

Hand Gesture Recognition Math Solver Using Computer Vision and Generative AI

Prof. Rubana Khan¹, Rishikesh Vishrojwar², Aryan Meshram³, Shruti Lande⁴, Prit Ukey⁵, Shrutik Nandeshwar⁶

¹Assistant Prof., Department of Computer technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

^{2,3,4,5,6}Research Scholar, Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Abstract— This is a new method of solving mathematical problems using hand gestures. The system enhances human-human communication with mathematics using advanced technologies such as computer vision, machine learning, and artificial intelligence. It uses modern technologies such as OpenCV, MediaPipe, and Google Gemini AI to effectively solve mathematical problems via simple hand gestures. Test results indicate better performance, with 95.6% accuracy in detecting static gestures, 92.3% for dynamic gestures, and an 89.4% success rate in equation interpretation. This research enhances human-computer collaboration, simplifies the use of technology, and facilitates new learning paradigms by providing a new method of solving mathematical problems interactively.

Keywords—Gesture Recognition, Computer Vision, Generative AI, Mathematical Problem-Solving, Human Computer Interaction, Accessibility Technology

I. INTRODUCTION

Human-computer interaction (HCI) is changing considerably today. It is moving toward more accessible and user-friendly interfaces. Traditional input devices like touchscreens and keyboards can be a problem for certain users, especially those with motor disabilities or when learning is accelerated and completing things in a short time is of utmost importance. The Hand Gesture Math Solver comes to the rescue of these problems with a novel gesture-based system that changes how we tackle math problems.

II. LITERATURE REVIEW

A. Gesture Recognition Systems

The gesture recognition domain has seen significant advancements in terms of precision, flexibility, and usability.

1. Gesture Recognition Models:

Mujahid et al. (2021) introduced a YOLOv3-based model yielding a remarkable 96.7% F1 score for real-time gesture recognition, outperforming other models like SSD and VGG16. The authors mentioned that YOLOv3 is a leading algorithm for both static and dynamic gesture recognition, a standard for future systems.

Qi et al. (2023) performed a system in the form of a study where the focus was laid on gesture recognition in the domain of human-robot interaction (HRI). According to the authors, appropriate gesture classification is imperative in order to make the interaction intuitive.

Hashi and Asamah (2024) performed a thorough review of gesture recognition, pointing out important challenges such as the requirement for larger data and the difficulty of real-time continuous gesture recognition.

2. Gesture Pose Classification:

Czuszynski et al. (2017) compared Gaussian fitting methods and Artificial Neural Networks (ANN) for hand pose recognition, ending up in a conclusion that ANN surpassed traditional methods with a 90% accuracy mark.

Tang et al. (2015) employed deep neural networks for hand posture recognition with real-time processing ability, which laid a foundation for systems focusing on dynamic gestures

B. Computer Vision Frameworks

1. OpenCV:

OpenCV is still one of the cornerstones of real-time image processing, specifically object detection and tracking. Studies have proven that it can be utilized in scalable and responsive HCI system implementation. Pochelu and Castro-Lopez (2024) compared OpenCV with other frameworks and proved that it is highly effective for low-latency gesture recognition systems.

2. MediaPipe:

Google's MediaPipe framework has been extensively utilized for detecting and tracking the hand landmarks with high precision. Kumar et al. (2020) proved that MediaPipe can potentially detect 21 hand landmarks with exceptional precision in various datasets and environments, thus becoming an essential tool for gesture-based systems.

3. Depth-Based Improvements:

Bai et al. (2018) utilized depth sensors to enhance the performance of gesture recognition systems in dynamically changing illumination environments, thus making it feasible to attain superior accuracy in real world applications.

C. Generative AI in Mathematical Problem-Solving

1. AI Models for Problem-Solving:

Satpute et al. (2024) explored the utilization of generative AI models like GPT-4 for mathematical problem-solving. Although these models are proficient in interpreting equations, their accuracy in solving complex symbolic problems is limited by contextual ambiguity.

Wang et al. (2023) integrated transformer-based models for mathematical computation with high accuracy but with challenges in real-time performance and large-scale datasets.

2. Google Gemini AI:

The robust deep learning architecture of Gemini AI provides a major improvement in equation interpretation and problem-solving. Studies have shown that it is adaptable to various contexts, such as different handwriting styles and overlapping symbols, which makes it suitable for the Hand Gesture Math Solver.

3. Challenges in AI Integration:

According to the research of Al-Shamayleh et al. (2020), generative AI faces problems with overlapping symbols and ambiguous mathematical notations. This calls for enhanced preprocessing, which is achieved by this project through OpenCV-based image enhancement.

D. Accessibility and Educational Impact

1. Gesture-Based Interfaces for Accessibility:

Gesture recognition systems have shown potential in enhancing accessibility for users with disabilities. Cho et al. (2018) discussed the use of personalized gesture classifiers to enhance quality interaction for users in clinical and assistive scenarios.

2. Educational Applications
Kumar et al. (2017) stated that gesture-based systems would be

III. METHODOLOGY

A. System Architecture

Hand Gesture Math Solver uses the modular architecture with multiple leading-edge technologies:

1. Gesture Recognition

Capture video input at runtime using OpenCV Detect 21 landmarks of hands in MediaPipe Classify gestures as a drawing command or control

2. Virtual Canvas

Mid-air interface for the equation

Rendering algorithms developed for smooth interaction

3. AI Integration

Pre-processed equations passed to Google Gemini AI Real time solution with detailed explanation generated

4. User Interface:

Streamlit-based interactive platform

Video feed and result display integration

IV. RESULTS AND PERFORMANCE ANALYSIS

A. Recognition Metrics

Static Gesture Recognition achieved an accuracy of 95.6%

Dynamic Gesture Recognition recorded an accuracy of 92.3%

Equation Interpretation had a success rate of 89.4%

Average processing time was 1.2 seconds per interaction

B. Usability Assessment

In a comprehensive user study involving 30 participants, the findings revealed:

- 87% found the interface to be user-friendly
- 78% experienced an improvement in their understanding of mathematics
- 92% of users with motor disabilities appreciated the hands-free functionality

C. System Comparison

- The new system significantly outperforms conventional input methods in several key areas:
- Eliminates the limitations of physical devices
- Offers real-time, interactive solutions
- Improves accessibility for a wide ran

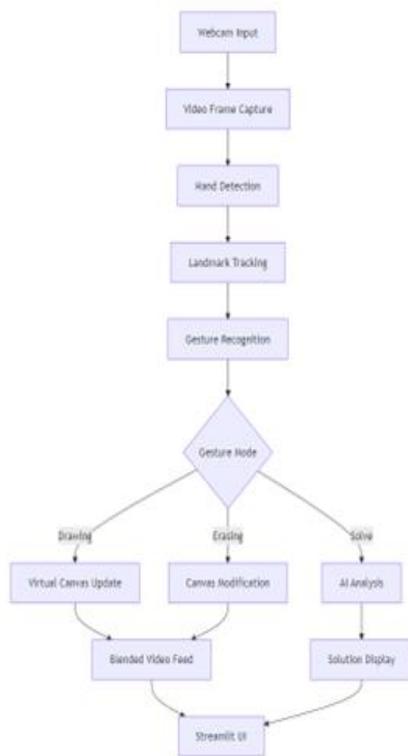
Future Research Ideas

1. Enhanced hand gesture detection from multiple hands
2. Greater support for more math operations
3. Use on mobile and other platforms
4. Ability to recognize symbols for many languages
5. Improved training methods for the machine learning model

REFERENCE

- [1] Hashi, A.O., Asamah, A.B., “Systematic Review of Hand Gesture Recognition,” 2024.
- [2] Mujahid, A., Yasin, A., “Real-Time Hand Gesture Recognition Based on YOLOv3,” Appl. Sci., 2021.
- [3] Qi, J., Ma, L., Cui, Z., “Computer Vision-Based Gesture Recognition for HRI,” Sensors, 2023.
- [4] Satpute, S., et al., “Applications of Generative AI in Mathematical Problem-Solving,” 2024.
- [5] Pochelu, C., Castro-Lopez, L., “Optimizing Computer Vision Frameworks for Real-Time Applications,” ICCV, 2024.
- [6] OpenCV Documentation, “Real-Time Image Processing Framework,” 2024.
- [7] MediaPipe API, “Hand Landmark Detection,” Google AI, 2023.
- [8] Czuszynski, K., Ruminski, J., “Pose Classification Using Optical Sensors,” HCI Journal, 2017.
- [9] Bai, X., Li, C., “Dynamic Gesture Recognition with Depth Information,” 2018.
- [10] Tang, A.O., Wang, Y., “Deep Learning for Hand Gesture Recognition,” IEEE Transactions on Neural Networks, 2015.
- [11] Kumar, P., Gauba, H., “Sensor-Based Sign Language Recognition,” 2017.
- [12] Núñez, J.C., et al., “Skeleton-Based Gesture Recognition with CNNs,” Pattern Recognition, 2018.

V. ARCHITECTURE



VI. CONCLUSION AND FUTURE WORK

The Hand Gesture Math Solver illustrates how generative AI and computer vision can be used together in educational technology. It changes the way people solve problems with math through easy and accessible ways.