COMP 4613 X2 – MACHINE LEARNING

PROJECT PROPOSAL

TRAFFIC SIGNS RECOGNITION

DEEPESH BELANI – 100145274

12th March 2021

Project Proposal – Traffic Signs Recognition

Index

Introduction

Problem Statement

The application will implement a Traffic Signs Recognition that will examine an image, extract its features, and then use it to identify the traffic sign that is present in the image. To get a good accuracy, the given set of images are used to train and test the model so that it recognizes the traffic sign correctly and with good accuracy.

Expected Success Criteria

CNNs are used for image-based Machine Learning Problems. We will be using CNN to examine and extract the features of the image.

Most problems of this type have proved to work better with CNNs. Expected accuracy is around 90%.

Approach

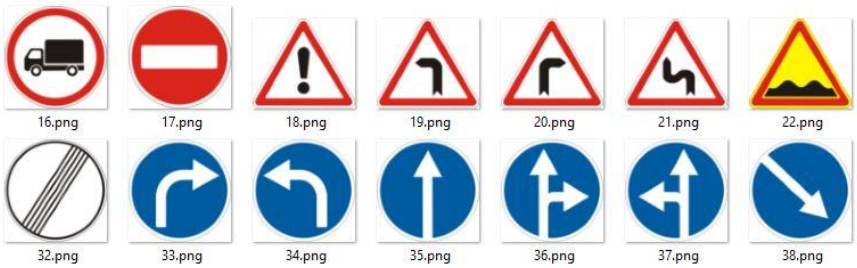
Dataset Description

The dataset that is going to be used is the Traffic Signs Dataset.

It contains more than 50,000 images of different traffic signs. It is further classified into 43 different classes.

All images are of different sizes. All images are in color (RGB).

Sample images from the Meta folder



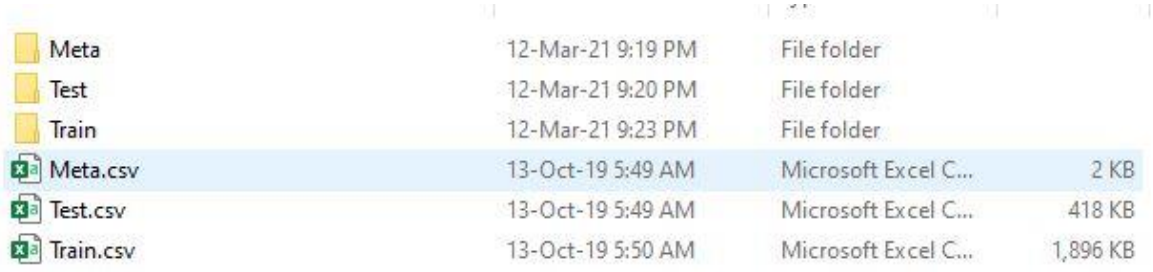
Dataset Preparation

The data is available on Kaggle. The file comes in a .zip format. The size of the dataset is around 300 mb.

https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign

After unzipping, the file is found to contain 3 folders and 3 csv files. The meta folder contains a few basic images. The test folder contains various traffic signs which will be used in our testing. The train folder contains many subfolders within it. Each sub folder contains various images which will be used to train our model.

Image of the Dataset Folder



Sample image of the Test folder Data (Images)

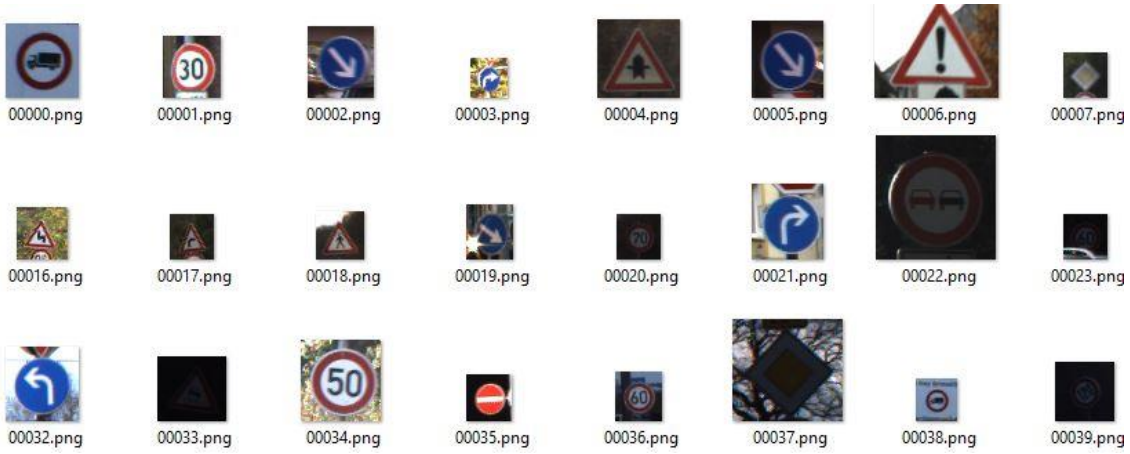
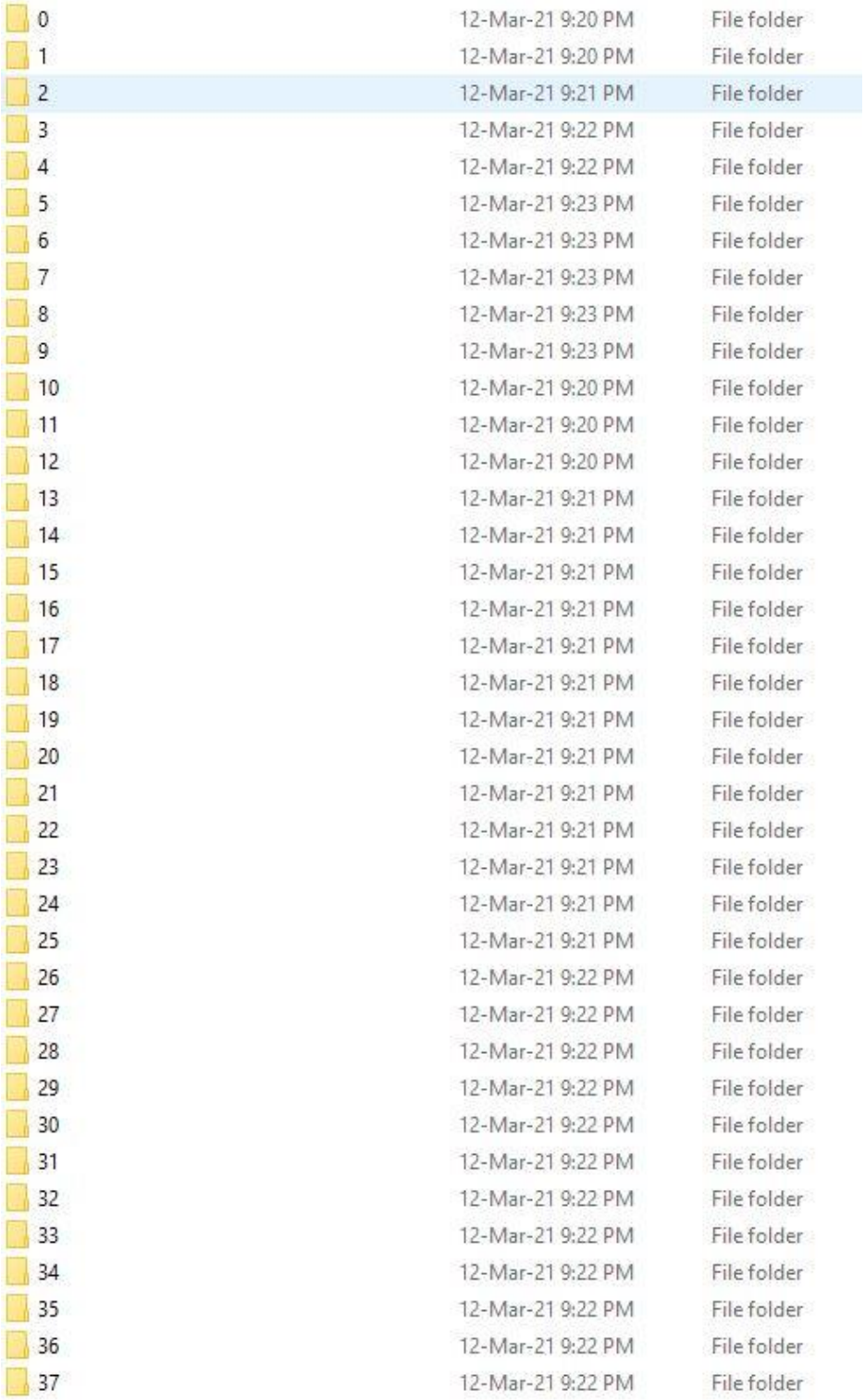
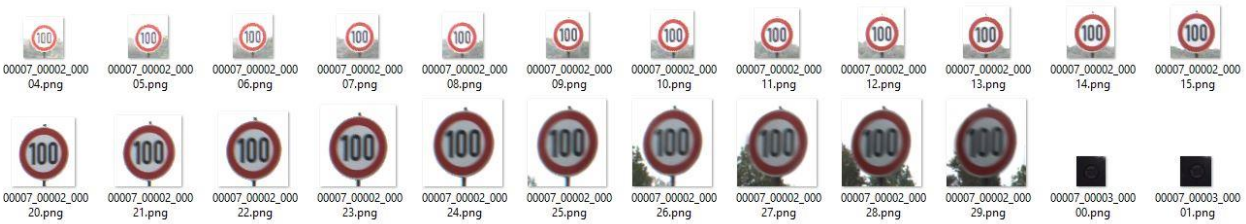


Image for the Sub folders Present in the Train Folder



Sample Image of the Data present in Sub Folder No.7



This data is already sorted out. So, the input for the problem can be taken either from the local directories or by saving it in the drive and then mounting the drive in Google Collab.

Machine Learning Methods to be Used

CNN

CNN is basically used for image classifications and identifying if an image is a bird, a plane, or a car etc. It scans images from left to right and top to bottom to pull out important features from the image and combines the features to classify images. It can handle the images that have been translated, rotated scaled and changes in perspective.

We will be using CNNs to classify the images into their respective categories. We will build our CNN model and then use it for testing and training.

Model Evaluations

Model Predictions

The model prediction will be done based on the test dataset. The models for a problem like this can be evaluated by comparing accuracies and tuning the parameters of the model to attain the accuracy that we expect.

The model can be evaluated using other visual metrics which allow us to determine how well the model is working based on the given architecture.

Outline of Project Report

1. Abstract
2. Introduction
    - 2.1 Problem Definition
    - 2.2 Expected Success Criteria
    - 2.3 Background
3. Overview of Approach
    - 3.1 Approach 1
    - 3.2 Method of Evaluations
4. Data Description
    - 4.1 Data Consolidation
    - 4.2 Data Preparation
    - 4.3 Training and Testing
5. Machine Learning Algorithms
    - 5.1 Used Algorithms
    - 5.2 Exploring Alternate Algorithms
6. Results
    - 6.1 Graphs
    - 6.2 Tables
    - 6.3 Comparison to Expected Results
7. Conclusion
8. References
9. Appendices