International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 12 | Dec - 2021 ISSN: 2582-3930

Real Time Sign Language Detection

Swojas Bhalunkar, Swapnil Dadamode, Siddhi Ekatpure, Divya Deshmukh

Computer Engineering,
Anantrao Pawar College of Engineering and Research, Pune, India

Abstract: Sign Language has been the most used medium of communication for deaf/mute community for many years. Using the machine learning methods, a system can be created which can interpret and show the sign language in readable format. This will allow disabled users to . We have attempted to build such a system which uses the SVM (Support Vector Machine) algorithm to analyze the hand gestures and predict the corresponding sentence/word/letter. As this system uses Google's Mediapipe, we were able to overcome some problems which old systems had (Most of these systems were using Convolutional Neural Network (ConvNet/CNN)). The results had an accuracy approximately equal to 98 percent.

Index Terms: Sign Language Detection, Hand Gesture Recognition, Machine Learning, SVM algorithm, OpenCV, Python, Flask, etc.

1. INTRODUCTION

There are over 5% of the world's population i.e. 430 million people (432 million adults and 34 million children) 'disabling' hearing loss. According to WHO (World Health Organization), by 2050 nearly 2.5 billion people will have some degree of hearing loss and at least 700 million will require hearing rehabilitation. Many people have difficulties in communication so they have a human-translator around them. These human-translator are individuals who know sign language and a spoken language. Currently, hiring a human-translator is expensive and cumbersome as there are very few service providers available.

SVM is one of the machine learning algorithms which can allow us to predict words using hand gestures. By providing appropriate training and validation dataset, one can build a ML model for the proposed system. MediaPipe customizable ML solutions for live and streaming media. MediaPipe Hands is part of this framework which enables us to get the 'X & Y coordinates' of all the landmarks on hand. We can use these handmark points as input for our SVM model. We have proposed a system based on the Flask framework which uses trained SVM model to interpret sign language using the hand gestures captured through webcam/camera.

2. RELATED WORKS:

Relatively hand gesture recognition is a difficult problem to address in the field of machine learning. Classification methods can be divided into supervised and unsupervised methods. Based on these methods the SLR system can recognize static or dynamic sign gestures of hands. Murakami and Taguchi in 1991, published a research article using neural network for the first time in sign language recognition. With the development in the field of computer vision, numerous researchers came up with novel approaches to help the physically challenged community. Using coloured gloves, a real-time hand tracking application was developed by Wang and Popovic. The colour pattern of the gloves was recognized by K-Nearest Neighbors

(KNN) technique but continuous feeding of hand streams is required for the system. However, Support Vector Mechanism (SVM) outperformed this algorithm in the research findings of Rekhaetal., Kurdyumovetal Tharwatetal. and Baranwal and Nandi. There are two types of Sign Language Recognition: Isolated sign recognition and continuous sentence recognition. Likewise, whole sign level modelling and subunit sign level modelling exist in the SLR system. Visual-descriptive and linguistic-oriented are two approaches that lead to subunit level sign modelling. Elakkiya et al. combined SVM learning and boosting algorithm to propose a framework for subunit recognition of alphabets. An accuracy of 97.6% was obtained but the system fails to predict 26 alphabets. To extract features of 23 isolated Arabic sign language Ahmed and Aly used the combination of PCA and local binary patterns. Despite getting an accuracy of 99.97% in signer dependent mode, due to the usage of threshold operator the system fails to recognize the constant grey-scale patterns in the signing area. In the field of machine learning, recognizing hand gestures is relatively problematic to solve. In most of the initial attempts, a conventional convolutional network is used that detects hand gestures from frames of images. R.Sharma et al., used 80000 individual numeric signs with more than 500 pictures per sign to train a machine learning model. Their system methodology comprises a training database of pre-processed images for a hand-detection system and a gesture recognition system. Image pre-processing included feature extraction to normalize the input information before training the machine learning model. The images are converted into grayscale for better object contour maintaining a standardized resolution and then flattened into a smaller amount of one-dimensional components. The feature extraction technique helps to extract certain features about the pixel data from images and feed them to CNN for easier training and more accurate prediction. Hand tracking in 2D and 3D space has been performed by W.Liu et al.,.They used skin saliency where skin tones within a specific range were extracted for better feature extraction and achieved a classification accuracy of around 98%. It is evident from all these previous methods that to recognize hand gestures precisely with high accuracy, models require a large dataset and complicated methodology with complex mathematical processing. Pre-processing of images plays a vital role in the gesture tracking process. Therefore, for our project, we used an open-source framework from Google known as Mediapipe which is capable of detecting human body part accurately.

2. ALGORITHM:

Support Vector Machine(SVM) is a supervised machine learning algorithm used for both classification and regression. When the frame is captured, the image is passed through the mediapipe function which gives hand_landmarks (X & Y Coordinates). Relevant data is extracted, filtered and normalized to make it compatible for the model. Trained model predicts a word/sentence/letter/number based on the input. This string is shown on the user interface.

3. CONCLUSION:

With an average accuracy of 98% in most of the sign language dataset using MediaPipe's technology and machine learning, our proposed methodology shows that MediaPipe can be efficiently used as a tool to detect complex hand gestures precisely. Although, sign language modelling using image processing techniques has evolved over the past few years but methods are complex with a requirement of high computational power. Time consumption to train a model is also high. From that perspective, this work provides new insights into this problem. Less computing power and the adaptability to smart devices makes the model robust and cost-effective. Training and testing with various sign language datasets show this framework can be adapted effectively for any regional sign language dataset and maximum accuracy can be obtained. Faster real-time detection demonstrates the model's efficiency better than the present state-of-arts. In the future, the work can be extended by introducing word detection of sign language from videos using Mediapipe's state-of-art and best possible classification algorithms

References:

[1] R. Sharma, R. Khapra, N. Dahiya. June 2020. Sign Language Gesture Recognition, pp.14-19

[2] W. Liu, Y. Fan, Z. Li, Z. Zhang. Jan 2015 . Rgbd video based human hand trajectory tracking and gesture recognition system in Mathematical Problems in Engineering,

[3] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C. L., & Grundmann, M. 2020. MediaPipe Hands: On-device Real-time Hand Tracking. arXiv preprint arXiv:2006.10214.

[4] Das, P., Ahmed, T., & Ali, M. F. 2020, June. Static Hand Gesture Recognition for American Sign Language using Deep Convolutional Neural Network. In 2020 IEEE Region 10 Symposium (TENSYMP) (pp. 1762-1765). IEEE.

[5] M. Taskiran, M. Killioglu and N. Kahraman. 2018 . A Real-Time System for Recognition of American Sign Language by using Deep Learning, 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, pp. 1-5, doi: 10.1109/TSP.2018.8441304.

[6] Nazmus Saquib and Ashikur Rahman. 2020. Application of machine learning techniques for real-time sign language detection using wearable sensors. In Proceedings of the 11th ACM Multimedia Systems Conference (MMSys '20). Association for Computing Machinery, New York, NY, USA, 178–189. DOI:https://doi.org/10.1145/3339825.3391869

[7] Dutta, K. K., & Bellary, S. A. S. 2017, September. Machine learning techniques for Indian sign language recognition. In 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC) (pp. 333-336). IEEE.

[8] Sahoo, Ashok. 2014. Indian sign language recognition using neural networks and kNN classifiers. Journal of Engineering and Applied Sciences. 9. 1255-1259.

[9] Raheja JL, Mishra A, Chaudary A. 2016 September . Indian Sign Language Recognition Using SVM 1. Pattern Recognition and Image Analysis.; 26(2).

[10] Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. 2014, September. Sign language recognition using convolutional neural networks. In European Conference on Computer Vision (pp. 572-578). Springer, Cham.