ISSN 2278-3091
Volume 13, No.1, January - February 2024
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse051312024.pdf
https://doi.org/10.30534/ijatcse/2024/051312024

Talkhands: Communicating With The Silenced Using A Machine Learning Powered Sign Language Converter

Bhoomika H D[1], Karunya M Siddi[2], Samiksha D Patil[3], Yashaswini M[4]
[1]Student, Department of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women India, bhoomikahd2002@gmail.com
[2]Student, Department of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women India, karunyakaju2511@gmail.com
[3]Student, Department of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women India, samikshadpatil20@gmail.com
[4]Student, Department of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women India, yashumalangi@gmail.com

Received Date : December 23, 2023 Accepted Date: January 25, 2024 Published Date: February 06, 2024

ABSTRACT

In our interconnected world, communication is a basic human right often hindered by barriers, particularly for the deaf and hard of hearing. TalkHands, an innovative project, employs cutting-edge machine learning to create a state-of-the-art Sign Language Converter. This technology aims to facilitate seamless communication between sign language users and non-users, fostering inclusivity and equality by giving silent voices a means to be heard and understood.

Key Words: Webcam, PyCharm, Python, OpenCV, Keras API, Tensorflow Framework, Mediapipe, pyttsx3 Library, CNN

1. INTRODUCTION

Introducing TalkHands, a ground breaking initiative with the goal of enriching the experiences of individuals facing deafness or hearing challenges. This innovative project leverages cutting-edge machine learning technologies to make a significant impact. At its core, TalkHands aims to create an advanced Sign Language Converter, paving the way for effortless communication between those proficient in sign language and those who are not. The ultimate aim is to develop a tool that narrows the communication divide between the hearing and non-hearing communities, providing a platform for silent voices to be acknowledged and comprehended. Through fostering inclusivity and equality, TalkHands strives to build bridges and amplify the voices that often go unheard.

2. EXISTING SYSTEM

The existing system can be useful for static signs only. The sign language recognizer system cannot be considered as a complete system, as for complete recognition of sign language, we have to include alphabets, words and sentences. Classifiers like multi class Support Vector Machine (SVM), is included in conducting experiments for the gesture recognition.

3. PROPOSED SYSTEM

We aim to create an inclusive communication solution that bridges the gap between people who are deaf or hard of hearing and those unfamiliar with sign language. Our approach involves developing a sign language analyzer driven by machine learning, which can interpret signs captured via a camera. The key function of the system is to identify and correlate the analyzed sign with a specific letter or word in a predefined dataset. The resulting text can be accessed by a broader audience through an audio feature, allowing the message to be spoken aloud. This innovation aims to dismantle communication barriers, fostering effective interaction for individuals with hearing impairments.

4. LITERATURE SURVEY

The approach to solving the classification problem involves three key stages. First, segment the skin part from the dataset as it is considered noise for character classification. Second, extract important features from the segmented skin images for subsequent learning and classification. Finally, use these features as input for training supervised learning models and employ the trained models for classification. [1]

Emotional voice conversion technology is pivotal for various applications, including expressive text-to-speech, speech emotion recognition (SER), and conversational agents. It has gained significant attention, with one-third of research papers in the prosody and emotion category focusing on it. These frameworks primarily rely on data-driven techniques and training data. This article provides an introduction to emotion in speech and a comprehensive overview of emotional voice conversion studies based on training data types. [2]

The methodology proposed in this research employs a combination of GRU and LSTM; given the known features of these methods, a combination may be expected to detect and identify sign language gestures from a video source and to generate the associated English word. In this method, the video source is provided as an input, consisting of ISL sign-language gesture sequences. [3]

A method utilizing palm and linear recognition models for the dactyl alphabet and sign language numbers was tested on magnetic positioning system data using a hand kinematic model. The results, with approximately 97% classification accuracy, support the feasibility of this approach for developing automated sign language translation systems to facilitate communication with the deaf and hard of hearing. [4]

The project methodology uses three key components: Convolution Neural Network (CNN) for feature extraction, OpenCV for camera input, and TensorFlow for deep learning. CNN extracts data from Indian Sign Language (ISL) symbols in recorded videos, while OpenCV captures and processes camera images for gesture detection. TensorFlow simplifies deep learning calculations for efficient ISL gesture recognition and communication. [5]

Sign language recognition (SLR) employs diverse datasets and techniques, encompassing traditional machine learning (ML) methods like classification and clustering, along with advanced deep learning approaches, including Convolutional Neural Networks (CNN). Deep learning stands out for its superior performance with extensive datasets, as it uses programmable neural networks to make decisions, reducing the need for human intervention. [6]

The Text-to-Speech converter is built using the pyttsx3 library, which is platform independent and works offline. It supports both Python 2.x and Python 3.x. The synthesis process uses speech synthesis techniques with phoneme-based English language units for natural-sounding speech generation. [7]

Using computer vision and deep learning methods to recognize hand gestures and translate them into appropriate text. The system was built using a combination of key point detection using MediaPipe, data pre-processing, label, feature generation and LSTM neural network training. [8]

The research involved collecting raw images with characters A to Z and numbers 1 to 9. It employed a combination of descriptive and experimental methods, including skin masking, edge detection, SURF feature extraction, clustering, and BoVW modeling. The system's effectiveness was analyzed using various classification algorithms like naive Bayes, logistic regression, K nearest neighbors, support vector machine, and convolutional neural network. [9]

Found the necessary datasets on Kaggle, used them to create a CNN model, and then developed a desktop application with a user-friendly GUI. The application captures live camera input, processes it to display character results, which are ultimately converted into words. [10]

5. COMPARISON TABLE

Table 1 below shows the literature survey comparison table where the survey has been made by referring to the papers and understanding the working and concept that is been described in the paper by the authors.

Table 1: Literature Survey

REFERENCE	TECHNOLOGIES	AUTHORS	YEAR OF PUBLICATION	EXISTING SYSTEM	PROPOSED SYSTEM
IRJET (International Research Journal of Engineering and Technology) [1]	ISL, Data set, Cross validation, Artificial Neural Network, HU's moments, Skin segmentation, SVM	Prof. Radha S Shirbhate, Mr. Vedant D Shinde, Ms. Sanam A Metkari, Ms. Pooja U Borkar, Ms. Mayuri A Khandge	2020	The system can be useful for static ISL numeral signs only.	The proposed system can be useful for real time alphabets and numeral signs.

ELSEVIER [2]	Emotional Voice Conversion Emotional Speech Database (ESD)	KunZhoua, Berrak Sisman, Rui Liu, Haizhou Lia	2021	Effectiveness of voice conversion techniques may vary depending on the diversity of speakers represented in the dataset.	The proposed system can be useful for effective and real time text to audio conversion.
MDPI (Multidisciplinary Digital Publishing Institute) [3]	ISL, Deep Learning, LSTM, GRU, Signs	Deep Kothadiya, Chintan Bhatt, Krenil Sapariya, Kevin Patel, Ana-Belén Gil-González, Juan M. Corchado	2022	Lack of specific details regarding how model accuracy could be improved, ambiguity in the approach to interpreting continuous sign language.	The proposed system can be useful for real time and continuous alphabets and numeral sign recognition with appreciable accuracy.
MDPI (Multidisciplinary Digital Publishing Institute) [4]	Sign Language, Handshape, Palm Definition Model, Media Pipe Face, Media Pipe Hands, SVM, Pattern Recognition, Multiple Classification	Nurzada Amangeldy, Saule Kudubayeva, Akmaral Kassymova, Ardak Karipzhanova, Bibigul Razakhova, Serikbay Kuralov	2022	The method was tested on a magnetic positioning system with a kinematic hand model. The effectiveness in recognizing a broader range of sign language gestures, especially dynamic signs, remains unexplored.	The proposed system can be useful for real time alphabets and numeral signs and can dynamically recognise the sign gestures.
IJRASET (International Journal for Research in Applied Science & Engineering Technology) [5]	Machine learning, Gesture Recognition, Sign Language, CNN, OpenCV Python Library, Tensor Flow	Vaibhav Shah, Nikhil Sharma, Prince Solanki, Prof. Hemalata Mote4	2022	The paper demanding significant data and mathematical expertise, may not always deliver real-time performance on various hardware.	The proposed system mainly focuses on the proper dataset in order to acquire, pre-process, identify and classify the dynamic sign gestures.
International Journal for Research [6]	Sign Language, Sign Language Recognition, Artificial Intelligence, CNN, Deep Learning	Ahmed Sultan, Walied Makram, Mohammed Kayed, Abdelmaged Amin Ali	2022	Traditional algorithms may struggle with the complexity of sign language recognition and may lack transparency in their decision-making.	The proposed system is effective in terms of sign language recognition and has transperency in the decision making process.
IJIRT (International Journal of Innovative Research in Technology) [7]	Text to Speech, Python, Audio	Dr.S.A. Ubale, Girish Bhosale, Ganesh Nehe, Avinash Hubale, Avdhoot Walunjkar	2022	The resulting speech is less than natural and pronunciation analysis from written text is a major concern.	The text to audio feature of the proposed system is more natural and the pronunciation of the text is been taken into consideration.
IJRASET (International Journal for Research in	Hand Gesture, Media Pipe, RNN, Machine Learning, TensorFlow, OpenCV,	Akash Kamble, Jitendra Musale, Rahul Chalavade,	2023	The availability of comprehensive and diverse language	The proposed system takes into consideration the

Applied Science & Engineering Technology) [8]	Python	Rahul Dalvi, Shrikar Shriyal		datasets can be limited.	comprehensive, dynamic and diverse sign gestures.
MDPI (Multidisciplinary Digital Publishing Institute) [9]	Classification Models, Gesture to Text, Skin Masking, Speeded-Up Robust Features, SURF, Feature Extraction	Kaustubh Mani Tripathi, Pooja Kamat, Shruti Patil, Ruchi Jayaswal, Swati Ahirrao, Ketan Kotecha	2023	Using a limited set of classification algorithms may restrict its adaptability in various real-world situations.	The proposed system implements proper classification of the data sets in order to provide appreciable accuracy.
ELSEVIER [10]	Computer Vision, CNN, American Sign Language (ASL), Sign Language Recognition	Yulius Obi, Kent Samuel Claudio, Vetri Marvel Budiman, Said Achmad, Aditya Kurniawan	2023	The process of forming letters into words, while essential for communication, have limitations related to speed and efficiency.	The proposed system takes into consideration and has appreciable speed of forming the words from the letters.

6. METHODOLOGY

6.1 Data Collection and Preparation

Collecting a diverse dataset of sign language gestures, signs, and variations. Annotating the dataset with corresponding letters or words for each sign. Preprocessing the data by normalizing hand positions, orientations, and other relevant features.

6.2 Machine Learning Model Development

Using Convolutional Neural Networks (CNNs) machine learning algorithm for sign language recognition. Training the model on the prepared dataset, optimizing it for high accuracy and real-time performance.

6.3 Real-time Video Analysis

Developing a module for capturing video input from a camera. Ensuring real-time processing of video frames to detect and track sign gestures as they are made.

6.4 Sign Language Recognition

Implementing the machine learning model to analyze each frame and recognize sign gestures in real-time. Extracting key features of sign language gestures, such as hand movements to improve recognition accuracy.

6.5 Text Generation

Matching recognized sign gestures to their corresponding letters or words from the annotated dataset. Converting the recognized signs into text format, forming sentences or phrases based on the sequence of signs.

6.6 Audio Output

Providing an audio output option for the generated text, allowing users to listen to the spoken version of the sign language message.

6.7 Testing and Validation

Conducting extensive testing with both sign language users and non-signers to assess the system's accuracy, real-time performance, and usability.

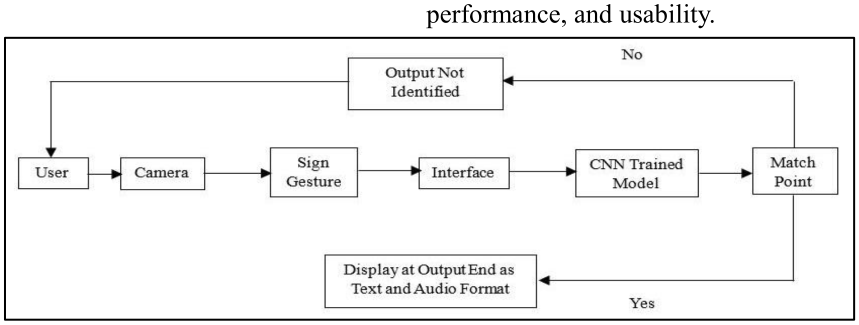

Figure 1: System Architecture

Figure 1 above shows the system architecture of the model and the system will work by following the below mentioned seven steps.

7. CONCLUSION

In summary, it is imperative to create a comprehensive solution that improves communication between individuals who are deaf or hard of hearing and those unfamiliar with sign language. This step is vital for promoting inclusivity and dismantling communication obstacles. The suggested solution incorporates a sign language analyzer driven by machine learning, offering a promising method to close the communication divide. Using camera input, the system can identify and correlate signs with letters or words in a predefined dataset. The inclusion of an audio button expands the reach of the recognized text, making effective communication accessible to a broader audience, especially benefiting individuals with hearing impairments.

REFERENCES

1. Prof. Radha S Shirbhate, Mr. Vedant D Shinde, Ms. Sanam A Metkari, Ms. Pooja U Borkar, Ms. Mayuri A Khandge, **Sign Language Recognition Using Machine Learning Algorithm**, International Research Journal of Engineering and Technology, 2020, Volume: 8 Issue: 8, avaliable at https://www.academia.edu/59078927/IRJET_Sign_language_Recognition_System_using_Machine_Learning . [1]

2. KunZhoua, Berrak Sisman, Rui Liu, Haizhou Lia , **Emotional Voice Convertion: Theory, Databases and ESD**, ELSEVIER, 2021, avaliable at https://www.researchgate.net/publication/357210011_Emotional_voice_conversion_Theory_databases_and_ESD . [2]

3. Deep Kothadiya, Chintan Bhatt, Krenil Sapariya, Kevin Patel, Ana-Belén Gil-González, Juan M. Corchado, **Deepsign: Sign Language Detection and Recognition Using Deep Learning**, Multidisciplinary Digital Publishing Institute, 2022, avaliable at https://doi.org/10.3390/electronics11111780 . [3]

4. Nurzada Amangeldy, Saule Kudubayeva, Akmaral Kassymova, Ardak Karipzhanova, Bibigul Razakhova, Serikbay Kuralov, **Sign Language Recognition Method Based on Palm Definition Model and Multiple Classification**, Multidisciplinary Digital Publishing Institute, 2022, avaliable at https://doi.org/10.3390/s22176621 . [4]

5. Vaibhav Shah, Nikhil Sharma, Prince Solanki, Prof. Hemalata Mote, **Deaf and Dumb Gesture Recognition System**, International Journal for Research in Applied Science & Engineering Technology, 2022, avaliable at https://doi.org/10.22214/ijraset.2022.41413 . [5]

6. Ahmed Sultan, Walied Makram, Mohammed Kayed, Abdelmaged Amin Ali, **Sign Language Identification and Recognition: A Comparative Study**, International Journal for Research, 2022, avaliable at http://dx.doi.org/10.1515/comp-2022-0240 . [6]

7. Dr.S.A. Ubale, Girish Bhosale, Ganesh Nehe, Avinash Hubale, Avdhoot Walunjkar, **A Review on Text to Speech Converter**, International Journal of Innovative Research in Technology, 2022, avaliable at https://ijirt.org/Article?manuscript=155350 . [7]

8. Akash Kamble, Jitendra Musale, Rahul Chalavade, Rahul Dalvi, Shrikar Shriyal, **Conversion of Sign Language to Text**, International Journal for Research in Applied Science & Engineering Technology, 2023, avaliable at https://doi.org/10.22214/ijraset.2023.51981 . [8]

9. Kaustubh Mani Tripathi, Pooja Kamat, Shruti Patil, Ruchi Jayaswal, Swati Ahirrao, Ketan Kotecha, **Gesture-to-Text Translation Using SURF for Indian Sign Language**, Multidisciplinary Digital Publishing Institute, 2023, avaliable at http://dx.doi.org/10.3390/asi6020035 . [9]

10. Yulius Obi, Kent Samuel Claudio, Vetri Marvel Budiman, Said Achmad, Aditya Kurniawan, **Sign Language Recognition System for Communicating to People with Disabilities**, ELSEVIER, 2023, avaliable at http://dx.doi.org/10.1016/j.procs.2022.12.106 . [10]