International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 12 | Dec 2023 www.irjet.net p-ISSN: 2395-0072

Real Time Crime Detection using Deep Learning

Kowshik, Shoeb, Dr.Y.Rama Devi

Kowshik & Student,Dept of AI&ML,Chaitanya Bharathi Institute of Technology, Hyderabad, Telanagana
Shoeb & Student,Dept of AI&ML,Chaitanya Bharathi Institute of Technology, Hyderabad, Telanagana
Dr. Y. Rama Devi& Professor,Dept of AI&ML,Chaitanya Bharathi Institute of Technology, Hyderabad, Telanagana

---***---

Abstract - *Crime detection and prevention have always been critical concerns for society. With the rapid advancements in technology, particularly in the field of deep learning and artificial intelligence, new opportunities have emerged to enhance real-time crime detection capabilities. Deep learning techniques have demonstrated remarkable potential in analysing various types of data sources, such as surveillance footage, sensor data, and social media feeds, to identify criminal activities, predict incidents, and aid law enforcement agencies in proactive responses. This survey paper seeks to provide a thorough overview of the application of deep learning in real-time crime detection. This paper digs into the various deep learning architectures and methodologies employed for this purpose, explore the challenges and limitations associated with these techniques, and discuss ethical and legal considerations.*

Key Words*:* ***crime detection, crime datasets, deep learning, CNN, LSTM.***

1. INTRODUCTION

The number of recorded criminal incidents is increasing, including robbery, vandalism, assault, murder, and kidnapping. However, the conventional methods of criminal investigation and prevention are often labor-intensive and inefficient. To address this challenge, a novel deep learning-based system named "Spot Crime" has been proposed. Spot Crime is a web application designed to enhance public safety and support law enforcement efforts by automating the monitoring of live CCTV footage and alerting police officials to suspicious activities. This innovative system employs custom convolutional neural networks (CNNs) and advanced models for behaviour classification, allowing it to analyze human activity in real-time video frames. The selection of the best-performing model is based on validation accuracy, and its purpose is to aid in the early detection of criminal incidents. In light of the increasing crime rates in India, especially during lockdowns, the need for such a system has become paramount. Spot Crime offers a solution to the challenges posed by manual surveillance, providing a more effective means of crime prevention and promoting public safety by promptly notifying authorities of potential criminal activity and its location [1].

YOLO,stands out as an effective object detection algorithm using a single convolutional neural network to predict bounding boxes and probabilities. It relies on a sizable training dataset for optimal performance. Google Collaboratory with an integrated Tesla K80 GPU is preferred for efficient GPU utilization. The trained YOLO model excels in detecting numerous objects even in complex scenes, striving to improve Mean Average Precision and minimize average loss for accurate object recognition. [2] The YOLOv5 architecture has emerged as a promising choice for real-time facial recognition. In this research paper, a comparative analysis was conducted, pitting YOLOv5 against its predecessors (v3 and v4) for real-time facial recognition systems. The experiments revealed an 87% accuracy rate for YOLOv5 on the FDDB dataset and an impressive 94% accuracy on a custom real-time face recognition dataset, significantly outperforming its predecessors. This underscores the potential of deep learning in the realm of facial recognition, suggesting that with the right direction and approach, deep learning and AI have much more to offer in the future. [3]. Public information resources are important in crime prediction. As a result, they gathered historical data using information resources such as news websites. They gathered information about occurrences that occurred around the country at a specific time and used classification to forecast future crimes. In the process, classification methods such as Support Vector Machine (SVM), Decision Tree, Random Forest, and logistic regression are used. Furthermore, the system uses all of the collected data to predict crime activity and visually displays regions with a higher risk of crime occurrence on a map. These forecasts can be used to improve emergency response by allocating resources depending on demand, as well as to prevent crime [4].

The present algorithms successfully detect items in certain annotated photographs, but they require locations, classes, and background distributions. When the objects were manually annotated, however, the assignment process became onerous and time-consuming. The former sliding window object identification approach's handcrafted characteristics had limitations that made it impossible to reliably recognise the items. [9] Furthermore, CNN outperformed the traditional technique in object detection. However, due to challenging situations such as object occlusion, increased variation in object scale, and weak lighting, the CNN detector was unable to achieve acceptable accuracy. [5]

[8] Deep Learning pre-trained models are built models that assist users in learning about algorithms or

experimenting with existing frameworks for better outcomes without explicitly developing. In addition, deep learning neural networks have five layers, which include input and output layers with Convolution, Max-Pooling, and Fully connected layers. Because of limited time, memory, and resources such as CPUs and Processors, many people prefer Deep Learning pre-trained concepts. And, when compared to machine learning, which requires us to construct explicitly, these pre-trained models will provide the best and most accurate outcomes. There is a significant amount of human interaction required to detect firearms in surveillance videos, which is prone to human error. It is required to design an automatic surveillance system that detects firearms with less computing time in order to limit crime incidence.[10]

2. Literature Survey

Samundeswari S[1] proposed Real-time fingerprint and criminal facial detection. Since the creation of CNN, real-time object identification has significantly advanced, leading to the development of other object detection techniques, including YOLO, SSD, and others. CNN performs a rather good job of detecting guns and criminal activity. The fastest performer is NN, taking 44 milliseconds to complete one image. During training, the accuracy of the model was 90.7 percent, and the loss was 23.88 percent. Validation accuracy for the model is 76.05 percent, and validation loss is 66.80 percent. Adam is the optimizer that's been used to make the model better. The model's epoch value is thirty.

P. Sivakumar[2] proposed the following system which make use of detecting crime scene objects like weapons. Based on that it will classify frames as suspicious or not.

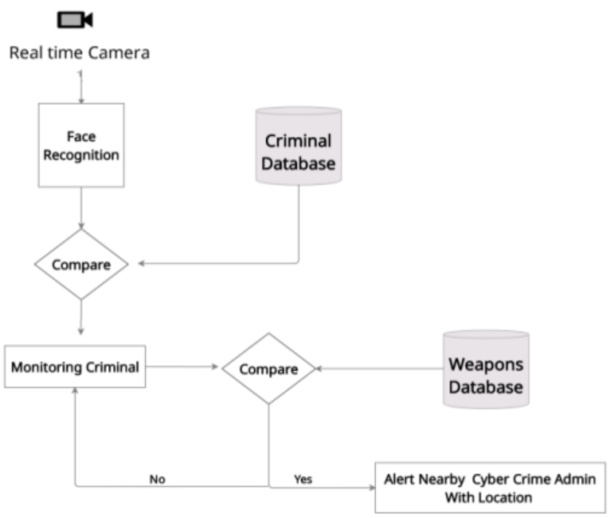


Fig.1. Architecture Diagram for proposed system

F. Majeed[3] predicted crimes using a combination of deep learning and machine learning. Regression analysis was employed to model the correlation among weather, demographic, and theft crime data. To estimate the likelihood of stealing offences in urban populations, the regression model uses two deep learning models: a Spatio-Temporal Graph Convolutional Network (ST-GCN) and a Long Short-Term Memory (LSTM) network.

C. Rajapakshe[4] used LSTM to encode the temporal features as RNN's are ideal solution for encoding the sequential data in videos.

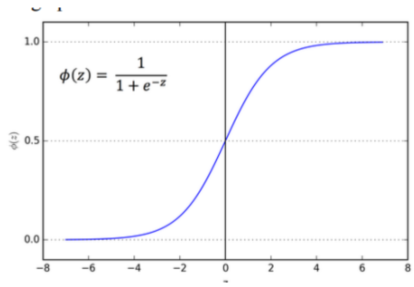


Fig.2. Sigmoid Function

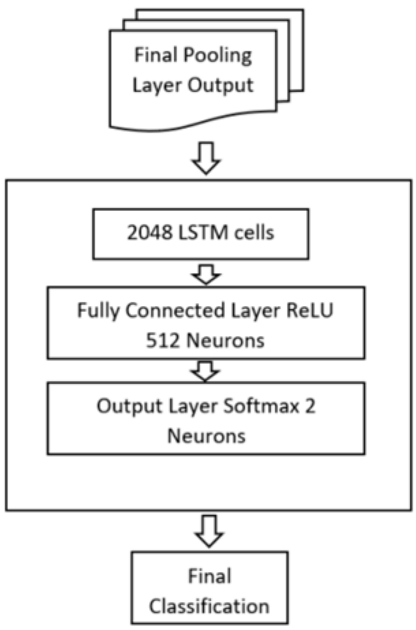
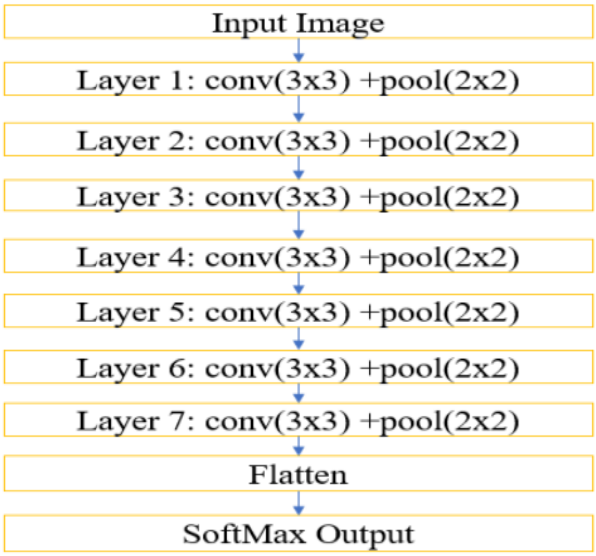


Fig.3. Architecture

They used sigmoid function for building suspicious human identification model.

Nandhini T J[5] proposed CNN architecture as displayed in below figure. Trained the model with 146 images and predicted the type of tool used for crime.



CNN NETWORK ARCHITECTURE

VARUN MANDALAPU[6] applied machine learning and deep learning algorithms to forecast crime. They modelled the association between stolen crime data, demographic data, and weather data using regression. To estimate the risk of theft offences in urban populations, the regression model employs two deep learning models: a Long Short-Term Memory (LSTM) network and a Spatio-Temporal Graph Convolutional Network (ST-GCN).

WAJIHA SAFAT[7] presented a crime prediction and analysis system that predicts crime in a given location using machine learning techniques. To convert the data into a format appropriate for analysis and detection, the authors adopted a pre-processing procedure. They concentrated on predicting the type of crime that may occur based on the location.

Sharmila Chackravarthy[8] introduced Intelligent Crime Anomaly Detection, which employs CNN, RNN, and Hybrid deep learning algorithms to extract high-performance characteristics in each frame. A facial recognition system with a high accuracy rating is built in phases. Using the object tracking method, abnormal behaviours are discovered using the DCNN and RNN networks. Using the HDL method, DCNN retrieves the higher performance attributes from the frames.

Uma. N[9] proposed Deep Convolutional Generative Adversarial Networks for Crime Scene Object Detection. The author used two-stage architecture for employing Perceptual GANs to detect Crime objects, consisting of Generator, Discriminator, Feature Pyramid Network (FPN), Attention Mechanisms.

Umadevi V Navalgund[10] proposed Crime Intention Detection

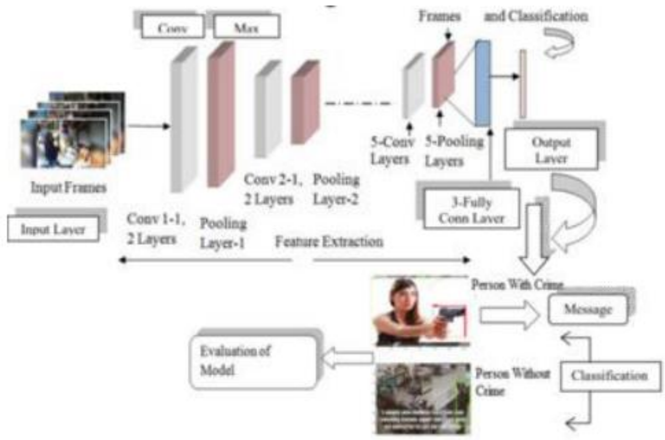


Fig.5. Architecture for Crime Classification

3. Methodology

The primary aim is to select best algorithm for the model building. On studying about features and implications of various we chosen 3-4 models. And the Evaluation Metrics defined in different papers are also carefully studied and examined and consider it for the evaluation for our model.

3.1 Data Collection

For this task we have chosen Kaggle UCF crime dataset of violent and non-violent situation dataset and UCF crimes. The first dataset consists of 1000 non-violence dataset and 1000 violence datasets. The second dataset consists of 14 types of crime each folder consists of 50 videos for each crime.

3.2 Preprocessing

- Clean and preprocess the data to remove noise and inconsistencies. This step might involve data augmentation, normalization, and other techniques to make the data suitable for deep learning models.
- Data Augmentation can also be implemented
- The input will be multiplied by a factor known as "rescale" before any further processing is performed. The RGB values of our original images range from 0 to 255, but at a typical learning rate, these values would be too high for our models to comprehend. To achieve values between 0 and 1, we scale our original photographs by a factor of 1/255.
- Shearing transformations can be applied at random with shear range

3.3 Training Model using Deep Learning Techniques

1) Customized CNN:

CNN (Convolution Neural Network) is a well-known algorithm for crime detection. It takes a picture as input and distinguishes distinct components and objects in the image by assigning different weights and biases to them. There are 32 filters in use. The spatial dimensions of the output volume are then minimised via max pooling. Then 64 filters are applied. Max pooling is utilised once again to reduce the spatial dimensions. Finally, there are 128 filters. The customised NN model with optimised parameters performed well in training and adequately in the test dataset.

2) R-CNN:

The Faster R-CNN approach is employed in a real-time crime scene evidence processing system that can detect items in an interior setting. The proposed system employs the Region Proposal Network and the VGG-16 network for object detection. Seven Convolution layers with Max-pooling, one Flatten, and the SoftMax activation function comprise the following architecture. The first convolution layer has 32 filters, while the next six hidden layers include 100 filters, each measuring 3 x 3. Max-pooling is also performed at each layer on a 2 x 2 scale. The convolutional layers are then flattened and normalised using the SoftMax step. The activation function "ReLu" is used by all layers, and the output range ranges from 0 to infinity. The shear and stride values are both 1.

3) VGGNET19 based on Fast RCNN & RCNN:

R-CNN (Region-based Convolutional Neural Network) and Fast R-CNN are two pivotal advancements in the field of computer vision, specifically designed for object detection tasks. They both address the challenge of identifying and localizing objects within images, but they differ in terms of their architecture and efficiency. R-CNN algorithm is adopted to detect multiple objects relevant to crime in real time videos and images. The Fast R-CNN architecture can be utilized to extract features from the proposed regions. These features can represent various aspects of the scene, such as objects, people, or vehicles. Crime detection often involves recognizing patterns and anomalies over time. R-CNN and Fast R-CNN can be applied to analyse video frames sequentially, allowing for the detection of events like break-ins, thefts, or altercations

4) ResNet50+LSTM (for Classification)

Residual Network, often known as ResNet, is a 152-layer deep neural network. It goes 8* deeper than VGG nets of lower complexity. It has been demonstrated to be more accurate than VGG and GoogLeNet. The picture input size is set to 224 x 224*3. RNNs are extended by LSTMs. It employs a hybrid of two forms of memory: long-term memory and short-term memory. This model is most effective at predicting crime in real time. In terms of feature extraction time, each video took roughly 8 seconds on average. To optimise the process, Transfer Learning methods were applied. Models were trained on the retrieved features for 30 epochs with a batch size of 16. Each epoch took an average of 66 seconds. Before each epoch, training data was randomised in batch-size chunks. Models were trained and evaluated using the same dataset and training circumstances.

5) YOLOv5:

YOLOv5 (You Only Look Once) is now regarded the benchmark for object detection and face recognition. It has four versions: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5xl. It is divided into three phases. The first step is the backbone network, which is a convolutional neural network that focuses on learning high-level features from input pictures. The neck is the second level of the network, and it is focused with learning characteristics at multiple scales in order to learn the same picture with varied sizes. In comparison to the backbone network, it has more layers. Finally, there is the head network, which is primarily responsible for identifying bounding boxes around objects and delivering final annotated pictures.

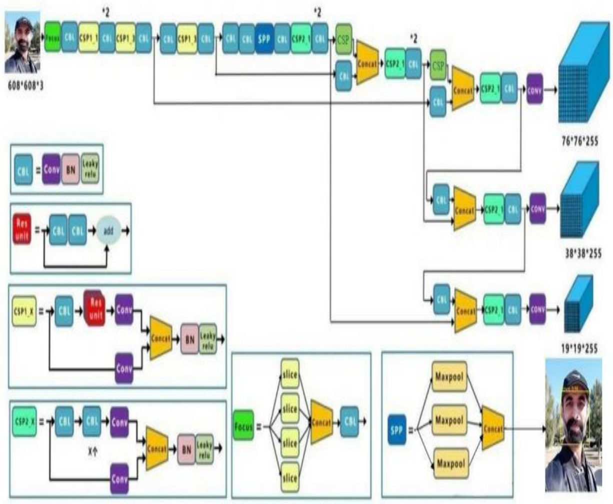


Fig. 3. YOLOv5s Network Architecture [19]

6) Simple YOLO

The You Only Look Once algorithm's primary goal is item recognition and categorization. The YOLO algorithm has four variations, each of which improves the model's performance. In comparison to RCNN, the YOLO algorithm employs a whole picture feature rather than a portion of the collected image. The YOLO method conducts bounding box and class prediction concurrently, which distinguishes it from other standard systems. Instead of a standard region proposal and classification, YOLO addresses the object detection problem as a regression problem. This allows YOLO to run more effectively in real-time while sacrificing some accuracy.

7) MobileNet

MobileNet is a cutting-edge deep learning architecture designed for efficient and real-time object detection tasks, making it an ideal candidate for enhancing crime detection capabilities. Leveraging the power of MobileNet, law enforcement agencies and security systems can efficiently process live video streams and images from surveillance cameras, smartphones, and drones to identify and classify criminal activities in real-time. Its lightweight and computationally efficient design allow for rapid analysis of diverse visual data, enabling swift response to incidents like theft, vandalism, or unauthorized intrusions. MobileNet's versatility makes it a valuable tool for enhancing public safety by providing accurate and timely alerts to law enforcement personnel, helping to mitigate potential threats and ensure the security of communities.

8)Xception

Xception, a state-of-the-art deep learning architecture, offers advanced capabilities for real-time crime detection applications. Its intricate design, characterized by a depth-wise separable convolutional network, enables highly accurate and efficient object recognition in surveillance videos and images. Xception's exceptional accuracy and speed make it a formidable tool for law enforcement agencies and security systems. By harnessing Xception's capabilities, authorities can swiftly analyse live video feeds, identify criminal activities, and respond promptly to incidents such as break-ins, suspicious behaviour, or vehicle theft. Its ability to process visual data with precision and efficiency empowers public safety efforts by delivering timely notifications to law enforcement, thereby enhancing overall security and crime prevention in communities.

9)InseptionV3+LSTM

InceptionV3 is a state-of-the-art CNN architecture known for its exceptional performance in image recognition tasks. It efficiently extracts hierarchical features from images, enabling accurate object detection and classification. InceptionV3 can process video frames to identify potential objects, individuals, or activities associated with criminal behaviour. LSTM can analyse sequential data from video frames to detect patterns of suspicious activities, such as loitering, trespassing, or violent actions. LSTM can analyze historical crime data alongside real-time video feeds to predict potential crime hotspots or times, aiding in resource allocation and proactive policing.

10) VGG16+LSTM

VGG16 is a well-known deep CNN architecture renowned for its simplicity and strong performance in image classification tasks. It is characterized by its deep and uniform architecture with 16 weight layers. LSTM can identify anomalies in video sequences, such as sudden changes in behaviour or the presence of unexpected objects or individuals. VGG16 can locate and track objects or persons of interest within a video feed, providing real-time information on their movements.

3.4 Evaluation Metrics

The models are evaluated using Precision, Accuracy Loss,etc

4.Results

MODEL	METRICS
Customized CNN	Accuracy of 90.7%, training loss of 0.4292
R-CNN	Precision- 0.82 , F-score- 0.86
VGGNET19 built on Fast RCNN and RCNN.	Accuracy- 85%
ResNet50+LSTM	Accuracy- 85%, loss- 0.0031
YOLOV5	Accuracy- 87%
Simple YOLO	Accuracy- 78.39%
MobileNet	Accuracy- 82%
Xception	Accuracy- 80%
InseptionV3+LSTM	Accuracy- 74.71%, Loss- 0.4820
VGG16+LSTM	Accuracy- 67.82%, Loss- 0.5023

5. CONCLUSION

This research study proposes a system to detect illegal activities in real time and warn police personnel in order to increase social security and cut crime rates. With a good training accuracy and less loss on training we are proposing a real-time criminal detection system employing the CNN algorithms and variants of CNN can successfully identify criminals even in crowded areas. With use of Web Terminology, we can develop a web application to identify suspicious activity and report to higher authorities.

REFERENCES

[1] S Samundeswari , Harini M, Dharshini, ”Real-time Crime Detection Using Customized CNN”, 2022 1st International Conference on Computational Science and Technology

[2] P. Sivakumar, and K. S, "Real Time Crime Detection Using Deep Learning Algorithm," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), 2021.

[3] F. Majeed, F. Z. Khan, M. J. Iqbal and M. Nazir, "Real-Time Surveillance System based Facial Recognition using YOLOv5," 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC).

[4] C. Rajapakshe, S. Balasooriya, H. Dayarathna, N. Ranaweera, N.Walgampaya and N. Pemadasa, "Using CNNs RNNs and Machine Learning Algorithms for Real-time Crime Prediction," 2019 International Conference on Advancements in Computing (ICAC)

[5] Nandhini T J and K Thinakaran "Detection of Crime Scene Objects using Deep Learning Techniques",2023 International Conference on Intelligent Data

[6] Varun Mnadalpu , Lavanya Elluri , Piyush , AND Nirmalya Royl," Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions",2023

[7] WAJIHA SAFAT, SOHAIL ASGHAR, AND SAIRA ANDLEEB GILLANI, "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques", 2021.

[8] Sharmila Chackravarthy, Steven Schmitt, Li Yang," Intelligent Crime Anomaly Detection in Smart Cities using Deep Learning" 2018 IEEE 4th International Conference on Collaboration and Internet Computing.

[9] Uma. N," Deep Convolutional Generative Adversarial Networks for Crime Scene Object Detection",2023 Second International Conference on Augmented Intelligence and Sustainable Systems.

[10] Umadevi V Navalgund, Priyadharshini.K," Crime Intention Detection System Using Deep Learning",2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)