Behavioural Cloning Algorithms For Autonomous Guided Vehicles

Archit jain GLA University ,Mathura , Indi**a** ,[**jainarchit088@gmail.com**](mailto:jainarchit088@gmail.com)

*Abstract*—With the rapid advancement of artificial intelligence and autonomous learning, Autonomous Guided Vehicles (AGVs) have seen increased popularity in logistics, manufacturing, and transportation sectors. One significant challenge in developing AGVs is behavior cloning, where the vehicle learns to imitate human driving behavior using inputs like camera images and steering angles. In this study, we explore two behavior cloning techniques: the NVIDIA model and transfer learning with VGG16. Our goal is to determine the optimal steering angles. The experimental results indicate that the NVIDIA model outperforms the VGG16 with transfer learning, exhibiting lower mean squared error and better performance in a simulator. The NVIDIA model's ability to generalize data leads to superior decision-making. Our study underscores the potential of deep learning techniques for AGVs and emphasizes the importance of evaluating multiple algorithms to find the best approach.

Keywords- Self driving car, mean square error, vgg16 model, transfer learning, behavioral cloning, NVIDIA model.

# Introduction

In recent times, self-driving vehicles have garnered significant attention due to their transformative potential in the transportation sector. Sivak and Schoettle (2015) highlight that autonomous cars could substantially alleviate traffic congestion, enhance road safety, and offer greater mobility to individuals unable to drive due to disabilities or other limitations.

This technology promises to make transportation more efficient, lower carbon emissions, and improve overall quality of life. Autonomous vehicles (AVs) have the capacity to revolutionize travel, making it safer, more efficient, and more convenient. The advantages of AVs include better road safety, decreased traffic congestion, enhanced mobility for the elderly and disabled, and reduced carbon emissions. AVs have diverse applications, including personal transport, freight delivery, public transit, and emergency response. With the advancement of sophisticated sensing and communication technologies, AVs can operate more safely and efficiently on roads, significantly reducing traffic accidents and congestion. However, achieving full autonomy poses several technical challenges. One method to enable self-driving capabilities is behavioral cloning, where a vehicle is trained to replicate human driving behavior using deep learning algorithms. Behavioral cloning has proven to be an effective technique for autonomous driving. NVIDIA, a prominent technology company, has developed a convolutional neural network (CNN) approach for self-driving cars known as the NVIDIA architecture, which has set the benchmark in autonomous driving, showcasing the potential of deep learning methods for complex real-world problems. This research aims to compare the effectiveness of different deep learning architectures for self-driving cars, specifically the NVIDIA architecture and VGG16. The NVIDIA model is a deep neural network tailored for autonomous vehicles. We utilized the Udacity simulator to simulate a self-driving car environment and assessed each model's performance based on steering angle predictions. Additionally, we employed a pre-trained VGG16 model using transfer learning on the ImageNet dataset.

## **Behavioral Cloning Techniques**

### Behavioral cloning is a method in machine learning where a model learns to replicate the behavior of a human or expert by observing and mimicking their actions. In the context of autonomous vehicles (AVs), behavioral cloning involves training a neural network to drive by imitating the behavior of a human driver. The model is trained using data collected from human drivers, such as camera images, steering angles, throttle, and brake inputs. Behavior cloning is well-suited for driving a self driving car because it allows the vehicle to learn from human driving behavior and adapt to the complex and dynamic environment on the road. By using behavior cloning, the vehicle can learn to make decisions and react to situations that are difficult to program directly, such as how to handle unexpected obstacles or other road hazards. In this study, a behavior cloning approach is used to develop an autonomous driving system.

The images of 3 positions center, left, right are collected with there corresponding steering angle, speed, throttle etc. And they are fed into the CNN network here we used NVIDIA and VGG16 architecture to do so. Once trained, the network cab generate steering from the video images of a single center camera. The training process involves minimizing the difference between the predicted steering angles and the ground truth steering angles.

### NVIDIA Architecture : The NVIDIA architecture is widely used in autonomous vehicle behavior cloning, employing deep learning to mimic human driving behavior. It processes sensor data, like camera images and vehicle speed, to predict actions such as steering and speed. The architecture has 9 layers: 1 normalization layer, 5 convolutional layers, and 3 fully connected layers. Input images are split into YUV planes and normalized in the first layer, which is optimized for GPU processing. The convolutional layers, empirically designed, extract features with strided convolutions in the first three layers and non-strided ones in the last two for classification or regression tasks.

#### **Transfer learning** is a widely used method in machine learning where knowledge gained from one task is utilized for a different but related task. For autonomous driving, transfer learning can enhance the performance of models by using pre-trained models on extensive image datasets like ImageNet. In this study, we employed VGG16 and VGG19 models, which were pre-trained on the ImageNet dataset, and applied transfer learning techniques. The following figure illustrates the structure of our transfer learning framework.

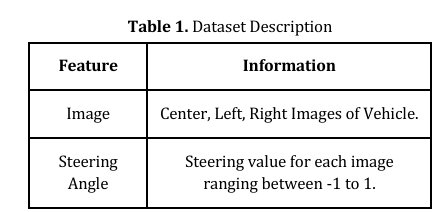
#### **VGG16** is a leading deep convolutional neural network (CNN) model, achieving second place in the ILSVRC (ImageNet) competition in 2014 . Unlike other models from ILSVRC, such as ResNet50 and Inception , the VGG16 model has a relatively lower number of parameters due to its specific arrangement of convolution filters. It uses a 3x3 filter with a stride of 1, followed by a 2x2 max pooling filter with a stride of 2. This pattern of convolution and max pooling is consistently applied throughout the network. The decision layer comprises two fully connected layers, resulting in a total of 138 million parameters. In our approach, we freeze the first four convolution blocks of the VGG16 model and fine-tune the fifth block to predict the appropriate steering

Fig 2. Architecture of vgg16[2]

angle based on the conditions captured in the frames.

## Methodology

In this research, data for training an autonomous vehicle through behavior cloning was gathered using the Udacity simulator. This simulator offers a virtual environment where the vehicle can drive, capturing data from various sensors. The collected data comprises images from three front-facing cameras, as well as records of the steering angle, throttle, and speed. The vehicle was driven in both clockwise and counterclockwise directions for four laps each. During these laps, steering angles, throttle, and speed were continuously recorded. We allocated 80% of the data for training purposes and 20% for testing. The inputs to the model are the images from the three front-facing cameras, and the output is the steering angle, which ranges from -1 to 1. A steering angle of -1 indicates a full left turn, 1 indicates a full right turn, and 0 represents driving straight ahead without any turn.



To enhance the quality and relevance of input data for the model, several image preprocessing methods are employed. These techniques aim to decrease computational complexity and improve both training stability and convergence. Key methods include .An example of a table can be seen in Table I, below.

* Image Cropping: This involves trimming away irrelevant parts of the image that do not contribute to the task.
* Normalization: Adjusting the pixel values to a specific range (such as 0-1) to enhance model training stability and convergence.
* Image Augmentation: Introducing random transformations to the images, like rotations, translations, or flips, which helps the model generalize better to variations in input data.
* YUV Format Conversion: Transforming images into the YUV color space.
* Image Resizing: Reducing the size of images captured by self-driving car cameras to 66x200x3 to lessen the computational load while maintaining essential information for the task.

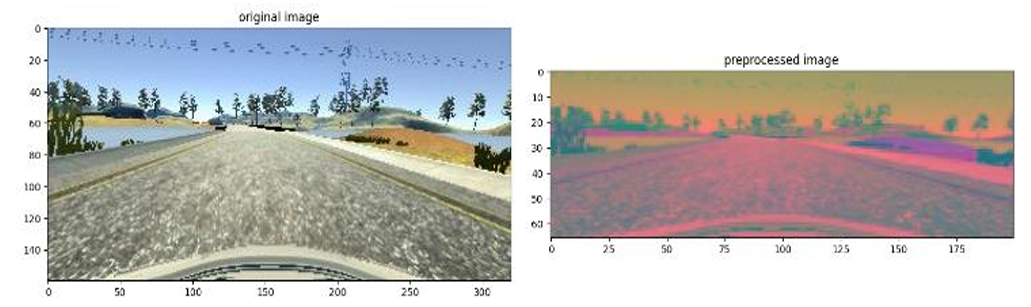


Fig 3. Preprocessed and original images

# Result and disscussions

Our study illustrates the effectiveness of behavior cloning for autonomous driving by applying three distinct deep learning algorithms: NVIDIA and VGG16. We trained these algorithms using a dataset consisting of images and their corresponding steering angles, captured from a human driver under various driving conditions. To assess the performance of each algorithm, we utilized the mean squared error (MSE) metric. This evaluation involved calculating the MSE on a separate test set of images and steering angles that were not included in the training phase.

The mean squared error (MSE) is a standard metric in deep learning for assessing regression models. It quantifies the average squared difference between predicted values and actual values. For autonomous driving, MSE evaluates how well a model predicts steering angles based on images from the car’s cameras. The MSE is determined by averaging the squared differences between predicted and true values. Mathematically, it can be represented as:

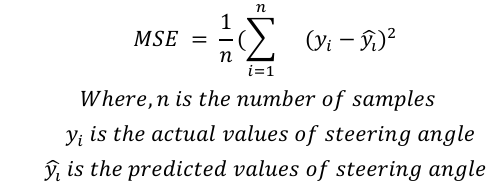


Fig 4. Mean Square Error[2]

The training process was performed for 20 epochs with a batch size of 100. The Adam optimizer was used with a learning rate of 1e-3. Obtained results are shown.

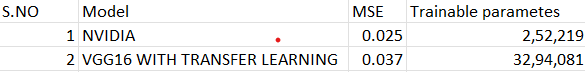


Table 2. MSE Value for each value

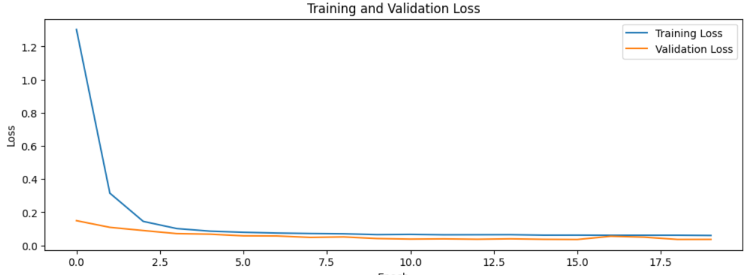


Fig 5.Vgg16 with transfer learning epochs vs loss

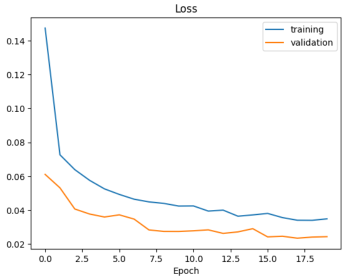


Fig 6. Nvidia Model epochs vs loss

# conclusion

Autonomous vehicles are set to transform transportation. This study examines behavioral cloning in self-driving cars, comparing the NVIDIA and VGG16 architectures with transfer learning. Results show that architecture choice significantly impacts model performance, with the NVIDIA model outperforming VGG16.

* **Optimization**: NVIDIA models are specifically tailored for autonomous driving, making them more suitable than general architectures like VGG16.
* **Efficiency**: NVIDIA models require fewer computational resources and have faster inference times.
* **Advanced Features**: Features like Tensor Cores and hardware optimizations enhance performance.
* **Training**: NVIDIA models benefit from advanced training techniques developed by their research.
* **Seamless Integration**: These models integrate effortlessly with NVIDIA's ecosystem, improving performance.
* **Pre-trained Models**: NVIDIA offers pre-trained models optimized for specific tasks.

Another key direction is the exploration of **multi-sensor data fusion**, combining data from various sensors like radar, lidar, and cameras to enhance the vehicle's perception capabilities. Additionally, research could focus on **real-time processing** and **edge computing** to reduce latency and improve the speed of decision-making in self-driving systems.

Future work can also explore the development of models that generalize better to diverse, unseen road conditions by using **unsupervised and reinforcement learning techniques**. Lastly, improving transfer learning methodologies, by fine-tuning models with larger and more diverse driving datasets, can significantly enhance their performance in real-world scenario

##### Acknowledgments

The authors would like to express their deepest gratitude to Mr. Vivek Kumar Kataria, a respected scientist at DRDO, for his invaluable mentorship and guidance throughout this project. His insights, expertise, and constant support were instrumental in the successful completion of this research. We are profoundly grateful for his time, encouragement, and dedication, which greatly contributed to the advancement of this work.

##### References

1. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to End Learning for Self-Driving Cars. *arXiv preprint arXiv:1604.07316*.
2. Md Wazid Ansari\*1, Sahil Yadav\*2, Nitin\*3, Amit Pal\*4 “COMPARATIVE ANALYSIS OF DEEP LEARNING BEHAVIOR CLONING ALGORITHMs” .DOI : <https://www.doi.org/10.56726/IRJMETS39781>.
3. . Uppala Sumanth, Narinder Singh Punn, Sanjay Kumar Sonbhadra and Sonali Agarwal “Enhanced Behavioral Cloning Based self-driving Car Using Transfer Learning” . arXiv:2007.05740v1 [cs.CV] 11 Jul 2020.