

COL780 Assignment 3 Report

Ananya Mathur - 2020CS50416

April 2024

1 Directory Structure

```
2020CS50416.zip
├── main.py
├── hog.py
├── roi.py
└── bonus.py
└── report.pdf
└── README.md
```

2 Tools and Technology Used

I have used sklearn for SVM classifier and metrics, opencv for basic image processing functions, mediapipe for hand detection as provided in roi.py, numpy for mathematical computation, os for creating new directories and reading files from given directories, pygame for playing and pausing music.

3 Algorithm

Feature detection is done using HOG and SVM is used as classifier.

Before extracting HOG features, the Region of Interest of hands in the image is detected using the roi.py file provided.

HOG Feature Extraction Algorithm :

- Calculate gradient magnitude and direction at each pixel.
- Group gradient directions of all pixels in a cell into a specified number of orientation bins.
- Group cells into blocks of same size with a given stride and normalise the cell histogram based on other cell histograms

- All the normalized histograms from all the blocks will be added up into a single feature vector. This feature vector is called the HOG descriptor.

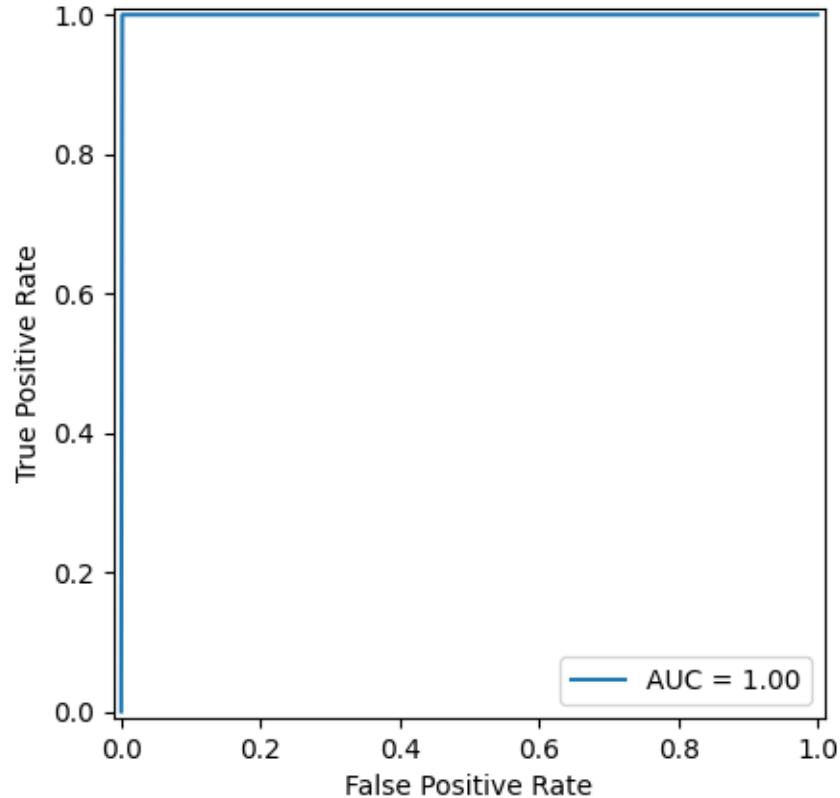
After feature extraction of all the train images using HOG Descriptor, a Linear SVM Classifier model is fit upon the train data and labels (0 for open hand , 1 for closed hand). HOG Descriptors for test data are also extracted similarly and predictions are made based on the SVM model for each test image.

4 Results and Conclusion

Classification Report

Evaluating classifier on test data ...				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	625
1	1.00	1.00	1.00	328
accuracy			1.00	953
macro avg	1.00	1.00	1.00	953
weighted avg	1.00	1.00	1.00	953

ROC Curve



Metrics for Hand Recognition Task

- Accuracy = 100 %
- Precision = 1
- Recall = 1
- F1 Score = 1
- True Positive Rate = $TP / (TP + FN) = 1$
- False Positive Rate = $FP / (FP + TN) = 0$
- AUROC = 1

The train and validation data is very similar hence the task is easy and we can easily achieve 100 % accuracy, which happens.

5 Problems Faced

There were several images with 2 hands and both hands were of different classes. This caused problems in assigning test labels and analysis.

6 Bonus

I have created a music player which runs with gestures as a real life application of the task. For this I have created my own dataset comprising of 8 different gestures : Unpause, Pause, Next, One, Two, Three, Four, Five. My model is trained using HOG descriptors and SVM classifier is used to classify images received as test inputs. A class from 1 to 8 is assigned to each image and an action accordingly is performed. Gestures 1 to 5 play respective song numbers from song database, pause gesture pauses the currently playing song, unpause plays the currently paused song while next plays the next song in line.



Figure 1: Gestures

7 References

<https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f>