

## Techniques for predicting dark web events focused on the delivery of illicit products and ordered crime

Romil Rawat<sup>1</sup>, Olukayode Ayodele Oki<sup>2</sup>, Sakthidasan Sankaran<sup>3</sup>, Hector Florez<sup>4</sup>,  
Sunday Adeola Ajagbe<sup>5</sup>

<sup>1</sup>Department of Computer and Information Technology, University of Extremadura, Badajoz, Spain

<sup>2</sup>Department of Information Technology, Walter Sisulu University, East London, South Africa

<sup>3</sup>Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, Chennai, India

<sup>4</sup>ITI Research Group, Universidad Distrital Francisco Jose de Caldas, Bogota, Colombia

<sup>5</sup>Department of Computer and Industrial Production Engineering, First Technical University Ibadan, Ibadan, Nigeria

### Article Info

#### Article history:

Received Aug 27, 2022

Revised Mar 10, 2023

Accepted Mar 12, 2023

#### Keywords:

Computer vision

Crime prediction

Cyber terrorism

Darkweb

Information security

Machine learning

### ABSTRACT

Malicious actors, specially trained professionals operating anonymously on the dark web (DW) platform to conduct cyber fraud, illegal drug supply, online kidnapping orders, CryptoLocker induction, contract hacking, terrorist recruitment portals on the online social network (OSN) platform, and financing are always a possibility in the hyperspace. The amount and variety of unlawful actions are increasing, which has prompted law enforcement (LE) agencies to develop efficient prevention tactics. In the current atmosphere of rapidly expanding cybercrime, conventional crime-solving methods are unable to produce results due to their slowness and inefficiency. The methods for accurately predicting crime before it happens "automated machine" to help police officers ease the burden on personnel while also assisting in preventing offense. To achieve and explain the results of a few cases in which such approaches were applied, we advise combining machine learning (ML) with computer vision (CV) strategies. This study's objective is to present dark web crime statistics and a forecasting model for generating alerts of illegal operations like drug supply, people smuggling, terrorist staffing and radicalization, and deceitful activities that are connected to gangs or organizations showing online presence using ML and CV to help law enforcement organizations identify, and accumulate proactive tactics for solving crimes.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



### Corresponding Author:

Sunday Adeola Ajagbe

Department of Computer and Industrial Production Engineering, First Technical University Ibadan

Km 15 Lagos-Ibadan Expy, 200255, Ibadan, Oyo, Nigeria

Email: sunday.ajagbe@tech-u.edu.ng

## 1. INTRODUCTION

In the field of artificial intelligence (AI), computer vision (CV) and image processing frameworks are used to identify and interpret the visual world, giving the machine a sense of awareness of its virtual cognitive surroundings [1]–[3]. The modeling of actual criminal patterns and signature loops is made easier by CV [4], [5]. By obtaining three-dimensional (3D) visuals in object detection, face and gesture recognition, image computation, criminal image identification, terrorist location and weapons recognition, illicit activity monitoring and alarming, geolocation tagging, and suspicious word scripts, mathematical approaches have been developed to retrieve and make it possible for automated processes (AS) to interpret data [6], [7]. VLFeat is a tool that can produce results much more quickly than anticipated [8], [9]. VLFeat was defined as a library of as a library of CV algorithms in an artificial intelligence-machine learning (AI-ML) study that was utilized

to carry out fast prototyping and identify the human posture using face detection and human identification [10]. A computer system may learn from past events despite needing to be expressly programmed using the machine learning (ML) approach [11] and ML understands the precise architecture and frameworks [12], [13]. Although the nature of various offenses and their motives often appear to be random, ML may aid with pattern identification [14], [15] and content modelling utilizing natural language processing (NLP) techniques based on CV.

Mahanolob is a cybercrime analysis and prediction tool with a dynamic time-wrapping technique that enables both the forecast of crime and the eventual perpetrator's apprehension, according to related research [1]. Furthermore, the law enforcement (LE) in the United States, United Kingdom, and other European nations use crime-predicting apps to monitor criminal activity on social media and in specific geographic areas [16]. National authorities and the government now encourage the merging of ML techniques with technological automated systems and criminal intelligence [17]. It provides the means of a brand-new, strong machine (a group of programs) to aid in the pursuit of criminal investigations. The main objective of crime prediction is to foresee incidents before they take place so that a prior plan may be developed in recognized terrorist and criminal hotspots, which helps to comprehend terrorist action plans. Forecasting, policing with a high degree of precision, government critical resources such as police manpower educated with cyber tools based on ML, detectives, and financial specialists at cyber network usage, to battle crime.

Figure 1 outlines the background behaviors of illicit activities containing terrorist cyber events, triggering modes, propagation modes, damaging factors, and structure of losses. Cyber vulnerabilities are planned and created by terrorists in a sequential manner, identifying the effects on online platforms. The cyber threat always triggers an evaluation of distributed factors. The purposes of triggering are to make the post-global and attract supporters to join terrorist camps using the online social networks (OSN) platform.

The remainder of the section is laid out as follows: section 2 discusses terrorism diagnosis using social media. Section 3 discusses crime anticipation using ML techniques. Section 4 discusses crime prediction approaches CV, ML, deep learning (DL). Section 5 discusses proposed concept and design for cybercrime prediction with crime statistics. Section 6 provides the results and discussion of this research while. And section 7 concludes the paper with future work. Contribution: i) to show crime prediction using ML, CV, and DL with crime statistics for tracing illicit events channels and criminals' associations; ii) to show the hidden criminal market business tracing; and iii) to help the law enforcement officials to trace criminal events on digital platforms, so that action can be taken.

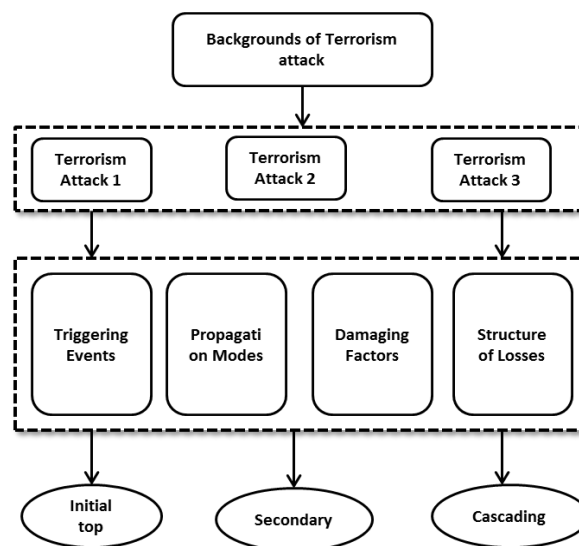


Figure 1. Terrorist cyber events triggering

## 2. TERRORISM DIAGNOSIS FROM SOCIAL MEDIA

Various techniques and automated engineers are being developed to detect terrorist content on social media [18], [19]. Malicious data in the form of text, pictures, videos, audio, likes, and re-sharing of posts spreads terrorist sentiments or infringements or messages for terror clusters, causing massive unrest and disruptions in the state or country, particularly in certain regions used for spreading propaganda and recruiting a terrorist army. Figure 2 shows the AI-based terrorist image behavior data prediction. Unethical posts related

to terrorism and data are collected from online platforms for creating data stores so that features can be extricated for further intelligent evaluation. The experimental data is collected by scarping the dark web platform to generate defined fingerprints and criminal activities associated with them. Based on the generated dataset, the model is trained for the prediction of all events relating to criminal activities, focusing on terrorist-related actions.

Figure 3 shows the labelling of terrorism-related post and contents. The online platform is surrounded by illicit activities, but it becomes difficult for normal users to identify and block them. So, terror-related content is selected for labelling and the results are modelled using intelligent algorithms convolutional neural network (CNN) and artificial neural network (ANN) [20], [21]. This helps the engines to automatically filter the malicious posts resembling terror activities and makes the modelled (group) vulnerable [22] as it helps the person sharing and resharing the post along with comments to highlight the information to the maximum audience.

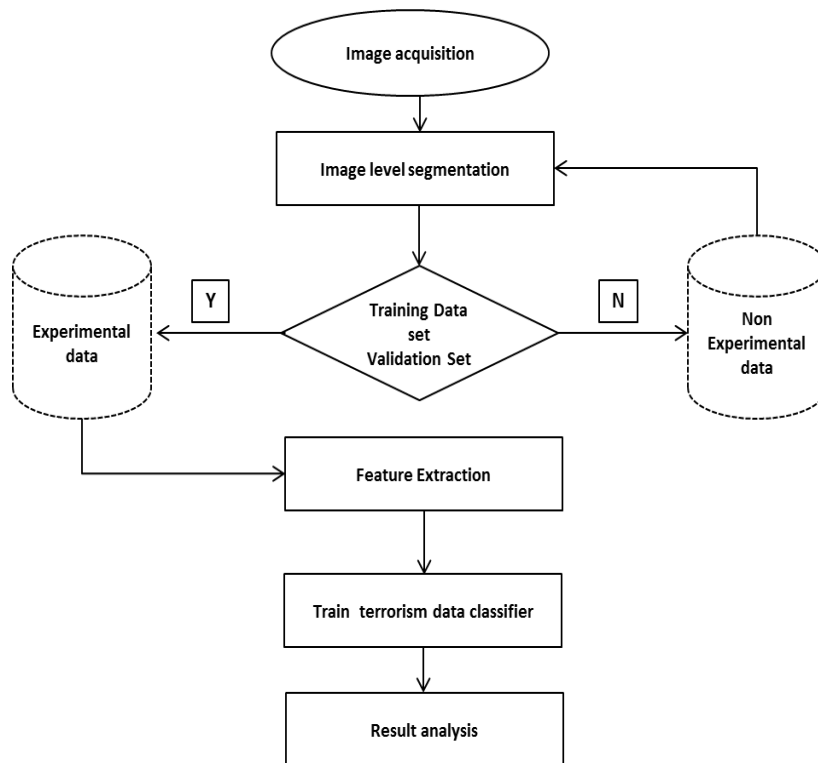


Figure 2. AI-based terrorist image behavior data prediction

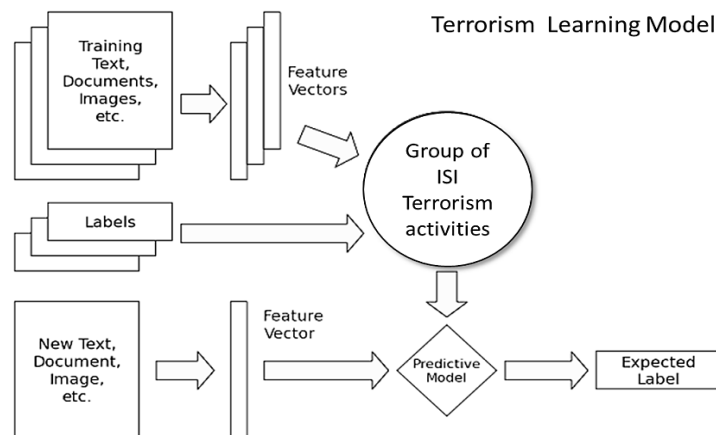


Figure 3. Labeling of terrorism-related posts and contents

### 3. CRIME ANTICIPATION USING ML TECHNIQUES

The comparative study was conducted using Weka, which is open an opensource tool for data mining. Violent crime trends from the dataset of communities and crime unnormalized and real-time crime statistical data based on three methods, namely linear regression (LR), additive regression (AR), and decision stump (DS), were constructed utilizing similar limited sets of characteristics for demonstrating the efficacy of ML approaches in predicting violent crime patterns of criminal hotspots, the test samples were chosen at random. LR algorithm shows appreciable results among the listed algorithms and tolerates unpredictability in the test data to some extent [23]. The crimes of house burglary, street robbery, and battery were examined retrospectively using an ensemble model to synthesize the findings of logistic regression and neural network (NN) frameworks using the predictive analytic approach to produce fortnightly and monthly forecasts (based on previous three years of cybercrime datasets) for the year [1]. ML was used to examine crime predictions. For the purpose of prediction, crime statistics from the previous 15 years in Vancouver (Canada) were studied. The accumulation of data, data categorization, pattern recognition, prediction, and visualization are all part of ML-based criminal investigations. The crime dataset was further analyzed using boosted decision tree (BDT) and k-nearest neighbor (KNN) methods. In a separate but similar research, [24], [25] looked at 560,000 crime statistics from 2003 to 2018 and found that using ML algorithms for crime prediction, the studies predicted crime with an accuracy of 44 per cent to 39 per cent respectively.

The crime dataset from Chicago, the United States. ML and data science (DS) approaches were applied to predict crime details consisting of parameters (scene positioning, type, date, time, and coordinates). decision trees (DT), random forest (RF), support vector machine (SVM), logistic regression (LR), and Bayesian techniques (BT) are used, with the most accurate model training. With an accuracy of about 0.787, the KNN classification proved to be the most accurate. The authors also utilized several graphics to assist in comprehending the various features of the Chicago crime dataset to better anticipate, identify, and solve crimes, resulting in a reduction in the crime rate. Data (taken from Chicago crime statistics, demographic and climatic data) accumulation, data preprocessing, predictive model development, dataset training, and testing are included in the proposed system to demonstrate the efficacy of the ML system to forecast violent behaviors, and crime incidences, and precise attributes of criminals. A deep neural network (DNN) forecasts crime attributes and occurrences by combining feature-based multi-model data from the environmental context. ML approaches like regression analysis (RA), kernel density estimation (KDE), and SVM is used in crime prediction systems [26], [27]. Figure 4 presents the dataflow diagram.

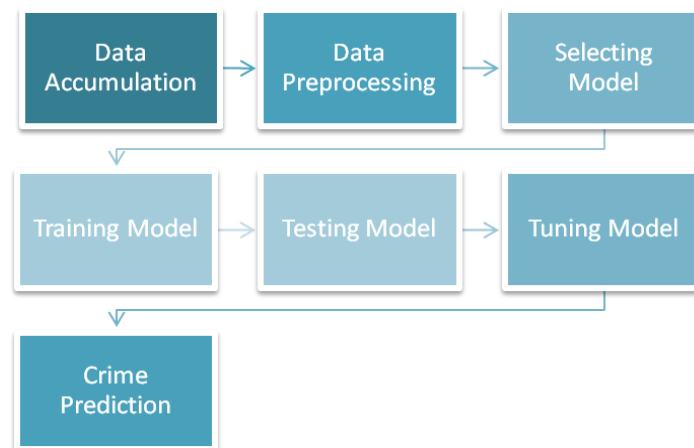


Figure 4. Dataflow diagram

The suggested DNN has an accuracy of 84.25%, whereas the SVM and KDE have an accuracy of 67.01% and 66.33%, indicating that the suggested DNN was much more accurate than the other prediction models in predicting crime occurrences [5]. The data were analyzed and interpreted using approaches such as Bayesian neural networks (BNN), and the Levenberg Marquardt algorithm (LMA) [12], and a scaled algorithm, with the scaled algorithm outperforming the other approaches. Statistical analysis revealed that using the scaled method, the crime rate could be reduced by 78%, implying an accuracy of 0.78. RapidMiner was used in a prediction study utilizing ML and historical crime trends in data collection, preparation, analysis, and visualization in the four primary visualization studies [9]. Big data (BD) offers a high throughput and fault tolerance, analyzing huge datasets and providing accurate findings, whilst the ML-based naive Bayes (NB)

method can make superior predictions with the existing datasets. Various data mining (DM) and ML methods utilizable singminal investigations are presented [6]. This study contributes by emphasizing the techniques utilized in crime data analytics. The grid-based crime forecasting framework created a series of spatial-temporal characteristics for a city in Taiwan based on 84 identified geographic locations for anticipating crime in the next slot (month) for every grid. DNN was determined to be the best model among the numerous ML techniques, particularly for a feature and attribute learning [28]. Furthermore, the suggested model architecture exceeded the baseline in terms of crime displacement testing. Figure 5 presents the functionality of the proposed approach.

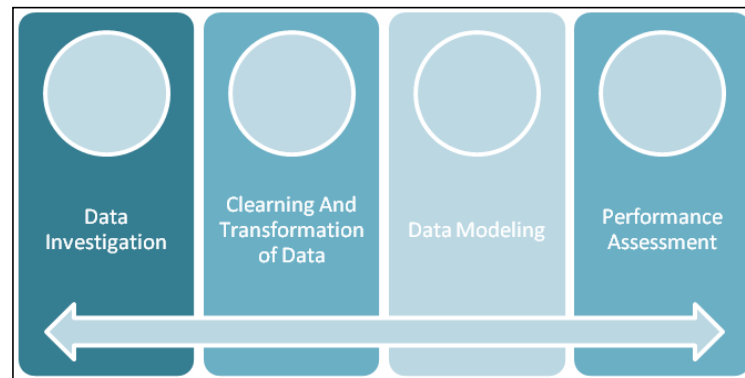


Figure 5. Functionality of the proposed approach

#### 4. CRIME PREDICTION APPROACHES (CV, ML, DL)

Alves *et al.* [29] demonstrated that integrating grey correlation analysis based on a new weighted k-nearest neighbor (GBWKNN) filling technique with KNN classification improves crime prediction accuracy. Using the suggested method, the study achieved a 67% accuracy rate. Obuandike *et al.* [30] classified crime data into two categories based on complexity, with the KNN method achieving an accuracy of approximately 87%.

Rajesh *et al.* [18] presented an insight into data mining and ML algorithms using an international database. With the help of Python and Jupyter Notebook, patterns and predictions were displayed as visualizations. This analysis aided in the development of suitable counter-terrorism measures, as well as increased investments, economic growth, and tourism. random forest regressor (RFR) outperformed all other ML algorithms considered in the study. Using the DT method, [31] obtained an accuracy of 84%. However, in both situations, a minor change in the data might result in a significant change in the structure. A novel crime detection approach known as naive Bayes (NB) is used for crime prediction and analysis [32]–[34]. Comes [11] only had an accuracy rate of 66% in predicting crimes and did not take into account computing speed, resilience, or scalability which are also important.

The multi-camera model of video surveillance was so well-designed that it can handle all three key tasks for normal police "stake-out", namely detection, representation, and recognition [35]–[37]. The detecting section combines video feed from numerous cameras to extract motion trajectories from videos quickly and accurately. The representation aids in the completion of raw trajectory data in order to create hierarchical, invariant, and content-rich motion event descriptions. Finally, the recognition section deals with event classification (such as robbery, as well as possible murder and molestation) and data descriptor identification. They created a sequence-alignment kernel function to perform sequence data learning to detect suspicious or possible criminal occurrences for effective recognition. A technique was proposed for distinguishing individuals for espionage using a novel feature called soft biometry, which incorporates a person's height, build, facial features, shirt and trousers color, motion behavior, and trajectory record to recognize and monitor passengers, as well as forecast crime pursuits and deal with some strange human error scenarios where the perpetrators get away with it [38]. They also conducted examinations with the findings being publicized. People's behaviors are captured, offering piggyback rides in increasingly remote locations with a given sequence from event footage. Table 1 summarizes the comparative study of crime prediction techniques with their accuracy and related findings. In Table 1, we summarized the evaluation models, further demonstrating qualitative analysis and accuracy.

Crime hotspots, known as severe-crime zones, have a high probability of crime occurrence and present abnormal events with a high likelihood of detecting criminals. They performed research on predicting crime

hotspots and implemented their model with google tensor flow. The emphasis is to produce higher value to demonstrate that the technique is more effective. with similar evaluation parameters, the gated recurrent unit (GRU) and long short-term memory (LSTM), achieved accuracy (81.5%), precision (86.5%), recall (75%), and an F1-score (0.8). Both outperform the standard recurrent neural network (RNN) version by a wide margin. The GRU version showed 2% better performance compared to RNN at receiver operating characteristic (ROC) area under the curve (AUC) findings. LSTM received the highest AUC score, which was 3% higher than the GRU version. A spatiotemporal crime network (STCN) is presented [36] which uses a CNN to predict crime before it happens. From 2010 to 2015, the authors used New York felony datasets (number-311) to test the STCN. The STCN outperformed the four baselines with an F1-score (88%) and an AUC (92%). Their suggested model outperformed the other baselines by F1-score and AUC values, and even when the time window approached 100, it was still better than the others in terms of the effectiveness of working in a densely populated region.

Table 1. Crime prediction techniques

No	Crime prediction techniques with references	Accuracy	Findings
1	RFR [18]	97%	High accuracy in previously recorded crimes.
2	DT [15]	83.95%	The DT shows good efficiency than NB, along with the same crime dataset implemented on Weka.
3	KNN (K=10) [39]	87.03%	Data has compared to five classification algorithms, finding that the NB, NN, and KNN algorithms have a better prediction rate than SVM and DT algorithms.
4	Decision tree (J48) [40]	59.15%	Experiments were done on J48 naïve Bayesian and ZeroR by comparing them.
5	NB [16]	65.59%	The comparative study is done based on the accuracy of k-NN, NB, and DT for the prediction of crimes and criminal behaviors.
6	Naïve Bayes classifier [28]	87.00%	NB is used for crime analysis and prediction.
7	SVM [29]	84.37%	Several models have been compared for analyzing the best chance of predicting hotspots.
8	KNN (K=5) [32]	66.69%	By combining GBWKNN and KNN classification approaches better accuracy is achieved.
9	Proposed word	89.50%	Focused on predicting the crime using ML, CV, and DL using crime statistics for tracing Illicit events channels and criminals' associations.

## 5. PROPOSED CONCEPT AND DESIGN FOR CYBERCRIME PREDICTION WITH CRIME STATISTICS

We assessed the relevance of each approach after discovering and comprehending numerous diverse ways utilized by security agencies for surveillance reasons. Every surveillance method generates appreciable results when found actively engaged in communication, like the sting ray used for detecting the geolocation of a user. So, to track the location based on replicating human approaches continually by self-updating modeling approach, even though communication is not made, a modern intelligent framework modeling DL, ML, and CV algorithms for conducting surveillance [41]–[45]. Table 2 contains the key components and processes of the proposed system. Table 2 contains the key components and processes of the proposed system.

By combining all these capabilities during a preliminary round, we would like to employ closed circuit television (CCTVs) connected to intelligent automated systems in real-world settings to comprehend the previously recorded crimes (collected Instances is 8,000), using ML and DL approaches for greater knowledge of criminality (explaining how, why, and where). We do not just propose building a world-class model to anticipate crimes; we propose teaching it to comprehend prior crimes in order to better assess and forecast them based on the utilization of scenario simulations. Following an analysis of the scene and the use of the key features listed above, the program should conduct at least 90 simulations of the current scenario in front of it, with the help of previously learned criminal records, to determine and recommend a plan of strategy for alerting LE personnel. In Figure 6, we provide the terrorist and criminals presence detection models.

- Input tracking: Data is collected from drones, static cameras, voice, and recording devices focused at suspicious places.
- Mapping with database: Containing profile and features of crime in security agency's databases relating to dark web (unusual weapon image, suspected criminal image, drug dealers, gangs' tattoos or marks, financial fraudulent agent).
- Automated engines: It will search the online presence of these criminals, for mapping with the site, so that the website and owner activity can be tracked.
- Alert of association: It is generated towards cyber cells or related authorities for collecting evidence.

- Dynamic database of security agencies connected with OSN: Containing CNN for crawling vulnerable posts, text, images, and video at OSN to map with Input tracking data [46], [47].

Table 2. Key components and processes of the systems

Components	Processes
Root analytics	<ul style="list-style-type: none"> <li>- Knowing the number of statistical methodologies able to anticipate future events.</li> <li>- The instance may range from behavioral intuition to robbing an organization in future timeframes.</li> </ul>
Neural networks	<ul style="list-style-type: none"> <li>- Consisting of a huge series of algorithms that assist in the discovery of data relationships by behaving and associating human cognition.</li> <li>- Replicating biological nerve cells, attempting to think for it.</li> <li>- Anticipating a crime scene.</li> </ul>
Automated intelligent engines	<ul style="list-style-type: none"> <li>- Engines that must fingerprint antivirus and viruses.</li> <li>- Improving the security of the system by identifying the type of threat and eliminating it using recognized antivirus.</li> <li>- Continuity of machine’s surveillance in case of broken down.</li> <li>- Prediction of anomaly time series prediction, and decisive approach with uncertainty.</li> </ul>
Cryptographic algorithms	<ul style="list-style-type: none"> <li>- Data mining in the detection of patterns in criminal’s activity.</li> <li>- Encrypting the known confidential criminal data in a secure manner.</li> <li>- Utilized to encrypt newly found possible criminal data.</li> </ul>
Cyber threat detection and classification	<ul style="list-style-type: none"> <li>- Classification of threats and criminal conduct like probable terrorist attacks can be anticipated based on the timeline.</li> </ul>
Forensic evidence NLP	<ul style="list-style-type: none"> <li>- Organize, analyze, and learn from the data once it has been collected.</li> <li>- Suspicious Speech print identification.</li> <li>- Identification of cyber criminal’s language and comprehension based on specific features represented using a mathematical formula.</li> </ul>
Data collection and analysis	<ul style="list-style-type: none"> <li>- Knowing previous crime attributes for casting future crime prediction rates.</li> </ul>
Gait analysis	<ul style="list-style-type: none"> <li>- To understand posture when walking and research human motion.</li> <li>- To gain a better understanding of a person's usual pace and body mark.</li> </ul>
Features	<ul style="list-style-type: none"> <li>- To determine an unusual visit to the criminal zone at a specific period, allowing the system to notify authorities.</li> </ul>

The scale of the dark web marketplaces (Silkroad, Alpha Bay, and Pandora) economy was difficult to determine and was growing all the time. Researchers estimated the Silkroad's sales volume at \$360,000 each day based on scrapes and comments, equating to more than \$120 million in a year [48]. The requirements for meeting the supply of illicit orders generated through dark web platforms are detailed in the Table 3. Our proposed model helps to track the activities of these associated criminals and agents contacting customers for delivery, thereby reaching out to the chain of order and criminal events. Table 3 presents the classification, dealers, agents and percentages of our system, the confusion matrix, and the outlines of graphical statistics of crime associated with the dark web environment are presented in Figures 7 and 8 respectively. The Table 4 is performance metrics and outcomes.

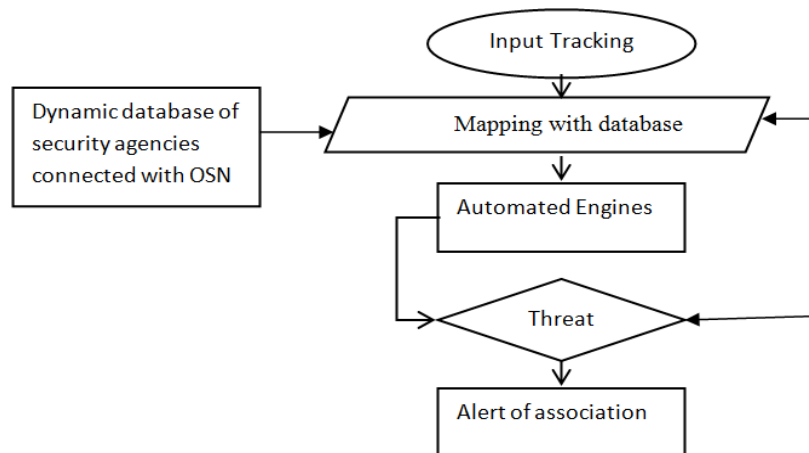


Figure 6. Terrorist and criminals’ presence detection model

Table 3. Dealers and agents meeting chart for illicit business trade and supply

Classification	Point of meeting - contact required (dealers and agents)	Percentage
Online gambling	No	1.7
Weapons trade	Yes	2.3
Criminal chat forums	May be	2.2
Pornography	Yes	3.5
Financial fraud	May be	4.9
Anonymity	May be	4.7
Ransomware	No	3.5
Prostitution	Yes	5.3
Human trafficking	Yes	5.8
Organ trafficking	Yes	5.1
Whistleblower	No	4.5
Drug trade	Yes	5.2
Financial fraud	May be	7.3
Contract killing	Yes	1.3
Gangs of Influence	Yes	2.3
Live streaming of criminals' events	Yes	3.8
Terrorism propaganda sharing	No	5.6
Terrorist recruitment and radicalization	Yes	3.4
Sale of antiques	Yes	2.8
cyber extortion	Yes	3.5
Hacking	No	5.2
Cyber-attack activation	No	5.3
Industrial applications controlling	May be	5.2
others	May be	5.6

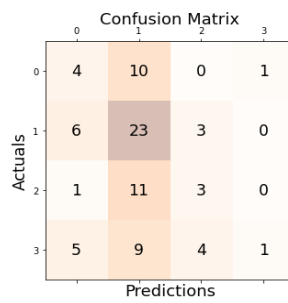


Figure 7. crime statistics confusion matrix

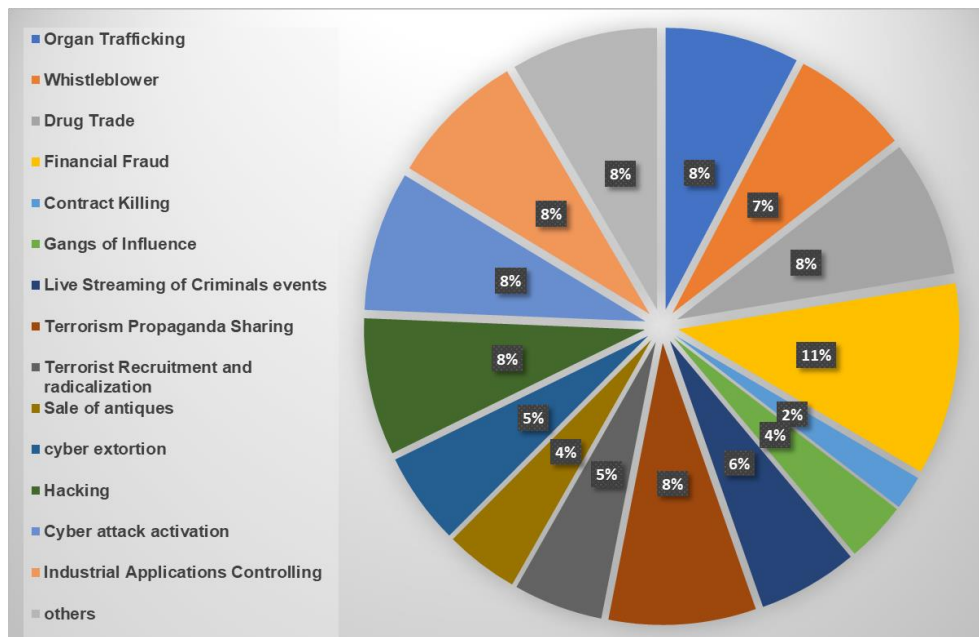


Figure 8. Crime statistics on dark web platform



Table 4. Performance metrics and outcomes

S/N	Measure	Descriptions	Outcomes
1	Sensitivity	$True\ positive\ rate\ (TPR) = True\ positive\ (TP)/(TP + False\ negative\ (FN))$	0.7383
2	Specificity	$SPC = True\ negative(TN)/(False\ positive\ (FP) + TN)$	0.9384
3	Precision	$Positive\ predictive\ value\ (PPV) = TP/(TP + FP)$	0.8650
4	Negative predictive value	$Negative\ predictive\ value\ (NPV) = TN/(TN + FN)$	0.9027
5	False positive rate	$False\ positive\ rate\ (FPR) = FP/(FP + TN)$	0.7116
6	False discovery rate	$False\ discovery\ rate\ (FDR) = FP/(FP + TP)$	0.7959
7	False negative rate	$False\ negative\ rate\ (FNR) = FN/(FN + TP)$	0.6817
8	Accuracy	$Accuracy\ (ACC) = (TP + TN)/(condition\ positive\ (P) + condition\ negative\ (N))$	0.8950
9	F1-score	$F1 = 2TP/(2TP + FP + FN)$	0.8001

## 6. RESULT AND DISCUSSION

The comparison of fortnightly projections of monthly analytical predictions with divides into day-night datasets, the researchers found, greatly improved the results. Due to its secrecy, the dark web has long been a target for criminals looking to make money illegally abroad. The current work uses ML, CV, and DL to forecast crime, and crime stats are offered to track criminal networks and compare the comparative research with the aspects of the suggested strategy that have been put into practice. The research is based on a fictitious model for locating terrorists and lawbreakers operating on the dark web who are engaged in drug dealing, human trafficking, staffing of terrorists, distribution of weapons, execution orders delivered online, and other illegal activities linked to gangs or organizations with active websites. Utilizing automated machine characteristics, modeling, and recognition. This experiment is about scraping the dark web site generates specific signatures and the illicit behaviors connected to them, which is how the exploratory data is gathered. The system is trained to forecast all criminal activity-related occurrences, with an emphasis on terrorist-related behaviors, using the provided dataset [49]. No such dataset exists contain records of criminals' events and channels like (drug supply, human trafficking, terrorist radicalization and recruitment, weapon delivery, online killing orders, and fraudulent activities associated with gangs or organizations showing online presence). The proposed focused on the work of hypothetical model and covered multidimensional illicit events channels with machine learning and computer vision technique [50].

Image processing technique and feature extraction utilizes ImageNet, one of the largest datasets of annotated pictures, CNN, a deep learning model that has been essential in enhancing computer vision, learns patterns that typically appear in images and is then equipped to adjust as new data is analyzed. Both a feature detector and a feature descriptor, spectrum feature transform (SIFT). SIFT splits an image into a vast number of localized characteristic vectors, all of which is somewhat robust to changes in light and affine or 3D projection as well as invariant to image translation, scaling, and rotation. Computer vision linking with image processing: AI and pattern identification methods for crime prediction are used in the domains of CV and image processing to acquire Illicit event sequences for extracting useful knowledge from photos, videos, and other visual inputs. One of the numerous methods used in CV is image synthesis, but other methods as well, including ML, CNN, and so on, are also used. One of the subfields in the science of CV is image processing and belongs to the subfield of image computing.

## 7. CONCLUSION AND FUTURE WORK

The authors concluded that comparing fortnightly forecasts of monthly analysis predictions with splits into day-night datasets improved the results significantly. Due to its anonymity, the dark web has always attracted the interest of criminals interested in generating illicit revenues across borders. The present work predicts crime using ML, CV, and DL with crime statistics to track criminal chains and compare the comparative study with the implemented features of the given approach. The work is based on a hypothetical model for tracking dark web criminals and terrorists involved in drug supply, human trafficking, terrorist radicalization and recruitment, weapon delivery, online killing orders, and fraudulent activities associated with gangs or organizations showing an online presence. The mapping and identification using automated machine features will help security agencies investigate the root suppliers of prohibited and illegal items. The anonymous dark web platform changes with hosting, so it takes time to track it. But criminals also use digital platforms for promotion or marketing tactics to supply or attract other criminals. Based on digital traces and evidence, security agencies can track the network. Our future research will begin with the creation of a machine that can predict and recognize patterns based on geo-location coordinates and the dates of similar crimes. We also hope to create software that can act as a universal security official, with eyes and ears everywhere.




## REFERENCES

- [1] N. Shah, N. Bhagat, and M. Shah, "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, no. 1, Apr. 2021, doi: 10.1186/s42492-021-00075-z.
- [2] S. A. Ajagbe, A. O. Adesina, T. J. Odule, and O. Aiyeniko, "Evaluation of computing resources consumption of selected symmetric-key algorithms," *Journal of Computer Science and Its Application*, vol. 26, no. 2, pp. 64–76, Feb. 2020, doi: 10.4314/jcsia.v26i2.7.
- [3] A. Razaque, B. Valiyev, B. Alotaibi, M. Alotaibi, S. Amanzholova, and A. Alotaibi, "Influence of COVID-19 epidemic on dark web contents," *Electronics*, vol. 10, no. 22, pp. 1–17, Nov. 2021, doi: 10.3390/electronics10222744.
- [4] A. L. Lira Cortes and C. Fuentes Silva, "Artificial intelligence models for crime prediction in urban spaces," *Machine Learning and Applications: An International Journal*, vol. 8, no. 1, pp. 1–13, Mar. 2021, doi: 10.5121/mlaij.2021.8101.
- [5] O. A. Adebisi, S. A. Ajagbe, J. A. Ojo, and M. A. Oladipupo, "Computer techniques for medical image classification: a review," in *Intelligent Healthcare*, Singapore: Springer Nature Singapore, 2022, pp. 19–36. doi: 10.1007/978-981-16-8150-9\_2.
- [6] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019, doi: 10.1038/s42256-019-0048-x.
- [7] F. Al-Fayez, M. W. A. El-Soud, and T. Gaber, "Thermogram breast cancer detection: a comparative study of two machine learning techniques," *Applied Sciences*, vol. 10, no. 2, Jan. 2020, doi: 10.3390/app10020551.
- [8] D. Thanh and S. Dvoenko, "Image noise removal based on total variation," *Computer Optics*, vol. 39, no. 4, pp. 564–571, 2015, doi: 10.18287/0134-2452-2015-39-4-564-571.
- [9] H.-W. Kang and H.-B. Kang, "Prediction of crime occurrence from multi-modal data using deep learning," *PLOS ONE*, vol. 12, no. 4, Apr. 2017, doi: 10.1371/journal.pone.0176244.
- [10] K. J. Hayward and M. M. Maas, "Artificial intelligence and crime: a primer for criminologists," *Crime, Media, Culture: An International Journal*, vol. 17, no. 2, pp. 209–233, 2021, doi: 10.1177/1741659020917434.
- [11] E. N. Comes, "Machine learning methods for predicting global and local crime in an urban area. Illinois Institute of technology," Illinois Institute of Technology, 2018.
- [12] M. Amiruzzaman, A. Curtis, Y. Zhao, S. Jamonnak, and X. Ye, "Classifying crime places by neighborhood visual appearance and police geonarratives: A machine learning approach," *Journal of Computational Social Science*, vol. 4, no. 2, pp. 813–837, Nov. 2021, doi: 10.1007/s42001-021-00107-x.
- [13] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, doi: 10.1016/j.heliyon.2018.e00938.
- [14] C.-C. Sun, C. Yao, X. Li, and K. Lee, "Detecting crime types using classification algorithms," *Journal of Digital Information Management*, vol. 12, no. 5, pp. 321–327, 2014.
- [15] R. Iqbal, "An experimental study of classification algorithms for crime prediction," *Indian Journal of Science and Technology*, vol. 6, no. 3, pp. 4219–4225, Mar. 2013, doi: 10.17485/ijst/2013/v6i3.6.
- [16] A. H. Wibowo and T. I. Oesman, "The comparative analysis on the accuracy of k-NN, naive Bayes, and decision tree algorithms in predicting crimes and criminal actions in Sleman regency," *Journal of Physics: Conference Series*, vol. 1450, no. 1, pp. 1–6, Feb. 2020, doi: 10.1088/1742-6596/1450/1/012076.
- [17] J. Abdollahi, B. Nouri-Moghaddam, and M. Ghazanfari, "Deep neural network based ensemble learning algorithms for the healthcare system (diagnosis of chronic diseases)," *Prepr. arXiv.2103.08182*, Mar. 2021, doi: doi.org/10.48550/arXiv.
- [18] P. Rajesh, D. Babitha, M. Alam, M. Tahermezadi, and A. Monika, "Machine learning and statistical analysis techniques on terrorism," in *Fuzzy Systems and Data Mining VI*, A. J. Tallón-Ballesteros, Ed. Andhra Pradesh: IOS Press, 2020, pp. 210–222. doi: 10.3233/FAIA200701.
- [19] M. Faisal, "We see you: terrorist prediction framework through psychological and social behaviors," in *2020 7th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, 2020, pp. 204–212. doi: 10.1109/ISCMI51676.2020.9311565.
- [20] C. Szegegy, A. Toshev, and D. Erhan, "deep neural networks for object detection," *Advances in neural information processing systems*, vol. 26, pp. 1–9, 2013.
- [21] Y. Zhou and Y. Su, "Development of CNN and its application in education and learning," in *2019 9th International Conference on Social Science and Education Research (SSER 2019)*, 2019, pp. 720–724.
- [22] R. Rawat, A. S. Rajawat, V. Mahor, R. N. Shaw, and A. Ghosh, "Dark web-onion hidden service discovery and crawling for profiling morphing, unstructured crime and vulnerabilities prediction," in *Innovations in Electrical and Electronic Engineering: Proceedings of ICEEE 2021*, Springer, 2021, pp. 717–734. doi: 10.1007/978-981-16-0749-3\_57.
- [23] A. I. Canhoto, "Leveraging machine learning in the global fight against money laundering and terrorism financing: an affordances perspective," *Journal of Business Research*, vol. 131, pp. 441–452, Jul. 2021, doi: 10.1016/j.jbusres.2020.10.012.
- [24] A. O. David and A. O. Eliahs, "A secured text encryption with near field communication (NFC) using Huffman compression," *International Journal of Engineering and Applied Computer Science*, vol. 04, no. 02, pp. 14–18, Mar. 2022, doi: 10.24032/IJEACS/0402/002.
- [25] R. Rawat, V. Mahor, A. Rawat, B. Garg, and S. Telang, "Digital transformation of cyber crime for chip-enabled hacking," in *Handbook of Research on Advancing Cybersecurity for Digital Transformation*, Pennsylvania: IGI Global, 2021, pp. 227–243, doi: 10.4018/978-1-7998-6975-7.ch012.
- [26] A. O. David and O. O. Oluwasola, "Zero day attack prediction with parameter setting using Bi direction recurrent neural network in cyber security," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 18, no. 3, pp. 111–118, 2020.
- [27] J. B. Awotunde *et al.*, "An ensemble tree-based model for intrusion detection in industrial internet of things networks," *Applied Sciences*, vol. 13, no. 4, 2479, Feb. 2023, doi: 10.3390/app13042479.
- [28] M. Jangra and S. Kalsi, "Crime analysis for multistate network using naive Bayes classifier," *International Journal of Computer Science and Mobile Computing*, vol. 8, no. 6, pp. 134–143, 2019.
- [29] L. G. A. Alves, H. V. Ribeiro, and F. A. Rodrigues, "Crime prediction through urban metrics and statistical learning," *Physica A: Statistical Mechanics and its Applications*, vol. 505, pp. 435–443, Sep. 2018, doi: 10.1016/j.physa.2018.03.084.
- [30] G. N. Obuandike, A. Isah, and J. Alhasan, "Analytical study of some selected classification algorithms in Weka using real crime data," *International Journal of Advanced Research in Artificial Intelligence*, vol. 4, no. 12, pp. 44–48, 2015.
- [31] S. Shojace, A. Mustapha, F. Sidi, and M. A. Jabar, "A study on classification learning algorithms to predict crime status," *International Journal of Digital Content Technology and its Applications*, vol. 7, no. 9, pp. 361–369, 2013.
- [32] H. A. H. Mahmood and H. A. Mengash, "A novel technique for automated concealed face detection in surveillance videos," *Personal and Ubiquitous Computing*, vol. 25, no. 1, pp. 129–140, Feb. 2021, doi: 10.1007/s00779-020-01419-x.
- [33] A. Mathew, F. A. Sa, H. Pooja, and A. Verma, "Smart disease surveillance based on internet of things (IoT)," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 5, pp. 180–183, 2015, doi: 10.17148/IJARCCE.2015.4541.




- [34] K. S. Adewole *et al.*, “Empirical analysis of data streaming and batch learning models for network intrusion detection,” *Electronics*, vol. 11, no. 19, Sep. 2022, doi: 10.3390/electronics11193109.
- [35] S. A. Ajagbe, H. Florez, and J. B. Awotunde, “AESRSA: A new cryptography key for electronic health record security,” in *Applied Informatics*, Arequipa, Peru: Springer International Publishing, 2022, pp. 237–251. doi: 10.1007/978-3-031-19647-8\_17.
- [36] S. Ojha and S. Sakhare, “Image processing techniques for object tracking in video surveillance—a survey,” in *2015 International Conference on Pervasive Computing (ICPC)*, Jan. 2015, pp. 1–6. doi: 10.1109/PERVASIVE.2015.7087180.
- [37] R. Rawat, V. Mahor, S. Chirgaiya, and A. S. Rathore, “Applications of social network analysis to managing the investigation of suspicious activities in social media platforms,” in *Advances in Cybersecurity Management*, Cham: Springer International Publishing, 2021, pp. 315–335. doi: 10.1007/978-3-030-71381-2\_15.
- [38] R. Rawat, V. Mahor, S. Chirgaiya, R. N. Shaw, and A. Ghosh, “Analysis of darknet traffic for criminal activities detection using TF-IDF and light gradient boosted machine learning algorithm,” in *Innovations in Electrical and Electronic Engineering*, Singapore: Springer Singapore, 2021, pp. 671–681. doi: 10.1007/978-981-16-0749-3\_53.
- [39] S. R. Nayak, J. Nayak, S. Vimal, V. Arora, and U. Sinha, “An ensemble artificial intelligence enabled MIoT for automated diagnosis of malaria parasite,” *Expert Systems*, vol. 39, no. 4, pp. 1–15, May 2022, doi: 10.1111/exsy.12906.
- [40] C. Fachkha and M. Debbabi, “Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1197–1227, 2016, doi: 10.1109/COMST.2015.2497690.
- [41] S. Farhana, A. Lajis, Z. A. Long, and H. Nasir, “Impact of big data congestion in IT: An adaptive knowledge-based Bayesian network,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 2, pp. 2031–2036, Apr. 2020, doi: 10.11591/ijece.v10i2.pp2031-2036.
- [42] S. O. Folorunso, J. B. Awotunde, F. E. Ayo, and K.-K. A. Abdullah, “Radiot: the unifying framework for IoT, radiomics and deep learning modeling,” in *Hybrid Artificial Intelligence and IoT in Healthcare*, S. Singapore, Ed. Singapore, 2021, pp. 109–128. doi: 10.1007/978-981-16-2972-3\_6.
- [43] T.-W. Sung, P.-W. Tsai, T. Gaber, and C.-Y. Lee, “Artificial intelligence of things (AIoT) technologies and applications,” *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–2, Jul. 2021, doi: 10.1155/2021/9781271.
- [44] S. A. Ajagbe, M. O. Ayegboyan, I. R. Idowu, T. A. Adeleke, and D. N. H. Thanh, “Investigating energy efficiency of mobile ad-hoc network routing protocols,” *Informatica*, vol. 46, no. 2, pp. 269–275, Jun. 2022, doi: 10.31449/inf.v46i2.3576.
- [45] K. Demestichas *et al.*, “Prediction and visual intelligence platform for detection of irregularities and abnormal behaviour,” in *CEUR Workshop Proceedings*, 2020, pp. 25–30, doi: 10.24406/publica-fhg-411356.
- [46] O. D. Adeniji, M. O. Ayomide, and S. A. Ajagbe, “A model for network virtualization with openflow protocol in software-defined network,” in *Intelligent Communication Technologies and Virtual Mobile Networks*, Singapore: Springer Nature Singapore, 2023, pp. 723–733. doi: 10.1007/978-981-19-1844-5\_57.
- [47] P. Patel and A. Thakkar, “The upsurge of deep learning for computer vision applications,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 538–548, Feb. 2020, doi: 10.11591/ijece.v10i1.pp538-548.
- [48] K. Thammarak, P. Kongkla, Y. Sirisathikul, and S. Intakosum, “Comparative analysis of Tesseract and Google cloud vision for thai vehicle registration certificate,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, pp. 1849–1858, Apr. 2022, doi: 10.11591/ijece.v12i2.pp1849-1858.
- [49] S. R. Tahhan, H. K. Aljobouri, and B. R. Altahan, “Anti-resonant based nested terahertz fiber design for illicit drugs detection,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 1588–1598, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1588-1598.
- [50] M. A. Al Noman *et al.*, “A computer vision-based lane detection technique using gradient threshold and hue-lightness-saturation value for an autonomous vehicle,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 347–357, Feb. 2023, doi: 10.11591/ijece.v13i1.pp347-357.

## BIOGRAPHIES OF AUTHORS






**Romil Rawat**    is a research scholar and attended several research programs and received research grants from USA, Germany, Italy, and UK. The author has research alignment towards cyber security, internet of things, dark web crime analysis and investigation techniques, and working towards tracing of illicit anonymous contents of cyber terrorism and criminal activities. He also chaired international conferences and hosted several research events including national and International Research Schools, PhD colloquium, workshops, training programs. He also published several research patents. He can be contacted at rawat.romil@gmail.com and rrawatna@alumnos.unex.es.






**Olukayode Ayodele Oki**    received his PhD from the University of Zululand, South Africa in 2019. He is a lecturer in the Department of Information Technology at Walter Sisulu University, South Africa. He has authored more than 30 articles. His research interests include biologically inspired computation, ICT4D, communication networks, internet of things, machine learning, data analytics and climate-smart agriculture. He has received several grants both for research and amp; development and to attend conferences. He is a recipient of the South Africa National Research Foundation (NRF) rated researcher award, an honorary rosalind member of the London journal press and a member of the IEEE South Africa subsection. He can be contacted at ooki@wsu.ac.za.






**Sakthidasan Sankaran**    is a Professor in the Department of Electronics and Communication Engineering at Hindustan Institute of Technology and Science, India. He received his B.E. degree from Anna University in 2005, M.Tech. Degree from SRM University in 2007 and Ph.D. Degree from Anna University in 2016. He is a senior member of IEEE for the past 10 years and a member of various professional bodies. He is an active reviewer in Elsevier journals and an editorial board member in various international journals. His research interests include image processing, wireless networks, cloud computing and antenna design. He has published more than 70 papers in Referred journals and International Conferences. He has also published three books to his credit. He can be contacted at sakthidasan.apec@gmail.com.



**Hector Florez**    obtained Ph.D. in Engineering, M.Sc. in Information and Communication Sciences, M.Sc. in Management, B.Sc. in Electronic Engineering, B.Sc. in Computing Engineering, and B.Sc. in Mathematics. He is a full professor at the Francisco Jose de Caldas District University, Bogota Colombia. He can be contacted at email: haflorezf@udistrital.edu.co.



**Sunday Adeola Ajagbe**    is a Ph.D candidate at the Department of Computer Engineering, Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria and a Lecturer, a First Technical University, Ibadan, Nigeria. He obtained MSc and BSc in Information Technology and Communication Technology respectively at the National Open University of Nigeria (NOUN), and his Postgraduate Diploma in Electronics and Electrical Engineering at LAUTECH. His specialization includes Artificial Intelligence (AI), Natural language processing (NLP), Information Security, Data Science, and the Internet of Things (IoT). He is also licensed by The Council Regulating Engineering in Nigeria (COREN) as a professional Electrical Engineer, a student member of the Institute of Electrical and Electronics Engineers (IEEE), and International Association of Engineers (IAENG). He has many publications to his credit in reputable academic databases. He can be contacted at email: sunday.ajagbe@tech-u.edu.ng.