International Journal of Engineering and Techniques - Volume 12 Issue 2, March-April - 2026

RESEARCH ARTICLE OPEN ACCESS

Transforming expression into language: American sign language recognition with image processing

Ms. Aparna R. Pingle, Ms. Rutuja V. Patil , Ms. Prerana A. Shinde , Ms. Riya B. Waje , Prof. P.J. Patel

Department of Artificial Intelligence and Data Science Engineering, Jawahar Education Society's Institute of Technology and Research, Nashik, India

Email: pingleaparna6@gmail.com Email: rutujapatil6030@gmail.com

Email: Preranaashokshinde2003@gmail.com Email: riyawaje1@gmail.com

Email: pooja.patel387@gmail.com

Abstract:

This article introduces "Transforming Expression into Language American Sign Language (ASL) Recognition System" American Sign Language (ASL) plays a crucial role in enabling communication for individuals with hearing and speech impairments. However, the lack of widespread understanding of sign language creates a communication barrier between deaf and hearing communities. This paper presents an intelligent and efficient system for real-time ASL recognition using advanced machine learning and computer vision techniques. The proposed model captures hand gestures through image or video input and processes them using a deep learning framework, specifically Convolutional Neural Networks (CNNs), to accurately classify ASL alphabets and gestures The system is trained on a diverse dataset of hand signs to ensure robustness against variations in lighting conditions, backgrounds, and hand orientations. Preprocessing techniques such as image normalization, segmentation, and feature extraction are applied to enhance recognition accuracy. The model demonstrates high performance in terms of accuracy, precision, and response time, making it suitable for real-time applications.

Keywords — **American Sign Language (ASL) , Hand Gesture Recognition, Machine Learning, Deep Learning, Convolutional Neural Network (CNNs), Image Processing, Real-Time Recognition**

I. INTRODUCTION

Communication is a fundamental aspect of human interaction, yet individuals with hearing and speech impairments often face significant challenges in expressing themselves and understanding others. American Sign Language (ASL) is one of the most widely used visual languages that relies on hand gestures, facial expressions, and body movements to convey meaning.

With the rapid advancement of artificial intelligence, particularly in machine learning and computer vision, automated sign language recognition systems have emerged as a promising solution to bridge this gap. These systems aim to translate hand gestures into text or speech, enabling seamless interaction. Among various techniques, deep learning models, especially Convolutional Neural Networks (CNNs), have

Traditional approaches to sign language recognition relied heavily on sensor-based methods, such as data gloves and motion trackers, which were

often expensive and less practical for everyday use. In contrast, vision-based systems using cameras are more accessible and cost-effective, making them suitable for real-time applications. However, challenges such as variations in lighting conditions, background complexity, and differences in hand shapes and orientations still affect system accuracy.

Recent research focuses on improving recognition accuracy and efficiency by utilizing large datasets, advanced preprocessing techniques, and optimized neural network architectures.

This paper aims to develop an efficient and accurate ASL recognition system using deep learning techniques, addressing the limitations of existing methods and contributing to the advancement of accessible communication technologies.

This project focuses on developing an efficient and reliable ASL recognition system that can recognize hand gestures in real time. By leveraging modern AI techniques, the system is designed to improve accuracy and usability, providing a practical solution for enhancing communication between sign language users and the wider community. The research not only highlights the technical aspects of gesture recognition but also emphasizes its social importance in promoting inclusivity and equal access to communication.

II. LITERATURE REVIEW

Sign language recognition has gained significant attention in recent years due to its importance in assisting individuals with hearing and speech impairments. Early studies focused on understanding the complexity of sign language, which involves hand gestures, facial expressions, and body movements, making it a challenging problem in the field of computer vision [6].

Traditional approaches to sign language recognition relied on sensor-based systems such as data gloves and motion trackers. While these systems provided accurate gesture detection, they were expensive, intrusive, and not suitable for real-time or large-scale applications. As a result, researchers shifted towards vision-based techniques using cameras, which are more practical and cost-effective.

With the advancement of deep learning, Convolutional Neural Networks (CNNs) have become the most widely used method for sign language recognition. CNN-based models automatically extract spatial features from images and have shown high accuracy in recognizing ASL alphabets and gestures. Studies have demonstrated that CNN-based systems can effectively classify hand gestures captured through webcams and convert them into text in real time.

Further improvements have been achieved by integrating multiple deep learning architectures. For example, hybrid models combining CNN with Long Short-Term Memory (LSTM) networks have been proposed to capture both spatial and temporal features in sign language videos. These models significantly improve performance for continuous sign recognition, where gestures occur in sequences rather than isolated images.

Digital job platforms that currently exist have been very effective in enhancing the efficiency of hiring in formal sectors but the design assumptions of these platforms are not in line with reality of informal workers. Studies have shown that a majority of these platforms demand organized resumes, daily internet connectivity, and some degree of digital literacy, ruling out much of the daily wage-earning population. In addition, such systems seldom have systems of testing practical skills, which can be more applicable than formal education in the informal sector.

With the rapid advancement of artificial intelligence and computer vision technologies, there is growing potential to bridge this communication gap. ASL recognition systems aim to interpret hand gestures and convert them into readable text or audible speech, making interactions more accessible and inclusive. These systems typically rely on machine learning algorithms, image processing techniques, and real-time data analysis to accurately detect and classify different signs

Recent research has also explored alternative modalities such as Wi-Fi signals, radar systems, and multimodal approaches to enhance recognition accuracy. These methods aim to overcome challenges such as background noise, lighting variations, and occlusion by capturing gesture information through different sensing techniques.

Additionally, modern systems incorporate preprocessing techniques like segmentation, normalization, and feature extraction to improve robustness and efficiency.

Overall, the literature indicates that deep learning-based approaches, particularly CNN and hybrid architectures, have significantly advanced the field of ASL recognition. However, further research is required to develop more accurate, real-time, and scalable systems for effective human-computer interaction.

Moreover, available studies reflect the necessity of incorporating career development assistance into job marketplaces. Long-term employability can be greatly improved by offering access to upskilling tools and training advice and insights on market demands. Nevertheless, the majority of existing solutions are aimed at quick job placement and the significance of ongoing skills upkeep is not taken into consideration.

Based on the literature reviewed, it is evident that the literature has been done in individual parts which include digital identity, accessibility, and job matching, but there has been no comprehensive solution that has incorporated these components into a single system. This gap inspires the creation of the improved Kamgar Connect platform that will provide an integrated, user-centric solution specifically designed to meet the needs of the informal workforce.

III. PROPOSED SYSTEM

The proposed architecture is a deep learning-based framework designed to recognize American Sign Language (ASL) gestures from images or real-time video input. It integrates computer vision and machine learning techniques to accurately detect, process, and classify hand gestures into meaningful outputs such as text or speech. The system is designed to be efficient, scalable, and suitable for real-time applications.

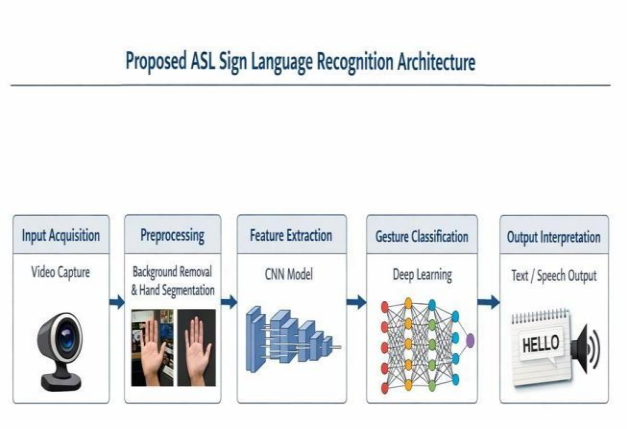


Fig .1 Proposed System

1. **System Overview:** The proposed ASL recognition system is designed to identify hand gestures and convert them into readable text in real time. The system uses a camera to capture live video input, which is then processed using image processing and machine learning techniques. First, the captured frames are pre-processed to remove noise and enhance the hand region. After that, feature extraction is performed to identify important patterns such as hand shape and movement.

2. **The important parts of the System:**

a) **Input Acquisition-**

The system captures hand gesture data using a camera or video input device. This can be in the form of live video streams or pre-recorded images. The system collects hand gesture data using a camera or video input device. It can capture gestures either in real time through a live video stream or from pre-recorded images for further processing.

b) **Pre-processing:**

In this stage, the captured images or video

frames are cleaned and prepared for analysis. Noise is reduced, the background may be removed, and the hand region is enhanced or isolated to make gesture detection more accurate.The captured input is processed to remove noise and improve quality. Techniques like background removal, resizing, normalization, and hand segmentation are applied to isolate the hand region.

c) **Feature Extraction:**

Important features from the hand gestures are extracted using deep learning models, mainly Convolutional Neural Networks (CNNs), which automatically identify patterns such as edges, shapes, and textures. Feature extraction is the process of identifying the most important details from the hand gesture image that help in distinguishing one sign from another. This includes characteristics like hand shape, edges, contours, and key points. These features are converted into a structured format so that the machine learning model can easily analyze and classify the gesture accurately.

d) **Gesture Classification :**

In this stage, the system uses a trained machine learning model to identify the hand gesture based on the extracted features. The model compares the input features with learned patterns and assigns the gesture to the correct class or label. This step is crucial for accurately recognizing different ASL signs.

The extracted features are passed to a trained classification model that predicts the corresponding ASL alphabet or gesture. This step uses deep learning algorithms to ensure high accuracy.

e) **Post Processing :**The predicted output is refined to improve reliability. This may include smoothing predictions, correcting errors, or combining multiple frames for better results. In this stage, the recognized gesture is refined and converted into a meaningful output. The system may correct minor errors, combine multiple gestures into words or sentences, and display the final result as text or speech for easy understanding.

f) **Output interpretation:**

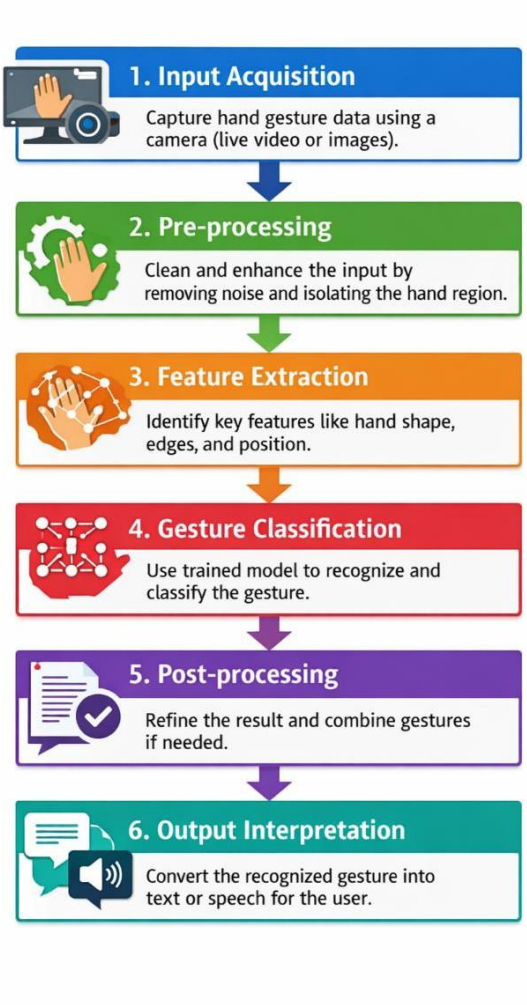
Finally, the recognized gesture is converted into readable text or audible speech, enabling communication between users. In this stage, the recognized gesture is translated into a clear and understandable form for the user. The system converts the classified sign into text or speech and presents it in a meaningful way, ensuring effective communication between the user and the system.

g) **Output:**

In this phase the user can get the proper sign output to transfer or get the language to other people it will be understand to them.

3. System Architecture (Conceptual Flow):

1. **Input Acquisition** – Capture hand gesture data using a camera (live video or images).
2. **Pre-processing** – Clean and enhance the input by removing noise and isolating the hand region.
3. **Feature Extraction** – Identify key features like hand shape, edges, and position.
4. **Gesture Classification** – Use a trained model to recognize and classify the gesture.
5. **Post-processing** – Refine the result and combine gestures if needed.
6. **Output Interpretation** – Convert the recognized gesture into text or speech for the user.



4. Advantages of Proposed System:

1. **Improves Communication** – Helps bridge the gap between deaf/mute individuals and non-sign language users.
2. **Real-time Interaction** – Enables instant translation of gestures into text or speech.
3. **User-Friendly** – Easy to use with just a camera, no special hardware required.
4. **Time Efficient** – Reduces the need for human interpreters in everyday situations.
5. **Cost Effective** – Once developed, it can be used widely without high operational costs.
6. **Increases Accessibility** – Makes communication more inclusive in public places like schools, hospitals, and offices.
7. **Scalable System** – Can be expanded to recognize more gestures or different sign languages.
8. **Supports Learning** – Can also be used as a tool for learning sign language.

IV. METHODOLOGY

This research paper follows a conceptual approach to the methodology, which involves modelling the ways in which a digital resource such as American Sign Language to understand people can enhance access to sign through the project in the informal economy. The system is perceived to be a linkage between normal people and deaf people and it tackles major issues like absence of structured skill representation, lack of trust and ineffective communication gaps matching. It is assumed that the action of the person effective or matching to the project are providers of skills and data sets are finding of opportunities and the lack of fit between the two is described as information incompleteness and poor visibility.

The proposed ASL recognition system follows a structured approach to accurately detect and interpret hand gestures. The process begins with data collection, where hand gesture images or video frames are captured using a camera. A dataset is created that includes different ASL signs under varying lighting conditions and backgrounds to improve the robustness of the system.

In the next stage, pre-processing is applied to enhance the quality of the input data. This includes noise removal, background subtraction, and segmentation to isolate the hand region from the image. The processed images are then passed to the feature extraction stage, where important characteristics such as hand shape, edges, and key points are identified and converted into a suitable format.

After feature extraction, a machine learning or deep learning model is trained using the prepared dataset. The model learns patterns associated with each gesture and is then used to classify new input gestures. Once the gesture is recognized, post-processing techniques are applied to refine the output and reduce errors. Finally, the recognized gesture is converted into readable text or speech, enabling effective communication between users.

$$= 1 + 2 + 3 + 4$$

Conceptualization of system architecture is in the form of a layered model made up of user interaction, application services, processing logic and data storage. This structure provides the flow of information between the user input and result generation with ease.

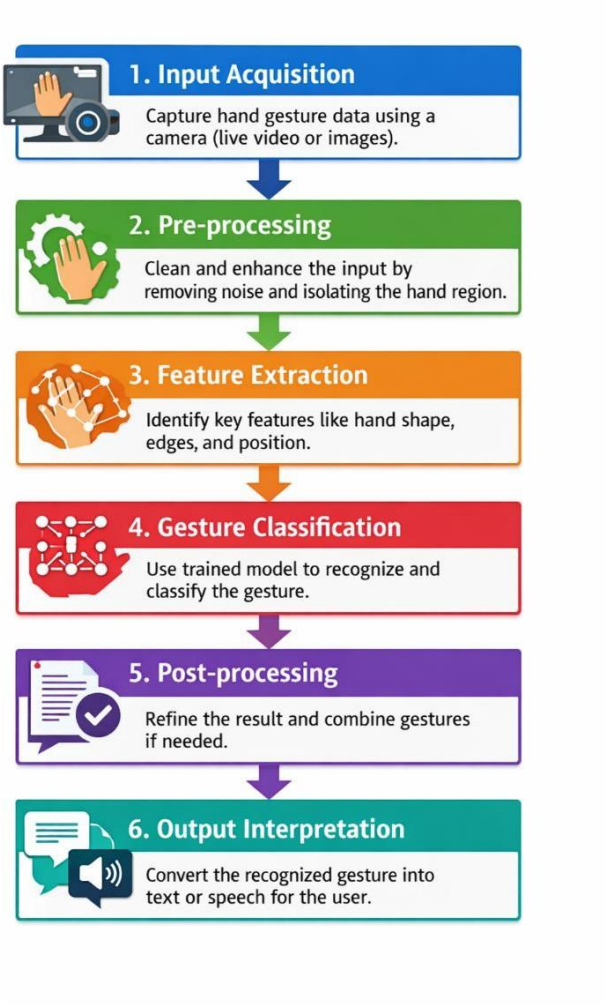


Fig.2:Architecture Diagram.

The running of the system can also be depicted in a sequence diagram, The sequence diagram represents the step-by-step interaction between different components of the American Sign Language (ASL) recognition system over time. It shows how the user's gesture is processed and converted into meaningful output.

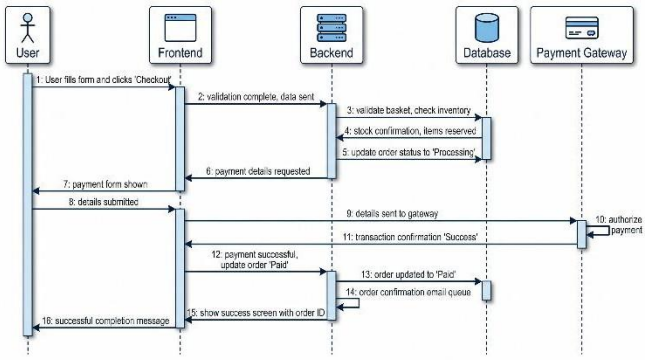


Fig.3:*Sequence Diagram*

All in all, this conceptual framework shows that structured information, smart matching, and mechanisms of trust can all be used to increase transparency, efficiency, and access in informal understanding of the communication gap which how many reduces .

V. FUTURE SCOPE

The future scope of the American Sign Language (ASL) recognition system is highly promising with continuous advancements in Artificial Intelligence and Computer Vision. One major improvement can be the expansion of the system to recognize not only alphabets but also full words, sentences, and complex expressions in real time. This would enable more natural and fluent communication for deaf and hard-of-hearing individuals.

Integration with mobile applications and wearable devices can make the system more portable and accessible in daily life. Additionally, combining ASL recognition with speech synthesis technologies can allow seamless two-way communication between sign language users and non-signers.

VI. CONCLUSIONS

IN CONCLUSION, THE AMERICAN SIGN LANGUAGE (ASL) RECOGNITION SYSTEM SUCCESSFULLY DEMONSTRATES HOW MODERN TECHNOLOGIES LIKE

ARTIFICIAL INTELLIGENCE AND IMAGE PROCESSING CAN BE USED TO BRIDGE THE COMMUNICATION GAP

BETWEEN DEAF AND HEARING INDIVIDUALS. THE SYSTEM CAPTURES HAND GESTURES USING A WEBCAM, PROCESSES THE IMAGES, AND ACCURATELY CLASSIFIES THEM USING A CNN-BASED MODEL TO GENERATE MEANINGFUL TEXT OR SPEECH OUTPUT.

THE PROJECT HIGHLIGHTS THE EFFECTIVENESS OF REAL-TIME GESTURE RECOGNITION WITH HIGH ACCURACY AND LOW LATENCY, MAKING IT SUITABLE FOR PRACTICAL APPLICATIONS. IT ALSO ENSURES A USER-FRIENDLY AND CONTACTLESS COMMUNICATION METHOD WITHOUT THE NEED FOR HUMAN INTERPRETERS.

OVERALL, THIS SYSTEM PROMOTES INCLUSIVITY, INDEPENDENCE, AND ACCESSIBILITY FOR THE DEAF
COMMUNITY. WITH FURTHER IMPROVEMENTS, IT HAS THE POTENTIAL TO BE WIDELY USED IN AREAS SUCH AS EDUCATION, HEALTHCARE, AND PUBLIC SERVICES, CONTRIBUTING TO A MORE CONNECTED AND INCLUSIVE SOCIETY.

REFERENCES

1) *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf — William C. Stokoe (1960/2005)*
2) *The Signs of Language — Edward S. Klima & Ursula Bellugi (1979)*
3) *Grammar, Gesture, and Meaning in American Sign Language — Scott K. Liddell (2003)*
4) *Sign Language and Linguistic Universals — Wendy Sandler & Diane Lillo-Martin (2006)*
5) *Language, Cognition, and the Brain: Insights from American Sign Language — Karen Emmorey (2001)*
6) *Psycholinguistic and Neurolinguistic Perspectives on Sign Languages — David P. Corina & Heather P. Knapp (2006)*
7) *Inside Deaf Culture — Carol A. Padden & Tom L. Humphries (2005)*
8) *Language Contact in the American Deaf Community — Ceil Lucas & Clayton Valli (1992)*
9) *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language — Susan Goldin-Meadow (2003)*
10) *Language Variation and Change in the American Sign Language Enterprise — Joseph C. Hill (2012)*