Perspectives in Communication, Embedded-Systems and Signal-Processing (PiCES) – An International Journal
ISSN: 2566-932X, Vol. 5, Issue 2, May 2021

A Survey on Sign Language Recognition

Abilash S
Department of Computer Science and Engineering, Jyothy Institute of Technology, Bengaluru, India, abisri220499@gmail.com

Ashish
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, ashish.mp139@gmail.com

Shreyas K S
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, shreyashebbar.k@gmail.com

Shreyas S Nadig
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, shreyasnadig02@gmail.com

Vallabh Mahale
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, vallabh.mahale@jyothyit.ac.in

Abstract: The modern day applications are more user experience-oriented. Every application tends to gather information specific to each person's operation and then provide customized features. The methods used to achieve such feats are a product of Machine Learning (ML) and Artificial Intelligence (AI) techniques. Using these techniques, real-time applications are reaching new heights in interacting with humans almost seamlessly. When living in such times, we can use these methods effectively to explore ways to overcome challenging problems. One such problem that we are taking upon is Sign Language Recognition (SLR). The project aims at bridging the communication gap with voice and hearing-impaired people, helping them to converse with the world more fluently. Using hand gestures as the primary input, we convert those into understandable language.

Keywords: Sign Language Recognition (SLR); Hand Gesture Recognition; Machine Learning (ML); Artificial Intelligence (AI); Real-Time Application

I. INTRODUCTION

Machine Learning (ML) is one among the fastest-growing areas in technology. Machine Learning is a subset of Artificial Intelligence that makes use of statistical methods that enables the computer to make data-driven decisions to carry out a specific task. Artificial Intelligence is a technique that enables machines to mimic human nature and behavior. AI models can be trained to accomplish a specific task from learning through experience and recognizing patterns in a large amount of data. ML models perform operations that humans currently can do better. The algorithms are designed in such a way that they learn and improve over time when exposed to new data.

In this paper, we survey works related to Sign Language recognition and the techniques involved in the process. Sign Language is the art of expressing oneself and communicating with others through hand gestures, facial expressions and figure moments. A mammoth task for a speech or hearing impaired person is to make the one they are communicating with understand sign language who may or may not be familiar with it. The idea is to convert hand gestures into text and sound outputs, and the given sound or text inputs to appropriate hand gestures.

II. LITERATURE SURVEY

[1] Building a gesture recognition system for translating Indian Sign Language (ISL) to resolve communication problems between hearing and hearing disabled people by using OpenCV for Skin Segmentation and fuzzy-c means for Clustering Algorithm. This system takes video of hand gestures as inputs and gives the accurate meaning of that gesture as output. This system achieves an approximated accuracy of 75%, which is higher than similar systems, this system can also work on dynamic data, and using fuzzy c-means makes it more efficient and reliable. Though this system is efficient, it requires more computation power than others.

[2] This paper proposes an improved threshold-based segmentation technique by adding a new constraint to Kovac's rule to limit the skin tone selection. This technique improves the segmentation results by correctly segmenting the apparel with the shade of purple, pale pink, and pinkish-yellow. This model was able to differentiate between pale brown and skin color, thus improving the accuracy score by 26% from Kovac's rule. The only disadvantage is it segments some pale brown clothing which is very close to skin color as skin tone.

[3] This system is an American Sign Language translator, which captures images of a person's hand under several different conditions and translate such hand gestures to their correct meaning with maximum efficiency. This is achieved using Canny Edge Detection for segmentation, Oriented Fast and Rotated Brief for feature extraction, and K-means clustering. The proposed technique outperforms all the other preprocessing techniques for Naive Bayes, Logistic Regression, and KNN classifiers. This system achieves an approximated

accuracy score of 95.81%. The drawback of this system is that it only works against static gesture images.

[4] This is an automation system that can convert sign language to spoken language which can help hearing-impaired people. This system can recognize sign language in complex backgrounds. Video of the signer is taken and is given as input to the system, parts of the hand with the sign is segmented concerning skin color. Features that identify the sign are extracted from the hand image and classified to recognize the sign. This is achieved using Support Vector Machine (SVM), YCbCr Color Space, and Histogram of Oriented Gradients for feature Extraction. The major disadvantages of this system are the lighting conditions in different situations, and it works only for static gestures.

[5] The system uses Viola and Jones algorithm with the AdaBoost classifier for image preprocessing, i.e., to detect and remove the face. HSV (Hue, Saturation, and Value) thresholding is used to isolate the hand by recognizing the largest region of similar skin tone. For the obtained image, the Center of Gravity is calculated and the fingertips are detected. This is done to identify if the input is static or dynamic. Zernike moments are used to extract features of static gestures. The SPHINX Speech Recognition Engine is made use of Speech Recognition. The system has an approximated accuracy of 93%. Disadvantages are such that it does not recognize facial cues or gestures with two hands and it's not accurate with unfavorable environmental conditions.

[6] The system uses a webcam to capture images in RGB. It is further converted to YCbCr and then to a binary image. Erosion and Dilation morphological operations are performed for noise removal and thresholding for image segmentation. Several algorithms like adaptive boosting algorithm is used for hand detection and haar classifier algorithm for classifier training. Finally, the convexity hull algorithm is used for finger-point detection and number recognition. The disadvantage of the system is that its application is limited to a small number of gesture inputs.

[7] The system capturing images with the help of a webcam. There are two classifiers used by the system, where one uses raw image features and the other uses histogram features and both use the Backpropagation Algorithm for Artificial Neural Network (ANN) learning technique to produce texts. The input dataset for the system was Marcel Static Hand Posture, and the data was split into training and testing. Python programming and SciPy libraries were used to implement multilayer perceptron neural network sign language recognition. The accuracy rates of Raw Features Classifiers and Histogram Features Classifiers were 70% and 85% respectively. The only drawback is that it did not produce satisfactory results when compared with other studies.

[8] The proposed system uses datasets provided by MNIST, the images are preprocessed using the PIL library. To build the model, advanced libraries such as Keras and TensorFlow were used and a Convolution Neural Network was developed. The captured images when processed and given as input to the trained model, it detected the gestures with an accuracy of 99.98%. In addition to sign-language translation, a bidirectional model was created, such that, speech or text can be converted to sign language. The disadvantage of this system is that the system does not recognize dynamic gestures and does not support video classification.

[9] The proposed system is designed to recognize hand gestures of programming keywords. Convolution Neural Network is used to train the classifier model, SciKit library is used in making a decision tree and a greedy algorithm is used to predict only the necessary frames, as it helps in avoiding redundancy in prediction. A syntax generation module is used to generate syntax for the predicted keywords. Image processing frameworks such as OpenCV and pillow are used for hand-detection and the video is captured at five frames per second. On providing the video input, the keywords were predicted with an accuracy of 61.58%. However, the only disadvantage is that the recognition is restricted to very few keywords of programming languages.

[10] Mudraksha uses OpenCV to convert video into frames. The obtained frames are then converted to greyscale images using NumPy and Matplotlib libraries, as greyscaling the image helps in reducing noise. The Canny edge detection technique is then applied to the processes images to detect the edges in hand gestures. A classifier is designed using Convolution Neural Network and is trained with sign language datasets, which are processed with the help of TensorFlow and Keras modules. The trained model is now used to detect hand gestures from the edge detected images after which suitable text and speech output is provided based on the gesture recognized. The proposed system has an accuracy of 98.07% for 10 epochs. The only disadvantage is that this system cannot recognize dynamic hand gestures.

[11] The System uses a combination algorithm to perform hand tracking. The contour following algorithm is used to extract the external boundary points of the hand gesture. The extracted features are used in the classifier which is designed using KNN and adaptive kernel matching. TensorFlow and Keras frameworks are used in training the model. The model scored an accuracy of 63%. The only disadvantage is that the system is implemented using a very small dataset.

[12] The proposed system uses a Convolutional Neural Network to predict hand gestures. Firstly, the images are transformed to Greyscale and an Augmentor Library is used to obtain multiple perspectives of the same image, which helps in refining the trained model. The training data had a set of 36000 images. The system got a validation accuracy of 99.22% for 200 epochs. A disadvantage of this system is that it needs more computation power, which might fail to work on devices with less computing power.

[13] In this system data glove is an important feature that consists of five flex sensors and an Accelerometer

MPU-6050, all these flex sensors will be connected to different pins of Arduino UNO i.e., the controlling unit. The data input will be provided by the flex sensors to the controlling element which displays the output in an LCD display and it converts the text into voice output using the speaker. Text-To-Speech is used to convert text into audio format.

[14] The system makes use of three different techniques. First, the Vision-based approach is used to compare 3D models or 2D projections of 3D models with the images stored in the database to recognize hand gestures. Second, in the Glove based approach, a sensor attached to the glove captures the position and movement of the hand, which provides near to perfect position and orientation of the hand, but the disadvantage is that it needs to be connected to the system to work. Third, the Color-marker approach uses different colored gloves to identify the position, movement, and motion of the hand. It is not regarded as the natural method for human computer interaction making it a flaw in the system.

III. PROPOSED SYSTEM

A novel system that takes in static or dynamic hand gestures as input and results the equivalent text or speech output.

IV. CONCLUSION

This paper mainly deals with methods to recognize hand gestures. The purpose is to ease the communication for people who have hearing/voice impairment and help those who are not familiar with sign language, understand the conversation. These systems see that they lubricate the process and help people to express their thoughts to others, using hand gestures.

Using ML and AI techniques we introduce a novel model that helps in Sign Language Recognition in real-time. So, the computer vision plays an important role by capturing the videos and then recognizing the hand gesture accurately. One of the advantages of the vision based method is that it can be implemented practically.

REFERENCES

[1] Muthu Mariappan H and Dr Gomathi V, “Real-Time Recognition of Indian”, IEEE, 2019.

[2] Novira Dwina,Fitri Arnia and Khairul Munadi, "Skin Segmentation based on Improved Thresholding", IEEE, 2018.

[3] Ashish Sharma, Anmol Mittal, Savitoj Singh and Vasudev Awatramani, "Hand Gesture Recognition using Image Processing and Feature Extraction Techniques", 2020.

[4] S. Reshna and M. Jayaraju, "Spotting and Recognition of Hand Gesture for Indian Sign Language Recognition System with Skin Segmentation and SVM", IEEE, 2017.

[5] Anup Kumar, Karun Thankachan and Mevin M., "DominicSign Language Recognition", IEEE, 2016.

[6] Ashish S. Nikam and Aarti G. Ambekar, "Sign Language Recognition Using Image Based Hand Gesture Recognition Techniques", IEEE, 2016.

[7] Tülay KarayÕlan and Özkan KÖlÕç, "Sign Language Recognition", IEEE, 2017.

[8] Lance Fernandes, Prathamesh Dalvi and Akash Junnarkar, "Convolutional Neural Network based Bidirectional Sign Language Translation System", 2020.

[9] Thanasekar B, Akshay V, Deepak Kumar G and Abdul Majeed Ashfaq, "Real Time Conversion of Sign Language using Deep Learning for Programming Basics", 2019.

[10] J Vaishnavi, M Vindhya, Mohammed Sadiq and Supreetha Vemala, "MUDRAKSHARA - A Voice for Deaf/Dumb People", 2020.

[11] Purva Chaithnya Badhe and Vaishali Kulkarni, "Artificial Neural Network based Indian Sign Language Recognition using hand crafted features", 2020.

[12] Md. Mehedi Hasan, Azmain Yakin Srizon and Md. Al Mehedi Hasan, "Classification of Bengali Sign Language Characters by Applying a Novel Deep Convolutional Neural Network", IEEE, 2020.

[13] Jenifer Nadar, Jerryman Lopes and Parth Dave, "Hand Gesture and Voice Recognition using Sign Language", 2018.

[14] Swapnil Chandel, Mrs. Akanksha Awasthi, "Hand Gesture Recognition for Sign Language Recognition", IJSTE, 2016.