TRAFFIC SIGN RECOGNITION USING DEEP LEARNING

A SECOND YEAR PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF B.Sc. IN COMPUTATIONAL MATHEMATICS

BY

1. Prashant Timalsina (Registration no. 028322-20)
2. Prastut Bhattarai (Registration no. 028300-20)
3. Susant Shrestha (Registration no. 028317-20)

SCHOOL OF SCIENCE
KATHMANDU UNIVERSITY
DHULIKHEL, NEPAL

8 June 2023

CERTIFICATION

This project entitled "Traffic Sign Recognition Using Deep Learning" is carried out under my supervision for the specified entire period satisfactorily, and is hereby certified as a work done by following students

1. Prashant Timalsina
2. Prastut Bhattarai
3. Susant Shrestha

in partial fulfillment of the requirements for the degree of B.Sc. in Computational Mathematics, Department of Mathematics, Kathmandu University, Dhulikhel, Nepal.

Dr.Samir Shrestha
Associate Professor
Department of Mathematics, School of Science
Kathmandu University, Dhulikhel, Kavre, Nepal
Date: 8th june, 2023

APPROVED BY:

I hereby declare that the candidate qualifies to submit this report of the Math Project (MATH-252) to the Department of Mathematics.

Head of the Department
Department of Mathematics
School of Science
Kathmandu University
Date:

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Traffic signs are vital for managing traffic on the road, making sure drivers follow the rules, and preventing accidents and damage. Intelligent Transportation Systems (ITS) use automatic detection and recognition to manage traffic signs. With the rise of self-driving cars, it's even more important to have this technology.With the use of traffic sign recognition technology, a car can recognizesigns like "speed limit" or "turn ahead" that have been placed on the road. The first TSR systems which recognized speed limits were developed in cooperation by Mobileye and Continental.
The ability to recognize traffic signs is essential for driverless vehicles and sophisticated driver support systems. For safe and effective driving, it entails real-time detection and classification of traffic signs. Traffic sign recognition is one of many computer vision problems where deep learning has demonstrated exceptional results. .The advancement of computer hardware has led to significant improvements in artificial neural networks, resulting in a golden age for the development of machine learning. Deep learning is a specific type of machine learning technique that utilizes a neural network model to simulate the structure of the human brain when processing information. This model is superior to conventional TSR algorithms in extracting effective features from road images, which has the potential to enhance the robustness and generalization of the algorithms.

1.1.1 Literature Review

1. "Traffic Sign Recognition with Multi-Scale Convolutional Networks"[5]: This influential paper introduced the use of convolutional neural networks (CNNs) for traffic

sign recognition. The authors proposed a multi-scale CNN architecture capable of handling traffic sign images at different resolutions, achieving state-of-the-art performance on benchmark datasets.

2. "Deep Neural Networks for Traffic Sign Recognition"[1]: This study presented a deep learning approach for traffic sign recognition using a deep neural network architecture. The authors introduced a novel data augmentation technique and achieved superior performance on the German Traffic Sign Recognition Benchmark (GTSRB) dataset.

3. "DeepTraffic: Driving Fast through Dense Traffic with Deep Reinforcement Learning"[2]: This research explored the use of deep reinforcement learning for traffic sign prediction in the context of autonomous driving. The authors developed an end-to-end deep reinforcement learning framework that learned to predict traffic signs and make appropriate driving decisions. The study demonstrated the potential of deep learning in addressing complex traffic scenarios.

4. "Real-Time Traffic Sign Recognition Based on Deep Learning and Semantic Segmentation" [3]: This study proposed a real-time traffic sign recognition system using deep learning and semantic segmentation. The authors utilized a combination of convolutional neural networks and recurrent neural networks to achieve accurate detection and classification of traffic signs in real-time scenarios. The research focused on improving the efficiency and robustness of traffic sign recognition algorithms.

5. "Traffic Sign Recognition with Transfer Learning: An Investigation" [4]: This research investigated the effectiveness of transfer learning in traffic sign recognition. The authors explored different pre-trained models, transfer learning strategies, and fine-tuning techniques to improve the performance of traffic sign prediction. The study emphasized the benefits of transfer learning for overcoming limited training data scenarios.

6. "Efficient Traffic Sign Recognition with Shallow and Deep Models"[6] : This study compared the performance of shallow and deep models for traffic sign recognition. The authors evaluated various shallow and deep learning architectures and analyzed their computational efficiency and accuracy. The research aimed to find a balance between model complexity and prediction performance.

1.2 Objectives

1. Enhancing Road Safety: To enhance road safety by precisely detecting and identifying traffic signs, enabling drivers to quickly and accurately comprehend the information conveyed by the signs. This facilitates adherence to traffic rules, reduces risks, and enables informed decision-making while driving.

2. Reducing Human Errors: Human mistakes, like not seeing or misunderstanding traffic signs, can result in accidents and breaking the law. Traffic sign recognition systems try to decrease such errors by detecting and understanding signs precisely and dependably, lessening the need for drivers to be constantly alert and attentive.

3. Enabling Intelligent Transportation Systems: Traffic sign recognition is an important part of intelligent transportation systems (ITS). It helps to collect useful data on traffic patterns, congestion, and road conditions by recognizing and analyzing traffic signs. This information can be used to manage traffic, plan infrastructure improvements, and optimize transportation systems.

4. Supporting Autonomous Vehicles: The recognition of traffic signs is extremely important for the functioning and progress of self-driving cars. These vehicles need to precisely detect and comprehend traffic signs to ensure safe navigation and take correct decisions while driving. Traffic sign recognition systems have a significant role to play in facilitating self-driving cars to perceive and react to traffic rules and situations.

1.3 Scope

The scopes of traffic sign recognition encompass various areas where this technology can be applied. Some of the key scopes of traffic sign recognition are as follows:

1. Driver Assistance Systems: Advanced driver assistance systems (ADAS) can include traffic sign recognition to aid drivers in real-time. The system can detect and interpret traffic signs, providing visual or audible alerts to ensure compliance with speed limits, stop signs, no-entry signs, and other traffic regulations. This feature enhances safety and adherence to traffic rules.

2. Autonomous Vehicles: Traffic sign recognition plays a crucial role in the functioning and advancement of autonomous vehicles. These vehicles heavily depend on precise detection and comprehension of traffic signs to effectively navigate and make informed decisions. The ability to recognize and interpret signs enables autonomous vehicles to regulate their speed, maneuver intersections, and respond suitably to varying road conditions.

3. Intelligent Transportation Systems (ITS): Traffic sign recognition plays a crucial role in intelligent transportation systems. These systems can collect significant data on traffic patterns, congestion, and road conditions by identifying traffic signs with precision. This data can be used for managing traffic, enhancing traffic flow, and ultimately improving transportation efficiency.

4. Traffic Data Analysis: Traffic sign recognition systems have the potential to aid in the gathering and evaluation of traffic data. These systems can identify and categorize traffic signs, which can offer important knowledge about traffic flow, road conditions, and infrastructure needs. Such data can be used to plan traffic, enhance transportation systems, and advance urban mobility.

1.4 Limitations

Some common limitations of traffic sign recognition using deep learning:

1. Large Amount of Training Data: Deep learning models usually need a considerable quantity of labelled training data to learn efficiently. Obtaining a diverse and comprehensive dataset of labelled traffic sign images can be a difficult and time-consuming task. Inadequate training data can result in overfitting or poor generalization, particularly for rare or region-specific traffic signs.

2. Limited Generalization to Unseen Conditions: Deep learning models face difficulty in recognizing objects when they encounter conditions that are significantly different from the training data. For instance, if the model is trained on images captured during the daytime, it may not be able to identify signs during nighttime or adverse weather. The process of generalizing the model to new scenarios such as different lighting, weather, or road environments is still a challenging task.

3. Computational Complexity:Traffic sign recognition using deep learning models can be computationally intensive and requires significant processing power and memory resources. Achieving real-time performance on resource-limited devices or embedded systems can be difficult. Therefore, it is important to balance accuracy and efficiency in model design and optimization.

4. Dependency on Large-Scale Training: Deep learning models require large-scale training with annotated data. When new traffic signs are introduced or existing signs are modified, it becomes necessary to update and retrain the models. This process can be resource-intensive and time-consuming, requiring continuous effort to keep the models up to date.

CHAPTER 2

METHODOLOGY

2.1 Theoretical/Conceptual Framework

1. Deep Learning: Deep learning forms the foundation of the conceptual framework, as it is the core technology used for traffic sign prediction. Deep learning involves the use of neural networks, specifically convolutional neural networks (CNNs), to automatically learn hierarchical representations from raw image data. CNNs excel in capturing complex visual patterns and have demonstrated remarkable performance in various computer vision tasks, including image recognition..

2. Traffic Sign Recognition: Traffic sign recognition is a subfield of computer vision that focuses on detecting and interpreting traffic signs in images or video streams. The conceptual framework revolves around the task of accurately detecting and classifying traffic signs using deep learning techniques. The objective is to develop a robust and efficient model capable of recognizing traffic signs in real-time scenarios, considering variations in shape, color, size, occlusion, and environmental conditions.

3. Data Acquisition and Preprocessing: The conceptual framework includes the acquisition and preprocessing of traffic sign data. This involves collecting a diverse and representative dataset of labeled traffic sign images, which may be obtained from public traffic sign databases, road surveillance systems, or simulation environments. Data preprocessing techniques such as resizing, normalization, and augmentation may be applied to enhance the quality and diversity of the dataset.

4. Model Architecture and Training: The model architecture represents the design and structure of the deep learning model used for traffic sign prediction. The conceptual

framework includes selecting an appropriate CNN architecture, such as VGGNet, ResNet, or InceptionNet, and customizing it for the task of traffic sign recognition. The model is trained using the acquired dataset, employing optimization techniques such as stochastic gradient descent and backpropagation to minimize the prediction error.

5. Evaluation Metrics and Performance Analysis: The conceptual framework involves the evaluation of the trained deep learning model's performance. Various evaluation metrics such as accuracy, precision, recall, and F1 score may be used to assess the model's ability to accurately predict traffic signs. The performance analysis may also include studying the model's robustness to variations, analyzing its computational efficiency, and comparing its performance against baseline methods or previous studies.

6. Deployment and Practical Applications: The conceptual framework considers the practical deployment and applications of the developed deep learning model for traffic sign prediction. This includes integrating the model into autonomous vehicles, advanced driver-assistance systems, or traffic management systems. The framework explores the implications, benefits, and challenges associated with implementing deep learning-based traffic sign recognition systems in real-world scenarios.

2.2 Model Architecture

An effective model architecture for traffic sign prediction using Convolutional Neural Networks (CNNs) is the following:

2.2.1 Input Layer

The input layer accepts the input traffic sign image, typically in the form of a 2D matrix representing pixel intensities.

2.2.2 Convolutional Layers

The convolutional layers perform the primary feature extraction. Each convolutional layer applies a set of learnable filters or kernels to the input image. These filters scan the image in a sliding window manner, performing convolutions to capture local patterns and features. The resulting feature maps highlight different aspects of the input image.

2.2.3 Activation Function

After each convolutional layer, an activation function such as ReLU (Rectified Linear Unit) is applied element-wise to introduce non-linearity and enhance the model's representational power. ReLU sets negative values to zero, while preserving positive values.

2.2.4 Pooling Layers

Pooling layers reduce the spatial dimensions of the feature maps, reducing computational complexity and providing a form of translation invariance. Common pooling operations include max pooling, which selects the maximum value within a pooling window, or average pooling, which computes the average value.

2.2.5 Fully Connected Layers

The flattened output from the last pooling layer is fed into one or more fully connected layers. These layers learn high-level representations by connecting every neuron from the previous layer to every neuron in the current layer. Each neuron in the fully connected layer computes a weighted sum of inputs and applies an activation function, typically ReLU.

2.2.6 Output Layer

The output layer produces the final predictions. For traffic sign prediction, the number of neurons in the output layer corresponds to the number of traffic sign classes. The output can be interpreted as class probabilities using a softmax activation function, where the highest probability indicates the predicted traffic sign class.

The specific architecture details, such as the number of convolutional layers, pooling layers, and fully connected layers, as well as the size of filters and pooling windows, can vary based on the complexity of the dataset and the available computational resources. Advanced architectures like VGGNet, ResNet, or InceptionNet can also be explored for improved performance, especially when dealing with larger and more challenging traffic sign datasets.

CHAPTER 3

RESULTS

3.1 Data Visualization

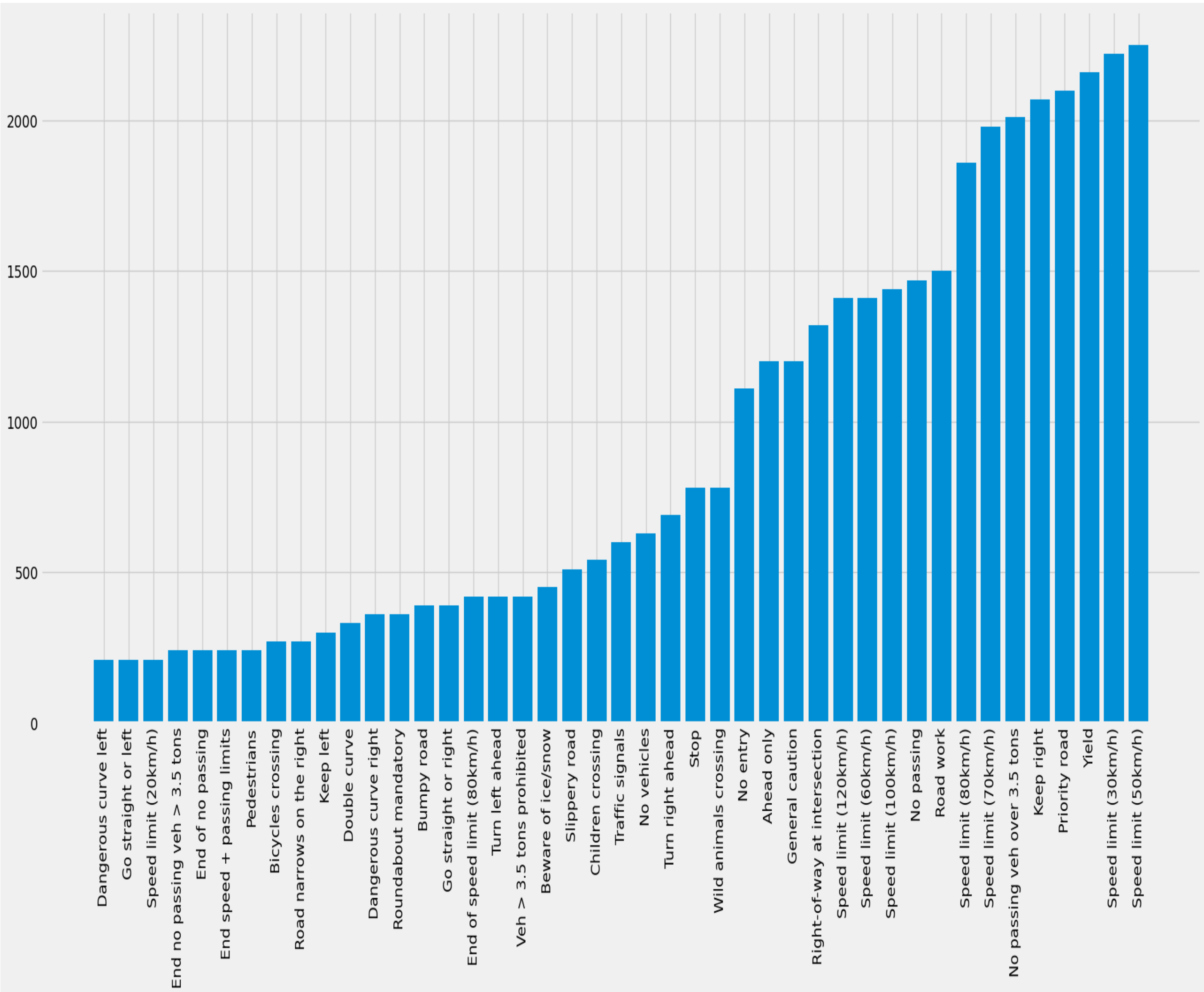

Figure 3.1: Data Visualization

3.2 Classification Summary

	precision	recall	f1-score	support
0	1.00	1.00	1.00	60
1	0.98	0.99	0.99	720
2	0.99	0.97	0.98	750
3	0.90	0.99	0.95	450
4	1.00	0.98	0.99	660
5	0.96	0.99	0.98	630
6	0.90	0.99	0.94	150
7	1.00	1.00	1.00	450
8	1.00	0.94	0.97	450
9	1.00	0.97	0.99	480
10	1.00	1.00	1.00	660
11	0.98	1.00	0.99	420
12	1.00	1.00	1.00	690
13	1.00	1.00	1.00	720
14	0.99	1.00	0.99	270
15	1.00	1.00	1.00	210
16	1.00	1.00	1.00	150
17	1.00	1.00	1.00	360
18	0.99	0.95	0.97	390
19	1.00	1.00	1.00	60
20	0.98	1.00	0.99	90
21	0.99	1.00	0.99	90
22	0.99	0.76	0.86	120
23	0.96	0.98	0.97	150
24	0.98	0.93	0.95	90
25	0.93	0.99	0.96	480
26	1.00	1.00	1.00	180
27	0.95	1.00	0.98	60
28	0.97	0.99	0.98	150
29	0.90	1.00	0.95	90
30	0.97	0.92	0.95	150
31	1.00	0.98	0.99	270
32	0.96	0.78	0.86	60
33	0.99	1.00	1.00	210
34	0.98	1.00	0.99	120
35	1.00	0.96	0.98	390
36	1.00	0.98	0.99	120
37	0.98	0.98	0.98	60
38	0.99	0.99	0.99	690
39	0.99	0.97	0.98	90
40	0.96	1.00	0.98	90
41	1.00	1.00	1.00	60
42	1.00	1.00	1.00	90
accuracy			0.98	12630
macro avg	0.98	0.98	0.98	12630
weighted avg	0.98	0.98	0.98	12630

Figure 3.2: Data Augmentation

3.3 Model Summary

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 28, 28, 16)        448

 conv2d_1 (Conv2D)           (None, 26, 26, 32)        4640

 max_pooling2d (MaxPooling2D  (None, 13, 13, 32)       0
 )

 batch_normalization (BatchN  (None, 13, 13, 32)       128
 ormalization)

 conv2d_2 (Conv2D)           (None, 11, 11, 64)        18496

 conv2d_3 (Conv2D)           (None, 9, 9, 128)         73856

 max_pooling2d_1 (MaxPooling  (None, 4, 4, 128)        0
 2D)

 batch_normalization_1 (Batc  (None, 4, 4, 128)        512
 hNormalization)

 flatten (Flatten)           (None, 2048)              0

 dense (Dense)               (None, 512)               1049088

 batch_normalization_2 (Batc  (None, 512)              2048
 hNormalization)

 dropout (Dropout)           (None, 512)               0

 dense_1 (Dense)             (None, 43)                22059

=================================================================
Total params: 1,171,275
Trainable params: 1,169,931
Non-trainable params: 1,344
_________________________________________________________________
```

Figure 3.3: Model "Sequential"

3.4 Accuracy

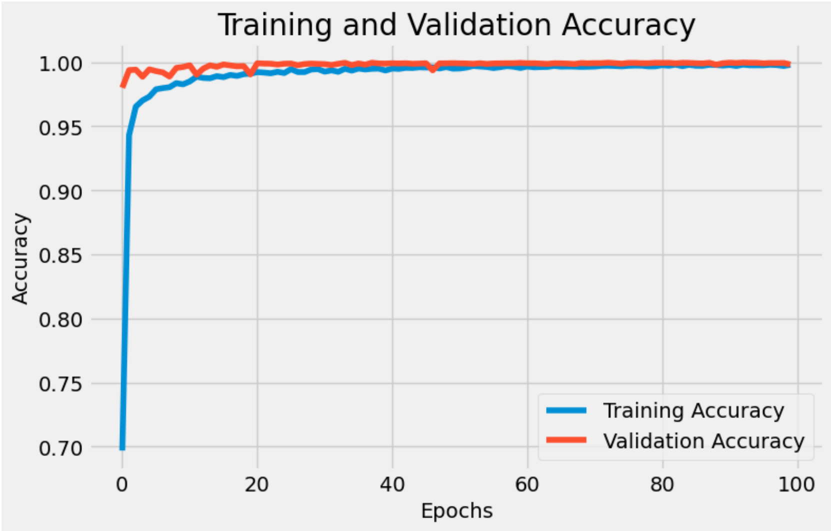

Figure 3.4: Accuracy

3.5 Loss

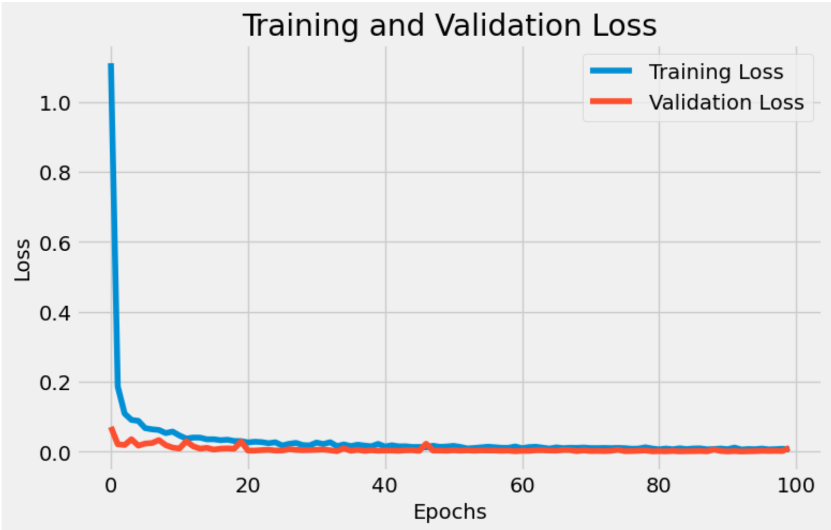

Figure 3.5: Loss

CHAPTER 4

CONCLUSIONS

In conclusion, traffic sign prediction using deep learning, specifically Convolutional Neural Networks (CNNs), offers a promising solution for accurate and efficient recognition of traffic signs in various real-world scenarios. Through the review of literature, it is evident that deep learning-based approaches have demonstrated significant advancements in traffic sign recognition, surpassing traditional methods and achieving high levels of accuracy. The objectives of this project were to develop a robust and efficient model for traffic sign prediction and to investigate the effectiveness of deep learning techniques in this domain. By following the theoretical and conceptual framework, a model architecture was designed, involving convolutional layers for feature extraction, pooling layers for dimension reduction, and fully connected layers for high-level representation learning.

The training and optimization process involved dataset preparation, data augmentation, hyper parameter tuning, choice of optimization algorithm, regularization techniques, and monitoring of the model's performance. Through proper training and optimization, the model can learn discriminative features from traffic sign images, improve generalization, and mitigate overfitting..

Theoretical studies and empirical evidence from the literature review support the effectiveness of CNNs for traffic sign prediction. The model's performance can be evaluated using various metrics such as accuracy, precision, recall, and F1-score, providing a quantitative assessment of its predictive capabilities.

The application of deep learning-based traffic sign prediction extends beyond research and has practical implications in real-world scenarios. The deployment of such models in autonomous vehicles, advanced driver-assistance systems, and traffic management systems can enhance road safety, improve traffic efficiency, and assist drivers in making informed decisions.

However, challenges such as occlusion, variations in environmental conditions, and real-time processing constraints need to be considered for the successful implementation of traffic sign prediction systems. Ongoing research and development efforts are focused on addressing these challenges and further improving the accuracy, robustness, and efficiency of deep learning models for traffic sign prediction.

In summary, traffic sign prediction using deep learning is a promising and rapidly evolving field. With advancements in deep learning techniques, availability of large-scale labeled datasets, and computational resources, the accuracy and reliability of traffic sign recognition systems continue to improve, paving the way for safer and more efficient transportation systems in the future.

REFERENCES

[1] Dan Claudiu Ciresan, Ueli Meier, Luca Maria Gambardella, and Jrgen Schmidhuber, *Deep, big, simple neural nets for handwritten digit recognition*, Neural Computation **22** (2012), no. 12, 3207–3220.

[2] Lex Fridman, Jack Terwilliger, and Benedikt Jenik, *Deeptraffic: Crowdsourced hyper-parameter tuning of deep reinforcement learning systems for multi-agent dense traffic navigation*, 2019.

[3] Kuo, *Real-time traffic sign recognition based on deep learning and semantic segmentation*, (2020).

[4] Chunmian Lin, Lin li, Wenting Luo, Kelvin Wang, and Jiangang Guo, *Transfer learning based traffic sign recognition using inception-v3 model*, Periodica Polytechnica Transportation Engineering **47** (2018).

[5] Pierre Sermanet and Yann Lecun, *Traffic sign recognition with multi-scale convolutional networks*, 09 2011, pp. 2809 – 2813.

[6] Yuchen Yang, Shuo Liu, Wei Ma, Qiuyuan Wang, and Zheng Liu, *Efficient traffic-sign recognition with scale-aware cnn*, 2018.