Learning to Detect Traffic Signs: Comparative Evaluation of Synthetic and Real World Data Sets

Prof. Abhinav V. Deshpande, Assistant Professor, Department of Electronics & Telecommunication Engineering, Prof. Ram Meghe Institute of Technology & Research, Badnera, Amravati-444701, India, 9370270054.

Abstract:
This study compares the performance of sign detection based on synthetic training data to the performance of detection which is based on real world training images. Viola-Jones detectors are created for four different traffic signs with both synthetic and real data, and varying numbers of training samples. The detectors are tested and compared. The result is that while others have successfully used synthetic training data in a classification context, it does not seem to be a good solution for detection. Even when the synthetic data covers a large part of the parameter space, it still performs significantly worse than the real world data.

Keywords: Traffic Sign Recognition, Machine Learning, Support Vector Machine, Distance to Bounding Box features

1. Introduction:

With the emergence of more advanced sensors embedded in cars, the field of Traffic Sign Recognition (TSR) has seen increasing interest over the last decade. The TSR systems can be used in a number of scenarios, ranging from Driver Assistance Systems (DAS’s)-as described in [14] to fully autonomous cars. Many sign detection systems rely on large amounts of training data to work. Over the past two years, a few traffic sign datasets has shown up. The GTSRB dataset [12][13], the Swedish Traffic Signs Dataset [6], and the KUL Belgium Traffic Signs Datasets. A common factor among these datasets is that they contain European Signs conforming to the Vienna Convention. Since the signs differ from region to region and in many cases from country to country, an interesting proposition is to use synthetically generated training data, saving a lot of time and effort in gathering the data. The synthetically generated training data has not yet been widely used in the field of TSR, but is worth a problem of doing research since very few datasets from outside of Europe exists. A recent survey [9] shows that the research work which is done on the problem of detection and recognition of traffic signs which are outside the countries conforming to the Vienna Convention on traffic signs is lacking in general. This research paper investigates if using data for the detection of traffic signs is feasible. The role and importance of high quality, representative datasets in the development of TSR systems cannot be overemphasized. The collection of such datasets is an expensive task (in time as well as effort). The issues in training, annotations in the real world and semi-supervised learning for object recognition is treated in [11]. Since the signs have a well defined appearance, the idea of using
The use of synthetic training in sign detection is not yet widespread, thereby prompting this research paper. The research paper is focused closely on the generation of synthetic training data for detection purposes. It is also the first of its kind which is dealing with the US signs. In [3][4], the generation of synthetic data specifically for classification is investigated. In [10], some of the aspects of detecting non-US signs with synthetic data is discussed. The detection task is somewhat harder than the classification due to the lack of knowledge about whether a sign is present, where it is and what size it has. The following section briefly covers the general workings of TSR systems followed by a section on how we generate the synthetic training data. Towards the end of the paper, the performance of synthetic training data is compared to the performance of real world training data when used to train a simple Adaboost cascade with Haar-like features [15].

2. TSR: General Approaches

The overviews of TSR can be found in [2][8][9]. The TSR can be split into two main stages: Sign Detection and Sign Classification as seen in Figure 1. Not all detection approaches require training as such, since they are using a theoretical model of the sign, which is based on e.g. the shape. With that said, many papers present Machine Learning (ML) based approaches. In [1], an Adaboost cascade which is similar to the one used in this paper was used, albeit on specific color channels. In [5], the image is segmented with a HIS threshold and then classifies the resulting blobs by using a linear support vector machine (SVM) on Distance to Bounding Box (DtB) features.

The DtB features are measurements of the distance between the edge of the blob and its rectangular bounding box.

Challenges in Traffic Sign Recognition:

The problems which are generally faced by a Traffic Sign Recognition system are given as below:

1. Lighting conditions:

Lighting conditions cannot be the same every time, it is changeable and not controllable. Lighting is different according to the time of the day, season, cloudiness and other weather conditions etc [2].

2. The presence of other objects:

Sometimes objects other than the traffic sign boards surround the traffic signs. This produces partial occlusions, shadows etc. [2].

Different approaches were adopted in the past for detecting road signs. In the older studies, eg [1,2], as well as in many recent ones, eg.[3,4], it was common to employ a heuristic encompassing prior knowledge about the traffic signs in order to (1) define how to pre-segment the scene to find the interest regions and (2) define the acceptable geometrical relationships between the sign parts with respect to shape and color. The major deficiency of these methods was a lack of solid theoretical foundations and high parameterization. In other studies, eg. [5] neural networks were used to model the above mentioned shape and appearance properties of traffic signs. A more convincing parameter-free method for
detecting road signs was proposed by Bahlmann et al [6].

Figure 1 Flowchart for ML-Based TSR Systems
3. Synthetic Training Data for Detection:

The question this research paper tries to answer is: Can we substitute real world training data with synthetic in ML based sign detection systems? The idea is to generate synthetic training images from a drawn template.

The goal is to emulate how signs of the given type might look on pictures from the real world. In order to do this, several transformations are made randomly to the templates.

A Hue variation emulates faded signs and color casts due to lighting of the natural scene which is done by adding to/subtracting from the hue parameter in the HSV color space.

The Lighting variation emulates shadows and variations in exposure which is done by adding to/subtracting from the value parameter in the HSV color space.

The Rotations around the X-, Y- and Z-axis with the origin in the center of the template. It emulates the signs which are captured from different perspectives.

The backgrounds which are taken from a real image are added to the template. This emulates the various backgrounds a sign might have in real life.

The Gaussian Blur is used to emulate an unfocussed camera. It should be noted that the Gaussian Blur does not really emulate the boken which is produced by an unfocussed lens, but emulating boken properly is not a straightforward task, and it would likely not give any notable detection benefit.

The Gaussian noise is used to emulate the sensor noise.

The Occlusions are added in the form of tree branches which are growing in front of some signs.

Each transformation should be applied with a random parameter within some realistic boundaries.

In order to evaluate whether the synthetic datasets cover the same variance in appearance as the real world data, we compare the distributions in intensity and blur values among the training sets. In figure 4 (a), a plot of the mean of the intensities in the training images is shown. Each point in the plot is a single image. The data for the detectors of two different signs is shown. In a few sets, the intensity span does not match, but the large 5000 image stop sign set is similar to the real world data. Another parameter which is shown in Figure 4 (b) is known as Blur which is calculated as:

\[ B = \frac{1}{n} \sum_{i=0}^{n} e_i \quad \text{Eqn. (1)} \]

Where B is the blur value, n is the number of vertical edges in an image and \( e_i \) is the edge width of a specific edge pixel which is given as the distance between the pixels with the local maximum and minimum intensity around the edge pixel. The measure is described briefly in [7]. This shows that the blur variance is covered well by the synthetic data.
Figure 2 Pedestrian Crossing Sign

Figure 3 Signal Ahead Sign

Figure 4 Stop Sign

Figure 5 Speed Limit Sign
Figure 6 Synthetic Training Images Template

Figure 7 Real World Training Images

Figure 8 Synthetic Training Images Template

Figure 9 Real World Training Images
4. Comparative Evaluation:

In order to compare the synthetic training data to training data which is obtained from a real footage, the simple Viola-Jones based detector [15] was trained for the four sign types which are illustrated in Figure 2. The choice of the detection algorithm is not crucial, as the purpose of this research paper is not to find a perfect sign detector, but rather look at the relative differences between the detectors trained with the synthetic and real world images. It was trained with an image size of 20 x 20 pixels in all cases, except for the rectangular speed limit sign, trained with 18 x 24 pixels. The detectors that are created with various numbers of training images was tested on a set of real world images which are collected from the cars in conjunction with this laboratory’s research work.

With all signs, the real world data performs significantly better than the synthetic data. Providing more training data in the synthetic case does help, but even a large increase (more than a doubling) of the training data does not make the synthetic data perform comparably to the real world data. All detectors were trained with a target of 20 stages, but some of them were terminated earlier due to a sufficiently good fit to the training data and others were lowered to give better detection performance at the cost of more false positives. It is indeed possible for the synthetic detector to find more true signs, but at a huge cost in false positives, and still not as good as the real world detector. Even in the cases like the stop sign detector with 5000/10000 training images where the synthetic data spans nearly the
same space as the real world detector, the synthetic detector fails to achieve a detection rate anywhere near the real world data.

5. Acknowledgements:

This research work was undertaken as a part of Technical Education Quality Improvement Program (TEQIP-2) which was sponsored by Maharashtra Human Resource Development (MHRD) in collaboration with Prof. Ram Meghe Institute of Technology & Research, Badnera, Amravati-444701 in order to promote and facilitate the research work which is carried out by the research scholars who are working in the field of Traffic Sign Detection and Recognition in order to curb the number of accidents which are caused due to the insufficient lighting conditions, poor visibility due to adverse weather conditions like snow, rain, dirt, dust, fog, dew etc. which tend to mislead the driver and the chances of accidents are more if proper care is not taken while driving the vehicle. In this research paper, a efficient algorithm to detect, recognize and classify a particular road traffic sign from the other traffic signs which fall under the same category with the help of color and shape features is presented and the merits and demerits of the algorithm are discussed. I am very much thankful to all my staff and faculty members as well as the Head of the Department, Department of Electronics & Telecommunication Engineering, PRMITR, Badnera, Amravati-444701 for their kind co-operation and valuable guidance and support.

6. References:


Books Referred:


Author’s Profile:

My name is Prof. Abhinav V. Deshpande. I have done S. S. C. in the stream of Science from Somalwar Nikalas High School, Nagpur-440022 in the year 2000 with an aggregate of 73%. I have done H. S. S. C. from Jupiter Science Junior College, Nagpur-440022 in the year 2002 with an aggregate of 54% in the branch of Electronics. I have done B.E. in Electronics and Telecommunication Engineering from G.H. Raisoni College of Engineering, Nagpur-440022 in the year 2002 with an aggregate of 63%. I have done M.Tech. in Electronics Engineering from G.H. Raisoni College of Engineering, Nagpur-440022 in the year 2012 with a CGPA of 7.90 on a scale of 10.00. I am currently working as an Assistant Professor on Contract Basis in the Department of Electronics & Telecommunication Engineering at Prof.
Ram Meghe Institute of Technology & Research, Badnera, Amravati-444701. I have published 17 research papers in different and reputed International Journals and 1 research paper in International Conference. I have also published 1 book in Saarbrucken, Germany in the year 2012. I am having memberships of different professional organizations like ISTE, IE (I), IEEE, UACEE, IETE, IET. My areas of research include different topics in the field of Electronics Engineering such as Digital Image Processing, Digital Signal Processing, Embedded Systems, VLSI, Soft Computing and Applications, Intelligent Transportation Systems etc. I have also passed the Ph.D. Entrance Test (PET) of RTM Nagpur University in the year 2013 with a valid score of 55 marks out of 100 and in the year 2015 with a valid score of 55.75 marks out of 100. I have also passed the Ph.D. Entrance Test (PET) of Gondwana University, Gadchiroli in the year 2015 with a valid score of 51 marks out of 100. I have also published a book titled “A Novel Method for the study of Content-based Image Retrieval-A New Algorithm for the extraction of different features from an image by using Gabor Filters and Zernike Moments” which was published by Lambert Academic Publishing (LAP) Company in Saarbrucken, Germany in the year 2012. I have participated in 10 workshops and STTP’s which were organized by AICTE, IIT Kharagpur, PRMITR, Badnera, Amravati-444701, GHRCE, Nagpur-440022.