Name: Ananta Uprety
Email: auprety6@gatech.edu

Road Sign Detection

Introduction

Detecting road signs is one of the interesting challenges in the computer vision space. The problem in itself is very interesting in comparison to human vision because humans are very capable of distinguishing between different signs even when presented with myriad of conditions. Humans can detect signs in sunny, cloudy or rainy weather. Humans can also detect in different lights, also can make prediction based on the shape of the signs. However, detection and classification of signs is challenging

While signs are well defined through laws and designed to be easy to spot, there are still plenty of challenges for a computer. Even though they are bright and easy to spot they similar across categories. This causes issues in creating classifiers. Some of the signs are also old and faded and might have lost their color. The signs might be occluded or bent. During different lighting conditions the colors look different. In urban environments a lot of objects look similar to signs. That is why applying a traditional classifier to do the job is not a simple solution. However, in this case I have taken to using combined approaches from both the computer vision space and the machine learning space. In order to verify my approach I have decided to use the LISA dataset which contains annotated images and videos of US traffic signs.

LISA Dataset

The LISA Traffic Sign Dataset contains videos and annotated frames containing US traffic signs. It is very good representation of the US traffic signs as it contains about 40 different traffic signs which have been annotated with the location of sign within the frame and also the name of the sign. It also represents the real world images as it was taken from a moving vehicle around busy streets. However, there is a difference in the frequency of the signs.After taking a sample of dataset for training and also hand validating the signs I was left with about 6500 signs that could be used for training. Some signs have lot more representation than others as can be seen from the image below. Stop, Speed Limit, Pedestrian, Signal Ahead signs were overrepresented in the dataset. Hence, I decided to use these signs for my training. I decided to us the videos portion of the dataset for testing. I also decided to focus on the signs that were the most present.

Fig i) Different signs in the dataset

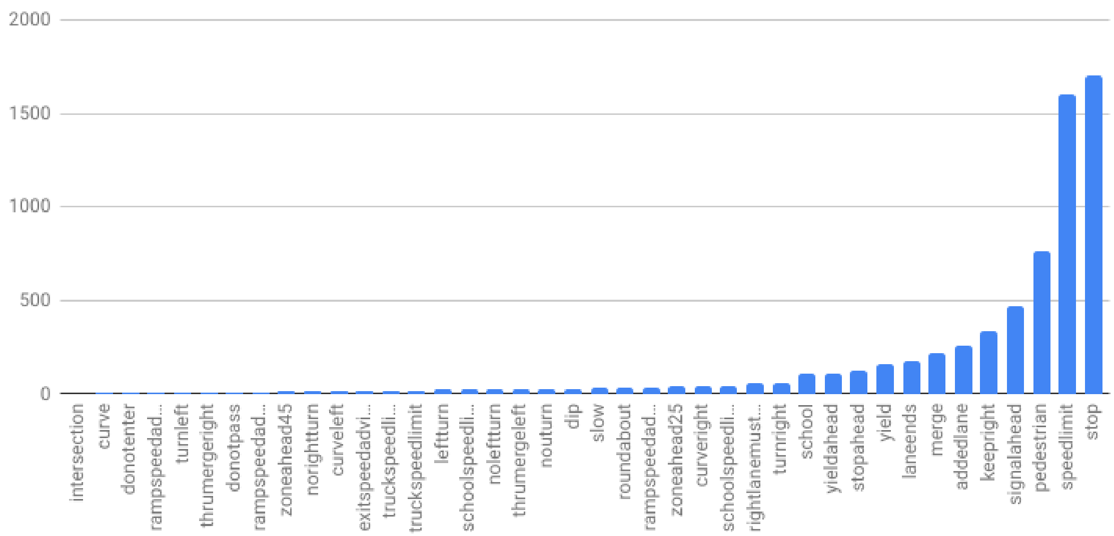

Fig ii) Frequency of the signs

Approach

After reading through the current day literature on the topic of traffic sign detection, it seemed that researchers have gotten more success from breaking down the task into two parts. The first task is to detect and identify where the sign might exist and the next part is to use a classifier to classify different types of signs. Hence, I have followed a similar approach as well. I have decided to use the detection method proposed by Viola and Jones for face detection using HAAR like features but, adapted for traffic sign detection. After the signs have been found on the images, I have used SVM classifier to classify the different signs using the HOG features of the images. The images are also put through some preprocessing to remove noise.

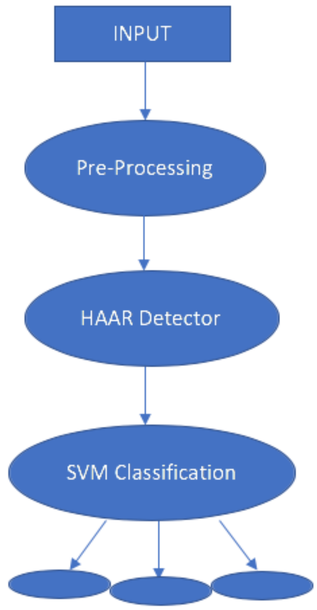

Fig iii) Outline of the process

i) Pre-processing Step:

It consists of cleaning the images of noise using the fastNlMeansDenoising algorithm. The process also consists of converting the image to grayscale such that it can be processed by the HAAR detector, another set of images is saved for the classification phase which relies on the color histogram.

ii) Detection :

Detection step is the most critical step and is also more important because the following steps work with the features from this stage. In this case, the detection step consists of training a HAAR detector that was proposed by Viola and Jones[1]. In this case, it uses the integral images and calculates the HAAR Features which are trained by using cascade of classifiers. Using HAAR features is better than using the template matching or searching through the image space because the signs could be translated, rotated or stretched. Template matching only works with the exact shape and size. However, HAAR features are invariant to size hence, I trained different signs with Cascade of classifiers. I also used the size of the HAAR features to be 32 X 32.

Initially, I trained all the signs with the HAAR classifier but, it did not get good results as can be seen from the below example. In the next stage I created different HAAR classifiers based on signs. I used one detector for stop signs which is good at detecting circular regions. I used another one to detect diamond like shapes which are good at detecting Signal Ahead, Left Turn, Right Turn, Merge and Roundabouts. I trained another one for the square signs like the Speed limit signs. However, the speed limit detector did not work as expected.

Detectors	Annotated labels	HAAR Detection	Possible signs Detected Per Image
Circular Detector	200	7471	37.35
Diamond Detector	200	3207	16.03
Square Detector	200	3250	16.25

Table i) Lots of false positives on HAAR Detection

iv) Image shows false positives for the HAAR detector and one true positive

As can be seen from the table the HAAR detector is not very precise as it gives a lot of false positives which depends on the surrounding environment. In an urban environment a lot of shapes look similar to a sign. As can be seen from the image. The detector has classified the window from the house to be a possible sign.

For another approach to detecting stop signs, I also tried the Fast Radial Symmetry algorithm but, it did not bear good results.

iii) Classification:

After the detection phase, the classification phase takes in the results fed from the detection phase and classifies the sub-images into different types of signs. The classifier I have chosen is the HOG based Linear SVM. The histogram of oriented gradients HOG feature allows for variations in shape even though it keeps the gradients of the dominant pixels intact. Using this as an input to SVM can be trained to classify multiple signs. I used the HOG feature to be of size 32X32 as well. I also used one vs. all SVM strategy by training each svm with a negative examples of the other svm. But, in order to break down the testing I broke apart the SVM into different ones for each sign. But, below is the confusion matrix for the signs that I have trained SVM for 5 different signs that I focused on and other signs in the dataset.

	stop	Turn right	Signal ahead	pedestrian	merge	Other signs
stop	369	1	0	4	5	444
Turn right	0	11	0	0	0	2
signal ahead	8	0	132	4	0	8
pedestrian	3	0	0	156	0	18
merge	0	0	0	0	35	1
Other signs	4	0	4	13	5	289

Table ii) Confusion Matrix for the SVM training with 65% Accuracy

Using the above process I was able to get some decent results on the signs that I trained on. I was able to filter out the noise and keep the sign as can be seen below.

I ran the test against videos of type Stop, Pedestrian, Signal Ahead, and KeepRight. For each sign I compare the number of signs detected and if they match to the annotation provided. The total accuracy for the following signs is 81%. The results are below:

Fig vii) Correct classification of the signal and the discarding the noise by the classifier.

Signs	Annotated Signs	Detected	Accuracy(%)
Stop	162	103	63.58
Signal Ahead	62	46	74.19
Pedestrian	27	16	59.25
Keep Right	7	2	28.57
Total	258	209	64.72

Table iii) Accuracy results for each sign tested

Video Link: https://youtu.be/y0OSgp93A7o

Conclusion:

Given the amount of time I think the approach I took generated good results even though it was not in par with the current research. I would use the cascaded version of the SVM with boosting instead of just many to one. I would invest more time in preprocessing the data and also try few more HAAR features with 20X20, and 30X30 feature sizes. I would also spend more time looking into Histogram based approach especially the Histogram Intersection kernel for image classification. I would also apply the tracking of the signs once the sign has been detected so as not to lose the sign after detection.

References:

[1] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.

[2] Y. Xie, L.-F. Liu, C.-H. Li, and Y.-Y. Qu, “Unifying visual saliency with HOG feature learning for traffic sign detection,” in Proc. IEEE Intell. Veh. Symp., Jun. 2009, pp. 24–29

[3] I. Creusen, R. Wijnhoven, E. Herbschleb, and P. de With, “Color exploitation in hog-based traffic sign detection,” in Proc. 17th IEEE ICIP, Sep. 2010, pp. 2669–2672.

[4] Chaiyakhan, Kedkarn & Hirunyawanakul, Anusara & Chanklan, Ratiporn & Kerdprasop, Kittisak & Kerdprasop, Nittaya. (2015). Traffic Sign Classification using Support Vector Machine and Image Segmentation. 52-58. 10.12792/iciae2015.013.

[5] http://answers.opencv.org/question/135094/hog-optimal-training-images/

[6] R. Lienhart and J. Maydt. “An extended set of Haar-like features for rapid object detection”. In Proc. of the IEEE Conf. on Image Processing (ICIP '02), pages 155 - 162, USA, 2002.

[7].G. Loy and A. Zelinsky. “Fast Radial Symmetry for Detecting Points of Interest”. IEEE Transactions on pattern analysis and machine intelligence. vol. 25, nr. 8, August 2003.

[8].D.M. Gavrila, V.Philomin. “Real-Time object detection for “Smart” Vehicles”. In Proc. of IEEE International Conference on computer vision, pages 87-93, Kerkyra, 1999

[9].Y. Freund. “An adaptive version of the boost by majority algorithm”. Proceedings of the 12th Annual Conference on Computational Learning Theory, 1999

[10]. Y. Freund and R.E. Schapire. “A Short Introduction to Boosting”. Journal of Japanese Society for Artificial Intelligence 14(5), pages 771-780, September, 1999 [12]. R.E. Shapire, Y. Freund, P. Bartlett and Wee Sun Lee. “Boosting the margin: A new explanation for the effectiveness of voting methods.”