Robotics, Artificial Intelligence
and Real-time Systems

Simulation-Based Autonomous Driving in Crowded City

Mahir Efe Kaya - Cem Şahin

Outline

Motivation

- Growing urbanization leads to increased traffic congestion and challenges for conventional driving systems.

- Autonomous driving offers potential solutions to improve traffic flow, reduce accidents, and optimize transportation.

Introduction

- Goal: Understand the current state of the art methodologies and develop a simulation-based autonomous driving system to navigate crowded city environments efficiently and safely.

- Contribution: Advancing the current state of research by proposing a novel simulation-driven approach for urban autonomous driving.

State of the Current Research

- The landscape of autonomous driving technologies spans from basic driver-assistance systems to the pursuit of fully autonomous vehicles.

- Research contributions have been significant in areas like perception systems, decision-making algorithms, and sensor fusion techniques.

- Despite progress, autonomous systems face challenges posed by intricate urban scenarios, dynamic pedestrian interactions, and the inherent unpredictability of city traffic.

State of the Current Research

- Nvidia’s ChaufferNet:

  a. Features the perception, mapping, and planning layers

  b. Uses diverse DNNs trained on high-quality, real-world driving data and synthetic data

  c. Generates an optimal trajectory through a series of model-based and data-driven analyses

YOLO (You Only Look Once)

- Unlike traditional object detection methods that require multiple passes over an image, YOLO performs detection in a single forward pass.
- YOLO divides the input image into a grid and predicts bounding boxes and class probabilities directly, all in one go.
- Remarkably fast inference times, making it well-suited for real-time applications like autonomous driving.

CARLA Leaderboard

- Pivotal platform for evaluating the performance of various autonomous driving algorithms within simulated urban environments

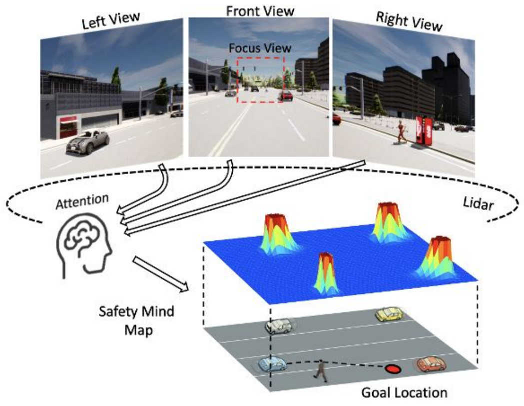
- Provides standardized scenarios and metrics, it facilitates fair comparisons between different approaches

- Accelerates the development cycle by enabling rapid prototyping, testing, and refining of autonomous algorithms

CARLA

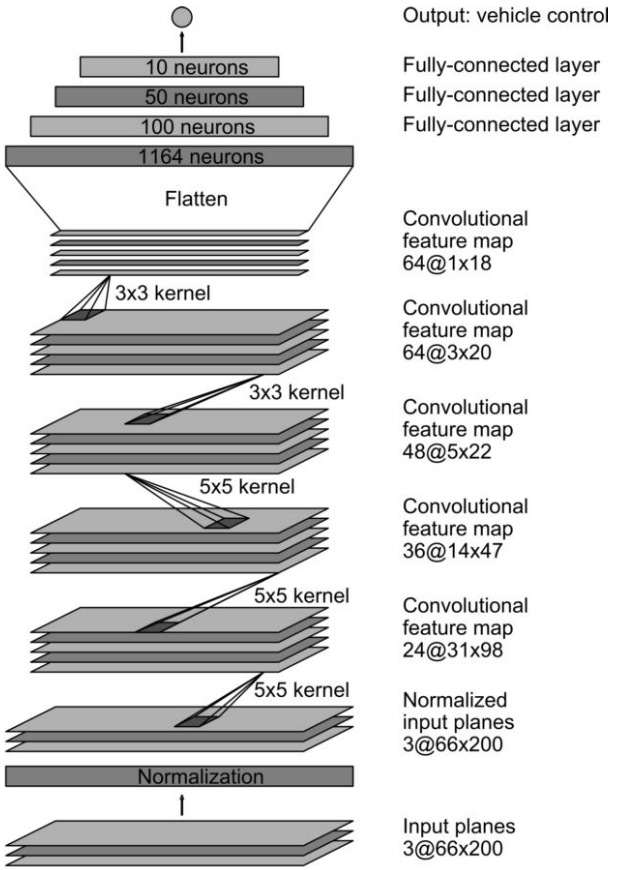
InterFuser

- #1 in CARLA Leaderboard
- Integrates information from multiple sensors like lidar, radar, and cameras, InterFuser enhances the overall perception system
- Addresses sensor limitations by compensating for each sensor's strengths and weaknesses, leading to improved reliability



Udacity Simulator

- Nvidia's CNN Architecture
  - ~50% completion rate on Track 1
  - ~40% completion rate on Track 2
- Udacity Source Code (Experimental)
  - Changing asphalt textures to increase dataset size
  - Trained U-Net Semantic Segmentation on road detection
  - Imitation Learning on top of Semantic Segmentation



Methodologies Used

- YOLO
    - Single Pass Detection
    - Anchor Boxes
    - Non-maximum Suppression
- Rules
- GRU

Methodologies Used

- OpenCV
  - Contour Detection and Marking Identification
  - Refinement through Contour Size Thresholding
  - Lane Marking Identification and Orientation Calculation

- Experimental:
  - Semantic Segmentation
  - Resnet
  - Semantic Segmentation and Resnet
  - Sift
  - Histogram
  - Ratio
  - Thresholding
  - Shape-matching

Initial Approach

- Aimed to develop a system akin to ChauffeurNet, a prominent end-to-end learning framework.
- YOLO was employed to detect proximate vehicles, enabling the creation of a social grid for our ego vehicle.
- Convolutional Neural Network (CNN) model was employed for lane following by the ego vehicle, also to derive the global path.
- Gated Recurrent Unit (GRU) model was trained to assimilate YOLO and CNN outputs, generating a secure and comfortable trajectory for the vehicle.

Revised Approach

- YOLO's capabilities were harnessed for the detection of nearby vehicles and traffic lights, enabling comprehensive scene perception.

- OpenCV's advanced tools, like FindContours, were integrated to discern lane lines accurately. Additionally, orientation was extracted from three images to guide the ego vehicle along the lane.
  - However, use of contours are computationally too expensive. Which lowered to FPS to a degree model wasn’t able to react.

- These perceptual inputs were utilized to direct the vehicle's behavior, enforcing established rules.

- Rule-based decision-making was introduced, encompassing critical actions such as throttling and braking in response to the presence of leading vehicles.

- Further, the system adhered to fundamental traffic regulations, halting at red lights and proceeding on green signals.

Experimental Results

- High confidence on YOLO car detection
  - Allows for a smooth throttle/break sequence depending on the cars ahead
- Mid confidence on YOLO traffic light detection
- Highly dependant on object detection
  - No red light detection may result in illegal passes

Experimental Results

- OpenCV FloodFill
  - Fell short on lane markings
- Limited sensor information bottleneck
  - Depending only on visuals
- Simulator FPS bottleneck
  - Frames are processed with 5 seconds delay

Discussion

- First Approach
  - Despite considerable efforts, first approach did not yield the desired performance
    - Minimal social grid (can only detect cars ahead)
    - Problems with lane detection on conjunctions
- Second Approach
  - Basic rules to handle basic scenarios
  - Hard to maintain rule-based approach to complex scenarios
  - Holistic approach to cover weaknesses of trained models

Outlook

- Advancing the field of autonomous driving will be guided by a concerted effort to refine and expand our methodologies

- Enhanced Dataset for Trajectory Planning
  - Augmentation of our training dataset for trajectory planning
  - By enriching our dataset with diverse and intricate driving scenarios
    - Cultivate a more robust and adaptable trajectory planning model.
  - Incorporating complex urban environments, various road layouts, and diverse traffic patterns will contribute to the efficacy of our system in real-world conditions.

Outlook

- Sidewalk and Pedestrian Detection
  - Addressing the safety of pedestrians and incorporating them into our system's perception framework is an imperative direction.
  - Our research will focus on the integration of sophisticated detection mechanisms to identify pedestrians and discern their presence on sidewalks.
  - The inclusion of pedestrian detection will not only enhance the overall safety of our autonomous driving system but also adhere to crucial ethical considerations.

Outlook

- Rule-Based Integration of Traffic Laws
    - An essential step towards enhancing the regulatory compliance of our autonomous driving system involves the integration of traffic laws.
    - This entails incorporating rules such as speed limits, overtaking restrictions, and yielding to emergency vehicles into our decision-making framework.
    - The integration of traffic laws not only improves the legal adherence of our system but also fosters safer interactions within the broader traffic ecosystem.

Outlook

- Continuous Learning and Adaptation:
  - In the dynamic landscape of autonomous driving, the capability for continuous learning and adaptation is paramount.
  - We aim to imbue our system with the capacity to learn from real-world driving experiences and adapt its behavior accordingly.
  - By leveraging reinforcement learning techniques and ongoing data collection, our system will evolve to navigate novel and complex scenarios.

References

1. Nvidia Blog - End-to-End Deep Learning for Self-Driving Cars
2. Nvidia - ChaufferNet
3. InterFuser - InterFuser
4. InterFuser - ReasonNet
5. Trajectory-Guided Control Prediction (TCP)

Thank you for listening!