UNLV HOWARD R. HUGHES COLLEGE OF ENGINEERING

U.S. Traffic Sign Recognition – Can We Bridge the Gap?

Author: Andres Graterol (University of Central Florida) | graterol.andres0@knights.ucf.edu

REU Smart Cities Research Program, Summer 2021

Research Mentor: Shahram Latifi, Ph. D

UCF

1. Introduction

- There are over 1,400 self-driving cars in the United States being tested on by over 80 companies.
- Traffic sign detection and recognition (TSDR) systems are especially important to the development of advanced driver assistance systems (ADAS). [3]
- The rise in prevalence of ADAS in the United States is going to require classification systems that are trained on U.S. Traffic Signs. This research hopes to determine whether we can begin to bridge the gap in image classification models that are trained based on these signs.

2. Research Focus

Dataset	Country	Classes	TS Scenes	TS Images	Image Size (px)	Sign Size (px)	Include Videos
GTSDRB (2012 and 2013)	Germany	43	9000	39,209 (training), 12,630 (testing)	15 × 15 to 250 × 250	15 × 15 to 250 × 250	No
KULD (2009)	Belgium	100+	9006	13,444	1628 × 1236	100 × 100 to 1628 × 1236	Yes, 4 tracks
STSD (2011)	Sweden	7	20,000	3488	1280 × 960	3 × 5 to 263 × 248	No
RUGD (2003)	The Netherlands	3	48	48	360 × 270	N/A	No
Stereopolis (2010)	France	10	847	251	1920 × 1080	25 × 25 to 204 × 159	No
LISAD (2012)	US	49	6610	7855	640 × 480 to 1024 × 52	6 × 6 to 167 × 168	All annotations
UKOD (2012)	UK	100+	43,509	1200 (synthetic)	648 × 480	24 × 24	No
RTSD (2013)	Russia	140	N/A	80,000+ (synthetic)	1280 × 720	30 × 30	No

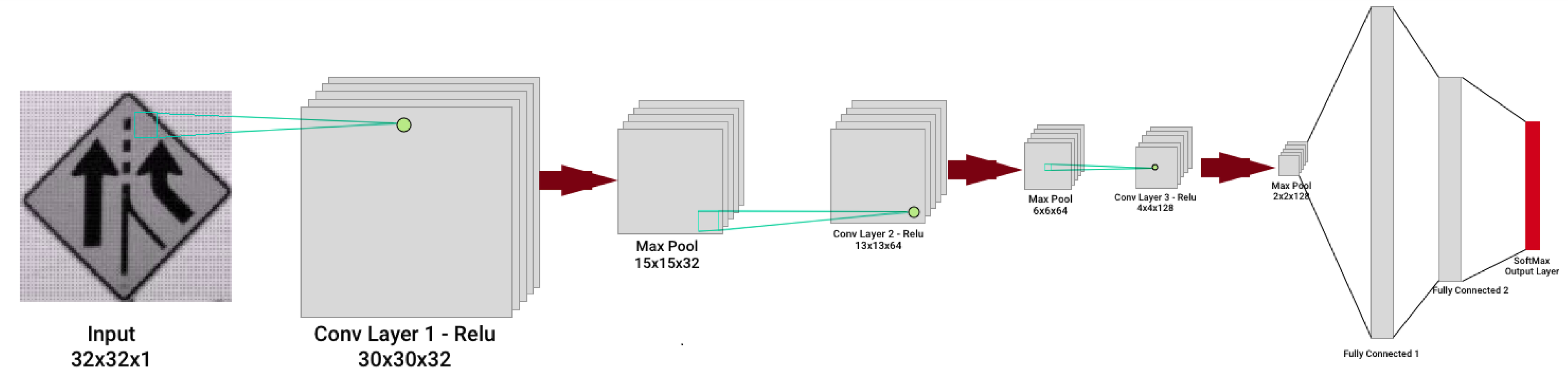
Publicly Available Traffic Sign Datasets

- Currently, most existing high-accuracy image classification models for traffic signs are trained on the German Traffic Sign Recognition Benchmark Dataset (GTSRB).
- How will the U.S. Traffic Sign Dataset, LISA, perform in training and testing an image classification model?
- Additionally, what would happen if we fed the classifier imperfect, yet realistic traffic sign images? How would it perform?

"Imperfect" Traffic Sign Images

- By measuring the testing accuracy, we can determine whether this dataset and the corresponding classification model would be efficient in training ADAS.
- We will be using the "EdLeNet"[2] 3x3 Convolutional Neural Network (CNN) architecture due to its high performance with classifying images in the GTSRB dataset.
- We are choosing to use this architecture over YOLOv5 (You Only Look Once version 5), an object detection model, because we want to make sure that we can accurately achieve image classification first.

3. Methodology



EdLeNet 3x3 CNN Architecture

- Split the LISA Dataset[1] 80/20 into training and testing sets.
- Used the EdLeNet 3x3 CNN architecture for the image classification model. Input image is a 32x32 grayscale image and the output is the model's prediction. Training and testing the model on the 32x32x1 images acted as our base case.
- Performed image preprocessing on the training/testing sets to reduce the noise in each image and as a result, increase model accuracy. After being tested on the base case, a series of progressively more complex transformation compositions was applied to observe how the model handles noise.
- Lastly, tested the model on the "extreme case", the imperfect traffic sign images.

Original Grayscale

Image Preprocessing for Producing Input Image

4. Training & Testing

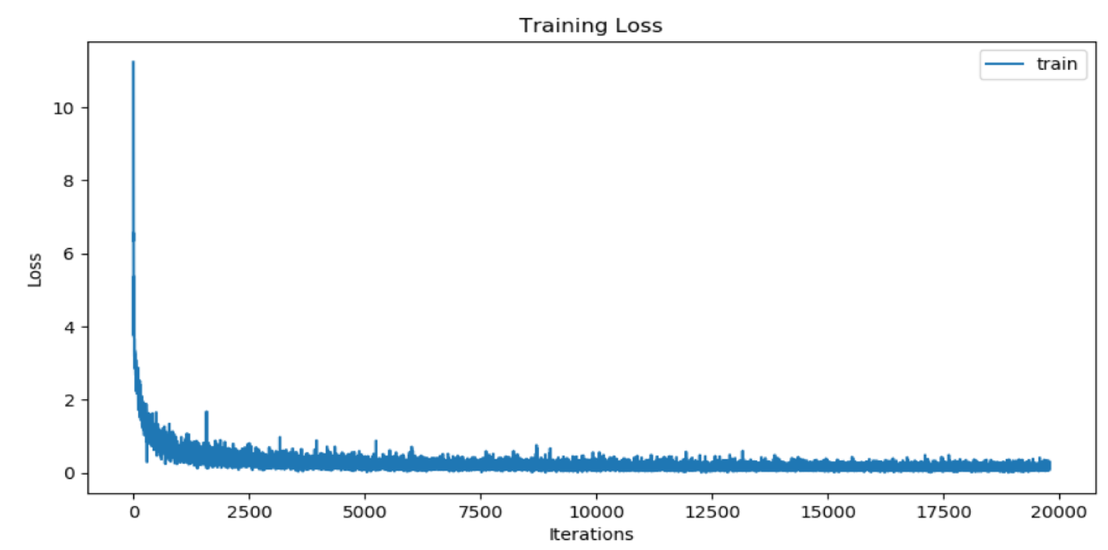


Figure 1 - Training Loss: Base Case Traffic Sign Images

- The model was trained on a random 80% split from the entire LISA dataset. Total of 5472 training images.
- This curve demonstrates how well the model is fitting the training data. Minimal loss leads to a more accurate model.

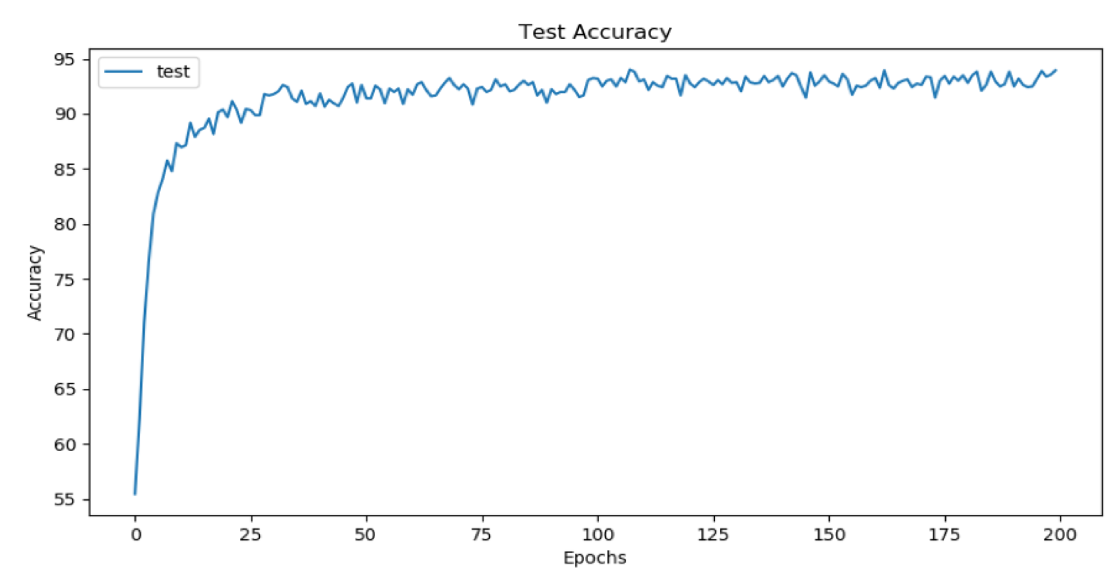


Figure 2 – Testing Accuracy: Base Case Traffic Sign Images

- The model was tested on a random 20% split from the entire LISA dataset. Total of 1515 testing images.
- This curve demonstrates how accurately the model is making its classification predictions. Over 200 epochs the average accuracy was 93%.

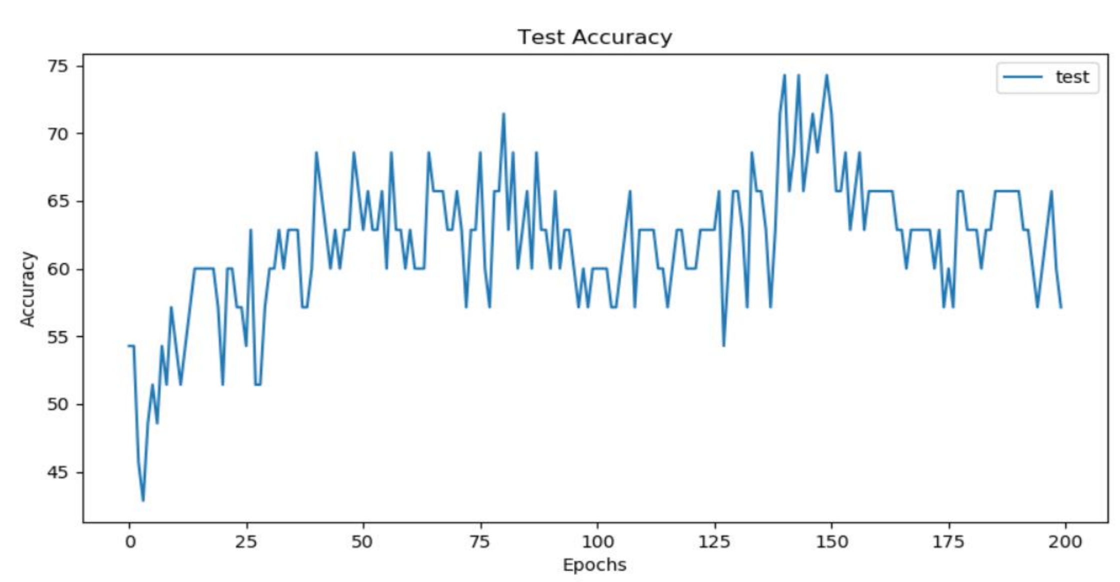


Figure 3 – Testing Accuracy: Extreme Case Traffic Sign Images

- The model was tested on 37 "imperfect" images that are representative of signs you may find in reality.
- Oscillations occur around 63% accuracy.

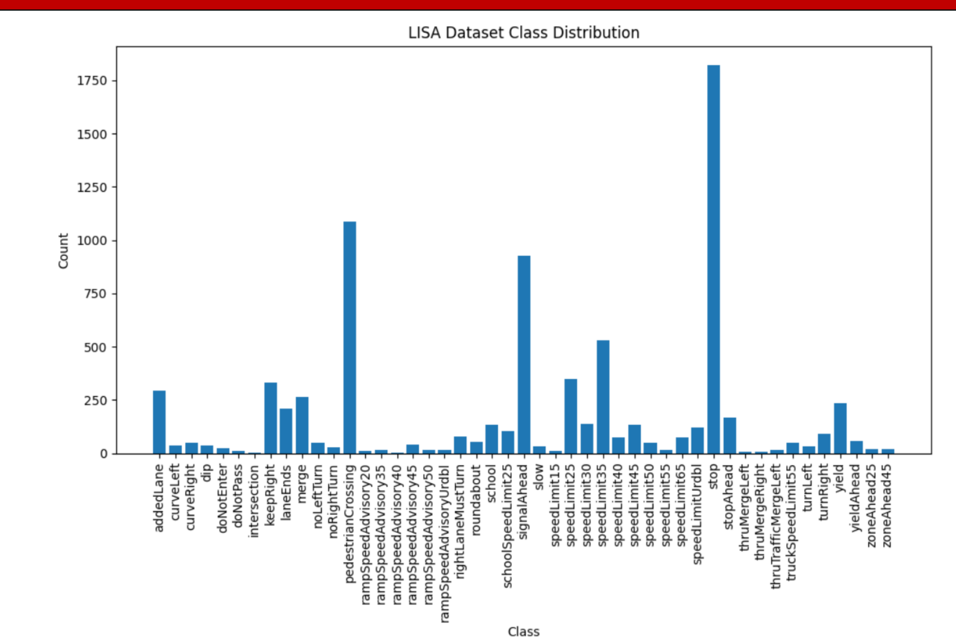
5. Results

Input	Transformations	Average Training Loss	Average Testing Accuracy
LISA Images	Resize, Grayscale, Gaussian Blur	0.31	93%
LISA Images	Resize, Grayscale, Random Rotation, Random Horizontal Flip	1.87	62%
LISA Images	Resize, Grayscale, Random Resized Crop, Random Affine	4.18	16%
Collection of Augmented Signs	Resize, Grayscale, Equalize, Gaussian Blur	2.69	63%

Model Output Results

- Average training loss and average testing accuracy both decreased as the transformations applied to the images became more complex.
- To compare, EdLeNet performed with 98% testing accuracy when using the GTSRB dataset to train and test the model.

6. Conclusions & Future Work



LISA Dataset Class Distribution

- This loss in testing accuracy can be attributed to bias in the class distribution of the LISA dataset. This bias contributes to overall loss of testing accuracy (*Figure 3).*
- These images are only from the San Diego, California area. Some signs are underrepresented, and some are not present at all. A more robust dataset is needed in order to more efficiently train and test an image classifier based on U.S. traffic signs.
- Plan on creating a testing set based on crisp images with missing information in order to obviate the need for image preprocessing. Will compare training loss and testing accuracy to the body of obtained data.
- In the future, I plan on testing the LISA Dataset against other, more modern CNN architectures such as ResNet. Also intend to move on to object detection models, possibly using YOLOv5.

6. Acknowledgements

[1] Andreas Møgelmose, Mohan M. Trivedi, and Thomas B. Moeslund, "Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey," IEEE Transactions on Intelligent Transportation Systems, 2012.

[2] Sermanet, Pierre, and Yann Lecun. "Traffic Sign Recognition with Multi-Scale Convolutional Networks." *The 2011 International Joint Conference on Neural Networks*, 2011.

[3] Wali, Safat B., et al. "Vision-Based Traffic Sign Detection and Recognition Systems: Current Trends and Challenges." *Sensors*, vol. 19, no. 9, 2019, p. 2093.

#1950872