Journal of Information Systems Engineering and Management
2025, 10(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

HearingHands : Amplifying the Voices of Deaf Community

Deepak T. Mane[1*], Devata R. Anekar[2], Yogeshwari V. Mahajan[3], Vinod S. Mahajan[4], Veena M. Kadam[5], Maqsood Ahmed Ansari[6]

[1,2]Vishwakarma Institute of Technology, Bibwewadi, Pune-411037, Maharashtra, India

[3]Pimpri Chinchwad College of Engineering & Research, Pune- 412101, Maharashtra, India

[4]Bharat Sanchar Nigam Ltd. Rajgurunagar, Pune-410505, Maharashtra, India

[5]4CRG solutions Pvt Ltd., Parvati, Pune-411009 Maharashtra, India

[6]JSPM Narhe Technical Campus, Pune-411041, Maharashtra, India

[1]dtmane@gmail.com*

[2]devata.anekar@vit.edu,

[3]yogeshwari.mahajan@pccoer.in

[4]vsmahajan70@rediffmail.com

[5]kadamveena1234@gmail.com

[6]maqans@gmail.com

Correspondence author: dtmane@gmail.com

ARTICLE INFO

Received: 30 Sept 2024

Revised: 02 Dec 2024

Accepted: 15 Dec 2024

ABSTRACT

Natural language, sign language employs several modes of expression for everyday communication. Researchers have paid less attention to ISL interpretation than other forms of sign language. This paper introduces an automatic technique for translating manual letter movements into Hindi sign language. Because it works with pictures of bare hands, the user may interact with the device naturally. System gives deaf individuals the chance to interact with hearing people without the use of an interpreter. This paper goal is to develop techniques and a system for automatically recognizing Hindi sign language. The creation of a Hindi Sign Language database is the initial stage in this system. The most important stage in any hand gesture recognition system is hand segmentation since higher recognition rates may be attained with better segmented output. To improve identification rates, the suggested system additionally incorporates reliable and effective hand segmentation and tracking algorithms. Using a large sample set, it has been shown to be feasible to recognise 43 different words from the Standard Hindi sign language. We want to be able to convert certain very basic parts of sign language into text in the suggested system and vice versa in Hindi. In this paper, we present a number of new features and prospective advancements:The emphasis on real-time systems that recognize and interpret sign language motions from live video input is one noteworthy advance. With the use of this feature, people who use sign language and others can communicate instantly without any delays or middlemen. Integration of Speech Synthesis and Hand Gesture Recognition: This innovative approach to communication accessibility combines speech synthesis and hand gesture recognition technology. The paper provides a comprehensive solution for promoting engagement and understanding between the deaf community and others who communicate primarily through spoken language by translating gestures from sign language into intonation-rich speech.

Keywords: Image Capturing, Pre-processing, Feature Extraction, Classification, Pattern Recognition/Matching.

INTRODUCTION

Individuals with hearing and speech disabilities face communication challenges, and for Deaf and Dumb (D&M) people, communication is primarily achieved through sign language. Sign language involves the nonverbal exchange of messages through hand gestures, allowing D&M individuals to express their ideas visually. Our paper centers around the development of a model designed to recognize finger-spelling-based hand gestures, enabling the construction of complete words by combining individual gestures. The specific gestures chosen for training correspond to those illustrated in the provided image. The primary objective is to enhance communication for individuals with hearing and speech impairments by accurately identifying and interpreting these targeted gestures. The gestures we intend to train are depicted in the image below.

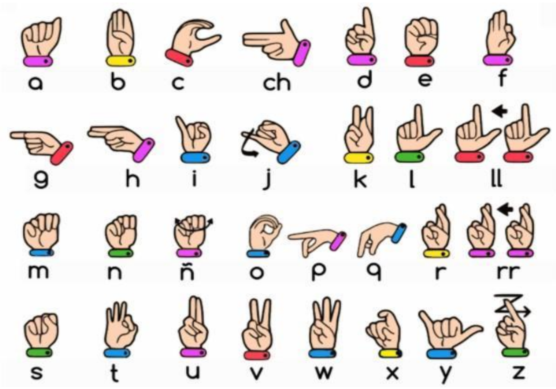

Figure1: Indian Sign Language(ISL)

The provided **Figure 1** illustrates the manual alphabet of Indian Sign Language (ISL). Each hand sign corresponds to a letter from 'a' to 'z'. These gestures are fundamental for communication in ISL. Understanding this manual alphabet is crucial for effective communication between hearing-impaired individuals and those who use ISL. By referring to this visual guide, learners can acquire the necessary skills to express themselves and engage in meaningful conversations using sign language.

Non-verbal communication stands as an integral facet of human interaction, allowing individuals to convey context, emphasize points, and express emotions alongside spoken language. Hand gestures, in particular, have been widely acknowledged as a fundamental mode of non-verbal communication, offering a means to communicate nuanced meanings, convey emphasis, urgency, and emotional nuances. Hand gestures are important for people who have trouble speaking and for interacting with computers. A new system has been created to help with this. It uses computer vision to see hand gestures in real-time and deep learning to turn these gestures into speech. The system quickly processes video to track and recognize gestures, then translates them into communication. It also changes spoken words into speech with the right tone and emotion, allowing for communication through both gestures and speech.

This paper can be a huge step for the people in terms of gestures as well as words

As this technology seems to easily integrate with the various devices and come out to be a good addition to society.

Over 70 million deaf individuals globally rely on sign languages for communication, enabling them to engage in learning, employment, accessing services, and participating in their communities. However, expecting everyone to learn sign language poses challenges in ensuring equal rights for people with disabilities. Therefore, the objective is to create a user-friendly Human-Computer Interface (HCI) that recognizes American Sign Language. This paper aims to enhance the lives of individuals who are deaf and mute by simplifying their daily activities.

An overview of the work is given below: Section 2 discusses the aim and objective.The literature review is given in section 3.The current system is described in section 4, along with the architecture of the system. The data gathering of different hand sample photos of sign language is the first step in Section 4. The dataset has 3500 photos altogether with a variety of hand gestures representing about 26 alphabets; methodology is discussed in section 5 and the dataset is divided into training, testing, and validation sets. We proceed with data pretreatment and data augmentation after separating our data. We developed our model and trained it on training data following the preprocessing of the data. Following the training phase, we used the test dataset to test our model and assessed it using several performance metrics and conducted an analysis of algorithms such as the Random Forest method in comparison.The experimental results were detailed in part 6, the future scope was discussed in section 7, and our research was ended in section 8.

AIM AND OBJECTIVE

Our objective is to develop a system capable of analyzing and identifying different alphabets from a collection of sign images. To achieve this goal, the application employs a range of Image Processing techniques, including segmentation and feature extraction.

In our research, we investigate specific methodologies for categorizing facial gestures with the aim of improving its efficiency and integrating it into any sign language or vision-based gesture recognition systems to facilitate accurate decision-making.

In our paper, we've developed a real-time system capable of detecting alphabets by identifying hand patterns through texture and shape analysis. This system operates with live video input, processing it instantaneously. We utilize the object detector feature from the computer vision toolbox to classify the hand from frames extracted from the video input.

The uniqueness of the paper is found in its real-time capabilities, speech synthesis and hand gesture detection integrated into one system, sophisticated image processing methods, emphasis on HCI, and overall objective of improving communication accessibility. Together, these features offer novel advances in the translation and identification of sign language.

LITERATURE SURVEY

[1] It is crucial to put in place a system that can correctly decipher and transmit hand signals, gestures, and communications made by people with speech impairments in order to close the communication gap between them and the broader population.where they detect hand gestures and improve the accuracy of the system.

[2] This paper describes to detect hand position and gestures from indian sign language in real time.There are 33 hand position where the goal is to improve communication between the general and deaf people.

[3] This paper describes the Current solutions as either not accurate enough or can't work in real-time. Where the goal is to improve the quality and accuracy .where hand position is detected through different position.

[4] People communicate with each other in order to exchange ideas, opinions, and life experiences. However, standard verbal communication options are not available to people who are deaf and mute. The use of sign language replaces auditory cues as a crucial tool for communication among the deaf and mute. The goal of this paper is to develop a system that can recognise and comprehend motions used in sign language.

[5] Most people connect with each other by speech and facial expressions, but new studies show that about 1% of Indians are deaf and mute. It can be difficult for people who are not familiar with these modes of communication to interact with these people because they primarily communicate through hand gestures and facial expressions. We suggest creating a static gesture classification system based on sign language standards in order to overcome this communication gap. By translating these motions into text and voice that match the regional accent, this method seeks to improve comprehension and communication for all parties.

[6] Verbal communication has always been the main way that people have interacted with one another. For people with disabilities, however, this conventional style of communication is inapplicable. One of the biggest problems facing society today is the gap between the disabled and the non-disabled. The only means of communication for those who are deaf and mute is sign language.

[7] Significant progress has been made in vision-based methods for sign language recognition in recent years. These developments cover a range of speech processing research papers that try to translate spoken words into text. Nonetheless, a still unexplored field is the vision-based categorization of facial gestures—like lip motions and eyebrow patterns—that are specifically designed for communication between people with diverse abilities.

[8] A naturally occurring method of human-computer communication, hand gesture recognition is a fast-growing area of computer vision and machine learning. This field provides users with a more practical way to interact with robots or system interfaces without needing extra equipment, and there are yet more uses to emerge. The main goal of gesture recognition research in the context of human-computer interaction (HCI) is to create systems that can recognise different human gestures and use them to transmit data or operate devices.

[9]Sign language (SL) recognition leverages many perspectives, histories, and skill sets to steer the progress of human-computer interaction and enhance communication within the deaf community. A broad spectrum of human talents and viewpoints are taken into consideration in hand gesture-based SL identification. However, efficient hand gesture identification is challenging to perform due to factors including shifting light levels, a diversity of motions, diverse views, self-referential characteristics, different forms and sizes, and complicated backdrop settings.

[10] For those who are deaf or hard of hearing, sign language is essential to their freedom as their primary form of communication. As a result, studies on automated interpretation of sign language have significantly increased. Numerous methods and algorithms in the field of image processing have been developed by the combination of artificial intelligence with image processing. All systems that recognise signs in sign language are trained to recognise signs and decipher them into the appropriate patterns. In 2023 [15], Mane D. et al. presented the compressive survey for the different web applications for the Devanagari languages.

OBSERVATION OF LITERATURE REVIEW

Observation: The combination of hand gesture recognition from video and its translation into intonation-rich speech is a cutting-edge field with potential to revolutionize human-computer interaction. While promising, it faces challenges such as complex hand gesture interpretation and real-time processing. However, advancements in computer vision and machine learning offer promising prospects for improving accessibility and communication interfaces.

The observation evolving the landscape of communication accessibility.Sign language system is a tool where they bridge the gap between deaf and normal people.This system consists of image ,text, and voice. And also use some machine learning.

A trend is focus on real time system capable of capturing the images and text through live video input.where its shows the real time integration of hand gestures and translating into intonation of rich speeches.

There are some problems also where the accuracy is not properly accurate.where the problems arise in some section like hand segmentation and feature extraction.

Overall the observation is about accurate system accuracy and trying to enhance the overcome problems by fostering the communication environment and advanced human paradigm.

EXISTING SYSTEM

Various recommendations have been developed and there are various approaches

that system was proposed.So basically they follow an approach from which vocal and deaf people will benefit from this technology.

User has to collect the dataset and it undergoes various pre-processing methods for extracting features. Following processing, the input image is converted to the appropriate text. Similar to this, preparation is carried out and the text is input into the textbox when translating text to a sign picture. Following processing, the text's sign picture appears on the screen. A database that is already existent in the database is used to aid with pattern recognition and matching when translating sign language to text and vice versa offline. Additionally, when recognising sign language via a web camera, an image of a hand gesture is captured by the device and analyzed to determine the proper text for the associated input hand motion image. The difficult aspect in the suggested system is identifying the input hand gesture image. From the input that the system is trained on, the suggested system will determine the appropriate output. The end user won't receive the output if the system receives incorrect or unknown input. Therefore, the user must submit a legitimate hand motion or text input.

Every nation has a specific sign language that is utilised across that nation. In a similar vein, deaf sign users in India speak Hindi Sign Language. The alphabets used in Hindi sign language have both vowels and consonants. Hindi sign Language alphabets are as follows:

Figure 2. Hindi Alphabets

The provided **Figure 2**.illustrates the Hindi alphabet in Devanagari script. It consists of two sections: 36 Letters in Hindi Consonants: The left section displays 36 gray squares, each containing a single Hindi consonant character. These consonants are essential building blocks for forming words and sentences in Hindi.

13 Letters in Hindi Vowels: The smaller section on the right features 13 gray squares, each representing a Hindi vowel character. Vowels play a crucial role in pronunciation and word formation.

Understanding the Hindi script is fundamental for reading, writing, and communicating effectively in Hindi. This visual guide serves as a valuable reference for learners and enthusiasts.

A particular sign language that may be utilized for communication is necessary in order to communicate in sign language. Sign language in Hindi is used to implement our suggested approach. The Indian sign language known as "Hindi" is used as a communication tool. The matching Hindi alphabets' sign language visuals are displayed in Figure 3.2. The proposed approach is intended to identify the Hindi sign, which is made up of consonants and vowels. Hindi sign language is converted into appropriate Hindi text upon recognition, and vice versa. Images in Hindi Sign Language.

The Existing systems for translating and recognising sign language have made great progress in closing the communication gaps that exist between people who are deaf or hard of hearing and the general public. To understand sign language motions and enable interaction, these systems mainly rely on computer vision, machine learning, and voice synthesis technology.

It has been observed that in order to facilitate communication accessibility, existing systems frequently concentrate on particular features, including hand gesture recognition or facial expression analysis. With real-time capabilities, which enable instantaneous interpretation of sign language motions from live video input, this capability has gained prominence. Communication efficacy is increased when gestures are translated into spoken language through integration with speech synthesis technology.

High precision and reliability are still difficult to achieve, nevertheless, especially in environments with different user viewpoints and situations. In order to improve recognition rates, hand segmentation and feature extraction techniques are essential. However, more developments are required to handle complicated backgrounds and lighting circumstances.

All things considered, current systems show encouraging progress in the areas of sign language translation and recognition; yet, there is still need for improvement in terms of precision, resilience, and usability. Subsequent advancements could center on optimizing algorithms, integrating multimodal inputs, and broadening language support to promote more diverse communication settings.

PROPOSED METHODOLOGY

In this paper we exhibit several novel aspects and potential innovations: One notable innovation is the focus on real-time systems that identify and decipher sign language gestures from live video input. With the use of this feature, people who use sign language and others can communicate instantly without any delays or middlemen. Integration of Speech Synthesis and Hand Gesture Recognition: This innovative approach to communication accessibility combines speech synthesis and hand gesture recognition technology.

This paper provides a comprehensive solution for promoting engagement and understanding between the deaf community and others who communicate primarily through spoken language by translating gestures from sign language into intonation-rich speech.

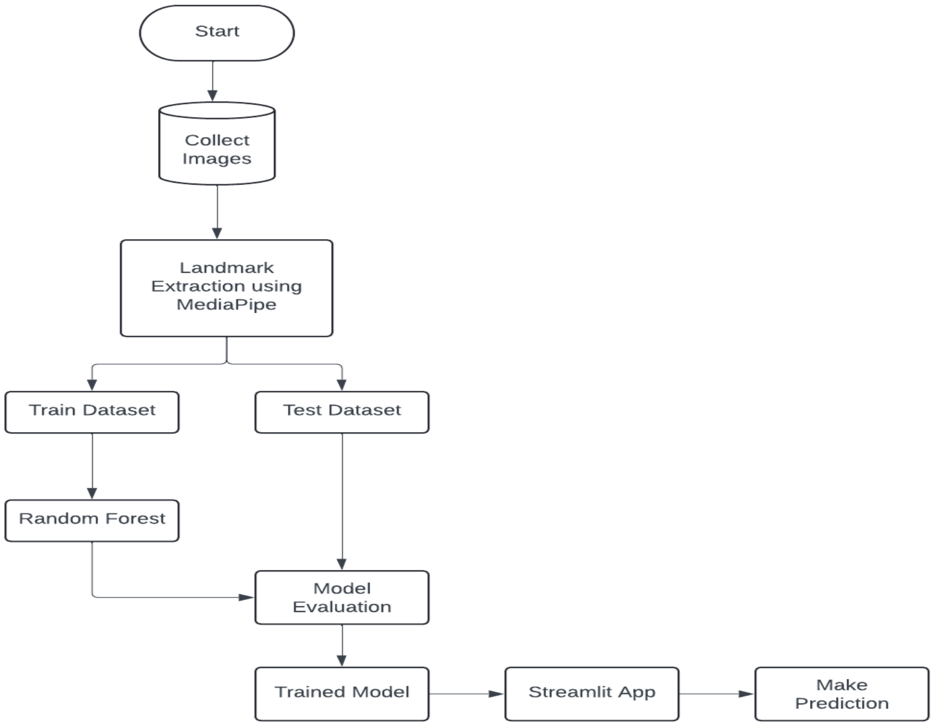

Figure 3. System architecture diagram

The suggested flow diagram The steps involved in developing and using a machine learning model are shown in Figure 3. First, data is gathered, and then the MediaPipe library is used to extract landmarks. Training and test datasets are created from the gathered photos. The training data is used to train the model, and then its performance is assessed. If successful, a Streamlit app may be used to use the trained model for prediction.

***Data Acquisition**:*

There are a number of methods that may be used to gather data on hand gestures:

Methods Based on Vision: The camera on a computer is the main instrument used by vision-based approaches to gather information about hands and/or fingers. These techniques, which are based on computer vision concepts, just require a camera to enable a conversational exchange between people and computers. This reduces expenses by doing away with the requirement for extra equipment. In vision-based hand detection, the main problem is to deal with the large variation in the look of the human hand. diverse skin color possibilities, and variations in viewpoints, scales, and camera capture speeds.

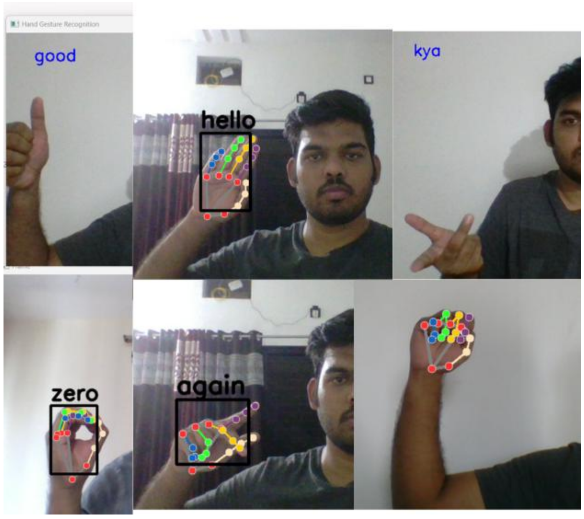

Figure 3.1 Representative Images from the Dataset

Figure 3.1 shows the image showcases a collage of five photographs, each capturing a person performing a distinct hand gesture. These gestures are labeled as "good," "hello," "kya," "zero," and "again." The overlay of colored dots and lines on the hands indicates that the system employs computer vision techniques to track and recognize these movements. Overall, this paper demonstrates the exciting potential of hand gesture recognition technology in various fields, including human-computer interaction, sign language translation.

Figure 3.2 Representative Images from the Dataset

Figure 3.2 shows the provided image showcases a compilation of six photographs, arranged in a 2x3 grid. Each photograph features an individual performing different hand gestures against a plain background.

Let's examine the gestures' specifics:

Thumbs-Up Gesture: One of the participants makes the gesture, which is frequently used to express the values.

Okay Sign: By putting the thumb and index finger together, one person makes the circular "okay" sign. This gesture frequently examine the approval.

Open Hand with Spread Fingers: The third person raises their hand, fingers widely apart.

This gesture could be useful for training machine learning models to recognize an extended hand.

Additional Poses: The remaining photographs depict various other hand poses, potentially including signs relevant to sign language or specific interactions.

Data pre-processing and Feature extraction:

Finding the hand in a webcam image is the first step in using this hand detection approach. The image-processing-oriented MediaPipe library is used for this. The image is cropped to acquire the area of interest (ROI) once the hand has been found. The image is first converted to grayscale using the OpenCV package, and then it is subjected to a Gaussian blur. The OpenCV library simplifies the process of applying this filter.

The grayscale picture must then be transformed into a binary image using the threshold and adaptive threshold techniques. These methods are crucial to image processing because they provide segmentation and augmentation of pertinent features.

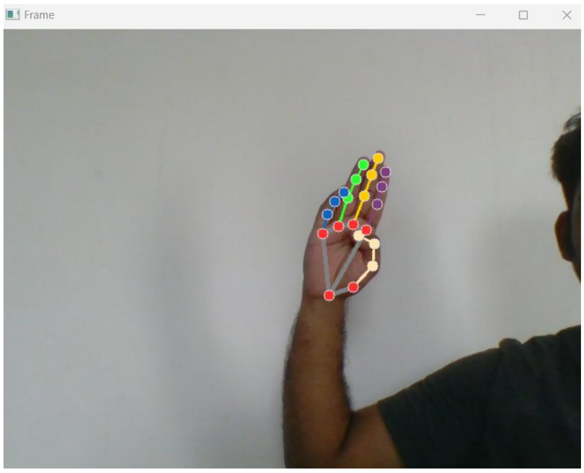

Figure4. Hand Anatomy Mapping for 'Hello'

Figure 4. The provided image depicts a human hand in a side profile, overlaid with colored dots and lines. These visual elements likely represent tracking points for finger joints and palm positioning. The purpose of this visualization is to demonstrate advanced gesture recognition, where individual finger movements are accurately tracked. Such technology finds practical applications in fields like robotics, virtual reality, and sign language interpretation.

For the sign letters A through Z, we have gathered pictures of various signs at various perspectives.

There are several flaws in this approach, such as the requirement that your hand be in front of a smooth, clean backdrop and that it be under optimal lighting circumstances in order for it to provide correct results. However, in the actual world, neither good background nor excellent lightning conditions are present everywhere.

To overcome this challenge, we experimented with different approaches until we devised an interesting solution. Employing a media pipe, we initially recognize the hand within the frame and extract the hand landmarks visible in the image. Subsequently, we depict and connect these landmarks on a simple white canvas.

Media pipe Landmark System:

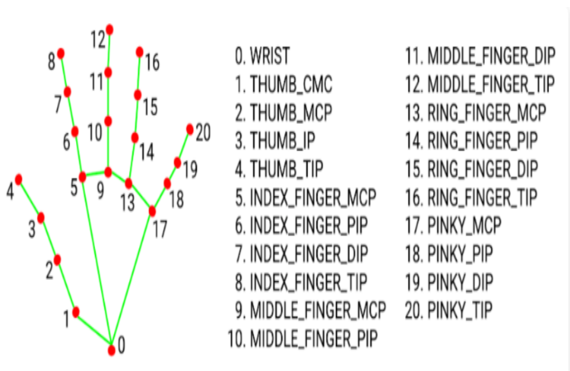

Figure5. Hand Landmarks

Figure 5. The suggested picture demonstrates how the hand landmark model bundle locates 21 hand-knuckle coordinates inside the identified hand areas via keypoint localization. To train the model, some 30,000 real-world images and many artificially created hand models overlayed on various backdrops were employed.

The hand landmarker model package includes two models: one for palm detection and the other for hand landmarks detection. While the palm detection model locates hands inside the input image, the hand landmarks detection model locates specific hand landmarks on the cropped hand picture supplied by the palm detection model.

In order to address background and illumination variance problems, we first address these difficulties by obtaining landmark points and using the OpenCV library to visualise them on a white backdrop.

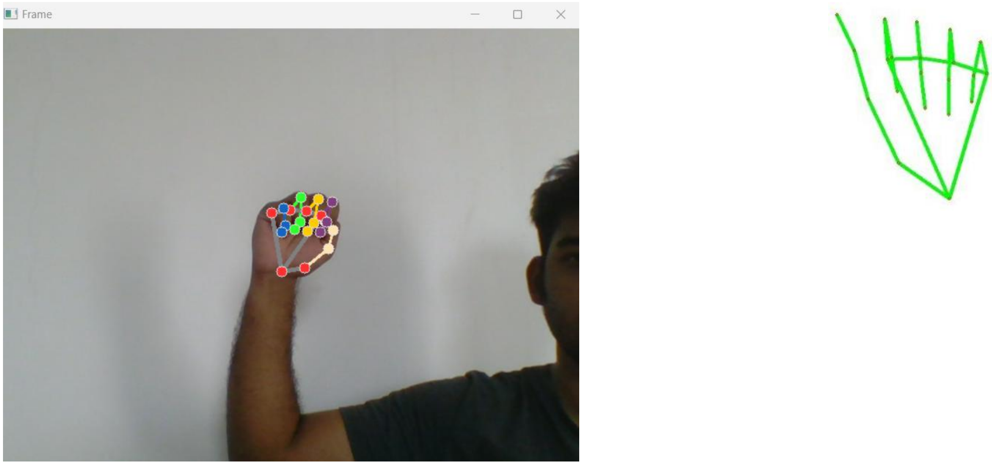

Figure.6. Hand Configuration for Closed Fist

Figure 6. We gathered different images as our custom dataset and this image is one of them .It denotes raised hand with infront of plain background .Coloured dots are there to map the hand structure to detection the points of the hand . All the lines indicates the relationships or movement in these points . It is essential to map the points for better accuracy for all the machine learning models to gestures in different computer vision techniques applications.

The above technique is effective in handling of different background and lighting conditions . Mediapipe gives us landmark options by neglecting the background and other lighting circumstances .

Random Forest Hand Gesture

Classification Model:

In this research, we used one classification technique name random forest . Random forest classifies different hand gestures captured images and categorizes it into different sign languages .Random forest does communication accessibility for those people who has hearing impairment.

Random Forest: Random forest has ability to harness the greatness of ensemble learning . Random forest works where multiple decision tree collaborate to make predictions.Decision tree are trained by different random subsets of the data to be trained. Then all decision trees average is taken into consideration, this reduces overfitting.

Gesture Recognition Application : The dataset that we have created comprises the hand gesture images which denotes different signs in sign languages . Images does preprocessing to get required features such as shapes and different movements . In training, The algorithm teaches to identify different patterns in features space in according to sign language gestures . By this ability the algorithm generalizes and classifies the data that is not seen by the algorithm also classifies it accurately .

Evaluation of performance : By using different datasets , the algorithm performance is calculated . The main thing that is required for the result is accuracy . Other parameters such as recall , the f1-score also important for classifying the dataset and increases efficiency

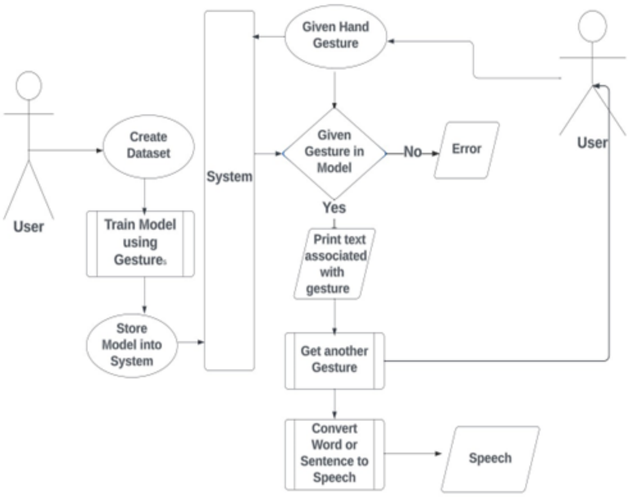

Figure 7. System Implementation Diagram

Figure 7. The images of the system Diagram shows the communication between user and the system that recognizes the hand . All the process begins with the dataset creation . After that we trains the model with different machine learning techniques . User provides the input and according to trained model hand gestures are recognized and the system gives the text .In case of any system error , it is displayed there.The system alo gives to converting of words into sentences

RESULTS

We have successfully implemented our paper and created a frontend using streamlit. The image of front page is added below

Figure8. Home page of HandSpeak

Figure 8: The provided imageHearingHands is a transformative initiative that bridges communication divides for the deaf community. Our platform understands sign language, translating it into written text and spoken words. Record your signs using a camera or webcam, and our technology converts them into your preferred language. Explore HearingHands and witness the power of inclusivity.

Our implementation is about the random forest algorithm for sign language .The classifier achieved 98.235% on testing data.

The level of accuracy shows the effectiveness of random forest approach with the hand gesture images.by ensemble decision tree the classifier works well to unseen data, and show new gestures accurately.

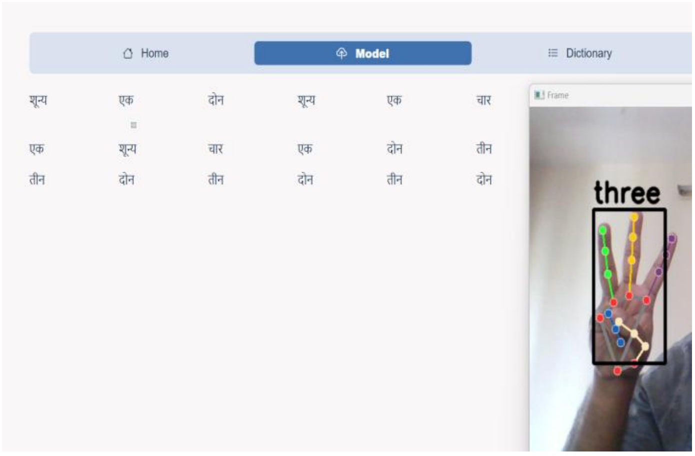

Figure9. Gesture Detection Process.

Figure 9. The picture detects the hand gesture of words Three and shows the colored dots. This mapping is essential for training data and shows the result of the output images.

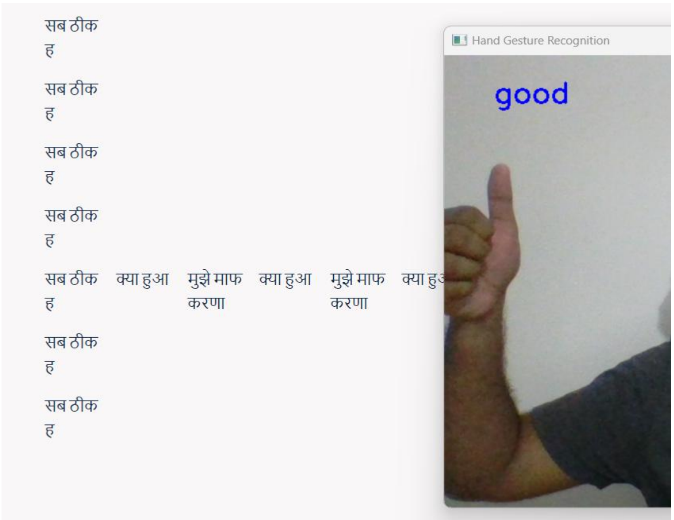

Figure 10. Gesture Detection Process.

Figure 10. The image shows the hand gestures of good and its detect the sentence. The classifier shows the output of the positive affirmation.

Accuracy Score:One popular assessment metric used in classification assignments is the accuracy score. Out of all the occurrences, it calculates the percentage of accurately predicted cases (samples). Accuracy may be computed mathematically in the following way:Total Number of Predictions / Number of Correct Predictions equals accuracy.

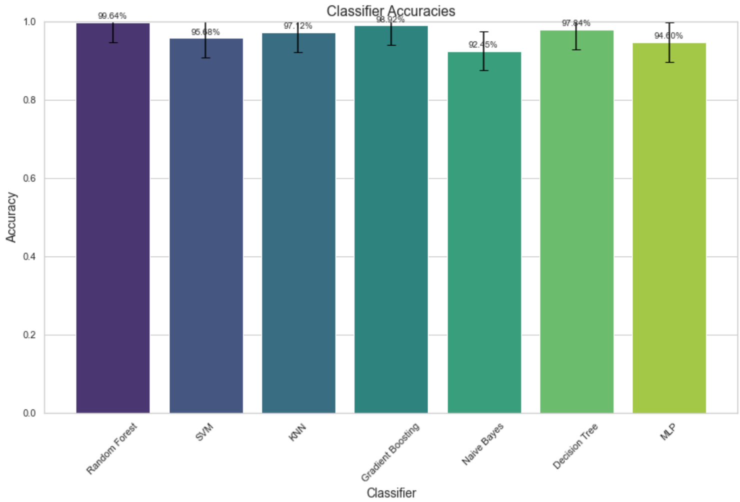

Figure11. Accuracy Metrics Visualization

Random Forest (RF): Achieved accuracy of approximately 99.64%.

SVM (Support Vector Machine): Achieved accuracy of about 90.46%.

KNN (K-Nearest Neighbors): Impressive accuracy of 97.12%.

Gradient Boosting: High accuracy at around 92.75%.

Naive Bayes: Respectable accuracy of 92.49%.

Decision Tree: Remarkable accuracy of approximately 97.49%.

MLP (Multi-Layer Perceptron): Solid accuracy score of about 94.10%.

F1 Score:

The **F1-score** is a crucial evaluation metric in machine learning, especially for binary classification tasks. It strikes a balance between two essential metrics: **precision** and **recall**. Here's what you need to know:

1. Precision:

- Precision represents the accuracy of positive predictions made by the model.
- It calculates how often the model correctly predicts positive instances
- Mathematically, precision is defined as:
  Precision=True Positives + False PositivesTrue Positives

2. Recall (Sensitivity or True Positive Rate):
The recall metric quantifies the model's ability to recognise real positive cases.

- It takes into account the ratio of real positive cases to the number of accurate positive forecasts.
- In terms of math, recall is provided by: True Positives + False Negatives equals recall.Real Advantages

3. F1-Score:

The F1-score integrates precision and recall into one measure. It is determined as the harmonic mean of precision and recall, using the formula.

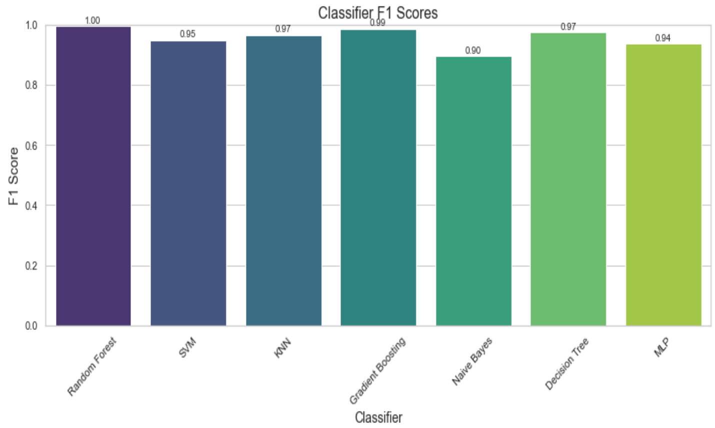

Figure12. Performance F1 Score Diagram

Random Forest (RF) achieves a perfect F1 score of 1.00.

SVM (Support Vector Machine) follows closely with an impressive score of 0.95.

Both KNN (K-Nearest Neighbors) and Gradient Boosting secured an F1 score of 0.97.

Naive Bayes performs slightly lower with an F1 score of 0.90.

Decision Tree also excels, matching the F1 score of 0.97.

Finally, MLP (Multi-Layer Perceptron) concludes the list with an admirable score of 0.94.

Matthews Correlation Coefficient (MCC):

In machine learning, the MCC is a performance statistic for binary classifiers. It takes into account each of the four components of a confusion matrix and calculates the correlation between the expected and actual binary results.

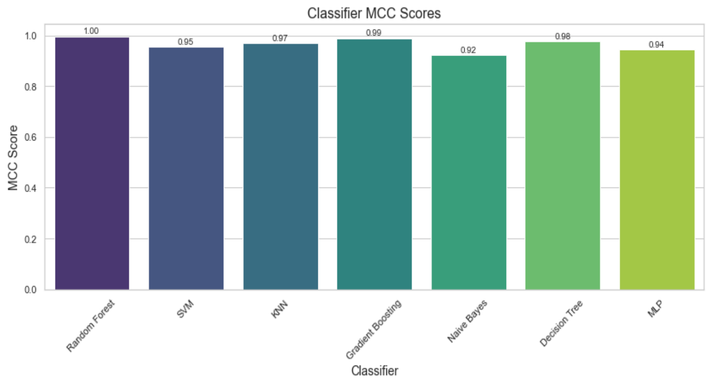

Figure.13. Model Performance MCC Plot

Random Forest (RF): Achieved an MCC score of 1.00 (perfect classification).

SVM (Support Vector Machine): Impressive MCC score of 0.95.

KNN (K-Nearest Neighbors): Strong performance with an MCC score of 0.97.

Gradient Boosting: Nearly perfect MCC score at 0.99.

Naive Bayes: Respectable performance with an MCC score of 0.92.

Decision Tree: Excellent MCC score of 0.98.

MLP (Multi-Layer Perceptron): Solid performance with an MCC score of 0.94.

Year	Researcher	Signs	Model	Accuracy
2018	Rao and Kishore [6]	18	ANN & MDC	90.58 %
2018	Rao et al. [12]	18	ANN	91%
2020	Uchil et al. [13]	20	CNN	85%
2022	Kothadiya et al. [14]	11	LSTM	95%
2024	Proposed Model	51	Random Forest	99.6%

Table 1. Comparative Analysis

The **table1** clearly shows the proposed model outperforms compared to existing algorithms. In the image compares various research studies conducted over different years, focusing on sign language recognition. Here are the key points:

Year and Researchers:

The studies span from 2018 to 2024.

Researchers include Rao, Kishore, Uchil, Kothadiya, and a proposed model.

Signs Analyzed:

The number of signs analyzed varies across studies, ranging from 11 to 51.

Models Used:

Different models were employed, such as ANN & MDCNN, CNN, LSTM, and Random Forest.

Accuracy: Accuracy percentages differ among studies, with the proposed model in 2024 achieving the highest accuracy (99.6%).

Overall, this analysis highlights advancements in sign language recognition research, emphasizing both model performance and the number of signs considered. However, factors beyond model choice, such as data quality and tuning, also play crucial roles in achieving high accuracy.

FUTURE SCOPE

Other than this issue, there are several applications for speech synthesis and gesture recognition. Future work should focus on strengthening the system's resilience and accuracy to enhance the lives of those who struggle with communication. These technologies also have the power to revolutionize ordinary human-computer interaction by providing more logical and organic ways for people to connect with technology.

In addition we are looking into a few more uses outside of this paper . This is a promising time for the field of hand gesture detection and voice translation .

CONCLUSION

In conclusion , hand gesture recognition from video and its translation into the text and voice is great human interaction with computer.and the way developed the system to bridge the gap between th deaf and normal people . This paper shows the possibilities of the advanced challenges and its challenges and also its understanding of further development. The potential here is that we are looking to the future where the system will be more accessible and accurate, capable of real time processing , therefore enhancing the quality of life for individual purposes. This technology helps everyday to transform people and they interact not only with limited resources but also explore more possibilities with technology.

REFERENCES

[1] "Sign Language Recognition in the Hindi Language Based on Computer Vision." Authorea Preprints (2023).Grover, Chhaya, Avni Rajpoot, Aditya Verma, and Ayush Patel.

[2] "Optimal video processing and soft computing algorithms for human hand gesture recognition from real-time video." Multimedia Tools and Applications (2023): 1-23.Uke, Shailaja N., and Amol Zade.

[3] "Real-time Indian sign language (ISL) recognition." In 2018 9th international conference on computing, communication and networking technologies (ICCCNT), pp. 1-9. IEEE, 2018.. Shenoy, Kartik, Tejas Dastane, Varun Rao, and Devendra Vyavaharkar.

[4] Vijayalakshmi, P., and M. Aarthi. "Sign language to speech conversion." In 2016 international conference on recent trends in information technology (ICRTIT), pp. 1-6. IEEE, 2016.

[5] Gauni, Sabitha, Ankit Bastia, B. Sohan Kumar, Prakhar Soni, and Vineeth Pydi. "Translation of gesture-based static sign language to text and speech." In Journal of Physics: Conference Series, vol. 1964, no. 6, p. 062074. IOP Publishing, 2021.

[6] G. Ananth Rao, and P.V.V. Kishore, "Selfie Video Based Continuous Indian Sign Language Recognition System," Ain Shams Engineering Journal, vol. 9, no. 4, pp. 1929-1939, 2018.

[7] Punekar, Aaditya C., Saranya A. Nambiar, Raunak Singh Laliya, and Saumya R. Salian. "A Translator for Indian Sign Language to Text and Speech." Int. J. Res. Appl. Sci. Eng. Technol 8 (2020): 1640-1646. .

[8] Das, Siddhartha Pratim, Anjan Kumar Talukdar, and Kandarpa Kumar Sarma. "Sign language recognition using facial expression." Procedia Computer Science 58 (2015): 210-216.

[9] Trigueiros, Paulo, Fernando Ribeiro, and Luís Paulo Reis. "Hand gesture recognition system based on computer vision and machine learning." Developments in medical image processing and computational vision (2015): 355-377.

[10] Rahim, Md Abdur, Jungpil Shin, and Keun Soo Yun. "Hand gesture-based sign alphabet recognition and sentence interpretation using a convolutional neural network." Annals of Emerging Technologies in Computing (AETiC), Print ISSN (2020): 2516-0281.

[11] Dutta, Kusumika Krori, and Anil Kumar GS. "Double handed Indian Sign Language to speech and text." In 2015 Third International Conference on Image Information Processing (ICIIP), pp. 374-377. IEEE, 2015.

[12] G. Anantha Rao et al., "Selfie Continuous Sign Language Recognition with Neural Network Classifier," Proceedings of 2nd International Conference on Micro-Electronics, Electromagnetics and Telecommunications: ICMEET 2016, vol. 434, pp. 31-40, 2018

[13] Aditya P. Uchil, Smriti Jha, and B.G. Sudha, "Vision Based Deep Learning Approach for Dynamic Indian Sign Language Recognition in Healthcare," International Conference on Computational Vision and Bio Inspired Computing, vol. 1108, 2020.

[14] Deep Kothadiya et al., "Deepsign: Sign Language Detection and Recognition Using Deep Learning," Electronics, vol. 11, no. 11, pp. 1- 12, 2022

[15] Mane, D., Patil, M., Chaudhari, V., Nayakavadi, R., Pandhe, S. (2023). A Survey on Chatbot in Devanagari Language. In: Pati, B., Panigrahi, C.R., Mohapatra, P., Li, KC. (eds) Proceedings of the 6th International Conference on Advance Computing and Intelligent Engineering. Lecture Notes in Networks and Systems, vol 428. Springer, Singapore. https://doi.org/10.1007/978-981-19-2225-1_31