Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

Smart Surveillance: Detecting Abnormal Human Activities with Deep Learning

Ragi Rakshitha Reddy, Muddam Sai Manish Yadav, Abhiram Venkata Daita
Department Of Computer Science and Engineering
Geethanjali College Of Engineering and Technology
Hyderabad, India

Under the Guidance of Dr. K. Srinivasa Reddy
Associate Professor, Department of Computer Science and Engineering
Geethanjali College of Engineering and Technology, Hyderabad, India

***Abstract*— We present a smart surveillance system to detect anomalous activities in video streams using deep learning. In the presented work, a Convolutional Neural Network (CNN) of Conv2D layers is trained to classify video frames as either normal or anomalous. Each frame is processed by the system, generating alerts when aberrant behavior is detected. The system features image and video file input and offers an extremely lightweight yet effective solution to enable automated surveillance operations.**

***Keywords*—Smart Surveillance, Anomaly Detection, Deep Learning, Convolutional Neural Network, Video Processing, Security System, Conv2D**

I. INTRODUCTION

The recent years have come to be linked with public security, and both industries-private and public-have been shaken in recent times with more reports of thefts, violence, and illegal intrusions. Conventional surveillance systems depend heavily on security personnel for manual observation, and therefore are ineffective in real-time detection of important events. Human frailties such as fatigue, distraction, and limited attention span serve as impediments to the processing of vast amounts of video data generated by multiple cameras. Therefore, important events may be false alarms or may be received too late to take preventive or corrective measures.

In resolving the above issues, the quest for intelligent surveillance systems has been of particular interest to researchers. Intelligent surveillance systems that utilize artificial intelligence (AI) and deep learning methods enable the video analysis process to be automated. They can continuously monitor the environment and detect suspicious or unusual behavior without human intervention by trained security officers. Merging automation with video surveillance would significantly enhance the speed and accuracy of threat detection, leading to more efficient security operations.

Deep learning, which is one of the machine learning subclasses inspired by the human brain-computing ability to look at visual data-, has brought a revolutionary change in how machines understand vision. Among the different classes of deep learning, Convolutional Neural Networks (CNNs) stand out as being extremely good at image recognition, object detection, and video classification. These frames are tailored to learn automatically the spatial hierarchies of features from the input images or videos using layers of filters and pooling mechanisms. When these models are employed in video surveillance, they can recognize normal activities from clinching ones based on visual patterns, even so, in an environment complicated by high-density crowds.

The paper introduces intelligent surveillance based on a CNN model developed from Conv2D layers in frame-by-frame video classification. The model is trained to detect anomalies in live video streams and raises an alarm whenever a human activity that is anomalous is detected. In comparison to a standard surveillance setup, the solution in this paper offers a lightweight, scalable, and automated solution that can be implemented in environments like schools, shopping malls, workplaces, parking spots, and other public transit place networks.

This infrastructure is intended to augment situational awareness with an offset decrease in reliance on manual surveillance, enabling accelerated human response to potential threats. Connecting the gap between conventional surveillance systems and intelligent automated security infrastructure with deep learning approaches, coupled with an easy-to-use interface and real-time alert system, is the objective of the suggested smart surveillance framework.

II. LITERATURE SURVEY

With the progress in computer vision and artificial intelligence, numerous research works have tried to leverage deep learning for autonomous surveillance. Among them, Convolutional Neural Networks (CNNs) have consistently emerged as extremely powerful in tasks such as human pose estimation, object detection, activity recognition, and anomaly detection. Their ability to learn automatically and extract hierarchical spatial features makes them ideal for high-level surveillance tasks.

CNN-Based Surveillance:

A few of the popular models like VGGNet, ResNet, and MobileNet have been adapted for surveillance due to their feature extraction capability. Researchers like Sultani et al. suggested anomaly detection using deep MIL (Multiple Instance Learning), where video clips are learned in an

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

unannotated frame-wise fashion. These approaches facilitate weakly supervised learning over real-world surveillance videos.

Temporal Pattern Recognition:
To incorporate temporal data, researchers have employed Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) units. The models can learn motion and sequence patterns over time and, from this, identify subtle or progressive anomalies. For instance, autoencoder-LSTM models have been used for frame prediction, where large reconstruction errors signal abnormal behavior.

Hybrid and Transformer Models:
Recent studies have also witnessed the utilization of hybrid models that integrate CNNs and attention-based mechanisms, i.e., Transformers. Such models improve the model's capability to attend to the important spatial-temporal locations, leading to enhanced performance on crowded or dense scenes. Nonetheless, the models are recognized to consume a considerable amount of computational resources and gigantic datasets, restraining their utilization to lightweight real-time systems.

Handmade Feature-Based Systems:
The earlier used methods were typically based on hand-designed features like Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and optical flow to identify motion. While these features offered interpretability and simplicity, they were not generalizable in dynamic scenes via sensitivity to light, noise, and occlusion.

Light CNN Models:
To counter the computational complexity of models, light CNN architectures with basic Conv2D layers have been proposed in recent studies. The light models are designed to be deployed on edge devices or low-hardware-capability systems, striking a balance between speed and accuracy. They enable real-time performance with efficient anomaly detection even with a small amount of training data.

Streamlit and Real-Time Systems: For deployment and user interface, Streamlit and Flask have been utilized to develop interactive platforms for intelligent surveillance systems. The combination of machine learning models and web applications improves usability and accessibility, allowing security personnel to interact with the system in real-time and be notified in real-time. Research Gaps and Motivation: Despite the progress, issues remain in increasing accuracy, reducing false alarms, and generalizing well to diverse surveillance settings. The majority of the current state-of-the-art models are either non-real-time or require training on extensive data sets. Our solution addresses these problems by introducing an effective, lightweight CNN-based system capable of processing anomalies in image and video inputs and implemented in a real-time alert-capable web application.

III. PROBLEM DEFINITION OR EXPERIMENTAL WORK

A. Problem Definition

Conventional surveillance systems lack the capability to automatically process visual data, with the possibility of failing to capture critical events due to human error. There is a need for a smart system that can autonomously detect suspicious behavior in video streams and notify concerned authorities in real-time, enhancing security and responsiveness.

B. Experimental Work

An abnormal human activity detection system using AI is suggested in this study.
It utilizes deep learning algorithms to scan surveillance videos in real-time to identify suspicious behavior in public spaces. The deployment is CNN-based and consists of a step-by-step pipeline of dataset selection, preprocessing, model initialization, training, testing, and deployment through a web interface.

a) Dataset and Preprocessing
The Dataset for Crowd Surveillance and Anomalous Scene Segmentation (DCSASS) is chosen for the experimental deployment. DCSASS is a set of annotated video samples with both normal and abnormal human behaviours in public places and is very well suited for training a supervised learning model. During preprocessing, OpenCV is utilized to capture frames from video samples at regular time intervals, avoiding redundancy with context preservation. Frames are resized to 64×64 pixels and are converted into grayscale to decrease computational complexity. Pixel values are normalized to the range [0, 1] to avoid non-uniform input to the CNN model. All processed frames are stored as NumPy arrays to facilitate efficient loading and processing during training.

b) Model Architecture Selection and Model Training
The framework utilizes a Convolutional Neural Network (CNN) to perform frame binary classification as normal or abnormal. The structure has some convolutional and pooling layers in order to learn spatial hierarchies of features, followed by fully connected layers and a sigmoid output layer.
It is optimized by the Adam optimizer and binary crossentropy loss, and accuracy and loss functions are used to check for performance. The model is trained on mini-batches of the processed frames along with labels to learn spatial features that identify normal and abnormal activities. Early stopping is also performed to prevent overfitting. The trained model obtained is stored in the.h5 format for further use.

c) Performance Appraisal
The trained CNN model is then tested with both image and video inputs. Under video testing, real-time frame extraction and processing are performed, with each frame passing through the same preprocessing pipeline and following classification. The identified anomalies trigger an alert.
The model is 83.18% precise, with the F1-score of 0.7946 and ROC AUC score 0.9082, respectively, which have good

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

balance as well as discriminating power. The confusion matrix is:

8406 true negatives

5397 true positives

1097 false positives

1694 false negatives

These results demonstrate the validity and usability of the model for real-time anomaly detection in surveillance environments.

d) Real-World Testing and Deployment

To provide a friendly interface, the trained model is embedded into a Streamlit web application. The application consists of user authentication to allow the system to accept inputs from just authorized users. Users can load images or video files for processing once they login. The media are processed employing the trained CNN model, with the results returned as clear visual cues.

When an anomaly is detected, real-time alert notifications are triggered by the application. The application is made light and lean with low hardware requirements to ensure ease of deployment in real-time surveillance applications. This study demonstrates the useful application of CNNs for live anomaly detection from videos of crowd scenes. Subsequent research can extend this by incorporating temporal models like ConvLSTM or 3D CNNs for enhanced context handling and by incorporating the system into IoT surveillance systems for mass deployment.

IV. RESULTS AND DISCUSSION

The Smart Surveillance system was successfully implemented as a complete web application using Streamlit, integrating machine learning and user-friendly interface elements. The results of the system are highlighted through various important functions supported by screenshots in this section.

User authentication is initiated with the secure system, where users must log into the dashboard. Login and logout features ensure that only authorized users utilize the surveillance functionalities thereby maintaining data privacy and access control (see fig. 1 & 2).

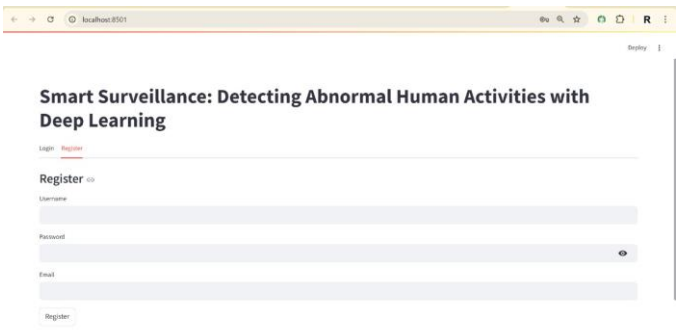

fig.1 The registration page for newcomers so as to create an account in the Smart Surveillance system

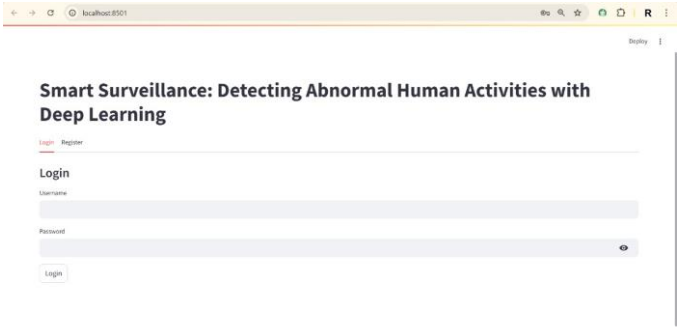

fig.2 The login interface of the Smart Surveillance system, where authentic users log in

After authentication, the user is redirected to the main interface from where users can upload either pictures or video files to be analyzed for the presence of anomalies. Once uploaded, the input is processed through the CNN model that gets trained in favorable scenarios. Normal and anomalous behavior are classified by the model, and the result is shown to the user. Screenshots in fig. 3 and 4 show sample results for image uploads where normal and anomalous actions are detected respectively; likewise, for video uploads, processing and classification were shown in fig. 5 and 6.

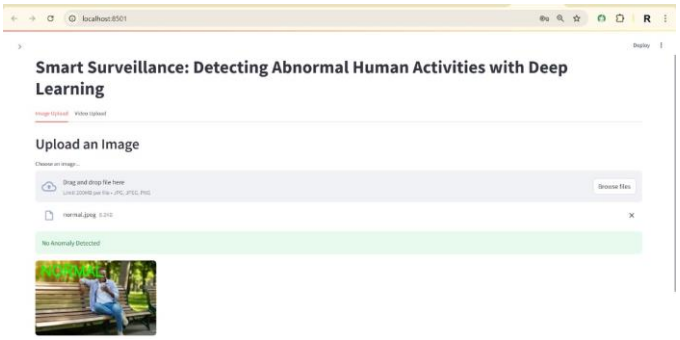

Fig.3 The image upload window showing normal activity detection.

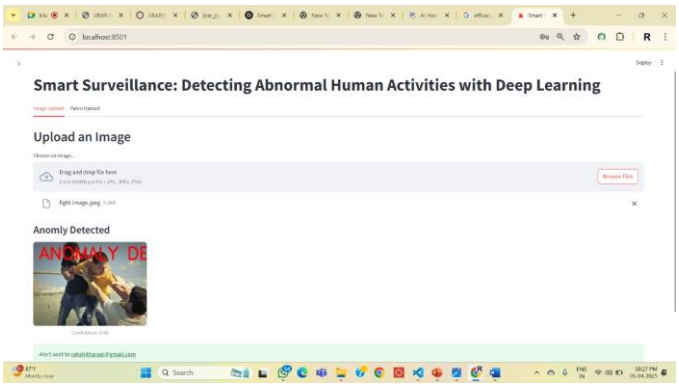

fig.4 The anomalous activity detection in the image upload window

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

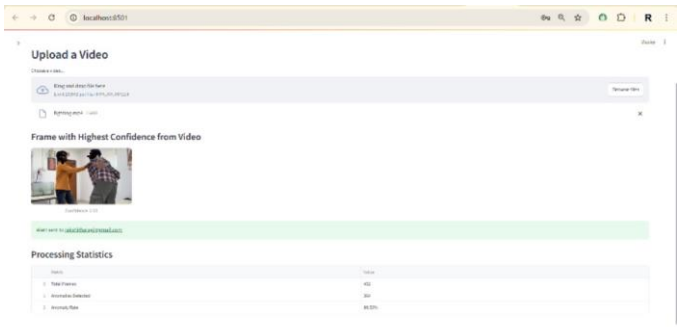

fig.5 The video upload window indicates anomalous behavior for the uploaded video

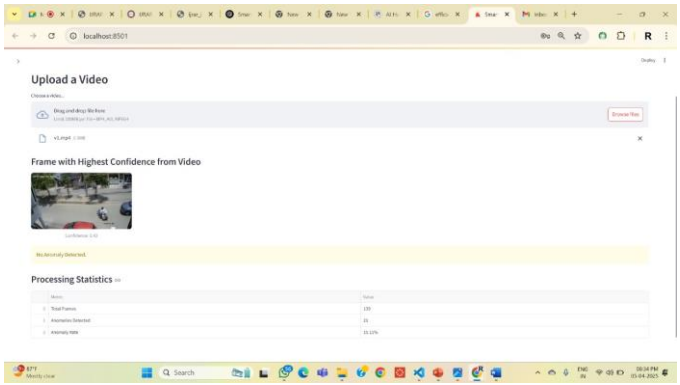

fig.6 The video upload window shows normal activity in the surveillance video

Uniquely, this system possesses a real-time email alert mechanism, whereby detection of any anomaly in the input media leads to an instant notification via email to a predefined recipient. This alert therefore acts as instant news for forthcoming security breaches or suspicious activities. In fig. 7 is a sample of the alert email showing the nature of the detected anomaly.

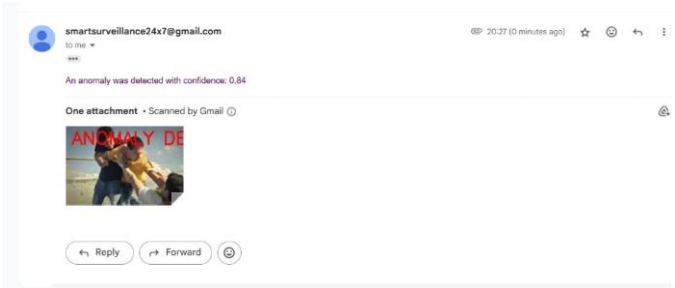

fig.7 An example of an email alert sent automatically for detected anomalies in uploaded media.

In conclusion, the results reveal good performance by the system in recognizing abnormal human behavior from both images and video data. The deep learning model has been seamlessly interfaced with an interactive web platform, which promotes user-friendliness and applicability for deployment in schools, public places, offices, or gated communities.

This demonstration of practical work through the Streamlit application thus confirmed the system not only to detect anomalies and respond quickly with alerts, showcasing the potential to being a smart, accessible, and responsive surveillance system.

V. CONCLUSION

Conclusion The smart surveillance system described in current working paper is an exciting and efficient technique for environmental control of monitoring through deep learning. It is also capable of detecting anomalies in images and video streams with high accuracy using convolutional neural network-based technology. The addition of enabling a Streamlit-based web interface concludes with the functions of registering and logging the users, uploading images and videos, as well as generating real-time alerts, enhances usability and accessibility further for application purposes by end users.

It is a pre-trained model tested on video data from the DCSASS dataset, demonstrated that it is already capable of identifying abnormal behaviors and sounding alarms in real time. This model has been tested, and it has shown an accuracy score of 83.18%, which means the system can be trusted in real-world anomaly detection applications. It also helped in sending alerts to users through mail for prompt action and a leap toward proactive surveillance management. The whole system is scalable and adaptable to most applications, for instance, public places, schools, offices, and restricted spaces.

The whole project thus combines deep learning with user-friendly frontend functions for an efficient and responsive surveillance solution, emphasizing automation as well as accuracy and efficiency in anomaly detection.

VI. ACKNOWLEDGMENT

The authors thank Geetanjali College of Engineering and Technology for the required infrastructure and resources and their unending support from the beginning to the end of this research.

The authors owe Dr. K. Srinivasa Reddy, Associate Professor, Department of Computer Science and Engineering, Geetanjali College of Engineering and Technology, for his valuable guidance, encouragement, and untiring support during the entire development of this project.

To our mentors and the staff in our department: We extend our sincere gratitude for their continuous help, constructive feedback, and encouragement through the entire course of our study.

The authors recognize the use of several open-source datasets, computational tools, and research publications, without which the development, experimentation, and performance evaluation of the model would not have been possible.

Finally, our profound gratitude goes to all those who contributed directly or indirectly to the successful completion of this work.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

REFERENCES

[1] YOLO-based anomaly activity detection system for human behavior analysis and crime mitigation. Springer. https://link.springer.com/article/10.1007/s11760-024-03164-7

[2] Chong, Y. S., & Tay, Y. H. Abnormal Event Detection in Videos using Spatiotemporal Autoencoder. Semantic Scholar. https://www.semanticscholar.org/paper/Abnormal-Event-Detection-in-Videos-using-Chong-Tay/527cc8cd2af06a9ac2e5cded806bab5c3faad9cf

[3] Suspicious Human Activity Recognition from Surveillance Videos Using Deep Learning. IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10620187

[4] Suspicious Activity Recognition for Monitoring Cheating in Exams. Springer. https://link.springer.com/article/10.1007/s43538-022-00069-2

[5] Real-Time Human Action Recognition Using Deep Learning. ResearchGate. https://www.researchgate.net/publication/366523828_Real-Time_Human_Action_Recognition_Using_Deep_Learning

[6] Transfer Learning Model for Anomalous Event Recognition in Big Video Data. Nature. https://www.nature.com/articles/s41598-024-78414-2

[7] Design of an Integrated Model with Temporal Graph Attention and Transformer-Augmented RNNs for Enhanced Anomaly Detection. Nature. https://www.nature.com/articles/s41598-025-85822-5

[8] Identification and Detection of Abnormal Activity in ATMs Using Deep Learning. IJRTI. https://ijrti.org/papers/IJRTI2302037.pdf

[9] A Framework for Anomaly Classification Using Deep Transfer Learning Approach. International Journal of Innovative Technology and Exploring Engineering (IJITEE). https://www.iieta.org/journals/ria/paper/10.18280/ria.350309

[10] Visually Explaining 3D-CNN Predictions for Video Classification with an Adaptive Occlusion Sensitivity Analysis. Semantic Scholar. https://www.semanticscholar.org/paper/Visually-explaining-3D-CNN-predictions-for-video-an-Uchiyama-Sogi/164cc3bb0621a2c3e9f42410a76cc540c88cdfd9

.