Sign Language Recognition from Hand gestures using GAN

Vipin Chaudhary
2014119
vipin14119@iiitd.ac.in

Praveen Kumar Jhanwar
2014078
praveen14078@iiitd.ac.in

Abstract

Hand gesture is one of the method which can be easily used to communicate. With the help of sign language these differently abled people can easily express their thoughts and emotions to other people.In order to achieve this purpose they need to keep a translating device along with them, whose accuracy depends on the accurate recognition of hand gesture. In this paper we will try to compare and analyze different methods to recognize hand gestures. We are going to recognize American Sign Language (ASL) digits and letters.

1. Introduction

The alternative to sign language is written language, which is problematic and impractical in case of emergency. In order to remove this obstacle and to enable communication with ease, we aim to build a real time system that translates a user's video of ASL signs into text. In order to achieve this goal, three tasks need to be done:

1. Input: video of user ASL signs
2. Recognizing the letter in each frame of input video
3. Output: Reconstructing the most likely word from the previous step.

There are certain challenges that we need to solve in order to solve the gesture recognition problem

1. Lighting, background and camera position
2. Sign boundary: Since one sign will persist over more than one frame so we need to be able to identify when one sign ends and the other begins
3. Occlusion: Some or all fingers or an entire hand

Since Neural networks have been applied to this problem previously also but most of them requires a 3D motion tracking component like Gloves or Kinect, So we are getting rid of this constraint by using simple input video of ASL signs

2. Related Work

According to this [3] They extract features images by discrete wavelet transform and then dimension of feature is reduced by Linear discriminant analysis. Then both LDA and SVM are applied to recognize sign language symbols and achieved an accuracy of 97.3% on random sign symbolic dataset of gestures.

According to [2],they use Naive Bayes approach to classify hand gestures in Indonesian Sign Language. they extracts 19 features from signs using Leap motion device and obtain an accuracy of around 80.5% in the trained data from ideal environment and 70.7% to the instant untrained dataset.

There are other linear methods too, that use different and sophisticated feature descriptors on preprocessed data that are mentioned in [1]

Some neural networks have been used to tackle this problem recently, but they use a significant amount of image preprocessing and feature extractions.Site obtained an accuracy of around 95% on 10 classes of hand gestures using KL transform. They use a linear classifier to distinguish between hand gestures including thumbs up, index finger pointing left and right, and numbers. They used SVM and KNN for classification.

[4] represented a real time HGR system based on ASL recognition with greater accuracy. This system acquires gesture images of ASL with black background from mobile video camera for feature extraction.They extract five features like fingertips, eccentricity, pixel segmentation, elongatedness and pixel segmentation in preprocessing. For feature extraction they use a new algorithm K convex hull method which can detect fingertip with high accuracy. They used feed forward Artificial Neural Network (ANN) for training a network using 30 feature vectors to recognize 37 ASL classes including alphabets and numbers.

3. Problem Statement

In this paper we will be solving sign language recognition problem using Deep learning techniques and will compare the results with other techniques and few prior work

papers as well. We will be specifically targeting American Sign Language(ASL). We will be solving this the problem using unsupervised techniques and then compare results with supervised version.

4. Approach

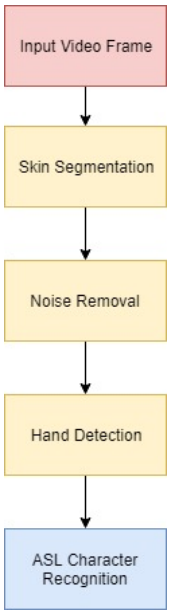

Figure 1. Archutecture Pipeline

Our approach to solve this paper is as follows

1. Extracting frames from input videos
2. Pre-processing of frames Contrast enhancement, Edge detection of hands and Image Segmentation
3. Use a linear model with feature extractors like SIFT, LDA etc to evaluate it's performance on out dataset and the Massey university ASL dataset
4. Transfer learning of a pretrained CNN and fine tune it according to our dataset
5. Extracting Vectors of features for every frames using fine tuned CNN
6. Object recognition and blob creation in the image using YOLO/faster-RCNN
7. Results comparisons

5. Experimental Setup

We used Massey University ASL dataset [5], which contains 2524 images of ASL signs in which each image has segmengted out hand sign on black background of all the digits and alphabets. We created our own American signs dataset as well and compared results on both of them in the end. We used OpenCV in python for dataset pre-processing purpose like segmentation and edge detection. We used Pytorch as our deep learning framework. We used Resnet18 as our pretrained CNN model that is trained on ImageNet dataset.

6. Result

Figure 2. Hand Segmentation

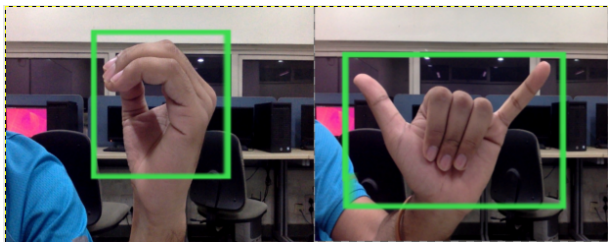

Figure 3. Hand Detection and Recognition

Model	Train Accuracy(%)	Test Accuracy(%)	Epochs
SURF	63.6	37.9	-
SIFT	57.6	33.5	-
Vanilla CNN-32	54.5	24.6	1000
Vanilla CNN-64	95.3	43.4	1000
ResNet18	89.7	78.8	500

7. Contribution

1. Pre-processing of dataset - Praveen
2. Dataset creation - Both
3. linear model - Praveen
4. Vanilla CNNs - Vipin
5. Transfer Learning - Both
6. Skin Segmentation - Both
7. Results compilation and comparison - Both

8. Discussion

We have built a end to end system for ASL character recognition. As far as our approach is concerned, we think we can make it faster by using better methods for hand detection like YOLO-based object detection methods and also since no deep learning method was there to segment skin, so that remains another field to research upon.

References

[1] S. A. V. Brandon Garcia. http://cs231n.stanford.edu/reports/2016/pdfs/214$_{r}eport.pdf$.

[2] R. S. L. Eko Pramunanto, Surya Sumpeno. Classification of hand gesture in indonesian sign language system using naive bayes. *IEEE*, 2017.

[3] N. Kumar. Motion trajectory based human face and hands tracking for sign language recognition. *IEEE*, 2017.

[4] . K. Tlay Karaylan. Sign language recognition. *Computer Science and Engineering (UBMK) 2017 International Conference on, pp*, 2017.

[5] M. university. http://www.massey.ac.nz/ albarcza/gesture$_{d}ataset2012.html$.