**Paper Review on:** Sign Language Recognition using Image-based Hand Gesture May 2019

**Abstract:**

In order to improve sign language education, immediate feedback on sign language gestures can be added. For that purpose, the authors of this paper classified images using webcam images. They can successfully classify up to 93% of the signs using linear or Gaussian Kernel Support Vector Machine or SVM. Implementing techniques in machine learning and image processing, they hope to achieve higher level of accuracy.

**Introduction:**

The author of this paper aims to increase the accuracy of their model by implementing techniques in machine learning and image processing. The current sign language literacy in India, like in most countries, is very low. People have to pay a large sum for teaching classes or practice by themselves. Neither is convenient, the authors aim to change that by creating a machine learning based tool that uses webcam to take input images and shows the accuracy of the input hand sign instantly. This is less costly and time consuming, and more convenient. They took webcam images as input and used a pyramid histogram of Oriented Gradient as a feature for SVM with 70-30 cross validation and Hidden Markov Model. They were able to achieve an accuracy rate of 97%.

**Literature Review:**

They used a custom method of taking pictures through webcams and teaching their model the corresponding ASL alphabets of those pictures. They used image processing before doing any machine learning and other methods. The author didn’t review any specific literature related to their work.

**Methods:**

Their dataset comprises of 1800 webcam images of hand gestures in ASL (American Sign Language). They had their hand rotated slightly while taking pictures to avoid having too similar data. They took a single image of the empty background to do background subtraction for the rest of the images in both training and test set.

They processed each image as, first converting it to greyscale from RGB. This loses color but makes the next steps much faster. Then they removed the background. They put a threshold value to each pixel and the background pixels were under a certain threshold. Each pixel under that threshold are removed, keeping only the foreground hand gesture part. Then for the next step they made a silhouette out of the hand, by binarizing the image. They turned all the background pixels to white and the rest of the portion of the image, the hand, into black.

They split their dataset into 85-15 for train-test sets respectively. They used multiple classifiers and features extraction methods. Then they used 10 Fold Cross Validation and training using KNN (K-Nearest Neighbors) to get their trained model. To recognize the test images, they used same image processing, feature extraction methods and KNN to classify and predict input image.

**Results:**

The authors used the dataset consisting of 15 letters, skipping J. They tested using Liner Kernel, Gaussian Kernel and KNN classifiers however, during testing, Gaussian Kernel had the highest accuracy of 93.5% for all 25 alphabets.

**Discussion:**

The authors of this paper used Gaussian Kernel Classifier, 10 Fold Cross validation and KNN. However, we used our custom CNN model, RestNet18 and VGG 16 models. The authors processed their images into greyscale and resized them to really small sizes, we didn’t do these processing. Their processing work was faster, however their accuracy suffered as a result. In all three of our models, our accuracy was always over 99.00%.

**Conclusion:**

This paper presents an overview on sign language using image based hand gestures captured from webcams. Since, it had an accuracy of 93.5%, there is still room for improvement. However, it was meant to be a fast and reliable method for people to test their ASL hand gestures. In that perspective, this work achieved that goal.

**Reference:**

URL:

https://www.researchgate.net/publication/347444229\_Sign\_Language\_Recognition\_using\_Image-based\_Hand\_Gesture