Convolutional Neural Network for Traffic Sign Identification

Cowan, Z.S., Reynolds, B.S., & Rice, M.J.

Bellarmine University

October 14th, 2023

Abstract

This paper discusses the use of convolutional neural networks (CNN's) for the identification of traffic signage. It seeks to compare the efficiency, accuracy, and precision of various different convolutional neural network architectures. Ideally the end goal of such an image classification system would be in the trustworthy and effective implementation in semi-autonomous or fully autonomous vehicles. This paper explores three main methods: one with VGG16 transfer learning pretrained with the ImageNet database, the similar VGG19 model, and a 19-layer architecture without pretraining.

Introduction

Within the last few years, the field of artificial intelligence has advanced greatly with many implementations becoming very practical to people of all walks of life. With the significant improvements of artificial intelligence, comes the desire to use this new technology to make life easier for people. A common and convenient use of artificial intelligence for quality-of-life improvements is automation of common, monotonous activities. A specific and key implementation of artificial intelligence is that of autonomous vehicles.

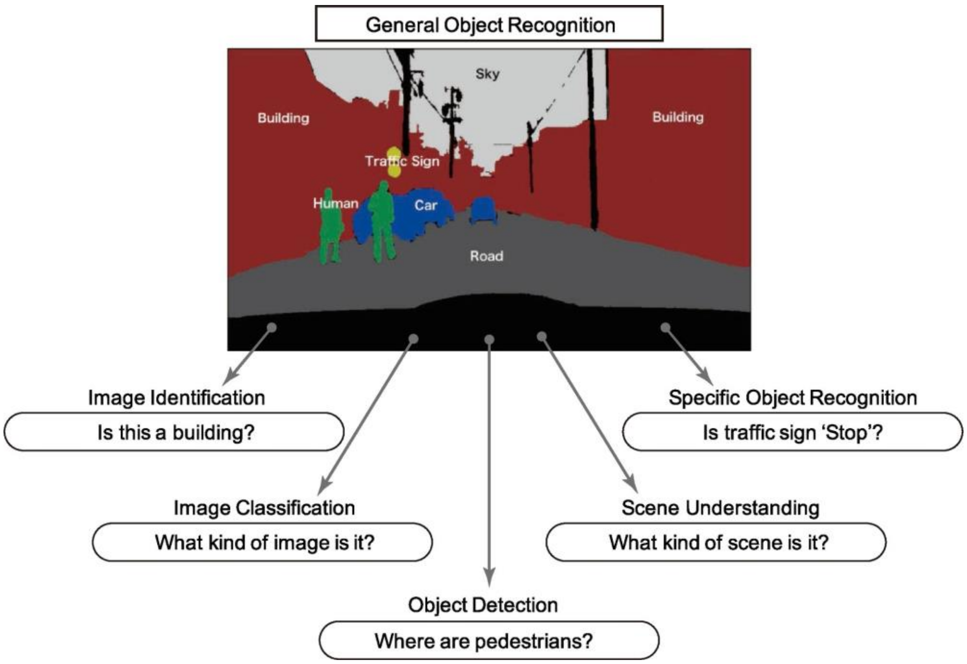
With any activity that has the potential to be dangerous, safety must be the number one priority when attempting to automate it. This is very true when it comes to automating cars as driving is very dangerous and has the potential to be lethal because of a miniscule error. When automating vehicles maximize safety, the vehicle must be able to identify obstacles, identify important signage, and make informed decisions based off the actions of other vehicles near it (Ni et al., 2020). This paper contains research and results of the use of artificial intelligence based off of the Convolutional Neural Network (CNN) structure to classify various important road signs that a car must be able to identify in real time.

Literature review

"A Survey on Theories and Applications for Self-Driving Cars Based on Deep Learning Methods" by Ni, J et al., details how the improvements in the field of artificial intelligence have led to many internet and automotive companies investing heavily into autonomous vehicles, the general structure of an autonomous vehicle system, the importance of deep learning when implementing autonomous vehicle systems, and current applications of deep learning in self-driving cars. The paper analyzes multiple approaches to deep learning but contains much detail on the successes of the CNN approach. Ni et al. state that the reason for the success of CNN is because of its local receptive fields, shared weights, and spatial sampling. Since many self-driving cars use image feature representations, the CNN approach allows for the system to classify obstacles, scenes, lanes, etc. very quickly and efficiently (2020).

In recent years, vision sensors have become very popular tools for obstacle detection in autonomous vehicles due to them having a fast-sampling speed, a large amount of information, and a low cost. This paper concludes that CNN algorithms perform quite well when it comes to recognizing traffic signs, detecting lanes, and detecting obstacles in the car's path using information from visual sensors due to it being good at distinguishing features from each other, having good spatial sense, and being able to process images very quickly. The paper concludes that deep learning will continue to grow more sophisticated and could possibly be the solution to creating fully autonomous cars.

"Deep learning-based image recognition for autonomous driving" by Fujiyoshi, H et al. begins by explaining the different parts that make up general object recognition that a deep learning algorithm must be able to perform in order to be able to drive autonomously. The main parts of object detection are image verification, where the distance of the feature vector of the reference image is calculated with the feature vector of the input image in order to determine if they are the same image, object detection, where a given object's location is found within a given image, image classification, where the object in the image must be classified into one of several predefined categories, scene understanding, which asks what is occurring in the image, and specific object recognition, which finds specific objects within an image. The following figure by Fujiyoshi et al. Simplifies the parts of object recognition (2019).



Their research found that CNN can perform image classification, object detection, and scene understanding because unlike other forms of deep and machine learning, CNN offers the flexibility to be applied to various tasks just by changing the network structure. CNN has also proven very effective in implementations of Advanced Driving Assistance Systems (ADAS) as it is very good at perceiving objects in the environment, especially because of its ability to handle multi-class objects (Fujiyoshi et al., 2019).

"Hyperparameter Optimization in Convolution Neural Network using Genetic Algorithms" by Aszemi and Dominic focuses on hyperparameters, variables that determine the network structure for a given learning algorithm, and how they can be optimized in order to increase a model's accuracy and the speed at which it can operate. Their research used a genetic algorithm (GA) which chooses the best performing model out of many random tests of hyperparameters and continues to test until the user decides to terminate it. They conclude that genetic algorithms will continue to be used for hyperparameter optimizing testing for Convolutional Neural Networks, but it comes at great computational cost as their tests often took three to five days to complete (Aszemi & Dominic, 2019).

Methodology

The aim of this project was to create and compare the accuracy of multiple CNN models to judge which model performed best at classifying various common road signs. The first step in this project was to determine the different CNN architectures and hyperparameters that could be used to compare the accuracy of each model. The main architectures that were chosen were a VGG16 transfer learning model pretrained with the ImageNet database, a VGG19 transfer learning model also pretrained with the ImageNet database, and a 19-layer architecture without pretraining. In addition to comparing the different architectures, the number of epochs the models were trained for were also analyzed to judge how it would affect the accuracy of the model. The activation functions of the layers were not analyzed since research shows that the ReLu activation function is best suited for hidden layers and the Softmax activation function is best suited for the output layers for multiclass image classification (Aszemi, Dominic). The main data that was analyzed to compare the models were the model accuracy and the confusion matrix of each model. The model accuracy was measured as a percentage of correctly classified images. The confusion matrix gave insight into the model performance including false negatives and false positives for each class of image.

Around 1000 images each of yield signs, speed limit signs, stop signs, and handicap signs were used for the models' dataset. Due to data constraints, image augmentation was needed to ensure that the dataset was large enough to accurately train the model. The dataset was divided into training and testing datasets to allow model training and evaluation. A prediction dataset of images was used to test if the models were able to accurately classify the images.

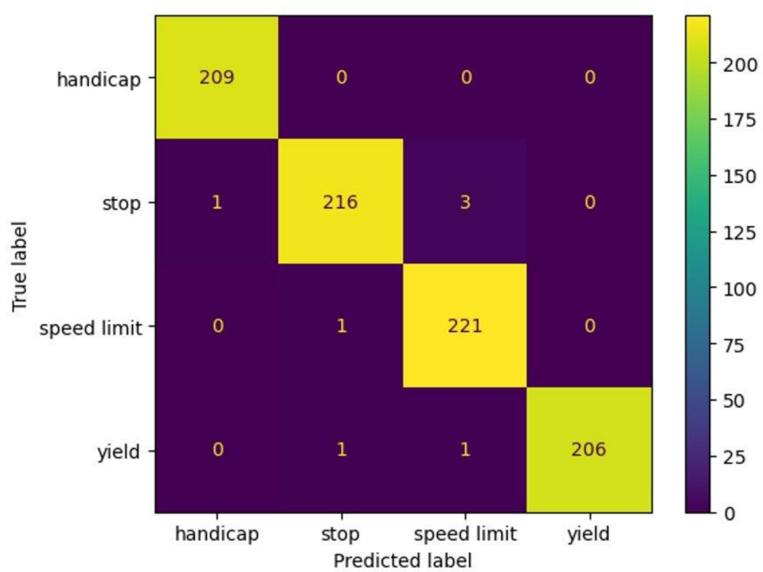
This methodology allowed for comprehensive evaluation of the performance of various multiclass CNN models in classifying various common road signs found in the United States. The results will be discussed in the following sections.

Results

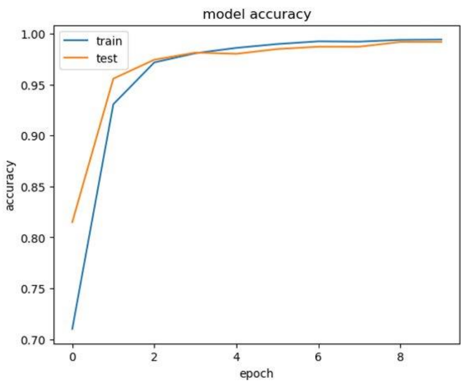
This section reports the results of the various different models this paper utilized. In general, the CNN models utilizing VGG transfer learning resulted in the most precise and accurate traffic sign identification.

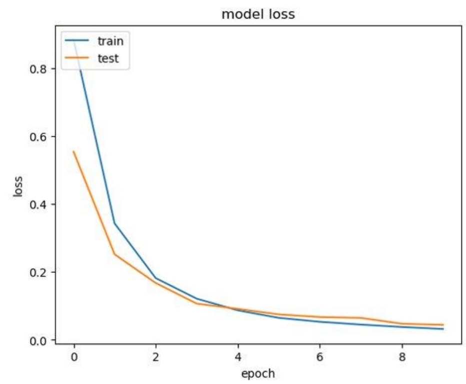
Our CNN with VGG16 transfer learning had results as follows:

- Confusion Matrix shows a great distribution, all but a few images were predicted true to their labeled identification. If any concerns were to be raised it would be the relatively higher number of stop signs being identified as a speed limit sign by the model.



- Accuracy and Precision Graphs are indicative of a model that is very close to a good fit. If anything, this specific model (trained to 10 epochs) may be tending towards overfitting.





Test Cases of VGG16 Model	
 Fig. VGG16.1	Prediction: Stop Sign

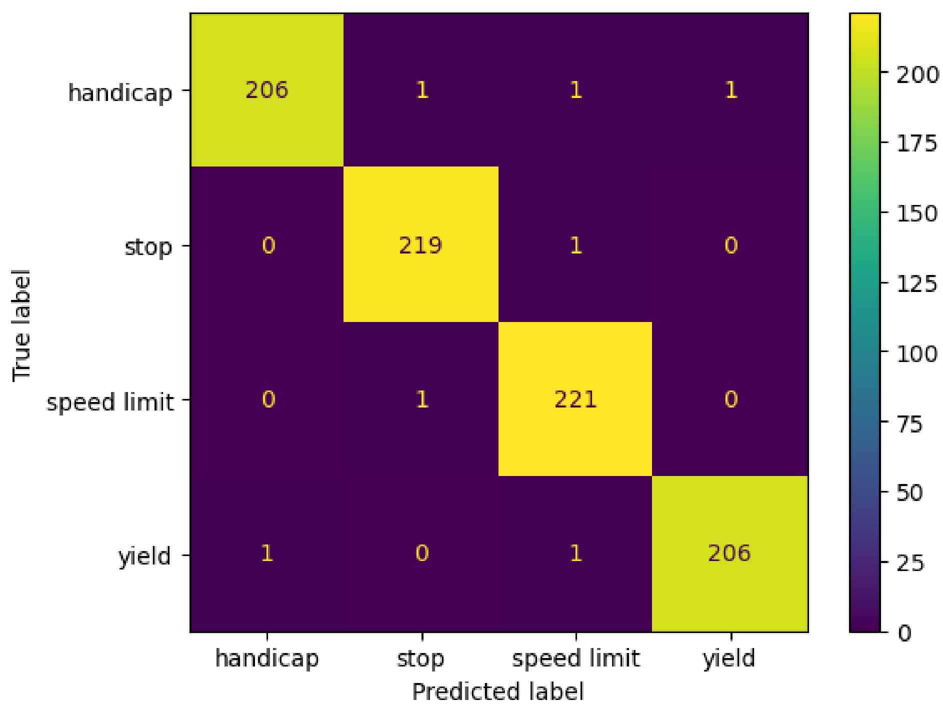
 Fig. VGG16.2	Prediction: Speed Limit Sign
 Fig. VGG16.3	Prediction: Yield Sign
 Fig. VGG16.4 (Cropped from Fig. VGG16.4)	Prediction: Stop Sign

Fig. VGG16.5

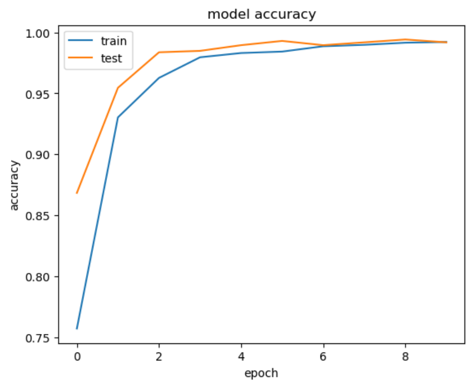
Prediction:
Stop Sign

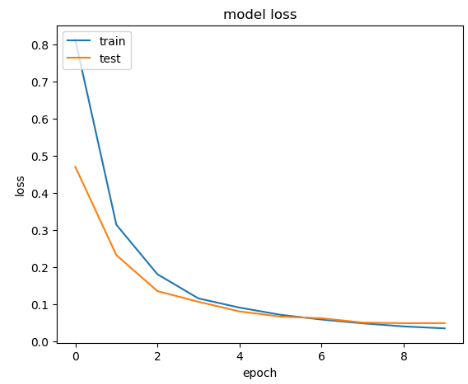
Our CNN with VGG19 transfer learning had results as follows:

- Confusion Matrix shows a great distribution, all but a few images were predicted true to their labeled identification. Unlike the VGG16 model, we don’t see the failed predictions grouping in any specific combination of true-predicted label.



- Accuracy and Precision Graphs are indicative of a model with a good fit.





Test Cases of VGG19 Model	
 Fig. VGG19.1	Prediction: Stop Sign
 Fig. VGG19.2	Prediction: Speed Limit Sign

 Fig. VGG19.3	Prediction: Yield Sign
 Fig. VGG19.4 (Cropped from Fig. VGG19.4)	Prediction: Stop Sign
 Fig. VGG19.5	Prediction: Stop Sign

Discussion

Overall, we found convolutional neural networks to be a valid pathway to traffic sign identification for self-driving vehicles. Although, before we would implement our specific models into a product that takes to the streets, we would certainly invest more time and energy into developing a dataset with more complex scenarios. As seen in the test cases of Fig. VGG16.3 and Fig. VGG19.3, our model struggled to properly identify a stop sign at a harsh angle and a large tree obstructing much of the foreground. While our model did properly identify the same stop sign from a less harsh angle and lesser distance, the model would ideally be able to identify the stop sign from approximately the same distance that a human driver would. This would ensure that if a given model was used in an automated driving scenario, the vehicle would have sufficient time to identify the stop sign and act accordingly.

Conclusion

Data collection is a delicate process. It is not as simple as collecting all images of a given search category online. In our collection process, we had to weed out many images that were auto downloaded as part of a search.

A quality dataset is crucial for training. A model needs to be trained in difficult cases like those with various foreground obstructions, tricky features, etc. At the same time, a dataset with too little quantity will find difficulties in the range of images it can properly identify. Additionally, the smaller the dataset, the easier it becomes to overfit a given model. On the other hand, data collection can be expensive and time-consuming. It can be just as detrimental to train a model on too many images as it is to train one on too little.

In general, for our specific application, the VGG16 and VGG19 ImageNet pretrained models performed better than the "in-house" architecture. They were overall less prone to overfitting and had much greater accuracy when compared to the “in-house” CNN model used.

CNN Image classification alone is surely a potential solution traffic sign identification in autonomous vehicles but fails to be able to classify in cases where multiple signs are present or similar features may "trick" the model into detecting a sign that does not exist.

References

Aszemi, N. M., & Dominic, P. D. D. (2019). Hyperparameter Optimization in Convolutional Neural Network using Genetic Algorithms. *International Journal of Advanced Computer Science and Applications*, *10*(6). https://doi.org/10.14569/ijacsa.2019.0100638

Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. *IATSS Research*, *43*(4). https://doi.org/10.1016/j.iatssr.2019.11.008

Ni, J., Chen, Y., Chen, Y., Zhu, J., Ali, D., & Cao, W. (2020). A Survey on Theories and Applications for Self-Driving Cars Based on Deep Learning Methods. *Applied Sciences*, *10*(8), 2749. MDPI AG. Retrieved from http://dx.doi.org/10.3390/app10082749

Acknowledgments

The authors of this paper would like to thank Dr. Sarkar of Bellarmine University for her help throughout this project. In addition, the authors would like to thank the authors of all references used in this paper.