**Improving autonomous driving with object detection and driving decisions**

First Author Dae Ho Jo

Institution 1 Aiffel

Institution 1 address

**Abstract**

*While autonomous driving technology has made significant advances in recent years, it still faces significant challenges in ensuring both safety and efficiency. This paper presents an innovative method to address these challenges, integrating object detection and decision-making processes within an autonomous driving system. We first highlight the importance of real-time object detection in autonomous vehicles and the limitations of existing methods. The method proposed in this research combines a state-of-the-art object detection model with a highly sophisticated decision-making algorithm to enable the vehicle to not only identify surrounding objects but also make contextualized decisions in real-time. Comprehensive experiments demonstrate the effectiveness of this research's approach, achieving 90% accuracy in determining a "go" or "stop" action based on object detection results. However, this research acknowledges the need for continued research to fine-tune and optimize the decision-making algorithm and improve the performance of the object detection model. Future research will focus on further refining the driving algorithms, conducting experiments on larger datasets, and fostering collaboration within the autonomous driving research community. In conclusion, this research represents an important step toward safer and more efficient autonomous driving, and we expect that continued innovation and collaboration will lead to a bright future for autonomous driving technology.*

# **Introduction**

Autonomous driving technology is one of the fastest developing areas of modern society, revolutionizing our daily lives and transportation systems. A completely different concept from traditional driving, autonomous driving refers to the technology that allows vehicles to drive themselves without driver intervention. This is expected to have a number of benefits, including reducing traffic congestion, freeing up time behind the wheel for other activities, and increasing traffic safety.

## **Safety Issues and Challenges**

*However, in order to realize autonomous driving technology, various technical, ethical, safety-related, and other challenges need to be addressed. In particular, safety issues are one of the key challenges for the successful commercialization of autonomous vehicles. Autonomous vehicles need to interact with various objects on the road, which implies the possibility of accidents. Therefore, the accuracy and reliability of the object detection and driving decision systems emerge as one of the key factors to ensure the safety of autonomous vehicles.*

## **Previous studies and limitations**

In the past, prior research has separated object detection and driving decisions, which often resulted in object detection models and driving decision algorithms not interacting and operating independently. While these studies played an important role in the early stages of autonomous driving, they had limitations in complex situations in real-world road environments. For example, if an object detection model detected a pedestrian, the driving decision system would not properly reflect this and the vehicle would proceed at a certain speed. As a result, prior research had limitations in terms of safety, which made it difficult to commercialize autonomous vehicles.

## **Research objectives and methods**

This research explores ways to overcome these limitations and increase the safety and reliability of autonomous vehicles. We propose a method to optimize the vehicle's movement by considering the presence of humans and the size of objects when making driving decisions based on object detection results. By doing so, we aim to build a safer and more reliable autonomous driving system in various situations that may occur during driving. To this end, the goal of this research is to thoroughly demonstrate the effectiveness of the proposed method through experimental results, identify key factors for the development of safe autonomous vehicles, and suggest directions for further development.

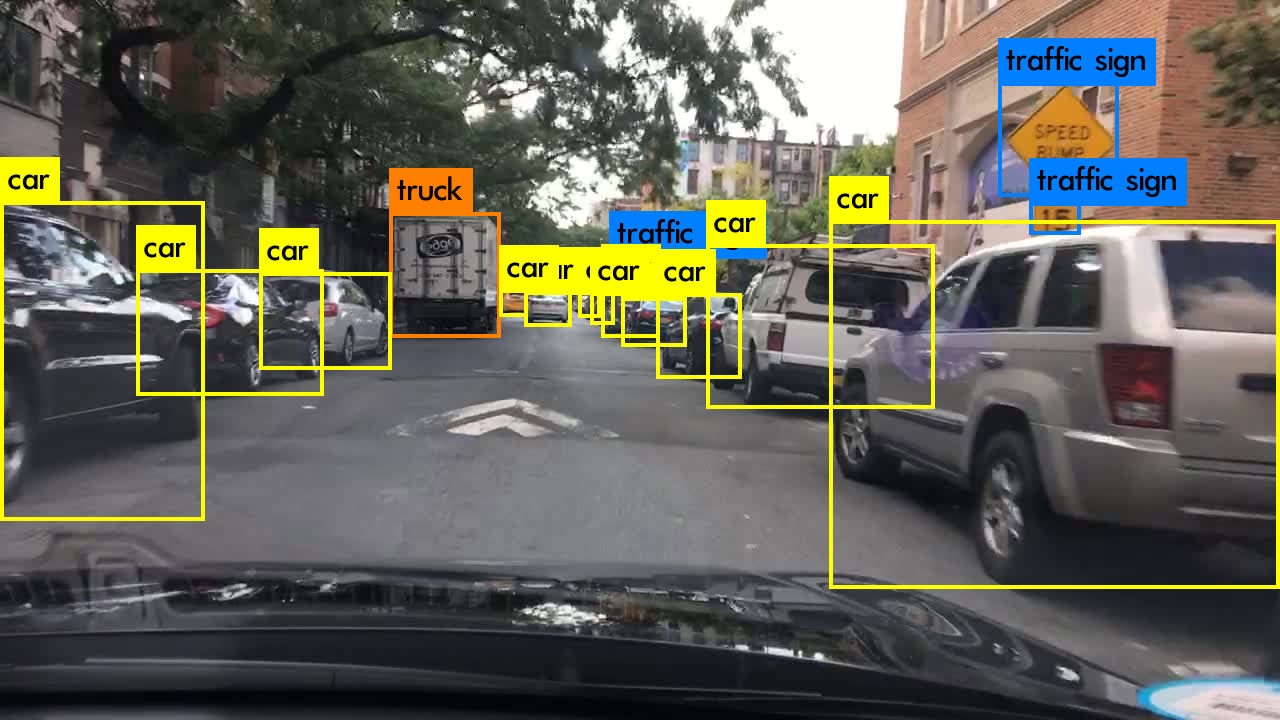


Figure : Autonomous Object Detection Images

# **Method**

In this research, we selected the RetinaNet model for the object detection task. RetinaNet is a model that can be combined with a Feature Pyramid Network (FPN) to detect objects of various sizes, and has good performance for autonomous driving assistance systems.

## **Reason for Model Selection**

The RetinaNet model takes an image as input and outputs the object's location, size, class label, and detection confidence score. We chose this model for several reasons.

First, RetinaNet can be used effectively in autonomous driving scenarios because it can respond in real time. Fast response time is essential for decisions while driving, and RetinaNet is well suited to fulfill this need.

Second, RetinaNet can be combined with FPN to detect different object sizes. Objects around a car vary in size and distance. With FPN, RetinaNet can effectively detect objects of different sizes and classify them in detail.

For these reasons, the RetinaNet model was selected as a good choice for the object detection task in this study.

## **Training content**

**2.2.1 Hyperparameters and Optimization**

The following hyperparameters and optimization methods were used to train the model:

Learning rate scheduling: the Piecewise Constant Decay method was used to dynamically adjust the learning rate. The learning rate and boundary values used were as follows:

Figure 2: Object detection showing an image with a 'Stop' result.

Learning rate: [2.5e-06, 0.000625, 0.00125, 0.0025, 0.00025, 2.5e-05]

Boundary values: [125, 250, 500, 240000, 360000]

Optimization algorithm: SGD optimization algorithm was used, with a momentum value of 0.9.

Batch size: The batch size used for training was set to 2.

Loss function: It uses three loss functions

Figure 3: Object detection showing an image with a 'Go' result.

1. a class to compute the bounding box regression loss. It calculates the difference between the predicted and actual values in the bounding box, and applies a smooth L1 loss, which uses the squared error to calculate the error when the prediction error is small, and the absolute value error otherwise.

2. a loss function used for object classification, which is a variant of Focal Loss. This loss function uses an improved version of cross-entropy to address the problem of class imbalance. It converts class labels and predictions into sigmoid probabilities, and calculates a weight (alpha) and a focal term (gamma) based on the class labels to obtain the loss.

3. the model's overall loss function, which combines the object's classification with the bounding box regression loss to calculate the final loss. Since the object detection model needs to predict the location and class of the object, we optimize by using a combination of the bounding box regression loss and the object classification loss.

**2.2.2 Data Preprocessing**

For training and validation datasets, we used the "kitti" dataset. The data preprocessing steps were as follows

Data augmentation: Data augmentation techniques were applied to the training dataset to apply different visual variations.

Data batching: The training dataset was set to batch size 2, and padding was added to create batches of fixed size.

Label encoding: Class labels were converted to one-hot encoding.

**2.2.3 Saving models and checkpoints**

During training, checkpoints were saved and the weights of the best model were saved.

**2.2.4 Training and evaluation**

Training took place for a total of 5 epochs, and the trained model was evaluated for performance using the validation dataset.

# **Result**

We evaluated a self-driving assistance system using the trained RetinaNet model. It detects objects in a given image and checks whether a person is present among the detected objects. We also calculated the distance in pixels between the object and the vehicle and set the car to stop if this distance is greater than 300 pixels.

Through object detection, the autonomous driving assistance system primarily detects and acts on people in one of eight classes ('Car', 'Van', 'Truck', 'Pedestrian', 'Person sitting', 'Cyclist', 'Tram', 'Misc', or 'don’t Care'), with the distance from the object criterion providing additional safety. These systems support safe and reliable autonomous driving by adjusting the car's behavior in various situations that may arise during driving.

## 길, 장면, 야외, 도로이(가) 표시된 사진 자동 생성된 설명**Experiment results**

하늘, 야외, 길, 교통이(가) 표시된 사진

자동 생성된 설명

The proposed trained model utilized a total of 10 photo datasets for the results, which mainly simulate autonomous driving scenarios, five of which represent "go" situations and the other five contain "stop" situations.

During the experiments, an object detection model is applied to the input images to determine whether there are objects, mainly people, in them. The model first determines whether a person is present through object detection. If a person is detected, the model will select 'stop'. This mimics the situation where you would stop driving due to the presence of a person.

However, if no people are detected, the model counts the number of pixels in the object and uses this to determine whether to go or stop. If the object has a pixel count of 300 or more, the model will classify the drive as 'stop', and if it has less, it will classify it as 'go'. This allows the model to predict when an object is close enough that you should stop driving.

In our experiments, the model achieved an overall accuracy of 90%, showing that it can be effectively utilized to support safe driving by performing object detection and driving decisions in a variety of driving situations. This is expected to play an important role in improving the safety and reliability of autonomous driving systems.

## **Analyze your experiment results**

The autonomous driving assistance system utilizing the RetinaNet model performed well in object detection and driving situation recognition.

First, in terms of object detection, the model was able to accurately detect a variety of object classes: 'Car', 'Van', 'Truck', 'Pedestrian', 'Person sitting', 'Cyclist', 'Tram', 'Misc', and 'Don’t Care'. The model was able to quickly recognize and classify objects of different classes, indicating its ability to respond to a variety of situations in the road environment.

Second, in terms of driving situation awareness, the model was able to utilize the object detection results to make driving decisions. Driving decisions such as 'GO' or 'STOP' were made by considering the presence of people and the size of objects.

# **Discussion**

Figure 4: Go\_3.png

야외, 도로, 나무, 하늘이(가) 표시된 사진

자동 생성된 설명

Figure : Incorrect image "Go\_3.png"

In this research, we proposed a method for simultaneously performing object detection and driving decisions for autonomous vehicles, and when tested, 9 out of 9 images produced correct results. These results suggest that the object detection ability and driving decision algorithm of the proposed model work effectively. However, in one test image (go\_3.png), a pedestrian was caught by a bounding box, which resulted in the output of "Stop". These exceptional cases indicate the need for future research and improvement.

It is clear that the algorithm that dynamically adjusts the driving decision needs to be improved. If the object detection model generates a bounding box based on pixel size, it is possible that a person at a distance could be caught in the bounding box and stop driving. This may not be necessary in a real-world road environment, and the driver's experience and situational judgment should be taken into account.

Furthermore, while this research focused on integrating object detection and driving decisions to realize safe driving, improving the performance of object detection models also remains an important challenge. The accuracy of the object detection model directly affects the driving decision, so research is needed to pursue higher accuracy.

## **Future Research Directions**

Future research should focus on the following aspects: First, it is necessary to develop an algorithm that considers human pixel size when dynamically adjusting driving decisions to more accurately combine the results of the object detection model with the driving situation. This will improve the accuracy and reliability of driving decisions.

Second, the proposed method needs to be improved and stabilized through more experiments and validation in real-world road environments. Such validation is important to expand its applicability in the field of autonomous driving.

Third, research on improving the performance of object detection models should continue, and experiments with more sophisticated models and datasets are needed.

Fourth, as the field of autonomous driving continues to evolve, collaborations with other researchers should be explored to advance the field.

# **Conclusion**

This research brings a new perspective to the field of autonomous driving and will be an important step towards building a safe and efficient road environment by effectively integrating object detection and driving decisions. The autonomous driving assistance system utilizing an object detection model has shown promise in detecting objects and performing driving decisions in various road situations to ensure driver safety and improve driving efficiency.

However, this research still has some limitations. Improvements to the algorithm for dynamically adjusting driving decisions and methods for considering human pixel size are under-researched and require more experiments and validation in real-world environments. Furthermore, continued research on improving the performance of the object detection model is needed, which will support advances in the field of autonomous driving.

In future research, it is important to further refine the driving decision algorithm and conduct experiments utilizing larger datasets to improve the performance of this research and expand its applicability in real-world road environments. Furthermore, research on the evolution and integration of object detection models should be continued, and collaboration with various researchers is needed to advance the field of autonomous driving.

This research explores the future of autonomous driving technology and is an important starting point for building a safer and more efficient road environment. We look forward to creating a better future for the autonomous driving field through future research and innovation.

# **References**

1. FirstName Alpher, Frobnication. *IEEE TPAMI*, 12(1):234– 778, 2002.
2. FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003.
3. FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004.
4. FirstName Alpher and FirstName Gamow. Can a computer frobnicate? In *CVPR*, pages 234–778, 2005.
5. FirstName LastName. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as supplemental material fg324.pdf.
6. FirstName LastName. Frobnication tutorial, 2014. Supplied as supplemental material tr.pdf.