Educational Administration: Theory and Practice
2024, 30(1) 4526-4532
ISSN:2148-2403
https://kuey.net/

Research Article

Computer Vision With DL/ML for Crime Data Analysis

Koffi Offori[1*], Dr. Mukesh Choudhary[2]

[1*]Student, PIET-MCA, Parul University, Vadodara, Gujarat, India, 2205112120039@paruluniversity.ac.in
[2]Associate Professor, PIET-MCA, Parul University, Vadodara, Gujarat, India,
mukesh.choudhary27542@paruluniversity.ac.in

Citation: Koffi Offori, et al (2024) Computer Vision With DL/ML for Crime Data Analysis, *Educational Administration: Theory and Practice, 30(1), 4526-4532*
Doi: 10.53555/kuey.v30i1.8226

ARTICLE INFO

ABSTRACT

Crime Analysis with Machine Learning and Deep Learning is considerably gaining researchers' attention and the incontestable integration of Computer Vision is currently taking another level of optimization to predict the crime of culprits intended to arm people.
These new features are the future of tracking crime perpetrators.
This work aims to showcase the interest in using machine learning, deep learning, and mostly computer vision to predict crime. Therein, it examines papers already showcasing an examination of plenty of articles that explore various machine learning, deep learning, and computer vision works. Plus, it highlights the accuracy of crime analysis by integrating computer vision with machine learning and/or deep learning.
This research will contribute to the rapidly evolving field of artificial intelligence to help law enforcement agencies in their daily difficult task.

Index Terms: deep learning, machine learning, crime analysis, and computer vision.

1.INTRODUCTION

Crime somewhat, is an illegal act or action which can lead to antisocial or wild behavior in a society. Criminal law interferes in this matter to punish the culprit to vehiculate a message as a beacon torch. Predicting crime comes in between to prevent such illegal actions from happening in the present time. The evolving ratio and the changing methods of crime require different and updated analytic methods to solve and intercept criminals and prevent their deemed ugly intentions. As the availability of datasets continues to rise for the benefit of good use researchers are being called on instinctively on the use of the rising technologies like machine learning, deep learning, and computer vision for a better crime analytic. [1-2] These papers examine the latest developments in machine learning, deep learning, and/or computer vision for crime prediction and how they are being utilized to identify criminal behaviors and prevent them.

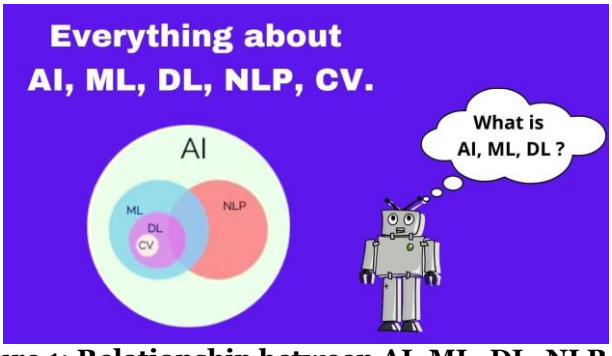

Figure 1: Relationship between AI, ML, DL, NLP, CV

(1) Machine learning, deep learning, and computer vision are subfields of artificial intelligence. Machine learning uses algorithms trained on datasets to create models to empower machines to operate tasks such as analyzing data, and predicting crime. On the other hand, deep learning and computer vision are both subsets

of machine learning. Machine learning enables computers to learn from data without being explicitly programmed. It uses mathematical methods to find patterns and make predictions or classifications based on data. (2) Deep Learning uses multiple-layered neural networks to mimic the human brain's ability to learn from data and patterns. The more layers a neural network has the more complex and abstract concepts it can learn and represent such as face speech and emotion recognition. Computer vision uses machine learning and neural networks to teach computers to gain meaningful information from digital images and videos. [3] Lately, many researchers have conducted experiments to predict crime harnessing lots of machine learning methods. Predicting crime has some common methodologies like Extra tree Classifier, K-Neighbor Classifier, Support Vector Machine (SVM), Decision Tree Classifier, and Artificial Neural Network (ANN). [2] Deep learning algorithms, such as convolution and recurrent neural networks are also successful in the crime prediction landscape. [4-5] Computer vision and deep learning algorithms are in the trade of crime prediction for predicting criminal acts or alerting law enforcement agencies in case of turbulent motion detected over video streams using methods like 3D Convolutional Neural Networks (3D CNN) to detect fight or violent senses. Therein, machine learning, deep learning, and computer vision are a great land of promise where researchers can get the necessary equipment to help law enforcement agencies in their endeavors against criminal actions. Thus this paper will be divided into three parts, first, showcasing the most recent trend studies of machine learning, computer vision, and deep learning from what others have already accomplished related to crime prediction. Second, showcasing what are the benefits of using machine learning via a small scalable prediction on the dataset Crime Data Set from Kaggle, a crime report from Seattle, on the focus on the AGGRAVATED ASSAULT-DV subcategory by leveraging a logistic regression model within the framework of machine learning. Third, the paper demonstrates a study based on deep learning and/or computer vision to intercept culprits.

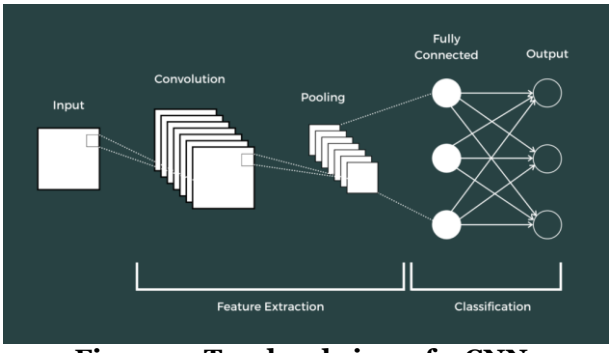

Figure 2: Top-level view of a CNN.

2.METHODOLOGY AND STRUCTURE

Our primary goal is to showcase how the use of machine learning, deep learning, and computer vision can help law enforcement to catch culprits. In that vision, we defined a title for this paper, discussed it with the supervisors, and had the right title which was "computer vision with AI/ML for crime data analysis. 1 Afterward, we searched for papers related to crime analysis, machine learning, computer vision, and deep learning of more than 20 papers and filtered only 9 fully read which brought us to the conclusion of redefining the title to "Computer Vision with DL/ML for crime data analysis". Based on that amount of information we made the introduction and the abstract and set our goal behind writing this research paper. Alongside defining the goal of this article we also Identified the different sub-titles after the abstract and introduction. Then our next focus was to find more papers in Google Scholar based on each subtitle and the information needed. This method is efficient in narrowing down the exact papers, not only the exact papers but also the information needed. **Table 1:** Papers read

Title	Authors	Journal/Conference	Year
A CNN-RNN Combined Structure for Real-World Violence Detection in Surveillance Cameras	Soheil Vosta and Kin-Choong Yow	MDPI	2022
A shallow 3D convolutional neural network for violence detection in videos	Naz Dündar, Ali Seydi Keceli, Aydin Kaya, Hayri Sever	Egyptian Informatics Journal	2024
Crime forecasting a machine learning and computer vision approach to crime prediction and prevention	Neil Shah, Nandish Bhagat, and Manan Shah	SpringerOpen	2021
Crime Prediction and Analysis	Pratibha, Akanksha Gahalot, Uprant, Suraina Dhiman, Lokesh Chouhan	ResearchGate	2020
Crime Prediction Using Machine Learning and Deep Learning A Systematic Review and Future Directions	Varun Mandalapu, Lavanya Elluri, Piyush Vyas, Nirmalya Roy	IEEE	2023
Illegal Activity detection on bitcoin Transaction using Deep Learning	Pranav Nerurkar	Springer Link	2023
Illegal Community Detection in Bitcoin Transaction Networks	Dany Kamuhanda, Mengtian Cui, Claudio J. Tessone	MDPI	2023
Machine Learning in Crime Prediction	Karabo Jenga, Cagatay Catal, Gorkem Kar	Springer Link	2023
Self-Supervised Graph Transformer for Deep fake Detection	Aminollah Khormali, Jiann-Shiun Yuan	arXiv	2023

3. AVAILABLE TECHNOLOGY

Policing Technology is a core of police operations. Law enforcement agencies need more accurate technologies to prevent and predict crimes, not only solve them. In [6] Nola M. Joyce explains the availability of technologies and even sends us to a future where police should be able to use those future technologies. Today police are using surveillance cameras, gunshot detection systems, automated license plate readers, facial recognition software, body cameras, drones, and numerous databases to investigate crime scenes. But that's not all, those data are a large amount of data to be used for future references. Those technologies cited before can help researchers apply their forensic knowledge via the data collected from these technologies utilizing some of their techniques. Therein, [7] in crime prediction applied machine learning we can identify some techniques with two basic ones: supervised and unsupervised learning. Moreover, supervised learning can be classified into two categories: classification and regression whereas unsupervised learning analyzes and clusters data into unlabeled datasets. In [7] they conducted a table research showing the tools used in their selected papers and WEKA is the most utilized tool after Python in second position then R-studio technologies and so on. [8] describes four types of machine learning: supervised, unsupervised, semi-supervised, and reinforcement instead of the two basic techniques supervised and unsupervised learning. They also describe numerous AI methods applied by scientists in their research.

4. MACHINE LEARNING FOR CRIME PREDICTION (TREND)

This section highlights recent papers related to machine learning and their results for their findings. Researchers concentrated on machine learning for their research and its algorithms but some study those research mostly on crime rates [9] to help crime researchers support their future research. They even show why crime prediction is being used so widely due to its relation with society. Also, machine learning algorithms are successfully able to predict spatial crime information. [9] they even surveyed crime prediction research works with machine learning shown on a table from 2014 to 2021 showing how machine learning is being successful in crime prediction. [10] is involved in sicking the most accurate algorithm that is capable of predicting crime in Dubai. Predicting crime is not an easy task but the most complicated is to have a very good accuracy with a large amount of data, sometimes a benchmark dataset to collect more information as needed. They have to operate comparison models on a dataset of sample crimes in the Emirate of Dubai, United Arab Emirates using the open-source data mining, software WEKA, which enabled them to use Random Forest, KNN, SVM, ANN, Naïve Bayes, and Decision Tree. [11] With classic forensic research on crime using machine learning methods to identify crimes based on previous datasets to predict the type of crime in a city most important is to identify the features that make these predictions. This research experiments supervised learning to classify the type of crime then decision tree (DT), random forest (RF), and K-nearest
neighbor (KNN) algorithms are utilized along with the Python programming language in the Jupyter Notebook environment where the random forest was the accurate algorithm with 86.07%. In the same vein, [12] main aim is to predict crime cases from 2017 to 2020 by using the dataset from 2001 to 2016 using machine learning techniques.

5. COMPUTER VISION AND/OR DEEP LEARNING FOR CRIME PREDICTION (TREND)

Like the first section highlighted recent papers related to machine learning this section will be focused on Computer vision and/or Deep learning which are two subsets of Machine learning. From IBM's definition, computer vision is the field of Artificial Intelligence that uses Machine learning and neural networks to teach computers and systems to derive meaningful information from digital images, videos, and other digital inputs and to make recommendations when they see defects or issues. [13] Along this line, scientists took upon themselves to address the issue of gun violence via CCTV by bridging the gap between criminology and computer vision. [14] Some researches are based on stalking detection to prevent crime from happening by leveraging CNN, LSTM, and MLP for a hybrid fusion model. Plus they did a great job by achieving a percentage of 89.58 testing accuracy. This approach helps prevent crime from happening to its start. [15] Using MLOps techniques to identify vehicle plate numbers, person detection through image and video analysis, and object behavior detection through image and video analysis the paper proposed a framework using MLOps techniques to prevent crime.

6. AGGRAVATED ASSAULT-DV PREDICTION USING MACHINE LEARNING

This part is subject to a subcategory of a dataset download from kaggle "Crime data set a crime report from seattle". This dataset contains approximately 481,000 crime reports from Seattle, WA covering a span of approximately 10 years. The dataset contains date and time information, crime categories and descriptions, and police department information including sector, beat, precinct, and neighborhood name. This dataset is used to predict the occurrence of crime AGGRAVATED ASSAULT-DV subcategory using machine learning with a model training Logistic regression model as machine learning is suitable for binary classification tasks.

6.1 METHODOLOGY, TECHNOLOGY

Data collection,
Dataset selection utilizes the Crime_Data.csv dataset.
Data cleaning,
(3) Handle missing values, categorical columns, and inconsistencies to ensure the dataset's quality.

```
[ ] # Handling missing values
    #numeric columns
    numerical_col=['Occurred Time']
    imputer_numeric=SimpleImputer(strategy='mean')
    df[numerical_col]=imputer_numeric.fit_transform(df[numerical_col])

    # categorical columns
    categorical_col=['Crime Subcategory','Sector','Beat','Precinct']
    imputer_categorical=SimpleImputer(strategy='most_frequent')
    df[categorical_col]=imputer_categorical.fit_transform(df[categorical_col])
```

Figure 3: missing values and categorical columns

Feature selection,
Identify and select features that contribute significantly to predicting Crime Subcategory_AGGRAVATED ASSAULT-DV.
Target definition,
Binary Target Variable: Define Crime Subcategory_AGGRAVATED ASSAULT-DV as the binary target variable, where 0 indicates the crime will not occur, and 1 indicates the opposite.
Model training,
Logistic Regression: Choose Logistic Regression as the machine learning algorithm, suitable for binary classification tasks.
Training set,
Split the dataset into training and validation sets, reserving a portion for model training.
Validation set,
Evaluate the model on the validation set to ensure its ability to generalize to unseen data.
Visualization,
(4) Developed a bar chart to visually represent the predicted probabilities of Crime Subcategory_AGGRAVATED ASSAULT-DV with different colors or shades to distinguish between predictions of 0 and 1.
Findings interpretation,
Interpret the results of the logistic regression model and the bar chart, emphasizing the probability distribution and the model's predictive performance.

Conclusion:

In the next picture we are going to discover a barchart to count the number of times Crime Subcategory_AGGRAVATED ASSAULT-DV could occur.
The bar chart shows the occurrence of Crime Subcategory_AGGRAVATED ASSAULT-DV is very low via the label 0.

```
import seaborn as sns
import matplotlib.pyplot as plt

ax=sns.barplot(x=lable_count.index,y=lable_count.values)
plt.title('ditribution of predicted Crime Subcategory_AGGRAVATED ASSAULT-DV')
plt.xlabel('Predicted lable')
plt.ylabel('Count')

ax.legend(labels=['Lable 0 : Will Not Occurre','Lable 1: Will Occurre'])

plt.show()
```

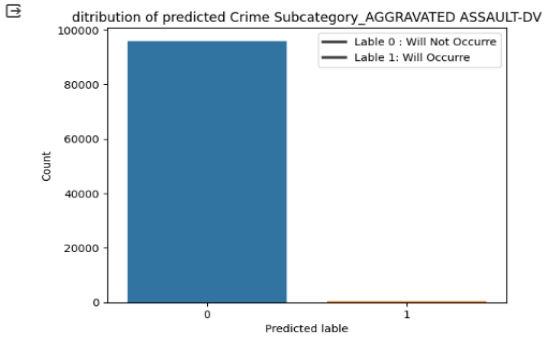

Figure 4: Distribution of predicted crime Subcategory_Aggravated Assault-DV

7. DEEP LEARNING AND/OR COMPUTER VISION FOR CRIME PREDICTION

Crime prediction using computer vision and/or deep learning is on the course of revolutionizing the game of solving crime in real life. In our research, we demonstrated the power of using computer vision to detect crime prediction. In this section we will showcase the latest tools and technologies used for crime prediction in computer vision and, by the end demonstrate with pictures or charts the training made on the dataset for Handgun detection using the YOLOv8 model on roboflow.
First, we used open images dataset v7 to collect data already labeled and the dataset used is the Handgun dataset. To collect the dataset you need to follow some steps but before we create a virtual environment, a virtual room to process our dataset.
Next, we clone the OIDv4 repository from GitHub to a directory of our choice in our machine then we download the dataset Handgun following a procedure (5).

```
create virtual environment
python -m venv virtual_environment_name

activate venv
.\virtual_environment_name\Scripts\activate

clone OIDv4 directory on your prefer derectory
git clone https://github.com/EscVM/OIDv4_ToolKit.git

change directory to OIDv7 and install all requirements
pip install -r requirements.txt //-r because we need to read the requirements inside of this file requirements.

download train dataset
python main.py downloader --classes Handgun --type_csv train --multiclass 1 --limit 800
or
python main.py downloader --classes Handgun --type_csv train --limit 400

download test dataset
python main.py downloader --classes Handgun --type_csv test --multiclass 1 --limit 160

go on classes.txt file and change
Apple,
Orange,
Light switch classes into the one you downoaded in my case it's Handgun

Then go on prompt
python convert_anotations.py
```

Figure 5: Process to clone and download the dataset from open image dataset V7

Figure 6: Sample of dataset used

(6). Moreover, we fetch that dataset to roboflow platform to convert it into yolov8 format and split it into train valid and test datasets. The training dataset stored 280 images with a percentage of 70% of data, the valid set got 80 images with a percentage of 20% of data and the test set got 40 images with a percentage of 10% of data, and an initial dataset of 400 images. After a training detection on the Handgun dataset using roboflow we got an accuracy of 86.9% mAP which is equal to the average of the Average Precision metric across all classes in a model (7),(8). A precision of 92.7%. Precision measures how often your model's predictions are correct. And a recall of 75.2%. Recall measures what percentage of relevant labels were successfully identified. We applied preprocessing using auto-orient and resized it to 640x640 but we did not apply augmentation.

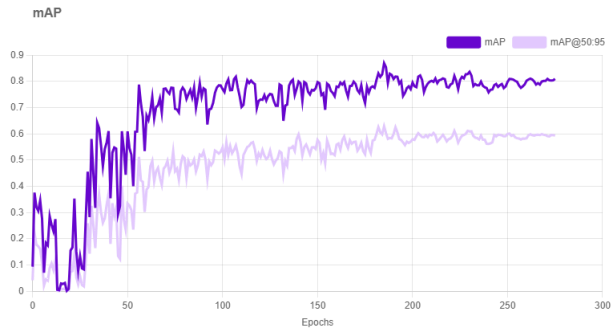

Figure 7: mAP visual chart evolution during training.

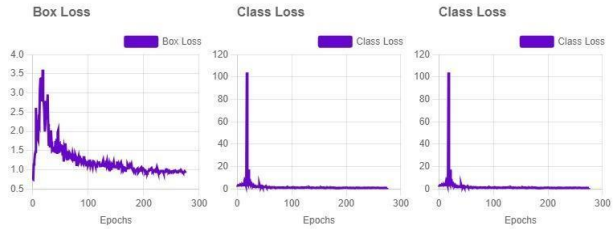

Figure 8: Loss of box, class, and object during training.

In the end, we can detect guns if there is a gun in some places the model can detect gun presence. One of the best things about roboflow is there is a QR code generated to use the model using a cell phone camera to detect any presence of a gun.

8. CONCLUSION FUTURE WORK

This study covers an overview of crime data analysis using machine learning computer vision research. In this paper, we reviewed the important factors of crime detection and prediction by the use of crime data analysis using machine learning (Logistic regression) and computer vision (YoloV8). This paper showcases the trends of Machine learning, Computer vision and/or Deep learning as well. With these two research based on machine learning (Logistic regression) and computer vision (YoloV8), we first conclude that machine can predict crime if we have access to previous crime scene datasets using a logistic regression model showing the occurrence of a crime following a bar chart 4. Then, using the YoloV8 model on roboflow we were able to detect gun presence in any area which can lead to automatic announcements to law enforcement agencies about a suspicious object in an area where the CCTV is visualizing.

This paper is limited by the various types of crime occurrence. [16] This paper extracts nineteen features from the Bitcoin network and proposes a deep learning-based graph neural network model using spectral graph convolutions and transaction features to detect illegal transactions on Bitcoin. The current article is also limited by the incapability of predicting [17] moral/psychological crimes. Plus, there are numerous objects to detect for criminal activity other than guns. Finally, the amount of dataset leverage for this research using YoloV8 can be considered not sufficient enough to conclude that this model is well accurate.

ACKNOWLEDGEMENT

We thank, Dr. Priya R. Swaminarayan, Dean - Faculty of IT and Computer Science, Director - Master of Computer Applications, for her visionary leadership and continuous encouragement. We also extend our heartfelt gratitude to Prof. Vivek Dave, HoD - Faculty of IT and Computer Science, for his insightful ideas and constructive feedback. Last but not least we thank Dr. Mukesh Choudhary Associate Professor, my research guide - Faculty of IT and Computer Science, for his expert supervision, patience, and mentorship.

REFERENCE

1. Varun Mandalapu, Lavanya Elluri, Piyush Vyas and Nirmalya Roy "Crime prediction using machine learning and deep learning: A systematic review and future Directions" *IEEE, Received 10 April 2023, accepted 11 June 2023, date of publication 14 June 2023, date of current version 20 June 2023.*
2. Neil Shah, Nandish Bhagat, and Manan Shah "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention" *Shah et al. visual computing for industry, biomedicine, and Art, 2021 4:9*
3. Pratibha, Akanksha Gahalot, Uprant, Suraina Dhiman, Lokesh Chouhan "Crime prediction and analysis" *ResearchGate February 2020*
4. Naz Dundar, Ali Seydi Keceli, Aydin Kaya, Hayri Sever, "A shallow 3D convolutional neural network for violence detection in videos", *Egyptian informatic journal of 26 (2024) 100455*
5. Simone Accattoli, Paolo Sernani, Nicola Falcionelli, Dagmawi Neway Mekuria & Aldo Franco Dragoni, "Violence Detection in Videos by Combining 3D Convolutional Neural Networks and Support Vector Machines", *Pages 329-344, Published online: 06 Feb 2020*
6. Technology and police operations written by Nola M. Joyce Fmr. Dep. Commissioner, Philadelphia Police Department
7. [7] Fatima Dakalbab, Manar Abu Talib, Omnia Abu Waraga, Ali Bou Nassif, Sohail Abbas, Qassim Nasir Artificial intelligence & crime prediction: A systematic literature review *Social Sciences & Humanities Open*
8. Abdulrahman Abdullah Alsubayhin, Bander Alzahrani, Muhammad Sher Ramzan Review of Crime Prediction Through Machine Learning *International Journal of Intelligent System and Application in Engineering*.
9. Ruaa Mohammed Saeed and Husam Ali Abdulmohsin A study on predicting crime rates through machine

learning and data mining using text *De Gruyter March 31, 2023*

10. Shaikha Khamis AlAbdouli, Ahmad Falah Alomosh, Ali Bou Nassif, Qassim Nasir Comparison of Machine Learning Algorithms for Crime Prediction in Dubai *ResearchGate*
11. Betelhem Zewdu Wubineh Crime analysis and prediction using machine-learning *SpringerLink*
12. approach in the case of Hossana Police Commission *Springer Link*
13. Sridharan, Srish, Vigneswaran and Santhi Crime Prediction using Machine Learning *ResearchGate Received on 01 December 2023, accepted on 04 February 2024, published on 15 February 2024*
14. Tyler E. Houser, Alan McMillan & Beidi Dong Bridging the gap between criminology and computer vision: A multidisciplinary approach to curb gun violence *Springer Link*
15. Murad Hasan, Shahriar Iqbal, Md. Billal Hossain Faisal, Md. Musnad Hossin Neloy, Md. Tonmoy Kabir, Md. Tanzim Reza, Md. Golam Rabiul Alam, Md Zia Uddin A Computer Vision Based Approach for Stalking Detection Using a CNN-LSTM-MLP Hybrid Fusion Model *arxiv Submitted on 5 Feb 2024*
16. Akash Kathole Sagar Shinde Lalitkumar Wadhwa Integrating MLOps and EEG Techniques for Enhanced Crime Detection and Prevention, *Malque*
17. [16] Pranav ajeet nerurkar Illegal Activity detection on bitcoin Transaction using Deep Learning *Research Gate*
18. Yilan Wu The impact of criminal psychology trend prediction based on deep learning algorithm and three-dimensional convolutional neural network *Springer Link*
19. *M. Choudhary, P. S. Solanki, V. Gamit and M. Joshi, "Machine Learning Classifier Used to Diagnosis of Liver Disorders," 2024 Parul International Conference on Engineering and Technology (PICET), Vadodara, India, 2024, pp. 1-6, doi: 10.1109/PICET60765.2024.10716100. keywords: {Analytical models;Accuracy;Machine learning algorithms;Liver diseases;Buildings;Medical services;Forestry;Classification algorithms;Random forests;Medical diagnostic imaging;K-Star;PART;Random forest;Random Committee;CNN},*