Assignment Overview

Title: Unsupervised Clustering and Self-Learning in Autonomous Driving

Objective: To develop and apply unsupervised learning models for clustering highway driving scenarios, utilizing both image data and telemetric data.

Data Provided:

> **Images**: 100 images from a Carla simulator highway driving scenario, stored in a folder named "Images". Each image is named after its frame.
> **CSV Data File**: Contains various parameters related to the vehicles in each frame, including positional data, velocities, lane associations, and the ego vehicle's status. See Appendix

Tasks

Framework Specification:

- Please use *PyTorch* for developing your clustering models.
- You are free to use any supporting code libraries as needed.

Task 1: Initial Clustering Model

- **Objective**: Create a model that clusters all frames into 5 distinct clusters.
- **Data Utilization**: You can choose which columns from the provided CSV file to use. The use of images is optional.
- **Deliverables**:
    a. **Clustered Frames**: Save each frame into one of *5* folders corresponding to the clusters.
    b. **Visualization**: Develop a method to visualize your clustering. This could include plots showing clusters in feature space or other creative visual representations.
    c. **Code**: Include all code used for the tasks (files or link to notebook). It should be well-commented and organized.
    d. **Documentation**: Provide a brief explanation of your approach, including the rationale behind the chosen features and clustering technique.
    e. **Submission Folder**: Save all deliverables under a folder named *task_1*.

Task 2: Advanced Clustering Approach

- **Objective**: Propose and implement a more sophisticated clustering approach. The number of clusters is at your discretion.
- **Motivation for Clustering**:
    a. **Explainability for Human Driver**: Clusters should provide insights that are understandable and useful for a human driver.
    b. **Training Multiple Agents**: The clusters should be beneficial for training autonomous driving algorithms.
- **Deliverables**:
    a. Same as above, with a folder per cluster. Call the main folder *task_2*.

Appendix - Data Column Explanation

- `data.csv` contains:
    - `frame`: Frame name, corresponding to image names.
    - `object_id`: Identifier for each vehicle across all frames (Ids are kept the same for each vehicle during all frames - but each frame might have a different amount of objects).
    - `lat_dist`: Lateral distance of the vehicle relative to the ego vehicle (positive to the right).
    - `long_dist`: Longitudinal distance from the vehicle to the ego vehicle.
    - `abs_vel`: Absolute longitudinal velocity of the vehicle.
    - `rel_vel`: Longitudinal velocity of the vehicle relative to the ego vehicle.
    - `is_cipv`: Indicator if the vehicle is the *Closest in Path Vehicle* (1 for yes).
    - `lane_association`: Lane position of the vehicle (1 for ego lane, 2 for left, 3 for right).
    - `lane_right_boundary`: Lateral distance of the right boundary of the ego lane from the center of the ego vehicle.
    - `lane_left_boundary`: Lateral distance of the left boundary of the ego lane from the center of the ego vehicle.
    - `ego_vel`: Longitudinal velocity of the ego vehicle.
    - `ego_acc`: Longitudinal acceleration of the ego vehicle.
    - `ego_yaw`: Yaw rate of the ego vehicle.
- *Note*: Units for all fields are considered irrelevant for this task.