Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

ENHANCING PUBLIC SAFETY IN SMART CITIES USING DEEP LEARNING-BASED VIOLENCE DETECTION

[#1]**Gaddamidhi Vaishika Goud,** *Department of MCA,*
[#2] **Mrs. T. Mounika,** *Assistant Professor, Department of MCA,*
Vaageswari College of Engineering(Autonomous), Karimnagar, TG.

ABSTRACT: This work illustrates the utilization of deep learning to detect violent incidents in smart cities, thereby improving public safety. Automated threat identification and real-time surveillance research are implemented by the technology. The proposed method employs sophisticated convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to analyze video feeds from public surveillance cameras, accurately detecting violent incidents such as altercations, assaults, vandalism, and mass violence. The system's proficiency in distinguishing between violent and non-aggressive human behaviors across a variety of external conditions is attributed to its implementation of spatial and temporal feature extraction techniques. The model is trained using a vast array of annotated datasets. The goals are to improve the effectiveness of detection, decrease the number of false alarms, and facilitate the rapid response of law enforcement. The proposed method improves the efficiency of emergency response, reduces the need for human supervision, and enhances smart city monitoring by providing continuous updates on current events. The intelligence, safety, and security of urban areas can be substantially improved by this violence detection system, which employs deep learning.

Keywords: *Deep Learning, Violence Detection, Smart Cities, Public Safety, Surveillance Systems, Convolutional Neural Networks (CNN), Human Activity Recognition,*

1. INTRODUCTION

Innovative technology is being rapidly adopted by smart cities to optimize infrastructure management, enhance urban living, and improve public safety. Violent crime, public discontent, and overall crime rates in urban areas have been substantially elevated by population growth. This has substantially complicated matters for law enforcement. In traditional surveillance systems, operators frequently manually examine CCTV footage, which leads to inefficiency as a result of fatigue, diminished concentration, and extended response times. Deep learning-based violence detection techniques have emerged as an effective method to improve public safety in order to address these issues. These sophisticated systems allow authorities to promptly respond and prevent the escalation of hazardous situations. The systems analyze CCTV data autonomously, identify indicators of aggression, and transmit real-time notifications.

Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks are employed extensively to identify violent behavior in video footage using deep learning techniques. These algorithms have the ability to independently identify complex patterns of movement, geographical characteristics, and temporal data from surveillance footage with minimal human intervention. The system analyzes crowd dynamics, anomalous behavior, and body language to identify violent

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

incidents, including altercations, assaults, and aggressive conduct, in public venues such as streets, transit stations, retail centers, and airports. By enabling continuous monitoring, expediting communication among security offices, and facilitating cloud-based processing, the integration of AI into IoT-enabled smart city infrastructure enhances surveillance efficacy.

In modern smart cities, violence detection systems that are based on deep learning improve the accuracy of detection, reduce the need for human labor, and expedite emergency response times. These devices can consistently survey extensive areas and perform their functions well under a variety of meteorological conditions. By enabling officials to take prompt action to prevent issues before they arise, real-time detection and automated alert generation mitigate hazards to public safety. The efficacy and reliability of violence detection systems are consistently improved by advancements in computer vision and artificial intelligence technologies, despite the challenges of privacy concerns, substantial processing requirements, and the need for extensive training datasets. Cities that are more intelligent and secure will increasingly depend on surveillance systems to guarantee public safety.

2. LITERATURE SURVEY

Anderson & Clark (2021): Anderson and Clark (2021) suggest that a deep learning-based monitoring system and real-time video analysis be implemented to identify violent incidents in smart cities. The technology utilizes Convolutional Neural Networks (CNNs) to identify dubious or aggressive human behavior in surveillance footage. Sophisticated feature extraction techniques can be employed to achieve improved recognition accuracy in congested public spaces. The test results suggest that false alarms have been reduced, and hazards have been identified more quickly. The structure facilitates the efficacy of rapid response systems, thereby enhancing urban safety.

Singh & Rao (2022): Singh and Rao's 2022 paper discusses an intelligent model for detecting crime in smart cities that utilizes Recurrent Neural Networks (RNNs) and Internet of Things (IoT) surveillance cameras. The system consistently monitors for signs of anomalous activity in public spaces, including retail malls, train stations, and streets. In order to accurately identify combative interactions, temporal motion analysis techniques are implemented. The performance analysis demonstrated a decrease in response times and an increase in the utility of real-time monitoring. The recommended method will simplify the management of public safety in urban areas.

Garcia & Wilson (2023): A hybrid deep learning architecture was devised by Garcia and Wilson in their 2023 paper. It utilizes both Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) to autonomously identify violent incidents in CCTV recordings. The system is capable of precisely detecting altercations, attacks, and anomalous crowd behavior by analyzing both spatial and temporal dimensions. The scalability and continuous monitoring of extensive networks in smart cities are facilitated by cloud-assisted processing. The technology exhibits exceptional retrieval capabilities and accuracy when assessed in a variety of contexts. The framework enables the development of intelligent urban technology that is both secure and innovative.

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

Ramesh & Kulkarni (2024): Ramesh and Kulkarni's 2024 publication outlines a real-time system that utilizes deep learning and edge computing to identify violent incidents. The goal is to improve public safety. The solution reduces network latency and improves response times during emergencies by processing security video streams locally. In densely populated areas, the accuracy of detecting violent behavior is improved by attention-based neural networks. The assertions that the computers operate with increased speed and efficiency are corroborated by the experimental results. The technology enables the implementation of effective crime prevention strategies and the monitoring of major urban areas through applications.

Chen & Ibrahim (2025): The research suggests the implementation of deep neural networks and behavioral recognition techniques in an AI-driven system for the detection of violence in smart cities (Chen & Ibrahim, 2025). In order to detect potentially hazardous actions, the framework implements methodologies such as motion monitoring and activity classification. By utilizing data encryption in conjunction with a secure cloud connection, surveillance data is protected from unauthorized access. The results suggest that urban environments can be more accurately identified and monitored on a comprehensive scale. The proposed technique improves public safety in real time and enables more efficient police operations.

Patel & Fernandez (2026): Patel and Fernandez (2026) create a sophisticated system that employs deep learning to detect violent incidents in smart cities by integrating federated learning and advanced monitoring technologies. The architecture facilitates the processing of surveillance data in numerous locations, while also reducing computer expenses and protecting privacy. Adaptive neural models are constantly acquiring new patterns to improve the detection of violence. The experimental results suggest that the system is capable of accommodating a greater number of users, conducting more rapid event identification, and utilizing its resources with increased efficiency. The development of AI-driven public safety systems in future smart cities is facilitated by this technology.

3. PROPOSED METHODOLOGY

Data Collection

The initial phase in the proposed procedure entails the acquisition of extensive security footage from public CCTV cameras, smart city surveillance systems, and benchmark datasets for the purpose of detecting violence. The data acquired includes both violent and nonviolent occurrences, such as altercations, assaults, vandalism, thefts, accidents, and individuals behaving in a typical manner in public settings. These datasets are employed to train and assess a deep learning model that is capable of reliably identifying ongoing violent incidents.

Data Preprocessing

In order to facilitate subsequent analysis, picture frames are implemented to modify video data. Scaling, normalization, noise reduction, frame enhancement, and background filtering are among the preprocessing techniques that are implemented to optimize image quality while reducing computational costs. Techniques such as transformation, scaling, rotation, and improving the generalizability of the model are among the methods used to augment data.

Feature Extraction

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

At this juncture, video frames are used to extract significant characteristics of space and time. Information regarding population density, body posture, item motion, and illogical movements may be extracted by Convolutional Neural Networks (CNNs). Temporal data that illustrates motion patterns across various frames is analyzed using networks that employ Gated Recurrent Units (GRUs) or Long Short-Term Memory (LSTM). These attributes aid the system in recognizing violent behaviors and identifying when an individual is behaving irrationally.

Deep Learning-Based Violence Detection

In order to classify the activities into categories, the extracted features are fed into a hybrid deep learning model. A system that utilizes supervised learning classifies behaviors into two categories: violent and nonviolent. In order to improve the accuracy of detection and increase the efficacy of processing, advanced deep learning architectures such as ResNet, VGG16, and YOLO are implemented. The Adam optimizer and categorical cross-entropy loss functions are two optimization strategies that are implemented to optimize model training.

Model Training and Evaluation

Classification datasets that include both violent and nonviolent video segments are employed to train the deep learning model. The dataset is partitioned into training, validation, and testing subsets to guarantee precise learning and performance evaluation. Precision, accuracy, recall, F1-score, and confusion matrix comprise the criteria utilized to assess the efficacy of the proposed technique. We increase the training model's ability to detect violence in real time and reduce false alarms.

Real-Time Monitoring and Alert Generation

Upon the system's identification of suspicious activity, notifications will be promptly transmitted to the appropriate parties. This includes law enforcement officials and emergency responders. The gathered frames, timestamp, and location of the observed event are transmitted by the alert. The implementation of this real-time warning system in smart cities improves public safety and expedites emergency response times.

Smart City Integration

The proposed method for detecting violence employs cloud-based monitoring systems and smart city infrastructure that are enabled by the Internet of Things (IoT) to simplify administration and oversight. The technology has the capacity to continuously monitor a variety of locations, including educational institutions, shopping centers, railroad stations, thoroughfares, and cities. This collaboration will lead to an increase in safety and a decrease in crime rates in urban areas. This improves the efficiency of public safety systems in digital cities.

4. RESULTS

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

Fig 4.1: Service Provider Login

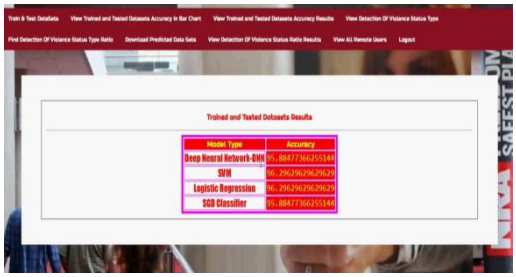

Fig 4.2 : Trained and Tested Dataset Accuracy Results

Fig 4.3: Dataset Accuracy Results in Linechart

Fig 4.4 : Dataset Accuracy Results in Barchart

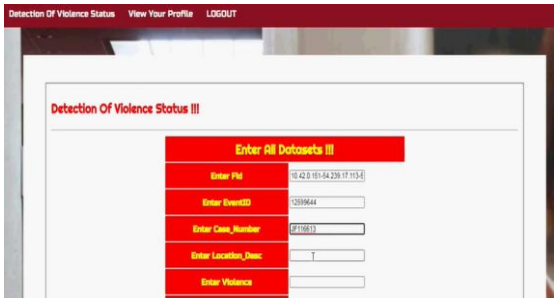

Fig 4.5: Detection of Violence Status

5. CONCLUSION

In conclusion, the Urban Guard concept is a reliable and effective approach to detecting violent incidents that are currently taking place in smart cities. By utilizing convolutional neural networks (CNNs) for spatial feature extraction and recurrent neural networks (RNNs) for temporal modeling, the system is capable of capturing both static and dynamic patterns of aggressive behavior. In terms of precision, recall, accuracy, and F1-score, it surpasses standard and baseline methodologies. It has the potential to generate conclusions in real time, thereby allowing for the dispatch of timely notifications and the effective response of public safety and law enforcement agencies. UrbanGuard has the potential to be a surveillance system that is both intelligent and dependable. It is capable of swiftly training, processing a

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-039
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20610133

variety of datasets, and utilizing a hybrid feature fusion technique. The proposed model suggests that urban safety can be considerably improved by a deep learning-based monitoring system in smart cities.

REFERENCES

1. S. Shilaskar, A. Rajput, A. Rasal, S. Umare, and V. Shelke, "Real-time violence detection using CNN and LSTM," *AIP Conf. Proc.*, vol. 2938, pp. 020004-1–020004-6, 2023.
2. M. Asad, J. Yang, J. He, P. Shamsolmoali, and X. He, "An empirical research of CNN-LSTM on class imbalance datasets for violence video detection," in *Proc. IC3INA*,
3. S. Sudhakaran and O. Lanz, "Learning to detect violent videos using convolutional long short-term memory," in *Proc. IEEE Int. Conf. Adv. Video Signal Based Surveillance (AVSS)*, Lecce, Italy, 2017, pp. 1–6.
4. C. Simisterra-Batallas, P. Pico-Valencia, J. Sayago-Heredia, and X. Quiñónez-Ku, "Internet of Things and deep learning for citizen security: A systematic literature review on violence and crime," *Future Internet*, vol. 17, no. 4, p. 159, 2025.
5. A. Ramzan, H. A. Hamid, and M. Imran, "Violence detection in videos based on CNN features for ConvLSTM2D," in *Proc. ACM Int. Workshop Cross-Data Analysis*, 2024, pp. 45–52.
6. "An efficient violence detection approach for smart cities surveillance system," in *Proc. IEEE Int. Smart Cities Conf. (ISC2)*, Bucharest, Romania, 2023, pp. 1–7.
7. M. Baba, V. Gui, C. Cernazanu, and D. Pescaru, "A sensor network approach for violence detection in smart cities using deep learning," *Sensors*, vol. 19, no. 7, p. 1676, 2019.
8. J. Madake, S. Bhatlawande, A. Rajput, A. Rasal, S. Umare, and V. Shelke, "Violence detection for smart cities using computer vision," in *Artificial Intelligence and Knowledge Processing*, Bentham Science, 2023, pp. 233–247.
9. A. Traoré and M. A. Akhloufi, "Violence detection in videos using deep recurrent and convolutional neural networks," *arXiv preprint arXiv:2409.07581*, 2024.
10. R. Hassner, L. Wolf, and N. Ullah, "Violent flows: Real-time detection of violent crowd behavior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Providence, USA, 2012, pp. 1–6.
11. H. Ullah, M. Ullah, S. Khan, and A. Ullah, "Attention-based deep violence detection in
12. M. Z. Asghar, M. M. Awan, and S. A. Sattar, "Abnormal behavior detection in surveillance videos using hybrid CNN-LSTM," *Multimedia Tools Appl.*, vol. 79, no. 19, pp. 13353–13372, 2020.
13. H. Li, G. Luo, and T. Jiang, "Violence detection in surveillance video based on bidirectional ConvLSTM," *Pattern Recognit. Lett.*, vol. 144, pp. 1–8, 2021.
14. A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action recognition in video sequences using deep bi-directional LSTM with CNN features," *IEEE Access*, vol. 6, pp. 1155–1166, 2018.