cancer ranked as the second most common cause of death worldwide in 2015. Cancer is the leading cause of mortality worldwide, accounting for 70 % of cases in low- and middle-income nations [23].

Abnormal cell development that spreads across the entire body of a human is the cause of cancer. There are several forms of cancer, including prostate cancer [24,25], breast cancer [26], and leukaemia [27]. Based on the literature survey, there have been no quantum level studies on TDPCA. These days, an important part of the analysis of structural and molecular interactions is played by Density Functional Theory, which makes use of the principles of quantum mechanical calculations and molecular simulation. Density functional theory has developed into a potent tool in recent years for studying the geometrical shapes of molecules [28]. The most effective technique for studying hybrid materials theoretically is DFT because it strikes a fair balance between calculation time and result precision [29].

In this work, the DFT/B3LYP approach was used to optimize the molecular structure of TDPCA. Theoretical calculations and practical observations were used to determine the molecule's particular vibrational wavenumbers. Using the VEDA 4 program, the potential energy distribution (PED) was calculated to determine the vibrational wavenumbers. Calculations were also performed for the Mulliken and natural atomic charge distributions. The molecular stability and reactivity of the compound were investigated using the Frontier molecular orbitals (FMOs) technique, as stated. A natural bond orbital (NBO) analysis was also performed in order to comprehend the molecule's bonding characteristics. The present work mainly focused to analyze the structural activity and docking ability of the title compound against the anticancer protein.

2. Materials and methods

Sigma Aldrich Chemical Company (USA) provided a powder sample of TDPCA, and this sample was used in the subsequent spectroscopic analysis without any further purification because its purity was certified to be greater than 98 %.Nicolet 6700 FTIR spectrometer records the Fourier transform infrared spectra of TDPCAat a temperature of 302.15 K in the range of 4000 – 400 cm-1. With the help of the Spectrum GX Fourier transform-infrared spectrometer, the FTIR spectrum of the provided substance was captured using the KBr pellet method. Nicolet Magna 750 Raman spectrometer with an InGaAs (Indium Gallium Arsenide) semiconductor detector was used to capture the FT-Raman spectrum of the molecule at a resolution of 4 cm-1 in the 3500-0 cm-1 range. The 1064-nm line of the Neodymium: Yttrium Aluminium Garnet laser was used as the excitation source. At the sample location, the laser output was typically 500 mW.

3. Computational details

The standard basis set 6–31++*G* was used to optimize the chemical structure, vibrational wave number, and related vibrational assignments of TDPCA acid using the Gaussian09 software package. The compound was improved using the B3LYP/6–31++*G*(d,p) calculation. Level ab initio calculations were done using the Gaussian 09 program and the 6–31G(d) and 6–31++*G*(d,p) basis sets of the density functional theory with the three-parameter hybrid functional (B3) for the exchange component and the Lee-Yang-Parr (LYP) correlation function [30,31].

The VEDA4 program was used to create a potential energy distribution, which served as the basis for the vibrational assignments [32]. NBO study on a molecule at the same level using second order perturbation revealed strong evidence of stabilization resulting from the

hyperconjugation of particular intramolecular interactions [33]. Gauss view09 program is used to locate the potential region by obtaining the highest occupied molecular orbital and lowest unoccupied molecular orbital maps (HOMO-LUMO) as well as molecular electrostatic potential maps [34]. RDG analysis for analyzing weak interactions, DOS (Density of States) analysis, and topological analysis were all carried out using Multiwfn, a multifunctional wave function analysis tool [35], and all isosurface maps were rendered using the VMD program [36]. The Auto dock tools 1.5.6 (ADT) program was used to simulate the interactions between three distinct active site proteins [37]. Many theories produce comparable results, however DFT theory's results were more accurate when compared to experimental data [38] hence we are just discussing DFT theory's results here.

4. Resuls and discussion

4.1. Molecular geometry

The number of atoms and geometric information such as bond length and bond angle are all obtained using Gaussian09 to determine the title compound's ideal molecular structure. The TDPCA molecule has a planar structure (Fig. 1). TDPCA can have the carboxylic group oriented in either an anti- or syn-orientation. The global minimum on the potential energy surface was found and identified, and the molecular geometry of TDPCA was fully optimized. According to Fig. 1, which is specific to the solid phase of orotic acid, the carboxyl group appears to be oriented in a syn-direction [39].

In TDPCA molecule average computed N2-C1 bond length is 1.407 Å, which is quite close to the typical N—C single bond (1.405 Å), as seen in Supplementary Table 1.The TDPCA molecule's anticipated N4-C3, N2-C3, and N4-C5 bond lengths vary depending on the method used, and range from 1.378 Å to 1.391 Å, which is halfway between the length of an N—C single bond (1.40 Å) and an N=C double bond (1.29 Å). The N4-C5 and N2-C3 bonds are the two N—C bonds that are the shortest, which suggests that the C6-C5-N4-C3 skeleton has experienced severe p-electron delocalization.

The bond distances between N4–H11 and N2–H13 in this molecule are estimated to be around 1.010 Å –1.012 Å Our theoretical values match actual evidence [40] and previous theoretical research on similar chemicals remarkably well [41].The C5=C6 bond's computed average length of 1.349 Å, which is somewhat longer than a C=C double bond and supports the p-electron delocalization, is a good indicator of this. A C—C single bond (1.54 Å) and a C=C double bond (1.34 Å) are more similar in length than the C5-C7 bond, which is 1.498 Å The average calculated O8-H10 bond length for TDPCA is 0.96 Å, which is very close
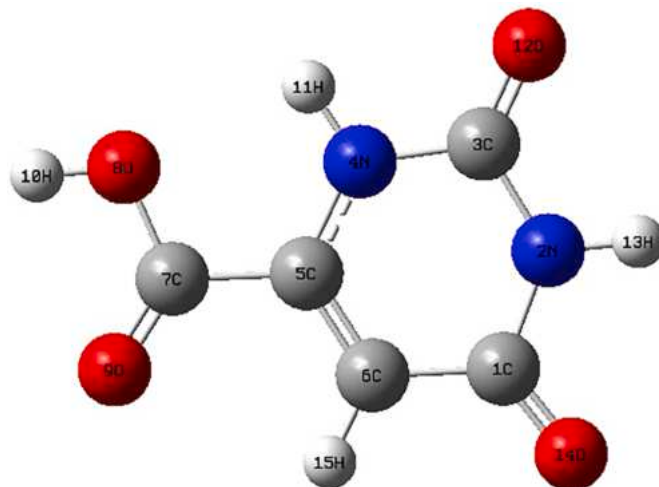

Fig. 1. Optimized structure of TDPCAcalculated at DFT level.

to the typical O—H single bond length (0.95 Å).

The computed length of the C7-O9 bond is 1.26 Å, longer than the typical length of a C=O double bond (1.20 Å), but shorter than the typical length of a C—O single bond (1.43 Å). The computed length of the C7-O8 bond is 1.353 Å, which is less than the average length of a C—O single bond (1.43 Å). The lengths of the carboxyl bonds at positions C7-O9 and C7-O8 are different from those of conventional single and double bonds, suggesting that the double bond and negative charge are delocalized [42–44].

The uracil ring has carbon-oxygen bond lengths of 1.210 for the C3-O12 bond and 1.214 for the C1-O14 bond, both of which are sufficient for a double C=O bond length (120 pm) (Fig. 1; Table 1). The two carbon-oxygen bonds in the carboxylate group have identical bond length values (Table 1), which suggests that the carbon-oxygen single bond is not coordinated [45].

The calculated dihedral angle for the carboxyl group O9-C7-O8-H10 is 0.0, indicating that the carboxyl group is planar and is in the same plane as the uracil ring. The computed bond angle for the carboxyl group O9-C7-O8-H10 is 123.95. The calculated dihedral angles C1–N2–C3–N4, N2–C3–N4– C5, C3–N4–C5–C6, and N4–C5–C7–O9 are 0.0 (Table 1) as well as O12–C3–N4–H11, O12–C3–N2–H13, O14–C1–N2–H13, and O14–C1–C6–H15. These images show how the TDPCA molecule is planar. The matching angles of the uracil ring, which are virtually identical to those documented in the literature, demonstrate the planarity of the 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acid molecule [46–48].

4.2. Natural bond orbital analysis

NBO analysis is a helpful method for studying charge transfer and conjugation in molecular systems, as well as intra- and intermolecular bonding interactions [49–51]. On the basis of second-order perturbation, the stabilization energy E2 is calculated.

$$E2 = \Delta E_{ij} = q_i F(i,j)^2 / Ei - Ej$$

The donor orbital occupancy is qi, the diagonal elements are Ei and Ej, and the Fock matrix elements are F(i,j).To ascertain the intramolecular rehybridization and delocalization of electron density, the molecule was subjected to NBO analysis at the DFT (B3LYP)/6–311 G(d, p) level. The estimated energy values for the interaction between the filled I and vacant j for the compound in the title have been compiled in Table 1.

Strong electron delocalization was visible in the electron densities of the conjugated bonds of the aromatic ring (0.19 to 0.28 e) and (1.82 to 1.97 e).This happens when molecules inside the ring interact through a charge transfer process. The shift from LP1(N2), (C5-C6), and (C5-C6) donor orbitals to (C3-C12), (C1-C14), and (C7-C9) acceptor orbitals,

with stabilization energies of 57.37, 19.91 and 16.47 kcal/mol, is crucial for the high stability of TDPCA molecule. The intramolecular hydrogen bonding is the result of the overlap between (C6-H15) and (N4-C5), which has a stabilization energy of 6.82 kcal/mol. An increase in the E (2) value indicates a stronger delocalization interaction between electron donors and electron acceptors.

4.3. Vibrational spectral analysis

The harmonic vibrational wavenumbers for TDPCA at the 6–311++G (d,p) level were calculated using the DFT method.1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidcontains 15 atoms with 39 distinct vibrational modes. Calculations were made to determine the potential energy distribution (PED) among the symmetry coordinates for each of the molecule's normal modes. Theoretical computations reveal that none of the frequencies have fictitious values. All DFT-calculated frequencies are scaled by 0.96 to compare to experimental frequencies.

Figs. 2 and 3 are the experimental and stimulated FT-IR and FT-Raman spectra of the title compound, respectively. The computed frequencies slightly differ from the experimental values, as shown in Table 2, as can be seen. The absence of anharmonicity in DFT computations may be the cause of these minor variances. Furthermore, the environment in which the molecule is placed has an impact on the vibrational frequencies. These variations are expected given that the computed vibrational modes are based on a single gas phase molecule.

4.3.1. C—C vibration

The benzene derivatives exhibit C—C stretching vibrations between 1650 cm^{-1}and 1590 cm-1 [52,53]. C—C stretching is thought to be responsible for the scaled frequency values at 1618 cm^{-1}, 1319 cm^{-1}in the compound uracil ring of the same name. The C—C stretching or strong band vibration, is shown in Raman with a wavelength of 1649 cm^{-1}. The assigned value is moved outside of the acceptable range. In these modes, the wavenumber shift caused by the mass of replacement has been reversed. This demonstrates that energy is gained rather than lost as in the earlier situations [54,55]. Due to N's attaching to the ring in this instance, ring C—C stretching vibrations were muted.

4.3.2. C- N vibration

Due to the possibility of multiple vibrational modes combining in this area, finding the ring C—N stretching vibration is a particularly challenging task. The C—N stretching vibration, which is always associated with other bands, is commonly seen between 1266 cm^{-1}, and

Table 1
Second order perturbation theory analysis of fock matrix in NBO basis.

Donor NBO (i)	ED	Acceptor NBO (j)	ED	E (2) Kcal/mol	E(j) – E (i) a.u	F(i, j) a.u
π(C1-O14)	1.9767	π*(C5-C6)	0.1990	5.39	0.39	0.043
π(C5-C6)	1.8215	π*(C1-O14)	0.2887	19.91	0.32	0.073
		π*(C7-O9)	0.2282	16.47	0.29	0.062
σ(C6-H15)	1.9725	σ*(N4-C5)	0.0305	6.82	0.96	0.072
LP(1)N2	1.6476	π*(C1-O14)	0.2887	47.88	0.29	0.107
		π*(C3-O12)	0.3473	57.37	0.28	0.113
LP(1)N4	1.6521	π*(C3-O12)	0.3473	52.65	0.29	0.110
		π*(C5-C6)	0.1990	40.24	0.30	0.102

The strong stabilization energy from LP1 (N2) to *(C3-O12) is 57.37 kcal/mol. The lone pair electrons of LP1(N4) to *(C3-C12), LP1(N2) to *(C1-C14), and LP1 (N4) to *(C5-C6) correspondingly have stabilization energies of 52.65, 47.88 and 40.24 kcal/mol.

Fig. 2. Experimental and simulated FT-IR spectrum of TDPCA.
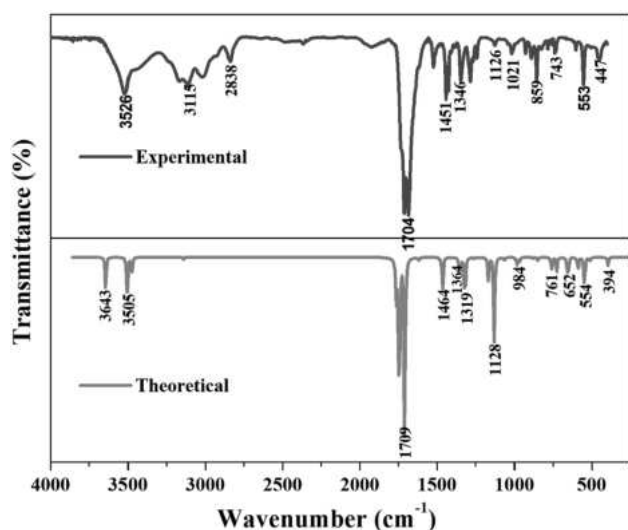
Fig. 3. Experimental and simulated FT-Raman spectrum of TDPCA.

Table 2
Comparison of the experimental (FT-IR, FT-Raman) wavenumbers (cm-1) and theoretical wavenumbers (cm-1) of TDPCA acid.

Wave Number			Assignments with PED(<10 %)
Experimental Values		Scaled Values	
FT-IR	FT-Raman		
3526	-	3643	ν(O8-H10)100 %
-	-	3505	ν(N4-H11)100 %
-	-	3476	ν(N2-H13)100 %
-	-	3143	ν(C6-H15)100 %
-	-	1764	ν(O9-C7)84 %
-	-	1744	ν(O12-C3)72 %
1704	1709	1709	ν(O14-C1)74 %
-	1649	1618	ν(C6-C5)66 %
1451	-	1464	ν(N4-C5)10 %+ν(N4-C5)15 % β(H11-N4-C5)32 %
1346	-	1364	β(H13-N2-C1)63 %
-	-	1350	ν(N4-C3)18 % + ν(N2-C1)12 % β(C6-C1-O14)22 % + β(H11-N4-C5)14 %
-	-	1319	ν(N4-C5)12 % + ν(C7-C5)12 % β(H10-O8-C7)26 % + β(H11-N4-C5)10 %
-	1250	1264	β(H10-O8-C7)11 % + β(H11-N4-C5)17 % + β(H15-C6-C5)38 %
-	-	1168	ν(N2-C3)50 % + β(H15-C6-C5)15 %
1126	-	1128	ν(N4-C5)23 % + ν(C7-C5)10 % β(H10-O8-C7)41 %
1021	1047	1064	ν(N4-C5)43 % + ν(H15-C6-C5)23 %
-	-	984	ν(N4-C5)21 % + β(C6-C1-O14)40 %
-	-	968	ν(N2-C3)17 % + ν(N4-C3)14 % β(C1-N2-C3)40 %
-	-	880	ν(N4-C3)23 % + ν(C7-C5)14 % β(C3-N4-C5)26 %
859	-	848	τ(H15-C6-C5-N4)68 % + γ(C7-C6-N4-C5)11 % + γ(O14-N2-C6-C1)12 %
743	-	761	γ(O14-N2-C6-C1)12 % +γ(O9-C5-O8-C7)69 %
-	731	727	γ(O12-N2-N4-C3)84 %
-	-	708	τ(H15-C6-C5-N4)21 % + γ(O14-N2-C6-C1)55 %
-	-	659	γ(O8-C7)13 % + β(O9-C7-O8)46 %
-	-	652	τ(H10-O8-C7-C5)10 % + τ(H13-N2-C1-C6)80 %
-	608	603	β(C5-C7-O8)33 %
-	-	590	τ(H10-O8-C7-C5)72 % + τ(H13-N2-C1-C6)12 %
553	-	554	γ (N2-C1)25 % + β(O9-C7-O8)14 % + β(C3-N4-C5)30 %
-	-	536	τ(H10-O8-C7-C5)81 %
-	-	513	β(C1-N2-C3)71 %
447	456	442	γ(C7-C6-N4-C5)47 % + τ(C5-N4-C3-N2)11 % +γ(O9-C5-O8-C7)11 %
-	-	415	β(C5-C7-O8)69 %
-	394	394	γ(N2-C1)16 % + β(C6-C1-O14)61 %
-	-	327	γ(N4-C13)10 % + γ(C7-C5)26 % β(C6-C1-O14)10 % + β(O9-C7-O8)13 % + β(C3-N4-C5)13 %
-	-	172	β(C5-C7-O8)82 %
-	-	168	γ(C7-C6-N4-C5)15 % + τ(C5-N4-C3-N2)65 %
-	-	142	γ(C7-C6-N4-C5)86 %
-	-	114	γ(C7-C6-N4-C5)77 %
-	-	57	τ(O8-C7-C5-N4)89 %

1382 cm^{-1} [56,57]. The C—N stretching mode is theoretically observed at 1364 cm^{-1}, 1350 cm^{-1},1319 cm^{-1}and 1168 cm^{-1}.In the current experiment, the medium band of the C—N stretching vibration is seen at 1346 cm^{-1}.

The C—N stretching vibration, which is also noted at the lower end of the anticipated range, may be connected to the interaction of the C—C vibration, whose frequency reaches this value [58]. In-plane bending of C—N measured at 330 cm^{-1}and out-of-plane bending measured at 580 cm^{-1}, respectively [59]. The C—N in plane bending is visible at 553 cm^{-1}in the infrared spectrum. At 447 cm^{-1}in the IR and 456 cm^{-1} in the Raman spectra, out of plane bending in C—N is observed. The measured spectrum and published results show excellent agreement with theoretically calculated C—N vibration wavenumbers produced using B3LYP/6–311++G (d, p) techniques.

4.3.3. C—H vibration

The hetero aromatic organic compounds usually exhibit several broad bands in the range of 3100 cm^{-1}–3000 cm^{-1}because of C—H stretching vibrations [60]. The 1000 cm^{-1}–1300 cm^{-1}range of the C–H in plane bending vibrations, which were detected, is particularly useful for characterization [61]. Vibrations in the C—H plane at 1264 cm^{-1}, 1168 cm^{-1},1128 cm^{-1},1064 cm^{-1}that were theoretically detected. The succession of weak bands in Raman at 1250 cm^{-1} that have been observed are thought to be caused by the C—H in plane bending vibration. The vibrations caused by the C—H out of plane bending occur in the 1000 cm^{-1}–750 cm^{-1} region.

4.3.4. N—H vibration

The hetero aromatic molecule with an N—H group absorbs stretching between 3500 and 3220 cm^{-1} [62]. The DFT level calculated vibration at 3505 cm^{-1} is assigned to N—H stretching vibration. This corresponds to interaction of the intermolecular hydrogen bond N—H…O. It was observed experimentally in IR at 2838 cm^{-1} [63,64]. The presence N—H hydrogen bond is evidenced from the elongation of N—H bond from the normal value. The title molecule's bending N—H vibration is very weakly visible at 1451 cm^{-1}in the IR spectra; the anticipated value for the N—H bending mode is 1464 cm^{-1}. The C—N stretching vibrations overlap with this vibration.

4.3.5. Hydroxyl vibration

The hydroxyl groups are likely to be the most sensitive to the environment and so they show pronounced shifts in the spectrum of the hydrogen bonded species. The hydroxyl group absorbs strongly in the 3700–3584 cm^{-1} region, whereas the existence of intermolecular

hydrogen bond formation can lower the O—H stretching wavenumber to 3550–3200 cm^{-1} region [63–65]. The IR spectrum in the high region shows a broad intense band at 3526 cm^{-1}attributed to hydrogen-bonded OH stretching. The in-plane OH deformation vibration usually appears as a strong band in the region 1440–1260 cm^{-1} [21]. The calculated band at 1350 cm^{-1} corresponds to the in-plane bending mode of the hydroxyl vibration.

4.3.5. C=O vibration

The carbonyl group is present in numerous different types of compounds and displays a strong band in the range of 1850 cm^{-1}–1550 cm^{-1}due to the C=O stretching vibration [66]. Due to hydrogen bond formation or conjugation, the intensity of these bands may rise. The carbonyl group's composition was also influenced by the oxygen's single

electron pair. Theoretically calculated wave numbers for C=O stretching vibrations in the current work are 1764 cm^{-1}, 1744 cm^{-1}, 1709 cm^{-1}, 1618 cm^{-1}. The frequency of the carbonyl stretching vibration in the TDPCA molecule is absorbed at 1704 cm^{-1} (IR), 1709 cm^{-1} (Raman) and 1709 cm^{-1}(DFT calculations), respectively. The observed FT-IR and FT-Raman bands are leads to the hydrogen presence in the title molecule [67]. It is anticipated that the in-plane and out-of-plane C=O vibrations will be present in the regions 725 cm^{-1} and 595 cm^{-1}, respectively [68]. In our most recent research, the bands at 447 cm^{-1}(IR), 456 cm^{-1}(Raman), and 442 cm^{-1}(DFT) have been distinguished as C=O in-plane and out-of-plane bending vibrations respectively.The in-plane and out-of-plane bending of the carbonyl group interacts heavily with the vibrations of the C—N and C—C bonds.

4.4. Frontier molecular orbitals (FMOs) analysis

The highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals—used to characterize how a molecule interacts with other species—are referred to as "FMOs" The HOMO and LUMO energies respectively, define the capacity to give and accept electrons. The energy gap between HOMO and LUMO, which is the result of a significant amount of ICT from the effective electron acceptor groups through the conjugated path to the end-capping electron donor groups, is what defines the molecule chemical stability and electron conductivity [69]. Using the DFT/B3LYP method and the 6311++G(d,p) basis set, Fig. 4 calculates the HOMO and LUMO orbitals of the TDPCA molecule.

The FMOs investigation of the TDPCA molecule reveals that modifications are delocalized within the molecule, increasing the molecule's molecular reactivity. Calculations are made for the title compound's additional significant properties, including its electron affinity, chemical potential, electronegativity, chemical hardness, ionization potential, chemical softness, and electrophilicity index. These qualities are listed in Supplementary Table 2. These tables demonstrate that electrons are transferred from the HOMO orbital, which contains many electrons with an energy of −7.6245 eV, to the LUMO orbital, which has fewer electrons with an energy of −3.0348 eV and a −4.5897 eV energy gap. The stability, internal charge transfer, and biological activity of TDPCA are all demonstrated by this bandgap. The molecule's anticancer activities are being supported by this charge transfer interaction.
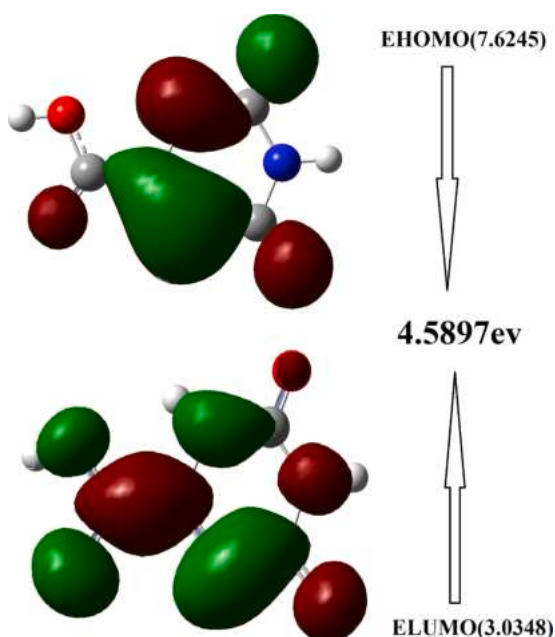
4.5. Molecular electrostatic potential

The molecular electrostatic potential (MESP) of a molecule is related to its dipole moments, partial charges, electro negativity and chemical reactivity. It denotes the net electrostatic impact created at a particular place by the overall charge distribution (electron + nuclei) of the molecule. It provides a visual representation of the molecule's relative polarity [70]. It demonstrates the degree of reactivity by the colours of the regions under electrophilic and nucleophilic assault.

The size, charge density, shape and position of a molecule's chemical reactivity are depicted on an electron density isosurface that has been generated by mapping an electrostatic potential surface. A mapped electrostatic potential surface has been created for the title molecule in the 6–311++ G(d,p) basis set using the computer application Gauss view [71]. Fig. 5 displays the MESP curve for TDPCA. The different electrostatic potential values at the surface are shown by different colours: blue indicates locations with the highest positive electrostatic potential, red indicates areas with the largest negative electrostatic potential, and green indicates places with zero potential. Red, orange, yellow, green, and blue have the potential for the most growth.

The structure and charge density distributions in the molecule's sites around the oxygen atom, which display regions with the majority of the negative electrostatic potential, have an effect on how the electrostatic potential surface looks in every case. The MEP map of the molecule under investigation clearly suggests that the negative potential regions (red) are electrophilic regions; these are largely caused by O atoms and are reflected as pale yellow glob. The highest positive potential zone (blue) is centered on the hydrogen atoms in the hydroxyl group as well as localized over the hydrogen C—H atoms, suggesting the probable sites for nucleophilic attack.

The MEP map reveals that the positive potential sites are concentrated around the hydrogen atoms, whereas the negative potential sites are found around the oxygen and nitrogen atoms. These active sites were discovered to be unmistakable proof of the compound's biological action.

4.6. Mulliken and natural population analysis

Atomic charges are crucial in the application of quantum chemical calculations to molecular systems because they affect the dipole moment, electronic structure, molecule polarisability and a variety of other features of molecular systems [72,73]. Supplementary Table 3 contains the Mulliken charges estimated at the B3LYP/6–31 G (d,p)
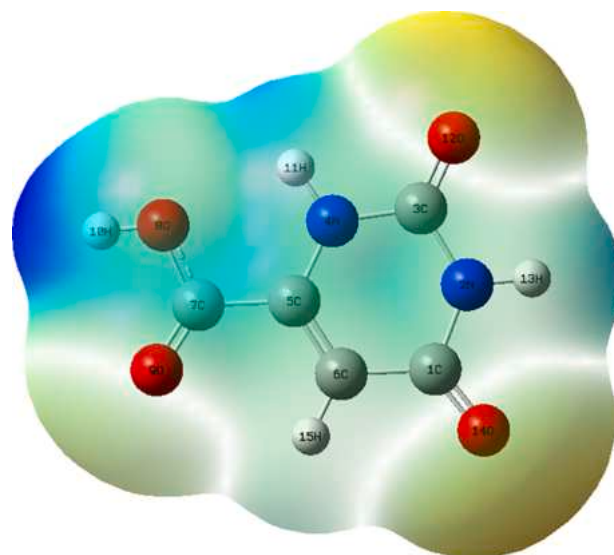

Fig. 4. Frontier molecular orbital of TDPCA.

Fig. 5. MESP of TDPCA.

basis set. However, the results can be better illustrated graphically, as seen in Supplementary Fig. 1. Because of polarizability, the charges alter depending on the base set. The basis set's definition of each atom's electron population was used to calculate the distribution of atomic charges.

The amount of positive and negative carbon atomic charges was discovered in the compound mentioned in the title. The oxygen atom has a donar atom, which is negatively charged. The acceptor atom, hydrogen, has a positive charge on its atom. The range of the nitrogen atomic charge's negative magnitude was found to be between -0.408252 and -0.418154.

An alternative to the traditional Mulliken population analysis is the natural analysis. It appears to show numerical stability and provide a better explanation for how electrons are distributed within molecules [74,75]. The lowest charge among hydrogen atoms is found on H15 atoms, while all carbon atoms except C6 are positively charged. Nitrogen N2 atoms have the largest negative charge (0.64009e), according to the charge study. All oxygen atoms have a negative charge, however O8 (-0.69409) has the highest charge of all oxygen atoms. The hydrogen atoms in the TDPCA molecule are all positively charged, however atoms H10 (0.49963e) and H11 (0.43624e) are positively charged more than the other hydrogen atoms because they are bound to an oxygen or nitrogen atom.

The largest positive and negative charges of the OA molecule indicate the potential for hydrogen bonding interaction in the natural population analysis, which is supported by the analysis of natural bond orbitals. Atoms C3 (0.80975e), shows a higher positive charge, and N2 (0.64009e), which displays a higher negative charge, are indicators of charge delocalization in the molecule.

4.7. Topology analysis

4.7.1. Atoms in molecule (AIM)

One of the most effective techniques is to take advantage of the topology analysis of charge density or electron density of molecules in the actual space boundaries of the atoms forming a system in order to analyze the properties of the atoms forming a system using the Bader theory of "Atoms in Molecules." AIM is a potent tool for learning about hydrogen bonding and chemical bonds in the molecular system [76,77].

The key interactions, such as those for powerful H-bonds ($\nabla^2 \rho$BCP) < 0 and covalent in nature, for weak H-bonds ($\nabla^2 \rho$BCP) > 0 and HBCP > 0 and electrostatic in nature and medium H-bonds ($\nabla^2 \rho$BCP) > 0, HBCP<0 and partially covalent, can be identified using topological analysis of the electronic charge density and their Laplacian at the bond critical point (BCPs).The AIM analysis was used to calculate the BCPs' parameters once the area of electron density was obtained. A larger electron density indicates that the bond is stronger.

The negative value of ($\nabla^2 \rho$) indicates the covalent nature of the covalent connection; in this instance, concentrated electron density in the area between the interacting atoms, leading to a highly constrained common electron pair. Positive values of ($\nabla^2 \rho$) calculated at the relevant BCP denote closed-shell interactions, which are typical of closed-shell interactions like hydrogen bonds and entail the depletion of electronic charge along the bond pathway. The cycle susceptibility to opening is specified by the bond ellipticity, which was looked into as a revealing indicator of the pi-bond character. The pi delocalization will be larger the higher the ellipticity value.

The AIM analysis's results for the electron density, Laplacian density, and ellipticity of the compound in the title were used to calculate the compound's properties, which are shown in Supplementary Table 4.The crucial points and associated graphic produced by the B3LYP/6–311++G(d,p) technique using the multiwfn program are shown in Supplementary Fig. 2. The C7-O9(0.4242) link in the TDPCA molecule has the highest electron density, and this bond's higher negative value, demonstrates its covalent character. The C5-C6 atom has a higher value than the others in ellipticity, which denotes the bond's instability.

4.7.2. ELF and LOL analysis

The Multiwfn program was used to do the topological analysis of the Localized Orbital Locator LOL and the Electron Localization Function (electrons are strongly localized) ELF [35]. Frequently, multiwfn programmes are used to reveal the atomic shell structure, categories chemical bonds, and validate charge-shift bonds on molecular surfaces. ELF allows for the local measurement of Pauli repulsion, which is directly connected to the kinetic energy of electrons. Greater values ELF [0.0 to 1.0] show that electrons are significantly localized, which implies the presence of a covalent bond, inner shells, or a lone pair of the atoms. Smaller values (0.5) represent locations where electrons are projected to be delocalized. Supplementary Fig. 3 shows the ELF and LOL color-filled map.

The ELF color-filled map (a) shows the delocalized electron cloud around a few carbon and nitrogen atoms in blue, the localized bonding around hydrogen atoms in considerable detail, and the non-bonding electrons in red. The LOL color-filled map(b)reveals that the centres of certain hydrogen atoms are white, indicating that a single localized orbital predominates the bonds (exceeds the upper limit 0.8).The valence shell, inner shells of a few carbon atoms, and oxygen atoms of the depleted region are represented by the blue circles, whereas the high LOL values of the covalent kind of electron depletion zone around the nitrogen atom are represented by the red circles.

4.7.3. RDG analysis

The RDG technique, which uses Multiwfn and the visual molecular dynamics (VMD) application, is a topological tool that exposes non-covalent interactions such as van der Waals, steric effects and hydrogen bonds [36]. The region of these interactions and their graphical representation are provided by the RDG analysis, which is based on the electron density and its derivatives [78]. The second eigen value of the electron density, sign (λ_2)ρ, can be used to build an RDG scatter graph between RDG and sign (λ_2)ρ, that provides crucial information on the nature and strength of the interactions. To describe the nature of interactions, use sign(λ_2)ρ. Repulsion is indicated by a sign (λ_2)ρ > 0, attraction is indicated by a sign (λ_2)ρ < 0, and van der Waals weak interaction is indicated by a sign (λ_2)ρ virtually 0 [79].

The steric effect can be seen in the red color scatter near the carboxylic group in the RDG isosurface plot, and it can also be seen in the scatter plot between the positive region between 0.05 and 0.01 a.u [80]. Green color scatter indicates the presence of non-covalent van der Waals force H…H inter scatter between 0.01–0.02.The red scatter in the isosurface density plots, which show the non-bonded interactions from the RDG graph, appears in a high range between 0.05 and -0.01 a.u. The RDG scatter graph demonstrates the significant steric effect and van der Waals force interaction that take place in the compound of TDPCA. The Fig. 6(a,b) displays the RDG scatter and isosurface plots of TDPCA.

4.8. Fukui function

One of the greatest ways to understand well-known chemical concepts like electronegativity, electron affinity, ionization potential, chemical potential etc., is through the use of DFT. The atomic descriptors were first used by Kolandaivel et al. to identify the molecule's regionally reactive areas [81]. Individual atomic charge calculated by NPA (Natural Population Analysis) and MPA (Mulliken Population Analysis) has been used to develop the Fukui function. The following expression is used to determine fukui functions [82]:

$$\text{The nucleophilic attack} = fj^+ = [qj(N+1) - qj(N)]$$

$$\text{The electrophilic attack} = fj^- = [qj(N) - qj(N-1)]$$

$$\text{and radical attack} = fj^0 = [qj(N+1) - qj(N-1)]/2$$

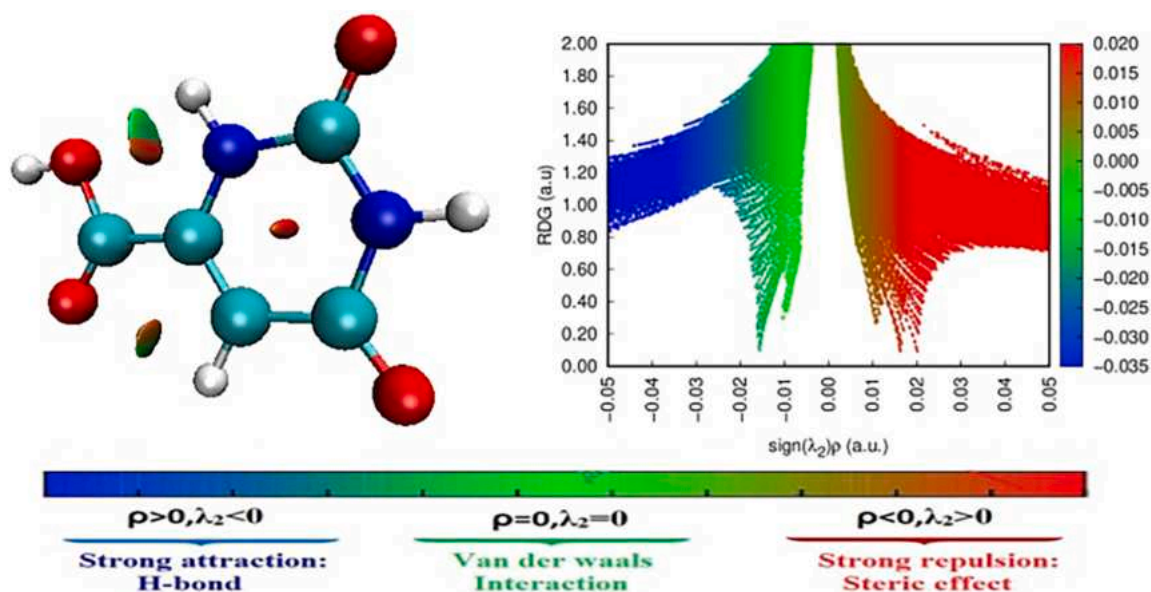where q stands for the atomic charge at the rth atomic site in chemical

Fig. 6. RDG scatter and isosurface plots of TDPCA.

species like neutral (N), anionic ($N + 1$), and cationic (N-1) species. The signs $+$, $-$, and 0, respectively stand for nucleophilic, electrophilic, and radical assaults. The formula below is used to calculate the Dual descriptor f(r).

$$\text{Dual descriptor} = \Delta F(r) = \left[f_j^+ - f_j^{--} \right]$$

Dual descriptors F(r)>0 and F(r)0, respectively, indicate that nucleophilic assaults are more advantageous at the site than electrophilic attacks. Both descriptors F(r) distinguishes between nucleophilic and electrophilic attack at a certain location with their sign. A visual representation of fukui functions and dual descriptors is shown in Supplementary Fig. 4(a,b,c).For each atom in the molecule, Supplementary Table 5 gives full details on the Fukui functions, local softness, Mulliken atomic charges and dual descriptor values. In the following order, the both descriptor for nucleophilic attack C7 > C1 > C3 is presence in the molecule. Dual adjective describes an electrophilic attack negatively as the order C6 > O14 > O12 > O9 > N4 > H15 > C5 > H11 > H13 > O8 > H10 >N2. The likelihood of radical attacks is indicated by numbers that are nearly equal to zero or equal to zero. Electrophilic attack is more likely to occur on negatively charged atoms. If the three types of assaults are compared, it may be discovered that the nucleophilic attack is more reactive than the electrophilic and radical attacks.

4.9. Molecular docking

The fields of structural molecular biology, pharmacogenomics, and computer-aided drug design all place a high value on molecular docking [83]. In the current investigation, molecular docking for a protein linked to anti-cancer activity was done. The composition of the target protein 2WXQ was retrieved from the RSCB protein data repository [84]. The Ramachandran plot, as illustrated in Fig. 11, is used to assess the protein's quality and confirm that all of the residues are present within the permitted range. To create the ligand PDB format, the optimal molecular structure using density functional theory is helpful [85]. TDPCA is docked with 2WXQ protein using the Auto Dock program. In the protein-ligand binding interaction, a binding energy of −5.45 kcal/mol was detected. Fig. 7 depicts this, and Table 3 lists the significant features. This chemical is a potent anti-cancer medication, as evidenced by the low value of binding energy.
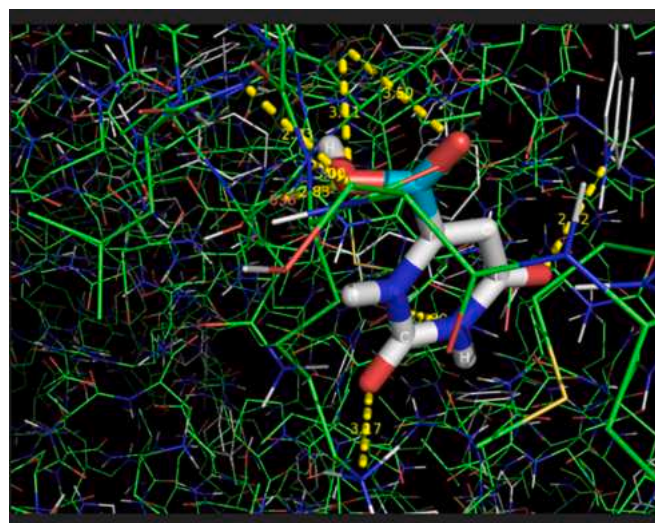
Fig. 7. Molecular docking of TDPCA.

Table 3
Protein, binding residue, bond distance, binding energy of TDPCA.

Protein (PDB: ID)	Binding residues	Bond distances (Å)	Binding energy (kcal/mol)
2WXQ	ILE	3.50	−5.6
	ILE	3.11	
	ASN	2.83	
	LYS	2.83	
	ASN	3.17	
	LYS	3.00	
	ASP	2.89	
	TRP	2.92	

5. Conclusion

The current study reports on the spectroscopic investigation of TDPCA utilizing FT-IR and FT-Raman, theoretically and experimentally. Overall, it was observed that there was good agreement between the typical modes of vibrations in theory and experiment. Strong

intramolecular hyperconjugative interaction and molecular stability are indicated by NBO analysis. The HOMO and LUMO values are computed and the energy gap is determined to be 4.5897 eV. The molecule's anti-cancer efficacy is supported by its low HOMO-LUMO energy, which also indicates the possibility of ICT interaction. Using MEP, Fukui, ELF, LOL, and RDG, the electron distribution and reactive sites on the surface of the compound were examined. The anti-cancer effect of the title molecule is further revealed by the strong activity of TDPCA molecule against IA PI 3-Kinase PDB (2WXQ), with a binding energy of −5.6 kcal / mol. Thus, the novel title compound TDPCA may be expected as a useful therapeutic target for cancer treatment. We expect that all of the study's findings will be useful to researchers who are analyzing and creating novel materials.

Funding

CRediT authorship contribution statement

M. Amalanathan: Writing – original draft, Funding acquisition. **T. Brintha:** Writing – review & editing, Investigation. **S. Sijana:** Conceptualization, Validation. **P.J. Jegan Babu:** Resources. **M. Sony Michael Mary:** Methodology, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financialinterestsor personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgment

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chphi.2023.100393.

References

[1] P.L. Panzeter, D.P. Ringer, DNA precursors are channelled to nuclear matrix DNA replication sites, Biochem. J. 293 (3) (1993) 775–779, https://doi.org/10.1042/bj2930775.

[2] M. Berthelot, G. Cornu, M. Daudon, M. Helbert, C. Laurence, Computer-aided infrared analysis of urinary calculi, Clin. Chem. 33 (11) (1987) 2070–2073, https://doi.org/10.1093/clinchem/33.11.2070.

[3] P. Banditt, Determination of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidin serum by high-performance liquid chromatography, J. Chromatogr. B Biomed. Sci. Appl. 660 (1) (1994) 176–179, https://doi.org/10.1016/0378-4347(94)00264-9.

[4] M.T. McCann, M.M. Thompson, I.C. Gueron, M. Tuchman, Quantification of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidin dried filter-paper urine samples by stable isotope dilution, Clin. Chem. 41 (5) (1995) 739–743, https://doi.org/10.1093/clinchem/41.5.739.

[5] A. Sarpotdar, J.G. Burr, The structures and hydrogen ion dependence of the complexes between 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidderivatives and the ions, Ni (II) and Eu (III), J. Inorg. Nucl. Chem. 41 (4) (1979) 549–553, https://doi.org/10.1016/0022-1902 (79)80443-1.

[6] D. Lalart, G. Dodin, J.E. Dubois, Study of the slow protonation of the complexes between divalent nickel ion and dianion of orotic-acid or its derivatives, J. Chim. Phys. 79 (1982) 449–453.

[7] E. Fernandez-Garcia, J.U. McGregor, Determination of organic acids during the fermentation and cold storage of yogurt, J. Dairy Sci. 77 (10) (1994) 2934–2939, https://doi.org/10.3168/jds.S0022-0302.(94)77234-9.

[8] A.S. Akalin, S. Gönç, The determination of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidin ruminant milks using HPLC, Milchwissenschaft (1996) (Germany).

[9] P. Ruas-Madiedo, J.C. Bada-Gancedo, E. Fernandez-Garcia, D.G. De Llano, C.G. de Los Reyes-Gavilan, Preservation of the microbiological and biochemical quality of raw milk by carbon dioxide addition: a pilot-scale study, J. Food Prot. 59 (5) (1996) 502–508.

[10] F. Takusagawa, A. Shimada, The crystal structure of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidmonohydrate (Vitamin B13), Bull. Chem. Soc. Jpn. 46 (7) (1973) 2011–2019, https://doi.org/10.1246/bcsj.46.2011.

[11] I.P. Mutikainen, X-ray structural studies on metal complexes of uracil and orotic acid: a survey of coordination induced changes in the uracil fragment, 1989.

[12] R. Letellier, M. Ghomi, E. Taillandier, Out-of-plane vibration modes of nucleic acid bases, Eur. Biophys. J. 14 (4) (1987) 227–241.

[13] P.K. Szczepaniak, M. Szczesniak, J.E.D. Bene, Experimental and ab initio quantum mechanical studies of the vibrational spectra of isolated pyrimidine bases. Recent Experimental and Computational Advances in Molecular Spectroscopy, Springer, Dordrecht, 1993, pp. 141–169.

[14] H. Rostkowska, K. Szczepaniak, M.J. Nowak, J. Leszczynski, K. KuBulat, W. B. Person, Thiouracils. 2. Tautomerism and infrared spectra of thiouracils, Matrix-isolation and ab initio studies, J. Am. Chem. Soc. 112 (6) (1990) 2147–2160, https://doi.org/10.1021/ja00162a016.

[15] J. Victor, G. LB, S. DL, Studies of the kinetic mechanism of orotatephosphoribosyltransferase from yeast, (1979) doi:10.1016/s0021-9258.(17)30121-7.

[16] H. Schmidbaur, H.G. Classen, J. Helbig, Aspartic and glutamic acid as ligands to alkali and alkaline-earth metals: structural chemistry as related to magnesium therapy, Angew. Chem. Int. Edit. Engl. 29 (10) (1990) 1090–1103, https://doi.org/10.1002/anie.199010901.

[17] K. Matsumoto, Antitumor activities of platinum blues containing α-pyrrolidone, 3-3-dimethylglutarimide, orotic acid, succinamic acid and oxamic acid, Inorganicachimicaacta 151 (1) (1988) 9–10.

[18] O. Kumberger, J. Riede, H. Schmidbaur, Preparation and crystal structure of zinc bis [orotate (1−)] octahydrate, Z. Naturforsch. B 48 (7) (1993) 961–964, https://doi.org/10.1515/znb-1993-0718.

[19] I. Bach, O. Kumberger, H. Schmidbaur, Orotate complexes. Synthesis and crystal structure of lithium orotate (—I) monohydrate and magnesium bis [orotate (—I)] octahydrate, ChemischeBerichte 123 (12) (1990) 2267–2271, https://doi.org/10.1002/cber.19901231207.

[20] O. Kumberger, J. Riede, H. Schmidbaur, Orotate complexes, II1) preparation and crystal structures of calcium and zinc orotate (2-) hydrates, ChemischeBerichte 124 (12) (1991) 2739–2742, https://doi.org/10.1002/cber.19911241215.

[21] I. Kostova, N. Peica, W. Kiefer, Theoretical and spectroscopic studies of new lanthanum (III) complex of orotic acid, Vib. Spectrosc. 44 (2) (2007) 209–219, https://doi.org/10.1016/j.vibspec.2006.11.004.

[22] M. Hejmadi, Introduction to Cancer Biology, Bookboon, 2014.

[23] I. WHO, World cancer report, Geneva, Lyon, (2003).

[24] F.C. Asogwa, E.C. Agwamba, H. Louis, M.C. Muozie, I. Benjamin, T.E. Gber, A. I. Ikeuba, Structural benchmarking, density functional theory simulation, spectroscopic investigation and molecular docking of N-(1H-pyrrol-2-yl) methylene)-4-methylaniline as castration-resistant prostate cancer chemotherapeutic agent, Chem. Phys. Impact 5 (2022), 100091.

[25] A.E. Owen, E.I. Ime, E.N. Mbim, H.O. Edet, I. Benjamin, G.I. Iniama, H. Louis, Exploring the anticancer potential of sulfate-hydroxy-butanone derivatives: insights from experimental and quantum chemical investigations, Z. Phys. Chem. 237 (10) (2023) 1643–1668.

[26] C.G. Apebende, P.S. Idante, H. Louis, U.S. Ameuru, T.O. Unimuke, T.E. Gber, F. C. Asogwa, Integrated spectroscopic, bio-active prediction and analytics of isoquinoline derivative for breast cancer mitigation, Chem. Afr. 5 (6) (2022) 1979–1995.

[27] B.R. Raajaraman, N.R. Sheela, S. Muthu, Quantum chemical, vibrational spectroscopic and molecular docking studies of 1-(Diphenylmethyl) piperazine, Polycycl. Aromat. Compd. 42 (5) (2022) 2672–2692.

[28] A.S. Kazachenko, Y.N. Malyar, N.Y. Vasilyeva, V.S. Borovkova, N. Issaoui, Biomass conversion and biorefinery, 1–10 (2021).

[29] O. Noureddine, N. Issaoui, S. Gatfaoui, O. Al-Dossary, H. Marouani, J. King Saud Univ.-Sci. 33 (2) (2021), 101283.

[30] A.D. Becke, A new mixing of Hartree–Fock and local density-functional theories, J. Chem. Phys. 98 (2) (1993) 1372–1377, https://doi.org/10.1063/1.464304.

[31] M.J. Frisch, et al., Gaussian 09 Revision D.01, Gaussian Inc., Wallingford, CT, 2009.

[32] M.H. Jomroz, Vibrational energy distribution analysis VEDA4 (Warsaw), 2004.

[33] E.D. Glendening, A.E. Reed, J.E. Carpenter, F. Weinhold, Nbo Version 3.1, Tci. University of Wisconsin, Madison, 65, 1998.

[34] R. Dennington, T. Keith, J. Millam, Semichem Inc. Shawnee Mission KS, GaussView, Version, 5, 2009.

[35] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunction analyzer, J. Comput. Chem. 33 (5) (2012) 580–592.

[36] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38.

[37] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A. J. Olson, Automated docking using a Lamarckian genetic algorithm and an

empirical binding free energy function, J. Comput. Chem. 19 (14) (1998) 1639–1662.

[38] A.S. Kazachenko, N. Issaoui, A. Sagaama, Y.N. Malyar, O. Al-Dossary, L. G. Bousiakou, Z. Xiang, Hydrogen bonds interactions in biuret-water clusters: FTIR, X-ray diffraction, AIM, DFT, RDG, ELF, NLO analysis, J. King Saud Univ.-Sci. 34 (8) (2022), 102350.

[39] R. Hilal, Z.M. Zaky, S.A. Elroby, Electronic structure of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidI. Geometry, conformational preference and tautomerism, J. Mol. Struct. Theochem 685 (1–3) (2004) 35–42.

[40] F.A.S Takusagawa, The crystal structure of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidmonohydrate (Vitamin B13), Bull. Chem. Soc. Jpn. 46 (7) (1973) 2011–2019.

[41] A. Hernanz, F. Billes, I. Bratu, R. Navarro, Vibrational analysis and spectra of orotic acid, Biopolym. Orig. Res. Biomol. 57 (3) (2000) 187–198.

[42] A.G. Schneider, H.W. Schmalle, F. Arod, E. Dubler, The interaction of 5-fluoro1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidwith transition metals: synthesis and characterisation of Ni (II), Cu (II) and Zn (II) complexes, J. Inorg. Biochem. 89 (3–4) (2002) 227–236.

[43] F.H. Allen, A.J. Kirby, Bond length and reactivity. variable length of the carbon-oxygen single bond, J. Am. Chem. Soc. 106 (21) (1984) 6197–6200.

[44] H.U. Suter, M. Nonella, A quantum chemical investigation of the C− O bond length and stretching mode of the phenolate anion, J. Phys. Chem. A 102 (49) (1998) 10128–10133.

[45] M. Kieninger, O.N. Ventura, S. Suhai, Density functional investigations of carboxyl free radicals: formyloxyl, acetyloxyl, and benzoyloxyl radicals, Int. J. Quantum Chem. 70 (2) (1998) 253–267.

[46] D.J. Darensbourg, J.D. Draper, D.L. Larkins, B.J. Frost, J.H. Reibenspies, Organometallic derivatives of orotic acid. CO− labilizing ability of the amido group in chromium and tungsten carbonyl complexes, Inorg. Chem. 37 (10) (1998) 2538–2546.

[47] S. Bekiroglu, O. Kristiansson, Hydrogen-bonded neutral and anionic lamellar networks: crystal structures of bis (O, O′, O″-hydroorotato) disilver (i) dihydrate, potassium hydroorotate and rubidium hydroorotate. Ab initio calculations on 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecarboxylic acidand the hydroorotate anion, J. Chem. Soc. Dalton Trans. (7) (2002) 1330–1335.

[48] P. Arrizabalaga, P. Castan, F. Dahan, Coordination sites of 5-nitro-6-carboxyuracil: UV study and x-ray structure determination of diammine (5-nitroorotato) copper (II) hydrate and hexaamminebis (5-nitroorotato) tricopper (II) pentahydrate, Inorg. Chem. 22 (16) (1983) 2245–2252.

[49] S. Shukla, A. Srivastava, P. Kumar, P. Tandon, R. Maurya, R.B. Singh, Vibrational spectroscopic, NBO, AIM, and multiwfn study of tectorigenin: a DFT approach, J. Mol. Struct. 1217 (2020), 128443.

[50] K. Sarojini, H. Krishnan, C.C. Kanakam, S. Muthu, Synthesis, X-ray structural, characterization, NBO and HOMO–LUMO analysis using DFT study of 4-methyl-N-(naphthalene-1-yl) benzene sulfonamide, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 96 (2012) 657–667.

[51] T.S.M. Rajamani, M. Karabacak, Electronic absorption, vibrational spectra, nonlinear optical properties, NBO analysis and thermodynamic properties of N-(4-nitro-2-phenoxyphenyl) methanesulfonamide molecule by ab initio HF and density functional methods, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 108 (2013) 186–196.

[52] F. Weinhold, C.R. Landis, Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective, Cambridge University Press, 2005.

[53] H.H. Perkampus, L.J. Bellamy, The Infrared Spectra of Complex Molecules, 3. Auflage, 1, Chapman and Hall Ltd., London, 1976, 1975433 Seiten32 Abb22 Tabellen, Preis:£ 8.—.

[54] D.N. Sathiyanarayanan, Vibrational Spectroscopy Theory and Application, 424, New Age International Publishers, New Delhi, 2004.

[55] E. Varsanyi, Vibrational Spectra of Benzene Derivatives, Acad. Kiado, Budapest, 1969.

[56] N.P. Singh, R.A. Yadav, Vibrational studies of trifluoromethyl benzene derivatives 1: 2-amino, 5-chloro and 2-amino, 5-bromo benzotrifluorides, Indian J. Phys. 75 (2001) 347–355.

[57] N. Sundaraganesan, C. Meganathan, M. Kurt, Molecular structure and vibrational spectra of 2-amino-5-methyl pyridine and 2-amino-6-methyl pyridine by density functional methods, J. Mol. Struct. 891 (1–3) (2008) 284–291.

[58] M.R. Silverstein, G.C. Bassler, C.T. Morrill, Spectrometric Identification of Organic Compounds, 7th Edn, I ohn Wiley and Sons, 1981.

[59] N. Sundaraganesan, C. Meganathan, B.D. Joshua, P. Mani, A. Jayaprakash, Molecular structure and vibrational spectra of 3-chloro-4-fluoro benzonitrile by ab initio HF and density functional method, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 71 (3) (2008) 1134–1139.

[60] V. Krishnakumar, V. Balachandran, T. Chithambarathanu, Density functional theory study of the FT-IR spectra of phthalimide and N-bromophthalimide, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 62 (4–5) (2005) 918–925.

[61] S. George, Infrared and Raman Characteristic Group Frequencies: Tables and Charts, Wiley, Chichester, 2001.

[62] L.J. Ballamy, The Infrared Spectra of Complex Molecules, Vol. 2, 1980.

[63] S. Gatfaoui, N. Issaoui, T. Roinsel, H. Marouani, J. Mol. Struct. 1225 (2021), 129132.

[64] A.S. Kazachenko, M. Medimagh, N. Issaoui, O. Al-Dossary, MJ. Wojcik, AS. Kazachenko, AV. Miroshnokova, YN. Malyar, J. Mol. Struct. 1265 (2022), 133394.

[65] N. Rekik, N. Issaoui, B. Oujia, M.J. Wójcik, J. Mol. Liq. 141 (2008) 104–109.

[66] V.A. Danie, B.J. Kumari, N.F. Nirmal, T.A.F. Reji, 2023 Molecular structure, vibrational assignment, HOMO-LUMO and mulliken analysis of 2-(4-amino-2-phenylaminothiazol-5-oyl)-N-methylbenzimidazole by DFT method.

[67] M. Medimagh, N. Issaoui, S. Gatfaoui, O. Al-Dossary, A.S. Kazachenko, H. Marouani, M.J. Wojcik, J. King Saud Univ. – Sci. 33 (2021), 101616.

[68] M. Barthes, G. De Nunzio, M. Ribet, Polarons or proton transfer in chains of peptide groups? Synth. Met. 76 (1–3) (1996) 337–340.

[69] K. Fukui, Role of frontier orbitals in chemical reactions, Science 218 (4574) (1982) 747–754.

[70] A.J. Garza, G.E. Scuseria, S.B. Khan, A.M. Asiri, Assessment of long-range corrected functionals for the prediction of non-linear optical properties of organic materials, Chem. Phys. Lett. 575 (2013) 122–125.

[71] R. Dennington, T. Keith, J. Millam, GaussView, version 5, 2009.

[72] R.P. Gangadharan, S.S. Krishnan, Natural Bond Orbital (NBO) population analysis of 1-azanapthalene-8-ol, Acta Phys. Pol., A 125 (1) (2014).

[73] I. Sıdır, Y.G. Sıdır, M. Kumalar, E. Taşal, Ab initio Hartree–Fock and density functional theory investigations on the conformational stability, molecular structure and vibrational spectra of 7-acetoxy-6-(2, 3-dibromopropyl)-4, 8-dimethylcoumarin molecule, J. Mol. Struct. 964 (1–3) (2010) 134–151.

[74] K. Jug, ZB Maksic in the Meaning and Distribution of Atomic Charges in Molecules, Part 3: Molecular Spectroscopy, Electronic Structure and Intermolecular Interactions Ed.: ZB Maksic, 1991.

[75] R.C. Sims, M.R. Overcash, Fate of polynuclear aromatic compounds (PNAs) in soil-plant systems. Residue Reviews, Springer, New York, NY, 1983, pp. 1–68.

[76] B. Silvi, A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature 371 (6499) (1994) 683–686.

[77] R.F.W. Bader, Atoms in molecules: a Quantum Theory Oxford University, Press: Oxford, n° UK, 1990.

[78] G. Saleh, C. Gatti, L.L. Presti, Non-covalent interaction via the reduced density gradient: independent atom model vs experimental multipolar electron densities, Comput. Theor. Chem. 998 (2012) 148–163.

[79] E.R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A.J. Cohen, W. Yang, Revealing noncovalent interactions, J. Am. Chem. Soc. 132 (18) (2010) 6498–6506.

[80] Z. Jia, H. Pang, H. Li, X. Wang, A density functional theory study on complexation processes and intermolecular interactions of triptycene-derived oxacalixarenes, Theor. Chem. Acc. 138 (9) (2019) 1–11.

[81] P. Kolandaivel, G. Praveena, P. Selvarengan, Study of atomic and condensed atomic indices for reactive sites of molecules, J. Chem. Sci. 117 (5) (2005) 591–598.

[82] Z. Demircioğlu, C.A. Kaştaş, O. Büyükgüngör, Theoretical analysis (NBO, NPA, Mulliken Population Method) and molecular orbital studies (hardness, chemical potential, electrophilicity and Fukui function analysis) of (E)-2-((4-hydroxy-2-methylphenylimino) methyl)-3-methoxyphenol, J. Mol. Struct. 1091 (2015) 183–195.

[83] T.J. Beaula, I.H. Joe, V.K. Rastogi, V.B. Jothy, Spectral investigations, DFT computations and molecular docking studies of the antimicrobial 5-nitroisatin dimer, Chem. Phys. Lett. 624 (2015) 93–101.

[84] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, P. E. Bourne, The protein data bank, Nucleic. Acids. Res. 28 (1) (2000) 235–242.

[85] A.R. Katritzky, L. Mu, V.S. Lobanov, M. Karelson, Correlation of boiling points with molecular structure. 1. A training set of 298 diverse organics and a test set of 9 simple inorganics, J. Phys. Chem. 100 (24) (1996) 10400–10407.

Full Length Article

Structural and spectroscopic investigation of 1-acetyl-2-(4-ethoxy-3-methoxyphenyl) cyclopropane and its NLO activity

D.P. Lydia Renj [a,h], R. Racil Jeya Geetha [b,h], J. Jeni James [c,h], M. Amalanathan [d,h,*],
M. Sony Michael Mary [e], Zoran Ratković [f], Jovana Muškinja [g]

[a] Research Scholar, Register Number: 19223112132013, Department of Physics &Research Centre, Nesamony Memorial Christian College, Marthandam - 629165, Tamil Nadu, India
[b] Department of Physics &Research Centre, Nesamony Memorial Christian College, Marthandam, Tamil Nadu 629165, India
[c] Research Scholar, Register Number: 21113102132001, Department of Physics & Research centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai-629 153, Tamil Nadu, India
[d] Department of Physics & Research centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai, Tamil Nadu 629 153, India
[e] Department of Physics, AnnaiVelankanni College of Engineering, Azhagappapuram, Kanyakumari, Tamil Nadu 629 401, India
[f] Department of Chemistry, Faculty of Sciences, University of Kragujevac, RadojaDomanovića 12, Kragujevac 34000, Serbia
[g] Department of Sciences, Institute for Information Technologies, University of Kragujevac, Jovana, Cvijića bb, Kragujevac 34000, Serbia
[h] Affiliated to ManonmaniamSundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu 627 012, India

ARTICLE INFO

ABSTRACT

To identify promising compounds and to develop a potent non-linear optical material, the molecule 1-acetyl-2-(4-ethoxy-3-methoxyphenyl) cyclopropane (AEMC) was selected. FTIR and FT-Raman spectroscopy techniques were employed to predict the functional groups and vibrational modes of AEMC. Gaussian 09 W software was utilised to analyse the parameters of the optimised title compound. Reactive sites were forecasted using MEP plots. To clarify the chemical significance of the molecule, ELF and LOL are utilised. Furthermore, the presence of interactions within the molecule is confirmed by RDG analysis. The strong and weak hydrogen bonds between the non-bonding atoms of AEMC are studied with the aid of AIM analysis. Additionally, the material's capacity to produce non-linear effects (NLO) was ascertained by examining the linear polarizability and first order hyper polarizability values.

1. Introduction

Nonlinear optical (NLO) materials respond to incident photon's electric field in a nonlinear approach, that may culminate in an array of optical phenomena including the creation of new light frequencies or changes to the material's optical characteristics [1]. Nonlinear optical materials have been applied in various fields such as atomic, molecular, solid-state physics, surfaces interface sciences, materials science, medicine, chemical dynamics, and biophysics. Due to recent developments in laser technology, Non-linear optics (NLO) has emerged as a prominent area of study. It is critical to develop non-linear optical materials with improved response. To create materials with improved non-linear optical properties, it is crucial to have a theoretical understanding of the factors impacting those properties [2].

Cyclopropane and its derivatives are unique amongst the cyclic

carbon compounds due to their unique structure, spectra, and chemical properties [3]. Cyclopropane is not only used in natural products, but it's also a common structure in drugs or drug development projects that are currently on the market [4]. The use of cyclopropane derivatives offers a unique opportunity to create structures with unparalleled synthetic potential. Furthermore, both natural and synthetic compounds with simple functions are imbued with a wide range of biological properties, including enzyme inhibition, antibiotics, antivirals, antitumor, and neurochemicals [5].

The past few decades have seen a significant advancement in bioinformatics and structural biology, with the advancement of computational tools for genomic data analysis, which has enabled the development of novel approaches for vaccine design [6]. The DFT calculation method under the B3LYP/6–311 $++G$(d, p) level of theory was utilised in this work to obtain the most stable structure in the

---

ground state [7]. Presently, the ab-initio community acknowledged the DFT method as a prevalent post-HF approach and precision with regards to the estimation of several molecular properties [8]. The best approach for studying hybrid materials theoretically is DFT as it impacts a good equilibrium between calculation time and result precision [9].In light of its significance in numerous chemical and biological processes, intra- or intermolecular proton transfer has drawn an abundance of attention in recent years [10].

In the search for promising compounds, 1-acetyl-2-(4-ethoxy-3-methoxyphenyl) cyclopropane (AEMC) is chosen as the title molecule for the study. This work set out to analyse the AEMC molecule's conformational analysis, energies of the molecular orbital, IR and Raman spectra theoretically and experimentally. The theoretical and experimental spectra of AEMC molecules have not been extensively studied to the fullest extent. Thus, this investigation was conducted in order to attain an extensive comprehension of the vibrational spectrum of AEMC molecules and to classify the different modes with higher wave number precision. By NBO (Natural Bond Orbital Analysis) the stability of AEMC is determined. The energy gap is determined by Frontier Molecular Orbital Analysis (Molecular Orbital Analysis). The electron distribution and reactive sites on surface of the analysed compound were determined by Molecular Electrostatic Potential (MEP), Electron localization function (ELF) and Localized Orbital locator (LOL).

2. Materials and methods

The AEMC was sourced from Sigma Aldrich Company with a purity of 99 % and was utilised without any additional purification. In the evacuation mode, the title compound's FTIR spectrum was recorded in a range of 4000–400cm^{-1}at a resolution of 1.0 cm^{-1}on a PERKIN ELMER FTIR spectrometer using KBR pellet technique. In the pure mode, the title molecule's FT-Raman spectra were recorded at 4000–100 cm^{-1} using an ND: YAG laser at 100 mW at a resolution of 2 cm^{-1} on a BRUCKER RFS 27.

3. Computational details

The present work's computations were carried out at B3LYP levels using the Gaussian 09 W [11] package program in conjunction with the 6–311++ G (d, p) basis set function of the density functional theory (DFT) through gradient geometry optimisation [12]. In order to comprehend intermolecular delocalisation or hyper conjugation,

Natural Bond Orbital (NBO) calculations were performed at the DFT/B3LYP level using the NBO 3.1 program as implemented in the Gaussian 09W package. To shape the Highest Occupied Molecular Orbital (HOMO), Lowest Unoccupied Molecular Orbital (LUMO) and Molecular Electrostatic Potential (MEP) the visualisation program of Gauss view 5.0.8 [13] was used. The VEDA4 program, developed by Jamroz, [14] was utilised to perform in-depth vibrational mode assignments through the use of percentage potential energy distribution (PED) analysis. The AIM, ELF, LOL, and NCI analyses [15–17] using Multiwfn software [18] were used to further look into the hydrogen bonding nature, and the VMD software package [19] was used to visualise the isosurfaces.

4. Result and discussion

4.1. Optimised parameter

Gaussian 09 W program package was used to optimize the molecular structural parameters of the title molecule, and to visualize the optimised molecular structure Gauss View was utilised.Fig.1 illustrates the optimised structure. Supplementary Table.1 displays the structural parameters that were determined using the DFT method. The bond lengths (C_1—C_{16}), (C_{10}—C_{11}) and (C_{12}- C_{13}) are larger than the C—C bond owing to the electron transfer from lone pair oxygen to the carbon atom. The presence of a hydrogen atom linked to the carbon atom accounts for the bond lengths (C_7–C_8) and (C_4–C_5) being shorter than the actual bond length. As a result of the single bond, it is found that (C_1—H_{14}), (C_1—H_{15}), and (C_{13}—H_{35}) are larger. 1.234 Å is the reported value that

Table 1
Calculated energy values of the title molecule using B3LYP++ G(d,p) basic set.

Molecular Properties	Mathematical Description	Energy(eV)
E_{HOMO}	Energy of HOMO	−8.1551
E_{LUMO}	Energy of LUMO	−4.4190
Energy Gap	$\Delta E_g = E_{HOMO}$- E_{LUMO}	3.7361
Ionization Potential (IP)	IP = -E_{HOMO}	8.1551
Electron Affinity (EA)	EA= -E_{LUMO}	4.4190
Electronegativity (χ)	χ = - 1/2(E_{LUMO} + E_{HOMO})	6.287
Chemical potential (μ)	μ = 1/2(E_{LUMO} + E_{HOMO})	−6.287
Global Hardness(η)	η= 1/2(E_{LUMO}–E_{HOMO})	1.8680
Softness (S)	S = 1/2η	0.2676
Electrophilicity index (ω)	ω = μ^2/2η	10.5797
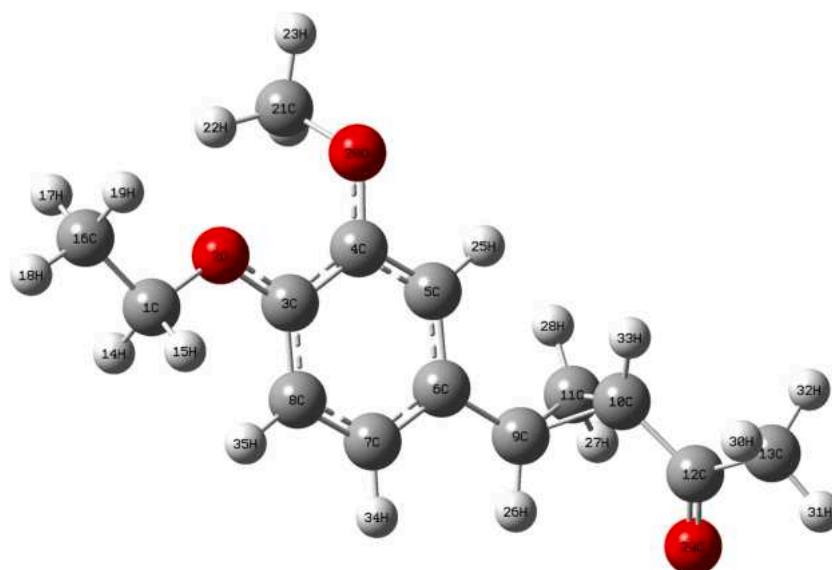
Fig. 1. Optimized structure of 1-acetyl-2-(4-ethoxy-3-methoxyphenyl) cyclopropane (AEMC).

corresponds to the calculated C = O bond length of 1.2172 Å [20]. Charge transfer within the molecule is the reason for the C_{12}—O_{29} (1.2172 Å) bond's shortening from its normal value. The values of (C_4—C_5—C_6) and (C_6—C_7—C_8) exceed 120°, implying the existence of hyper conjugate interaction. A bond angle of (C_6—C_5—H_{25}) greater than 120° signifies the delocalisation of charge within the methyl group. Since methyl groups are typically associated with electron-donating substitution, the bond lengths C_{10}–C_{11} increase with the value of 1.5151 Å Conversely, the bond lengths of C_1—O_2 (1.4307 Å) and C_4—O_{20} (1.373 Å) both increase as a result of the methyl group that donates electrons being attached to the carbon atoms of C_1 and C_4. It is also well known that DFT optimised bond lengths are usually longer, due to the inclusion of electron correlation [21].The methylene group exhibits a greater degree of tilt from the para substituted phenyl ring, as seen by the torsion angles C_8—C_3—C_4—O_{20} = −177.0°, C_3—C_4—O_{20}—C_{21} = −69.3°, C_6—C_5—C_4—O_{20} = 176.6°, and C_5—C_4—O_{20}—C_{21} = 113.1°

As the cyclopropane is attached to the C_9 carbon atom in the phenyl ring, the bond lengths of the C_6—C_7 (1.394 Å) and C6—C7 (1.402 Å) carbon atoms are slightly longer than those of the other carbon atoms in the phenyl ring. It is evident that there is hyperconjugation (n→σ*) between the lone-pair electrons of the O_{29} atom and the C_{10}—C_{12} and C_{12}—C_{13} bonds by lengthening the bond distances C_{10}—C_{12} (1.4909 Å) and C_{12}—C_{13} (1.5147 Å) and contracting the internal bond angle C_{10}—C_{12}—C_{13} (116.3°).The dihedral angles for the title molecule C_1—O_2—C_3—C_4 (−175.9°), C_3—C_4—O_{20}—C_{21} (−69.3°), and C_1—O_2—C_3—C_8 (2.147°), C_5—C_4—O_{20}—C_{21} (113.1°) shows folded conformations of the molecule.

The geometric and electronic structure of the molecules can be significantly influenced by the type of substituent, which can either donate electrons (CH_3) or accept them (Cyclopropane). The charge distribution on the ring's carbon atom provides an elucidation for slight variations in bond length in terms of the substitution effect [22]. The ground state energy of the optimised structure is obtained as −769.955657 Hartree and the dipole moment is 3.2340 Debye.

4.2. Natural bond orbital

The DFT level of calculation is utilised to examine the various second order interactions between filled orbitals from one subsystem and empty orbitals from another subsystem, which serves as a measure of the hyperconjugation or delocalization of the electrons [23]. In order to attain more extensive understanding of the molecular electronic structure, the principal natural orbital interactions of the molecule were evaluated with the aid of the Non-Binding Organization (NBO) 5.0 program [24]. The strong intramolecular hyperconjugative interactions of the molecule are unveiled by the second-order perturbation theory analysis of Fock matrix in the NBO basis, which are presented in Supplementary Table 2.

Intramolecular Hyperconjugative Interactions (ICTs) are formed by orbital overlap between the orbitals of the σ and π C—C bond and σ* and π* C—H, C—C bond orbitals of the two hydrogen bond orbitals. When ICT is induced, it stabilises the system. The most important feature of the NBO of a hydrogen bond system is the transfer of charge between the lone pair of protons (proton acceptor) and the antibonds (proton donor). The NBO result validate the existence of hydrogen bonds within molecules due to the interaction between the lone oxygen pair and the antibond orbitals of the two molecules, LP_2O_2 – σ*(C_1 -H_{14}), LP_2O_2 – σ * (C_1—H_{15}), respectively. The C—H bond gets stronger and weaker as a result of rehybridising [25], Which is disclosed by the lower value of ED 0.02683e and 0.0263e in the σ*(C_1 -H_{14}) and σ*(C_1—H_{15}) orbitals respectively. The blue shift of hydrogen bonding due to improper bonding is predicted by the NBO results with optimised geometric parameters and the corresponding up shift of C—H stretching wavelength (C—H = 3054 cm^{-1}).The higher the E (2) value, the more intense the interactions between the electron donors and the electron acceptors are Intramolecular Hyperconjugative Interactions (IGI) are formed by

orbital overlap between the orbitals of the electron donor and the electron acceptor. The orbital overlap between the orbital of the donor and the bond orbitals of the π(C—C) and π*(C—C) bond results in Intramolecular Charge Transfer (ICT) which stabilises the system. The phenyl ring exhibits strong electron delocalisation at conjugated bonds ED at π bonds (1.67–1.97 e) and π* bonds (0.35–0.39 e), resulting in stabilisation of (≈ 17–20 kJ/mol) energy. When NBO is conjugated with π(C_6—C_7) → π*(C_4—C_5), it results in a huge amount of stabilisation (20.02 KJ/mol) for the title molecule. This strong stabilization indicates greater delocalization. The presence of a strong interaction between the lone O_2 orbital containing a π type electron pair and the neighbouring phenyl ring antibonding orbital of a pair of π*(C_3—C_8) has been observed. This interaction causes the lone pair's orbital occupancy to be lower than the other orbital occupancy, and can cause O_2 to be hyper-conjugated with the phenyl ring.

Intramolecular C–H…O hydrogen bonding is induced by the overlap of the orbital LP_2O and σ* (C–H), resulting in Intramolecular Charge Transfer (ICT) which stabilizes H-bonded systems. As a result of this interaction, the C—H anti-bonding orbital has an increased electron density, thereby reinforcing the C—H bond. C–H…O intramolecular hydrogen bonding is also confirmed between lone pair LP_2O_2, $LP_2O_{29}$and antibonding orbitals σ* (C_{14}–H_{15}), σ* (C_{12}–H_{13}) with stabilisation energy of 19.85 kcal/mol. This confirms the intrinsic C–H…O hydrogen bonding.

The atom label and the hybrid label, which indicates the hybrid orbital (sp^x) composition (s character and p character), of the title molecule, were identified and tabulated in Supplementary table. 3. The presence of a p character of ≈100% was observed in atoms of π bonding in π (C_3–C_8), π (C_4–C_5) and π (C_6–C_7. The occupancy of Nucleobases in the molecule is indicative of their profound dependency on the chemical system. i.e.), the bonding orbital of C_3- C_8 with 1,6852 electrons contains 99.99% C_3 characters in a $sp^{1.00}$hybrid, while the corresponding C8 hybrid orbital contains 99.95%. C_4—C_5 bonding orbital (1.7034) contains 99.14% of C4 in $sp^{1.11}$and 99.75% of C_5 in $sp^{1.09}$. C_6—C_7 bonding (1.6746) contains 99.96% of C_6 in $sp^{1.00}$and 99.97% of C_7 in hybrid orbital.C_9—C_{10} bonding orbitals possessed 81.05% C3 character in a $sp^{4.29}$ hybrid and 80.12% C8 character in a $sp^{4.04}$ hybrid orbital with 1.8900 electrons.

Natural atom hybrid orbital energies possess a high p-character and low occupation number. Natural hybrid orbital energies are equal to −0.3999 a.u and −0.3199 a.u respectively. The hybrid orbital energies of the p-character are found to be ~99.9% and its lower occupancy numbers are ~1.8470 and 1.889 e respectively. Conversely, LP_1O_{20} and LP_2O_{20} possessed with occupancy number of 1.9511 and 1.9081e with energies of −0.4992 and −0.3556 a.u.

4.3. Mulliken charge analysis

Mulliken atomic charge distribution is a major determinant of a molecule's dipole moment and polarisability. It is also influential in the electronic structure of the molecule and the vibrational modes of the molecule. This distribution has been utilised to elucidate the electrostatic potential equalisation process and model of the electrostatic potential external to the molecular surface [26]. These elements play a critical role in the qualitative analysis of both inorganic and organic reactivities [27]. The atomic charge distributions were determined by calculating the electron number of each atom according to the basis set. The total atomic charge is calculated using Mulliken's population analysis with an optimised geometry, and the natural charges are calculated using natural bond orbital analysis. These are shown in Supplementary Table 4.The graphical representation of the analysis is shown in Supplementary fig. 1.

The distribution of carbon atoms in the compound indicates that the positively charged carbon atoms are those of the substituted carbon group (C_6, C_8, C_9, C_{10} and C_{12}) and the remaining carbon atoms are negatively charged. This is indicative of the existence of a positive

charge in the carbon atom, which is attributed to the delocalisation of electrons from both the oxygen and the carbon atoms of the methyl group. C_{10} and C_{12} show the transfer of charge from LP_2O_{29} to $\sigma^*(C_{10}-C_{12})$ (18.82 kJmol^{-1}) and $\sigma^*(C_{12}-H_{13})$ (19.85 kJ mol^{-1}). All hydrogen atoms bewitched a positive charge. In addition, the carbon atom C_{11} (- 0.8047 e) has a higher electronegativities than other carbon atoms due to the fact that it is surrounded by two positively charged hydrogen atoms. The results indicate that the oxygen atom O_{29}, has more negatively charged atoms, while the hydrogen atom H_{26}, has more positively charged atoms. This implies that the hydrogen atom is bonded to the oxygen atom, and the oxygen atoms are all electron acceptors, resulting in the transfer of charge from hydrogen to oxygen.

4.4. HOMO-LUMO

The excitation energies can be calculated in a variety of ways. The most basic method entails the difference between a neutral system's highest and lowest occupied molecular orbital values (HOMO and LUMO respectively). This form is analogous to the frozen orbit approximation, as ground state characteristics are utilised to determine excitation parameters. The HOMO-to-LUMO energy gap is one of the most important parameters in determining the molecular electrical transport characteristics and contributes to the electron conductivity measurement [28]. Calculation of the HOMO, LUMO and energy gap of the title molecule is based on the B3LYP level using 6–311++G(d, p). The wave function analysis suggests that the electronic absorption is associated with the transition from ground to the initial excited state, and is primarily characterized by a single electron excitation between the highest occupied molecular orbit (HOMO) of the molecule and the lowest unoccupied molecular orbital (LUMO) of the molecule. In general, the energy of the LUMO molecule and the energy of the HOMO

molecule, as well as their energy gap, are indicative of the chemical functioning of the molecule, which is a key factor in its stability [29]. The title compound may have more interactive properties due to the evenly distributed electron density of its LUMO and HOMO orbitals [30].

Table 1 shows the title compound's frontier molecular orbitals energies (E_{HOMO}, E_{LUMO}), energy gap (ΔE), electronegativity (χ), the chemical potential (μ), global hardness (η), global softness (S) and global electrophilicity index (ω). Recently, the quantities of reactivity have been determined in order to get brief overview about the different pollutants that are toxic in terms of reactivity and selectivity of the site. The molecular stability of a compound is determined by the overall hardness and flexibility of the compound, which is a reliable indicator of its chemical stability. The HOMO-LUMO energy gap value of a molecule is used to determine whether it is a soft molecule or a hard molecule. Hard molecules are one which has a high energy gap, while a molecule with a less energy gap is classified as a soft molecule. Fig. 2 illustrates the HOMO and LUMO energy levels and their distribution.

The DFT calculation of the title molecule indicates that the HOMO cloud distribution, as well as the LUMO cloud distribution, is centred around the phenyl ring and the methoxy group, with partial coverage of the ethoxy and cyclopropane groups, and the O_{29} atom. The LUMO, on the other hand, concentrates on phenyl ring, cyclopropane and acetyl group. The hardness (η) of the title compound is determined and it is calculated to be 1.886 eV, which means it is a hard material. The energy gap (ΔE) between HOMO and LUMO energy of the compound is 3.7361 eV The lower value of the HOMO-LUMO gap indicates that the molecule can be analysed intramolecularly for the transfer of charge and also indicates that the molecule is bioactive and stable. The value of the chemical potential of the title molecule $\mu = -6.287$ eV is negative, thus indicating the stability of the molecule. Also, band gap energy of the title

$E_{LUMO} = -4.4190$

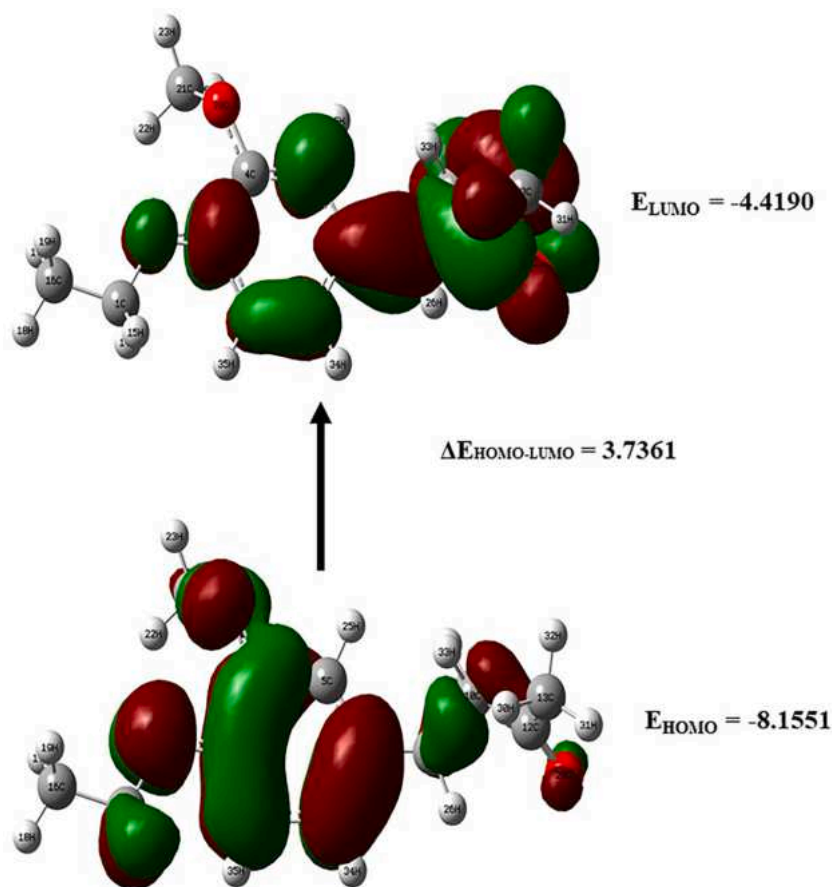$\Delta E_{HOMO-LUMO} = 3.7361$

$E_{HOMO} = -8.1551$

Fig. 2. Energy distribution between HOMO and LUMO.

molecule is found to be low, which indicates the stability of the molecule.

4.5. Molecular electrostatic potential (MEP)

MEP plays an essential role in studying molecular interactions, predicting nucleophile and electrophile sites, molecular clusters, and predication, as well as a wide variety of microscopic properties [31,32]. The electrophilic and nucleophilic attack on the reactive sites for AEMC is predicted by MEP at the B3LYP/6–311++ G(d,p) method. The negative electrostatic potential refers to the proton's attraction due to the high concentration of electrons in the molecule (represented by red on the surface of the MEP), while the positive correlate with the proton's repulsion due to atomic nuclei in areas where there is low electron density and the nuclear charge isn't fully shielded (represented by blue). For a better clarity, reactive sites of the molecule have been identified with different colours [33].The electrostatic potential of the negative atom decreases and the positive region of the other atom becomes less positive when there is an intra- or intermolecular interaction [34] .MEPs allow the evaluation of the capacity of a molecule to interact Electrostatically with a binding site and visualization [35].The mapping scheme was based on a colour-coded system, with the red region indicating the region of the highest electro-negative potential (electrophilic attack), and the blue region indicating the highest electro-positive potential(nucleophilic attack).MEP surface map is shown in Fig. 3.

In this analysis, the negative potential lies close to the carbonyl group and the acetyl group, specifying electrophilic reactions of the title molecule, while the positive potential is centred on the phenyl ring and all of the hydrogen atoms, specifying nucleophilic reaction of the title molecule. These interactions and active sites on the title molecule provide an indication of the biological activity of the compound. The electrostatic potential is found to lie in the range of $-5.393\ e^{-2}$ to $5.393\ e^{-2}$.

4.6. Vibrational analysis

The spectroscopic properties of AEMC have not been extensively investigated to the fullest extent. Since AEMC has 35 atoms, it has 99 modes of vibration (3N-6) in addition to three translational modes and

three rotational modes. The vibrational wave numbers of FT-IR and FT-Raman were theoretically determined theoretically by the DFT method using the B3LYP/6–31++G (d,p) basis set. Calculated frequencies are naturally harmonic while experimental frequencies are unharmonic in nature. This is the primary reason for the deviation in methylene group. The deviation in other groups is due to the approximations that were carried out in the computational method. To address these systematic errors, the frequencies calculated were scaled by an appropriate scaling factor (0.961) to reduce the difference in theoretical and experimental performance.Combined experimental and theoretical FTIR and FT-Raman spectra are depicted in Fig. 4 and Fig. 5. Table.2 shows the calculated scaled wave numbers with the experimental FTIR and FT-Raman wave numbers of the title compound.

4.6.1. Phenyl ring vibration
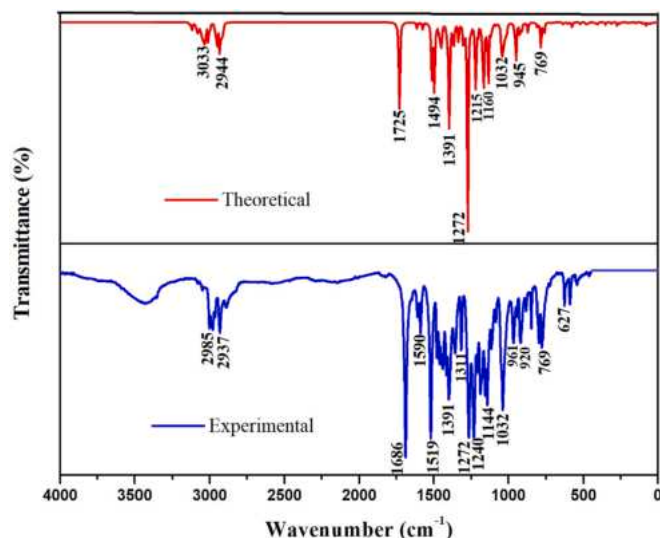AEMC is made up of a tri-substituted benzene ring. Carbon stretching

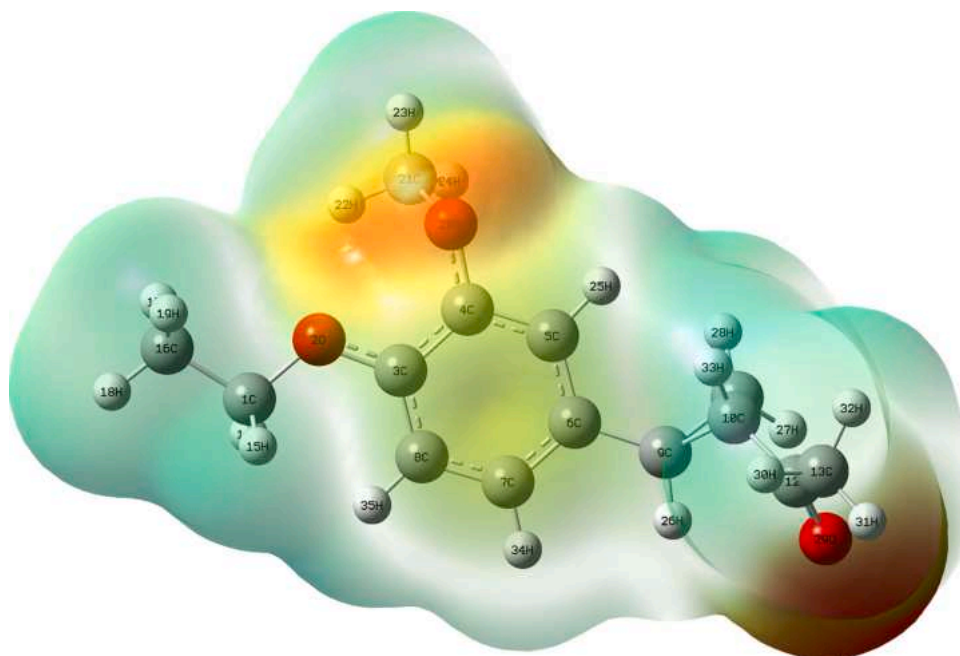**Fig. 4.** Combined theoretical and experimental FTIR spectra of the title molecule.

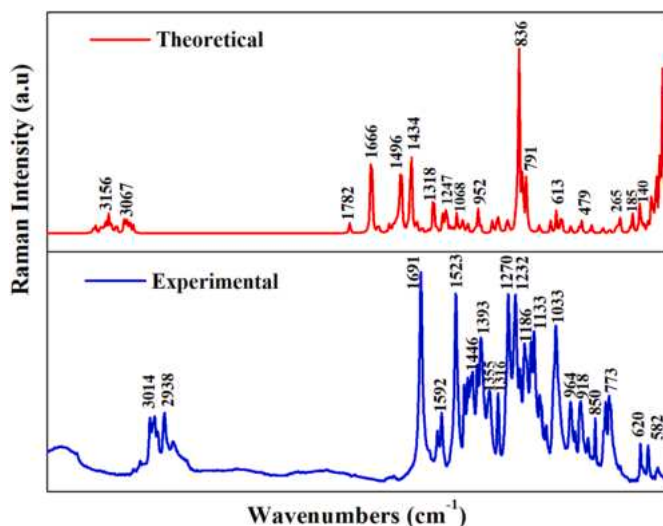**Fig. 3.** Molecular Electrostatic Potential surface map of AEMC.

Fig. 5. Combined theoretical and experimental FT-Raman spectra of the title molecule.

is usually very important in vibrational spectra for benzene derivatives. The phenyl group exhibits stretching modes in the C—C range, ranging from 1650–1100 cm^{-1} [36]. In the present molecule the peaks perceived at 1592 cm^{-1}, 1393 cm^{-1} and 1316 cm^{-1} are assigned to Raman spectrum and the peaks at 1590 cm^{-1}, 1391 cm^{-1} 1,311 cm^{-1}, 1270 cm^{-1} and 920 cm^{-1} are assigned to IR spectrum to C—C stretching vibration. The electro-negativity of the oxygen atom causes slight variations in frequency for these modes. The phenyl ring's bending modes were characterized by medium-intensity bands at 1592 cm^{-1} for Raman and 1590 cm^{-1} for IR spectrum. The other vibrations observed in the fingerprint area correspond closely to the spectra observed.

4.6.2. CH$_3$ vibrations

When the electron-donor methyl group (CH$_3$) is directly connected to the oxygen atom, the stretching and bending bands of the C–H pair can undergo energetic changes due to the electronic interactions involved. This results in the expansion of the O—CH$_3$ vibrational bands across a wider area than the C—CH$_3$ group does. In the present work, a methoxy (OCH$_3$–) group attached to the phenyl ring of the title compound.The CH$_3$ stretching predicted at 2983 cm^{-1} lies under the observed values at 3014 cm^{-1}, 2938 cm^{-1} in Raman and 2985 cm^{-1}, 2937 cm^{-1} in IR spectra, respectively. The wavenumber is deviated due to the slight deviation in the orientation of the methoxy group. AEMC contains two methyl groups in its side chain. Vibrations of CH$_3$ in the form of stretching and deformations are relatively localised and result in satisfactory group wavenumbers. The most stable vibrations in the spectrum are C—H stretching vibrations. Stretching vibrations for the CH$_3$ and C—H groups gives a series of sharp bands from 2980 to 2875 cm^{-1} [37, 38].For the title compound, at 2985 cm^{-1} a medium peak is observed, which corresponds to the CH$_3$ stretching mode vibration.

4.6.3. C═O, C—O vibrations

Generally, Ketones, Aldehydes, and Amides demonstrate IR absorption at a range of 1750–1650 cm^{-1}. Carbonyl group vibrations have resulted in distinct bands in the vibrational spectrum, which have been extensively studied. The increase in intensity of these bands can be due to the induction of conjugation or hydrogen bonding. The combination of conjugation, ring size and hydrogen bonding, as well as steric and electronic reactions, can often lead to substantial changes in the C═O absorption frequency [39].

The intensity of the carbonyl vibration bands in ketones is typically very high and is predicted to be in the range of 1715–1680 cm^{-1}. The double bond in carbon–oxygen is formed by pπ–pπ between oxygen and

Table 2

Scaled and Observed wave frequencies with PED contribution of AEMC using method B3LYP++ G(d,p) basic set.

Wave Numbers Experimental Values		Scaled values	Assignments with PED (>10 %)
FT-IR	FT-Raman		
–	–	3134	$\upsilon C_{11}H_{27}$ (99)
–	–	3119	υC_8H_{35} (95)
–	–	3093	υC_5H_{25} (99)
–	–	3084	υC_7H_{34} (92)
–	–	3074	$\upsilon C_{10}H_{33}$ (89)
–	–	3068	υC_9H_{26} (88)
–	–	3062	$\upsilon C_{13}H_{31}$ (85)
–	–	3049	$\upsilon C_{21}H_{22}$ (91)
–	–	3048	$\upsilon C_{11}H_{27}$ (98)
–	–	3038	$\upsilon C_{16}H_{17}$ (89)
–	–	3031	$\upsilon C_{16}H_{17}$ (90)
–	3014	3011	$\upsilon C_{21}H_{22}$ (89)
–	–	3001	$\upsilon C_{13}H_{30}$ (99)
2985	–	2962	$\upsilon C_{16}H_{17}$ (98)
–	–	2953	υC_1H_{14} (97)
–	–	2945	$\upsilon C_{13}H_{31}$ (14) + $\upsilon C_{13}H_{30}$ (86)
2937	2938	2933	$\upsilon C_{21}H_{22}$ (95)
–	–	2916	υC_1H_{14} (98)
1686	1691	1731	$\upsilon O_{29}C_{12}$ (87)
–	1666	1613	υC_4C_5 (55) + $\beta H_{23}C_{21}O_{20}$ (15)
1590	1592	1574	υC_3C_8 (53) + $\beta C_3C_4C_5$ (13)
1519	1523	1514	$\beta H_{14}C_1H_{15}$ (36)
–	–	1498	$\beta H_{14}C_1H_{15}$ (56) + $\beta H_{17}C_{16}H_{19}$ (12)
–	–	1490	$\beta H_{22}C_{21}H_{24}$ (78)
–	–	1480	$\beta H_{14}C_1H_{15}$ (16) + $\beta H_{17}C_{16}H_{19}$ (66)
–	–	1466	$\beta H_{23}C_{21}H_{24}$ (68)
–	–	1464	$\beta H_{18}C_{16}H_{19}$ (64) + $\beta H_{23}C_{21}H_{24}$ (11) + $\tau H_{17}C_{16}C_1O_2$ (12)
–	–	1455	$\beta H_{27}C_{11}H_{28}$ (27)
–	–	1454	$\beta H_{30}C_{13}C_{12}$ (27) + Out $C_{13}H_{30}C_{12}H_{31}$ (63) + Out $C_{21}H_{22}C_{24}H_{23}$ (25)
–	–	1450	$\beta H_{27}C_{11}H_{28}$ (29) + Out $C_{21}H_{22}C_{24}H_{23}$ (37)
–	1446	1445	$\beta H_{30}C_{13}C_{32}$ (78)
–	–	1415	υC_4C_5 (19) + $\beta H_{27}C_{11}H_{28}$ (16) + $\beta H_{25}C_5C_4$ (10) + Out $C_{21}H_{22}C_{24}H_{23}$ (10)
–	–	1403	Out $C_{16}H_{17}H_{19}H_{18}$ (56)
1391	1391	1395	υC_4C_5 (10) + $\beta H_{26}C_9C_6$ (24) + $\beta H_{27}C_{11}H_{28}$ (12)
–	–	1374	$\beta H_{15}C_1C_{16}$ (32) + $\beta H_{18}C_{16}C_1$ (11) + Out $C_{16}H_{17}H_{19}H_{18}$ (28)
–	1355	1364	Out $C_{13}H_{31}H_{32}H_{30}$ (75)
–	–	1334	$\beta H_{26}C_9C_6$ (46)
1311	1316	1301	υC_3C_8 (45)
1270	–	1274	υC_5C_6 (17) + υO_2C_3 (18) + $\upsilon O_{20}C_4$ (12) + $\beta H_{15}C_1C_{16}$ (26) + $\tau H_{14}C_1C_{16}H_{19}$ (28) + $\tau H_{17}C_{16}C_1O_2$ (15)
1272	–	1273	$\beta H_{15}C_1C_{16}$ (11)
–	–	1253	$\beta H_{25}C_5C_4$ (56)
1240	1232	1220	υO_2C_3 (17) + $\beta H_{23}C_{21}O_{20}$ (17)
–	–	1204	υC_9C_{11} (18) + $\upsilon C_{10}C_{11}$ (15) + $\beta H_{26}C_9C_6$ (12)
–	1186	1181	$\beta H_{23}C_{21}O_{20}$ (61)
–	–	1167	$\upsilon C_{10}C_{11}$ (13) + $\upsilon C_{12}C_{13}$ (15) + $\beta H_{26}C_9C_6$ (13) + $\beta H_{31}C_9C_6$ (12)
–	–	1157	$\beta H_{18}C_{16}H_{19}$ (14) + $\tau H_{14}C_1C_{16}H_{19}$ (17) + $\tau H_{17}C_{16}C_1O_2$ (48) + $\tau H_{19}C_{16}C_1O_2$ (10)
1144	–	1146	υC_6C_9 (13) + $\beta H_{14}C_1H_{15}$ (12)
–	–	1140	$\beta H_{23}C_{21}H_{24}$ (12) + Out $C_{21}H_{22}C_{24}H_{23}$ (82)
–	1133	1134	υC_1C_{16} (10) + $\beta H_{18}C_{16}C_1$ (12) + $\tau H_{27}C_{11}C_{10}C_9$ (36) +
–	–	1108	υC_1C_{16} (14) + $\beta H_{18}C_{16}C_1$ (22) + $\beta C_{10}C_{12}C_{13}$ (10) + $\tau H_{27}C_{11}C_{10}C_9$ (12)
–	–	1086	$\beta H_{27}C_{11}C_{10}$ (34) + Out $C_{10}C_{11}C_{12}H_{33}$ (36) + $\tau H_{27}C_{11}C_{10}C_9$ (11)
–	–	1051	$\tau H_{27}C_{11}C_{10}C_9$ (78)
–	–	1043	υC_1C_{16} (64)
1032	1033	1031	υC_1C_{16} (15) + $\upsilon O_{20}C_{21}$ (17)
–	–	1028	$\beta H_{30}C_{13}C_{12}$ (39) + Out $C_{13}C_{30}C_{12}C_{31}$ (17) + $\tau C_9C_{11}C_{10}C_{12}$ (13)
–	–	1013	$\upsilon O_{20}C_{21}$ (40) + $\beta H_{26}C_9C_6$ (10) + $\tau H_{27}C_{11}C_{10}C_9$ (11)

(continued on next page)

Table 2 (*continued*)

Wave Numbers Experimental Values		Scaled values	Assignments with PED (>10 %)
FT-IR	FT-Raman		
961	964	950	$\upsilon C_{10}C_{12}(13) + \upsilon C_{10}C_{11}(10) + \beta H_{31}C_{13}C_{12}$ (33)
–	–	925	$\beta H_{27}C_{11}C_{10}$ (11) + $\tau H_{27}C_{11}C_{10}C_9(20)$ + Out $C_{10}C_{11}C_{12}H_{33}(19)$
920	–	923	$\upsilon C_5C_6(13) + \upsilon C_{12}C_{13}(13)$
–	918	912	$\upsilon C_1C_{16}(33) + \beta H_{18}C_{16}C_1$ (11)
–	–	894	$\tau H_{34}C_7C_8H_9(10)$
–	850	864	$\upsilon C_9C_{11}(10) + \tau H_{25}C_5C_6C_9(17) + \tau H_{27}C_{11}C_{10}C_9(16)$
–	–	822	$\beta H_{18}C_{16}C_1$ (10) + $\tau H_{34}C_7C_8H_9(16)$ + $\tau H_{14}C_1O_2C_3(11) + \tau H_{18}C_{16}C_1O_2(20)$
–	–	806	$\beta C_9C_{11}C_{10}(17) + \tau H_{34}C_7C_6C_9(41)$
769	773	767	$\upsilon C_{10}C_{12}(19) + \beta H_{27}C_{11}C_{10}$ (14)
–	–	764	$\upsilon O_{29}C_{12}(87) + \upsilon O_2C_3(13) + \beta C_5C_6C_7(15)$
–	–	713	$\tau C_5C_4C_3C_8(46)$
–	–	693	$\upsilon C_6C_9(10) + \beta C_6C_7C_8(30)$
627	620	635	$\tau C_5C_6C_7C_8(43)$
–	–	604	$\beta C_5C_6C_7(30)$
–	582	582	$\beta H_{30}C_{13}C_{12}$ (10) + $\tau C_9C_{11}C_{10}C_{12}(45)$
–	–	576	$\upsilon C_{12}C_{13}(24) + \beta C_{10}C_{12}C_{13}$ (48)
–	–	480	$\beta C_3C_4O_{20}$ (13) + $\tau C_3C_8C_7C_6(10)$
–	–	466	$\beta C_3C_4C_5$ (16) + $\tau C_3C_8C_7C_6(17)$
–	–	424	$\beta C_6C_9C_{11}$ (21) + $\tau C_5C_4C_3O_2(10)$
–	–	410	$\beta C_{10}C_{12}C_{13}$ (11)
–	–	352	$\beta C_{10}C_{12}C_{13}$ (43)
–	–	330	$\beta C_{10}C_{12}C_{13}$ (12) + $\beta C_4O_{20}C_{21}$ (21) + $\tau C_3C_8C_7C_6(28)$
–	–	313	$\beta C_3C_4O_{20}(16) + \beta C_1O_2C_3$ (13)
–	–	275	$\beta C_{10}C_{12}C_{13}$ (18) + $\beta C_4O_{20}C_{21}$ (16) + $\tau C_5C_4C_3O_2(15)$
–	–	257	$\beta C_4C_3C_8(13) + \beta C_3C_4O_{20}(11) + \beta C_7C_6C_9(11)$ + Out $O_{29}C_{10}C_{13}C_{12}(11)$
–	–	249	$\tau H_{17}C_{16}C_1O_2(60)$
–	–	203	$\beta C_9C_{10}C_{12}(23) + \beta C_{10}C_{12}C_{13}$ (12)
–	–	187	$\beta C_9C_{10}C_{12}(14) + \tau H_{24}C_{21}O_{20}C_4(24)$
–	–	152	$\tau H_{24}C_{21}O_{20}C_4(37)$ + Out $O_{29}C_{10}C_{13}C_{12}(15)$
–	–	145	$\tau H_{24}C_{21}O_{20}C_4(37) + \tau C_1O_2C_3C_4(52)$
–	–	135	$\beta C_1O_2C_3(12) + \tau H_{32}C_{13}C_{12}O_{29}(27)$
–	–	127	$\tau H_{32}C_{13}C_{12}O_{29}(63)$
–	–	104	$\tau C_1O_2C_3C_4(47)$
–	–	102	$\tau C_1O_2C_3C_4(41) + \beta C_7C_6C_9(11)$
–	–	82	$\tau C_1O_2C_3C_4(17) + \tau C_1O_2C_3C_4(46)$
–	–	65	$\tau C_3O_2C_1C_{16}(62)$
–	–	54	$\tau C_6C_9C_{11}C_{10}(71)$
–	–	37	$\tau C_1O_2C_3C_4(41) + \tau C_1O_2C_3C_4(17)$
–	–	22	$\tau C_7C_6C_9C_{11}(74)$
–	–	13	$\beta C_9C_{11}C_{10}(17) + \tau H_{34}C_7C_6C_9(41)$

ν-stretching, β – in plane bending, γ -out of plane bending, τ -torsion.

carbon atoms and the electrons of lone pair on oxygen, which also determines the nature of carbonyl group [40]. In the AEMC molecule the frequency of the carbonyl stretching vibration is absorbed as a strong peak at 1691 cm^{-1} (IR) and very strong peak at 1686 cm^{-1}(Raman) and DFT calculations give this mode at 1790 cm^{-1}. The high intensity of the C=O stretching vibrational mode indicates the presence of intramolecular charge transfer within the molecule, resulting in the bioactivity of the molecule [41].

The bands caused by the ester C—O stretching vibration are intense (partly caused by C—C vibration) and are found in the range of 1300–1100 cm^{-1}. The observed bands in IR at 1272 cm^{-1}, 1240 cm^{-1},1032 cm^{-1}and Raman at 1232 cm^{-1},103 cm^{-1} were assigned for C–O stretching vibration. C—OH bending is another form of vibration associated with carboxylic groups. Depending on the presence of monomer, dimer or other hydrogen-bonded species, they are expected to be found in 1150–1450 cm^{-1} ranges [42]. The medium Raman band at 1232 cm^{-1}, weak Raman bands at 1186 cm^{-1},1133 cm^{-1}, 773 cm^{-1} and medium IR band at 1240 cm^{-1}shows substantial O—H bending character which were generated from intermolecular hydrogen bonding

interaction.

4.6.4. C—C vibrations

C—C stretching vibrations are predicted to be in the range of 650 to 1100 cm^{-1} and are not significantly affected by the composition of the substituted compound [43].The C—C stretching vibration of this compound were observed at 1033 cm^{-1}, 918 cm^{-1},850 cm^{-1}, 773 cm^{-1} in FT-IR and 1144 cm^{-1}, 1032 cm^{-1}, 920 cm^{-1},769 cm^{-1}in FT-Raman respectively.

4.6.5. C—H vibrations

It has been entrenched that a carbon-hydrogen (C—H) group is a hydrogen bond donor. Although the interactions between C—H…O are generally regarded as weak, they account for approximately 20–25 % of the total hydrogen bond number, making them the second most significant group [44].Observed C—H bands in plane bending for the title molecule are characterized by a high degree of sharpness; however, the intensity of the bands is weak to medium. The calculated scaled down frequencies 1395, 1273, 1220, 1146, 912, 769and 582 cm^{-1}at B3LYP were attributed to C—H in plane bending vibrations. Interestingly the recorded FT-Raman absorptions at 1393, 918, 773, 582 cm^{-1} and FTIR absorptions at 1391, 1272, 1144, 769 cm^{-1} attributed to C—H in plane bending vibrations relies satisfactorily on theoretical calculated values. A very weak band obtained in FT-Raman at 1355 cm^{-1} is assigned to C—H out of plane bending vibrations.

4.7. AIM

AIM (Atoms in Molecules) is an interesting theory to illustrate molecular topology as it elucidates the bond critical points (BCPs) between adjacent atoms [45,46]. The bond critical points can be localized, allowing for the calculation of various properties at their respective positions in space. The most significant of these properties is the bond critical point charge density (ρBCP), which is the minimum charge density along the bond's length. The properties of BCPs were obtained from the topological parameters like the kinetic energy densities G(r), electron density ρ(r), the eigenvalues (λ1, λ2, λ3), the λ1/ λ3 ratio, the Laplacian $\Delta^2 \rho$(r), the potential V(r) the total energy densities H(r) and the bond energy E. The electron density ρ(r) of hydrogen and its Laplacians $\Delta^2 \rho$(r), can be used to measure the strength of hydrogen interactions. Generally, the high values of $\Delta^2 \rho$(r) and ρ(r) indicate a high degree of interaction strength. Supplementary Figure 2 shows the optimised AIM structure of AEMC. The ellipticity of the molecule calculated using AIM analysis is shown in supplementary table. 5.

AIM results shows that the AEMC is characterized by two BCPs describing a C—O…H type hydrogen bonding interaction. At the hydrogen bond C_{12}—O_{29}…H_{30}, the BCP has an electron density (0.4005 a. u) and a Laplacian value (0.2342 a. u). This positive Laplacian value indicates the loss of electrons along the binding path. The energy of the hydrogen bond C—O…H has been determined by the correlation $E_{HB} = V(r_{BCP})/2$ as outlined by Espinosa et al. [47].AEMC has intramolecular hydrogen bond energy of −18.46 kcal/mol between C_{12}—O_{29}…H_{30}.

4.8. Electron localization function (ELF)

The Electron localization function (ELF) topological analysis is one of the most powerful tools for determining the probability of electron pair localisation. The Electron localization function is a measurement of the electron's spatial location in relation to an electron of equivalent spin. It is used to map a pair of electrons and measure the spatial localization of the corresponding reference electron. [48]. It is an effective way to measure the number of electrons in the nucleus system [49].It also assists in the identification of electron localization in multielectron systems. Supplementary Fig..3 displays a two-dimensional representation of the Electron Localization Function. The scale of the function ranges from 0.0 to 1.0. The colour codes are indicated by a red value for a high

ELF value, a yellow to green colour for a medium ELF value, and a blue colour for a low ELF value. Higher levels of electron localisation in a specific region correspond to higher ELF, while lower ELF corresponds to lower levels of repulsion [50,51].

The region of depletion between the valence shell and inner shell are represented by a blue colour circle surrounding few carbon atoms close to oxygen atoms where lone pair formation can occur. Critical points and their trajectories, chemical bonds and chemically significant regions (reds and oranges) in the vicinity of hydrogen atoms provide a clear indication of the localized bonding, or non-bonding electrons represented by the colour red is shown in ELF map. This is where the electrons are located (high Fermi holes), so the covalent bond must exist at this point, where the lone pair of AEMC atoms is involved.

4.9. Located orbital locator (LOL)

Located orbital locator (LOL) is one of the most common electron localization descriptors used to describe molecular bonds, reactivities and chemical structures. LOL and ELF are similar in that they both depend on kinetic energy density [52].The LOL map is also simpler and clearer than the ELF map. In addition, ELF elucidates the electron pair density, and LOL demonstrates the highest localized orbitals that overlap due to the orbital's gradient. Supplementary fig.4 illustrates a two-dimensional image of localized orbital locators (colour shaded and contour maps), with a scale range between 0.0 and 0.8. The distribution of localised orbital locators is derived from B3LYP/6–311++ G (d, p). Due to the overlapping of localized orbitals, the gradients of localized orbitals increase to the maximum value as indicated by red colour label in colour shade map. [53]

In the LOL colour filled map, the core of hydrogen atom is indicated by a white colour, indicating that the bonds are influenced by a localized orbital (above the upper limit of 0.8).The blue colour circles indicate the impoverished region's valence shell and the inner shells of the few carbon and oxygen atoms respectively. The green colour indicates the majority of the areas where the carbon and hydrogen atoms are covalently bonded.

4.10. Reduced density gradient (RDG)

The Noncovalent Interactions Analysis (NCI) is a visualization tool to visualize the site in the molecule in which the non-covalent interactions occur in the physical space. In numerous chemical, biological, and physical phenomena, weak interactions are important. They have a wide range of application, including the storage of hydrogen for renewable energy sources [54] . The NCI index is based on the Reduced Density Gradient (RDG) and detailed in several research publications [55] . Additionally, it provides the most significant interaction data, such as Van der Waals, hydrogen bond, and steric relationships within the molecule [56]. RDG spikes that are significantly negative then, sign (λ2) ρ are the sign of attractive interactions (e.g. dipole and dipole interactions, hydrogen bonding, etc.). If the value of sign(λ2) ρ is large and positive, the interaction occurs is non-bonding (steric effects).When the values are close to zero, it means that the interactions are very weak, i.e. van der Waals interactions (VdW) [57] . The strength of the bond can be determined from the sign (λ2) ρ of the molecule. For instance, the hydrogen bond has a density which is negative than the Van der Waals interaction, while for positive values, the repulsive zone will be observed. This allowed for the classification of the interactions and their strength.

Scatter plot and reduced density gradient structure of AEMC is show in Fig. 6. In the title molecule Red spike indicates steric repulsion in the centre of phenyl ring in RDG isosurfaces. The red contour from 0.01–0.02 au indicates the repulsive exchange contribution in the RDG scatter chart. This plot shows the steric repulsion between aromatic carbon atoms in the rings.The combination of red and green peaks indicates the presence of a weak non-covalent interaction between H…H, which is also reflected in the RDG graph, where more green peaks are observed between the values of 0.005 and 0.010 and a.u.

4.11. Nonlinear optical (NLO) effects

Investigating the nonlinear optical properties of molecules is critical due to the core functions of frequency shifting, optical modulation, optical switching, optical logic, and optical memory for emerging technologies such as telecommunications, signal processing, and optical interconnections [58]. The total dipole moment (tot), linear polarizability (ij), and first-order hyper polarizability (ijk) from the Gaussian output file have previously been published and discussed in detail [59–61]. The total molecular dipole moment (μ), linear polarizability (α), and the first-order hyper polarizability (β) were calculated by the B3LYP method with 6–311++G(d,p) basis sets to investigate the effect of basis sets on the NLO properties. DFT level is used to compute the
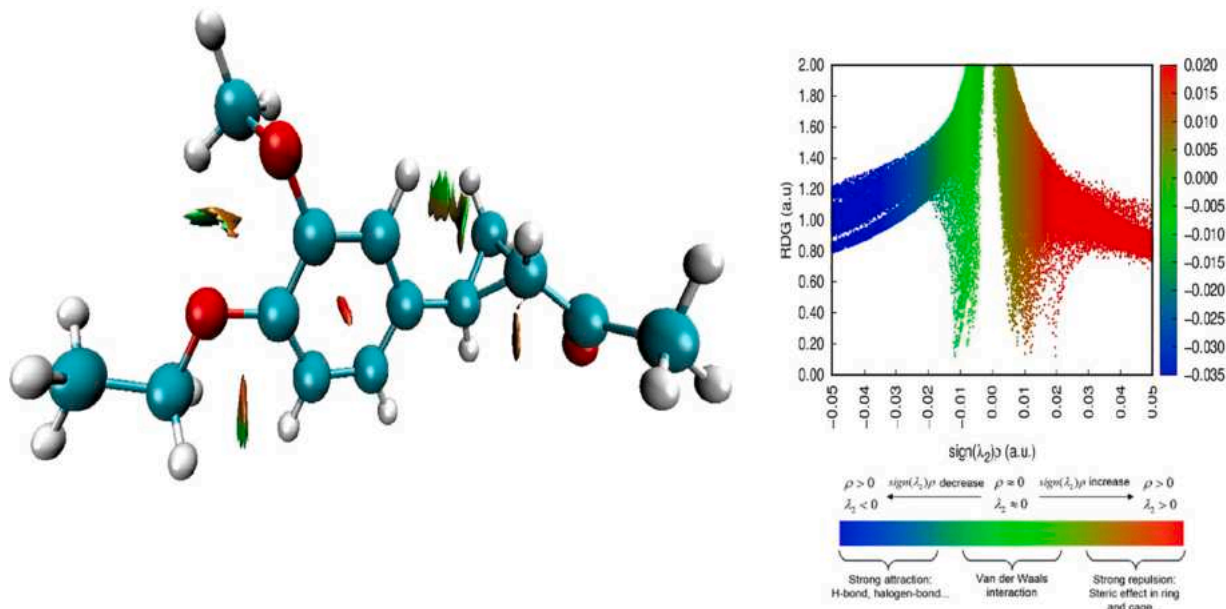

Fig. 6. Reduced Density Gradient structure and scatter plot of the title molecule.

dipole moment of AEMC, Urea, and KDP. Supplementary Table 6 shows the dipole moment components of AEMC, Urea, and KDP, and the total dipole moment is found to be 1.315 Debye, 1.527 Debye, and 5.911 Debye, respectively. The additional polarizability components (xy, xz, etc.) are not required to achieve the isotropic quantity. Supplementary Table 7 shows the polarizability components of AEMC, Urea, and KDP. The static polarizabilities of Urea, KDP, and AEMC are computed using the polarizability component and are found to be $22.29 \times 10{-}23$ e.s.u, $5.04 \times 10{-}24$ e.s.u, and $7.28 \times 10{-}24$ e.s.u, respectively. The dynamic polarizabilities are $13.99 \times 10{-}23$ e.s.u, $6.30 \times 10{-}24$ e.s.u, and $12.65 \times 10{-}24$ e.s.u, respectively. According to the foregoing results, the static and dynamic polarizability of AEMC material is greater than that of the NLO reference materials Urea and KDP. It confirms that heavy intramolecular charge transfer (ICT) improves their linear polarizability more than others. The presence of ICT contributes to an increase in the mean polarisation and also leads to an increase in the longitudinal polarisation of µxx. From these findings, it is inferred that strong charge transfer substitutions are good candidate for broad polarisation systems [62,63].

Supplementary Table 8 lists the 10 hyper polarizability components of AEMC, Urea, and KDP. At the DFT level, the first order hyper polarizability of Urea, KDP, and AEMC is calculated to be $2.184 \times 10{-}29$ e.s.u., $0.78 \times 10{-}30$ e.s.u., and $7.390 \times 10{-}30$ e.s.u., respectively. When compared to Urea and KDP, the values for AEMC are high. The study also suggests that the amount of a molecule's first hyper polarizability is dependent on the availability of a single Pair of nitrogen atom electrons to be conjugated with the Phenyl ring [64]. From the above result it confirms that the increase of polarizability increase the NLO activity of the molecule. The increase of polarizability is due to the increasing of charge transfer interaction inside the molecule, which leads the greater NLO activity of AEMC than Urea and KDP. The theoretical second order hyper polarizability was calculated using Gaussian 09 software. The components of second order hyper polarizability are given in Supplementary Table 9. The calculated average second order hyper-polarizability of AEMC, Urea and KDP are found to be $32.22 \times 10{-}39$ e.s.u, $7.301 \times 10{-}39$ e.s.u and $8.937 \times 10{-}39$ e.s.u respectively.

4.13 S. harmonic generation (SHG) test

In order to confirm nonlinear optical (NLO) property, microcrystalline form of 4DS was packed and placed between two transparent glass slides (sample cell). A fundamental laser beam of 1064 nm from a Nd: YAG laser was made to fall on the sample cell and Second Harmonic Generation (SHG) was confirmed by emission of green light (λ=532 nm). The green output (532 nm) was collected by a photomultiplier tube and finally measured on the storage oscilloscope (CRO) as output voltage. Output signal of 7.36 mV is obtained for the titled compound while the urea and KDP is 3.5 mV and 15 mV respectively. It is concluded that title compound is 0.49 times greater than that of reference KDP and 2.01 times of Urea. Output SHG intensities of Urea, KDP, AEMC and the comparative SHG outputs are given in Supplementary Table 10

5. Conclusion

In this work, for the first time, quantum computation and spectroscopic vibrational analysis was performed on the title compound. NBO analysis determined the interactions between donors and acceptors, as well as the stability of the AEMC compound. Stability of the compound is also shown by the AIM analysis. The analysis of the Mulliken population and the mapping of Molecular Electrostatic Potentials (MEP) provided a more comprehensive picture of the chemical active areas. ELF and LOL evaluated the electron distribution and localization of the title compounds on the surface. Colour-coded plot RDG iso surface indicated the strength of the interactions that result in the structural stability of the molecule. FMOs analysis indicates the density of delocalized electrons

within the molecule and it shows the higher reactivity of the molecule. NLO property of the title compound is confirmed by the increased polarizability of the molecule. As a result the material demonstrated good non-linear optical behaviour and could be used for NLO devices, according to the calculated first order hyper polarizability.

CRediT authorship contribution statement

D.P. Lydia Renj: Writing – original draft, Funding acquisition. **R. Racil Jeya Geetha:** Supervision, Writing – review & editing. **J. Jeni James:** Writing – review & editing, Investigation. **M. Amalanathan:** Conceptualization, Validation. **M. Sony Michael Mary:** Methodology, Formal analysis. **Zoran Ratković:** Resources. **Jovana Muškinja:** Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chphi.2023.100420.

References

[1] L. Dalton, Nonlinear optical polymeric materials: from chromophore design to commercial applications, Polymers Photonics Appl. I (2002) 1–86.
[2] D.F. Eaton, Nonlinear optical materials, Science 253 (1991) 281–289.
[3] A. De Meijere, Bonding properties of cyclopropane and their chemical consequences, AngewandteChemie Int. Edition in Eng. 18 (11) (1979) 809–826.
[4] A. De Meijere, S.I. Kozhushkov, H. Schill, Three-membered-ring-based molecular architectures, Chem. Rev. 106 (12) (2006) 4926–4996.
[5] J. Salaün, Rearrangements involving the cyclopropyl group, Cyclopropyl Group 1 (1987) 809–878.
[6] L. Backert, O. Kohlbacher, Immunoinformatics and epitope prediction in the age of genomic medicine, Genome Med. 7 (2015) 1–12.
[7] M. Medimagh, N. Issaoui, S. Gatfaoui, O. Al-Dossary, A.S. Kazachenko, H. Marouani, M.J. Wojcik, Molecular modeling and biological activity analysis of new organic-inorganic hybrid: 2-(3, 4-dihydroxyphenyl) ethanaminium nitrate, J. King Saud Univ.-Sci. 33 (2021), 101616.
[8] M. Tahenti, S. Gatfaoui, N. Issaoui, T. Roisnel, H. Marouani, A tetrachlorocobaltate (II) salt with 2-amino-5-picolinium: synthesis, theoretical and experimental characterization, J. Mol. Struct. 1207 (2020), 127781.
[9] O. Noureddine, S.G N.Issaoui, O. Al-Dossary, H. Marouani, Quantum chemical calculations, spectroscopic properties and molecular docking studies of a novel piperazine derivative, J. King Saud Univ.-Sci. 33 (2) (2021), 101283.
[10] N.B. Arslan, N. Özdemir, O. Dayan, N. Dege, M. Koparır, P. Koparır, H. Muğlu, Direct and solvent-assisted thione–thiol tautomerism in 5-(thiophen-2-yl)-1, 3, 4-oxadiazole-2 (3H)-thione: experimental and molecular modeling study, Chem. Phys. 439 (2014) 1–11.
[11] M.J. Frisch, Gaussian09, http://www.Gaussian.Com/, (2009).
[12] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. rev. 136 (3B) (1964) 864.
[13] R. Dennington, T. Keith, J. Millam, Gaussview Version 5.0. 8, Gaussian, Inc, wallingford CT 235 (2009).
[14] M.H. Jamroz, Vibrational energy distribution analysis VEDA4 (2004).
[15] R.F.W. Bader, A quantum theory clarendon press, 1994.
[16] P.L.A. Popelier, F.M. Aicken, S.E. O`Brien, Atoms in Molecules, 188, Prentice Hall, Manchester, 2000.
[17] R.F. Bader, Atoms in molecules, Acc. Chem. Res. 18 (1) (1985) 9–15.

[18] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunctionanalyser, J. Comput. Chem. 33 (5) (2012) 580–592.

[19] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38.

[20] J. Lukose, C.Y. Panicker, P.S. Nayak, B. Narayana, B.K. Sarojini, C. Van Alsenoy, A. A. Al-Saadi, Synthesis, structural and vibrational investigation on 2-phenyl-N-(pyrazin-2-yl) acetamide combining XRD diffraction, FT-IR and NMR spectroscopies with DFT calculations, SpectrochimicaActa Part A: Molec. Biomolec. Spectrosc. 135 (2015) 608–616.

[21] S. Demir, S. Cakmak, N. Dege, H. Kutuk, M. Odabasoglu, R.A. Kepekci, A novel 3-acetoxy-2-methyl-N-(4-methoxyphenyl) benzamide: molecular structural describe, antioxidant activity with use X-ray diffractions and DFT calculations, J. Mol. Struct. 1100 (2015) 582–591.

[22] S. Sebastian, N. Sundaraganesan, B. Karthikeiyan, V. Srinivasan, V, Quantum mechanical study of the structure and spectroscopic (FT-IR, FT-Raman, 13C, 1H and UV), first order hyperpolarizabilities, NBO and TD-DFT analysis of the 4-methyl-2-cyanobiphenyl, SpectrochimicaActa Part A: Molecul. Biomolec. Spectrosc. 78 (2) (2011) 590–600.

[23] H.W. Thompson, P. Torkington, The vibrational spectra of esters and ketones, J. Chem. Soc. (Resumed) 171 (1945) 640–645.

[24] E.D. Glendening, J.K. Badenhoop, A.E. Reed, J.E. Carpenter, J.A. Bohmann, C. M. Morales, F. Weinhold, NBO, Version 5.0, 68, Theoretical Chemistry Institute, University of Wisconsin, Madison, WI, 2001.

[25] I.V. Alabugin, M. Manoharan, S. Peabody, F. Weinhold, Electronic basis of improper hydrogen bonding: a subtle balance of hyperconjugation and rehybridization, J. Am. Chem. Soc. 125 (19) (2003) 5973–5987.

[26] J. Gao, A molecular-orbital derived polarization potential for liquid water, J. Chem. Phys. 109 (6) (1998) 2346–2354.

[27] F. Jensen, An introduction to the state of the art in quantum chemistry, Annu. Rep. Comput. Chem. 1 (2005) 3–17.

[28] C. BİLKAN, Determination of structural properties of some important polymers used as interfacial layer in fabrication of schottky barrier diodes (SBDs), J. Inst. Sci. Technol. 10 (1) (2020) 225–233.

[29] D.F.V. Lewis, C. Ioannides, D.V. Parke, Interaction of a series of nitriles with the alcohol-inducible isoform of P450: computer analysis of structure—Activity relationships, Xenobiotica 24 (5) (1994) 401–408.

[30] W. Guerrab, I.M. Chung, S. Kansiz, J.T. Mague, N. Dege, J. Taoufik, Y. Ramli, Synthesis, structural and molecular characterization of 2, 2-diphenyl-2H, 3H, 5H, 6H, 7H-imidazo [2, 1-b][1, 3] thiazin-3-one, J. Mol. Struct. 1197 (2019) 369–376.

[31] J.S. Murray, K. Sen, Molecular electrostatic potentials: concepts and applications (Eds.), 1996.

[32] L.G. Wade Jr, Organic Chemistry, Pearson Prentice Hall, New Jersey, 2006 sixth ed.

[33] E. Scrocco, J. Tomasi, Electronic molecular structure, reactivity and intermolecular forces: an euristic interpretation by means of electrostatic molecular potentials, Adv. quant. chem. 11 (1978) 115–193.

[34] G. Demirtaş, N. Dege, H. İçbudak, Ö. Yurdakul, O. Büyükgüngör, Experimental and DFT studies on Poly [di-μ 3-acesulfamato-O, O: O′; O′: O, O-di-μ-acesulfamato-O, O; N-di-μ-aqua-dicalcium (II)] complex, J. Inorg. Organomet. Polym. Mater. 22 (2012) 671–679.

[35] S.D. Kanmazalp, M. Macit, N. Dege, Hirshfeld surface, crystal structure and spectroscopic characterization of (E)-4-(diethylamino)-2-((4-phenoxyphenylimino) methyl) phenol with DFT studies, J. Mol. Struct. 1179 (2019) 181–191.

[36] S.M. Hiremath, N.R.P A.Suvitha, S.S.K.C.S. Hiremath, K.O S.K.Pattanayak, Molecular structure, vibrational spectra, NMR, UV, NBO, NLO, HOMO-LUMO and molecular docking of 2-(4, 6-dimethyl-1-benzofuran-3-yl) acetic acid (2DBAA): experimental and theoretical approach, J. Mol. Struct. 1171 (2018) 362–374.

[37] B.C. Smith, Infrared Spectral interpretation: a Systematic Approach, CRC press, 2018.

[38] P. Hobza, Z. Havlas, Blue-shifting hydrogen bonds, Chem. Rev. 100 (11) (2000) 4253–4264.

[39] D.L. Pavia, G.M. Lampman, G.S. Kriz, J.A. Vyvyan, Introduction to spectroscopy, Cengage learn. (2014).

[40] B. Smith, Infrared Spectral Interpretation: A Systematic Approach, CRC Press, Washington, DC, 1998.

[41] C.C. Moreno, R. Delgado, V. Hernandez, S. Hotta, L. Navarrete, Efficiency of the π conjugation in a novel family of α,α′-bisphenyl end-capped oligothiophenes by means of Raman spectroscopy, J. Chem. Phys. 116 (2002) 10419–10427.

[42] M. Amalanathan, I.H.J V.K.Rastogi, R.T M.A.Palafox, Density functional theory calculations and vibrational spectral analysis of 3, 5-(dinitrobenzoic acid, SpectrochimicaActa Part A: Molec. Biomolec. Spectrosc. 78 (5) (2011) 1437–1444.

[43] N. Sundaraganesan, B.D.J S.Ilakiamani, Vibrational spectroscopy investigation using ab initio and density functional theory analysis on the structure of 3, 4-dimethylbenzaldehyde, SpectrochimicaActa Part A: Molec. Biomolec. Spectrosc. 68 (3) (2007) 680–687.

[44] M.S. Weiss, J.S M.Brandl, D. Pal, R. Hilgenfeld, More hydrogen bonds for the (structural) biologist, Trends Biochem. Sci. 26 (9) (2001) 521–523.

[45] R. Bader, Atoms in molecules: a Quantum Theory, Oxford Univ. Press, Oxford, 1990.

[46] R.F. Bader, A quantum theory of molecular structure and its applications, Chem. Rev. 91 (5) (1991) 893–928.

[47] E. Espinosa, C.L E.Molins, Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities, Chem. Phys. Lett. 285 (1998) 170–173.

[48] A.D. Becke, K.E. Edgecombe, A simple measure of electron localization in atomic and molecular systems, J. Chem. Phys. 92 (9) (1990) 5397–5403.

[49] C.S. Abraham, S.M J.C.Prasana, M. Raja, Quantum computational studies, spectroscopic (FT-IR, FT-Raman and UV–Vis) profiling, natural hybrid orbital and molecular docking analysis on 2, 4 Dibromoaniline, J. Mol. Struct. 1160 (2018) 393–405.

[50] O. Christiansen, J. Gauss, J.F. Stanton, P. Jorgensen, The electronic spectrum of pyrrole, J. Chem. Phys. 111 (2) (1999) 525–537.

[51] B. Silvi, A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature 371 (6499) (1994) 683–686.

[52] A. Sagaama, O.Al-D N.Issaoui, M.J.W A.S.Kazachenko, Non-covalent interactions and molecular docking studies on morphine compound, J. King Saud. Univ.-Sci. 33 (8) (2021), 101606.

[53] H. Jacobsen, Localized-orbital locator (LOL) profiles of chemical bonding, Can J Chem 86 (7) (2008) 695–702.

[54] A.S. Kazachenko, M. Medimagh, N. Issaoui, O. Al-Dossary, M.J. Wojcik, A. S. Kazachenko, Y.N. lyar, Sulfamic acid/water complexes (SAA-H2O (1-8)) intermolecular hydrogen bond interactions: FTIR, X-ray, DFT and AIM analysis, J. Mol. Struct. 1265 (2022), 133394.

[55] F. Hammami, N. Issaoui, S. Nasr, Investigation of hydrogen bonded structure of urea-water mixtures through Infra-red spectroscopy and non-covalent interaction (NCI) theoretical approach, Comput. Theoret. Chem. 1199 (2021), 113218.

[56] G. Saleh, L.LoP C.Gatti, J. Contreras-García, Revealing non-covalent interactions in molecular crystals through their experimental electron densities, Chemistry– Eur. J. 18 (48) (2012) 15523–15536.

[57] G.M. Morris, R.S.H D.S.Goodsell, R. Huey, W.E. Hart, A.J.O R.K.Belew, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (14) (1998) 1639–1662.

[58] D.M. Burland, R.D. Miller, C.A. Walsh, Second-order nonlinearity in poled-polymer systems, Chem. Rev. 94 (1) (1994) 31–75.

[59] C.H. Bosshard, R. Spreiter, L.&. Degiorgi, P. Gunter, Infrared and Raman spectroscopy of the organic crystal DAST: polarization dependence and contribution of molecular vibrations to the linear electro-optic effect, Phys. Rev. B 66 (2002) 205107–205112.

[60] M. Amalanathan, T.S. Xavier, I.H. Joe, V.K &Rastogi, Normal coordinate analysis and Nonlinear Optical Response of cross-conjugated system 4, 4-DimethylBenzophenone, SpectrochimicaActa Part A: Molecul. Biomole. Spectrosc. 116 (2013) 574–583.

[61] M. Amalanathan, G.F. Jasmine, S.D.D. Roy, Comparative studies on molecular structure, vibrational spectra and hyperpolarizabilies of NLO chromophore Ethyl 4-Dimethylaminobenzoate, J. Mol. Struct. 1141 (2017) 400–416.

[62] M. Springborg, K. Schmidt, H. Meider, L.D. María, Theoretical studies of electronic properties of conjugated polymers, in: G.G R.Farchioni (Ed.), Organic Electronic Materials – ConjugatedPolymers and Low Molecular Weight Organic Solids, Springer-Verlag Berlin Heidelberg, Germany, 2001.

[63] T. Vijayakumar, I.H. Joe, C.R. Nair, V.S Jayakumar, Efficient π electrons delocalization in prospective push–pull non-linear optical chromophore 4-[N, N-dimethylamino]-4′-nitro stilbene (DANS): a vibrational spectroscopic study, Chem. Phys. 343 (1) (2008) 83–99.

[64] N.S. Labidi, A. Djebaili, I. Rouina, Substitution effects on the polarizability (α) and first hyperpolarizability (β) of all-trans hexatriene, J. Saudi Chem. Soc. 15 (1) (2011) 29–37.

_____

Relationship Between Sunspot and Solar Wind Parameters During the Ascending Phase

[1*]Abisha S Santham , [2]A. Iren Sobia, [3]Mariya Shaniya , [4]S.S. Bidhu

[1*] Research scholar Reg.No: 21113092132005, Department of Physics and Research Centre, Muslim Arts College, Thiruvithancode,Tamilnadu, India..

[2]Assistant Professor, Department of Physics and Research Centre, Muslim Arts College, Thiruvithancode, Tamilnadu, India.

[3]Research scholar Reg.No: 21213092132006, Department of Physics and Research Centre, Muslim Arts College, Thiruvithancode,Tamilnadu, India

[4]Assistant Professor, Department of Physics, Nanjil Catholic College of Arts and Science, Kaliyakkavilai.

Affiliated to Manonmaniam Sundaranar University, Abishekapatti-627012,Tirunelveli,Tamilnadu,India.

*Corresponding Author E-mail: abishassantham01@gmail.com.

Abstract: The characteristics of the solar wind, such as the magnetic field and plasma velocity, are dependent on the sunspot activity cycle overall. One of the most significant indicators for determining the sun's overall activity levels is the sunspot cycle. In the plot, we examine solar wind measurements made during solar cycle 25's rising phase and contrast them with comparable information from cycles 24 before. To determine the association between solar wind plasma properties and sunspots, information collected by the ACE and STEREO A spacecraft and NOAA (National Oceanic and Atmospheric Administration) was utilised. Analysis was done on the solar wind details during the 24 and 25 solar increasing activity phases, including sunspot, magnetic field, temperature, density, and wind speed. Many correlations between different features of the sunspot cycle have been discovered. A softer form of the solar wind emerges from the corona during solar minimum due to the Sun's low activity.

Keywords: sunspot, magnetic field, proton density, speed, temperature

1. Introduction

One of the primary objectives of space study is to understand the solar wind (SW), which is created when the heated solar corona expands into the interplanetary medium. On the one hand, research on the solar wind helps us comprehend the characteristics of the solar atmosphere and the mechanisms behind its plasma outflow [1-2]. It is essential to consider the solar magnetic field and how it interacts with plasma properties when elucidating specific reactions in the solar atmosphere. The magnetic field within the sun is produced by the solar dynamo process [3]. In the solar atmosphere, this expression is seen as a range of events [4-5]. An important component that sheds light on solar corona dynamics is the shift in profile of density and temperature during the solar cycle. The changes in temperature, emission, and density are caused by the solar magnetic field's structure evolving. Density changes with both solar and interplanetary sources are seen in solar wind data at 1 AU across a wide variety of time periods. The largest scale fluctuations show the wind's solar origin: high densities are found in the sluggish flow from the area of the streamer belt, while low densities are found in the rapid flow from coronal holes [6]. The quantity and area of sunspots produced during a sunspot cycle serve as indicators of its intensity. Trends in the average features of individual sunspots during a sunspot cycle or between succeeding sunspot cycles have been hard to find, despite the fact that the quantity of sunspots varies significantly from one cycle to the next, suggesting an intrinsic variability in the mechanism that creates sunspots. Additionally, we may anticipate that at a cycle's maximum rather than its rising or descending phase, stronger, larger sunspots would emerge more frequently [7]. The border between rising chromospheric temperature and falling density is marked by the transition area. Similarly, rising temperature and falling density are observed through the solar corona. Solar wind velocity (V) multiplied by magnetic field (B), or V * B, is the most significant indicator of the degree of solar wind connection with the terrestrial atmosphere and magnetosphere [8]. The solar cycle is one of the most important indicators of the sun's overall activity. Radars,

_____

high-frequency signals, radio transmissions, ground power lines, spaceflights, the geospace environment, and life on Earth are all significantly impacted by solar activity [9]. For this reason, predicting solar activity is essential to both comprehending the solar activity process and safeguarding contemporary systems.

Due to the interdependence between solar wind characteristics, this work has evaluated many solar wind and sunspot parameters, including temperature, density, magnetic field, and wind velocity. Several parameters change in this case: magnetic field, proton density, temperature, and speed of the solar wind.

2. Data Analysis

For the current investigation, these two rising eras were taken into consideration. The spacecraft's ACE and STEREO A provided the daily averaged parameters of the solar wind values, while NOAA provided the sunspot data. In the study, measurements of the sunspot and solar wind parameters during the ascending phase of 24 and 25 solar cycles were obtained. In mid-2008, the solar cycle 24 started, and it concluded in 2019, and the ascending phase relates to the end of 2011–2013. Solar cycle 25 began in December 2019 and is still going on, with the ascending phase spanning from 2020 to 2022. For the current investigation, these two rising eras were taken into consideration. The spacecraft's ACE and STEREO A provided the daily averaged values of the solar wind parameters, while NOAA provided the sunspot data. The Solar Wind Electron, Proton, and Alpha Monitor (SWEPAM) instrument for ACE and the Solar Wind Experiment (SWE) instrument for STEREO A studied the speed, temperature, and density of the solar wind. The Sun-Earth Connection Coronal and Heliospheric Investigation (SECCHI) sensors, the magnetometer (MAG), and the STEREO/WAVES (SWAVES) instruments provided the magnetic field data for ACE. IMPACT will offer the regional magnetic field and plasma parameters of solar energetic particles. PLASTIC will offer plasma characteristics for protons, alpha particles, and large ions in the solar wind.

3. The Relation Between Solar Wind Parameters

Changes in temperature and density profiles over the solar cycle are substantial and provide an understanding of the dynamics of the solar corona. A velocity vs. temperature graph for the 24 and 25 solar ascending phases is shown in Figure 1. The solar wind temperature and velocity as noticed by the ACE (red) and STEREO A (green) spacecrafts are shown in the graphic. Solar wind speeds vary throughout the rising phase of the solar cycle due to the sun's increasingly distorted magnetic field. These figures also show that the solar wind velocity stays between 400 and 600 km/s for most of the data set. During solar cycle 24, the highest recorded speed was 1200 km/s at a temperature of 300,000 K. At a temperature of 400000K, the solar cycle 25 achieved its maximum velocity of 750 km/s. Since the particle speed and temperature of the solar wind are related, temperature changes should be observed in parallel with variations in solar wind speed. It is often noticed that a stream moving at a very rapid speed has relatively high temperatures; slower streams are typically cooler. [10] concluded that the solar wind's acceleration near the sun is related to temperature fluctuations caused by the stream's growth in the interplanetary medium. The component event on the sun's surface caused the huge amplitude of the solar wind's temperature and speed in Figure 1. It is believed that this stream of particles is extremely tenuous due to its rapid and high temperature.

_____
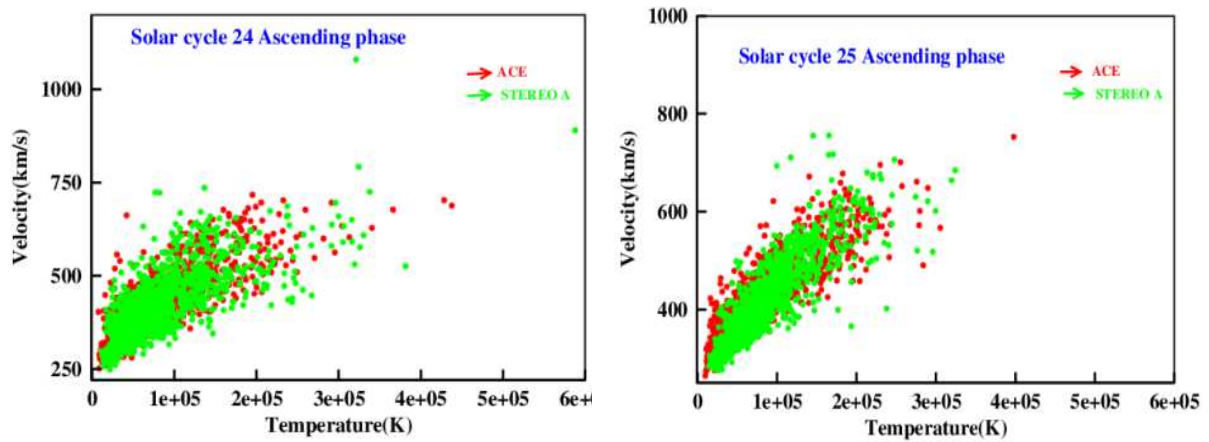
Fig 1: solar wind velocity Vs .temperature for 24 and 25 solar cycle ascending phase

For the rising phase of solar cycles 24 and 25, the density versus velocity charts is displayed in Figure 2. While most of the highest densities were seen at lower velocities, the combined plot of the density vs. velocity values for the observation time showed an increase in the extreme values of density over the velocity range. The increased density was caused by the lower velocity streams' faster journey, which packed the particles within. The square root of the medium's density and velocity are negatively correlated. This implies a reduction in velocity with an increase in density. The greatest densities, which range from 300 km/s to 500 km/s, are observed for most data series during the ascending portion of solar cycles 24 and 25. Since the results of tomographical density measurements reveal a very narrow belt of high-density (low-speed) streamers during the solar minimum and because the belt's small latitudinal meanderings permit regions of extremely low density to stretch from the poles to the equator, there is a periodic presence of very high-speed solar wind streams that originate from polar coronal holes close to the equator [11-12]. Due to the increasing quantity of high-velocity streams leaving the solar surface during the solar rising phase, the density values are abnormally low. The particles spread out at very fast speeds, making the stream extremely flimsy.
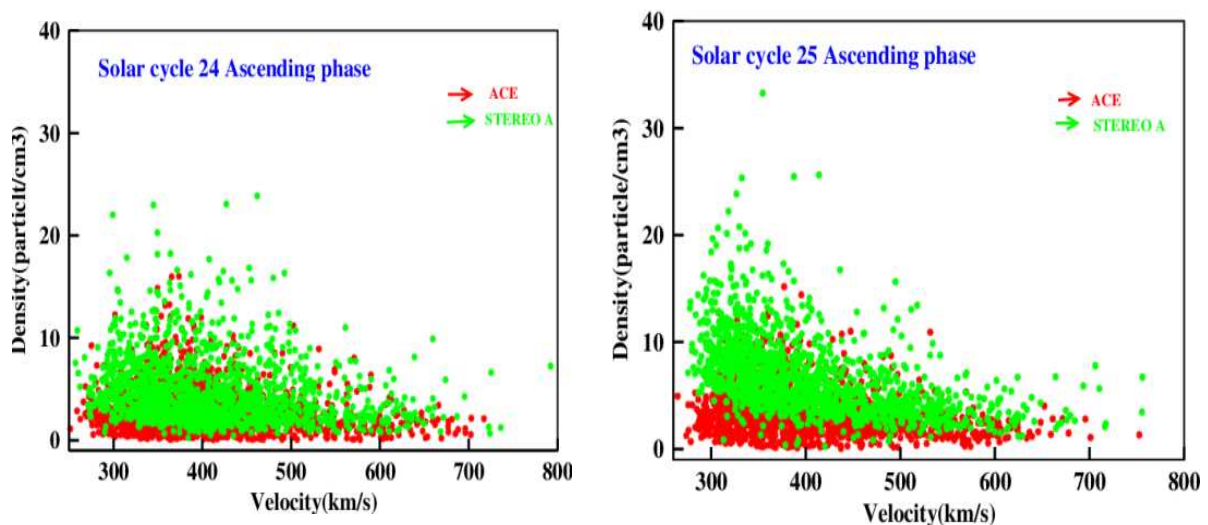
Fig 2: solar wind density Vs. velocity for 24 and 25 solar cycle ascending phase

_____

Thus, it is widely recognised that higher solar wind velocities originate from lower-density areas of the corona and vice versa [13-14]. A spacecraft experiences a decrease in density as its velocity increase. [15-16] the quick solar wind is characterised by less density, smaller charge state ratios, and greater proton temperatures when compared to the slow solar wind.
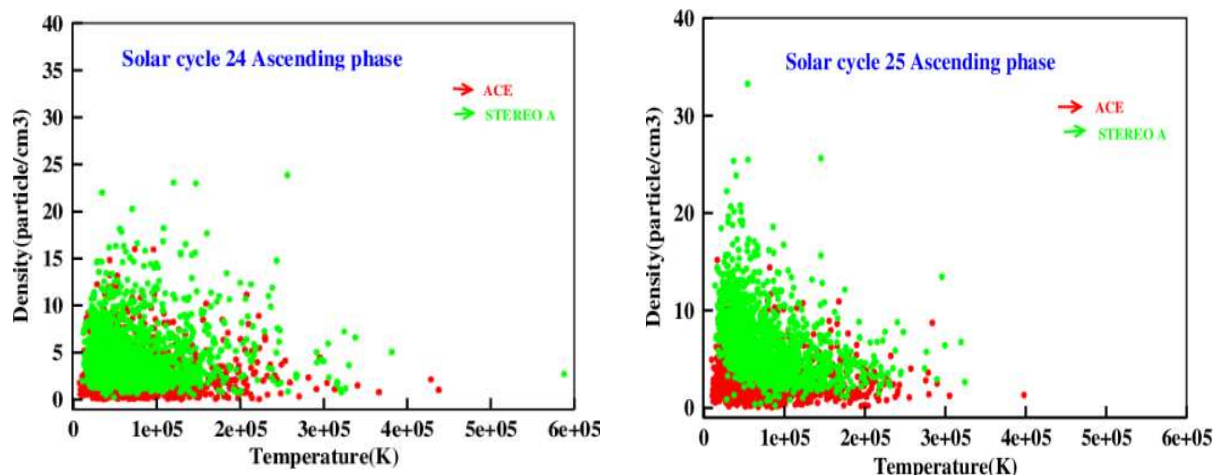
Fig 3: solar wind density Vs. temperature for 24 and 25 solar cycle ascending phase

The solar wind's temperature and velocity both affect its particle density. The magnetic field is linked to the temperature and density profiles. The density vs. temperature graph for stages 24 and 25 of the solar ascending process is displayed in Figure 3. The values on the density axis in this instance increase to 24 p/cm^3 for solar cycle ascending phase 24 and to 35 p/cm^3 for solar cycle ascending phase 25. In solar cycle 25, the temperature axis moved to 600,000 K, while in cycle 24, it reached up to 400,000 K. Because of a greater rate of density, higher temperatures result in a decrease in pressure across the solar cycle. As solar activity increases, temperatures rise somewhat and density decreases. The link between temperature and density is obviously inverse, with rising temperatures resulting in declining densities. Because of the higher solar wind density in both solar ascending phases, a lower temperature profile was observed. Where the lines of magnetic field are open as well as the solar wind is free to circulate, these are known as coronal holes in the sun. The temperatures are lower and the concentrations are lower in these locations. The temperature and density patterns are significantly dependent on the intensity of the magnetic field since the density distribution is directly proportional to the magnetic field while the temperature profiles are inversely proportional [17]. During solar cycle 25, as opposed to cycle 24, greater values in the density profiles are more common.
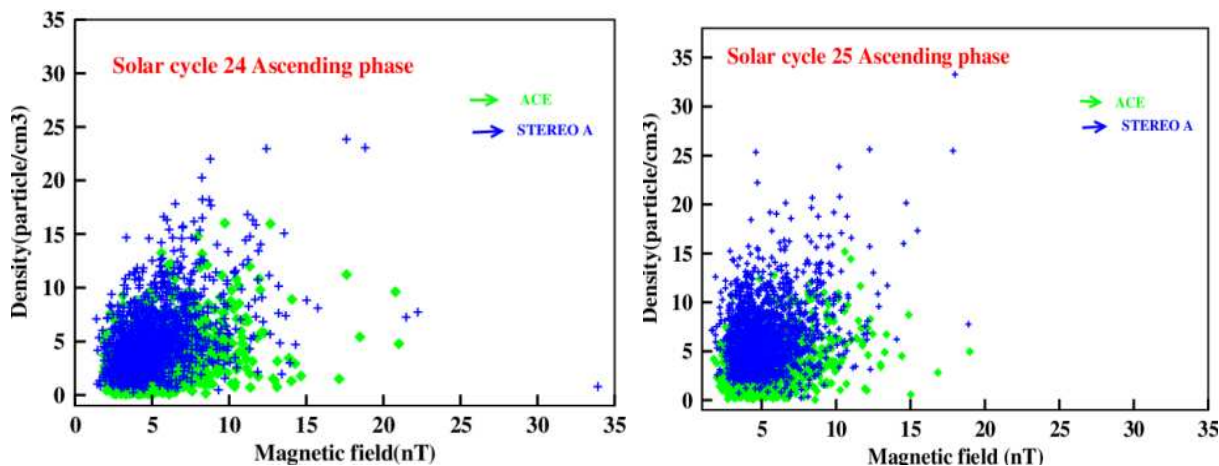
Fig 4: solar wind density Vs. magneticfield for 24 and 25 solar cycle ascending phase

_____

Density was lowest at higher magnetic field strengths, and vice versa. Peak magnetic field values were recorded during both solar cycles, with density falling between 15 n T and 20 n T. While density varies sharply, no event lasts long enough to generate several hours in a row with extremely high or low density values. The regions with higher density values in the areas with lower magnetic fields are shown in the figure 4. A passing CME can be identified by its typical reduction in density. There is an active region connected to these places. Structures like coronal mass ejections (CMEs) and solar flares, which originate on the solar surface, produce large-scale changes in the density of the solar wind [18]. The density variations' distribution, which coronal holes are likely to influence, is thus linked to the activity of the sun's magnetic field.
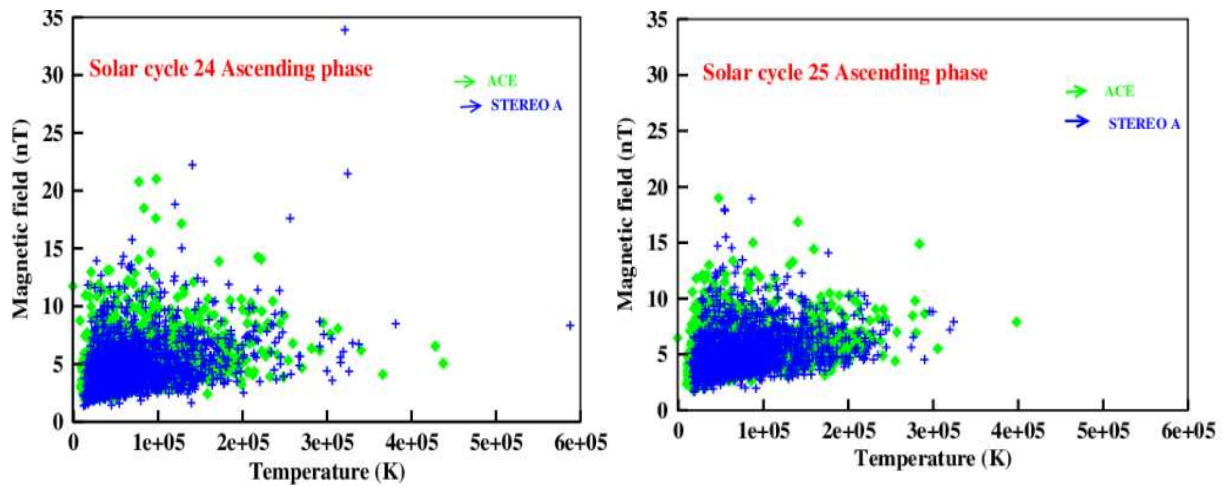
Fig 5: sun wind magneticfield Vs. temperature for 24 and 25 solar cycle ascending phase

Plots of the magnetic field vs. temperature for the 24 and 25 ascending phases are displayed in Figure 5. In both cycles, the solar wind was dispersed disorderly for greater levels of magnetic field and temperature and tightly distributed for lower values. The magnetic fields were stronger in the 24 solar cycles when the maximum was seen for the same temperature range in two solar cycles. In both solar ascending phases, its magnetic field weakened as the temperature of the sun wind rose. MCs have plasma values much lower because of the powerful magnetic field and low proton temperatures.
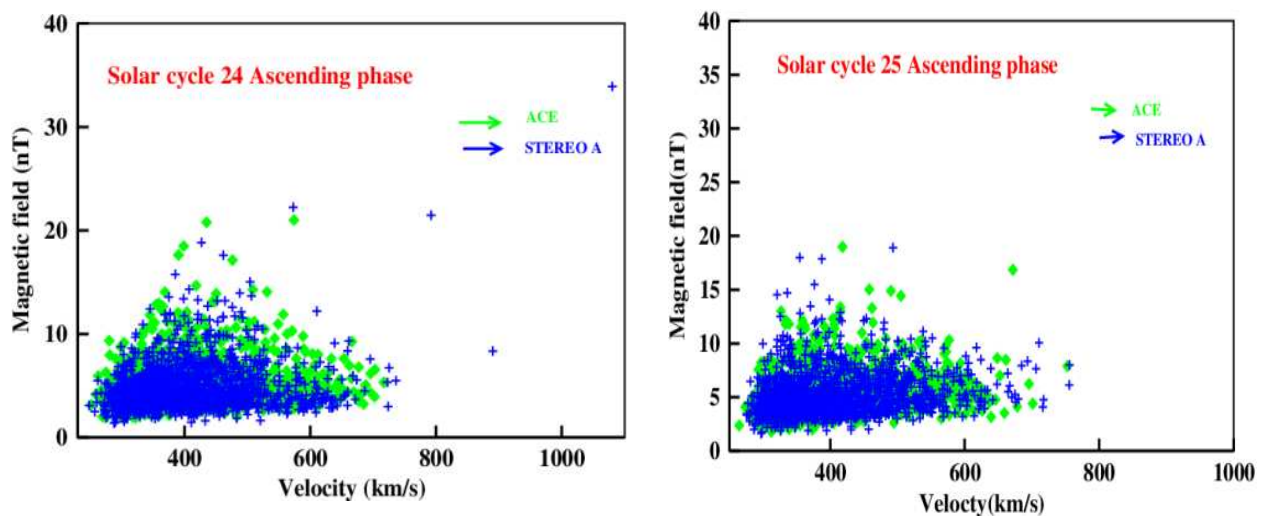
Fig 6: solar wind magneticfield Vs. velocity for 24 and 25 solar cycle ascending phase

_____

The magnetism of the sun becomes more and more distorted throughout the rising part of the solar cycle, which causes variations in solar wind speed. A depiction of the solar wind magnetic field against velocity for the ascending periods of solar cycles 24 and 25 is presented in Figure 6. The greatest magnetic field of 35 nT was reported during solar cycle 24, and just one magnetic field was observed after 1000 km/s. The greatest magnetic field of 20 n T was reported during solar cycle 25. The maximum magnetic field that may be produced is 400–600 km/h. Lower-speed streams in the solar cycle flow at the rate at which the particles inside them compress and cause a drop in solar activity. Interplanetary disturbances were most likely unconnected to changes in velocity and magnetic fields. Consequently, [19] found that at times of low solar activity, there is a strong connection between the radial and vertical components of the magnetic field that exists between planets and that indicates the relationship coefficient relies on both mean magnetic field directions.

4. The Relationship Between Sunspot And Solar Wind Characteristics In The Solar Wind Rising Phase

On the surface of the sun, activity varies in tandem with changes in magnetic fields. Sunspot counts are one method of monitoring the solar cycle. A solar lowest number of sunspots on the solar, marks the start of a solar cycle. Both solar activity and the number of sunspots increase with time. We examine charts showing sunspot numbers in relation to several solar wind characteristics that the spacecraft measured.
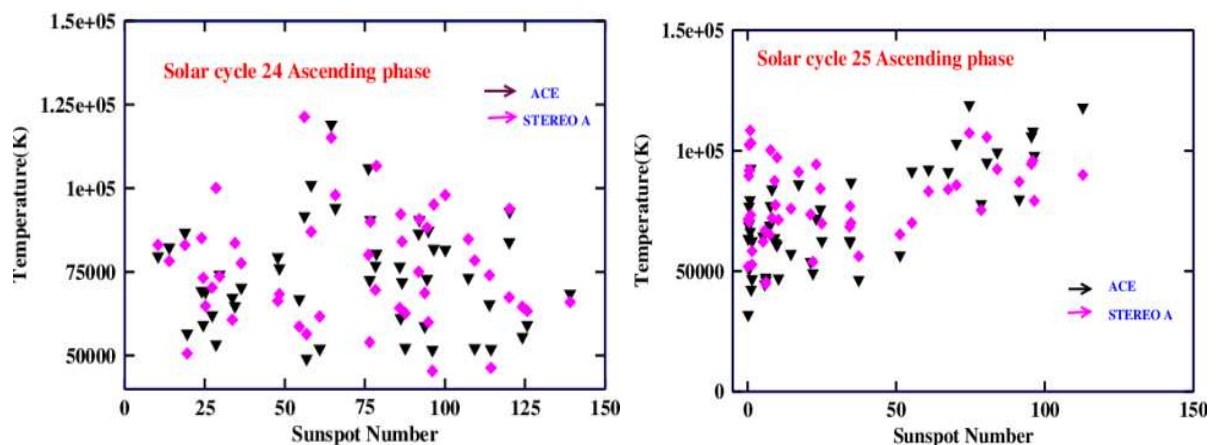
Fig 7: Solar wind temperature Vs .sunspot number in 24 and 25 solar ascending

A plot of the sunspot number against speed for the minimum-year solar cycle 24 and another plot for the solar cycle 25 are shown in Figure 7. The graph depicts the spread rates of the solar cycle 24 and the more condensed values of the solar cycle 25. On both the vertical and horizontal axes, the values are more dispersed. The sunspot numbers during solar cycles 24 and 25 ranged from 10 to 140 and 1 to 120, respectively. 125000K was the hottest temperature recorded throughout both cycles. The temperatures increased because more sunspots produced more energy. According to Brooks and Warren (2012)[20], the figure illustrates that active area outflows have high coronal temperatures. Additionally, [21] have shown that these outflows have open field lines that let plasma slip into the heliosphere.

_____
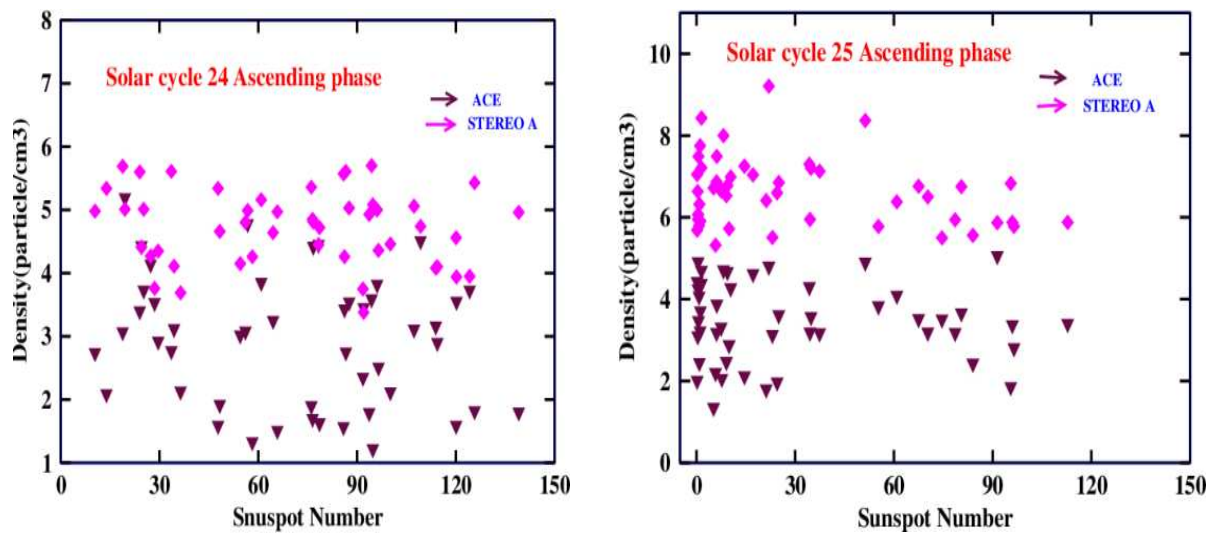
Fig 8: Solar wind density Vs .sunspot number in 24 and 25 solar ascending

A composite plot of particle density vs. sunspot number for the ascending phase is shown in Figure 8. Values on the vertical and horizontal axes are spread more equally during the solar minimum. For solar cycles 24 and 25, respectively, the range is 1 to 6 and 1 to 9 p/cm3. The greatest density occurrences recorded by the STEREO A satellite were both solar cycles 24 and 25. Given that the magnetic field and the density profile are proportionate. [17] noted that the density profiles correspond to the sunspot movement over the solar cycle, along with fluctuations in the magnetic field. The density of particles rises in colder, slower-moving streams, whereas it decreases in hotter, faster-moving streams. According to [22] the solar wind's maximum temperatures and speeds really happen during the cycle's dropping phase, whereas its maximum densities happen during its rising phase. Between 1 and 75 au, there is a strong agreement between the theoretical and measured solar wind densities, which decline with increasing heliocentric distance [23].
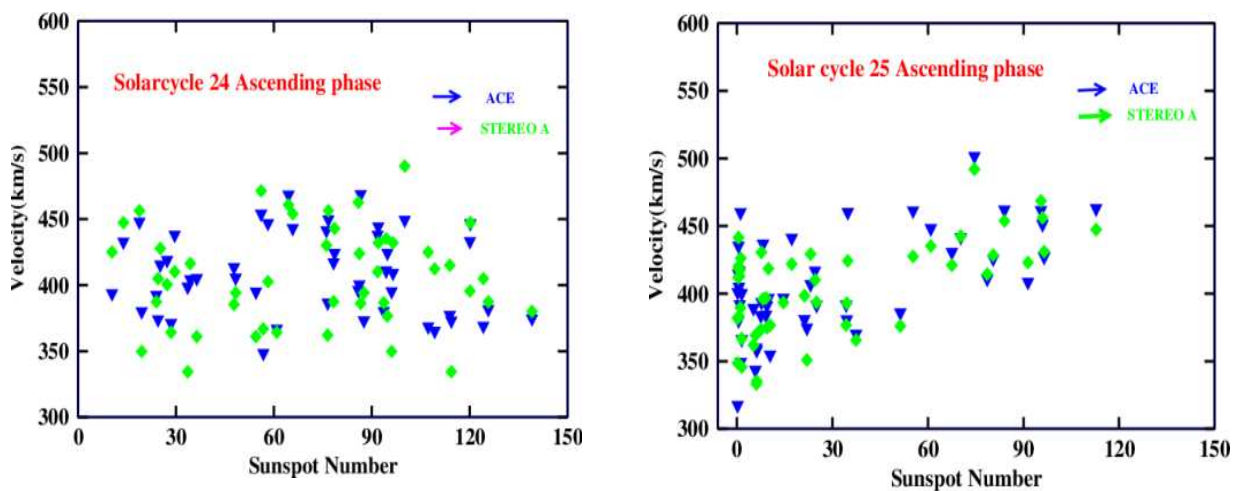
Fig 9: Solar wind velocity Vs .sunspot number in 24 and 25 solar ascending

_____
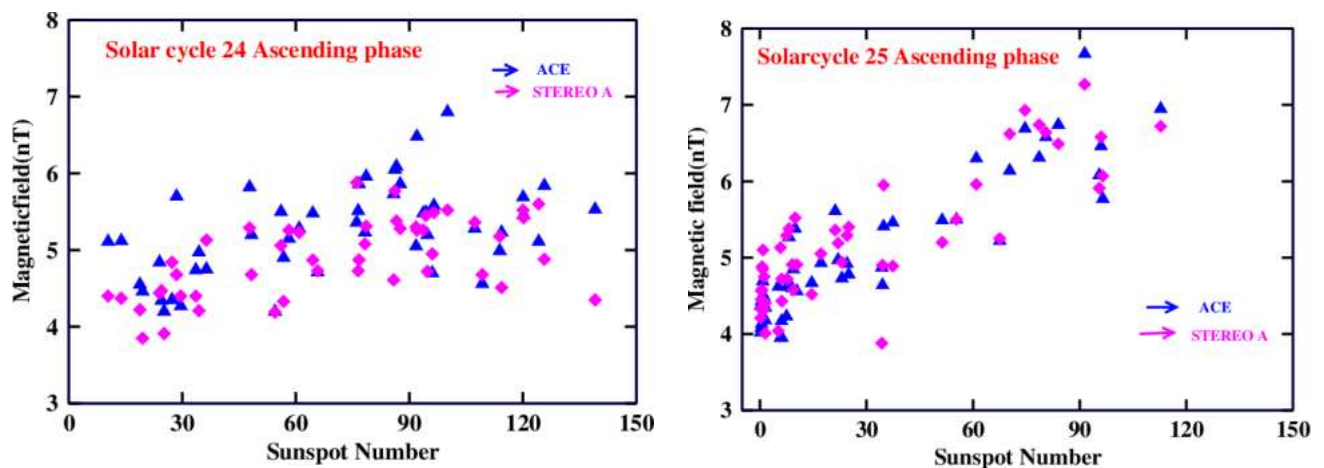
Fig 10: Solar wind magnetic field Vs sunspot number in 24 and 25 solar ascending.

The dynamics of the sun's outer layer are significantly impacted by the fluctuation of the magnetic field, which is measured by a variety of solar metrics, including the sunspot number, frequency of CME occurrences, solar wind plasma temperature, and solar wind speed. Figure 10 depicts the sun during the start of solar cycles 24 and 25, when it is quite calm and has fewer sunspots and solar activity. The magnetic field in 24 SC had a vertical range of around 7 nT during the solar ascending phase, while the magnetic field in 25 SC had a vertical range of about 8 nT. The magnetic field progressively rises with the number of sunspots throughout the rising phase. At the beginning of the rising phase, the magnetic field showed predictable fluctuations with the sunspot number. Low activity during the first part of solar cycles 24 and 25 was seen. The magnetic field is weak, and the sunspot region is modest right now. Here, the sunspot movement and area are at their highest, and the magnetic field is strongest. The relative sunspot number correlates with an increase in the solar magnetic field's strength. However, when the sunspot grows during the latter portion of the rising phase, the magnetic field weakens. The intensity of the solar cycle is established by the quantity, darkness, and field strength of sunspots, according to [24]. Because of how powerful these magnetic fields are, some of the sun's heat is prevented from rising to the surface.

In comparison to the parameters in the solar cycle 25 ascending, all of the parameters in the solar cycle 24 ascending were significantly smaller in quantity and magnitude, including solar wind velocity, temperature, density, magnetic field, and sunspot number. This is most likely caused by the 24 solar cycles' weaknesses. According to [25], solar cycle 24 is incredibly feeble, as shown by sunspot numbers, and little solar activity is seen on the Sun during this rising period of the solar cycle. In terms of disturbances seen in the heliosphere and on the solar surface, Cycle 24 is shown to be weaker than Cycle 23 [26-27].

5. Conclusion

Variations in the sun's surface activity lead to differences in the solar wind produced around the corona. The solar wind characteristics during the 24th and 25th solar ascending phases were obtained by cross-analysing the interdependence of solar wind parameters and sunspots. ACE and STEREO spacecraft observed similar types of variations in solar wind parameters in the 24 and 25 solar ascending phases. In comparison to the solar wind parameters in the ascending phase of solar cycle 25, the parameters related to solar wind, including temperature, density, magnetic field, and velocity, exhibited lower levels of activity in solar cycle 24. The smaller activity was due to the weak solar cycle. Owing to the density profile's proportionality to the magnetic field, variations in the magnetic field are associated with the sunspot trend observed during the solar cycle.

The characteristics of solar wind speed, magnetic field, density, temperature, and sunspot number did not depend on each other when compared to the data recorded by the 24 solar ascending phases. The traits shown in the 25 solar ascending were more numerous than those in the 24 solar ascending, suggesting that the 24 solar cycle is weaker. Along with the charge exchange rates of interstellar neutral species within the heliosphere, the solar wind representation depicts the development of solar wind density and speed. A weak

_____

cycle with few sunspots and little solar activity is predicted for the 24th solar cycle. A low quantity of sunspots was found in the equatorial area, where the number was uniformly distributed during the ascending phase. When comparing solar cycles 24 and 25, the total magnetic activity of cycle 24 was less.

References

[1]   R.Schwenn, Solar Wind Sources and Their Variations over the Solar Cycle. Space Sci. Rev. **124**, 51–76 (2006)

[2]   R. Schwenn, Solar Wind Sources and Their Variations over the Solar Cycle. In Solar Dynamics and Its Effects on the Heliosphere and Earth; Baker, D.N., Klecker, B., Schwartz, S.J., Schwenn, R., Von Steiger, R., Eds.; Space Sciences Series of ISSI; Springer: New York, NY, USA,; Volume 22 (2007)

[3]   M. Dikpatiand, P. A. Gilman, Flux-Transport Solar Dynamos. Ssr, **144,** 67–75, Apr. 2009.doi:10.1007/s11214-008-9484-3.

[4]   S. K.Solanki, , B. Inhester, and M. Schüssler, The solar magnetic field, Rep. Prog. Phys., **69**, 563–668 (2006), doi:10.1088/0034-4885/69/3/R02.

[5]   H.D . Mackay, R.A. Yeates, and francois-xavier bocquet, Impact Of An L5 Magnetograph On Nonpotential Solar Global Magnetic Field Modeling. The Astrophysical Journal, 825:131 (16pp), 2016. http://dx.doi.org/10.3847/0004-637X/825/2/131

[6]   N.U. Crooker Solar and heliospheric geoe ective disturbances Journal of Atmospheric and Solar-Terrestrial Physics 62 (2000) 1071–1085

[7]   A.A. Norton, E.H. Jones2 and Y. Liu How do the magnetic field strengths and intensities of sunspots vary over the solar cycle? Eclipse on the Coral Sea: Cycle 24 Ascending (GONG 2012, LWS/SDO-5, and   SOHO   27) Journal of Physics: Conference Series 440 (2013) 012038. http://doi:10.1088/1742-6596/440/1/012038.

[8]   Aslam O. P. M., Badruddin, Study of Cosmic-Ray Modulation during the Recent Unusual Minimum and Mini Maximum of Solar Cycle 24, 2015, Sol. Phys. 290, 233359.

[9]   D. Nandy, Progress in solar cycle predictions: sunspot cycles 24–25 in perspective: invited review. Sol. Phys. **296** (3) (2021). https://doi.org/10.1007/s11207-021-01797-2.

[10]   L. F. Burlaga, N. F. Ness, F. Mariani,. Magnetic fields and flows 3168 between 1 and 0.3 AU during the primary mission of Helios 1. J. Geo- 3169 phys. Res., **83**(A11), 5167–5174 (1978), https://doi:10.1029/JA083iA11p05167.

[11]   H. Morgan and Anthony C. Cook, The Width, Density, and Outflow of Solar Coronal Streamers, The Astrophysical Journal, **893:**57 (15pp) (2020) https://doi.org/10.3847/1538-4357/ab7e32.

[12]   H. Morgan, and S.R. habbal, observational aspects of the three-dimensional coronal structure over a solar activity cycle, The Astrophysical Journal, **710,**1–15   (2010). http://dx.doi.org/10.1088/0004-637X/710/1/1

[13]   R. C. Allen, D. Lario, D. Odstrcil, G. C. Ho,   L. K.Jian, C. M. S.Cohen, (2020) Solar wind streams and stream interaction regions observed by the parker solar probe with corresponding observations at 1 AU. The Astrophysical Journal - Supplement Series, **246**(2), 36. https://doi.org/10.3847/1538-4365/ab578f.

[14]   R. Schwenn, Space weather: The solar perspective. Living Reviews in Solar Physics, 3. (2006). https://doi.org/10.12942/lrsp-2006-2.

[15]   F. Xu, J.E. Borovsky A new four-plasma categorization scheme for the solar wind. J Geophys Res **120**(1), 70–100. (2015). https://doi.org/10.1002/2014JA020412.

[16]   Y.-M. Wang, The Oxygen Charge-State Ratio As An indicator Of Footpoint Field Strength In The Source Regions Of The Solar Wind, The Astrophysical Journal, **833**, 1, https://doi.org/ 10.3847/1538-4357/833/1/121.

[17]   M.R. Gomez L. Vieira, A. Dal Lago, and J. Palacios, Coronal Electron Density Temperature and Solar Spectral Irradiance during Solar Cycles 23 and 24, The Astrophysical Journal, **852**:137 (11pp), 2018. https://doi.org/10.3847/1538-4357/aa9f1c.

[18]   S.K. Bisoi, P. Janardhan, M. Ingale, P. Subramanian, S. Ananthakrishnan, M. Tokumaru, and K. Fujiki,, A Study Of Density Modulation Index In The Inner Heliospheric Solar Wind During Solar

_____

Cycle 23, The Astrophysical Journal, **795**:69 (8pp), 2014, http://dx.doi.org/10.1088/0004-637X/795/1/69.

[19]   W. Lyatsky, A.Tan, S. Lyatskaya, Effect of Solar magnetic field polarity on interplanetary magnetic field Bz. Geophys. Res. Lett. **30** (24), 2258 (2003). doi:10.1029/2003GL017431.

[20]   H. David, D.H Brooks, and Harry P. Warren, The Coronal Source Of Extreme-Ultraviolet Line Profile Asymmetries In Solar Active Region Outflows, The Astrophysical Journal Letters, **760**:L5 (6pp). http://dx.doi.org/10.1088/2041-8205/760/1/L5.

[21]   V.Slemzin , L. Harra , A. Urnov, S. Kuzin , F. Goryaev , D. Berghmans , Signatures of slow solar wind streams from active regions in the inner corona, Sol. Phys., **286**, 157 (2013)

[22]   I. S. Veselovsky, A.V. Dmitriev, A.V. Suvorova, M.V. Tarsina, Solar Wind Variation with the Cycle, Indian Academy of Sciences, J. Astrophys. Astr. **21**, 423-429, (2000).

[23]   M.Cameron R, SchüsslerThe crucial role of surface magnetic fields for the solar dynamo. Science **347**,1333–1335, (2015). doi:10.1126/science.1261470.

[24]   G. P. Adhikari, L. Zank, L. Zhao, and G. M. Webb, Evolution of Entropy and Mediation of the Solar Wind by Turbulence The Astrophysical Journal, **891**:34 (7pp) (2020).https://doi.org/10.3847/1538-4357/ab7010

[25]   N. Gopalswamy, H. Xie, S. Akiyama, P. Makela, S. Yashiro, and G. Michalek, The Peculiar Behavior of Halo Coronal Mass Ejections in Solar Cycle 24. Astrophysical J. **804**, L23 (2015) http://dx.doi.org/10.1088/2041-8205/804/1/L23.

[26]   H. M. Antia , S. Basu S, solar rotation rate during the cycle 24 minimum in activity Apj, **720,** 494 (2010), http://dx.doi.org/10.1088/0004-637X/720/1/494.

[27]   I.G. Richardson Geomagnetic activity during the rising phase of solar cycle 24. J Space Weather Space Clim, **3,** A08. (2013), https:// doi:10.1051/swsc/2013031.

Investigation of three events of solar parameters and Interplanetary Coronal Mass ejections during the maximum phase of solar cycle 24

M. S. Dhaiya[1], S. S. Bidhu[2], A. Iren Sobia[3]

[1] Research scholar (Reg. No:20215092132011), Department of Physics and Research Centre, Muslim Arts College, Thiruvithancode – 629174, Tamilnadu, India. Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli – 627012, Tamil Nadu, India
deenamasha@gmail.com

[2] Assistant Professor, Department of Physics and Research Centre, Nanjil Catholic College of Arts and Science, Kaliyakkavilai - 695502, Tamilnadu, India.

[3] Assistant Professor, Department of Physics and Research Centre, Muslim Arts College, Thiruvithancode - 629174, Tamilnadu, India.

Abstract. This paper analyzes the Three events of Solar Parameters and Interplanetary Coronal Mass Ejections in the maximum phase of Solar Cycle 24 and focuses on the magnetic activity of interplanetary coronal mass ejection during the solar cycle 24. We investigate the magnetic field magnitude (B), Proton temperature (Tp), Proton density (Np). From this study, we find the highest peak of IP (Interplanetary) shocks on disk center MC (Magnetic cloud) events during the solar cycle 24 and also we investigate the magnetic activity of the solar cycle. In this study, we find that the ACE spacecraft shows the fastest coronal mass ejection and highest interplanetary shock wave on solar cycle 24. It is important to note that extreme events can happen at any time during a cycle. In solar cycle 24, from July 13, 2012 to July 15, 2012 largest storm occurred because the magnetic field was -52nT and linear speed was $1500 \, \mathrm{km \, s^{-1}}$ observed.

Key words: Interplanetary Coronal Mass Ejection, Magnetic activity, Magnetic cloud, IP shocks

Introduction

The Sun is the main source of heat and light in the solar system. The number of sunspots that can be seen on the Sun's surface varies over the duration of an 11-year solar cycle. Along with other factors, it has something to do with the solar cycle, magnetic field, coronal mass ejections, solar outbursts and radio emission.

The large amount of ionized plasma that can come out of the upper atmosphere of the Sun is called a Coronal Mass Ejection (CME). CMEs are very important events for the Sun. The view of the White-light corona graph field is different and extremely bright, showing that the CME was moving away on the timescale of minutes to hours (Munro et al., 1979).

The CME comes from corona of the Sun. The fastest mode shocks speed up the charged particles and also produce very powerful geomagnetic storm. The most energetic phenomenon in the heliosphere are interplanetary coronal mass ejections (ICMEs) and is detected in the corona graph. This presents an overview of the observational features of ICMEs in respect to the normal solar wind and their ancestor CMEs.

We looked at how the speed of the CMEs, which propel the shocks towards the Sun and into the IP medium, affected the IP shocks mentioned by Gopalswamy et al. (2010). We discover that SC/SIs were responsible for

nearly 91 percentage of IP shocks. The CMEs connected to sudden commencement(SCs) and sudden impulses(SIs) have an average speed of 1015 km s^{-1}. This is roughly double the speed of a conventional CME.

It is well known that CMEs, which are large-scale structures, are connected to coronal slow/fast streamers, coronal gaps, and other CMEs. Advanced radio emission, particle acceleration, solar wind bulk flow speed, proton density, CME deflection, and proton temperature are a few of these interactions' prominent consequences. CMEs can deviate from the Sun-Earth axis and still strike the planet. The Sun-Earth line may finally be struck by CMEs that have been diverted, resulting in powerful geomagnetic storms.

The CMEs are made up of two occurrences: halo and partial halo events with a width larger than 120°. Because of how they appear in corona graph photos during the solar cycle's maximum and rising phases, these CMEs occurred throughout the solar cycle's rising and maximum phases, which lasted from 2009 to 2019. Because of how they appear in corona graph photos, coronal mass ejections directed at Earth are referred to as "halo events". The Sun looks to be being enriched by the growing cloud of an Earth-directed CME, which is encircling our star. CMEs with a halo have an apparent angular width of 360 degrees. In contrast to limb halos, which have an apparent angular width between 45°and 90°, CMEs have an apparent angular width between 120°and 360°.

Based on data from the Advanced Composition Explorer (ACE) spacecraft (Richardson & Cane, 2010) and numerous thorough ICME catalogues, measurements are provided. It is impossible to analyse ICMEs using a single, entirely objective method (Huttunen et al., 2005), hence some subjective judgments may influence how ICMEs are detected and how accurately their boundaries are specified. Chi (2016) found that despite some variances, the results of the RC and Chi catalogues are comparable.

To carry out our current studies, first ICMEs in the background solar wind must be identified according to a number of criteria, such as increased magnetic field strength, smoothly changing field direction, low proton temperature, etc. (Zurbuchen & Richardson, 2006; Wu & Lepping, 2011; Song & Yao, 2020). Then, each ejecta should be given an MC. The progressive change in field direction is the characteristic of MCs that is most noticeable. Here, in order to separate the concerns of ICME identification and statistics and to narrow the emphasis of our investigation, we decided to use the ICME catalogue offered by the specialists in this field instead of recognising the events ourselves.

Magnetic clouds are huge interplanetary flux ropes that grow as they travel through the solar wind from the Sun. We look at how in situ data and magnetic cloud models, which are represented as cylindrical shape magnetic flux ropes with expansion, correspond. We continue our earlier investigation of the Rise, Maximum, and Declining phases of solar cycle 24 in this paper. Because each ICME is located nearer to the disk's core, both the Data section and the Data Analysis Method were assessed.

1 Data Selection and Method of Analysis

The data used in this study are taken from ACE. The IP CMEs and Magnetic clouds we examined are listed at the below link[4]. Richardson & Cane (2010) and the MCs of cycle 24 published before (Gopalswamy et al., 2008; Lepping et al., 2015) are used. We use proton temperature and magnetic field data from the ACE' MAG-SWEPAM level 64-sec averages[5]. For each ICME, there is a corresponding CME observed by the Large Angle and Spectrometric Corona-graph (LASCO) (Brueckner et al., 1995) on board the Solar and Heliospheric Observatory (SOHO) mission.

2 Result and Discussion

2.1 Magnetic cloud during solar cycle 24

Table 1. Number of Magnetic cloud during 24th solar cycle

Year	MC count
2008	04
2009	12
2010	17
2011	36
2012	44
2013	35
2014	24
2015	03
2016	18
2017	12
2018	11
2019	10

SOHO has recorded 255 MCs during Solar Cycle 24. We began with the MC events that were reported between December 2008 and December 2019. Due to the Sun's low activity and small range of CMEs, Solar Cycle 24, 2008, only saw four CMEs. Due to the Sun's activity being higher than other occurrences in solar cycle 24 in 2012, 44 MCs were detected by the LASCO sensor on board SOHO. The MCs with the highest speed and largest overall magnitudes in this solar cycle are located nearer to the centre of the solar disc, which makes little sense when hyphenated across lines.

2.2 ICME Identification for Selected Event

Gopalswamy (2010) used the following data to identify the MCs in charge of IP shocks: (Stone et al., 1998) employed the ACE's magnetic field magnitude

---

[4] *www.srl.caltech.edu/ACE/ASC/DATA/ level3/icmetable2.htm*
[5] *www.srl.caltech.edu/ACE/ ASC/level2/lvl2DATA_MAG − SWEPAM.html*

Table 2. Disk center MC Events during the three phases of solar cycle 24

Phase	Shock		MC	
	Date	Time(UT)	Start time(UT)	Duration(h)
Rising phase	07/02/10	1425	07/1800	28
	12/02/10	0413	02/0800	19
	07/03/11	1542	06/0900	45
	05/06/11	0052	05/0200	17
	05/08/11	1920	05/0500	09
	09/09/11	1216	10/0300	26
	26/09/11	1308	26/2000	31
	25/10/11	0904	24/2200	18
	29/11/11	0256	29/0000	08
	29/11/12	0412	29/2200	15
Maximum phase	17/06/12	0120	16/2300	08
	15/07/12	0720	15/0600	13
	03/09/12	1156	04/1100	45
	07/03/13	0740	17/1000	15
	02/10/13	0420	02/2300	09
	08/10/13	0216	09/0900	23
	09/11/13	6504	08/2200	39
	08/12/13	0216	08/0100	09
	16/02/14	0400	16/0500	33
	06/04/14	0212	05/2200	11
	20/08/14	1104	19/1600	31
	12/09/14	1812	12/2200	37
	23/12/14	1045	22/0400	28
	17/03/15	1336	17/1300	13
	20/09/15	1536	21/0800	16
Declining phase	19/07/16	0700	19/1500	08
	24/07/16	1500	25/1600	25
	12/10/16	2212	13/0600	08
	08/09/17	1100	10/2100	38
	30/06/18	2000	02/1000	14
	16/05/19	2100	18/0800	11

(B) and its Y and Z components (By, Bz), solar wind bulk flow speed (V), proton temperature (Tp), and proton density (Np) data in GSE coordinates.

The relationship between ICMEs and the associated CMEs detected by SOHO/LASCO was made by looking at all CMEs that took place during a window of 0.5 to 5 days prior to the shock arrival time at 1 AU (Gopalswamy, 2010).

Figures 2, 3, and 4 each depict one instance of an ICME being driven by shock: an MC on July 15, 2012, at 07:20 UTC. The MC event took place during the decline stage. Whether the ICME signal is present or not, it is clear that the Tp is high in the sheath region between the shock and the start of the ejecta (this region is typical of all shocks). Remember that the Tp depression and the bulk of the other traits persist past the designated MC period, which may suggest that there were one or more non-MCs following the first MC.
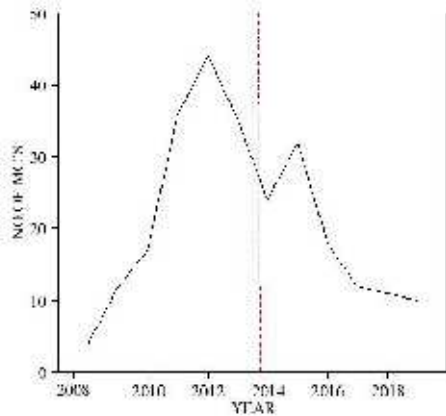
Fig. 1. Number of magnetic cloud during 2008-2019, X-axis indicates year and Y-axis indicates Number of MC. The data is taken from ACE. The Red grid line indicates the large number of MCs.

There are typically a few small variances in the identification of ICME limits by different authors, although these differences only last a few hours.

CME's heliospheric equivalents are ICMEs. When they are seen on corona graphs, CMEs are sometimes referred to as ICMEs. In Table 2, the first column shows the three phases data of solar cycle 24 and the three phases are rising phase, maximum phase and declining phase. The second column and the third column shows the IP shock occurring date and time. The time is denoted in Universal time. Fourth and fifth column shows the MC start time and duration and in Table 3, the first column shows the three phases data of solar cycle 24. Second and third column shows the CME event starting time in Universal time and duration. Fourth column shows the CME speed in kilometre per second. The fifth column shows the Measurement position angle (MPA). In this Table we take three important ICME and MC events, on that day the solar wind emitted from the Sun's outer atmosphere was very fast, so we have the giant magnetic flux rope, therefore IP shock reaches the high peak and also a more active MC was occurred.

2.3 ICME and MC event on 15 July 2012

The ICME shock's surface location is S17 W08, and it lasted from July 14 at 17:00 through July 17 at 5:00.The most favourable Magnetic cloud occurred on July 15, 2012, between 19:00 and 06:00 UT. The shock on July 15, 2012, was another instance of an extremely geo-powerful shock happening inside a CME (Liu et al., 2014). According to the radiation belt estimations from the Van Allen Tests, there has been a significant decrease in the velocity of energetic particles in the external radiation belt.

The magnetic activity of the 23rd solar cycle was greater than the 24th solar cycle at the time of observation, according to Tripathy et al. (2015). As a result, the number of CMEs was higher than during the 24th solar cycle. Despite a 40 percentage reduction in sunspot numbers, Gopalswamy et al.
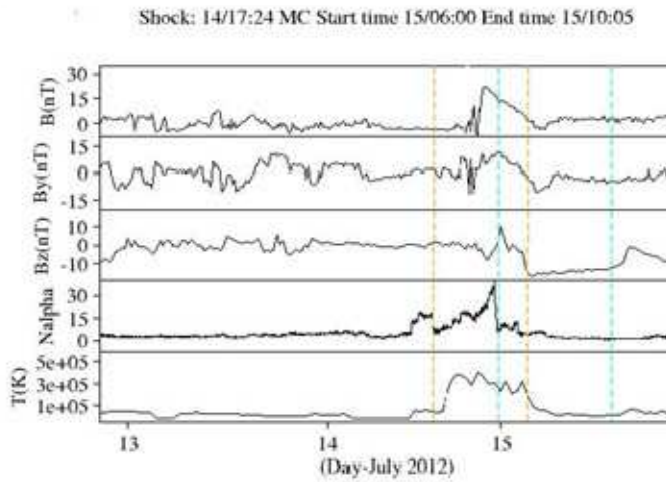
Fig. 2. ICME observed by ACE following the IP shock on 15 July 2012.
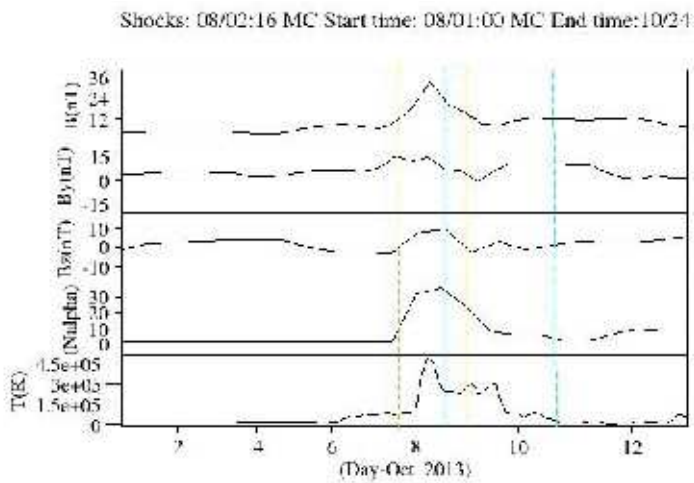
Fig. 3. ICME observed by ACE following the IP shock on 08 Oct. 2013.

Table 3. Disk center MC Events during the three phases of solar cycle 24

Phase	CME			
	Date	Time(UT)	V(km s^{-1})	MPA(deg)
Rising phase	02/07	0354	421	113
	02/12	1342	509	44
	03/07	2000	2125	313
	06/05	2205	2425	300
	08/05	0412	1315	298
	09/09	2305	575	300
	09/26	1248	1915	78
	10/25	1200	570	54
	11/29	1400	455	100
	11/29	1312	1064	134
Maximum phase	06/17	1412	987	144
	07/15	0624	873	329
	09/03	0400	538	90
	03/07	0712	1063	112
	10/02	0709	964	110
	10/08	1443	567	10
	11/09	1700	532	253
	12/08	0736	1085	274
	02/16	1000	634	227
	04/06	2312	514	115
	08/20	1112	600	359
	09/12	1800	1267	175
	12/23	1212	669	189
	03/17	0148	719	240
	09/20	1812	1239	219
Declining phase	07/19	0312	331	291
	07/19	1325	607	357
	10/12	0224	179	38
	09/08	1200	1385	72
	06/30	1248	185	261
	05/16	0912	305	79

(2015) observed that the rate of halo CMEs has increased during cycle 24 compared to cycle 23. It is crucial to remember that extraordinary occurrences can occur on 15 July 2012, with a magnetic field of -52nT and a speed of 1500 km s^{-1}, as a result of the greatest storm of solar cycle 24, which struck from 13 to 15 July 2012.

During the event, the magnetic field (Bz) is amplified and rotates smoothly. The blue vertical faint lines represent the magnetic clouds, and the orange vertical solid lines show the shock's arrival time at 1 AU. The ejecta were positioned at the conclusion of the study to allow for the measurement of the ICME interval. Two examples of ICMEs that induce shocks are shown in Fig. 2: on June 15, 2012, at 6:00 UT, an MC. The Tp depression is our main determinant of the ICME limitations. There are only two hours between our end time and the time of this report. It's intriguing to observe that most of the indications were present on June 15, 2012, up until about 6:00 UT, indicating
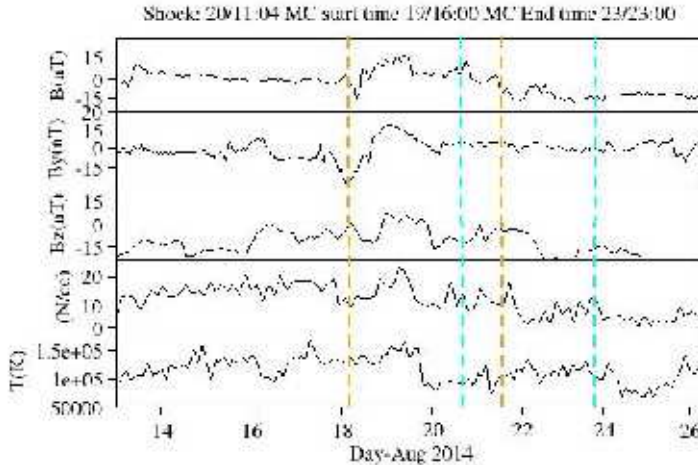
Fig. 4. Example of an ICME observed by ACE following the IP shock on 20 August 2014.

that further ejecta occurred after this MC event. For disk-centre events, the number of MCs is lowest during the rising phase and highest during solar maximum. The enhanced solar activity during solar maximum is the direct cause of this. A thorough examination of the five disk-centre CMEs linked to the solar origins of IP shock was provided by Gopalswamy 2010.

The entire magnetic field, Bx, which is roughly depicted in Fig. 2, abruptly spikes during the leading shock, which is how you can identify it. The in situ measurements and the computer operate essentially simultaneously. The peak values of velocity, total magnetic field, temperature, and density are remarkably close to the genuine values during the peak time. A rough estimate of the shocks passage time of 48 hours is also provided.

A decrease in magnetic field and a lower than expected proton temperature are indicators of the presence of the ICME magnetic cloud (Cane & Richardson, 1995, 2003). As it gets colder, the magnetic cloud starts to form. The magnetic cloud started on July 15, 2012, at 6:00 UT, according to this graph. The maximum plasma temperature that has been recorded throughout the time period under consideration is known as the maximum temperature. When the solar wind enters the heliosphere at a radial velocity faster than the solar wind it is going through, the solar wind is compressed on the leading edge of an ICME. The temperature increase during this compression is probably due to temperature gains along the leading edge of the ICME, which rely on the proton density and magnetic field (Light et al., 2020).

2.4   The Shock arrival time at 1 AU on 15 July 2012

Based on in-situ measurements, the shock's arrival time at 1 AU is shown by the blue-vertical solid lines (Mostl et al., 2014; Hess & Zhang, 2014). On July 14, at 17:24 UT, a shock and sheath area arrive at 1 AU simultaneously. The sheath region ended on July 15 at about 06:00 UT, and a magnetic cloud

started (Mostl et al., 2014). The shock arrival time from the simulation was only a little over an hour sooner than the data.

2.5   ICME and MC event on 08 Oct 2013

The ICME shock's surface position is S15 W12, and it occurred between 20:20 on October 8 and 00:00 on October 11.Between 0:00 and 24:00 on October 8, 2013, the magnetic cloud was more active. On October 8, 2013, the IP Shock occurred, causing the MC value to increase to a new peak. There aren't many active locations during solar minimum, and the filament disappearances take place around the equator. It has also been demonstrated that CMEs that are in close proximity to the solar minimum are frequently directed equatorially by the rapid solar wind flow that originates from huge pole coronal holes (Cremades & Bothmer, 2004). This shows that the majority of solar minimum CMEs, which have an MC structure, intersect the axis of the Earth. The majority of filament eruptions that take place at high latitudes during the solar optimum are either completely undeflected or mistakenly directed towards the poles (Cremades & Bothmer, 2004).

The magnetic field (Bz) is strengthened and rotates smoothly during the event. The magnetic clouds are depicted by the blue vertical faint lines, while the orange vertical solid lines indicate when the interplanetary shock will arrive. The list includes shocks generated by either MC shocks or IP shocks between December 2008 and December 2019. Each ICME's corresponding CME is monitored by the LASCO instrument on board the SOHO mission. The collected ICMEs are listed in Table 1 (MCs). Using information from Gopalswamy (2010), we divide the list of ICMEs in the declining phase into MCs.

MCs, which are among the most often used examples of ICMEs, are defined as ICMEs when the field intensity is high and the proton density and magnetic field are low mentioned by Bothmer & Schwenn (1998) and Sheeley et al. (1985). According to this diagram, interplanetary shocks start when the proton temperature starts to rise because as it does, so do the proton density and magnetic field. The IP Shock Event begins in this graph on October 7, 2013, at 20:24 UT. Ahead of them, shocks may be produced by ICMEs with sufficiently high proton temperatures in comparison to the background solar wind. Individual signatures, however, might not always be recognised in ICMEs (Cane & Richardson, 2003). An ICME may contain multiple signatures, but it does not mean that they are all coincident. When compared to other signatures, such as proton temperature depressions, some signs have been recorded comparatively infrequently (Cane & Richardson, 1995, 2003).

2.6   ICME and MC event on 20 August 2014

The surface position of the ICME shock in August 2014 was S17 W10, and it occurred between August 19 and August 21, both at 7:00 am. A long magnetic cloud lifetime is noticed on August 20, 2014. The results are a $440 \, \mathrm{km \, s^{-1}}$ (nine hours) and 16:30 UT (August 20, 2014) ICME appearance time and speed at 1 AU. IP Shock's high peak first emerged on August 20, 2014. Activities for ICME and MC on August 20, 2014. ICME and MC activities on August 20, 2014. The more active MC and IP shock on August 20, 2014, are depicted in

this graph. The data for the magnetic field Bx, By, Bz, proton density Nalpha, and proton temperature Tp used in this graph were obtained from the ACE satellite. The IP shock occurs between August 19 at 11 o'clock and August 21, 2014. The IP shock can be seen as an orange line. One of the biggest shock peaks has occurred in solar cycle 24. The magnetic field (Bz) is strengthened and rotates smoothly during the event. The magnetic cloud's arrival time is shown by the blue vertical faint lines and the IP shocks by the orange vertical solid lines. The largest peak occurred during the Sun's maximum phase, when it is at its most active. As the sun's temperature rises, its activity rises as well. If the Sun is more active, the proton temperature increases, the IP shock intensifies, and on August 20, 2014, when this particular event is moving at a speed of 860 to 960 $\mathrm{km\,s^{-1}}$, a more active and long-lasting MC forms. The wider CMEs in a slower background solar wind (McComas et al., 2013) may be subject to more drag in the IP medium, resulting in a slower MC speed at 1 AU, which is one explanation. In cycle 24, there are also fewer MCs that are shock drivers.

This graph depicts the MC that existed between August 18 and August 22, 2014. The proton temperature is quickly falling during that time. Since solar wind plasma and magnetic field data have been easily accessible since the dawn of the space age, the low solar wind proton temperature zones are suitable starting points for an ICME list for this ACE mission. We utilise the criterion of abnormally low proton temperature as the major ICME identification signature since we are interested in detecting all sorts of ICMEs (Richardson & Cane, 1993).

Conclusion

This article presents an in-depth analysis of the solar parameters, CMEs, and ICMEs that happened from 2008 to 2019. The 24th solar cycle's declining phase took place between 2016 and 2019, its ascending phase between 2012 and 2015, and its ascending and maximum phases between 2009 and 2012. Between the end of 2012 and the beginning of 2013, Solar Cycle 24 reached its highest peak.

There are 208 ICME events in total and 255 MCs in the current work. We have looked into the interactions between CMEs and MCs and IP shocks as they are observed by the LASCO instruments on board the SOHO spacecraft. The ICME catalogue is produced using a different methodology than other catalogues.

According to the study's analysis of current data, the 23rd solar cycle's CME number is higher than the 24th solar cycle's. The MC event happened during the Maximum period. As evidenced by two geomagnetic storms, the determined shock, and the MC, CMEs usually involve two subsequent collisions. Here, we report two cases where two successive impacts were blended into one, leading to a particularly unpleasant geomagnetic effect on an average strength driver. This kind of driver (consistent, increasing Bz) combines the traits of shocks and CMEs. The shock on July 15, 2012, had a number of ICME and shock features, as well as IP conditions. This result raises the possibility that fluctuations in the ICME's magnetic field with respect to the observer at 1 AU may be to blame for variations in the structure that have been observed between MC occurrences.

References

Bothmer, V., & Schwenn, R., 1998, *Ann. Geophysicae*, 16, 1-24

Brueckner, G. E., Howard, R. A., Koomen, M. J., et al., 1995, *Solar Physics*, 162, 357-402

Cane, H.V., & Richardson, I., 1995, *J. Geophys. Res.*, 100 (A2), 1755-1762

Cane, H.V., & Richardson, I., 2003, *J. Geophys. Res.*, 108 (A4), 1156

Chi, Y., 2016, *Solar Physics*, 291, 2419-2439

Cremades, H., & Bothmer, V., 2004, *Astron. & Astrophys.*, 422(1), 307-322

Gopalswamy, N., 2010, *Proceedings of the 20th National Solar Physics Meeting*, held 31 May - 4 June, 2010 in Papradno, Slovakia, p. 108-130

Gopalswamy, N., Akiyama, S., Yashiro, S., et al., 2008, *Journal of Atmospheric and Solar Terrestrial Physics*, 70(2-4), 245-253

Gopalswamy, N., Makela, P., Xie, H., Akiyama, S. & Yasshiro, S., 2010, *AIP Conf. Proc.*, 1216, 452-458

Gopalswamy, N., Yashiro, S., Xie, H., et al., 2015, *J. Geophys. Res.*, 120(11), 9221

Hess, P., & Zhang, J., 2014, *ApJ*, 792, 49

Huttunen, K.E.J., Schwenn, R., Bothmer, V., & Kooskinen, H.E.J., 2005, *Ann. Geophysicae*, 23, 625-641

Lepping, R. P., Wu, C.-C., & Berdichevsky, D. B., 2015, *Solar Physics*, 290, 553–578 comparison of magnetic cloud parameters, sunspot number, and interplanetary quantities for the first 18 years of the wind mission,

Light, C., Bindi, V., Consolandi, C., et al., 2020, *ApJ*, 896, 133

Liu, Y. D., Luhmann, J. G., Kajdič, P., et al., 2014, *Nature Communications*, 5(1), 3481

McComas, D. J., Angold, N., Elliott, H. A., 2013, *ApJ*, 779, 2

Möstl, C., Amla, K., Hall, J. R., et al., 2014, *ApJ*, 787(2), 119

Munro, R. H., Gosling, J. T., Hildner, E., et al., 1979, *Solar Physics*, 61, 201-215

Richardson, I., & Cane, H. V., 1993, *J. Geophys. Res.*, 98(A9), 15295-15304

Richardson, I., & Cane, H. V., 2010, *Solar Physics*, 264, 189-237

Sheeley Jr., N. R., Howard, R. A., Koomen, M. J., et al., 1985, *J. Geophys. Res.*, 90(A1), 163-176

Song, H., & Yao, S., 2020, *Science China Technological Sciences*, 63 (11), 2171-2187

Stone, E. C., Frandsen, A. M., Mewaldt, R. A., et al., 1998, *Space Sci. Rev.*, 86(1-4), 1-22

Tripathy, S. C., Jain, K., & Hill, F., 2015, *ApJ*, 812, 20

Wu, C. C., & Lepping, R. P., 2011, *Solar Physics*, 269, 141-153

Zurbuchen, T. H., & Richardson, I., 2006, *Space Sci. Rev.*, 123, 31-43

Full Length Article

Density functional theory calculations, vibrational spectral analysis and topological analysis of 1-acethyl-2(4-isopropoxy-3-methoxyphenyl) cyclopropane with docking studies

D.P. Lydia Renj [a,i], R. Racil Jeya Geetha [b,i], A. Benifa [c,i], M. Amalanathan [d,], [*], M. Sony Michael Mary [e], Zoran Ratković [f], Jovana Muškinja [g], Jayasekar Micheal [h]

[a] Research Scholar, Reg. No: 19223112132013, Department of Physics & Research Centre, Nesamony Memorial Christian College, Marthandam, Tamil Nadu 629165, India
[b] Department of Physics & Research Centre, Nesamony Memorial Christian College, Marthandam, Tamil Nadu 629165, India
[c] Research Scholar, Reg. No: 23111172132003, Rani Anna Government College for Women, Gandhi Nagar, Tirunelveli, India
[d] Department of Physics & Research Centre, Nanjil Catholic College of Arts & Science, Kaliyakkavilai-629 153, Tamil Nadu, India
[e] Department of Physics, Annai Velankanni College of Engineering, Kanyakumari District, Azhagappapuram, Tamil Nadu 629 401, India
[f] Department of Chemistry, Faculty of Sciences, University of Kragujevac, Radoja Domanovića 12, Kragujevac 34000, Serbia
[g] Department of Sciences, Institute for Information Technologies, University of Kragujevac, Jovana Cvijića bb, Kragujevac 34000, Serbia
[h] PCB Design Engineer, Intel Corporation, Santa Clara, CA 95054, USA
[i] Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli, Tamil Nadu 627 012, India

ARTICLE INFO

ABSTRACT

A systematic spectroscopic investigation of 1-acethyl-2(4-isopropoxy-3-methoxyphenyl) cyclopropane was performed by utilizing Density functional theory approaches at B3LYP level using Gaussian 09 W software package. The FT-IR and FT-Raman techniques were utilized to assign the spectral properties of the title compound. On the basis of Natural Bond Orbital (NBO) analysis, the transfer of second order perturbation energies and Electron Density (ED) from filled lone pairs of Lewis base to unfilled Lewis acid sites were analysed. The chemical stability, distribution of energy and energetic behaviour of the compound were calculated from the Highest Occupied and Lowest Unoccupied Molecular Orbital (HOMO-LUMO) Analysis. The nucleophilic and electrophilic locales of the molecule was perceived by the Molecular electrostatic potential (MEP). NCI investigation gives data around the inter and intra non covalent interlinkages. By using the Multiwavefunction software the topological analysis of ELF and LOL were performed. The chemical reactivity sites were determined by means of the fukui function. The assignments of vibrational spectra were computed using the vibration energy distribution analysis (VEDA). Drug similarity factors were intended to understand the biological aspects. Vadar software was used to generate the Ramachandran plot. The bioactivity of the title compound was confirmed from the molecular docking studies.

Introduction

Cyclopropane derivatives are fascinating compounds for theoretical studies because of their similarity to biological species [1]. The biological and pharmaceutical properties of cyclopropane and its derivatives have recently acquired prominence due to their notable antimicrobial activities [2]. Cyclopropane ring systems are present in several natural products, pesticides and medicamental drug candidates

as they are omnipresent in nature. Recent studies had proved that the cyclopropane analogues showed various biological executions like anti-HIV, anticancer, antibacterial, antifungal, antiviral, antitumor, COX-II inhibitory properties [3].

The acethyl group is sometimes referred to as a fraction, incorporated by a methyl group linked to a carbonyl group. The carbonyl group is made up of a carbon atom and an oxygen atom doubly bonded with each other. Studies had proved that the presence of acethyl group makes
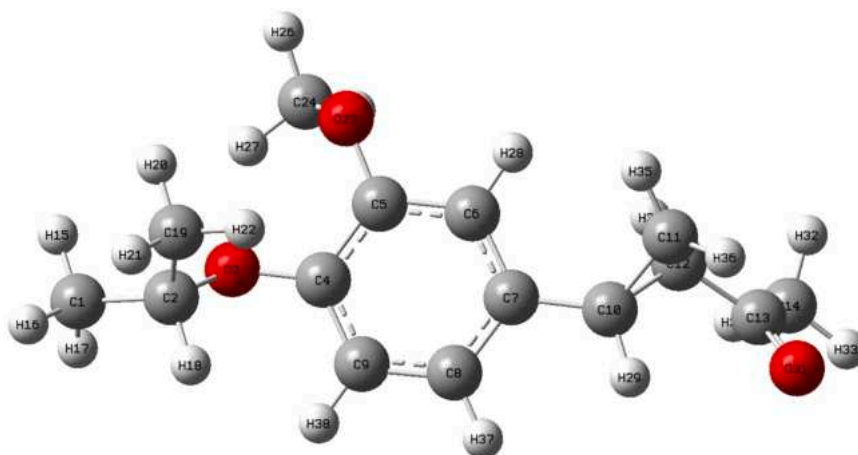
Fig. 1. Optimized molecular structure of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

the drug's effects more intense and increases the effectiveness of a given dose [4].

Numerous natural,organic, synthetic and chemical substances contain phenyl groups [5]. Similar to all other aromatic compounds, phenyl groups are significantly more stable than aliphatic (non-aromatic) groups. The unique characteristics of molecular orbitals in aromatic form are responsible for this enhanced stability [6]. Phenyl groups has pharmacological qualities and is utilised as an anti-inflammatory, analgesic, choleretic, etc. [7].

The emergence of swift computers accompanied by advanced computational techniques has simplified the obligation of resolving numerous chemical and structural issues [8]. Ab initio DFT estimations had transformed into a useful method for predicting biological complex molecular structure, harmonic force fields, vibrational wavenumbers, and FT-IR &Raman activities. ground state [9].

This study used the DFT calculation approach to determine the most stable structure.Currently, the ab-initio community recognises the density functional theory (DFT) approach as a reliable post-HF strategy for estimating molecular characteristics [10]. DFT assists in attaining a suitable balance between computational duration and result accuracy [11]. Specifically, using DFT methods for polyatomic compounds results in more precise molecular structure predictionsand vibrational wavenumbers than the conventional operations [12].

In the present study, the title compound, 1-acethyl-2(4-isopropoxy-3-methoxyphenyl) cyclopropane(AMC) have been investigated using B3LYP calculations with 6- 311++G(d,p) basis set. A comprehensive investigation of the literature reveals that there hasn't been any in-depth empirical and conceptual research on the title compound until now.

The Redistribution of electron density (ED) in different bonding, antibonding, and E(2) energies is computed using the natural bond orbital (NBO) approach. Frontier molecular orbital (FMO) energies are accustomed to examine the electronic characteristics of the title compound [13]. Topology analyses were obtained from Multiwfn software which is a wave function analyser software.

Fukui parameters and chemical response descriptors were used to study the reactivity of molecules and the majority of reactive sites [14]. Molecular docking analysis is a practical and efficient computational technique for anticipating a ligand's binding mechanismand propensity for proteins [15]. Molecular docking was performed using pyrx and pymol software.

Materials and methods

The titled compound 1-acethyl-2(4-isopropoxy-3-methoxyphenyl) cyclopropanewas purchased from Sigma-Aldrich with a 99 percent purity level. It was utilized without being processed further. Using a

PERKIN ELMER FTIR spectrometer in discharge mode and KBr pellet method, the Fourier Transform Infrared spectrum of the substance under investigation in the solid phase was acquired at 1.0 cm^{-1} resolution within the 4000–450 cm^{-1} range. Also, FT-IR and FT-Raman spectra were theoretically calculated using the DFT software.

Computational details

The use of computational chemistry as a tool for designing and developing organic pharmaceuticals has greatly increased with the emergence of density functional theory (DFT).The Gaussian 09 W software package was employed to complete computational calculations utilising B3LYP correction with the 6–311G++ (d, p) basis set [16]. The basic sets B3LYP 6- 311++G (d, p) was augmented by a 'd' polarization function for heavy atoms and 'p' polarisation for hydrogen atoms [17]. Bond length, bond angle, and dihedral angles were determined as well using the DFT-B3LYP level with the 6–311++G(d,p) basis set. Donor-acceptor interactions in natural bond orbitals (NBO) have been determined using a secondary order Fock matrix,by employing NBO version 3.1 which is launched in the Gaussian 09 W package [18]. The molecular structure was visualised with the help of Gauss View 5.0.8 software. Multiwfn 3.7 software was employed to carry out AIM, Electron Localization Function (ELF), Localized Orbital Locator (LOL), fukui and Non-Covalent Interaction (NCI) analysis [19]. Open Babel, Py Rx and Pymolsoftwares were used to perform the molecular docking studies.Potential Energy Distribution (PED) Analysis was performed with the aid of the VEDA software conceptualised by Jamroz [20]. Volume Area Dihedral Angle Reporter, or VADAR, is an extensive web server for quantitative analysis of protein structures [21]. It was used to generate the Ramachandran plot.

Results and discussion

Molecular geometry

Molecular geometry refers to the spatial arrangement of atoms contained in a molecule.Molecular geometry offers crucial insights into the overall shape of the molecule, as well as key geometric variables such as bond lengths, bond angles, torsional angles, and the positions of individual atoms. Molecular geometry has a profound impact on various properties of compounds like its reactivity, orientation, the state of matter, hue, magnetism and biological functions [21].

The title compound, 1-acethyl-2(4-isopropoxy-3-methoxyphenyl) cyclopropane(AMC), is a cyclopropanecompound with specific substituents. This compound is characterized by the presence of three distinct groups: acethyl, isopropoxy, and methoxyphenyl.Before doing a

Table 1

Second order perturbation theory analysis of Fock matrix in NBO basis of 1-acetyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

Donor (i)	ED	Acceptor (j)	ED	E (2) Kcal/mol	E(j) –E (i) a.u	F(i, j) a.u
$\pi C_4 - C_9$	1.67317	$\pi^* C_5 - C_6$	0.36620	20.07	0.29	0.069
$\pi C_4 - C_9$	1.67317	$\pi^* C_7 - C_8$	0.37292	18.89	0.30	0.068
$\pi C_5 - C_6$	1.67931	$\pi^* C_4 - C_9$	0.37999	19.33	0.29	0.067
$\pi C_5 - C_6$	1.67931	$\pi^* C_7 - C_8$	0.37292	20.49	0.30	0.070
$\pi C_7 - C_8$	1.65615	$\pi^* C_4 - C_9$	0.37999	21.49	0.27	0.06
$\pi C_7 - C_8$	1.65615	$\pi^* C_5 - C_6$	0.36620	19.99	0.28	0.067
$\sigma C_{14} - H_{32}$	1.97111	$\pi^* C_{13} - O_{31}$	0.13240	5.00	0.54	0.047
$LP(1)O_3$	1.95345	$\sigma^* C_4 - C_9$	0.02860	6.27	1.07	0.073
$LP(2)O_3$	1.90031	$\sigma^* C_2 - H_{19}$	0.02232	5.73	0.74	0.059
$LP(2)O_3$	1.90031	$\pi^* C_4 - C_9$	0.37999	12.02	0.38	0.065
$LP(1) O_{23}$	1.95231	$\sigma^* C_4 - C_5$	0.04445	6.10	1.04	0.071
$LP(2) O_{23}$	1.90676	$\pi^* C_5 - C_6$	0.36620	9.37	0.39	0.058
$LP(2) O_{23}$	1.90676	$\sigma^* C_{24} - H_{25}$	0.01810	5.69	0.74	0.059
$LP(2) O_{31}$	1.88807	$\sigma^* C_{12} - C_{13}$	0.06597	19.20	0.69	0.104
$LP(2) O_{31}$	1.88807	$\sigma^* C_{13} - C_{14}$	0.05457	20.21	0.65	0.104
$\sigma C_{10} - C_{12}$	1.89000	$\sigma^* C_{11} - C_{12}$	0.03255	6.41	0.84	0.067
$\sigma C_{10} - C_{12}$	1.89000	$\pi^* C_{13} - O_{31}$	0.13240	5.23	0.54	0.048
$\sigma C_{11} - C_{12}$	1.97782	$\sigma^* C_{10} - C_{11}$	0.03075	5.33	0.89	0.062
$\sigma C_{11} - C_{12}$	1.97782	$\sigma^* C_{10} - C_{12}$	0.04948	6.85	0.84	0.068
$\pi^* C_4 - C_9$	0.39416	$\pi^* C_7 - C_8$	0.37292	291.24	0.01	0.081

straight away geometry optimization, a conformational analysis has to be carried out. The most stable conformer must be subjected to geometry optimization. Here, the conformational analysis is performed using PES scan method. The minimum energy conformation is used for the geometry optimization.

The idealised arrangement of the compound AMC was produced with the aid of Gaussian09 software is shown in Fig. 1. The optimal structural variables such as bond lengths, bond angles and dihedral angles calculated by the Becke,3-parameter, Lee-Yang-Parr method DFT–B3LYP level with the 6–311++G(d,p) basis set are displayed in Supplementary Table 1. The sequence of atom numbering used in Fig. 1 is followed to label the atoms in the table [22].

The AMC compound comprised of 15 Carbon atoms, 3 Oxygen atoms and 20 Hydrogen atoms having Carbon-Carbon, Carbon-Oxygen and Carbon-Hydrogen bonds. Based on the derivedvalues, the bond lengths of C–C atoms within the phenyl ring are discovered to be close around 1.4 A° (C$_4$–C$_5$, C$_5$–C$_6$, C$_6$–C$_7$, C$_7$–C$_8$) which signifies the indistinctness of the single bond as well as the double bond due to the π electron conjugation [23].

The slight increase in the bond length of the C–C atoms from 1.4 A° to 1.49 A° in the cyclopropane ring (C$_7$–C$_{10}$, C$_{10}$–C$_{11}$, C$_{12}$–C$_{13}$) is due to the heavy substitution of hydrogen atoms around cyclopropane ring. The increase in the bond length of the C–C atoms from 1.4 A° to 1.52 A° (C$_1$–C$_2$, C$_2$–C$_{19}$, C$_{10}$–C$_{12}$, C$_{13}$–C$_{14}$) is due to the impact of neighbouring hydrogen atoms.

The C—H bonds in the isopropoxy and acethyl groups are found to be 1.09 Å, which indicate that there is no difference in electronic distribution in these groups [24]. The CH bonds in the phenyl group is found to be unaffected due to these substitutions. The slight change in the bond length of the C$_{10}$–H$_{29}$(1.0846 A°), C$_{11}$–H$_{35}$(1.0834 A°), C$_{11}$–H$_{36}$(1.0833 A°), C$_{12}$–H$_{30}$(1.0839 A°) atoms is due to the cyclopropane ring structure.

The decrease in length of the bond of C$_{13}$–O$_{31}$ (1.2169A°)atoms

shows the presence of double bond within the carbon and oxygen atoms. The increase in the bond length from 1.23 A° to 1.36 A° (O$_3$ – C$_4$, C$_5$ – O$_{23}$) is because of the consequences of double bonded atoms in the ring arrangement.

The bond angle lowers as the dimension of the central atom rises, and improves when the dimension of the ligand atom increases [25]. Endocyclic C—C-C bond angles in the phenyl ring have been estimated to be about 120.0°, with a little reductionin C$_4$ – C$_5$ – C$_6$ (119.8263°),C$_6$–C$_7$ – C$_8$ (119.8807°)and small increases in C$_7$ – C$_8$ – C$_9$(121.9772°), C$_5$–C$_6$ – C$_7$ (121.6839°)are witnessed as a result of ring substitution that does not share electrons.

The bond angle of C$_{12}$ – C$_{13}$ – O$_{31}$ (121.6882°) is greater than C$_{12}$ – C$_{13}$ – C$_{14}$ (116.4218°) which shows that the oxygen being more electronegative than carbon atom [26]. The isopropoxy and acethyl group bond angles are in the range 110° because of the positive hydrogen atoms. The decrease in the exocyclic bond angles of the phenyl ring is due to the repulsion between the atoms.

The hyperconjugation (n→σ*) between the lone-pair electrons of the O$_{23}$ atom and the cyclopropane group contracts the internal bond angle of C$_{10}$–C$_{12}$–C$_{13}$ (117°).The torsional angles of the phenyl ring (C$_4$ – C$_5$ – C$_6$ – C$_7$, C$_4$ – C$_5$ – C$_6$ – H$_{28}$, C$_5$ – C$_6$ – C$_7$ – C$_8$, C$_5$ – C$_6$ – C$_7$ – C$_{10}$, C$_7$ – C$_8$ – C$_9$ – H$_{38}$, H$_{37}$ – C$_8$ – C$_9$ – C$_4$, C$_6$ – C$_7$ – C$_8$ – C$_9$) of AMC compound is planar which indicates that no twisting occurs within the phenyl ring and substituted groups [27].The torsional angles of H$_{15}$ – C$_1$ – C$_2$ – H$_{18}$(−178.626°), C$_5$ – O$_{23}$ – C$_{24}$ – H$_{26}$ (−179.5801°)and C$_1$ – C$_2$ – O$_3$ – C$_4$ (−165.3647°) shows folded conformations of the molecule [28].

Natural bond orbital (NBO) analysis

Natural bond orbital (NBO) assessment is a powerful aid in the field of computational chemistry that can derive a molecule's precise Lewis structure. It is performed using Non-Binding Organization (NBO) 5.0 Program.It allows for the identification and quantification of the finest feasible proportion of electron density throughout an orbital. It is a sensitive instrument for internal and intermolecular interaction analysis by taking into account of all the potential interactions between full donor orbitals and empty acceptor orbitals [29].

The E$^{(2)}$ value plays a crucial role in characterizing the degree to which the electron donors interact with the acceptors.It provides valuable insights onto the degree of conjugation in a molecular system.As the E$^{(2)}$ value becomes larger, the electron donor-electron acceptor interaction becomes more intense [30].

For each of the donor NBOs (i) and the acceptor NBOs (j), the electron delocalization stabilization E$^{(2)}$ between the donor and the acceptor is approximated to be

$$E^{(2)} = \frac{q_i \left(F_{i,j}\right)^2}{\epsilon_j - \epsilon_i}$$

Where, q_i = orbital occupancy, ε_i = diagonal element, ε_j = off diagonal element, $F_{i,j}$= off diagonal NBO Fock matrix element [31]. The result of NBO analysis of the title compound analysed with the aid ofthe GAUSSIAN 09 software is summarized in Table 1.

The overlapping of orbitals between the π(C—C) and π*(C—C) bonds produce the intra-molecular hyper-conjugation which results in the transfer of intramolecular charges and stability of the compound [32]. In the title compound, the hyper-conjugative interactions of π(C$_4$ – C$_9$)→ π* (C$_5$ – C$_6$), π(C$_4$ – C$_9$) → π*(C$_7$ – C$_8$), π(C$_5$ – C$_6$)→ π*(C$_4$ – C$_9$),π(C$_5$ – C$_6$)→ π*(C$_7$ – C$_8$),π(C$_7$ – C$_8$) →π*(C$_4$ – C$_9$),π(C$_7$ – C$_8$) →π*(C$_5$ – C$_6$) are found to be 20.07, 18.89, 20.49, 21.49 and 19.99 kcal/mol, respectively. Strong electron delocalization is observed at conjugated bonds ED at π bonds and π* bonds in the phenyl ring, leading to the energy stabilisation.

Contributions from other interactions σ→σ* were also observed. Furthermore, the creation of other bonds that result in the stable crystal structure of all compounds is greatly influenced by these interactions [33]. Also, the stabilisation of the compound is greatly aided by the
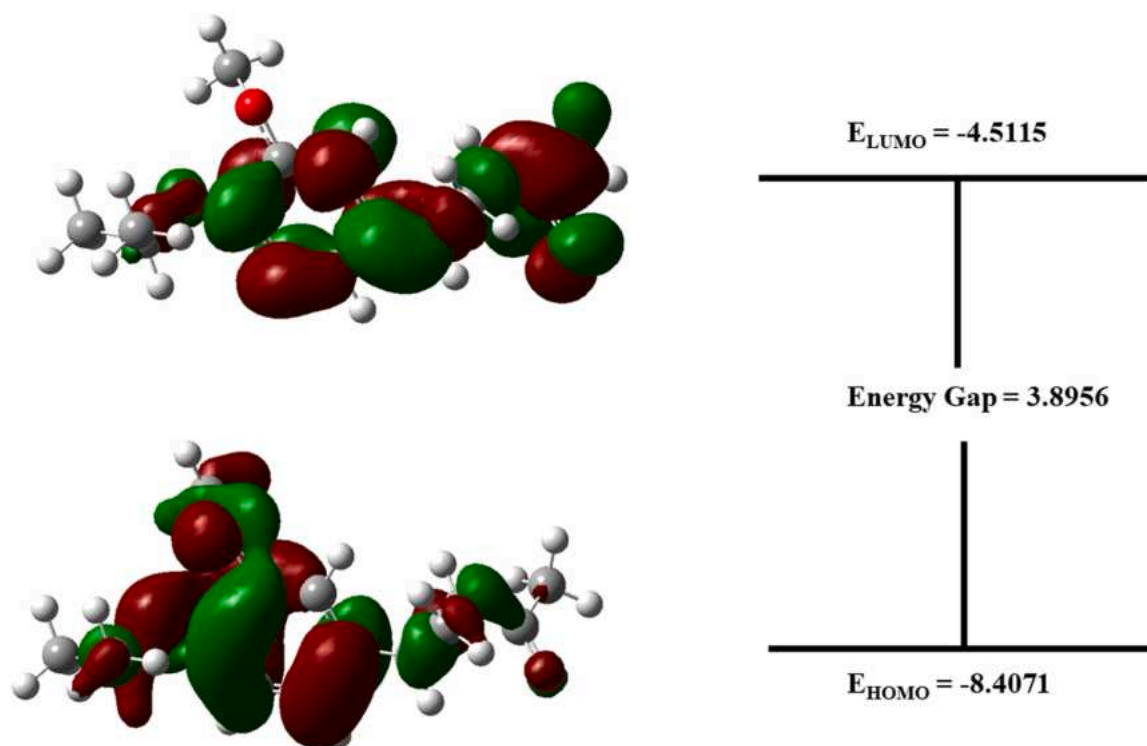
$E_{LUMO} = -4.5115$

Energy Gap = 3.8956

$E_{HOMO} = -8.4071$

Fig. 2. HOMO-LUMO plot of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

interlinkage of the electron lone pair on the oxygen atoms. It is evident by the hyper-conjugative interactions of LP(2) $O_3 \rightarrow \pi^*(C_4 - C_9)$, LP(1) $O_{23} \rightarrow \sigma^*(C_4 - C_5)$ and LP(2) $O_{23} \rightarrow \pi^*(C_5 - C_6)$.

Mulliken atomic chargesand natural population analysis

Mulliken population analysis is a widely used technique in theoretical chemistry to compute Mulliken atomic charges. This analysis assists inunderstanding the distribution of electron concentration within a molecule and provides insights into the reactivity and chemical behaviour of atoms [34]. It enables the analysis and prediction of various molecular properties, including the dipole moment, polarizability, electronic structure, and so on [35].

The reactive behaviour of numerous chemical systems can be deciphered and predicted using the Mulliken Population Method in both electrophilic and nucleophilic reactions [36]. InAMC compound, the carbon atoms in the cyclopropane ring structure (C_{10} and C_{12}) and attached to the oxygen atom O_{31} are positive, whereas other carbon atoms remain negative. As the oxygen atoms are electronegative, they possess negative mulliken atomic charges ($O_3 = -0.0521$; $O_{23} = -0.116752$; $O_{31} = -0.289791$) and acts as donor atoms.

Mulliken atomic charge obtained shows that $H_{29}(0.276953)$ atom has acquired a higher positive atomic charge relative to the other carbon atoms, oxygen atoms and hydrogen atoms due to more negatively charged neighbouring atoms.Furthermore, all hydrogen atoms have positive charges and serve as acceptor atoms. The carbon atoms in benzene rings are negative as they share the electrons within the ring equally due to conjugation [37].

The natural population analysis (NPA) of an organic compound shows the distribution of electrons in various atomic orbital subshells [38]. One of the key advantages of Natural population Analysis over Mulliken population is that it can detect reactive sites as well as other molecular interactions (including charge transfer) by treating natural bond orbitals instead of treating all orbitals and being affected by basis sets [39].

The calculated values of Mulliken charge distribution and natural

population Analysis of individual atom by NBO method utilizing the GAUSSIAN 09 software B3LYP/6–311++ G(d,p) basis set are tabulated in Supplementary Table 2. The visual representation of Mulliken atomic charges and Natural population analysis is shown in Supplementary Fig. 1.

According to the electrostatic perspective, more electro positivity atoms are able to absorb electrons, and more electronegativity atoms can readily give electrons. In the title compound, carbon (C_{13}) atom has the more electropositive charge and oxygen atom (O_{31}) has the more electronegative charge. This is due to the donation and acceptance of electrons between those atoms.

The natural energies of the AMC compound focused on the sub-shells implies that:

Core: 35.98547 (99.9597 % of 36)
Valence: 97.64575 99.6385 % of 98)
Rydberg: 0.36878 (0.2752 % of 134)

Frontier molecular orbital analysis

One of the finest chemical stability hypotheses is the FMOs theory using HOMO and LUMO [40]. HOMO is the maximum occupied molecular orbital whereas, LUMO is the minimum occupied molecular orbital. The energy of the LUMO and HOMO molecules, along with their energy gap, indicate the molecule's chemical function and stability [41].

The title compound's LUMO and HOMO orbitals exhibit an even distribution of electron density, potentially leading to more interactive characteristics [42]. The HOMO and LUMO energies offer details on the distribution of energy and energetic behaviour. The E_{HOMO} and E_{LUMO} values having negative magnitude gives the stability of compounds [43].

The HOMO and LUMO energy describe the possibility of donating an electron and receiving an electron.The higher the HOMO energy, the more reactive the molecule is in the electrophilic reactions while the lower the LUMO energy is for the molecular reactions with nucleophile [44]. The energy difference between the HOMO and LUMO is the HOMO–LUMO gap also called the energy gap. It also symbolises the molecule's biological activity.

Table 2

Global Reactivity descriptors of 11-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

Molecular Properties	Mathematical Description	Energy(eV)
E_{HOMO}	Energy of HOMO	−8.4071
E_{LUMO}	Energy of LUMO	−4.5115
Energy Gap	$\Delta E_g = E_{HOMO} − E_{LUMO}$	3.8956
Ionization Potential (I_P)	$IP = − E_{HOMO}$	8.4071
Electron Affinity (E_A)	$EA = − E_{LUMO}$	4.5115
Electronegativity (χ)	$\chi = − 1/2(E_{LUMO} + E_{HOMO})$	6.1513
Chemical potential (μ)	$\mu = 1/2(E_{LUMO} + E_{HOMO})$	−6.1513
Global Hardness (η)	$\eta = 1/2(E_{LUMO} − E_{HOMO})$	1.9478
Softness (S)	$S = 1/2\eta$	0.2566
Electrophilicity index (ω)	$\omega = \mu^2/2\eta$	9.7131

The HOMO-LUMO energies of AMC compound was −8.4071 eV and −4.5115 eV, respectively. The energy gap was found to be 3.8956 eV. The HOMO cloud is focused on the phenyl ring and methoxy group, with limited coverage for isopropoxy and cyclopropane groups. Similarly, the LUMO cloud concentrates on phenyl ring, cyclopropane, and acethyl group.The graphical representations of HOMO and LUMO are presented in Fig. 2.

Molecules with a large HOMO - LUMO gap are characterized as hard molecules, small molecules, and highly non-polarizable. Similarly, molecules with a small HOMO - LUMO gap are characterized as soft molecules, large molecules, and highly polarizable [45]. A small HOMO-LUMO gap is associated with anti-aromaticity, whereas a large gap depicts a high level of molecular stability and a low level of chemical reactivity. [46].

Global reactivity descriptors

In DFT the commonly used global reactivity parameters are Electronegativity(χ), electrophilicity index(ω), hardness (η), softness (S) etc. Recently, reactivity quantities were established to provide a concise review of harmful contaminants and their selectivity at the location. A compound's molecular stability is determined by its hardness and flexibility, which are dependable indicators of chemical stability.

The chemical hardness parameter(η) is helpful in understanding how the chemical systems behave.Electro negativity refers to an atom's tendency to strongly entice electrons to form covalent bonds with it(χ). The term chemical potential(μ) refers to the escaping tendency of electrons. The higher the electronic chemical potential, the more unstable or reactive the compound is [47].

Ionization Potential (I_p) is the amount of time it took to liberate an electron from the molecular system and convert it into a positively-charged gas ion. Electron affinity (E_A) is defined as the rate of energy evolved when electrons are added to a neutral gas atom to convert it into a negatively charged gas ion [48]. For a molecular system, the global softness (S) is the reciprocal of the global hardness (η) [49].

Global hardness is a measure of the electron distribution resistance of a group of nuclei and electrons [50]. Electrophilicity is used as a structural depicter for chemical reactivity analysis of molecules.It's a measure of how likely a species is to accept the electrons. If the value of (ω) is low then the compound is nucleophile, similarly a high value of (ω) gives good electrophile [51]. The compound's Global Reactivity descriptorsare listed in Table 2.

For the title compound, the electrophilicity index(ω) is calculated as 9.7131. The higher electrophilicity index value, indicates that it can act as an electrophilic species and has a greater ability to bind to bio-molecules [52]. The Global hardness (η) and Chemical potential (μ) values are calculated as 1.9478 and −6.1513, respectively. The title compound is a soft compound with strong polarizability, as indicated by the lower global hardness value and high negative chemical potential [53].
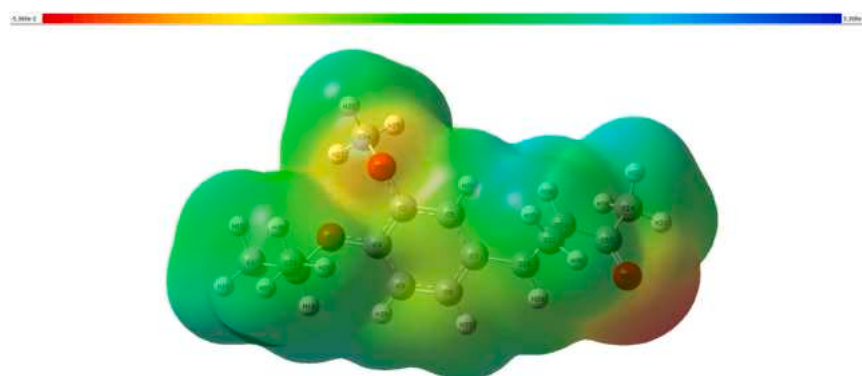
Molecular electrostatic potential

Molecular electrostatic potential (MEP) is a map of the electrostatic potential on a constant electron density surface [54]. It is primarily used for site prediction and electrophilic/nucleophilic reactivity prediction. It is also usedin the prediction of biological signals and molecular interactions. [55]. Intermolecular interactions cause the negative atom's electrostatic potential to diminish and the positive area of the other atom to become less positive [56].

Different hues represent different electrostatic potential levels. Red and green shades indicate the least electrostatic potential, caused by a large concentration of electrons, which attracts the proton.While the highest electrostatic potential is revealed by the blue regionsdue to proton repulsion in locations with low electron density and partial shielding of the nuclear charge [57]. DFT calculations employing the optimised structure with B3LYP/6–31G+(d,p) basis set was utilised to estimate the MEP surface analysis of AMC compound. The MEP map is shown in the Fig. 3.

The range in which the title compound has its colour code is of $−5.369e^{−2}$ to $+5.369e^{−2}$.Oxygen atoms are more electronegative in the title compound as indicated by the red colour and is prone to the nucleophilic attack. Similarly, the blue colour regions indicate the electrophilic attack.The title molecule's interactions and active sites indicate the compound's biological activity.

Non covalent interaction (NCI) analysis

The non-covalent interactions sites in actual space are graphically visualised using the NCI analysis approach. A distinction was made between hydrogen bonding, Van der Waals interaction, and steric repulsion interactions with the aid of NCI approach [58]. Johnson et al. examined theRDG techniques along with real spatial weak interactions based on electron densities and their derivatives [59]. The RDG function

Fig. 3. Molecular electrostatic potential map of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.
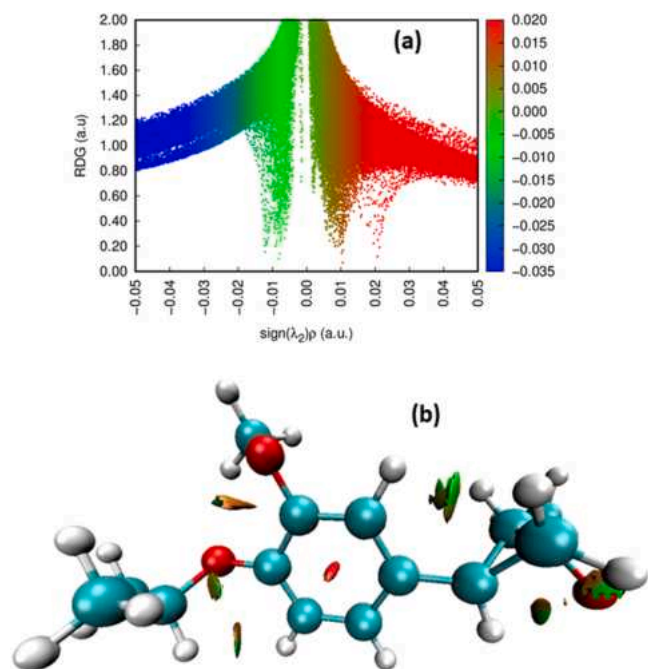
Fig. 4. (a) Reduced Density Gradient scatter plot of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane. (b) Reduced Density Gradient structure of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

can be written as

$$RDG = \frac{1}{2(3\pi^2)^{1/3}} \frac{\nabla\rho(r)}{\rho(r)^{4/3}}$$

where $\rho(r)$ denotes the electron density [60]. Multiwfnsoftware and VMDprogram were used to calculate the results.Gnuplot software was used to plot the RDG graph. Thegradient isosurfaces and RDG coloured scatter graph were shown in the Fig. 4.

The NCI plot is calculated by plotting the RDG (reduced density gradient) against the sign of $\lambda 2$ (eigenvalues of the electron density of the Hessian matrix) multiplied by the density,ρ [61]. The nature of interaction is predicted from (sign $\lambda 2)\rho$ value. The value of (sign $\lambda 2)\rho$for

a repulsive interaction is greater than 0 [(sign $\lambda 2)\rho > 0$] and it is denoted by red colour. For Vander Waals interaction the value of (sign $\lambda 2)\rho$ is equal to 0 [(sign $\lambda 2)\rho = 0$] and it is denoted by green colour. For hydrogen bonding the value of (sign $\lambda 2)\rho$ is less than 0 [(sign $\lambda 2)\rho < 0$] and it is denoted by blue colour [62].

For the title compound, the red coloured interactions in the RDG isosurface map depicts the steric influence in the range of (0.01–0.05) a. u. Similarly, the blue and green colour interactions show the hydrogen bonding (−0.05 a.u. to −0.02 a.u.) and Vander Waals effect (−0.02 a.u. to 0.01) respectively.RDG isosurfaces exhibit a red spike indicating steric repulsion in the middle of the phenyl ring. The existence of the combined red and green peaks suggests a modest non-covalent interaction between the bonds.

Electron localization function (ELF) and localized orbital locator (LOL) analysis

The topological analysis of ELF and LOL were performed by the Multiwfn software. The Electron localization function (ELF) was used with the relief map and the LOL maps were utilised to carry out surface analysis based on covalent bonding [63]. Each atom's electron environment was outlined by a relief map with a large or small peak area. These raise the prospect of discovering an electron pair on the molecular surface [64].

Colour codes represent ELF values as red for high, yellow to green for medium, and blue for low [65]. Increased electron localization in a location result in increased ELF, while lower ELF leads to reduced repulsion [66]. The reds and oranges hues, critical locations, trajectories, chemical bonds, and chemically relevant regionsnear hydrogen atoms in the ELF map depicts localised bonding or non-bonding electrons.

ELF and LOL share the same chemical conflict due to their reliance on kinetic energy density. However, ELF is derived from the electron pair density. LOL merely displays the gradient of localized orbitals, and is utilised when localized orbits overlap [67]. The bonded and non-bonded electrons represented in the ELF and LOLcolour shade map for the title compound is given in Fig. 5.

The ELF map is formed between the values of 0.0 and 1.0. The LOL earns greater values > 0.5 in regions where electron density is dominated by electron localization [68]. The delocalized electron domainis of the range below 0.5. The delocalized electron cloud around carbon
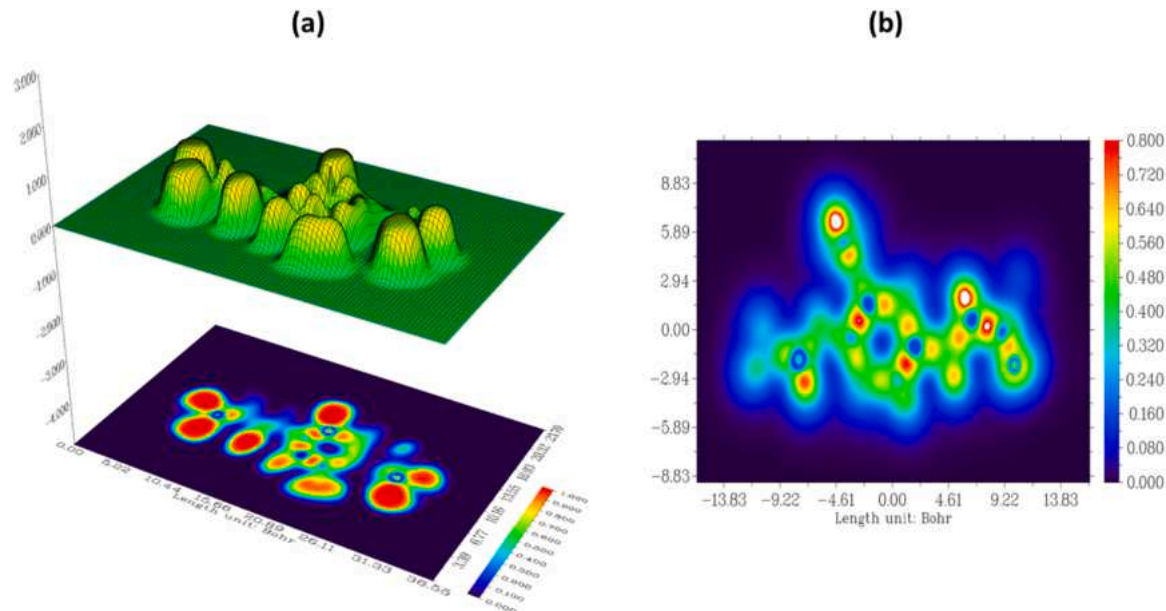

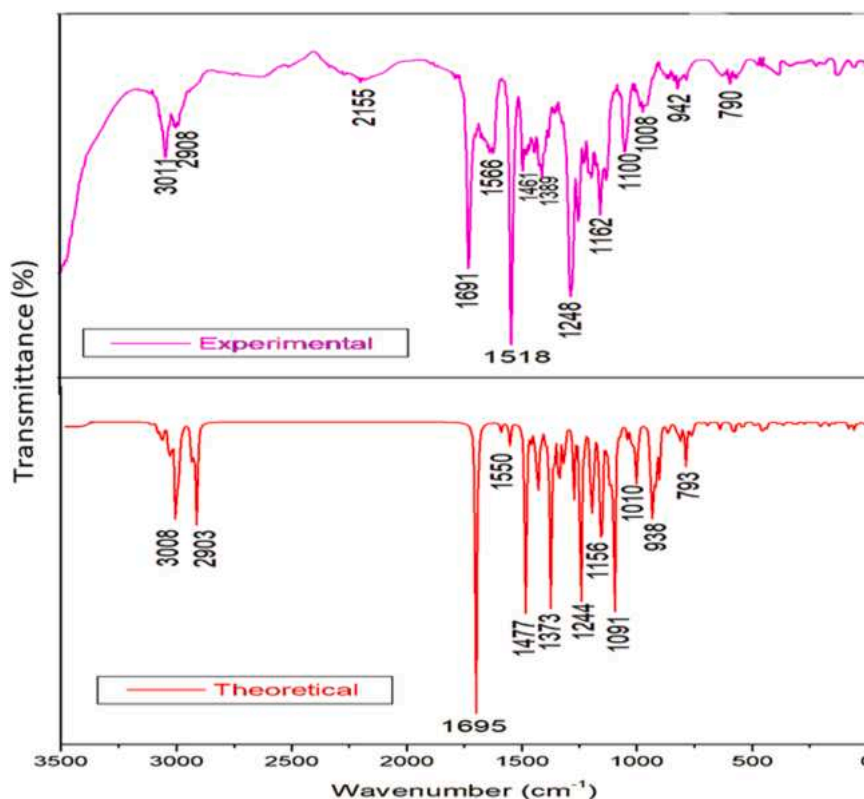**Fig. 5.** (a) Colour filled map for ELF. (b) Colour filled map for LOL.

Fig. 6. Theoretical and experimental FT-IR spectra of 1-acetyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

atoms is shown by the blue coloured regions. In the LOL plot, electron density in the centre of the few hydrogen atoms is greater than the limit of the colour scale as indicated by the white colour.

AIM analysis

The types of distinct interactions in various molecular systems have been studied in detail using the atoms in molecules theory (AIM). Also, the concept of Hydrogen bonding is effectively explained by the AIM theory [69].The bond critical points can be localised to calculate attributes at specific positions in space.

The chemical bonds and strength between atoms are recognised using the Bond critical point (BCP),which is the minimal charge density over the bond's strength [70]. Topological metrics such as kinetic energy densities (G(r), electron density (ρ(r),Laplacian $\Delta 2$ ρ(r), potential V (r), total energy density H(r), and bond energy Ewere used to determine BCP characteristics.The calculated values of electron density, Laplacian values and ellipticity are tabulated in Supplementary Table 3. Supplementary Figure 2 shows the AIM analysis in pictorial form.

Rozas et al. reported that there are three sorts of interactions: (i) for strong H-bonds ($\rho\nabla^2$) < 0; H < 0 and covalent in nature, (ii) for medium H-bonds ($\rho\nabla^2$) > 0; H<0 and partially covalent in nature and (iii) for weak H bonds ($\rho\nabla^2$) > 0; H>0 and electrostatic in nature, where $\rho\nabla^2$denotes the Laplacian of electron density and H denotes the total electron energy density [71]. The negative values of ($\rho\nabla^2$) shows the covalent bond interactions between the atoms.

In the title compound, except C_{13} – O_{31} bond all the other atoms exhibit covalent bond interactions and has strong hydrogen bonds between them as shown by the negative values of H and ($\rho\nabla^2$). For the C_{13} – O_{31} bond, the value of H and ($\rho\nabla^2$) are −0.6753 and 0.1053 respectively. It shows the presence of medium hydrogen bond and partially covalent character. Also, there are possibilities for intra hydrogen bonding interactions between the O_3 – H_{27} atoms. Their Laplacian of electron and overall electron energy density statistics are found to be

0.04548 and 0.001585 respectively. It shows the possibility for the presence of electrostatic interactions and strong covalent bonds.

Fukui function

A prominent localised density functional descriptor for chemical reactivity modelling and site selectivity modellingis the Fukui function. It is frequently employed in reactive site prediction [72]. Fukui indices indicates the atoms in a molecule that has a strong propensity to either lose or absorb an electron. It enables us to anticipate the molecule's most susceptible atom to nucleophilic or electrophilic attack [73]. The Fukui function is outlined as

$$f(r) = \left[\frac{\delta\rho(r)}{\delta N}\right]_v$$

Where, ρ(r), N and vdenotes the electron density, number of electrons and external potential, respectively [74].

Condensed Fukui functions are studied to determine each atom's ability to operate as a reactive site in a molecule. Fukui function f^+(r) is also referred as nucleophile attack index, which is the index which arises when a molecule receives an electron.Similarly, the Fukui functionf^-(r) is referred to as electrophilic attack index which occurs when a molecule has the tendency to lose an electron [75]. Condensed Fukui functions on the kth atomic site is calculated as

For Nucleophilic attack: $f_j^+ = q_j (N + 1) − q_j(N)$
For Electrophilic attack: $f_j^- = q_j(N) − q_j (N-1)$
For Radical attack: $f_j^0 = \frac{1}{2}[q_j (N + 1) − q_j (N-1)]$ where, $q_j$refers to the charge of atom at the jth atomic site.

A novel dual descriptor, Δf(r), has been proposed by Morrel et al. to identify the chemical locations vulnerable to electrophilic and nucleophilic attack [76].The difference between the nucleophilic and electrophilic Fukui functions is defined as the dual descriptor, Δf(r) given by the relation

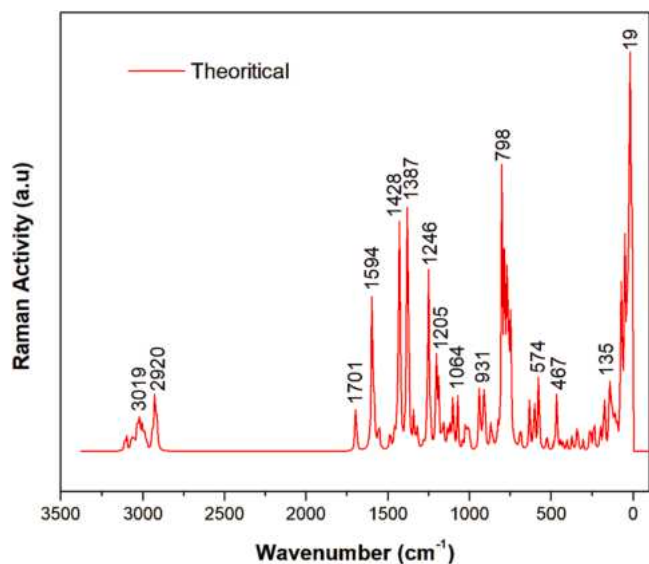For dual descriptor: $\Delta f(r) = [f_k^+ - f_k^-]$

Fig. 7. Theoretical FT-Raman spectra of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane.

If the dual descriptor value is greater than zero [Δf(r)> 0], then that site is susceptible to nucleophilic attack. Similarly, when the dual descriptor value is lesser than zero [Δf(r)< 0], that site is prone to electrophilic attack [77].The visual representation of the isosurface map for f^-, f^+, f^0 and dual descriptor obtained from the multiwfn software are displayed in Supplementary figure 3 The fukui indices calculated using the multiwfn software are tabulated in Supplementary Table 4.

For the title compound, the possible nucleophilic sites are C_1, O_3, C_7, C_8, C_9, C_{11}, C_{12}, C_{13}, C_{14}, C_{19}, O_{23}, C_{24}, H_{25}, H_{28}, H_{30}, H_{32} and H_{36} as they have positive dual descriptor values. Similarly, the possible electrophilic sites are C_2, C_4, C_5, C_6, C_{10}, H_{15}, H_{16}, H_{17}, H_{18}, H_{20}, H_{21}, H_{22}, H_{26}, H_{27}, H_{29}, O_{31}, H_{33}, H_{34}, H_{35}, H_{37} and H_{38} as they have negative dual descriptor values.

The order of vulnerability to nucleophilic attack is given by C_{19}>C_1>C_9>C_8>C_{24}>C_7>C_{11}>O_3>C_{12} >H_{28}>O_{23}>H_{25}>H_{36}>H_{14}>H_{30}>H_{32} >C_{13}. Whereas, the order of vulnerability to electrophilic attack is given by C_2>H_{21}>H_{16}>C_5>H_{22}>H_{20}>H_{38}>C_6 >H_{17}>H_{15}>H_{27}>C_{10}>H_{26}>H_{35}>H_{18}>C_4>H_{34}>H_{37}>H_{29}>O_{31}>H_{33}. The local behaviour of the molecule determines how it responds to electrophilic and nucleophilic attacks throughout the reaction [78].

Vibrational spectral analysis

Depending on the distinctive vibrations of the methyl, phenyl, carbonyl, and acethyl groups, the vibrational spectral analysis is carried out [79]. It is now feasible to calculate theoretical vibrational spectra by applying a wide range of quantum chemical techniques, including variational, perturbational, density functional, and linked clusters methods [80]. Among them PED analysis is a more precise method that allows one to quantify the amount of movement of a specific set of atoms in a normal mode [81].

The AMC compound contains 38 atoms and has 108 vibrational modes. The modes of vibrations were specified by utilising the VEDA4 softwaredepending on the potential energy distribution (PED) analysis. Thetheoretical and experimental FT-IR and FT-Raman spectra of AMC compound was shown in the Figs. 6 and 7, respectively. According to theoretical calculations, there are no hypothetical values for any of the frequencies. Also, the theoretically calculated vibrational modes match up with a significant percentage of the observed vibrational patterns. In order to compare with the frequency investigations obtained, all the calculated wavenumbers are scaled by 0.967. The computed vibrational

Table 3

Vibrational assignments of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropane using VEDA.

Mode Nos.	Wavenumber (cm^{-1})		Assignments with PED
	Experimental FT-IR	Scaled	
1	–	3110	$\upsilon C_{11}H_{35}(99)$
2	–	3002	$\upsilon C_9H_{38}(95)$
3	–	3068	$\upsilon C_9H_{28}(99)$
4	–	3062	$\upsilon C_8H_{37}(92)$
5	–	3057	$\upsilon C_{12}H_{30}(91)$
6	–	3048	$\upsilon C_{10}H_{29}(90)$
7	–	3035	$\upsilon C_{14}H_{33}(85) + \upsilon C_{14}H_{32}(15)$
8	–	3017	$\upsilon C_{24}H_{26}(76) + \upsilon C_{24}H_{25}(17)$
9	3011	3019	$\upsilon C_{11}H_{35}(99)$
10	–	3005	$\upsilon C_1H_{15}(12) + \upsilon C_1H_{17}(84)$
11	–	3004	$\upsilon C_1H_{15}(12) + \upsilon C_1H_{16}(85)$
12	–	2999	$\upsilon C_1H_{15}(75) + \upsilon C_{19}H_{20}(21)$
13	–	2992	$\upsilon C_{19}H_{20}(76) + \upsilon C_1H_{15}(23)$
14	–	2990	$\upsilon C_{24}H_{25}(82) + \upsilon C_{24}H_{26}(14)$
15	–	2978	$\upsilon C_{14}H_{32}(100)$
16	–	2934	$\upsilon C_1H_{15}(86) + \upsilon C_1H_{16}(11)$
17	–	2930	$\upsilon C_1H_{15}(85) + \upsilon C_1H_{16}(15)$
18	–	2924	$\upsilon C_{14}H_{33}(15) + \upsilon C_{14}H_{32}(85)$
19	–	2918	$\upsilon C_2H_{18}(95)$
20	2908	2911	$\upsilon C_{24}H_{25}(11) + \upsilon C_{24}H_{26}(10) + \upsilon C_{24}H_{27}(78)$
21	1691	1697	$\upsilon O_{31}C_{13}(88)$
22	–	1589	$\upsilon C_4C_9(64)$
23	1556	1553	$\beta\ C_5C_6C_7(12) + \upsilon C_4C_9(44)$
24	1518	1485	$\upsilon C_4C_9(11) + \beta\ H_{28}C_9C_8(22) + \beta\ H_{28}C_6C_7(12)$
25	1461	1461	$\beta\ H_{15}C_1H_{16}(68) + \tau H_{15}C_1C_2C_{19}(24)$
26	–	1459	$\beta H_{25}C_{24}C_{27}(73) + \tau H_{25}C_{24}O_{23}C_5(20)$
27	–	1444	$\tau H_{15}C_1C_2C_{19}(20) + \beta H_{15}C_1H_{16}(64)$
28	–	1438	$\beta H_{15}C_1H_{16}(70) + \tau H_{16}C_1C_2H_9(10)$
29	–	1436	$\beta H_{25}C_{24}H_{26}(68) + \tau H_{26}C_{24}O_{23}C_5(13)$
30	–	1432	$\beta H_{35}C_{11}H_{36}(37) + \beta H_{15}C_1H_{16}(15)$
31	–	1431	$\beta H_{35}C_{11}H_{36}(14) + \beta H_{15}C_1C_{16}(52)$
32	–	1426	$\beta H_{25}C_{24}H_{26}(68) + \tau H_{32}C_{14}C_{13}C_{12}(19)$
33	–	1425	$\beta H_{26}C_{24}H_{27}(73)$
34	–	1415	$\beta H_{32}C_{14}H_{33}(69) + \tau H_{32}C_{14}C_{13}C_{12}(22)$
35	1389	1394	$\beta H_{35}C_{11}H_{36}(32) + \upsilon C_5C_6(21) + \beta H_{29}C_{10}C_7(11) + \beta H_{37}C_8C_9(12)$
36	–	1375	$\upsilon C_5C_6(15) + \upsilon C_{12}C_{13}(12) + \beta H_{29}C_{10}C_7(17)$
37	–	1372	$\beta H_{15}C_1H_{16}(78)$
38	–	1357	$\beta H_{15}C_1H_{16}(69) + \tau H_{18}C_2O_3C_4(11)$
39	–	1339	$\beta H_{18}C_2O_3(53) + \tau H_{18}C_2O_3C_4(11)$
40	–	1334	$\beta H_{18}C_2O_3(68)$
41	–	1322	$\beta H_{18}C_2O_3(10) + \beta H_{18}C_2O_3(13) + \tau H_{18}C_2O_3C_4(43)$
42	–	1315	$\beta H_{18}C_2O_3(12) + \beta H_{30}C_{12}C_{11}(42)$
43	–	1272	$\upsilon C_4C_9(54)$
44	1248	1243	$\upsilon O_3C_4(51) + \beta H_{38}C_9C_8(17)$
45	–	1230	$\upsilon O_3C_4(11) + \beta H_{28}C_6C_7(39)$
46	–	1198	$\beta H_{37}C_8C_9(11) + \beta H_{29}C_{10}C_7(11) + \upsilon O_3C_4(20)$
47	–	1189	$\upsilon O_3C_4(12) + \upsilon C_{11}C_{10}(19) + H_{25}C_{24}O_{23}C_5(13)$
48	1162	1167	$\tau H_{25}C_{24}O_{23}C_5(42)$
49	–	1151	$\beta H_{15}C_1H_{16}(14) + \tau H_{15}C_1C_2C_{19}(40)$
50	–	1141	$\beta O_{31}C_{13}C_{14}(10) + \tau H_{32}C_{14}C_{13}C_{12}(17) + \upsilon C_{12}C_{13}(22) + \beta H_{30}C_{12}C_{11}(15)$
51	–	1130	$\upsilon C_7C_{10}(27) + \beta H_{38}C_9C_8(13)$
52	–	1128	$\beta H_{25}C_{24}H_{26}(17) + \tau H_{25}C_{24}O_{23}C_5(42) + \tau H_{26}C_{24}O_{23}C_5(32)$
53	–	1116	$\beta H37C8C9(14)$
54	–	1113	$\beta C_1C_2O_3(10) + \tau H_{17}C_1C_2C_{19}(11) + \upsilon C_1C_2(35)$
55	1100	1099	$\upsilon O_3C_2(10) + \tau H_{15}C_1C_2C_{19}(17) + \tau H_{29}C_{10}C_7C_6(20)$
56	–	1095	$\tau H_{15}C_1C_2C_{19}(21) + \tau H_{29}C_{10}C_7C_6(16)$
57	–	1072	$\tau H_{30}C_{12}C_{13}C_{14}(47) + \beta H_{35}C_{11}C_{12}(16)$
58	–	1041	$\tau H_{30}C_{12}C_{13}C_{14}(71)$
59	–	1021	$\upsilon O_{23}C_{24}(12) + \beta H_{28}C_6C_7(12) + \tau H_{35}C_{11}C_{12}C_{10}(11)$

(continued on next page)

Table 3 (*continued*)

Mode Nos.	Wavenumber (cm^{-1})		Assignments with PED
	Experimental FT-IR	Scaled	
60	1008	1012	$\tau H_{30}C_{12}C_{13}C_{14}(11) + \tau H_{32}C_{14}C_{13}C_{12}(41) + outO_{31}C_{12}C_{14}C_{13}(18)$
61	–	998	$\upsilon C_{23}C_{24}(45) + \tau H_{35}C_{11}C_{12}C_{10}(10)$
62	942	935	$\upsilon C_{12}C_{13}(13) + \upsilon O_3C_2(12) + \tau H_{32}C_{14}C_{13}C_{12}(18)$
63	–	928	$\upsilon O_3C_2(25) + \tau H_{32}C_{14}C_{13}C_{12}(14)$
64	–	919	$\tau H_{37}C_8C_9C_4(27)$
65	–	915	$\tau H_{18}C_2O_3C_4(11) + \tau H_{17}C_1C_2C_{19}(45) + \upsilon C_1C_2(12) + \tau H_{15}C_1H_{16}(10)$
66	–	912	$\upsilon C_{11}C_{12}(37)$
67	–	904	$\upsilon C_1C_2(13) + \beta H_{15}C_1H_{16}(11) + \tau H_{16}C_1C_2C_{19}(36)$
68	–	902	$\tau H_{37}C_8C_9C_4(34) + \tau H_{16}C_1C_2C_{19}(10)$
69	–	866	$\tau H_{28}C_6C_7C_{10}(56)$
70	–	859	$\tau H_{28}C_6C_7C_{10}(13) + \tau H_{35}C_{11}C_{12}C_{10}(12)$
71	–	826	$\upsilon C_1C_2(52)$
72	–	811	$\beta C_{12}C_{11}C_{10}(15) + \tau H_{37}C_8C_9C_4(55)$
73	790	786	$\beta C12C_{11}C_{10}(49) + \beta H_{37}C_8C_9C_4(17)$
74	–	764	$\upsilon C_{12}C_{13}(32) + \beta H_{35}C_{11}C_{12}(24)$
75	–	756	$\beta C_8C_7C_6(13) + \upsilon C_7C_{10}(14)$
76	–	715	$\tau C_4C_9C_8C_7(60)$
77	–	691	$\upsilon C_7C_{10}(20) + \beta C_9C_8C_7(22)$
78	–	637	$\tau C_6C_7C_2C_9(10) + outC_{10}C_6C_8C_7(35)$
79	592	586	$\beta O_3C_4C_9(17)$
80	–	583	$outO_{31}C_{12}C_{14}C_{13}(29) + \tau O_{31}C_{13}C_{14}(14)$
81	–	574	$outO_{31}C_{12}C_{14}C_{13}(10) + \upsilon C_{12}C_{13}(10) + \beta O_{31}C_{13}C_{14}(37)$
82	–	541	$\beta C_8C_7C_6(15)$
83	–	487	$\tau C_4C_9C_8C_7(14) + outO_3C_5C_9C_4(14)$
84	–	461	$\tau C_4C_9C_8C_7(17) + outO_3C_5C_9C_4(12) + \beta O_{23}C_5C_6(10)$
85	–	448	$outC_{19}C_1O_3C_2(49)$
86	–	438	$\beta C_{11}C_{10}C_7(15)$
87	–	388	$outC_{10}C_6C_8C_7(11) + \beta C_{14}C_{13}C_{12}(12)$
88	–	368	$\beta C_{14}C_{13}C_{12}(10) + \beta C_{19}C_2C_1(32)$
89	–	359	$\beta C_{24}O_{23}C_5(16) + \tau C_4C_9C_8C_7(24) + \beta C_1C_2O_3(16)$
90	–	321	$\beta C_1C_2O_3(11) + \beta C_{14}C_{13}C_{12}(14) + \beta C_{19}C_2C_1(15)$
91	–	306	$\tau H_{16}C_1C_2C_{19}(10) + \beta C_1C_2O_3(16)$
92	–	274	$\beta C_{24}O_{23}C_5(12) + outC_{10}C_6C_8C_7(19)$
93	–	245	$\beta C_{19}C_1C_2(13) + \beta C_4C_9C_8(10) + \beta O_{23}C_5C_6(12)$
94	–	237	$\tau H_{16}C_1C_2C_{19}(76)$
95	–	206	$\tau H_{15}C_1C_2C_{19}(72)$
96	–	198	$outC_{13}C_{10}C_{11}C_{12}(17) + \beta C_{14}C_{13}C_{12}(14)$
97	–	177	$outC_{13}C_{10}C_{11}C_{12}(27)$
98	–	145	$\tau C_6C_7C_8C_9(27) + \tau C_1C_2O_3C_4(16)$
99	–	135	$\tau H_{25}C_{24}O_{23}C_5(11) + \tau H_{26}C_{24}O_{23}C_5(11) + outO_3C_5C_6C_4(12)$
100	–	125	$\tau H_{26}C_{24}O_{23}C_5(12) + \beta C_2O_3C_4(13) + \tau H_{34}C_{14}C_{13}C_{12}(11)$
101	–	115	$\tau H_{34}C_{14}C_{13}C_{12}(73)$
102	–	101	$\beta C_{10}C_7C_8(16) + \tau C_1C_2O_3C_4(16)$
103	–	74	$\tau H_{25}C_{24}O_{23}C_5(11) + \tau C_1C_2O_3C_4(41)$
104	–	64	$\tau C_1C_2O_3C_4(62)$
105	–	54	$\tau C_2O_3C_4C_5(47)$
106	–	48	$\tau C_7C_{10}C_{11}C_{12}(73)$
107	–	31	$\tau C_2O_3C_4C_5(46) + \tau C_{11}C_{10}C_7C_6(12)$
108	–	17	$\tau C_{11}C_{10}C_7C_6(61) + \tau C_2O_3C_4C_5(11)$

ν – stretching vibrations; β – bending vibrations; τ – torsion-in-plane vibrations; out - out-of-plane vibrations.

wave numbers, Experimental values of FT-IR and FT-Raman in addition to the PED Assignments are documented in Table 3.

C—C vibrations

The phenyl ring of the title compound is tri-substituted.For benzene compounds, carbon stretching is typically quite significant in vibrational spectra [82]. The projected range for C–C stretching vibrations is 650–1100 cm^{-1}. The substituted compound's composition has no

Table 4
Drug likeness parameters of the title compound.

Descriptors	Calculated	Expected
Molecular mass (Dalton)	248.32	<500
Hydrogen bond donor	0	<5
Hydrogen bond acceptor	3	<10
Log P	2.88	<5
Molar refractivity	71.32	40–130

discernible impact on these vibrations. The C–C stretching vibration forthetitle compound was observed at 592, 790, 942, 1008 and 1100 cm^{-1} in experimental FT-IR. The theoretical values of 793, 938, 1010 and 1091 cm^{-1} are also in the expected region.

The bending mode vibrations for the phenyl ring is found in the region of wavelength 1000–1300 cm^{-1}. It is proved from the peaks in the FT-IR experimental spectra at 1008 and 1100 cm^{-1}. The scaled values of 1243, 1230, 1099 and 1021 cm^{-1} confirmed it. The typical region for the ring breathing vibration modes is in the wavelength between 1620 and 1400 cm^{-1} [83]. During the ring "breathing" mode, all of the hydrogen and carbon atoms move in phase towards and away from the ring's centre i.e., there may be stretching and contraction of C—C and C—H bonds in the phenyl ring structure [84]. For the title compound, the scaled vibrations in the region 1594–1515 cm^{-1} confirms the ring breathing vibrations. The torsional vibrations are found to be in the region of wavelength 900–600 cm^{-1}. It is validated from the sharp peak in the FT-IR experimental spectra at 790 cm^{-1} and also 793 cm^{-1} in FT-IR theoretical spectra.

C—H vibrations

The C–H stretching mode vibrations are reported to be found in the area having wavelength 3100–3000 cm^{-1} [85]. The stretching vibrations of the phenyl C—H ring seem to be relatively faint due to effective conjugation induced via steric effect [86]. This is proved from the RDG plot of the NCI analysis for the title compound. The stretching vibrations for the AMC compound are found in 3070–2900 cm^{-1} region. It is confirmed from the sharp peak in the theoretical spectra at 3008 cm^{-1} and 2903 cm^{-1}. In theeexperimental FTIR spectra, it is evident from 3011 to 2908 cm^{-1} values.

When a compound has a methoxy group, the lone pair of electrons on the oxygen atom donates the electrical charge back to the σ * orbital of Carbon–Hydrogen bonds, diminishing the C–H bond [87]. As the result, the intensities of Infrared band of the Carbon-Hydrogen stretching modes are boosted [88]. It is evident from the strong peak at 2908 cm^{-1} in FT-IR spectra. Also, the theoretically scaled values of 3048, 2990 and 2911 cm^{-1} represents the stretching vibrations. For the title compound, the scaled values of 3035 and 2095 cm^{-1} represent the C—H stretching vibrations of the carbonyl group.

The CH$_3$ functional groups in the acethyl group have one symmetric and two asymmetric CH$_3$ stretching vibrations [89]. A sharp peak was detected at 2920 cm^{-1} in the FT-IR spectra showed the acethyl group stretching vibrations [90]. In the title compound, the scaled values of 2930 and 2924 cm^{-1} confirms the above data.

Normally, there are several strong, sharp bands in the areas of 1000–750 cm^{-1} and 1300–1000 cm^{-1}, representing the C—H out-of-plane and in-plane bending vibrations respectively [91]. In this work, it is evident from the sharp peaks in the experimental FT-IR spectra. Also, the scaled theoretical values of 1372, 1357, 1339, 1334, 1322, 1198, 1151, 1116, 919, 904, 786, 764 and 756 cm^{-1} showed good agreement to this. The vibrations of the C—H bonds are strongly influenced by the in-plane and out-of-plane bending of the carbonyl group [92]. The substituent's nature has no obvious effect on the bands in this region [93]. Also, the scaled values agreed with the experimental data.

The torsion vibrations for the AMC compound are observed in the range of 1500–1100 cm^{-1}. The scaled values for torsional mode of vibrations are observed in 1357, 1339, 1322, 1157 and 1116 cm^{-1}. The
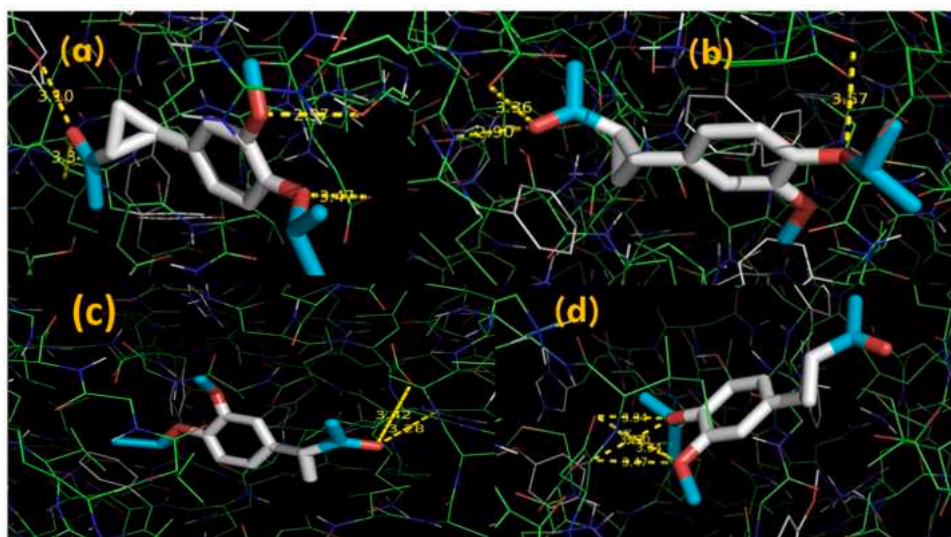
Fig. 8. Molecular Docking of the title compound with (a) antiviral protein, (b) antibacterial protein, (c) Anticanceral protein (d) antimalarial protein.

Table 5
The target proteins and docking activity of the title compound.

Protein (PDB: ID)	Binding residues	Bond distances (Å)	Binding energy (kcal/mol)
1VZV	THR113	3.10	−5.6
	TYR155	3.34	
	SER103	2.97	
	GLU106	3.47	
1JIJ	GLU302	3.36	−6.0
	PHE273	2.90	
	LYS305	3.57	
1GII	ASP86	3.42	−5.3
	ASP86	3.28	
1LDG	ASP53	3.27	−6.1
	ASP53	3.34	
	ASP53	3.30	
	ASP53	3.47	
	IUE54	3.34	

sharp peaks in the experimental FT-IR spectra at 1389 and 1162 cm^{-1} confirms the torsional modes of vibrations.

C—O, C=O vibrations

The carbonyl group is found in a wide variety of compounds, and because of the C=O stretching vibration, it exhibits an intense band in the 1850 cm^{-1}–1550 cm^{-1} region [94]. For the carbonyl group, C=O stretching vibrations are found in 1695 cm^{-1}. The sharp peak found in the experimental FT-IR spectra in the wavelength of 1692 cm^{-1} confirms the carbonyl stretching vibrations.

The C—O stretching is anticipated to be in the range of 1050–1000 cm^{-1}. Electronic influences can cause the C—H stretching and bending bands to change positions when the CH$_3$ group has a direct bond with an oxygen atom. Because of this, the O—CH$_3$ stretching bands cover a wider area than the C—CH$_3$ group [95]. It is obvious from the scaled values of 1021 and 1012 cm^{-1} and the experimental FT-IR values of 1008 cm^{-1}.

The C—O-CH$_3$ bending and torsional vibrations are look forward to be obtained in the range of 1200–1100 cm^{-1}. It is evident from the scaled values of 1198, 1167 and 1128 cm^{-1} and the experimental FT-IR values of 1162 cm^{-1}. The scaled values of C=O bending vibrations for the AMC compound is observed in 1157 and 1012 cm^{-1}. The regions from 725 cm^{-1} to 595 cm^{-1} are predicted to exhibit torsion in-plane and out-of-plane vibrations [96]. For the title compound, the scaled values of torsional vibrations are 583 and 574 cm^{-1}.

Drug likeness

Drug-likeness isone of the most significant variables in the early phases of drug discovery. Drug likeness a qualitative property of chemicals that describes how closely a compound's qualities resemble those of already-approved drugs [97]. The most popular way to judge drug likeness is through rules, the most well-known of which is Lipinski's Rule of Five (Ro5) [98].

The relationship between physicochemical and pharmacokinetics indices is illustrated by this rule-of-five [99]. According to Lipinski's rule, an orally active drug-like compound cannot violate more than one of the following conditions: its molecular weight shouldn't exceed 500 Da, its hydrogen bond donors shouldn't exceed 5, its hydrogen bond acceptors shouldn't exceed 10, its molar refractivity should lie between 40 and 130, and its log P shouldn't exceed 5 [100].

A few key variables, including the compound's rotatable bonds, logP, molar refractivity, hydrogen bond acceptors (HBA) and donors (HBD) numbers, are utilised to determine the drug-likeness quality of the molecule [101]. The drug likeness properties are calculated by using the SwissADME software and the obtained results are organized in Table 4.

For title compound, the Molecular mass is calculated as 248.32 Da. The hydrogen bond donors and acceptors are calculated as 0 and 3 respectively. The molar refractivity is calculated as 71.32. The log P value is obtained as 2.88 which is analysed to be within the suitable range.

From the drug likeness results, the title compound complies with Lipinski's rule of five with no violations and it can be used as an active prospective drug. It is also confirmed from the docking results, as the title compound has high binding affinity (−6.1 kcal/mol) towards the antimalarial protein and exhibits good antimalarial activity.

Molecular docking

Molecule docking has grown to be a significant step in the drug discovery process due to its perceived ease of use and relatively low-cost consequences [102]. When proteins, enzymes and nucleic acids interact with tiny molecules to form supramolecular complexes, their biological function may be enhanced or inhibited [103]. Docking is a computational molecular modelling technique used to forecast how an enzyme will interact with ligands, or tiny molecules. The interaction of the protein-ligand complex and binding energy are predicted via molecular docking [104].

For the molecular docking, the title compound optimized with
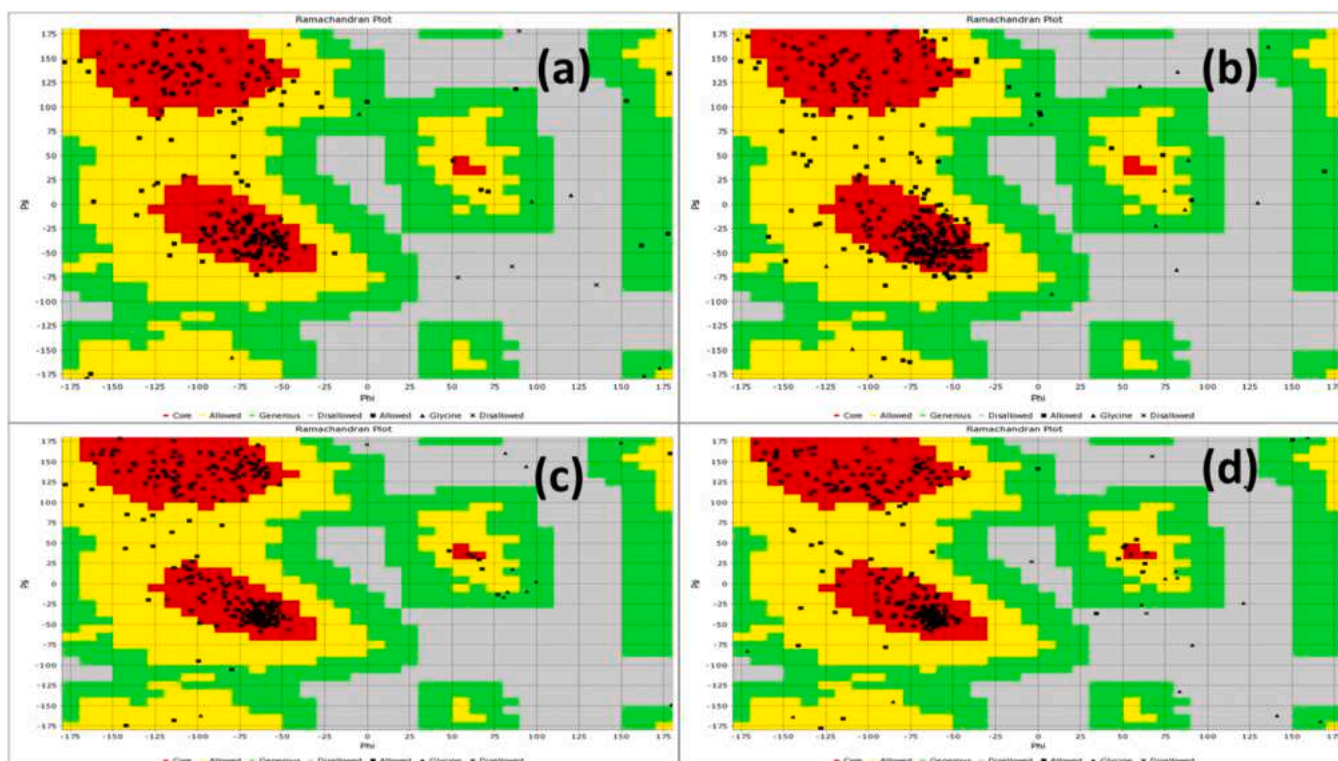
Fig. 9. Ramachandran plot for (a) antiviral protein, (b) antibacterial protein, (c) Anticanceral protein (d) antimalarial protein.
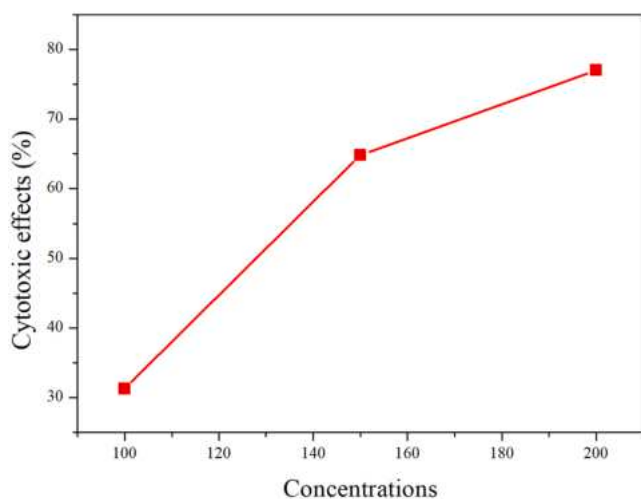
Fig. 10. The cytotoxic effects of 1-acethyl-2-(4-isopropoxy-3-methoxyphenyl) cyclopropaneonHeLa cells after 48 h of treatment with various concentrations.

minimum energy at the B3LYP/6–311++G(d,p) level of basis set is prepared as the ligand in the pdb format using open babel software. The target proteins are downloaded in pdb format from the RCSB Protein data bank.By using the pymol software, the target proteins were prepared by removing the solvents, co-crystallized ligands, organic compounds and water molecules. The molecular docking is done using the PyRx application.

For the preset study, the targeted proteins are antiviral [PDB: 1VZV], antibacterial [PDB: 1JIJ], anticanceral [PDB: 1GII] and antimalarial proteins [PDB: 1LDG]. The molecular docking of the title compound with the targeted proteins is displayed in the Fig. 8. The docking parameters like binding residues, binding energy and bond distances of the protein-ligand interactions are calculated using the PyRx application. The docking parameters are tabulated in Table 5.

The title compound is docked with the antiviral protein 1VZV. It creates four interactions with the amino acids namely THR113, TYR155, SER103 and GLU106. The amino acids in the AMC compound reacted with the oxygen atoms. The bond distance between the amino acids THR113, TYR155, SER103 and GLU106 are 3.10, 3.34, 2.97 and 3.47, respectively. The protein ligand binding affinity is calculated as −5.6 kcal/mol. Based on the acquired results, the title compound exhibits good antiviral activity.

Then the title compound is docked with the antibacterial protein
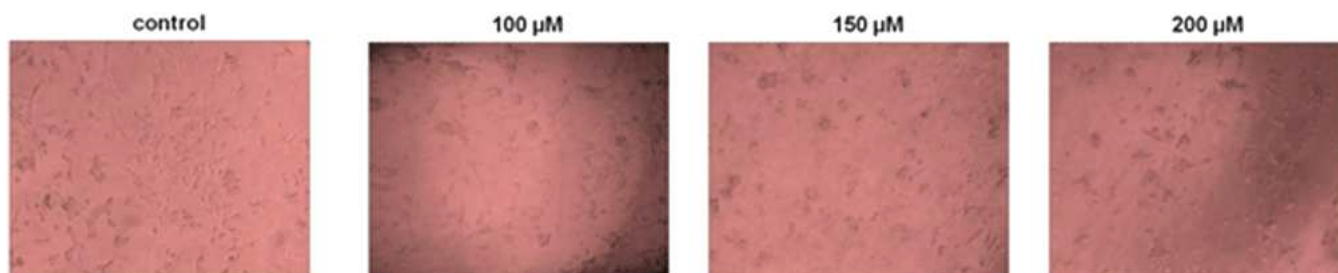
Fig. 11. Changes of the HeLa cells morphology after 48 h treatment with 1-acethyl-2-(4-isopropoxy-3-methoxyphenyl)cyclopropane.

1JIJ. It generates three interactions with the amino acids namely GLU302, PHE273 and LYS305. The amino acids in the AMC compound reacted with the oxygen atoms. The bond distance between the amino acids GLU302, PHE273 and LYS305 are 3.36, 2.90 and 3.57, respectively. The protein ligand binding affinity is calculated as -6.0 kcal/mol. In light of the obtained outcomes, the title compound exhibits good antibacterial activity.

For the anticanceral activity, the title compound is docked with the anticanceral protein 1GII. It produces four interactions with the amino acids namely ASP86 and ASP86. The amino acids in the AMC compound reacted with the oxygen atoms. The bond distance between the amino acids ASP86 and ASP86 are 3.42 and 3.28, respectively. The protein ligand binding affinity is calculated as -5.3 kcal/mol. Considering the obtained effects, the title compound exhibits good anticanceral activity.

For the antimalarial activity, the title compound is docked with the antimalarial protein 1LDG. It produces five interactions with the amino acids namely ASP53, ASP53, ASP53, ASP53 and IUE54. The amino acids in the AMC compound reacted with the oxygen atoms. The bond distance between the amino acids ASP53, ASP53, ASP53, ASP53 and IUE54 are 3.27, 3.34, 3.30, 3.47 and 3.34, respectively. The protein ligand binding affinity is calculated as -6.1 kcal/mol. Taking into account the acquired accomplishments, the title compound exhibits good antimalarial activity.

Ramachandran plot for the targeted proteins is displayed in the Fig. 9. From the Ramachandran plot, the stability of the protein selected for the binding interaction is indicated by the fact that 90 % of the amino acids are located in the allowed region, which is indicated by the colour red, and very few are located in the disallowed region.

From the molecular docking results, the docking arrangement of the ligand with the protein discloses that the only location undergoing bond formation is the electronegative oxygen atom. This interaction verifies the title compound's bioactivity, which was evident from the charge transfer from LP1(O_{23}) $\rightarrow$ σ^* (C_4–C_5) in NBO analysis. Also, it is evident that the title compound has high binding affinity (-6.1 kcal/mol) towards the antimalarial protein compared to other proteins. Thus, it exhibits good antimalarial activity.

Cytotoxic effect

Cytotoxic effect of the title compound been tested on HeLa cells. Cells were seeded in a 96-well plate at 3×10^3 cells/well. The cells were incubated at 37 °C in 5 % CO2 atmosphere and absolute humidity for 48 h with concentrations (100, 150 and 200 μM) in triplicate. Then, the cell culture media was removed, and 100 μL of MTT solution (3-(4,5-dimethylthiazol-2-yl)–2,5-diphenyltetrazolium bromide, 0.5 mg/mL) was added to each well. After 2 h incubation at 37 °C, MTT solution was removed and 150μl of DMSO was added to the wells.

The absorbance was measured at 595 nm using a multiplate reader (Zenyth 3100, AnthosLabtec Instruments).

The percentage of cytotoxic cells was calculated using the formula:

$$Cytotoxicity(\%) = (1 - (\text{exp.group}(ABS)) / (\text{control group}(ABS)) \times 100$$

of 1-acethyl-2(4-isopropoxy-3-methoxy phenyl) cyclopropanehave cytotoxic activity against HeLa cells line. The Compounds showed better cytotoxic effects with (100 μM- 31.26 %; 150 μM–64.82 %; 200 μM–77.02 %). In order to determine the cytotoxic effects of the title compound on the morphology of treated and untreated HeLa cells, the phase-contrast microscope is used. The HeLa cells were seeded in a 24-well plate and incubated for 48 h with different concentrations pyrazolines (100, 150 and 200 μM). Morphological changes of both, experimental and control HeLa cells, were visualized with phase contrast microscopy under 100 X magnification on Olympus microscope (model BX51). Figs. 10 and 11 shows the cytotoxic effects and Changes of the HeLa cells morphology after of 1-acethyl-2-(4-isopropoxy-3-methoxyphenyl)cyclopropane. The results showed that the compound 1-acethyl-2-(4-isopropoxy-3-methoxyphenyl)cyclopropane induced morphological changes of HeLa cells in dose dependent manner compared to the control group. Morphological changes (loss of shape cells and decrease in the number cells) were observed by phase-contrast microscopy.

Conclusion

Using the quantum chemical computations, the structural, vibrational and topological characteristics of the compound was studied. The $6–311{++}G$(d,p) basis set was used to optimise the Molecular geometry employing DFT-B3LYP and the structural parameters were identified. The intra- and inter-molecular interactions calculated from the NBO analysis confirms the bioavailability of the compound. The HOMO-LUMO energy values and the global reactivity descriptors were also calculated. The HOMO-LUMO energy gap is calculated as 3.7434 eV which also accounts for the biological activity of the molecule. The vibrational spectra were thoroughly interpreted with the aid of VEDA software. FT-IR and FT-Raman spectra have recorded and the values agreed with the theoretical vibrational assignments. The steric repulsion, weak and strong attractions of the tile compound was analysed by using the NCI approach. MEP and the topological analysis like ELF, LOL, Fukui, AIMetc, were utilised to evaluate the title compound's reactive sites and electron distribution. The drug likeness studies strengthens that the title compound can be utilised as an active medication to treat specific illnesses. For the molecular docking studies, the target protein is docked with antiviral, antibacterial, anticanceral and antimalarial target proteins. Also the Ramachandran plot is potted for the targeted proteins. Among them the title compound has high binding affinity (-6.1 kcal/mol) towards the antimalarial protein and exhibits good antimalarial activity.

CRediT authorship contribution statement

D.P. Lydia Renj: Writing – original draft, Funding acquisition. **R. Racil Jeya Geetha:** Supervision, Writing – review & editing. **A. Benifa:** Writing – review & editing, Investigation. **M. Amalanathan:** Conceptualization, Validation. **M. Sony Michael Mary:** Methodology, Formal analysis. **Zoran Ratković:** Resources. **Jovana Muškinja:** Resources. **Jayasekar Micheal:** Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chphi.2024.100524.

References

[1] M.P. Toraskar, V.J. Kadam, V.M. Kulkarni, Synthesis and antimicrobial activity of functional analogues of fluconazole, Int. J. Pharm. Pharm. Sci. 2 (2) (2010) 132–133. ISSN: 2656-0097.

[2] K. Ajay Kumar, Brief review on cyclopropane analogs: synthesis and their pharmacological applications, Int. J. Pharm. Pharm. Sci. 5 (1) (2013). ISSN-0975-1491.

[3] W. Cao, H. Zhang, J. Chen, H. Deng, M. Shao, L. Lei, J. Qian, Y. Zhu, A facile preparation of trans-1,2-cyclopropanes containing p-trifluoro methyl phenyl group and its application to the construction of pyrazole and cyclopropane ring fused pyridazinone derivatives, Tetrahedron 64 (2008) 6670–6674, https://doi.org/10.1016/j.tet.2008.05.016.

[4] M. Sumithra, N. Sundaraganesan, K. Venkata Prasad, R. Rajesh, V. Vetrivelan, V. Ilangovan, A. Irfan, S. Muthu, Effect of green solvents physical, chemical, biological and bonding nature on 5-acethyl-thiophene-2-carboxylic acid by DFT and TD-DFT approach – An antiviral agent, J. Indian Chem. Soc. 100 (2023) 100867, https://doi.org/10.1016/j.jics.2022.100867.

[5] J. Buus Laursen, J. Nielsen, Phenazine natural products: biosynthesis, synthetic analogues, and biological activity, Chem. Rev. 104 (2004) 1663–1685, https://doi.org/10.1021/cr020473j.

[6] F.H. Hendrik, Computation of the structures of the phenyl and benzyl radicals with the UHF method &quot, J. Org. Chem. 52 (22) (1987) 5025–5026, https://doi.org/10.1021/jo00231a035.

[7] J. March, Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 4th ed., Wiley, New York, 1992. ISBN 978-0-471-60180-7.

[8] C. James, A.A. Raj, R. Reghunathan, V.S. Jayakumar, J. I.H, Structural conformation and vibrational spectroscopic studies of 2, 6-Bis (p-N, N-Dimethyl benzylidene) cyclohexanone using density functional theory, J. Raman Spectrosc. 37 (12) (2006) 1381–1392, https://doi.org/10.1002/jrs.1554.

[9] M. Medimagh, N. Issaoui, S. Gatfaoui, O. Al-Dossary, A.S. Kazachenko, H. Marouani, M.J. Wojcik, Molecular modeling and biological activity analysis of new organic-inorganic hybrid: 2-(3, 4-dihydroxyphenyl) ethanaminium nitrate, J. King Saud Univ.-Sci. 33 (2021) 101616.

[10] M. Tahenti, S. Gatfaoui, N. Issaoui, T. Roisnel, H. Marouani, A tetrachlorocobaltate (II) salt with 2-amino-5-picolinium: synthesis, theoretical and experimental characterization, J. Mol. Struct. 1207 (2020) 127781.

[11] O. Noureddine, S.G.N. Issaoui, O. Al-D, H. Marouani, Quantum chemical calculations, spectroscopic properties and molecular docking studies of a novel piperazine derivative, J. King Saud Univ.-Sci. 33 (2) (2021) 101283.

[12] M. Amalanathan, V.K. Rastogi, I.H Joe, M.A. Palafox, R. Tomar, Density functional theory calculations and vibrational spectral analysis of 3,5-(dinitrobenzoic acid), Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 78 (5) (2011) 1437–1444, https://doi.org/10.1016/j.saa.2011.01.023.

[13] J. Christina Jebapriya, J.C. Prasana, S. Muthu, B. Fathima Rizwana, Spectroscopic (FT-IR and FT-Raman), quantum computational (DFT) and molecular docking studies on 2(E)-(4-N,N-dimethylaminobenzylidene)-5-methylcyclohexanone, in: Proceedings of the Materials Today 50, 2022, pp. 2695–2702, https://doi.org/10.1016/j.matpr.2020.08.221.

[14] J. Sánchez-Márquez, Correlations between Fukui indices and reactivity descriptors based on Sanderson's principle, J. Phys. Chem. A 123 (40) (2019) 8571–8582, https://doi.org/10.1021/acs.jpca.9b05571.

[15] S. Premkumar, T.N. Rekha, R.M. Asath, T. Mathavan, A.M.F. Benial, Vibrational spectroscopic, molecular docking and density functional theory studies on 2-acethylamino-5-bromo-6-methylpyridine, Eur. J. Pharm. Sci. 82 (2016) 115–125, https://doi.org/10.1016/j.ejps.2015.11.018. Volume.

[16] M. Ali, A. Mansha, S. Asim, M. Zahid, M. Usman, N. Ali, DFT study for the spectroscopic and structural analysis of p-dimethyl aminoazo benzene, J. Spectrosc. (2018), https://doi.org/10.1155/2018/9365153.

[17] L. Dian, I. Marek, Asymmetric preparation of polysubstituted cyclopropanes based on direct functionalization of achiral three-membered carbocycles, Chem. Rev. 118 (18) (2018) 8415–8434, https://doi.org/10.1021/acs.chemrev.8b00304.

[18] D. Roy, T. Keith, J. Millam, GaussView, Semichem Inc., Shawnee Mission, KS, 2009.

[19] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunction analyzer, J. Comput. Chem. 33 (2012) 580–592.

[20] M.H. Jamroz Vibrational energy distribution analysis VEDA, 4 (2004).

[21] L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R.F. Boyko, B.D. Sykes, D.S. Wishart, VADAR: a web server for quantitative evaluation of protein structure quality, Nucleic. Acids. Res. 31 (13) (2003) 3316–3319, https://doi.org/10.1093/nar/gkg565.

[22] M.C. Murry, E. John, Organic Chemistry, 3rd ed, Belmont, Wadsworth, 2024.

[23] P. Senthil Raj, D. Shoba, S. Ramalingam, S. Periandy, Spectroscopic (FT-IR/FT-Raman) and computational (HF/DFT) investigation and HOMO/LUMO/MEP analysis on 1,1-difluoro-2-vinylcyclopropane, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 147 (2015) 293–302, https://doi.org/10.1016/j.saa.2015.03.084.

[24] P. Muniappan, R. Meenakshi, G. Rajavel, M. Arivazhagan, Vibrational spectra and theoretical calculations (dimerization, UV–Vis, multinuclear NMR and pes analyses) of 3,4-dimethylbenzamide and 3,4,5-trihydroxybenzamide, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 117 (2014) 739–753, https://doi.org/10.1016/j.saa.2013.08.049.

[25] S. Xavier, S. Periandy, Spectroscopic (FT-IR, FT-Raman, UV and NMR) investigation on 1-phenyl-2-nitropropene by quantum computational

[26] S. Chand, FA.M.Al- Omary, AA. El-Emam, VK. Shukla, O. Prasad, L. Sinha, Study on molecular structure, spectroscopic behavior, NBO, and NLO analysis of 3-methylbezothiazole-2-thione, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 146 (2015) 129–141, https://doi.org/10.1016/j.saa.2015.03.068.

[27] P. Aiswarya, T. Jayavarthanan, S. Periandy, S. Suresh, S. Soundhariya, Molecular structural analysis, conformers and spectral (FT-IR, FT-Raman, NMR and UV–Visible), Importance of solvent role in molecular, ADME and molecular docking investigation on alpha-cyano-4-hydroxycinnamic acid, Chem. Phys. Impact 7 (2023) 100353, https://doi.org/10.1016/j.chphi.2023.100353.

[28] G. Velraj, S. Soundharam, C. Sridevi, Investigation of structure, vibrational, electronic, NBO and NMR analyses of 2-chloro-4-nitropyridine (CNP), 2-chloro-4-methyl-5-nitropyridine (CMNP) and 3-amino-2-chloro-4-methylpyridine (ACMP) by experimental and theoretical approach, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 137 (2015) 790–803, https://doi.org/10.1016/j.saa.2014.08.075.

[29] D.L Renj, R.RJ Geetha, J.J James, M. Amalanathan, M.SM Mary, Z. Ratkovi, J. Mu̇skinja, Structural and spectroscopic investigation of 1-acetyl-2-(4-ethoxy-3-methoxyphenyl)cyclopropane and its NLO activity, Chem. Phys. Impact (8) (2024) 100420, https://doi.org/10.1016/j.chphi.2023.100420.

[30] L. Li, C. Wu, Z. Wang, L. Zhao, Z. Li, C. Sun, Tiemin Sun Density functional theory (DFT) and natural bond orbital (NBO) study of vibrational spectra and intramolecular hydrogen bond interaction of L-ornithine–L-aspartate, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 137 (2015) 790–803, https://doi.org/10.1016/j.saa.2014.08.153.

[31] P.G Rubarani, S. Sampath Krishnan, Natural Bond Orbital (NBO) population analysis of 1-Azanapthalene-8-ol, Acta Phys. Polonica Ser. A 125 (1) (2014) 18–22, https://doi.org/10.12693/APhysPolA.125.18.

[32] G. Collin, H. Höke, Ullmann's Encyclopedia of Industrial Chemistry, Wiley-VCH, Weinheim, 2005.

[33] H. Marshan Robert, D. Usha, M. Amalanathan, R. Racil Jeya Geetha, M. Sony Michael Mary, Spectroscopic (IR, Raman, UV, NMR) characterization and investigation of reactive properties of pyrazine-2-carboxamide by anti-bacterial, anti-mycobacterial, Fukui function, molecular docking and DFT calculations, Chem. Data Collections 30 (2020), https://doi.org/10.1016/j.cdc.2020.100583.

[34] K.N Prathap, L. Neratur, Three novel coumarin-benzenesulfonylhydrazide hybrids: synthesis, characterization, crystal structure, Hirshfeld surface, DFT and NBO studies, J. Mol. Struct. 1171 (2018), https://doi.org/10.1016/j.molstruc.2018.06.022.

[35] R. Rameshkumar, N. Santhi, Synthesis, Geometry optimization, Mulliken, MEP, HOMO-LUMO and NLO properties of 2-aryl-3-(2,6-diisopropylphenyl) thiazolidin-4-one based on DFT calculations, World Sci. News 91 (2018) 59–72. EISSN 2392-2192.

[36] M. Krishna Priya, B.K. Revathi, V. Renuka, S. Sathya, P. Samuel Asirvatham, M. Structure, Spectroscopic (FT-IR, FT-Raman, 13C and 1 H NMR) analysis, HOMO-LUMO energies, Mulliken, MEP and thermal properties of new chalcone derivative by DFT calculation, Mater. Today Proc. 8 (Part 1) (2019) 37–46, https://doi.org/10.1016/j.matpr.2019.02.078.

[37] P.P. Zamora, K. Bieger, A. Cuchillo, A. Tello, J.P. Muena, Theoretical determination of a reaction intermediate: fukui function analysis, dual reactivity descriptor and activation energy, J. Mol. Struct. 1227 (9–10) (2020), https://doi.org/10.1016/j.molstruc.2020.129369.

[38] W. Guerrab, I.M. Chung, S. Kansiz, J.T. Mague, N. Dege, J. Taoufik, Y. Ramli, Synthesis, structural and molecular characterization of 2, 2-diphenyl-2H, 3H, 5H, 6H, 7H-imidazo [2, 1-b][1, 3] thiazin-3-one, J. Mol. Struct. 1197 (2019) 369–376.

[39] M.A. Mumit, T.K. Pal, MdA Alam, MdAl-A-Al-A Islam, Subrata Paul, and Md Chanmiya Sheikh, DFT studies on vibrational and electronic spectra, HOMO–LUMO, MEP, HOMA, NBO and molecular docking analysis of benzyl-3-N-(2,4,5 trimethoxyphenyl methylene) hydrazinecarbodithioate, J. Mol. Struct. 1220 (2020) 128715, https://doi.org/10.1016/j.molstruc.2020.128715.

[40] A. Reed, R. Weinstock, F. Weinhold, Natural population analysis, Chem. Phys. 83 (1985) 735–746, https://doi.org/10.1063/1.449486, 10.1063/1.449486.

[41] R.S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions, J. Chem. Phys. 23 (1955) 1833–1840, https://doi.org/10.1063/1.1740588.

[42] C. BILKAN, Determination of structural properties of some important polymersused as interfacial layer in fabrication of schottky barrier diodes (SBDs), J. Inst. Sci. Technol. 10 (1) (2020) 225–233.

[43] Y. Sushma Priya, K.R. Rao, P. Venkata Chalapathi, A. Veeraiah, Vibrational and electronic spectra of 2-phenyl-2-imidazoline: a combined experimental and theoretical study, J. Mod. Phys. 9 (No.4) (2018) 149, https://doi.org/10.4236/jmp.2018.94049, 53.

[44] O. El-Gammal, G.A. El-Reash, S. Ghazy, A. Radwan, Synthesis, characterization, molecular modeling and antioxidant activity of (1E,5E)-1,5-bis(1-(pyridin-2-yl) ethylidene)carbonohydrazide (H2APC) and its zinc(II), cadmium(II) and mercury (II) complexes, J. Mol. Struct. 1020 (2012) 6, https://doi.org/10.1016/j.molstruc.2012.04.029.

[45] R. Bahnasawy, E.S.T. Kashar, Thermal and temperature dependence of electrical conductivity studies on Zn, Cd and Hg hydrazone complexe, J. Therm. Anal. Calorim 39 (1993) 65–74, https://doi.org/10.1007/BF02235447.

[46] V.K. Choudhary, A.K. Bhatt, D. Dash, N. Sharma, DFT calculations on molecular structures, HOMO–LUMO study, reactivity descriptors and spectral analyses of newly synthesized diorganotin(IV) 2-chloridophenylacetohydroxamate complexes, J. Comput. Chem. 40 (27) (2019) 2354–2363, https://doi.org/10.1002/jcc.26012. Volume.

[47] MJ.S. Dewar, Carlos De Llano Ground states of conjugated molecules. XI. Improved treatment of hydrocarbons, J. Am. Chem. Soc. 91 (4) (1969) 789–795, https://doi.org/10.1021/ja01032a001.

[48] R. Meenakshi, Spectral investigations, DFT based global reactivity descriptors, Inhibition efficiency and analysis of 5-chloro-2-nitroanisole as π-spacer with donor-acceptor variations effect for DSSCs performance, J. Mol. Struct. 1127 (2017) 694–707, https://doi.org/10.1016/j.molstruc.2016.08.030.

[49] V. Arjunan, S. Thillai Govindaraja, A. Jayapraksh, S. Mohan, Structural, vibrational and nuclear magnetic resonance investigations of 4-bromoisoquino-line by experimental and theoretical DFT methods, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 107 (2013) 62–71, https://doi.org/10.1016/j.saa.2013.01.037.

[50] N.S. Babu, D. Jayaprakash, Global and reactivity descriptors studies of cyanuric acid tautomers in different solvents by using of Density Functional Theory (DFT), Int. J. Sci. Res. 4 (2015) 615–620.

[51] G. Banuppriya, R. Sribalan, V. Padmini, Synthesis and characterization of curcumin-sulfonamide hybrids: biological evaluation and molecular docking studies, J. Mol. Struct. 1155 (2018) 90–100, https://doi.org/10.1016/j.molstruc.2017.10.097.

[52] S.W. Qader, A. Suvitha, M. Ozdemir, I. Benjamin, A.S.R. NSA, MU. Akem, A. E. Frank, EC. Eluwa, Investigating the physicochemical properties and pharmacokinetics of curcumin employing density functional theory and gastric protection, Chem. Phys. Impact (5) (2022) 100130, https://doi.org/10.1016/j.chphi.2022.100130.

[53] RA. Costa, P.O. Pitt, MLB. Pinheiro, KM.T. Oliveira, Kahlil Schwanka Salomé, Andersson Barison, Emmanoel Vilaça Costa, Spectroscopic investigation, vibrational assignments, HOMO-LUMO, NBO, MEP analysis and molecular docking studies of oxoaporphine alkaloid liriodenine, Spectrochim. Acta A Mol. Biomol. Spectrosc. 174 (2017) 94–104, https://doi.org/10.1016/j.saa.2016.11.018. Mar 5.

[54] J. Murray, K. Sen, Molecular Electrostatic Potentials Concepts and Applications, 1st Edition, 3, Elsevier Science BV, Amsterdam The Netherlands, 1996. ISBN: 9780080536859.

[55] V. Vidhya, A. Austine, Arivazhagan, Molecular structure, aromaticity, vibrational investigation and dual descriptor for chemical reactivity on 1- chloroisoquinoline using quantum chemical studies, Results Mater. 6 (2020) 100097, https://doi.org/10.1016/j.rinma.2020.100097.

[56] K. Fukui, Role of frontier orbitals in chemical reactions, Science 218 (1982) 747–754.

[57] G. Demirtas¸, N. Dege, H. Içbudak, ¨O. Yurdakul, O. Büyükgüngör, Experimental andDFT studies on Poly [di-μ 3-acesulfamato-O, O: O′; O′: O, O-di-μ-acesulfamato-O, O; N-di-μ-aqua-dicalcium (II)] complex, J. Inorg. Organomet. Polym. Mater. 22 (2012) 671–679.

[58] A. Sethi, R. Prakash, Novel synthetic ester of Brassicasterol, DFT investigation including NBO, NLO response, reactivity descriptor and its intramolecular interactions analyzed by AIM theory, J. Mol. Struct. 1083 (2015) 72–81, https://doi.org/10.1016/j.molstruc.2014.11.028.

[59] B. Silvi, A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature 371 (6499) (1994) 683–686.

[60] A. Sagaama, O.Al-D N.Issaoui, M.J.W. A.S.Kazachenko, Non-covalent interactions and molecular docking studies on morphine compound, J. King Saud. Univ.-Sci. 33 (8) (2021) 101606.

[61] N.S. Venkataramanan, Y.K Suvitha, Intermolecular interaction in nucleobases and dimethyl sulfoxide/water molecules: a DFT, NBO, AIM and NCI analysis, J. Mol. Graph. Model. 78 (2017) 48–60, https://doi.org/10.1016/j.jmgm.2017.09.022.

[62] E.R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A.J. Cohen, W. Yang, Revealing noncovalent interactions, J. Am. Chem. Soc. 132 (2010) 6498–6506, https://doi.org/10.1021/ja100936w.

[63] E. Nemati-Kande, R. Karimian, V. Goodarzi, E. Ghazizadeh, Feasibility of pristine, Al-doped and Ga-doped Boron Nitride nanotubes for detecting SF4 gas: a DFT, NBO and QTAIM investigation, Appl. Surf. Sci. 510 (2020) 145490, https://doi.org/10.1016/j.apsusc.2020.145490.

[64] A.D. Isravel, J.K. Jeyaraj, W.J.J. Thangasamy, DFT, NBO, HOMO-LUMO, NCI, stability, Fukui function and hole – Electron analyses of tolcapone, Comput. Theor. Chem. 1202 (2021) 113296, https://doi.org/10.1016/j.comptc.2021.113296.

[65] D.L. Lande, S.A. Bhadane, S.P. Gejji, Noncovalent interactions accompanying encapsulation of resorcinol within azacalix[4]pyridine macrocycle, J. Phys. Chem. 121 (2017) 1814–1824, https://doi.org/10.1021/acs.jpca.6b12912.

[66] K. Arulaabaranam, G. Mani, S. Muthu, Computational assessment on wave function (ELF, LOL) analysis, molecular confirmation and molecular docking explores on 2-(5-Amino-2- Methylanilino)-4-(3-pyridyl) pyrimidine, Chem. Data Collections (2020) 100525, https://doi.org/10.1016/j.cdc.2020.100525.

[67] B. Silvi, A. Savin, Classification of chemical bonds based on topological analysis of electron localization functions, Nature 371 (1994) 683–686, https://doi.org/10.1038/371683a0.

[68] F. Akman, A. Demirpolat, AS. Kazachenko, AS. Kazachenko, N. Issaoui, O. Al-Dossary, Molecular structure, electronic properties, reactivity (ELF, LOL, and Fukui), and NCI-RDG studies of the binary mixture of water and essential oil of phlomisbruguieri, Molecules. 28 (6) (2023) 2684, https://doi.org/10.3390/molecules28062684.

[69] J. Heiko, Localized-orbital locator (LOL) profiles of chemical bonding, Can. J. Chem. 86 (7) (2008) 695–702, https://doi.org/10.1139/v08-052.

[70] I. Rozas, I. Alkorta, J. Elguero, Behavior of ylides containing N, O, and C atoms as hydrogen bond acceptors, J. Am. Chem. Soc. 122 (45) (2000) 11154–11161, https://doi.org/10.1021/ja0017864.

[71] K. Srivastava, MR. Shimpi, A. Srivastava, P. Tandon, K. Sinha, SP. Velaga, Vibrational analysis and chemical activity of paracetamol-oxalic acid cocrystal based on monomer and dimer calculations:DFT and AIM approach, RSC. Adv. 6 (12) (2016) 10024–10037, https://doi.org/10.1039/C5RA24402A.

[72] H.R. Belhouchet, T. Abbaz, A. Bendjedou, A. Gouasmia, D. Villemin, A computational study of the inclusion of β-cyclodextrin and nicotinic acid: DFT, DFT-D, NPA, NBO, QTAIM, and NCI-RDG studies, J. Mol. Model. 28 (11) (2022) 348, https://doi.org/10.1007/s00894-022-05342-1.

[73] M.L. Beatrice, S.M. Delphine, M. Amalanathan, M.S.M. Mary, Robert, molecular structure, spectroscopic, Fukui function, RDG, anti-microbial and molecular docking analysis of higher concentration star anise content compound methyl 4-methoxybenzoate-DFT study, J. Mol. Struct. 1238 (2021) 130381, https://doi.org/10.1016/j.molstruc.2021.130381.

[74] M. H, D. Usha, M. Amalanathan, R.M Geetha, M. S. Spectroscopic (IR, Raman, UV, NMR) Characterization and investigation of reactive properties of pyrazine-2-carboxamide by anti-bacterial, anti-mycobacterial, fukui function, molecular docking and DFT calculations, Chem. Data Collections (2020) 100583, https://doi.org/10.1016/j.cdc.2020.100583.

[75] P.W. Ayers, R.G. Parr, Variational principles for describing chemical reactions: the Fukui function and chemical hardness revisited, J. Am. Chem. Soc. 122 (2000) 2010–2018, https://doi.org/10.1021/ja9924039.

[76] A.O. Zacharias, A. Varghese, K.B. Akshaya, M.S. Savitha, L. George, (DFT, spectroscopic studies, NBO, NLO and Fukui functional analysis of 1-(1-(2, 4-difluorophenyl)-2-(1H-1, 2, 4-triazol-1-yl) ethylidene) thiosemicarbazide, J. Mol. Struct. 1158 (2018) 1–13, https://doi.org/10.1016/j.molstruc.2018.01.002.

[77] C. Morell, A. Grand, A. Toro-Labbé, New dual descriptor for chemical reactivity, J. Phys Chem A 109 (1) (2005) 205–212, https://doi.org/10.1021/jp046577a. Jan 13.

[78] H.M. Robert, D. Usha, M.G Amalanathan, M.S.M. Mary, Vibrational spectral, density functional theory and molecular docking analysis on 4-nitrobenzohydrazide, J. Mol. Struct. 128948 (2020), https://doi.org/10.1016/j.molstruc.2020.128948.

[79] B. Fathima Rizwana, C. Prasana, S. Muthu, C.S. Abraham, Molecular docking studies, charge transfer excitation and wave function analyses (ESP, ELF, LOL) on valacyclovir: a potential antiviral drug, Comput. Biol. Chem. 78 (2019) 9–17, https://doi.org/10.1016/j.compbiolchem.2018.11.014.

[80] D.M. Suresh, M. Amalanathan, S. Sebastian, D. Sajan, I. Hubert Joe, V. Bena Jothy, I. Nemec, Vibrational spectral investigation and natural bond orbital analysis of pharmaceutical compound 7-Amino-2,4-dimethylquinolinium formate – DFT approach, J. Mol. Struct. 1223 (2020) 128948, https://doi.org/10.1016/j.saa.2013.06.077.

[81] J. Neugebauer, M. Reiher, C. Kind, B.A. Hess, Quantum chemical calculation of vibrational spectra of large molecules – Raman and IR spectra for buckminsterfullerene, J. Comput. Chem. 23 (9) (2002) 895–910, https://doi.org/10.1002/jcc.10089.

[82] C.S. Casari, M. Tommasini, R.R. Tykwinskic, A. Milani, Carbon-atom wires: 1-D systems with tunable properties, Nanoscale 8 (2016) 4414, https://doi.org/10.1039/c5nr06175j.

[83] B. Smith, Infrared Spectral Interpretation. A Systematic Approach, DOI, CRC Press, Washington, DC, 1999, https://doi.org/10.1201/9780203750841.

[84] M. Rumi, G. Zerbi, Conformational dependence of vibrational and molecular nonlinear optical properties in substituted benzenes: the role of π-electron conjugation and back-donation, J. Mol. Struct. 11 (1999) 509, https://doi.org/10.1016/S0022-2860(99)00207-0.

[85] N. Sundaraganesan, B.D.JS. Ilakiamani, Vibrational spectroscopy investigation using ab initio and density functional theory analysis on the structure of 3, 4-dimethylbenzaldehyde, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 68 (3) (2007) 680–687.

[86] V.A. Danie, B.J. Kumari, N.F. Nirmal, T.A.F. Reji, Molecular structure, vibrational assignment, HOMO-LUMO and mulliken analysis of 2-(4-amino-2-phenylaminothiazol-5-oyl)-N-methylbenzimidazole by DFT method, J. Res. Sci. 2 (2014) 21–24. ISSN: 2278-9073.

[87] M. Medimagh, N. Issaoui, S. Gatfaoui, O. Al-Dossary, A.S. Kazachenko, H. Marouani, M.J. Wojcik, Impact of non-covalent interactions on FT-IR spectrum and properties of 4-methylbenzylammonium nitrate. A DFT and molecular docking study, Heliyon 7 (10) (2021) e08204, https://doi.org/10.1016/j.heliyon.2021.e08204.

[88] M. Amalanathan, T. Brintha, S. Sijana, P.J. Jegan Babu, M. Sony Michael Mary, Identification of interaction of 1,2,3,6-Tetrahydro-2,6-dioxo-4-pyrimidinecar-boxylic acid using DFT studies, molecular docking, biological activity and topology analysis for biological applications, Chem. Phys. Impact 7 (2023) 100393, https://doi.org/10.1016/j.chphi.2023.100393.

[89] G. Varsanyi, Vibrational Spectra of Benzene Derivatives, Academic Press, New York, 1969. ISBN: 9780323150668.

[90] C. James, G.R. Pettit, O.F.N Nielsen, V.S. Jayakumar, I. Hubert Joe, Vibrational spectra and ab initio molecular orbital calculations of the novel anti-cancer drug combretastatin A-4 prodrug, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 70 (5) (2008) 1208–1216, https://doi.org/10.1016/j.saa.2007.10.052.

[91] J.D. Deephlin Tarika, X.D. Divya Dexlin, S. Madhankumar, D.D Jayanthi, T. J Beaulae, Tuning the computational evaluation of spectroscopic, ELF, LOL, NCI analysis and molecular docking of novel anti COVID-19 molecule 4-dimethyla-mino pyridinium 3, 5-dichlorosalicylate, Spectrochim. Acta A Mol. Biomol. Spectrosc. 259 (2021) 119907, https://doi.org/10.1016/j.saa.2021.119907.

[92] A. Sakamoto, M. Tasumi, Symmetry of the benzene ring and its normal vibrations: the "breathing" mode is not always a normal vibration of a benzene ring,

J. Raman Spectrosc. 52 (12) (2021) 2282–2291, https://doi.org/10.1002/jrs.6131.

[93] C.S.C. Kumar, C. Parlak, H.K. Fun, M. Tursun, G. Kesan, S. Chandraju, C.K. Quah, Experimental and theoretical FT-IR, Raman and XRD study of 2-acethyl-5-chlorothiophene, Spectrochim. Acta 127A (2014) 67–73, https://doi.org/10.1016/j.saa.2014.02.033.

[94] A. Atilgan, S. Yurdakul, Y. Erdogdu, M.T. Güllüoğlu, DFT simulation, quantum chemical electronic structure, spectroscopic and structure-activity investigations of 4-acethylpyridine, J. Mol. Struct. 1161 (2018) 55–65, https://doi.org/10.1016/j.molstruc.2018.01.080.

[95] K.C. Medhi, The vibrational spectra of 2-, 3-and 4-acethylpyridine, Indian J. Phys. 51A (1977) 399–413.

[96] M.M. El-Nahass, M.A. Kamel, A.F. El-deeb, A.A. Atta, S.Y. Huthaily, Ab initio HF, DFT and experimental (FT-IR) investigation of vibrational spectroscopy of P-N,N-dimethylaminobenzylidenemalononitrile (DBM), Spectrochim. Acta A Mol. Biomol. Spectrosc. 79 (3) (2011) 443–450, https://doi.org/10.1016/j.saa.2011.02.055.

[97] 1and O. Belaidi, 1.T. Bouchaour, U. Maschke, Structural preferences and vibrational analysis of 2-hydroxy-2-methyl-1-phenylpropan-1-one: a computational and infrared spectroscopic research, J. Struct. (2013) 1–9, https://doi.org/10.1155/2013/942302.

[98] MS. Almutairi, S. Muthu, JC. Prasana, B. Chandralekha, AR. Al-Ghamdi, MI. Attia, Comprehensive spectroscopic (FT-IR, FT-Raman, 1H and 13C NMR) identification and computational studies on 1-acetyl-1H-indole-2,3-dione, Open. Chem. 15 (2017) 225–237, https://doi.org/10.1515/chem-2017-0026.

[99] O. Ursu, A. Rayan, A. Goldblum, TI. Oprea, Understanding drug-likeness Wiley interdisciplinary reviews, Comput. Mol. Sci. 1 (5) (2015) 760–781, https://doi.org/10.1002/wcms.52.

[100] G. Richard Bickerton, GV. Paolini, J. Besnard, S. Muresan, AL. Hopkins, Quantifying the chemical beauty of drugs, Nat. Chem. 4 (2) (2012) 90–98, https://doi.org/10.1038/nchem.1243.

[101] C.A. Lipinski, Lead-and drug-like compounds: the rule-of-five revolution, Drug Discov. Technol. 1 (4) (2004) 337–341, https://doi.org/10.1016/j.ddtec.2004.11.007.

[102] A. Ghose, Knowledge-based chemoinformatic approaches to drug discovery, Drug Discov. Today 11 (23–24) (2006) 1107–1114, https://doi.org/10.1016/j.drudis.2006.10.012.

[103] H. Tijjani, A. Olatunde, A.P. Adegunloye, A.A. Ishola, In silico insight into the interaction of 4-aminoquinolines with selected SARS-CoV-2 structural and nonstructural proteins, Coronavirus Drug Discov. 3 (2022) 313–333. Druggable Targets and In Silico UpdateID: covidwho-2149151.

[104] M. Berry, B. Fielding, J. Gamieldien, Practical considerations in virtual screening and molecular docking, emerging trends in computational biology, Bioinform. Syst. Biol. (2015) 487–502, https://doi.org/10.1016/B978-0-12-802508-6.00027-2.

Preparation And Characterization Of Polyacrylic Acid

K Reema Raj[a] and R. Murali*

[a]PG Student, PG and Research Department of Chemistry, Nanjil Catholic College of Arts and Science, Kaliyakkavilai- 629153, Tamilnadu, India
*Assistant Professor and Head, PG and Research Department of Chemistry, Nanjil Catholic College of Arts and Science, Kaliyakkavilai- 629153, Tamilnadu, India

ABSTRACT

Polyacrylic acid is an acrylic polymer. It is a synthetic polymer with a high molecular weight. Polyacrylic acid is synthesized using Free radical polymerization, using hydrogen peroxide as initiator as well as an oxidizing agent and ascorbic acid as reducing agent and with the monomer acrylic acid to give polyacrylic acid. It's like a free radical redox reaction.

The characterization of Polyacrylic acid can be done through spectroscopic studies like Fourier Transform Infrared Spectroscopy (FTIR) and UV spectroscopy. FTIR determines the functional group of polyacrylic acid and the UV analysis determines the metal present in the polymer. PAA can be used as thickening agents in pharmaceuticals and cosmetics, paint additives, adhesives and sealant chemicals etc. PAA is a superabsorbent, non-toxic, non-irritating and biodegradable synthetic polymer.

Keywords: Polymer, Molecular Weight, FTIR, Uv-Visible

INTRODUCTION

Polymers are amorphous solid consisting of very large molecules known as macromolecules, containing many repeating subunits. The word polymer arise from Greek words called 'Poly' as "many" and 'mer' means "unit". Polymer molecules are very large and has a high molecular weight. Polymers are classified as natural, bio, and synthetic Polymers. Polymers has no specific length and they don't form crystalline structure. Polymers have unique properties depending on the type of molecules being bonded and how they are bonded ie, some Polymers have properties like bending and stretching eg, rubber and polystyrene, but others are hard and brittle and tough eg, epoxies and glass (1).

Polymers has wide variety of applications in our modern and developing world, such as Polymers are used for making plastics many household things also. Polymers are used in industries in pharmaceuticals and in many other fields (2-4).

Polymerization is a method where the synthetic Polymers are produced by combaining smaller molecules that is monomers into a chain which is held together by the means of covalent bond.

Now in this modern world Polymers are used on every single area ie,for example grocery bags, soda and water bottles, textiles, paints, food packaging, pharmaceuticals and cosmetics etc (5).

Polyacrylic Acid

Polyacrylic acid is an amorphous synthetic polymer with a high molecular weight. PAA is a non-toxic, non-irritating, biocompatible and 26% biodegradable polymer, has many variety of applications in many fields which have been found in recent years. It is hygroscopic when absorbs water and brittle under normal condition

Polyacrylic acid has the ability to absorb and hold water molecules even if it expands to many times from their original volume. This behavior of polyacrylic acid makes a base for varieties of applications. Its approximate molecular weight is 450,000 and it cannot be dissolved in its monomer but is soluble in water, dioxane, ethanol, methanol and it's viscosity is 700cp. PAA is water sensitive polymer to serve as plastics.

MATERIALS:

Acrylic acid as monomer, hydrogen peroxide, ascorbic acid, and sodiumhydroxide and chromiumoxide.

METHOD:

Into an 200 ml clean and dry beaker 10ml of acrylic acid is taken and about 2.5g of ascorbic acid is weighed and transferred in an another clean dry 100ml beaker and is dissolved in 10ml of distilled water. This 10ml ascorbic acid is then combined with the 10ml of acrylic acid taken in the 200ml beaker and this mixture is kept on a magnetic stirrer. 2ml of hydrogen peroxide was added drop by drop while stirring.

Figure 1. Polyacrylic Acid

From the prepared polymer the metal addition is done. The metal addition is done in two ways, direct method and indirect. For direct method, a small amount of polymer about 6g is weighed and about 5ml of 1M chromiumoxide is added and a polymer metal complex is obtained. In indirect metal addition, about 3g of polyacrylic acid is weighed and dissolved in 200ml water taken in a 250ml beaker and kept in a magnetic stirrer. After dissolving completely 25ml of chromiumoxide is added and continued stirring for 30 minutes after that polymer metal complex is formed.

RESULT AND DISCUSSION:

Polyacrylic acid was successfully synthesised via free radical redox polymerization. It also undergoes redox reaction as the hydrogen peroxide act as an oxidizing agent and the ascorbic acid act as reducing agent. It's an uncontrollable free radical Polymerization reaction.

Obtained product that is Poly acrylic acid is in solid form with a pale yellow transparent and on oxidation with atmospheric oxygen it becomes dark brown color. On adding water to the polymer it absorbs maximum water and swell. Due to this water absorbing property it can be used for many applications one of its applications is water treatment. And of capable of forming gels it is used for varieties of applications such as in pharmaceuticals, cosmetics, skin care products etc.

The molecular structure is determined using FTIR Spectroscopy (NIIST Trivandrum) and metal addition is analyzed by UV analysis taken from (NMCC College,Marthandam).

Fourier Transform Infrared Spectroscopy:
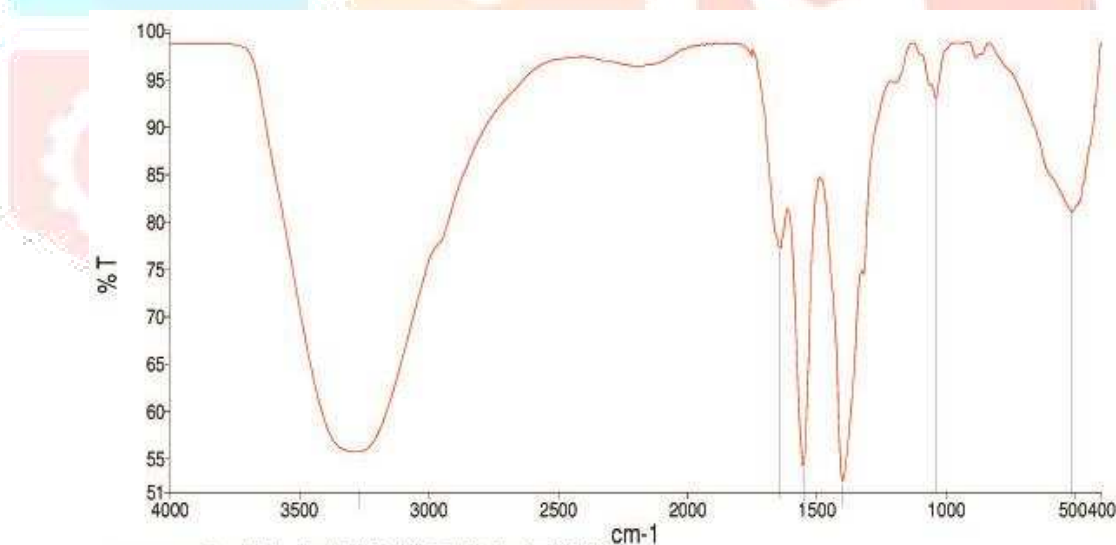
List of Peak Area/Height		
Peak Number	X (cm-1)	Y (%T)
1	3270.72	55.72
2	1640.04	77.26
3	1552.24	54.20
4	1399.26	52.63
5	1039.48	93.16
6	512.43	81.10

Figure 2: FTIR of Polyacrylic acid

In FTIR analysis it shows the presence of six peaks which represents the functional group present. Here $3270.72cm^{-1}$ indicates the presence of –OH group. $1640.04cm^{-1}$ represents the presents of mono-substituted alkenes, $1552.24cm^{-1}$ shows the presence of carboxylic acids group, $1399.26cm^{-1}$ indicates the presence of -CH aliphatic bending group and $1039.48cm^{-1}$ represent that it contains polysaccharide and also has a peak range at $512.43cm^{-1}$ which indicates the presence of a specific chemical bond or functional group and contains a C=O bond stretching.

UV analysis of Metal addition in Polyacrylic acid:

Chromium metal in Poly(acrylic) acid is determined by UV analysis. The UV analysis shows an peak at the range of 372.16nm, 0.597abs for sample 'A' and it shows 373.17nm, 1.580abs for sample 'B'. That is the sample may contain nano-particles and may have surface Plasmon resonance.

Direct Metal Addition

Indirect Metal Addition

Figure 3: Uv spectra of Polyacrylic acid after metal addition

CONCLUSION

The preparation and characterization of polyacrylic acid is done successfully via series of experiments and analysis. It is synthesized via free radical redox polymerization and characterized using FTIR spectrum and UV spectrum. It acts as a good absorbing compound so called superabsorbant. It is 26% biodegradable and biocompatible synthetic polymer.

REFERENCES

[1] Gowariker, V. R, Viswanathan, N V, Jayadev Sreedhar. Polymer Science. New age International Publishers, 4th Edition, 2018.

[2] Awad, S A. and Khalaf, E M. 2019. Investigation of improvement of properties of polypropylene modified by nanosilica composities. Compos. Commun, 12: 59-63.

[3] Marta Worzakowska. 2015. Thermal and mechanical properties of polystyrene modified with esters derivatives of 3-phenylprop-2-en-1-ol. J Therm Anal Calorim, 121: 235-243.

[4] Arvanitoyannis, I. and Biliaderis C. G. 1999. Physical properties of polyol-plasticized edible blends made

   of methyl cellulose and soluble starch. Carbohydr Polym, 38:47-58.

[5] Matyjaszewski, K. 1996. Curr. Opin. Solid State Mater. Sci. , 1, 769.

Cloud-Based Top-Down and Bottom-Up Approach for Agriculture Data Integration

Mr. Abhilash Sam Paulstin K. C.[1], Dr. Angeline Prasanna G.[2]

Abstract: The big data revolution and the growth of information technology have had a profound influence on many areas of our life. Many farms today employ precision farming techniques and capture enormous volumes of data. Making the most of these datasets for decision support requires integrating data from several sources, doing analysis rapidly, and generating conclusions based on the results. In order to analyse agricultural data from several sources, this study presents a framework that utilises cloud-computing. This approach is more scalable, adaptable, cost-effective, and easy to maintain than existing alternatives. Based on extensive interviews, surveys, and literature reviews, the framework offers a workable architecture for cloud-based services for data integration, analysis, and visualisation. This skeleton architecture was used to construct many programmes; as we learned more and faced tougher problems, we tweaked the plan to make it work better. We demonstrate the framework's value with several example applications. Each use case has its own specific requirements for data integration; therefore, it makes use of a different set of services from the suggested architecture.

Keywords: Agriculture, Big Data, cloud computing, data analysis, data integration, data visualisation, decision support systems.

1. Introduction

Precision farming techniques have significantly increased as a result of the usage of IT and big data in agriculture. (also known as Agriculture 3.0). Sensors, satellites, and other technologies are used by a wide variety of agricultural information systems (IS) to gather information, such as soil moisture, nitrogen concentrations, and temperature gradients between the ground and the air. In addition, databases spanning many years are increasingly becoming accessible, allowing researchers to examine data trends across time. These advancements have led to a "data overflow" problem, For effective data retrieval and analysis, the article emphasises the significance of cognitive storage and processing abilities. Another issue is that many farms still use incompatible data management systems, meaning they are missing out on opportunities to enhance the quality of their decision-making tools.

Information gathered and analysed from many sources is crucial for better agricultural decision-making. For example, in addition to weather, topological terrain, irrigation, and agricultural yield data, information on beehive treatment, plants growing in the area of the hive, and honey production levels might be useful for making management decisions. Integrating data is a great first step toward better decision making, but it is not sufficient on its own. Users also need quick and simple access to the data and the ability to make actionable conclusions.

A major deficit in the capacity to combine data from different agricultural resources was already apparent in the early days of recent years. The large amounts of data generated by precision agriculture constitute a "data overflow" concern. They emphasise the need of standardised data transfer and tools for managing and analysing information. There have been subsequent presentations of other strategies for integrating and managing agricultural data.

In order to analyse agricultural data from several sources, this study presents a framework that utilises cloud-computing. This approach is more scalable, adaptable, cost-effective, and easy to maintain than existing alternatives.

The framework establishes a practical architecture of cloud-based services for the purposes of data integration, analysis, and visualisation. Users can take use of the framework's included services, or they can utilise those services as a foundation to create their own. It should be noted that the framework is meant to include and process data from both new and current external databases, while making use of existing software services whenever practical.

Numerous applications and demonstrations of the framework's effectiveness are provided. Each use case has its own specific requirements for data integration; therefore, it makes use of a different set of services from the suggested architecture.

[1] Research Scholar, Department of Computer Science, AJK College of Arts and Science, Coimbatore, Tamilnadu, India.
ORCID ID : 0000-3343-7165-777X
[2] Former Associate Professor and Head, Department of Computer Science, AJK College of Arts and Science, Coimbatore, Tamilnadu, India.
ORCID ID : 0000-3343-7165-777X
* Corresponding Author Email: abhi.sam83@gmail.com

The study makes use of cloud computing, which has benefits including cheaper prices, on-demand computational capabilities, flexibility, and outsourced IT infrastructure management, to analyse and aggregate agricultural data from diverse sources. This is especially crucial in agriculture, where farmers sometimes lack the tools for digitalization and vast amounts of data must be monitored and analysed since data is constantly being collected.

The research approach, design science, requirement analysis stage, background knowledge on cloud computing and data integration techniques are all included in the article. Along with two use cases and a functional framework to illustrate its use and many features, it also offers a full analysis of the results. The conclusions are discussed in Section 4.

2. Literature Survey

The study's two key goals are:

In order to improve farm management and decision-making, it would be useful for farmers to have access to a framework outlining a cloud-based services architecture. The infrastructure services allow for the development of novel agricultural offerings.

Explain how you may use cloud computing to integrate and analyse data in the agricultural sector. Here, we provide two programmes that might serve as examples when developing novel data-integration services.

Our research strategy was motivated by the design science paradigm [1]. Based on 1) Through a study of the literature, unstructured interviews with Indian farmers, and agricultural research, a framework for cloud-based services for agriculture was created.

Based on the requirements of diverse farms, a framework is created for various farm situations. On top of the framework, cloud-based services are developed, with illustrations shown for each situation.

- Developing a system to compile data from many sources on the RPW's geographic distribution and display it on a map.
- The development of an ontology-based pest control DSS.
- Models can optimise irrigation schedule, estimate yield, and predict disease by using data from sources including soil and plant sensors and weather stations.
- Abeehive management system that keeps track of treatment data for all of the beehives is essential for analysing productivity in connection to treatments and bee genetic lines.
- We built a dashboard that aggregates data from several sources and presents several key performance indicators (KPIs) pertinent to farming in order to better help farmers in making educated decisions.
- A major field crop farm is now putting a cloud-based farm management system into place.
- Development of an agricultural machinery selection aiding system.
- Making a smartphone app to gather and analyse complaints of insect damage from a broad population.

Due to space limitations, we will only be able to cover the first two conceivable uses. An ETL process that involves data collection, transformation, and loading into a single database is shown in the first use case (RPW). The second example (DSS for pest control) exemplifies knowledge management and reuse while highlighting the Semantic Web's and ontologies' possibilities. (Irrigation recommendation), we see how machine learning technology combined with data fusion can forecast how much water should be utilised for irrigation.

As we developed applications and services to fulfil the required use cases, we improved and extended the framework.

Following this is an introduction to the issue of data integration, followed by detailed explanations of how to do a needs assessment and make use of cloud computing for agricultural purposes.

2.1. Needs Assessment

As mentioned above, the first framework was developed with the use of interviews and surveys to gather information on farmers' and experts' needs from a variety of agriculturally-related agencies. In all, 117 farmers were polled [2]. In the polls, farmers were asked whether or not they were currently using certain agricultural treatments, whether they were familiar with them but had no intentions to employ them, and whether or not they were familiar with them at all. Different application-specific responses are shown in Figure 1's frequency distribution. With these results in mind, it's evident that many agricultural applications are crucial for farmers right now.
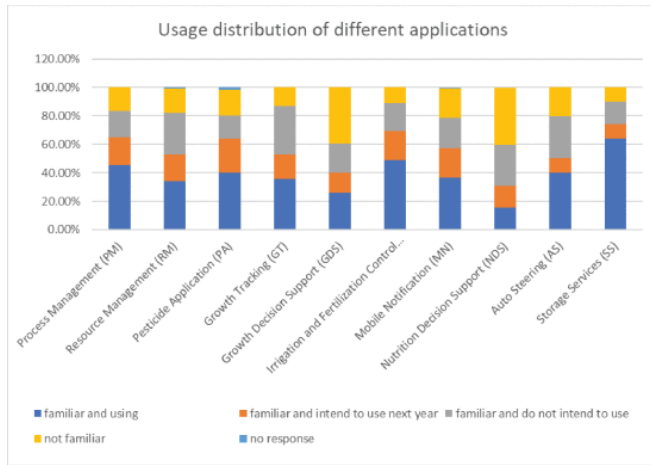
Fig 1. Proportions of the population that have used, are familiar with, or intend to use a certain app.

According to [2]'s research, by conducting interviews with farmers and specialists from different agricultural organisations, we were able to compile the following broad areas of functionality required in agricultural software.

- There is an urgent need for individualised management and decision-making resources for rural professionals.
- The interface of a system should be easy to use and flexible enough to accommodate different sorts of users. Software with a basic interface is ideal for those with less technical knowledge.
- Automated and straightforward software is required for efficient data management. The user should be able to simply add and programme new automated operations into the system.
- Moreover, the user should have complete control over the processing and analytic parameters. Advanced users may want to try out new methods.
- It's essential to have specialised knowledge, particularly rule-based expertise, to adapt systems to particular environments and take into account users' abilities, habits, and preferences. (such as risk profiling, for instance).
- More standardised, top-notch computer systems are urgently required to cut down on learning time and effort and guarantee excellent technological performance.
- We have read research on cloud-based agricultural systems and precision agriculture, as well as extra needs like:
- Assist with cordless, unobtrusive spying [2].
- The interface is offered in the farmer's native tongue, and the same information may be supplied in other systems, native languages. [3].
- storage capacity [7].
- Multiple types of information must be measured or gathered to give useful decision assistance. Measuring data (together with timestamps) is better

collected over time than it is collected quickly and then forgotten.
- The text promotes the use of open data standards, interfaces, and protocols for easy interoperability with other software applications and data sources, both locally and remotely over the internet, which is essential for facilitating compatibility with more antiquated systems and decentralised architectures. [5].
- The capacity to support a sizable and growing number of users and data sources [6].
- Meta-data support for data portability across programmes; this feature allows for easier sharing of information across programmes.
- Cost Effective.

2.2. Services using Cloud Computing

On-demand self-service, which enables users to provision resources as needed, broad network access, which encourages the use of a variety of client platforms, including smartphones, tablets, and personal computers, and elastic resource provisioning, which enables resource access over the network are all features of cloud computing.

As a result of these features, cloud computing offers several advantages. As a first step, cloud computing makes it possible for customers to pay for computing resources and services on an as-needed basis [3], [7]. This reduces or eliminates the need for customers to make large initial expenditures in computer hardware and software. The second advantage is that it enhances the scalability and responsiveness of programmes to altering business requirements by making computer resources accessible swiftly (within minutes). In addition, virtualization and other cloud-specific parallel processing technologies (like Spark and Hadoop) boost application performance [4].

Users may access networked computers from any place by outsourcing IT infrastructure management to external cloud service providers, which can lower IT maintenance costs and remove the need for internal IT staff. [4].

While infrastructure-as-a-service (IaaS) offers processing, storage, and network resources on demand for software deployment, software-as-a-service (SaaS) is the cloud-based deployment and operation of software. Platform-as-a-Service (PaaS) enables programmers to create and deploy software using the IDEs, libraries, APIs, and tools provided by cloud computing, which is rapidly being investigated for use in the agricultural sector.

With the use of precision agricultural technology, farmers and scientists may gather enormous volumes of data on variables such as crop production, topographical characteristics, moisture content, and temperature, allowing efficient reaction to unforeseen conditions.

The majority of farmers, however, do not have access to powerful computers, which restricts their capacity to analyse this data online.

Farmers without the resources to invest in IT infrastructure can benefit from cloud computing services, which offer on-demand access to agricultural data, faster processing, and thorough analysis, including details on disease transmission, pest control, weather information, and best practises.

In order to help farmers make educated choices, the study provides a cloud-based framework for creating SaaS applications with an agricultural emphasis. This framework combines data from many sources, does analysis, and graphically shows results.

2.3. Tools for Data Integration

The suggested architecture unifies data from several sources, including relational databases (where data are obtained using SQL), no-SQL databases (such as those based on JSON), CSV files, This section contains background knowledge on HTML files on the Web, which are less prevalent but vital, and ontologies and triplestores, which are required for implementing the second use case.

An ontology is a precise, computer-readable description of domain knowledge. Formal, computer-interpretable ideas known as ontologies specify domain entities, properties, and links. [9] They are increasingly used as a foundational tool in knowledge management systems, DSSs, and other intelligent systems to describe domain information.

An ontology is a formal description of domain knowledge that computers may use to make inferences about the world. For instance, the insect Red mite may be controlled using Agriron's active component, abamectin. [9] If the ontology states that a pesticide is successful if one of its compounds is effective against that bug, then this conclusion may be drawn. The foundation of knowledge bases for DSSs and expert systems may be formed via ontologies.

With the use of ontologies, databases and application programmes may work together without exchanging data structures. This makes it possible to automatically extract and aggregate data from two applications or websites that share the same taxonomy. The ideas presented in one app can be transferred to the corresponding ideas presented in the other app.

Last but not least, ontologies enable the reuse of specialised knowledge. When an ontology is made available to the public, it may be utilised by applications across numerous industries. An further advantage of ontologies is that they may be joined to create a more thorough ontology of a domain [8]. A few of open ontologies that have found application in agriculture include AGROVOC [6] and DBpedia. The United Nations' Food and Agriculture Organization (FAO) publishes a multilingual thesaurus called AGROVOC, DBpedia, an important ontology, adopted from the well-known Wikipedia, gives definitions and features of agricultural subjects in 17 different languages, making it simple to combine material from different languages. As a result, it provides access to billions of unique words (Things) and cross-references to several other dictionaries.

When it comes to building out the so-called Semantic Web, ontologies are a must. The Semantic Web, also known as the Web of Data or the Web of Linked Data, is a set of standards developed by the World Wide Web Consortium (W3C) to provide a formal representation of the information available on the World Wide Web, facilitating data sharing and reuse in various contexts.

HTML pages, which can be read by both people and machines, currently make up the overwhelming majority of online content. Machines can only do a keyword search of these papers and cannot understand the text itself. The Semantic Web project aims to fix this problem by using ontologies to make text in online publications machine-readable. Ontologies offer a formal vocabulary to describe the relationships between digital information and physical things (i.e., instances of classes). Standard ontologies make it possible to combine different Web-based data sets into a single database, such AGROVOC and DBpedia, which may then be used to improve communication across diverse agricultural efforts on the Semantic Web.

The World Wide Web Consortium's (W3C) guidelines and tools enable the Semantic Web. A paradigm for publishing and exchanging data, the Resource Description Framework (RDF), is extended by the Web Ontology Language (OWL) to support ontologies with richer semantics and knowledge inference. RDF data stored in international triplestores may be retrieved using the all-purpose query language SPARQL. Triplestores[10] are a specific type of graph database designed specifically for storing and retrieving RDF. Triples, which have a subject, a predicate (or verb), and an object, are used to hold the RDF data.

3. Results and Methodology

The framework's structure was developed after the data collected via questionnaires and interviews was thoroughly analysed. This skeleton architecture was used to construct many programmes; as we learned more and faced tougher problems, we tweaked the plan to make it work better. This section describes the top layer of the functional architecture as well as two more application scenarios that are incorporated into it.

A. System Design for a Framework

The framework architecture is depicted in Figure 2 with its four-tiered structure of functional services.

Relational databases, triple stores, XML files, papers, etc. are all part of the Data Layer's collection of data storage technologies. At this stage, data can be made either accessible to the public or kept confidential Only the farm's employees have access to the data it collects. Our second-layer IoT data extraction services are augmented by third-party Web services that grant access to external databases and sensors.

- The process of data ETL include gathering, converting, and loading data from numerous sources before storing it in a centralised system with a range of services.
- Provides SQL query services for accessing relational databases.
- SPARQL may be used to query triplestores as an ontology query language.
- Creators and shapers of ontologies.
- Services for managing access to information in accordance with users' roles and permissions inside an organisation.
- Data may be automatically extracted from many sources, cleaned, and fed into a centralised database using any number of existing ETL solutions, both commercial and bespoke.
- Providers of IoT data gathering services, with a focus on soil and remote sensing devices.

The Data Analysis Layer provides a variety of services for processing data and presenting findings in a form that facilitates decision making. Among the many available offerings are:

- Businesses that offer GIS services for spatially analysing agricultural data.
- In order to better inform suggestions and projections, several businesses now provide data mining services. Clustering of data is another service provided by machine learning algorithms Learning models, such as regression, classification, and clustering, identify related data and treat them appropriately to forecast the goal value of an observation, such as the predicted yield of a field, and categorise observations, such as pest photos or land plots.
- Business intelligence (BI) services use data acquired and combined at data access and integration levels to provide real-time analysis and clever visualisations.These kinds of services are invaluable to ranchers and farmers, allowing them to construct dashboards with more intelligence and effectiveness.
- Deep learning and other approaches are used by image processing services to automatically analyse

pictures, such as satellite photography, to comprehend their qualities or meaning.

The Applications Layer is a group of SaaS applications that provide assistance to farmers. At this stage, applications are constructed on the infrastructure laid out in Levels 1-4.
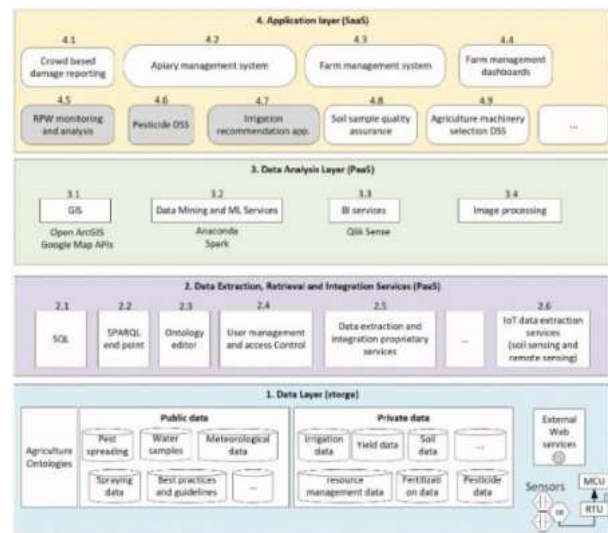
Fig 2. This is a schematic showing how the future service infrastructure is planned out.

B. Example(s) of Application

1) Monitoring and analysing data on the spread of the red palm weevil (RPW)

This use case is intended to demonstrate how data from many sources may be integrated to monitor an issue's development.

Snout beetle Rhynchophorusferrugineus (RPW) is a serious danger to Indian palm farms. The fast growth of RPW has increased the number of unhealthy trees, presenting hazards to both safety and the economy. Two organisations in India have been looking into the matter, but because of their disparate databases, it is difficult to provide a thorough analysis. In order to assist stop the spread of RPW, this use case attempts to provide a cloud-based solution that integrates data from these organisations with analytical tools that are visual and map-based.

The following requirements for the envisioned service were uncovered through interviews with many RPW monitoring groups.

- The RPW distribution shouldn't just be shown by area, but rather at the national level, hence data from several information systems should be merged to demonstrate this. Specify the infected palm tree count and the number of RPW caught and released.

- The service shouldn't interfere with the numerous procedures an organisation uses to accomplish its goals.
- The service should include tools for statistical analysis, such as data aggregation and filtering, as well as the ability to see the combined data on a map.
- Global, always-on access to information is a must.
- Due to planned actions to improve awareness of and encourage public reporting of sick trees, the system must be scalable and capable of managing additional data and bandwidth.

Through an ETL process, the service will combine data from diverse sources, enabling visualisation and analysis without altering current practises. Due to consumer needs and a lack of inter-firm collaboration, this strategy is favoured. Instead of developing a comprehensive information system that covers user interfaces and data input processes, the service focuses on data integration since it is quicker and less expensive to do so.

You'll need to take care of a visual representation of the combined data as part of the answer. To achieve this objective, we employ ArcGIS. ArcGIS from Esri (Esri.com) is a GIS that can be used to create and alter maps that depict information on the planet's topography [11]. After the mapping process is complete, researchers have many options for examining the aggregated data on the spread of RPW.

The solution architecture is comprised of the following parts:

- The consolidated database, depicted as the Data layer in FIGURE 2, requires the use of a relational database management system for its administration (RDBMS).
- The Data Extractor service (component 2.5 in FIGURE 2)is a Java software that collects data on RPW distribution using Web services, storing it in a unified database in CSV and JSON formats from different databases.
- • The creation of new maps and the modification of existing ones is made easier by ArcGIS Online, the web-based ArcGIS desktop GIS programme (see section 3.1 in FIGURE 2). In contrast to the offline version, which requires installation on client servers or rented cloud infrastructure, it enables users to share their maps with others. ArcGIS Online is accessible to users who pay as they go. Maps may be made using ArcGIS Online and the many data layers included in the data warehouse. The user has the freedom to dynamically select which layers to display from the complete set of database tables and views.
- Converting data from the centralised database to a comma-separated value format (CSV) for use in ArcGIS mapping applications (see component 4.5 of FIGURE 2). Thus, we developed some code to

retrieve the required data from the centralised database (through SQL - part 2.1 in FIGURE 2) and save it as CSV files on the cloud server. The software may export global and yearly summaries of palm trees, palm tree health, palm tree sickness, traps, and captures as comma-separated value (CSV) files. It was decided to automate the program's daily execution, thus the corresponding batch file was created and uploaded to the cloud server. Layers on the map are dynamic, reflecting updates to the related CSV files.

In Fig. 3, we can see how all the different components of the suggested design are linked together.

- The Data extractor service (3) receives real-time updates from two sources:
- The data is a comma-separated values (.csv) file from the Eden farm1 (1.1).
- The Agriculture Research Organization's data is provided via a web service that delivers the data in JSON format (1.2).
- It is possible to immediately import static data from Plant Protection and Inspection Services (2) into the central repository.
- The Data Extractor collects information from many sources and deposits it into a centralised database (4)
- The storage capacity of the unified database is increased by a wide variety of triggers.
- Database to ArcGIS Converter, the fifth component, transfers data from the centralised database into the ArcGIS format (6).
- ArcGIS Online is a data monitoring and analysis tool (7).
- Furthermore, users can make direct connections to the database from remote locations and run queries.
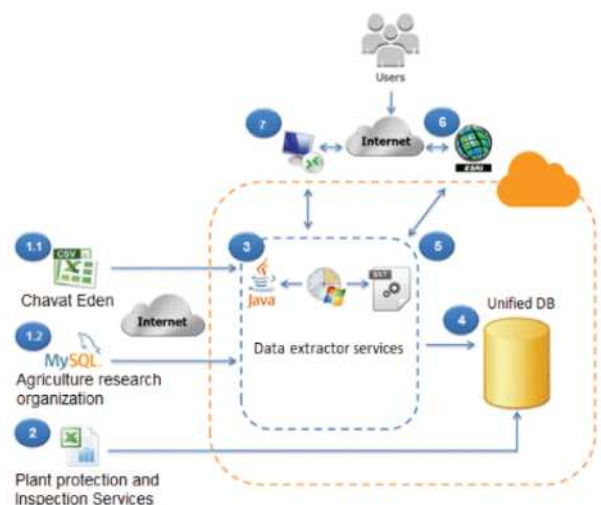
Fig 3. Description of the proposed methodology used.

With SQL Server serving as the primary database, the product is currently accessible on the AWS cloud. It is hosted on Amazon Elastic Compute Cloud. Machine

instances with the user-selected operating system, programmes, libraries, and data are started using the EC2 infrastructure service.[4] Using an Amazon Aurora DB instance in a VPC[12] and providing data extraction services as Lambda functions are two alternative deployment techniques..[13] When it comes to scalability, performance, and availability [14], go no farther than Amazon Aurora, a serverless relational database management system (RDBMS) that maintains data across several DB instances.Data from many sources must be combined in a single ArcGIS database using the comma-separated values (CSV) format.

2. Insecticide Application Decision Support System

This use case aims to demonstrate the integration of online data to develop a DSS for pest control.

A farmer must consider the pesticide's effectiveness, the pesticide's safe application, and the pesticide's cost when determining which pesticide to use to combat an insect that is damaging his crop. Farmers find it challenging to acquire information and make decisions since pest control data is often housed in many places and languages. An ontology-based Web system for pest control is recommended as a solution to this problem in order to handle the various pesticide regulations in the nations where crops are sold. enhancing efficiency and effectiveness. By employing a common terminology, this system would facilitate the coordinated presentation of data from many web resources.

Better decision support for farmers can help them use pesticides responsibly and in accordance with the regulations of the countries to which they export their goods.

The most efficient pesticides should be recorded by the pest-control DSS so that they may be used to exterminate the many pests that cause harm to agricultural products. It must take into consideration the maximum residue limit (MRL) for a given pesticide in the given nation, the number of days prior to harvest during which that pesticide may be applied, and the amount of pesticide that may be used on a given crop before reaching dangerous levels. Given that this information is dispersed across several resources and languages, an ontology seems like the best way to bring it all together.

It is hoped that the ontology will serve as the foundation for a Web service that may be used by farmers, the PPIS pesticide recommendations, and anyone involved in pest control.

Reusing ontology ideas from other ontologies and publishing the ontology to the Semantic Web are both necessary. Consequently, it must be formalised in accordance with Semantic Web standards.

In [15], we detail the steps taken to develop this ontology as well as its final product. Figure 4 represents the user's web browser; it communicates with web server. A SPARQL Server's triplestore now provides access to the ontology. When the results of a user's query are returned to them, the browser presents them to the user.
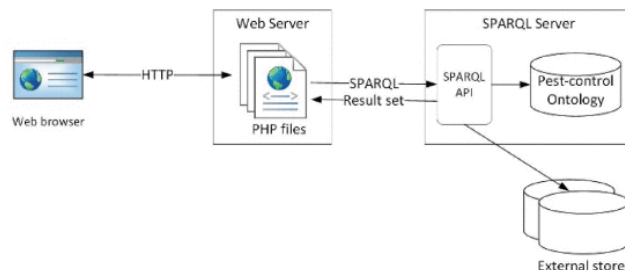
Fig 4. A diagram of the DSS used to spray for insects.

Apache Jena Fuseki Server originally kept the ontology in a triplestore, however Amazon Neptune provides a scalable, manageable deployment alternative for PHP applications that can be set up using Amazon's Elastic Container Service.

The application serves as an example of how Semantic Web ontologies may be used to reuse ideas across different databases. It creates hierarchical menus of crops and pests using pest pictures from DBpedia and AGROVOC, showing how concepts are connected online. so that we can quickly locate the pest or crop for which we are looking.

In conclusion, the two examples presented above demonstrate the framework's utility.

4. Conclusion

Using cloud computing, the system combines and analyses agricultural data from numerous sources, boosting the scalability, flexibility, cost-effectiveness, and maintainability of solutions. Incorporating, analysing, and visualising data is simplified with the use of the framework's predetermined cloud-based service infrastructure. The first layer is the Data layer, Data including insect spreading data, weather data, water sample data, linked data, and document files are stored in the second layer. In order to provide effective data management and retrieval, it also comprises services like IoT, data extraction, and unique data extraction and integration service.Several use cases have been used to demonstrate the framework and assess its efficacy. This study focuses on two such use cases, each of which is based on a different set of services from the proposed architecture and reflects a different set of data integration requirements. Take the case of an India programme designed to monitor the spread of RPW as an illustration. This example shows how an ETL process may use a GIS

service for data integration and how online data can be utilised with a decision support system to help with pesticide consumption. Additionally, sensor data analysis and plot-specific watering recommendations may be done using ML models.

Some cloud-based frameworks [16[The proposed framework] attempts to give a thorough description of the functional aspects required for data integration and analysis, while as much as possible abstracting particular technologies. This practical information comes from both the ground up and previous books. (through actual experiences of applying the framework to real-world use cases) and the top down (via a comprehensive literature analysis, interviews, and surveys). We believe the framework's features will be adequate for developing the vast majority of cloud-based agriculture apps.

Our efforts double the original worth. To begin, more scenarios needing multi-source data integration can benefit from the proposed architecture and its implementation, as indicated by the two use examples. The framework encourages farmers to work together and makes it possible to integrate data across agricultural systems. Secondly, our research adds to the body of knowledge on how cloud computing can be used to unify different types of agricultural data. The proposed approach utilises cloud computing for the integration of data. When it comes to cloud computing and storage, customers may "pay as they go," eliminating the need to fork over a large sum of money up front for costly gear and software. Since cloud-based solutions allow for the nearly immediate addition of resources like processor speed, data storage space, and network throughput, they are more adaptive and expandable as well. Cloud computing offers advantages beyond the pest control and sensor-based irrigation sectors by enabling customers to outsource IT infrastructure maintenance to other vendors. It may incorporate information from agricultural information systems, including plant disease monitoring, satellite crop yields, and weather predictions, improving overall effectiveness and efficiency., it is vital to have a system that can scale to accommodate growing data volumes while being cost-effective and easy to use. Keep in mind that farmers can already be using agricultural cloud services (SaaS) that facilitate data collecting and decision making. These services often have their own application programming interfaces (APIs) that facilitate integration. By integrating and using current services, cloud computing may considerably help in the development of new agricultural data integration services, improving farmers' capacity to share data and facilitating supply chain cooperation. The ability to integrate data is a major factor in this. Several common farming situations might benefit economically from this ([17], [18]). The design of cloud-based agricultural services may provide network effects that raise

the platform's value for all users. The framework will generate new agricultural services, data products, and information as more people join and utilise it. The rise of cloud-based agricultural services is anticipated to increase their use and advantages for farmers and stakeholders. Positive feedback is anticipated to contribute to this growth.

References

[1] A. R. Hevner, S. T. March, J. Park, and S. Ram, ``Design science in information systems research,'' MIS Quart., vol. 28, no. 1, pp. 75, 2004.

[2] I. Izhaki, L. Fink, and O. Raphaeli, ``Identifying the factors that affect the adoption of cloud computing technology among farmers,'' M.S. thesis, Dept. Ind. Eng., Ben Gurion Univ. Negev, Beersheba, Israel, 2018. J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, no. 1, pp. 34–39, Jan. 1959, DOI: 10.1109/TED.2016.2628402.

[3] A. Kaloxylos, A. Groumas, V. Sarris, L. Katsikas, P. Magdalinos, E. Antoniou, Z. Politopoulou, S. Wolfert, C. Brewster, R. Eigenmann, and C. M. Terol, ``A cloud-based farm management system: Architecture and implementation,'' Comput. Electron. Agricult., vol. 100, pp. 168179, Jan. 2014. E. H. Miller, "A note on reflector arrays," *IEEE Trans. Antennas Propagat.*, to be published.

[4] A. Bahga and V. Madisetti, Cloud Computing: A Hands-On Approach. Scotts Valley, CA, USA: CreateSpace Independent Publishing Platform, 2013. J.H. Davis and J. R. Cogdell, "Calibration program for the 16-foot antenna," Elect. Eng. Res. Lab., Univ. Texas, Austin, TX, USA, Tech. Memo. NGL-006-69-3, Nov. 15, 1987.

[5] P. M. Mell and T. Grance, The NIST Denition of Cloud Computing. Gaithersburg, MD, USA: National Institute of Standards and Technology, 2011.

[6] AGROVOC. AGROVOC Thesaurus. Accessed: Jul. 22, 2022. G. O. Young, "Synthetic structure of industrial plastics," in Plastics, vol. 3, Polymers of Hexadromicon, J. Peters, Ed., 2nd ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15-64.

[7] M. S. Goraya and H. Kaur, ``Cloud computing in agriculture,'' Int.J. Technol. Innov. Res., vol. 16, 2015. The Terahertz Wave eBook. ZOmega Terahertz Corp., 2014.

[8] N. F. Noy and D. L. McGuinness. (2001). Ontology Development 101: A Guide to Creating Your First Ontology, Knowledge Systems Laboratory. Accessed: Jun. 3, 2021.

[9] W3C. The Semantic Web Made Easy. Accessed: Feb. 1, 2017.

[10] Ontotext. What is an RDF Triplestore? Accessed: Jul. 18, 2022.

[11] ArcGIS. (Oct. 9, 2021). Powerful Mapping and Analytics Software.Accessed: Jul. 22, 2022.

[12] AWS. Working With a DB Instance in a VPC. Accessed: Jul. 22, 2022.

[13] (2022). AWS. What is AWS Lambda? J.O. Williams, "Narrow-band analyzer," Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, USA, 1993.

[14] [14] AWS. What is Amazon Aurora? Accessed: Jul. 17, 22.

[15] [15] A. Goldstein, L. Fink, O. Raphaeli, A. Hetzroni, and G. Ravid, ``Addressing the `tower of babel' of pesticide regulations: An ontology for supporting pest-control decisions,'' J.Agricult. Sci., vol. 157, no. 6, pp. 493-503, Aug. 2019.

[16] [16] M. S. Farooq, S. Riaz, A. Abid, K. Abid, and M. A. Naeem, ``A surveyon the role of IoT in agriculture for the implementation ofsmart farming,'' IEEE Access, vol. 7, pp. 156237156271, 2019.

[17] [17] S. Zhang, Y. Mu, and Y. Sun, ``Early impact of COVID-19 pandemic on the molluscan shellsh supply chain in China,''Ocean Coastal Manage., vol. 213, Nov. 2021.

[18] [18] V. N. Erasmus, T. Kadhila, N. N. Gabriel, K. L. Thyberg, S. Ilungu, andT. Machado, ``Assessment and quantication of Namibian seafoodwaste production,'' Ocean Coastal Manage., vol. 199, Jan. 2021.

Stochastic Modelling and Computational Sciences

REVOLUTIONIZING AGRICULTURE THROUGH MOBILE CROWD SENSING: ENHANCING DATA COLLECTION FOR SMART FARMING

Abhilash Sam Paulstin K.C[1] and **Dr. G. Angeline Prasanna[2]**
[1]Research Scholar in Computer Science, AJK College of Arts and Science
[2]Former Associate Professor & Head Department of Computer Science, AJK College of Arts and Science
[1]abhi.sam83@gmail.com and [2]drgangelineprasanna@gmail.com

ABSTRACT

In the realm of smart agriculture, the pursuit of efficiency and intelligence in physical farm management has paved the way for transformative advancements. While this paradigm shows immense promise, its progress has been hampered by the constraints of conventional data collection methods. In response, a shift towards innovative solutions is imperative to overcome the existing challenges. One such solution, Mobile Crowd Sensing (MCS), offers a trifecta of advantageous attributes: cost-effectiveness, scalability, and robust mobility.

As the Internet of Things (IoT) matures into a tangible reality, smartphones are permeating even the remotest corners, contributing to the widespread availability of necessary technology infrastructure. This convergence of MCS attributes and accessible plug-and-play resources has catalysed a new era in smart agriculture. This synthesis extends a multitude of novel avenues at the application level, effectively redefining agricultural practices.

This paper undertakes a comprehensive evaluation of Agriculture Mobile Crowd Sensing (AMCS), casting a discerning light on data collection paradigms specific to agriculture. Through meticulous scrutiny, we present a juxtaposition with existing agricultural data collection solutions, ultimately demonstrating that AMCS offers pronounced advantages. These advantages include heightened flexibility, implicit data aggregation, and cost-effectiveness. Nonetheless, it's essential to acknowledge that while AMCS holds great potential, there exist considerations regarding data integrity and quality that must be addressed in future endeavours.

With an in-depth analysis of the challenges and prospects that characterize MCS-enabled agriculture, we present six prospective applications of AMCS in this domain. Each application marks a distinctive pathway towards the fusion of technology and agriculture, culminating in a more sustainable and effective landscape. In conclusion, this study advocates for an exploration of the evolving agricultural landscape, with a strategic focus on key characteristics such as seasonality and regionality. Through such research, a future enriched by data-driven cultivation practices awaits, providing a blueprint for redefining agriculture's future trajectory.

Keywords: Mobile Crowd Sensing, Data Collection, Smart Agriculture, Internet of Things

1. INTRODUCTION

In the domain of smart agriculture, the convergence of Big Data technology and mathematical models holds the potential to revolutionize agricultural production. By adeptly analyzing substantial volumes of data, this technology offers farmers invaluable insights without necessitating the intervention of dedicated specialists. This approach proves particularly promising in addressing persistent challenges, such as the predicament of unmarketable agricultural products due to a lack of seamless information exchange between farmers and consumers [1].

Within the realm of agricultural data acquisition, two predominant methods have emerged: a) Site survey conducted by specialized professionals [2], and b) Sensing technology hinged on the Space-Air-Ground Integrated Network (SAGIN) [3]. However, both these approaches, despite their merits, prove to be inadequately feasible and scalable within the contemporary and future agricultural landscape. For instance, the site survey method, albeit valuable, falls short of compatibility with the tenets of Big Data technology. It's labour-intensive, time-consuming, and primarily samples localized data, rendering it unsuitable for large-scale insights. On the

Stochastic Modelling and Computational Sciences

other hand, while the SAGIN approach exhibits potential, it grapples with significant limitations. The deployment and operational costs are prohibitively high, and it lacks the flexibility and scalability crucial for addressing the evolving needs of modern agriculture.

In response to these challenges, this study introduces a platform that evaluates and proposes novel data collection methodologies to surmount the aforementioned limitations. At the heart of this endeavour lies the concept of Mobile Crowd Sensing (MCS), a technique hinging on the participation of a multitude of individuals equipped with mobile devices such as smartphones and wearables. These devices are imbued with the capability to sense and share pertinent information, collectively engaging in the execution of extensive and intricate sensing tasks. MCS embodies three pivotal attributes: cost-effectiveness, scalability, and mobility [4]-[7].

Recent years have witnessed a surge in research interest surrounding existing MCS systems, wherein various approaches are explored to enhance the accessibility of MCS technology for diverse application scenarios. In this context, this work endeavours to contribute by evaluating and proposing innovative data collection methodologies that transcend the limitations of current approaches. Through a fusion of technological prowess and agricultural insights, this study envisions a harmonious future where the barriers between data, technology, and agricultural progress are transcended, and smart farming realizes its full transformative potential.

The realm of Mobile Crowd Sensing (MCS) has found diverse and impactful applications across multiple sectors, as categorized below:

Environmental Monitoring: MCS is harnessed to monitor environmental data critical for sustainable urban development. Solutions like CrowdRecruiter [8] optimize participant selection to minimize incentive payments while ensuring probabilistic coverage. GRC-Sensing [10] tracks noise pollution, while Urban Safety [11] gathers information about damaged urban infrastructure. SenSquare [24] evolved into a versatile system that handles heterogeneous data and offers visual programming plugins [26].

Disaster Prediction: MCS aids in predicting large magnitude earthquakes to mitigate potential devastation. MyShake [15][16] employs smartphone accelerometers for earthquake early-warning systems.

Social Networking: MCS data aids in enhancing social network services. TrackMaison [17][18] monitors smartphone users' social behavior through data usage, location, usage frequency, and session duration.

Living Service: MCS enhances citizens' quality of life by providing real-time insights. CrowdQTE [9] leverages sensor-enhanced mobile devices to offer queue time information. Mobibee [19] enables indoor localization by incorporating user-contributed data. WasteApp [22] aids in recycling waste, and MCNet [23] measures wireless performance.

Urban Management: MCS improves urban management, including traffic control. The CREAM system [12] facilitates timely traffic management responses, and SafeStreet [21] detects road anomalies for safer driving.

Health Care: MCS contributes to healthcare with solutions such as Track-YourTinnitus [13][14], revealing insights on tinnitus treatment via data analysis and visualization. CovidSens [28] monitors COVID-19 propagation through GPS, microphone, and camera.

Other: The diversification of crowdsensing applications led to the creation of an operating system, CrowdOS [27], addressing challenges in maximizing sensing resource utility.

Through comprehensive evaluation, it becomes evident that MCS has permeated various domains, resulting in research advancements focusing on participant selection, task allocation, incentive strategies, data mining, visualization, and privacy protection. This involvement predominantly encompasses the smart city domain, wherein citizens play a pivotal role in sensing tasks. Notably, the agricultural sector has not fully embraced MCS technology, and farmers are yet to participate in MCS campaigns. However, with smartphone ownership on the

Stochastic Modelling and Computational Sciences

rise even in underdeveloped countries, farmers equipped with smart devices possess the potential to contribute valuable agricultural data, owing to their participation throughout the agricultural production process.

Motivated by these insights, there is a compelling aspiration to bridge the gap in agricultural data collection systems, particularly addressing the shortcomings of the Space-Air-Ground Integrated Network (SAGIN). Given the prevalence of smartphones among farmers and their active involvement in agricultural endeavours, the integration of MCS technology holds promise for filling the void left by existing systems. The aim is to harness MCS to amplify agricultural data collection, embracing the dynamic landscape of smart agriculture and facilitating the seamless exchange of valuable insights.

2. Data collection in smart agriculture
The diverse landscape of agricultural industries encompasses various factors of production, including farming, stockbreeding, forestry, aquaculture, and sideline activities. While each industry follows a distinct production process, they collectively generate a plethora of data. In this section, we delve into the context of farming to explore the manifold data types and dissect the existing data collection systems.

2.1. Data Varieties in Farming
Agricultural production forms a comprehensive chain, yielding copious data types that underpin informed decision-making. This process unfurls across three pivotal phases: Pre-Production, In-Production, and Post-Production. These phases encompass an array of production links, such as production preparation and farmland management, forming the bedrock of agricultural data diversity.

Pre-Production Phase: The inherent time constraints in this phase impel swift decisions, including crop selection and planting plans. These choices significantly hinge on the feedback gleaned from previous year's market sales, wielding substantial influence on the ultimate harvest outcomes.

In-Production Phase: With dynamic weather changes dictating the course, farmers wield Agricultural Technology to dynamically manage crops in response to shifts in rainfall, temperature, and humidity. Strategies like irrigation and pesticide application are meticulously employed to optimize crop yield.

Post-Production Phase: The culmination of the agricultural process sees harvested crops navigating through transportation, storage, and sales, culminating in consumption or disposal, especially when shelf life is exceeded.

2.2. Existing Data Collection Methods
Space-Air-Ground Integrated Network (SAGIN): The advancement of sensing technologies, encompassing remote sensing and wireless sensing, has ushered in the era of the Space-Air-Ground Integrated Network (SAGIN). This comprehensive system significantly enriches agricultural information acquisition, enhancing the accuracy of farmland management through precise growth data. SAGIN comprises three integral components:

Space: Remote Sensing Satellites (RSSs) capture agricultural environmental data through imagery. This involves the adoption of 3S technology, comprising Remote Sensing (RS), Geography Information Systems (GIS), and the Global Positioning System (GPS).

Air: Unmanned Aerial Vehicles (UAVs), including specialized agricultural aircraft, play a pivotal role in gathering specific information from target areas. These vehicles are equipped with sensors like hyperspectral cameras to yield detailed insights.

Ground: The realm of Wireless Sensor Networks (WSNs) encompasses deployed wireless sensors utilizing diverse protocols like Zigbee, Bluetooth, and Lora. These networks collect on-ground data, encompassing temperature and humidity among other parameters.

Crowdsourcing (CS): In regions constrained by resources, the power of community-driven participation comes to the fore through Crowdsourcing (CS). This approach leverages extensive user engagement for data collection. In settings with limited surveillance capacity, as seen in developing countries, CS provides real-time surveillance

Stochastic Modelling and Computational Sciences

data on disease and pest incidence. A prime example is a study that employed CS to furnish critical insights into viral disease and pest prevalence. This data underpins the eventual development of an automated diagnostic tool for cassava diseases, alongside a real-time disease map.
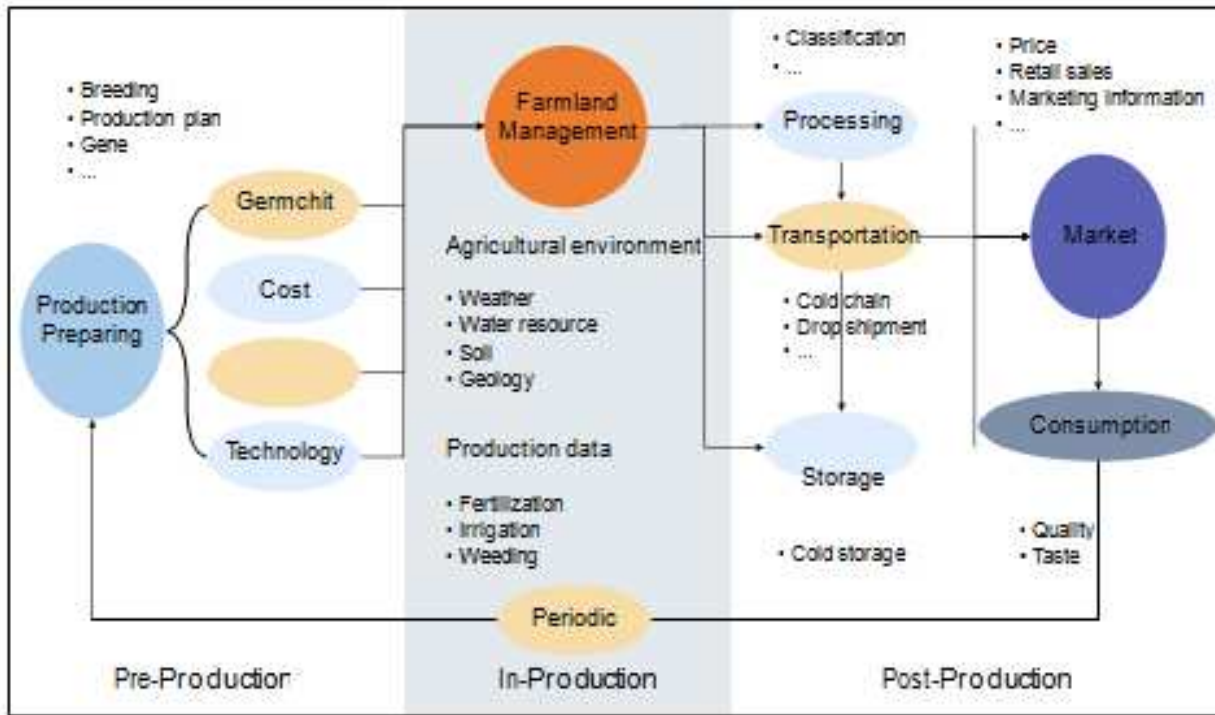
Figure 1: Types of data used in agricultural production

2.3. Comparison
An all-encompassing comparison of data collection methods unfolds through a prism of seven pivotal factors.

Comparing AMCS with Existing Methods: A comprehensive assessment involving AMCS (Agriculture Mobile Crowd Sensing), alongside other prevalent methods, encompasses crucial considerations:

Data Granularity: The resolution of sensing equipment profoundly influences data granularity. WSNs, CSs, and AMCSs excel in obtaining fine-grain data within close ranges. In contrast, RSSs and UAVs exhibit comparatively lower resolution, attributed to remote sensing technology.

Flexibility: The realm of flexibility, encompassing mobility and expansibility, is illuminated. While RSSs offer global coverage, their deployment remains inflexible. UAVs exhibit dynamic potential through adaptable sensor configurations. WSNs, though expandable, lack mobility. In contrast, CSs and AMCSs boast flexibility enabled by smart device interfaces, and participant mobility facilitates data acquisition in specific areas.

Data Integrity: RSSs, UAVs, and WSNs ensure comprehensive data collection through their respective technologies. In the context of CS and AMCS campaigns, incomplete tasks can lead to partial data gaps.

Data Quality: Professional equipment underpins the reliability of data collected by RSSs, UAVs, and WSNs. Data collected through CS and AMCS campaigns may be subject to user-driven variations, challenging data quality assurances.

Implicit Data Collection: CS and AMCS both facilitate unstructured data collection, notably in the pre-production and post-production phases. This bridges the physical and digital realms, facilitating the seamless sharing of data across the entire agricultural industry chain.

Cost: AMCS stands as a cost-effective approach, leveraging equipped smart devices to complete sensing tasks without the burden of deployment and maintenance costs. It is imperative to note that CS campaigns may necessitate the provision of extra equipment for users, introducing additional expenses compared to AMCS.
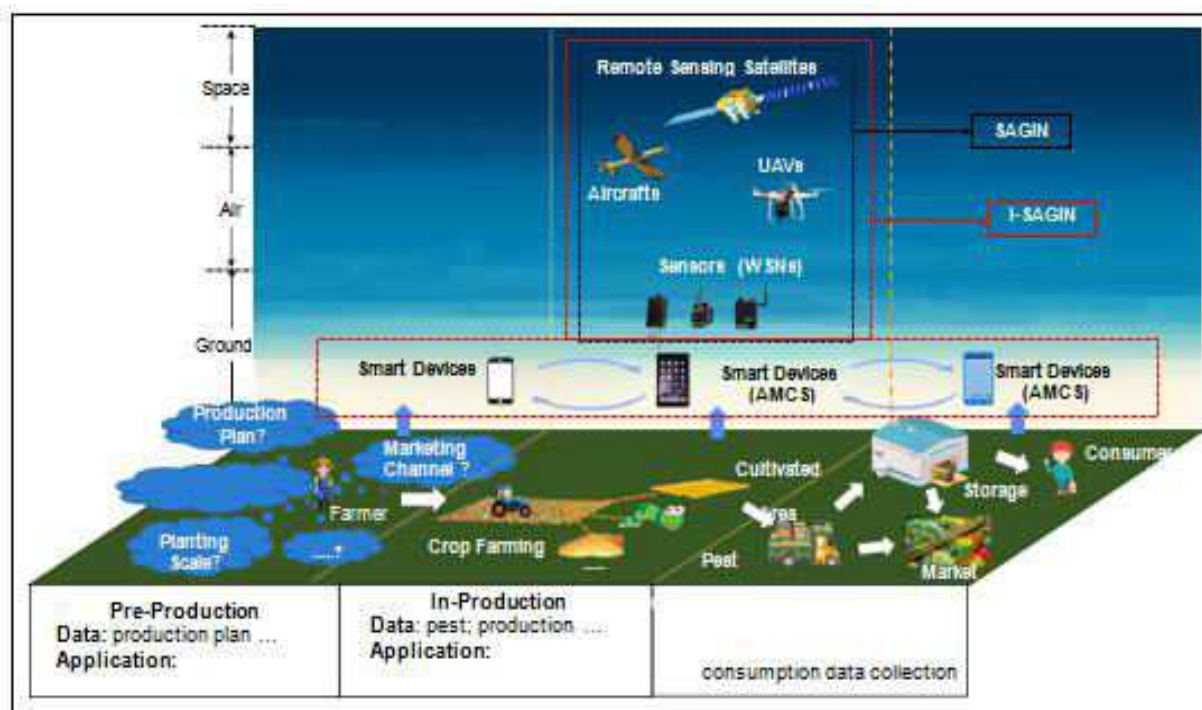

Figure 2: Incorporating AMCS into SAGIN to enhance the effectiveness of data gathering in agricultural production.

In juxtaposition with RSSs, UAVs, and WSNs, AMCS shines brightly, exemplifying enhanced flexibility, implicit data acquisition, and cost-effectiveness. However, it remains vital to acknowledge AMCS's limitations in terms of data integrity and quality. Hence, AMCS emerges as a crucial tool for acquiring agricultural data, augmenting the efficiency of data collection across varied agricultural applications. Rather than entirely replacing the SAGIN system, AMCS's strengths can complement the existing structure. Notably, comparing AMCS with CS highlights two significant advantages: participant selection based on location and the convenience of participation through users' mobile devices. In contrast, CS campaigns often require artificial offline recruitment and equipment provision, presenting a stark contrast to AMCS's dynamic and user-centric approach. Moreover, employing farmers' smartphones for data collection promises further cost reduction.

The comparison presented in this section underscores the importance of AMCS in revolutionizing the data collection landscape within the agricultural domain. While each method brings unique advantages, the flexibility, cost-effectiveness, and implicit data gathering prowess of AMCS make it a pivotal tool for enhancing agricultural efficiency and decision-making.

3. Enabling the Synthesis of AMCS with Agriculture
While the benefits of AMCS over existing data collection systems have been illuminated, it's imperative to delve into the practical integration of AMCS within the agricultural realm. This section meticulously examines the pivotal facets of combining AMCS with agriculture, scrutinizing the fourfold factors delineated in Figure 3.

Stochastic Modelling and Computational Sciences
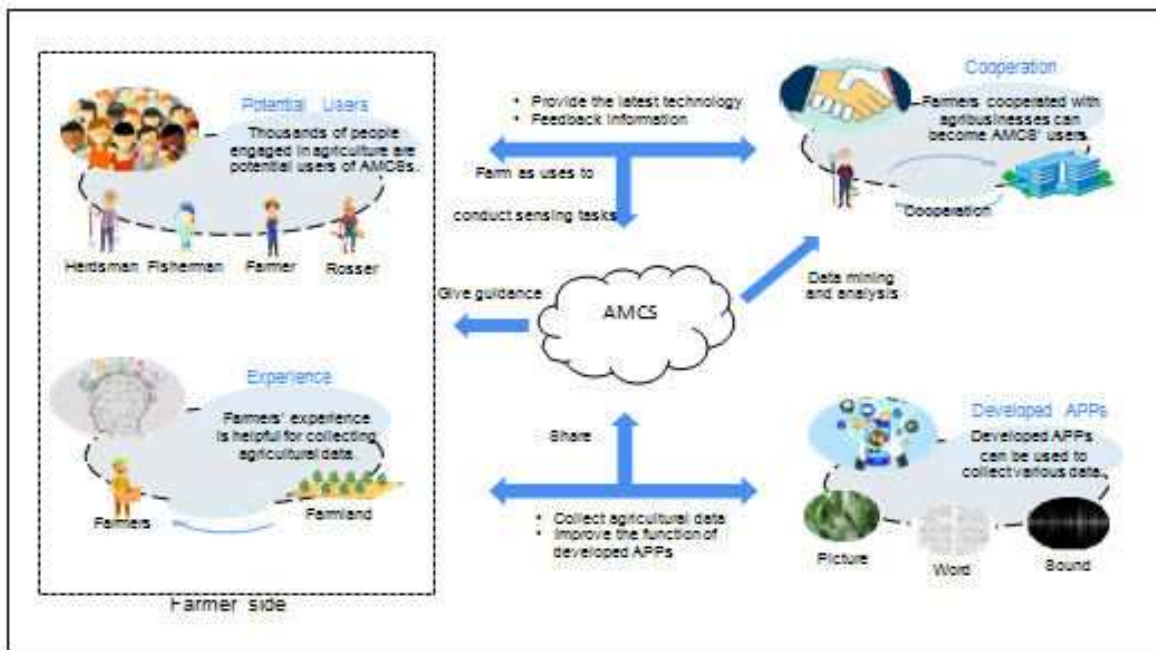
Figure 3: The pivotal element of integrating AMCS into agricultural operations

3.1. Abundance of Potential Users

Recent data from The World Bank in 2019 underscores the substantial employment within the agriculture sector, particularly pronounced in developing countries where the proportion of agricultural employment often exceeds 25% of the total workforce. Remarkably, even in highly developed agricultural nations like the USA, over 3 million individuals are engaged in agriculture, many of whom possess smartphones. Consequently, this vast workforce emerges as a prime pool of potential users for AMCS to engage in the collection of agricultural data.

3.2. Flourishing Agriculture-Related APPs

A proliferation of agriculture-related applications, predominantly based on Android and iOS operating systems, has pervaded agricultural production. These applications facilitate the utilization of smartphones for data collection across various facets of agriculture. Illustratively, a study engineered an application to analyze brightness through smartphone cameras, while another harnessed mobile phones as soil colour sensors. The user-friendly nature and practicality of these applications enhance their adoption, generating substantial datasets. Hence, the presence of well-developed applications forms an indispensable prerequisite for the effective deployment of AMCS.

2.3. Harnessing Farmer Expertise

Farmers, steeped in the intricacies of agriculture, possess a wealth of professional knowledge that lends itself to precise agricultural data collection. Their profound understanding of agricultural nuances empowers them to contribute accurate and meaningful data. For instance, when capturing images of novel plant diseases, farmers can enrich data sets by adding descriptions drawn from their experiential insights. In this symbiotic exchange, the synergy of AMCS and farmer expertise fosters a deep integration of technology into agricultural practices.

2.4. Synergistic Agribusiness-Farmer Collaboration

Agribusinesses, acting as primary custodians of agricultural data, frequently establish symbiotic relationships with farmers to propel new product and technology adoption while bolstering farmer incomes. These collaborations serve as a conduit to convert farmers into AMCS users. Building on existing partnerships, farmers exhibit greater enthusiasm in sharing valuable data, underpinned by the assurance of mutual benefits. In turn, agribusinesses glean reliable insights into their products and technologies, establishing a virtuous feedback loop. This

cooperative interplay between agribusinesses and farmers serves as a potent catalyst in facilitating the application of AMCS within agriculture.

The synthesis of AMCS within the agricultural domain hinges on strategic factors. By tapping into the extensive workforce of potential users, capitalizing on the proliferation of agriculture-related applications, harnessing farmer expertise, and fostering collaborative ties between agribusinesses and farmers, AMCS gains a robust foothold within agriculture. This symbiotic fusion not only revolutionizes data collection but also catalyses the evolution of smart agriculture into an even more dynamic and efficient sphere.

4. Unleashing the potential of AMCS: application prospects

In light of the current landscape of agricultural production, we envision a spectrum of six compelling applications propelled by AMCS, as illustrated in Figure 4.
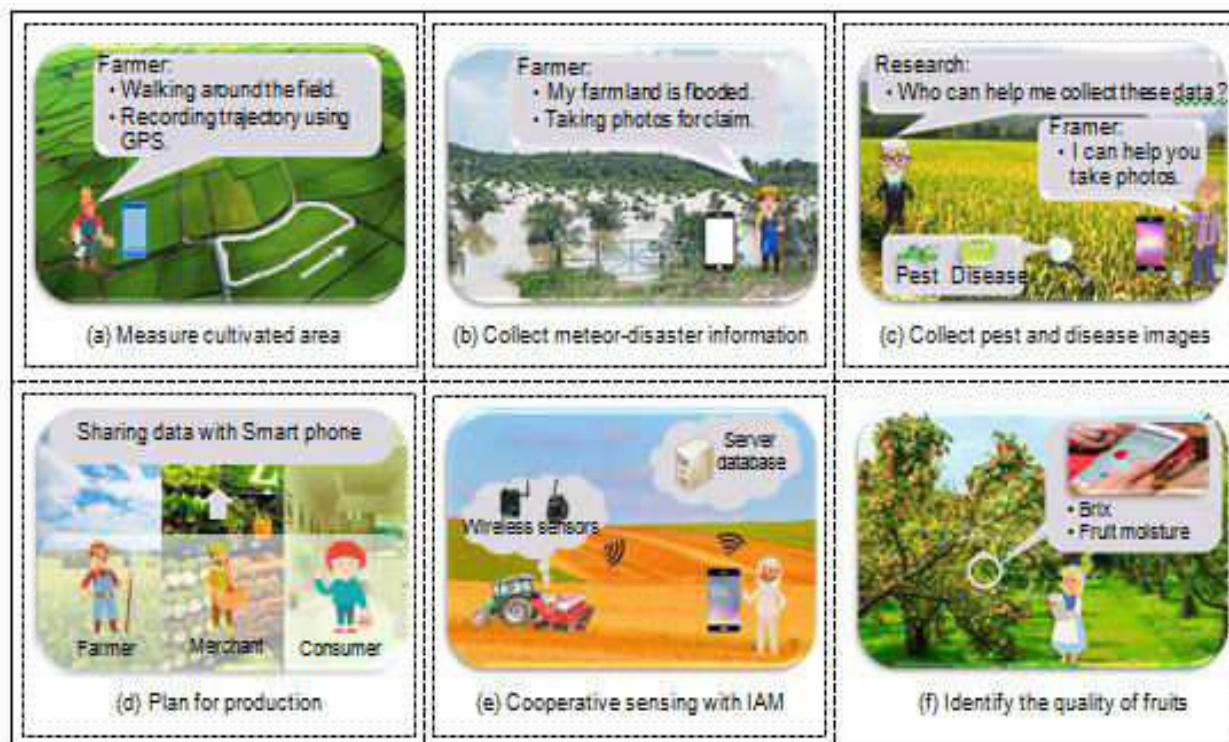
Figure 4: Six possible uses of the AMCS

4.1. Precision Cultivated Area Measurement

A paramount concern for agricultural policy formulation is the accurate prediction of cultivated area expansion or contraction. While 3S technology enables rough measurements of large-scale farmland, it falters in small-scale scenarios due to error intolerance and exorbitant costs. Enter AMCS-equipped farmers, who, traversing the perimeters of cultivated land, can perform fine-grained measurements. Subsequent calibration ensures precision, thereby enhancing the determination of planting area, crop yield, and variety, yielding more accurate insights into agricultural landscapes.

4.2. Empowering Meteorological Disaster Insights

Agricultural meteorological catastrophes, encompassing floods, frosts, snowstorms, and hail, exact severe tolls on crop productivity, resulting in substantial economic losses. During these crises, comprehensive emergency information assumes paramount importance for government intervention and agricultural insurance compensation. Farmers, functioning as eyewitnesses, offer detailed disaster insights, augmenting the quality of data garnered by Remote Sensing Satellites (RSSs) and fostering more informed disaster management decisions.

4.3. Crowdsourced Pest and Disease Surveillance

Pests and diseases incessantly plague crop production, incurring substantial losses. Traditionally, researchers invest significant resources in field visits to procure images of these afflictions. Harnessing the expertise of farmers in identifying these issues can revolutionize data collection. Furthermore, existing Android-based apps, such as, designed to recognize pests and diseases, could synergize with MCS technology, optimizing the collection of requisite images. This fusion not only expands the dataset but also enables the study of pest migration patterns through recorded photographs.

4.4. Informed Production Planning

A glaring issue in agriculture is the generation of unmarketable produce, straining resources and finances. Farmers, often driven by the anticipation of high demand, produce surplus, resulting in wasted resources. Similarly, disjointed information flow among farmers, merchants, and consumers hampers the supply-demand equilibrium. Herein lies the significance of shared data during production. Collaborative data exchange between stakeholders could mitigate the problem of unmarketable agricultural goods, streamlining production for optimal resource utilization.

4.5. Intelligent Agricultural Machinery (IAM) Synergy

In contrast to traditional farming machinery, Intelligent Agricultural Machinery (IAM) integrates wireless sensors for precision operations. This paradigm shift allows the collection of vital data—fertilization and seeding quantities, for instance—enabling farmers to assess IAM performance. Capitalizing on smartphone computational prowess, farmers can visualize received data from wireless sensors, while archiving it for historical analysis and storage. This convergence of IAM and AMCS ushers in a new era of cooperative sensing.

4.6. Fruit Quality Authentication

The evolution of sensing technology has endowed smartphones with formidable capabilities. Notably, smartphones with house miniature molecular spectroscopy sensors that capture spectral data, facilitating assessments of fruit attributes like sweetness and moisture. Capitalizing on these parameters, researchers can delve deeper into fruit quality evaluation, enhancing both the understanding of fruit parameters and early growth monitoring. Farmers, empowered with timely insights, can dynamically adapt management strategies for optimal yield enhancement.

AMCS's potential applications are both diverse and transformative. Through precision measurement, disaster insights, pest surveillance, production optimization, IAM collaboration, and fruit quality authentication, AMCS seamlessly integrates into diverse facets of agriculture, amplifying productivity and efficiency while elevating agricultural practices to new heights.

5. Unveiling unresolved AMCS research agendas in agriculture

Intricately interweaving practical applications with the unique nuances of agriculture and rural settings unveils a range of specific research conundrums within the realm of AMCS.

Stochastic Modelling and Computational Sciences
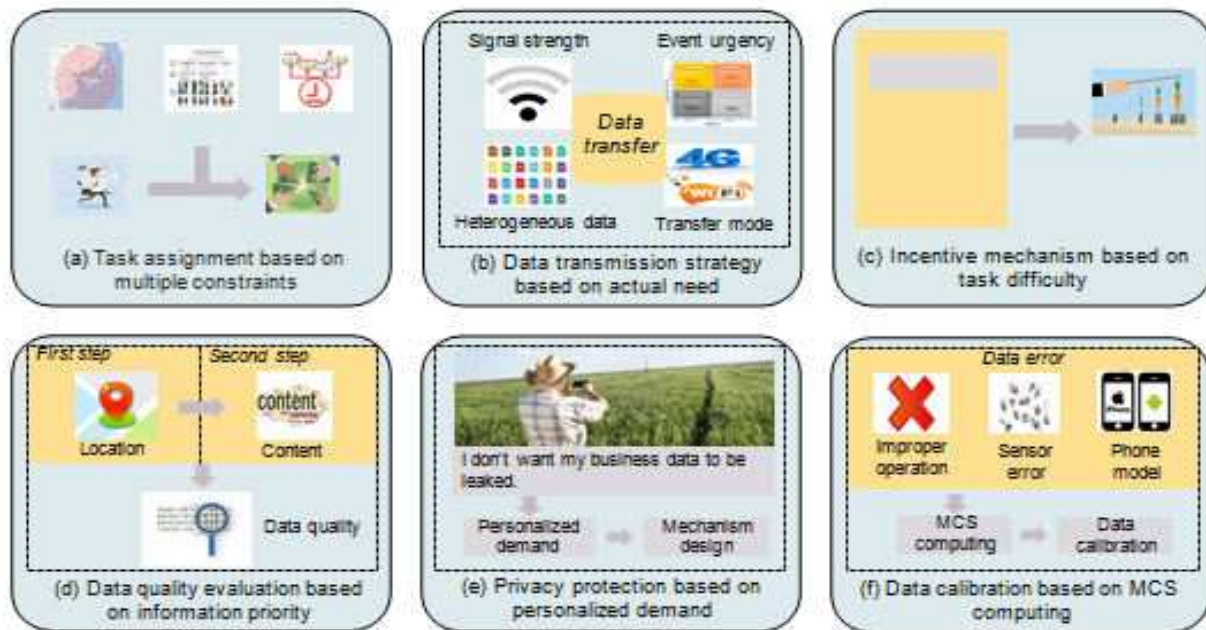
Figure 5: Unresolved research challenges

This section encapsulates the agrarian fabric's quintessence before expounding on research connotations, as depicted in Figure 5.

5.1. Unveiling Agricultural Traits

The distinctive features characterizing agriculture underpin the contrasts between AMCS and other MCS paradigms, notably:

Seasonality and Regionality: Agriculture dances to nature's tune, marked by the ever-changing environment—summer's rainfall, winter's snowfall. The geographical disparities further paint unique patterns—wheat in China's north, rice in its southern expanse.

Diverse Species and Multifaceted Data: The array of species heralds a multidimensional data landscape, sparking diverse application contexts. This heterogeneity poses challenges in data transmission, visualization, and amalgamation.

Uneven Base Station Deployment: Sparse base station distribution in rural terrains stems from low population densities, culminating in irregular network coverage and unstable connections. This scarcity negatively reverberates onto GPS precision.

Villager Residences: Farmers' domicile in rural hamlets engenders intermittently unreachable areas during off-farming periods. This temporal-spatial skew imperils uniform sensing task completion, hindering data sufficiency.

5.2. Pervasive Research Pursuits

Task Allocation in Multiconstraint Scenarios: Task quality constitutes a linchpin in MCS, embracing coverage-cost equilibrium and duration management. In the agricultural spectrum, fresh constraints emerge—monitoring migratory pest trajectories necessitates extensive coverage and data finesse, constrained by seasonal pest migration and multi-objective task alignment.

Data Transmission Optimization: The convergence of species diversity and irregular 4G base stations invokes factors influencing data transmission strategy: signal strength variations (average, weak, or null), urgency-

Stochastic Modelling and Computational Sciences

mandated real-time or delay-tolerant transmission, optimal transmission mode selection (4G, 5G, Wi-Fi), and accommodation of heterogeneous data types (text, sound, image).

Incentivizing Mechanisms in Accordance with Task Complexity: Task intricacy manifests in remote, inaccessible locations, and data-intensive tasks. Incentives aligning with complexity—remote tasks and data-rich ventures—propel farmer involvement, while calibrated rewards encourage participation in demanding endeavours.

Data Quality Appraisal Weighing Information Priority: In the absence of benchmarks, prior research gauged MCS data quality based on context, disregarding GPS accuracy. Given agriculture's dependency on precise location data, GPS data quality evaluation becomes pivotal, considering factors like satellite coverage, base station distribution, and signal interference.

Personalized Privacy Measures in Sensitive Scenarios: In this transition from citizen-centric MCS to farmer-engaged AMCS, privacy safeguards assume distinct forms. The overlap of field monitoring and data collection jeopardizes privacy, potentially deterring farmer participation. Striking a balance between data necessity and privacy preservation becomes pivotal.

Data Calibration via MCS-Driven Computing: Smartphones, while practical, introduce data aberrations due to non-standard user conduct, sensor accuracy, and device variability. Augmenting smartphone processing power with calibration algorithms holds promise to match traditional equipment's precision, ensuring data accuracy and reliability.

In Essence, inextricably linked to agrarian dynamics, AMCS unravels a tapestry of research puzzles. Seasonality, diversity, technology limitations, rural habitats—each facet paints a distinctive canvas. From task allocation and data transmission to privacy preservation and data calibration, these challenges call for innovative solutions in the context of modern agriculture.

6. Implementing Agriculture Mobile Crowd Sensing (AMCS)
Let's discuss how we could implement the concept of Agriculture Mobile Crowd Sensing (AMCS) using Python:

Create Sensor Class:
Define a Sensor class to represent individual sensors. Each sensor should have attributes like sensor_id and methods to collect data.

Create Farm Class:
Define a Farm class to represent farms. Farms can have multiple sensors. Implement methods to collect data from all sensors in a farm.

Create AMCS Class:
Create an AMCS class that manages multiple farms. Implement methods to collect data from all farms.

Data Collection Simulation:
Simulate data collection by creating instances of the AMCS class and using its methods to collect data. You can use random data generation for demonstration purposes.

Data Aggregation:
Implement data aggregation logic to aggregate data collected from different farms and sensors. This could involve processing, filtering, and combining data.

Communication and Storage:
Implement data communication and storage mechanisms. This could involve sending data to a server or a database for further processing and analysis.

Stochastic Modelling and Computational Sciences

User Interface:

Developing a user interface that allows users (farmers or researchers) to interact with the system, view collected data, and manage tasks.

Error Handling and Quality Control:

Implement mechanisms to handle errors, ensure data quality, and address data integrity issues that might arise during data collection.

Testing and Refinement:

Test your implementation thoroughly, identify any issues or bugs, and refine the system based on feedback and testing results.

6.1. AMCS python script (AMCS.py):

```python
class Sensor:

def __init__(self, sensor_id, sensor_type):

self.sensor_id = sensor_id

self.sensor_type = sensor_type

class Farm:

def __init__(self, farm_id, location):

self.farm_id = farm_id

self.location = location

self.sensors = []

def add_sensor(self, sensor):

self.sensors.append(sensor)

class AMCS:

def __init__(self):

self.farms = []

def add_farm(self, farm):

self.farms.append(farm)

def collect_data(self):

# Simulate data collection process

data = []

for farm in self.farms:

for sensor in farm.sensors:

data.append(f"Data from Farm {farm.farm_id}, Sensor {sensor.sensor_id}: {sensor.sensor_type}")

return data

def add_sensor_to_farm(self, farm_id, sensor):

for farm in self.farms:
```
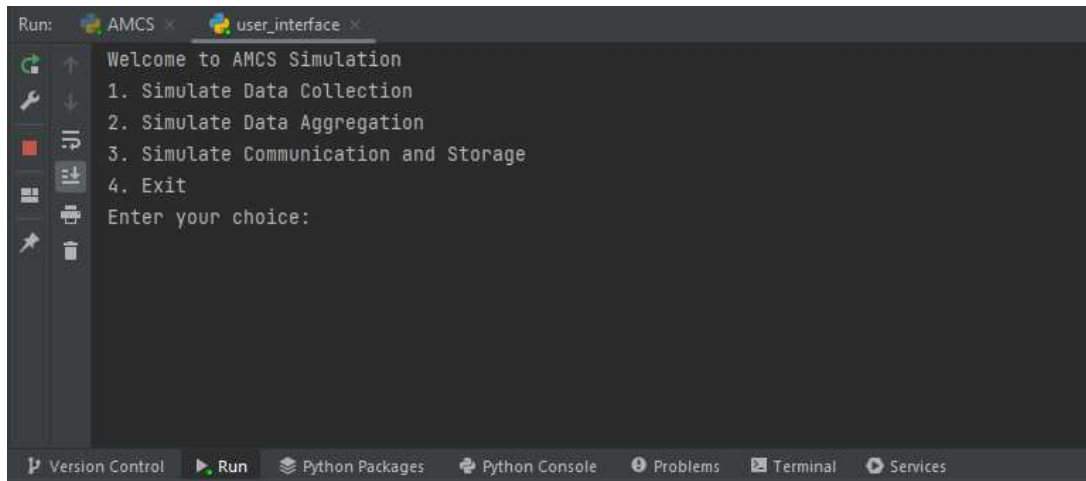
```python
if farm.farm_id == farm_id:
    farm.add_sensor(sensor)
    break
class UserInterface:
    def __init__(self, amcs):
        self.amcs = amcs
    def run(self):
        print("Welcome to the AMCS User Interface!")
        while True:
            choice = input("Select an option:\n1. Add Farm\n2. Collect Data\n3. Exit\n")
            if choice == "1":
                farm_id = input("Enter Farm ID: ")
                location = input("Enter Farm Location: ")
                farm = Farm(farm_id, location)
                self.amcs.add_farm(farm)
                print(f"Farm {farm_id} added successfully!")
            elif choice == "2":
                data = self.amcs.collect_data()
                print("Collected Data:")
                for item in data:
                    print(item)
            elif choice == "3":
                print("Exiting...")
                break
            else:
                print("Invalid choice. Please select a valid option.")
if __name__ == "__main__":
    amcs_system = AMCS()
    user_interface = UserInterface(amcs_system)
    user_interface.run()
```
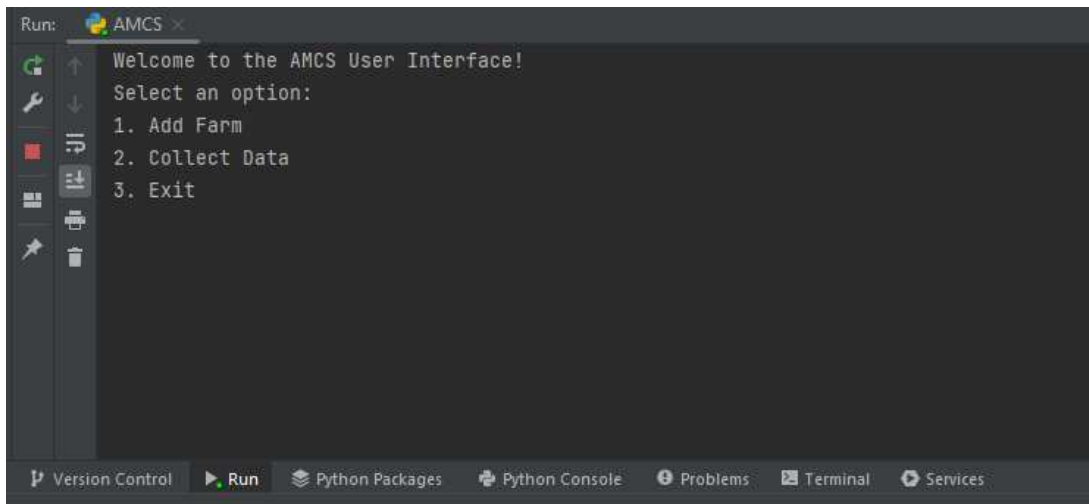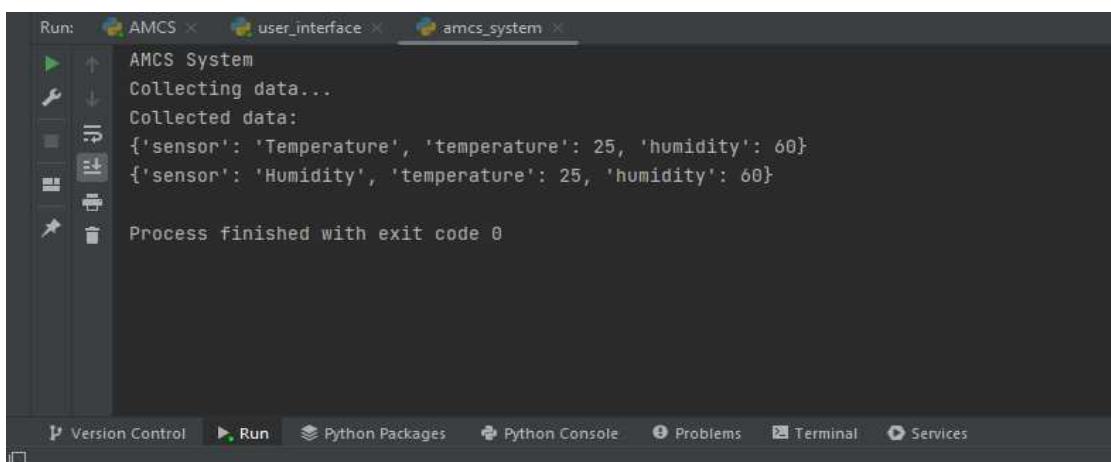
Stochastic Modelling and Computational Sciences

6.2 Module implementation

Figure 6: AMCS UI module

Figure 7: AMCS UI for data collection

Figure 8: AMCS data processing module
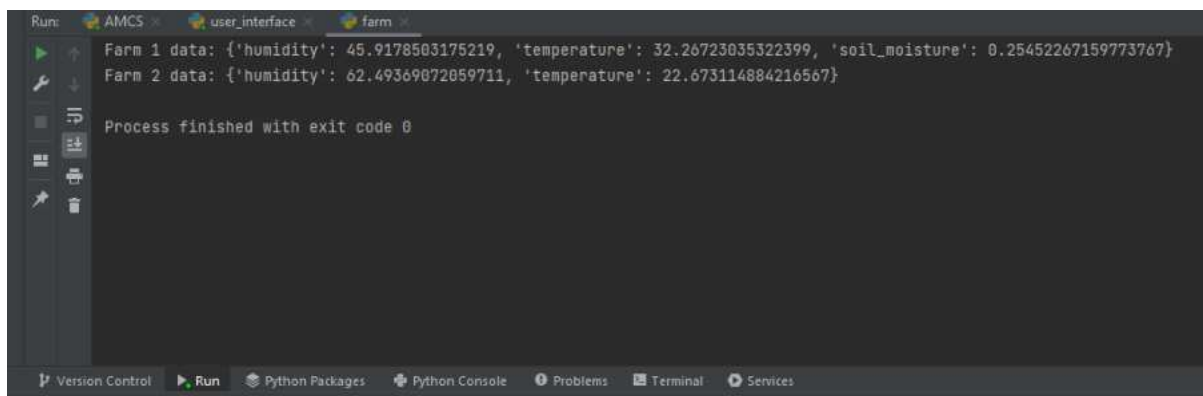
Stochastic Modelling and Computational Sciences
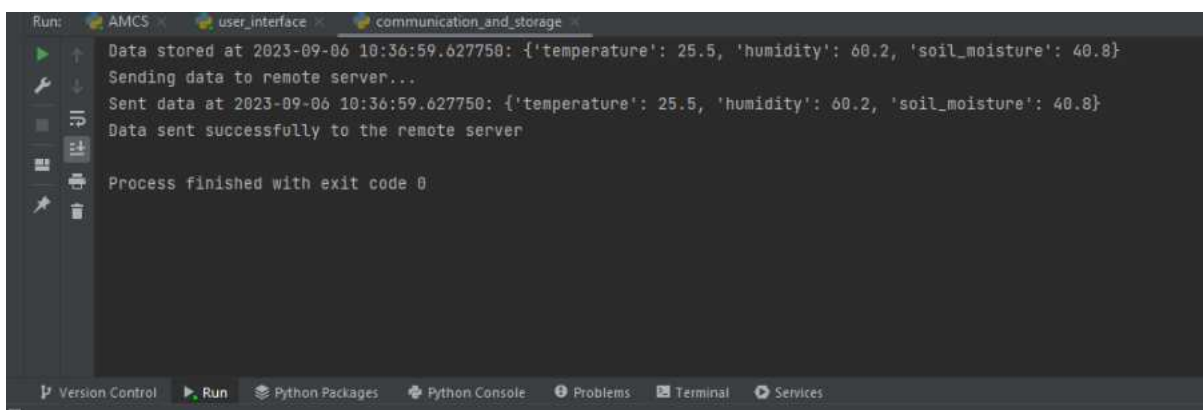
Figure 9: AMCS sensor data representation module

Figure 10: AMCS farm data module

Figure 11: AMCS communication and storage module for remote communication

7. Conclusion and insight: forging smarter agriculture with AMCS

In the landscape of smart agriculture, where intelligence converges with agriculture's core, the invention of advanced technologies such as Big Data and the Internet of Things augments agricultural production and begets a realm of research and applications—Agricultural Big Data being one such front. Yet, the domain grapples with pertinent limitations in data collection, spanning cost, scalability, data granularity, and flexibility. This paper, in

Stochastic Modelling and Computational Sciences

its resolute quest for solutions, illuminates the integration of Mobile Crowdsensing (MCS) within the smart agriculture domain—a novel trajectory ripe with promise.

The journey finds AMCS not confined to smartphones' traditional communication realm but metamorphosed into a potent agricultural management tool, catalysing production management and augmenting farmer livelihoods. The discourse expands and rejuvenates AMCS, bolstering its utility and broadening its horizons. Through the prisms of abundant potential users, the ubiquity of developed agriculture-centric apps, farmers' experiential insights, and the symbiotic rapport between agribusiness and farmers, AMCS's viability and impact gain validation.

In culmination, this expedition not only envisions the manifold potential applications of AMCS but, in identifying nascent research challenges, propels the endeavour towards an intellectually rich trajectory. The voyage engenders the following insights:

Farmers' Tech Eminence: The farmer's toolkit encompasses more than just implements—it resonates with innovation and connectivity. Smartphones metamorphose into tools for empowerment and informed decision-making.

AMCS's Amplified Concept: The paper casts AMCS in a broader light, infusing it with heightened relevance and effectiveness. Through its keen insights, AMCS stands poised as an indispensable agricultural tool.

Feasibility & Collaborative Synergy: By underscoring the confluence of potential users, app-driven accessibility, farmers' expertise, and synergistic industry collaboration, the paper firmly underscores AMCS's promise.

Embarking on Pathways of Innovation: This journey doesn't merely present solutions—it beckons to a future rife with innovative applications and burgeoning research frontiers. It beckons a concerted exploration into the domain's potential.

Thus, this endeavour breathes life into the vision of not just a smarter agriculture, but one imbued with collaboration, empowerment, and a profound technological tapestry. As this paper concludes, it ushers in an era where AMCS stands poised to harness collective intelligence, drive precision in agricultural practice, and usher in a new age of symbiotic growth—catalysing a harmony between human acumen and technological prowess.

REFERENCES:
[1]   S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big data in smart farming: a review," Agricultural Systems, Vol. 153, pp. 69-80, Feb. 2017.

[2]   J. Hammond, S. Fraval, J. Van Etten, J. G. Suchini, L. Mercado,T. Pagella, and V. D. albuena, "The rural household multi-indicator survey for rapid characterisation of households to inform climate smart agriculture interventions: description and applications in east africa and central america." Agricultural Systems, Vol. 151, pp. 225-233, 2017.

[3]   J. Liu, Y. Shi, Z. D. Fadlullah, and N. Kato, "Space-Air-Ground inte- grated network: a survey," IEEE Communications Surveys & Tutorials, Vol. 20, No. 4, pp. 2714-2741, May. 2018.

[4]   H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," IEEE Communications Magazine, Vol. 52, No. 8, pp. 29-35, Aug. 2014.

[5]   R K. Ganti, F. Ye, and H. Lei, "Mobile crowd sensing: current state and future challenges," IEEE Communications Magazine, Vol. 49, No. 11, pp. 32-39, Nov. 2011.

[6]   L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. M'hamed, "Sparse mobile crowdsensing: challenges and opportunities." IEEE Communications Magazine, Vol.54 No. 7, pp. 161-167, 2016.

Stochastic Modelling and Computational Sciences

[7] Z. Zhou, H. Liao, B. Gu, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "Robust mobile crowd sensing: when deep learning meets edge com- puting." IEEE Network, Vol. 32, No. 4, pp. 54-60, 2018.

[8] D. Zhang, H. Xiong, and L. Wang, "CrowdRecruiter: selecting partic- ipants for piggyback crowdsensing under probabilistic coverage con- straint," Proc. 2014 Acm International Joint Conference on Pervasive and Ubiquitous Computing, pp. 703-714.

[9] J. Wang, Y. Wang, and D. Zhang, "Real-time and generic queue time estimation based on mobile crowdsensing," Frontiers of Computer Science, Vol. 11, No. 1, pp. 49-60, 2017.

[10] W. Zamora, E. Vera, and C. Calafate, "GRC-Sensing: an architecture to measure acoustic pollution based on crowdsensing," Sensors, Vol. 18, No. 8, 2018.

[11] X Zhao, N Wanga, and R Han, "Urban infrastructure safety system based on mobile crowdsensing," International Journal of Disaster Risk Reduction, Vol. 27, pp. 427-438, 2018.

[12] X. Wang, Z. Ning, and X. Hu, "A city-wide real-time traffic management system: enabling crowdsensing in social internet of vehicles," IEEE Communications Magazine, Vol. 56, No. 9, pp. 19-25, Sept. 2018,

[13] Navatha S et al. 2023. Multitask Learning Architecture for Vehicle Over Speed as Traffic Violations Detection and Automated Safety Violation Fine Ticketing Using Convolution Neural Network and Yolo V4 Techniques. Chinese Journal of Computational Mechanics. 5 (Oct. 2023), 431–435.

[14] R. Pryss, W. Schlee, and B. Hoppenstedt, "Applying machine learning to daily-life data from the TrackYourTinnitus mobile health crowdsensing platform to predict the mobile operating system used with high accuracy: longitudinal observational study," Journal of Medical Internet Research, Vol. 22, No. 6, 2020.

[15] Q. Kong, R M. Allen, and L. Schreier, "MyShake: a smartphone seismic network for earthquake early warning and beyond," Science Advances, Vol. 2, No. 2, Feb. 2016.

[16] Q. Kong, A. Inbal, and R. M. Allen, "Machine learning aspects of the MyShake global smartphone seismic network" Seismological Research Letters, Vol. 90, No. 2A, PP. 546-552, 2019.

[17] F. Anjomshoa, M. Catalfamo, and D. Hecker, "Mobile behaviometric framework for sociability assessment and identification of smartphone users," Proc. 2016 IEEE Symposium on Computers and Communication, pp. 1084-1089.

[18] Z. Rauen, F. Anjomshoa, B. Kantarci, "Empowering human-computer interaction in securing smartphone sensing," Proc. 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 1-6.

[19] Q. Xu, R. Zheng, "Mobibee: A mobile treasure hunt game for location- dependent fingerprint collection," Proc. 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, pp.1472- 1477.

[20] Q. Xu, R. Zheng, and E. Tahoun, "Detecting location fraud in indoor mobile crowdsensing," Proc. 2017 the First ACM Workshop on Mobile Crowdsensing Systems and Applications, pp. 44-49.

[21] V. Singh, D. Chander, and U. Chhaparia, "SafeStreet: an automated road anomaly detection and early-warning system using mobile crowd- sensing," Proc. 2018 10th International Conference on Communication Systems & Networks (COMSNETS), pp. 549-552.

[22] Kumar, E. B., & Thiagarasu, V. (2017). Comparison and Evaluation of Edge Detection using Fuzzy Membership Functions. International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE), ISSN, 2454-4248.

Stochastic Modelling and Computational Sciences

[23] S. Rosen, S. Lee, and J. Lee, "MCNet: crowdsourcing wireless per- formance measurements through the eyes of mobile devices," IEEE Communications Magazine, Vol. 52, No. 10, pp. 86-91, 2014.

[24] Reddy, C. S., Yookesh, T. L., & Kumar, E. B. (2022). A Study on Convergence Analysis of Runge-Kutta Fehlberg Method to Solve Fuzzy Delay Differential Equations. Journal of Algebraic Statistics, 13(2), 2832-2838.

[25] F. Montori, L. Bedogni, and L.Bononi, "A collaborative internet of things architecture for smart cities and environmental monitoring," IEEE Internet of Things Journal, Vol.5, No. 2, pp. 592-605, 2017.

[26] F. Montori, L. Bedogni, and G. Iselli, "Delivering iot smart services through collective awareness, mobile crowdsensing and open data," Prof. 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 1-6.

[27] Y. Liu, Z. Yu, and B. Guo, "CrowdOS: a ubiquitous operating system for crowdsourcing and mobile crowd sensing," IEEE Transactions on Mobile Computing, 2020. (Early Access)

[28] M. T. Rashid, D. Wang, "CovidSens: a vision on reliable social sensing for COVID-19," Artificial Intelligence Review, Vol. 54, No. 1, pp. 1-25, 2021.

_____

Using Semi-Supervised Learning to Predict Weed Density and Distribution for Precision Farming

Mr. Abhilash Sam Paulstin K.C.
Research Scholar
Department of Computer Science
AJK College of Arts and Science
Coimbatore, Tamilnadu, India.
e-mail: abhi.sam83@gmail.com

Dr. Angeline Prasanna G.
Former Associate Professor and Head
Department of Computer Science
AJK College of Arts and Science
Coimbatore, Tamilnadu, India.
e-mail: drgangelineprasanna@gmail.com

Abstract— If weed growth is not controlled, it can have a devastating effect on the size and quality of a harvest. Unrestrained pesticide use for weed management can have severe consequences for ecosystem health and contribute to environmental degradation. However, if you can identify problem spots, you can more precisely treat those areas with insecticide. As a result of recent advances in the analysis of farm pictures, techniques have been developed for reliably identifying weed plants. . On the other hand, these methods mostly use supervised learning strategies, which require a huge set of pictures that have been labelled by hand. Therefore, these monitored systems are not practicable for the individual farmer because of the vast variety of plant species being cultivated. In this paper, we propose a semi-supervised deep learning method that uses a small number of colour photos taken by unmanned aerial vehicles to accurately predict the number and location of weeds in farmlands. Knowing the number and location of weeds is helpful for a site-specific weed management system in which only afflicted areas are treated by autonomous robots. In this research, the foreground vegetation pixels (including crops and weeds) are first identified using an unsupervised segmentation method based on a Convolutional Neural Network (CNN). There is then no need for manually constructed features since a trained CNN is used to pinpoint polluted locations. Carrot plants from the (1) Crop Weed Field Image Dataset (CWFID) and sugar beet plants from the (2) Sugar Beets dataset are used to test the approach. The proposed method has a maximum recall of 0.9 and an accuracy of 85%, making it ideal for locating weed hotspots. So, it is shown that the proposed strategy may be used for too many kinds of plants without having to collect a huge quantity of labelled data.

Keywords- Artificial intelligence, convolutional neural networks, machine learning, precision agriculture, semi-supervised learning.

I. INTRODUCTION

When considering the importance of industries for human existence, agriculture remains at the top of the list. There has been significant progress in farming equipment in recent years. In farming, weeding refers to the process of removing or treating undesired vegetation. Weeds can compete with your crops for resources like water, fertiliser, and natural light, so it's essential to keep weeds under control. For this reason, their deliberate removal is required to guarantee a high-quality harvest [1, 2]. But the usual practise of using agrochemicals to treat all farms the same way to kill weeds is ineffective and can hurt soil biodiversity, the quality of fresh water, and human

health. Rather than using substances to kill the weeds could try hand-weeding. Even though this way of carrying out things gets the job done, it takes a lot of time and work. By definition, "precision agriculture" is a "management approach that takes into account temporal and geographical variability to improve agricultural productivity over the long term." [3]. Precision agriculture is often used to find weeds, check the health of crops and soil, control operations like tillage, sawing, mechanical weeding, and fertiliser distribution, estimate crop output, find fruits and vegetables, and pick them[4]. It has been demonstrated that autonomous robots can be used for chemical weeding of weed plant patches [5,6]. These robots utilise

277

_____

machine vision and other detection and localization methods to focus in on and eradicate undesirable plant life.

There are four primary steps make up a traditional image processing-based weed detection method: pre-processing, segmentation, feature extraction, and classification. In pre-processing, the input image is prepared for segmentation by applying various image enhancement techniques, such as a change to the colour space. Afterward, a segmentation technique is used to separate the enhanced image into a foreground and a background. The two most common kinds of this procedure are index-based segmentation and learning-based segmentation. The index-based technique differentiates between plant and background by comparing the intensity value of each pixel to a threshold setting. This approach is notoriously unreliable due to factors such as overlapping crop and weed plants and fluctuating lighting conditions [7,8]. In order to accurately identify the vegetation, it has been shown that learning-based techniques are better [2]. In segmentation, a plant mask is made, and its parts could be crop pixels or weed pixels. This is why a feature vector is made by hand using the biological shape, spectral properties, visual textures, and geographical settings of food plants and weed plants. The feature vectors are then sent into a classifier, which decides whether or not the segmented plant in question is in fact a weed. Conventional weeding methods have a number of drawbacks, including the necessity for a lot of manual labour and the expense and inconvenience of chemical spraying. Because they rely on human-created features, however, these techniques can only be employed with a limited number of plant species or invasive weeds. Deep learning-based techniques [10] have been proposed in recent years as a means of eliminating the need for such characteristics.

Most of these methods, however, are supervised, which means that they need a lot of training data. This means that they can only be used on a small group of crops and weeds. It is hard to create a strong, scalable vision system for the independent robots as a result of factors such as 1) various kinds of lighting, 2) weed and crop plants that overlap and hide each other, 3) different weed densities, and 4) different types of crops and weed plants. For the supervised learning method to work, labelled data is also very important. By looking at the sorts of weeds included in the image, a species name may be assigned to the complete picture [11]. This approach may identify weeds in the field, but it can't tell how many there are. This study comes up with a way to figure out how many weeds there are and where they are without having to label each pixel. Our study looks at a semi-supervised method for weed localization and density estimates with the goal of reducing the amount of human annotation needed to train deep networks. By using less data-intensive segmentation networks, we may be able to speed up adoption for a wider range of crop/weed types and settings.

Our main goal is to make a semi-supervised decision-support system that can successfully predict where and how many weeds are from a single colour picture taken by a self-driving robot. Our main focus is not on pixel-level segmentation, but on the more basic question of whether or not pesticides deserve to be used specifically in specific regions. Either the weed's expected spread or its location and density can be used to figure this out. By applying the proposed approach, sources of weed damage can be found accurately. Find out how many weeds there are in the affected areas.

Because it doesn't depend on pixel-by-pixel annotations like standard end-to-end deep learning segmentation networks, it is more scalable and can be used in more places.

To determine if a picture is background or vegetation, the suggested technique uses an unsupervised Convolutional Neural Network. Neither only that, but the proposed approach identifies as background every pixel that is not ground or a plant. The vegetation mask is applied on top of the tiled parts of the input colour image. Next, the algorithm labels each tile that is covered by plants as either a weed or a crop. Unlike prior image-based approaches to weed classification, the proposed method does not rely on manually created attributes. Also, the proposed answer does not need a lot of segmentation tagging of crop and weed plant pixels, like the methods in [10], [12].

II. RELATED WORKS

Here, we'll take a quick look back at both traditional and modern approaches to weed categorization in photos using deep learning. The latest advancements in deep learning have been applied to precision agriculture, allowing farmers to avoid the pitfalls of older methods. In [13], the most recent applications of deep learning to agricultural problems such weed identification, land cover classification, and fruit counting are outlined.

A. Supervised Technique

Recent years have seen state-of-the-art outcomes for applications such as autonomous driving achieved through the application of deep learning techniques to challenging datasets. Yet, they are generic in that they may be used with many different kinds of things. Class management is minimal while doing weed identification and mapping. Several research [10] propose an end-to-end semantic segmentation network based on earlier efforts like SegNet to distinguish crop plants from weed plants. Networks are trained using 465 multispectral images, and impressive F1 scores (>0.95) are produced in another research paper. Although not a large number of training images are utilized, high-priced multispectral sensors are required for reliable results. Training networks on the same set of 10,000 RGB images as the authors of [9], [10] resulted in

278

_____

comparable performance (F1 score > 0.90). These results support the feasibility of using deep learning models to distinguish between crop plants and weeds in a training environment. However, a big, manually annotated dataset is required to train a supervised learning model's network. This issue is less pressing in contexts where models can adequately generalise to numerous environments without suffering a performance hit (such as object detection for common items such as chairs, humans, etc). The authors of [9] further demonstrate the need for adaptable networks by testing the trained network on a new set of plants and achieving success. Vegetation in the background seems pretty similar in both datasets, though. Instead, our study offers an alternative to pixel-wise segmentation algorithms. Using a scatter transform to build feature vectors, the authors of [14] classify cultivated plants. When applied to domesticated vegetation, the accuracy of this method—which is taught using a made-up data set—is around 85%. Another example of a supervised learning technique from [15] is presented. Using artificial IDs for planted crops, this strategy achieves 99.7 percentile results in computer vision detection accuracy. Instead, then using photographs that have been edited in any manner or marked with physical markers, our solution relies on raw RGB images as input.

Object detection is advocated as a means of identifying weeds. A deep neural network is trained on the data, and then it generates coverage maps and bounding boxes to find the locations of plants and weeds. Since accurate findings from this approach need manual annotation of covering maps and bounding boxes, it is highly data-intensive. Making a multispectral orthomosaic map [16] involves projecting a 3D point cloud onto a 2D plane. They provide a possible answer to the challenge of scanning a large area without losing fine-grained information on plant distribution. These maps are then used as input for a modified SegNet model, which extracts the weeds from the background noise. Data-intensive and requiring sensors able to generate point clouds and the training of an end-to-end segmentation model (the study employed a dataset with more than 10,000 pictures). As shown in [17], a binary vegetation mask is produced initially by employing an end-to-end segmentation network. The generated landscape mosaic is then sent into a sophisticated VGG-16 network for labelling. Two-stage pipelines are effective, but both networks must be trained on the relevant agricultural domains. Our study applies unsupervised learning for vegetation segmentation based on the idea of a two-step method for weed identification (which is the first stage). Making tile labels is now all that's needed to begin training a classifier. These modules may be used to reduce dependency on data, and they can be easily adapted to fit other crop/weed scenarios.

B. Semi-Supervised Techniques

Semi-supervised and unsupervised methods of learning have also been explored for their potential use in weed identification. As an illustration, in [1] we see a comparison of the deep unsupervised learning algorithms JULE and DeepCluster [18], in conjunction with a deep network like VGG-16 or ResNet-50; they are used for weed categorization and automated labelling. K-Means pre-training is used to fine-tune the network weights before a LeNet-5 model is used for weed classification. These algorithms, in contrast to others, do not produce a precise map of weed or weed pixels but rather a general forecast regarding the image's categorization. They need to know the weed density in order to use variable herbicide spraying, which increases application efficiency and minimises environmental damage, but they have no method of doing so. An unsupervised technique for classifying plants is provided. They achieve competitive performance if there is no overlap between weeds and crops. When dealing with such a wide variety of plant species, that is not a reasonable assumption to make. Another challenging aspect of the unsupervised technique is determining how many clusters to use when dividing the image. The suggested approach solves this problem by relying solely on an unsupervised method to partition the vegetation mask, resulting in a maximum of two distinct groups.

The approach described in [17] is the one that most closely resembles the one that is being proposed. The authors employ a deep learning approach to weed identification. A two-stage network was implemented, with a convolutional neural network (CNN) doing the initial mask extraction to differentiate weeds from crops. On the other hand, there is a significant difference between these parts and the ones used in the proposed job. The proposed method requires significantly less data for training than the supervised learning networks utilized in [17]. (Vegetation segmentation is unsupervised while the classifier is trained with a small number of region labels). Comparing the 2000 pictures needed to train a network in [17] with the 90 and 500 images used to test the proposed method (including the upgraded versions) reveals a significant reduction in the number of images required for testing. In contrast to the networks in [17], which require pixel-wise annotations for training, the proposed pipeline does not require them. Using instances of previously undiscovered plant species and their binary labels, only the classifier has to be adjusted in the proposed study. They also point out that the overlap between plants is a common source of errors. It has been shown that the proposed approach is robust under low-light settings, occlusions, and high-plant densities, and that it is also adaptable enough to work with a wide variety of plant species. The proposed technique provides more accurate estimates of

_____

weed density and dispersion from RGB images than previous semi-supervised algorithms.

C. Estimation of Weed Density

The number of weeds is a good sign of which places need to be treated with chemicals. There are ways to figure out how many weeds are in a row crop in [19] and [20]. The number of weeds per unit of land and the number of weeds per unit of crop are used to measure the quantity of weeds. Positional histograms, which are mentioned in [20], are used to figure out where the weeds are. By counting the number of white pixels in a binary vegetation mask down each column, we can get the horizontal pixel distribution, which is then shown as a histogram. It involves finding out the weed density (the number of weed pixels at a certain interval as a percentage of the size of the whole picture) for a set number of time intervals. This strategy only works for weeds growing in the spaces between rows, and it doesn't account for the potential of weeds and crops growing in the same space. When trying to estimate weed densities, it is also necessary to make assumptions regarding the positioning of the crop rows. As a result, it can only be employed in a limited range of agricultural contexts.

III. PROPOSED METHODOLOGY

To ensure that only severely infested regions are treated, the proposed technique seeks for and evaluates weed density. All that is needed to feed into the pipeline is a single RGB picture.
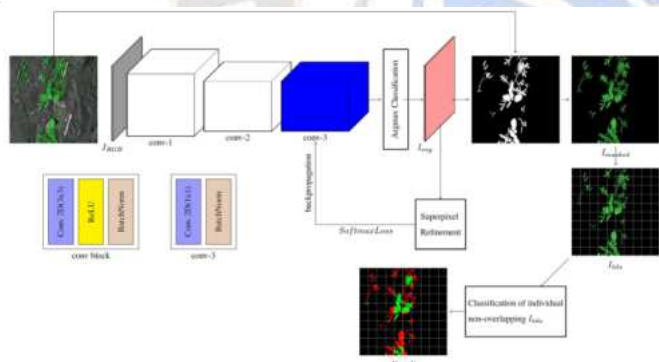

Figure 1. Proposed Method Overview

Figure 1 describes about getting a better grasp on the intended pipeline layout. Using an unsupervised deep learning-based segmentation system, each pixel in a picture is initially classified into one of two groups: vegetation or background. Two masks are made: one for the plants in the foreground and other is the background. Imasked defines the area of interest (ROI) by applying the vegetation mask, Iveg, on top of the original RGB picture. Itile then cuts this into square tiles that are smaller and smaller. Each tile has a feature vector (Itile) that describes the plants shown by the pixels on that tile. These vectors are used by a binary classifier to decide

whether Itile is a food plant or a weed. We also look at how well a trained CNN (ResNet50) can sort things on Itile. From where "weed" Itile zones are, you can tell where weeds are a problem. We can use the number of plants per square metre as a proxy for the number of weeds. By dividing the number of pixels in an area by the total number of pixels in that area, we can get an idea of how many crops and weeds are in that area. The proposed method is easily scalable and applicable to a broad variety of weed and agricultural plant kinds since just a tiny percentage of it is trained under supervision. The steps are described in further depth below.

A. Segmentation of Vegetation

First, the bicubic interpolation method from the OpenCV package is used to stretch the IRGB picture to 500x500 square pixels. The image's pixels need to be split into two different groups: the background and the centre. Here, we use the unsupervised segmentation technique with convolutional neural networks (CNNs). However, the piece stands on its own, thus just the most vital information about the work is provided. This iterative approach consists of two parts: the forward pass of the network (label prediction) and the back-propagation (learning network parameters while assuming the labels are fixed). The following restrictions are suggested by the approach for identifying the most likely cluster or class to which each pixel belongs: The first limitation is in the necessity of feature similarity. Similar pixels tend to group together in clusters. Each pixel requires its own response map, which is generated for this purpose. Each pixel is assigned to a cluster with its neighbours based on the response map. Second, there can be no gaps in the continuity between distinct sites. The authors use an image's superpixels and randomly give each of those labels the same cluster label. The term "superpixel" refers to a group of pixels that are all physically similar in some way, such as their proximity to one another or their brightness levels. The network employs the Simple Linear Iterative Clustering (SLIC) method to extract the superpixels from a five-dimensional space (three channels of CieLabcolorspace and two-dimensional image coordinates (x, y)). The number of individual clusters into which the image is divided is then constrained. If a maximum of q clusters is employed, under segmentation can be prevented even with a large number of classes. This necessitates doing intra-axis normalization on the response map prior to applying cluster labels.

It is based our methodological choice on the restrictions imposed by pixel-wise segmentation. This technique prioritizes spatially continuous pixels and allows us to set the bare minimum number of clusters at two, which is useful for identifying weeds and agricultural plants thanks to their closed-loop structures (background and vegetation). Pixel-by-pixel segmentation is refined in this way until (1) most pixels can be

_____

separated into two groups, or (2) the maximum number of classification rounds has been reached. This limits the time it takes for the segmentation to converge, preventing either under- or over-segmentation. The cluster with the fewest pixels is utilized as a mask for the plants when the image is divided. This is so because there will be more pixels in the background than there will be plants.

To improve the performance of the unsupervised segmentation, we randomly choose 30% of the data from each dataset and use that to fine-tune the network's parameters. The ideal values for the network's parameters, including the number of superpixels, their density, and their rate of learning, are determined by a thorough examination of the data. By adjusting one setting at a time, we can determine which is optimal. Furthermore, during this time, all other variables are ignored. Mean intersection over union (mIOU) values are selected as the optimal choices for each parameter. The following experimental parameters were determined: The parameters are as follows: (1) learning rate = 0.1, (2) number of superpixels = 2500, and (3) superpixel compactness = 25. Photos in the experimental group have had vegetation masks applied to them using the ideal parameters (Iveg). To compare the effectiveness of the supervised and unsupervised methods, we additionally trained U-Net on the training subset of the datasets. U-efficacy Nets have been shown to be a supervised learning solution for pixel-wise segmentation in several uses, such as medical image segmentation and autonomous driving.

Using an encoder-decoder architecture, the network first downsamples the picture to get the prediction, and then upsamples the image to get the real data. At each step in the "upsampling" process, the feature map from the most recent "downsampling" is added to the original. Even though traits are lost when downsampling, the network may still be able to learn something from them. The network was taught to put binary class labels on the white pixels in the centre to show that there were plants there.

B.   Tile Classification

The masked image Imasked is the result of applying the vegetation mask Ivegon the input picture IRGB. This masked image consists of just the RGB pixels for the vegetation, ensuring that only plant-related attributes are used in the classification process (crops and weeds). The masked image (Imasked) is then divided into even smaller pieces (called Itile) of 50x50 pixels. Quite a few places could have few or no greenery pixels. So, in Itiles, if the percentage of land covered by plants (measured in terms of the number of vegetation pixels) is less than 10%, weeds are not considered to be an issue (in pixels). Figure 2 depicts image Imasked, region rejection owing to a lack of vegetation pixels, and area selection for classifier training.
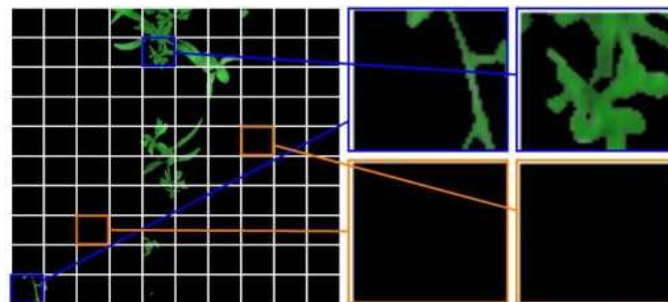

Figure 2. I_{masked} is divided into smaller tiles (I_{tile})

Various machine learning techniques, including SVMs, RFs, and MLPs, have been developed to help with this problem. A few examples of well-known machine learning methods have been used to classification issues. In this study, we evaluate the classifiers' capacity to accurately categorize Itile as either a weed or a crop. In this part, we first go through feature vector-based classifiers for determining if Itile is a weed or a crop. Furthermore, we discuss an alternate image-based classifier for Itile, one that use a trained convolutional neural network rather than calculating the feature vector directly.

C.   Weed Density Estimation

After identifying weed-infested regions (Itiles with the weed attribute), weed density may be determined by measuring the total area covered by plants in those locations. Here, a cluster rate (CR) from is used to quantify and estimate weed density ([19]). Estimates of weed densities are crucial for site-specific weed management [2]. It's possible that this population density estimate will help when choosing where to spray herbicides in the field. This choice would be influenced by factors such as the sorts of crops and weeds to be grown and the distance between plants.

CR=Weed plant coverage in the region (in pixels) /Total land area of the region (in pixels)

Algorithm: Weed Distribution and Density Estimation

Input: Color image (I_{RGB}) of the field acquired from an autonomous robot;

Output: Weed density and distribution;

Given (I_{RGB}), Generate the vegetation mask (I_{veg}) using CNN based unsupervised segmentation;

Overlay I_{RGB} with I_{veg} to get I_{masked};

Divide the image I_{masked} into smaller regions I_{tile}(squaretiles);

For (I_{tile} in $_{Imasked}$)do

Classify I_{tile} into crop, weed or background;

If I_{tile} is weed then

Estimate weed density

end

end

281

_____

IV. RESULTS AND DISCUSSION

The evaluations are carried out through qualitative analysis and quantitative analysis.

A. Evaluation by Qualitative Analysis

Figure 3 displays the segmentation outcomes for some sample instances in both datasets. The unsupervised segmentation network outperforms the supervised segmentation network in distinguishing the vegetative pixel data from the background, according to the inferred performance (U-Net). Figure 3 shows how the unsupervised segmentation method may be used to identify and distinguish between various plant structures (3rd row of the image). It's important to find U-Net can classify vegetation even whether it's represented by a single pixel or a very sparse collection of them. This is not a typical behaviour for the unsupervised approach. That's because our approach gives greater weight to the spatial continuity of the vegetation clusters, while U-Net focuses more on the neighbourhood of a single pixel (down sampling using max pooling). This pattern was considerably more apparent in the Sugar Beets dataset (which has less contrast than the CWFID).
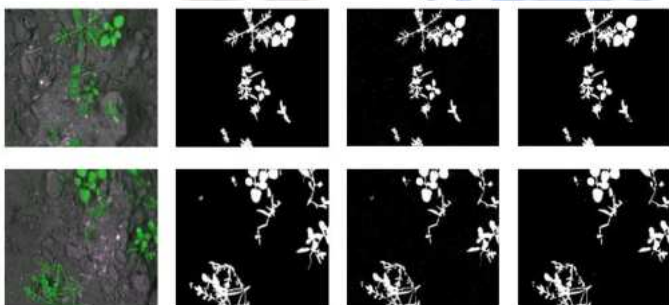

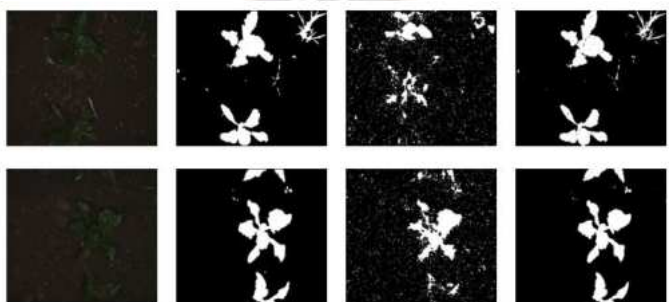Figure 3(a). Crop/Weed Field Image Dataset

Figure 3(b). Sugar Beet Dataset

B. Evaluation by Quantitative Analysis

The test splits of both datasets are used to get the mean intersection over union (mIoU) number, which is then used to compare the performance of the two networks. Table 1 shows how things turned out. The unsupervised network did much better than U-Net on the Sugar Beets dataset, and it also did better than U-Net by a small amount on the CWFID dataset.

This might be because, unlike U-Net, the unsupervised method doesn't depend on learning mapping from a limited set of features to tell the difference between pixels in the foreground and pixels in the background. Both supervised and uncontrolled algorithms do better on the CWFID dataset than on the Sugar Beets dataset. This is because the pictures in the Sugar Beets collection have bad lighting and, as a result, not much contrast. The results of this study support the idea that an unsupervised network can be used to pull vegetation pixels from shots of different plants in different places.

TABLE I. QUANTITATIVE EVALUATION OF VEGETATION

Model	Dataset	mIoU
Unsupervised Segmentation	CWFID (Crop/Weed Field Image Dataset)	0.9
Unsupervised Segmentation	Sugar Beet Dataset	0.8
UNet	CWFID	0.9
UNet	Sugar Beet Dataset	0.7

C. Estimation of Weed Density

Once the weedy spots have been found, the cluster rate per tile can be found by using the weeded vegetation pixels. In Table 2, the rate of clusters that was seen in the weedy areas is compared to the rate that was expected. These results show that for both types of data, it is possible to make accurate estimates of the number of weeds. There are four main reasons for the loss of weed density pixels: 1) ignoring tiles or areas where plants cover less than 10% of the total area, 2) incorrect vegetation segmentation, 3) mislabeling weed-infested areas as crop plants, and 4) plants in a given tile that overlap. Possible mistakes could start with the lack of plants in some places. For the goals of this study, a 10% threshold was chosen because it can be changed enough to fit a wide range of crop plants and weed plants. When used on the CWFID and Sugar Beets datasets, the proposed method gets a mean absolute error of 5% for vegetation segmentation and 1% for weed spread. This backs up the idea that the proposed method might be a good way to fix mistakes that come from the above sources. Since the RMSE between datasets for two different crop/weed species is less than 8%, it is clear that the proposed method can be used for any crop/weed species. Before choosing where to carefully apply agrochemicals, it is important to find out where and how many weeds there are.

TABLE II. WEED DENSITY ESTIMATION ACCURACY

Dataset	Mean Accuracy (%)	MAE	RMSE
CFWID	75	5	7.5
SugarBeets	85	1	3

282

_____

V. CONCLUSION

Once the weedy spots have been found, the weeded vegetation pixels can be used to figure out the rate of clusters per tile. In Table 2, the rate of clusters that was seen in the weedy areas is compared to the rate that was expected. These results show that for both types of data, it is possible to make accurate estimates of the number of weeds. With the help of a computer vision system, we might be able to treat only certain places and cut down on the amount of chemicals we use by a lot. This study suggests a semi-supervised method that could be used to help precision agriculture get more accurate estimates of weed densities and locations.

For the suggested method to work, it needs colour pictures as input. A binary flora mask can be made after the background pixels have been taken out. Using an unsupervised neural network, the bits that make up the background and the plants are put together. Second, we use the mask to divide the original colour picture into smaller parts (50-pixel-square tiles). Then, each piece is labelled as either a crop or a weed. In this study, a fine-tuned ResNet50 is compared to several other classifiers, such as SVM, Gaussian Naive Bayes, Neural Network, and Random Forest, which use a pre-trained ResNet50 as a feature generator. The suggested method is tried on two sets of images called Crop/Weed Field Image and Sugar Beets, which show a wide range of crops and weeds. With a maximum memory of 0.90, weedy areas can be found, and the number of weeds in those areas can be estimated with an accuracy of 85%. One of our goals is to reduce the need for elaborately annotated data sets. In order to identify weeds without the need for new features, the proposed work employs unsupervised segmentation and a pre-trained ResNet50. A pixel-wise segmentation network is shown to be unnecessary for estimating weed distribution and density, in contrast to previous techniques. The recommended pipeline is adaptable enough to process low-contrast images, images with overlapping plants, and images of distinct plant species. This strategy could be advantageous for agricultural organizations looking for cost-effective implementations due to the low data needs for training and tuning. A standard RGB camera is sufficient so long as a platform is in place from which to photograph the plants from above.

REFERENCES

[1] A. dos Santos Ferreira, D. M. Freitas, G. G. da Silva, H. Pistori, and M. T. Folhes, ''Unsupervised deep learning and semi-automatic data labeling in weed discrimination,'' Comput. Electron. Agricult., vol. 165, Oct. 2019, pp. 104963.

[2] A. Wang, W. Zhang, and X. Wei, ''A review on weed detection using ground-based machine vision and image processing techniques,'' Comput. Electron. Agricult., vol. 158, pp. 226–240, Mar. 2019.

[3] International Society of Precision Agriculture, Monticello, IL, USA. Precision Ag Definition. Accessed: Feb 9, 2021. [Online]. Available: https://www.ispag.org/about/definition

[4] Y. Xiong, C. Peng, L. Grimstad, P. J. From, and V. Isler, ''Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper,'' Comput. Electron. Agricult., vol. 157, pp. 392–402, Feb. 2019.

[5] W. McAllister, D. Osipychev, A. Davis, and G. Chowdhary, ''Agbots: Weeding a field with a team of autonomous robots,'' Comput. Electron. Agricult., vol. 163, Aug. 2019, Art. no. 104827.

[6] M. N. Abd. Kharim, A. Wayayok, A. R. M. Shariff, A. F. Abdullah, andE. M. Husin, ''Droplet deposition density of organic liquid fertilizer at low altitude UAV aerial spraying in rice cultivation,'' Comput. Electron. Agricult., vol. 167, Dec. 2019, Art. no. 105045.

[7] M. Hassanein, Z. Lari, and N. El-Sheimy, ''A new vegetation segmenta- tion approach for cropped fields based on threshold detection from hue histograms,'' Sensors, vol. 18, no. 4, p. 1253, Apr. 2018.

[8] S. Sabzi, Y. Abbaspour-Gilandeh, and G. García-Mateos, ''A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms,'' Comput. Ind., vol. 98, pp. 80–89, Jun. 2018.

[9] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, ''Fully convolutional networks with sequential information for robust crop and weed detection in precision farming,'' IEEE Robot. Autom. Lett., vol. 3, no. 4, pp. 2870–2877, Oct. 2018.

[10] A. Milioto, P. Lottes, and C. Stachniss, ''Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs,'' in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), May 2018, pp. 2229–2235.

[11] A. Olsen, D. A. Konovalov, B. Philippa, P. Ridd, J. C. Wood, J. Johns,W. Banks, B. Girgenti, O. Kenny, J. Whinney, B. Calvert, M. R. Azghadi, and R. D. White, ''DeepWeeds: A multiclass weed species image dataset for deep learning,'' Sci. Rep., vol. 9, no. 1, Dec. 2019, Art. no. 2058.

[12] P. Lottes, J. Behley, N. Chebrolu, A. Milioto, and C. Stachniss, ''Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming,'' J. Field Robot., vol. 37, no. 1, pp. 20–34, Jan. 2020.

[13] A. Kamilaris and F. X. Prenafeta-Boldú, ''Deep learning in agriculture: A survey,'' Comput. Electron. Agricult., vol. 147, pp. 70–90, Apr. 2018.

[14] P. Rasti, A. Ahmad, S. Samiei, E. Belin, and D. Rousseau, ''Supervised image classification by scattering transform with application to weed detection in culture crops of high density,'' Remote Sens., vol. 11, no. 3, p. 249, Jan. 2019.

[15] R. Raja, D. C. Slaughter, S. A. Fennimore, T. T. Nguyen, V. L. Vuong,N. Sinha, L. Tourte, R. F. Smith, and M. C. Siemens, ''Crop signalling: A novel crop recognition technique for robotic weed control,'' Biosyst. Eng., vol. 187, pp. 278–291, Nov. 2019.

[16] I. Sa, M. Popović, R. Khanna, Z. Chen, P. Lottes, F. Liebisch, J. Nieto,C.Stachniss, A. Walter, and R. Siegwart, ''WeedMap: A

_____

large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming,'' Remote Sens., vol. 10, no. 9, p. 1423, Sep. 2018.

[17] M. Fawakherji, A. Youssef, D. Bloisi, A. Pretto, and D. Nardi, ''Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation,'' in Proc. 3rd IEEE Int. Conf. Robotic Comput. (IRC), Feb. 2019, pp. 146–152.

[18] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, ''Deep clustering for unsupervised learning of visual features,'' in Proc. ECCV, 2018, pp. 132–149.

[19] Y. Xu, R. He, Z. Gao, C. Li, Y. Zhai, and Y. Jiao, ''Weed density detection method based on absolute feature corner points in field,'' Agronomy, vol. 10, no. 1, p. 113, Jan. 2020.

[20] Y. Xu, Z. Gao, L. Khot, X. Meng, and Q. Zhang, ''A real-time weed mapping and precision herbicide spraying system for row crops,'' Sensors, vol. 18, no. 12, p. 4245, Dec. 2018.
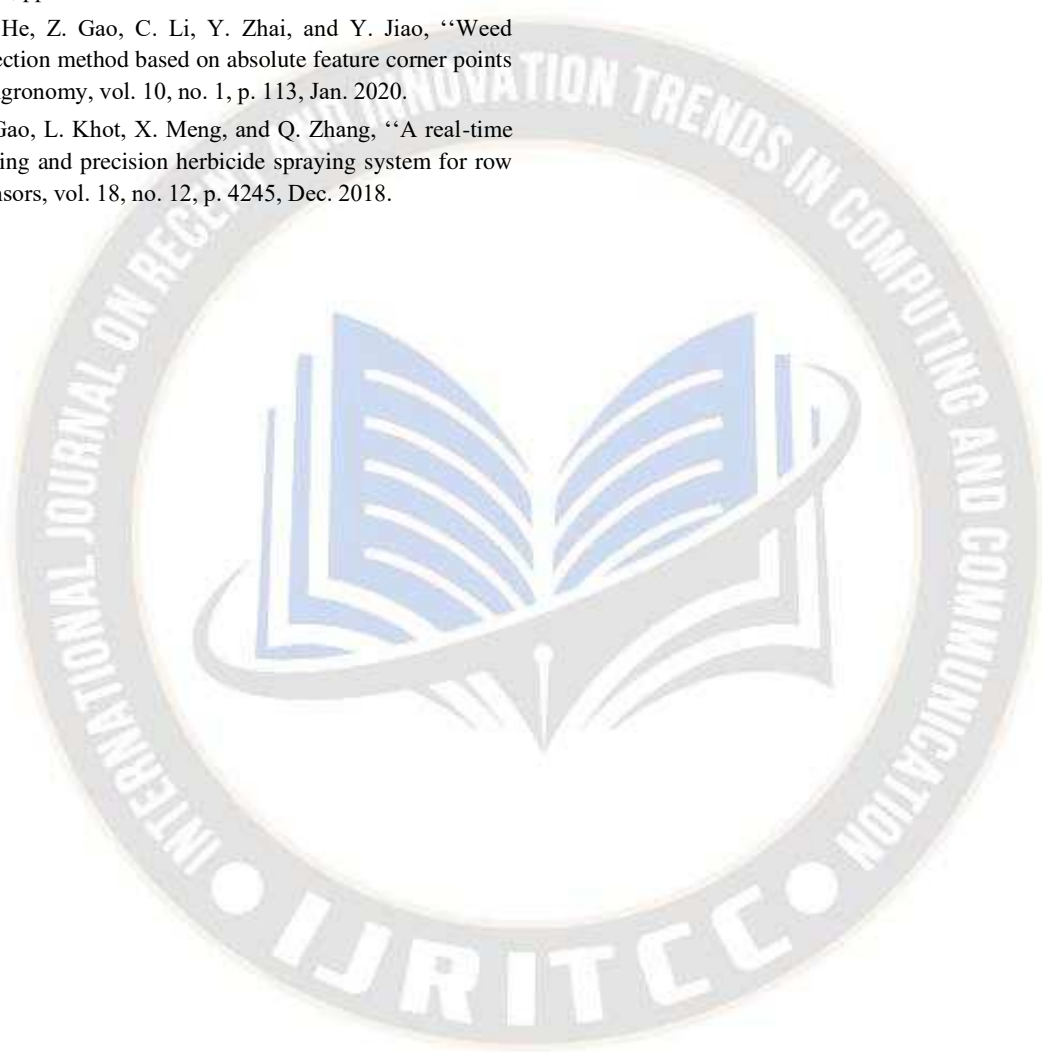
Enhancement of Criminal Facial Image Using Multistage Progressive V-Net for Facial Recognition by Pixel Restoration

S.S. Beulah Benslet[1,*] and P. Parameswari[2]

[1]Research Scholar, Department of Computer Science, Karuppannan Mariappan College, Muthoor, Chettiyarpalayam
[2]Principal, Palanisamy College of Arts, Perundurai-Erode Road, Tamil Nadu 638052, India

Abstract

INTRODUCTION: Criminal activity is expanding exponentially in modern society, which leads towards a great concern about security issues. Facial recognition technology (FRT) is a powerful computer-based system that increasingly being used for recognize and match faces to solve crimes and investigations.

OBJECTIVES: Due to poor image clarity and noisy pixels, the detection of criminal faces tends to be inaccurate. Hence, image enhancement techniques are required to recognize criminals with better accuracy. In the proposed model, a multistage progressive V-net based image quality enhancing technique is employed to improve accuracy.

METHODS: The Convolutional Neural Network (CNN) for restoring images called MPRV-Net has three stages for a difficult balance between spatial data and highly contextualized information for image restoration tasks while recovering images.

RESULTS: For image restoration tasks, including denoising, deblurring, and deraining, MPRV-Net has provided considerable performance benefits on a number of datasets. The suggested network is significant as it eliminates all three types of deviations using a single architecture. The proposed model's performance is tested using performance metrics such as accuracy, precision, recall, and specificity, obtaining 94%, 96%, 93%, and 95%.

CONCLUSION: Thus, the proposed Multistage Progressive V-Net model for effectively improves the criminal Facial image for detecting criminals in public places with greater accuracy.

1. Introduction

Criminal identification is the most significant but challenging and time-consuming task for the police. It will be challenging in crowded environments, such as cities or public places and in some cases, manual identification techniques provide the possibility to acquire more details about offenders. Still, it's time-consuming and involves the risk of missing criminals. Public supervision is not necessary for an automated identification system (AIS) [1]. This will help the police in tracking down and apprehending criminals in public areas. Relevant face detection and recognition algorithms are included in automated identification systems [2]. Face recognition for criminal identification uses a distinct biometric method. This approach identifies and verifies a person's identity based on video or image frames, including their face. Face recognition preprocessing is integrated systems that enhance an input face image to improve the quality of the image by increasing the visibility of facial features and improving prediction performance. The

---

*Corresponding author. Email: beulahabraham15@gmail.com

efficiency of face recognition systems is improved by pre-processing [3]. Additionally, the preprocessing stage improves distorted images and gathers regions of interest in an image for further feature extraction.

The pre-processing step of image enhancement is necessary for enhancing the performance of face recognition systems. Face image enhancers are more effective approaches used before the recognition process and are commonly included in the design of most face recognition systems [4]. Face recognition systems employ a wide range of image enhancement techniques. Each of these approaches has advantages as well as disadvantages. Some recent methods modify an input image resulting in a more precise or detailed output image [5]. However, in uncontrolled contexts, the quality of images in face recognition systems may be significantly decreased for numerous reasons, including lighting conditions, i.e., in dark or overly bright environments. Restoring a clear image from an imperfect one is known as image restoration [6]. CNN-based methods outperform other approaches primarily owing to model development. Among the computational modules and functional units designed for image recovery include encoder-decoders [11], generative models [12], dilated convolutions [8, 9], dense connections [10], and recursive residual learning [7].

Recently, there haven't been many attempts to use the multi-stage method for image deraining and deblurring [13]. Investigate these methods to discover the architectural constraints limiting their efficacy [14]. In order to begin, current multi-stage techniques either rely on a single-scale pipeline [16], which produces spatially accurate but semantically less reliable outputs, or an encoder-decoder architecture [15], which is reliable in preserving broad contextual information but unreliable in preserving spatial image details. However, the conventional approach is insufficient to differentiate between criminal and non-criminal behaviour. The majority of the accessible models for low-level visual issues are developed in a single stage. In high-level performance, however, it has been proved that multi-stage networks outperform single-stage analogues. To increase accuracy, the proposed model has introduced multistage progressive v-net based image restoration. The following is presented as the proposed model's contribution:

- Multistage progressive V net had been proposed for pixel restoration, effectively enhancing criminal images.
- The collected dataset for criminal recognition is gathered from public websites and split into criminal and non-criminal categories for training classifiers.
- Vnet architecture is used in multistage progressive networks for extracting the pixel level feature based on encoding and decoding.
- The final stage includes the original resolution sub network (ORSNet) to achieve spatially relevant outcomes for a pixel-to-pixel correspondence between the input and output images.
- An efficient Supervised Attention Module (SAM) that uses the restored image at every level to enhance incoming features before propagating them further.

The remaining portion of the manuscript is sectioned as follows: Section 2 denotes some related work related to criminal prediction. Section 3 presents the proposed methodology along to predict criminals. Section 4 presents the result and discussion of the work, and Section 5 defines the conclusion of the work.

2. Related Work

According to development, several techniques have been introduced to detect criminals in various scenarios and its states. Among them, a few recently developed work which is related to criminal detection is reviewed below.

Ratnaparkhi et al. [17] had designed an embedding method that maps facial features to a compact Euclidean face map, which can be used to detect differences in the face and a deep convolution neural network approach to detect criminals. This technique works well; however, it doesn't work well on multiple faces due to blurry or cropped images.

Sandhya et al. [18] had developed a system that uses a feed-forward neural network termed an autoencoder, whose input and output values are identical, to implement the deep learning neural network approach. One of the most well-known functions of auto-encoders is reconstructing the input image. As a result, use the encoder's output, which can be utilized for obtaining the input image by the decoder for identification verification rather than comparing the perpetrator's overall face. This technique uses a simple algorithm to assess how closely the input image equals images stored in the criminals' database. Although this method is efficient, it is time-consuming and overfitting inappropriate for multiple images.

Kumar et al. [19] had suggested a face recognition and criminal identification system utilizing a multi-task cascading network. This system will be capable of automatically distinguishing criminal faces. The system's one-shot learning mechanism would also require a single image of the criminal to identify. In this method, the criminal's face is recognized, the data for the identified criminal is gathered from the database, and a notification is sent to the police staff with every relevant detail and the position where the criminal was being watched by the camera. However, this method required a high level of time-consuming and model execution needs to be enhanced.

Venkatesh et al. [20] had evolved a classifier that used Fisher face and Local Binary Pattern Histograms (LBPH). It has been demonstrated that this technique works well for identifying criminal faces. The Fisher face classifier is useful for lowering dimensionality, whereas the LBPH classifier is known for its robustness to changes in illumination and subtle fluctuations in facial expressions. Combining these two techniques makes it possible to identify faces even in difficult lighting situations. However, this approach is not appropriate for rotating image prediction and consumes more time.

Ganji et al. [21] had used the well-known Principal Component Analysis method to develop an automatic facial recognition system for a criminal database. Automatic face

detection and recognition will be possible with this system. In the absence of a thumbprint being present at the scene, this will aid law enforcement in identifying or detecting the criminal in the case. This technique works well, but it needs to be enhanced.

Amjad et al. [22] invented a facial recognition method based on deep learning that can identify or predict whether a person is a criminal or not and identify the probability they would be. A ResNet50 model that uses CNN and an SVM classifier to extract features from a dataset was utilized for training. This method is effective, but it has to be improved. Aherwadi et al. [23] had developed a method for face detection employing Face Encodings. In addition to providing the Police tremendous ease in identifying criminals, this version of the criminal detection system also saves them time because processes are automated. Although this strategy is efficient, it occasionally fails to identify appropriate people.

The above literature has numerous techniques and algorithms that were introduced to predict criminal. The outcome of these methods provide well performance but have some limitation that reduces the system's accuracy. The impacts that are having in the convolution model are reduced efficiency [21], pixel mismatch [17], overfitting [18], process consume more time as well as poor image quality [20], [22] and lack of robustness [19],[23]. To overwhelm these drawbacks, a novel method is proposed that enhances the facial image of criminal and non-criminal and improve accuracy. The next section presents a thorough explanation of the suggested methodology.

3. Proposed Methodology

Criminal identification and detection are slow and challenging processes. Nowadays, criminals are more clever than ever before, leaving no traces of biological evidence or fingerprints at the scene of the crime. Using advanced facial identification technology is a quick and simple solution. The majority of buildings and traffic lights now have CCTV cameras installed for surveillance purposes due to advancements in security technology. The camera's video footage can be utilized to identify suspects, criminals, runaways, and missing persons, among other things. These images are usually insufficient, making it challenging to recognize criminals and poor quality. Suitable enhancement methods are required for pre-processing the images. A multistage progressive V net image restoration technique is used in the proposed model to improve image quality. The proposed Multistage Progressive V-Net model process flow is depicted in Figure 1.
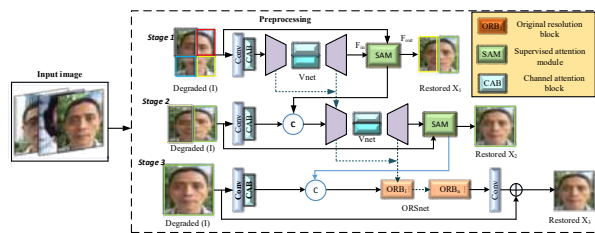

Figure 1. Architecture of proposed model.

The dataset used for this model includes both criminal and non-criminal images. Preprocessing the dataset is necessary due to the image's poor quality and difficulty in recognizing criminals. A Multistage progressive V-net based preprocessing technique is used in the proposed model. A three-stage CNN (convolutional neural network) called the multi-stage progressive V network (MPRV-Net) is used to restore images. For various types of Image restoration challenges such as deblurring, deraining and denoising, eliminated using an MPRV-Net and providing good performance on the dataset. The below section clearly explains the proposed methodology.

3.1. Dataset collection

The dataset used in this model was gathered from CK+ and fbi.gov websites which consist of both criminal and non-criminal images. To enhance the quality of the images as well as prediction, these datasets need to be preprocessed. The following section clearly explains the proposed preprocessing technique.

3.2. Preprocessing

Image restoration is a process of removing distortions like blur, noise, and rainfall from images to develop new pixels with clear details. For instance, the blur effect may be caused by the camera moving, the subject moving, the distortion of scattered light, the lack of proper depth of field, and the lens's softness. Noise is a term used to describe unwanted fingerprints and variations in the brightness or colour of the data. The unwanted effects of rainfall could impede the performance of the processing method. Therefore, it is crucial to employ pre-processing evaluators to eliminate these effects and improve the accuracy of the proposed model. The MPRV-Net-based preprocessing technique is used.

Multi-Stage Progressive V Net
Three steps combine to form MPRV-Net's progressive image restoration process. The first two stages are developed using a traditional V-net-based encoder-decoder model, which learns every contextual detail from the input image. The result uses the original resolution sub-network (ORSNet), which proceeds at the original image resolution resulting in spatially accurate results for pixel-by-pixel relation between the input and output images.

Include a supervised attention module between each stage rather than simply cascading several steps. The module rescales the feature maps from earlier stages before transferring them to the following stage, all under the surveillance of ground truth images. Introduce a method for integrating intermediate features from different stages of a network by using the intermediate multi-scale contextualized features of the earlier sub-network. The input image is accessed at each stage of the MPRV-Net stacking process. According to Fig. 1, the final result is divided into three distinct, non-overlapping stages: the initial phase using four patches, the second phase using two patches, and the final phase utilizing the original image. This division is typical of current restoration techniques. Similar to current restoration techniques, the multi-patch structure on the input image splits the image.

At each stage C, a diminished input image that I is included in the residual image R_C predicted by the proposed model, which produces X_C:

$$X_C = I + R_C \qquad (1)$$

The following loss equation is used in the MPRV-Net's from start to finish optimization:

$$L = \sum_{C=1}^{3} \left[L_{char}(X_C, Z) + \lambda L_{edge}(X_C, Z) \right] \qquad (2)$$

Where L_{char} is the char bonnier loss, and Z stands for the ground truth image.

$$L_{char} = \sqrt{\|X_C - Z\|^2 + \varepsilon^2} \qquad (3)$$

With constant ε empirically adjusted to 10^{-3} for all experiments. Additionally, L_{edge} is the edge loss and is defined as:

$$L_{edge} = \sqrt{\|\Delta(X_C) - \Delta(Z)\|^2 + \varepsilon^2} \qquad (4)$$

Where Δ represents the Laplacian operator, the relative relevance of the two loss factors is controlled by the value in equation (2), which is set to 0.05. The following section describes the key elements of the suggested method individually.

Key elements of MPRV-Net

The key element of the MPRV-Net model is,
- Processing of Complementary Features
- Cross-stage Feature Fusion
- Supervised Attention Module

In below section clearly explain each key feature,
- *Processing of Complementary Features*

Decoder-encoder and pipeline with only one scale of features are typically the architecture designs employed by single-stage CNNs for image restoration [24]. The input is gradually converted into low-resolution representations by the encoder-decoder networks, and after that, progressively reverse map and restore the original resolution. Although these models effectively encode multistate data, their frequent usage of

down sampling techniques makes them susceptible to losing spatial characteristics. However, approaches that use a single-scale feature processing can be used to obtain images with excellent spatial features. However, the restricted receptive field makes their outputs less semantically robust. This demonstrates the inherent limitations of the aforementioned architectural design decisions, which can only have spatially accurate and contextually reliable effects. The following is a proposed multi-stage architecture that employs encoder-decoder networks in earlier stages and a network that works with the original input resolution in the final stage in order to utilize the advantages of both designs:

Encoder-Decoder Subnetwork: The encoder-decoder sub-network employs channel attention blocks (CABs) to acquire features at every stage. Figure 2 illustrates an encoder-decoder subnetwork developed using a conventional V net.
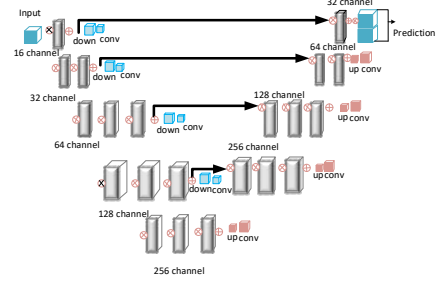

Figure 2. Architecture of V-net model.

The decompression path is positioned in the lower section of the V-Net, and the compression path is obtained in the higher section. A residual function is learned at each step of the calculation to speed it up, and it is then processed by ReLU before being added to the output of the final convolutional layer.

The bottom portion of the V-Net collects features and enhances the spatial support of lower resolution feature maps in order to gather and combine the necessary data and present volumetric segmentation. The foreground and background regions of the low- and high-resolution feature maps are probabilistically segmented using the soft-max function. Using a de-convolution process, the input complexity is increased. Features gathered from the initial stages of the compression path are transferred down to the bottom segment to improve the precision of contour prediction.

Using a combination of contour recovery and an approximation method, coordinates and the bounding box size are recovered and refined throughout the fusion stage. A distance ratio of 1:1 is used to determine whether two candidates are actually one discovery or two independent ones for candidates whose centres are too near to one another. The distance between the centres of two identified candidates is divided by the expected side of the bounding box from larger candidates to obtain the distance ratio. The results from each V-Net combine to determine whether the final forecast will be accurate. High-resolution features are computed by

the MPR V-Net stage using a number of original resolution blocks (ORBs), each of which also contains CABs.

Original Resolution Subnetwork: Introduce the original-resolution subnetwork (ORSNet) in the final stage to maintain fine features from the input to the output image. While developing spatially enriched, high-resolution features, ORSNet does not use downsampling operations. There are several original-resolution blocks (ORBs) in it, each of those ORBs also contains CABs. ORSnet model effectively develops multiscale information, it has the ability to lose some spatial detail information by continuously down sampling. In the final stage, present the original resolution network (ORSNet) and include the enhancedd features from the input to the output image. Original resolution sub-network to generate rich spatial high-resolution features to make up for the loss of spatial information. At the network's end, the high-resolution rain line features obtained from the original resolution sub-network are combined with the original rain map to obtain the final rain removal image. ORSNet generates rich, high-resolution spatial features without down sampling operations, which consist of many original resolution blocks (ORBs). The structure of the ORB is shown in Figure 1,

- *Cross-stage Feature Fusion*

The CSFF module is integrated between two encoder-decoders and ORSNet in the proposed architecture. Instead of being transmitted to the following stage for aggregation, the characteristics from one stage are first improved by 11convolutions. The proposed CSFF has several advantages. Starting with the encoder-decoder, it continuously performs upward and down-sampling operations in order to reduce a network's sensitivity to information loss. Second, the multistate characteristics of one stage assist in enhancing the characteristics of the following stage. Third, by facilitating information flow, the network optimization process becomes more stable, allowing to expand the overall design by a number of steps.

- *Supervised Attention Module*

A clear image will be identified at each level by recent multi-stage image restoration networks, which is then passed on to the stage after that. To significantly boost performance, add a supervised attention module (SAM) between each pair of steps. The SAM schematic diagram is depicted in Figure 1 and details both of its contributions. For starters, it generates real-time supervisory signals useful at each progressive image restoration stage. Second, develop attention maps using locally supervised predictions that only allow the most efficient attributes to proceed to the following stage, suppressing the less valuable ones at this stage.

Figure 1 depicts the circumstances, the input parameters $F_{in} \in R^{H \times W \times Y}$ from the previous stage are processed by SAM, which then provides an unprocessed image $R_C \in R^{H \times W \times 3}$ using an automated 1×1convolution, where Y is the total number of channels, and $H \times W$ is the space dimension. The degraded input image I blended with the residual image resulting in the modified image $X_C \in R^{H \times W \times 3}$. Provide explicit supervision with the ground-truth image for the predicted image X_C. Using a 1×1 convolution and sigmoid activation, the image X_C is then used to create per-pixel

attention masks$M \in R^{H \times W \times Y}$. Following the application of these masks to the modified local features, F_{in} provides attention-guided features that are subsequently included in the identity mapping path. Last but not least, problems with image restoration are eliminated by the attention enhanced representation F_{out} developed.

4. Result and Discussion

A novel Multistage Progressive V net for pixel restoration that significantly enhanced criminal images has been designed in this model. The objective manner of the dataset is collected from public websites divided into criminal and non-criminal categories for criminal recognition. These images are usually poor in quality and inadequate, making them difficult to recognize. As a result, it is crucial to employ pre-processing techniques to remove and restore the noise in the criminal image. A novel, multi-stage methodology that can result in outputs that are both contextually rich and spatially accurate is used in the proposed model. Due to its multi-stage structure, a framework is used to break down the difficult task of image restoration into easier steps and progressively restoration a diminished image. The proposed model performance is analyzed using Intel Core i7 CPU, NVIDIA GeForce RTX 3070 GPU, and 64GB RAM software.

4.1. Dataset Description

In the designed model, the dataset is collected through CK+ and fbi.gov websites [25] [26]. Overall, 934 image data are collected, split into criminal and non-criminal groups. In this group, 339 images are criminal, and 595 images are non-criminal considered.

The images are frequently insufficient, poor quality and difficult to identify. For pre-processing the criminal and non-criminal images, suitable enhancing techniques are required [27]. The proposed model enhances the image quality using a multistage progressive v net image restoration technique. This algorithm consists of 3 stages, which split the image and learn every feature using the V-net architectures. Figure 3 represents pre-processing results of the proposed dataset [28].

| Original image	Preprocessing image
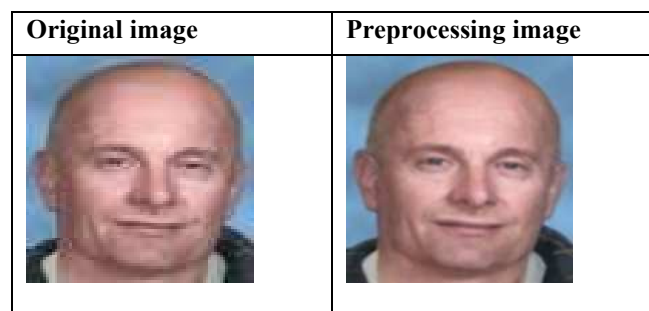
Figure 3. The outcome of the proposed model pre-processing.

MPRV-net image performance metrics are compared with a few other current models in the following part to analyze image quality. Average gradient, mean square error (MSE), root mean square error, peak signal-to-noise ratio, signal-to-noise ratio, and correlation coefficient are the metrics used for comparison. Each metric is individually analyzed, and the values are observed to analyze the performance of the pre-processing approach.
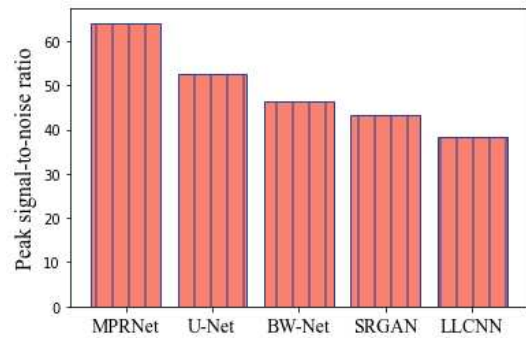
Figure 4. Peak signal-to-noise ratio for the proposed and existing model.
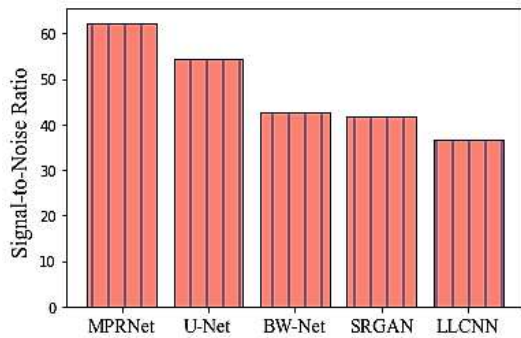
Figure 5. Signal-to-noise ratio for the proposed and existing model.
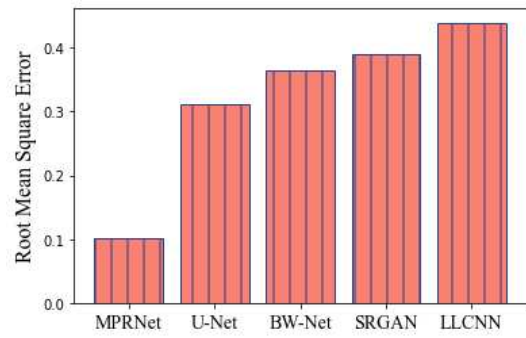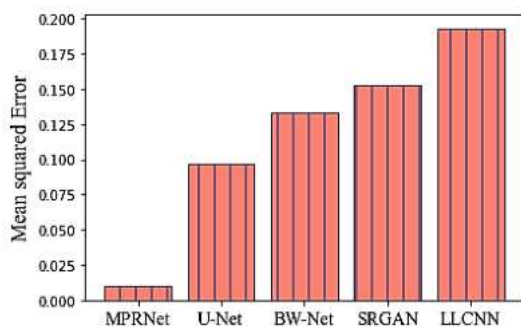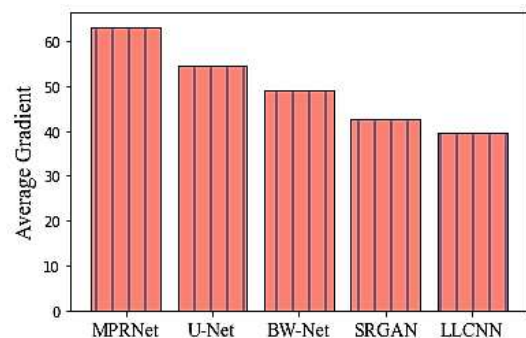

Figure 7. Root mean square error for the proposed and existing model.

A comparison of the peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR) of the proposed MPRV-net model provides for testing the methodology shown in figure 4 and figure 5. The PSNR is the ratio of a signal's highest practicable value (power) to the highest possible amount of distorted noise that compromises the accuracy of its representation. The proposed MPRV-net model provides 64.1 PSNR, U net have 52.6 PSNR, BW-Net have 46.3 PSNR, SRGAN have 43.2 PSNR and LLCNN model have 38.4 PSNR. Thus, the proposed model provides a higher PSNR value than the traditional approaches shown in figure4. Then the signal-to-noise ratio is analyzed, and its comparison is shown in figure5. The SNR will rise with a larger field of vision because a larger pixel size results in more signals being received by each individual pixel. Large pixels will pick up more signals, resulting in high SNR images. The proposed MPRV-net model provides 62.3 SNR, Unet have 54.3 SNR, BW-Net have 42.8 SNR, SRGAN have 41.7and LLCNN model have 36.6 SNR. It demonstrates the proposed model provides a better SNR value than another models.

The proposed MPRV-net, MSE and RMSE are compared with the methods that are already in use, as shown in figure 6 and figure 7. To validate the procedure that was illustrated in figure 6, the MSE is analyzed, and the observed values are compared to the convolution model. The proposed MPRV-net model has 0.0105MSE, U net has 0.0968MSE, BW-Net has 0.1332MSE, SRGAN has 0.1524MSE, and the LLCNN model has 0.1932MSE. Thus, show that the proposed model has a lower error value than the traditional approaches. The method depicted in figure 7 is then validated by doing an MSE analysis and comparing the observed values to the convolution model. The proposed MPRV-net model has 0.10246RMSE, whereas U net, BW-Net, SRGAN, 0.3649RMSE, and the LLCNN model each have 0.39038RMSE, 0.3649RMSE, and 0.31112RMSE, respectively.

Figure 8. Average gradient for the proposed and existing model.

Figure 6. Mean square error for the proposed and existing model.
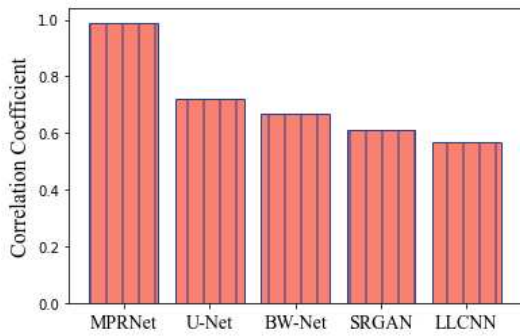
**Figure 9.** Correlation coefficient for the proposed and existing model.

The proposed MPRV-Net average gradient and correlation coefficient are compared with the existing methods, as shown in Figures 8 and 9. The average gradient develops the relationship between the image's clarity and the pattern's finely detailed variance. The proposed MPRV-net model provides a gradient of 63.2, whereas Unet provides a gradient of 54.6, BW-Net a gradient of 49.2, SRGAN a gradient of 42.8, and the LLCNN model provides a gradient of 39.5. It indicates that when compared to existing approaches, the proposed MPRV-net model has high gradient values in figure 8. Figure 9 shows the analysis of the correlation of the image. When performing correlation, a filter mask, frequently referred to as a kernel, is moved over the image, and the sum of the products is calculated at each place. The proposed MPRV-net model provides 0.99 correlation, Unet have 0.72 correlation, BW-Net have 0.67 correlation, SRGAN have 0.61 correlation, and the LLCNN model have 0.57 correlation. It demonstrates that the proposed model provides a better average gradient and correlation coefficient value than other models.
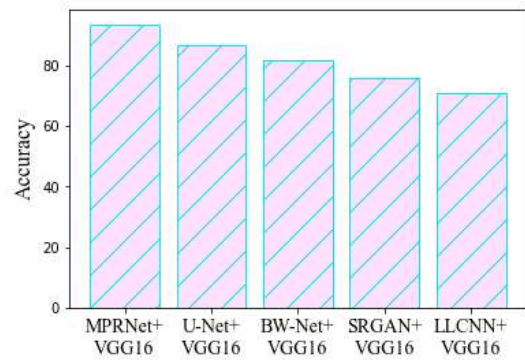
Confusion matrix
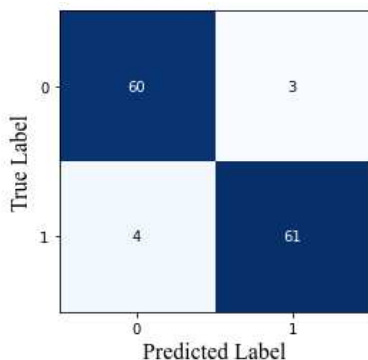
Figure 10. Confusion matrix of the proposed method.

Figure 10 shows the proposed model confusion matrix model. A classification algorithm performance is determined using a confusion matrix technique, which is a collection of

data that is organized according to both actual and expected data. For the performance analysis, data obtained for such a system are evaluated. The predictive analysis technique provides a confusion matrix that includes positive and negative rates (true and false). Figure 10 makes it clear that 61 and 60 images are correctly predicted in 0 and 1 class, respectively, whereas 3 and 4 images are wrongly predicted. The positive and negative rates are also used to measure the accuracy, sensitivity, specificity, and null error rates. The performance metrics are compared to existing approaches like U Net, BW-Net, SRGAN and LLCNN models.

Figure 11. Comparison of accuracy for the proposed and existing model.

Figure 11 shows the accuracy values, which were calculated using the confusion matrix. ADD some points about performance metrics and accuracy. The system that predicts a value with the least degree of error is said to be accurate. The proposed method has a 94% accuracy rate, compared to Unet at 87%, BW-Net at 82%, SRGAN at 76%, and LLCNN at 71%.
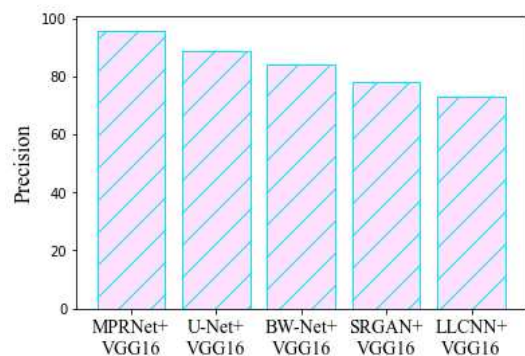
Figure 12. Comparison of precision for the proposed and existing model.
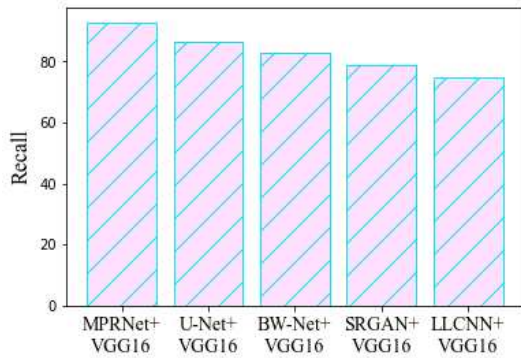
Figure 13. Comparison of recall for proposed and existing model.

Figure 15. Comparison of error for the proposed and existing model.

A comparison of the intended and existing approaches' precision is shown in Figure 12. The quantity of anticipated favourable events is what is meant by accurate measurement. In comparison to other existing approaches, such as Unet, BW-Net, SRGAN and LLCNN, with corresponding precision values of 89%, 84%, 78%, and 73%, the proposed method's precision value was discovered to be 96%. As shown in figure 13, recall is calculated by dividing the total number of components that actually fall into the positive class by the number of true positives. In comparison to other existing approaches, such as Unet, BW-Net, SRGAN and LLCNN, with corresponding recall values of 86.4%, 83%, 79% and 75%, the proposed method's recall value was discovered to be 93%.

Figure 14 shows a comparison of the specificity of the proposed and current techniques. The degree to which a model can predict the real negatives of every imaginable sort is known as specificity. In comparison to various approaches, such as Unet, BW-Net, SRGAN and LLCNN, with corresponding specificity values of 84%, 86%, 77% and 76%, the proposed method's specificity value was determined to be 95%. The system operates worse when the error is high and better when the error is low. Figure 15 illustrates the error contrast among the proposed and existing methodologies. The proposed method's error rate is 6%, Unet's is 13%, BW-Net's is 18%, SRGAN's is 24% and LLCNN's is 25%. It demonstrates that the proposed model has lower error rates than the conventional models.

Figure 14. Comparison of specificity for proposed and existing model.

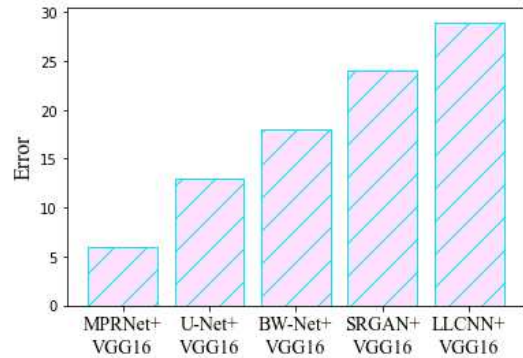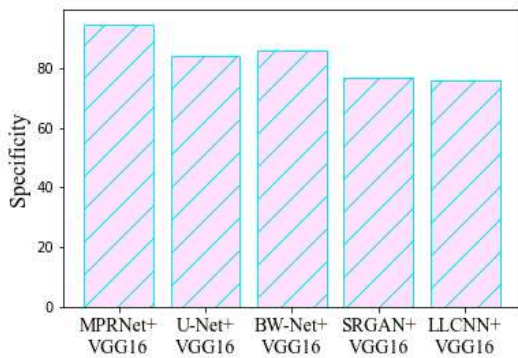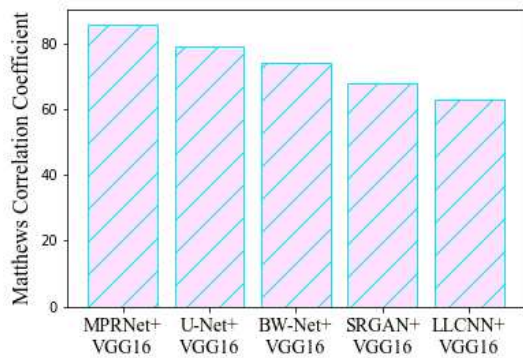**Figure 16.** Comparison of Matthews's correlation coefficient for the proposed and existing model.
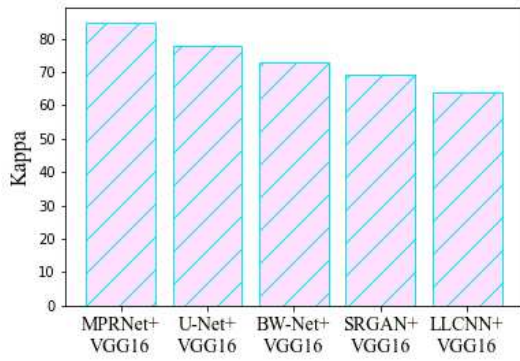
Figure 17. Comparison of kappa for the proposed and existing model.

Figure 19. Analysis of testing time for the proposed and existing model.

Figure 16 illustrates an MCC comparison of the proposed and existing methods. The purpose of MCC is to evaluate or quantify the discrepancy between expected and actual values. MCC in the proposed method has a value of 86%, Unet of 79%, BW-Net of 74%, SRGAN of 68% and LLCNN of 63%. Kappa comparison of the suggested and existing approaches is shown in Figure 17. A statistical indicator of the consistency of different variables across rates is called kappa. The proposed method's kappa value is 85%, Unet is 78%, BW-Net is 73%, SRGAN is 68% and LLCNN is 64%.

4.2. Time comparison analysis

In this section, the overall working process time comparison was provided. Process Cycle Time refers to the duration of the overall process. It may be computed instantly by adding the total sum of time spent on each process step and the delay time. Figure 18 and figure 19, 20 shows the training time, testing time, and execution time. The numbers are compared to Unet, BW-Net , SRGAN and LLCNN models to confirm the procedure.
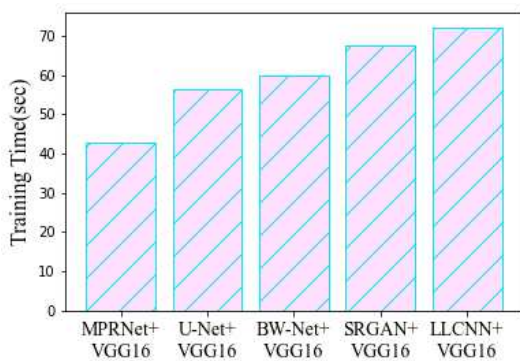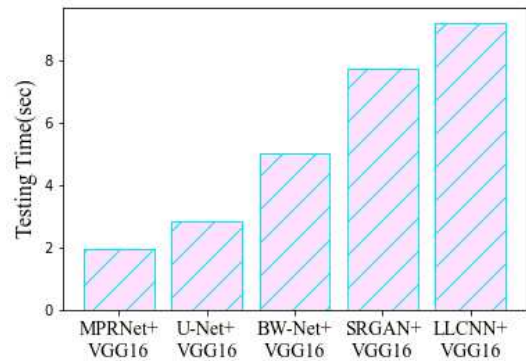

Figure 18. Analysis of training time for the proposed and existing model.

Figure 20. Analysis of execution time for the proposed and existing model.

During the training period, 80% of the data were used to train the MPRV-net model for effective prediction, as the existing models were also trained using 80% data. The training period causes some time to complete the training process. In figure 18, the training time of the proposed model is 42.72 ms, the Unet model consumes 35 ms, the BW-Net model takes 41 ms, the SRGAN model consumes 45 ms, and the LLCNN consume 50 ms. The testing time is also analyzed and plotted in figure 19. It demonstrates that the proposed MPRV-Net takes 1.96 ms, the existing Unet model consumes 2.84 ms, the BW-Net model takes 5.02 ms, SRGAN model takes 7.73 ms, and the LLCNN model takes 79.21 ms. Then the proposed model overall execution time is analyzed and compared to some other models that are shown in figure 20. The overall execution time is termed the addition of training and testing time. The proposed model consumes 44.68 ms, the Unet model consumes 59.14 ms, the BW-Net model takes 64.82 ms, the SRGAN model consumes 75.34, and the LLCNN consume 81.51 ms to complete the process. The above comparison demonstrates that the proposed model effectively improves the prediction accuracy of any input image.

5. Conclusion

A multistage progressive V-net image restoration technique is used in the proposed model to improve image quality. Performance is limited by unwanted traces and fluctuations in the images' brightness or colour, making recognition challenging. So, using a revolutionary image restoration technique, it is proposed to eliminate distortions from images like blur, noise, and rainfall in order to develop new, clear ones. At first, the criminal images with different styles of each criminals' and non-criminals' datasets are collected through public websites. The multi-stage progressive image restoration network (MPRV-Net) has three steps for progressively restoring images. The input criminal and non-criminal images have been employed in the first two steps to develop an encoder-decoder model based on a common V-net. The original resolution sub-network (ORSNet) generates spatially appropriate results for a pixel-to-pixel consistency between the input and output images in the final stage at the original image resolution. Because of their hierarchical multi-scale representation and computational effectiveness, encoder-decoder-based V-net architectures have been mostly used for recovery in convolutional networks. MPRV-Net revealed an enhanced multi-layer architecture developed to accommodate high-level global characteristics and local specifics. Channel attention blocks (CABs) extract the encoder-decoder sub-networks characteristics at each level, followed by bi-linear interpolation and a convolution layer. Multiple original resolution blocks (ORBs), each including CABs, are used to compute detailed features. A cross-stage feature fusion (CSFF) module is implemented between each pair of stages. As a result of multiple up- and down-sampling, the network as a whole becomes more robust and stable, besides improving information flow.

Additionally, the efficacy of the proposed MPRV-net approach was analyzed and evaluated against other conventional approaches like U-net, BW-net, SRGAN and LLCNN. The proposed approach has 94% accuracy, 93% recall, 95% specificity, 86% MCC and 85% kappa. The comparative analysis demonstrates that the suggested model provides a more efficient result than the present approaches. In future work, the enhanced images are used for detecting the facial landmarks even with different sizes, lighting and poor pixel quality that effectively predict the criminal face using neural network.

Reference

[1] Apoorva P, Impana HC, Siri SL, Varshitha MR, Ramesh B. Automated criminal identification by face recognition using open computer vision classifiers. In2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) 2019 Mar 27 (pp. 775-778). IEEE.

[2] Hasan BM, Abdulazeez AM. A review of principal component analysis algorithm for dimensionality reduction. Journal of Soft Computing and Data Mining. 2021 Apr 15;2(1):20-30.

[3] Payal P, Goyani MM. A comprehensive study on face recognition: methods and challenges. The Imaging Science Journal. 2020 Feb 17;68(2):114-27.

[4] Porcu S, Floris A, Atzori L. Evaluation of data augmentation techniques for facial expression recognition systems. Electronics. 2020 Nov 11;9(11):1892.

[5] Kortli Y, Jridi M, Al Falou A, Atri M. Face recognition systems: A survey. Sensors. 2020 Jan 7;20(2):342.

[6] Sun S. Application of fuzzy image restoration in criminal investigation. Journal of Visual Communication and Image Representation. 2020 Aug 1;71:102704.

[7] Liu H, Cao F, Wen C, Zhang Q. Lightweight multi-scale residual networks with attention for image super-resolution. Knowledge-Based Systems. 2020 Sep 5;203:106103.

[8] Lei X, Pan H, Huang X. A dilated CNN model for image classification. IEEE Access. 2019 Jul 8;7:124087-95.

[9] Amirian S, Rasheed K, Taha TR, Arabnia HR. Automatic image and video caption generation with deep learning: A concise review and algorithmic overlap. IEEE access. 2020 Dec 4;8:218386-400.

[10] Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y. Residual dense network for image restoration. IEEE transactions on pattern analysis and machine intelligence. 2020 Jan 22;43(7):2480-95.

[11] Brooks T, Mildenhall B, Xue T, Chen J, Sharlet D, Barron JT. Unprocessing images for learned raw denoising. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2019 (pp. 11036-11045).

[12] Zhang H, Sindagi V, Patel VM. Image de-raining using a conditional generative adversarial network. IEEE transactions on circuits and systems for video technology. 2019 Jun 3;30(11):3943-56.

[13] Suin M, Purohit K, Rajagopalan AN. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020 (pp. 3606-3615).

[14] Ren D, Zuo W, Hu Q, Zhu P, Meng D. Progressive image deraining networks: A better and simpler baseline. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019 (pp. 3937-3946).

[15] Tao X, Gao H, Shen X, Wang J, Jia J. Scale-recurrent network for deep image deblurring. InProceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 8174-8182).

[16] Zamir SW, Arora A, Khan S, Hayat M, Khan FS, Yang MH, Shao L. Multi-stage progressive image restoration. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2021 (pp. 14821-14831).

[17] Ratnaparkhi ST, Tandasi A, Saraswat S. Face detection and recognition for criminal identification system. In2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) 2021 Jan 28 (pp. 773-777). IEEE.

[18] Sandhya S, Balasundaram A, Shaik A. Deep Learning Based Face Detection and Identification of Criminal Suspects. Computers, Materials & Continua. 2023 Feb 1;74(2).

[19] Kumar KK, Kasiviswanadham Y, Indira DV, Bhargavi CV. Criminal face identification system using deep learning algorithm multi-task cascade neural network (MTCNN). Materials Today: Proceedings. 2023 Jan 1;80:2406-10.

[20] Venkatesh M, Dhanalakshmi C, Adapa A, Manzoor M, Anvesh K. Criminal Face Detection System.

[21] Ganji A, Kamat R, Biradar SC, Gadivaddar S. Criminal Face Detection Using Machine Learning.

[22] Amjad K, Malik AA, Mehta S. A technique and architectural design for criminal detection based on lombroso theory using deep learning. Lahore Garrison University Research Journal of Computer Science and Information Technology. 2020 Sep 25;4(3):47-63.

[23] Aherwadi NB, Chokshi D, Pande DS, Khamparia A. Criminal Identification System using Facial Recognition. InProceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021 Jul 12.

[24] Haque IR, Neubert J. Deep learning approaches to biomedical image segmentation. Informatics in Medicine Unlocked. 2020 Jan 1;18:100297.

[25] Dataset 1: https://www.fbi.gov/wanted/fugitives

[26] Dataset 2: https://paperswithcode.com/dataset/ck

[27] Ratnaparkhi ST, Singh P, Tandasi A, Sindhwani N. Comparative analysis of classifiers for criminal identification system using face recognition. In2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) 2021 Sep 3 (pp. 1-6). IEEE.

[28] Jhanani R, Harshitha S, Kalaichelvi T, Subedha V. Mobile Application for Human Facial Recognition to Identify Criminals and Missing People Using TensorFlow. Journal of Research in Engineering, Science and Management. 2020;3(4):16-20.
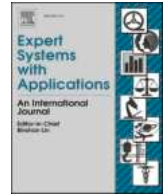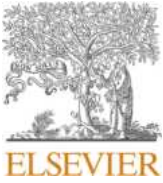
Network intrusion detection: An optimized deep learning approach using big data analytics

D. Suja Mary [a,1], L. Jaya Singh Dhas [b,2], A.R. Deepa [c,3,*], Mousmi Ajay Chaurasia [d,4], C. Jaspin Jeba Sheela [e,5]

[a] *Assistant Professor, Department of Computer Science, Nanjil Catholic College of Arts and Science, Kaliyakkavilai, Tamil Nadu, India*
[b] *Assistant professor, Department of Computer Science Scott Christian College (Autonomous), Affiliated to Manonmaniam Sundaranar University, Tirunelveli Nagercoil 629003, India*
[c] *Associate Professor, Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur District, Andhra Pradesh, India*
[d] *Muffakham Jah College of Engineering and Technology, Hyderabad, India*
[e] *Assistant Professor Department of PG Computer Science, Nesamony Memorial Christian College, Marthandam, India*

ARTICLE INFO

ABSTRACT

Managing enormous amounts of data, such as big data, and detecting network traffic intrusions are inefficiently handled by current computing technologies. Traditional analytical techniques cannot manage the incursions in continuous internet traffic and the enormous log data of server activity, leading to many inaccurate results and a prolonged training period. As a result, this research provides an efficient deep learning-based approach to enhance the attack identification task by addressing the basic big data complexity linked to many heterogeneous security data types. This framework employs a novel feature selection method incorporating the Aquila Optimizer (AO) and Fuzzy Entropy Mutual Information (FEMI) algorithms to pick distinctive characteristics. Subsequently, a modified canonical correlation-based technique is applied to combine selected characteristics. Then, the intrusion identification and categorization are carried out using the optimized ResNet152V2 method. Additionally, data augmentation using Auxiliary Classifier Generative Adversarial Network (ACGAN) is performed. Finally, we used the CICDDoS2019 and ToN-IoT datasets to validate the suggested methodology. By comparing the presented approach to several baseline methods, the effectiveness of the suggested methodology is assessed using various performance measures, including F1-score, recall, precision, accuracy, confusion matrix, and ROC curve. Finally, simulation results show that the suggested strategy is superior to other existing techniques and demonstrate that it is a resilient solution for network intrusion detection.

1. Introduction

In the modern world, the development of Internet of Things (IoT) devices and technologies has accelerated at a never-before-seen pace. With growing transmission bandwidth and speed, ioT devices can gather, send, and process massive amounts of data. These data, which could be essential for forecasting, making decisions, competing in the marketing industry, etc., are constantly at risk (Wang et al., 2023; Dasgupta and Saha, 2022; Altunay and Albayrak, 2023). Additionally, the rise in the amount and categories of intrusions (like malicious attacks, illegal recordings, network viruses, etc.) poses a significant risk to people's data safety and asset protection. A corrupted IoT device could transmit incorrect cloud data to servers or allow illegal access to confidential company records, financial estimates, and business data. Equipment malfunctioning and loss of money could result from this (Ramya et al., 2023; Gu et al., 2023; Ponmalar and Dhanakoti, 2022).

* Corresponding author.
E-mail addresses: drdsujamary@gmail.com (D. Suja Mary), jayasinghdhas@scottchristian.org (L. Jaya Singh Dhas), deepaamuth@kluniversity.in (A.R. Deepa), mousmi.chaurasia@mjcollege.ac.in (M.A. Chaurasia), jaspinjebasheela@nmcc.ac.in (C. Jaspin Jeba Sheela).
[1] ORCID: 0000-0002-3690-076x.
[2] ORCID: 0009-0006-3271-9791.
[3] ORCID: 0009-0006-7963-2212.
[4] ORCID: 0000-0003-4904-6643.
[5] ORCID: 0000-0002-2565-0814.

Therefore, the importance of communications and information security has increased for both people and society.

Especially some critical networks, such as military bases and government buildings, need a high level of security, and firewalls are frequently utilized as a fundamental security measure. However, it is no longer adequate owing to the challenges of people's adaptation and the introduction of unfamiliar intrusions (Sarkar et al., 2023; Al Moaiad et al., 2022; Ragab et al., 2023). Furthermore, identifying intrusions is not much simpler in a massive data environment (big data). Therefore, to discover network intrusion detections in a big data environment, an efficient intrusion detection technique is required.

For various networking assaults, numerous intrusion-detection systems (IDS) models have been provided; nevertheless, many of these models have identified previously undisclosed attacks (Pustokhina et al., 2022; Ahakonye et al., 2023; Alhabshy et al., 2022). In addition to the conventional detection system, machine learning techniques such as Support Vector Machines (SVM), Neural Networks, and Hidden Markov Models were developed and applied to support intrusion detection systems. However, the big data environments often involve datasets with a large number of features or dimensions. Therefore, traditional machine learning models may struggle with high-dimensional data as they may be prone to overfitting or computational inefficiency. Moreover, the security threats in big data environments often exhibit complex and nonlinear relationships that are challenging for traditional models to comprehend. Furthermore, conventional machine learning often relies heavily on manual feature engineering, where domain experts need to identify and extract relevant features (Venkatesan, 2023; Awotunde et al., 2023; Polepally et al., 2022). Therefore, identifying incursion in the big data environment is very difficult using these machine learning techniques. Moreover, the attacks carried out by unidentified attackers, the real-time application has also become challenging.

Unlike machine learning, deep learning approaches yield superior outcomes and have a higher number of computing layers. In addition, it may identify complicated trends and nonlinear interactions in the data set. This flexibility is essential for identifying intricate and dynamic patterns of intrusion in large data sets. Additionally, they effectively manage the huge data (Zhang, 2022; Yin et al., 2023; Adewole et al., 2022). Apache Spark and deep learning algorithms are employed in big data frameworks to improve the overall efficacy and adaptability of the system. The continuously increasing amount of network traffic necessitates using a modern intrusion detection system (IDS) in conjunction with Apache Spark, Hadoop, and Apache (Park et al., 2023; Aldwairi and Alansari, 2022). As a result, this work uses the Pyspark framework to create an efficient deep learning-based technique for intrusion detection in a big data context.

The framework includes a highly effective hybrid deep feature selection technique based on the Aquila optimizer (AO) and fuzzy entropy mutual information (FEMI). The AO algorithm, recognized for its optimization capabilities, may find potential feature subsets with efficiency. This can be used in conjunction with FEMI, which excels at collecting non-linear and fuzzy relationships, to produce a powerful hybrid method that minimizes processing complexity and converges to significant feature subsets. After that, they are integrated using the modified canonical correlation technique (MCC). Ultimately, the big data architecture intrusions are detected and categorized using the optimized ResNet152V2. The techniques under consideration were trained on two different dataset such as CICDDoS2019 and ToN-IoT to achieve a high degree of accurateness in network intrusion detection.

The main contributions of the suggested framework are listed below:

- This study's feature selection process combines AO and FEMI to offer a novel approach. This hybridization increases the effectiveness of feature selection, which is crucial to improving the security breach detection procedure.
- When MCC is utilized for feature fusion, the feature engineering procedure assumes a new dimension. This technique contributes to

optimizing the selected features, further improving the model's overall performance.
- The introduction of an optimized ResNet152-based classifier contributes to accurate and robust classification of intrusion patterns. Moreover, the Wildebeest Herd Optimization (WHO) algorithm based hyperparameter optimization enhance its performance.
- Testing our method on two different datasets shows that it works well in various situations and with different types of intrusion data. This helps prove that our approach is adaptable and reliable.

The paper's subsequent sections are organized as follows. A summary of several pertinent concepts put out recently is covered in Section 2 of this article. Section 3 presents the preliminary concepts. The approach we suggest for efficient intrusion detection is described in Section 4. The evaluation findings are covered in Section 5. The paper is concluded in Section 6.

2. Literature review

In the field of cybersecurity, overcoming big data problems and enhancing the detection of attacks are essential. To manage the complexity of various security data types, researchers are investigating novel approaches that combine conventional and cutting-edge techniques. A brief description of a few of them is provided below, and Table 1 provides an overview of the literature review. Furthermore, in order to tackle the challenges posed by heterogeneous security data, this review examines a variety of approaches.

Ponmalar and Dhanakoti (Ponmalar and Dhanakoti, 2022) addressed primary challenges related to heterogeneous security data by proposing a method that improves the intrusion detection process. Their approach combines Ensemble Support Vector Machines (SVM) and Chaos Game Optimisation (CGO) algorithms. Through statistical analysis and the evaluation of performance indicators such as confusion matrix, ROC curve, precision, accuracy, F1-score, and recall, the effectiveness of their methodology was assessed. The model demonstrated proficiency in categorizing various intrusions on the UNSW-NB15 dataset, outperforming several baseline models.

An intrusion detection framework with deep learning techniques was presented by Cui et al. (Cui et al., 2023). The three components of this system were classification, data enrichment, and feature extraction. Initially, the stacked autoencoder-based component was offered for obtaining the characteristics from the input data to get an improved description of the data. Secondly, the Wasserstein generative adversarial network and the clustering technique derived from the Gaussian mixture model were used for the data augmentation procedure. Thirdly, the categorization step was created using CNN and LSTM. Using the UNSW-NB15 and NSL-KDD datasets, the performance of the proposed method was evaluated and contrasted with the state-of-the-art methods.

In a comparable manner, Fu et al. (Fu et al., 2022) employed an attention mechanism coupled with bidirectional LSTM to detect anomalies within the network. The creators of this structure first use the CNN approach to extract key characteristics from the input data. The network weights were then redistributed using the attention technique. Ultimately, the Bi-LSTM acquired the attributes extracted to categorize the incursions. The authors used the adaptive synthetic sampling (ADASYN) technique to gather more data in smaller categories to address the unbalanced data issue. Additionally, an improved stacked autoencoder has been employed for dimensionality reduction. Various performance metrics were utilized for performance evaluation, and the results were contrasted with previous standard methods.

To develop an intrusion detection system, Ramkumar et al. (Ramkumar et al., 2022) utilized a Spark environment with deep learning techniques. In this framework, Deep Residual Network (DRN) was hybridized with the Exponential Sea Lion Optimization and RV coefficient. The composite feature fusion based on the RV coefficient was applied to choose the distinctive features. It was created by combining the

Table 1

Summary of the literature review.

Study	Techniques	Dataset	Performance metrics	Drawback
Ponmalar and Dhanakoti (Ponmalar and Dhanakoti, 2022)	Ensemble SVM, CGO	UNSW-NB15	Confusion matrix, ROC curve, Precision, Accuracy, F1-score, Recall	This approach is slightly affected by changes made to the conditions determining when the algorithm stops its iterations (referred to as "halting function revisions"). These changes might impact the performance of the solution, especially in terms of energy consumption during the classification process.
Cui et al. (Cui et al., 2023)	CNN– LSTM	UNSW-NB15, NSL-KDD	Accuracy, Precision, Recall, and F1 Score	The integration of complex modules may introduce computational complexity
Fu et al. (Fu et al., 2022)	Bi-LSTM	NSL-KDD	false positive rate, F1 score, recall, precision, and accuracy	Attack detection accuracy is low
Ramkumar et al. (Ramkumar et al., 2022)	DRN with Spark, Exponential Sea Lion Optimization	MQTT-IoT-IDS2020, Apache Web Server dataset	Precision, Recall, and F1 Score	Deploying and integrating a DRN-based intrusion detection system into a real-world environment can be complex.
Hagar and Gawali (Hagar and Gawali, 2022)	Apache Spark, LSTM, CNN, Random Forests	CSE-CIC-IDS2018	Confusion matrix, f1-score, accuracy	This approach failed to utilize more datasets for handling various types of attacks
Pandey et al. (Pandey et al., 2023)	ExpSSOA with Deep Maxout network	Apache Web Server, MQTT-IOT-IDS2020	Accuracy, precision, recall, and F1 score	Training time of this network is very high and may require high computational resources
Talukder et al. (Talukder et al., 2023)	Various deep and machine-learning techniques	KDDCUP'99, CIC-MalMem-2022	Accuracy, precision, recall, F1 score, RMSE, confusion matrix, and ROC	Limited analysis of various types of attacks, performing only binary classification on the CIC-MalMem-2022 dataset.

Canberra distance, class-wise information gain (CIG), and wrapper in the secondary node. After the distinguishing feature selection, over-sampling was employed during the data augmentation phase, which makes the data more suitable for use in the secondary node's further processing. Finally, the DRN classifier has been employed on the primary node, trained using the created Exponential Sea Lion Optimization method to detect intrusions.

To increase the detection of all forms of assaults in big data, Hagar and Gawali (Hagar and Gawali, 2022) suggested three models: Apache Spark, LSTM, and two CNN. Random forests (RF), which selected 19 out of 84 characteristics, were used to minimize the dimensionality and choose the key features. The dataset is unbalanced, hence approaches for over- and under-sampling were used to reduce the imbalance ratio. The CSE-CIC-IDS2018 was used for performance evaluation.

A hybrid algorithm named Exponential Shuffled Shepherded Optimization Algorithm (ExpSSOA) based intrusion detection in big data was presented by Pandey et al. (Pandey et al., 2023). This approach integrated a shuffled shepherded optimization algorithm and an exponentially weighted moving average technique. A CIG measure-based feature selection process was implemented to choose essential characteristics from the dataset. Afterward, the data augmentation was applied to increase the number of significant attributes. Finally, the ExpSSOA technique trained the Deep Maxout network to perform intrusion classification. For performance evaluation, the Apache Web Server dataset and MQTT-IOT-IDS2020 dataset were utilized.

Similarly, Talukder et al. (Talukder et al., 2023) analyzed several deep and machine-learning techniques for attack identification. Initially, the authors implemented several preprocessing techniques, such as label encoding, standardization based feature scaling, and SMOTE based augmentation, to prepare the raw input data. Afterwards, the significant features were selected by XGBoost based machine learning technique. Finally, the selected features are forwarded to Artificial Neural Network, CNN, Multilayer Perceptron, K-Nearest Neighbor, Decision Tree (DT), and RF to perform final classification. RMSE, MSE, f1-score, MAE, ROC Curve, precision, recall, and accuracy were employed for performance assessment.

3. Preliminary studies

3.1. AO algorithm

AO is a recently introduced meta-heuristic algorithm based on swarm intelligence, draws inspiration from the sophisticated foraging habits of the Aquila which lives in Northern Hemisphere. The Aquila's

hunting prowess involves a combination of breakneck flight speeds and powerful claw attacks tailored to different prey species. These techniques include Elevated Soaring with Vertical Dive, Navigating Contours with Swift Gliding Approach, attacking at Low Altitude with a Slow Declining Speed, and walking and grabbing prey. When these techniques are translated into the mathematical framework of the AO approach, the first two methods stand for the exploration stage, and the following two reflect the exploitation stage. The algorithm smoothly switches amongst local exploitation and global exploration depending on the greatest amount of iterations (T) and the current cycle (t). If $t \leqslant \left(\frac{2}{3}\right) * T$ then the exploration stages are implemented; if not, the exploitation phase is activated. The four techniques that make up the mathematical foundation of the AO algorithm are thoroughly defined in the section that follows.

Extended Exploration: Elevated Soaring with Vertical Dive

Aquila seeks its prey from the sky in its initial move, much like a bird searching for food. Once it spots the best hunting spot, it swoops down quickly to catch its prey. This action is translated into our algorithm to mimic how Aquila carefully explores and then swiftly goes in for the hunt during the exploration phase.

$$X_i(t+1) = X_b(t) \times \left(1 - \frac{t}{T}\right) + X_m(t) - X_{best}(t) \times r_1 \qquad (2)$$

In this context, $r1$ is a arbitrary number falling among 0 and 1, The $X_b(t)$ stands for the best solution found thus far; $X_i(t+1)$ represents the potential location of the i-th Aquila in the upcoming iteration $t+1$. Additionally, $X_m(t)$ indicates the mean location value derived from all individuals in the population, determined using Equation (2).

$$X_m(t) = \frac{1}{N} \sum_{i=1}^{N} X_i(t) \qquad (3)$$

Here, $X_i(t)$ signifies the position vector of the ith Aquila in the ongoing iteration t and N signifies the population size.

Focused Exploration: Navigating Contours with Swift Gliding Approach

This is Aquila's primary predation approach. After pinpointing the prey area, Aquila transitions from high-altitude soaring to a hovering stance above the target. In this poised state, it carefully scans for an opportune moment to initiate an attack. In this stage, the formula for updating the location is expressed as:

$$X_i(t+1) = X_b(t) \times LF(D) + X_r(t) + (y - x) \times r_2 \qquad (4)$$

Here, r_2 is a randomly generated number between 1 and 0; The problem's dimension size is denoted as D; $X_r(t)$ represents the position of a

randomly selected Aquila individual; LF(•) signifies the Lévy flight distribution function, defined in the following equation:

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left(\frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(1+\beta) \times \beta \times 2(\frac{\beta-1}{2})} \right)^{\frac{1}{\beta}}$$ (5)

Here, β is the constant value 1.5, (•) denotes the gamma function and v and u represents the arbitrary numbers within 0 and 1. The variables x and y in Equation (4) are used to illustrate the contour spiral shape, determined by the calculations defined in Equation (6).

$$\begin{cases} x = (R + U \times D_1) \times \sin\left(-\omega \times D_1 + \frac{3 \times \pi}{2} \right) \\ y = (R + U \times D_1) \times \cos\left(-\omega \times D_1 + \frac{3 \times \pi}{2} \right) \end{cases}$$ (6)

Here, ω = 0.005, U designates a minor value fixed at 0.00565, The integer numbers represented by $D1$ range from 1 to the dimension size (D), and A fixed amount of search steps ranging from 1 to 20, is represented by R.

Expanded Exploitation: Attacking at Low Altitude with a Slow Declining Speed

In the third strategy, Aquila's descent begins once the prey's location is approximately determined, initiating an initial attack to observe the prey's reaction. This hunting behavior is mathematically represented by Equation (7).

$$X_i(t+1) = (X_b(t) - X_m(t)) \times \alpha - r_3 + ((\text{UPB } - LRB) \times r_4 + LRB) \times \delta$$ (7)

Here, the exploitation adjustment factors, α and δ, are set at 0.1. Random numbers between 0 and 1 denoted by r3 and r4. The search domain's lower and upper boundaries are denoted by LRB and UPB, correspondingly.

Focused Exploitation: Ground Approach and Seizing Prey

During this phase, Aquila transitions to the ground, mimicking the target prey's escape trajectory in a random manner, and ultimately executes a precise attack. This behavior is described below:

$$X_i(t+1) = QF \times X_b(t) - G_1 \times X_i(t) \times r_5 - G_2 \times LF(D) + G_1 \times r_6$$ (8)

$$\begin{cases} QF(t) = t^{\frac{2 \times r_7 - 1}{(1-T)^2}} \\ G_1 = 2 \times r_8 - 1 \\ G_2 = 2 \times \left(1 - \frac{t}{T} \right) \end{cases}$$ (9)

Here, G1 and G2 are parameters governing Aquila's motion and flight slope during the pursuit of the evading prey; G2 linearly decreases from 2 to 0; G1 is a arbitrary number within − 1 to 1; the random numbers r5, r6, r7, r8 are ranging between 0 and 1; QF represents the quality function employed to regulate the search approach.

4. Proposed methodology

Due to the arrival of new big data applications, intrusion detection has been promoted as an essential duty in security-related concerns. In order to detect network breaches and improve safety, a deep learning based reliable intrusion detection mechanism is developed in this study. There are four primary phases in the suggested intrusion detection method. They are preprocessing, data augmentation, feature selection, and classification. The raw data is first put through several preprocessing procedures to prepare it for further analysis. After the preprocessing, the data augmentation is conducted based on Auxiliary Classifier Generative Adversarial Network (ACGAN) to increase the minority class samples. Then the significant features are selected from the dataset using

a hybrid feature selection technique. This hybrid approach selects the important attributes using FEMI and AO algorithm. Afterward, the selected attributes are fused using the MCCA to obtain the optimal feature set. Then, optimal features are forwarded to the ResNet152V2 network for final intrusion detection and classification. The network's hyperparameters are optimized using the WHO technique to raise the classifier's efficiency. The overall system framework of this work is depicted in Fig. 1.

The forthcoming sections will look into the suggested approach for detecting intrusions in huge data. The Section 4.1, "Preprocessing," delineates the essential stages of data cleansing and information translation, demonstrating how these procedures improve the dataset for successful deep learning model training. In the next Section 4.2, the unbalanced data issue is addressed and ACGAN-based data augmentation is introduced to create synthetic data. Using the AO and FEMI algorithms, we carried out a hybrid feature selection procedure and the features are then fused using the MCCA method in Section 4.3. In section 4.4, ResNet152v2 carried out intrusion categorization using the final fused features. Lastly, section 4.4.1 uses the WHO algorithm to adjust the ResNet152v2 hyperparameters.

4.1. Preprocessing

In the classification of big data, data preprocessing plays a crucial role, involving two interrelated steps: information transformation and data cleaning. During the data cleaning step, the redundant, irrelevant, and noisy data are removed from the massive data sets that have been trained on ToN-IoT and CICDoS2019. Afterwards Information transformation step is performed. It ensures that the data is appropriately formatted and prepared for deep learning tasks. This phase includes steps such as Label Encoding and data normalization. By transforming the information, the data becomes more conducive to effective model training, enabling algorithms to learn patterns and make accurate predictions.

In the Label Encoding procedure, the categorical values in the data set are assigned to numerical values. Then, to address the computational difficulties, the biggest and least values in the datasets are obtained and normalized using Equation (10). By supporting the normalized dataset, the values fall inside the range [0, 1].

$$v' = \frac{v - v_{\min}}{v_{\max} - v_{\min}}$$ (10)

Here, the normalized value of v is denoted by v′, and the least and highest value of the dataset is represented by $v_{\min}$ and $v_{\max}$, respectively.

4.2. ACGAN based data augmentation

In this study, we performed synthetic data augmentation using an enhanced variant of GAN termed ACGAN that enables the discriminator to be imposed by recreating auxiliary data, like predicting the classification label of a data rather than acquiring this as input. Every produced sample has noise z and appropriate class label c.

In ACGAN, the fake data produced by the generator $X_{fake} = G(c,z)$ and the discriminator D produces a distribution of probabilities across class labels and sources, denoted as P(SR|X), P(CS|X) = D(X). The log-likelihood for the correct source (LSE) and the log-likelihood for the correct class (LCS) make up the objective function of the ACGAN, which are expressed in equations (12) and (13).

$$LSE = E[\log P(SR = real|X_{real})] + E\left[\log P\big(SR = fake|X_{fake}\big)\right]$$ (12)

$$LCS = E[\log P(CS = cs|X_{real})] + E\left[\begin{array}{c}\log P(CS = \\ cs|X_{fake})\end{array}\right]$$ (13)

In Equation (12), $\log P(SR = real|X_{real})$ denotes the Log-likelihood of the discriminator correctly classifying real samples' source as real;
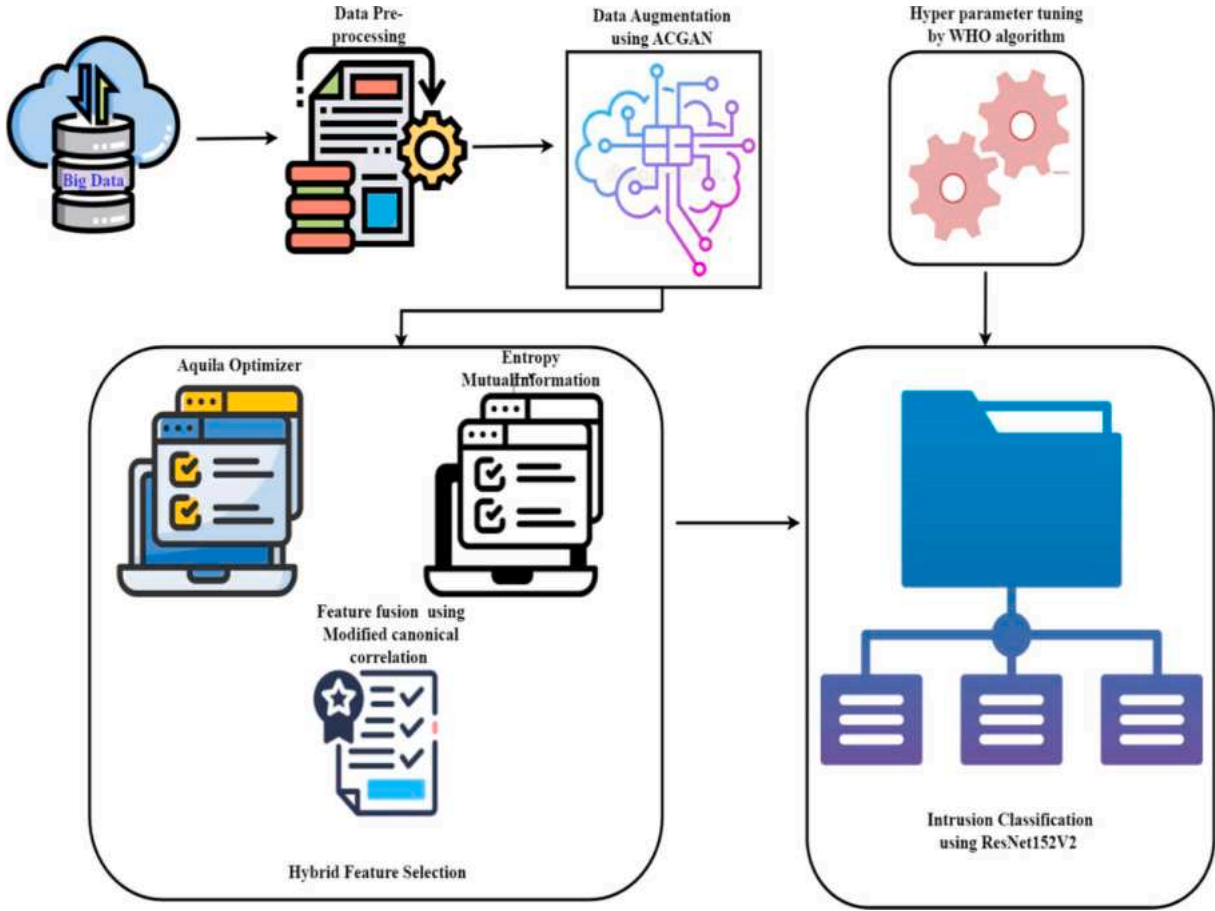
Fig. 1. System Architecture.

$\log P(SR = fake|X_{fake})$ denotes the Log-likelihood of the discriminator correctly classifying fake samples' source as fake;E[.] denotes the average over all samples.

In Equation (13), $\log P(CS = cs|X_{real})$ represents the Log-likelihood of the discriminator correctly classifying real samples with the correct class label; $\log P(CS = cs|X_{fake})$ represents the Log-likelihood of the discriminator correctly classifying fake samples with the correct class label.

The discriminator and the generator have been trained to maximize LSE + LCS and LCS-LSE, respectively. The LSE measures how well the discriminator can correctly classify whether the sample is real or fake. This helps the discriminator become proficient in identifying the source (real or generated). Also, LCS ensures that the generated samples not only look real but also carry the correct class information. It guides the generator to produce diverse and realistic samples across different classes. By incorporating these log-likelihood terms into the loss functions, the ACGAN optimizes both the discriminator and generator networks simultaneously.

Three 1D transposed convolutional layers, hyperbolic tangent (tanh) layers, two batch normalization layers, and three ReLU activation layers are the layers that make up the generator portion of the model. Four 1D convolutional layers, four batch normalization layers, and four LeakyReLU activation layers with a 0.2 slope comprise the discriminator. Every layer contains a 0.5 dropout except the final one. After flattening the output, the second layer employs the softmax function to forecast the label of the class. The sigmoid function's flattened final output is used by the first output layer to forecast the reality of the generated data. The general architecture of ACGAN is given in Fig. 2.

4.3. Hybrid feature selection approach

Selecting suitable characteristics is an essential process that increases classification accuracy and decreases computational time. Two strategies are used in this work to choose the finest features. They are the AO algorithm and FEMI. The features selected by these techniques are finally fused to obtain the best optimal features using MCCA.

4.3.1. Fuzzy entropy mutual information (FEMI)
In this subsection, FEMI based feature selection process is presented to handle the feature inconsistency. To make the final choice, the original feature vector's fuzzy entropy is first calculated and then included in the mutual information equation. Following is a definition of fuzzy entropy:

$$Fuz(Er) = -K \sum_{i=1}^{n} \{\overline{x_i}\log(\overline{x_i}) + (1 - \overline{x_i})\log p(1 - \overline{x_i})\} \quad (14)$$

Here, Fuz(Er) represents the fuzzy entropy of Er; K is constant; n denotes number of elements in vector E; x_i denotes the ith element of the vector E; $\log p(1 - \overline{x_i})$ is the probability of the complement of x_i

The following formula is used to calculate the join entropy of two variables:

$$J(Er) = \sum_{x_i,x_j} p(x_i, x_j)\log(x_i, x_j) \quad (15)$$

$$J(x_i|x_j) = -\sum_{x_i,x_j} p(x_i, x_j)\log p(x_i|x_j) \quad (16)$$

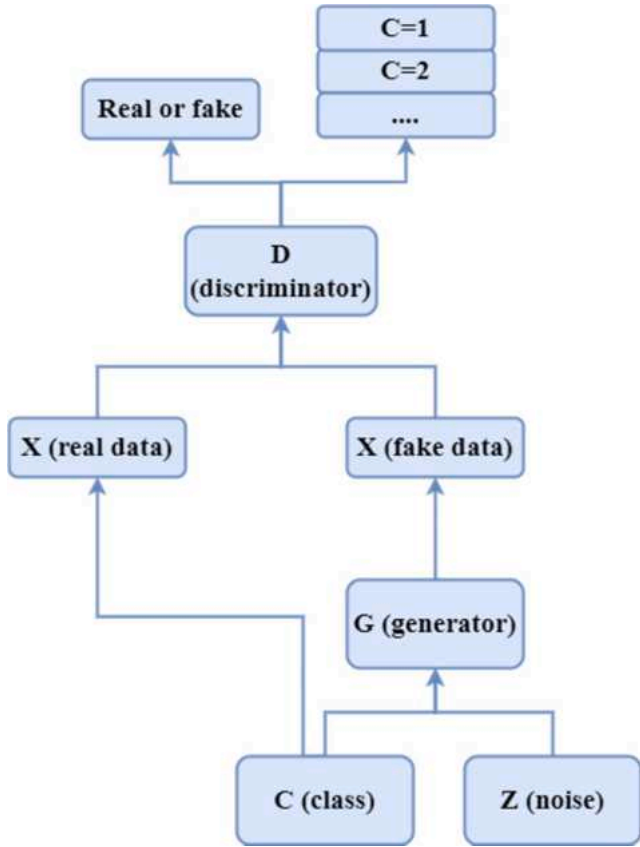Here, J(Er) denotes the Joint entropy of the variable Er; x_i,x_j are values

Fig. 2. ACGAN.

of the variables E; $p(x_i, x_j)$ represents the Joint probability distribution function of x_i, x_j; $J(x_i|x_j)$ denotes the conditional entropy of x_i, x_j

Then the mutual information between xi and xj is computed using the following Equation.

$$MI(x_i; x_j) = -\sum_{x_i, x_j} p(x_i, x_j) \log\left(\frac{p(x_i, x_j)}{p(x_i), p(x_j)}\right) \tag{17}$$

Here, $MI(x_i; x_j)$ denotes the mutual information among x_i and x_j; $p(x_i)$, $p(x_j)$ denotes the marginal probability distribution functions; $p(x_i, x_j)$ denotes the Joint probability distribution function; $\log\left(\frac{p(x_i, x_j)}{p(x_i), p(x_j)}\right)$ is the natural logarithm of the ratio of the joint probability to the product of the marginal.

Finally, Using the FEMI approach $MI(x_i; x_j)$ is the final generated vector.

4.3.2. Aquila optimizer (AO)

Using this approach, Agent N is initially distributed across x individuals. The agents in the present population are upgraded depending on the best solution of AO. The following Equation generates the initial population X, which comprises N solutions.

$$X_i = LRB + radm(1, NF) \times (UPB - LRB) \tag{18}$$

Here, radm (1, NF) stands for the random vector containing features, UPB and LRB stand for the upper and lower search space borders, and number of features represented by NF. Following that, at the step of population updating, the populations are converted into the Boolean value BX_{ij}, denoted by the following Equation.

$$BX_{ij} = \begin{cases} 1 \text{ if } X_{ij} > 0.5 \\ 0 \ otherwise \end{cases} \tag{19}$$

The characteristics with zero values are removed from the feature collection after the Boolean conversion. As a result, the feature set no longer includes unnecessary features. Then the following equations are used to calculate the fitness of each feature.

$$Fit_i = \lambda \times \gamma_i + (1 - \lambda) \times \left(\frac{|BX_i|}{D}\right) \tag{20}$$

Here, the weights are $\lambda \in [0, 1]$ included in the algorithm to regulate the balance between the appropriate features ratio $\left(\frac{|BX_i|}{D}\right)$ and γ_i error in the classification. The best Fit and its related agent Xb are then identified. At this stage, the stopping conditions are evaluated; if they are not satisfied, the updated stage is repeated. Without this, Xb is utilized as the output of the selected features, and iteration is finished.

Finally, the resultant feature vector Xb is fused with the obtained FEMI based feature selection vector using MCCA to obtain the best optimal feature set.

4.3.3. MCCA based feature fusion

In this algorithm, the most compelling features of FEMI and AO are eventually combined into one matrix. The original CCA is expressed mathematically as,

$$\{x_i\}_{i=1}^n \in S^j, \{y_j\}_{j=1}^n \in S^k, \{z_k\}_{k=1}^n \in S^l, \tag{21}$$

Here, n is the observation size, whereas l, k, and j represent the sample space's dimensions. Projections directories are what MCCA is looking for.

$$a_x \in S^j, a_y \in S^k, a_z \in S^l \tag{22}$$

This increases the relationship among the sample matrices $a_x^T X, a_y^T Y, a_z^T Z$. here; the sample matrices are Z = [z$_1$, z$_2$, … z$_n$], Y = [y$_1$, y$_2$, … y$_n$] and X = [x$_1$, x$_2$, … x$_n$]. In mathematical terms, it is resolved by,

$$\rho = \max \frac{a_x^T a_y Z_{xyz} a_z}{\sqrt{(a_x^T Z_{xx} a_x)(a_y^T Z_{yy} a_y)(a_z^T Z_{zz} a_z)}} \tag{23}$$

Here, $Z_{xx} = XX^T$, it $Z_{yy} = YY^T$, $Z_{zz} = ZZ^T$ reflects covariance within three feature sets, and the covariance matrix between feature sets is defined by $Z_{xyz} = XYZ^T$. Generalized Eigen-problem computation can determine MCCA when the matrices contained in feature sets are not singular.

$$\begin{bmatrix} Z_{xyz} Z_{zz}^{-1} Z_{zxy} & 0 & 0 \\ 0 & Z_{zxy} Z_{yy}^{-1} Z_{yzx} & 0 \\ 0 & 0 & Z_{yxz} Z_{xx}^{-1} Z_{xzy} \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix}$$
$$= \lambda \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \tag{24}$$

Let's define three projection directories matrices as A$_x$ = [a$_{x1}$, a$_{x2}$, … a$_{xn}$], A$_y$ = a$_{y1}$, a$_{y2}$, … a$_{yn}$, A$_z$ = [a$_{z1}$, a$_{z2}$, … a$_{zn}$]. In this case, the vector pairs $(a_{xi}, a_{yi}, a_{zi})_{i=1}^d$ match the highest d generalized Eigenvalue. The fused feature, which is described below, can be obtained from the three modalities:

$$F(i) = \begin{bmatrix} A_x^T x \\ A_y^T y \\ A_z^T z \end{bmatrix} \tag{25}$$

In this case, the fused vector F (i) was arranged into decreasing order, and redundant characteristics were removed using a comparative approach. The final results features are finally transferred to ResNet152 based classifier to perform intrusion detection.

4.4. ResNet152v2 based classification

A more recent member of the ResNet family, ResNet152V2 outperforms ResNet152V1. The main concept of the V2 model is the introduction of a shortcut path that permits data to be sent not just among residual blocks but also across the whole network. This model, with 152 layers, is aptly named ResNet152V2. To effectively classify attacks, the network includes components like a dense layer with forged frames, a flatten layer, a dropout layer, reshape layer, and a softmax activation with a dense layer. The depth of this deep network poses a challenge in appropriately adjusting the weights between layers, given the layer growth. To overcome this, the network employs a small parallel link through regular convolution layers, addressing the issue of performance degradation.

ResNet152V2 unfolds in four sequential stages, each characterized by a specific number of repetitive residual blocks. The initial stage, Stage 1, encompasses three repeated residual blocks, with each block featuring three 1D convolution layers. Moving to Stage 2, the network incorporates eight recurring residual blocks, followed by a substantial expansion to 36 repetitive residual blocks in Stage 3. The final stage, Stage 4, concludes with three repetitive residual blocks. The backbone of ResNet152V2 lies in its residual blocks, governed by Equation (26), which serves as the foundational equation guiding the architecture's construction and ensures the seamless flow of information through the network.

$$y_1 = f(x_1, w_1) + h(x_1), x2 = f(y1) \tag{26}$$

Here, the residual function is denoted by f; the residual unit weight is denoted by $w_1, w_2 \ldots$ i, and ith residual unit are represented by $x_1, x_2, \ldots$ i. If $x_2 \equiv y_1$ then

$$x_2 = x_1 + f(x_1, w_1) \tag{27}$$

$$x_3 = x_2 + f(x_2, w_2) = x_1 + f(x_1, w_1) + f(x_2, w_2) \tag{28}$$

$$x_4 = x_3 + f(x_3|w_3) = x_1 + f(x_1, w_1) + f(x_2, w_2) + f(x_3, w_3) \tag{29}$$

$$x_i = x_1 + \sum k = 1^{i-1} f(x_k, w_k) \tag{30}$$

The backpropagation process is affected by Equation (30) which is presented in the following Equation.

$$\frac{\partial \phi}{\partial x_1} = \frac{\partial \phi}{\partial x_i} \cdot \frac{\partial x_i}{\partial x_1} = \frac{\partial \phi}{\partial x_i} \left(1 + \frac{\partial}{\partial x_1} \sum k = 1^{i-1} f(x_k, w_k)\right) \tag{31}$$

Here, the loss function is denoted by ϕ. Based on the above two equations (Eq. (30) and (31)), it is identified that the signal is easily forwarded in both directions. Furthermore, the bi-directional gated recurrent unit (Bi-GRU) is implemented in this study to reduce the model's sequential difficulty. The concealed state of forward and backward travels are combined as the following reserved representation J_j^λ in the Bi-GRU.

$$J_j^\lambda = \left[GR\overset{\mu}{U}(\sigma_j); GR\overset{\smile}{U}(\sigma_j) \right] \tag{32}$$

In the above equation, σ_j is computed based on weights and token $t = \sum_{y=1}^{Y} 2^y$ s, expressed in the following Equation.

$$\sigma_j = \sum_{y=1}^{Y} \frac{b_y J^y j}{t} \tag{33}$$

In this case, the weights of the yth layer are denoted by b_y

The integration of Bi-GRU in conjunction with ResNet152V2 is a strategic move to mitigate data noise during processing. By considering both the preceding and succeeding sequences of data, Bi-GRU contributes to an efficient and prolonged operation, accelerating the model's performance within a short timeframe. The collaborative approach of ResNet152V2 and Bi-GRU empowers the model to discern and categorize network attacks, culminating in an output generated through the final layer of the architecture. To further optimize the network's functionality, the Wildebeest Herd Optimization (WHO) method is employed to fine-tune hyperparameters, thereby increasing the whole performance of the classifier.

4.4.1. Parameter optimization using wildebeest Herd Optimization (WHO) algorithm

We have optimized our classifier network using the WHO algorithm. The WHO algorithm was based on the wildebeests' behavior when seeking food. The friendly, versatile mammal known as the wildebeest hunts for food. To entice females to reproduce, males engage in competition with opponents. It begins by randomly initializing a number of populations (wildebeests) as candidates. In this case, the population is constrained between the lower ($W_{\min}$) and upper ($W_{\max}$) borders.

$$W_i \in [W_{\min}, W_{\max}] \tag{34}$$

Here, i = 1, 2...N

During the local milling process, the wildebeest engage in strategic movements within their immediate surroundings. This phase is emulated by a search for an optimal position, incorporating a consistent yet small random movement denoted by the variable "n" across solution spaces. Candidates within the location "W" utilize an arbitrary position "Zn," actively exploring minute random stages. The introduction of a random step size, adjustable in length, contributes to the adaptability of this local exploration. The formula employed for determining the local experimental phase "Z" encapsulates the nuanced dynamics of this localized movement, enhancing the algorithm's ability to navigate the solution landscape effectively.

$$Z_n = W_i + \varepsilon \times \theta \times v \tag{35}$$

Here, a random uniform value among 0 and 1 is represented as θ the random unit vector is designated as v, the contestant number is represented as W_i, and the ε denotes the learning rate.

In response to its surroundings, the wildebeest adapts its position, seeking an optimal and randomized location within its vicinity.

$$W_i = \alpha_1 \times Z_n^* + \beta_1 \times (W_i - Z_n^*) \tag{36}$$

Here, the candidate's local movement is directed by the leader variables α_1 and β_1

Modeling the wildebeests' swarm instinct is the final step. Once the other candidates are set up in a spot with the right kind of food source, it is activated. That is described in the following Equation.

$$W_i = \alpha_2 \times W_i + \beta_2 \times W_h \tag{37}$$

Here, a random candidate denoted by W_h

To avoid candidates venturing into regions with limited resources, an additional term is incorporated into the WHO algorithm, introducing a mathematical representation that guides candidates away from such areas:

$$W_i = W_i + \theta \times (W_{\max} - W_{\min}) \times \bar{v} \tag{38}$$

Here $\bar{v}$ is the vector of random units.

Simulating a crowded region is the next phase of this algorithm. When there is a large scale of grass available, there is a crowd. The phase is known as "individual pressure." The strongest competitor applies the following formula to annihilate the others during the competition phase:

$$if(\|W^* - W_i\|) < \eta, (\|W^* - W_i\|) > 1 \tag{39}$$

Then

$$W_i = W^* + \varepsilon \times \widehat{n} \qquad (40)$$

Here η represents a cutoff to avoid overcrowding in the area and the quantity of exploitable sectors adjacent to the optimal location denoted by $\widehat{n}$.

For better placements, the final step simulates the swarm social memory. The following formula is used to obtain it.

$$W = W^* + 0.1 \times \widehat{v} \qquad (41)$$

At last, the ideal optimal result of the WHO method is used to initialize the hyperparameters of the ResNet152V2, such as learning rate, epoch, batch size, weight decay rate, momentum, and dropout rate.

5. Result and discussion

This part describes the investigational examination of the presented work and the performance measures applied to evaluate the classification effectiveness of our suggested method. The research used the Pyspark tool, which supports the Python programming language on the Apache Spark big data framework within the Google Colab setting. All testing used Windows 10 64-bit, a Core i7 processor running at 2.70 GHz, 16 GB of RAM, and the Python programming language. The dataset is distributed into testing and training portions of 30 % and 70 %, respectively, and the 0.001 learning rate, weight decay rate at 1e-06, momentum at 0.8, batch size at 8, and dropout rate at 0.9 were optimized using the WHO method.

5.1. Dataset description

CICDDoS2019 dataset: The Canadian Institute for Cybersecurity (CIC) has made the CIC-DDoS 2019 dataset available for use. It was created in a suitable test environment and contains the findings of actual internet traffic reports. The 30,480,823 records in the CICDDoS2019 dataset are divided into 30,423,960 recordings of DDoS attacks and 56,863 records of benign activity. In addition, twelve different types of DDoS attacks are separated from regular network traffic. Eighty-six features are used to describe every record. It is produced using more sophisticated technologies and has a wider variety of DDoS attack traffic. This dataset is accessible as PCAP and CSV files with labeling.

ToN-IoT dataset: It is designed to gather and examine diverse forms of IIoT and IoT data. It includes the information gathered from various resources, such as monitoring data from linked gadgets, Windows and Linux log files, and computer traffic from the network. Additionally, it includes concurrent sets of legal and illegal actions in network architecture, software platforms, and IoT applications. This dataset is available in CSV with nine categories of intrusions (injection, backdoors, ransomware, MITM, reconnaissance, password cracking attacks, DoS, DDoS, and XSS).

5.2. Performance metrics

Using standard evaluation metrics like F-score, Recall, Precision, and Accuracy, the efficacy of the suggested intrusion detection method is assessed.

$$Accuracy = \frac{tp + tn}{tp + fp + tn + fn} \qquad (42)$$

$$precision = \frac{tp}{tp + fp} \qquad (43)$$

$$recall = \frac{tp}{fn + tp} \qquad (44)$$

$$f1 - score = 2 \times \frac{precision \times recall}{precision + recall} \qquad (45)$$

The stability among the true positive and false positive rate is plotted using the area under the receiver operating characteristics (ROC) Curve (AUC-ROC) based on the following Equation.

$$AUC_{roc} = \int_0^1 \frac{tp}{tp + fn} d\,\frac{fp}{tn + fp} \qquad (46)$$

The symbols fn, fp, tn, and tp denotes for false negative, false positive, true negative, and true positive, correspondingly. The recall can be utilized to assess how accurate the suggested classifier is by dividing the proportion of true positives from the total positive events. High precision and recall indicate the presence of necessary true positives and true negatives. The F-score represents the harmonic mean of recall and precision.

5.3. Performance evaluation on CICDDoS2019 dataset

The findings from the thorough analysis of the suggested method on the CICDDoS2019 dataset are shown here. The DoS and DDoS are basic but successful attack types that have become highly popular among network attackers. As a result, network security researchers are paying more attention to them than other kinds of attacks. On the CICD-DoS2019 dataset, a comparative analysis of various DDoS attack types is shown in Table 2.

Our research shows that the suggested method performs best at detecting "Benign" traffic (99.98 %) and poorly at detecting "WebDDoS" traffic (98.82 %). Due to the "WebDDoS" sample being extremely rare across the entire dataset, our model performs poorly. Moreover, the behavior pattern of traffic type "UDP-Lag" is similar to that of regular network traffic; it is more challenging for our model to correctly identify "UDP-Lag" attacks, which achieves only 98.97 % accuracy compared to other techniques.

Table 3 contrasts the effectiveness of the suggested strategy with several IDS models from a big data perspective. We introduced an enhanced deep learning technique based on ResNet152V2 to categorize incoming traffic into good and harmful categories. The suggested model was evaluated and contrasted with the following models that were also used on the same dataset: Random forest, energy-based Online Sequential Extreme Learning Machine (e-b OSELM), Optimised LSTM, LightGBM, and Kafka Streams-based DDoS (KS-DDoS).

According to Table 3 and Fig. 3, our approach performed better than existing techniques regarding several standard metrics. The proposed deep learning technique has the potential to detect intrusions effectively due to its ability to handle a high degree of complex nonlinear interactions. It might be applied to get beyond the drawbacks of traditional classification methods, which rely on classical feature encoding to find abnormal traffic. The Random forest model showed inferior F1-score, recall, precision, and accuracy compared to all other methods.

Table 2
Comparison of various attack categories on CICDDoS2019.

Intrusion Categories	F1-score	Recall	Precision	Accuracy
Benign	99.93	99.91	99.95	99.89
WebDDoS	99.32	99.13	99.51	98.82
UDP-Lag	99.42	99.25	99.59	98.97
NTP	99.87	99.84	99.9	99.81
LDAP	99.84	99.8	99.89	99.78
SSDP	99.5	99.38	99.63	99.19
UDP	99.66	99.55	99.77	99.39
NetBIOS	99.6	99.47	99.73	99.27
MSSQL	99.73	99.63	99.84	99.49
SNMP	99.83	99.79	99.88	99.75
TFTP	99.79	99.75	99.83	99.67
DNS	99.77	99.73	99.81	99.52

Table 3
Comparative Analysis of the CICDDoS2019 dataset.

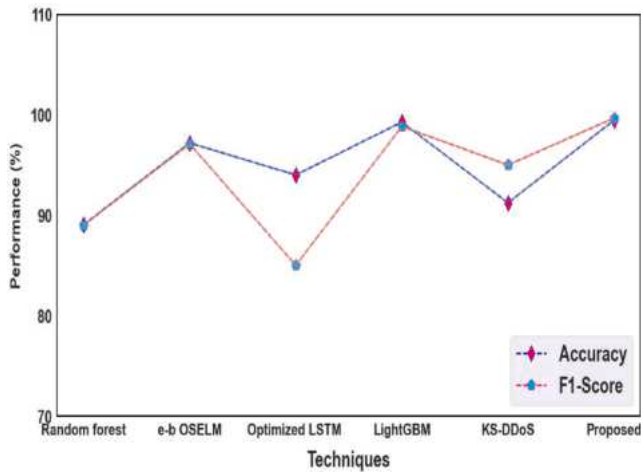| Techniques	AUC	F1-score	Recall	Precision	Accuracy
Random forest (Patil et al., 2022)	–	89	89	89.14	89.05
e-b OSELM (Wang et al., 2022)	0.985	97.1	–	–	97.19
Optimized LSTM (Packialatha, 2023)	–	85.02	85.02	85.02	94.01
LightGBM (Rani et al., 2023)	0.995	98.83	99.26	99.71	99.27
Kafka Streams-based DDoS (KS-DDoS) (Patil et al., 2022)	–	0.95	0.92	0.98	91.23
Proposed	0.998	99.68	99.60	99.77	99.46

Fig. 3. Comparison of accuracy and F1-score on the CICDDoS2019 dataset.

Moreover, LSTM's f1-score, recall, and precision values are lower than other models. Because future contextual information that would aid in understanding the meaning of the examined text is not retained by this method; instead, it is ignored. As a result, among all the models, it performed the worst.

In order to examine the stability among the False Positive Rate and True Positive Rate of all attack categories, ROC curves are created on the CICDDoS2019 datasets, depicted in Fig. 4. The chart shows that the outcomes of our suggested approach are close to the ideal position,
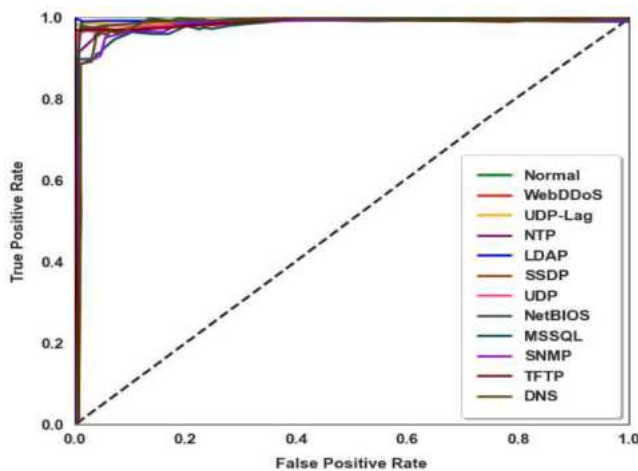
where the true positive's number is one, and the false positive's number is zero. This indicates that the suggested technique can accurately classify 99.8 % of DDoS and regular classes. This shows that the hybrid approach can effectively identify and classify malicious payloads that incorporate DDoS attacks.

The confusion matrix-based classification for various assault types is displayed in Fig. 5, along with the precise number of predictions. On the cicddos2019 dataset, the proposed approach accurately identifies benign and various attack types, and the tp and tn rates are quite high. In contrast, fp and fn rates remain very low (i.e., are marked by various color squares). Compared to other samples, the benign samples have a considerably higher proportion, although there is no color square surrounding tn cases in those samples. We also see the 'WebDDOS' attack type perform poorly because its test samples are much fewer than those for other classes.

Additionally, we discover through experiments that the attributes of the "WebDDoS" and the "SSDP" share similar qualities. Due to this, certain WebDDoS data are mistakenly labeled as "SSDP" by the proposed approach, and we can observe that fp and fn are significantly highlighted for this attack type. However, this has no impact on the performance of the overall result.

5.4. Performance analysis on ToN-IoT dataset

To analyze the performance of the proposed approach, several examinations are conducted on the ToN-IoT dataset, and the performance result for individual attack types is shown in Table 4. The "XSS" and "scanning" attack types contained the least quantity of IoT device data compared to the other attack types. However, the suggested model could still detect them with a detection rate of 98.87 % (XSS) and 98.92 % (Scanning) accuracy due to our ACGAN based data augmentation technique.

The performance study of the suggested model on the ToN-IoT dataset is shown in Table 5, along with comparing its performance with five leading models. The table shows that the proposed model performs with 99.22 % accuracy better than the others. This is because the suggested hybrid feature selection technique offers more robust optimized features than existing approaches. Additionally, an optimized classifier outperforms a single classifier in terms of accuracy. Among other models, Convolutional Neural Networks (CNN) and Artificial Neural Networks (ANN) obtained the second and third greatest accuracy with 98.87 % and 98 %, respectively. Furthermore, CNN and ANN achieved the second and third highest accuracy with 98.87 % and 98 %, respectively, than other existing models. However, it is not more than our proposed approach.

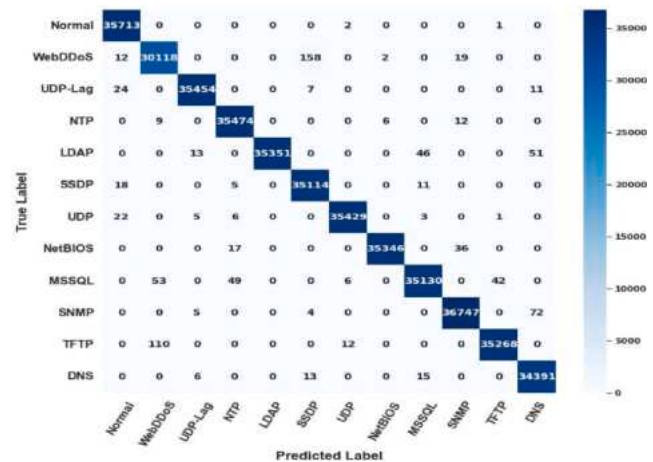Table 5 and Fig. 6 display the Deep Neural Network (DNN) model's

Fig. 4. ROC curve on the CICDDoS2019 dataset.

Fig. 5. Confusion matrix on the CICDDoS2019 dataset.

Table 4
Comparison of various attack categories on ToN-IoT.

Intrusion Categories	F1-score	Recall	Precision	Accuracy
Benign	99.68	99.71	99.65	99.75
Backdoor	98.95	98.98	98.92	99.05
DoS	99.57	99.61	99.53	99.65
DDoS	99.41	99.43	99.4	99.47
Injection	99.26	99.29	99.24	99.33
Scanning	98.88	98.9	98.86	98.92
MITM	99.21	99.22	99.2	99.26
Ransomware	99.09	99.11	99.08	99.14
Password	99.13	99.15	99.11	99.19
XSS	98.83	98.84	98.82	98.87

Table 5
Comparative Analysis of the ToN-IoT Dataset.

Techniques	AUC	F1-score	Recall	Precision	Accuracy
KNN (Thaseen et al., 2021)	94.95	98.10 %	96.89 %	96.77 %	97.73 %
ANN (Komisarek et al., 2021)	0.9681	98	98	98	98
DNN (Vishwakarma and Kesswani, 2022)	–	61.96 %	69.53 %	56.84	69.53 %
XGBoost (Awad et al., 2022)	–	87.36	84.81	83.18	94
CNN (Ferrag et al., 2021)	–	82.5	83.3	81.6	98.87
Proposed	0.992	99.20	99.22	99.18	99.26

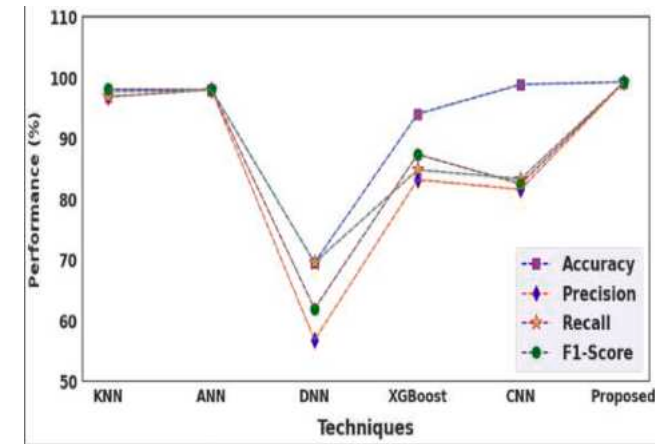**Fig. 6.** Comparison of accuracy, precision, recall, and f1-score.

unsatisfactory precision, recall, F1-score, and accuracy results. DNN models would be unable to remember information for very long periods since they cannot handle very long-term sequencing. As a result, the DNN model generated less than ideal-results. Additionally, if the data contains many characteristics, the model gets more complicated and tends to overfit the training set. Therefore, this study uses an efficient hybrid feature selection technique to maintain the most crucial aspects while eliminating unnecessary features. As a result, the model will be more accurate.

The AUC values on the ToN-IoT dataset are also analyzed to assess the performance of the suggested model, which are depicted in Fig. 7. A performance indicator for classification issues at different threshold levels is the ROC curve. A higher AUC number suggests greater accuracy in predictions. Fig. 7 shows the x-axis and y-axis for the erroneous and accurate categorization rates, respectively. We utilize the one-vs.-rest technique to show the ROC curve and AUC value for every category in a multi-class classification strategy. All AUC values are higher than 0.9,

Fig. 7. ROC Curve for ToN-IoT dataset.

confirming that the suggested model yields superior classification outcomes.

Fig. 8 shows the confusion matrix for the suggested technique on the TON-IOT dataset. The summarized output of the incursions that our suggested classifier correctly or incorrectly identified is known as the confusion matrix in classification. This matrix offers precise information on the performance of the suggested model. It is clear from the graphic that the suggested strategy correctly categorizes the majority of the occurrences from the testing set.

Fig. 8 also demonstrates a more precise classification of the data affected by different assaults. The true negative (with intrusion) and true positive (without intrusion) rates categorized by the suggested approach are shown in the values in the confusion matrix's diagonal; in contrast, other columns and rows indicate the classified data's false positive and false negative values. Misclassifying DDoS as Injection, Injection as DoS and DDoS, and XSS as password provide some false classifications. However, each assault type was correctly classified most of the time, so these misclassifications do not affect the classifier's overall performance. The uppermost true negative and true positive values are displayed in that matrix's diagonal values. The diagonal values in the matrix display the high true negative and true positive values. Therefore, we may conclude that the suggested strategy successfully categorizes intrusion detection on big data.
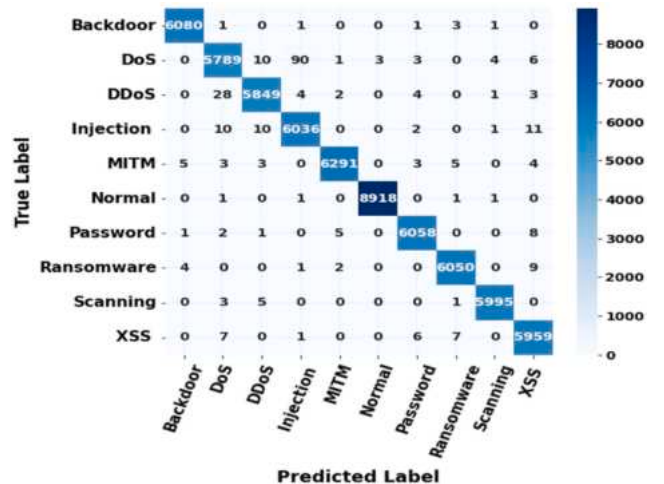

Fig. 8. Confusion matrix of the ToN-IoT dataset.

5.5. Conclusion

This study represents a significant stride in increasing the reliability of intrusion detection in the challenging domain of big data. Our novel framework presented a powerful deep learning method based on ResNet152, demonstrating its capacity to precisely classify incursions. To effectively tackle imbalanced datasets, minority class data are boosted by ACGAN-driven data augmentation. Another significant effort is the hybrid feature selection method, which combines the AO algorithm with FEMI to ensure that relevant qualities are found and included. For the best feature fusion using the MCCA technique, this creates the prerequisites. In experimental evaluation, our proposed model demonstrates exceptional performance on the CICDDoS2019 and ToN-IoT datasets, outperforming previous intrusion detection strategies with impressive accuracy rates of 99.46 % and 99.26 %, respectively. These results underscore the efficacy and superiority of our developed algorithm. Future enhancements may involve creating a hybrid deep learning-based method to address processing challenges in big data intrusion detection, potentially expanding coverage to include other attack types like Trojan, rootkit, etc. for classification.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Wang, Y. C., Houng, Y. C., Chen, H. X., & Tseng, S. M. (2023). Network Anomaly Intrusion Detection Based on Deep Learning Approach. *Sensors, 23*(4), 2171.

Dasgupta, S., & Saha, B. (2022). HMA-ID mechanism: A hybrid mayfly optimization based apriori approach for intrusion detection in a big data application. *Telecommunication Systems, 80*(1), 77–89.

Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+ LSTMbased intrusion detection system for industrial IoT networks. *Engineering Science and Technology, an International Journal, 38*, Article 101322.

Ramya, S., Sakthi Devi, R., Senthil Pandian, P., Suguna, G., Suganya, R., & Manimozhi, N. (2023). Analyzing Big Data challenges and security issues in data privacy. *International Research Journal of Modernization in Engineering Technology and Science, 5*, 421–428.

Gu, S., Liu, C., Li, Y., & Kan, X. (2023). Optimization Strategy of Computer Network Security Technology in Big Data Environment. *Frontiers in Business, Economics, and Management, 7*(1), 1–4.

Ponmalar, A., & Dhanakoti, V. (2022). Hybrid Whale Tabu algorithm optimized convolutional neural network architecture for intrusion detection in big data. *Concurrency and Computation: Practice and Experience, 34*(19), e7038.

Sarkar, A., Sharma, H. S., & Singh, M. M. (2023). A supervised machine learning-based solution for efficient network intrusion detection using ensemble learning based on hyperparameter optimization. *International Journal of Information Technology, 15*(1), 423–434.

Al Moaiad, Y., Tarshany, Y. M. A., Algeelani, N. A., & Al-Haithami, W. (2022). Cyber Attack Detection Using Big data analysis. *International Journal of Computer Science and Information Technology Research, 3*(10), 26–33.

Ragab, M., Alshammari, S. M., & Al-Ghamdi, A. S. (2023). Modified Metaheuristics with Weighted Majority Voting Ensemble Deep Learning Model for Intrusion Detection System. *Computer Systems Science & Engineering, 47*(2).

Pustokhina, I. V., Pustokhin, D. A., Lydia, E. L., Garg, P., Kadian, A., & Shankar, K. (2022). Hyperparameter search based convolution neural network with Bi-LSTM model for intrusion detection system in the multimedia big data environment. *Multimedia Tools and Applications, 81*(24), 34951–34968.

Ahakonye, L. A. C., Nwakanma, C. I., Lee, J. M., & Kim, D. S. (2023). SCADA intrusion detection scheme exploiting the fusion of modified decision tree and Chi-square feature selection. *Internet of Things, 21*, Article 100676.

Alhabshy, A. A., Hameed, B. I., & Eldahshan, K. A. (2022). An Ameliorated Multiattack Network Anomaly Detection in Distributed Big Data System-Based Enhanced Stacking Multiple Binary Classifiers. *IEEE Access, 10*, 52724–52743.

Venkatesan, S. (2023). Design an Intrusion Detection System based on Feature Selection Using ML Algorithms. *Mathematical Statistician and Engineering Applications, 72*(1), 702–710.

Awotunde, J. B., Ayo, F. E., Panigrahi, R., Garg, A., Bhoi, A. K., & Barsocchi, P. (2023). A Multi-level Random Forest Model-Based Intrusion Detection Using Fuzzy Inference System for Internet of Things Networks. *International Journal of Computational Intelligence Systems, 16*(1), 31.

Polepally, V., Jagannadha Rao, D. B., Kalpana, P., & Nagendra Prabhu, S. (2022). Exponential Squirrel Search Algorithm-Based Deep Classifier for Intrusion Detection in Cloud Computing with Big Data Assisted Spark Framework. *Cybernetics and Systems*, 1–20.

Zhang, Z. (2022). Class Image Processing Application of XSS Intrusion Intelligent Detection for Big Data in Campus Network Construction. Security and Communication Networks, 2022.

Yin, Y., Jang-Jaccard, J., Xu, W., Singh, A., Zhu, J., Sabrina, F., & Kwak, J. (2023). IGRF-RFE: A hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 Dataset. *Journal of Big Data, 10*(1), 1–26.

Adewole, K. S., Salau-Ibrahim, T. T., Imoize, A. L., Oladipo, I. D., AbdulRaheem, M., Awotunde, J. B., & Aro, T. O. (2022). Empirical Analysis of Data Streaming and Batch Learning Models for Network Intrusion Detection. *Electronics, 11*(19), 3109.

Park, N. E., Lee, Y. R., Joo, S., Kim, S. Y., Kim, S. H., Park, J. Y., & Lee, I. G. (2023). Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks. *Computers and Electrical Engineering, 105*, Article 108548.

Aldwairi, M., & Alansari, D. (2022). n-Grams exclusion and inclusion filter for intrusion detection in Internet of Energy big data systems. *Transactions on Emerging Telecommunications Technologies, 33*(3), e3711.

Ponmalar, A., & Dhanakoti, V. (2022). An intrusion detection approach using ensemble support vector machine based chaos game optimization algorithm in big data platform. *Applied Soft Computing, 116*, Article 108295.

Cui, J., Zong, L., Xie, J., & Tang, M. (2023). A novel multi-module integrated intrusion detection system for high-dimensional imbalanced data. *Applied Intelligence, 53*(1), 272–288.

Fu, Y., Du, Y., Cao, Z., Li, Q., & Xiang, W. (2022). A deep learning model for network intrusion detection with imbalanced data. *Electronics, 11*(6), 898.

Ramkumar, M. P., Reddy, P. B., Thirukrishna, J. T., & Vidyadhari, C. (2022). Intrusion detection in big data using hybrid feature fusion and optimization enabled deep learning based on spark architecture. *Computers & Security, 116*, Article 102668.

Hagar, A. A., & Gawali, B. W. (2022). Apache Spark and Deep Learning Models for High-Performance Network Intrusion Detection Using CSE-CIC-IDS2018. Computational Intelligence and Neuroscience, 2022.

Pandey, B. K., Veeramanickam, M. R. M., Ahmad, S., Rodriguez, C., & Esenarro, D. (2023). ExpSSOA-Deep maxout: Exponential Shuffled shepherd optimization based Deep maxout network for intrusion detection using big data in cloud computing framework. *Computers & Security, 124*, Article 102975.

Talukder, M. A., Hasan, K. F., Islam, M. M., Uddin, M. A., Akhter, A., Yousuf, M. A., & Moni, M. A. (2023). A dependable hybrid machine learning model for network intrusion detection. *Journal of Information Security and Applications, 72*, Article 103405.

Patil, N. V., Krishna, C. R., & Kumar, K. (2022). SSK-DDoS: Distributed stream processing framework based classification system for DDoS attacks. *Cluster Computing*, 1–18.

Wang, X., Tu, S., Zhao, W., & Shi, C. (2022). A novel energy-based online sequential extreme learning machine to detect anomalies over real-time data streams. *Neural Computing and Applications, 34*(2), 823–831.

Packialatha, A. (2023). Hybrid classification model with tuned weight for cyber-attack detection: Big data perspective. *Advances in Engineering Software, 177,* Article 103408.

Rani, S. J., Ioannou, I., Nagaradjane, P., Christophorou, C., Vassiliou, V., Charan, S., & Pitsillides, A. (2023). Detection of DDoS attacks in D2D communications using machine learning approach. *Computer Communications, 198*, 32–51.

Patil, N. V., Krishna, C. R., & Kumar, K. (2022). KS-DDoS: Kafka streams-based classification approach for DDoS attacks. *The Journal of Supercomputing*, 1–31.

Thaseen, I. S., Mohanraj, V., Ramachandran, S., Sanapala, K., & Yeo, S. S. (2021). A hadoop based framework integrating machine learning classifiers for anomaly detection in the Internet of Things. *Electronics, 10*(16), 1955.

Komisarek, M., Pawlicki, M., Kozik, R., Hołubowicz, W., & Choraś, M. (2021). How to Effectively Collect and Process Network Data for Intrusion Detection? *Entropy, 23*(11), 1532.

Vishwakarma, M., & Kesswani, N. (2022). DIDS: A Deep Neural Network based real-time Intrusion detection system for IoT. *Decision Analytics Journal, 5*, Article 100142.

Awad, M., Fraihat, S., Salameh, K., & Al Redhaei, A. (2022). Examining the Suitability of Net Flow Features in Detecting IoT Network Intrusions. *Sensors, 22*(16), 6164.

Ferrag, M. A., Friha, O., Maglaras, L., Janicke, H., & Shu, L. (2021). Federated deep learning for cyber security in the internet of things: Concepts, applications, and experimental analysis. *IEEE Access, 9*, 138509–138542.

A STUDY ON BIOMETRIC VERIFICATION BASED CRIME IDENTIFICATION

Dr. D. SUJA MARY., M.Sc.,M.Phil., Ph.D Assistant Professor, Department of Computer Science, Nanjil Catholic Arts & Science College Kaliyakkavilai, KanyaKumari Dist.

Dr. M. SURIAKALA., M.Sc.,M.Phil.,Ph.D., Research supervisor and Assistant Professor, Department of Computer Science Government Arts College for Men(Autonomous) Nandanam, Chennai -35.

ABSTRACT
Availability of relevant and timely information is of utmost importance in conduct of business by Police, particularly in investigation of crime and in tracking & detection of criminals. Police organizations everywhere have been handling large amounts of information and huge volume of records pertaining to crime and criminals. Information Technology (IT) can play a very vital role in improving outcomes in the areas of Crime Investigation and Criminals Detection and other functioning of the Police organizations, by facilitating easy recording, retrieval, analysis and sharing of the pile of Information. Quick and timely information availability about different facets of Police functions to the right functionaries can bring in a sea change both in Crime & Criminals handling and related Operations, as well as administrative processes. Bio Enable offers a range of criminal identification products that make use of Biometric data and help police forces at the crime scene. In this paper, Fingerprint based biometric verification system is used for crime scene is important clues to solve serial cases. Further the system can be used by the people to register the complaints and is helpful to the police department in identifying the criminals. The main purpose of the application is to improve the effectiveness and efficiency of interaction procedures between the police officials and common people. It would be an outstanding tool to monitor and track the criminals around the country and also have a complete online record of crime related information.

Keywords: *Information Technology, criminals investigation, fingerprint*

INTRODUCTION
Rules and regulations are paramount to all aspects of life and it accommodates both how one wishes to live, and how others should accommodate one's lifestyle. Certain proponents have asserted that crime which is a violation against laws of the society, is integral to the human nature and hence the society can never be completely free from it. Modern society is characterized by increasing levels of risk posed by internal and external security threats. Within this context, security driven by technology is increasingly being used by government, corporate bodies and individuals to monitor and reduce risk.

Forensics techniques are being used in the investigation of criminal activities as traditional methods. "Forensic science" begins with the effective identification, documentation (collection of notes, photographs, sketching and videos of crime scene), collection and preservation of physical (covers items of non-living origin such as fingerprints, footprints, fibers, paint, tire or shoe impression and weapons) and biological evidence (originates from a living source and includes DNA, other bodily fluids, hair, skin and bone material) at the crime scene. The evidence is then subjected to scientific analysis in the forensic laboratory and the results of the examinations yield forensic evidence for consideration by court. Ultimately, the evidence will be presented as proof that a crime was committed and will prove the identification of the criminal
  Biometrics is one of the most fascinating ways to solve the crime. It is an automated way to establish the identity of a person on the basis of his or her physical finger print, face, hand/finger geometry, iris, retina, ear, etc.) and behavioral characteristics (signature, voice, gait, odor, etc.). Biometric technology makes a contribution to crime detection by associating the traces to the persons stored in the database, ranking the identity of persons and selecting subdivision of persons from which the trace may originate.

A biometric system is a pattern recognition device that acquires physical or behavioral data from an individual, extracts a salient feature set from the data, compares this feature set against the features set stored in the database and provides the result of the comparison.

This paper proposed the system enhances the crime recording operations of the National Police Force. The data used by the Crime information system is stored in a centralized database which holds information about criminals, crime and users of the system. The database is the basis for all actions in the system and can be easily updated and used to aid in all of the system's processes, that is, all of the required information is stored in one central location and thus is easily accessible. Furthermore, the correctness of the centralized database will allow functions such as crime report generation and statistical analysis of crime data. This is a more effective storage method than a paper-based file system.

BACKGROUND WORK

Biometrics is a rapidly developing and yet emerging field of technology, with immense potential, which helps to make our lives easier and safer. Biometrics is widely used in many areas such as security monitoring, database access, forensic applications, and for verification and identification[1]. Accurate and efficient identification have become a vital requirement for forensic application due to increasing criminal activities. Identification approaches in forensic science are being replaced by recent advancement in the biometric technology which is equipped with computational intelligence techniques[2]. Almost all biometric modalities like, face, fingerprint, iris, denture etc. are used in different forensic identification areas. The emergence of forensic biometrics covers a wide range of applications which include identification of criminals, mass disaster victim identification, and identification of fire victims etc. Forensic Biometrics also overcomes the loopholes in traditional forensic identification systems that used manual ways for person identification. It is considered as a fundamental shift in the way criminals were detected[4].

Fingerprints have been used in criminal investigations as a means of identification for centuries because of their robustness and uniqueness. A fingerprint is the pattern of friction ridges and valleys on the surface of a fingertip[5]. In order to match a print, a fingerprint technician digitalizes or scans the print obtained at a crime scene and computer algorithms of a biometric system locate all the unique minutia and ridge endings and bifurcation points of a questioned print[6]. These unique feature sets are then matched against those stored in the fingerprint database.

IMPLEMENTATION

This research paper was aimed at the implementation of a Criminal Records Management System. It is a database system in which the police keep the record of criminals who have been arrested, to be arrested, or escaped. This will help the police department in enhanced management of information. The main entities in the whole process include; the petitioner (the person who files a First Incident Report (FIR)), victim, accused or criminal, case, and investigating officer. The CRMS keeps records of the petitioner, victim, accused, FIR, case and investigation officer entities.

A criminal's data is entered into a database where it is used for referential purposes and can be updated or modified regularly. Also, the ability to compare fingerprints whose unique patterns assist in redundancy control is an added attribute of the system. Advantages of the system include reduction of redundancies and inconsistencies in criminal information, ensures user defined rules to promote data integrity, enables sharing of data across all applications, and ensures proper access authorization for users. Its weaknesses are that the system was unable to generate reports and focuses more on criminal information.

3.1 Investigation Information

The system implements a biometric-based crime investigation system for the Police officers. A criminal's data is entered into a database where it is used for referential purposes and can be updated

or modified regularly. Also, the ability to compare fingerprints whose unique patterns assist in redundancy control is an added attribute of the system. Advantages of the system include reduction of redundancies and inconsistencies in criminal information, enable sharing of data across all applications, and ensure proper access authorization for users.

Figure 1 Investigated Police information

The figure1 shows the database system in which the police keep the record of criminals who have been arrested, to be arrested, or escaped. This will help the police department in enhanced management of information. The main entities in the whole process include; the petitioner, victim, accused or criminal, Fingerprint of accused, case, and investigating officer. The system's strengths lay in that it allows for storage of multiple data for a criminal

3.2 Fingerprint Verification

This module acquires the finger biometric data from a user and claims his identity. Feature extraction, processes the acquired biometric data and extracts a feature set using Minutiae Extractor and Orientation. Fingerprints can be classified as weakly-order textures exhibiting a dominant ridge orientation at each point. The orientation field provides a rough description of the fingerprint pattern that can be estimated with reasonable accuracy even from noisy input images. Here, characterize the location of each minutia with respect to the input fingerprint pattern based on a descriptor that comprises information about the orientation field in a broad region around the minutia point. Since the ridge orientation typically exhibits small spatial variations between neighbourhood pixels, a large area of the orientation field can be reconstructed from the orientation angles estimated in a relatively small number of sampling points.
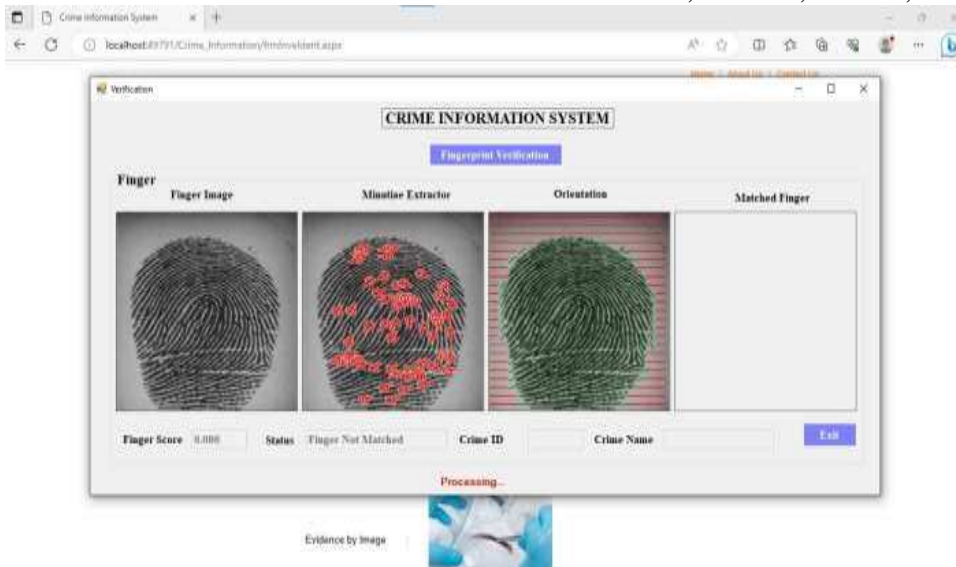
Figure 2 Finger print verification

The figure2 describe the sampling points assigned to each minutia can be organized in a circular pattern around the minutia position. Then the system compares the traits with the templates of the claimed identity provided at enrollment phase. It produces finger match score using matching algorithm.

3.3 Criminal Record Information

This process was aimed at the implementation of a storing, editing and deleting criminal records. It is a database system in which the police keep the record of criminals who have been arrested, to be arrested, or escaped. This will help the police department in enhanced management of information. The main entities in the whole process include; the petitioner, victim, accused or criminal, Fingerprint of accused, case, and investigating officer. The system's strengths lay in that it allows for storage of multiple data for a criminal

Figure 3 Crime Charge sheet Information

The data used by the Crime information system is stored in a centralized database which holds information about criminals, crime and users of the system. The database is the basis for all actions in the system and can be easily updated and used to aid in all of the system's processes, that is, all of the required information is stored in one central location and thus is easily accessible.

CONCLUSION

The paper titled as "Online Crime information System" is a web based application. This software provides facility for reporting online crimes, complaints. This software is developed with scalability in mind. Additional modules can be easily added when necessary. The software is developed with modular approach. All modules in the system have been tested with valid data and invalid data and everything work successfully.

A successful implementation of the Biometrics-Based Police Investigation System will greatly increase the efficiency and will help to ensure that criminal records are managed properly monitoring of criminal suspects in the country. the problem of delay in retrieving criminal suspects records for reference purposes and for appropriate court action or prosecution to be taken can be reduced drastically and the efficiency in the management of criminal records and investigation of criminal case being rendered by the Nigerian Police Force will greatly be improved upon.

FUTURE ENHANCEMENT

In future work, implementation of Automated Multi-modal Biometric Identification System (AMBIS). AMBIS is an advanced crime detection system, for detection of crime based on the available fingerprint, palmprint and iris scan data of criminals on record as well as accidental finger/ palm-prints collected from crime scenes.

Implementation of GPS can be used to get the coordinates of the crime area, which makes investigation even more adequate and easier.

References

[1] Jason Butlre and tony Caudil, "ASP.NET Database Programming" Fifth Edition, Hungry minds, Inc.Publicating Company Limited, New Delhi.

[2] Glen Johnsonwiley, Learning ASP.NET , Fourth Edition, Hungry minds, Inc. Publicating Company Limited, Uttar Pradesh.

[3] Jason Butlre and tony Caudil, " K.M. Hussian and Donna Hussaian, Information Systems: Analysis, Design and implementation, Second Edition, Tata McGraw-Hill, Delhi.

[4] Edward Yourdon, Managing the System Life Cycle, Second Edition, Englewood cliffs & N.J, Yourdon Press, US.

[5] Edward Yourdon and Larry L. Constantine, Structured Design: Fundamentals and Applications in Software Engineering, Second Edition, Englewood Cliffs & N.J, Yourdon Press, US.

[6] Shaku Atre, Data Base: Structured Techniques for Design, Performance and Management, Third Edition, Wiley, New York.

USER'S PLEASURE REGARDING THE SERVICE RENDERED BY TAMIL NADU STATE TRANSPORT CORPORATION (TOWN BUSES) WITH SPECIAL REFERENCE TO KANYAKUMARI DISTRICT

Dr.G.Rajesh Babu - Assistant Professor, Nanjil Catholic College of Arts and Science Kaliyakkavilai – 629 153, Kanyakumari District. Affiliated by Manonmaniam Sundaranar University, Tirunelveli - 627 012 : rajeshbabuji@gmail.com
Dr.R.Jolly Rosalind Silva - Assistant Professor, Nanjil Catholic College of Arts and Science Kaliyakkavilai – 629 153, Kanyakumari District.Affiliated by Manonmaniam Sundaranar University, Tirunelveli - 627 012 :silva.jolly@gmail.com

ABSTRACT

Tamil Nadu State Transport Corporation Limited is is a public transport bus operating service of Tamil Nadu. It is the largest bus corporation in the world and operates buses along intra and intercity routes. Kanyakumari is the southern part of India. In Kanyakumari district there are several villages. According to the National Transport Policy Committee (1980), "Future Nationalization should be guided by efficiency of operation of existing undertaking and to the extent to which they can provide consumer pleasure. Only the better financial and operational performance of any State Transport Undertaking will not fulfill the expectations of the people. The major objective is to study the bus user profile and to measure the bus user's pleasure of the sample respondent with regard to the TNSTC town bus operations. The study is an empirical research based on survey method. The study is based on primary data and has been collected from a sample of 150 respondents by means of questionnaires. The sample is made on the basis of cluster sampling. The information so collected has been analyzed using tools like percentage and arithmetic mean. The major findings are the age group between 25 to 50 years user's pleasures is more than other age group. The mean score of concessional facilities is 4.81 (rank 1) and experienced driver is 4.2 (rank 2). The private employee user's pleasures are more than other occupation. Improving and latest fuel technologies by adopting for a long term policy of the Transport Department of the Government will be more effective. The recommendation is the Government increases more and more facilities at the bus stands for attracting users and for giving them maximum comforts. The study of users' pleasure reveals that the users are satisfied with the operation of buses of TNSTC (town buses). Passenger's demand for travelling in India is increasing every day, and more and more passengers are in search of transportation information.

Key words: User's Pleasure, Transport, Operational Performance

INTRODUCTION

Transportation is a major activity in this world as people need to travel, goods have to be moved from place to places. So, transportation act as a catalyst and forms the backbone of the economic growth. In India the state transport undertakes ply more than 3.5 lakh buses serving more than 300 million people. Tamilnadu state transport's undertaking hold 17,284 buses and serve more than 15 million users daily. The objectives of the state owned transport undertakings are to provide economic, effective, efficient and well coordinated user road service network in Tamilnadu.

Transport is the servant of the people – the genie which makes their dreams come true – the magic carpet of their future – mobility of their desires, their aspirations, and their resources. Road transport being a service industry has to provide a quality service with reference to customer pleasure. Transport plays a significant role in the overall development of a nation's economy.

The road network has increased almost eightfold during the same period. India has an extensive network of major and minor roads as well as good number of well maintained networks of national highways connecting all major cities and tourists' destinations. The road ways provide transportation to millions of people every day.

Tamil Nadu State Transport Corporation Limited is established in the year 1972. It is a public transport bus operating service of Tamil Nadu. It is the largest bus corporation in the world and operates buses along intra and intercity routes. Kanyakumari is the southern part of India. In Kanyakumari district there are several villages.

At present Tamil Nadu State Transport Corporation (Tirunelveli) Limited Nagercoil region, (TNSTC) has 11 branches such as Ranithootham I, Ranithootham II, Ranithootham III, Kanyakumari, Vivenkanandapuram, Monday market, Colachel, Thiruvattar, Marthandam, Kuzhithurai, Chettikulam.

REVIEWS:-

Vanniarajan.T and Mr. Meenakshinathan (2007) conducted a study on A topology analysis of quality , customer satisfaction and behaviour intention in Indian rural banking service to analyse the impact of service quality factors on customer pleasure . The study reveals that the degree of impact of various quality factors on customer satisfaction is greater than on the behavioural intention.

Vijayakumar.T and Velu (2007) conducted a study on critical determinants of customer satisfaction in retail banking in India with the objective of identifying the determinants of quality dimensions, service problems, product used and the intention of switching over to other banks. The study provided insights and implications for managers in retail banks who want to improve customer satisfaction and retention rates.

Friman et al. (2001) established an evaluation model to evaluate the customer satisfaction of public transport and concluded that the overall pleasure is positively correlated with the cumulative pleasure.

Kennedy et al. (2005) believed that the public should participate in public transport management and analyzed the influence of the public on the pleasure of public transport.

Mahesh Chand in his research study entitled "Current Issues in Public Road Transport Management" has tried to evaluate critically the pros and cons of Nationalization taking into consideration the constraints of public transport undertakings also.

STATEMENT OF THE PROBLEM

According to the National Transport Policy Committee (1980), "Future Nationalization should be guided by efficiency of operation of existing undertaking and to the extent to which they can provide consumer pleasure. Only the better financial and operational performance of any State Transport Undertaking will not fulfill the expectations of the people, it is essential for the policy makers's to build such a transport system that concentrates in the people's orientation, giving them all amenities and comforts and satisfying them in all respects. The present research work has been to study the quality of service, the TNSTC (town buses) provides its users keeping in mind the level of pleasure of its users in turn of money they pay. The present study entitled user's pleasure regarding the service rendered by TNSTC (town buses) with special reference to Kanyakumari district will be of immense help to the policy makers and other authorities to arrive at proper transport policy.

OBJECTIVE OF THE STUDY

To study the bus user profile and to measure the bus user's pleasure of the sample respondent with regard to the TNSTC town bus operations.

METHODOLOGY

The present study intents to examine the issues framed in the objectives. The study is an empirical research based on survey method. The primary data are collected a fresh for the first time and thus happen to be original in character. The study is based on primary data and has been collected from a sample of 150 respondents by means of questionnaires. The sample is made on the basis of cluster sampling. The information so collected have been analyzed using traditional tools

like percentage and arithmetic mean.

USER PROFILE OF THE SAMPLE RESPONDENTS

S.No.	Place of Residence	Percentage	Gender	Percentage
1	Urban	37	Male	33
2	Rural	63	Female	67
	Age		**Marital Status**	
1	Below 25 years	18	Married	59
2	25 to 50 years	53	Unmarried	41
3	Above 50 years	29		
	Educational Qualification		**Occupation**	
1	Upto HSC	39	Government Employee	16
2	Diplomo	17	Private employee	42
3	Graduates	28	Home maker	13
4	Postgraduates	12	Self employee	8
5	Professionals	4	Students	21
	Monthly Income		**Bus Usage**	
1	Lessthan Rs.10000	21	Daily	62
2	Rs. 10000 to 20000	48	Frequently	26
3	Rs20001 to 30000	13	Occasionally	12
4	Above Rs. 30000	18		

PLEASURE OF USER AND OPERATIONAL ASPECTS OF BUS OPERATIONS IN TNSTC

S.No	Services	Total Score	Mean Score	Rank
1	Quick Service	589	3.92	IV
2	Availability during convenient hour and place	472	3.14	X
3	Maintenance and appearance of the bus	452	3.01	XI
4	Safety and reliability	568	3.78	V
5	Not satisfied with the operation of other bus	549	3.66	VI
6	Cooperative attitude of Crew members	602	4.01	III
7	Concessional facilities	627	4.81	1
8	Easy to carry luggage	546	3.64	VII
9	Experienced driver	631	4.2	II
10	Many stopping	503	3.35	IX
11	Noise Pollution	529	3.52	VIII
12	User Comforts	450	3	XII

The mean score of concessional facilities is 4.81 (rank 1), score of experienced driver is 4.2 (rank 2), score of cooperative attitude of crew members is 4.01 (rank 3), score of quick service is 3.92 (rank 4), score of safety and reliability is 3.78 (rank 5), score of not satisfied with the operation of other bus is 3.66 (rank 6), score of easy to carry luggage is 3.64 (rank 7), score of noise pollution is 3.52 (rank 8), score of many stopping is 3.35 (rank 9) and score of availability during convenient hour and place is 3.14 (rank 10)

.

AGE AND USER PLEASURE

S.No	Age Group	Percentage of Respondents	Mean Score	Rank
1	Below 25 years	18	3.5	III
2	25 to 50 years	53	4.21	I
3	Above 50 years	29	3.98	II

The mean score of age between 25 to 50 years is 4.21 (rank 1), score of age above 50 years is 3.98 (rank 2) and score of below 50 years is 3.5 (rank 3). So the age group between 25 to 50 years user's pleasures is more than other age group.

GENDER AND USER PLEASURE

S.No	Gender	Percentage of Respondents	Mean Score	Rank
1	Male	33	2.88	II
2	Female	67	4.42	I

The mean score of female is 4.42 (rank 1). The mean score of male is 2.88 (rank 2).. So the female gender   user's pleasures are more than male.

USER'S PLEASURE AND THEIR MARITAL STATUS

S.No.	Marital Status	Percentage of Respondents	Mean Score	Rank
1	Married	59	4.38	I
2	Unmarried	41	4.06	II

The mean score of married is 4.38 (rank 1). The mean score of unmarried is 4.06 (rank 2). So the married user's pleasures are more than unmarried.

USER'S PLEASURE WITH THEIR LEVEL OF EDUCATION

S.No.	Education Level	Percentage of Respondents	Mean Score	Rank
1	Upto HSC	39	3.96	II
2	Diplomo	17	3.88	III
3	Graduates	28	3.99	I
4	Postgraduates	12	3.44	IV
5	Professionals	4	3.21	V

The mean score of graduates is 3.99 (rank 1), score of upto HSC is 3.96 (rank 2), score of diplomo is 3.88 (rank 3), score of postgraduate is 3.44 (rank 4) and score of professionals is 3.21 (rank 5). So the graduate's user's pleasure is more than other educated people.

USER'S PLEASURE AND THEIR OCCUPATION

S.No.	Occupation	Percentage of Respondents	Mean Score	Rank
1	Government Employee	16	4.16	III

2	Private employee	42	4.83	I
3	Home maker	13	4.58	II
4	Self employee	8	3.0	V
5	Students	21	4.11	IV

The mean score of private employee is 4.83 (rank 1), score of home maker (rank 2), score of Government employee is 4.16 (rank 3), score of students is 4.11 (rank 4) and score of self employee is 3 (rank 5). So the private employee user's pleasures are more than other occupation.

MONTHLY EARNINGS AND USER'S PLEASURE

S.No.	Monthly Income	Percentage of Respondents	Mean Score	Rank
1	Lessthan Rs.10000	21	4.10	II
2	Rs. 10000 to 20000	48	4.22	I
3	Rs20001 to 30000	13	3.56	III
4	Above Rs. 30000	18	3.12	IV

The mean score of Rs.10000 to 20000 is 4.22 (rank 1), score of lessthan Rs.10000 is (rank 2), score of Rs.20001 to 30000 is 3.56 (rank 3) and score of above Rs.30000 is 3.12 (rank 4). So the monthly income Rs.10000 to 20000 user's pleasures is more than other income groups.

PURPOSE OF TRAVEL AND USER'S PLEASURE

S.No.	Usage	Percentage of Respondents	Mean Score	Rank
1	Daily	62	3.72	II
2	Frequently	26	3.98	I
3	Occasionally	12	3.6	III

The mean score of frequently user is 3.98 (rank 1), score of daily user is 3.72 (rank2) and score of occasionally user is 3.6 (rank 3). So the frequently user's pleasures are more than other users.

SUGGESTIONS
  - The Government is supposed to improve the functioning of Research and Development.
  - The Government is supposed to look for improving its operational policy by making drastic changes in the transport policy.
  - Improving and latest fuel technologies by adopting for a long term policy of the Transport Department of the Government will be more effective.
  - The Government is supposed to increase more and more facilities at the bus stands for attracting users and for giving them maximum comforts.

CONCLUSION

The study of users' pleasure reveals that the users are satisfied with the operation of buses of TNSTC (town buses). Passenger's demand for travelling in India is increasing every day, and more and more passengers are in search of transportation information.

BIBLIOGRAPHY

1. Abd-El-Fattah Sabry M. (2010), Longitudinal Effects on customer Satisfaction: Journal of International Social Research, Year 2010, Vol. 3, Issue 10, and pp 11-21.

2. Alam Muhammad Masroor, Jamilha Fakir Mohammad (2010), Level of Customer Satisfaction and Intent to Leave Among Malaysian, Business Intelligence Journal, Vol. 3, Issue 1, pp 123-137.

3. Ali Taskina, Ireen Akhter (2009), Job Satisfaction of Staffs Members in Private Universities -In Context of Bangladesh, International Business Research, Vol. 2, Issue 4.

4. B Krnjajic Stevan (2003), The executives under stress, Zbornik Instituta za Pedagoška Istraživanja, Vol. 35, pp 222-244.

5. Bennis, W. (1989). On becoming a leader. New York: Addison Wesley Publishing Company.

6. Bhattacharyya, S., & Chatterjee, L. (2005). Organizational downsizing: From concepts to practices. *Vikalpa*, 30(3), 65-78.