IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Conversion of Sign Language into Text Using Machine Learning Technique

Eshwari [1],Dr.Savitha C[2], Mangala[3], Nayana[4], Pooja [5]

[1,3,4,5] Student,[2] Professor, Department of Artificial intelligence and Machine Learning, Sri Siddhartha Institute of Technology, SSAHE, Tumkur, India

DOI: https://doi.org/10.56975/ijcrt.v14i7.309916

ABSTRACT

Communication is basic form to share our information, or feelings through talking, signing, or any other method. It is essential for human who are not able communicate in the usual way. Some individuals are born deaf and mute, or they may not able to speak due to health issues later in life. These individuals often rely on sign language. This highlights the need for technological solutions that can assist in bridging communication gaps. The system described here captures video from a camera, detects hand movements using computer vision, identifies key points on the hands with MediaPipe, and applies machine learning to interpret gestures and sign language. After processing, the system displays words on the screen, making conversation easier for users.

Technical Keywords: Convolution Neural Network (CNN), Machine learning, OpenCV, Sign conversion, Sign Language.

I.INTRODUCTION

American Sign Language (ASL) serves as a widely used communication system among individuals with hearing and speech impairments, particularly in the United States. “Instead of relying on spoken words, it convey meaning through a combination of hand gestures, facial expressions, and body movements.” However, it’s not a universal language, which complicates communication between signers and hearing individuals. This disadvantage creates a need for interpreters, which are not always available. The system aims to solve this issue by developing a model in Python that can watch live video, recognize gestures, match them to signs, and display the corresponding text. It employs real-time processing and AI to foster inclusion, allowing deaf individuals to communicate more independently. There is often a dependency where sign language users rely on others for translation, isolating them from regular social interactions. Literature reviews suggest various technologies for this purpose. One study notes that using the Neural network, Naives bayes classification and also decision trees, and SVM achieves around 88 to 90 percent accuracy in translating signs to speech, primarily to assist mute individuals in communicating with the public. Another study discusses how converting signs to text and speech improves connections within communities, that will be easy for deaf individuals. Additional work main goal to use real-time tools for speech-impaired individuals to facilitate two-way conversations, even in noisy environments. Generative AI with Computer

vision and NLP helps non-users understand signs better. CNNs have become popular as they effectively recognize image based on body gestures and movements. Some systems employ RNN LSTM for gesture recognition, achieving 100 percent accuracy for 26 letters, though results be not too perfect sometime it may change. OpenCV and Media Pipe enable real-time conversion, that help for people to communicate easily.

II. LITERATURE REVIEW

Deaf and Dumb people mainly communicate with hand gestures and body movements. The different technology used in this paper mainly Neural networks, decision Tree, Naives bayes classification and (SVM) Support Vector Machine that provide accurate results about 80% to 90%. It is introduced for the people to communicate by translate the sign language to speech.before this an electronic machine is needed to convert the sign to text for communication between any two people [1]. The integration of sign to text conversion and also speech represents a transformative technological advancement that act as bridge for the people who not able to communicate. This majorly helps for the deaf people to share their thoughts and express their feelings more fluently in various everyday scenarios [2]. Real-Time Conversion of gesture to text the work focuses on introducing a tool that should bridge the communication gap among speech impaired community. The work that includes system that helps to communicate in two paths between people with speech disorders [3]. Real-time sign language translation systems use computer vision (CV), natural language processing (NLP), and generative AI these techniques help to build communicate users and non-users.[4]. In old years, Convolutional Neural Networks (CNNs) has become a widely adopted approach in image classification and gesture recognition tasks due to their superior capability in learning spatial features. CNNs can effectively capture hand shapes, finger orientations, and subtle variations in gesture movements, making them well-suited for sign language recognition [5]. system capable of real-time gesture recognition and output generation [6]. Real-Time Gesture to Speech Converter Using OpenCV and MediaPipe it facilitates effective communication between deaf and hearing individuals [7]. LSTM networks can develop long term dependencies, they were investigated and employed for the classification of gesture data. The constructed model exhibited a 100% accuracy rate while categorizing 26 motions, highlighting the usefulness of LSTM-based neural networks for sign language translation [8].

METHODOLOGY

Sign language to text conversion is based on three important components, each contributing to accurate interpretation of gestures.

Fingerspelling: is used to represent words letter by letter using specific hand gestures. It is mainly applied for spelling names, places, or words that do not have a standard sign. In a conversion system, the model identifies individual hand shapes and maps them to corresponding alphabets to form words.

Word-level sign vocabulary is used for the majority of communication, where a single gesture represents a complete word or concept. This makes communication faster and more natural compared to fingerspelling. In the system, these gestures are directly recognized and translated into meaningful words in text form.

Non-manual features include facial expressions, lip movements, and body posture, which provide additional meaning and context to the signs. These features help in understanding emotions, sentence structure, and intent. In sign language to text systems, capturing these cues improves the correctness and clarity e generated text**.**

It breaks down to fingerspelling for letters, word signs for quicker talk, and non-manual cues like faces for context. The goal is a simple interface where the computer gets human signs, especially ASL fingerspelling, with neural nets. The hand goes through a filter first, then a classifier predicts the gesture. Overall, it's about turning gestures into text in real-time, recognizing hands, fingers, and sometimes

expressions. The dataset has diagrams of letters, I guess, but basically, input from the camera via OpenCV, processed through a CNN. Convolution layers pull features, pooling shrinks it down, more layers for better accuracy, flatten to a vector, fully connected, then softmax for output.

The goal of this model to develop a user-friendly Human Computer Interface (HCI) where the system analyses the gesture shown by human. Sign language it is the older form of method which is used to communicate. And nowadays, many people do not have knowledge about the signs and how to convey their thoughts for people who not able to speak and hear. For this the model use American sign language method that has mainly 3 ways. In that choose first method that is fingerspelling where it takes the hand gesture image that the image will be filtered then it predicts the gesture shown by person to opposite person.

Systems that can translate sign expressions into a format easily understood by everyone. Sign Language to Text Conversion is one such innovative forms that help to build communication gap. This technology focuses on recognizing hand gestures, finger movements, and sometimes facial expressions, and converting them into readable text in real time. By doing so, it enables smooth, natural, and inclusive communication between deaf and dumb people.

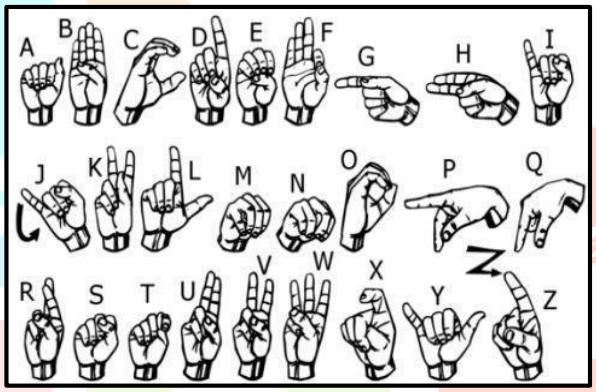

Fig 1: Data set diagram

The proposed recognition technique relies on a convolutional neural network model (CNN) with a feature-mapped output layer. The camera captures the input using OpenCV, which is then sent through the CNN classifier. The captured input first passes through the convolution layer; from the image we extract the filters the number of parameters is reduced in pooling layer when the images are large. As more number of convolutional and pooling layers can improve accuracy. After several convolution and pooling layers, the results of images are flattened into a vector and sent to a fully connected layer. The last layer is the softmax layer, which generates the output.

- *PYTHON LIBRARIES*

- OpenCV: OpenCV means Open Computer Vision It is used for image processing in python. This is required to perform various actions and processing of the captured images.
- Matplot library: Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.
- Numpy: In python numpy is one the library tool that used in multi-dimensional arrays and matrices, also used to perform the high-level mathematical functions to operate on these arrays

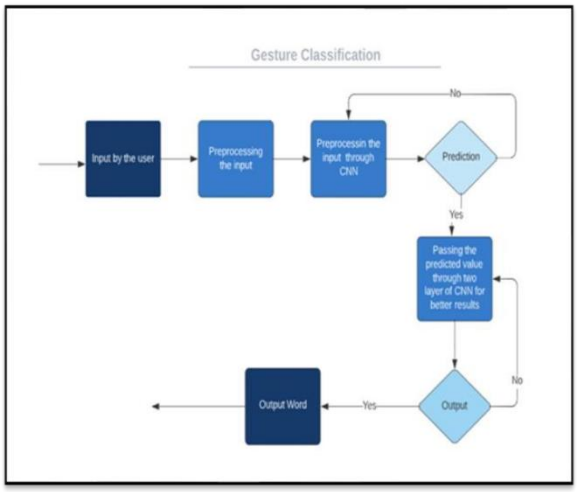

Fig 2 : Gesture Classification

Input by the User: The user performs a hand gesture in front of the camera, capturing this gesture as an image or video frame for processing.

Preprocessing the Input: The captured image is cleansed and enhanced to eliminate background noise. Techniques like resizing, normalization, and filtering are applied.

Preprocessing the Input through CNN: The processed image is passed through a Convolutional Neural Network (CNN), which extracts important visual features.

CNN Model:

CNN model takes image as an input, process it and classify it under certain categories. Through the several layers in CNN like the convolution layer, pooling layer, fully connected layers the input image can moved and these layers consists of activation functions

- *RESULTS AND DISCUSSION*

It trains on A to Z and translates to the English alphabet. Testing shows it works for communication, cutting the gap for the deaf and mute. They can chat easier without barriers. This decreases isolation, but it might not catch everything perfectly yet.

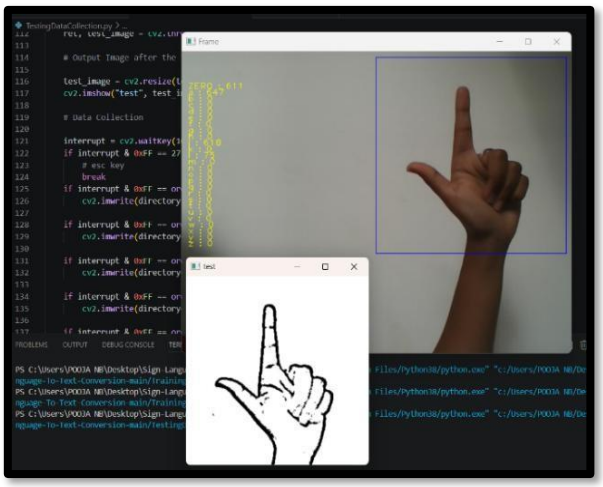

Fig 3 : Model Training

The above image is the training of dataset where the model train our model by each letter by making gesture.

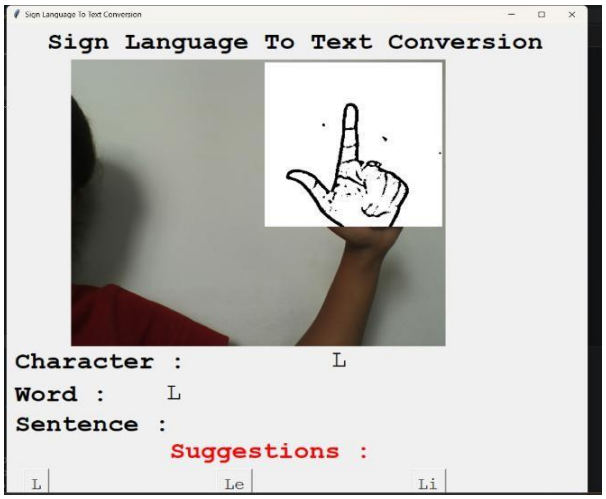

Fig 4 : Predicting the letters

This is testing of data once the model train then run the model the web camera will open by that the model show some sign to that then the model will predict the letters. And based on the letter it gives some suggestion that help to analyze faster.

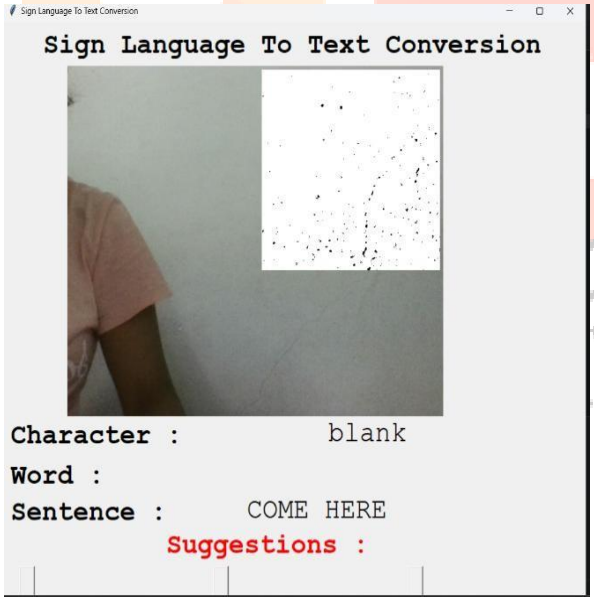

Fig 5 : Text conversion

This is image of final outcome that predict by model as it predicts each wor`ds and give complete sentence that helps the people to understand.

CONCLUSION

System built a vision-based ASL recognizer for the alphabets in real time. It uses CV, processing, deep learning to turn gestures to text, helping interactions. Modern AI tackles these challenges, making talks smoother, though some parts like full sentences or expressions could use more work. It seems promising for inclusion. In American Sign Language recognition for D&M people have been developed for asl alphabets. In this project the model got accuracy of 90% to 92% based on the model training dataset. a clear background and the background should be noise resistant. And also model try to improve the accuracy by using model and train the model in two layers that helps the model to analyse the gesture properly and show the proper results.

8. REFERENCES

- Ashwini M. Deshpande, Snehal R.Kalbhor "Video-Based Marathi Sign Language Recognition and text Conversion using Convolutional Neural Network", Conference Paper 2019.
- Balakrishna Reddy; Ahmed Shahebaaz; Ratnam Dodda; C. Raghavendra, "Transformative Advancements: Sign Language Conversion to Text and Speech", International Conference on Emerging Smart Computing and Informatics (ESCI),2024.
- Aishwarya D Shetty; Jyothi Shetty; Karthik K; Rakshitha; Shabari Shedthi B," Real-Time Translation of Sign Language for Speech Impaired", 7th International Conference on Computing Methodologies and Communication (ICCMC),2023 .
- Parth Upadhaya; Devansh Chamoli; Aman Chopra; Supriya Raheja," Real Time Sign Language Translation by Leveraging Generative AI", International Conference on Computing and Communication Technologies (ICCCT),2025
- Sangeetha Allam; Allam Balaram; Yedavalli Venkata Raghava Rao; Ajmeera Kiran; Hema Valpadasu; Soleti Navya," SIGN LANGUAGE RECOGNITION USING CNN", International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC),2024.
- Oum Tipare; Saniya Pathre; Parth Pathak; Deepak Karia," GestureSpeak: A Real-Time Sign to Speech Translation", 3rd International Conference on Inventive Computing and Informatics (ICICI),2025.
- M. Kowsigan; Rahul Dhawan; Ankan Kundu," An Efficient Speech to Sign Language Conversion and Text Recognition through Live Gesture", IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE),2024.
- Madhu Chaudhary; Saurav Gaur; Riya Sharma; Krishna Bihari Dubey," Sign Language to text Conversion Using Machine Learning (For Deaf and Mute People)", 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N),2024.
- Pushpam Vaidya; Devashish Sathawane; Kartik Kalsarpe; Aditya Narule; Srujal Taksande; Chetan Puri," Empowering Communication: Converting Sign Language to Text Using CNN Technology", 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI),2024.
- Shivani Deshpande; Rajashree ShettarIn, "Hand Gesture Recognition Using MediaPipe and CNN for Indian Sign Language and Conversion to Speech Format for Indian Regional Languages", 7th International Conference on Computation
- System and Information Technology for Sustainable Solutions (CSITSS),2023.
- Anvisha Pathak; Niraj Patil; Prashant Padhy; Adarsh Jadhav; Smita Rukhande; Lakshmi Gadhikar ,"Real-Time Sign Language to Speech Converter Using OpenCV and MediaPipe",2025 International Conference on Emerging Systems and Intelligent Computing (ESIC),2025 .
- Amutha Seviappan; Kothai Ganesan; Anna Anbumozhi; A. Sameer Reddy; B. Vamsi Krishna; D.Sampath Reddy, "Sign Language to Text Conversion using RNN-LSTM International Conference on Data Science, Agents and Artificial Intelligence (ICDSAAI),2023.