Multi-Agent Multipath:

Multiple Probabilistic Anchor Trajectory Hypotheses for Behavior Prediction for multiple agents

Team Members:
Anirudh Swaminathan, Keshav Rungta, Saikiran Komatineni, Derek Yang

Mentors:
Nachiket Deo, Kaouther Messaoud, Ross Greer, Bowen Zhang, Prof. Mohan Trivedi
Motivation

- MultiPath -> paper we saw in lectures
- MultiPath predicts trajectories for a single agent
- PRECOG -> another paper we saw in lectures
- PRECOG -> predicts continuous distribution for trajectory prediction for multiple agents
- Motivation - Modify PRECOG to implement MultiPath for multiple agents
Problem Statement

- Predict future trajectory of multiple vehicles in a scene given past trajectories
- Expand MultiPath -> which predicts single agent future trajectories
- Cannot directly apply MultiPath to all agents
  - Trajectories that seem to be possible individually need not be compatible with each other
- Model Joint distribution -> Similar to PRECOG ESP module, but discrete
- Discrete distribution => discrete variables correspond to anchors/intents
Relevant Papers

- **MultiPath** - Yuning Chai*, Benjamin Sapp*, Mayank Bansal Dragomir Anguelov from Waymo LLC
  - A model which predicts parametric distributions of future trajectories for agents in real-world settings
  - Dataset: Custom, CARLA

- **PRECOG** - Nicholas Rhinehart¹ Rowan McAllister² Kris Kitani¹ Sergey Levine² from CMU and UC Berkeley
  - Multi-Agent forecasting method - ESP
  - PRECOG algorithm - forecasts conditioned on agent goals which improved joint-agent and per-agent predictions
  - Dataset: nuScenes, CARLA

- **SocialGAN** - Agrim Gupta¹ Justin Johnson¹ Li Fei-Fei¹ Silvio Savarese¹ Alexandre Alahi¹,² from Stanford University¹ Ecole Polytechnique Féde´rate de Lausanne²
  - GAN based encoder decoder framework for trajectory prediction
Methodology

- **Objective** - Trajectory generative model for multiple agents based on discrete intentions.

- **Methodology** -
  - PRECOG, models agent-agent interactions, decide whether their intentions/trajectories are compatible.
  - MultiPath, extracts discrete trajectory anchors (intentions).
  - Social GAN, handle variable sized inputs via max pooling.

- **Expected Goal** -
  - Discrete intention sampling for each agent.
  - Scoring network that scores the sampled set of intentions.
    - Variable sized inputs
  - Intention-conditioned trajectory generation

![Diagram](image-url)
Significance and novelty

- Trajectory prediction in literature - 2 approaches
  - Approach 1
    - Past trajectory -> Future trajectory (PRECOG, R2P2, Social GAN)
  - Approach 2
    - Past trajectory -> Discrete intention -> Future trajectory (MultiPath, Convolutional Social Pooling)

- Intentions usually outperform latent variables (Classification VS. Regression).
- We apply the second approach to multi-agent scenario.
- We also allow number of vehicles to vary.
Datasets, software libraries and compute resources

- Pytorch
- DataHub (additional GPU Clusters if available)
- Python Dev kits - Scipy, Numpy, Matplotlib, etc

Datasets

- nuScenes (Primary)
- No longer have time to consider Carla Dataset
Current progress

Network (Anirudh, Derek)
- Identified and read few relevant papers
- Multiple meetings with multiple ideas narrowed down to current one
- Ideated the use of scoring functions, with scene context and anchors as inputs with Global Maxpooling to handle multiple agents before softmax
- Created GitHub repository for our project

Dataset (Keshav, Sai)
- Downloading NuScenes - 70% done
- Access script to download on pods - 100% done
Current Progress

NUSCENES Prediction Challenge
NuScenes

- Successful Data loader implementation
- Input data structure - temporally aligned across agents
  - Dictionary with keys: coordinates, heading_rate_change, and rendered map
  - Coordinates:
    - 1st dim: total number of data points
    - 2nd dim: total number of agents
    - 3rd dim: fixed number of coordinates for time interval = 2.5
NuScenes

- We only get histories and future trajectories of 2 agents in any given scene. We filter out all other agents.
  - We do this to simplify the scene complexity and get initial results to show proof of concept
  - The trajectories so rendered on the maps are annotated for 2.5 seconds at 2 Hz
Toy Dataset

- Provide Proof of concept for the joint discrete distribution generative model
- Avoid computational time and complexity of implementing anchor mining
- Control number of agents, scene structure, frequency of sampling
- Developed code to generate physically probable trajectories
- Can shift code to nuScenes when algorithm works on toy dataset
Toy Dataset

- **Vehicle:**
  - Keep track of velocity and coordinate
  - $x(t+1) = x(t) + v(t) \cdot dt$
  - $v(t+1) = v(t) + a(t) \cdot dt$
  - Control $a(t)$ only
  - Intenions:
    - FW: Go straight.
    - LT: Left turn to a specific goal location with a quarter-circle trajectory. ($a = v^2 / r$)
    - RT: Right Turn.
    - BK: Brake, gradually decrease $v(t)$.

- **Environment:**
  - Initialize vehicles with different locations and intentions
  - Determine whether collision happens
  - Save as dataset
Toy Dataset

- Initially generate scenes with 2 agents.
- To minimize the load on the scene parsing CNN, only keep 90° intersections.
- 4 anchors for 4 intents.
- 500*500 image size to be compatible with nuScenes dataset.
- Intersections -> square with 50*50 size.
- Generate frames at 2Hz, same as nuScenes.
- Random starting points in mutually exclusive lanes.
- Random initial velocities; 90° turns.
- Do not bother with traffic rules as of now.
Sample Gifs
Input to Network

Input -> Past trajectories + scene context
Network Details and Discussions

Input:
- Past trajectories: numpy arrays/binary marks.
- Toy dataset: 4 intentions rather than 4 anchors.
- nuScenes: generated anchors with all its scene variability.

Network details:
- Toy dataset:
  - Scene features: lightweight resnet, or the first few layers of the VGG nets.
  - Past trajectory encoding: 32-dimensional GRU
Network Details and Discussions

Scoring and Loss:

- Use maxpooling to allow variable input channels (number of vehicles)
- Return logits for each of the $k^n$ possible modes (in the toy dataset $n=2, k=4$).
- Softmax and compute cross-entropy with ground truth (one-hot).
Roles and Assignments

- All hands on deck to generate the two datasets because the original paper was using a custom dataset
- Sai and Keshav will be working on visualisation and optimisation of the project
- Anirudh and Derek will be working on the scoring function and the implementing the network.
- This project is taking much longer than we originally thought because we are implementing our own solution from scratch with no prior implementation of Multipath available.
Milestones and timeline for the next 3 weeks

Network (Derek, Anirudh)

- Finish Implementation for MultiPath Anchor generation by Wednesday next week
- Convert anchors to global coordinates and design and start implementation of a scoring function by week 10
- Complete implementation of scoring function for 2 agents in nuScenes, for 5 anchors with MultiPath for final trajectory generation (finals week Wednesday)

Dataset (Keshav, Sai):

- Figure out actual inputs (Goal to finish by saturday latest)
- Dataloader - Next (Goal to write template by monday)
- Keep tweaking dataloader as needed and then help out with the network
Milestones and timeline for the next 2 weeks

Network (Derek, Anirudh):
- Generate toy dataset
- Finish Scoring function
- Write the multipath implementation

Dataset (Keshav, Sai):
- Finish dictionary creation of feature vectors
- Generate Nuscenes dataset
- Help with the network side as much as possible
- Write report
Acknowledgements

- We would like to thank all our mentors for meeting with us, ideating and helping to formulate the ideas and implementation techniques
- Thank you for being patient with our emails 😊
Respected Professor and project mentors,

This is Anirudh Swaminathan, and I am writing this email on behalf of my team, team 6, consisting of Derek Yang, Keshav Rungta and Saikiran Komatineni.

Since the Professor was unavailable for the meeting on Monday, May 18, 2020, I am also including a small summary of that meeting in this email.

**Monday Meeting Summary**

After the our meeting with you on Friday, our team came up with two main ideas that we wanted to pursue. Both our ideas were formulated on the basis of utilizing the nuScenes dataset.

- Our first idea was to modify the papers "Scene Induced Multi-modal image forecasting" and "Trajectory forecasting in unknown environments" by Nachiket on the nuScenes dataset. Our question was whether the nuScenes map dataset would sufficiently replace the SDD dataset drone images. Since Nachiket is currently working on porting his paper to different datasets now due to his reviewer comments, he felt that it would be redundant for us to implement the same.

- Our second idea was to implement Convolutional Social Pooling paper by Nachiket on the nuScenes dataset. We initially believed that this would involve significant work, as we felt that we would have to track object instances, and then utilize that information to predict trajectories and then work on the dataset. Nachiket though, pointed out that this would not be significant progress, as nuScenes already contains the locations of the object in the global coordinate frame already.

**Further ideas to decide upon**
hello team 6: hope you all are doing well and staying healthy..when you get five mins do send me an update of your project..enjoy..thanks

Mohan Trivedi
to me, Keshav, Anirudh, Derek

Sun, May 31, 5:57 PM (4 days ago)

Anirudh Swaminathan
to Nachiket, Kaouther, Mohan, me, Keshav, Derek

Respected Professor and Mentors,

This is Anirudh Swaminathan. Thank you for your email and your enquiries professor. I'm writing on behalf of team 6 here. To refresh your memory, on Thursday, May 28, 2020, we had a meeting to discuss the next steps and to clarify a few questions that we had back then.

Meeting Summary - May 28, 2020
In our meeting, we decided on the next objectives for the project and clarified our questions.
The objective for Derek and I was to create a toy dataset that we could then use with MultiPath.
The objective for Saikiran and Keshav was to continue improving the DataLoader for nuScenes.

Algorithm to follow
The key takeaway for the structure of the algorithm was formulated as follows:-
1. Process BEV scene image with agents + their past trajectories through a CNN.
2. Generate k*n anchor representations using RNNs.
Thank you
Final Project Presentation Template

The final group presentations for Week 10 should cover the following:

- **Project title + group members + mentors (1-2 slides)**
- **Methodology, expected goals, significance and novelty (2 slides)**
  
  Include a figure and/or flowchart (can be preliminary, you can make changes later on)

- **Current progress & results (5-8 slides)**
  
  Include detailed progress reports and early results from experiments

- **Detailed breakdown of individual workload (1-2 slides)**
  
  Information on who has done what so far

- **Milestones and timeline for the next 2 weeks (1 slide)**

The presentation can be coordinated such that one member shares their screen while each member presents a few slides. **Each group will get to present for 20 minutes followed by 10 minutes of Q&A.**