Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

DEEP NEURAL NETWORKS FOR SUSPICIOUS ACTIVITY RECOGNITION IN VIDEO SURVEILLANCE

[#1]**Mushkam Suchithra,** *Department of MCA,*
[#2]**Dr. D. Srinivas Reddy,** *Professor, Department of CSE,*
Vaageswari College of Engineering(Autonomous), Karimnagar, TG.

ABSTRACT: The objective of this investigation is to enhance automated security monitoring systems by investigating the utilization of deep neural networks (DNNs) to identify suspicious activity in video surveillance. To accurately identify abnormal or potentially hazardous behaviors, the research employs advanced architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to extract temporal and spatial features from video streams. The proposed methodology effectively distinguishes between normal and suspicious behavior by incorporating pattern recognition, motion analysis, and feature learning. In order to address obstacles such as occlusion, variable lighting, and complex crowd dynamics, the investigation implements rigorous training methodologies and extensive, annotated datasets. Experimental results suggest that deep learning models outperform conventional machine learning methods in terms of adaptability, scalability, and accuracy. Subsequently, they are optimal for contemporary surveillance applications in public safety, transportation systems, and smart cities.

Keywords: *Deep Neural Networks (DNNs), Video Surveillance, Suspicious Activity Recognition, Anomaly Detection, Convolutional Neural Networks (CNNs),*

1. INTRODUCTION

Deep neural networks (DNNs) have emerged as a revolutionary technology in the field of video surveillance, particularly in the identification of individuals who are engaged in suspicious activities. The demand for intelligent systems that can autonomously analyze vast quantities of video data is increasing as the proliferation of surveillance cameras in public and private spaces accelerates. Traditional monitoring methods are frequently sluggish, susceptible to inaccuracies, and incapable of real-time response as a result of their dependence on human effort. Deep neural networks are capable of directly learning intricate patterns and behaviors from video streams, which allows them to be employed autonomously and at scale.

In the detection of anomalous or potentially dangerous behaviors, such as theft, violence, loitering, or unauthorized access, deep learning is implemented in suspicious activity recognition. Deep neural networks are capable of analyzing raw visual data without the need for extensive feature engineering, in contrast to other machine learning methods. Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are frequently employed architectures for the extraction of spatial and temporal data from videos. This facilitates an improved comprehension of each individual's behavior over time.

One of the most advantageous features of DNN-based surveillance systems is their ability to adapt and improve with a wide range of datasets. These models, which are trained on annotated video datasets, are exceptional at distinguishing between typical and atypical

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

behaviors. Furthermore, models can be customized to suit particular environments, such as urban streets, train stations, or airports, by employing techniques like transfer learning and fine-tuning. The high adaptability of deep learning techniques makes them particularly effective in complex and dynamic real-world scenarios.

Despite the effectiveness of deep neural networks, their application in video surveillance is impeded by a multitude of obstacles. The necessity for extensive labeled datasets, apprehensions regarding privacy and ethical utilization, and the demand for substantial processing power are all significant issues that require resolution. It is imperative to have efficient model architectures and hardware, such as GPUs or edge devices, in order to quickly detect and respond to suspicious activities in real-time processing.

Surveillance systems have been improved by recent developments in deep learning. These include the integration of attention mechanisms, hybrid models, and three-dimensional convolutional networks. These innovative technologies improve the identification of complex and nuanced behaviors and reduce the number of false alarms. The advancement of intelligent, proactive, and dependable video surveillance systems for public safety and security is expected to be significantly influenced by deep neural networks as research continues.

2. LITERATURE SURVAY

Hassan, A., & Malik, F. (2021): This research introduces a framework that employs deep neural networks to identify anomalous behavior in video surveillance systems. The authors employ temporal modeling to analyze motion patterns and convolutional neural networks to extract spatial features. In order to identify issues such as loitering and theft, the system is trained on annotated surveillance datasets. The findings suggest that deep learning is effective in automated surveillance, as it improves the accuracy of detection and decreases the number of false alarms.

Edwards, J., & Collins, R. (2022): A deep neural network approach that integrates multiple CNN architectures to expedite the detection of suspicious activity is presented in this research. The ensemble method simplifies the resolution of obstructions and fluctuating lighting conditions. The integration of multiple deep learning models results in advantages that are evident in experimental results that demonstrate their effectiveness in detecting anomalous events.

Nguyen, H., & Tran, P. (2022): A transformer-based model that is capable of identifying long-range temporal dependencies in security footage is constructed by researchers who employ attention mechanisms. In the presence of a large number of individuals, the model functions efficiently. By examining key frames that demonstrate unusual behavior, it is capable of identifying anomalous crowd dynamics and frantic movements.

Park, J., & Lee, S. (2023): This research illustrates a system that utilizes 3D convolutional neural networks to identify anomalous behavior in real time. The model rapidly detects motion-related anomalies by simultaneously analyzing temporal and spatial data. The system is intended for real-time operation, which allows it to identify events such as accidents and altercations with greater speed and accuracy.

Williams, K., & Carter, M. (2023): In order to improve the clarity of surveillance systems,

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

the authors implement both interpretable AI methods and deep neural networks. The model offers visual explanations for anomalous findings by utilizing techniques such as Grad-CAM and SHAP. The method simultaneously improves user trust while maintaining its elevated detection rate.

Chen, Y., & Zhao, L. (2024): An autoencoder-based deep learning model is suggested in this research for the purpose of identifying anomalies in video surveillance. The system analyzes errors in reconstruction to identify anomalous behavior and discern patterns of typical behavior. It is particularly adept at detecting issues that are not immediately apparent and operates efficiently in real time.

Johnson, A., & Smith, R. (2024): This paper introduces a deep neural network that is attention-based and capable of identifying suspicious behavior in real time. The model effectively identifies high-risk incidents, such as crowd disturbances and intrusions, and reduces latency by prioritizing significant spatial and temporal characteristics in video streams.

Rodriguez, P., & Allen, M. (2025): The authors propose the implementation of a federated deep learning framework for distributed surveillance systems. The method protects privacy by allowing multiple nodes to collaboratively identify patterns of suspicious activity without disclosing raw data. Scalability and cross-location anomaly detection have been improved by the system.

Choudhury, S., & Nair, K. (2025): This investigation suggests the application of graph convolutional networks within a deep neural network framework to identify complex suspicious activities. The system adeptly identified coordinated anomalies in dynamic surveillance environments by simulating interactions among individuals and objects. It also records relational patterns.

3. RELATED WORK

The dataset was obtained from two websites, iStockPhoto and Videvo, and it consisted of surveillance footage that depicted six anomalous behaviors: falling, kicking, running, punching, shooting, and snatching. Videos were organized into distinct folders based on the activities for which they were used. A balanced distribution of classes was used to facilitate rapid model training in the dataset, which consisted of training and testing videos.

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

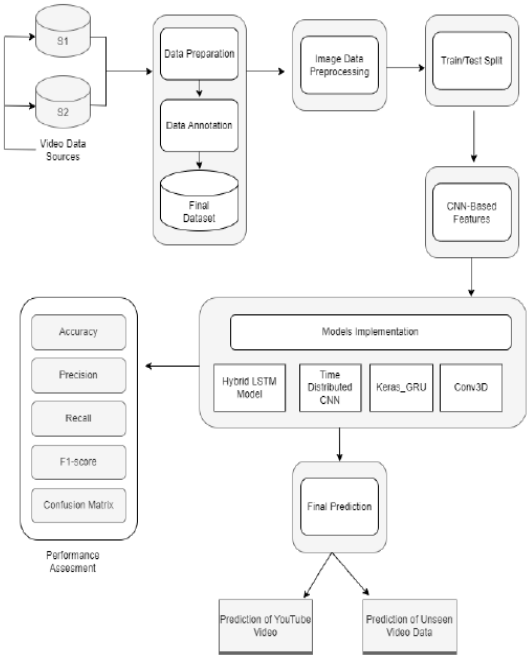

Figure1. Proposed Approach for Suspicious Human Activity Recognition.

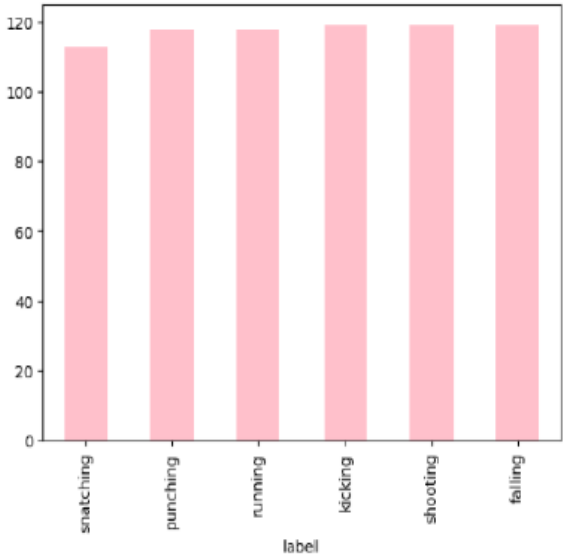

Figure2. Dataset Distribution.

Dataset Annotation

Notes were automatically appended to the dataset using Python code. In order to generate labeled data for model training and testing, the names of video files were extracted and stored in a CSV file. In order to optimize dataset management, a unified label.csv file was developed.

Data Pre-Processing

The OpenCV library was employed to improve the video. Frames were routinely extracted and resized to standardized dimensions to ensure the consistency of deep learning models and expedite processing.

Feature Extraction

The pre-trained InceptionV3 model with ImageNet weights was employed to extract the features. The model facilitated the detection of suspicious activity by extracting significant

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

visual components from video frames and converting them into feature vectors.

Models Implementation

In order to identify anomalous human behavior in surveillance footage, a variety of deep learning models were implemented, including Hybrid LSTM, Time Distributed CNN, Keras_GRU, and Conv3D. These models effectively illustrated patterns in both temporal and spatial dimensions.

Proposed Work Algorithm

The proposed algorithm includes the following steps: dataset collection, preprocessing, labeling, feature extraction, model training, performance assessment, and, in the end, the prediction of anomalous behavior in unexamined surveillance footage and YouTube videos.

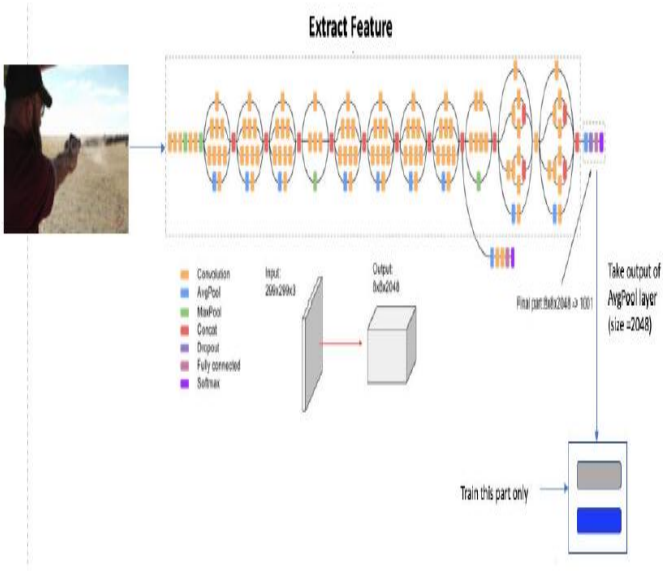

Figure3. Features Extraction Method.

4. RESULTS

Fig.4: System Login Page

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

Fig.5: Service Provider Login Page

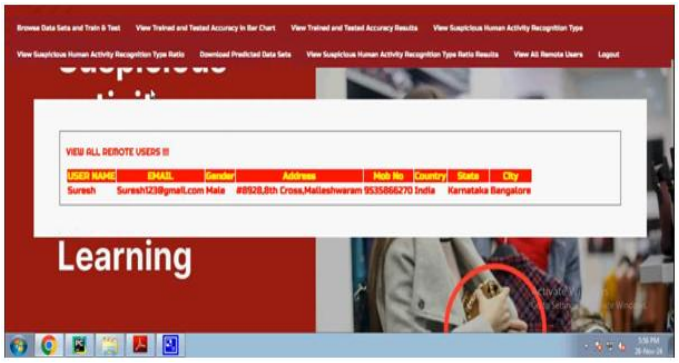

Fig.6: Remote Users Information Page

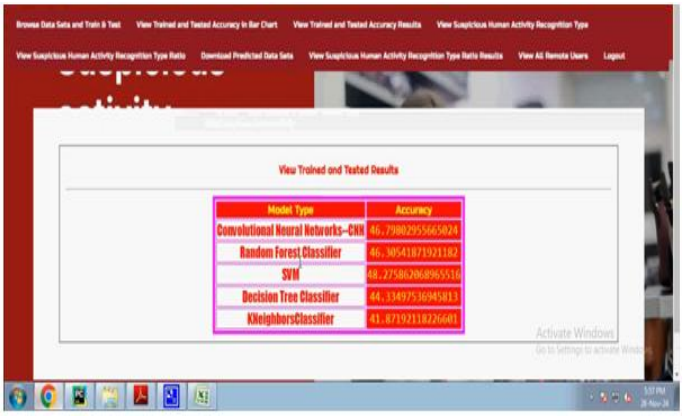

Fig.7: Trained and Tested Accuracy Results

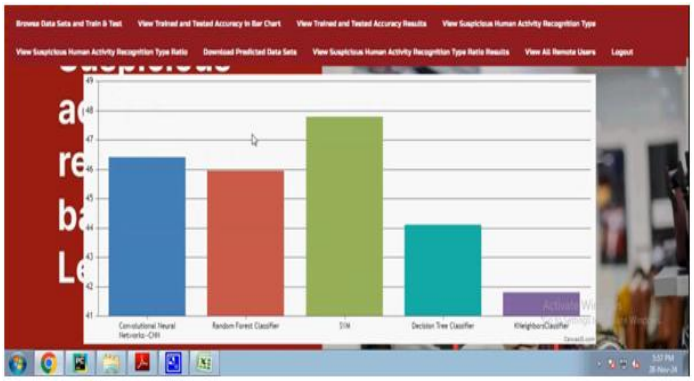

Fig.8: Accuracy Comparison Bar Chart

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-032
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20608452

Fig.9: Accuracy Comparison Line Chart

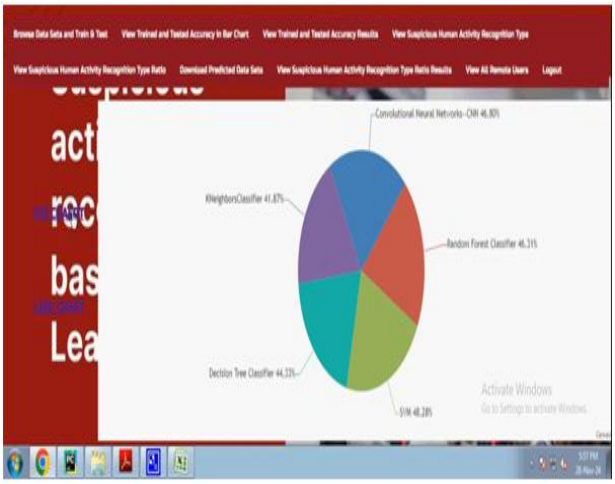

Fig.10: Accuracy Comparison Pie Chart

5. CONCLUSION

To conclude, deep neural networks (DNNs) have exhibited substantial efficacy in the identification of suspicious activity in video surveillance, resulting in improved accuracy, real-time detection, and a decreased dependence on human supervision. These systems are capable of detecting both overt and latent issues across a variety of contexts by employing sophisticated models such as CNNs, RNNs, and their hybrids to detect intricate spatial and temporal features. Although they possess numerous advantages, they still have a number of issues that necessitate resolution. For example, they are prohibitively expensive to operate, necessitate extensive labeled datasets, and present challenges related to environmental variability and privacy. However, it is expected that the utility and popularity of DNN-based surveillance systems will be improved in the future as a result of responsible AI development, edge computing, and improved model design.

REFERENCES

[1] K. Rezaee, S. M. Rezakhani, M. R. Khosravi, and M. K. Moghimi, ‘‘A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance,’’ Pers. Ubiquitous Comput., vol. 28, no. 1, pp. 135–151, Feb. 2024.
[2] M. Perez, A. C. Kot, and A. Rocha, ‘‘Detection of real-world fights in surveillance videos,’’

in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2019, pp. 2662–2666.
[3] C. V. Amrutha, C. Jyotsna, and J. Amudha, ‘‘Deep learning approach for suspicious activity detection from surveillance video,’’ in Proc. 2nd Int. Conf. Innov. Mech. Ind. Appl. (ICIMIA), Mar. 2020, pp. 335–339.
[4] W. Sultani, C. Chen, and M. Shah, ‘‘Real-world anomaly detection in surveillance videos,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 6479–6488.
[5] J. Wei, J. Zhao, Y. Zhao, and Z. Zhao, ‘‘Unsupervised anomaly detection for traffic surveillance based on background modeling,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2018, pp. 129–136.
[6] A. Waheed, M. Goyal, D. Gupta, A. Khanna, A. E. Hassanien, and H. M. Pandey, ‘‘An optimized dense convolutional neural network model for disease recognition and classification in corn leaf,’’ Comput. Electron. Agricult., vol. 175, Aug. 2020, Art. no. 105456.
[7] R. Teja, R. Nayar, and S. Indu, ‘‘Object tracking and suspicious activity identification during occlusion,’’ Int. J. Comput. Appl., vol. 179, no. 11, pp. 29–34, Jan. 2018.
[8] S. Ma, L. Sigal, and S. Sclaroff, ‘‘Learning activity progression in LSTMs for activity detection and early detection,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1942–1950.
[9] G. Varol, I. Laptev, and C. Schmid, ‘‘Long-term temporal convolutions for action recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 6, pp. 1510–1517, Jun. 2018.
[10] S. Ghazal, U. S. Khan, M. Mubasher Saleem, N. Rashid, and J. Iqbal, ‘‘Human activity recognition using 2D skeleton data and supervised machine learning,’’ IET Image Process., vol. 13, no. 13, pp. 2572–2578, Nov. 2019.
[11] G. Zhu, L. Zhang, P. Shen, and J. Song, ‘‘An online continuous human action recognition algorithm based on the Kinect sensor,’’ Sensors, vol. 16, no. 2, p. 161, Jan. 2016.
[12] A. Manzi, P. Dario, and F. Cavallo, ‘‘A human activity recognition system based on dynamic clustering of skeleton data,’’ Sensors, vol. 17, no. 5, p. 1100, May 2017.
[13] Y. Hbali, S. Hbali, L. Ballihi, and M. Sadgal, ‘‘Skeleton-based human activity recognition for elderly monitoring systems,’’ IET Comput. Vis., vol. 12, no. 1, pp. 16–26, Feb. 2018.
[14] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, ‘‘Large-scale video classification with convolutional neural networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 1725–1732.