iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.77803

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Touchless Gesture-Based Human-Machine Interaction for Industrial Safety and Automation

Ms. Rutuja Gharat[1], Mr. Neej Patel[2], Ms. Rashmi Chaudhari[3], Mrs. Hiral Patel[4]
Department of Data Science, St. John College of Engineering and Management, Palghar

Abstract: Recent research explores the integration of computer vision to advance industrial automation, specifically focusing on touchless gesture control and safety monitoring. The sources describe systems designed to replace physical interfaces with hand gestures for controlling Programmable Logic Controllers (PLCs) and collaborative robots, thereby enhancing hygiene and worker safety in hazardous zones [1][4]. These technological advancements utilize deep learning models like LSTM and 3D Convolutional Neural Networks to ensure real-time recognition accuracy [1][4]. Nevertheless, user-centered evaluations highlight barriers such as the absence of tactile feedback and social concerns regarding the professional perception of gesturing [2]. Complementing these control interfaces, the sources also propose action-aware safety frameworks that automatically verify Personal Protective Equipment (PPE) compliance based on the specific task being performed [3]. Collectively, these contributions illustrate a shift toward more intuitive, hands-free industrial environments that prioritize both operational efficiency and proactive risk mitigation.
Keywords: Industry Automation, Human Mahine Interaction, Industrial Safety, Gesture recognition

I. INTRODUCTION

The shift toward Industry 4.0 has introduced collaborative robots into shared workspaces, necessitating natural communication methods that do not rely on traditional control panels [1][4]. Traditional physical interfaces, such as switches and buttons, are often cumbersome, restrictive, and susceptible to wear, which increases maintenance downtime [1]. In hazardous environments, touchless interaction improves safety by eliminating the risk of electric shocks, burns, or injuries associated with direct contact [1][3]. Hygiene-sensitive sectors like healthcare and food processing also benefit from hands-free operation to prevent the spread of germs and bacteria [1][2]. Furthermore, industrial environments often involve residues or materials that can damage touchscreens or smear physical controls [4]. Touchless gesture systems also address ergonomic challenges; for example, in factory loading stations, they allow operators to control machinery from a safe distance without constant movement between the equipment and fixed panels [2]. By providing real-time feedback and accessibility for workers with limited mobility, gesture control enhances system responsiveness and operational efficiency [1][4].

II. METHODOLGY

Data acquisition across the sources ranged from logging MediaPipe landmark coordinates and 3D skeleton tracking via Kinect sensors to curating 320 hours of industrial surveillance footage. Preprocessing techniques involved motion isolation through inter-frame subtraction to eliminate background noise and human-detection filtering to standardize clip quality. Recognition was primarily facilitated by hybrid architectures, combining 3D Convolutional Neural Networks (3D CNN) or Multilayer Perceptrons (MLP) with Long Short-Term Memory (LSTM) and ConvLSTM2D layers to capture spatio-temporal dependencies [1][4]. Safety-centric frameworks specifically integrated SlowFast networks with YOLOv9 for action-aware PPE verification [3]. Hardware integration featured direct communication with Mitsubishi iQ-R PLCs using the MC protocol and Python-based real-time interfaces (Tkinter/OpenCV) for cobot interaction [1]. Evaluation methods employed MS-COCO metrics, such as mAP and IoU, alongside multidimensional scaling (MDS) to analyze emotional user experience factors like competence and frustration [2]. Finally, comparative human studies were used to validate algorithmic recall against professional safety standards.

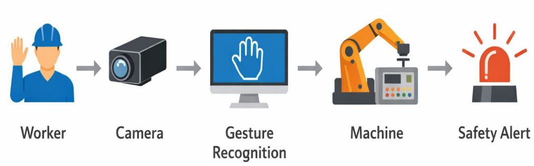


Fig 1: Workflow of model

III. LITERATURE ANALYSIS

Research on touchless gesture control in industrial environments demonstrates a clear progression from technical development toward practical deployment and safety-oriented applications. A vision-based approach for human–machine interaction has been proposed using MediaPipe to extract 21 hand landmark coordinates, which are processed through Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP) models [1]. The system achieves 92% recognition accuracy and directly interfaces with Mitsubishi iQ-R PLCs through the MC communication protocol, enabling real-time control of industrial servo mechanisms. This work establishes the feasibility of replacing traditional physical control panels with gesture-driven automation.

Real-time hand gesture recognition for collaborative robotic systems has also been explored through alternative architectures [4]. Instead of landmark tracking, this approach employs a hybrid deep learning model combining 3D Convolutional Neural Networks and ConvLSTM layers to capture spatio-temporal features from raw video frames. The system preprocesses input by isolating motion from static backgrounds using grayscale frame differencing and delivers predictions every 2.4 seconds via an OpenCV-based interface. This method highlights the importance of robust preprocessing and temporal modeling for reliable industrial gesture recognition in complex environments.

While the above studies emphasize technical accuracy and system implementation, human-centered aspects of gesture interaction have also been critically examined [2]. Using Microsoft Kinect sensors and template-based recognition, user acceptance of gesture interfaces was evaluated in real factory environments. Although touchless interaction improves safety and hygiene, significant barriers were identified, including lack of tactile feedback and social perceptions that gesturing appears unprofessional. This work underscores that successful industrial adoption depends not only on algorithmic performance but also on operator comfort, ergonomics, and perceived reliability.

Gesture and vision-based systems have further been extended into the domain of industrial safety monitoring [3]. An action-aware framework combining SlowFast networks for activity recognition with YOLOv9 object detection has been proposed for task-specific PPE verification. By first recognizing the worker's action—such as welding or material handling—the system checks only for the relevant protective equipment, reducing false alarms and improving F1-score by 23%. This demonstrates how action recognition can enhance automated safety compliance in dynamic industrial settings.

Collectively, the reviewed studies reveal that effective touchless automation requires a balance between recognition accuracy, real-time performance, hardware integration, and human usability. While deep learning models enable precise gesture interpretation, practical deployment must also address user acceptance and safety requirements. The literature therefore highlights the need for integrated systems that combine robust technical frameworks with human-centered design principles.

A concise comparison of the reviewed studies is presented in Table 1:

Paper	Methodology	Application Area	Key Outcome
[1]	MediaPipe + LSTM/MLP	PLC-based gesture control	92% recognition accuracy
[2]	Kinect-based HMI	Usability evaluation	Identified human-factor barriers
[3]	SlowFast + YOLOv9	Action-aware PPE monitoring	Recall up to 93%
[4]	3D CNN + ConvLSTM	Real-time cobot interaction	Robust real-time performance

IV. DISCUSSION

A. *Comparison of Strengths*

1) Precision and Technical Accuracy: The PLC-based gesture control approach demonstrates high technical reliability, achieving a 92% recognition accuracy by utilizing a combination of MediaPipe for landmark detection and LSTM models [1]. Similarly, the PPE violation detection system provides a significant 23% improvement in F1-score over traditional methods by making safety checks task-specific only verifying equipment like welding helmets when the worker is actively welding [3].
2) Robust Pre-processing: The cobot interaction method excels in environmental adaptation by isolating movement between grayscale frames to effectively remove background noise, allowing the model to focus purely on the gesture [4].
3) User-Centered Design: The factory loading station study is the only one to leverage year-long workshops with industry experts to ensure the gesture vocabulary is both robust and intuitive for real-world workers [2].

B. Comparison of Weaknesses

1) Human Factors and Social Barriers: User centered evaluations identify a major psychological barrier: the lack of physical feedback (tactile buttons) causes operator frustration and a perceived lack of control [2]. Furthermore, large-scale gesturing was often socially perceived as "fooling around" rather than professional work.
2) Environmental and Visual Limitations: Safety monitoring frameworks struggle with Field of View (FOV) issues; in large manufacturing plants, small PPE items like gloves or glasses are difficult to detect from a distance on a standard 2D feed [3].
3) Algorithmic Confusion: Real-time gesture recognition systems report lower overall accuracy (84%) and note that the model often confuses similar gestures, such as "opening" versus "closing" a hand, due to their subtle visual differences [4].

C. Practical Usability

1) Direct Machine Control: Landmark-based gesture systems are highly usable for direct actuation, having established a practical communication link via the MC protocol to control Mitsubishi iQ-R PLCs and servo drives [1].
2) Safety Auditing: Action aware safety monitoring is best suited for automated safety monitoring and root cause analysis, achieving a 93% recall rate that ensures nearly all violations are flagged for review by safety officers [3].
3) Collaborative Robotics: Real time gesture interfaces are designed for Human-Robot Interaction (HRI), providing a friendly interface where operators can even add and retrain new gestures in real-time to suit specific tasks [4].
4) Operational Ergonomics: Gesture-based control enhances usability by allowing operators to control machinery from a safe distance, eliminating the need for constant movement between the workpiece and a fixed control panel [2].

V. RESEARCH GAPS

1) Lack of 3D and Depth Information for Complex Safety Scenarios: A significant limitation in current surveillance-based safety monitoring is the reliance on 2D RGB video feeds, which lack the depth information required for a comprehensive understanding of the working environment [3]. This makes it difficult to accurately calculate the distance between workers and hazardous machinery, such as the safe distance a worker must maintain from a moving crane.
2) Absence of Tactile and Physical Feedback: A critical shortcoming of touchless interaction is the lack of tactile feedback, which significantly impacts an operator's "feeling of control" and professional self-efficacy. Operators used to direct physical feedback from buttons find the transition to non-physical controls "strange," suggesting a need for alternative feedback methods to compensate for this absence [2].
3) Insufficient Realism and Multi-Actor Complexity in Datasets: Many benchmark datasets are recorded in controlled laboratory settings with single actors performing single actions. These fail to capture the dynamic and chaotic nature of real-world industrial environments where multiple individuals perform different activities simultaneously in the same camera view [4].
4) Small Object Detection in Large Fields of View (FOV): In large manufacturing plants, cameras are typically positioned to cover the maximum possible area. This creates a large FOV where small but essential safety items—such as gloves, safety glasses, or specific finger gestures—become extremely difficult for computer vision models to detect accurately from a distance [3].
5) Social Acceptability and Workplace Norms: There is a gap in understanding how to make gesture-based interaction socially acceptable in a professional context. Currently, performing elaborate body gestures is often perceived by workers as "fooling around" or like "games," which conflicts with social norms and the perceived professionalism of the industrial workplace [2].
6) Limited Functionality in Gesture-to-Machine Mapping: Current research is often limited to singular inputs and basic on/off or directional commands. There is a noted need for future work to accommodate multi-input gestures and the ability to configure analog values (like specific speeds or temperatures) through complex finger tracking [1].

VI. CONCLUSION

Current research highlights a transformative shift toward touchless industrial interaction, demonstrating the efficacy of hybrid deep learning architectures—specifically LSTM, 3D CNN, and SlowFast networks—in achieving high-precision recognition for PLC and cobot control [1][3][4]. Key findings indicate that landmark-based models can achieve 92% accuracy in machine actuation, while action-aware frameworks improve safety compliance by reducing false alarms through task-specific PPE verification, yielding recall rates up to 93% [1][3]. However, field evaluations reveal significant barriers to adoption, as the absence of tactile feedback correlates with operator frustration and a perceived lack of professional self-efficacy [2].

To address these limitations, future research must prioritize the development of alternative feedback mechanisms to restore a sense of control without physical contact. Furthermore, expanding system functionality to support multi-input gestures and analog value configurations is essential for complex operations [1]. Finally, transitioning toward depth-aware monitoring and utilizing more realistic, multi-actor datasets will be critical to ensuring environmental robustness within the chaotic field of view typical of real-world industrial shop floors [3][4].

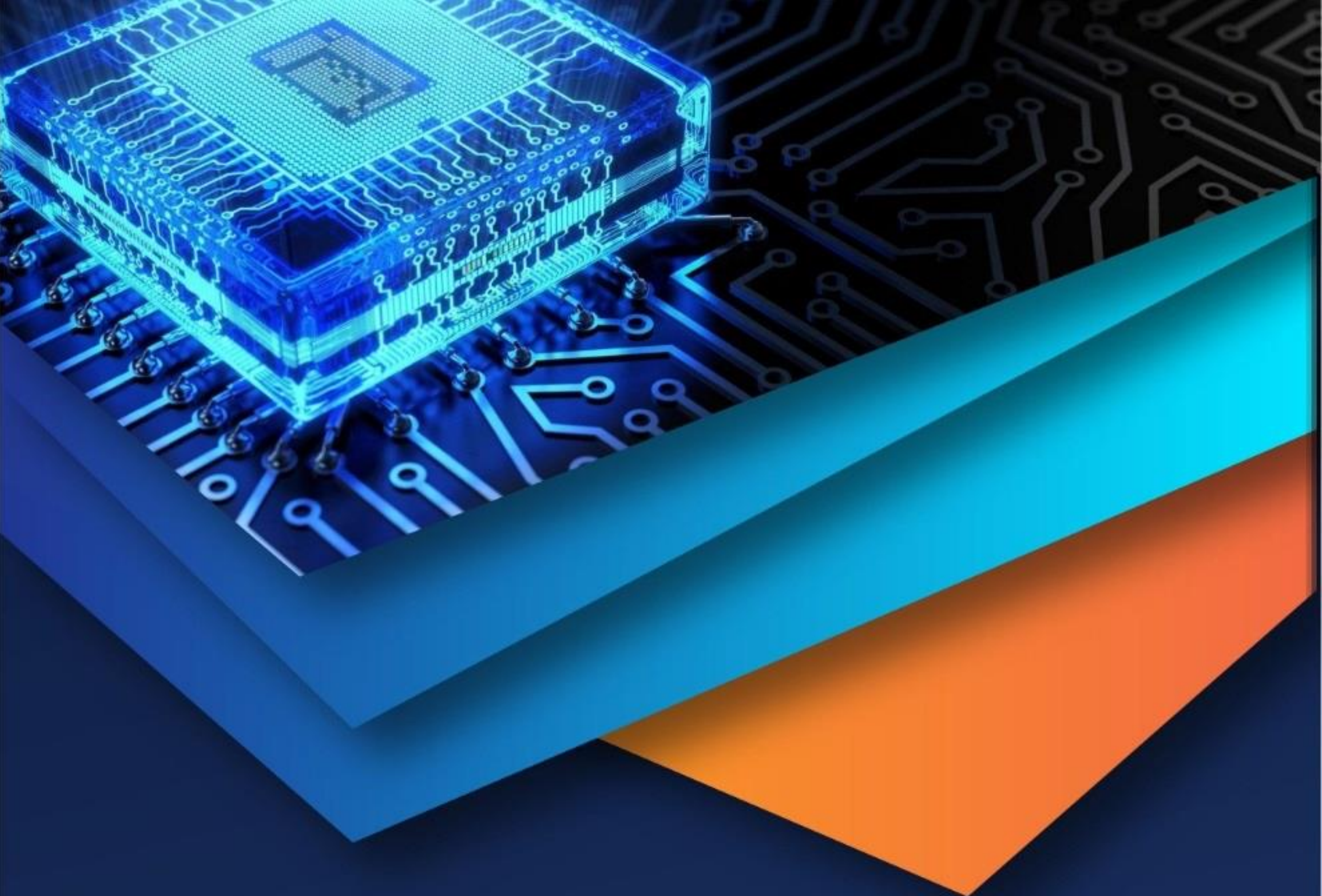
REFERENCES

[1] Yashika, A. Patra, and N. S. Das, "Human-Machine Interaction in Industrial Automation: Gesture-Based PLC Control," in Proc. 15th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), 2024, pp. 1–13.

[2] T. Heimonen, J. Hakulinen, M. Turunen, J. P. P. Jokinen, T. Keskinen, and R. Raisamo, "Designing Gesture-Based Control for Factory Automation," in Human-Computer Interaction – INTERACT 2013, Part II, LNCS 8118, P. Kotzé et al., Eds. Heidelberg, Germany: Springer, 2013, pp. 202–209.

[3] S. N. Reddy, V. Kurrey, M. Nagar, and G. R. Gupta, "Action Recognition based Industrial Safety Violation Detection," in Proc. 8th Int. Conf. Data Sci. Manage. Data (CODS-COMAD), Jodhpur, India, Dec. 2024, pp. 1–10. doi: 10.1145/3703323.3703722.

[4] W. Dumoulin, N. Thiry, and R. Slama, "Real Time Hand Gesture Recognition in Industry," in Proc. 3rd Int. Conf. Video, Signal Image Process. (VSIP), Wuhan, China, Nov. 2021, pp. 1–11. doi: 10.1145/3503961.3503962



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY