Journal of Information Systems Engineering and Management
2025, 10(7s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Intelligent Surveillance System for Real-Time Detection of Anomalous Activities in Video Streams

Aarya Santosh Gadekar[1], Nidhi Vijay Surve[2], Dr. Ekta Sarda[3], Pooja Yogesh Patil[4], Sakshi Sunil Sawant[5]

[1]Department of Computer Engineering, Ramrao Adik Institute of Technology D. Y Patil Deemed to be University, India.
aaryagadekar1405@gmail.com

[2]Department of Computer Engineering, Ramrao Adik Institute of Technology D. Y Patil Deemed to be University, India.
survenidhi05@gmail.com

[3]Department of Computer Science and Engineering, Ramrao Adik Institute of Technology D.Y Patil Deemed to be University, India.
ekta.sarda@rait.ac.in

[4]Department of Information Technology, Ramrao Adik Institute of Technology D.Y Patil Deemed to be University India.
poo.pat.rt19@rait.ac.in

[5]Department of Information Technology , Ramrao Adik Institute of Technology D Y Patil Deemed to be University India.
Sak.saw.rt19@rait.ac.in

ARTICLE INFO

Received: 14 Oct 2024

Revised: 13 Dec 2024

Accepted: 26 Dec 2024

ABSTRACT

The increasing complexity of surveillance structures necessitates advanced techniques for monitoring big volumes of video statistics. This record opinions the utility of convolutional neural networks (CNN) and deep getting to know techniques for detecting suspicious pastime in video streams. The technique involves preliminary video statistics processing to extract key features, observed by training a CNN version to distinguish between normal and abnormal behaviours through recognizing spatial and temporal patterns within the video frames. techniques consisting of transfer gaining knowledge of and statistics augmentation are employed to enhance the model's generality and robustness. The effectiveness of this approach was validated via various checks, including experiments on datasets like u.s.a. Pedestrian and road, where the CNN-primarily based technique established high accuracy and go back quotes in figuring out suspicious activities. The scalability and actual-time processing abilities of the version make it adaptable to various tracking environments. These findings are great for the advancement of the surveillance era, offering a dependable technique for real-time detection of suspicious sports in video streams. The proposed CNN-based approach is promising for bolstering safety in public spaces, transportation structures, and important infrastructure, thus contributing to better safety features.

Keywords: Surveillance structures ,Convolutional Neural Networks (CNN), Deep learning techniques, Suspicious activity detection, Video streams, Feature extraction, Transfer learning , Real-time processing.

1.INTRODUCTION

The paper investigates the utility of deep getting to know strategies, mainly convolutional neural networks (CNNs), for detecting suspicious activities across various domain names which include surveillance, cybersecurity, and monetary systems. conventional techniques, frequently rule-primarily based, struggle to detect new or diffused patterns of suspicious conduct efficiently. CNNs, through their ability to autonomously study complicated hierarchies of features from uncooked records, offer full-size improvements in both accuracy and performance. The studies leverage various data sources inclusive of video, photographs, community traffic logs, and financial transactions to educate deep studying models. Those fashions excel in obligations such as photo reputation, natural language processing, and anomaly detection, making them ideally fitted for figuring out complex patterns indicative of suspicious sports. Moreover, the integration of synthetic intelligence (AI), gadget mastering (ML), and net of things (IoT) technologies allows actual-time applications on cellular and other related devices. The paper presents a method that makes use of ML to research sequences of video frames

for abnormal activity detection, with a focal point on dynamic environments like childcare and eldercare facilities. strange activities detected are without delay communicated to give up-users via IoT notification devices, enhancing the reaction capabilities to potential threats or emergencies.

Key contributions of this work consist of the development of a robust model for detecting both static and dynamic ordinary sports, the use of a deep neural community, and the usage of multi-type techniques for correct activity prediction. This method appreciably advances the sector through addressing the constraints of previous techniques that centred predominantly on static activity prediction.

2. RELATED WORK

The prominent growth of digital data and the maturation of malicious activities over the last decade have drawn attention to the urgent demand for effective suspicious activity detection systems. Subsequently, deep learning, a subset of machine learning methods, has become an extremely favourable tool for learning patterns and extracting key features from vast amounts of data. This literature review investigates the latest advances, methodologies, and issues surrounding the implementation of deep learning techniques to identify suspicious activities in various areas such as cybersecurity, finance, health care and other domains. The intention of this work is to present the current state-of-the-art and study the most recent experimental approaches and work with deep learning to identify trends, potential gaps and challenges of suspicious activity detection.

The review of existing systems introduces a variety of deep learning applications to identify suspicious activities in various settings. The work presented in [1] focused on children's safety in parks and daycare centres using an innovative Random Forest Evolution (RFKD) method with adaptive motion compensation and kernel density, coupled with an MQTT protocol for real-time alerts. [2] adopts a semantic approach to distinguish between living and non-living objects in videos, applying the Haar algorithm for classification and using motion features to detect suspicious activities.[3] Classification of human activities into normal and suspicious using convolutional neural networks for feature extraction and activity prediction.[4] uses MATLAB to design a cost-effective home security system and integrate Pi cameras and sensors managed via GSM for monitoring. [5,14] developed a fall detection system using a PCANet trained on different video sequences and established an SVM-based action model. [6] implemented an anomaly detection framework at ATM locations using video analysis techniques such as MHI and Hu Moments, enhanced by PCA and SVM for event reporting.[7,10] reviews background modelling techniques to improve object detection accuracy in video surveillance.[8] discusses the complexity of medical data and highlights the need for reliable abnormality detection in radiotherapy. [9] introduced a two-part image analysis method for advanced behavioural tracking. Finally, [10] describes a real-time violence detection system in football stadiums that uses video analytics in the Spark environment to maintain safety by proactively alerting security personnel. Collectively, these studies demonstrate the innovative integration of deep learning technologies to enhance safety and security in various domains, [4,8, 11] introducing multi-object detection systems and recognizing abnormal behaviour to prevent ATM crime. It recognises human approach utilised features and access the functions used, while a classifier should be used for human detection. When there is case of partial occlusion, this system may not detect a person. [12] presented a new method for the detection of abandoned objects. The method involved using kernel-based object tracking to track suspicious human activity and object tracking. They used a histogram of the forward-backward ratio and a finite-element machine to identify transmission conditions, achieving 100% accuracy in detecting abandoned objects. [13] In this study, a fall detection system framework with automatic feature learning was proposed. The training set contains images where people are captured from different viewpoints in video sequences. After training PCANet with all samples, the label of each image was predicted.

Existing ATM vulnerability detection systems using machine learning face several significant limitations that hinder their effectiveness. First, scalability issues are evident as these systems often struggle to process real-time transactions, making fraud detection capabilities slow or ineffective. In addition, the common problem of data imbalance favours legitimate transactions over fraudulent transactions in the training data, leading to high false positive rates and lower overall accuracy. Adapting these models to new and evolving fraud tactics remains a major challenge as fraud schemes continue to evolve and outpace current detection technologies. Finally, there is a great need for more transparent and explainable machine learning models. Current systems lack sufficient explanation capabilities necessary to gain user trust and understand the reasons behind flagged suspicious activity. To address these gaps, new approaches can improve real-time processing, improve data imbalance,

adapt more dynamically to emerging threats, and increase the transparency of the decision-making process in ATM fraud detection systems.

3. PROPOSED WORK

The proposed research aims to integrate machine learning and deep learning techniques to develop an automatic surveillance system that detects suspicious activities in video feeds. This initiative addresses the ineffectiveness of traditional human surveillance in the face of the growing volume of video data in the public and private sectors. The new system aims to reduce problems such as high false alarm rate, rigidity in different environmental conditions and delayed response. The paper explores and evaluates different deep learning models to effectively detect suspicious behaviour under different scenarios, such as changing camera angles and lighting conditions. The goal is to provide a dynamic and responsive system that increases surveillance efficiency, reduces response time, and supports crime prevention efforts in a variety of environments.

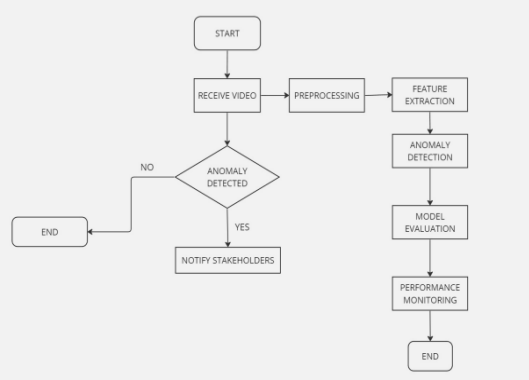

FIGURE 3.1: Proposed System

In the above Figure 3.1 the flowchart depicts how a video processing system finds anomalies. It begins by obtaining a video, which is then preprocessed and evaluated for feature extraction. When anomalies are found, stakeholders are notified, followed by model evaluation and performance monitoring, which brings the process to a conclusion. If no anomalies are discovered, the process concludes.

The proposed system focuses on the detection of abnormalities within the ATM vestibule. It is configured to detect abnormal activity by dividing its operation into two main phases: the training phase and the detection phase. The method, shown in Figure 1, follows standard procedures for detecting abnormal behaviour and is specifically adapted to the ATM context to enhance security and effectively detect anomalies.

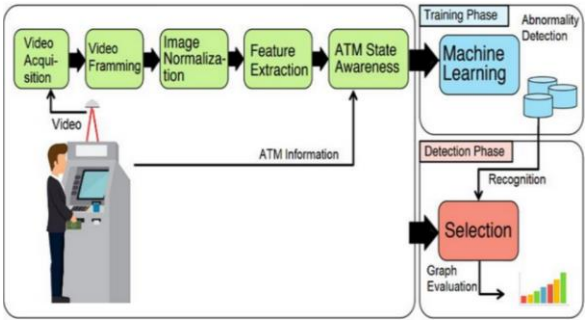

FIGURE 3.2 : SYSTEM ARCHITECTURE

Figure 3.2 shows that the proposed system for ATM atrial abnormality detection consists of several main steps designed for efficient processing and analysis of video data.

a) Video Ingest: This first step involves capturing video data, importing legacy formats, and synchronising multiple video streams into a common timeline for continuous playback and processing.

b) Video Framing: In this step, frames are extracted from the input video and duplicate frames are removed to simplify the analysis process. This step reduces workload by focusing on individual images instead of the entire video sequence.

c) Image normalisation: This process adjusts the pixel intensity values of an image within a standard range and makes the image more suitable for processing and analysis by standardising the visual input to a more common scale.

d) Feature extraction: This important step involves extracting various features from the image, such as shape, edges and scale invariant feature transform (SIFT) features. These features are important for classifying and identifying anomalies in videos.

e) Graph evaluation: Once an anomaly is detected using a machine learning algorithm, it is important to evaluate the performance of the algorithm. This is done by evaluating metrics such as precision, recall, f-measure, and accuracy on a test dataset to investigate the effectiveness of anomaly detection.

The system also uses advanced feature extraction techniques.

- Contextual Feature Extraction: Capture spatial and temporal information from video frames to understand scene context and detect deviations from normal behaviour.
- Spatial feature extraction: describes the static content of a scene without temporal dynamics, focusing on capturing visual patterns and structure within the frame.
- Motion Feature Extraction: Analyses motion patterns in frames to identify unusual movements and behaviours that are important for dynamic anomaly detection.

4. RESULTS AND DISCUSSIONS

[14] and [15] ATM Dataset is a collection meant for training and testing machine learning models for detecting anomalies in ATM vestibules. The dataset provides temporally annotated movies that distinguish between regular and abnormal activities, making it excellent for designing systems to enhance ATM security. The dataset is ideal for researchers and developers that want to test real-time monitoring and anomaly detection systems.

The model design includes several input layers that can handle a variety of data formats, such as RGB pictures, sequences, and time series. A pretrained ResNet50v2 CNN is used to extract features from images. These characteristics, along with other input data, are concatenated and then processed through dense layers with dropout to avoid overfitting. The final output layer represents a two-class classification task. The model has about 42 million parameters, with a large fraction of them non-trainable, most likely due to pre-trained layers such as ResNet50v2. The complicated architecture implies a multi-input task that combines image processing with other inputs to make comprehensive predictions or classifications, with the goal of producing finer and more context-aware outcomes.

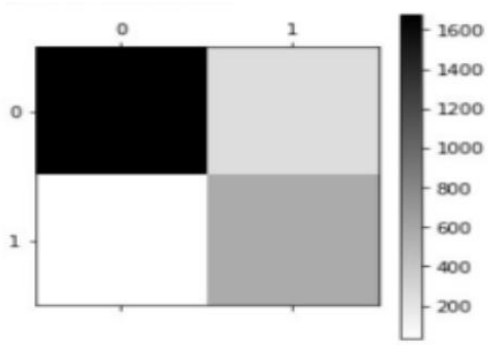

Figure 4.1: CONFUSION MATRIX

Accuracy	0.8835482610394686
Precision	0.6903914590747331
Recall	0.9402261712439418
F1 Score	0.7961696306429549
False Positive Rate	0.1345360824742268
ROC AUC	0.9260513298802525

Table No.1 : Performance metrics for CNN Model

1. ACCURACY

$$\text{Accuracy} = \frac{P + N}{P + N + FP + FN}$$

2. PRECISION

$$\text{Precision} = \frac{P}{P + FP}$$

3. RECALL

$$\text{Recall} = \frac{TP}{TP + FN}$$

4. F1 SCORE

$$F1\ \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Substituting Precision and Recall

$$F1\ \text{Score} = 2 \times \frac{\frac{P}{P+FP} \times \frac{P}{P+FN}}{\frac{P}{P+FP} + \frac{P}{P+FN}}$$

5. FALSE POSITIVE RATE

$$\text{False Positive Rate} = \frac{FP}{FP + N}$$

6. ROC AUC

ROC AUC=Area Under the ROC Curve

Here, P for True Positives (TP), N for True Negatives (TN), FP for False Positives, FN for False Negatives.

Figure 4.1 and Table 1 shows that the recall value has the highest percentage to find the anomalies as compared to false positive and precision having less as compared to other parameters to find the video stream.

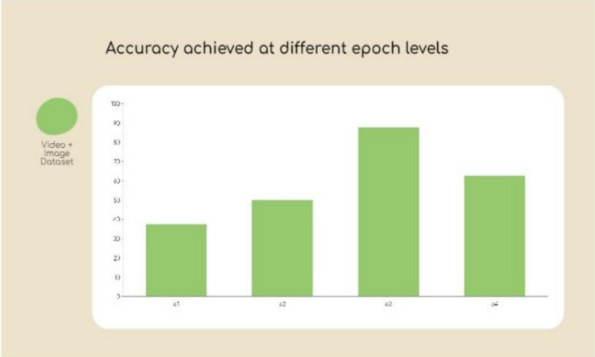

Figure 4.2 : Accuracy achieved over different epoch levels

Graph presents the accuracy tiers carried out at numerous epoch degrees at some point of a training process for a model using a video and photograph dataset. The accuracy possibilities boom notably from epoch 1 to epoch three, peaking at epoch 3, earlier than slightly decreasing at epoch four. The respective epoch levels are: epoch 1 = 10, epoch 2 = 7, epoch 3 = 5, epoch 4 = 3 as shown in figure 4.2.

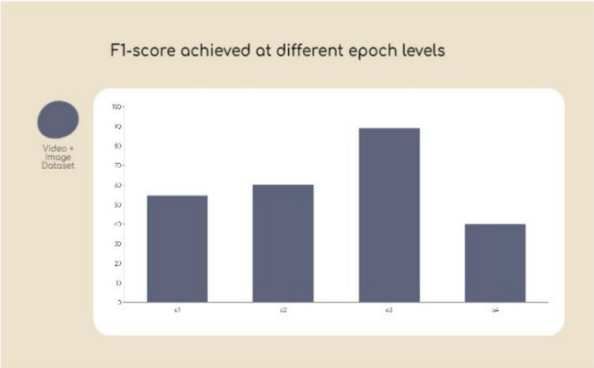

Figure 4.3: Score achieved over different epoch levels

The above figure 4.3 graph shows the F1-rankings done at various epoch levels throughout a model education using a video and photo dataset. Epoch 1 uses 10, epoch 2 uses 7, epoch 3 uses 5, and epoch 4 uses 3.

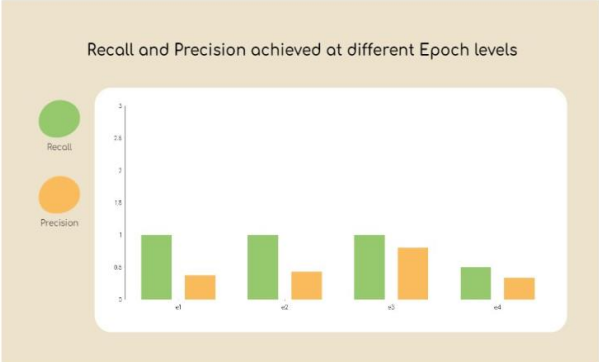

Figure 4.4: F1 Score achieved over different epoch levels

The above figure 5.4 shows the recall and precision values for a model at different epochs, labelled as e1 to e4. Green bars represent recall, and orange bars represent precision .

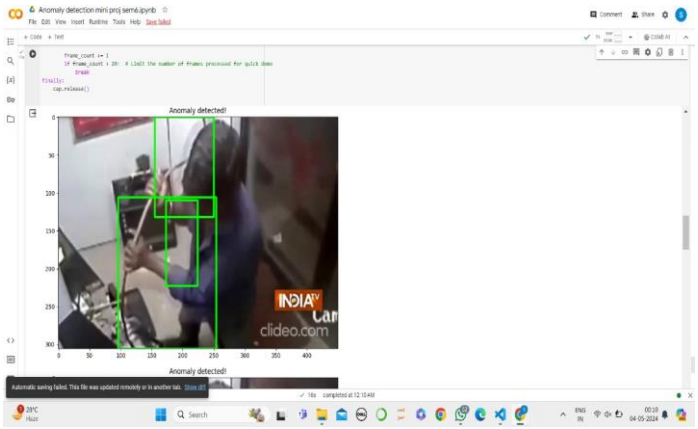

Figure 4.5: Anomaly Detected

The figure 4.5 shows the analysis of videos with anomalies detected using the CNN model. The detected anomalies are highlighted with a green border, indicating that suspicious activity is present.

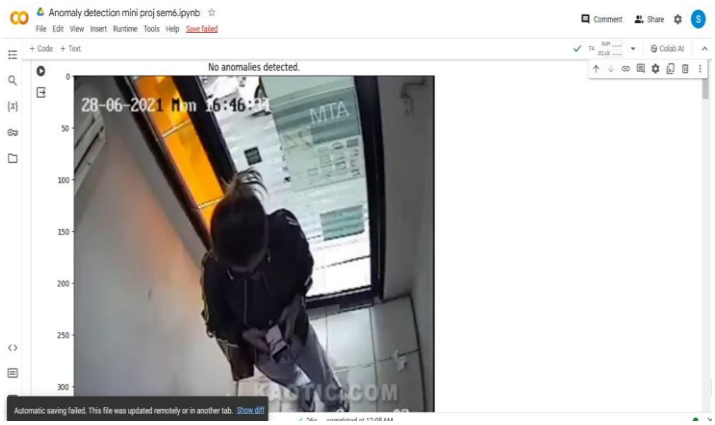

Figure 4.6: Anomaly not detected

The above figure 4.6 image shows analysis of videos that is not significantly different from the CNN-based model. The absence of empty boxes indicates that the behaviour is always appropriate.

5. CONCLUSION

The potential of deep learning, and CNNs in particular, in detecting suspicious activity in the areas of regulatory, cyber security, and finance. Using the integration of ML and IoT enables real-time monitoring in various environments such as daycares and nursing homes, and helps develop robust diagnostic tools. Experimental results show that this approach is superior to traditional methods. Future research should focus on diversifying datasets, improving real-time processing, and investigating advanced anomaly detection algorithms such as GANs and autoencoders. Increasing compatibility between domains and incorporating user feedback mechanisms are also important to improve detection accuracy. Ethical considerations regarding privacy and legal implications must be taken into account to ensure a reliable monitoring system. Addressing these challenges will lead to more complex and adaptable systems, thereby increasing security in a wide variety of applications. The future of intelligent analytics for real-time detection of abnormal activity in video streams includes many new advances. First, developing different knowledge will lead to better models for different areas such as public spaces, transportation facilities and important systems, making it possible to do well in different situations. Integrating multiple deep learning algorithms, including artificial neural networks (GANs) and autoencoders, can improve the accuracy and efficiency of anomaly detection. In addition, improving the operating capacity of the end equipment will increase the strength and performance of the airframe, making it suitable for large-scale deployment. Collaboration with users will help continuously improve the model and align its performance with real-world needs. In addition, expanding its use in other areas such as health, security and finance will increase its benefits and impact. Resolving ethical issues, especially privacy and legal implications, is important for the acceptance and trust of the system. By exploring these future directions, the program can lead to smart, flexible and secure monitoring solutions that can help improve public safety and security.

REFERENCES

[1] Vallathan, G., John, A., Thirumalai, C. *et al.* Suspicious activity detection using deep learning in secure assisted living IoT environments. *J Supercomput* 77, 3242–3260 (2021).

[2] Dinama, D. D. M., A'yun, Q., Syahroni, A. D., Sulistijono, I. A., & Risnumawan, A. (2020). A semantic approach to real-time fire detection and suspicious activity identification.

[3] Barsagade, K., Tabhane, S., Satpute, V., & Kamble, V. (2023). Suspicious activity detection using deep learning approach.

[4] Faisa, Z. G. (2020). The proposed system is implemented using software and hardware components in MATLAB to achieve the goal of home security.

[5] Wang, et al. (2019). A fall detection system framework with automatic feature learning.

[6] Tripathi, et al. (2019). A framework was developed to identify unusual behaviour at ATM locations such as money theft, assaulting customers, arguing with customers and alerting the police in case of such events.

[7] Bouwmans, T. (2019). Survey of background modelling techniques for foreground detection in videos from stationary cameras. *Journal of Background Modelling*.

[8] V. I. IMRT. (2021). Investigation of current systems and practices in radiation therapy: Data inaccuracies and anomaly detection in medical records.

[9] A. C.V., C. Jyotsna, and A. J. Amudha, "Different techniques are used to help distinguish different types of suspicious behaviour by tracking depending on the images," 2020.

[10] Fenil, E., Manogaran, G., Vivekananda, G. N., Thanjai Vadivelu, T., Jeeva, S., & Ahilan, A. J. C. N. (2019). Real time violence detection framework for football stadium comprising big data analysis and deep learning through bidirectional LSTM. *Computer Networks*, *151*, 191-200.

[11] Mani, D. R., Ninan, N. D., Kumar, C. S., & Ramachandran, V. (1992). Implementation of active power filter using hybrid neural network.

[12] Sujith, B. (2014). Crime detection and avoidance in ATM: A new framework. *International Journal of Computer Science and Information Technologies*.

[13] Chiu, C., Zhan, J., & Zhan, F. (2017). Uncovering Suspicious Activity From Partially Paired and Incomplete Multimodal Data.

[14] ATM-I Dataset: https://doi.org/10.34740/kaggle/ds/2080545.

[15] ATMA-V Dataset: https://doi.org/10.34740/kaggle/dsv/3455016.