



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Hand Gesture Controlled Presentation Using Machine Learning

S. REVATHI¹, A SHAJAN², K SHIVAPRASAD³, V VINOTH KANNA⁴,

¹ Ms. S Revathi, Assistant Professor, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, India.

² A Shajaan, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

³ K Shivaprasad, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India.

⁴ V Vinoth Kanna, CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract:

This project presents a gesture-controlled smart presentation system designed to enhance user interaction during slide-based presentations. By integrating computer vision and hand-tracking technologies using OpenCV, CVZone, and Mediapipe, the system enables users to navigate PowerPoint slides, annotate, zoom, undo, and highlight content via natural hand gestures. PowerPoint (.pptx) files are automatically converted into PDF and then to PNG images using LibreOffice and pdf2image, allowing seamless slide integration. A semi-transparent webcam overlay enhances presenter visibility, while gesture mapping allows real-time control—such as using specific finger patterns to switch slides, zoom, or activate a laser pointer. The system is built in Python and features a clean interface with minimal dependencies, optimized for real-time performance on standard webcams. This hands-free, intuitive approach to presentations is especially useful for educators, professionals, and remote presenters, providing an interactive and immersive experience without the need for clickers or manual input.

[1] INTRODUCTION

In today's digital age, enhancing audience engagement during presentations is crucial. Traditional input devices like remotes or keyboards can hinder smooth interaction. This project introduces a gesture-controlled presentation system that leverages computer vision and hand-tracking to offer an intuitive, hands-free experience. Built using Python, OpenCV, and CVZone, the system allows users to control PowerPoint slides through natural hand gestures—enabling navigation, zoom, annotations, and a laser pointer. By converting .pptx files to image slides, the application ensures compatibility and performance. This innovation is ideal for educators, professionals, and presenters seeking seamless, interactive, and dynamic control over their presentations.

1.1 General Introduction

This project presents a smart, gesture-controlled presentation system designed to enhance interactivity and convenience. Utilizing Python, OpenCV, and CVZone, it enables users to control slide navigation, zooming, annotations, and a laser pointer using simple hand gestures. By converting PowerPoint files into images, the system ensures smooth operation and compatibility. It also integrates a semi-

transparent webcam feed for added engagement. Ideal for educators and professionals, this tool offers a modern, hands-free solution for delivering impactful and dynamic presentations.

1.2 Problem Statement

Traditional presentation tools rely heavily on physical input devices like keyboards, mice, or clickers, which can disrupt the flow and reduce the interactivity of presentations. Presenters often face difficulty managing slides, annotating content, or emphasizing points without interrupting their delivery. These limitations hinder smooth communication, especially in educational or professional environments. This project addresses the problem by introducing a gesture-controlled system that allows seamless navigation, zooming, annotation, and highlighting using hand gestures. It eliminates the need for physical contact, offering a more natural, intuitive, and engaging presentation experience through real-time computer vision and smart slide management.

1.3 Algorithm

1. Hand Detection Model
2. Gesture detection
3. hand Detection & tracking Model

[2] LITERATURE SURVEY

A comprehensive framework for hand gesture recognition using hybrid-metaheuristic algorithms and deep learning models

Hassan Mohyuddin's 2023 study, *A Comprehensive Framework for Hand Gesture Recognition Using Hybrid-Metaheuristic Algorithms and Deep Learning Models*, presents an advanced approach to improving gesture recognition systems. The framework integrates hybrid-metaheuristic optimization techniques with deep learning to enhance the accuracy and responsiveness of gesture classification. By combining different optimization algorithms, the system avoids common issues like getting stuck in local minima and slow convergence. This hybrid approach is used to fine-tune

parameters such as feature selection and classification thresholds. Deep learning models are employed to interpret complex hand gestures under varying conditions—including changes in hand shape, orientation, lighting, and occlusion—making the system robust in real-world scenarios. The study shows that this combination significantly improves recognition accuracy and processing speed compared to traditional methods. The insights and techniques from this paper directly support the goals of the current project, which aims to create a real-time, accurate, and user-friendly hand gesture-controlled presentation system for interactive environments.

An improved hand gesture recognition system using keypoints and hand bounding boxes

Tuan Linh Dang's 2022 study, **An Improved Hand Gesture Recognition System Using Key points and Hand Bounding Boxes**, presents a refined approach that significantly enhances gesture recognition performance. The system uses bounding boxes to isolate the hand region from the background, effectively reducing noise and improving detection reliability, especially in dynamic or cluttered environments. Key point detection further refines the model by identifying critical hand landmarks such as finger joints and the palm center, allowing for a highly accurate representation of various hand gestures. This precise modeling enables the system to differentiate between even complex gestures with high confidence. Dang's framework also addresses real-world challenges such as partial hand occlusion and inconsistent lighting, making it suitable for real-time, practical applications. Compared to older recognition techniques, the study demonstrates superior accuracy and adaptability. These innovations are closely aligned with the goals of the current project, which seeks to implement a gesture-controlled presentation system that is both accurate and responsive in real-time scenarios.

Study on Hand Gesture Recognition by using Machine Learning

The study *Study on Hand Gesture Recognition by Using Machine Learning* by Mohanarathinam investigates the effectiveness of machine learning techniques in enhancing hand gesture recognition systems. The paper explores various models to identify those best suited for real-time gesture detection and classification, addressing limitations of traditional methods—such as reliance on static gesture datasets and vulnerability to environmental changes like lighting fluctuations and background interference. By leveraging adaptive learning, the study finds that machine learning algorithms, particularly convolutional neural networks (CNNs), excel in recognizing a wide variety of hand gestures with high accuracy and efficiency. CNNs offer significant advantages in processing speed and adaptability, making them suitable for dynamic, real-world scenarios. The research demonstrates that machine learning reduces false positives and increases robustness, even in complex conditions. These findings are directly applicable to the hand gesture-controlled presentation system, guiding the integration of machine learning for reliable, hands-free operation in classrooms, offices, and other professional environments where intuitive user interaction is critical.

Slideshow Presentation Control through Hand Pattern Recognition using web camera

The study *Slideshow Presentation Control through Hand Pattern Recognition Using Web Camera* by Del Mundo explores the implementation of a cost-effective, hands-free slideshow control system using hand pattern recognition. Relying solely on a standard web camera, the system applies image processing techniques to detect and interpret simple gestures—such as swipe and tap—for navigating presentation slides. This method eliminates the need for physical remotes or keyboards, offering an accessible and practical alternative for controlling slideshows. The study effectively addresses real-world challenges, including variations in hand size, gesture positioning, and lighting, by employing adaptive

algorithms that improve recognition across diverse environments. However, Del Mundo also notes limitations such as reduced accuracy with complex gestures and occasional lag in response time—common drawbacks in camera-based systems. This research serves as a strong foundation for the current hand gesture-controlled presentation project, which aims to overcome these limitations using advanced machine learning models. The goal is to improve gesture recognition accuracy, speed, and flexibility for a seamless, hands-free presentation experience.

Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network

This study presents a real-time hand gesture recognition framework leveraging fine-tuned convolutional neural networks (CNNs), tailored for applications requiring rapid and precise gesture detection. Due to the difficulty of acquiring large, labeled datasets for static gestures, the researchers adopt an end-to-end fine-tuning approach using pre-trained CNN models. This strategy allows the models to adapt effectively to small gesture datasets, preserving high accuracy without the need for extensive training data. To further enhance recognition performance, the system uses score-level fusion techniques that combine multiple model predictions. This approach increases robustness against variations in hand shape, orientation, and environmental backgrounds. The framework's effectiveness is validated through leave-one-subject-out and standard cross-validation methods, with particularly strong results in American Sign Language (ASL) classification. For the hand gesture-controlled presentation system project, this study highlights the potential of using fine-tuned CNNs to deliver accurate, real-time gesture detection, even in data-scarce scenarios, thereby enabling seamless, intuitive control during live presentations.

3. EXISTING METHODOLOGY

Traditional presentation control systems rely on physical devices like keyboards, remotes, and clickers, which limit the presenter's movement and flexibility. These devices require the user to stay close to the computer or rely on battery-

powered remotes, potentially interrupting presentation flow. While touch-based and voice-activated controls offer alternatives, they have limitations such as restricted range, noise sensitivity, and accessibility challenges, particularly for users with mobility impairments. Additionally, these systems do not fully support dynamic, real-time interaction, making them less ideal for engaging presentations. These limitations highlight the need for a hands-free, gesture-based control solution to enhance user experience.

3.1. DISADVANTAGES

Traditional presentation control methods, such as keyboards, remotes, and clickers, often restrict the presenter's movement and require staying near the computer, disrupting the natural flow of presentations. These devices may also face issues like battery failure or signal interference. Touch and voice controls offer some flexibility but are limited by environmental factors like noise and lighting, and may not be suitable for users with disabilities. Furthermore, these systems lack the dynamic interaction and responsiveness needed for engaging, real-time presentations. Such limitations highlight the need for an intuitive, hands-free solution like gesture-based control to improve user experience and accessibility.

4. PROPOSED METHODOLOGY

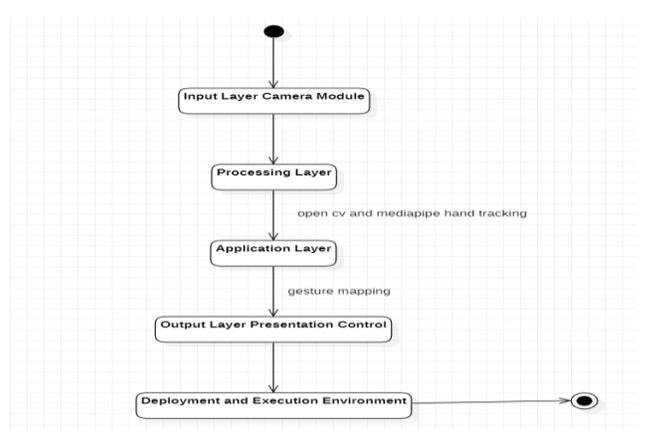
The proposed hand gesture-controlled presentation system introduces a user-friendly, hands-free alternative to conventional presentation tools. Instead of relying on remotes, keyboards, or touch-based controls, it utilizes computer vision and machine learning to detect and interpret hand gestures in real-time. With Mediapipe's Palm Detection and Hand Landmark models, the system achieves precise gesture recognition—such as swipes and taps—with up to 95.7% accuracy and performance between 30 to 60 frames per second. This dynamic approach enhances mobility and accessibility, allowing presenters to interact naturally and engage audiences more effectively. It is especially beneficial in professional, educational, and interactive environments, supporting users with mobility

challenges and promoting a more seamless presentation experience.

4.1. ADVANTAGES

The hand gesture-controlled presentation system offers enhanced mobility, allowing presenters to move freely without physical devices. It improves accessibility for users with disabilities and enables real-time, seamless interaction. The system increases audience engagement, operates efficiently with high accuracy, and reduces reliance on battery-powered tools, making it ideal for professional, educational, and dynamic presentation environments.

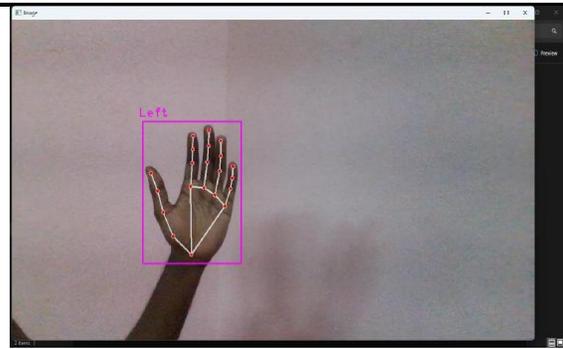
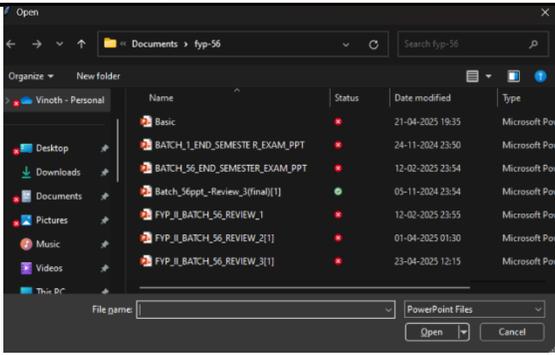
4.2 BLOCK DIAGRAM



5. RESULTS

5.1. FILE SELECTION

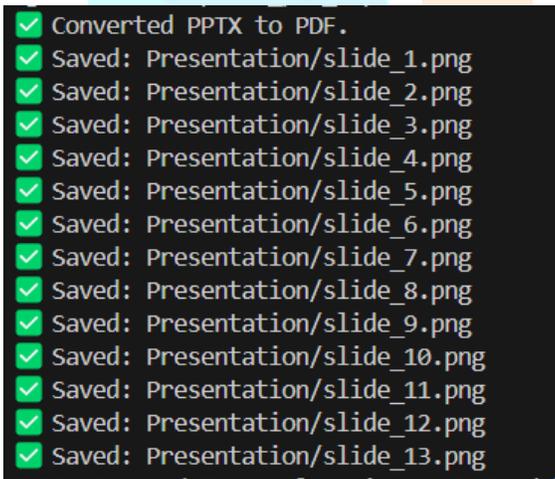
In this project, file selection is managed through a Tkinter-based file dialog, making it easy for users to browse and select a PowerPoint (.pptx) file from their local system. After selection, the chosen file undergoes a conversion process to extract individual slides, which are later used in the gesture-controlled presentation system. This method ensures a smooth and intuitive experience for users, allowing them to quickly load and present their desired content.



5.2. IMAGE CONVERSION

After the user selects a PowerPoint file, it is initially transformed into a PDF format using LibreOffice. Following that, the pdf2image library is used to convert the PDF pages into individual PNG images, saved in a specific "Presentation" directory. This dual conversion process guarantees that the slides are preserved with high quality, ensuring smooth loading and clear visuals for the gesture-driven presentation system.

Formula:



5.3 HAND DETECTION

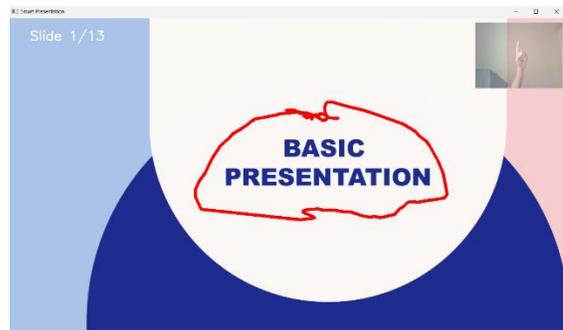
Hand detection in this project is handled through the HandDetector module provided by the cvzone library, which uses OpenCV underneath. It captures the user's hand from the webcam feed and monitors crucial landmarks like fingertips. The system interprets finger patterns to recognize gestures, allowing users to control the presentation with commands such as zooming, annotating, navigating slides, and activating a laser pointer — all through natural, touchless interaction.

5.5 GESTURE DETECTION

The project's gesture detection works by examining the finger positions identified by the HandDetector module. Distinct finger configurations are mapped to actions such as zoom adjustments, slide transitions, annotation drawing, undoing strokes, and laser pointer activation. A gesture cooldown timer ensures that commands are not repeatedly triggered by a single gesture, resulting in stable and efficient presentation control through simple, intuitive hand movements.

5.4. GESTURE ACTIONS

Gesture actions are precisely assigned to different finger patterns in the project. Three-finger gestures allow users to zoom in and out, while one-finger gestures control slide navigation. Separate hand poses initiate annotation drawing, undo recent marks, jump to the first or last slide, and enable the laser pointer. Every gesture is thoughtfully designed to correspond with a particular action, making the entire system user-friendly and responsive during presentations.



6. CONCLUSION

The hand gesture-controlled presentation system marks a significant step forward in intuitive, hands-free human-computer interaction. By integrating Mediapipe for real-time hand tracking, OpenCV for image processing, and CNN-based models for gesture recognition, it offers a seamless, responsive experience with up to 95.7% accuracy. Presenters can control slides naturally without remotes or keyboards, improving flow and enhancing audience engagement. The web-based interface built with Streamlit offers instant visual feedback, making gesture control user-friendly and accessible. This system particularly benefits users with limited mobility, promoting inclusivity in tech applications. Its real-time adaptability makes it ideal for educational and professional settings. Looking ahead, enhancements like expanded gesture sets, improved lighting resilience, and voice integration could broaden its functionality. Moreover, the technology has potential in areas such as smart homes, gaming, and AR/VR interfaces. This project showcases how natural user interfaces can revolutionize interaction, making it more immersive, efficient, and inclusive across diverse environments and industries..

7. REFERENCES

- [1] H. Mohyuddin, "A Comprehensive Framework for Hand Gesture Recognition Using Hybrid-Metaheuristic Algorithms and Deep Learning Models," ScienceDirect, 2023.
- [2] T. L. Dang, "An Improved Hand Gesture Recognition System Using Keypoints and Hand Bounding Boxes," ScienceDirect, 2022.
- [3] A. Mohanarathinam, "Study on Hand Gesture Recognition by Using Machine Learning," IEEE Xplore, 2020.
- [4] J. S. Del Mundo, "Slideshow Presentation Control Through Hand Pattern Recognition Using Web Camera," IEEE Xplore, 2017.
- [5] A. J. Prakash, P. Pławiak, and S. Samantray, "Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network," MDPI Sensors, 2022.

[6] A. Jaramillo-Yáñez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review," MDPI Sensors, 2020.

[7] J. Zhang and J. Li, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," IEEE Xplore, 2021.

[8] S. Pearson, D. Bochtis, et al., "An Integrated Real-Time Hand Gesture Recognition Framework for Human-Robot Interaction in Agriculture," MDPI Applied Sciences, 2022.

