Capstone Proposal

2016 October 26th
Tomoaki Tsuzuki

Domain Background

Detecting hand written or crafted multi digit numbers from real world image or picture has number of applications such as Google street view housing number detection[*1]. This type of problem was really hard to be solved computationally before the rise of deep learning. With vast number of data becomes available and emerging of the deep learning technology, it is now possible to address this type of the problem with computers. Multi digit number or multi letters character recognition enables new application which cannot be created before, such as real time translation of road signs[*2]. It can also find lots of other applications such as vehicle plate recognition by image taken by security camera for example.

Problem Statement

In this project, recognition of multi digit number from real world image will be addressed. The result is quantifiable by comparing prediction of the numbers with actual data.

Datasets and Inputs

Street View House Number data is used in this project which is available from this site[*3]. The data is digital images with labels associated with it. Each image has multi digit numbers in it. The project is to use those images to train neural network model such that it can predict numbers in the images. The

test data is also available from same site, so trained model performance can also be measured with the data available.

Solution Statement

The solution to this problem is to use convolutional neural network. Convolutional neural network model is trained by training data. Its performance will be measured by predicting unseen test data.

Benchmark Model

Paper from Google “Multi digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks”[*1] is the benchmark for this project.

Evaluation Metrics

Evaluation will be done by comparing prediction of test data that model will produce against real numbers in the picture.

Project Design

Project will be designed as follows.

1. Problem statement
2. Data analysis
3. Explanation of algorithm
4. Implementation
5. Result and improvements
6. Conclusion

References

*1 Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks https://arxiv.org/pdf/1312.6082.pdf

*2 http://newatlas.com/google-translate-update/35605/

*3 The Street View House Numbers Dataset: http://ufldl.stanford.edu/housenumbers/