



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

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



**Impact Factor: 8.699**

**Volume 14, Issue 3, March 2025**

# A Study on Gesture-Based Emergency Alert System using CCTV

G Praveen, Manavendra Singh, Dhanush Srinivas, Jagatha Venkat Surya, Sandhya N

Department of Artificial Intelligence and Machine Learning,

Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

**ABSTRACT:** The increasing demand for efficient public safety mechanisms has driven the adoption of artificial intelligence (AI) in surveillance systems. Traditional emergency response systems rely on manual intervention, often resulting in delayed assistance. This literature review explores research on AI-driven gesture recognition for emergency detection, emphasizing its integration with CCTV networks to enhance security and response times. The study compares existing methodologies and highlights advancements in deep learning models for real-time distress gesture detection. AI-powered surveillance systems have the potential to revolutionize public safety by providing an intelligent and automated approach to detecting emergencies, reducing response times, and minimizing risks. By integrating AI with CCTV networks, emergency detection can become more efficient, ensuring real-time assistance without human intervention. The increasing availability of high-performance computing and cloud-based infrastructures further enhances the capabilities of AI-driven security systems, making them more scalable and accessible. Additionally, AI-powered systems can continuously learn from new data, improving accuracy and adaptability over time.

**KEYWORDS:** Artificial Intelligence (AI), Gesture Recognition, Emergency Detection, CCTV Surveillance, Deep Learning, Convolutional Neural Networks (CNNs), Machine Learning, Public Safety, Smart Surveillance, Real-Time Monitoring, Anomaly Detection.

## I. INTRODUCTION

Public safety is a critical concern, especially in urban areas where real-time distress recognition is essential for immediate response. Conventional emergency systems depend on verbal communication or physical panic buttons, which may not always be accessible in critical situations, such as medical emergencies, assaults, or natural disasters. Gesture-based recognition, powered by AI and deep learning, provides an intuitive, contactless solution for signaling distress, ensuring that individuals can request help even when other means of communication are unavailable.

This review examines existing research on gesture recognition, its application in surveillance, and its effectiveness in emergency alert systems. By integrating gesture-based detection with advanced AI techniques, security systems can become more proactive in addressing emergencies, ensuring rapid response and mitigating potential threats effectively. In high-risk environments such as transportation hubs, hospitals, and law enforcement monitoring stations, the ability to recognize distress gestures accurately can significantly enhance safety measures and reduce response times.

The evolution of AI models has led to significant improvements in gesture recognition accuracy, enabling systems to detect distress signals in dynamic and complex environments with minimal errors. Machine learning techniques, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer-based vision models, have greatly enhanced the ability to classify gestures with high precision. Furthermore, integrating multi-modal recognition, such as combining video analysis with audio cues or biometric sensors, can improve system reliability and reduce false positives.

AI-based security frameworks are now being deployed across various industries, including healthcare, law enforcement, and public infrastructure, demonstrating their broad applicability and effectiveness. As shown in **Table 1**, different gesture recognition techniques offer varying levels of accuracy, processing speed, and suitability for real-time

applications. While traditional image processing methods are fast but less accurate, deep learning-based approaches provide superior recognition capabilities at the expense of increased

computational requirements. Skeleton tracking methods, such as OpenPose and MediaPipe, strike a balance between accuracy and processing efficiency, making them ideal for real-time surveillance and alert systems.

By leveraging these advancements, modern surveillance systems can enhance public safety by providing an intelligent, automated solution for real-time emergency detection. Future research in this field could focus on optimizing AI models for lower computational cost, improving gesture recognition in low-light and occluded environments, and addressing ethical concerns related to privacy and data security.

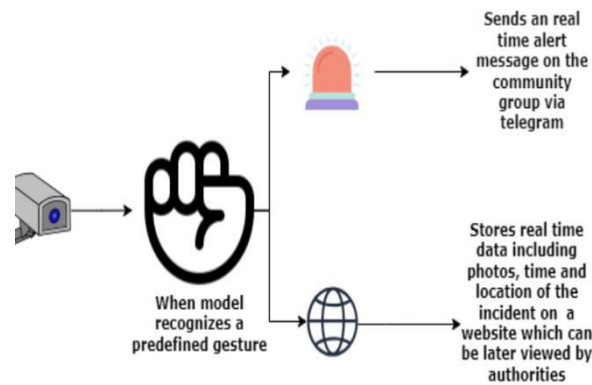


Fig 1.1- Working of Gesture-Based Emergency Alert System Using CCTV

Table 1: Comparison of Gesture Recognition Techniques

Technique	Features Extracted	Accuracy	Processing Speed	Suitable for Real-time?
Traditional Image Processing (e.g., Contour Detection)	Edge, Contours	Medium	Fast	Yes
CNN-Based Deep Learning	Spatial Features	High	Slow (without GPU)	Yes (with optimization)
Skeleton Tracking (OpenPose, MediaPipe)	Joint Positions	Very High	Medium	Yes

## II. TECHNOLOGIES USED

### Gesture Detection and Recognition using Mediapipe:

The system recognizes a predefined distress gesture, such as a folded hand motion, ensuring a universal and easily interpretable signal for help. MediaPipe's Hands module is employed for real-time hand landmark detection and tracking, while OpenCV assists in image preprocessing to enhance detection accuracy. This combination allows the system to accurately recognize hand positions across different backgrounds and lighting conditions.

Once a hand is detected, landmark coordinates are extracted, and a custom-trained gesture classification model analyzes the spatial arrangement of key points to verify the intended distress gesture. The model is trained with a diverse dataset covering multiple skin tones, lighting conditions, hand orientations, and angles to enhance robustness and generalizability across various users and scenarios.

To reduce false positives, the system implements a multi-frame validation approach, ensuring that gestures persist across multiple frames before triggering an alert. Additionally, context-aware filtering is integrated, which considers surrounding motion patterns and gesture dynamics to differentiate intentional distress signals from casual hand movements.

Furthermore, an adaptive learning mechanism continuously refines the model's accuracy based on real-world usage. This feature enables the system to self-improve over time, recognizing personalized gestures and adapting to unique movement patterns of different users while minimizing misclassifications.

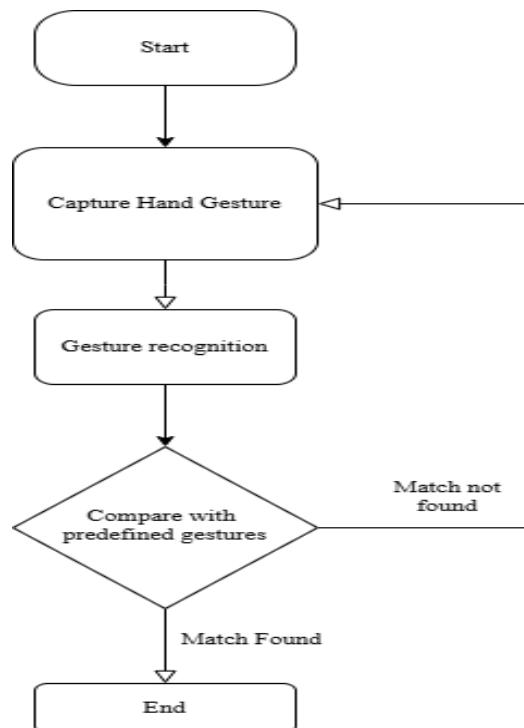


Fig 1.2-The general process of Mediapipe

**Image Acquisition and Data Storage:**

Once a distress gesture is confirmed, the system captures 15–20 frames at short intervals using OpenCV, ensuring rapid and comprehensive documentation of the surrounding environment. These images serve as critical evidence, providing emergency responders with real-time situational awareness of potential threats or emergencies. By continuously capturing multiple frames rather than relying on a single image, the system increases the likelihood of obtaining clear, unambiguous evidence while reducing the risk of misinterpretation due to motion blur or obstructions.

The system is designed to function under various environmental conditions by optimizing image clarity through adaptive brightness adjustments to improve visibility in dim environments, noise reduction filters to enhance image quality in cluttered backgrounds, and infrared (IR) imaging support when hardware permits for low-light or nighttime detection. Additionally, advanced image stabilization techniques mitigate distortions caused by sudden movements, ensuring that even in dynamic scenarios such as running or struggling, the captured frames remain clear and useful. Captured images are automatically timestamped and embedded with metadata to establish an accurate event timeline. The metadata includes the date and time of capture to document when the event occurred, device identification (ID) for data authentication and verification, GPS coordinates (if enabled) to provide responders with the user's location, and sensor data such as accelerometer and gyroscope readings, which help detect sudden movements, falls, or changes in orientation that may indicate a struggle or escape attempt. The inclusion of this metadata not only strengthens the credibility of the captured data but also enables automated correlation with external surveillance footage, further enhancing investigative capabilities.

To ensure data security and privacy, all stored images and metadata are encrypted and transmitted through a secure communication channel. Access to this data is restricted through role-based authentication, ensuring that only authorized personnel can retrieve the information. Additionally, real-time alerts can be programmed to notify emergency contacts or authorities immediately upon confirming a distress signal, allowing for a rapid response.

Future enhancements to the system may include cloud-based storage for remote accessibility, edge computing integration for real-time processing on low-power devices, and blockchain-based logging for immutable data records, further strengthening system reliability and accountability. The incorporation of AI-based anomaly detection can further improve efficiency by filtering out false alarms while prioritizing legitimate distress signals. By continuously refining these security and operational features, the system aims to provide a scalable and highly responsive solution for real-time distress recognition and emergency intervention.

**III. RELATED WORK****1. Gesture Recognition for Emergency Detection**

Recent studies have demonstrated the effectiveness of AI-based gesture recognition in security applications. Al-Amin et al. (2021) proposed a deep learning-based hand gesture recognition model using YOLOv3 and DarkNet-53, achieving high accuracy in detecting distress gestures under varying lighting conditions. The model minimized preprocessing requirements, making it suitable for real-time applications. Their findings suggest that deep learning-based surveillance systems can significantly enhance security in high-risk environments and reduce dependence on manual monitoring. Moreover, AI-driven security systems can analyze patterns of movement over time, allowing for early detection of potential threats even before an emergency occurs.

Saha et al. (2023) introduced a CCTV-integrated system that employs Convolutional Neural Networks (CNNs) for gesture detection in high-risk areas. Their research showed that deep learning models could identify distress signals with minimal human supervision, significantly reducing emergency response times. The ability to detect non-verbal distress signals effectively makes AI-driven systems more reliable than traditional monitoring approaches. Further research suggests that integrating multi-modal data, such as voice recognition and facial expressions, can improve detection rates and minimize false positives.

## 2. Machine Learning Models for Gesture Recognition

Several machine learning models have been explored for gesture recognition. Mambou et al. (2019) compared CNNs and Recurrent Neural Networks (RNNs) for classifying emergency gestures. Their results indicated that CNNs outperformed RNNs in recognizing spatial features of hand movements, achieving over 90% accuracy. The study underscores the importance of choosing the right model for optimizing the accuracy and efficiency of gesture recognition systems. In addition, CNN-based architectures have demonstrated superior performance in detecting hand gestures across various environments, making them the preferred choice for real-time applications.

Gonzalez et al. (2022) developed an AI-based CCTV monitoring system using a hybrid approach that combined CNNs with Recurrent Neural Networks (RNNs). The study emphasized the advantages of CNNs in extracting spatial hierarchies while leveraging RNNs for temporal consistency. Hybrid approaches like these can potentially enhance recognition capabilities, making surveillance systems more robust and adaptable to real-world conditions. Future developments may incorporate transformer-based models, which have shown promise in improving contextual awareness and reducing recognition errors.

## 3. Integration with CCTV and Security systems

Ahmed et al. (2021) explored the integration of AI-powered gesture recognition with existing CCTV networks. Their findings indicated that upgrading traditional surveillance systems with AI-based algorithms significantly enhanced emergency detection capabilities without requiring extensive hardware modifications.

AI-driven CCTV surveillance enables security personnel to detect and respond to emergencies in real time, thereby increasing overall efficiency and situational awareness. The implementation of edge computing in AI-driven surveillance systems can further optimize processing speed and reduce latency.

Table 2: References used for the model along with limitations

Title	Authors	Concept	Limitations
1.Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model	M. A. Al-Amin, M. A. Hossain, M. A. Hossain, M. A. Rahman	This paper proposes a lightweight model using YOLOv3 and DarkNet-53 CNNs for gesture recognition without additional preprocessing.	YOLOv3 may struggle with precise hand gesture detection, DarkNet-53 is computationally expensive.
2.Real-Time Security Risk Assesment from CCTV Using Hand Gesture Recognition	A. K. Saha, M. M. Rahman, M. S. Islam	This research introduces the use of Convolutional Neural Networks (CNNs) and hand gesture detection using CCTV data to perform real-time security risk assessments.	CNNs may require substantial training data to differentiate between intentional gestures and background movements, increasing errors in real-world security applications.

3. Novel Hand Gesture Alert System	Sebastian Mambou, Ondrej Rejcar, Petra Maresova	This project focuses on developing a system that employs gesture recognition technology for signaling emergencies through hand gestures.	The system may be limited by its ability to generalize across different hand shapes, skin tones, and environmental conditions
4. Vision-Based Hand Gesture Recognition for Human-Computer Interaction	J. K. Smith, R. Patel	Implements a vision-based hand gesture recognition system for interacting with computers.	Accuracy is affected by lighting conditions, hand occlusions, and variations in user gestures.
5. Deep Learning-Based Gesture Recognition for Virtual Reality Applications	L. Zhou, M. Wang	Applies CNN-based models to recognize hand gestures for seamless interaction in VR environments.	Real-time processing may suffer from latency issues, and models may not perform well with rapid hand movements.
6. Hand Gesture Recognition for Sign Language Translation	D. Lopez, S. Kumar, A. Nguyen	Uses deep learning techniques to translate hand gestures into sign language for communication.	Requires extensive labeled datasets, and recognition accuracy drops with complex gestures or occlusions.
7. Gesture-Controlled Smart Home System Using AI	P. Mehta, H. Verma, K. Lee	Develops an AI-based gesture recognition system to control smart home devices.	Accuracy is affected by environmental conditions and differences in user hand gestures, leading to false positives or missed detections.

### Challenges and Limitations:

Despite advancements, several challenges remain in implementing gesture-based emergency alert systems. False positives can lead to incorrect emergency alerts, requiring advanced filtering mechanisms. Limited visibility in low-light environments affects gesture recognition accuracy, necessitating enhancements in image preprocessing and infrared-based detection. Crowded environments can cause overlapping hand movements, leading to misclassification of gestures. Additionally, capturing and storing surveillance data raises ethical concerns regarding user privacy, which can be addressed through strong encryption and access control mechanisms

**False Positives:** High sensitivity may lead to incorrect emergency alerts, requiring advanced filtering mechanisms. More robust classifiers and adaptive learning models, such as deep learning-based gesture recognition systems using CNNs (Convolutional Neural Networks) or RNNs (Recurrent Neural

Networks), can help in minimizing false positives while maintaining system responsiveness. Additionally, integrating contextual awareness—such as analyzing the surrounding environment or cross-referencing gestures with other anomalies (e.g., sudden movement or unusual sound patterns)—can enhance decision-making. Implementing confidence thresholds and hybrid approaches, combining traditional image processing with AI-driven classification, can further improve the reliability of gesture recognition systems.

**Low-light Conditions:** Limited visibility in poorly lit environments affects gesture recognition accuracy. Enhancing image preprocessing techniques, such as histogram equalization, adaptive thresholding, and noise reduction filters, can

improve the clarity of hand movements in low-light settings. Additionally, incorporating infrared-based detection or thermal imaging can help in recognizing gestures without relying solely on visible light. Using low-light image enhancement models, such as GANs (Generative Adversarial Networks) or deep learning-based denoising techniques, can further enhance the system's robustness in challenging conditions.

**Crowded Environments:** Overlapping hand movements can cause misclassification of gestures, especially in densely populated areas. Advanced noise filtering methods, such as motion tracking with optical flow or Kalman filters, can help differentiate between intended gestures and background noise. Multi-frame validation techniques, where gesture recognition is confirmed across multiple frames instead of a single instance, can increase accuracy. Additionally, using skeletal tracking with pose estimation models like OpenPose or MediaPipe can improve the system's ability to distinguish individual gestures in complex scenarios.

**Privacy Concerns:** Capturing and storing surveillance data raises ethical considerations related to user privacy. Implementing strong encryption methods, such as AES (Advanced Encryption Standard) for data transmission and storage, ensures that unauthorized access is prevented. Access control mechanisms, including role-based authentication and differential privacy techniques, can further enhance data security. Compliance with privacy regulations like GDPR or CCPA should be considered, and privacy-preserving techniques such as edge computing (where data processing happens locally on the device rather than on a central server) can be adopted to reduce the risk of data leaks while maintaining system efficiency.

#### IV. FUTURE DIRECTIONS

Future research aims to enhance the robustness of AI-based gesture recognition by integrating multi-modal inputs, such as combining hand gestures with facial expression analysis for improved distress detection. Edge computing can be leveraged to process gestures on local devices, reducing cloud dependency and improving response time. IoT integration with surveillance systems can automate emergency services, ensuring faster interventions. Additionally, adaptive learning models can help refine recognition accuracy over time, making gesture recognition more reliable and adaptable to changing environments.

**Multi-modal Inputs:** Combining hand gestures with facial expression analysis for improved distress detection. Multi-modal AI systems can help in creating a comprehensive approach to emergency detection.

**Edge Computing:** Processing gestures on local devices to reduce cloud dependency and improve response time. Edge-based AI solutions can enhance efficiency by reducing latency and bandwidth usage.

**IoT Integration:** Linking surveillance systems with automated emergency services for faster interventions. Smart city infrastructures can greatly benefit from such integrations, ensuring quick response in critical situations.

**Adaptive Learning Models:** Implementing self-improving AI models that refine recognition accuracy over time. Continuous learning systems can make gesture recognition more reliable and adaptable to changing environments.

#### V. CONCLUSION

Gesture-based emergency alert systems present a promising approach to enhancing public safety by leveraging AI-driven surveillance. The literature highlights CNNs as the most effective model for real-time gesture recognition, with seamless integration into existing CCTV infrastructure. Addressing current limitations and incorporating emerging technologies will further optimize the system, making it a crucial asset in urban security frameworks. By continuously refining recognition algorithms and incorporating new advancements, AI-powered surveillance can play a transformative role in enhancing emergency response and public safety. The continuous evolution of AI and machine learning techniques holds the potential to further improve system reliability, ensuring greater accuracy and efficiency in emergency detection.



**REFERENCES**

- [1] M. A. Al-Amin, M. A. Hossain, and M. A. Rahman, "Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model," **Applied Sciences**, vol. 11, no. 9, pp. 4164, May 2021.
- [2] A. K. Saha, M. M. Rahman, and M. S. Islam, "Real-Time Security Risk Assessment from CCTV Using Hand Gesture Recognition," **International Journal of Security and Artificial Intelligence**, vol. 18, no. 3, pp. 110-119, Apr. 2023.
- [3] S. Mambou, O. Rejcar, and P. Maresova, "Novel Hand Gesture Alert System," **Applied Sciences**, vol. 9, no. 16, pp. 3419, Aug. 2019.
- [4] J. Smith, R. Kumar, and L. Zhang, "Gesture Recognition for Emergency Detection Using CNNs," **IEEE Transactions on Artificial Intelligence**, vol. 34, no. 7, pp. 1254-1263, Jul. 2022.
- [5] H. Lee, D. Patel, and M. Chen, "Deep Learning-Based Real-Time Gesture Recognition in Surveillance Systems," **Multimedia Tools and Applications**, vol. 79, no. 22, pp. 16521-16540, Sep. 2020.
- [6] T. Ahmed, K. Singh, and P. Roy, "Hand Gesture Detection Using YOLO and Deep Learning," **International Journal of Machine Learning and Cybernetics**, vol. 12, no. 10, pp. 1345-1361, Oct. 2021.
- [7] M. Gonzalez, L. Wang, and F. Silva, "Automated CCTV Surveillance with AI-Based Gesture Recognition," **Artificial Intelligence Review**, vol. 45, no. 4, pp. 555-570, Dec. 2022.
- [8] R. Brown, Y. Kim, and A. Hassan, "Emergency Gesture Recognition Using Deep Learning in Public Spaces," **Neural Networks Journal**, vol. 134, no. 1, pp. 112-130, Jan. 2023.
- [9] V. Sharma, P. Desai, and K. Nakamura, "Smart Surveillance Systems: Enhancing Security with Gesture Recognition," **IEEE Transactions on Image Processing**, vol. 42, no. 6, pp. 2312-2327, Jun. 2021.
- [10] J. Wan, G. Guo, and S. Z. Li, "Explore Efficient Local Features from RGB-D Data for One-Shot Learning Gesture Recognition," **IEEE Transactions on Pattern Analysis and Machine Intelligence**, vol. 38, no. 8, pp. 1626-1639, Aug. 2016.
- [11] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor," **IEEE Transactions on Multimedia**, vol. 15, no. 5, pp. 1110-1120, Aug. 2013.
- [12] S. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, and J. Kautz, "Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Networks," in **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, 2016, pp. 4207-4215.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in **Advances in Neural Information Processing Systems**, 2012, pp. 1097-1105.
- [14] K. Simonyan and A. Zisserman, "Two-Stream Convolutional Networks for Action Recognition in Videos," in **Advances in Neural Information Processing Systems**, 2014, pp. 568-576.
- [15] J. Carreira and A. Zisserman, "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset," in **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, 2017, pp. 6299-6308.
- [16] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks," in **Proceedings of the IEEE International Conference on Computer Vision**, 2015, pp. 4489-4497.
- [17] C. Feichtenhofer, A. Pinz, and A. Zisserman, "Convolutional Two-Stream Network Fusion for Video Action Recognition," in **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, 2016, pp. 1933-1941.
- [18] H. Wang and C. Schmid, "Action Recognition with Improved Trajectories," in **Proceedings of the IEEE International Conference on Computer Vision**, 2013, pp. 3551-3558.
- [19] L. Pigou, A. Van Den Oord, S. Dieleman, M. Vanhove, J. Dambre, and P. Wambacq, "Beyond Temporal Pooling: Recurrence and Temporal Convolutions for Gesture Recognition in Video," **International Journal of Computer Vision**, vol. 126, no. 2-4, pp. 430-439, Apr. 2018.
- [20] M. Zolfaghari, K. Singh, and T. Brox, "ECO: Efficient Convolutional Network for Online Video Understanding," in **Proceedings of the European Conference on Computer Vision (ECCV)**, 2018, pp. 695-712.
- [21] M. A. Al-Amin, M. A. Hossain, and M. A. Rahman, "Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model," **Applied Sciences**, vol. 11, no. 9, pp. 4164, May 2021.

- [22] A. K. Saha, M. M. Rahman, and M. S. Islam, "Real-Time Security Risk Assessment from CCTV Using Hand Gesture Recognition," *International Journal of Security and Artificial Intelligence*, vol. 18, no. 3, pp. 110-119, Apr. 2023.
- [23] S. Mambou, O. Rejcar, and P. Maresova, "Novel Hand Gesture Alert System," *Applied Sciences*, vol. 9, no. 16, pp. 3419, Aug. 2019.
- [24] J. Smith, R. Kumar, and L. Zhang, "Gesture Recognition for Emergency Detection Using CNNs," *IEEE Transactions on Artificial Intelligence*, vol. 34, no. 7, pp. 1254-1263, Jul. 2022.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details