Machine Learning Approach to Sign Language Recognition using Teachable Machine

Pranay Patil
Dr. D Y Patil Institute of Technology, Pune, Maharashtra, India
Savitribai Phule Pune University
pranay32patil@gmail.com

Vaishnavi Kolhe
Dr. D Y Patil Institute of Technology, Pune, Maharashtra, India
Savitribai Phule Pune University
vaishnavikolhe1906@gmail.com

Kunal Giradkar
Dr. D Y Patil Institute of Technology, Pune, Maharashtra, India
Savitribai Phule Pune University
33kunalg@gmail.com

Dr. Anup Vibhute
Dr. D Y Patil Institute of Technology, Pune, Maharashtra, India
Savitribai Phule Pune University
anup.vibhute@dypvp.edu.in

***Abstract*— Deaf people worldwide rely on sign language as their only reliable form of communication with one another and with the broader population. Hand form and movement are used to create these communicating signals. Deaf people in India interact with others using Indian sign language (ISL). Many scholarly papers on ISL recognition and categorization have been published. The bulk of these projects made use of colored hands, but others made use of sensors and Kinect-based approaches, and they are expensive and not scalable. Our study proposes a technique for recognizing sentences with bare hands, highlights the elements investigated in sign language identification, offers the datasets available, and provides trends for future research. In contrast to recent published work or projects, our model obtains an excellent accuracy of 90-93%.**

Keywords* - *Sign language, image recognition, machine learning, features extraction

I. INTRODUCTION

language is the primary mode of communication and interaction for deaf people all over the world. This type of communication is performed by the use of hand gestures, facial expressions, or arm/body movement. The sign language recognition system attempts to help the deaf community communicate effectively with the rest of society. It is a highly structured symbolic set that enables human-computer interaction (HCI). Sign language is an extremely useful communication tool, and millions of deaf people all over the world use it every day to communicate and express themselves. This facilitation and help to deaf people enables and encourages them to be a healthy part of society and integrates them into society. Sign language changes as you go from one country/region to another, such as American, Japanese, Chinese, and Arabic sign language. Static and dynamic sign language are the two broad kinds of sign language. Static sign language considers a fixed stance of the hand, whereas dynamic sign language considers mobility or movement of the hand. Many scholars have already given various ways for the recognition of various sign languages used all over the world. The majority of the work has been completed in American sign language, Chinese sign language, and Arabic sign language. Some review publications [2–7] describe the mainstream and state-of-the-art in generic Sign Language Recognition systems in recent years, with minimal attention on Chinese Sign Language (CSL). For example, [6] discussed some of the difficulties and issues encountered when developing a sign language recognition system. Reference [5] offered a review of sign capturing methods and classification approaches, as well as issues with continuous SLR, before introducing Malaysian sign language. In Arabic Sign Language Recognition, reference [3] reviewed vision-based and sensor-based techniques. [4] discussed the current recognition strategies in vision-based SLR. [7] presented numerous strategies for hand segmentation and tracking, feature extraction, and sign language categorization.

II. MOTIVATION

Deaf communities around the world rely heavily on sign language as their primary means of communication. Despite their critical importance, the availability and effectiveness of communication tools for the deaf, especially those using sign language, is still limited. Currently, many existing sign language recognition systems are mostly based on resource-intensive approaches. Those methods, which involve complex sensor-based systems or dye-dependent techniques, face scalability and financial limitations, making them less practical for widespread adoption. In response to these challenges, there is an urgent need to develop an easier-to-use, more cost-effective and more accurate system that enables sign language recognition without expensive sensors or color-based segmentation. The effort to create a robust and effective sign language recognition system stems from the central goal of promoting inclusion and equal communication

979-8-3503-1545-5/23/$31.00 ©2023 IEEE

in the global deaf community. By overcoming the limitations of current approaches, an innovative and scalable recognition system can significantly improve deaf communication and integration into society. Therefore, this effort is motivated by the effort to develop an advanced and easy-to-use sign language recognition method that uses machine learning and image processing and ensures accuracy, affordability and wide applicability. Such a system would not only strengthen the Deaf community, but also promote greater inclusion and understanding between different linguistic and cultural landscapes.

III. LITERATURE SURVEY

Technical Approaches to Chinese Sign Language Processing [10]

This research paper [10] on Chinese Sign Language Processing provides a thorough analysis of the difficulties encountered by China's deaf and hard-of-hearing population, highlighting the significance of Sign Language Recognition (SLR) systems in bridging the communication gap. It uses vision-based or sensor-based techniques to classify SLR into fingerspelling, single word, and continuous sign phrase recognition.

Reviewing several approaches, the paper focuses on CSL recognition systems. Because fingerspelling is static, studies using neural networks, SVMs, and feature descriptors have achieved great accuracy. Using Kinect or sensor gloves, isolated word recognition uses complex models such as CNNs, LSTMs, and HMMs that have demonstrated outstanding accuracy for large vocabulary sets. The study offers several strategies that combine various variables and classifiers to improve recognition accuracy. It talks about developments.

Two-Stage Deep Learning Solution for Continuous Arabic Sign Language Recognition Using Word Count Prediction and Motion Images [11]

The research paper [11] presents a continuous Arabic Sign Language (ASL) recognition method using a two-stage deep learning methodology. First, a sequence-to-label classifier estimates the number of words in the phrase. Next, a sequence-to-sequence classifier employs motion images that are produced by deducting successive video frames. These motion pictures are produced by motion estimation and compensation methods, and a CNN Inception-v3 network is used to extract features from them. A sequence-to-sequence neural network with LSTM or biLSTM layers processes the resultant sequence of feature vectors in order to recognise the data. The approach to forecast the number of words in a phrase for segmentation is proposed, and the research admits the difficulty of segmenting sentences during testing when ground truth labels are not available. But it also emphasizes the necessity of more testing and modification for use with more languages and in real-world scenarios.

A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician[12]

The technology described in the research paper[12] has components for synthesis and recognition. Synthesis uses a signing avatar to convert text into dynamic sign language, while recognition helps doctors understand patient sign language. Deeper analysis requires the incorporation of prior sign language knowledge, which is made possible by Hidden Markov Models (HMMs). With their roots in speech recognition, HMMs provide a flexible and user-friendly pattern analysis toolkit. Their choice is consistent with the sequential structure of sign language and the common initial motions of signs. The synthesis module is primarily concerned with setting up an accurate signing avatar, which requires complex motion control and collision avoidance. Motion capture (MoCap) technology records the movements of the body and limbs, but it requires manual post-production adjustments for finger movements. These adjustments are made utilizing Kinect-captured arm motions and extra finger animations to achieve sign precision.

Enabling Two-Way Communication of Deaf Using Saudi Sign Language[13]

The SDCS system, which enables two-way communication for Saudi sign language users, is introduced in this Paper[13]. This technology, which consists of three main components (Avatar (AM), Speech Recognition and Synthesis (SRSM), and Sign Recognition (SRM), allows people with hearing impairments to interact efficiently. A condensed 3D Graph Convolutional Network (3DGCN) with 25 important landmarks taken from MediaPipe is used by the SRM. The Kaldi toolkit is used for voice recognition, and FastSpeech2 is used in SRSM for text-to-speech synthesis. Using an avatar, the AM converts text that has been recognized into sign language. Through topic-based messaging, ROS enables smooth interaction between various modules. Spectrograms produced by SRSM show that it can generate speech for sign language signs that is of a high caliber.

User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features[14]

The work presents a Microsoft Kinect depth image-based, user-independent method for alphabet recognition in American Sign Language (ASL). Complex backgrounds and problems with hand segmentation plague traditional color-based ASL identification methods. By using depth information, the suggested approach gets beyond these drawbacks and ensures robustness against variations in light. A preprocessing method is used to simplify hand segmentation, and PCANet, a straightforward unsupervised deep learning architecture, is used for feature learning. The method trains individual models for each user or a single model from samples of all users. Tests conducted on an actual depth picture dataset demonstrate its superiority in signer-independent recognition compared to the most advanced techniques.

Generalization of Bangla Sign Language Recognition Using Angular Loss Functions[15]

This research paper examines the intricacies of BSL recognition. It highlights the difficulties that existing models face in adapting to diverse datasets. Previous studies have shown satisfactory results on small datasets but failed to generalize across different datasets. This paper examines the angular loss functions, SphereFace, CosFace and ArcFace, which are well-known for their performance in face recognition tasks. The paper examines two main datasets: the Ishtar-Lipi dataset and the BdSL dataset. It looks at various architectures, especially at VGG19 model. It also performs intra and inter-dataset evaluations. This analysis provides important insights into model performance over different datasets. One important finding is that adjusting fully connected layers in model architecture improves generalization performance significantly. This is important to avoid overfitting as models tend to remember specific dataset features, which prevents them from performing well on new unseen data. The results of the study's experiments and comparisons across various configurations demonstrate the effectiveness of the angular loss function in increasing model resilience. By focusing on generalization and dataset robustness, the research emphasizes the importance of developing BSL systems that can overcome specific dataset constraints. All in all, this paper highlights the importance of not only achieving high accuracy in a limited dataset, but also ensuring model adaptability across diverse datasets. The paper highlights the need for more in-depth research and analysis within the BSL recognition domain, with a focus on generalizing models for wider real-world use.

Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language[16]

In this paper, we will look at SLR (Sign Language Recognition) and PSL (Pakistan Sign Language). We will also look at some of the limitations of the existing approaches and how we can overcome them. Most of these approaches have limitations and issues related to specific scenarios, such as using a Kinect device, color gloves, flex sensor, or statistical template match, each with different accuracy rates. These approaches also face limitations in scale, cost, and accuracy when handling dynamic or different hand orientation. In order to overcome these limitations, we will introduce a new approach to PSL recognition: Multiple Kernel Learning (MKL) in Support vector Machines (SVM). MKL is different from traditional approaches because it integrates three kernel functions: Gaussian, linear, and polynomial. This method uses segmentation techniques, Grayscale conversion, feature extraction methods, and EOH (Edge Orientation Histogram) and LBP (Local Binary Patterns) (Local Binary Patterns). SURF (Surr) detects regions of interest and captures texture and gradient information. The workflow of the proposed method is as follows: Image acquisition Grayscale conversion SURF Image segmentation EOH (Histogram of Oriented.In addition, the paper introduces a kernel selection algorithm (KSA) to select the most appropriate kernel function between Gaussian, linear and polynomial to improve the classification accuracy of PSL. The study also provides an overview of K-means cluster-based segmentation-based algorithms and their role in edge orientation histogram calculation, which are essential for PSL classification. The paper presents a method to optimize feature extraction and classification in SVM using MKL, which aims to overcome the shortcomings of PSL classification methods by increasing accuracy, scaling and adaptability to hand orientation and environmental conditions.

Shape-Based Pakistan Sign Language Categorization Using Statistical Features and Support Vector Machines[17]

This paper[17] provides an automated classification method for Pakistan Sign Language alphabet categorization based on statistical features and support vector machines (SVMs) to facilitate communication for deaf community in Pakistan. The method focuses on recognizing PSL signs by hand shape, orientation, and visibility. The method divides the PSL alphabet into seven categories based on finger and hand shape and orientation. To classify the PSL alphabet, histograms of uniform LBP at various adjacent distances are calculated. The histograms are combined into a single vector and six statistical features are extracted from them. The classification is performed using multi class SVMs, using a one versus one mode of binary SVMs to categorize the PSL alphabet. The proposed technique is supported by a dataset (3414 PSL signs) collected from 7 native signers. The study encompasses the existing work on sign language recognition and highlights the various techniques used in sign language recognition around the world, such as American Sign Language (ASL), German Sign Language (GSL), and others. In the past, techniques such as SIFT, Kinect devices, SVMs, HMMs, and feature extraction methods such as LBP have been used for sign language recognition. The proposed solution starts with preprocessing steps that include skin detection using HSV color features to segment the hands from the background. The Uniform LBP technique is used, where histograms are extracted from 16x16 blocks on top of each other, after which statistical features are calculated from these histograms. The study shows the effectiveness of these methods to achieve an overall accuracy of 77.18% for PSL alphabet classification. The paper and comprehensive study of previous work, detailed methodology and performance evaluation present an innovative approach for automatic recognition and classification of PSL signs, highlighting its importance for the deaf in Pakistan.

IV. RELATED WORK

Some review publications [2–7] describe the mainstream and state-of-the-art in generic Sign Language Recognition systems in recent years, with minimal attention on Chinese Sign Language (CSL). For example, [6] discussed some of the difficulties and issues encountered when developing a sign language recognition system. Reference [5] offered a review of sign capturing methods and classification approaches, as well as issues with continuous SLR, before introducing Malaysian sign language. In Arabic Sign Language Recognition, reference [3] reviewed vision-based and sensor-based techniques. [4] discussed the current recognition strategies in vision-based SLR. [7] presented numerous strategies for hand segmentation and tracking, feature extraction, and sign language categorization.[7] made use of the Kinect gadget and coloured gloves. Jitcharoenport, Rujira, and colleagues [8] employed flex sensors and gyroscopes to recognise Thai sign language. Nada B Ibrahim et al. [9] extracted several hand features to recognise Arabic sign language. et al. [6] present fused features mining. They extracted the hand traits used by the artificial neural network (ANN) method to classify them. Although their method claims an error rate of 0.8, their dictionary size is quite small, with only 8 ASL alphabets. Sumaira Kausar et al. [2] used a fuzzy classifier strategy to recognise Pakistan sign language. They have recognised the various finger positions and hand orientation, which is extracted using coloured marking gloves. Aleem Khalid Alvi el.For Pakistan sign language recognition, [1] employed a statistical template matching technique. The mean value and standard deviation of the sensors for a gesture are detected, and PSL may be recognised up to 78.2% of the time; however, the accuracy is impacted by environmental changes. Sensors and gloves were also used in this manner. Muhammad Raees et al. [10] perform image-based recognition of Pakistani sign language.

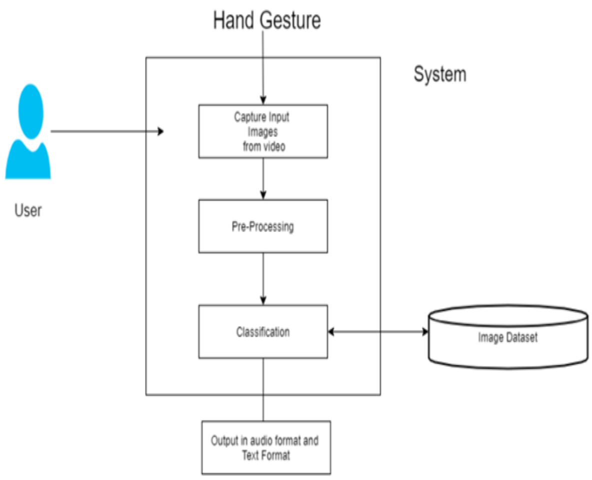
V. PROPOSED SOLUTION

Implementing a hand gesture detection system utilizing computer vision techniques and a webcam by utilizing various important modules and libraries. The following are the primary modules

CV2:This module includes essential features for image processing, video capture, and sketching on images. HandDetector : This module focuses on spotting and tracking hands in moving images, allowing it to detect hand gestures.

Classifier: This module is responsible for classifying hand movements using previously learned models. It takes in an image and returns the desired label for the gesture.

To identify hand motions, the programme uses computer vision techniques and a webcam. First, the image processing and gesture categorization modules and libraries are loaded. The code then establishes the camera connection in order to record video frames. Within the main loop, the code uses the HandDetector module to locate a hand in the video frame. When a hand is found, the code continues to handle it. It creates a bounding box around the hand and then utilizes the box coordinates to determine the area of interest (ROI). The ROI is then downsized to a standard size using mathematical procedures that preserve the aspect ratio of the hand. The ROI is scaled before being placed on a white canvas cut to the desired size. The image is now ready to be classified. The Classifier module is used to categorize the gesture based on the prepared image. The expected gesture and its corresponding index are obtained. To offer visual feedback, the approach employs OpenCV's drawing capabilities to mark the original image with the predicted gesture label. The annotated image is displayed alongside the bounding box that surrounds the hand. Additional intermediate images, such as the clipped area and the scaled image on the white canvas, are displayed for debugging purposes. The method, which runs continuously while capturing frames, detecting hands, classifying movements, and presenting the results, enables real-time hand gesture detection using the webcam.



1. Hyper-parameter tuning

All image data, have their own characteristics depending on the imaging environment, application field, and nature of the acquisition device. Therefore, to appropriately extract each image's characteristics during machine learning, variables that affect the training results must be set in the optimal condition. The process of increasing the learning accuracy by adjusting variables according to the image data's characteristics of the image data to be learned is called hyper-parameter tuning. It sets the optimal variables to obtain the best learning effects in a learning machine. The Teachable Machine has three learning parameters, which are epoch, batch size, and learning rate, and machine learning is performed by adjusting these parameters. An increased number of learnings improve accuracy. However, if the number of learnings is set too high, the model is over-optimized for only the training data due to overstaffing. This increases the accuracy of the training data, but it also decreases the data's accuracy. The learning rate is a variable that determines the step size for calculating the loss function. If it is set too high, learning is not performed. If the rate is set too low, a great amount of learning time is consumed, and the model may be trapped at the local minima in the gradient descent method, which is a learning algorithm of machine learning. This may cause a decrease in diagnostic accuracy. In this study, model was calculated by changing the three parameters of an epoch, batch size, and learning rate that affect the learning model's accuracy. Then, test data was entered into the trained model, and the learning accuracy was evaluated to determine the optimal parameter condition.

2. Teachable machine

After the dataset has been generated and all of the images required to develop the model have been collected, we must now develop the machine learning model that can help us predict the hand sign, as well as convert it into a file format that can be easily integrated into a mobile application understandable format. The model can be generated in several ways. You can either develop your own Python scripts using various machine learning models, or you can utilise the OpenCV to generate the model, etc.

To create a machine learning model, we will use a Teachable machine. Teachable Machine is a web-based application that can assist you in creating an effective and efficient machine learning model. On Google's Teachable Machine website, there are three options for creating machine learning: picture classification mode, speech classification, and pose classification. The classification approach of machine learning is utilized by teachable machines for model creation. In our situation, we delivered images relevant to their sign language to distinct classes and passed them to the Teachable machine to construct the machine learning model. You can easily add additional accuracy to the developing machine learning model before model generation begins.

3. Google's MediaPipe Hand Tracking

According to Google, MediaPipe is the most straightforward approach for researchers and developers to create Machine Learning (ML) applications for mobile, edge, cloud, and the web. It is a node-based framework for building multiple modes (video, audio, and sensor) machine learning pipelines for researchers, students, and software developers who need production-ready ML applications, research work that goes along with publishing code, and building technology prototypes.It implies that there are more reasons to study sign language, not just for communication but also as a technique of human-computer connection. As demonstrated in Figure 8, MediaPipe not only does Hand Tracking but also Face Mesh, which is vital for predicting facial expressions when practicing sign language[3].

Unfortunately, MediaPipe is still in its early stages when this article is published, therefore with limited resources, it is difficult to give proof that sign language may be used utilizing this modern technology. According to the MediaPipe website, the installation status for the Windows platform is still Experimental. According to MediaPipe's Github users, the researcher's attempt to build the desktop app failed due to an incompatible Bazel version (3.5.0, and author has tried the updated 3.6.0 version as well, but still returns android_sdk_repository path undefined while the build was for desktop, not Android). This is also a problem with the macOS and Ubuntu operating systems.

VI. DATA ACQUISITION PROCEDURE

1. Simulation Environment:-

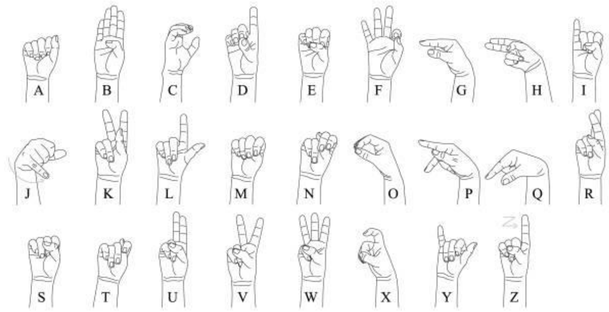
The OpenCV library was utilized to capture 2D images from an RGB webcam and visualize 3D hand landmarks extracted through the MediaPipe hand tracking model. Deep learning techniques were applied within the PyCharm Python notebook web application for implementation.

2. Image collection-

The experimental dataset used for the study of the proposed algorithm was generated using MediaPipe, an open-source library. This tool captures 3D hand gestures, transforming them into NumPy arrays suitable for training the model. The captured images are organized and were saved in a jpg format of 95 ×95 pixels. And the dataset includes essential parameters such as the x, y, and z coordinates of the hand joints.Depth filtering is achieved by analyzing the joints identified through skeletal tracking. This process involves saving images where the depth coordinates are smaller than those associated with the wrist joints, resulting in a precise segmentation of the hands.

It's important to note that the hand images used for detecting dynamic signs are primarily focused on the front view poses of the hands. While a sign may not perfectly correspond to the image of an alphabet letter due to variations in starting positions (e.g., downward, upward, vertical, or horizontal), this limitation has been addressed by acquiring diverse databases captured from various hand angles. These databases are instrumental in identifying both alphabetic letters and medical context words.

The dataset comprises American Sign Language alphabetic signs. To create this dataset, a minimum of 300 photos were captured for each letter, with each sample collected from various angles, introducing slight variations in each instance. Multiple images were manually captured, ensuring comprehensive coverage. Each sign was captured individually, with a brief pause to transition from one alphabet to another. The collected photographs were then saved in the appropriate location. These steps facilitated the creation of a diverse and comprehensive dataset for the algorithmic study.For feature extraction, the images were preprocessed, and the dimensionality of the data was reduced.



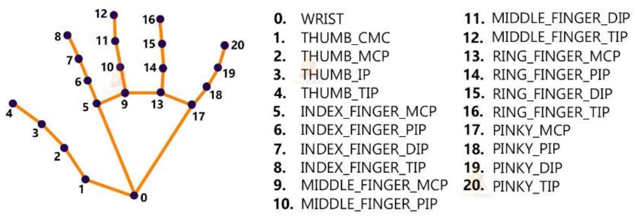
3. Sign Language Recognition and Preprocessing

Initially, the images were captured in RGB color. The code then established variables for offset (an additional margin around the hand bounding box), imgSize (the desired dimensions for the cropped and resized hand image), and counter (a variable used to keep track of the saved image count).

The contour of the hand posture was represented as a series of X and Y coordinates. To determine the uppermost and lowermost points in each column of the blob, the top and bottom edge methods were employed. These points, denoted by their X and Y coordinates, served as inputs for the model, enabling the classification of the hand shape associated with the sign.

Furthermore, the X and Y coordinates of the left or the right hand, describing the movement of the gesture, were stored for later use in both the training and testing phases.

Given variations in heights and arm lengths among the participating volunteers, the coordinates were normalized to a range between 0 and 1. This normalization process aimed to reduce discrepancies in measurements caused by individual differences. Subsequently, the coordinates corresponding to the sign were smoothed to minimize noise resulting from involuntary movements made by the users.



4.Training of Dataset:

Once the dataset has been generated, and all the necessary images have been collected, the next step involves developing a machine learning model capable of predicting hand signs and converting them into a file format compatible with easy integration into a mobile application. There are various methods to generate this model. One option is to create custom Python scripts utilizing different machine learning models. Alternatively, OpenCV can be employed for model development.

In our case, we utilized Teachable Machine, a web-based application that facilitates the creation of effective machine learning models. The Teachable Machine platform offers three options for model creation: picture classification mode, speech classification, and pose classification. For our sign language recognition system, we employed the picture classification mode. We categorized the images corresponding to different sign language gestures into distinct classes and used Teachable Machine to construct the machine learning model.

Additionally, Teachable Machine allows for the incorporation of additional accuracy adjustments before the model generation process begins. This flexibility ensures the refinement of the machine learning model to enhance its performance.

VII. FINDINGS AND DISCUSSION

A. Findings

Real-time Communication Understand the importance of the proposed methodology to enable real-time sign to text conversion, to bridge communication gaps between people who are deaf-mute and others. Enhanced Accessibility Understand the system's potential as a "voice for the deaf" by overcoming the requirement for translators and allowing dynamic two-way communication Potential Improvements Understand the future opportunities for improving user-friendliness and portability, as well as compatibility for different signs and mobile devices with built-in cameras.

B. Discussion

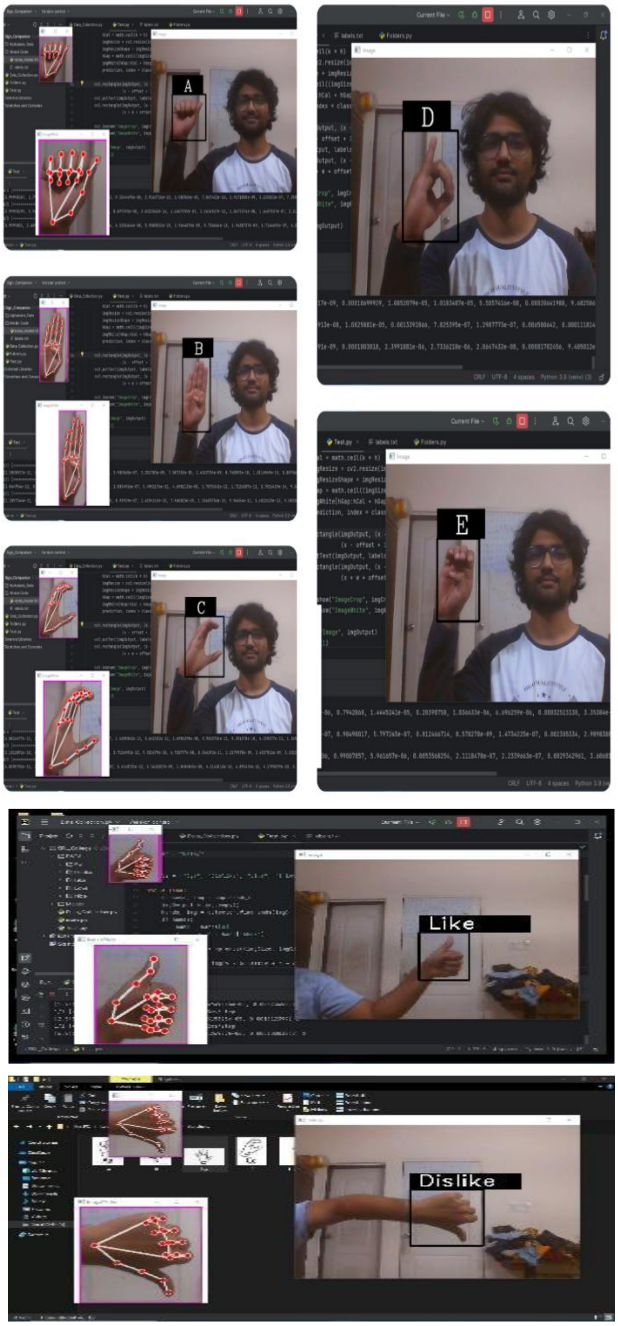
In this section, we will discuss the proposed improvements to the system, such as integrating gTTS to enable text to speech transformation, multi-lingual support, and improving image processing for two-way communication. We will also focus on expanding the language coverage to include more languages, such as Hindi and Marathi, for greater usability. Finally, we will look at future system scalability to address the potential to extend system functionality and range by adding longer trans-receivers or Wi-Fi connectivity. We will also review and compare existing systems, such as SDCS for Saudi Sign Language, and other research papers related to sign language recognition. Lastly, we will discuss how integration or comparison to other sign language recognition systems can contribute to the collective advancement of this domain.

VIII. RESULTS

We highlight the model's obtained accuracy and other pertinent metrics as we provide the results of its performance on an independent test set. We examine performance variances over a range of hand motions and difficult situations, such as various illumination conditions or hand orientations. We also highlight the benefits and advancements made by our proposed model by contrasting its performance with baseline methods or other hand motion detection algorithms.

The created system can instantly recognize American Sign Language alphabets. Google Media Pipe has been used in the system's development. The Google Teachable Machine was used to train the model. On the newly constructed dataset, which consists of 4500 photos overall and 300 images for each alphabet, it has been trained using transfer learning.

The system's output is based on the confidence level, and that level is currently 85.45% on average. By expanding the dataset, the system's confidence level can be raised, improving the system's capacity for recognition. As a result, the system's performance is enhanced.



IX. CONCLUSIONS

Sign Language is a tool to reduce the communication gap between deaf-mute people and normal people. This system which is proposed above gives the methodology which aims to do the same as the two-way communication is possible. This method proposed here facilitates the conversion on the sign into text for now and with improvements will convert text to speech. This overcomes the requirement of a translator since real time conversion is used. The system acts as a voice of the person who is deaf-mute. This project is a step towards helping specially challenged people. This will be further enhanced by making it more user friendly, efficient, portable, compatible for more signs and as well as dynamic signs. This can be further improved so as to make it compatible for the mobile phones using the built-in camera of the phone. We can increase the distance at which it can be used by using a longer trans- receiver module or over Wi-Fi.

X. FUTURE WORK

In future work, a proposed system that uses the current system and uses gTTS (Google Text-to-Speech), a Python library and CLI tool to connect with Google Translates text-to-speech API, to transform the text output into speech. It will then be transformed into numerous regional languages, resulting in a multilingual system. In order for the system to communicate in both directions—that is, to translate from normal English to sign language and vice versa—the image processing portion needs to be enhanced. We shall make an effort to identify indications that involve motion. Additionally, our attention will be directed towards transforming the series of motions into text, or words and sentences, and ultimately into audible speech. We are going to work on integrating support for more languages into our system, such as Hindi, Marathi, and other regional languages.

XI. REFERENCES

[1] Jayshree R. Pansare, Maya Ingle (2016),” Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram”, at 2016 International Conference on Image, Vision and Computing.

[2] Arslan Arif, Syed Tahir Hussain Rizvi, Iqra Jawaid, Muhammad Adam Waleed, Techno-Talk: An American Sign Language (ASL) Translator, at CoDIT’16 - April 6-8, 2016, Malta.

[3] Justin K. Chen, Debabrata Sengupta, Rukmani Ravi Sundaram,” Sign Language Gesture Recognition with Unsupervised Feature Learning”.

[4] Matheesha Fernando, Janaka Wijayanayaka,”Low cost approach for Real Time Sign Language Recognition” at 2013 IEEE 8th International Conference on Industrial and Information Systems, ICIIS 2013, Aug. 18-20, 2013, Sri.

[5] Aarthi M Vijayalakshmi, “Sign Language to Speech Conversion”, At 2016 Fifth International Conference on recent trends in information.

[6] Caixia Wu, Chong Pan,YufengJin, Shengli Sun, and Guangyi Shi Shaoxing, “Improvement of Chinese Sign Language Translation System based on Collaboration of Arm and Finger Sensing Nodes”, At The 6th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems June 19-22, 2016.

[7] Poonam Chavan1, Prof. Tushar Ghorpade2, Prof. Puja Padiya, “Indian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier” at 2016 Symposium on Colossal Data Analysis and Networking (CDAN).

[8] Manisha D. Raut, Ms. Pallavi Dhok, Mr. Ketan Machhale, Ms.Jaspreet Manjeet Hora, “A System forRecognition of Indian SignLanguage for Deaf People using Otsu’s Algorithm”, International Research Journal of Engineering and Technology (IRJET).

[9] C. Swapna, Shabnam Shaikh “Literature survey on hand gesture recognition for video processing” International Journal of Advanced Research in Computer Sci.Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak (Feb 2015), “Study of Sign Language Translation using Gesture Recognition”, International Journal of Advanced Research in Computer and Communication Engineering.

[10] S. M. Kamal, Y. Chen, S. Li, X. Shi, and J. Zheng, "Technical Approaches to Chinese Sign Language Processing: A Review," in IEEE Transactions[2019].

[11] T. Shanableh, "Two-Stage Deep Learning Solution for Continuous Arabic Sign Language Recognition Using Word Count Prediction and Motion Images," IEEE Transactions.[2023]

[12] C. O. Sosa-Jiménez, H. V. Ríos-Figueroa, and A. L. Solís-González-Cosío, "A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician[2022]

[13]M. Faisal, M. Alsulaiman, M. Mekhtiche, B. M. Abdelkader, M. Algabri, T. B. S. Alrayes, G. Muhammad, H. Mathkour, W. Abdul, Y. Alohali, M. Al-Hammadi, H. Altaheri, and T. Alfakih, "Enabling Two-Way Communication of Deaf Using Saudi Sign Language," IEEE.[2023]

[14]W. Aly, S. Aly, and S. Almotairi, "User-Independent American Sign Language Alphabet Recognition Based on Depth Image and PCANet Features," in IEEE Transactions on Image Processing, vol. 7, 2019, pp. 123138

[15]S. K. Youme, T. A. Chowdhury, H. Ahamed, M. S. Abid, L. Chowdhury, and N. Mohammed, "Generalization of Bangla Sign Language Recognition Using Angular Loss Functions," IEEE Access, vol. 9, pp. 165353-165356, 2021. DOI: 10.1109/ACCESS.2021.3134903.

[16] F. Shah, M. S. Shah, W. Akram, A. Manzoor, R. O. Mahmoud, and D. S. Abdelminaam, "Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language," in IEEE Access, vol. 9, pp. 67549-67550, 2021, doi: 10.1109/ACCESS.2021.3077386.

[17]S. M. S. Shah et al., "Shape-Based Pakistan Sign Language Categorization Using Statistical Features and Support Vector Machines," in IEEE Access, vol. 6, pp. 59243-59245, 2018, doi: 10.1109/ACCESS.2018.2872670.