CS333 Mini Project

Group 11 - Object Detection in Autonomous Vehicles

September 21, 2021

Contents

1 Introduction

The rapid advancement of artificial intelligence has significantly aided the advancement of unmanned driving. Self-driving technology has been under development for the past couple of decades, and has considerable potential to reshape traditional transportation systems. They have been increasingly popular among them due to its significant economic and social impact [7].

As shown in a recent World Health Organization statistics report, 1.35 million people die each year as a result of road traffic accidents all around the world. Among them, 94 percent of vehicle accidents are attributed to human drivers' irresponsibility, such as breaking the speed limit and driving while intoxicated. Self-driving technology is seen as a solution as well as a cost-effective method of transportation for customers. Self-driving technology has been in development for several decades and has the potential to drastically transform current transportation infrastructure.[7]

Image recognition is one of the topics that has been quickly developing and advancing with the use of machine learning to identify impediments and signs in the surroundings, and we will dive in-depth with object categorization and object identification algorithms. [4] In this report, We mention the Convolution Neural Network (CNN) and You Only Look Once (YOLO) Algorithm due to their comparative strong learning ability and desired features. And finally, we discuss the issues and regulation of Self-driving Vehicles.

2 Context

Self-driving automobiles are becoming more popular in terms of market demand and commercial benefit. More and more businesses and scientific research institutes, like Google, Tesla, Apple, Nissan, and Audi, etc. are investing in the field. In addition to the enterprises mentioned above, increasing amounts of Internet and car companies around the world have recently focused on the self-driving car industry.[5]

The four main components of self-driving technology are sensing, perception, planning, and control. Sensing refers to a self-driving vehicle's capacity to capture real-time data from a wide range of sensors (such as cameras, ultrasonic sonars, LiDARs, and radars) in order to extract usable data from the surroundings. [7]

Information supplied by the sensing module, the perception module builds a situational awareness of the surroundings. Through the acquisition of relevant external data, the environment perception system uses existing knowledge of the area to create an environmental model that includes impediments, road structures and signs. [7]The environment perception system's primary purpose is to implement operations like lane detection, traffic signal detection, and obstacle identification using devices such as cameras and optical radars. [5]

Deep learning has a strong capacity to adapt with various issues in the field of self-driving cars since it can learn comprehensive and efficient feature representation through layer-by-layer feature transformation of the original signal automatically. [5]

3 Algorithms

We will first discuss convolutional neural networks in general, then talk about You Only Look Once, an algorithm built on a convolutional neural net.

3.1 Convolution Neural Networks:

3.1.1 Overview

In this section, we will give a brief overview of convolutional neural networks (CNNs) in the field of deep learning. CNNs are generally considered state of the art for image classification by accuracy. The general intuition for CNNs is that they capture features present in pixels of the images. The image is broken into blocks and similarity is measured by performing convolution with kernels. The kernels can be thought of as filters. This process is repeated over and over and max pooling (taking the maximum value in each block) is also performed to reduce the dimensionality of the image. The image is thus transformed from a pixel space to a feature space in this fashion. After transforming the image into a feature space, a standard neural network can be applied or also any other machine learning classification algorithm as well to produce an output vector. [9]

3.1.2 Training

CNNs are trained by updating the parameters of the network such as the aforementioned kernel and the hyperparameters of the neural network that is run after image transformation. Such hyperparameters are generally trained using backpropagation. [3] Backpropagation basically takes the error from the model output and uses that to adjust the activation function on each individual neuron. It does this by basically using the chain rule from calculus and calculating the gradient of the error with respect to each neuron. [9] This reduces error in the network. [3]

3.1.3 Applicability to Self Driving Cars Image Classification

Convolutional neural networks have been used for various tasks in the algorithms for self-driving cars including object detection, semantic segmentation, end to end vehicle control. We will now describe each in turn.

In object detection, we are using the CNN to detect objects present in the image like pedestrians and other cars. We perform this by using a classification based neural net or other algorithm after our convolutions that map an image into a feature space. We also keep track of where the features are in the image as well by dividing the picture into grids and scanning it that way.

Semantic segmentation basically segregates the image into its requisite components such as people, buildings, or roads. It accomplishes this by repeatedly applying convolutions to learn similarities in the picture, and outputting this back on the original image size in the form of a probability map.

Now, neural networks have been combined with the CNNs to achieve end to end car control. Here, the neural network outputs a throttle and steering input in order to control the car. It does this by combining data from the aforementioned CNN process with other sensor data from the car. This could be LIDAR or other sensor. Current work is trying to reduce the necessity of the other sensors, and simply infer the direction/steering of the car simply from the image, but there is still work to be done on that front. [3]

3.2 You Only Look Once (YOLO) Algorithm

3.2.1 Overview

You Only Look Once (YOLO) is an image recognition algorithm used across many self driving companies because of its effectiveness in comparison to other image recognition algorithms; YOLO is one of the few algorithms capable of real time detection with minimal error. The algorithm was initially introduced in a research paper by Joseph Redmon; the key element is that it is a regressive algorithm versus the usual classification type of algorithm used for image recognition. [8]

3.2.2 Performance

A typical image recognition algorithm, such as R-CNN will generate image bounding boxes, run a classifier on the boxes, refine the output, and rescore boxes based on surrounding boxes. This is a very long process that does not hold up well in real-time object detection, especially at the speeds that self-driving systems require. Meanwhile, YOLO just uses a single convolutional network (discussed above) that predicts bounding boxes and class probabilities all at once. Thus, the name You Only Look Once. YOLO is very performant; as Redmon highlights "YOLO model processes images in real-time at 45 frames per second. A smaller version of the network, Fast YOLO, processes an astounding 155 frames per second while still achieving double the mAP of other real-time detectors"[8]. The pipeline process for YOLO is very simple, allowing for blistering speeds. It is critical for vehicles moving at high speeds to be able to quickly detect surrounding objects in real-time: YOLO is perfect for this. Another advantage of YOLO is that it views the entire image at once, allowing for instant contextual information that aids in classification. [8]

3.2.3 How YOLO Works

YOLO operates using a single neural network, reasoning globally about the image and all the objects contained within. The input image is divided into an SxS grid; each cell predicts bounding boxes and confidence scores for the boxes where the score indicates how confident the algorithm is that the box contains an object and how accurate the box is itself. A bounding box consists of 5 attributes: x, y, w, h, and confidence. X and y represent the center of the box relative to the bounds of the cell, w and h are coordinates relative to the whole image, and confidence is the intersection over union between the predicted and ground truth box. A class probability map is also created, allowing for the final classification of a bounding box.[8]

3.3 Neural Networks and Adversarial Examples

Now we will briefly talk about an issue with convolutional neural networks, namely their susceptibility to adversarial attacks. Adversarial attacks are defined as an input x' that is semantically similar to the original input x but is misclassified or basically $f(x) \neq f(x')$. We note this is problematic as a stop sign could be misinterpreted as a speed limit sign, causing a car to speed up when it should slow down. [1] While research has main been on simpler image classification tasks, researchers have shown that adversarial examples are able to be discovered against state of the art neural networks for all types of networks in both white-box and black-box attacks where the attacker has varying information about the neural network. [2] This susceptibility is due to the highly nonlinear decision boundaries of neural networks and the fact that they are non-interpretable black-box algorithms. [1] Thus, the comments made in [3] about strengthening neural networks and making them more interpretable to a human user is highly relevant.

4 Issues/Ethics

4.1 The Regulation of Self-driving Vehicles

Since the inception of the autonomous vehicle, new rules and standards had to be created in order to accommodate the new technology. In 2014, the Society of Automative Engineers established a standard for assessing the level that a self-driving car can have. This assessment was a scale that would judge the autonomy of a car, with 0 being a car with full driver control and no assistance and 5 being full automation with no human intervention. [11] In the years that followed, with the use of automated cars becoming more widely tested and eventually released to the public, more statistics and regulations began to emerge. The first driver death caused by a crashing self-driving car occurred in 2016 and the first pedestrian death by a self-driving car occurred in 2018. [6] The issue that arises in situations like these is that there are several factors that a self-driving car is not able to precalculate with absolute certainty even when all the cameras/sensors and computational power are at peak performance and condition. This is due to factors like the path the car is on, environmental conditions, the driver's condition, etc. Although these facts and statistics may seem alarming, it is important to keep in mind that most accidents that the Google car experienced in 2.3 million miles were caused by other drivers for non-autonomous vehicles. [11]

4.2 Trolley Dilemma

The well-known philosophical scenario known as the "Trolley Dilemma" is a major point of concern when discussing the ethics of self-driving cars. The dilemma is as follows: a group of five people are tied onto a railroad track. A second connected railroad track has only one person tied to a railroad track. Is it morally right to switch an oncoming train to hit and kill just the one person instead of killing all five? Is it better to kill an elderly person or a child? Save the passenger or the pedestrian? The programmers for self-driving cars have to confront and decide uncomfortable ethical scenarios like this when they preprogram crash optimization algorithms. Programmers are thus confronted with two main problems: how to program self-driving cars in advance as well as ethical research on situations like the Trolley Dilemma so that the car can make the "ethically right" decision every time. This would consist of instructions for the self-driving car on how to act in predictable situations as well as set guiding rules for unpredictable situations. [11]

4.3 Legal Issues with Driving Under the Influence

In December of 2018, a man was arrested for sleeping in a self-driving car driving at 70 mph on a Californian highway. In California as well as many other states, a driver can be charged with a DUI if they are in "actual physical control of the vehicle" while intoxicated. The phrase "actual physical control" pertains to one or more of the following being evident: the person is in the driver's seat, the keys are in the ignition, the engine is running, the tires are warm, or the vehicle is parked on the side of the road/highway. This is a reasonable standard for human drivers in normal cars, but not necessarily for self-driving cars. If a self-driving car is working as intended, law enforcement technically has no reason to stop a car or suspect driving under the influence unless there is a sobriety checkpoint/the officer specifically sees the driver drunkenly stumble out of the car. This legal loophole has led some auto companies to consider/implement sensors that can detect whether the driver is incapacitated or not in order to keep drunk drivers off the road. However, others consider self-driving cars to be the solution to drunk drivers rather than being an issue. Legislation and technologies are still in a relatively infant stage, so the current outlook on this specific topic is generally cautious but optimistic.

5 Individual Contributions

Billy - I researched and wrote sections 3.1 and 3.3 and the associated slides with those sections.
Nathan - I researched and wrote section 4 and the corresponding slides.

Ahmad - I researched and wrote section 3.2 and its associated slides.
Liuren - I researched and wrote section 1 and 2 and the associated slides.

6 References

References

[1] Carlini, N., & Wagner, D. (2017, May). Towards evaluating the robustness of neural networks. In 2017 ieee symposium on security and privacy (sp) (pp. 39-57). IEEE.

[2] Chen, J., Jordan, M. I., & Wainwright, M. J. (2020, May). Hopskipjumpattack: A query-efficient decision-based attack. In 2020 ieee symposium on security and privacy (sp) (pp. 1277-1294). IEEE.

[3] Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. IATSS research, 43(4), 244-252.

[4] Kumar, A. (2020). Self-driving care using a simulator, International Journal of Advanced Research in Computer Science, 11, 158-162.

[5] Ni, J., & Chen, Y., & Zhu, J. (2020). A survey on theories and applications for self-driving cars based on deep learning methods. Applied Sciences, 10(8), 2749.

[6] Nyholm, Sven. The ethics of crashes with self-driving cars: A roadmap, II. Philosophy Compass 2018; e12506.

[7] Rasib, M. (2021). Are self-driving vehicles ready to launch? an insight into steering control in autonomous self-driving vehicles. Mathematical Problems in Engineering, 2021

[8] Redmon, J., & Divvala, S. (2016). You Only Look Once: Unified, Real-Time Object Detection. CVPR, (pp. 779-788).

[9] Rudin, Cynthia. Intuition for the Algorithms of Machine Learning, Self-pub, eBook, 2020

[10] Wilfert, Jarrod. Can You Get a DUI While Operating a Self-Driving Vehicle?, Wilfert Law P.C., 2018.

[11] Wolkenstein, A. What has the Trolley Dilemma ever done for us (and what will it do in the future)? On some recent debates about the ethics of self-driving cars. Ethics Inf Technol 20, 163–173 (2018).