**Traffic Sign Recognition**

**Build a Traffic Sign Recognition Project**

The goals / steps of this project are the following:

\* Load the data set (see below for links to the project data set)

\* Explore, summarize and visualize the data set

\* Design, train and test a model architecture

\* Use the model to make predictions on new images

\* Analyze the softmax probabilities of the new images

\* Summarize the results with a written report

**Writeup / README**

**Data Set Summary & Exploration**

**1. Provide a basic summary of the data set and identify where in your code the summary was done. In the code, the analysis should be done using python, numpy and/or pandas methods rather than hardcoding results manually.**

The 2nd code cell of the IPython notebook contains the code for this step.

I used python and numpy library to calculate summary statistics of the **original** traffic signs data set:

Number of training examples = 34799

Number of valid examples = 4410

Number of testing examples = 12630

Image data shape = (32, 32, 3)

Number of classes = 43

**2. Include an exploratory visualization of the dataset and identify where the code is in your code file.**

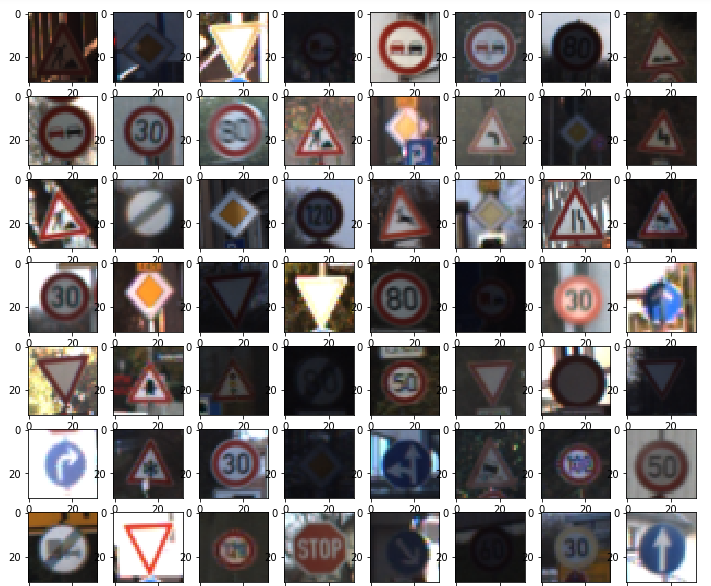
The 3rd code cell of the IPython notebook contains the code for this step.

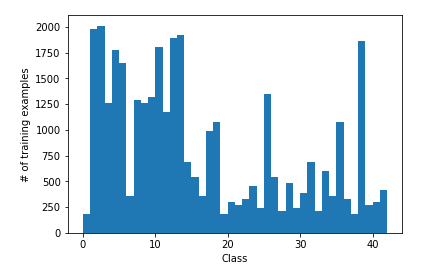
We can observer than:

1º There are many dark images.

2º We have an imbalanced dataset. There are classes with many samples and classes with little samples.

Here is an exploratory visualization of the data set.





**Design and Test a Model Architecture**

**1. Describe how, and identify where in your code, you preprocessed the image data. What techniques were chosen and why did you choose these techniques? Consider including images showing the output of each preprocessing technique. Pre-processing refers to techniques such as converting to grayscale, normalization, etc.**

I’ve tested several models with different preprocess and hyperparameters.

1st Model:

Pipeline:

Initialization of weights:

I initialize the weights with truncated\_normal function with very low values:

mean=0 and std=0.1

Preprocessing:

Normalization of RGB images between values [0 - 1]

Model:

+ CNN1 layer [32 kernels] >> RELU activation >> MaxPooling >> Dropout

+ CNN2 layer [64 kernels] >> RELU activation >> MaxPooling >> Dropout

+ FC1 layer >> RELU activation

+ FC2 layer >> RELU activation >> Dropout

Hyperparameters:

Keep\_prob = 0.5

Results: ~ 92 - 93% accuracy

2nd Model:

Pipeline:

Initialization of weights:

I get some tips of the paper winner in IJCNN 2011 competition:

Multi-Column Deep Neural Network for Traffic Sign Classification

<http://people.idsia.ch/~juergen/nn2012traffic.pdf>

I initialize the weights with truncated\_normal function with very low values:

mean=0 and std=0.05

Preprocessing:

Normalization of images between values [0 - 1]

Model:

+ CNN1 layer [100 kernels] >> RELU activation >> MaxPooling >> Dropout

+ CNN2 layer [150 kernels] >> RELU activation >> MaxPooling >> Dropout

+ FC1 layer >> RELU activation

+ FC2 layer >> RELU activation >> Dropout

Hyperparameters:

Keep\_prob = 0.5

Results: ~ 93 - 95% accuracy after 40 epoch.

3rd model:

Pipeline:

Initialization of weights:

I initialize the weights with truncated\_normal function with very low values:

mean=0 and std=0.05

Preprocessing:

I convert images dataset from RGB to HSV.

I use this function from TF: tf.image.rgb\_to\_hsv(images, name=None).

This function returns values between [0,1].

I follow the same reasoning than I made in Project 1: “Finding Lane Lines on the Road ”.

In Wikipedia article about HSV, we can read: “HSL and HSV models are based more upon how colors are organized and conceptualized in ***human vision*** in terms of other color-making attributes ”. In the same way, if we visualize the hidden layers from a CNN network, we can observe similarities with ***human vision.***

Refs:

http://docs.opencv.org/3.2.0/df/d9d/tutorial\_py\_colorspaces.html https://en.wikipedia.org/wiki/HSL\_and\_HSV https://en.wikipedia.org/wiki/Convolutional\_neural\_network

Model:

+ CNN1 layer [100 kernels] >> RELU activation >> MaxPooling >> Dropout

+ CNN2 layer [150 kernels] >> RELU activation >> MaxPooling >> Dropout

+ FC1 layer >> RELU activation

+ FC2 layer >> RELU activation >> Dropout

Hyperparameters:

Keep\_prob = 0.5

Results: ~ 92 - 93% accuracy

**Final System:**

**Preprocessing Techniques:**

I observed than:

1º There are many dark images.

2º We have an imbalanced dataset. There are classes with many samples and classes with little samples.

To get more accuracy I need:

A. Improve the images.

B. Increase the number of images.

C Balance the dataset

**A. Improve the images.**

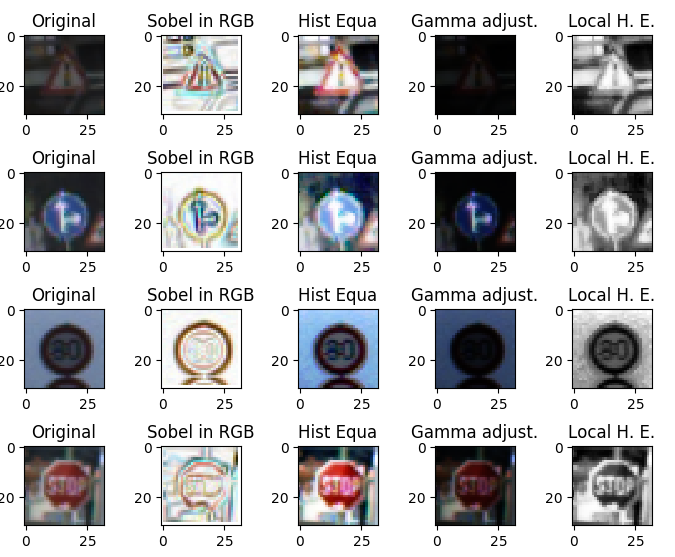
To get better images, I test several methods to preprocessing the images.

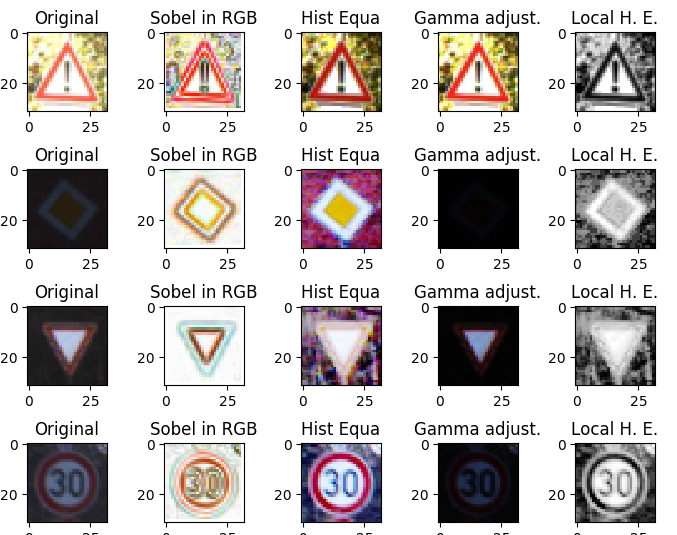
I try to follow the methods of the paper “Multi-Column Deep Neural Network for Traffic Sign Classification”.

I use scikit-image library.

The 5th code cell of the IPython notebook contains the code for this step.

Samples of preprocessed Images:





I’ve applied these filters:

+ Sobel to RGB images

+ Histogram Equalization[¶](http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_equalize.html#histogram-equalization)

+ Gamma and log contrast adjustment[¶](http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_log_gamma.html#gamma-and-log-contrast-adjustment)

+ Local Histogram Equalization[¶](http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_local_equalize.html#local-histogram-equalization)

Refs:

<http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_adapt_rgb.html#sphx-glr-auto-examples-color-exposure-plot-adapt-rgb-py>

<http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_equalize.html#sphx-glr-auto-examples-color-exposure-plot-equalize-py>

<http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_log_gamma.html#sphx-glr-auto-examples-color-exposure-plot-log-gamma-py>

<http://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_local_equalize.html#sphx-glr-auto-examples-color-exposure-plot-local-equalize-py>

**B. Increase the number of images. Data Augmentation**

I apply data augmentation before training and during training.

For DA before training I reuse code from @navoshta.

For DA during training I use ImageDataGeneratorfrom Keras.

Ref:

#Data Augmentation

<http://navoshta.com/traffic-signs-classification/>

<https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html>

https://keras.io/preprocessing/image/

**C. Balance the dataset**

What can we do when we have imbalanced data?:

* Ignoring the problem.
* Undersampling the majority class.
* Oversampling the minority class.

I code functions to balance dataset in Xnd cell. Using Keras I can weight the classes during the training. But I don’t apply weights to classes during my training because I get bad results during my tests.

**Ref:**

##Balance of DATASET

#http://www.marcoaltini.com/blog/dealing-with-imbalanced-data-undersampling-oversampling-and-proper-cross-validation

#http://datascience.stackexchange.com/questions/13490/how-to-set-class-weights-for-imbalanced-classes-in-keras

#https://github.com/fchollet/keras/issues/1875

#https://github.com/scikit-learn/scikit-learn/blob/master/sklearn/utils/class\_weight.py

**Model:**

I use 3 CNN with 32, 64 and 128 filters.

Kernels size equal to 3x3

Max pooling 2x2

-------------------------------------------------------------------------------

model = Sequential()

#CNN1

model.add(Conv2D(32, kernel\_size=(3, 3), padding='valid', activation='relu',input\_shape=input\_shape))

model.add(MaxPooling2D(pool\_size=(2, 2)))

model.add(Dropout(do[0]))

#CNN2

model.add(Conv2D(64, (3, 3), padding='valid', activation='relu'))

model.add(MaxPooling2D(pool\_size=(2, 2)))

model.add(Dropout(do[1]))

#CNN3

model.add(Conv2D(128, (3, 3), padding='valid', activation='relu'))

model.add(MaxPooling2D(pool\_size=(2, 2)))

model.add(Dropout(do[2]))

#Flatten

model.add(Flatten())

model.add(Dense(1024, activation='relu'))

model.add(Dropout(do[3]))

#Output

model.add(Dense(num\_classes, activation='softmax'))

------------------------

Hyperparameters:

#Init Variables

batch\_size = 512

num\_classes = 43

epochs = 50

dropout = [1.,1.,1.,1.]

# input image dimensions

img\_rows, img\_cols = 32, 32

input\_shape = (img\_rows, img\_cols, 1)

**2. Describe how, and identify where in your code, you set up training, validation and testing data. How much data was in each set? Explain what techniques were used to split the data into these sets. (OPTIONAL: As described in the "Stand Out Suggestions" part of the rubric, if you generated additional data for training, describe why you decided to generate additional data, how you generated the data, identify where in your code, and provide example images of the additional data)**

Pipeline for Dataset transformations:

1. First load train, valid and test dataset form files. Code in cell number 1.
2. Data Augmentation.
   1. Before training: Code in cell number 6 & 7.
   2. During training: Code in cell number 17.
3. Preprocessing the data and save to files. Code in cell number 9 & 10.

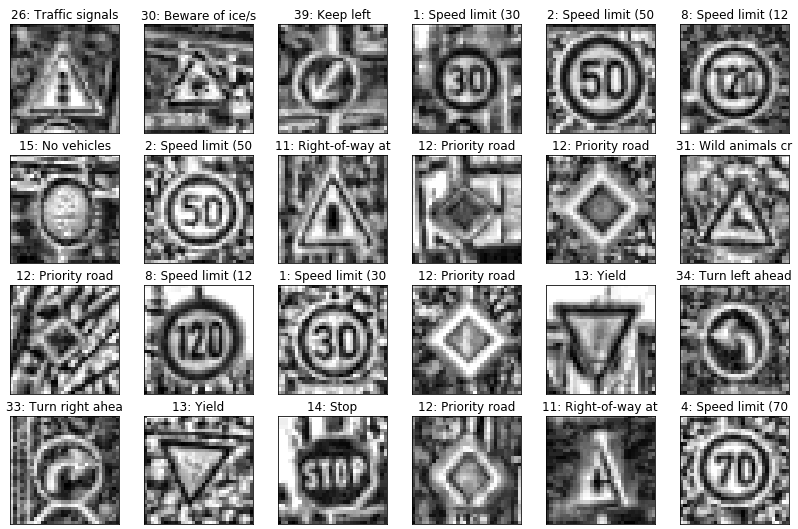
My final dataset:

Number of training examples = 59788

Number of valid examples = 7590

Number of testing examples = 21720

Here is an example of augmented images:



Data Augmentation

**3. Describe, and identify where in your code, what your final model architecture looks like including model type, layers, layer sizes, connectivity, etc.) Consider including a diagram and/or table describing the final model.**

The code for my final model is located in the cell number 19.

My final model consisted of the following layers:

|  |  |
| --- | --- |
| **Layer** | **Description** |
| Input | 32x32x1 Gray image |
| Convolution 3x3 | 32 filters, valid padding, RELU activation |
| Max pooling | 2x2 stride |
| Dropout | Value = float between 0 and 1 |
| Convolution 3x3 | 64 filters, valid padding, RELU activation |
| Max pooling | 2x2 stride |
| Dropout | Value = float between 0 and 1 |
| Convolution 3x3 | 128 filters, valid padding, RELU activation |
| Max pooling | 2x2 stride |
| Dropout |  |
| Flatten |  |
| Fully connected | output= 1024 |
| Softmax | 43 classes |

**4. Describe how, and identify where in your code, you trained your model. The discussion can include the type of optimizer, the batch size, number of epochs and any hyperparameters such as learning rate.**

The code for training the model is located in the 21th cell of the ipython notebook.

+ Optimizer: Adam. With values: lr=0.001, beta\_1=0.9, beta\_2=0.999, epsilon=1e-08, decay=0.0).

+ Batch Size: 512

+ Epochs: 50

I trained the model in several steps:

Step 1:

I generate random training data during training process.

To do this process, I use Keras class: **ImageDataGenerator.**

I set dropout values to 1. No dropout.

I train the model during 50 epochs and I get an accuracy of 0.98419.

I save the model

Step 2:

I load the model from step 1.

I generate random training&valid data during training process.

I set dropout values to 1. No dropout.

I train the model during 50 epochs and I get an accuracy of 0.xxxx.

I save the model

Step 3:

Repeat step 2 with learning rate = 0.0001

I train the model during 50 epochs and I get an accuracy of 0.xxxx.

I save the model

**5. Describe the approach taken for finding a solution. Include in the discussion the results on the training, validation and test sets and where in the code these were calculated. Your approach may have been an iterative process, in which case, outline the steps you took to get to the final solution and why you chose those steps. Perhaps your solution involved an already well known implementation or architecture. In this case, discuss why you think the architecture is suitable for the current problem.**

The code for calculating the accuracy of the model is located in the 21th cell of the Ipython notebook.

My final model results were:

\* training set accuracy of ?

\* validation set accuracy of ?

\* test set accuracy of ?

*If an iterative approach was chosen:*

*\* What was the first architecture that was tried and why was it chosen?*

My first architecture was the model proposed by the team that won the final phase of the German traffic sign recognition benchmark in the paper: “Multi-column deep neural network for traffic sign classification”.

\* What were some problems with the initial architecture?

No problem. But it is not simple. It has many filters.

I need many hours to train the model.

I don’t like work with this model.

\* How was the architecture adjusted and why was it adjusted?

First apply data augmentation over Lenet 5 to get a fast feedback.

Second, when DA is working ok, I change Lenet5 for a more deep architecture. I follow the recommendations of 2 papers and several blogs:

+ “Systematic evaluation of CNN advances on the ImageNet”

+ “Traffic Sign Recognition with Multi-Scale Convolutional Networks ”

+ “[Building powerful image classification models using very little data](https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html)”

This will be my final model.

I train this model in several incremental steps.

\* Which parameters were tuned?

I tuned: dropout, learning rate and validation data.

\* How were they adjusted and why?

I down learning rate a decade in the last step of training steps. I need to get a fine optimization of my model.

I use random generator in the training and validation data to get more “variability” in my data.

I don’t apply dropout. During my experiments I get very bad results.

\* What are some of the important design choices and why were they chosen?

**The most import design choice is use Keras:**

+ To define the model.

+ To generate random data .

+ To select the best hyperparameters by default.

Tensorflow is an unnecessarily complicate. Keras like scikit-learn is simple and powerful.

For example, why might a convolution layer work well with this problem? How might a dropout layer help with creating a successful model?

\* How does the final model's accuracy on the training, validation and test set provide evidence that the model is working well?

**Test a Model on New Images**

**1. Choose five German traffic signs found on the web and provide them in the report. For each image, discuss what quality or qualities might be difficult to classify.**

Here are five German traffic signs that I found on the web:

The first image might be difficult to classify because ...

**2. Discuss the model's predictions on these new traffic signs and compare the results to predicting on the test set. Identify where in your code predictions were made. At a minimum, discuss what the predictions were, the accuracy on these new predictions, and compare the accuracy to the accuracy on the test set (OPTIONAL: Discuss the results in more detail as described in the "Stand Out Suggestions" part of the rubric).**

The code for making predictions on my final model is located in the tenth cell of the Ipython notebook.

Here are the results of the prediction:

|  |  |
| --- | --- |
| **Image** | **Prediction** |
| Stop Sign | Stop sign |
| U-turn | U-turn |
| Yield | Yield |
| 100 km/h | Bumpy Road |
| Slippery Road | Slippery Road |

The model was able to correctly guess 4 of the 5 traffic signs, which gives an accuracy of 80%. This compares favorably to the accuracy on the test set of ...

**3. Describe how certain the model is when predicting on each of the five new images by looking at the softmax probabilities for each prediction and identify where in your code softmax probabilities were outputted. Provide the top 5 softmax probabilities for each image along with the sign type of each probability. (OPTIONAL: as described in the "Stand Out Suggestions" part of the rubric, visualizations can also be provided such as bar charts)**

The code for making predictions on my final model is located in the 11th cell of the Ipython notebook.

For the first image, the model is relatively sure that this is a stop sign (probability of 0.6), and the image does contain a stop sign. The top five soft max probabilities were

|  |  |
| --- | --- |
| **Probability** | **Prediction** |
| .60 | Stop sign |
| .20 | U-turn |
| .05 | Yield |
| .04 | Bumpy Road |
| .01 | Slippery Road |

For the second image ...