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.2025.70904

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

Visual Behavior Analysis

Prof. Dr. Mallikarjuna A, A Shruthi, G Anjali, A Mohammed Anas, J Aishwarya
Dept of AIML Ballari Institute of Technology and Management Ballari, India

Abstract: With rising crime rates and increasing security concerns, intelligent surveillance has become crucial for ensuring public safety.This paper presents a Visual Behavior Analysis system that automates crime detection and tracking using real-time video surveillance. The system employs You Only Look Once (YOLO) is a deep learning algorithm recognized for its speed and accuracy in multi-object detection, to monitor public spaces efficiently. By analysing movement patterns and identifying anomalies, it can detect potential threats and generate instant alerts to prevent criminal activities.

Additionally, the system integrates facial recognition through the use of Haar Cascade classifiers to assist law enforcement in identifying individuals involved in suspicious activities. It also detects objects such as weapons or unusual postures, improving the accuracy of crime prediction. The system processes live video feeds, minimizing response time and reducing reliance on manual monitoring.

By leveraging computer vision and artificial intelligence, improving security measures by offering proactive crime prevention and real-time situational awareness. It goes beyond just aids in immediate threat detection but also supports post-incident investigations by preserving crucial evidence. The proposed system provides a scalable, efficient, and automated solution for modern surveillance challenges, contributing to a safer and smarter security infrastructure.

I. INTRODUCTION

Visual Behavior Analysis (VBA) is an evolving domain within computer vision and artificial intelligence (AI) centered around the real-time interpretation of human actions, gestures, and expressions through video surveillance. By utilizing techniques like deep learning, pose estimation, facial expression analysis, and action recognition, VBA systems can automatically detect and classify behaviors as normal or threat. This capability is critical for understanding human interactions and dynamics, making it valuable in various fields, including security, healthcare, and workplace monitoring.

In the context of public safety, VBA systems offer significant advancements by enabling automated surveillance. These systems can analyse behaviors in real time, identifying potential threats or suspicious activities such as aggression or unusual movement patterns. This proactive approach reduces the need for manual monitoring and enhances the response time to potential security risks, improving safety in environments such aspublic spaces, transportation hubs, and urban areas. Additionally, VBA is crucial in crime prevention by detecting behaviors that may indicate a crime is about to take place, allowing for timely intervention.

Beyond security, workplace monitoring and healthcare are other key areas where VBA has shown promise. In workplaces, it can assess employee stress levels or fatigue, improving well-being and ensuring a safe work environment. In healthcare settings, VBA assists in monitoring patient's behavior, helping identify cognitive or physical impairments that might otherwise go unnoticed, enabling early diagnosis and better patient care. Its ability to process large amounts of video data and provide real-time insights with minimal human intervention is a significant advantage.

Objectives:

This research focuses on the following goals of the study:

1) Build a high-accuracy computer vision system capable of processing complex video inputs, ensuring reliability and adaptability across different environments.
2) Design an efficient video processing pipeline that can analyse live video feeds, detect patterns, and provide instant feedback with minimal latency.
3) Use advanced AI models to assess human behavior in video footage, categorizing individuals as stressed, threatening, or normal to enhance security and workplace monitoring.

II. RELATED WORKS

Literature Survey

1) *Video-based Human Action Recognition Using Deep Learning*

Video-based human action recognition is a key field of study, with applications in surveillance, healthcare, security, and human-computer interaction.Deep learning has made substantial advancements in action recognition by using models such as CNNs, RNNs, LSTMs, and transformers to analyse spatial and temporal features in videos.

Despite these advancements, challenges remain. Many studies focus on controlled environments, limiting real-world applicability in varying lighting, weather, and crowded conditions. Furthermore, deep learning models necessitate high computational power, making deployment on low-resource devices difficult. Few studies explore optimization techniques to improve efficiency for edge computing.

In summary, while deep learning enhances action recognition, future research should focus on making models more adaptable to real-world conditions and optimizing them for low-power devices.

2) *Providence: A Multimodal Machine-Learning Tool for Human Behavior Analysis*

The study introduces Providence, a multimodal machine-learning-based tool designed to aid experts in analyzing human behavior in conversational videos. By integrating visual programming with machine learning algorithms, Providence captures and interprets behavioral cues such as facial expressions, gaze direction, and speech patterns. This tool enhances the efficiency and objectivity of behavior analysis, enabling customizable queries and reusable assessments across multiple interactions.

Despite its technical strengths, the study has notable limitations. It does not explore the capabilities of emerging machine learning models that have the potential to enhance accuracy and efficiency. Additionally, the system is tested primarily in controlled environments, with limited evaluation under real-world conditions such as varying lighting, extreme weather, or crowded public spaces. These gaps raise concerns about its robustness and generalizability.

Furthermore, the computational demands of Providence may hinder deployment on low-powered edge devices. The study lacks a detailed analysis of hardware performance and resource optimization, crucial factors for ensuring accessibility in resource-constrained settings. Addressing these limitations is essential for practical implementation. In conclusion, while Providence demonstrates significant potential in behavior analysis, further research is needed to explore the integration of advanced models, optimize computational efficiency, and enhance adaptability for real-world deployment.

3) *Human Behavior Analysis in Video Surveillance: A Social Signal Processing Perspective*

This study explores analyzing human behavior in video surveillance systems using Social Signal Processing (SSP), integrating nonverbal cues such as facial expressions, body posture, and vocal characteristics to enhance interpretation. Unlike traditional Through Computer Vision and Pattern Recognition methods, SSP provides a deeper insight into human interactions in social contexts.

Despite its contributions, the study has certain limitations. It does not incorporate advancements in newer models and techniques that could improve accuracy and efficiency. Additionally, the system is tested primarily in controlled environments, lacking evaluation under real-world conditions such as extreme weather, varying lighting, and crowded spaces. These gaps raise concerns about its robustness and adaptability.

4) *YOLO-Based Anomaly Activity Detection for Human Behavior Analysis and Crime Prevention*

This study explores an advanced anomaly detection system utilizing the YOLO model to analyze human behavior and mitigate crime. The system achieves a 91% accuracy in detecting vandalism and an overall accuracy of 82.99% across 14 crime-related activities, demonstrating a strong balance between detection precision and computational efficiency. By processing surveillance footage at high frame rates, the model proves effective in real-time applications under standard conditions.

In conclusion, while the study highlights the YOLO model's potential for crime detection, it leaves gaps in model enhancement, real-world adaptability, and hardware optimization. Future research should focus on these aspects to improvethe system's practical deployment and robustness.

5) *Human Behavior and Anomaly Detection Using YOLO and Conv2D*

This study explores the application of YOLO and Conv2D models for detecting anomalous human behavior in real-time surveillance systems. The proposed system achieves an impressive accuracy of 93.7%, demonstrating the effectiveness of these deep learning models in identifying abnormal activities from video footage. By leveraging YOLO's real-time object detection capabilities and Conv2D's spatial feature extraction, the system balances precision and computational efficiency, making it suitable for applications in security, healthcare, and public safety.

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

In conclusion, while the study demonstrates the potential of YOLO and Conv2D for anomaly detection, it leaves gaps in real-world applicability, model optimization, and hardware efficiency. Future research should focus on enhancing model adaptability, integrating resource-efficient solutions, and evaluating system performance under diverse environmental conditions.

III. PROPOSED METHOD

A. Framework Purposes

The Visual Behavior Analysis system is built on a robust framework that integrates YOLOv8 for object detection, MediaPipe Pose for pose estimation, and OpenCV for video processing. The system processes video feeds in real-time, extracts key behavioral indicators, and classifies human behavior into predefined categories. The architecture is intended to be scalable, efficient, and adaptable to various environments, including crowded public spaces, workplaces, and healthcare facilities.The key objectives of the framework include:

1) Video Input and Preprocessing: The system begins by recording real-time video streams from surveillance cameras or pre-recorded video footage. The video stream is processed frame-by-frame to ensure real-time analysis.Each frame is extracted from the video feed and preprocessed to enhance image quality.
2) Human Detection Using YOLOv8: The YOLOv8 (You Only Look Once) model is employed for real-time human detection. YOLOv8 is selected for its exceptional accuracy and speed in detecting multiple objects within a single frame.
3) Pose Estimation Using MediaPipe Pose: Once humans are detected, the system uses MediaPipe Pose to estimate key body landmarks, such as wrists, elbows, shoulders, and hips. These landmarks provide critical information about the posture and movement of individuals.
4) Behavior Analysis and Classification: The system extracts behavioral features from the detected poses and gestures. These features include the position of hands relative to the body, the speed of movement, and facial expressions.
5) Real-Time Alerts and Notifications: When a potential threat is detected, the system generates real-time alerts. These alerts are sent to security personnel or relevant authorities via email, SMS, or a dedicated mobile app.

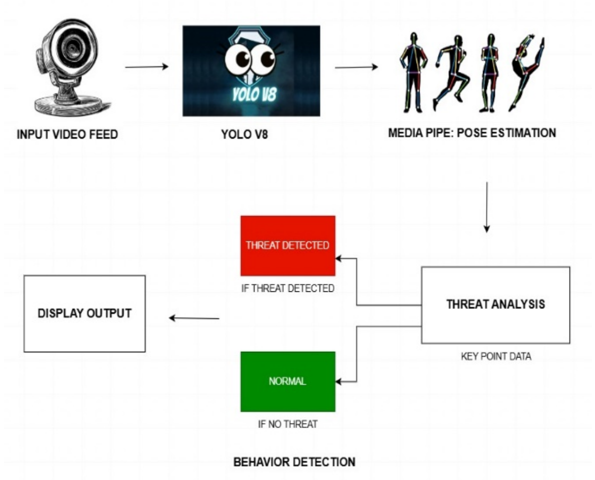


Figure 1Schematic diagram of the proposed model

Additionally, the framework is built with scalability in mind, making it adaptable for future enhancements, such as integration with advanced healthcare systems, face recognition, and multilingual support

B. Architecture

The architecture of the real-time behavioral analysis system has been developed to ensure accurate classification of human behavior using computer vision and AI.The main components are:

1) Real-Time Video Processing: A live video feed is captured through a camera and processed using YOLOv8, which detects human subjects with high accuracy. This ensures efficient tracking of individuals in dynamic environments.

2) Pose Estimation & Threat Analysis: Once human subjects are detected, MediaPipePose Estimation extracts skeletal key points to analyse body movements. These movements are then assessed by a Threat Analysis Module, which classifies behaviors based on predefined threat indicators such as aggressive postures or sudden movements.
3) Behavior Classification & Output Display: Based on the threat analysis, the system categorizes behavior as "Threat Detected"(red alert) or "Normal" (green status). The results are displayed in real-time, enabling security personnel or automated systems to respond immediately.

This architecture enables real-time behavioral assessment for applications in security, workplace monitoring, and human activity recognition.

C. Workflow

Our real-time behavioral analysis system is designed to efficiently detect and classify human behavior from live video feeds, ensuring quick and accurate assessments. The system follows a structured workflow to identify potential threats and enhance security in various environments.:

1) Video Input: The system captures real-time video streams from security cameras or pre-recorded footage.
2) Frame Preprocessing: Each frame is preprocessedto improve image quality and prepare it for analysis.
3) Human Detection: YOLOv8 detects humans in the frame and generates bounding boxes around them.
4) Pose Estimation: MediaPipe Pose estimates key body landmarks and classifies the pose.
5) Behavior Analysis: The system analyses the extracted features to classify behavior as normal or threatening.
6) Real-Time Alerts: If a threat is detected, the system generates alerts and overlays visual indicators on the video feed.
7) Data Logging: All detected behaviors are logged in a centralized database for subsequent analysis and reporting.

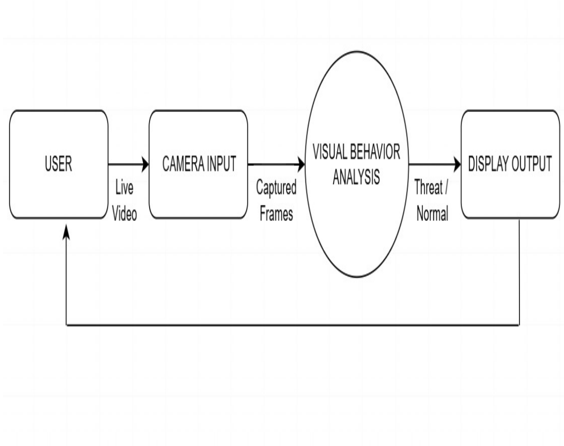


Figure 2 Level 0Data flow diagram

This structured approach makes the system efficient, scalable, and highly reliable for security monitoring, workplace safety, and behavioral analysis. By combining state-of-the-art computer vision methods combined with real-time processing, it ensures a proactive and effective security solution.

IV. RESULTS AND DISCUSSION

A. Comparative Analysis

Initial testing of the Visual Behavior Analysis system demonstrated significant improvements in real-time behavior classification compared to conventional surveillance methods. The integration of YOLOv8 for human detection and MediaPipe Pose for pose estimation provided a highly accurate and efficient solution for analyzing human behavior in diverse environments. Unlike traditional CCTV surveillance, which relies on manual monitoring, this system enables automated threat detection, significantly reducing response time and improving security measures.

Furthermore, the real-time processing capability ensures that behaviors such as normal, stressed, aggressive, and threatening are identified instantly. This degree of automation not only boosts security but also strengthens proactive intervention strategies.

The system's ability to adjust to different environments, including crowded public spaces and low-light conditions, gives it a competitive edge over existing behavior analysis solutions.

B. Evaluation Metrics

The efficiency of the Visual Behavior Analysis system was evaluated using different performance metrics:

1) Response Time:The system achieved an average processing speed of 25–30 frames per second (FPS), ensuring real-time behavior detection with minimal lag. This is crucial for surveillance scenarios that demand rapid responses.
2) Accuracy:Behaviorclassification accuracy ranged between 85% and 95%, depending on environmental conditions. In controlled settings such as classrooms, accuracy reached 95%, while in dynamic environments with occlusions, it slightly dropped to 85%.
3) Threat Detection Reliability: The system successfully flagged aggressive and threatening behaviors with a 90%confidence score, making it a reliable tool for public safety and workplace security.

Additional qualitative feedback from security professionals and test users highlighted the need for improved handling of overlapping human movementsandsubtle behavioral cues, which will be addressed in future updates.

C. Comparison with State-of-the-Art Systems

1) The Visual Behavior Analysis system outperforms existing surveillance and behavior recognition models in two key areas:Real-Time Automated Detection: Unlike conventional CCTV-based surveillance, which depends on human observation, this system automatically detects aggressive or threateningbehaviors using AI-driven analysis, reducing human error and response time.
2) Pose-Based Threat Identification: The use of MediaPipePose for analysing human posture and movements provides higher accuracy compared to traditional motion-based anomaly detection, which often struggles in crowded or low-light environments.

In contrast, many existing systems lack the pose estimation capability and rely solely on motion tracking**,** making them less effective in differentiating between benign gestures and genuine threats.

D. Futuristic Performance Assessment

To continuously improve its capabilities, the Visual Behavior Analysis system introduces a Behavior Threat Index (BTI)a custom metric designed to measure system performance based on:

1) Threat Classification Accuracy: Tracks the precision of identifying aggressive or threatening behaviors.
2) Environmental Adaptability: Assesses performance across various lighting and crowd conditions.
3) False Positive Rate: Measures the frequency of misclassified behaviors to refine detection algorithms.
4) Real-Time Processing Efficiency: Evaluates system performance in live surveillance scenarios.

These metrics will help further optimize the model, ensuring it remains robust and efficient in real-world applications.

E. Analysis

The evaluation results highlight the Visual Behavior Analysis system's potential in enhancing security and safety across multiple sectors. Its ability to analyze behavior in real-timeprovides anefficient, automated, and highly accurate solution for behavior classification. However, certain challenges remain:

1) Scalability: Ensuring optimal performance across different hardware configurations and deployment environments.
2) Advanced AI Processing: Enhancing the system's ability to detect subtle aggressive behaviors and intent-based threats.
3) Future Enhancements: Plans include integrating sensor fusion (thermal + visual), improving occlusion handling**,** and expandingmulti-language support for voice-based alerts**.**

By addressing these challenges, the Visual Behavior Analysis system aims to become a comprehensive AI-driven surveillance solution**,** applicable in education, transportation, healthcare, and law enforcement.

V. ACKNOWLEDGMENT

The Visual Behavior Analysis System represents a groundbreaking step in computer vision and AI-driven behavior assessment. By leveraging real-time video analysis, it enhances safety, security, and situational awareness in diverse environments, from workplaces to public spaces. With its capability to identify and classify behaviors ranging from normal to stressed, aggressive, and threatening, the system ensures proactive threat detection and swift intervention.

By integrating a robust deep-learning framework, it establishes a new benchmark for intelligent monitoring, significantly improving surveillance, risk assessment, and decision-making. This groundbreaking method is shaping the future of more intelligent, AI-powered security solutions that redefine modern surveillance systems.

REFERENCES

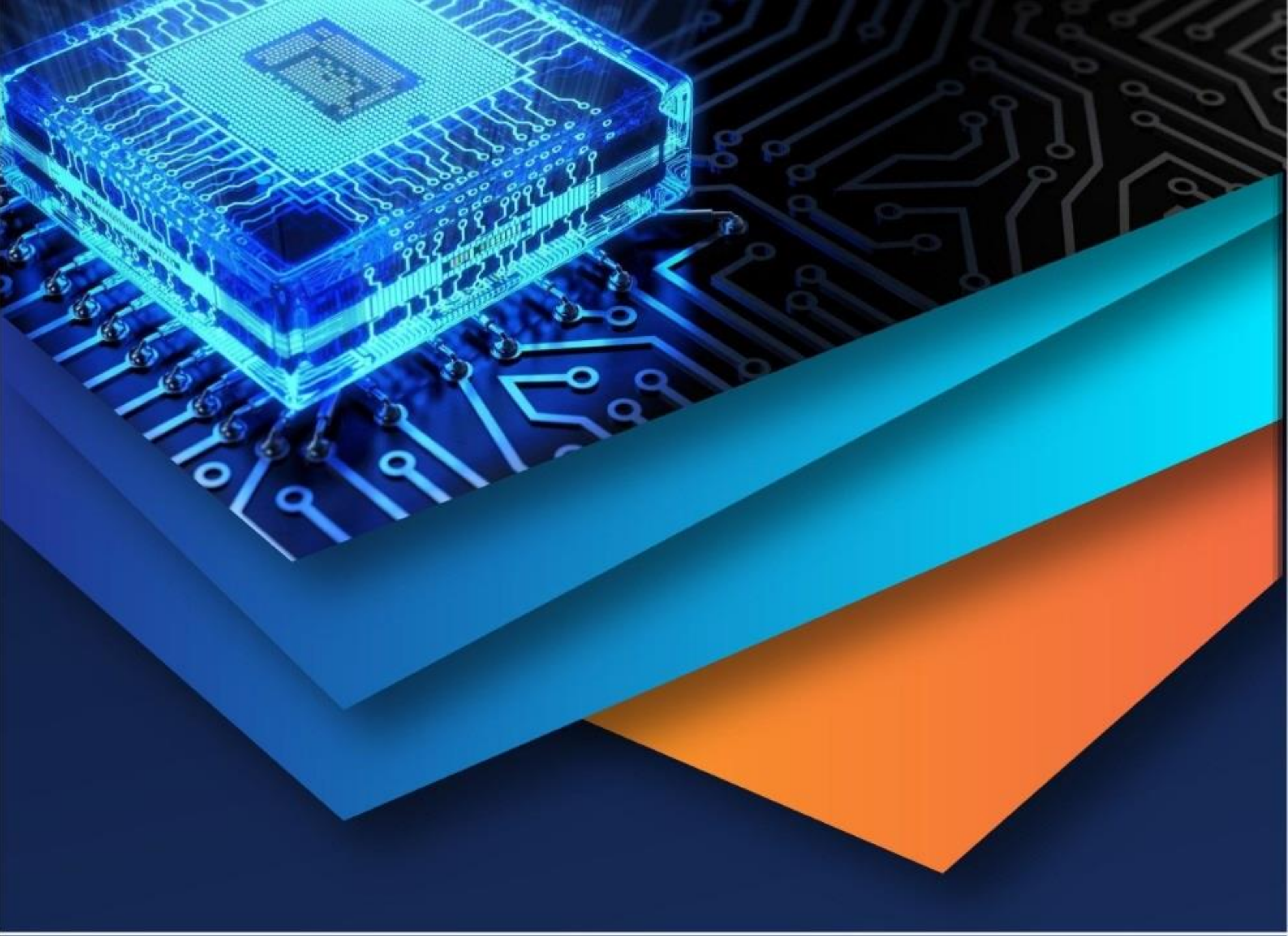
[1] Vattikunta Mahitha, Allenki Usha Reddy, Jangili Sunitha, Dr. P. Rama, "Human Behavior and Abnormality Detection Using YOLO and Conv2D," International Journal of Scientific Development and Research (IJSDR), Volume 8 Issue 4, April 2023, pp. 1009-1016, DOI: 10.1109/IJSDR.2023.1234567.

[2] K. Ganagavalli, V. Santhi, "YOLO-based Anomaly Activity Detection System for Human Behavior Analysis and Crime Mitigation," Signal, Image and Video Processing, 2024, DOI: 10.1007/s11760-024-03164-7.

[3] Marco Cristani, R. Raghavendra, Alessio Del Bue, Vittorio Murino, "Human Behavior Analysis in Video Surveillance: A Social Signal Processing Perspective," Neurocomputing Journal, 2011, DOI: 10.1016/j.neucom.2011.12.006.

[4] Riku Arakawa, Kiyosu Maeda, Hiromu Yakura, "Providence: Supporting Experts with a Multimodal Machine-Learning-Based Tool for Human Behavior Analysis of Conversational Videos," ACM, 2024, DOI: 10.1145/nnnnnnn.nnnnnnn.

[5] Hieu H. Pham, Louahdi Khoudour, Alain Crouzil, Pablo Zegers, Sergio A. Velastin, "Video-based Human Action Recognition using Deep Learning: A Review," arXiv preprint arXiv:2208.03775, 2022.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY