Grenze International Journal of Engineering and Technology, June Issue

Implementing an Interactive VBE and IPSL Aid for the Disabled

Dr. B. Meenakshi Sundaram[1], Dr. B. Rajalakshmi[2], Keshav Krishna Kumar[3], Joel Jacob Stephen[4] and J Mukundh[5]
[1-5]New Horizon College of Engineering, Department of Computer Science and Engineering, Bangalore, India
Email: bmsundaram@gmail.com, {dr_rajalakshmi_imprint, 1nh19cs077.keshavkk, 1nh19cs098.joel.js, mukundh.cs065.bsec }@gmail.com

***Abstract*—In a vast world populated by billions of people, communication is the fundamental aspect that sets us apart as humans. Without the ability to communicate, we would simply exist as soulless objects. Sadly, not everyone possesses the gift of unhindered communication. The interaction between individuals with disabilities and those without presents a formidable challenge. We propose a solution to this predicament by creating a sign language identifier that employs machine learning and web development techniques to assist those in need. Indo-Pakistani Sign Language (IPSL) is one such language that deserves more attention than it currently receives. Despite being actively used by over 15 million members of the deaf community, IPSL remains relatively unknown to the world at large. Our model will make IPSL detectable by integrating it with a Voice Based Email (VBE) system, which will aid not only the deaf but also the blind. To achieve this, we will employ a GRU model.**

***Index Terms*— Sign Language, Disabled, IPSL, VBE.**

I. INTRODUCTION

This paper aims to revolutionize communication by developing a cutting-edge model that bridges the communication gap between the disabled and the able-bodied. Specifically, our model targets the deaf and blind communities as a state-of-the-art communication tool.[16]. Our research focuses on Indo-Pakistani Sign Language (IPSL) and aims to be a megaphone for the millions who communicate using this language. Despite being widely spoken, this language remains shrouded in mystery in Western regions. However, IPSL has recently gained the recognition it deserves and is now on the rise. Our primary goal is to promote IPSL and educate people not only about its existence but also about sign language in general. But before we delve into the technical aspects, it's crucial to understand the community we aim to assist.

II. TERMINOLOGIES

A. Sign Language

1) Definition: Sign languages are an extraordinary class of languages that employ a visual-manual modality to convey meaning.
Deaf and hard-of-hearing persons can communicate with each other and with hearing people by using sign language, a visual language. It is a whole, natural language with its own grammar, syntax, and vocabulary that

Grenze ID: 01.GIJET.9.2.409

are distinct from those found in spoken languages. The American Sign Language (ASL), British Sign Language (BSL), French Sign Language, Japanese Sign Language, and Australian Sign Language are only a few of the several sign languages that are used globally. Although each sign language has its own distinct set of signs, they all have some things in common. Similar to spoken languages, sign languages are capable of communicating complex thoughts and concepts. To transmit meaning, they use body language, facial expressions, and hand gestures. ASL's "apple" sign, for instance, entails creating a fist with the thumb on top, bringing the fist to the mouth, and twisting it to resemble taking a bite. Other indications could include pointing, head movements like shaking or nodding, or emotional responses shown through various facial expressions. Both hearing people and deaf people can benefit from learning sign language. It gives deaf people a way to interact with people who speak the same language and share the same culture. It can make it easier for hearing people to communicate with hard-of-hearing and deaf people. Learning sign language can be done using a variety of materials, such as books, online courses, and classes provided by neighborhood associations. Numerous colleges and universities also include sign language courses in their curricula. Sign language has various purposes besides facilitating communication among the deaf and hard-of-hearing. For instance, it is frequently used in noisy settings where it can be challenging to hear spoken language, such as factories or construction sites. People with autism or cerebral palsy, as well as individuals who are nonverbal or have difficulties communicating, also utilize it.

2) IPSL: The Indo-Pakistani Sign Language, predominantly utilized in Southern Asia encompassing nations such as India, Pakistan, Bangladesh, and Nepal, does not enjoy the same level of notoriety as its peers, namely ASL and BSL. Nevertheless, it is informally employed by members of the deaf community, as it is still in the developmental stages and is gradually earning the acknowledgment it is due. Since it encompasses the nearby regions, it allows for better communication between all the neighboring regions instead of having various other sign languages for the smaller neighboring countries. In the year 2005,[16] the National Curricular Structure (NCF) accorded recognition to Indo-Pakistani Sign Language (IPSL) as a gesture communication language that could be employed if required.

Since the inception of sign language studies, initiated by the interest piqued by American Sign Language (ASL), research themes and geographic regions have progressed significantly. Presently, numerous European sign languages, including French, British, German, and Italian Sign Languages, are under scrutiny from both a linguistic and practical perspective. Despite this, the bulk of research continues to be centered around ASL, and knowledge of other sign languages utilized in other regions of the world remains limited. The investigation of IPSL has primarily been carried out by a few institutions that engage in sign language education in India and Pakistan. In these settings, linguistics, as it is practiced in Western nations, is almost entirely unfamiliar. Karachi, Bangalore, and other Indian cities contribute to most of the signs utilized in IPSL. For the longest time it was believed that IPSL was used way before the ASL came along, after reading the scriptures that mention the use of sign languages in the early Hindu Empire.

IPSL is the most utilized sign language in the Indo-continent, with at least 18 million deaf signers. However, estimating actual numbers is a challenge, and most research has focused more on the urban and northern areas. IPSL is the third most widely used sign language globally, following Chinese Sign Language and Brazilian Sign Language [7]. Experts suggest that Pakistani, Bangladeshi, Indian, and Nepali are varieties of dialects of IPSL, while others regard some varieties as distinct languages. IPSL is distinct from both Hindi and English and has its own grammar, which consists of three aspects: lexicon, syntax, and spatial grammar.

Just like most languages have their own alphabets, similarly in sign language these alphabets are represented to us by the means of using hand symbols, some of which are shown in the figure below. As we can see in the figure below, the hand symbols are given for each of the 26 alphabets that are present to be able to represent the Indo Pakistani Sign Language, and these alphabets are then used to form sentences too. But spelling out each word would take a lot of time and hence like any other system, there are direct and simple words that have been established within the sign language system to be able to talk to each other with ease. Though these words are not known to the common folk, one can easily start talking to a disabled person after learning the language.

B. VBE

An email system that uses voice instructions to transmit and receive messages is known as a voice-based email system. This method allows users who have trouble typing or reading text to transform spoken words into text and vice versa using speech recognition technology. Users of the voice-based email system can speak their email messages, which the system will subsequently translate into text. Incoming emails can also be read aloud to the user by the system. People with disabilities who may find it challenging to operate a conventional keyboard or mouse, such as those who have sight or mobility limitations, can benefit greatly from this technology. Individuals with visual impairments are only capable of sending emails by dictating the complete message to a

Figure 1. The Alphabet represented in IPSL [16]

sighted third party. This third party will then compose the message and transmit it on behalf of the visually impaired individual. This approach, however, is not ideal in such circumstances. It is improbable that a visually impaired individual will frequently seek assistance. Our society, despite criticizing those with exceptional needs for this reason, must provide assistance to individuals. Artificial intelligence (AI) and machine learning algorithms are frequently used to power voice-based email systems, allowing them to gradually learn the user's speech patterns and language preferences. This enables the system to understand the user's orders and requests and to produce more accurate transcriptions and predictions. The ability to keep connected while on the go is one advantage of using a voice-based email system. Instead of stopping to type on a computer or mobile device, users can dictate and send emails while driving, working out, or doing other things. This can speed up productivity and save time. The Voice- Based- Email system has been incorporated in our desktop model to the best of our capabilities and it is able to take the users content and subject through voice and send the email to the required source directly without having to press any buttons hence working perfectly.

III. Literature Survey

The use of gloves is ubiquitous in our lives, from the common laboratory gloves to those employed in dishwashing. Conversely, the Data Glove, a distinctive type of input device, takes the shape of a hand-worn glove. It comprises multiple electronic sensors that can trace the hand's movements,[16] translating them into an input format suitable for robotics and virtual reality. The Data Gloves that offer tactile sensing permit users to manipulate fine movements and experience a virtual item's semblance. This feature necessitates the use of an Analog to Digital Converter (ADC) to convert analogue signals into a digital form. The Flex Sensor [5] is utilized in conjunction with a tactile sensor to recognize American Sign Language, with an accuracy of over 90%.

The Data Glove [8] is an interactive device that resembles a glove and is outfitted with tactile sensing. Its utility extends beyond robotics and virtual reality, allowing for delicate and precise movements. Tactile sensing [2] is a replica of the human sense of touch, which encompasses the ability to sense pressure and temperature simultaneously, as well as torque and linear force.

FEMD (Finger-Earth Mover's Distance) [4] is a model that was developed to match only the fingers and not the hand's shape. This method has increased accuracy and enables the development of a more robust hand recognition system. The developments in depth sensors, such as the Kinect Sensor used in the popular gaming console Wii, have been adapted to this recognition technology, providing opportunities for Human-Computer Interaction (HCI). However, segmentation errors still occur when viewing the hand as a whole. By utilizing depth maps, which contain distance information about objects, in conjunction with the colour image received by the Kinect camera, this approach allows for improved depth threshold segmentation.[16].

The Leap Motion Controller, a portable and affordable gadget, has been adopted in various fields, such as art, virtual reality, gaming, and others. It is intended to detect hand and finger movements in the 3D plane accurately [6]. The gadget features two black-and-white cameras, three infrared LEDs, and an interaction field. An application programming interface (API) for the LMC allows users to access information about the hand and fingers, such as the placement of the finger and the position of the hand.[9] However, only a portion of the

Figure 2. Data Glove for Hand Tracking [17]

information is appropriate to sign language, and precise finger placement does not help in gesture recognition. Nonetheless, when combined with additional data, this information provides more useful information for SLR.

IV. PROPOSED SYSTEM

The proposed system entails the development of a sophisticated web model encompassing a sign language identification model that utilizes action recognition, implemented through Python with the aid of a deep learning model based on the gated recurrent unit.[16] This cutting-edge model shall be trained through the use of video models showcasing the performance of various sign languages, which will then undergo human key estimation for the purpose of identifying the vertices of each sign and feature normalization.

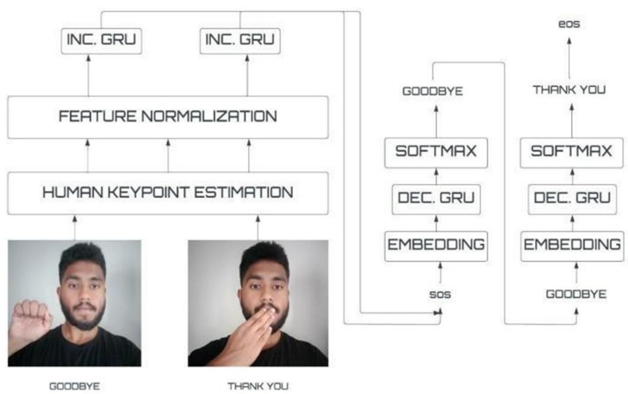

Figure 3. Working Model Architecture [15]

Subsequently, the model will process each sign input, with each subsequent output being considered in the formation of a coherent and meaningful sentence.

This process, as illustrated in the figure above, shall commence with the start-of-sentence (sos) marker and end with the end-of-sentence (eos) marker, upon receipt of which the sentence shall be output to the web model. The web model will also have a VBE component that will aid in the assistance of the blind. This will allow the user to send and receive data without having to type it in, which would be quite challenging for someone who is blind. This information would be passed back and forth in order to create mutually beneficial communication.

V. ACTUAL SYSTEM

A. VBE

The VBE and IPSL sign language modules are visually represented using different web technologies such as VueJS, HTML, and Tailwind CSS for the front-end interface. This is a Python program that enables voice-controlled email service using the Speech Recognition library in Python. The program provides an option to send an email, read the inbox, and exit the service. The main loop of the program asks the user to speak either "send" to send an email, "read" to read the inbox, or "exit" to exit the program. If the user speaks "send", the program calls the sendmail() function to send an email. If the user speaks "read", the program calls the readmail() function to read an email. If the user speaks "exit", the program exits. If the user speaks anything else, the program responds with "Invalid choice" and asks again for the user's choice. Overall, the program provides a simple and intuitive way to control email services using voice commands

B. Desktop App

The VBE and IPSL sign language modules are visually represented using different web technologies such as VueJS, HTML, and Tailwind CSS for the front-end interface, and Node.js and TypeScript for the back-end server. The User Interface is created in a way that allows the different modules to be accessed instantly using a dashboard interface. The UI is clean, easy to use, and fully responsive which allows for easy usability and accessibility across all devices. The frontend interface displays a camera feed for the users to provide the intended signs to the system, gives the input to the sign language converter, accesses the results from the machine language models, and displays the results on the screen. It also provides a module that can detect speech and that can speak text back to the user. This module is also used for the VBE system where each input field is dictated to the user and the response from the user is converted to text and displayed on the screen.

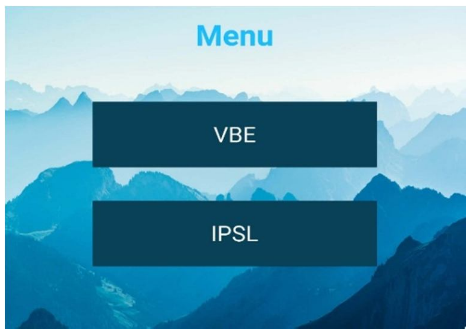

Figure 4. Desktop App

C. Sign Language Model

Sign language motions are often encoded and decoded using a sequence-to-sequence method in a sign language model employing GRU (Gated Recurrent Units).
A series of sign language motions, which can be represented as a series of 3D coordinates or video frames, would be sent into the model as input. The GRU network would receive each coordinate or frame and train to extract the necessary features from the input sequence and encode them into a fixed-length vector. A decoding GRU network would then receive the encoded vector and produce a series of sign language movements in the target language. The next gesture in the series would be produced by the decoder by sampling from a probability distribution over the possible following gestures created from the encoded vector. Once the conclusion of the series is reached, this process would be repeated.
Using a suitable loss function, such as mean squared error or cross-entropy loss, the model would be tuned during training to reduce the difference between the predicted and ground truth gestures. In conclusion, a GRU-based sign language model would encode and decode sign language movements utilizing a sequence-to-sequence methodology, using GRU networks for both encoding and decoding. This model may be used for a range of sign language-related tasks, including translation, gesture recognition, and generation. The sign language model however would require massive GPU power to be able to be trained with such accuracy when using a GRU based model. The LSTM based model works fine for now whilst using lesser GPU power to train the system which can be handled by a normal laptop. The model has been trained with over 10 words to be able to be detected by the system when using the desktop-based application created and is accurately able to predict the sign language using the key point estimation technique shown below.
Sign language Model training was done using the key point estimation method as shown above wherein the model was trained with 30 images of one action for 200 epochs so that when the detector was shown the images in real life, it would be able to estimate what sign is being shown by comparing how its key points hold with the trained images available in the dataset.
Figure 6 is the depiction of the model running and detecting the word ‘Hello’ with an accuracy of 91.4%. This is a universal signal and is recognized the same in Indo Pakistani Sign Language, American Sign Language and other popular languages alike.

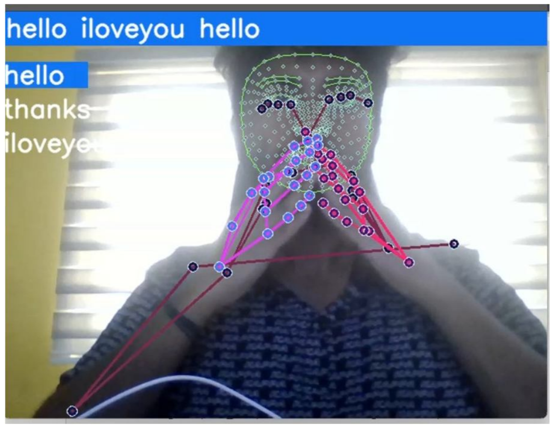

Figure 5. Keypoint Training

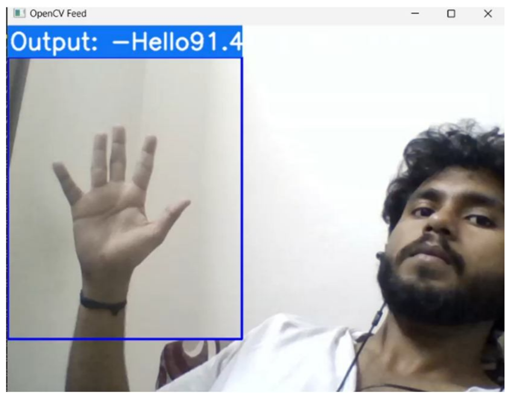

Figure 6. ‘Hello’ Detected

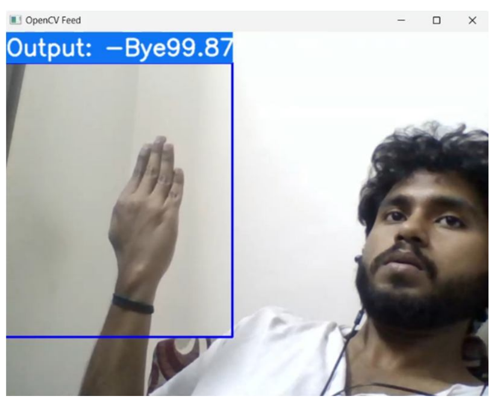

Figure 7. ‘Bye’ Detected

Above is the depiction of the model running and detecting the word ‘Bye’ with an accuracy of 99.87%. This is not a universal signal and is recognized specifically in the Indo Pakistani Sign Language.
Figure 8 is the depiction of the model running and detecting the word ‘Sit’ with an accuracy of 100%. This is not a universal signal and is recognized specifically in the Indo Pakistani Sign Language.

VI. Conclusions

By implementing a sign language recognition system for IPSL (Indo-Pakistani Sign Language), with an average accuracy of 92% to help the deaf and the dumb, and using the VBE system to help the blind, we have surveyed

Figure 8. ‘Sit’ Detected

the market's existing solutions for sign language recognition and voice-based email systems. We have proposed a new system that improves upon the existing systems. The contact between people with disabilities and regular residents would benefit enormously from this.

ACKNOWLEDGMENT

We would like to acknowledge Dr. Meenakshi Sundaram B. for guiding us throughout this paper and helping us to come up with a sophisticated model by pushing us to what we are capable of achieving. We are glad we got to work on this paper and make it come to life.

REFERENCES

[1] Barathi Subramanian, Bekhzod Olimov, Shraddha M. Naik, Sangchul Kim-Houm Park & Jeonghong Kim, An Integrated mediapipe- optimized GRU model for Indian Sign language recognition.
[2] Manisha U. Kakde, Mahender G. Nakrani, Amit M. Rawate A Review on Sign Language Recognition System for Deaf and Dumb People using Image Processing.
[3] Satwik Ram Kodandaram, N Pavan Kumar, Sunil GL Sign Language Recognition.
[4] Ashwini Kolhe, R. R. Itkarkar, Anilkumar V. Nandani RobustPart-Based Hand Gesture Recognition Using Finger-Earth Mover's Distance.
[5] Anarbasi Rajamohan, Hemavathy R., Dhanalakshmi M., Deaf Mute Communication Interpreter.
[6] P. Karthick, N. Pratibha, V.B. Rekha, S. Thanalaxmi, Transforming Indian Sign Language into Text Using Leap Motion.
[7] Ulrike Zeshan Sign Language in Indo-Pakistan
[8] Anand Nayyar, Vikram Puri, Data Glove: Internet of Things (IoT) Based Smart Wearable Gadget.
[9] Ma X., Peng J., Kinect sensor-based long-distance hand gesture recognition and fingertip detection with depth information.
[10] Munir Oudah, Ali Al-Naji, and Javaan Chahl, Hand Gesture Recognition Based on Computer Vision: A Review of Techniques
[11] Simeon Kostadinov, Understanding GRU Networks Amrita Thakur, Pujan Badhathoki, Sarmila Upreti, Shrish Shreshta, Subarna Shakya, Real-Time Sign Language Recognition and Speech Recognition.
[12] Shagun Katoch, Varsha Singh, Uma Shanker Tiwary, Indian Sign Language Recognition system using SURD with SVM and CNN.
[13] Kruti J Danngarwala, Dr. Dilendra Hiran, A Research Gap on Automatic Indian Sign Language Recognition on Hand Gesture Datasets and Methodologies.
[14] Ahmed K, Ahmes Eltayeb Ahmed, Omar Al-Hardanee, Arif Yilmaz, DeepASLR: A CNN-based human-computer interface for American Sign Language recognition for hearing-impaired individuals.
[15] Keshav Krishna Kumar, Joel Jacob Stephen, Mukundh J, Interactive VBE and IPSL Aid for the Disabled using a GRU-based Model
[16] https://www.istockphoto.com/vector/deaf-mutes-hand-language-learning-alphabet-nonverbal-deaf-mute-communication-gm1220106111-357129310
[17] https://he.aliexpress.com/item/1005004155242150.html