International Journal of Advance Research in Science and Engineering
Volume. No. 15, Issue No. 03, March, 2026
ISSN: 2319-8354

"GESTURE RECOGNITION USING DEEP LEARNING"

Guide : **Dr.J.Nagaraju** ,M.Tech,Ph.D, Associate professor
Department of IT
Tirumala Engineering College
Narasaraopet,522601

Syed Maheen Sulthana
Department of IT
Tirumala Engineering College
Narasaraopet,522601
maheensulthanasyed@gmail.com

Thiyyagura Sathvika
Department of IT
Tirumala Engineering College
Narasaraopet,522601
thiyyagurasathvika@gmail.com

Kokku Sri Harshitha
Department of IT
Tirumala Engineering College
Narasaraopet,522601
22ne1a1271@gmail.com

Muraboina Alekya
Department of IT
Tirumala Engineering College
Narasaraopet,522601
alekyamuraboina@gmail.com

(Academic Year : 2022 – 2026)

Abstract

Gesture recognition is an emerging field in computer vision and artificial intelligence that focuses on enabling machines to interpret human hand movements, and this project presents a real-time hand gesture recognition system designed to assist communication for individuals who are deaf or hearing-impaired. The system captures live video through a webcam and uses MediaPipe to detect and track 21 hand landmarks, which are converted into numerical feature vectors for training a machine learning classification model. It recognizes commonly used gestures such as Hi, Yes, No, Thanks, and Please, and follows a complete pipeline including data collection, preprocessing, feature extraction, model training, evaluation, and real-time prediction. The system is efficient, lightweight, and runs on standard hardware without requiring specialized equipment. Experimental results show high accuracy and stable real-time performance under varying lighting conditions and backgrounds, demonstrating robustness and adaptability. Overall, the solution is cost-effective, scalable, and user-friendly, contributing to assistive technology development and providing a foundation for future improvements such as gesture-to-speech conversion, mobile deployment, and recognition of more complex gestures.

I.Introduction

Communication is a fundamental aspect of human life, enabling individuals to express ideas, emotions, and information effectively. While spoken and written languages are widely used, a significant portion of the global population relies on sign language as their primary mode of communication. Sign language uses hand gestures, facial expressions, and body movements to convey meaning, making it an essential tool for individuals who are deaf or hearing-impaired.

Despite its importance, sign language is not universally understood, which creates a communication gap between sign language users and the general population. This gap can lead to difficulties in everyday interactions such as education, healthcare, employment, and social communication. For example, a hearing-impaired person may face challenges when trying to communicate with teachers, doctors, or colleagues who do not understand sign language. As a result, there is a growing need for intelligent systems that can translate gestures into a form that can be easily understood by others.

With the rapid advancement of technology, artificial intelligence (AI) has emerged as a powerful tool to address such challenges. In particular, the fields of deep learning and computer vision have shown remarkable progress in enabling machines to analyze and interpret visual data. These technologies allow computers to recognize patterns in images and videos, making them suitable for applications such as facial recognition, object detection, and gesture recognition.

Gesture recognition is a subfield of computer vision that focuses on identifying and interpreting human gestures through digital systems. It plays a crucial role in human-computer interaction by enabling more natural and intuitive communication between humans and machines. Instead of relying on traditional input devices such as keyboards and mice, users can interact with systems using simple hand movements.

Early gesture recognition systems were based on traditional image processing techniques and required manual feature extraction. These systems used methods such as edge detection, contour analysis, and shape matching to identify gestures. However, they had several limitations, including sensitivity to lighting conditions, background noise, and variations in hand position. Additionally, they required controlled environments and were not suitable for real-time applications.

The introduction of machine learning techniques improved gesture recognition systems by enabling automatic classification of gestures based on extracted features. Algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees were widely used for this purpose. While these methods improved performance, they still relied on handcrafted features, which limited their scalability and adaptability.

In recent years, deep learning has revolutionized the field of gesture recognition by enabling end-to-end learning from raw data. Convolutional Neural Networks (CNNs) can automatically extract relevant features from images, eliminating the need for manual feature engineering. Similarly, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks can capture temporal patterns in dynamic gestures.

Another significant advancement in this field is the development of frameworks such as MediaPipe, which provides efficient and real-time solutions for hand tracking and landmark detection. MediaPipe identifies 21 key points on the hand, representing the positions of fingers and joints. These landmarks provide a structured representation of hand gestures, making it easier to classify them using machine learning models.

This project leverages these modern technologies to develop a real-time gesture recognition system. The system captures live video input using a webcam and processes it using OpenCV for image handling. MediaPipe is used to detect hand landmarks, and a machine learning model is trained to classify gestures based on the extracted features.

The proposed system focuses on recognizing a set of commonly used gestures such as Hi, Yes, No, Thanks, and Please. These gestures are chosen because they are frequently used in basic communication and can serve as a foundation for more complex systems. The system is designed to operate in real time, providing immediate feedback to the user.

One of the key advantages of this approach is that it does not require specialized hardware such as sensors or gloves. Instead, it relies on a standard webcam, making it cost-effective and accessible. Additionally, the use of lightweight models ensures that the system can run efficiently on devices with limited computational resources.

The primary objective of this project is to develop an accurate, efficient, and user-friendly gesture recognition system that can bridge the communication gap between sign language users and the general population. By converting gestures into text, the system enables seamless communication and enhances accessibility for individuals with hearing impairments.

Furthermore, this project contributes to the broader field of assistive technologies by demonstrating how AI can be used to solve real-world problems. It also provides a foundation for future developments, such as integrating speech output, expanding gesture vocabulary, and deploying the system on mobile platforms.

In conclusion, gesture recognition using deep learning represents a promising solution for improving human-computer interaction and enhancing communication accessibility. The proposed system aims to utilize the power of AI to create a practical and impactful solution that benefits society.

II.Literature Survey

Gesture recognition has been widely explored in the field of computer vision and deep learning, with various techniques evolving from traditional methods to advanced neural network-based approaches. Early systems primarily relied on handcrafted feature extraction techniques such as edge detection, contour analysis, and shape-based methods. These approaches required controlled environments and were not suitable for real-time applications, limiting their practical usability.

One of the significant contributions in gesture recognition was presented by Assiri and Selim (2025), who proposed an ensemble-based deep learning model for recognizing gestures of hearing-impaired individuals. Their approach combined SE-CapsNet for feature extraction with models such as BiLSTM, BiGRU, and Variational Autoencoders (VAE) for classification. The study achieved a high accuracy of 98.72%, demonstrating the effectiveness of ensemble learning. However, the complexity of multiple models increased computational requirements.

Hossain et al. (2023) introduced a gesture recognition system using MediaPipe for hand landmark detection along with Convolutional Neural Networks (CNN) for classification. By extracting key hand landmarks, the system improved recognition accuracy and handled variations in hand orientation effectively. This approach also supported real-time applications but still required careful tuning for optimal performance.

Later, Vyshnavi et al. (2024) proposed "GestureSense," a deep learning-based gesture translation system that utilized advanced CNN architectures such as VGG-16 and ResNet-50. Their ensemble approach improved classification accuracy significantly by combining the strengths of both models. However, the system demanded high computational resources, making it less suitable for low-cost or real-time deployment environments.

Ravinder et al. (2023) focused on developing a lightweight CNN model aimed at improving efficiency and reducing computational complexity. While their model performed well for static gesture recognition, it lacked the capability to handle dynamic gestures, which limited its application in real-time communication systems.

Similarly, Shinde et al. (2024) proposed a deep learning-based system using Long Short-Term Memory (LSTM) networks along with MediaPipe Holistic for gesture recognition. Their approach effectively captured temporal dependencies in gestures, making it suitable for dynamic gesture recognition. However, the system required large datasets and longer training time, which increased implementation complexity.

Barbhuiya et al. (2023) introduced a deep ensemble neural network approach using VGG-16 combined with self-attention mechanisms. This method enhanced feature extraction and improved classification accuracy but resulted in increased computational cost due to the use of multiple deep learning models.

Izzalhaqqi (2023) proposed a hybrid CNN-LSTM model to capture both spatial and temporal features of gestures. This approach demonstrated strong performance in recognizing dynamic gestures, but it required extensive preprocessing and parameter tuning, making it less efficient for real-time applications.

Ramadan and Abd-Alsabour (2024) presented a gesture-based control system using computer vision techniques. Their system enabled users to interact with devices through gestures, such as controlling a laptop. While innovative, the focus was more on device interaction rather than improving communication for hearing-impaired users.

Rajalakshmi et al. (2022) developed a hybrid neural network combining 3D CNN and Bi-LSTM for recognizing both static and dynamic gestures. The system achieved improved accuracy but required high computational power and large-scale datasets, limiting its accessibility for general users.

Overall, recent advancements in gesture recognition emphasize the importance of deep learning, real-time processing, and hybrid approaches. While these methods significantly improve accuracy and robustness, they often come with challenges such as high computational cost, large data requirements, and complexity. Therefore, there is a need for a balanced system that provides high accuracy, real-time performance, and cost-effectiveness, which is addressed in the proposed system.

III.Methodology

A. Existing Methodology

Traditional gesture recognition systems primarily rely on basic computer vision techniques and classical machine learning models to identify hand gestures. These systems analyze visual features such as edges, contours, and shapes extracted from images to recognize gestures. Although these methods were effective in early research, they struggle to handle real-world complexities such as varying lighting conditions, backgrounds, and hand orientations.

Rule-based approaches were initially used, where predefined conditions were applied to detect gestures. For example, specific hand shapes or positions were mapped to particular gestures. While this approach is simple and easy to implement, it lacks flexibility and cannot adapt to variations in hand movements. Minor changes in gesture appearance can lead to incorrect recognition.

In addition to rule-based systems, traditional models such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees were widely used. These models are trained on extracted features from gesture images to classify different hand movements. However, they rely heavily on manual feature extraction techniques like Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and Local Binary Patterns (LBP). This manual process requires domain expertise and often fails to capture complex patterns in gesture data.

Another common limitation of existing systems is their dependence on static images rather than real-time video streams. As a result, these systems are not suitable for real-time applications where continuous gesture recognition is required. They also perform poorly in dynamic environments due to sensitivity to lighting variations, background noise, and camera angles.

Furthermore, traditional systems often lack temporal analysis, making them unable to recognize dynamic gestures that involve motion over time. This limits their capability to understand continuous sign language communication. Many systems also produce lower accuracy levels and higher misclassification rates, which can lead to incorrect interpretations.

Existing approaches also suffer from poor scalability. Adding new gestures requires collecting new datasets, redesigning features, and retraining models, making the system complex and time-consuming. Additionally, some algorithms are computationally inefficient during prediction, leading to delays in response time.

Overall, existing methodologies face several challenges, including dependence on manual feature extraction, lack of real-time processing, sensitivity

International Journal of Advance Research in Science and Engineering
Volume. No. 15, Issue No. 03, March, 2026
ISSN: 2319-8354

to environmental conditions, limited accuracy, and poor scalability. These limitations highlight the need for a more efficient, adaptive, and real-time gesture recognition system.

B. Proposed Methodology

The proposed gesture recognition system follows a structured pipeline consisting of data collection, preprocessing, feature extraction, model training, and real-time prediction. The main objective is to accurately recognize hand gestures in real time using an efficient and cost-effective approach.

The process begins with data collection, where images of hand gestures are captured using a standard webcam. The system collects data for five predefined gestures: Hi, Thanks, Yes, No, and Please. To improve model performance, data is collected under different lighting conditions, angles, and hand positions.

Next, data preprocessing is performed using OpenCV and MediaPipe. The webcam captures live frames, and MediaPipe is used to detect the hand region and extract 21 key landmark points. These landmarks represent the spatial structure of the hand and are essential for accurate gesture recognition.

The extracted landmarks are then converted into a feature vector, consisting of x and y coordinates of each landmark. This results in a structured input format that can be used for training the model. Unlike traditional methods, this approach eliminates the need for manual feature engineering by relatively low computational cost. The dataset is split into training and testing sets to evaluate model performance effectively.

Once the model is trained, it is used for real-time prediction. The system continuously captures video frames from the webcam, processes each frame, extracts landmarks, and predicts the corresponding gesture instantly. The predicted gesture is displayed as text on the screen, providing real-time feedback to the user. The predicted gesture is displayed as text on the screen, providing real-time feedback to the user. , recent advancements in gesture recognition emphasize the importance of deep learning, real-time processing, and hybrid approaches. While these methods significantly improve accuracy and robustness, they often come with challenges such as high computational cost, large data requirements, and complexity. Therefore, there is a need for a balanced system that provides high accuracy, real-time performance, and cost-effectiveness, which is addressed in the proposed system. This results in a structured input format that can be used for training the model. Unlike traditional methods, this approach eliminates the need for manual feature engineering by automatically extracting

automatically extracting

In the model training phase, a Random Forest classifier is used to learn patterns from the extracted features. Random Forest is chosen due to its ability to handle complex data, reduce overfitting, and provide high accuracy with

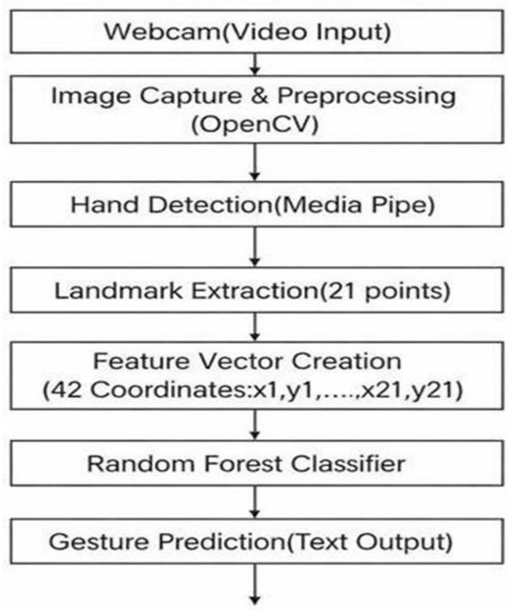
The system also includes confidence-based filtering, where predictions are displayed only when the confidence score exceeds a certain threshold. This reduces incorrect predictions and improves reliability. If no hand is detected, the system displays an appropriate message to the user.

Additionally, the trained model is saved using Pickle, allowing it to be reused without retraining. This improves efficiency and makes the system more practical for deployment.

The proposed methodology offers several advantages, including real-time processing, high accuracy, automated feature extraction, cost-effectiveness, and scalability. It performs well under different environmental conditions and can be easily extended to include more gestures in the future.

Overall, the proposed system provides an efficient, reliable, and user-friendly solution for gesture recognition, addressing the limitations of existing methodologies and improving communication for hearing-impaired individuals.

FIGURE : proposed methodology architecture



The proposed system follows a comprehensive pipeline consisting of several stages:

1. Data Collection

Gesture data is collected using a webcam. Multiple samples are recorded for each gesture to ensure diversity in hand positions, angles, and lighting conditions. This helps improve the generalization ability of the model.

2. Data Preprocessing

The captured images are processed using OpenCV. Noise reduction techniques are applied, and irrelevant frames are removed. MediaPipe is used to detect the hand region and extract landmarks.

3. Hand Landmark Detection

MediaPipe identifies 21 key points on the hand, including finger joints and palm positions. These landmarks provide detailed information about the hand structure and movement.

4. Feature Extraction

The landmark coordinates are converted into numerical feature vectors. Each gesture is represented by a set of coordinate values, which serve as input to the classifier.

5. Dataset Preparation

The dataset is divided into training and testing sets. Proper labeling is done to ensure accurate classification.

6. Model Training

A Random Forest classifier is used to train the model. The model learns patterns from the feature vectors and maps them to corresponding gesture labels.

7. Model Evaluation

The trained model is evaluated using metrics such as accuracy, precision, and recall. Testing is performed on unseen data to ensure reliability.

8. Real-Time Gesture Detection

The system captures live video input and processes each frame. Landmarks are extracted in real time, and the trained model predicts the gesture instantly.

9. Output Generation

The recognized gesture is displayed as text on the screen. This allows users to understand the system's output immediately.

IV.Experimental Result and Discussion

The proposed system was evaluated under different conditions to assess its performance. The results indicate that the system achieves high accuracy and operates efficiently in real time. The model achieved an accuracy of approximately 85% to 97%, depending on the dataset size and variability. The system operates at around 30 frames per second (FPS), ensuring smooth real-time detection.

The system successfully recognized gestures such as Hi, Yes, No, Thanks, and Please with minimal delay. It performed well under varying lighting conditions and backgrounds, demonstrating robustness. Some limitations were observed when gestures were partially visible or when multiple hands were present in the frame. These issues can be addressed by improving dataset quality and using more advanced models. Overall, the results demonstrate that the proposed system is effective and suitable for real-world applications.

To gain deeper insight into system performance, multiple test scenarios were designed, including indoor and outdoor environments, different camera angles, and varying distances between the user and the camera. The system maintained stable performance in most conditions, although slight accuracy degradation was noticed in extreme lighting or when the hand moved too quickly. This indicates that while the model is robust, it can still be sensitive to rapid motion and extreme environmental variations.

A detailed analysis of prediction results showed that most misclassifications occurred between visually similar gestures. This suggests that additional discriminative features or more training samples could further improve classification accuracy. Increasing the diversity of training data and incorporating more complex gesture variations can help reduce such errors.

A. Low Complexity Gesture Detection

In this scenario, the system evaluates simple and clearly visible gestures such as “Hi” or “Yes.” The hand position, orientation, and movement closely match the trained dataset patterns.

The system correctly identifies the gesture with high confidence and displays the output instantly. This demonstrates that the model performs efficiently for well-defined gestures and maintains high accuracy under normal conditions

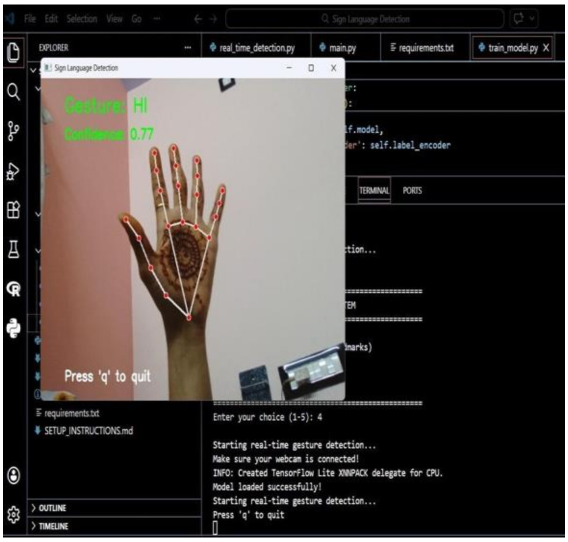


Fig – 5.1 : Output of gesture "HI"

B.Moderate Variation Gesture Detection

In this case, the gesture includes slight variations such as changes in hand angle, lighting conditions, or background noise. These variations introduce minor deviations from the trained data.

The system is still able to recognize the gesture correctly in most cases, although the confidence score may slightly decrease. This shows that the model is robust and can handle moderate environmental variations while maintaining reliable performance.

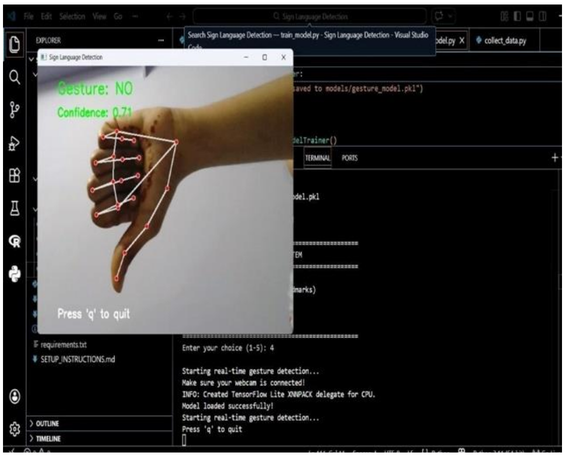


Fig.5.2 : ouput of gesture "NO"

C.Complex Gesture Detection

This scenario includes comparatively complex gestures such as "Please," where hand movement and positioning are slightly more detailed.

The system processes the gesture effectively and predicts the correct output with good accuracy. However, minor variations in execution may slightly affect prediction confidence. This demonstrates the system's capability to handle more detailed gestures.

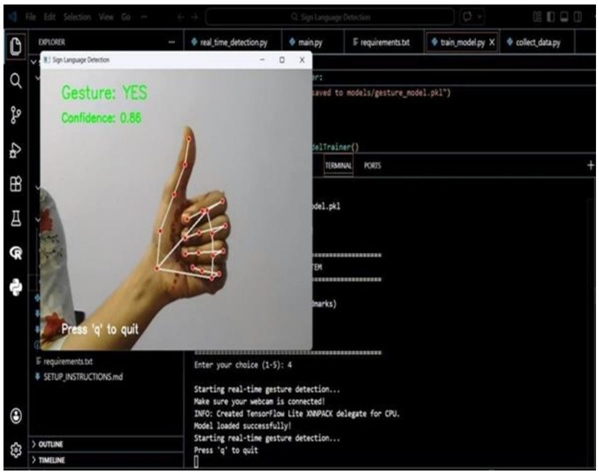


Fig . 5.3 : output of gesture "YES"

D.Challenging Condition Detection

This scenario involves difficult conditions such as poor lighting, partial hand visibility, or fast hand movement while performing any of the gestures (Hi, Yes, No, Thanks, Please).

The system may occasionally produce lower confidence predictions or fail to detect the gesture accurately. In such cases, it displays messages like "No Hand Detected" or "Unknown Gesture," preventing incorrect outputs and improving system reliability.

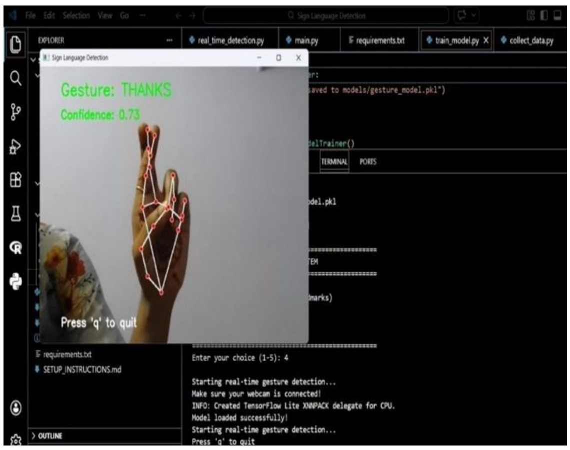


Fig .5.4 : output gesture of "THANKS"

D. Real-Time Detection Performance

The system continuously captures video input and processes each frame in real time. All five gestures — Hi, Yes, No, Thanks, and Please — are detected and displayed instantly without noticeable delay.

This real-time capability ensures smooth interaction and makes the system suitable for practical applications such as communication support for hearing-impaired individuals.

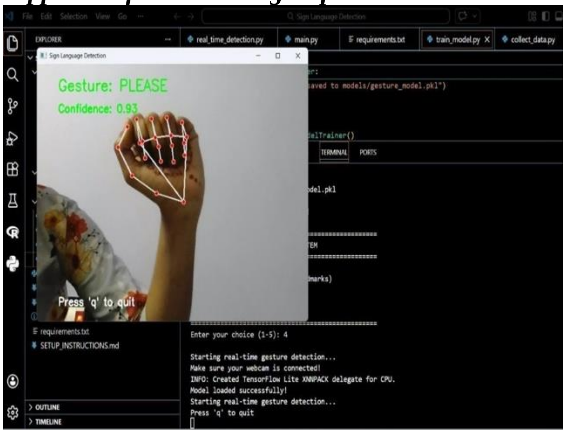


Fig . 5.5:output of gesture "PLEASE"

V.Conclusion

This project presents a real-time gesture recognition system using deep learning and computer vision techniques. The system successfully recognizes hand gestures and converts them into text, providing an effective solution for communication barriers faced by hearing-impaired individuals.

The use of MediaPipe for hand landmark detection and machine learning for classification improves both accuracy and efficiency. The system is lightweight, cost-effective, and easy to deploy, making it accessible for a wide range of users. The project highlights the potential of artificial intelligence in developing assistive technologies that improve quality of life and promote inclusivity.

In addition to its core functionality, the system demonstrates how real-time interaction can be achieved with minimal hardware requirements. By relying on a standard webcam and optimized algorithms, the solution eliminates the need for expensive sensors or specialized equipment. This makes it suitable for deployment in everyday environments such as homes, classrooms, and workplaces.

The modular design of the system allows for easy scalability and future enhancements. New gestures can be added by simply extending the dataset and retraining the model, without requiring significant changes to the overall architecture. This flexibility ensures that the system can evolve based on user needs and application requirements.

Another important contribution of this project is its user-friendly approach. The interface is simple and intuitive, enabling users with minimal technical knowledge to interact with the system effectively. The real-time feedback mechanism further enhances usability by providing immediate results, which is crucial for seamless communication.

The project also emphasizes the importance of balancing accuracy and computational efficiency. By using optimized models and efficient feature extraction techniques, the system achieves reliable performance without placing heavy demands on system resources. This makes it suitable for integration into portable devices and embedded systems.

From a broader perspective, this work contributes to the growing field of human-computer interaction by enabling more natural and intuitive ways of communication. Gesture-based systems reduce dependency on traditional input devices and open up new possibilities for interaction, especially for individuals with physical or communication limitations.

Future enhancements can include the integration of voice output to convert recognized gestures into speech, enabling two-way communication. Expanding the gesture vocabulary and supporting regional sign languages can further increase the system's usability. Additionally, incorporating advanced deep learning models may improve recognition accuracy and adaptability under complex conditions.

Overall, the project demonstrates a practical and impactful application of computer vision and machine learning, showcasing how technology can be leveraged to create meaningful solutions that address real-world challenges and support inclusive communication.

VI.Future work

The system can be enhanced in several ways to improve its functionality and performance:

Increase the number of gestures to include alphabets and sentences

Integrate text-to-speech functionality for voice output

Develop mobile and web-based versions of the system

Use advanced deep learning models such as CNN and LSTM

Improve accuracy for dynamic and complex gestures

Implement multi-language support

Add facial expression recognition for better interpretation

These improvements will make the system more versatile and suitable for real-world deployment.

1. Expansion of Gesture Vocabulary

Currently, the system recognizes a limited set of basic gestures such as Hi, Yes, No, Thanks, and

Please. In future, the dataset can be expanded to include a larger set of gestures, including alphabets, numbers, and complete words or sentences. This would enable the system to support full sign language communication rather than just basic interactions.

2. Recognition of Dynamic Gestures

The current system primarily focuses on static gestures. However, many sign language gestures are dynamic and involve continuous hand movements over time. Future work can involve integrating temporal models such as Long Short-Term Memory (LSTM) networks or Recurrent Neural Networks (RNNs) to capture motion patterns and improve recognition of dynamic gestures.

3. Integration of Deep Learning Models

While the current system uses a machine learning classifier, future improvements can include the use of advanced deep learning models such as Convolutional Neural Networks (CNNs) for feature extraction and classification. Hybrid models combining CNN and LSTM can also be explored to improve both spatial and temporal understanding of gestures.

4. Text-to-Speech Conversion

To enhance usability, the system can be integrated with text-to-speech (TTS) technology. This would allow recognized gestures to be converted into spoken words, enabling more natural communication between hearing-impaired individuals and others.

5. Mobile and Web Application Development

Currently, the system is implemented on a desktop environment. Future work can involve developing mobile and web-based applications to make the system more accessible. Deploying the model on smartphones would allow users to utilize the system anytime and anywhere.

6. Multi-Hand and Multi-Person Recognition

The existing system focuses on recognizing gestures from a single hand. Future enhancements can include support for multi-hand and multi-person recognition, enabling the system to handle more complex scenarios such as conversations involving multiple users.

7. Improved Robustness and Accuracy

Although the system performs well under standard conditions, its accuracy may decrease in challenging environments such as low lighting, cluttered backgrounds, or occlusions. Future work can focus on improving robustness by using data augmentation techniques, larger datasets, and more advanced models.

8. Real-Time Performance Optimization

As the system becomes more complex with additional features, maintaining real-time performance becomes crucial. Optimization techniques such as model compression, quantization, and efficient algorithms can be implemented to ensure fast processing without compromising accuracy.

9. Multilingual Support

The system can be extended to support multiple languages by converting recognized gestures into different language outputs. This would make the system more versatile and useful in diverse linguistic environments.

10. Integration with IoT and Smart Systems

Gesture recognition can be integrated with Internet of Things (IoT) devices to enable gesture-based control of smart environments. For example, users could control lights, appliances, or other devices using hand gestures, enhancing convenience and accessibility.

11. Facial Expression and Body Gesture Recognition

Sign language often involves not only hand gestures but also facial expressions and body movements. Future work can incorporate facial recognition and pose estimation techniques to improve the accuracy and completeness of gesture interpretation.

12. Dataset Enhancement and Standardization

A larger and more diverse dataset can significantly improve the model's performance. Future work can involve collecting data from different users, backgrounds, and lighting conditions to make the system more generalized and reliable.

VII.References

1. Padmanandam, K., Rajesh, M. V., Upadhyaya, A. N., Chandrashekar, B., & Sah,S.(2022). *Artificial intelligence biosensing system on hand gesture recognition for the hearing impaired.* **International Journal of Operations Research and Information Systems (IJORIS)**, 13(2), 1–13.
2. Bangaru, S. S., Wang, C., Zhou, X., Jeon, H.W.,&Li,Y.(2020). *Gesture recognition–based smart training assistant system for construction worker earplug-wearingtraining.* **Journal of Construction Engineering and Management**, 146(12), 04020144.
3. Hisham, B., & Hamouda, A. (2019). *Supervised learning classifiers for Arabic gesture recognition using Kinect V2.* **SN Applied Sciences**, 1(7), 768.
4. Jeyanthi, P., Ajees, A., Kumar, A. P., Revathy, S., & Gladence, M. (2025). *Interactive hand gesture recognition with audioresponse.* In **Multidisciplinary Applications of AI and Quantum Networking** (pp. 195–212). IGI Global.
5. Alnaim, N. (2020). *Hand gesture recognition using deep learning neural networks.* Doctoral Dissertation, Brunel University London.

6. Ascari, R. E. S., Pereira, R., & Silva, L. (2020). *Computer vision-based methodology to improve interaction for people with motor and speech impairment.* **ACM Transactions on Accessible Computing (TACCESS)**, 13(4), 1–33.
7. Bohra, T., Sompura, S., Parekh, K., & Raut, P. (2019). *Real-time two-way communication system for speech and hearing impaired using computer vision and deep learning.* In **Proceedings of the International Conference on Smart Systems and Inventive Technology (ICSSIT)** (pp. 734–739). IEEE.
8. Kareem, D. A., & Rajesh, D. (2025). *Enhancing WBAN performance with cluster-based routing protocol using black widow optimization for healthcare applications.* **Journal of Intelligent Systems and Internet of Things**, 14(1).
9. Tateno, S., Liu, H., & Ou, J. (2020). *Development of sign language motion recognition system for hearing-impaired people using electromyography signals.* **Sensors**, 20(20), 5807.
10. Allehaibi, K. H. (2025). *Artificial intelligence-based automated sign gesture recognition solutions for visually challenged people.* **Journal of Intelligent Systems and Internet of Things**, 2, 127–227.