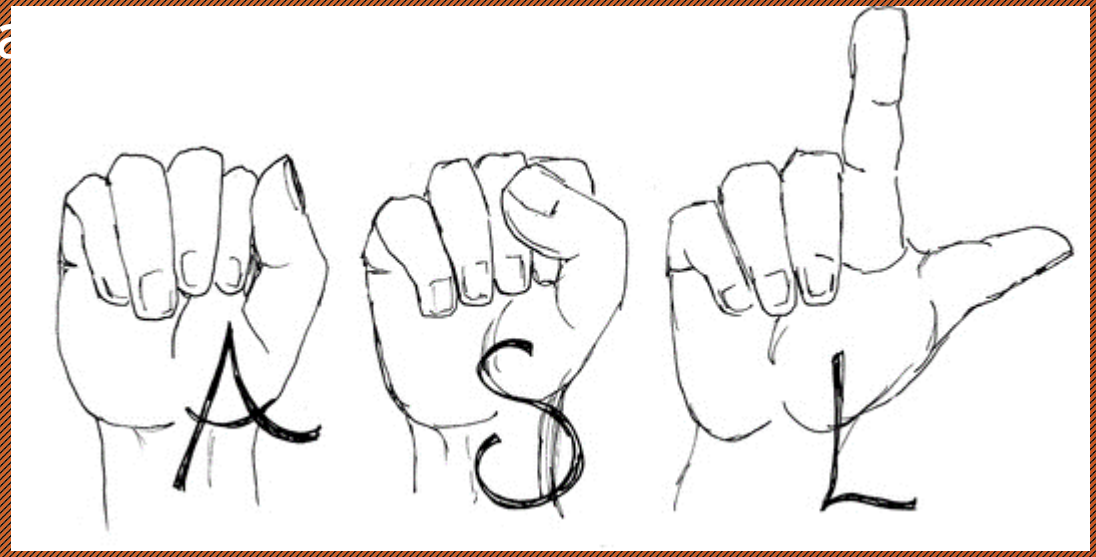
Machine Perception End-Term Project-2016

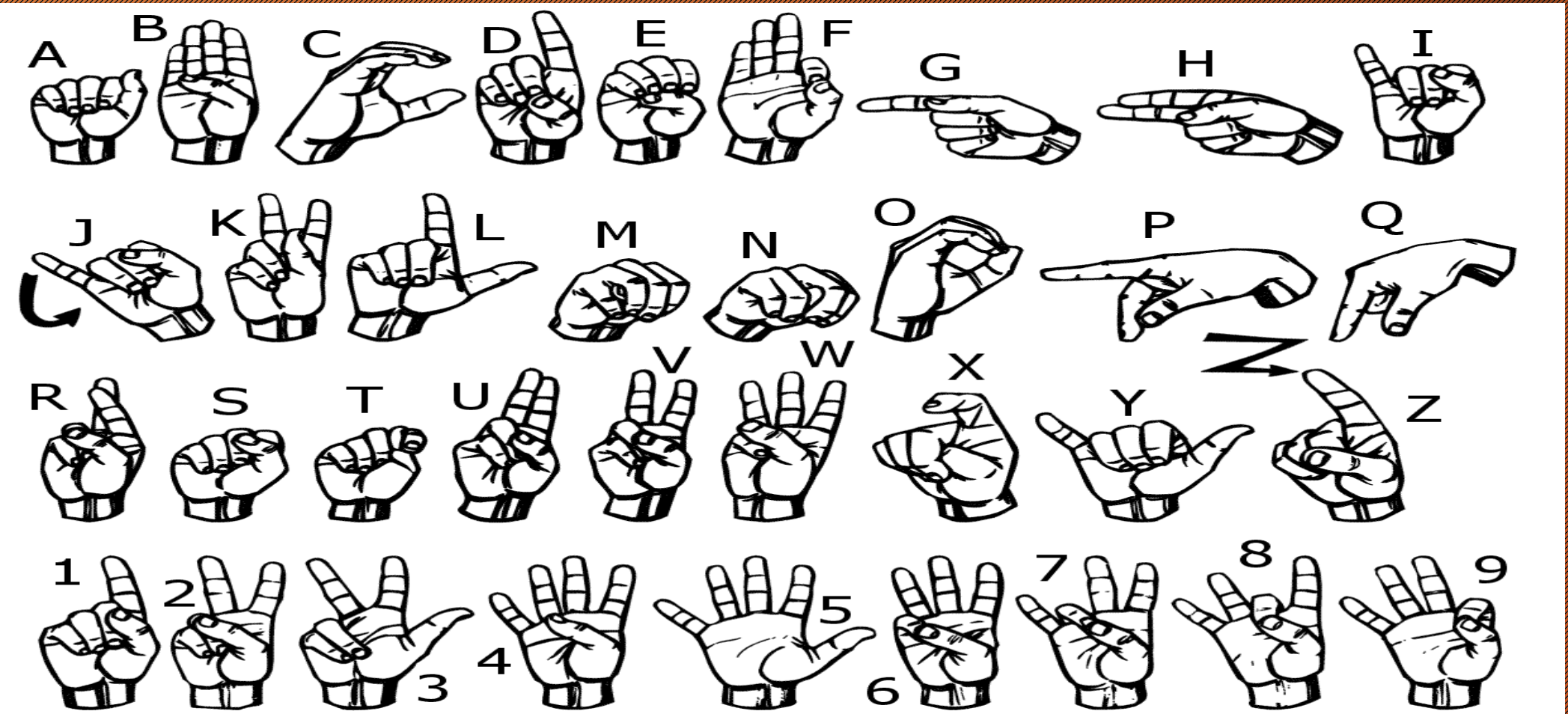
ASL Translator

What is ASL ?

- American Sign Language (ASL) is the predominant sign language of Deaf communities in the United States and most of Anglophone Canada.
- ASL use has propagated widely via schools for the deaf and Deaf community orga

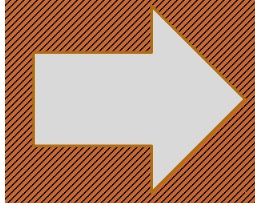


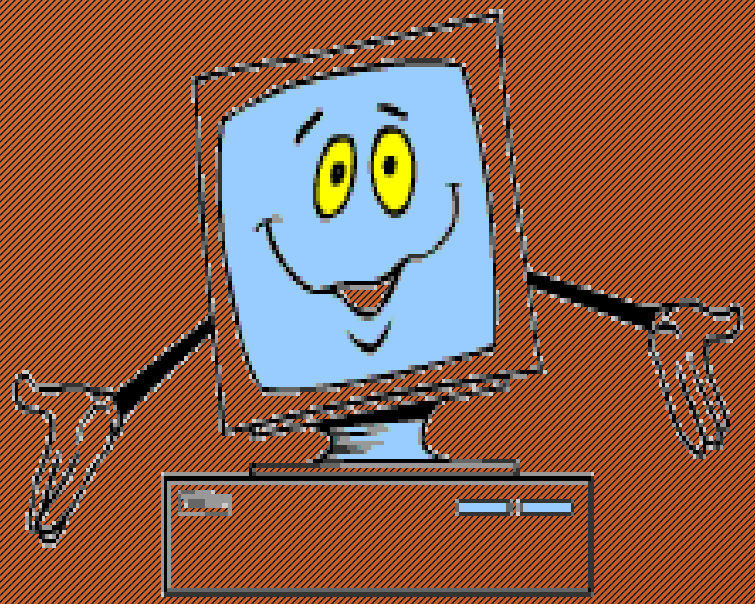
ASL at a Glance

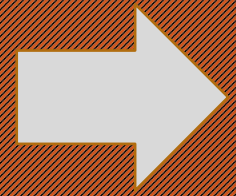


Proposed System

- A software that detects ASL using external hardware like a video camera and converts it into text.

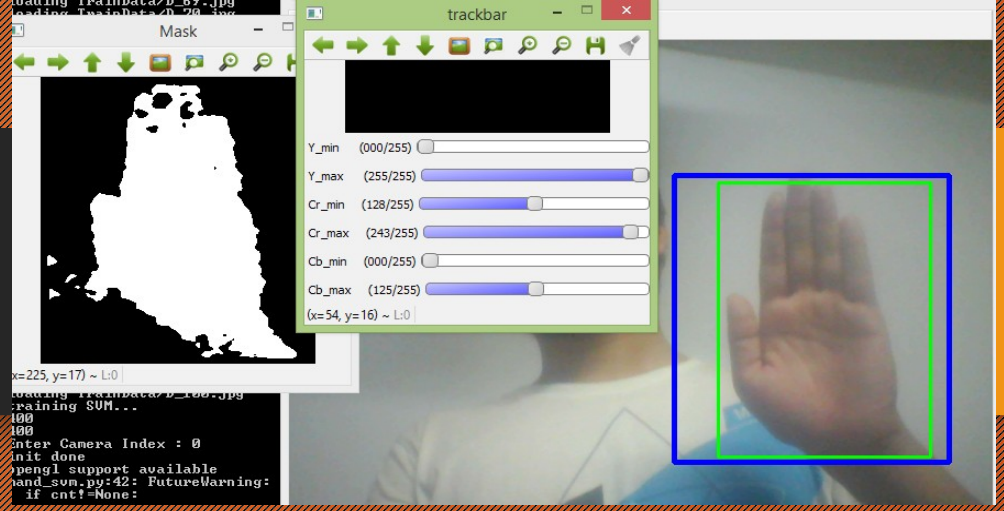






I Love You

Previous Work Done



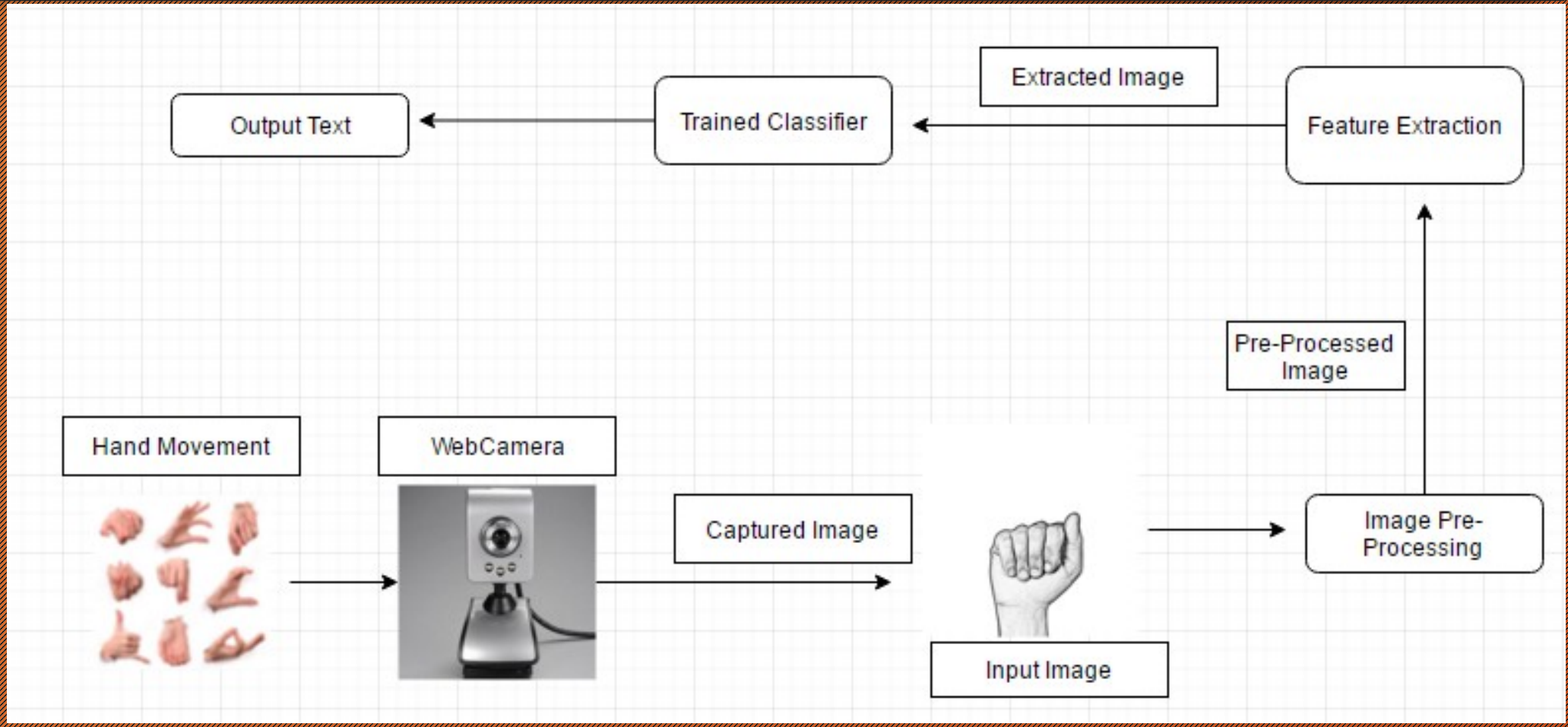
- We used the YCbCr colorspace to detect hand gestures and were able to get training data from them.With the trackbar we were able to get good dataset with variety of people with light skin or dark skin color.
- Used the maxContour approach to get the contour of hand and thus avoiding any other skin pixels apart from hand to come into picture.
- Problems we had:
  1. Overlapping of skin pixels.
  2. Max contour could be anywhere in the whole frame.

  eg. Face contour can be more than Hand contour.

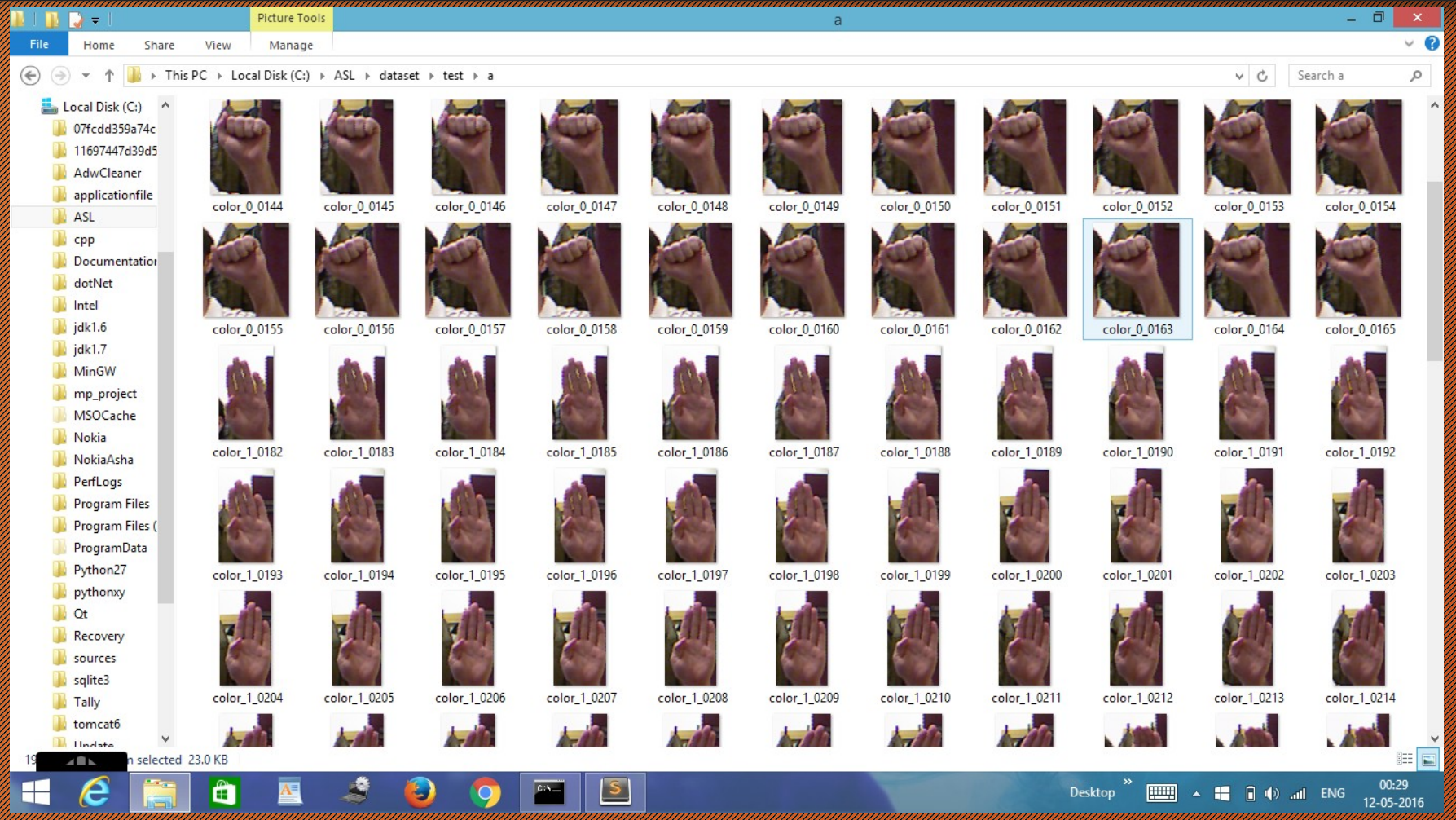
Solution to the Problem

- We used the Region of interest approach.
- We provided a rectangle on the screen for the user to allow him to fit his hand in that box.
- With this approach skin pixel overlapping , etc. problems got resolved as user knows where his hand should go in.
- Nevertheless , the hand detection part can be improved by other approaches like depth sensors , Kinect , Haar Cascade ,etc.
- But we wanted an easy approach where your webcam can become a powerful tool rather than providing extra hardware or doing extra computations just to detect hand.

Workflow


Output Text
Trained Classifier
Extracted Image
Feature Extraction
Pre-Processed Image
Hand Movement
WebCamera
Captured Image
Input Image
Image Pre-Processing

DataSet



Training- 5000
Test - 1000

Approach 1: Bag of Words

- We used the principles of SIFT , where the features and their descriptors are extracted and a Histogram of Feature words is created.
- Why this model?

  Unlike our color space based segmentation we wanted to play with natural images with backgrounds and different sizes.
- We trained the SVM classifier with dataset consisting of 4000 images and got a model .
- This model was used to test the dataset consisting of 500 images.

Approach 1: Result

- Result:
  - 355 correct and 145 wrong
  - Accuracy ~ 71%

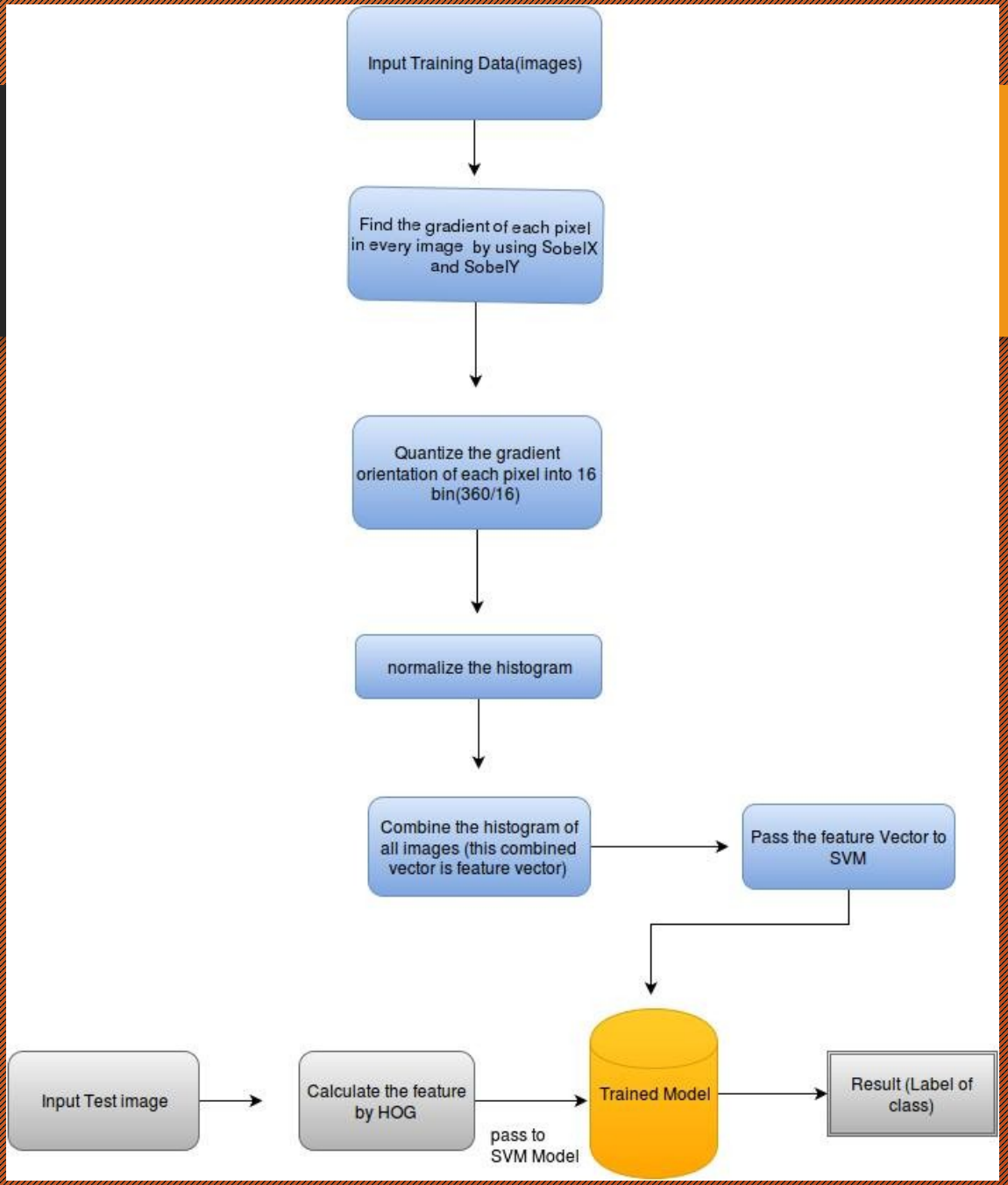
Approach 2: HOG

- It is a feature descriptor method just like SIFT
- Counts occurrences of gradient orientation in localized portion of an image.
- This approach creates bins of say 16 , and clubs the histogram of different samples based on magnitude and angle.
- Here we are segmenting the hand first using YCbCr color space and then processing the image through HOG and providing to model.
- We trained the SVM classifier with dataset consisting of 5000 images and got a model .
- This model was used to test the dataset consisting of 1000 images.

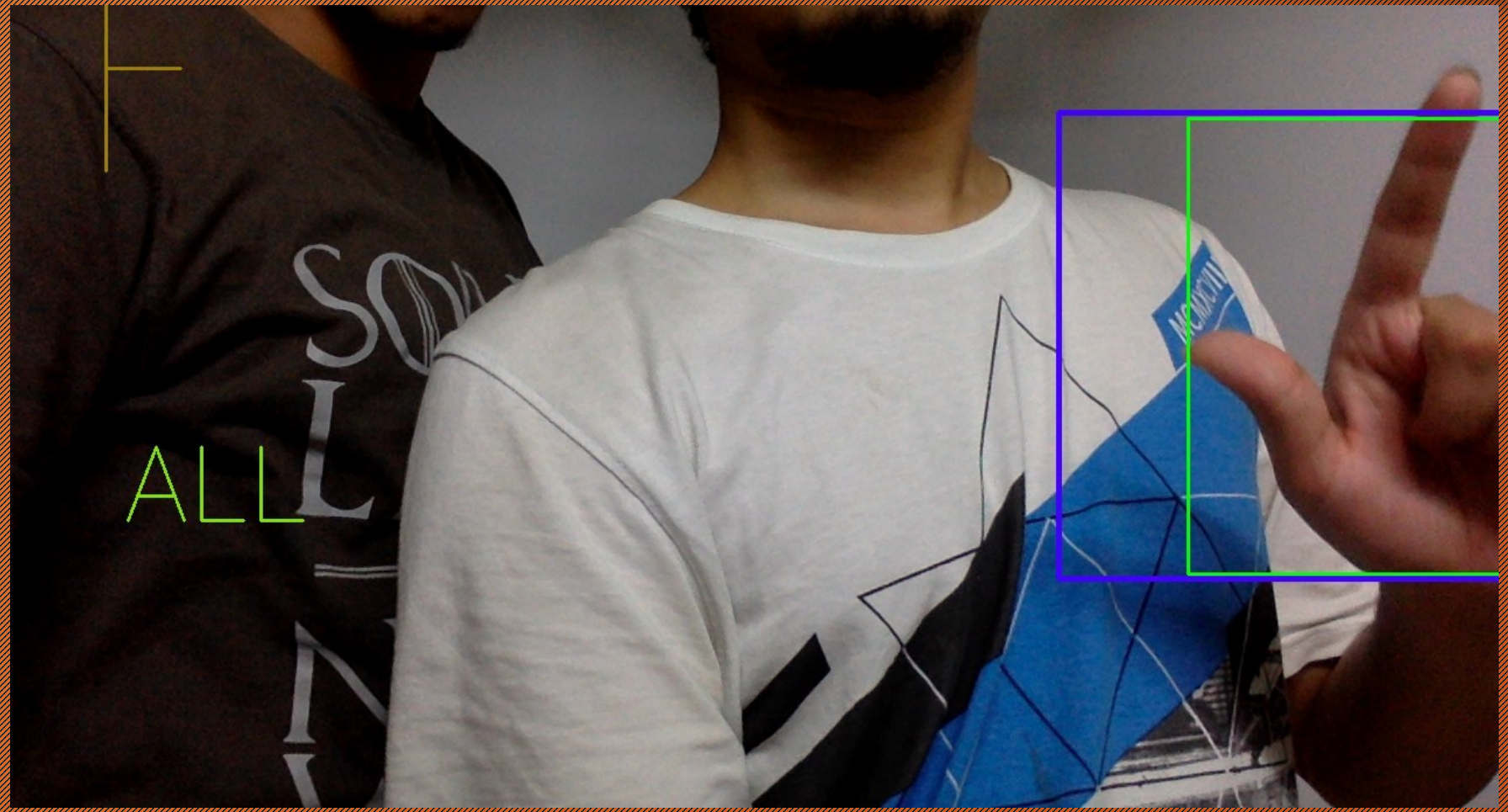
Approach 2: Result

- Trained for 15 alphabet classes from A to O.
- So far giving good results.
- Result:
  - 880 correct and 120 wrong
  - Accuracy ~ 88.7%

Approach 2: FlowDiagram



Approach 2: Output Results



We are able to predict what a person is trying to say by forming sentences through alphabets.

Future Work

- Enhanced prediction by using other approaches like SIFT with ADABOOST or applying HMM to predict finger spelled words.

- Implementing it as a Smartphone app.

Reference

- Wikipedia
  https://en.wikipedia.org/wiki/American_Sign_Language
- Python OpenCV doc 2.4.10
- Static Hand Gesture Recognition Based on HOG with Kinect - Hui Li , Lei Yang, Xiaoyu Wu,Shengmiao Xu ,Youwen Wang
- Hand Gesture Recognition using Histogram of Oriented Gradients and Partial Least Squares Regression - Arindam Misra , Abe Takashi , Takayuki Okatani