Group 2 Proposal

Title: Training Reinforcement Learning Self-Driving Car using Generative Model of the World (i.e, World Model)

Specific: Using World Models as a Backbone for training Neural-Network-Based Controllers in Autonomous Driving: Simulation and Evaluation under Controlled Environments in OpenAI Gymnasium

Problem Motivation

The rapid rise of autonomous driving technology has been largely driven by advances in supervised learning and reinforcement learning (RL). However, most existing control approaches depend heavily on dense supervision, extensive real-world data collection, and large numbers of interactions with costly or unsafe environments.

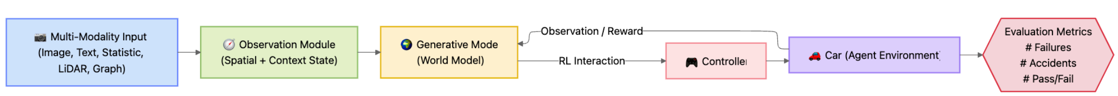
Despite the success of behavior cloning, end-to-end RL, and model-predictive control (MPC), these systems still suffer from **sample inefficiency, limited generalization, and a lack of interpretability**.

Also, implementing the Autonomous Vehicle (A.V.) is needed to train in real-life for getting a realistic environment observation, which is impossible due to safety & economic reasons (price & expense of the physical hardware & the data-collection)

Humans, in contrast, learn to drive safely and adaptively from relatively few experiences by building **internal predictive models of the world** — mental simulations of "what will happen if I act this way." This ability to *imagine outcomes before acting* is precisely what *world models* in AI attempt to replicate.

Therefore, the motivation of this research is to explore whether **a learned generative world model**, trained to predict environmental dynamics from high-dimensional sensor inputs, can serve as a differentiable backbone for an autonomous vehicle controller, allowing for safe, data-efficient learning and robust decision-making in a simulated, controlled setting.

Problem



- We have to resolve 3 problems:
  - Training `Observation module` for encoding the multi-modality inputs from the real world (Image of the road, text on road description, statistic on car velocity & environment, road-graph) for the latent-generative model on the world (World Model)
  - Training the `World Model` for generating the latent-encoding version of the world (can be images)
  - Training the `self-driving car controller` using the generative data from the world model

Dataset (Tentative for using - can expand / not using all of those based on realistic case) - with the article using them

🧩 1. Hu et al., 2023 — GAIA-1

Domain: Real-world multimodal autonomous driving

Datasets:

- **UK Urban Driving Dataset –** large-scale real-world driving videos with multimodal labels (video, text, and actions) used to train GAIA-1.

  → *(dataset not public yet; described as "large corpus of real-world UK urban driving data")*

- **Reference Datasets Used for Comparison** (cited for related work):
  - **TrafficSim** (CVPR 2021) — realistic multi-agent traffic simulation dataset https://github.com/boris-belousov/trafficsim
  - **Waymo Open Sim Agents Challenge 2023** — interactive simulation dataset from Waymo Open Dataset https://waymo.com/open/simulation
  - **CommonRoad Benchmark** (IVS 2017) — composable motion-planning benchmark https://commonroad.in.tum.de

---

🧠 2. Hafner et al., 2019 — PlaNet

Domain: Continuous-control reinforcement learning from pixels

Datasets / Simulated Environments:

- **DeepMind Control Suite** — physics-based MuJoCo environments (Cheetah, Walker, Finger, etc.)

  🔗 https://github.com/deepmind/dm_control

- **MuJoCo** simulator (Todorov et al., 2012) — physics engine for locomotion control
- **Atari Learning Environment** — for early action-conditional video prediction baselines

---

🚗 3. Vasudevan et al., 2024 — AdaptiveDriver / Reactive World Model for AVs

Domain: Autonomous vehicle motion-planning benchmarks

Datasets:

- **DARPA Urban Challenge (2007–2008) –** autonomous driving competition dataset (Boss, Junior, Odin, Stanley entries).

  🔗 https://www.darpa.mil/about-us/darpa-urban-challenge

- **DeepDriving (ICCV 2015) –** learning affordances from camera input for direct perception.

  🔗 https://github.com/sermanet/DeepDriving

- **Baidu Apollo Open Dataset (2018) –** EM motion-planning dataset.

🔗 https://apollo.auto/dataset.html

- **CommonRoad** and **Waymo Open Sim Agents**, also referenced here for validation

---

🧩 4. Liu et al., 2019

Domain: Reinforcement-learning world-model evaluation

Datasets / Environments:

- **Standard MuJoCo benchmarks:** Ant-v2, HalfCheetah-v2, Humanoid-v2, Walker2d-v2, HumanoidStandup-v2
- **Flickering MuJoCo** — custom partial-observability (POMDP) extension of MuJoCo where frames are randomly hidden

---

🌍 5. Ha & Schmidhuber (2018) / World Models

Domain: Visual RL benchmark for compact world models

Datasets / Environments:

- **CarRacing-v0 (OpenAI Gym)** – 2D top-down driving simulation.

  🔗 https://gym.openai.com/envs/CarRacing-v0/

- **DoomTakeCover-v0 (VizDoom via Gym) –** first-person vision task for reactive control.

  🔗 https://gym.openai.com/envs/DoomTakeCover-v0/

Algorithm & Model Designing (Proposing in using based on the SOTA)

- For the `Observation module` the using: Variational Autoencoder , Graph-Convolutional Neural Network, Auto-Encoder is usable
- World Model architectural using: (idealistic generative model have to be light-weight
  - Auto-regression Transformer neural network
  - Recurrent neural network + Mixture Gaussian network - the original WorldModel Paper
  - Graph Convolutional Network (for graph-encoding)
- The Decoder of the data
- Controller of the self-driving car on guessing the moving state (training the moving policy) - follow the Reinforcement learning
  - Simple perceptron-layer / Multi-layer perceptron

Input

- Multi-modality input on the transportation: Image, Text-Description of the image & environment (the obstacle & the surrounding enviorment) , physical statistic (velocity, etc.) of the car, graph of the road (if possible)

Output

- For the generative-module → the latent representation of the world for the controller training
- For the controller → the car-automation that follow a specific goal that need to optimize (# failure case, # accident, # pass or fail case)

Reference:

- Hu, A., Russell, L., Yeo, H., Murez, Z., Fedoseev, G., Kendall, A., Shotton, J., & Corrado, G. (2023). *GAIA-1: A Generative World Model for Autonomous Driving*. ArXiv.org. https://arxiv.org/abs/2309.17080
- Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2023, January 10). *Mastering Diverse Domains through World Models*. ArXiv.org. https://doi.org/10.48550/arXiv.2301.04104
- Vasudevan, A. B., Peri, N., Schneider, J., & Ramanan, D. (2024). *Planning with Adaptive World Models for Autonomous Driving*. ArXiv.org. https://arxiv.org/abs/2406.10714v1
- Liu, J., Gu, X., & Liu, S. (2019). *Reinforcement learning with world model*. ArXiv.org. https://arxiv.org/abs/1908.11494
- Ha, D., & Schmidhuber, J. (2018). World Models. *World Models*, *1*(1), e10. https://doi.org/10.5281/zenodo.1207631