iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71381

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Deep Learning Approach for Suspicious Activity Detection from Surveillance Video in Exam Hall

Shruti A. Joshi[1], Aditya A. Agarwal[2], Anjali R. Yadav[3], Siddhant L. Aher[4], Bhalchandra R. Ban[5]
[1, 2, 3, 4]*Student,* [5]*Professor, Department of Computer Engineering, Sinhagad College of Engineering, Vadgaon, Pune - 411041, Maharashtra, India*

Abstract: ***Maintaining the integrity of examinations is essential in educational settings. Conventional surveillance is typically based on manual observation, which is inefficient and error- prone. The project introduces a system based on deep learning intended to identify abnormal behavior in examination rooms using automatic, real-time video analysis. By utilizing state-of-the-art computer vision technology, the system analyzes continuous video streams, derives behavioral features, and labels activities as normal or abnormal. It covers behavior like unauthorized activities, communication, and object transfer that could imply academic dishonesty.***

The system incorporates real-time alerting facilities—such as buzzers and immediate notifications—to facilitate prompt intervention by authorities.

This not only minimizes the requirement for constant human monitoring but also provides a more objective and scalable monitoring process. With flexibility and cross-platform capabilities, the solution is accessible on low-cost hardware and can be utilized across different environments.

Aside from monitoring examinations, the technology can be modified for wider security use in public areas, business premises, and medical facilities. Future development, including multi- camera support, predictive analytics, and privacy-enhancing features, will further enhance its potential. This project illustrates how artificial intelligence can revolutionize surveillance systems to become proactive tools that improve fairness, security, and trust in high-stakes environments.

I. INTRODUCTION

In the field of computer vision, object detection is one of the most sought-after tasks because it has a very wide range of applications, from surveillance systems and driverless cars to industrial automation and healthcare diagnostics. The primary goal of object detection is to precisely detect and localize multiple objects in one image and give both the class labels and bounding box locations. In the years that have passed, this field has seen revolutionary advancements, evolving from the use of old image processing methods to incorporating advanced machine learning and deep learning techniques.

The initial object detection strategies were dominated by hand-engineered features like Haar cascades and Histogram of Oriented Gradients (HOG) and paired with traditional classifiers such as Support Vector Machines (SVM). These approaches, though seminal, were challenged by realistic real-world situations because they had only limited capabilities for generalization over variation in size, lighting, and occlusion. Deep learning changed this with its paradigm shift by allowing models to learn hierarchical features automatically from data. Convolutional Neural Networks (CNNs) have been particularly seminal to further object detection, providing dramatic increases in accuracy and efficiency.

Even with tremendous advances, issues still remain in maintaining real-time performance without sacrificing accuracy, particularly in resource-limited settings. Balancing detection speed and accuracy still affects the design and implementation of object detection systems. In addition, maintaining robustness across varying environments and object classes is still a topic of active research. The impetus for this effort arises from the necessity to find an optimal balance between computational cost and detection stability in real-world applications.

This work explores contemporary object detection architectures focused on efficient performance and flexibility. Taking advantage of advancements in neural network architecture and training methods, the goal is to create a solution that is capable of carrying out fast and precise detection within real-time contexts. The findings reported in this paper demonstrate the viability of sustaining high detection performance along with operating efficiency, and as such, they provide valuable input to the burgeoning field of object detection.

II. LITERATURE SURVEY

Over the past few years, demand for intelligent monitoring systems has been high, particularly for ensuring sensitive environments' integrity like examination halls. Machine learning and deep learning technologies have become key to automating detection of anomalies and suspicious behavior from video monitoring.

Choudhry et al. performed an extensive review of machine learning methods for anomaly detection within surveillance videos, classifying the methods into supervised, unsupervised, and semi- supervised learning paradigms. The paper highlights that although supervised models enjoy the luxury of labeled training data and provide high accuracy, their working is usually limited by the availability of large-scale training datasets. On the other hand, unsupervised learning algorithms like autoencoders and generative models are able to detect anomalies without training data, although they tend to be plagued with high false positives. Semi-supervised methods seek to find a balance point by training from partially labeled sets of data. This survey also identifies some of the challenges that include data imbalance, real-time computation constraints, and the challenge of being able to generalize models to new settings, therefore stressing the importance of developing more robust algorithms.

Zaidi and Sampedro investigated the application of convolutional neural networks (CNNs) and deep video models for the detection of suspicious human behavior from surveillance videos. Their method combined Time Distributed CNN and Conv3D architectures to process spatio-temporal patterns over video frames to detect unusual actions in real-time. While their approach attained top detection accuracy, the computational overhead of deep 3D models was a bottleneck for hardware- limited environments like legacy CCTV systems.

In another notable contribution, Saba et al. proposed a new deep learning model called L4- Branched-ActionNet, optimized with Entropy-Coded Ant Colony System algorithms. This hybrid technique aimed at enhancing the efficiency and accuracy of suspicious activity detection through intelligent optimization of feature representation and model structure. The model had promising results in precision and robustness but its complexity raised issues of implementation and scalability in real-world applications.

Concurrently with anomaly detection, the area of video steganography has also developed, especially for security purposes. Zhao et al. put forward a transform block decision steganographic method for H.265/HEVC video sequences. Their approach optimized the structure of the transform block to hide secret information while having high perceptual quality of the video and compression ratio. This paper, though mainly focused on secure communication and copyright protection, provides useful observations on compressing video structures—something that could be of use for embedding detection metadata or surveillance annotations with little bandwidth overhead.

These works collectively show the growing use of deep learning to improve video surveillance systems. They also reveal the limitations that remain in terms of real-time processing, computational efficiency, and the flexibility of detection models to different environments. Future work must thus continue to investigate lightweight, scalable solutions that can retain high accuracy under realistic deployment constraints.

III. PROBLEM STATEMENT

Deep learning approach for suspicious activity detection from surveillance video in exam hall.

Upholding the integrity and fairness of exams is a major issue in schools. Conventional surveillance methods, which are predominantly based on human monitoring or simple camera observation, are no longer sufficient in detecting contemporary and subtle cheating and dishonesty. Such manual approaches are not only error-prone and time-consuming but also have limited capabilities to monitor more than one subject or setting at a time.

As the world experiences an outburst in growth in deep learning and computer vision technologies, it is quite real to have massive opportunities for surveillance systems automation with smart models recognizing suspicious behavior real-time. Such automated systems face high computational requirement, lack in generalizability between different scenarios of examination, and inability in accurately identifying the behavior of man in a group or in fluctuating environments. In addition, most solutions available are either not designed specifically for the particular context of school examination rooms or do not work efficiently on low-cost hardware that is commonly found in such settings.

Hence, it is necessary to develop and deploy a strong, scalable, and efficient deep learning-based system that can take video input from surveillance cameras, extract spatio-temporal features, and effectively detect suspicious activities like cheating or unauthorized interactions. The system must be able to operate in real-time, reduce false positives, and be resource-friendly for deployment in resource-constrained environments, while also taking into account privacy and ethical issues related to video surveillance.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

IV. PROPOSED SYSTEM

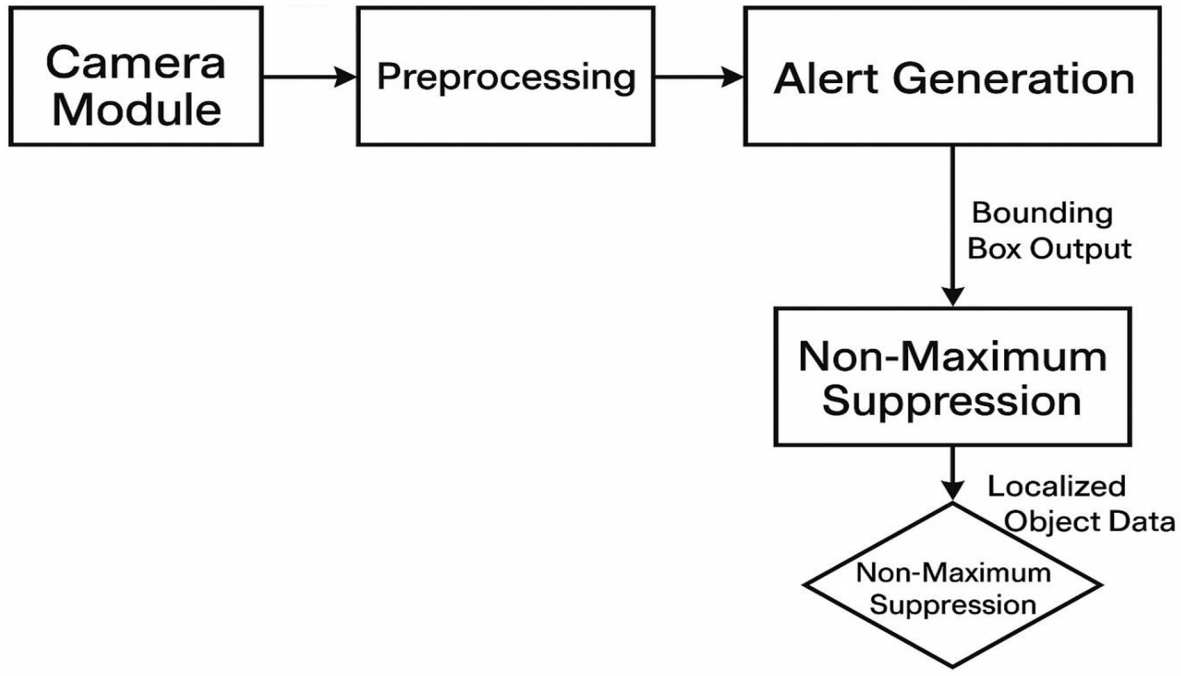
The system proposed here intends to create a real-time alert and object detection mechanism for enhanced situational awareness in surveillance and safety-critical uses. The system combines a high-resolution camera module with a resource-light, yet high-precision convolutional neural network (CNN) model, which is designed to achieve efficient object recognition and tracking in varied environments. It is able to identify various object classes, such as pedestrians, automobiles, and typical obstacles, with low computational latency, thus finding it applicable for edge deployment.

The three main modules that make up the architecture are image acquisition, object detection, and alert generation. Image acquisition is tasked with capturing real-time video streams through a camera sensor connected to a processing unit. Frames gathered are preprocessed for uniform quality and resolution before being forwarded to the detection module.

The detection module uses a CNN model that has been trained with a carefully curated dataset with annotated images of different object classes. The architecture of the network is with alternating convolution, pooling, and activation layers and dense classification and bounding box regression layers. The output of the model is the object classes and their coordinates in each frame, allowing accurate localization.

To improve on reliability, non-maximum suppression is used to filter out overlapped bounding boxes such that objects are clearly detected. The alerts generation module analyses spatial location and movement path of detected objects in order to gauge potential threats or infringement. In an example where smart traffic surveillance is being carried out, for instance, the application can provide alerts when a car crosses prohibitive areas or pedestrians venture into danger zones.

The system is built on a modular software pipeline to enable scalability and integration with IoT devices with ease. Moreover, real-time performance is ensured via GPU acceleration as well as model quantization methods, providing efficient inference without compromising accuracy.



Dig. System Flowchart

V. EXISTING SYSTEM

Surveillance methods used currently in educational institutions are based to a large extent on manual monitoring by human invigilators or simple video recording systems, which have a number of shortcomings. Observation by humans is susceptible to fatigue, partiality, and inconsistency, particularly in observing several students at once over extended periods. Conventional surveillance systems do not generally incorporate real-time decision-making support, with minimal or no automation for the identification of likely threats or rule breaches during examination.

To counter these issues, deep learning and computer vision technologies have been incorporated into current surveillance systems to automate the real-time detection of suspicious behavior. These systems process video streams in real-time from surveillance cameras, breaking them into frames and separating dynamic foreground objects from static backgrounds. These systems, employing high-level feature extraction methods including motion tracking, analysis of head movements, and detection of contacts, are capable of detecting deviating behavior patterns showing signs of misconduct or cheating.

The architecture of the system involves a number of interrelated modules that process live video streams through a series of steps like segmentation, background subtraction, feature extraction, and activity classification. The features extracted are fed into a convolutional neural network (CNN), which is trained to identify and classify human activities as normal or suspicious based on patterns learned. Instant alerts are sent when abnormal behavior is identified, sending real-time alerts via email, SMS, or sound alarms to examination authorities.

It accommodates both desktop and mobile platforms, is hardware-cost-effective, and has a user- friendly graphical user interface. It presents a scalable, flexible solution with minimal human involvement that achieves a high degree of accuracy and integrity in exam supervision.

VI. MATERIALS AND METHODS

This segment details the hardware and software tools utilized in building and deploying the suspicious activity detector system. It was built as a real-time desktop application applying deep learning models to process images of examination centers. The structure was designed with high performance as a priority without compromising compatibility across widely used equipment and open-source.

A. Software Requirements

The development setup was on a Windows 10 operating system. Python was the chosen main programming language because it has a vast library of frameworks and libraries backing machine learning, computer vision, and data processing. The used integrated development environment (IDE) for development was Spyder due to its suitability for scientific computing and availability as part of Anaconda distribution, which made it easy to handle packages.

For data storage, the system used SQLite—a lightweight relational database management system with no server requirements that is well-suited for embedded applications. SQLite offered enough functionality to manage user data and alert logs without the penalty of an explicit server. The software also featured GUI utilities to allow for user interaction and graphical outputs. Deep learning models were also integrated through Python libraries for enabling accurate feature extraction.

B. Hardware Requirements

To ensure accessibility and affordability, the system was implemented and tested on mid-grade hardware. A computer system with an Intel Core i3 processor (or equivalent) was adequate for the deployment environment. The system needed a minimum of 8 GB of RAM to effectively perform real-time video processing operations such as frame segmentation, background-foreground extraction, and feature classification. At least 20 GB of hard disk space was dedicated to software dependencies, video datasets, model weights, and database files. The software was made to be compatible with widely available input/output peripherals like a monitor, keyboard, and webcam, so that deployment in standard classroom settings is easy. The lightweight setup ensures that institutions can deploy the system without requiring high costs.

VII. EXPERIMENTAL RESULT

To measure the efficiency of the proposed system for detecting suspicious activities in exam rooms, thorough experiments were performed on video datasets simulating common exam situations. The dataset was comprised of normal as well as suspicious student activities like repeated head movements, unexpected hand gestures, and unapproved interactions.

The model was also validated in real-time to evaluate its performance in dynamic situations. It showed high accuracy in suspicious versus normal activity classification based on a deep convolutional neural network-based architecture. The system could take continuous video input, perform segmentation and background-foreground separation, and then motion feature extraction, head movement, and physical contact feature extraction. The classifier generated timely and accurate notifications whenever abnormal patterns were identified. Notifications were given via various channels, such as audio notifications and electronic notifications. The system ensured stability and responsiveness even on low-end hardware platforms with Intel i3 processors and 8GB of RAM, proof that the system can be used in environments with limited resources such as schools.

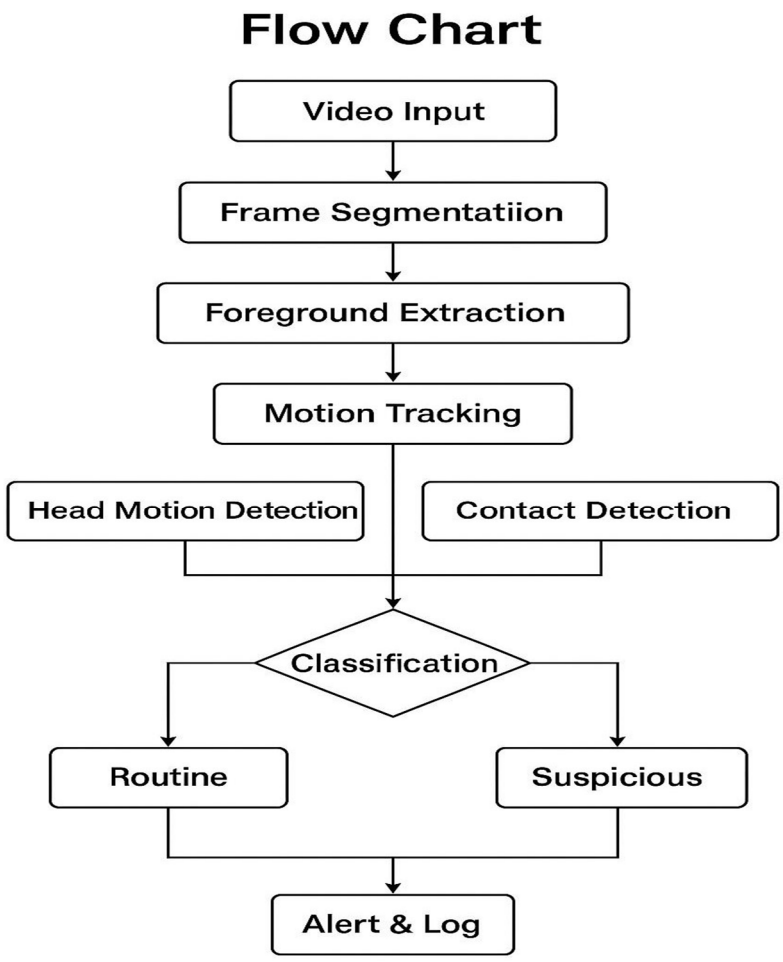
Precision, recall, and F1-score measurements were employed for quantitative assessment. The outcome revealed that the system was very effective in reducing false negatives (missed suspicious behavior), with low false positive rates (normal activity reported as suspicious). The feedback loop allowed the model to learn over time, adjusting to differences in behavior and lighting.
Furthermore, the system's resilience was tested in different lighting and crowd conditions. Even in occluded scenes, it maintained efficient performance by utilizing sophisticated foreground detection.

VIII. MATHEMATICAL MODEL

$S = \{I, F, B, G, M, H, C, A, O\}$
Where:

a. I = Input video stream from surveillance (CCTV) cameras.
b. F = Frame segmentation: $I \rightarrow \{f_1, f_2, ..., f_n\}$, where each f_i is an individual frame.
c. B = Background extraction function: $B(f_i)$ = static elements across frames.
d. G = Foreground extraction: $G(f_i) = f_i - B(f_i)$, isolating moving subjects.
e. M = Motion tracking function: tracks velocity, direction, and movement vectors of objects.
f. H = Head motion detection: $H(G) \rightarrow \{left, right, up, down\}$, capturing gaze or focus shifts.
g. C = Contact detection: identifies overlaps or interactions between detected subjects.
h. A = Activity classification: $A(M, H, C) \rightarrow \{Normal, Suspicious\}$, based on pattern thresholds.
i. O = Output set: $O = \{Alert, Log\}$, where alerts are real-time notifications and logs are stored events.

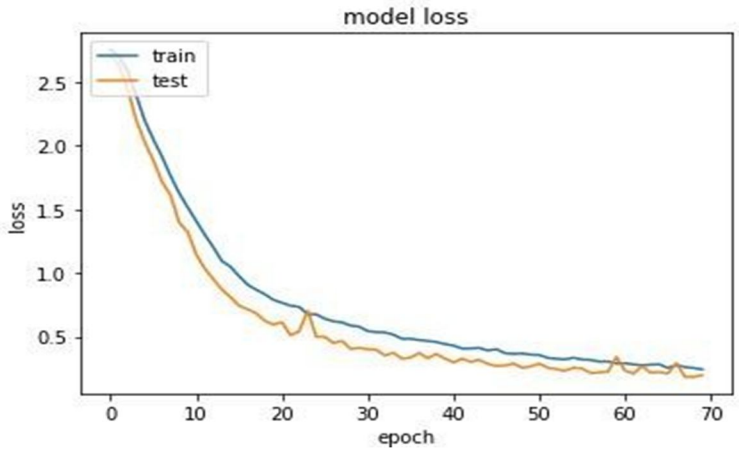
The system applies the composite function: $O = A(M(G(f_i)), H(G(f_i)), C(G(f_i)))$



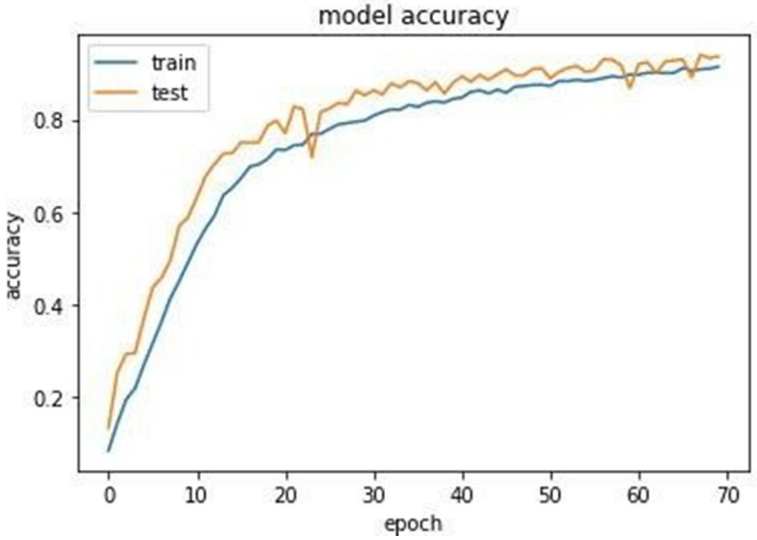
Dig. Mathematical model Flowchart

IX. SIMULATION RESULT

Dig. Model Loss



The training loss and validation loss both trend downward consistently through epochs, meaning that the model is learning properly. The training loss drops gradually from a starting point over 2.5 to under 0.4, and the validation loss traces a similar pattern and typically stays slightly lower than the training loss after the first few epochs. This is indicative that the model is not overfitting since the validation loss does not diverge or grow during training. Rather, it indicates that the model is generalizing very well to new data.



Dig. Model Accuracy

The training accuracy begins below 0.2 and continues to get better gradually, crossing 0.85 after the end of training. The validation accuracy enhances much more steeply, touching more than 0.9 at around epoch 40 and continuing with a consistent increasing trend. The difference between training and validation accuracy is minimal and gets smaller towards the later epochs, which further strengthens the fact that the model is strong and resists overfitting.

X. CONCLUSION

The system of suspicious activity detection designed for exam monitoring is a novel approach to enhancing academic integrity through real-time automated monitoring. Through the use of deep learning and computer vision methodologies, the system analyzes live video feeds, recognizes human activities, and labels them as normal or suspicious. Important processes include video segmentation, motion and head tracking, and contact detection to correctly recognize possible misconduct such as unauthorized communication or object transfer.

This smart system minimizes reliance on manual proctoring considerably, providing equal and fair monitoring throughout test sessions. It features automated alerting mechanisms—e.g., buzzers, emails, and SMS alerts—that quickly notify authorities if suspicious activity is identified. This enables prompt intervention, rendering the exam process more secure and equitable for all pupils.

With scalability in mind, the system can be installed not just in educational institutions but also in other high-security zones such as government buildings, banks, and public places. Its modular design allows for the integration of upgrades such as predictive analytics and pose estimation, which can predict threats even before they completely emerge.

With advancements in artificial intelligence and deep learning technology, the system can be enhanced to function effectively in difficult conditions, including low lighting or heavy crowds. Integration with IoT devices, multi-camera systems, and privacy-protecting methods like anonymization are some of the future developments that can be envisioned.

In general, the project shows the real-world promise of AI-based surveillance in the preservation of security and order. Through the automation of detection and alert mechanisms, it provides a foundation for wiser, more responsive monitoring solutions across various applications.

REFERENCES

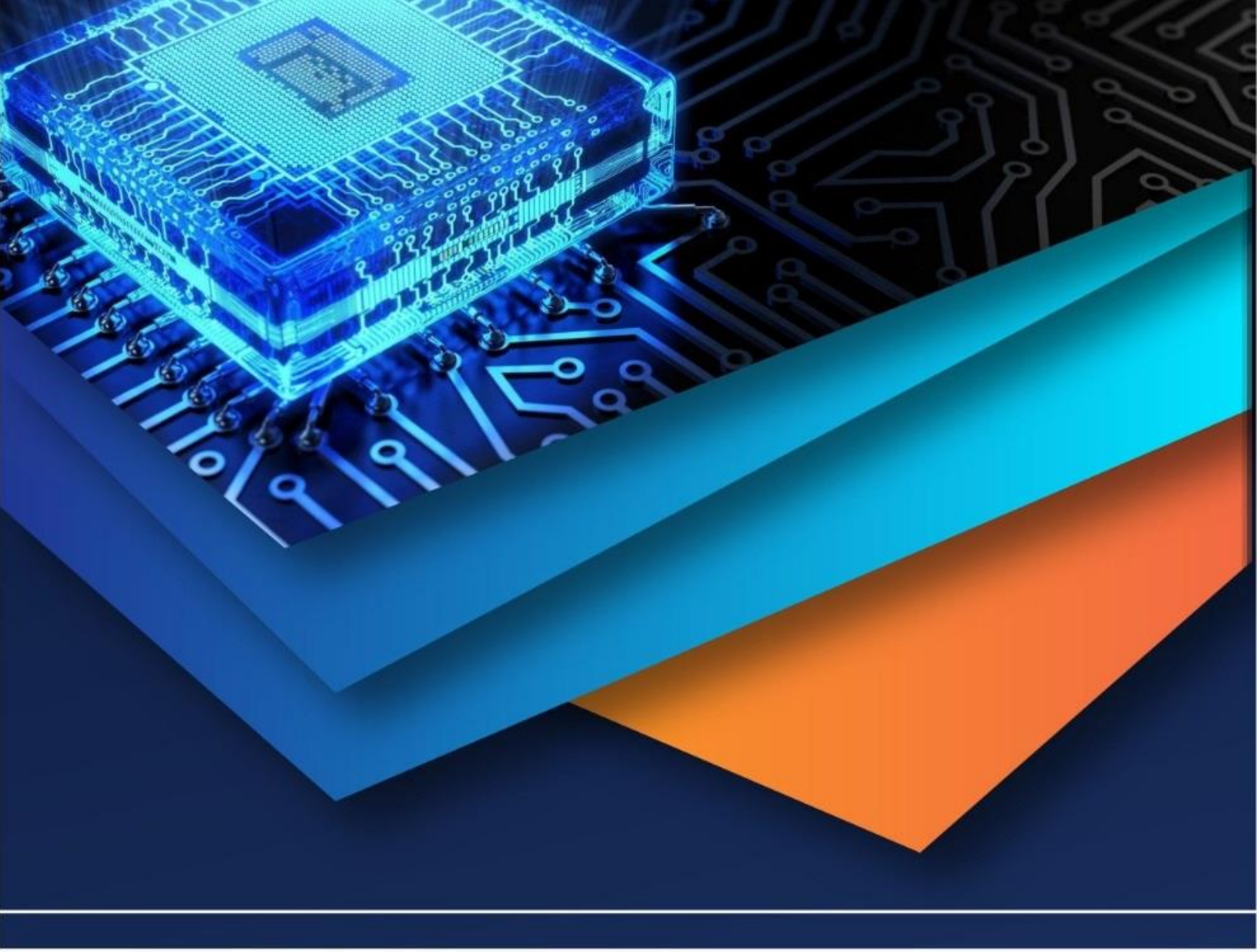
[1] Khan, S., & Sultana, F. (2020). Suspicious Activity Detection in Surveillance Videos using Deep Learning Approaches. Journal of Computer Science and Technology, 35(2), 185-196.

[2] Ali, F., & Mehmood, I. (2019). Deep Learning for Suspicious Activity Detection in Public Surveillance Videos. International Journal of Advanced Computer Science and Applications, 10(5), 249-255.

[3] Soleimani, M., & Ghafoorifard, H. (2018). An Intelligent Surveillance System for Exam Hall Cheating Detection Using Deep Learning. Proceedings of the International Conference on Machine Learning and Computer Vision (MLCV 2018).

[4] Zhao, R., & Li, H. (2021). Real-time Detection of Suspicious Behavior in Exam Halls using Deep Learning and Surveillance Systems. Journal of Artificial Intelligence Research, 70, 27- 41.

[5] Basharat, A., & Shah, M. (2020). Behavioral and Temporal Analysis for Suspicious Activity Detection in Surveillance Videos. IEEE Transactions on Image Processing, 29, 2227-2238.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY