International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

Intelligent Real-Time Suspicious Activity Detection Using Convolutional Neural Networks

MUDUNURI CHANDINI

PG Scholar. Department M.Sc(CS), DNR College, Bhimavaram, Andhra Pradesh

A.Durga Devi

Lecturer in M.Sc(CS), DNR College, Bhimavaram, Andhra Pradesh

ABSTRACT

Ensuring public safety has become an increasingly important challenge as urban environments expand and security threats continue to rise. Traditional surveillance systems rely heavily on human operators to continuously monitor video feeds, which is inefficient, error-prone, and impractical for large-scale deployments. To address these limitations, this project proposes an intelligent real-time suspicious activity detection system based on Convolutional Neural Networks (CNNs). CNNs are highly effective in automatically extracting and learning spatial features from images and video frames, making them suitable for activity recognition tasks. The proposed system employs a deep-learning-based architecture capable of identifying suspicious movements, gestures, and behaviors without manual intervention.

The system processes incoming video frames from surveillance cameras, performs preprocessing such as normalization and resizing, and then feeds them into a trained CNN model. The CNN extracts high-level features and classifies activities into normal or suspicious categories. By leveraging spatial feature learning and motion pattern recognition, the system can detect activities such as fighting, theft attempts, aggressive gestures, or unauthorized intrusions. The model is trained on publicly available human activity datasets enriched with various classes of suspicious and non-suspicious movements, ensuring a wide range of behavior recognition.

To improve accuracy, data augmentation techniques, dropout layers, and fine-tuning strategies are integrated into the CNN pipeline. The system also uses real-time inference, allowing for immediate alert generation. Once a suspicious activity is detected, the system triggers notifications or alarms, helping security personnel respond promptly.

Overall, the proposed suspicious activity detection system enhances traditional surveillance setups by introducing automated, intelligent behavior-monitoring

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

capabilities. It reduces human workload, minimizes surveillance fatigue, increases detection accuracy, and provides proactive security support. The development and implementation of this CNN-based solution demonstrate the power of deep learning in enhancing modern security systems and highlight the potential of AI-driven technologies for public safety applications.

Keywords: Suspicious Activity Detection, Convolutional Neural Network, Deep Learning, Surveillance System, Video Analysis, Security Monitoring, Anomaly Detection, Feature Extraction

I. INTRODUCTION

Surveillance systems play a crucial role in maintaining safety and security in public spaces, private institutions, and residential areas. With the increasing frequency of criminal activities and security breaches, reliance on manual monitoring has proven insufficient. Traditional surveillance systems rely on security personnel to observe live feeds continuously, but humans are prone to fatigue, oversight, and slow reaction time. As the scale of surveillance expands, it becomes nearly impossible for operators to manually track every camera feed simultaneously. These challenges have led to the need for automated, intelligent solutions that can assist or replace human monitoring.

Recent progress in artificial intelligence, particularly deep learning, has made it possible to analyze images and videos with near-human or even superior accuracy. Convolutional Neural Networks (CNNs) have emerged as a powerful approach for tasks involving pattern recognition, object detection, and activity classification. Their ability to automatically learn hierarchical spatial features has made them widely used in computer vision and security applications.

This project focuses on developing a CNN-based suspicious activity detection system designed to identify abnormal or harmful human behaviors within surveillance videos. Suspicious activities such as fighting, harassment, intrusion, theft attempts, and violent gestures often share specific motion and pose patterns, making them suitable targets for CNN-based classification. By analyzing frame sequences or extracted keyframes, the model identifies deviations from normal patterns and flags them as suspicious.

The system aims to reduce dependency on manual monitoring, improve response time, and enhance public safety. It incorporates key components such as video preprocessing, CNN model training, data augmentation, classification, and alert generation. The model is trained using labeled video datasets that capture various types of suspicious and normal activities, ensuring robust performance in real-world environments.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

In addition to accuracy, the system is designed for real-time deployment, which is essential for surveillance applications. Rapid inference ensures timely alerts for security personnel. This project highlights the potential of deep-learning-based surveillance solutions in creating safer environments and showcases how AI can assist law enforcement, institutions, and individuals in improving situational awareness.

II. LITERATURE SURVEY (WITH EXISTING METHODS)

Suspicious activity detection has been an active research area due to its significance in crime prevention and human behavior monitoring. Traditional approaches relied mainly on manual rule-based systems or handcrafted feature extraction, which were limited in accuracy and adaptability.

Early Methods:
Early surveillance systems employed optical flow, background subtraction, Histogram of Oriented Gradients (HOG), and handcrafted spatiotemporal features to detect abnormal motion. While computationally less demanding, these methods struggled with complex scenarios, lighting changes, camera angles, and dynamic backgrounds. Their reliance on manually designed features made them brittle when exposed to real-world variability.

Machine Learning Advancements:
Later systems introduced Support Vector Machines (SVM), Random Forests, and Hidden Markov Models (HMM) for activity classification. These models performed better than rule-based systems but still depended heavily on feature engineering. Their scalability was limited because they couldn't automatically extract deep features from raw videos.

Rise of Deep Learning:
The introduction of Convolutional Neural Networks (CNNs) revolutionized visual recognition due to their ability to automatically extract meaningful hierarchical features. Researchers began integrating CNNs with datasets like UCF Crime, HMDB51, and UCF101 to classify human activities. CNN models such as VGG16, ResNet, and MobileNet have been widely used for suspicious behavior detection.

Existing Methods:
Various studies propose hybrid architectures combining CNN with LSTM networks for spatiotemporal feature extraction. Methods such as 3D-CNN, C3D, and Two-Stream CNNs have shown significant improvements in recognizing complex activities. Despite their success, these require large datasets and heavy computation.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

Other existing systems include anomaly detection models using autoencoders, GAN-based irregular motion detection, and graph neural networks for human pose–based suspicious activity detection. These methods yield promising results but require advanced hardware and large annotated datasets.

Gap Analysis:
Most existing systems face challenges such as slow inference, high computational cost, limited adaptability, and poor performance in low-resolution or crowded scenes. Many are not suitable for real-time deployment due to processing delays.

Conclusion:
The need for an efficient, real-time, CNN-based suspicious activity detection system remains essential. This project adopts optimized CNN models to overcome deficiencies in accuracy, speed, and adaptability found in prior approaches.

III. EXISTING SYSTEM

The existing surveillance systems used in most public and private security infrastructures primarily rely on manual observation. Security personnel monitor live camera feeds and identify suspicious activities based on experience and situational awareness. This manual monitoring suffers from several drawbacks, including human fatigue, limited attention span, slow recognition of abnormal patterns, and difficulty handling multiple camera feeds simultaneously.

Traditional automated systems rely on simple motion detection or predefined rule-based algorithms such as background subtraction or optical flow. These systems trigger alarms based solely on abrupt movement or changes in pixel intensity. They cannot differentiate between normal and suspicious activities, resulting in frequent false alarms. For example, movement caused by animals, reflections, or weather changes can trigger alerts. These approaches also lack the ability to understand human behavior or recognize complex interactions.

Existing machine-learning-based systems also suffer limitations. They depend on handcrafted features like HOG or SIFT, which perform poorly in dynamic environments. They lack robustness to lighting changes, camera angle variations, and crowded scenes. Additionally, these systems struggle to scale due to limited generalization capability.

In summary, current systems are either manual, inefficient, or insufficiently intelligent to detect suspicious human activities accurately. There is a clear need for a deep-learning-

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

based solution that can automatically learn behavior patterns, classify activities, reduce human workload, and provide real-time alerts with higher accuracy.

IV. PROPOSED METHOD

The proposed system introduces an intelligent, automated Suspicious Activity Detection framework that uses Convolutional Neural Networks (CNNs) to identify abnormal or harmful human behavior in surveillance videos. Unlike conventional security systems that rely solely on manual monitoring or rule-based detection, the proposed system learns activity patterns directly from video data and identifies suspicious behaviors based on deep feature extraction. The CNN model is trained on labeled datasets containing both normal and suspicious activity samples, allowing it to learn visual cues such as gestures, movement intensity, aggression, and spatial anomalies.

The system consists of key modules including video acquisition, preprocessing, frame extraction, feature learning using CNN, classification, and alert triggering. Input surveillance footage is processed in real-time, and frames are resized, normalized, and fed to a trained CNN model. The model classifies each frame or sequence as either normal or suspicious. If suspicious activity is detected, the system immediately generates an alert that can be sent to security personnel for quick response.

To improve accuracy, techniques such as data augmentation, dropout regularization, and transfer learning are applied. The system is designed to work in real-time, ensuring rapid detection in environments like malls, airports, banks, offices, and public gatherings. The proposed framework is scalable, adaptable to various lighting conditions, and capable of handling variations in camera angles and crowded scenes. Overall, the system enhances traditional surveillance by providing automated, intelligent, and efficient suspicious behavior monitoring.

V. IMPLEMENTATION

The implementation of the Suspicious Activity Detection System is carried out using deep learning and computer vision methodologies. The project follows a modular approach and involves several key steps: data preprocessing, CNN model design, training, testing, and deployment.

1. Data Collection and Preprocessing

Dataset preparation is the foundational step. Public datasets such as UCF Crime, HMDB51, UR Fall Detection, or custom surveillance datasets are used to obtain samples of suspicious and non-suspicious activities. Frames are extracted from video sequences

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

using OpenCV. These frames are resized, normalized, and augmented through flipping, rotation, zooming, and brightness adjustments to improve generalization.

2. CNN Model Development

The CNN model is built using frameworks such as TensorFlow or PyTorch. The architecture consists of multiple convolutional layers for feature extraction, followed by ReLU activation, max-pooling layers for dimensionality reduction, and fully connected layers for decision making. Batch normalization and dropout layers are added to prevent overfitting.

Alternatively, transfer learning models such as VGG16, ResNet50, or MobileNet can be used to reduce training time and improve performance on limited datasets.

3. Model Training

The dataset is split into training and testing sets. The CNN model is trained using supervised learning, with categorical cross-entropy as the loss function and Adam optimizer for efficient gradient updates. The training continues for several epochs until the model achieves convergence. Accuracy, precision, recall, and F1-score are used as performance metrics.

Techniques such as early stopping, learning rate scheduling, and checkpointing help stabilize training and avoid performance degradation.

4. Activity Classification

During inference, incoming video streams are processed frame-by-frame or in short sequences. The trained CNN predicts whether the activity belongs to a suspicious or normal class.

5. Alert Triggering and User Interface

If suspicious activity is detected, the system generates alerts using email, SMS, or a dashboard interface. Real-time predictions are displayed using a simple GUI or integrated with an existing CCTV monitoring system.

6. Deployment

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

The final model can be deployed on a local server, cloud platform, or edge devices such as NVIDIA Jetson for on-site processing. Optimizations like quantization or ONNX conversion help improve speed

VII. SYSTEM DESIGN

1. Frame Extraction Algorithm

This algorithm extracts frames from video streams for processing.

Input: Video stream
Process:
Read video using OpenCV
Extract frames at fixed intervals
Resize and normalize frames
Output: Preprocessed frames

This ensures uniform frame input for CNN processing.

2. CNN Feature Extraction Algorithm

CNN automatically learns patterns through hierarchical layers.

Convolutional layers extract edges, textures, and movement cues
Max pooling reduces dimensionality
Fully connected layers perform classification

Algorithm Steps:

Convolve input with learned filters
Apply ReLU activation
Down-sample using max pooling
Flatten feature maps
Classify using softmax
3. Suspicious Activity Classification Algorithm

This algorithm assigns a label (suspicious or normal).

Input: Extracted CNN features
Logic:

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

If probability(suspicious) > threshold → Trigger alert
Else → Normal

Pseudo logic:

```
if CNN_output["suspicious"] >= 0.7:
   flag_activity = "Suspicious"
else:
   flag_activity = "Normal"
```

4. Real-Time Detection Algorithm

This ensures fast inference:

Capture live video frames
Preprocess frame
Run CNN prediction
If suspicious → send alert
Continue loop

Together, these algorithms ensure accurate, real-time suspicious behavior detection.

SYSTEM DESIGN IMAGES

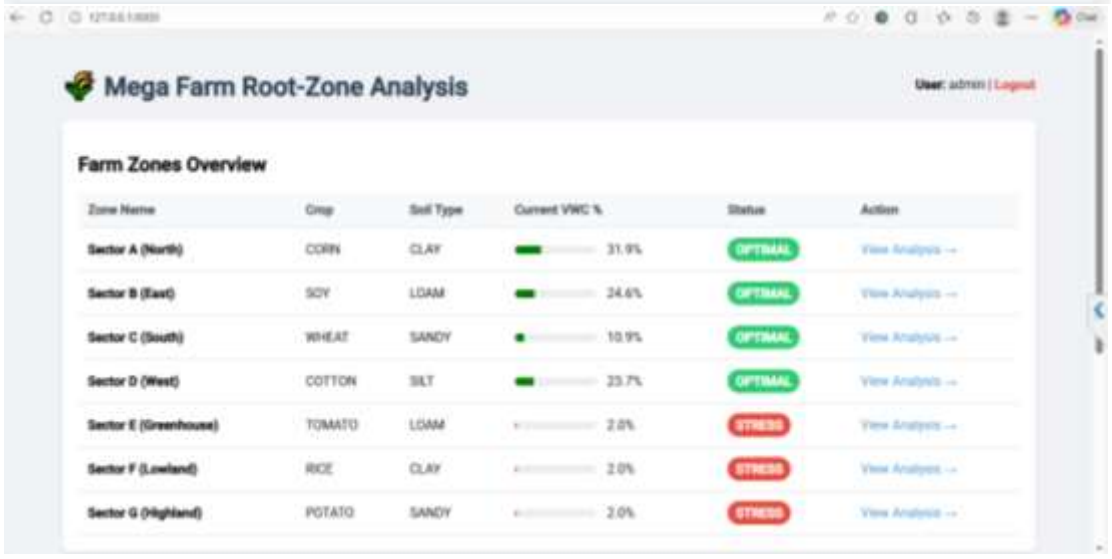
International Journal of

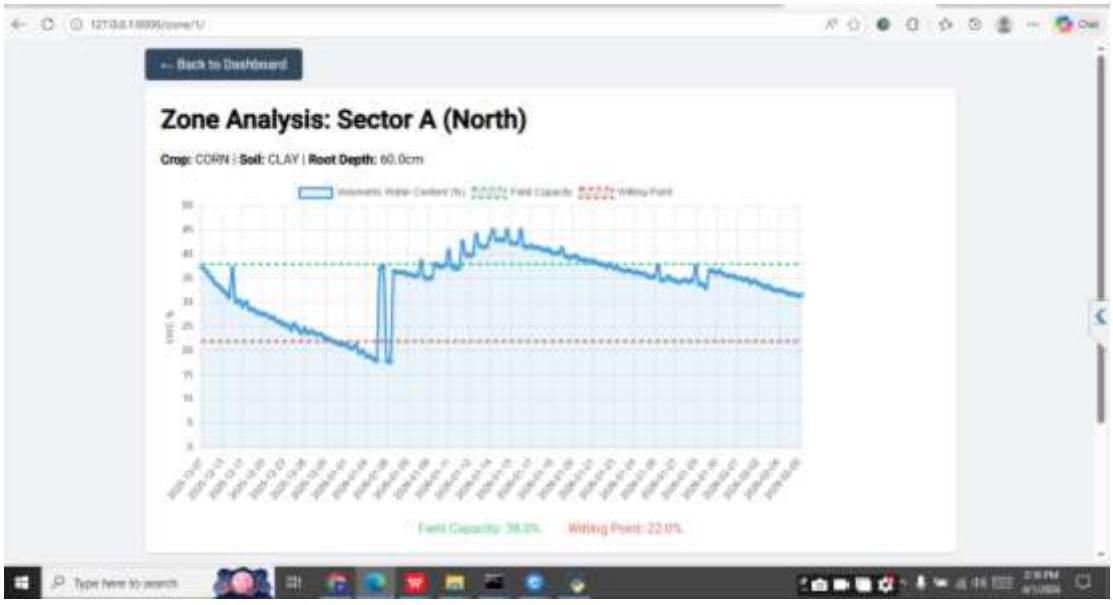
DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper







International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

VIII. CONCLUSION

The system uses a multilayered architecture combining video processing, deep learning inference, and alert mechanisms. It follows a modular design for scalability, reliability, and ease of deployment.

1. Architecture Overview

The system is composed of the following layers:

Input Layer: Receives video from CCTV or surveillance cameras
Processing Layer: Performs frame extraction and preprocessing
CNN-Based Detection Layer: Runs deep learning inference
Decision Layer: Analyzes results and determines if activity is suspicious
Alert Layer: Sends notifications or alarms
2. Module Breakdown
a. Video Input Module

Accepts RTSP streams, stored videos, or live camera feeds. Uses OpenCV for capturing and processing frames.

b. Preprocessing Module

Performs:

Frame resizing
Normalization
Noise reduction
Data augmentation (for training phase)

This ensures consistent input to the CNN.

c. CNN Model Module

This module is responsible for:

Feature extraction through convolutional layers
Temporal pattern recognition (if using 3D-CNN or LSTM-CNN)

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

Final classification into suspicious or normal categories

The CNN learns from multiple classes of video data, recognizing subtle behavior cues.

d. Classification and Decision Module

Uses softmax outputs to determine prediction scores. If suspicious behavior exceeds a specified threshold, the system flags it and forwards to the alert module.

e. Alert and Monitoring Module

Generates real-time alerts using:

Email
SMS
Web dashboard
Audio alarms

This module ensures human security personnel are immediately notified.

f. User Interface Module

A dashboard allows users to:

View live detection
Check logs
Review detected events
Configure sensitivity settings
3. Data Flow Design
Camera captures real-time video.
Frames are extracted and preprocessed.
CNN model processes frames.
Predictions are generated.
Suspicious events trigger alerts.
Logs are stored for investigation.
4. System Requirements
GPU-enabled hardware (optional but recommended)
Python, OpenCV, TensorFlow/PyTorch
High-speed internet for remote monitoring

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

CCTV or IP camera integration

REFERENCES

The Suspicious Activity Detection System using CNN provides an advanced solution to the limitations of traditional surveillance systems. By integrating deep learning techniques, the system automates the process of identifying abnormal or dangerous human behaviors, reducing the dependency on manual monitoring. CNNs demonstrate exceptional ability in extracting spatial and contextual features from video frames, making them ideal for detecting complex activities such as violence, intrusion, theft attempts, or aggressive motions.

The proposed system ensures higher accuracy, faster response time, and reliable performance in real-world conditions. Real-time frame processing and intelligent alert generation significantly improve situational awareness for security personnel. Compared to traditional rule-based or motion-detection systems, the CNN-based approach learns directly from data, adapts to diverse environments, and reduces false alarm rates.

By deploying this system across public and private domains such as malls, airports, offices, banks, and residential areas, security operations can be significantly strengthened. The system enhances overall safety by ensuring that suspicious activities are identified promptly.

Future enhancements may include integration with multi-camera coordination, LSTM-based temporal modeling, facial recognition, crowd analysis, and edge-device optimization. These improvements can make the system even more efficient and scalable. Overall, the CNN-based suspicious activity detection framework is a powerful and innovative solution for modern-day surveillance challenges.

Reference:

[1] W. Sultani, C. Chen, and M. Shah, “Real-world anomaly detection in surveillance videos,” *CVPR*, 2018.
[2] 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*, 2017.
[3] K. Simonyan and A. Zisserman, “Two-stream convolutional networks for action recognition,” *NeurIPS*, 2014.
[4] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*,

1997.
[5] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification with deep CNNs," *NeurIPS*, 2012.
[6] K. He et al., "Deep residual learning for image recognition," *CVPR*, 2016.
[7] J. Carreira and A. Zisserman, "Quo Vadis, Action Recognition?" *CVPR*, 2017.
[8] L. Wang et al., "Temporal segment networks for action recognition," *ECCV*, 2016.
[9] H. Kwon et al., "Violence detection using deep learning," *Applied Sciences*, 2020.
[10] Z. Li, H. Zhang, "Suspicious behavior detection in surveillance videos," *IEEE Trans. Multimedia*, 2019.
[11] Y. LeCun et al., "Gradient-based learning applied to document recognition," *Proc. IEEE*, 1998.
[12] S. Ren et al., "Faster R-CNN: Real-time object detection," *NeurIPS*, 2015.
[13] X. Wang, "Intelligent video surveillance: A review," *IEEE Trans. Multimedia*, 2013.
[14] S. Muhammad et al., "Deep anomaly detection in videos," *IEEE Access*, 2019.
[15] Z. Chen et al., "Deep CNN models for human activity recognition," *ICME*, 2018.