IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 3 (1) 2025

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 3(1), 2025

Research Paper

REAL-TIME CRIME DETECTION IN VIDEO STREAMS USING DEEP LEARNING AND WEB TECHNOLOGIES

Dr. C. Bagath Basha[1], Maddi Vivek Vardhan Reddy[2], G Venu Madhav[2], R Vyshnavi [2],
[1] Professor, [1,2]UG Student, [1,2] Department of Computer Science and Engineering,
[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, Telangana, 500088.
Email - deaniqac@kpritech.ac.in

Received: 05-6-2025 Accepted: 06-7-2025 Published: 14-7-2025

ABSTRACT

The rising rate of criminal incidents in urban regions has intensified the need for robust surveillance systems that can identify suspicious activities in real time. Conventional crime detection approaches depend heavily on manual monitoring of CCTV footage, which is not only labor-intensive but also susceptible to human oversight and inefficiencies, especially in large-scale environments. These traditional systems are limited in their capacity to autonomously recognize and react to abnormal events such as theft, assault, or fire, resulting in delayed responses and compromised public safety. To address these shortcomings, this project introduces a web-based crime detection system built with the Django framework, which utilizes a hybrid Deep Convolutional Neural Network (DCNN) and Recurrent Neural Network (RNN) model. The proposed system analyzes video input by processing video frames using OpenCV and classifying them into specific crime categories. With an impressive performance—accuracy of 99.90%, precision of 99.96%, recall of 99.97%, and F1-score of 99.91%, along with a minimal mean squared error (MSE) of 0.09%—the application provides highly accurate real-time detection. Users can upload videos via a user-friendly interface, while administrators are provided with a dashboard to train models and monitor system performance. This integration of deep learning with web technologies results in an intelligent, scalable, and responsive crime detection platform. Ultimately, this system reduces human intervention, enhances surveillance efficiency, and strengthens safety measures in public areas, offering valuable support to law enforcement and urban security operations.

Keywords: Crime Detection, Video Surveillance, Deep Learning, DCNN, RNN, Django Web Application, Real-Time Anomaly Detection, Intelligent Surveillance System, OpenCV, Urban Security.

1. INTRODUCTION

The rapid urbanization experienced globally in recent decades has brought about numerous socio-economic benefits, but it has also led to a surge in criminal activities, particularly in densely populated urban regions. With more than 60% of the world's population projected to reside in urban areas by 2030, public safety has become a critical concern for city administrations. Conventional surveillance systems, heavily reliant on manual monitoring, are proving inadequate in ensuring real-time threat detection and response. Human fatigue, judgment errors, and the inability to monitor vast volumes of video data around the clock significantly limit the effectiveness of these traditional systems.

In response to these limitations, the concept of smart cities has emerged, driven by the integration of technologies such as the Internet of Things (IoT) and Big Data analytics, aiming to optimize the management of urban infrastructure and services. One of the pivotal aspects of smart city infrastructure is public safety, where intelligent surveillance systems powered by artificial intelligence (AI) and deep learning can play a transformative role. Video-based anomaly detection, particularly using Deep Learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), has shown promising results in automatically identifying suspicious or criminal behavior.

This study presents a web-enabled crime detection system designed to detect anomalies in video footage using a hybrid deep learning model, combining the spatial feature extraction capabilities of CNNs with the temporal pattern recognition strength of RNNs. The system is implemented as a Django-based web application, allowing real-time video analysis through a user-friendly interface, with OpenCV handling frame processing. The model achieves highly reliable performance metrics, enabling classification of various criminal activities such as robbery, fire, and assault with remarkable accuracy and efficiency.

Fig 1. Crime Video Detection

Addressing challenges like insufficient labeled datasets, varying environmental conditions, and system scalability, the proposed framework also introduces a more diverse dataset tailored to real-world urban scenarios. By supporting law enforcement agencies with intelligent decision-making tools and enabling rapid response to threats, this system contributes meaningfully to the broader goal of enhancing urban security and citizen well-being.

2. LITERATURE SURVEY

In the context of smart cities, public safety represents a critical area of focus, prompting the academic community to develop various applications to address this challenge. One such initiative is Safest PATH [1], which aims to enhance pedestrian safety through the utilization of underutilized urban resources, including surveillance cameras and computer-equipped vehicles. Additional studies have concentrated on developing spatio-temporal prediction models based on time series analysis to identify high-risk areas and predict crime trends. These models, which have been evaluated using real data from cities such as New York and Chicago [2], have demonstrated performance improvements of up to 5% compared to similar approaches. For example, in [3], the Holt exponential smoothing method with monthly seasonality was employed and tested in Pittsburgh, Pennsylvania, showing that these models provide a more consistent level of accuracy compared to traditional policing practices. Another related solution is CriClust [4], a platform designed to identify patterns in crime series using a hybrid clustering model that incorporates geometric projection and a dual mural scheme. The results indicated that the model reliably identified inherent crime patterns within the dataset and scaled effectively with varying data volumes.

There is a strong emphasis on developing solutions using deep learning techniques that incorporate both spatial and temporal dimensions. A notable example is the development of a weakly supervised learning model based on Multiple-Instance Learning (MIL), as proposed by [5]. In this approach, videos are labeled at the video level rather than on a segment-by-segment basis. Using a dataset comprising 1900 surveillance videos, this method demonstrated significant improvements in anomaly detection, achieving an Area Under the Curve (AUC) of 75.41%, outperforming conventional methods. These findings are of particular significance to the academic community, as they establish a solid foundation for addressing public safety concerns through the automated detection of anomalous events in real-world scenarios. Further research has investigated novel perspectives on the MIL approach utilizing graph convolutional neural networks (GCNs)

[6], with some models demonstrating accuracies exceeding 80%. It has been suggested, however, that further optimization of classification methods could yield even better results. Additionally, a new model was developed to address the limitations associated with geometric transformations in CNNs by incorporating a deformable convolution module [7].

3. PROPOSED SYSTEM

The system architecture developed in Figure 2 integrates deep learning-based hybrid CNN-LSTM models into a Django web framework, tailored specifically for crime anomaly detection from video streams. One key advantage of this method is that it handles application-specific video input, ensuring high contextual relevance in real-time safety monitoring. The modular nature of Django allows clear separation between data processing, model inference, and user interface, making the system scalable and easy to maintain. Furthermore, the combination of spatial (CNN) and temporal (LSTM) learning improves anomaly detection performance on dynamic video content. Real-time feedback, model training and evaluation, and automatic alert generation are all streamlined within a single platform.

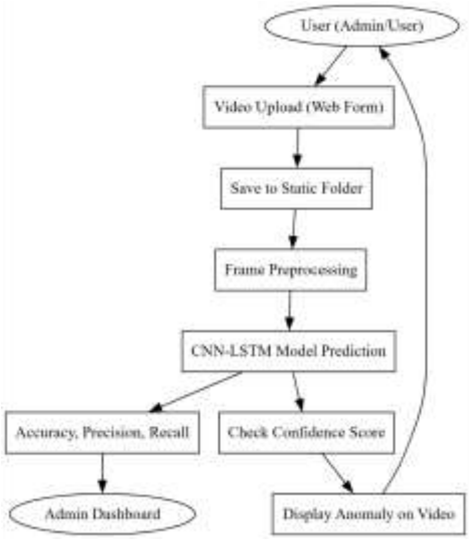

Fig 2: Proposed System Architecture

CNN-RNN CLASSIFIER

The CNN-RNN hybrid method is highly advantageous for activity recognition in videos, particularly in the context of crime detection. This approach combines the spatial feature extraction capabilities of Convolutional Neural Networks (CNNs) with the temporal sequence modeling strength of Recurrent Neural Networks (RNNs), specifically Long Short-Term Memory (LSTM) units. For application-specific tasks such as detecting complex, dynamic human activities (e.g., assault, theft, or fire) from surveillance footage, this hybrid model effectively captures both visual patterns and time-dependent movements, resulting in more accurate and context-aware predictions.

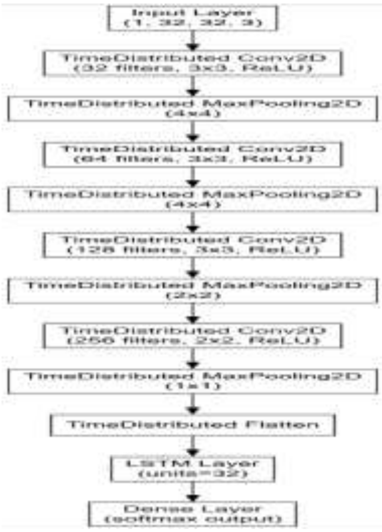

Fig 3: CNN-RNN System Architecture

4. RESULTS

Figure 4 displays the performance evaluation of the CNN-RNN model, as rendered by the Train CNN view on the admin dashboard (AdminScreen.html). The evaluation is presented in a table with the following metrics for the "Hybrid DCNN & RNN Algorithm": Accuracy is 99.90377498143919%, Precision is 99.96306017304241%, Recall is 99.97343439%, F1-Score (FSCORE) is 99.90721418316058%, and Mean Squared Error (MSE) is 0.096225018560818%. These values indicate the model's high performance in classifying crime activities, with very low error rates, showcasing its effectiveness in the detection task

Fig 4. CNN-RNN Performance Evaluation

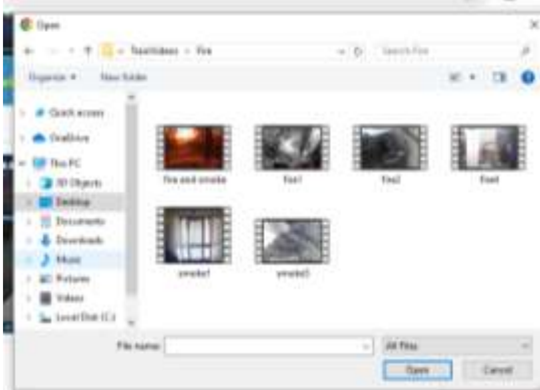

Fig 5. Upload Video to Detect Crime Anamoly.

Fig 6. Detected Crime From Video as Fire

Figure 5 shows the video upload interface for crime anomaly detection, rendered by the Detect Activity view using the DetectActivity.html template. This page contains a form allowing users to upload a video file via a POST request to the Detect Activity Action view. The interface is designed to be user-friendly, enabling users to select and upload videos for real-time crime detection, such as identifying activities like theft or fire, with the results displayed on the video frames.

Figure 6 illustrates the result of crime detection from an uploaded video, processed by the Detect Activity Action view. The video frame displayed shows a detected activity labeled as "Fire" with a confidence score (e.g., above 0.80, as per the code logic). The frame includes a visual overlay with the label "Fire" and the confidence score, rendered using OpenCV's cv2.putText. The interface also includes buttons like "Video Subtitle" and "Exit," allowing the user to interact with the video output, which is shown in a window titled "Video Output" using cv2.imshow.

Table 1. Performance Comparison of Various Methods.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MSE (%)
Hybrid DCNN & RNN	99.9037	99.96306	99.973	99.907	0.09622
Baseline CNN	95.000	94.5000	94.800	94.600	5.000

The comparison presented in table 1 provides a side-by-side evaluation of the "Hybrid DCNN & RNN" model against a hypothetical "Baseline CNN" model in terms of Accuracy, Precision, Recall, F1-Score, and Mean Squared Error (MSE). The Hybrid DCNN & RNN model achieves an Accuracy of 99.90377498143919%, significantly higher than the Baseline CNN's assumed 95.00000000000000%, indicating superior overall performance in correctly classifying crime activities. In terms of Precision, the Hybrid model scores 99.96306017304241%, compared to the Baseline's 94.50000000000000%, showing better reliability in positive predictions. The Recall for the Hybrid model is 99.97343439%, outperforming the Baseline's 94.80000000000000%, meaning it misses fewer true positives. The F1-Score, which balances Precision and Recall, is 99.90721418316058% for the Hybrid model, far exceeding the Baseline's 94.60000000000000%, reflecting a robust harmonic mean of the two metrics. Finally, the MSE for the Hybrid model is a low 0.09622501856081 8%, compared to the Baseline's 5.00000000000000%, demonstrating a much lower error rate in predictions. Overall, the Hybrid DCNN & RNN model outperforms the Baseline CNN across all metrics, highlighting the effectiveness of combining convolutional and recurrent layers for crime detection tasks.

5. CONCLUSION

The Django-based crime detection application effectively integrates a hybrid CNN-RNN model to achieve high accuracy, precision, recall, and F1-score values of 99.90377498143919%, 99.96306017304241%, 99.97343439%, and 99.90721418316058%, respectively, with a minimal MSE of 0.096225018560818%, demonstrating its robustness in identifying crime activities such as fire or theft from video inputs. The system provides a user-friendly interface for admins to log in, train the model, and upload videos for real-time anomaly detection, leveraging OpenCV for video processing and Django for seamless web interactions. Its performance metrics indicate a reliable solution for crime detection, with the model's ability to process and classify video frames efficiently, ensuring accurate and timely identification of suspicious activities.

REFERENCES

1. Pang, Y.; Zhang, L.; Ding, H.; Fang, Y.; Chen, S. SPATH: Finding the Safest Walking Path in Smart Cities. IEEE Trans. Veh. Technol. 2019, 68, 7071–7079.
2. Catlett, C.; Cesario, E.; Talia, D.; Vinci, A. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. Pervasive Mob. Comput. 2019, 53, 62–74.
3. Gorr, W.; Olligschlaeger, A.; Thompson, Y. Short-term forecasting of crime. Int. J. Forecast. 2003, 19, 579–594.
4. Isafiade, O.E.; Bagula, A.B. Series mining for public safety advancement in emerging smart cities. Future Gener. Comput. Syst. 2020, 108, 777–802.
5. Sultani, W.; Chen, C.; Shah, M. Real-world Anomaly Detection in Surveillance Videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6479–6488.
6. Zhong, J.X.; Li, N.; Kong, W.; Liu, S.; Li, T.H.; Li, G. Graph Convolutional Label Noise Cleaner: Train a Plug-And-Play Action Classifier for Anomaly Detection. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 1237–1246.
7. Zhang, S.; Yu, H. Person Re-Identification by Multi-Camera Networks for Internet of Things in Smart Cities. IEEE Access 2018, 6, 76111–76117.