Object Motion prediction in Autonomous Driving

Abstract

Self-driving technology has become one of the hottest fields during 2020. As of today's record, Billions of Dollars are invested into self-driving research. Hundreds of startups are working hard to find their way out in the self-driving industry. Here are some current record of self-driving technology. "In 2014, Waymo revealed a prototype of a driverless car without any steering wheel, gas pedal or brake pedal, thereby being 100 percent autonomous. By the end of last year, more than 2 million miles had been driven by Google's autonomous car."[1] Another catching eye news related to autonomous driving is their fundraising. Cruise has raised 1.2B during 2019, it is one of their continuous fundraising since 2014. Waymo has raised 2.25 billion in their first ever fundraising in 2020[3]. Autonomous driving companies are raising a lot of money to speed up their research progress, and trying to bring their technology into mass production as early as possible.

Object Motion prediction is one of the hardest research fields in Autonomous Driving. Because of different environments like time of day, season of the year, speed limits, and traffic congestion level different people will have different behavior on their motion planning. They will also react differently based on the environment change. Some of the existing algorithms are background subtraction, template matching and Speeded Up Robust Features[4]. But none of them are good enough to become industry standards. The core of study autonomous driving is to continuously develop learning models and algorithms to reach the highest prediction/detection correctness. So people can take the ride from autonomous driving vehicles with the highest safety score.

Our team is planning to use machine learning technology on a large size of dataset to train the model. After comparing with different newly created models and algorithms, Bring out one good solution for object motion prediction in autonomous driving that suits for most autonomous driving systems. We started with learning existing research papers, try to recreate the experiment they mentioned in the paper. For most Cases, the recreate process is very hard. Most of them will fail due to data access limitation, detailed steps missing, and account authorization. We are able to find some open sourced experiments with the help of our project professor. We are also planning to build a model that can call out the correctness of the motion prediction. In use to evaluate the different models and algorithms we generated. It is a very important step for us, For most cases, the differences with different models and algorithms are very low. They can be creating similar percentages of correctness on path prediction. With the large number of models and algorithms we need to test, It is crucial to have an automated testing logic to identify the correctness of the path prediction model we generated.

Our team has faced some issues with data preprocessing. While we are working on 3 of the available object detection models. The ideal way is to use the same dataset. So we can compare the results from 3 different testing models. However we found a large amount of data

preprocessing work requested. As this semester, we are not requested to do a large amount of coding to solve this issue. The professor is more interested in whether we are able to bring 3 models online. So we skipped the data sync process to save some time to bring 3 research models online.

I. **Introduction**

Automotives have been widely used as a transportation method in the world. Some of the companies used autonomous driving as an extension of vehicle cruise function. For me, Cruise helped me a lot while I drove on the freeway. It does not require a human step on the gas pedal to keep the vehicle moving speed. It helps humans drive without an uncomfortable position. It is a key function to allow humans to drive long distances on a quiet freeway. The Idealed L5 autonomous driving are looking to allow human drive vehicles without operating on them in both freeway and local environment.

Object motion prediction is to track or predict the moving direction of large and small objects. It can be predicted using physics or math models. It also can be predicted using machine learning models or multiple models combined. In Autonomous driving field, object motion predictions are tracking and predicting the movements of objects surrounding the vehicles. It can be traffic lights, pedestrians, other vehicles and random moving in objects. Autonomous driving vehicles are required to not only follow the standard traffic rules but also react based on the change of surroundings. To allow maximum level of safety on the road.

In autonomous driving Technology, people often use the correctness of identification level on the images to measure the tech level of autonomous driving. However we can also use the safety level of the vehicle to measure the level of it. Most of the autonomous driving companies or vehicles can reach the safe level of 99%. Only the top 5 companies can achieve beyond that level to 99.9% - 99.99999%. Thus those companies can be worth billions of dollars. Only them may get approved by the government and allow their vehicle drive on the road and used for operation business. Governments are making sure autonomous driving is much safer than regular human driving before they can be used. Since the safety level is the most important parameter for the success of autonomous driving, object motion prediction is very important in Autonomous driving Technology.

I am writing this paper myself to describe the study of object motion prediction. However we have 4 people from San Jose State University who formed a team for CMPE 295 Final Project. We are planning to research the state of the art of this study during spring 2020, attend one of the machine learning competitions during Summer 2020, and finish the improved object motion prediction technology during Fall 2020. Some major milestones including: 1. success form a team by 15 Feb 2020. (exact 4 engineers are taking CMPE 295A during Fall and CMPE 295B during Fall, one qualified research professor, and one good research idea) 2. Finish research all the state-of-the-art technology by 15April 2020. 3. Complete all Lab related practise on the state of art open source demo by 1 May 2020. 4. Finish first 2 chapters of research paper by 20 May 2020. 5. Complete Kargo machine learning competition by 15 July 2020 6. Finish improved object motion prediction lab by 1 Oct 2020. 7. Last, Finish the final research paper by 1 Dec 2020. Before we start the research, I have taken the CMPE 297 baidu autonomous driving class provided by SJSU. Something we could do after the research is, join a company work on autonomous driving technology, start a company to continue work on the technology, or apply for the patent for future use.

Currently, Autonomous Driving Vehicles are installed with many sensors. Computers get data from sensors and process them into vehicle movement or reaction commands. Some of the processing sensors include: Ladar, Lidar, and Camera. Most common technologies are using combined sensor data and map data. Tesla is using an Ultrasonic sensor instead of a Ladar sensor to reduce the cost of Autonomous Driving Vehicles. However the most common technologies are proving better accuracy with $500,000 plus in cost for average Autonomous Driving Vehicles. One of the cool Autonomous Driving Company Aptiv has succeeded in completing 100,000 Rides in Las Vegas ordered by regular customer order from Taxi App Lyft. It is level 4-5 autonomous driving technology, driven without human interactions. My research focus will impact all level Autonomous Driving technology.

II. **Literature/State-of-the-Art Review**

A. Introduce related components like LiDAR, RADAR, Camera and GPS/INS

There are many sensors used in Autonomous Driving Including LiDAR, RADAR, Camera and GPS/INS. LiDar Sensors are the most reliable sensor but also very expensive to use. It produces a huge amount of data, a perception algorithm cannot be carried out in real-time without simplifying the sensor information. Unlike prior works, it also can be done by using hybrid ground classification and the Region of Interest (ROI) identification method in order to filter out the amount of unwanted raw data for the actual tracking. And the environment is also abstracted based on an occupancy grid map. [5]

Radar normally can provide reliable information on straight lanes, but fail in curves due to their restricted field of view. On the other hand, Lidar sensors are able to cover the regions of interest in almost all situations, but do not provide precise speed information. We can combine the advantages of both sensors with a sensor fusion approach in order to provide permanent and precise spatial and dynamical data. The results in highway experiments with real traffic are looking great. [6]

By using Camera sensor, we can provide RGB and Pixel data. Object recognition known as Decision Tree and Decision Fusion based Recognition System (D2TFRS) for autonomous driving. We can combine two separate feature sets, which are RGB pixel values and spatial points X, Y of each pixel to form our dataset. The D2TFRS is based on the intuition that reclassification of pre-identified misclassified objects in a driving environment can give better prediction accuracy. D2TFRS can outperform AdaBoost with respect to speed. D2TFRS can have better parallelization performance. [7]

B. State-of-the-Art: talk about the algorithm will be developed during the research

During the reasearch our main job is to develop a customized algorithm using combined sensor data provide a more accurate object motion prediction. We will based from YOLO algorithm. It has been growing rapidly recently. Deep Learning is the-state-of-art algorithm for object detection in many situations. In addition, the YOLO algorithm is a good way to handle object detection for Autonomous driving. It has a fast object detection rate (10-50 fps) to handle the fast changes in the surroundings during driving. It also learns features from the continuous information of videos.

My research topic is motion predictions in Autonomous Driving. It doesn't have as much research as object detection, but it is one of the critical modules for Autonomous Driving. It directly connects with the level of safety and optimization in Autonomous Vehicles. I learned that there are many existing theories related to object control, but those theories are not being explored for Autonomous Driving. There will be more findings for this topic as Autonomous Driving companies improve the safety and optimization level of their vehicles.

III. Project Justification

In this project, we are going to do research on the latest object detection and trajectory prediction algorithms and implement an approach to perform motion prediction in autonomous driving. Nowadays, autonomous driving is one of the greatest ambitions in the technology field since it's a revolutionary technology that can change the future. If fully autonomous driving is achieved, benefits will include less traffic collision, less traffic congestion, less fuel consumption and less carbon dioxide emission. Autonomous vehicles can not only improve people's quality of life by saving lots of time spent on commuting, but also help to protect our environment because they are more eco-friendly than traditional cars. The market of autonomous vehicles is expected to expand rapidly. Therefore, companies are competing with each other to seize the market. Without doubt, the first batch of companies that can achieve fully autonomous driving are going to make huge profits.

While technology is developing fast, safety is still the major concern for autonomous vehicles. Driving is a complex task that involves various kinds of scenarios. For example, there are several vehicles and pedestrians at a four way stop intersection. In this scenario, it's essential for the autonomous vehicle to detect other objects and people, and predict their moving paths in order to find a safe path to avoid collision. However, since human behaviors cannot be analyzed using a set of rules, we need to use deep learning based algorithms to make predictions.

There are many approaches that have been developed to make path prediction. However, accuracy is not good enough for fully autonomous vehicles. The purpose of our project is to build our prediction model and improve accuracy. We will use KITTI datasets as our training dataset. Then we will build our model using state-of-the-art techniques such as You Only Look Once(YOLO), Convolutional Neural Network(CNN) and Extended Kalman Filter Algorithm. Finally, our model can use data coming from sensors as input, and make path predictions based on the input data.

IV. **Statement of the Need**

A. The safety level of autonomous driving vehicles

In order to put autonomous driving vehicles on the road, we need to ensure the safety of the vehicle. The drive behavior should at least not damage any of the following elements: the other vehicles on the road, any objects on the road and the passengers in the vehicles. Most of the states or regions in the US have published guidelines for autonomous vehicles. Government and other societies are putting the safety level into the most important factor to consider autonomous driving vehicles on the road.

The average U.S. driver has one accident roughly every 165,000 miles. Here's how to get that figure: the average mileage per year is 16,550, according to the federal highway administration; the average length of time it goes between traffic accidents is 10 years, according to Allstate. (In particularly safe cities such as Fort Collins, Colo., that number can rise to 14 years). However Google announced earlier that its self-driving cars have completed 300,000 miles of test-drives, under a "wide range of conditions," all without any kind of accident. (The project has seen a few accidents in the past - but only with humans at the wheel.)[8]

*B. The response speed of autonomous driving vehicle*s

While we drive on the road, passengers can feel differently relevant to a lot of factors: the road type, the quality of the road, the weather, and the vehicles etc. A lot of passengers have complained that autonomous driving makes them feel uncomfortable while driving. It made wide turns, bouncing left and right, or didn't break in time and cause accidents etc. All of the above situations can be caused by the response speed of autonomous driving.

For example, the vehicle sensors gathered all the data from the environment and passed it to the computer to process. Computers then use algorithms to convert the data into drive command. If the turn left command got delayed by 0.5 second due to the large volume of data and algorithm process, with the speed of 20 kmh it can go over 2.8 meters. Which is half of the vehicle length, it will cause an accident or cleared validation on traffic rule. Thus we have to ensure the response time is smaller than 0.1 second. Sometimes we can use algorithms to send out predicted commands to improve this situation.

C. The large test need in Autonomous Driving

Automotive has been used for more than 100 years. Compared with it, Autonomous Driving is only tested for less than 10 years. Based on the tested mileage, It can be less than 1 over 100,000 percent of all the mileage drove by regular vehicles. In order for Autonomous Driving Technology can be accepted by the world. It needs to be test drived millions miles and tested in all different situations before it can be fully trusted by the government, passengers, and producers.

In other words, the amount of autonomous driving companies are still very small compared with regular vehicle producers. It makes it difficult to be tested in a large number of miles. In recent 5 years, a lot of companies have made some of their code open source. Largely reduced the minimum requirement to allow companies to have their own autonomous driving vehicles. However since most of the technology is using Lidar sensors which is very expensive compared to the vehicle price. Made the autonomous driving vehicles have large test data to be difficult to accomplish.

V. **Goals of the Project**

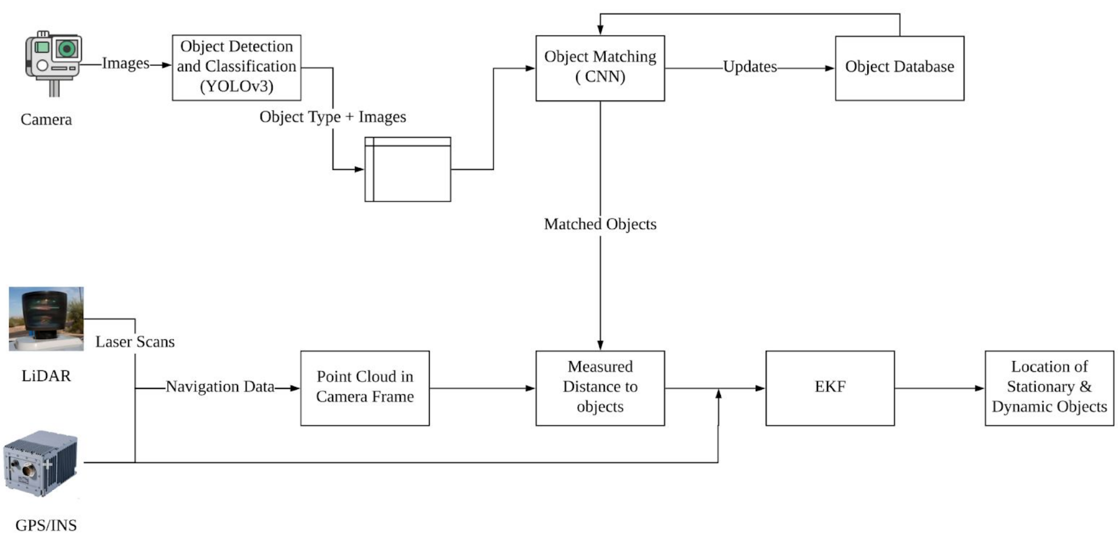
A. An Improved algorithm provide prediction faster and better

In autonomous driving and other software fields, algorithms are being widely used. It is the way developers define/require computers to process certain data. It can create a large user experience gap. One of the common examples is search engines. Google search as the core product of google, The UI can simply be an input box and provide service to billions of users on the world. The fast and accurate search speed is generated by their customized algorithm on a large dataset. Similar companies like Yahoo are having the same functions, but their algorithms are far behind Google. They have to file a bankruptcy due to failed user experience or product.

Most of the algorithms in autonomous driving can reach the correct/safe level of 99%. But only top 5 companies can push the level to above 99.99% and above. In the importance of safety in autonomous driving vehicles, 99% can be considered as failure. Government and society will not allow surrounding vehicles to have 1% incorrect decision making. It will bring danger to other people.

VI. Technical Aspect

A. Project Architecture



1). One of the most famous datasets, the KITTI datasets, are used in this project which contain the data from different sensors (LIDAR, RADAR, Camera and GPS/INS) and python development kit to assist with data preprocessing.

2). Detecting and classifying static and dynamic objects by using state-of-the-art network YOLOv3.

3). Projecting 3D point clouds into images which are taken by camera in KITTI dataset by using python development kit.

4). Matching the same object feature from one image frame to another by implementing the algorithm from Convolutional Neural Network (CNN).

5). Predicting the objects next motion around the autonomous vehicle by fusing the above information with GPS/INS data using Extended Kalman Filter algorithm

B. Evaluation Technology

The evaluation methodology is established on the basis of implementation of object tracking and motion forecasting in Autonomous Driving by using Argoverse API. The algorithm developed in Argoverse is based on the datasets: HD Map, Argoverse Tracking Dataset and Argoverse Forecasting Dataset. The Argoverse detailed map includes Vector Map of Lane Geometry, Rasterized Driveable Area Map and Rasterized Ground Height Map. The baseline experiments have been implemented by installing argoverse module and mayavi under Python 3 environment and running on Jupyter Notebook. The argoverse provides both the full dataset and the sample version of the dataset for testing purposes.

C. Implementation and Evaluation of Object Tracking using Argoverse-Tracking Dataset

The Argoverse-Tracking dataset provides track annotations and raw data from camera (@30hz) and lidar sensors (@10hz) as well as two stereo cameras (@5hz). A total 113 scenes/logs are released, separated into 65 logs for training, 24 logs for validating, and 24 logs for testing. The training data is separated into smaller files and 1 log is provided as sample data.

We leverage standard evaluation metrics commonly used for multiple object tracking (MOT). The MOT metric relies on a distance/similarity metric between ground truth and predicted objects to determine an optimal assignment. Once an assignment is made, we use three distance metrics for MOTP: MOTP-D (centroid distance), MOTP-O (orientation error), and MOTP-I (Intersection-over-Union error). MOTP-D is computed by the 3D bounding box centroid distance between associated tracker output and ground truth, which is also used in MOTA as detection association range. Our threshold for "missed" tracks is 2 meters, which is half of the average family car length in the US. MOTP-O is the smallest angle difference about the z (vertical) axis such that the front/back object orientation is ignored, and MOTP-I is the amodal shape estimation error, computed by the 1 - IoU of 3D bounding box after aligning orientation and centroid. For all three MOTP scores, lower scores indicate higher accuracy.

In our experiments, we run our tracker over the 24 logs in the Argoverse 3D Tracking test set. We are also interested in the relationship between tracking performance and distance. We apply a threshold (30, 50, 100 m) to the distance between vehicles and our ego-vehicle and only evaluate annotations and tracker output within that range. The results show that our baseline tracker performs well at short range where the LiDAR sampling density is higher, but struggles for objects beyond 50 m.

Our baseline tracking pipeline clusters LiDAR returns in driveable region to detect potential objects, uses Mask R-CNN to prune non-vehicle LiDAR returns, associates clusters over time using nearest neighbor and the Hungarian algorithm, estimates transformations between clusters with iterative closest point (ICP), and estimates vehicle pose with a classical Kalman Filter using constant velocity motion model.

VII. Conclusion

Safety level is critical for Autonomous Driving technology. We work hard to develop something better than current object motion prediction algorithms to improve the safety level of autonomous driving. By using an enhanced YOLO algorithm, it will reduce the time we need to recognize the objects path and predict its movement. Imagine one day autonomous driving processors can recognize objects from an image within 0.01 second, Also it can predict the movement of surrounding objects correctly. It will become the preferred transportation method due to its safety level and convenience.

If everyone can save 1 hr per day for changing daily transportation methods to autonomous driving. As a whole, It will help increase the country's GDP by 5-10%, and bring more happiness to society. Some other possible solutions to improve object motion prediction are use a better processing speed processor (10x-100x), use more accurate and react faster sensors, or improve the speed of internet support during driving to get better described map data. In future, researchers can use any better technology to improve the correctness of object motion prediction. And hoping to see Autonomous Driving technology can be widely used within 20 years.

VIII. Reference

1. Conrad A. Buchler, Jr., “Where We’re Going, We Don’t Need Drivers,” 19 Loy. J. Pub. Int. L. 1 (2017)
2. Christian Laugier. Impact of AI on Autonomous Driving. *WRC 2019 - WRC 2019- IEEE World Robot Conference*, Aug 2019, Beijing, China. pp.1-27.
3. K. Korosec, “Waymo brings in $2.25 billion from outside investors, Alphabet,” TechCrunch, 02-Mar-2020. [Online]. Available: https://techcrunch.com/2020/03/02/waymo-brings-in-2-25-billion-from-outside-investors-alphabet/. [Accessed: 10-Mar-2020].
4. A. Raj, A. Sivaraman, C. Bhowmick and N. K. Verma, "Object tracking with movement prediction algorithms," *2016 11th International Conference on Industrial and Information Systems (ICIIS)*, Roorkee, 2016, pp. 285-290.
5. J. Choi, S. Ulbrich, B. Lichte and M. Maurer, "Multi-Target Tracking using a 3D-Lidar sensor for autonomous vehicles," *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, The Hague, 2013, pp. 881-886, doi: 10.1109/ITSC.2013.6728343.
6. D. Göhring, M. Wang, M. Schnürmacher and T. Ganjineh, "Radar/Lidar sensor fusion for car-following on highways," *The 5th International Conference on Automation, Robotics and Applications*, Wellington, 2011, pp. 407-412, doi: 10.1109/ICARA.2011.6144918.
7. Alam F., Mehmood R., Katib I. (2018) D2TFRS: An Object Recognition Method for Autonomous Vehicles Based on RGB and Spatial Values of Pixels. In: Mehmood R., Bhaduri B., Katib I., Chlamtac I. (eds) Smart Societies, Infrastructure, Technologies and Applications. SCITA 2017. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 224. Springer, Cham
8. Taylor, C., 2020. *Google's Driverless Car Is Now Safer Than The Average Driver*. [online] Mashable. Available at: <https://mashable.com/2012/08/07/google-driverless-cars-safer-than-you/>
9. P. J. Glavine, O. De Silva, G. Mann, and R. Gosine, “GPS Integrated Inertial Navigation System Using Interactive Multiple Model Extended Kalman Filtering,” 2018 Moratuwa Engineering Research Conference (MERCon), May 2018.

10. Y. Maalej, S. Sorour, A. Abdel-Rahim, and M. Guizani, “Tracking 3D LIDAR Point Clouds Using Extended Kalman Filters in KITTI Driving Sequences,” 2018 IEEE Global Communications Conference (GLOBECOM), Dec. 2018.
11. A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The KITTI dataset,” The International Journal of Robotics Research, vol. 32, no. 11, pp. 1231–1237, Aug. 2013.
12. M.-F. Chang, D. Ramanan, J. Hays, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, and S. Lucey, “Argoverse: 3D Tracking and Forecasting With Rich Maps,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2019.
13. M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz, and H. M. Gross. Complexer-YOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds. CVPRW, 2019.