1. Team Name: Deep3D
2. Is this a Facebook project? No
3. Project Title: 3D Deep Learning on Point Cloud Data Representation
4. Project summary (4-5+ sentences). Robot perception and augmented reality require more accurate and efficient 3D models for object classification, part segmentation and semantic segmentation. 3D deep learning research has practical values for these emerging applications. Point cloud as one representation of 3D data structure has many unique advantages compared to others. We’re interested in doing a toy project in the area of 3D deep learning.
5. What you will do (Approach, 4-5+ sentences) - We’ll build the project based on the idea of Pointnet proposed by Stanford researchers. Each of the team members will cover different areas such as object classification, part segmentation, semantic segmentation, and data visualization. We’ll compare the Pointnet architecture based on point cloud data presentation with panorama and rotation net based on multi-view data representation and explore potential orthogonal approaches between the two.
6. Resources / Related Work & Papers (4-5+ sentences).
   1. [https://arxiv.org/pdf/1905.12683.pdf (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Farxiv.org%2Fpdf%2F1905.12683.pdf&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738557312%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=FkNbiXcVmq8SByPn0Mz6fTTNzgL7CAcOGaK22FAjRq8%3D&reserved=0)
   2. [https://arxiv.org/pdf/1808.01462.pdf (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Farxiv.org%2Fpdf%2F1808.01462.pdf&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738567310%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=Pc3qNLhjVlaupTpOdZkVuDwg4k%2BfL5deYQHe9aGuH%2B0%3D&reserved=0)
   3. [https://arxiv.org/pdf/1612.00593.pdf (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Farxiv.org%2Fpdf%2F1612.00593.pdf&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738577304%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=j7ZvhOc7VUIDWuONRMazEGjlAHjeD3G3GlReLKibeYg%3D&reserved=0)
   4. [https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2595507/PANORAMA-NN-Preprint.pdf?sequence=1&isAllowed=y (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Fntnuopen.ntnu.no%2Fntnu-xmlui%2Fbitstream%2Fhandle%2F11250%2F2595507%2FPANORAMA-NN-Preprint.pdf%3Fsequence%3D1%26isAllowed%3Dy&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738577304%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=MZc4MecTvqWedVKILz8A15sR5z%2FPZN9SKvRcF%2FcNviY%3D&reserved=0)
   5. TBA.
7. Datasets (Provide a Link to the dataset): we’ll be using the following three datasets for 3D object classification, part segmentation and semantic segmentation.
   1. .[https://modelnet.cs.princeton.edu/ (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Fmodelnet.cs.princeton.edu%2F%23&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738587302%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=teqcDfXgIsy2YssxHLTJIV4s936zM5WvhCxXZgCrm7Q%3D&reserved=0)
   2. [https://shapenet.cs.stanford.edu/ (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Fshapenet.cs.stanford.edu%2F&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738587302%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=xBBDlZ0BJzAwTO71c6LdKPdqpCTQNMIzO6sTFtlsZ%2F0%3D&reserved=0)
   3. [http://buildingparser.stanford.edu (Links to an external site.)](https://nam12.safelinks.protection.outlook.com/?url=http%3A%2F%2Fbuildingparser.stanford.edu%2F&data=04%7C01%7Czhangsu%40gatech.edu%7Cb4448f3bec72444d111b08d8ea7bcc68%7C482198bbae7b4b258b7a6d7f32faa083%7C0%7C0%7C637517164738597295%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=Xbftn%2FkNJosZkZ81rUCi7W2KPN9GvHvrMsxhxtF4nUA%3D&reserved=0)
8. List your Group members.
   1. Su Zhang
   2. Chenghao Wang
   3. Rui Meng
   4. Oscar Chen

**Team Name:**Roadway Precognition

**Facebook Project:** No

**Project Title:** Sequence Prediction of Multi-agent Systems in Autonomous Driving

**Project Summary:**

Forecasting the positions of dynamic actors poses one of the greatest challenges facing SAE level 4/5 automated driving. Given a sequence of observations from the past, the task is to learn a multi-modal behavior distribution across some time horizon into the future. This lends the question of how to choose those observations. Typically, they are selected as a combination of raw sensory information, states (ex. position, velocity, acceleration, etc), and top-down grid maps. However, one piece of information that these sources cannot explain is the underlying traffic rules governing motion decisions. Factors such as traffic light modes, road-edge types, traffic signs, etc. play a significant role in determining those paths, yet seem to be left out in recent work.

Using timestamped meta-data from high-definition maps and state sequences of vehicles, pedestrians, and bicyclists one second into the past, our goal is to develop and validate an end-to-end model for predicting up to eight actor trajectories for eight seconds into the future.

**Approach:**

* Since the order of the input data matters, we would want to explore the various RNN architectures that have had successful track records with this problem. Then, we would want to implement one of those as a baseline.
* We would also look into using a combination of an RNN and CNN, where the first can encode sequence data into spatial maps tensors and the second can encode features from those tensors.
* As a stretch goal, we’d like to explore the use case of graph neural network architectures for this problem, given the growing number of publications for this technique.

**Resources/Related Work:**

[1] Sajjad Mozaffari, et. Al, “Deep Learning-based Vehicle Behaviour Prediction For Autonomous Driving Applications: A Review”, 2019

[2] A. Zyner, S. Worrall, and E. Nebot, “A recurrent neural network solution for predicting driver intention at unsignalized intersections,” IEEE Robotics and Automation Letters, July 2018

[3] T. Zhao, Y. Xu, et. al, “Multi-agent tensor fusion for contextual trajectory prediction,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2019

[4] X. Li, X. Ying, and M. C. Chuah, “Grip: Graph-based interaction-aware trajectory prediction,”

**Datasets:**

Waymo Open Motion Dataset: [https://waymo.com/open/data/motion/ (Links to an external site.)](https://waymo.com/open/data/motion/)

**Team Members:**

Alexander Greene, Pushkar, Justin Manuel

**Looking For More Members:**

Not actively looking but open to one more person depending on level of interest

Team: Synapse

Facebook Project: No

Project Title: 3D Object Detection - Classification

Project Summary:

The ability to identify 3D objects has many applications such as autonomous driving, robotics, human interaction, video surveillance, medical imaging, and pedestrian detection. Compared to 2D images, 3D data can have multiple representations: multi-view images, volumetric data, point clouds, and 3D meshes. Based on these representations, different methodologies have been developed. For example, 3D convolutions can be performed on 4D kernels (e.g., AlexNet, 3D ShapeNets). Another method is the Bird’s Eye View approach, trying to translate the 3D problem to a 2D problem. Point clouds are the most common representation of sensor data. There are several approaches that covert point clouds to other representations that are fed to Neural Networks (e.g., Voxelization 3D CNN and Projection/rendering 2D CNN). PointNet is an approach that performs end-to-end learning using point clouds which respects the permutation invariance of points in the input. PointNet can be used for object classification, object part segmentation and semantic scene parsing. Complex Yolo is an approach to 3d object detection in real time. If time allows, we will try a Yolov4 implementation of Complex Yolo and/or implement a Scaled yolo version.

Approach:

- Based on preliminary research outlined above, we will explore a few of the existing approaches outlined above. Our goal is to achieve similar performance to the current literature benchmarks and at least be able to replicate PointNet using existing code. We expect a fair amount of time will be devoted to gaining familiarity with the data.

If time and compute budget allow:

- We want to explore ways to improve on exiting approaches.

- We want to use visualization methods to understand the features that the model has learnt.

- Replicate Complex Yolo, likely v4 using existing code. If time allows, we will look at Scaled Yolo and try to implement. As of writing there is no code base which is a Scaled Complex Yolo model.

Resources/Related Work: [1] “PV-RCNN++: Point-Voxel Feature Set Abstraction With Local Vector Representation for 3D Object Detection”, Shi, S, et. al.

[2] “"Pointnet: Deep learning on point sets for 3d classification and segmentation", Charles, Qi et al. [3] “3D ShapeNets: A Deep Representation for Volumetric Shape Modeling”, Wu, Z., et al.

[4] “Voxel R-CNN: Towards High Performance Voxel-based 3D Object Detection”, Deng J., et. al.

[5] “Spatial Transformer Networks”, Jaderberg M., et. al.

[6] “Complex-YOLO: Real-time 3D Object Detection on Point Clouds”, Simon M. et al.

[7] “YOLOv4: Optimal Speed and Accuracy of Object Detection”, Bochkovskiy A. et al.

[8] “Scaled-YOLOv4: Scaling Cross Stage Partial Network”, Chin-Yao W. et al.

Datasets:

ModelNet (CAD models from the 10 or 40 categories). https://modelnet.cs.princeton.edu/

The 3D object detection benchmark. http://www.cvlibs.net/datasets/kitti/eval\_object.php?obj\_benchmark=3d

Team Members:

Carlos Brioso

Travis Ruddy

Looking for more members: No

Team: BamzNet

Facebook Project: No

Project Title: Waymo Motion Prediction

Project Summary:

Waymo is a leading company in the field of autonomous vehicle research. The company is currently hosting a competition where research teams can access Waymo’s autonomous vehicle dataset and submit their own networks with the intent to predict object motion within the sensor suite’s field of view. This project aims to develop or apply a competitive network to the latest set of Waymo inputs, both for the purpose of assessing the viability of the model’s architecture and researching methods for evaluating best fit for networks based on the underlying data to be analyzed.

Approach:

Based on the team’s research, there are a number of architectures which could be applied to motion prediction. Our team plans to implement a basic encoder-recurrent-decoder model which will be our baseline. The team will then look to extend the model as time allows using attention mechanics, social pooling [6], and other concepts from state of the art networks such as those found in MFP [5], PECNet [1], DESIRE [2], or other techniques included in Resnik et. al [3], Paravarzar et. al [4], and others found over the course of the project.

Resources/Related Work:

[1] “It is not the Journey but the Destination: Endpoint Conditioned Trajectory Prediction”, Mangalam et al

[2] “DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents”, Lee et al

[3] “Vehicle motion prediction for autonomous driving”, Resink et al

[4] Paravarzar, Shahrokh, and Belqes Mohammad. "Motion Prediction on Self-driving Cars: A Review." arXiv preprint arXiv:2011.03635 (2020).

[5] “Multiple Futures Prediction”, Tang et al.

[6] “Social LSTM: Human Trajectory Prediction in Crowded Spaces”, Alahi et al.

Datasets:

[https://github.com/waymo-research/waymo-open-dataset (Links to an external site.)](https://github.com/waymo-research/waymo-open-dataset)

Team Members:

Brendon Pierson

Matthew Lemons

Adam Pierce

Zachary Jacob

Team Name

**Team Driverless**

Is this a Facebook project?

**No**

Project Title

**Lyft 3D Object Detection for Autonomous Vehicles**

Project summary (4-5+ sentences). Fill in your problem and background/motivation (why do you want to solve it? Why is it interesting?). This should provide some detail (don't just say "I'll be working on object detection")

**The field of autonomous vehicles is rapidly growing and has the potential to create a massive positive impact on our society. However, there are many issues that need to be solved before we can see large scale production of fully autonomous vehicles.**

**There are 5 levels of self-driving cars, and the current vehicles being tested around the world are at Level 4, which means the self-driving cars right now though autonomous, still need a human driver available for more complex driving scenarios. The reasons why these vehicles cannot reach Level 5 can be summarized into two parts: perception and prediction, i.e. to identify and classify objects near the vehicle and then make predictions on the next positions of the objects when time goes on.**

**Therefore, the detection of 3D objects using data obtained from the vehicles such as LIDAR, RADAR and camera feeds has become an important problem to be solved. An accurate model that can classify and assign bounding volumes to such objects would allow for safer and better autonomous vehicles.**

What you will do (Approach, 4-5+ sentences) - Be specific about what you will implement and what existing code you will use. Describe what you actually plan to implement or the experiments you might try, etc. Again, provide sufficient information describing exactly what you'll do. One of the key things to note is that just downloading code and running it on a dataset is not sufficient for a description or a project! Some thorough implementation, analysis, theory, etc. has to be done for the project.

* **3D object prediction model involves two main steps: object classification and 3D box estimation**
* **Object classification can be achieved using CNN for ImageNet as a backbone via Resnet. After that, explore and implement Inception Network or EfficientNet to improve the performance.**
* **3D box estimation or region detection can be done using YOLO (You Only Look Once) method**
* **Upsampling and downsampling on different blocks in the neural network may be involved**
* **Anchors may be used to adjust coordinates based on size and shape of objects detected**
* **Explore several classification loss calculation techniques such as cross-entropy, hard negative mining, and focal loss.**
* **Optimizing the model using various methods**
* **One stretch goal is to explore multiple techniques to improve the 3D object detection performance.**
* **Another stretch goal is to optimize the task in term of of runtime and resources**

Resources / Related Work & Papers (4-5+ sentences). What is the state of art for this problem? Note that it is perfectly fine for this project to implement approaches that already exist. This part should show you've done some research about what approaches exist.

[1]  “An Overview of 3D object detection”, [Yilin Wang (Links to an external site.)](https://arxiv.org/search/cs?searchtype=author&query=Wang%2C+Y), [Jiayi Ye (Links to an external site.)](https://arxiv.org/search/cs?searchtype=author&query=Ye%2C+J) ([https://arxiv.org/pdf/2010.15614.pdf (Links to an external site.)](https://arxiv.org/pdf/2010.15614.pdf))

[2] “Center-based 3D Object Detection and Tracking”,  [Tianwei Yin (Links to an external site.)](https://arxiv.org/search/cs?searchtype=author&query=Yin%2C+T), [Xingyi Zhou (Links to an external site.)](https://arxiv.org/search/cs?searchtype=author&query=Zhou%2C+X), [Philipp Krähenbühl (Links to an external site.)](https://arxiv.org/search/cs?searchtype=author&query=Kr%C3%A4henb%C3%BChl%2C+P) ([https://arxiv.org/pdf/2006.11275.pdf (Links to an external site.)](https://arxiv.org/pdf/2006.11275.pdf))

[3] ”Lidar 3D Object Detection methods”, Mohammad Sanatkar ([https://towardsdatascience.com/lidar-3d-object-detection-methods-f34cf3227aea (Links to an external site.)](https://towardsdatascience.com/lidar-3d-object-detection-methods-f34cf3227aea))

[4] “You Only Look Once: Unified, Real-Time Object Detection” Joseph Redmon

([https://www.cv-foundation.org/openaccess/content\_cvpr\_2016/html/Redmon\_You\_Only\_Look\_CVPR\_2016\_paper.html (Links to an external site.)](https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html))

[5] “U-Net: Convolutional Networks for Biomedical Image Segmentation”, Olaf Ronneberger, Philipp Fischer, Thomas Brox ([https://arxiv.org/pdf/1505.04597.pdf (Links to an external site.)](https://arxiv.org/pdf/1505.04597.pdf))

[6] “3D Bounding Box Estimation Using Deep Learning and Geometry”, Arsalan Mousavian, Dragomir Anguelov, John Flynn, Jana Kosecka ([https://arxiv.org/pdf/1612.00496.pdf (Links to an external site.)](https://arxiv.org/pdf/1612.00496.pdf))

[7] “”3D Bounding Box Estimation for Autonomous Vehicles by Cascaded Geometric Constraints and Depurated 2D Detections Using 3D Results”, Jiaojiao Fang, Lingtao Zhou, Guizhong Liu ([https://arxiv.org/pdf/1909.01867.pdf (Links to an external site.)](https://arxiv.org/pdf/1909.01867.pdf))

[8] “3D-Object Detection for Autonomous Vehicles”, Sijo Vm ([https://towardsdatascience.com/3d-object-detection-for-autonomous-vehicles-b5f480e40856 (Links to an external site.)](https://towardsdatascience.com/3d-object-detection-for-autonomous-vehicles-b5f480e40856))

[9] “3D Object Detection Using Lidar Data for Self Driving Cars”, Abhinav Sagar ([https://medium.com/swlh/3d-object-detection-using-lidar-data-for-self-driving-cars-ee0eb0e6389e (Links to an external site.)](https://medium.com/swlh/3d-object-detection-using-lidar-data-for-self-driving-cars-ee0eb0e6389e))

[10] “Monocular 3D Object Detection in Autonomous Driving — A Review” Patrick Langechuan Liu ([https://towardsdatascience.com/monocular-3d-object-detection-in-autonomous-driving-2476a3c7f57e (Links to an external site.)](https://towardsdatascience.com/monocular-3d-object-detection-in-autonomous-driving-2476a3c7f57e))

Datasets (Provide a Link to the dataset). This is crucial! Deep learning is data-driven, so what datasets you use is crucial. One of the key things is to make sure you don't try to create and especially annotate your own data! Otherwise the project will be taken over by this.

Kaggle 3D Object Detection Dataset ([https://www.kaggle.com/c/3d-object-detection-for-autonomous-vehicles/data (Links to an external site.)](https://www.kaggle.com/c/3d-object-detection-for-autonomous-vehicles/data) )

List your Group members.

**Tam Nguyen**

**MinhTrang (Mindy) Nguyen**

**Lai Chung Sing, Kevin**

**1. Team Name**

Team 42

**2. Is this a Facebook project?**

Our project is not one of the Facebook projects.

**3. Project Title**

3D Object Detection via PointPillars

**4. Project summary (4-5+ sentences). Fill in your problem and background/motivation (why do you want to solve it? Why is it interesting?). This should provide some detail (don't just say "I'll be working on object detection")**

We will be replicating and attempting to improve upon state-of-the-art methods for autonomous vehicle image recognition using PointPillars. Object detection and computer vision is paramount to making the leap to completely autonomous driving.  This is an interesting area of research by major institutions and companies in industry such as Uber, Tesla, and Google.  This will also show our complete understanding of CNNs and extend it into computer vision.

**5. What you will do (Approach, 4-5+ sentences) - Be specific about what you will implement and what existing code you will use. Describe what you actually plan to implement or the experiments you might try, etc. Again, provide sufficient information describing exactly what you'll do. One of the key things to note is that just downloading code and running it on a dataset is not sufficient for a description or a project! Some thorough implementation, analysis, theory, etc. has to be done for the project.**

We will be attempting to replicate 3D object detection using PointPillars. We will be reading research papers that have been published in recent years and attempt to achieve state-of-art performance on the KITTI data set, which is standard for this problem. We will use a standard dataset to focus on the research and use a standard benchmark. We will be writing our Neural Network using PyTorch and training on our graphics processing units (GPUs).  We have three different servers each with their own GPU so we can train three different models at once testing different architectures. This will enable us to tune our model three times as quickly. PointPillars is an approach developed by researchers at Aptiv/Nutonomy. The approach consists of a three stage method, a pillar feature head, backbone 2D CNN, and a detection head. The incoming pointcloud is transformed into pillar feature embeddings organized as a pseudo image This is then passed into the 2D CNN to extract the required features for the detection head. The detection head performs regression, outputting anchor box size, angle, and position deltas, as well as occupancy and class ID.

**6. Resources / Related Work & Papers (4-5+ sentences). What is the state of art for this problem? Note that it is perfectly fine for this project to implement approaches that already exist. This part should show you've done some research about what approaches exist.**

The state-of-the-art for this problem is PointVoxel-RCNN. Unlike PointPillars, PV-RCNN is based on 3D Voxel CNN and has achieved 83.9% AP on the Kitti Cars dataset.

We will continue to explore existing methods, but the related works that we will focus on initially include the following:

[1] "PointPillars: Fast Encoders for Object Detection from Point Clouds", Lang et al.

[2] "LiDAR Point-Cloud Based 3D Object Detection Implementation with Colab,” Adusumilli, Gopalakrishna.

[3] "Optimisation of the PointPillars Network for 3D Object Detection in Point Clouds", Stanisz et al.

[4] “Implementing Point Pillars in TensorFlow”, Tyagi, Anjul.

[5] “VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection”, Zhou et al.

[6] "Point-Voxel CNN for Efficient 3D Deep Learning", Liu et al.

[7] "PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection", Shi et al.

[8] "PV-RCNN++: Point-Voxel Feature Set Abstraction With Local Vector Representation for 3D Object Detection", Shi et al.

**7. Datasets (Provide a Link to the dataset). This is crucial! Deep learning is data-driven, so what datasets you use is crucial. One of the key things is to make sure you don't try to create and especially annotate your own data! Otherwise the project will be taken over by this.**

For this project we will be utilizing the KITTI dataset used in the “PointPillars: Fast Encoders for Object Detection from Point Clouds" paper. The KITTI Vision Benchmark Suite includes 29 GB of Velodyne Lidar Point Clouds as well as labeled ground truth 3D bounding boxes. ([http://www.cvlibs.net/datasets/kitti/eval\_object.php?obj\_benchmark=3d (Links to an external site.)](http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d))

**8. List your Group members.**

Our group members are: Christopher Klapperich, Daniel Martinez, Christopher Scully, and Luke Wileczek

**9. Are you looking for more members?**

We are not presently looking for additional team members.

Team: The Church of Agnostic Weights in Motion Prediction

Facebook Project: No

Project Title: Motion Prediction for Autonomous Driving Using the Waymo Open Dataset

Project Summary:

We seek to complete 2021 Challenge 1 on the Waymo Open Dataset which has the stated goal of predicting the positions of up to 8 agents 8 seconds into the future given their tracks for the past 1 second on a map. Motion prediction is particularly interesting because it combines the problems of object detection with tracking and memory. It will be informative to see how much techniques traditionally used for machine translation such as RNNs, LSTMs, and Transformers can help in predicting future locations of moving objects. In a bigger picture sense, motion prediction is a crucial component of autonomous driving, which has potential for widespread use in the coming years.

Approach:

* One technique we would like to implement is Weight Agnostic Neural Networks to come up with the best network structure, allowing us to skip manual tuning of the architecture. We could also use WANN as an initialization strategy to reduce the searching space for network optimization.
* We will attempt to use evolving neural network techniques such as NEAT to see if there is any performance gain over fixed topologies.
* We would also like to explore network structures that use memory such as RNNs, LSTMs, and Transformers that use attention to see how memory can be used in motion prediction.
* We seek to conduct motion prediction using more computationally cheap methods instead of 3D point clouds.

Resources/Related Work:

[1] Gu Z, Li Z, Di X, Shi R. An lstm-based autonomous driving model using a waymo open dataset. Applied Sciences. 2020 Jan;10(6):2046.

[2] Gaier, A. and Ha, D., 2019. Weight agnostic neural networks. *arXiv preprint arXiv:1906.04358*.

[3] Mandal, S., Biswas, S., Balas, V.E., Shaw, R.N. and Ghosh, A., 2020, October. Motion Prediction for Autonomous Vehicles from Lyft Dataset using Deep Learning. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (pp. 768-773). IEEE.

[4] Pang, S. and Radha, H., 2021. Multi-Object Tracking using Poisson Multi-Bernoulli Mixture Filtering for Autonomous Vehicles. *arXiv preprint arXiv:2103.07783*.

[5] Hu, H., Yang Y., Fischer, T., Darrell T., Yu, F., Sun, M. Monocular Quasi-Dense 3D Object Tracking. arXiv preprint arXiv: 2103.07351.

[6] Stanley, K. O., Miikkulainen, R., Evolving Neural Networks through Augmenting Topologies. The University of Texas at Austin. 2002.

Datasets:

Waymo Open Dataset[**https://waymo.com/open/data/motion/** (Links to an external site.)](https://waymo.com/open/data/motion/)

Team Members:

Jingquan Wang

Changcheng Li

Steven Pan

Looking for more members:

No