Machine Learning Engineer Nanodegree

Capstone Proposal

Tanwani Rahul
Jan 1st, 2018

Proposal

Traffic Signs Classification for Automonous Vehicles.

Domain Background

Traffic sign detection and classification is one of the preliminary requirement for autonomous driving to be useful in the real world context. Prior work in this field has mainly focussed on building the solution using image processing techniques, and hand coded features followed by a classification model[2][3] with few exceptions[4]. The relevant work in the field has been well summarized in [1]. With the recent advancements in machine learning, specifically in the areas of computer vision using deep learning techniques, we aim to build the robust traffic sign detection system that can accurately identify the traffic sign in real time without requiring the prior knowledge of traffic sign locations.

Problem Statement

Traffic signs recognition has direct real world applications in the autonomous driving and in driver assistance systems. Traffic signs are intended to be visible for drivers and have very little variability in appearance. Natual variations and calamities such as viewpoint variations, lightning conditions, sun glare, occlusions, physical damage and color fading etc contribute additional complexity and make the problem even more challenging. In this project, we propose to build the system for detecting the correct traffic sign given the complex natural images. Given we have more than 2 images, we will build the model for multi-class classification.

Datasets and Inputs

We plan to use GTSRB - German Traffic Sign Recognition Benchmark, the traffic signs dataset hosted and maintained by INI benchmarks[5]. This dataset represent a multi-class classification challenge with more than 40 classes. The dataset has more than 50,000 images in total.

Dataset specifications:

- 43 traffic signs in total
- More than 50,000 images in total (39,209 training images, and 12630 testing images)
- Reliable ground-truth data due to semi-automatic annotation

Image format

- The images contain one traffic sign each
- Images contain a border of 10 % around the actual traffic sign (at least 5 pixels) to allow for edge-based approaches
- Images are stored in PPM format (Portable Pixmap, P6)
- Image sizes vary between 15x15 to 250x250 pixels
- Images are not necessarily squared = The actual traffic sign is not necessarily centered within the image.
- The bounding box of the traffic sign is part of the annotatinos.

Annotation format

Annotations are provided in CSV files. Fields are separated by ";" (semicolon). Annotations contain the following information:

- Filename: Filename of corresponding image
- Width: Width of the image
- Height: Height of the image
- ROI.x1: X-coordinate of top-left corner of traffic sign bounding box
- ROI.y1: Y-coordinate of top-left corner of traffic sign bounding box
- ROI.x2: X-coordinate of bottom-right corner of traffic sign bounding box
- ROI.y2: Y-coordinate of bottom-right corner of traffic sign bounding box
- The training data annotations will additionally contain ClassId: Assigned class label

Class Distribution:

To precisely define the overall solution approach and evaluation metrics, it is important to understand the underlying distribution for the target variable. Following image shows the distribution across 43 traffic signs represented in the dataset:

- Normal Scale:
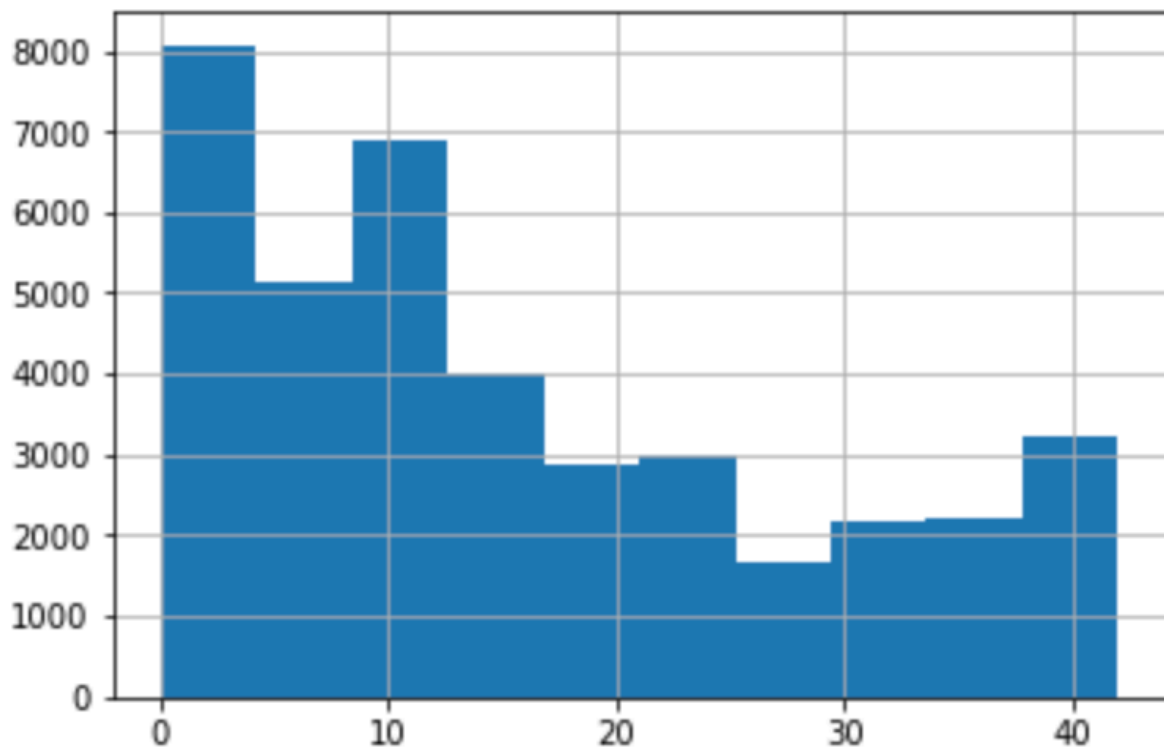
- Log Scale:
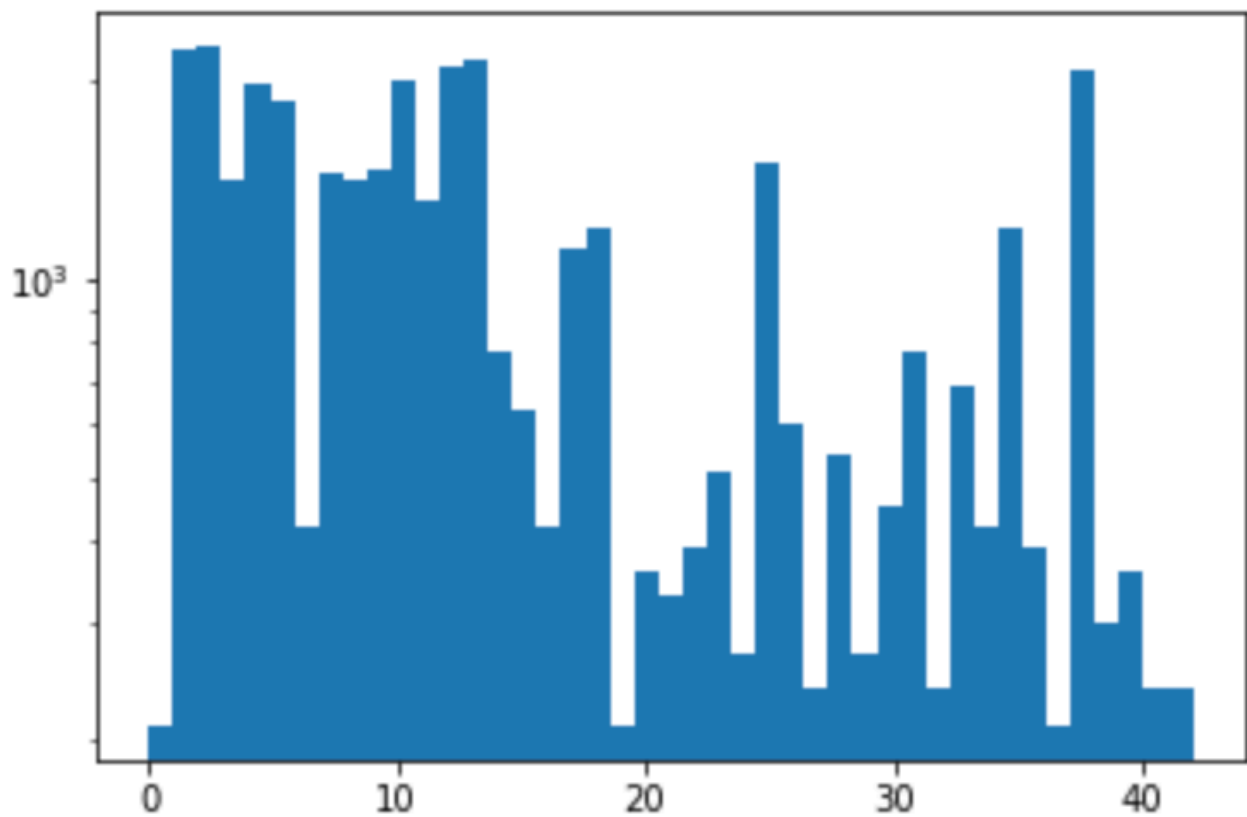
We can clearly see the distribution is uneven / imbalanced. This suggests we should pay attention on how to effitiently use the data for training and validation set, need for balancing and the effective measure for model

evaluation.

Solution Statement

Most of the prior algorithms and solutions for this problem have employed traditional image processing techniques, where the relavent features are hand-coded, and global signs are first detected using various heuristics and color thresholding techniques, the detected windows are then feed into the classifier for the final classification [2][3][6][7]. We would like to follow a learning based approach to effitiently learn the relevant features and detect the appropriate traffic signs. We plan to use the convolutional neural network based model to detect and classify the traffic signs. To make sure the model is robust, we will apply various data/image augmentation techniques before learning the model.

Benchmark Model

The benchmarks for this problem are maintained and made available by INI benchmarks[5]. Peformance is measure for each class individually and also the overall peformance across all traffic signs. Current best score across all traffic signs stands to be 99.46%. Yann Lecun et. al. used the CNN based deep learning approach for the same problem, their architecture and performance is descriobe in [4]. To evaluate our model against the benchmarks, we can do the submission on INI benchmarks portal.

Sumission instructions

- The results will be submitted as single CSV.
- It contains two columns and no header. The separator is ";"(semicolon).
- There is no quoting character for the filename.
- First columns is the image filename, second column is the assigned class id.

Evaluation Metrics

The algorithms are evaluated and ranked based on the correct classification rate (CCR). CCR could be defined as 1 - misclassification error rate. In addition to overall performance of the model, each traffic sign is evaluated separately and performance for each of the sign is measured. The samples are un-weighted which implies all the traffic signs are euqally important.

Like we have seen above, we have an imbalanced dataset, so CCR (accuracy) alone may not be the best way to evaluate the performance of the model. We should rather use the measure that takes into account the precision and recall both to evaluate. F1-Score would be a better evaluation metric to use. In addition to the single metric, it is importatnt to analyze the peformance for each of the class separately. Confusion matrix would also be able to understand how the model is performing for each of the traffic sign we have in the dataset.

Project Design

At a hgher level, there are few important elements for designing the overall solution. I would break this up in the following broad areas:

1. Data Preparation and Analysis
2. Data Augmentation
3. Model Building and Validation
4. Performance and Error Analysis

Data Preparation and Analysis:

Here, we focus on preparing the data into necessary data structures from their raw PPM format. Once we have the images and their corresponding class / label attached, we will look at the distribution among each traffic sign. If we notice the high error for a under represented class, balancing the data may be helpful.

Data Augmentation:

To build the generalized, robust model which is not baised towards the limitation of the dataset, it is very important to perform data augmentation. The techniques and impact of data augmentation on computer vision problems have been well discussed and summarized in [8][9]. In this project, we will explore various data augmentation techniques and measure their impact on the performance. Some of the techniques we will consider:

- Sample-wise standardization.
- Feature-wise standardization.
- ZCA whitening.
- Random rotation, shifts, shear and flips.

Model building and validation:

With the recent advancements in nueral networks based learning algorithms in general, computer vision algorithsms in specific, we aim to build the robus model to achieve the best results. For this probelm, we will consider convolutional neural neteworks (CCN) based architecture and algorithms to achive the best results. In addition, we will use the validation set for tuning the parameters and model validation.

Performance and error analysis:

We will use CCR measure to evaluate the performance on the validation and test set. The confusion matrix would gives us insights on any confusion among traffic signs that the model is struggling with. We will sample the images in error from the validation set to see if other data augmnentation techniques would be helpful in the

context.

References:

1. [Detection and Recognition of Road Traffic Signs - A Survey](#)
2. [Road and traffic sign detection and recognition](#)
3. [Traffic Road Sign Detection and Recognition for Automotive Vehicles](#)
4. [Traffic Sign Recognition with Multi-Scale Convolutional Networks](#)
5. [Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition](#)
6. [Automatic detection and classification of traffic signs](#)
7. [Traffic sign shape classification evaluation I: SVM using distance to borders](#)
8. [The effectiveness of data augmentation in image classification using deep learning](#)
9. [Smart Augmentation - Learning an Optimal Data Augmentation Strategy](#)