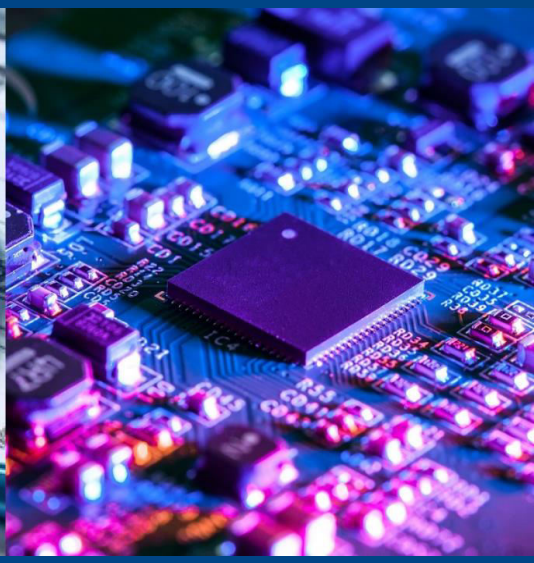
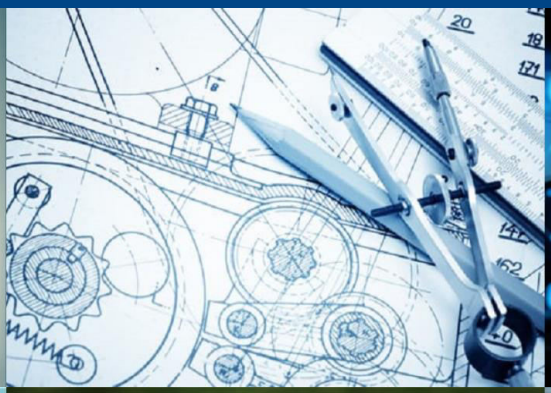
ISSN: 2582-7219

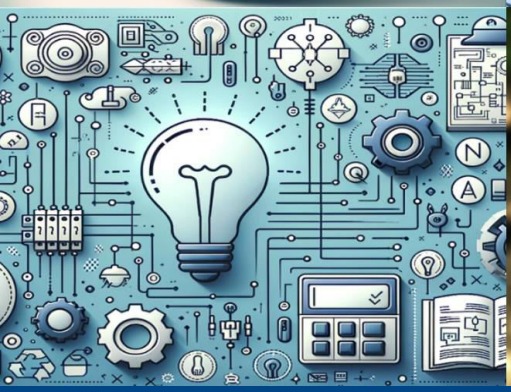




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 9, Issue 4, April 2026**

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI-Powered Drone Surveillance System for Student Activity Monitoring

D. Vamsi [1], B. Praneeth Koundinya[2], G. Harika[3], M. Mahendra[4]

UG Student, Department of ECE, R.V.R & J.C College of Engineering, Chowdavaram, India[1]

UG Student, Department of ECE, R.V.R & J.C College of Engineering, Chowdavaram, India[2]

UG Student, Department of ECE, R.V.R & J.C College of Engineering, Chowdavaram, India[3]

UG Student, Department of ECE, R.V.R & J.C College of Engineering, Chowdavaram, India[4]

ABSTRACT: Safety and security in public areas, such as a college campus, has become a significant issue with increasing security concerns. This paper proposes an AI-based drone surveillance system which is capable of identifying humans and recognizing their actions in real-time. The system uses a drone to capture live video, which is then transmitted and processed through a ground station and cloud-based machine learning algorithms. Face recognition is done through classification using encoding-based techniques, while human activity recognition is done through a hybrid CNN and LSTM model. The system can detect activities such as boxing, running, walking, and hand movements, and sends real-time alerts through the monitoring system. The system is cost-effective and eliminates the need for human surveillance.

KEYWORDS: Drone Surveillance, Face Recognition, Action Recognition, CNN, LSTM, Real-Time Monitoring.

I. INTRODUCTION

With an increased number of people and an expansion in areas like college campuses, safety and security have become a major concern. Current surveillance systems rely on fixed camera systems and human monitoring, which can be quite inefficient and time-consuming. These systems lack intelligent capabilities to understand and respond to situations, leading to delayed actions in critical situations.

Recently, Artificial Intelligence (AI) and computer vision have emerged as powerful techniques for building intelligent surveillance systems. These systems can process images and videos in real-time to identify objects, people, and human activities. However, existing intelligent surveillance systems can either perform face recognition or human activity detection. These systems often rely on fixed camera systems with limited viewing areas.

The integration of Unmanned Aerial Vehicle (UAV) technology, commonly referred to as Drones, has brought a new dimension to surveillance systems. These systems can cover large areas and offer dynamic surveillance capabilities, unlike fixed camera systems. These systems can be further enhanced with Artificial Intelligence techniques to offer efficient intelligent surveillance systems.

This paper proposes an surveillance system based on Artificial Intelligence techniques and Drones. The proposed system integrates face recognition and human activity detection for efficient decision-making capabilities. The proposed system can be quite efficient in offering intelligent surveillance capabilities without relying on human supervision. The proposed system utilizes a drone for capturing video streams, which are further processed at a ground station and in the cloud using machine learning techniques for face and human activity detection. The proposed system can be quite efficient in offering intelligent surveillance capabilities without relying on human supervision.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE REVIEW

Here we propose an AI-based drone surveillance system, which is capable of real-time human identification and activity recognition. The proposed system would utilize computer vision, deep learning, and drones, making it an intelligent surveillance system. Various research works have been carried out on AI-based surveillance, focusing more on improving accuracy.

Face recognition systems have been widely utilized in various applications, including human identification. Face recognition systems use encoding techniques to identify individuals.

Some of the advantages of face recognition systems are:

- Highly accurate in identifying human faces
- Fast, as encoding techniques are employed
- Can be utilized in real-time applications

Some of the disadvantages of face recognition systems are:

- Not capable of recognizing human activities or behavior
- Accuracy may be reduced depending on light conditions

Human activity recognition systems have been proposed using deep learning techniques, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM). These systems utilize video analysis to identify human activities, including walking, running, or fighting.

The advantages of the activity recognition system are as follows:

- Multiple human actions can be classified
- Efficient use of temporal and spatial features
- Suitable for behavior monitoring

However, the system has some challenges as follows:

- Does not give the identity of the individuals
- Requires a large dataset
- Computational intensive

Surveillance systems using drones are important due to the ability of the system to offer a dynamic and wide area of surveillance. This is because the system can move freely and capture video images in real time from different angles.

However, the advantages of the drone-based surveillance system are as follows:

- Wide area of surveillance
- Can capture video images in real time
- Does not depend on infrastructure

However, the system has some challenges as follows:

- Does not integrate with intelligent analysis
- Depends on other systems

With the recent advancements in AI-based surveillance systems, the integration of multiple technologies can offer a more efficient system. This can be achieved by integrating face recognition and activity recognition with drone-based surveillance systems to offer a more efficient system.The proposed system integrates multiple technologies to offer a more efficient system in the real world. This system can efficiently identify humans and monitor behavior in real time. This can improve the efficiency of the system and the decision-making ability of the system. This can reduce the need for supervision by humans.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. PROPOSED METHOD

Problem Statement:
The traditional surveillance system has been seen to be less efficient, especially due to its limitation to the coverage area and its dependency on human intervention. The traditional system lacks intelligence and hence cannot perform an analysis on the current situations, especially when identifying individuals and their actions.

In addition, most of the existing intelligent surveillance systems using AI are found to be focused only on the recognition of faces or the actions performed, and this is found to be incomplete. The limitation is the major factor that reduces the efficiency of the system in handling the real-time situations.

Objectives
The objectives of the proposed system are as follows:

- Design and develop an intelligent system using drones for video surveillance
- Design and develop an intelligent system using face recognition and encoding
- Design and develop an intelligent system using action recognition and CNN/LSTM
- Real-time alerts at the monitoring station
- Cost-effective and efficient surveillance system

System Architecture Overview
The proposed system includes three main components:

1. Drone Unit:
   It records real-time video feed and covers dynamic areas.

2. Base Station:
   It receives video data from the drone and acts as an interface for monitoring.

3. Cloud-Based Processing:
   It executes machine learning models for face and action detection.

Working Methodology
The working methodology of the proposed system is explained below:

1. The drone records video feed in real-time from the surveillance area.
2. It sends video data to the base station.
3. It processes video frames and sends them for face and action detection.
4. Face detection recognizes individuals based on stored encodings.
5. Action detection recognizes human actions using a CNN-LSTM model.
6. It generates an alert message based on detection outcomes.
7. It continuously executes this process for real-time video monitoring.

Architecture Diagram

1. **System Initialization:** The system is configured by powering the drone and linking it to the ground station. The required modules are now ready for processing.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

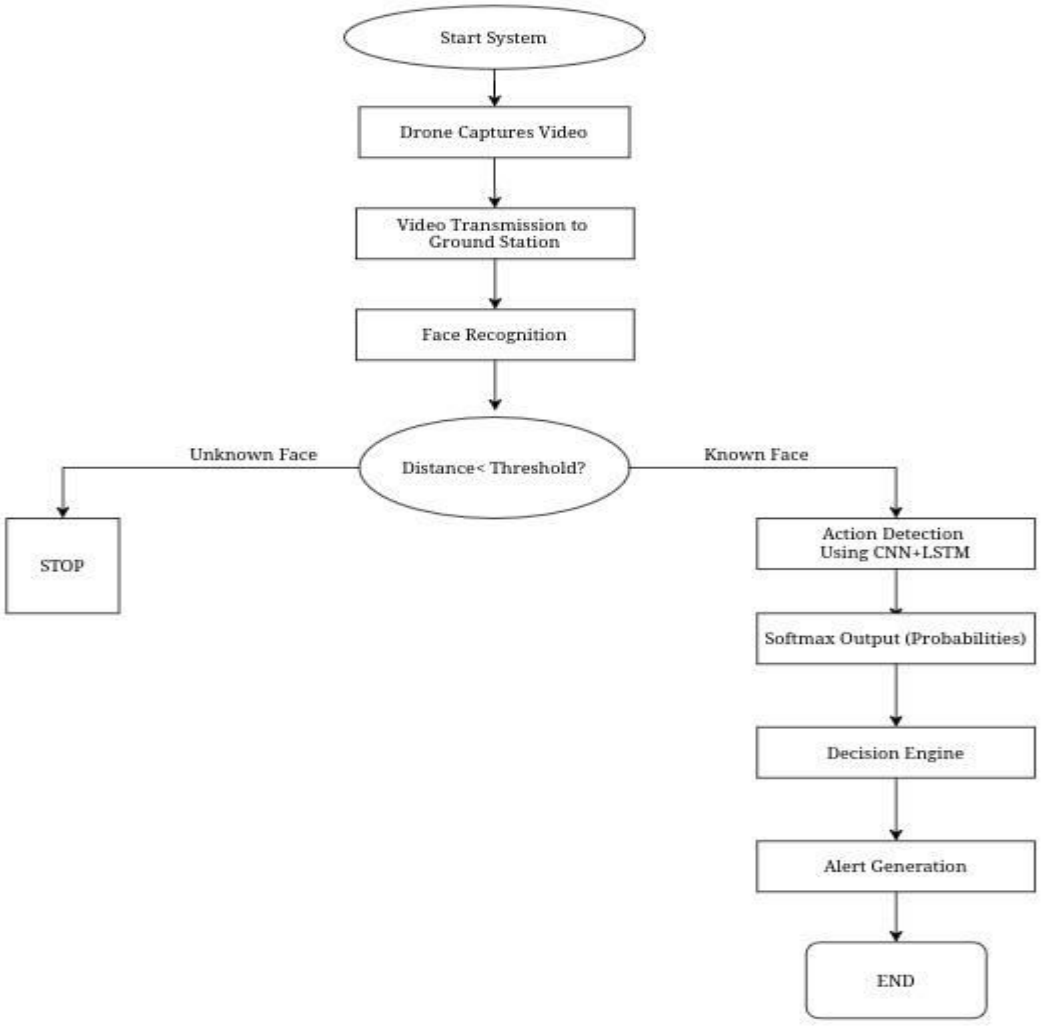


Fig.1. Architecture diagram

2. **Video Capture:** The drone captures continuous live video feed from the surveillance region using the camera installed on it.
3. **Video Transmission:** The video feed is now transmitted to the ground station (laptop) via ESP32 S3 module.
4. **Frame Processing:** The system converts the received video feed into individual frames then the frames are resized and normalized for processing.
5. **Face Recognition:** The faces are recognized within individual frames and then compared for identity recognition.
6. **Decision Making :** The recognized face is then checked for matching identity based on a threshold value. If there is no match, then continuous processing occurs.
7. **Action Recognition:** If a match is found for a recognized face, then human activity recognition occurs based on individual frames using a CNN-LSTM model.
8. **Activity Classification:** The activity is then recognized based on individual probabilities for different actions. The action for which the probability is the highest is selected as the recognized activity.
9. **Decision Engine:** The recognized identity and activity are then used to make a decision based on individual information.
10. **Alert Generation:** A real-time alert message is displayed on the display panel for the activity detected within the region.
11. **Continuous Monitoring:** The above steps are repeated for continuous processing.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. IMPLEMENTATION OF THE SYSTEM

Requirements for Software and Hardware
Software Requirements:

Python and Google Colab:
The proposed system has been implemented using Python as a programming language. The system runs on Google Colab. Google Colab has enough computational capabilities to train and execute the machine learning models. Google Colab also facilitates easy handling of datasets, execution of models, and real-time inference.

OpenCV:
The system uses OpenCV for video processing. OpenCV facilitates capturing images from videos, resizing images, etc. OpenCV ensures that images are properly processed for execution through the machine learning models.

Face Recognition Library:
The face recognition system is performed using a library called *face_recognition*, which uses a pre-trained deep learning model. The system uses a technique called Histogram of Oriented Gradients (HOG) for detecting faces. The system extracts features from faces as 128D vectors. The system stores these vectors for execution. During execution, it uses Euclidean distance to compare vectors for identifying faces.

TensorFlow and Keras:
The system uses TensorFlow and Keras for implementing the action recognition system. The system uses a hybrid model that integrates CNN with LSTM for recognizing human activities. CNN extracts spatial features from images, while LSTM extracts temporal features from a sequence of images.

Face Recognition Implementation:
The dataset for known people is created using multiple images. The images are processed to get facial encodings, which are stored for future use. During the execution phase, the faces in each video frame are identified, and encodings are obtained. The encodings obtained from the video frames are compared with the stored encodings using the distance method. If the distance is less than a threshold, the individual is identified; else, the face is marked as unknown.

Action Recognition Implementation:
The action recognition model is implemented using a sequence-based deep learning approach. The video data is divided into frames. Each frame is resized and normalized. A specified number of frames are used for the model.
The CNN layers identify the spatial features for each frame. The LSTM layer recognizes the sequence of frames to identify the action. The output is acquired using a softmax layer. Output of Action Recognition is the probabilities for each action: boxing, running, walking, hand movement, etc. The action with the maximum probability is selected as the action being performed.

System Integration:
The face recognition as well as action recognition modules are integrated with a unified system. The video feed from the drone is processed at the ground station, where face recognition occurs first. If a known person is recognized, then action recognition follows. According to the recognized action, a real-time warning message pops up on the monitoring screen.

Hardware Requirements

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. ESP32-S3 Module:

The primary hardware component in this system is ESP32 S3 Module. It is used for receiving and transmitting video data from the camera module to the ground station. The module is also used for Wi-Fi connection, which is necessary for real-time communication with the drone.

1. **Drone Platform:**
   The drone is used as a mobile surveillance component in this system. The drone is used for carrying the ESP32-S3 module with the camera module for real-time video transmission from different angles. The drone is useful for covering more area with respect to static systems.
2. **In-built Camera Module:**
   The camera module is used for receiving live video feed from different sources and is used for providing inputs to the machine learning models for action and face detection.
3. **Ground Station (Laptop):**
   The laptop is used as a receiving component in this system, which is connected to the drone using Wi-Fi for receiving live video feed from the drone and executing face detection and action detection models.

V. RESULT

To test the efficacy of the proposed drone surveillance system based on artificial intelligence (AI), several tests were carried out with video footage. The AI-based surveillance system showed positive results in human recognition and classification of human activities.

In order to test the performance of the face recognition module of the proposed drone surveillance system, several tests were carried out with faces of multiple individuals. The system successfully detected and identified people by encoding faces. The output result of the test is shown in Fig. 3, in which multiple faces have been detected successfully by the system.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

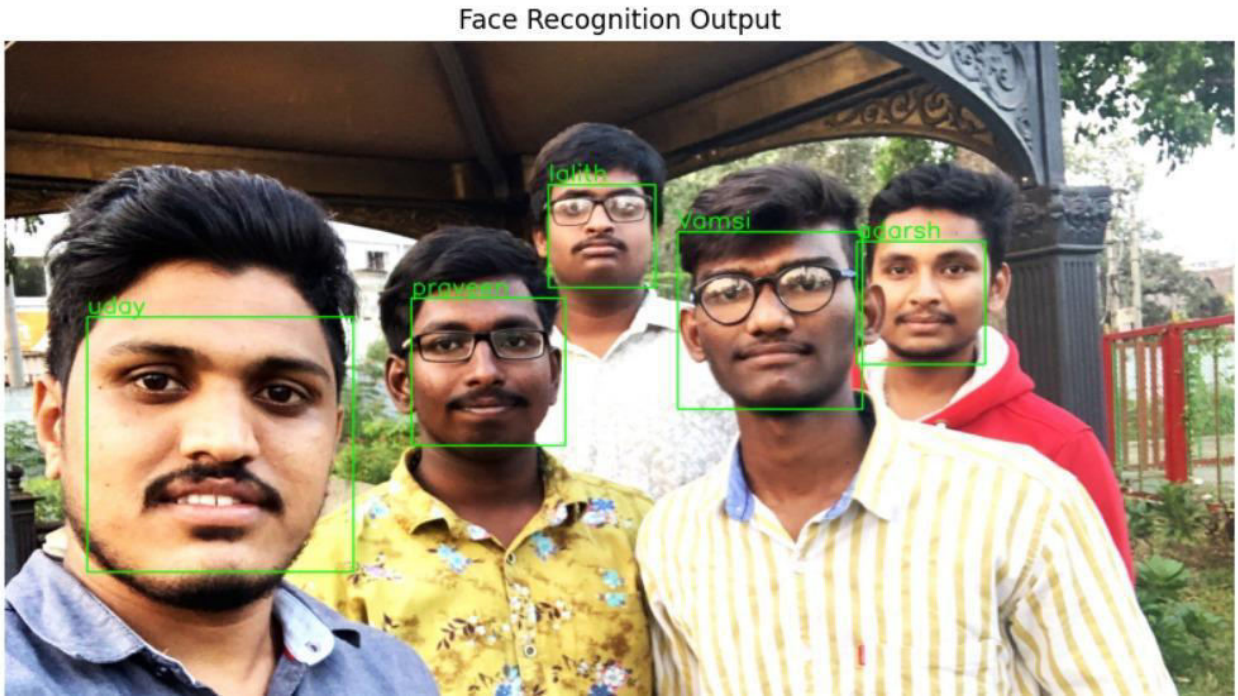


Fig. 3: Output Result from the Face Recognition Module (Multiple Individuals)

Additionally, the face recognition module was also tested with input from a single person. The system was able to recognize the single person with high efficiency, as illustrated in Fig. 4.

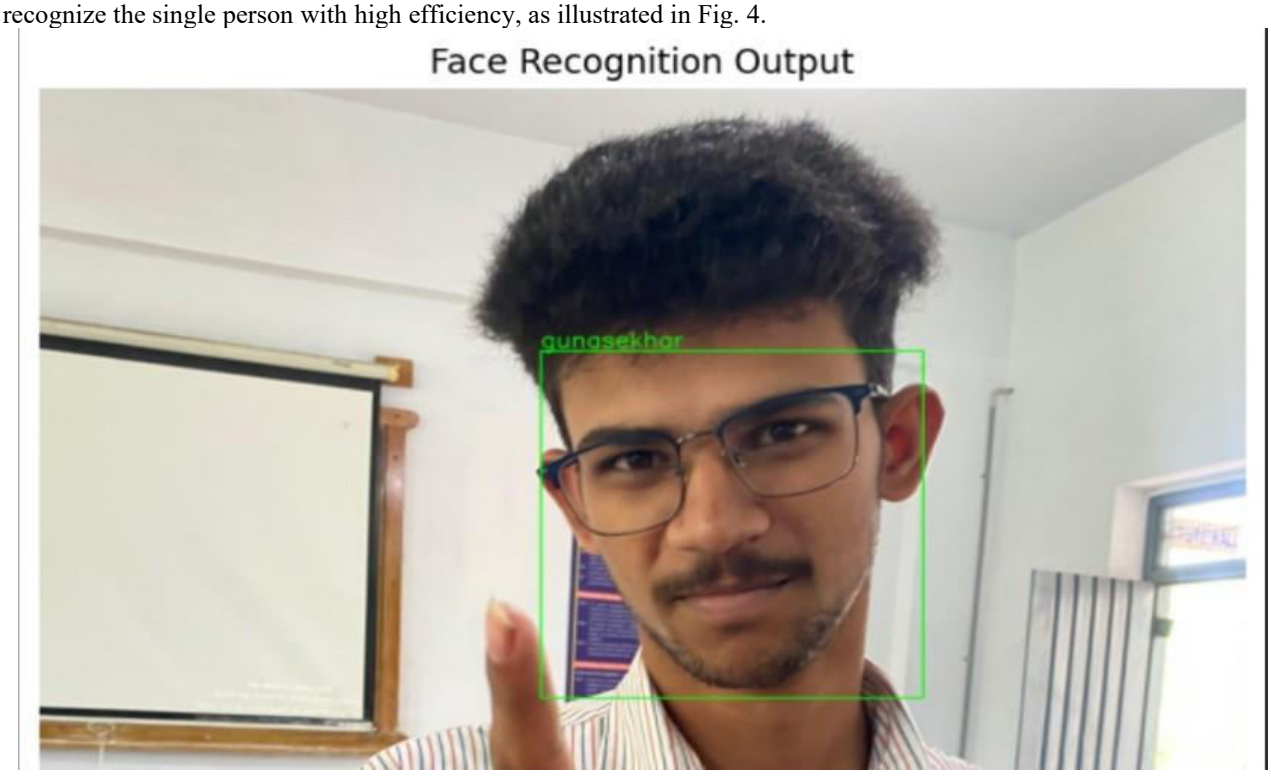


Fig. 4: Output Result from Face Recognition Module (Single Person Detection)

© 2026 IJMRSET | Volume 9, Issue 4, April 2026| DOI:10.15680/IJMRSET.2026.0904131

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Action Recognition

The performance of the action recognition module has been tested with action samples. The CNN-LSTM model classified human actions such as boxing, handwaving, running, and walking successfully, as presented in Fig. 5.

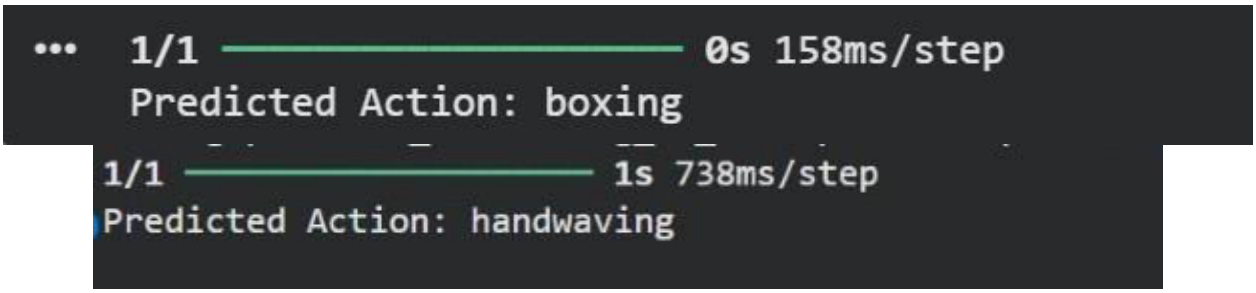
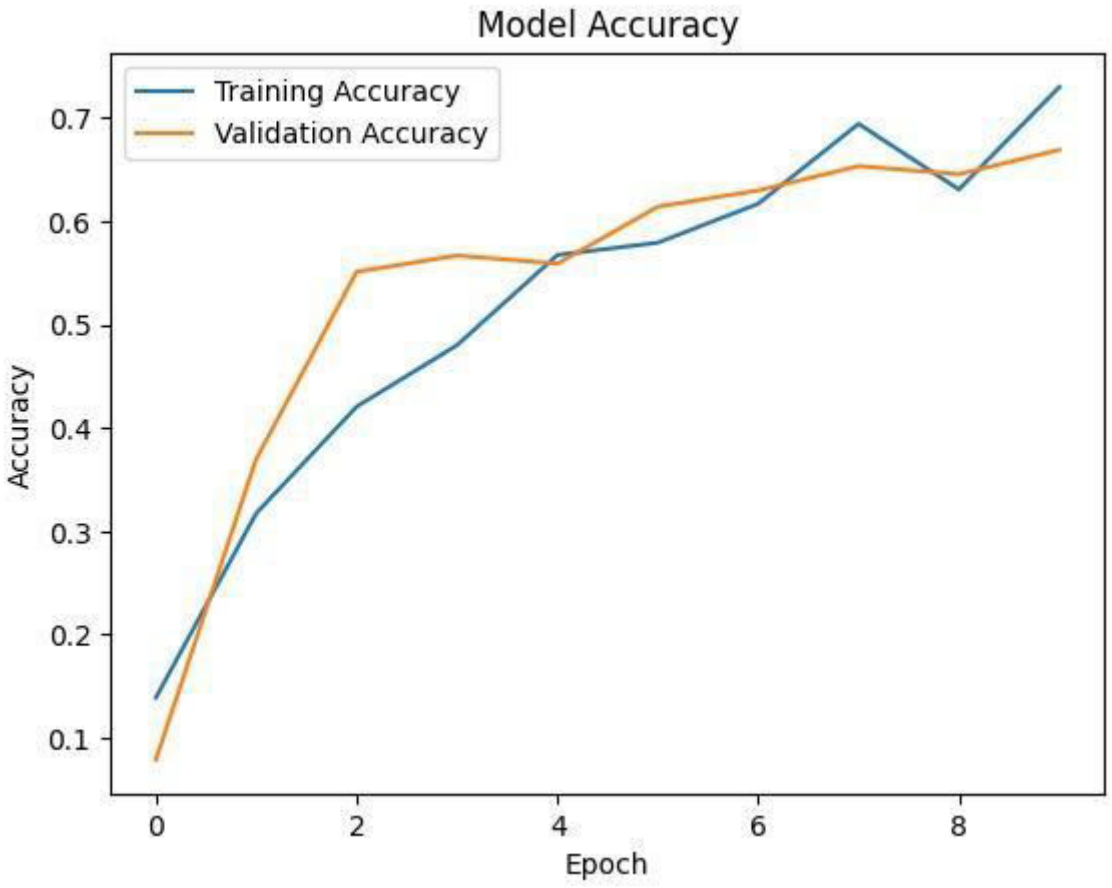


Fig. 5: Output of Action Recognition System for Identified Activity

The training performance of the action recognition system was evaluated based on the performance metrics of accuracy. The accuracy of the model was increasing in both the training and validation sets with respect to various epochs, which is depicted in Fig. 6 below.



ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

	precision	recall	f1-score	support
handclapping	0.80	0.91	0.85	22
walking	0.71	0.68	0.70	22
running	0.76	0.45	0.57	29
boxing	0.94	0.65	0.77	26
handwaving	0.53	0.80	0.64	10
jogging	0.39	0.67	0.49	18
accuracy			0.67	127
macro avg	0.69	0.69	0.67	127
weighted avg	0.73	0.67	0.68	127

Fig. 6: Training Performance of the Action Recognition System with Respect to Accuracy

The system performed well under live situations. Combining face recognition and action recognition resulted in intelligent monitoring and hence could be used in areas such as campus safety surveillance.

VI. CONCLUSION

This paper demonstrates an AI-based drone surveillance system that is capable of real-time identification of humans and recognition of their activities. The proposed system has improved situational awareness, as opposed to traditional means of surveillance. The use of a drone for surveillance is advantageous, as opposed to camera-based systems.
The face recognition component of the proposed system is effective in identifying individuals through encoding-based approaches. Similarly, the action recognition component of the proposed system is effective in recognizing activities of humans through a CNN-LSTM-based approach.

The proposed system is cost-effective, efficient, and has potential applications in campus security, public security, and other areas. In conclusion, the proposed system has demonstrated the potential for artificial intelligence to be used in conjunction with drones for intelligent and automated systems.

VII. RECOMMENDATIONS FOR FUTURE WORK:

The proposed system could be improved to function efficiently in different environments. Future work could be directed toward improving the accuracy of the system. In addition, future work could be directed toward implementing real-time processing on edge devices, as opposed to cloud-based systems and automating the operation of Drone surveillance.

REFERENCES

[1] S. Kumar et al., “AI-Based Smart Surveillance System for Real-Time Monitoring,” International Journal of Innovative Research in Technology (IJIRT), (2023).
[2] T. Manikandan et al., “IoT-Based Smart Surveillance System Using Computer Vision,” IEEE Conference Proceedings, (2024).
[3] J. Miao et al., “Deep Learning-Based Human Activity Recognition System,” Journal of Engineering Science, (2024).
[4] R. Patel et al., “Real-Time Face Recognition System Using Deep Learning Techniques,” International Journal of Computer Applications, (2023).
[5] A. Sharma et al., “CNN-LSTM Based Human Activity Recognition Using Video Data,” IEEE Access, (2025).

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[6] N. Jain et al., "AI-Based Video Surveillance System for Smart Cities," IEEE Conference Proceedings, (2024).
[7] S. Verma et al., "Drone-Based Surveillance System with Real-Time Monitoring," International Journal of Advanced Research in Engineering and Technology (IJARET), (2023).
[8] T. Delwar et al., "Real-Time Surveillance and Activity Detection Using Deep Learning," Future Internet Journal, (2025).
[9] D. E. King, "Dlib-ml: A Machine Learning Toolkit," Journal of Machine Learning Research, (2009).
[10] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, (1997).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com