

SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING

S.Saravana Kumar¹, Vedant L. Iyengar²

¹ Professor, Department of Computer Science and Engineering , Karpagam College of Engineering, Coimbatore.

² Student, Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore.

ABSTRACT

This is a proposal for a dynamic Sign Language Recognition System. The end user can be able to learn and understand sign language through this system. Machine Learning is an up and coming field which forms the basis of Artificial Intelligence. Machine Learning has been widely used for optical character recognition that can recognize characters, written or printed. The principles of supervised learning have allowed us to come up with many classification, prediction and identification systems. One such example is the existing Facial Recognition technology in Apple's iPhone X. This proposal aims at providing a learning tool for differently abled people to identify and learn Sign Language Symbolisation, specifically, the ASL (American Sign Language). Though existing systems identify sign language with sufficient accuracy, this proposal incorporates the recognition from a live video feed. As a result it provides more interactivity than existing systems.

Keywords: sign language, image recognition, machine learning, Support Vector Machines (SVM)

1.INTRODUCTION

Humans have a variety of methods for communicating with each other. This includes actions like bodily gestures, face expressions, spoken words, etc. However, people who are hard on hearing are limited to communicate with hand motions. People with hearing disabilities and/or speech

disabilities use a standard sign language which cannot be understood by people who do not know it. Also, learning sign language is hindered by their disability. A modern learning and translation tool for sign language implemented in Machine Learning can significantly affect the ease of Sign Language Communication. This tool will aim to do the following:

- Obtain a video feed from the camera
- Classify and display the equivalent English Alphabet for the American Sign Language Alphabet.

However, the following factors pose some challenges to this system:

1. The conditions of the illumination in the place where this system is used has to be considered, as it plays an important role in the accuracy of the recognition.
2. It has to be made sure that the hand gesture is at an ideal distance from the camera
3. The camera must capture the hand upto the wrist at least, and it must not focus on background subjects
4. Identifying the features of the symbols that can be used in the system to get a better accuracy

This system aims to consider these challenges and recognize Sign Language Symbols (except those which require movement of hands)

II.PROBLEM STATEMENT

In our progressive society, it is necessary to socialize with all people to whether for recreation or for a purpose. Communication is important for every human being. However, people who have a hearing disability and/or a speech disability need a

different way to communicate other than vocal communication. They resort to sign language to communicate with each other. However, Sign Language requires a lot of training to be understood and learn and not every person may understand what the sign language gestures mean. Learning sign language is also time consuming as there are no effective, portable tool for recognizing sign language. Hearing or Speech disabled people who know Sign Language require a translator who also knows Sign Language to explain their thoughts to other people in an effective manner. To help overcome these problems, this system helps hearing or speech disabled people to learn as well as translate their sign language.

III.RELATED WORK

Sign language recognition is a topic which has been addressed multiple times and is not new. Over the last few years, different classifiers have been applied to solve this problem including linear classifiers, neural networks and Bayesian networks. Linear models are easy to work with, but require complex feature extraction for increased accuracy. Work done by Singha and Das helped them achieve accuracy of 96% on 10 classes for images of gestures of one hand using Karhunen-Loeve Transforms Real-time American Sign Language Recognition with Convolutional Neural Networks [1]. These transformations rotate the input frame

feed and set up a new coordinate system based on the variance of the data. This is preceded by applying image pre-processing to extract the significant part of the images. They use a classical linear classifier to identify fingers pointing in directions. Work done by Sharma uses a combination of Support Vector Machines and k-Nearest Neighbours to recognize the symbol after background subtraction and noise removal [2]. They make use of contour tracing, which represents hand contours. This system attained an accuracy of 61%. Bayesian networks like Hidden Markov Models have also achieved high accuracies [3]. These capture temporal patterns accurately, but they require clearly defined models that are defined prior to learning. Starner and Pentland achieved a highly accurate system (99%) using a Hidden Markov Model though the high accuracy was due to the usage of a sensorized glove [4]. The glove helped them acquire accurate 3D spatial details. Neural networks have also been used to achieve Sign Language Translation. Neural networks have an advantage, which however, comes with a trade-off. They improve the accuracy by determining which model to make use of but this comes with the cost of compromising slower speed of training and usage. To date, most have been relatively shallow. Admasu and Raimond classified Ethiopian Sign Language correctly in 98.5% of cases using

a feed-forward Neural Network [5]. They use a significant amount of image preprocessing, including image size normalization, image background subtraction, contrast adjustment, and image segmentation. Admasu and Raimond extracted features with a Gabor Filter and Principal Component Analysis. L.Pigou used a Microsoft Kinect to identify full-body features of a person showing the gesture [6]. The Kinect allows capture of depth features, which aids significantly in classifying the signs. Indian Sign Language Recognizer System, developed by Sanil Jain and KV Sameer Raja [7] extracted features from coloured images and used Gaussian random and the Histogram of Gradients (HoG). They used 3 subjects to test the system, after which an accuracy of 54 % was achieved when used by a different person.

IV.PROBLEM WITH EXISTING SYSTEM

Systems exist where Neural Networks are used to detect Sign Language. However these require more computational power which cannot be carried out by, say a handheld device like a mobile phone or a Tablet. Further, some other systems require a High-Tech glove with motion sensors to capture the gesture in 3 dimensions or even a Microsoft Kinect, both of which are pretty

expensive. These also impose scalability issues due to the equipment dependency. Benefits of portability has to be compromised, with the equipment having to be carried everywhere where the system is to be used.

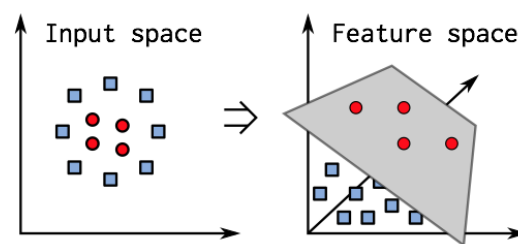
V. PROPOSED SYSTEM

The proposed system consists of a camera which captures a video feed. This video feed is processed frame by frame. A library called OpenCV [9] is used to process this video feed. The contours for the frames in the video are identified by darkening the image and obtaining the white border of the hand. This border is used to identify the contours of the hand. The contours are then used to identify the type of symbol provided in the video feed.

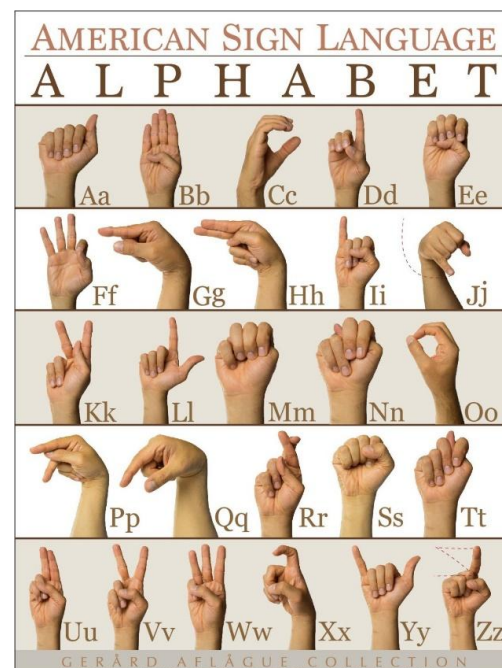
Before actually using this system as mentioned above, it is required to train the system with the dataset of images which contain alphabets of the Sign Language. These are mapped to their equivalent English alphabet and fed to the system. The system gets trained on this data and stores the training results as file. A support vector machine model [8] is used. In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. Given a set of training examples, each

marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier.

A library called scikit-learn [10] provides the necessary SVM Model ready for training.



Feature based classification by using an SVM



The American Sign Language Alphabet

VI. CONCLUSION

An American Sign Language Interpreter has been implemented based on a Support Vector Machine Classifier (SVM). The system requires some constraints, like a white background and the palm of the hand to face the camera. Further, similar symbols were sometimes misinterpreted for one another. This could be due to the limited training provided to the system. The system might overcome these limitations if a more detailed dataset in different environmental conditions is provided for training. Also letters that needed movement of the hands could not be properly identified.

VII. FUTURE WORK

In future, the following enhancements can be used to improve the system:

Pre Training the model with detailed dataset:

The system can be trained with a more detailed dataset with thousands of samples for each letter of the alphabet, spanning different environmental conditions, lighting, different hand positions, and skin tones for optimizing.

Image enhancement:

The process of interpretation could be made much easier if there were a pre-processing

stage for each frame that performed contrast processing, background subtraction, and identifying and cropping the main focal subject.

VIII. REFERENCES

- [1] Singha, J. and Das, K. "Hand Gesture Recognition Based on Karhunen-Loeve Transform", Mobile and Embedded 232 Technology International Conference (MECON), January 17-18, 2013, India. 365-371.
- [2] R. Sharma et al. Recognition of Single Handed Sign Language Gestures using Contour Tracing descriptor. Proceedings of the World Congress on Engineering 2013 Vol. II, WCE 2013, July 3 - 5, 2013, London, U.K
- [3] R.Karthikeyan, C.Jothi Kumar, "Improved Reputation System for Wireless Sensor Networks (WSNS) ", International Journal of Innovations in Scientific and Engineering Research (IJISER), Vol.1, no.3, pp.191-195, 2014.
- [4] M. Jeballi et al. Extension of Hidden Markov Model for Recognizing Large Vocabulary of Sign Language. International Journal of Artificial Intelligence & Applications 4(2); 35-42, 2013
- [5] T.Starner and A. Pentland. Real-Time American Sign Language Recognition from Video Using Hidden Markov Models.

Computational Imaging and Vision, 9(1);
227-243, 1997.

[6] Y.F. Admasu, and K. Raimond,
Ethiopian Sign Language Recognition
Using Artificial Neural Network. 10th
International Conference on Intelligent
Systems Design and Applications, 2010.
995-1000

[7] L. Pigou et al. Sign Language
Recognition Using Convolutional Neural
Networks. European Conference on
Computer Vision 6-12 September 2014

[8] Indian Sign Language Character
Recognition by Sanil Jain and K.V.Sameer
Raja

[9] Video on Support Vector Machines
(Instructed by Patrick Winston):
https://www.youtube.com/watch?v=_PwhiWxHK8o

[10] OpenCV Image processing library
www.learnopencv.com

[11] Scikit Learn - Documentation
<http://scikit-learn.org/stable>

