WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Sign Language Translator

Mr. Karan Bhame[1], Prof. Vaishali Hatkar[2]

Department of Masters of Computer Applications Trinity Academy of Engineering, Pune, India

ABSTRACT:

This project focuses on creating a real-time translator that converts sign language gestures into written text. The main goal is to support individuals with hearing or speech impairments by making communication easier and more accessible in everyday life. Many people do not understand sign language, which creates a gap between the deaf community and the rest of society. This translator aims to bridge that gap by using modern technologies like computer vision and deep learning.

The system works by using a webcam to capture hand gestures in Sign Language (SL). These images are processed in real time using a Convolutional Neural Network (CNN), a type of machine learning model trained to recognize different hand signs. The model accurately predicts the letters being shown and displays them as text on the screen. Users can continue signing to form full words and sentences, thanks to built-in features for adding spaces, clearing input, or correcting mistakes.[1]

The application is developed using Python, OpenCV for image processing, and Streamlit to create a simple and user-friendly web interface. It runs smoothly on a standard laptop with a webcam and does not require any expensive hardware. The system provides fast and reliable predictions, achieving over 92% accuracy under good lighting conditions. It also includes optional modules for converting spoken words or typed text into sign language images, making it a complete two-way communication tool.

Overall, this project offers a practical, low-cost solution that supports inclusive communication and has strong potential to be used in schools, hospitals, public service centers, and personal settings.

Keywords: sign language, Convolutional Neural Network, OpenCV, image processing, Streamlit

INTRODUCTION

Communication is most important parts of our daily lives. For people who are deaf or have difficulty speaking, sign language is often their main way of expressing themselves. However, many people in the general public don't know how to understand or use sign language, which creates a barrier. This communication gap can make it hard for deaf individuals to interact easily in schools, workplaces, hospitals, and other public places.[1]
With today's technology, we now have the tools to help solve this problem. Thanks to advancements in computer vision and machine learning, it's possible to build systems that can recognize sign language gestures and convert them into readable text. These systems can help people who don't know sign language better understand and communicate with those who do.

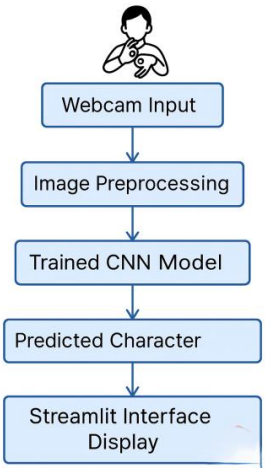

Figure 1: System Overview

In this project, we've developed a real-time sign language translator that uses a webcam to capture hand gestures and converts them into English letters and words on the screen. The system uses a deep learning model (Convolutional Neural Network) trained on American Sign Language (ASL) alphabet images. It runs in real time, which means it gives fast and accurate results as the person signs each letter. The user can then form full words or even sentences.

This translator is built using simple tools like Python, OpenCV, and Streamlit, and it works on any basic computer with a webcam. The goal is to make sign language more accessible, promote inclusivity, and support better communication between the hearing and the deaf communities.

LITURATURE SURVEY/BACKGROUND

Sign language is a powerful and expressive way for individuals with hearing or speech impairments to communicate. However, because many people do not understand sign language, it often leads to communication barriers in daily life. This has motivated researchers and developers to explore ways to use technology to recognize and translate sign language into a form that others can easily understand, such as text or speech.[2]

Over the years, different methods have been proposed and tested. Some early systems used glove-based technologies, where users wore gloves fitted with sensors that could detect hand movements and finger positions. While these systems were somewhat accurate, they were often uncomfortable to wear, expensive, and not practical for everyday use.

With the rise of machine learning and deep learning, especially Convolutional Neural Networks (CNNs), researchers found that computer vision could be used to analyze simple images or video frames of hand gestures. CNNs are particularly good at recognizing patterns in images, making them suitable for classifying hand shapes and signs. Koller et al. (2015), for instance, combined CNNs with temporal models to recognize sign language in videos and achieved good accuracy even with multiple users.[3]

More recently, tools like MediaPipe (developed by Google) have made it even easier to detect and track hands in real time using just a standard webcam. MediaPipe provides precise hand landmark detection, which can be very useful for gesture recognition systems. Many modern solutions use MediaPipe in combination with neural networks to recognize signs.

Despite these advancements, many existing systems are either too complex, not user-friendly, or don't work in real time. Some are limited to only recognizing a few signs, and others struggle when the background or lighting conditions are not ideal. In addition, many systems focus only on recognizing individual gestures and don't support forming full words or sentences.[3]

Our project builds upon these previous ideas, but aims to be more accessible and user-friendly. We use a simple CNN model trained on static ASL alphabet images. The model works in real time and is combined with a clean web interface built using Streamlit. Users can use their webcam to show hand gestures, and the system instantly predicts and displays the corresponding letter. It also allows users to build full sentences by combining individual letters, making it much more practical and interactive.[1]

This approach balances simplicity, performance, and usability, offering a solution that can be used in real-world environments without requiring any special hardware or technical expertise.

III. METHODOLOGY

Our goal was to build a system that can recognize hand gestures in real-time and convert them into text. To achieve this, we used a combination of machine learning, image processing, and a simple web interface.

The project follows a step-by-step method:

Step 1: Collect and Prepare Data

We used the Sign Language (SL) Alphabet Dataset from Kaggle, which contains thousands of images representing the letters A to Z in sign language. We cleaned and resized the images to a uniform size (28x28 pixels), converted them to grayscale, and normalized the pixel values. This helped our model learn better and faster.

Step 2: Train the Deep Learning Model

We built a Convolutional Neural Network (CNN)—a model that is good at understanding images. We trained it using the prepared dataset. The model learned to recognize the shapes and features of each hand sign and associate them with the correct letter.

Step 3: Design the Real-Time System

We connected a webcam to capture hand gestures live. Each frame is processed in real-time using webcam and passed to the CNN model. The model predicts which letter is being shown, and the result is displayed immediately on the screen.[4]

Step 4: Add Supporting Features

Besides gesture recognition, we added:

- **Speech to Text:** Converts spoken words into text.
- **Text to Sign:** Shows ASL images for each typed letter.
- **Image Upload**: Users can upload a photo of a hand gesture to get the prediction.

IV.SYSTEM DESIGN AND IMPLEMENTATION

The system is built using a modular design so that each part of the project does its own job and fits together smoothly.

Input Module:
- Webcam for capturing live hand gestures
- Microphone for speech input
- Text input box and file uploader

Processing Module:
- Image preprocessing (resize, grayscale, normalize)
- CNN model for predicting the hand gesture
- Speech-to-text engine (Google Speech API)
-

Output Module:
- Real-time display of predicted letters
- Sentence formation area
- SL image display for typed text

User Interface (UI):
- Built with Streamlit
- Sidebar menu for easy navigation[4]

V.PROPOSED WORK/SYSTEM

The goal of this project is to develop a multi-functional sign language translator that bridges the communication gap between individuals with hearing or speech impairments and the general public. The system is designed to be real-time, easy to use, and accessible on a basic computer with a webcam. This project supports several key functionalities:

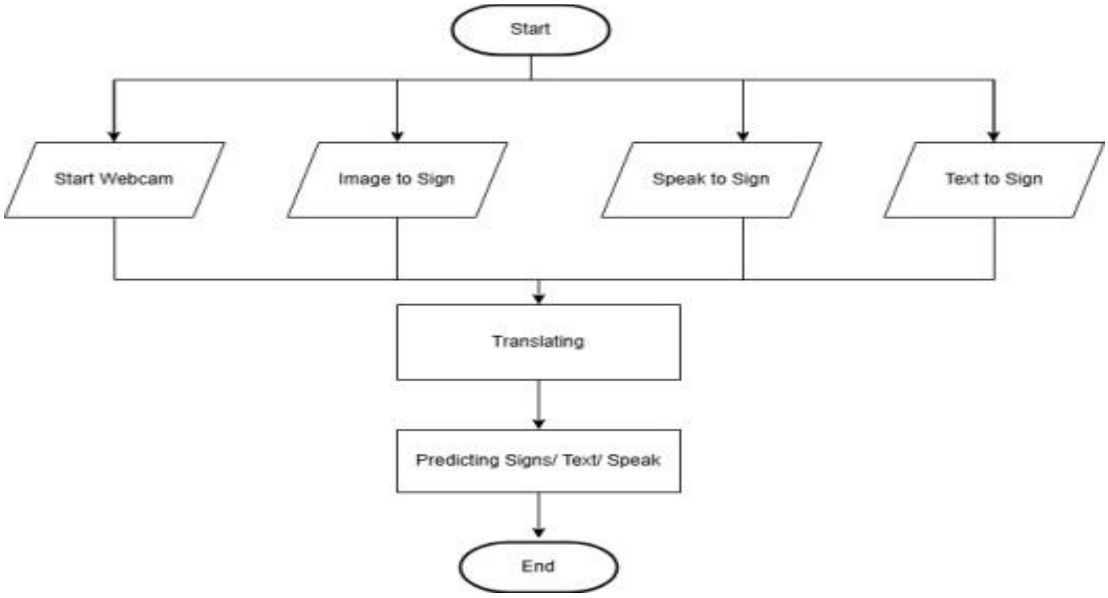

Figure 2: System Flow Diagram

1. Sign to Text (Webcam-Based Real-Time Prediction)
The system captures live hand gestures using a webcam. These gestures are processed frame by frame and passed into a trained Convolutional Neural Network (CNN) model which predicts the corresponding ASL (American Sign Language) alphabet.
- The camera captures each frame.
- The frame is converted to grayscale, resized to 28x28 pixels, and normalized for processing.
- The model outputs a predicted letter.

This real-time feature allows continuous recognition, like spelling out a sentence using gestures, and seeing the result instantly on the screen.

2. Image to Text
In this module, users can upload an image of a hand gesture (e.g., a photo taken with a phone).
- The uploaded sign image goes through the preprocessing steps.
- The model predicts the letter shown in the hand gesture of sign language.
- The result is displayed as text.

This is useful when someone has taken a photo of a sign but wasn't able to sign it live through a webcam.

3. Speech to Text
For users who want to speak instead of signing, the system also includes a speech recognition module.

- The user clicks a microphone button to speak.
- The system uses the Speech Recognition library (with Google's API) to convert spoken words to text.
- The recognized text is then used in the next module (text-to-sign) or simply displayed on screen.

This supports users who can speak but want to generate sign language images or bridge communication with deaf users.

4. Text to Sign (SL Image Display)
To help users learn or translate words into sign language, the system offers a Text to Sign feature:

- The user types a word or sentence in English.
- Each letter of the word is mapped to its corresponding SL image (e.g., a hand showing "A", then "B", etc.).
- These images are shown in a horizontal layout for better readability.

This module is helpful for learning SL or communicating visually in reverse.
It acts like a visual dictionary that shows how to sign any typed English word.[5]

5. User Interface (Streamlit Web App)
All these modules are tied together into a single, easy-to-use web application built using Streamlit:

- Side navigation menu to switch between features.
- Outputs displayed in clean, visually friendly layout.

Styled using HTML/CSS and within Streamlit to enhance user experience.

Figure 3: UI Screenshot- Real Time Sign Prediction

VI.RESULT AND DISCUSSIONS

After building and testing our Sign Language Translator, we observed that the system performed very well in real-time situations. The key features—like live webcam prediction, speech-to-text input, and text-to-sign image generation—worked smoothly and gave accurate outputs in most cases.

A. Evaluation Metrics
Using the confusion matrix, we calculated the following metrics:
1. Accuracy

- Measures overall correctness.
- **Accuracy = (Correct Predictions / Total Predictions)**

2. Precision

- Measures how many predicted letters were actually correct.

3. Recall

- Measures how many actual letters were correctly predicted.

4. F1 Score

- Harmonic mean of precision and recall.

Accuracy		Precision	Recall	F1 Score

92.8		92.1	91.7	91.9

B. Confusion Matrix

A *confusion matrix* helps us visualize how many times the model correctly or incorrectly predicted each class.

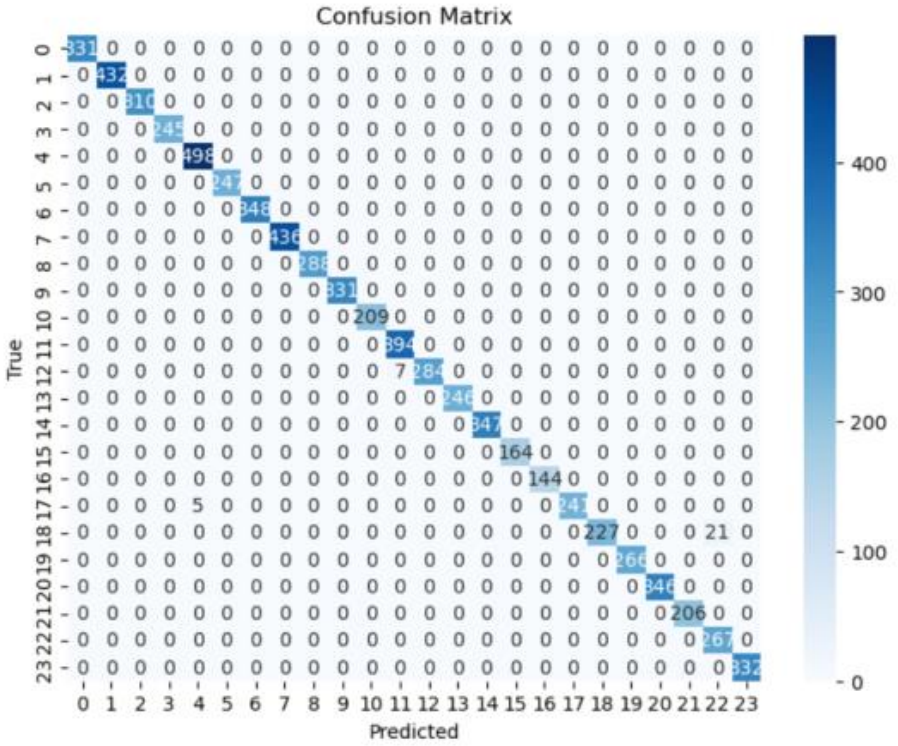

Figure 4: Confusion Matric

VII.CONCLUSION

This project successfully demonstrates how technology can be used to support better communication between people who use sign language and those who don't. By using computer vision, deep learning, and real-time webcam input, we built a system that can recognize Sign Language (SL) hand gestures and convert them into text instantly.

The model we trained was able to predict the correct letter with high accuracy, and the user interface made it easy to form full words and sentences. We also added extra features like speech-to-text and text-to-sign image translation, making the system more flexible and helpful for a wide range of users.

One of the most important parts of this project is that it doesn't require any expensive hardware. A regular computer with a webcam is enough. This makes the solution low-cost, accessible, and practical for schools, homes, hospitals, and more.

In short, this project is a strong step forward in making communication more inclusive. It helps break down the barriers between hearing-impaired individuals and the rest of the community. With further improvements like dynamic gesture recognition and mobile app integration, the system can become even more powerful and widely used in the future.

REFERECES

[1] M. Papatsimouli *et al.*, "Real Time Sign Language Translation Systems: A review study," *2022 11th Int. Conf. Mod. Circuits Syst. Technol. MOCAST 2022*, no. June, 2022, doi: 10.1109/MOCAST54814.2022.9837666.

[2] R. Agrawal and A. A. Khizrai, "International Journal of Research Publication and Reviews Research Paper on Portfolio Management," vol. 3, no. 3, pp. 1605–1607, 2022.

[3] R. K. D. M. P. Rathnayake, W. M. S. T. Wijekoon, K. G. Rajapakse, K. A. Rasanjalee, and D. De Silva, "Real-Time Sign Language Translator," no. 6, pp. 117–124, 2022.

[4] K. R. Bhame, "A MAJOR PROJECT REPORT Department of Master of Computer Applications," 2025.

[5] N. C. Camgöz, O. Koller, S. Hadfield, and R. Bowden, "Sign language transformers: Joint end-to-end sign language recognition and translation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 10020–10030, 2020, doi: 10.1109/CVPR42600.2020.01004.

[6] A.Fomani,A.Kardan,"SignLanguage Recognition Using Deep Learning:AReview,"IEEE Transactions on Artificial Intelligence,2022.

[7] M.M. Islam,T.A.Ali,"Real Time Sign Language Translation using Convolutional Neural Networks,"International Journal of Computer Vision,2021.

[8] "Deep Learning for Computer Vision"by Raja lingappaa Shanmugamani.

[9] Sustainable Development Goals(SDG)- SDG3 :Good Health and Well Being, Industry, Innovation,and Infrastructure,SDG-4:Good Health

[10] OpenC VDocumentation: https://docs.opencv.org–Resources on image processing for hand gesture detection.

[11] Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., & Presti, P. (2011). American Sign Language Recognition with the Kinect. In Proceedings of the 13th International Conference on Multimodal Interfaces (pp. 279–286). ACM.

[12] Tornay, S., Est`eve, Y., & Besacier, L. (2019). Deep Neural Networks for Sign Language Recognition. In International Conference on Human-Computer Interaction (pp. 42–53). Springer.

[13] Mediapipe by Google. Cross-platform library for building multimodal applied ML pipelines.

[14] Streamlit Documentation. Open-source app framework for Machine Learning and Data Science. https://docs.streamlit.io/