Stefan Merbele

Udacity Capstone Proposal

Machine Learning Engineer

GTSRB - German Traffic Sign Recognition Benchmark

Domain Background

In general, humans are very good at recognizing visual input. But everyone knows the problem. People just can't focus all the time, to get and process all visual input they get throughout the day. Especially, when there are many things to concentrate on at the same time. Usually, that's not a big problem. But there are areas, where it would be very helpful to get some help from technology.

One of that areas is recognizing traffic signs. Just think about you are in a city with a lot of traffic and pedestrians. You have to know your environment all the time and decide fast, what's the best action at the moment. It gets even harder when there are other people like children in your car, where you have to have an eye on.

That's why researchers in Germany wanted to tackle the problem and collected a huge database to find out if convolutional neural networks could be as good as humans at that task. If that's the case, it could help drivers around the world and take another little step to automated driving.

Link to the paper abstract:
https://www.sciencedirect.com/science/article/abs/pii/S0893608012000457?via%3Dihub

Problem Statement

The challenge that will be solved is a multi-class, single-image classification challenge held at the International Joint Conference on Neural Networks (IJCNN) 2011. That means, that there are images with their respective label that holds information about what kind of sign is seen at the picture. Based on that, a model will be trained and subsequently evaluated to find out, if it can classify new pictures in the right way.

One approach to master that challenge would be to create a convolutional neural network to classify these images with high accuracy.

Datasets and Inputs

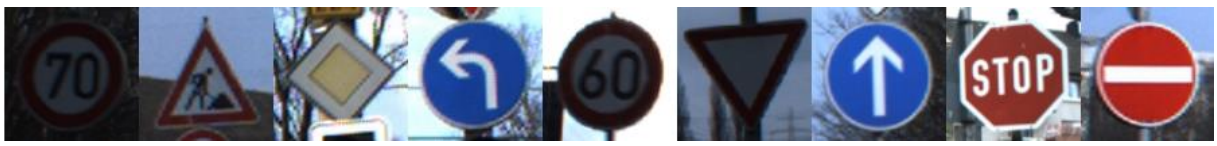

Figure 1: Example images of traffic signs

As you can see in Figure 1, there are different traffic signs. These pictures were made in everyday driving situations. So there are different weather conditions (rainy, cloudy, winter, summer). That adds a huge variance to every class of traffic sign. Furthermore, the images in the dataset are in different shapes and distances to the sign. That makes the problem more challenging.

There are 43 classes of traffic signs in more than 50.000 images in total. Details can be found here: [1].

Solution Statement

In this challenge, I will use a convolutional neural network created by my own with Tensorflow and Keras. I think that is an appropriate approach to show, that I understand data processing as well as the architecture of neural networks. Furthermore, it is a great chance, to compare my solutions with many others and will see, if the performance of my creation is as good as the competitors.

Benchmark Model

At the following links, there are some benchmark models. They all use CNN's as well. Their accuracy on test data reaches from 94 to 99%. Where neural networks trained from scratch and self-created architecture is between 94 and 97%. The ResNet-34 model, which is pre-trained on the ImageNet dataset reaches 99%.

https://www.kaggle.com/basel99/gtsrb-model-with-accuracy-97

https://www.kaggle.com/marinovik/recognizing-traffic-signals-with-keras-cnn

https://github.com/surmenok/GTSRB

Evaluation Metrics

As discussed at the "Benchmark Model" the success of my model will be evaluated through the Metric "accuracy". That means, that every predicted label will be compared to the ground truth label. If all estimations are right, there will be an accuracy of 1, if all estimations would be false, there will be an accuracy of 0. So the goal is, to get an accuracy of nearly 1.

Project Design

1. Upload the data and have a more detailed look. How do the pictures look like? What's their shape? Is the data balanced? Then transform the data into a Type that the machine learning model can process as well and efficiently as possible.
2. Convolutional neural networks have proven, that they perform very well on image data. That's why I will use CNN to solve that challenge. Furthermore, I want to develop an own network and train that from scratch with Keras and Tensorflow. I think, that this will help me to understand neural networks better, than using a prepared one from the Model Zoo, etc.
3. Have a look on the test data. How high is the accuracy? Are there any striking problems in classifying different signs?

I hope, that this proposal helps you to understand the challenge I want to tackle clearly.

Best regards,

Stefan Merbele