

Traffic Signs Classification

Submitted by:

Narendran Suresh

**ACKNOWLEDGMENT**

I would like to take the opportunity to thank the organisation ‘FlipRobo Technologies”, “Datatrained “and my mentor Ms. Astha Mishra for their immense support and guidance for making this feat achievable. I would also like to thank my dad Mr. Suresh Kumar S who is a statistics subject expert for guiding me along the right path and techniques to be used.

FlipRobo Technologies who are specialised in making ML / AI models provided me the Data set as a part of my internship.

References were made to several articles among Medium, KdNuggets, towardsdatascience, realpython, machinelearningmastery, python and sklearn documentation for the successful completion of the project.

**INTRODUCTION**

* Business Problem Framing

To classify the images (road signs) into different category road signs, to be able to ultimately use the techniques for autonomous cars.

* Conceptual Background of the Domain Problem

For the concept of autonomous cars like tesla to be a reality, it is required to create a deep learning model that can predict the different objects on the road. More importantly, it should also be able to classify road signs so as to abide by the local road rules.

* Review of Literature

Extensive reseach was done on CNN and Residual Network in order to find the solution. Web scraping domain knowledge and research was done to enhance the techniques used for data retrieval and image format conversion.

* Motivation for the Problem Undertaken

A lot of data is available in the form of images / videos and the capability of machines to predict / classify objects based on the image properties is a domain where deep learning and neural networks can play a major role in saving time and efforts for humans.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

The use of Convolutional neural networks was critical for arriving at an accurate solution for the problem in hand.

Image conversion to arrays, label encoding and the normalization techniques were critical for arriving at the solution.

* Data Sources and their formats

Data was provided by Fliprobo. Data was provided in the form of photos and a few csv files which characterised the images and the categories for the dataset.

* Data Preprocessing Done

Images were loaded to the notebook and the data was then converted to numpy arrays and was saved for future retrieval.

Data was split in the ratio of 80:20 against training and validation.

* Data Inputs- Logic- Output Relationships

The data was pretty trivial and the classification prediction task was attained by the use of Convolutional neural networks.

State the set of assumptions (if any) related to the problem under consideration

N/A

* Hardware and Software Requirements and Tools Used
* Hardware:
* Inter Core (i7) – 5500U, clock speed at 2.40GHz
* RAM – 12.0 GB
* Software:
* Jupyter Notebook (Anaconda 3) – Python 3.7.6
* Microsoft Excel

Libraries & Packages used – Pandas, numpy, sklearn, matplotlib, seaborn, sklearn, scipy, imblearn, tensorflow, keras, Image Data Generator, Maxpool2D

**Model/s Development and Evaluation**

* Identification of possible problem-solving approaches (methods)

The dataset has been provided in training and test folders with train containing subfolders as per the category. The test data was then to be validated and classified.

* Testing of Identified Approaches (Algorithms)

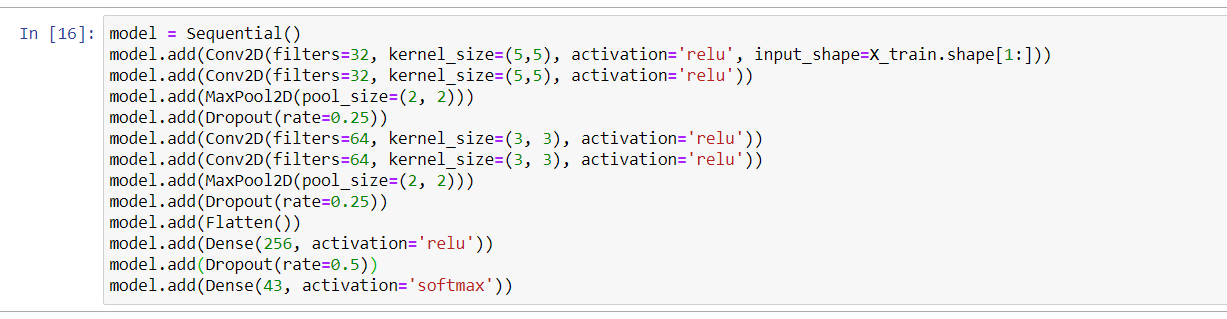
**CNN Optimizers:**

Adam

RMSProp

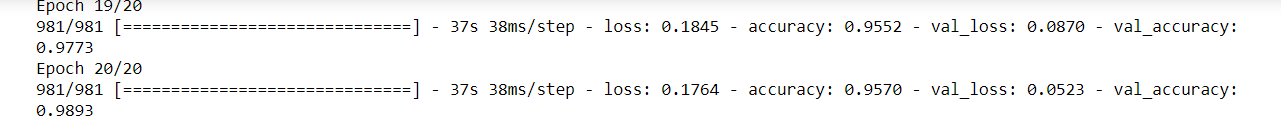
**ResNet50**

Final model was set as CNN **with parameters as below:**



Run and Evaluate selected models

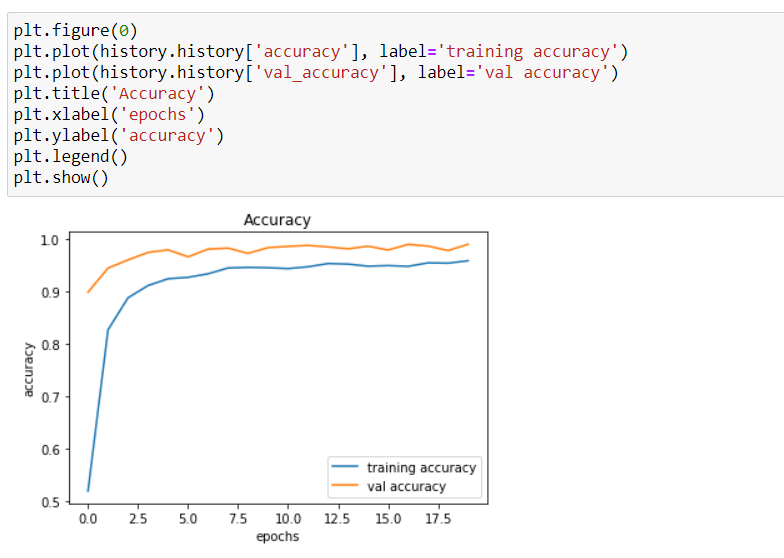
CNN was used to arrive at the solution.

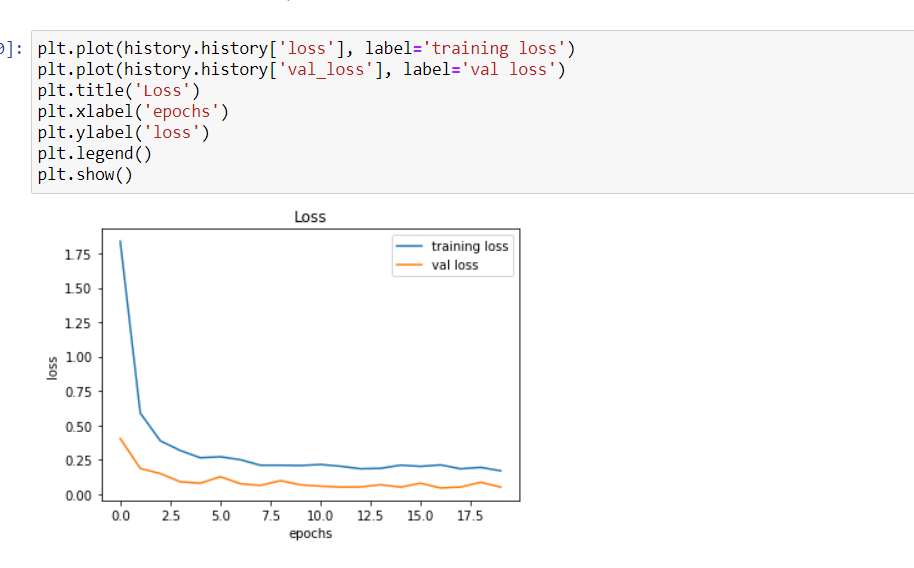


* Key Metrics for success in solving problem under consideration

Accuracy was the key metric for the solution to the problem. Visualizations

The accuracy and loss functions for the train and test sets were plotted and were found to be impressive.





* Interpretation of the Results

The model does a very good job in classifying the signs and can be used for predictions. Addition of more categories would mean that the model can be used for a wide range of applications.

**CONCLUSION**

* Key Findings and Conclusions of the Study

The training and testing images given in different folder structure was an initial challenge but was addressed later.

* Learning Outcomes of the Study in respect of Data Science

CNN, layers and folder structure formatting expertise has been improved.

* Limitations of this work and Scope for Future Work

The data for some categories were limited and hence model may not perform as well if more data is added.

Addition of more data would mean the number of batches will have to be increased and hence the computational time might increase drastically for a decent number of epochs.