IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

COMPARATIVE STUDY OF AUTOENCODER AND DENSENET – 121IN CRIME PREDICTION

Ronisha Salome. A[1], Vijayanandh.T[2]

[1]Student, Master of Engineering, Department of ECE, National Engineering College, Kovilpatti, India.

[2]Assistant Professor, Department of ECE, National Engineering College, Kovilpatti, India.

ABSTRACT

A purposeful action that results in harm to one's person, harm to one's property, or both constitutes a crime. Depending on the offense's seriousness, the state or other authorities may punish the offender. Authorities are being forced to devise effective preventive measures due to the worrisome rise in both the quantity and types of criminal activity. In the current climate, where crime has sharply increased, traditional crime resolution approaches are sluggish and ineffectual. Hence, it will lessen the strain on the police and aid in crime prevention if you can figure out a means to foresee crime, predict it in detail before it occurs, or develop a "machine" that can support police personnel. The sole goal of this research is to determine how government agencies, including law enforcement, may use computer vision and machine learning to spot, stop, and combat crime. In conclusion, advancements in computer vision and machine learning can speed up the growth of law businesses. The proposed research uses neural networks with autonomous coding, decoding, and Artificial Intelligence to learn condensed versions of raw inputs, autoencoder-style neural network kinds can be employed. The branch of AI known as computer vision teaches computers to extract and decipher information from picture and video data. By altering the typical CNN design and streamlining the connecting pattern between layers, DenseNet - 121 finds a solution to this issue. Based on a comparison of autoencoder and DenseNet - 121 for crime prediction, this suggested workwas created.

Keywords: Crime, Preventing Measures, Traditional crime resolution techniques, Machine Learning, Artificial Intelligence, Autoencoder, Neural Network.

I. INTRODUCTION

By teaching them to comprehend and grasp the visual world, computers may learn to understand their surroundings thanks to the field of artificial intelligence known as computer vision. Due in large part to the fact that it analyses the camera's environmental data, it has significant application possibilities. It may be used for object identification, geolocation, augmented reality, license plate recognition, and mixed reality. Research is now being done to create mathematical methods for reproducing 3D images and making them understandable to computers. Having a 3D visualisation of an object allows for the effective use of eigenfaces and 3D shape models, as well as instance sensing, geometric placing, database technology, location detection, category detection, vocabulary, part-based models, recognition by segmentation, smart photo editing, context and scene understanding, extensive image collection and learning, image search, and image recognition. These are merely the most fundamental usage; the aforementioned categories can be looked into further.

Machine learning is a tool that enables machines to autonomously learn from the past and get better over time without explicit programming. It is not always easy to pinpoint certain patterns or bits of information after evaluating the data. Use ML to comprehend the precise pattern and information in such circumstances. By providing the correct data to computers, ML extends the notion that machines can learn and resolve intricate maths

issues as well as particular difficulties. Unsupervised machine learning and supervised machine learning are frequently used to categorise machine learning. Supervised learning trains a computer using a set of planned training examples to enhance its ability to make precise and accurate judgements when presented with new data. Unsupervised learning requires that machines analyze a dataset for recurring patterns and connections. Autoencoders use the unsupervised learning methodology.

II. RELATED WORKS

Researchers have put up a variety of strategies for foretelling crime. Nonetheless, there are still certain restrictions on this tactic. Hotspot Prediction may utilise the supported vector machine method to specify a crime rate level and provide the percentage of data. To classify each of the chosen data points in accordance with the predetermined degree of crime rate, a subset of the crime datasets is picked (based on a percentage or the total number of data crimes). Hotspots are defined as data points that surpass the predetermined rate, whereas non- hotspot classes are defined as data points that fall below the established rate. The selected parameter will be used to train an SVM for classification. To determine how effectively one-class SVM can forecast hotspot crime of location, linear, polynomial, and gaussian kernel functions have been utilised. The information comes from free public web databases. Despite the changes, computing this method still requires a lot of time and money.

To identify trends in the neighbourhood's crime rate. To cut processing costs, this method employs a streamlined procedure and model that make use of straightforward arithmetic operations. This strategy made use of crime statistics from the previous seventeen years. instances of murder in Delhi City). The variables included are therefore real reported instances and years. Segmenting the data into five intervals (Scheme I), ten intervals (Scheme II), and twenty intervals (Scheme III) produced three separate sets of data (Scheme III). The outcomes of Schemes II and III can be predicted with some degree of accuracy, while Scheme I has a tendency to overpredictwith an usual absolute error. (-0.323).

The ANN approach is formed by concentrating on geographic regions that perform better at forecasting crime than conventional police limits. To help with predictive modelling, an ANN may be learned using geographic groups of crime data. Clusters with relatively high levels of crime hotspots were found using a scanning technique based on neighbourhood crime incidents. A general use of ANN is to express the patterns linked to each cluster. In this investigation, a dataset of 18,498 violent incidents was used (criminal harm, violence against the person). It contained information on the time of day, the month in which the incident occurred, the site, and the weather. The outcome demonstrates how the ANN (9.94) and random walk's mean standard error (MSE) (22.50). Because of this, ANN predicts events more accurately than a random walk.

One of the Bayesian types of learning is used to forecast crime. With this technique, it is possible to create solid mathematical models that take repeat offenders' behaviour into consideration. Several elements might potentially have an impact on which offenders would commit crimes in the future. Using the serial murder dataset from Gansu, China, the model was evaluated. So, in addition to the crime hotspots, the parameters are those of the victims (age, gender, employment, and ethnicity) (residence, school, bus stop, hotel, and hospital). Also, the forecasted regions (coloured yellow, green, and red) will aid authorities in apprehending offenders depending on the geographical parameters chosen. Nevertheless, the success of the strategy is entirely dependent on the choiceof the parameter.

III. PRESENT TECHNOLOGY USED IN PREDICTING CRIMES

Crime forecasting is the process of foreseeing crimes before they take place. You need the necessary equipment to foresee crimes before they happen. Nowadays, police have resources at their disposal to help them with certain tasks, such as using a body camera to film an odd illegal behaviour or listening in on a suspect's phone call. We've included a few of them below to let you see how these technologies may do with more modern aid. The Stingray, a cutting-edge advancement in police surveillance, is an effective tool for keeping tabs on phones because it can be used to ascertain a cell phone's position by designed to simulate cell phone towers and broadcasting the signals to trick nearby cell phones into communicating their location and other information. The fourth amendment is one justification for opposing the use of stingrays in the US. In addition to the District of Columbia, 23 other jurisdictions use this technology. The authors address concerns about privacy violations by explaining how this is more than just a surveillance device. Additionally, the Federal Communications Commission got involved and eventually requested that the manufacturer fulfil two requirements in order to qualify for a grant: (1) The limitation on marketing and sales of these items is "Federal, State, Local Community Security and Law Enforcement Officers Only." (2) "State, municipal, and other law enforcement entities shall collaborate with the FBI before acquiring and using the equipment authorised under this paragraph." Although still a contentious topic, its application is helpful. The method known as "the stakeout" has been used frequently since the invention of the spy.

The most common sort of observation used by police to gather information on different suspect categories is stakeouts. According to the writers, who go on to highlight the importance of a stakeout, police officers are obligated to produce reports about a range of activities they witness. During stakeouts or patrols, these criminal activities are seen; during home searches, firearms, narcotics, and other evidence are discovered; and during arrests, descriptions of the suspect's and the officers' actions are provided. The most common sort of observation used by police to gather information on different suspect categories is stakeouts. According to the writers, who go on to highlight the importance of a stakeout, police officers are obligated to produce reports about a range of activities they witness. During stakeouts or patrols, these criminal activities are seen; during home searches, firearms, narcotics, and other evidence are discovered; and during arrests, descriptions of the suspect's and the officers' actions are provided. Stakeouts, in which the police personally observe crucial occurrences, are extremely beneficial and often regarded as being completely dependable. But are they totally accurate? Being human, all police officers experience the effects of fatigue.

Drones may be used, among other things, to map cities, find criminals, look into accident and crime scenes, control traffic, and support search and rescue efforts after a disaster. The distribution of airspace and legal problems related to drone use are examined. The public's privacy concerns and the police's expanding authority and influence both leads to legal issues. How high a drone can fly prompts questions about airspace distribution. Body cameras, license plate identification, and face recognition are additional surveillance techniques. In order to learn more about suspects, the authors claim that face recognition may be used to collect their profiles and analyse them using information from other sources. A licence plate scanner may also be used to locate a vehicle that may have been used in a criminal activity. The reader watches and records what the police officer sees since it's possible that they utilise body cameras to catch things that the naked eye can't. Typically, we can't recall everything we see as it happens. Body cams' effects on police misconduct and domestic violence during arrests have been researched. But, the use of body cameras does not stop here. One of the main reasons to always wear a body camera is to record what is going on in front of the user in order to capture significant or useful happenings while performing routine duties. While it's true that each of these strategies works well, they all have one thing in common: they all function independently. Even if the police can employ any of these strategies alone or simultaneously, having a device that can combine the advantages of all these technologies would be quite helpful.

IV. ML TECHNIQUES USED IN PREDICTING CRIMES

Using the freeware data mining programme Waikato Environment for Knowledge Analysis, trends for violent crime from the Societies and Crime Unnormalized Dataset were compared to real crime statistics (WEKA). Using datasets comprising actual crime and community data, three algorithms—linear regression, additive regression, and decision stump—were used. Participants in the experiment were chosen at random. The linear regression approach fared better than the other two methods because it could manage some unpredictability in the test data. The project's goal was to demonstrate how machine learning (ML) algorithms may be employed for a variety of tasks, including identifying criminal hotspots, developing criminal profiles, and identifying criminal trends, as well as to predict violent crime patterns.

Consider Complete Guide, a cutting-edge graphical user interface that can replace Internet Explorer, when thinking about WEKA. It offers a simpler way to comprehend the information. process orientation mixed with mining, where a particular data flow is graphically depicted utilising discrete learning components (represented by Java beans). The authors go on to provide an additional graphical user interface called the experimenter, which is intended to evaluate the effectiveness of different learning approaches on different data sets.

A predictive research application for predicting crime in an urban setting is being investigated. In the past, 200 m and 250 m grids of burglaries, house invasions, and robberies on the street were combined and analysed. Based on the crime data gathered over the previous three years, an ensemble model was utilised to incorporate the outcomes of logistic regression and neural network simulations in order to create biweekly and monthly projections for the year 2014. The straight hit rate, accuracy, and predictive index were used to assess the forecasts. The results of the twice-weekly projections show how accurate prediction is possible when using apredictive analytic method to the data. They came to the conclusion that by comparing the fortnightly predictions with the monthly projections with a distinction between day and night, the results would be considerably improved.

Using ML, predictions for crime were examined. Vancouver, Canada, looked at crime statistics from the previous 15 years to create estimates. Data collection, data classification, pattern recognition, prediction, and visualisation are all included in this machine learning-based crime research. The criminal dataset was further examined using boosted decision trees and K-nearest neighbour (KNN) algorithms. During their study, they looked at a total of 560,000 crime data from 2003 to 2018 and discovered that they could anticipate crimes with an accuracy range from 39% to 44% by applying ML algorithms. Despite the prediction model's poor performance, the authors came to the conclusion that it may be enhanced for some applications by changing the algorithms and crime data.

For the purpose of predicting crime-related data in Philadelphia, Pennsylvania, an ML approach is investigated. Three categories—determining whether a crime happens, criminal occurrence, and crime most likely to occur—were created. The datasets were trained using algorithms including logistic regression, KNN, ordinal regression, and tree methods in order to provide more accurate quantitative crime projections. Also, they provided a map for crime prediction that showed different crime categories in different Philadelphia neighbourhoods throughout the course of a certain time period, with different colours signifying each type of crime. To depict the overall trend of crime in Philadelphia during a certain time period, a variety of crimes, including assaults and cyber fraud, were included. They found that their approach had a remarkable 69% accuracy in predicting the possibility of a crime occurring and a 47% accuracy in predicting the number of crimes occurring, which could be anywhere between 1 and 32.

A number of crimes are included in a dataset that predicted the types of crimes that would occur in the near future based on a variety of factors. A criminal dataset from Chicago, Illinois, in the United States was used to anticipate crime using ML and data science methodologies. The crime dataset includes details about the occurrence, such as the kind of crime, the date, the time, and its specific location. The most accurate model was selected for training after experimenting with a variety of model combinations, including KNN classification, logistic regression, decision trees, random forests, an SVM, and

Bayesian methods. KNN classification proved to be the most accurate, with an accuracy of around 0.787. To comprehend the many features of the Chicago crime dataset, they also employed a number of visualisations. This article's major goal is to teach law enforcement organisations how to utilise machine learning (ML) to anticipate, detect, and solve crimes much more quickly, which will lead to a decrease in crime.

To properly predict the possibility of crime, it is advised to fuse multi-model data from several domains with outside context data using a deep neural network (DNN)-based feature-level data fusion approach. The collection includes information from an online database that includes statistics on Chicago crime, images, demographic and climatic information, and more. Many ML approaches, such as regression modelling, kernel density estimation (KDE), and SVM, are used in crime prediction strategies. They mainly used three steps in their strategy: data collection, statistical analysis of the relationship between criminal incidences and gathered data, and finally, accurate crime prediction. The DNN model is composed of spatial attributes, temporal characteristics, and surroundings. Fig. 4.1 depicts the data flow of the ML approaches.

FIG 4.1 DATAFLOW DIAGRAM

V. ALGORITHM DESCRIPTION

1. AUTO-ENCODERS

With neural networks, supervised learning is a typical application. It utilises training information with an output identification. The neural network tries to determine how to match the input to the label for the output. What if the output name was instead the input vector? After then, the network will try to determine a mapping from the input to itself. This transformation is the identity function, which is simple. But, if the network is not allowed to simply duplicate the input, it will be forced to only record the prominent features. This limitation exposes a potential use for neural networks that was not before known. Dimensionality reduction and particular data compression are the two main uses. First, the network is trained using the given input. Based on the information it has collected, the network creates an approximation of the input. the stages of training when the error rate and backpropagation are calculated. An auto-traditional encoder's design resembles a bottleneck.

The encoder component of the web is used for encoding and sometimes even for compression algorithms reasons, despite the fact that it is less efficient than other popular compression techniques like JPEG. Encoding is accomplished via the network's encoder component, which contains a decreasing number of hidden units in each layer. As a result, this section can only contain the most significant and representative aspects of the data. The second part of the network is responsible for decoding. This section makes an effort to reconstruct the original input by decoding the encrypted data using an increasing number of hidden units in each layer. As a result, auto-encoders are a technique for autonomous learning.

2. COMPUTER VISION ALGORITHM

It seems rather simple to give a machine-human intelligence and instincts. We frequently have a tendency to neglect the limitations of computers in comparison to our biological capabilities, perhaps because it can be solved by very young children as well. Even with just human intellect, the complexity of visual perception is infinitely variable and constantly changing. Our brain has the capacity to recognize the object, process the information, and make a decision, finishing a complicated task in a fraction of a second. The goal is to make it possible for computers to perform similar tasks. Because of this, it is a discipline that blends artificial intelligence and machine learning. It requires the teaching of methods and specific methods to understand what the computer sees. In the ever-expanding field of computer vision, specialized bespoke tasks and methodologies are utilised to focus on certain application areas. I envision it having a market value that grows along with its capabilities. Thanks to our brains and determination, we will soon be able to merge our abilities with computer vision and achieve newheights.

3. DENSENET - 121

A kind of convolutional neural network called a DenseNet employs Dense Blocks to build dense connections between layers by directly connecting every layer (with matching feature-map sizes) to every other layer. Each layer receives extra inputs from all levels that came before it and send its own map of characteristics to all layers that came after it in order to maintain the system's feed-forward nature. Convolutional and completely connected layers are examples of layers with trainable weights that fall under category 121. (batch norm excluded). To maximise data flow between network tiers, this is done. Each layer receives input from all stages that came before it and transmits its own image features to all layers that will follow it in order to maintain the feed-forward nature.

VI. RESULTS

Our main objective is to find the crime before it happens. Each image in the dataset used to create the model shows a unique form of criminal activity. The test datasets include representations of 14 different types of crimes; these datasets are gathered and put into the algorithm. 8.1% of the remaining data were used for testing, leaving 91.9% for training. The separation of the testing and training data is depicted in Fig. 6.1.

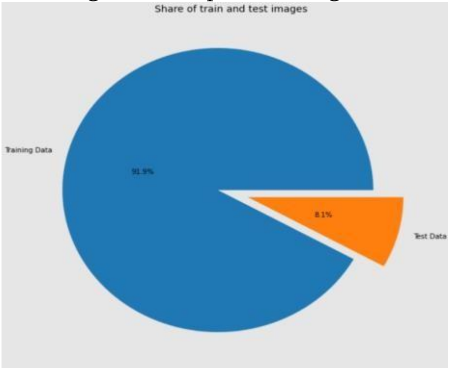

FIG 6.1 SHARE OF TRAIN AND TEST DATA

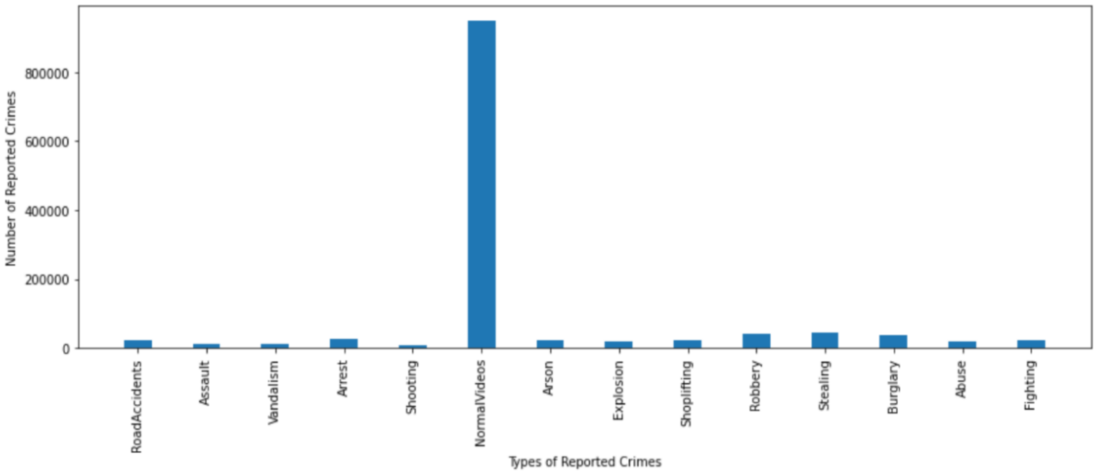

FIG 6.2 TYPES OF CRIMES IN TRAIN AND TEST DATA

As you can see in the above Output graph and chart, Fig 6.2 describesthe different types of reported crime datasets that are taken for our testingand training purposes. The above pie chart shows how much data are collected on particular reported crimes. Further, the AUC score will be calculated for all the reported crimes.

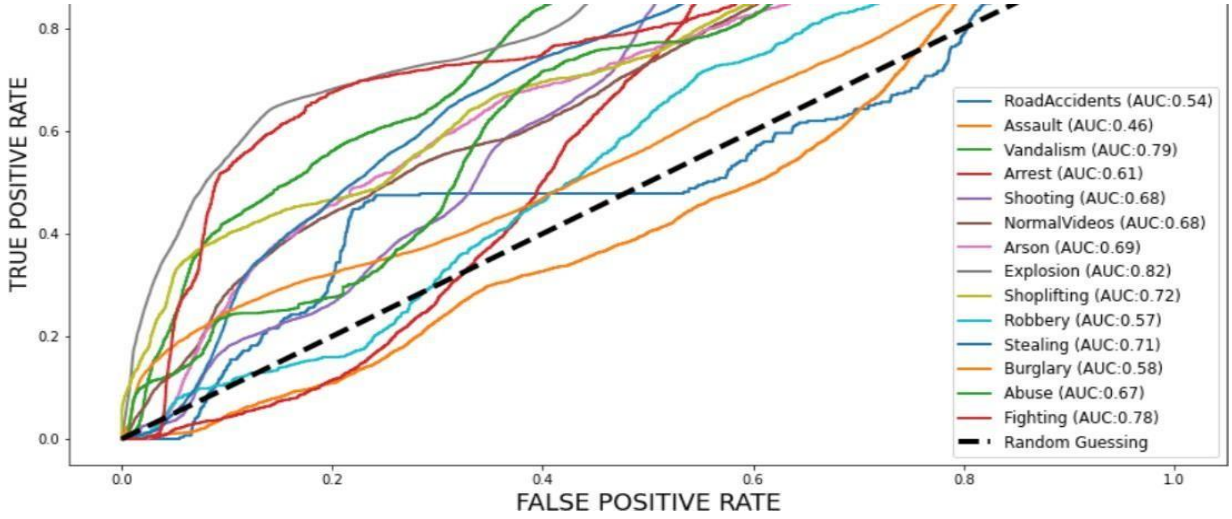

FIG 6.3 AUC SCORE GRAPH OF AUTO ENCODER ALGORITHM

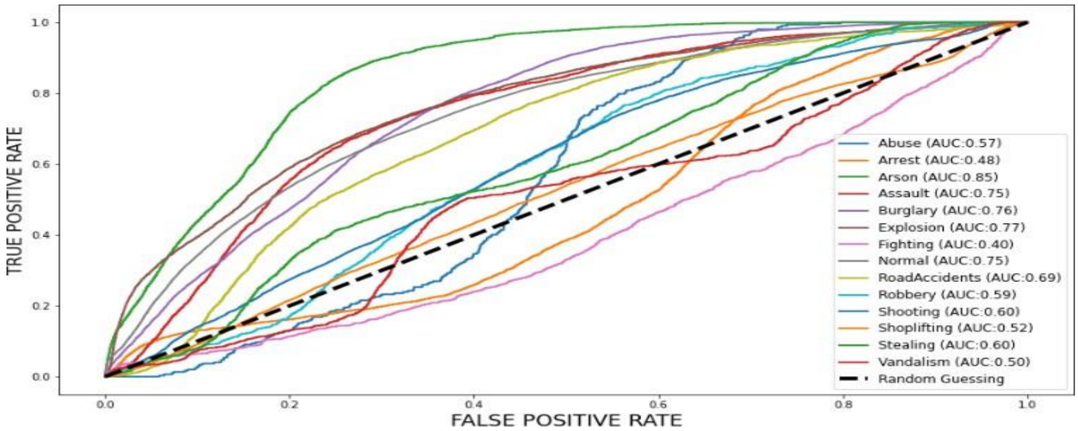

FIG 6.4 AUC SCORE GRAPH OF DENSENET-121 ALGORITHM

AUC represents the probability that a randomly chosen positive example (green) will be offered before a randomly chosen negative example (red). AUC ranges in value from 0 to 1. The AUC of a model with 100% erroneous predictions is 0.0, whereas the AUC of a model with 100% correct predictions is 1.0. The results of the AUC tests were assessed as exceptional for AUC values between 0.9 and 1, good for AUC values between 0.8 and 0.9, decent for AUC values between 0.7 and 0.8, terrible for AUC values between 0.6 and 0.7, and unsuccessful for AUC values between 0.5 and 0.6.

VII. CONCLUSION

The AUC score for the recorded crimes is displayed in the graph above. The end goal is to create a modelwith an overall AUC value between 0.8 and 0.9. Although the Auto-encoder's AUC score of 0.73 indicated thatthe model was reasonable, it cannot be used for general use. With DenseNet-121, the AUC score value was 0.84, which is a solid model by any standard. We can raise the AUC score number while enhancing the dataset.When Auto-encoder and DenseNet-121's AUC score graphs are compared, DenseNet-121's AUC score value isgreater than Auto-encoders. The DenseNet-121 can therefore be used to forecast crime.

REFERENCES

[1] Shah, N., Bhagat, N. & Shah, M. Crime forecasting: a machine learningand computer vision approach to crime Prediction and prevention. Vis. Comput. Ind. Biomed. Art 4, 9 (2021). https://doi.org/10.1186/s42492-021-00075-z

[2] P. Chen, H. Yuan and X. Shu, "Forecasting Crime Using the ARIMAModel," 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 2008, pp. 627-630, doi: 10.1109/FSKD.2008.222.

[3] A H Wibowo and T I Oesman 2020 *J. Phys.: Conf. Ser.* 1450 012076

[4] Mrinalini Jangra Ms. Shaveta Kalsi, “Crime Analysis for Multistate Network using Naive Bayes Classifier”, IJCSMC, Vol. 8, Issue. 6, June2019, pg.134 – 143

[5] Luiz G.A. Alves, Haroldo V. Ribeiro, Francisco A. Rodrigues, Crimeprediction through urban metrics and statistical learning, Physica A: Statistical Mechanics and its Applications, Volume 505,2018, Pages 435-443, ISSN 0378-4371, https://doi.org/10.1016/j.physa.2018.03.084.

[6] X. Zhang, L. Liu, L. Xiao and J. Ji, "Comparison of Machine LearningAlgorithms for Predicting Crime Hotspots," in IEEE Access, vol. 8, pp. 181302-181310, 2020, doi: 10.1109/ACCESS.2020.3028420.

[7] A. Shukla, A. Katal, S. Raghuvanshi and S. Sharma, "Criminal Combat: Crime Analysis and Prediction Using Machine Learning," 2021 InternationalConference on Intelligent Technologies (CONIT), 2021, pp. 1-5, doi: 10.1109/CONIT51480.2021.9498397.

[8] Chamola, Vinay & Hassija, Vikas & Gupta, Sakshi & Goyal, Adit & Guizani, Mohsen & Sikdar, Biplab. (2020). Disaster and Pandemic Management Using Machine Learning: A Survey. IEEE Internet of Things Journal. PP. 10.1109/JIOT.2020.3044966

[9] Shraddha Ramdas Bandekar, C. Vijayalakshmi, Design and Analysis of Machine Learning Algorithms for the reduction of crime rates in India, Procedia Computer Science,Volume 172, 2020, Pages 122-127, ISSN 1877- 0509, https://doi.org/10.1016/j.procs.2020.05.018

[10] Hossain, S., Abtahee, A., Kashem, I., Hoque, M.M., Sarker, I.H. (2020). Crime Prediction Using Spatio-Temporal Data. In: Chaubey, N., Parikh, S., Amin, K. (eds) Computing Science, Communication and Security. COMS2 2020. Communications in Computer and Information Science, vol 1235. Springer, Singapore. https://doi.org/10.1007/978-981-15-6648-6_22

[11] P. Sivakumar, J. V, R. R and K. S, "Real Time Crime Detection Using Deep Learning Algorithm," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), Puducherry, India, 2021, pp. 1-5, doi: 10.1109/ICSCAN53069.2021.9526393.

[12] U. V. Navalgund and P. K., "Crime Intention Detection System Using Deep Learning," 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 2018, pp. 1-6, doi: 10.1109/ICCSDET.2018.8821168.

[13] V. V. Nojor et al., "Design of a Deep Learning-based Detection System for Criminal Activities," 2022 3rd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey, 2022, pp. 1-5, doi: 10.1109/IISEC56263.2022.9998276.

[14] P. Mahesha, K. J. Royina, S. Lal, Y. Anoop Krishna and M. P. Thrupthi, "Crime Scene Analysis Using Deep Learning," 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 2021, pp. 760- 764, doi: 10.1109/ISPCC53510.2021.9609350.

[15] O. Llaha, "Crime Analysis and Prediction using Machine Learning," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 496-501, doi: 10.23919/MIPRO48935.2020.9245120.

[16] S. Chackravarthy, S. Schmitt and L. Yang, "Intelligent Crime Anomaly Detection in Smart Cities Using Deep Learning," 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), Philadelphia, PA, USA, 2018, pp. 399-404, doi: 10.1109/CIC.2018.00060.

[17] M. S. Nikhila, A. Bhalla and P. Singh, "Text Imbalance Handling and Classification for Cross- platform Cyber-crime Detection using Deep Learning," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-7, doi: 10.1109/ICCCNT49239.2020.9225402.

[18] H. Verma, S. Lotia and A. Singh, "Convolutional Neural Network Based Criminal Detection," 2020IEEE REGION 10 CONFERENCE (TENCON), Osaka, Japan, 2020, pp. 1124-1129, doi: 10.1109/TENCON50793.2020.9293926.

[19] J. Azeez and D. J. Aravindhar, "Hybrid approach to crime prediction using deep learning," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, 2015, pp. 1701-1710, doi: 10.1109/ICACCI.2015.7275858.

[20] P. Sivakumar, J. V, R. R and K. S, "Real Time Crime Detection Using Deep Learning Algorithm," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), Puducherry, India, 2021, pp. 1-5, doi: 10.1109/ICSCAN53069.2021.9526393.