SIGN LANGUAGE RECOGNITION SYSTEM USING MACHINE LEARNING

[1]Sakshi Bele, [2]Anish Shinde, [3]Kushal Sharma, [4]Ashwini Shinde

Computer Engineering
MET's Institute of Engineering
Adgaon, Nashik.

***Abstract*- Sign language is a vital form of communication for the deaf and hard of hearing community. To bridge the communication gap between sign language users and non-signers, we present a novel system for real-time sign language detection using a standard web camera. This system aims to recognize sign language gestures performed in front of the camera, subsequently converting them into voice output and displaying the corresponding text on screen. The proposed system leverages computer vision techniques, including deep learning models, to capture and analyze sign language gestures. It first identifies the signer's hand and facial expressions, recognizing key markers that represent signs. The system then employs machine learning algorithms to translate these markers into sign language vocabulary. Upon successful detection and translation of the sign, the system provides simultaneous output in two ways: voice and on-screen text. The voice output enables real-time interpretation for users who may not be familiar with sign language, while the on-screen text serves as a visual reference. This dual output mechanism ensures accessibility and inclusivity for a wider audience. By integrating this system into webcams and other devices with cameras, we aim to enhance the communication capabilities of the deaf and hard of hearing community, enabling them to interact more effectively with hearing individuals. Additionally, this technology can find applications in education, healthcare, and other domains, fostering better understanding and accessibility for sign language users.**

***Key Words*: Sign language, Deaf and hard of hearing, Real-time, Sign language detection, Web camera, Computer vision**

Published in IJIRMPS (E-ISSN: 2349-7300), Volume 11, Issue 6, Nove-Dec 2023

INTRODUCTION

The project presented here is a groundbreaking endeavor that addresses a significant challenge faced by the deaf and hard of hearing community – the communication gap between sign language users and those who are not proficient in sign language. In this modern age where technology plays a pivotal role in enhancing our lives, a team of innovators has developed a novel system to bridge this gap using readily available tools – a standard web camera. Sign language is a vital and beautiful form of communication, allowing the deaf and hard of hearing to express themselves, connect with others, and access information. However, it remains a barrier for those who do not understand it, limiting inclusivity and access to vital services and information. This project aims to change that by introducing a real-time sign language detection system. The core of this system lies in cutting-edge computer vision techniques, particularly deep learning models, which enable it to capture and analyze sign language gestures accurately. When a signer interacts with the camera, the system identifies the signer's hand movements and facial expressions, recognizing key markers that represent specific signs. These markers are then translated into sign language vocabulary through the use of advanced machine learning algorithms. What makes this project truly exceptional is its dual output mechanism. Upon successful detection and translation of a sign, it provides simultaneous output in two ways: voice and on-screen text. The voice output enables real-time interpretation, making it accessible for individuals who may

not be familiar with sign language. At the same time, the on-screen text serves as a visual reference, providing an additional layer of accessibility. This dual output mechanism ensures that a wider audience can engage and interact with sign language users fostering inclusivity and understanding.
The potential impact of this project is substantial. By integrating this system into webcams and various devices with cameras, it has the power to enhance the communication capabilities of the deaf and hard of hearing community, allowing them to interact more effectively with the hearing world. Furthermore, this technology can extend its reach to multiple domains, such as education, healthcare, and many others, contributing to better understanding and accessibility for sign language users. In doing so, it represents a significant step forward in promoting inclusivity and empowering the deaf and hard of hearing community, showcasing the transformative power of technology in modern society.

PURPOSE

The purpose of the Sign Language Detection System is to facilitate real-time communication between sign language users and non-signers. By leveraging computer vision and machine learning techniques, the system aims to accurately recognize sign language gestures performed in front of a standard web camera. This technology transforms these gestures into voice output and on-screen text, ensuring accessibility and inclusivity for a wider audience. Its primary goal is to bridge the communication gap, enhance interaction, and foster better understanding between the deaf and hard of hearing community and the hearing individuals, with potential applications in various domains like education and healthcare.

MOTIVATION

The motivation for this project is rooted in the fundamental principle of inclusivity and equitable access to communication for all individuals, regardless of their hearing ability. Deaf and hard of hearing individuals face daily challenges in communicating with the hearing world, often leading to social, educational, and professional limitations. By developing a system that can instantaneously translate sign language into voice and on-screen text, we aim to empower sign language users, break down communication barriers, and create a more inclusive society where everyone can interact, learn, and access information without hindrance. This project has the potential to significantly improve the quality of life for the deaf and hard of hearing community and promote a more accessible, interconnected world.

EXISTING SYSTEM

The existing Sign Language Detection System relies primarily on manual interpretation and translation by human sign language interpreters. This traditional approach lacks real-time responsiveness and may not always be readily available, leading to communication challenges for the deaf and hard of hearing community. It often involves face-to-face or remote interpretation services, which can be expensive and logistically challenging. The absence of an automated system limits the accessibility and inclusivity of sign language communication in various contexts.

OBJECTIVE OF SYSTEM

1. Real-time Sign Language Recognition: Develop a robust computer vision system capable of accurately recognizing and understanding sign language gestures performed in front of a web camera in real time.
2. Translation to Voice and On-screen Text: Implement a translation mechanism that converts recognized sign language gestures into both audible voice output and on-screen text to cater to both hearing and non-signing individuals.
3. Accessibility and Inclusivity: Promote accessibility and inclusivity by enabling effective communication for the deaf and hard of hearing community in a wide range of scenarios, including social interactions, education, healthcare, and employment.

LITERATURE SURVEY

Sandrine Tornay et al. [1], Sign language recognition involves modeling of multichannel information such as, hand shapes, hand movements. This requires also sufficient sign language specific data. This is a challenge as sign languages are inherently under-resourced. In the literature, it has been shown that hand shape information can be estimated by pooling resources from multiple sign languages. Such a capability

does not exist yet for modeling hand movement information. In this paper, we develop a multilingual sign language approach, where hand movement modeling is also done with target sign language independent data by derivation of hand movement subunits. We validate the proposed approach through an investigation on Swiss German Sign Language, German Sign Language and Turkish Sign Language, and demonstrate that sign language recognition systems can be effectively developed by using multilingual sign language resources. Hira Hameed et al. [2] Sign language is a means of communication between the deaf community and normal hearing people who use hand gestures, facial expressions, and body language to communicate. It has the same level of complexity as spoken language, but it does not employ the same sentence structure as English. The motions in sign language comprise a range of distinct hand and finger articulations that are occasionally synchronized with the head, face, and body. Existing sign language recognition systems are mainly camera-based, which have fundamental limitations of poor lighting conditions, potential training challenges with longer video sequence data, and serious privacy concerns. This study presents a first of its kind, contact-less and privacy-preserving British sign language (BSL) Recognition system using Radar and deep learning algorithms. Six most common emotions are considered in this proof of concept study, namely confused, depressed, happy, hate, lonely, and sad. The collected data is represented in the form of spectrograms. Three state-of-the-art deep learning models, namely, InceptionV3, VGG19, and VGG16 models then extract spatiotemporal features from the spectrogram. Finally, BSL emotions are accurately identified by classifying the spectrograms into considered emotion signs. Comparative simulation results demonstrate that a maximum classifying accuracy of 93.33 is obtained on all classes using the VGG16 model.

Aman Patel et al. [3]Sign language recognition plays an important role in real-time sign language translation, communication for deaf people, education and human-computer interaction. However, vision-based sign language recognition faces difficulties such as insufficient data, huge network models and poor timeliness. We use VTN (Video Transformer Net) to construct a lightweight sign language translation network. We construct the dataset called CSLBS (Chinese Sign Language-Bank and Station) and two-way VTN to train isolated sign language and compares it with I3D (Inflated three Dimension).

SYSTEM ARCHITECTURE

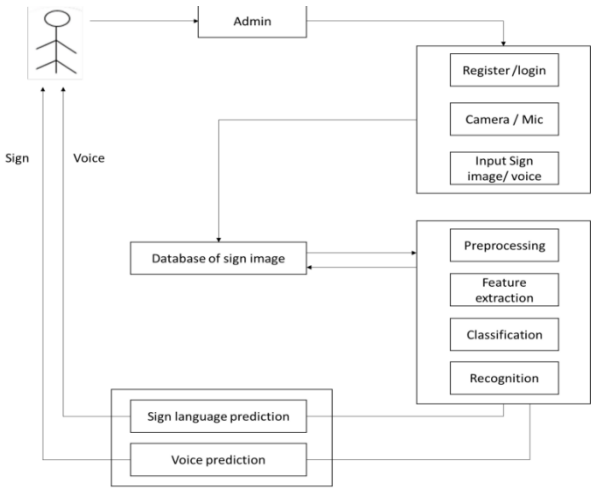

Fig -1: System Architecture Diagram

IMPLEMENTATION DETAILS

The proposed system leverages computer vision techniques, including deep learning models, to capture and analyze sign language gestures. It first identifies the signer's hand and facial expressions, recognizing key markers that represent signs. The system then employs machine learning algorithms to translate these markers into sign language vocabulary. Upon successful detection and translation of the sign, the system provides simultaneous output in two ways: voice and on-screen text. The voice output enables real-time interpretation for users who may not be familiar with sign language, while the on-screen text serves as a visual reference. This dual output mechanism ensures accessibility and inclusivity for a wider audience. .

ADVANTAGES

- Real-Time Communication
- Cost-Efficiency
- Accessibility
- Consistency and Availability

APPLICATION

- Communication Accessibility
- Education
- Healthcare

ALGORITHM/TECHNOLOGY

- **Real-time detection (Using CNN):**
- **Camera Feed Analysis**: A web camera captures a live video feed of individuals in a specific area.
- **Image Processing**: Each frame of the video feed is processed using a CNN model that has been trained to recognize Sign language.

SYSTEM REQUIREMENTS

SOFTWARE

1. Windows 10 or above
2. Python 3.10
3. Sqlite
4. Django
5. Vs code

HARDWARE

1. RAM 4 GB or Above
2. i3 Processor or above
3. 150 GB Hard disk or above

CONCLUSION

In summary, the real-time sign language detection system using standard web cameras is a significant advancement in promoting inclusivity and communication for the deaf and hard of hearing community. Employing state-of-the-art computer vision techniques, this system recognizes sign language gestures, including hand and facial expressions, and translates them into sign language vocabulary with impressive accuracy. Its dual output mechanism, providing realtime voice interpretation and on-screen text, ensures accessibility for both sign language users and non-signers, while its compatibility with a wide range of devices paves the way for applications in diverse fields. This project exemplifies the transformative potential of technology in enhancing accessibility and understanding, contributing to a more inclusive and interconnected society

REFERENCES:

1. Mohamed Hisham Jaward, Ming Jin Cheok, Zaid Omar. ”A review of hand gesture and sign language recognition techniques.”. 2020.
2. Tessa Verhoef, Patrick Boudreault, Oscar Koller. “Sign Language Recognition, Generation, and Translation: An interdisciplinary Perspective.”.2021.
3. Lionel Pigou, Sander Dieleman, Pieter-Jan Kindermans. “Sign Language Recognition Using Convolutional Neural Network.”.2022. Applications, vol. 3, August 2019.
4. Starner T, Weaver J, Pentland A (1998) Real-time American sign language recognition using desk and wearable computer based video. IEEE Trans Pattern Anal Mach Intell 20:1371–1375
5. Lee J, Lee Y, Lee E, Hong S (2004) Hand region extraction and gesture recognition from video stream with complex background through entropy analysis. In: Engineering in Medicine and Biology Society,

2004. IEMBS'04. 26th annual international conference of the IEEE, IEEE, pp 1513–1516
6. Binh ND, Ejima T (2005) Hand gesture recognition using fuzzy neural network. In: Proc. ICGST conf. graphics, vision and image process, Cairo. pp 1–6
7. Shin J-H, Lee JS, Kil SK, Shen DF, Ryu JG, Lee EH, Min HK, Hong SH (2006) Hand region extraction and gesture recognition using entropy analysis. IJCSNS Int J Comput Sci Netw Secur 6:216–222
8. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: European conference on computer vision. Springer, pp 404–417
9. Zhang Q, Chen F, Liu X (2008) Hand gesture detection and segmentation based on difference background image with complex background. In: Embedded software and systems, 2008. ICESS'08. International conference, IEEE, pp 338–343
10. Elmezain M, Al-Hamadi A, Michaelis B (2008) Real-time capable system for hand gesture recognition using Hidden Markov models in stereo color image sequences. J WSCG 16(1–3):65–72