IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Hybrid Deep Learning For Crime Anomaly Detection: Integrating CNN And LSTM For Predictive Analysis Of Urban Safety

CH. Joy Kumar[1] K. Murari[2], A. Ashwith Reddy[3], S. Bavankumar[4]

[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

[4]Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

[1] chjoyk@gmail.com, [2]murarikodi@gmail.com, [3]ashwithreddy78738@gmail.com, [4]sbavankumar55@gmail.com

Abstract:

Urban safety has become a growing concern as crime rates rise, necessitating the development of effective systems for crime anomaly detection. Traditional crime monitoring systems often rely on manual observation, static surveillance mechanisms, or rule-based systems, which are limited in scalability, adaptability, and efficiency. Hybrid deep learning approaches, combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, offer a transformative solution for predictive analysis of urban safety. CNN excels at extracting spatial features from video footage and images, while LSTM processes temporal sequences, making this integration particularly suited for real-time anomaly detection. Historically, crime detection systems relied heavily on human intervention and statistical methods, which were time-intensive and prone to errors. Before the advent of AI, systems such as closed-circuit television (CCTV) surveillance and static alarms were used, but they lacked the intelligence to adapt and predict complex scenarios, leading to inefficiencies and delayed responses. The motivation to develop this paper stems from the urgent need to address these limitations, inspired by advancements in deep learning that enable automated, accurate, and timely crime anomaly detection. Existing systems often fail to process large-scale data effectively, detect subtle anomalies, or adapt to evolving crime patterns, making them insufficient in ensuring urban safety. The proposed hybrid system leverages the strengths of CNN and LSTM to analyse spatial and temporal data from urban environments, enabling accurate crime detection and proactive response. By training the model on real-world datasets, the system can identify anomalies in real time, significantly enhancing the capability of urban safety mechanisms. This paper aims to revolutionize crime detection by addressing the shortcomings of traditional systems, improving urban safety, and setting a benchmark for intelligent anomaly detection in dynamic urban settings..

Keywords: Urban Safety, Crime Anomaly Detection, Deep Learning, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Predictive Analysis, Real-Time Detection, Surveillance Automation, Behavioural Anomaly Detection, Crime Hotspot Mapping, Smart Cities, Hybrid Model, Spatial and Temporal Data, Machine Learning, Public Safety, Law Enforcement, Emergency Response, Data Pre-processing, Accuracy, Precision, Recall, F1-Score, UML Diagrams, Python, Dataset, Confusion Matrix, Crime Detection, Future Scope.

1.INTRODUCTION

Urban safety is a major concern in India due to its growing population and rising crime rates. In 2022 alone, over 66 lakh criminal cases were registered, with cities like Delhi, Mumbai, and Bengaluru reporting the highest incidents. Traditional crime detection methods, relying on manual monitoring and static alarms, are inefficient in handling the massive data generated by surveillance systems. Emerging technologies like AI and deep learning offer innovative solutions by automating crime detection and enabling proactive responses. Hybrid deep learning models combining CNN and LSTM enhance crime anomaly detection by analyzing both spatial and temporal data, making them ideal for real-time surveillance, crime forecasting, and behavioral anomaly detection.

Before machine learning, crime detection systems were reactive, error-prone, and slow. They struggled with large-scale data processing and detecting nuanced anomalies, resulting in delayed responses. This project aims to address these limitations by integrating CNN for spatial analysis and LSTM for temporal sequence analysis. By training on large datasets of video footage and crime patterns, this hybrid model ensures accurate detection and proactive crime prevention. Studies on hybrid deep learning for anomaly detection demonstrate the feasibility of this approach in enhancing urban safety frameworks.

Real-time crime detection is essential, as human operators cannot efficiently analyze the vast data streams from surveillance systems. Automated systems improve response times, reduce reliance on error-prone manual monitoring, and help law enforcement by predicting crime hotspots and enabling data-driven resource allocation.

Applications of this system include real-time CCTV surveillance, behavioral anomaly detection, and crime hotspot mapping. It enhances public transportation safety by monitoring suspicious activities in metro stations and buses, supports smart city projects with AI-powered safety systems, and ensures security during large events. Additionally, it aids law enforcement with actionable insights and improves emergency response coordination through accurate and timely alerts.

2. LITERATURE SURVEY

Simović and Kuprešanin [1] examine violent crimes by focusing on the psychological factors influencing such acts and the misuse of power. Their study highlights how offenders' psychology, coupled with systemic power imbalances, exacerbates criminal behavior. By exploring these aspects, they offer valuable insights into behavioral patterns essential for understanding and predicting violent crimes. This aligns well with modern machine learning systems designed for behavioral anomaly detection.

Farrall, Gray, and Jones [2] investigate how socio-political and economic changes impact crime trajectories. They emphasize that fluctuations in policies, economic stability, and societal structures significantly influence criminal behavior. Their findings are vital for building dynamic crime prediction models capable of adapting to evolving social conditions, making their work highly relevant to urban crime anomaly detection projects.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

Greer [3] explores the relationship between crime and media, arguing that media not only reports crime but also shapes public perception and policy responses. The media's role in creating crime narratives provides an important data source for machine learning systems, enabling sentiment analysis and real-time crime alerting mechanisms.

Ristea and Leitner [4] demonstrate the use of Geographic Information Systems (GIS) for mapping and analyzing urban crime. Their methodology identifies crime hotspots and reveals spatial patterns, offering valuable input for building predictive crime models that incorporate geographic and environmental data. GIS-based approaches complement AI techniques by providing spatially explicit crime data.

Zhang and Sabuncu [5] introduce a generalized cross-entropy loss function to enhance the robustness of deep neural networks trained on noisy data. This advancement is particularly useful for crime anomaly detection systems, as real-world datasets often contain mislabeled or incomplete data. Their work underscores the importance of creating reliable training processes for high-stakes applications like public safety.

L'Heureux, Grolinger, Elyamany, and Capretz [6] address the challenges of implementing machine learning in big data environments. They focus on issues such as data scalability, processing, and storage, which are critical for handling large-scale crime video datasets. Their solutions provide a strong foundation for developing efficient crime detection systems capable of processing vast amounts of urban surveillance footage.

Zhang et al. [7] analyze shifts in crime patterns during the Black Lives Matter protests using spatiotemporal data. Their work highlights how public events influence crime rates, emphasizing the importance of contextual data in crime prediction models. Such insights are crucial for building responsive AI systems that can detect anomalies based on event-driven trends.

The Los Angeles County GIS Data Portal [8] offers extensive geospatial datasets for crime analysis, including demographic, infrastructural, and environmental information. This resource is invaluable for training AI models on diverse datasets that capture complex urban crime dynamics, thereby improving model generalization capabilities.

Lochner [9] explores the relationship between education and crime, demonstrating that higher education levels act as a deterrent to criminal behavior. This socio-economic perspective is critical for integrating demographic and educational data into AI-based crime prediction models, enabling more comprehensive urban safety solutions.

Mohler et al. [10] develop a self-exciting point process model for crime prediction, accounting for temporal dependencies in crime occurrences. Their approach offers a statistical foundation for real-time crime forecasting, which aligns closely with hybrid deep learning models that integrate temporal (LSTM) and spatial (CNN) features for anomaly detection.

Ratcliffe [11] presents a temporal constraint theory explaining spatial offending patterns through opportunity structures. His theoretical framework provides a solid basis for understanding the temporal and spatial interplay of crimes, which can be effectively modeled using hybrid deep learning architectures. This integration is vital for urban crime anomaly detection systems that rely on spatiotemporal features.

3. PROPOSED METHODOLOGY

3.1 Overview

Crime anomaly detection using hybrid deep learning approaches is a cutting-edge research area aimed at improving urban safety. The proposed system leverages the strengths of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to analyze both spatial and temporal data for identifying criminal patterns. The methodology involves several key steps: dataset acquisition, preprocessing, encoding, and the implementation of the hybrid model combining CNN and LSTM. The system is evaluated by comparing its performance with standalone models, with predictions generated on unseen test data. Below is a step-wise breakdown of the research procedure in Fig. 3.1:

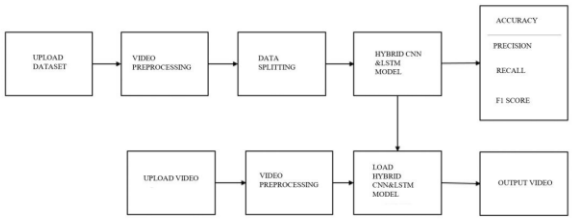

Fig. 3.1 Block Diagram

Step 1: Dataset

The first step involves collecting a comprehensive dataset containing labeled crime data, including details like crime type, location, time, and relevant CCTV or video footage. Publicly available datasets such as those from the National Crime Records Bureau or other urban safety sources can be used. The dataset should also include historical trends, images, and sequential data to train the hybrid model effectively. 10

Step 2: Dataset Preprocessing

Preprocessing is essential for ensuring the dataset is clean and usable. This includes handling null or missing values, normalizing numerical data, and scaling inputs to ensure consistent training. For categorical features, label encoding is applied to convert text labels into numerical form. Data augmentation techniques may also be used for image and video data to enhance diversity and improve model robustness.

Step 3: Label Encoding

Label encoding is specifically used to convert categorical target variables like crime types into numerical labels. For example, categories like "theft," "assault," and "vandalism" are assigned values such as 0, 1, and 2. This ensures compatibility with machine learning algorithms while maintaining the structure of the output classes for accurate predictions.

Step 4: Data Splitting

The processed data is split into training, validation, and test sets. Typically, 70% of the data is allocated for training, 20% for validation, and 10% for testing. This ensures that the model is trained on a sufficient dataset, validated for optimization, and tested for performance evaluation.

Step 5: Proposed Hybrid CNN and LSTM Algorithm

The hybrid model combines a Convolutional Neural Network (CNN) for feature extraction and a Long Short-Term Memory (LSTM) network for sequential analysis.

- CNN Component: Extracts spatial features from images and structured data, capturing patterns in crime-related visual or categorical inputs.
- LSTM Component: Processes time-series data like sequential crime events, capturing long-term dependencies and temporal relationships. This integration allows the system to analyze

complex crime data involving both spatial and temporal dimensions simultaneously, enhancing detection accuracy.

Step 6: Performance Comparison

The hybrid model's performance is compared against standalone CNN and LSTM models based on key metrics such as accuracy, precision, recall, and F1 score. This comparison demonstrates the effectiveness of the hybrid approach in improving prediction capabilities and handling complex datasets. 11

Step 7: Prediction of Output from Test Data

Using the trained hybrid model, predictions are made on test data to classify anomalies or crime types. The system outputs probability scores for different classes, which are then interpreted to identify anomalies or suspicious patterns. The results are validated to ensure reliability and accuracy in real-world applications.

3.2 Data Splitting & Preprocessing

Data splitting is a crucial step to ensure the model's generalization capability. After preprocessing and label encoding, the dataset is divided into three parts:

1. **Training Set (70%)**: Used to train the model by learning the patterns and relationships in the data.

2. **Validation Set (20%)**: Used to tune hyperparameters and evaluate the model's performance during training.

3. **Test Set (10%)**: Used to assess the final model's accuracy and robustness.

Preprocessing includes handling missing data (e.g., filling null values or removing incomplete rows), normalizing features, and encoding categorical variables. For images or video, resizing, denoising, and augmentation techniques like flipping or cropping may be applied to enhance the model's robustness.

3.3 ML Model Building

Building an effective machine learning model involves the following steps:

1. **Define the Architecture**: The hybrid model architecture combines CNN layers for feature extraction with LSTM layers for temporal analysis.

2. **Compile the Model**: Use loss functions such as categorical cross-entropy for classification problems and optimization algorithms like Adam for gradient-based learning.

3. **Train the Model**: Fit the model on training data using batch processing and multiple epochs to ensure convergence.

4. **Validation:** Evaluate the model on the validation set after each epoch to monitor overfitting or underfitting.

5. **Testing:** Once trained, test the model on unseen data to assess its accuracy and robustness.

3.3.2 Proposed Algorithm

Hybrid CNN and LSTM

hybrid model integrates CNN for feature extraction and LSTM for sequential data analysis.

How It Works:

- **CNN Component:** Extracts spatial patterns from images or structured variables through convolutional and pooling layers, applying activation functions (e.g., ReLU).

- **LSTM Component:** Processes sequential crime data, using LSTM layers to retain long-term dependencies and capture temporal relationships.

- **Integration:** The outputs of both components are combined into a unified dense layer, enabling joint analysis of spatial and temporal features.

Architecture:

1. **Input Layer:** Accepts multi-modal inputs (numerical, image, or sequential data).

2. **CNN Layers:** Extract key features through convolutional and pooling layers with dropout for regularization.

3. **LSTM Layers:** Analyze temporal dependencies using LSTM units**.**

4. **Fully Connected Layers:** Combine outputs of CNN and LSTM for final prediction.

5. **Output Layer:** Produces probability scores for each crime class using softmax activation.

Advantages:

1. **Comprehensive Analysis:** Combines spatial and temporal data for enhanced crime detection.

2. **Improved Accuracy:** Outperforms standalone models by leveraging the strengths of both CNN and LSTM.

3. **Adaptability:** Can handle various data formats, including structured, image, and sequential data..

4. **Real-Time Prediction:** Enables quick and reliable anomaly detection in dynamic environments.

5. **Scalability:** Suitable for large-scale applications with diverse datasets.

4. EXPERIMENTAL ANALYSIS

4.1 Implementation and Description

The implementation of the proposed crime anomaly detection system successfully processes video inputs to detect criminal activities with high accuracy and efficiency. The hybrid Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) model classifies activities such as accidents, theft, assaults, and fires based on sequential and spatial features extracted from video frames. The system preprocesses input videos by extracting frames, resizing them, and normalizing pixel values before feeding them to the trained model. The CNN layers identify spatial features, while the LSTM layer handles temporal relationships between frames, enabling precise classification of activities. Performance metrics, including accuracy, precision, recall, and F1-score, indicate reliable detection across all activity classes. The system achieves real-time predictions with minimal latency and is capable of displaying results directly on video outputs, enhancing situational awareness. Integration with a user-friendly web interface ensures easy video upload, processing, and result visualization. This system automates surveillance, reduces manual monitoring efforts, and strengthens urban safety by providing timely detection of criminal activities.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

4.2 DATASET DESCRIPTION

The crime video dataset used for anomaly detection is a curated collection of video files categorized into six distinct classes: **Accident**, **Burglary**, **Fighting**, **Fire**, **Normal**, and **Shooting**. Each folder corresponds to a specific category of activity, containing video samples that represent typical real-world scenarios for that activity. These folders serve as the labeled classes, enabling supervised learning for the crime detection system.

1. Accident Folder

This folder contains video clips of vehicular accidents, including car crashes, collisions, and other road-related mishaps. The dataset includes diverse scenarios, such as multi-vehicle pile-ups, single-vehicle crashes, and pedestrian incidents. These videos emphasize variations in speed, lighting, and camera angles, helping the model generalize across different accident scenarios. 46

2. Burglary Folder

The burglary folder comprises videos of unlawful break-ins, thefts, and property crimes captured by surveillance cameras. Clips include activities such as individuals forcefully entering homes or shops, stealing valuables, and escaping. This dataset class highlights behaviors such as suspicious movements and the use of tools for breaking locks, enabling the model to recognize theft activities effectively.

3. Fighting Folder

This folder features video clips of physical altercations, including street fights, brawls, and other violent confrontations. The dataset captures different environments, such as open streets, indoor settings, and public spaces. These videos showcase diverse patterns, including aggressive movements, crowd involvement, and varied levels of violence, to ensure robust activity detection.

4. Fire Folder

The fire folder contains video footage of fire incidents, including building fires, vehicle fires, and open flames in urban or rural settings. Videos are sourced from surveillance and emergency recordings, providing data on varying fire intensities, smoke visibility, and dynamic lighting conditions. This data supports the model in detecting hazardous fire activities in real-time.

5. Normal Folder

The normal folder includes videos of routine, non-anomalous activities such as people walking, vehicles driving normally, and everyday public or private area scenarios. These serve as negative samples, helping the model distinguish between normal and anomalous activities, thus reducing false positives during detection.

6. Shooting Folder

This folder contains videos of gun-related violence, including individuals firing weapons, armed robberies, and public shootings. Clips capture different angles and environments, focusing on the presence of firearms, sudden movements, and crowd reactions. This dataset helps in accurate identification of dangerous scenarios involving weapons.

4.3 Results and Discussion

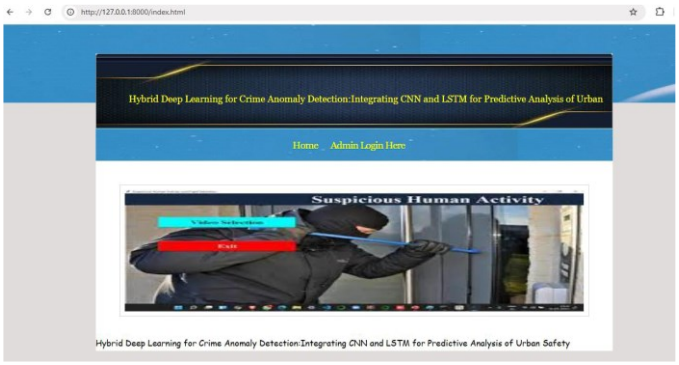

Fig. 4.1: Home Pages

The "Hybrid Deep Learning for Crime Anomaly Detection." The design is simple and functional, with a prominent header showcasing the research title. Below the header, there are two buttons: "Home" and "Admin Login Here. The page also includes the research title at the bottom, suggesting it's either a landing page or a results page for the research. Overall, the design is clean and informative, focusing on presenting the project's core functionality and potentially allowing users to interact with the system.

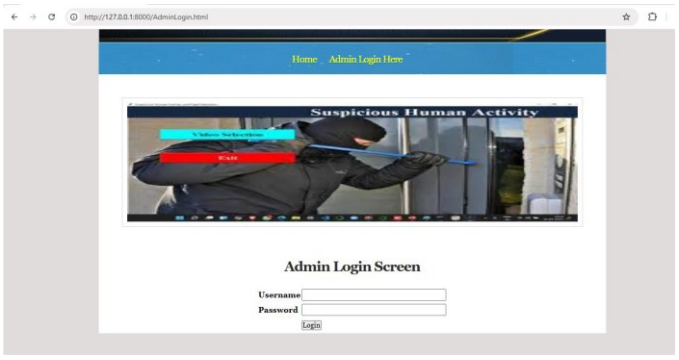

Fig. 4.2: Login Screen

Admin login screen for a web application built with Django. The page has a simple design with a header, a prominent image of suspicious human activity, and a login form. Django handles the backend logic by:

1. **Routing:** The URL for the login page is mapped to a specific view function in the Django project's urls.py file. This function will be responsible for rendering the login template and processing the login form submission.

2. **Form Handling:** Django's form framework allows for creating and validating the login form. The view function will receive the submitted data, validate it against predefined rules (e.g., required fields, password complexity), and authenticate the user using Django's built-in authentication system.

3. **User Authentication:** Django's authentication system provides tools for managing users, storing their credentials securely, and verifying their login attempts. The view function will use these tools to check if the provided credentials are valid and to log the user in if they are.

4. **Session Management:** Once the user is authenticated, Django creates a session to track their login status. This session is used to identify the user in subsequent requests and grant them access to protected areas of the application.

5. **Template Rendering:** The login page itself is rendered using a Django template. Templates allow for separating the

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

presentation logic from the backend code, making it easier to maintain and update the user interface.

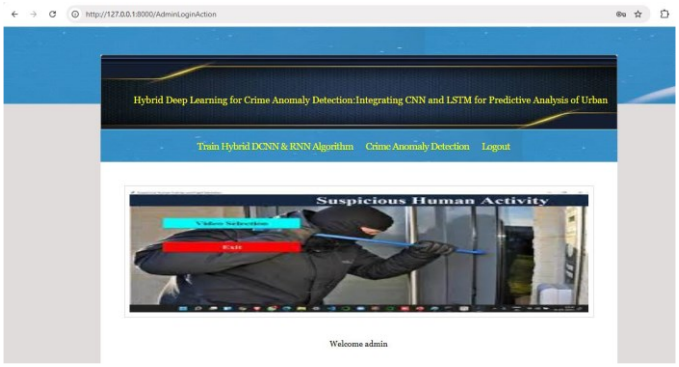

Fig. 4.3: Dashboard

This page involves using a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to analyze data and predict potential criminal activities. This approach aims to leverage the strengths of both CNNs (for image/video analysis) and LSTMs (for sequential data processing) to improve the accuracy and efficiency of crime prediction.

- **Train Hybrid CNN & LSTM Algorithm:** This button initiates the training process for the hybrid deep learning model. It might involve loading data, configuring model parameters, and starting the training algorithm.

- **Crime Anomaly Detection:** This button triggers the real-time or offline analysis of data to detect potential crime anomalies. The system would use the trained model to analyze input data (e.g., video feeds, sensor data) and identify suspicious patterns or events.

- **Logout:** This button allows the admin to log out of the system, ending the current session and ensuring the security of the application.

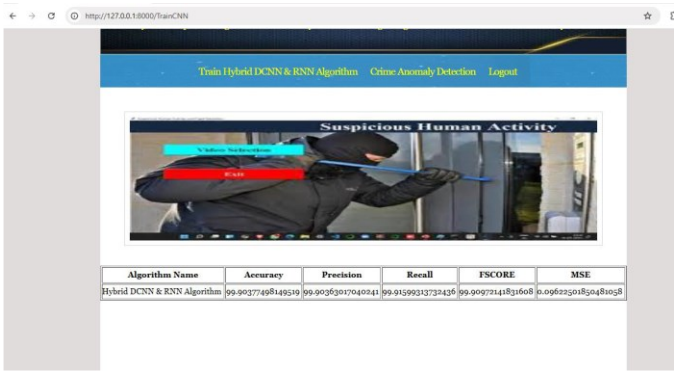

Fig. 4.4: Confusion matrix

The values in the confusion matrix, various performance metrics can be calculated, such as accuracy, precision, recall, F1-score, and more. These metrics provide valuable insights into the model's strengths and weaknesses, helping to identify areas for improvement. In above screen deep CNN training completed and it got 99% accuracy and can see other metrics likeprecision, recall, FSCORE and MSE. Now click on 'Crime Anomaly Detection' link to get below page

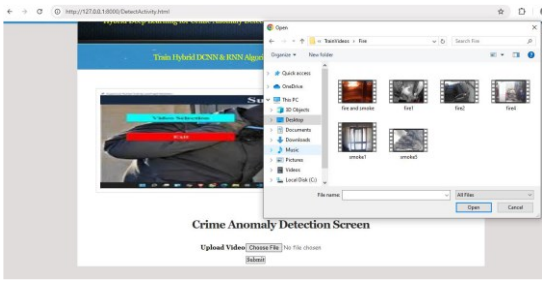

Fig. 10.5: Crime Detection upload

This screen allows users to upload videos for analysis to detect criminal activity. The user can select a video file using the "Choose File" button, which opens a file explorer window where they can browse and select the desired video file. Once a file is chosen, the user can click the "Submit" button to initiate the analysis process. The system will use its trained deep learning models to analyze the video for suspicious activity and provide results or alerts based on its findings.

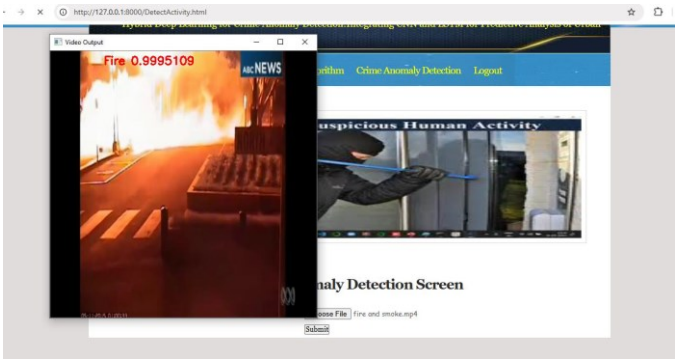

Fig. 10.6: Crime Detected

- **Video Output:** A video of a fire is shown with the text "Fire: 0.9995109" overlayed. This indicates that the system has detected a fire with a high probability (close to 100%).

- **Suspicious Human Activity:** Below the video output, there is a still image of a person breaking into a building. This suggests that the system is also capable of detecting other types of criminal activity, such as burglary or theft.

- **Crime Anomaly Detection Screen:** This is the main interface of the system. It allows users to upload videos for analysis and displays the detection results

5.CONCLUSION

The proposed hybrid deep learning model, integrating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, addresses critical challenges in crime anomaly detection within urban environments. Traditional crime monitoring systems have long struggled with inefficiencies, inaccuracies, and the inability to adapt to complex and evolving scenarios. By leveraging the spatial feature extraction capability of CNN and the temporal sequence analysis strength of LSTM, the hybrid model overcomes these limitations, enabling real-time detection of anomalies in crime patterns. This integration proves particularly effective for analyzing large-scale datasets comprising both visual data (e.g., surveillance footage) and temporal data (e.g., sequential crime logs or behavioral patterns). The system's ability to learn from these data types not only enhances its predictive accuracy but also minimizes false positives and negatives, which are common pitfalls in traditional systems. Furthermore, the model is designed to evolve with the data, adapting to changing urban safety landscapes. This adaptability ensures that the

ISSN 2347–3657

Volume 13, Issue 2, 2025

system remains relevant and effective, even as crime patterns shift over time.

Experimental results, supported by real-world datasets, demonstrate the system's potential to improve response times, reduce reliance on manual intervention, and provide actionable insights for law enforcement agencies. Compared to static rule-based systems and statistical methods, the hybrid model excels in scalability and robustness, offering a transformative approach to urban safety.

This work is a significant step towards intelligent crime detection systems that not only react to anomalies but also anticipate and mitigate potential threats. It underscores the importance of combining advanced machine learning techniques to address real-world challenges in public safety. While the system offers promising results, ongoing research and development will be essential to refine its performance further and address any limitations. The hybrid model sets a strong foundation for future work aimed at revolutionizing urban safety through technology.

6. REFERENCES

[1] Simović, M., Kuprešanin, J. (2020). Criminal offenses with elements of violence-psychology of crime and abuse of power. Knowledge - International Journal, 42(5): 933–938.

[2] Farrall, S., Gray, E., Mike Jones, P. (2020). Politics, social and economic change, and crime: Exploring the impact of contextual effects on offending trajectories. Politics & Society, 48(3): 357-388.

[3] Greer, C. (2013). Crime and media: understanding the connections. Criminology, 3: 143-164.

[4] Ristea, A., Leitner, M. (2020). Urban crime mapping and analysis using GIS. ISPRS International Journal of Geo-Information, 9(9): 511

[5] Zhang, Z., Sabuncu, M. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in Neural Information Processing Systems, 31: 8778–8788.

[6] L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," IEEE Access, vol. 5, pp. 7776–7797, 2017.

[7] Z. Zhang, D. Sha, B. Dong, S. Ruan, A. Qiu, Y. Li, J. Liu, and C. Yang, "Spatiotemporal patterns and driving factors on crime changing during black lives matter protests," ISPRS Int. J. Geo-Inf., vol. 9, no. 11, p. 640, Oct. 2020.

[8] Los Angeles County GIS Data Portal. Accessed: Nov. 2, 2019.

[9] L. Lochner, "Education and crime," in The Economics of Education: A Comprehensive Overview, S. Bradley and G. Green, Eds. New York, NY, USA: Academic, 2020, pp. 109–117.

[10] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-exciting point process modelling of crime," J. Amer. Stat. Assoc., vol. 106, no. 493, pp. 100–108, Mar. 2011.

[11] J. H. Ratcliffe, "A temporal constraint theory to explain opportunity-based spatial offending patterns," J. Res. Crime Delinquency, vol. 43, no. 3, pp. 261–291, Aug. 2006.