Project Report on AI-based Autonomous Driving

D Yaswa Anirwin

October 16, 2024

Introduction

The objective of this project is to develop an AI-based autonomous driving system that focuses on critical aspects of road safety and navigation, including lane detection, traffic light detection, semantic segmentation, and object detection. By leveraging advanced computer vision techniques, the project aims to enhance the decision-making capabilities of autonomous vehicles, ensuring they can operate effectively in diverse driving conditions. The progress made in the first two weeks is foundational for the subsequent stages of model development.

As the world moves towards increased automation in transportation, this project holds significance in contributing to safer and more reliable autonomous vehicles. The integration of AI in navigation systems is a critical step towards achieving this goal. Therefore, the focus on developing robust algorithms for understanding complex driving scenarios is paramount.

Dataset and Methodology

Dataset Overview

The project utilizes the GT Fine dataset, which contains a variety of high-resolution images annotated for various driving scenarios. This dataset is pivotal for training machine learning models due to its comprehensive representation of real-world conditions. Given the dataset's size, a sample dataset was created to facilitate efficient processing on limited hardware. This sample includes images that represent the full dataset, focusing on critical scenarios relevant to lane detection, traffic light detection, semantic segmentation, and object detection tasks.

In addition to the primary dataset, several secondary datasets were considered to augment the training process. These datasets include variations of traffic environments, different weather conditions, and times of day to ensure the models are robust and capable of generalizing across diverse scenarios.

Methodology

The methodology employed for this project includes several critical steps, which are detailed below.

Data Loading

- Implemented functions to load images and their annotations from the specified directory structure. This ensures a seamless workflow for accessing image data required for preprocessing and model training.
- Successfully extracted images from the zipped dataset, organizing them in a way that is conducive to easy access and further processing. This organization allows for quick retrieval of images during training and evaluation phases, improving overall efficiency.

Image Preprocessing

Basic preprocessing steps were applied, including:

- **Resizing**: All images were resized to a standard output size (e.g., 256x256 pixels) to maintain uniformity for model input. This step is essential to ensure that the input dimensions are consistent across the dataset, allowing for batch processing during training.
- **Normalization**: Pixel values were normalized to a range between 0 and 1 to improve model training efficiency. Normalization helps the model to converge faster during training by ensuring that all input features contribute equally.
- **Noise Reduction**: Gaussian blur was applied to reduce noise in the images, helping to refine edges and improve the performance of edge detection algorithms. Noise reduction is particularly important in real-world driving scenarios, where sensor noise can impact the accuracy of detection tasks.
- **Cropping**: Relevant regions of interest (ROIs) were identified and cropped to focus on the areas most relevant to driving tasks, such as lanes and traffic lights. Cropping helps reduce computational overhead and allows models to concentrate on important features in the image.

This preprocessing is essential for the performance of semantic segmentation, lane detection, traffic light detection, and object detection models, as it ensures consistency and reduces computational overhead.

Results

Progress Overview

- Successfully downloaded the GT Fine dataset and created a sample dataset for initial experimentation. The sample dataset provides a balanced representation of different scenarios encountered in real driving conditions.
- Implemented data loading and preprocessing functions that effectively handle image resizing and normalization. These functions are modular and can be reused in future phases of the project.
- Enhanced the dataset through various preprocessing techniques, significantly improving the quality of input images for model training. The improvements in image quality are expected to lead to better performance during the model evaluation phase.

Challenges Encountered

Handling Large Dataset Files

- Encountered memory issues when working with the full dataset. The creation of a sample dataset alleviated some of these challenges and allowed for more manageable processing. The efficient handling of memory is crucial, especially when dealing with high-resolution images in deep learning tasks.
- The need for additional preprocessing steps became apparent during initial testing phases, where certain images did not perform well under specific conditions. This led to the exploration of more advanced preprocessing techniques to improve model robustness.

Model Performance Evaluation

- Initial evaluations indicated that while the data loading and preprocessing steps were successful, additional tuning of parameters would be necessary to achieve optimal results. Establishing benchmarks for model performance will guide future iterations of model training.

Conclusion

The achievements during Weeks 1 and 2 set a solid foundation for the autonomous driving project. The successful implementation of data loading and preprocessing functions ensures that the next phases of the project can proceed smoothly. This progress provides a clear pathway toward developing models for lane detection, traffic light detection, semantic segmentation, and object detection.

As the project progresses, continuous evaluation and refinement of preprocessing steps and model parameters will be essential to enhance performance. By focusing on comprehensive preprocessing and robust model development, the project aims to contribute significantly to the field of autonomous driving.

Future Objectives

- **Complete Dataset Exploration**: Further explore the sample dataset to understand its characteristics better and identify any additional preprocessing steps required. Comprehensive exploration will inform decisions regarding model architecture and training strategies.
- **Model Development**: Initiate the development of models for lane detection, traffic light detection, semantic segmentation, and object detection based on the preprocessed dataset. Each model will require tailored approaches based on the specific tasks and dataset characteristics.
- **Evaluation Metrics Establishment**: Define evaluation metrics to measure the performance of developed models, such as Intersection over Union (IoU) for segmentation tasks and precision-recall metrics for object detection. Establishing clear evaluation criteria will facilitate performance assessment and comparison across models.
- **Iterative Improvement**: Plan for iterative improvements based on model evaluation results, focusing on enhancing accuracy and reducing inference time. Continuous improvement cycles will ensure that the models remain state-of-the-art and meet practical application requirements.

References

- Documentation for libraries used (e.g., PIL, NumPy). `https://pillow.readthedocs.io/en/stable/`

- Tutorials on semantic segmentation, lane detection, traffic light detection, and object detection. `https://pytorch.org/tutorials/`
- Relevant research papers in the field of computer vision and autonomous driving systems. `https://arxiv.org/`
- Datasets for autonomous driving tasks. `https://www.cityscapes-dataset.com/`Cityscapes Dataset