INTEGRATED OBJECT DETECTION, ROAD SEGMENTATION, AND PATH FINDING FOR AUTONOMOUS VEHICLES WITH EXPLAINABLE AI

22AIE313 - Computer vision & Image processing

By, Team - 17, Batch A

TEAM MEMBERS

- Bhanushray Gupta [CB.EN.U4AIE22009]
- Gnanasabesan [CB.EN.U4AIE22018]
- Naga Koushik [CB.EN.U4AIE22046]
- Praneeth Tvs [CB.EN.U4AIE22054]

INTRODUCTION

Autonomous vehicles must be able to see, understand, and act—all in real time and with high reliability. But just accuracy isn’t enough. In safety-critical applications like self-driving, we need systems that are also explainable.

PROBLEM

- Traditional perception systems often operate as black boxes, offering little to no insight into how decisions—such as avoiding a pedestrian or selecting a driving path—are made. This lack of transparency poses serious risks in dynamic, real-world scenarios where safety is paramount.
- It addresses this challenge by designing an integrated vision-based system that combines:
  - **Object Detection** using YOLOv10 and Faster R-CNN,
  - **Road Segmentation** with EfficientPS and Mask2Former
  - **Real-Time Path Planning** with Hybrid A* with Cubic Spline Smoothing.
- What sets our approach apart is the integration of Explainable AI (XAI) techniques like SHAP and Grad-CAM, which allow us to visualize the model's focus areas and reasoning at every decision point. This not only boosts user trust but also enhances model accountability, which is essential for real-world deployment in autonomous vehicles.

RESEARCH GAP

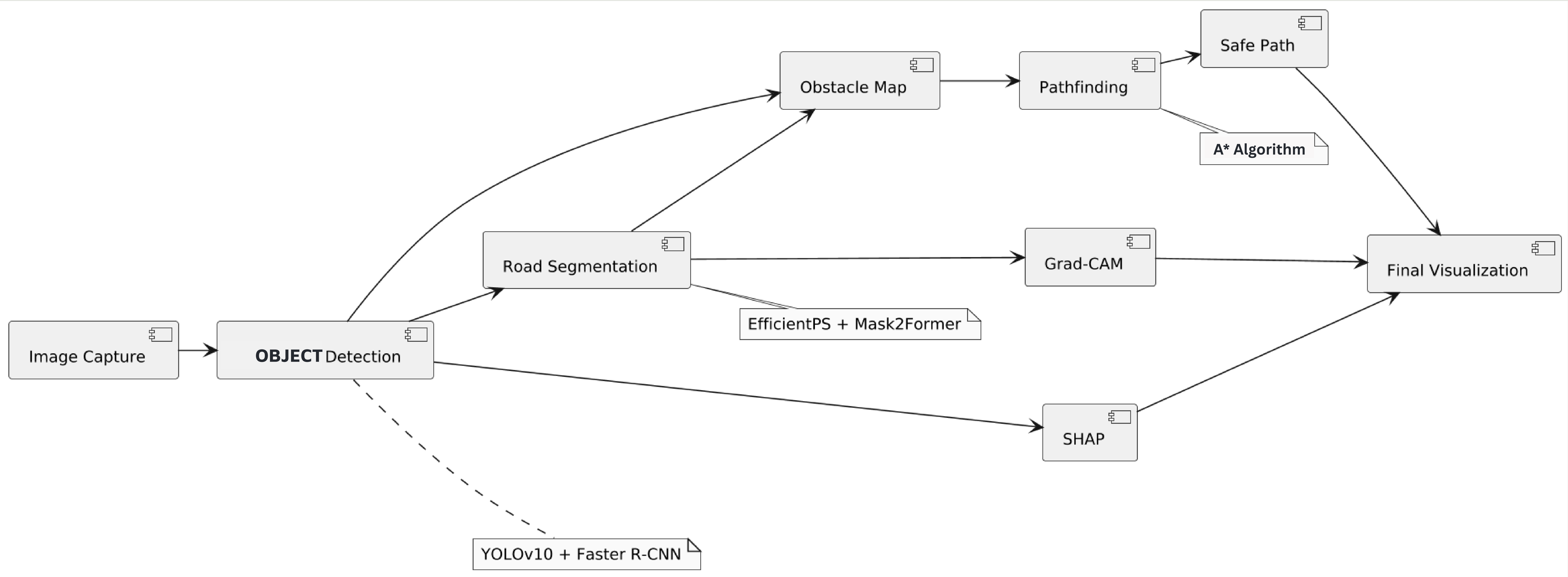
- Integrating explainable artificial intelligence (XAI) methods into autonomous driving systems to enhance transparency and societal acceptance
- Developing lane detection methods that effectively address challenges posed by road curvature and discontinuities.
- Creating hierarchical frameworks that refine trajectory planning for autonomous vehicles in intricate traffic environments to ensure safety and efficiency.
- Enhancing deep learning-based Object detection systems to overcome issues like occlusion and low-quality imaging.

DATASET

- **Object Detection:** We employed the KITTI dataset, renowned for its comprehensive real-world urban driving scenarios, to train and evaluate our pedestrian detection models.
  - This dataset contains 7481 training images with labels and 7518 testing images
  - In kaggle : https://www.kaggle.com/datasets/klemenko/kitti-dataset

- **Road Segmentation:** For road segmentation tasks, we used the Cityscapes dataset, which offers high-resolution images with fine-grained annotations of urban street scenes.
  - This dataset contains 2975 train and 500 val images with their masks stitched in the image
  - In Kaggle : https://www.kaggle.com/datasets/dansbecker/cityscapes-image-pairs

FLOW DIAGRAM



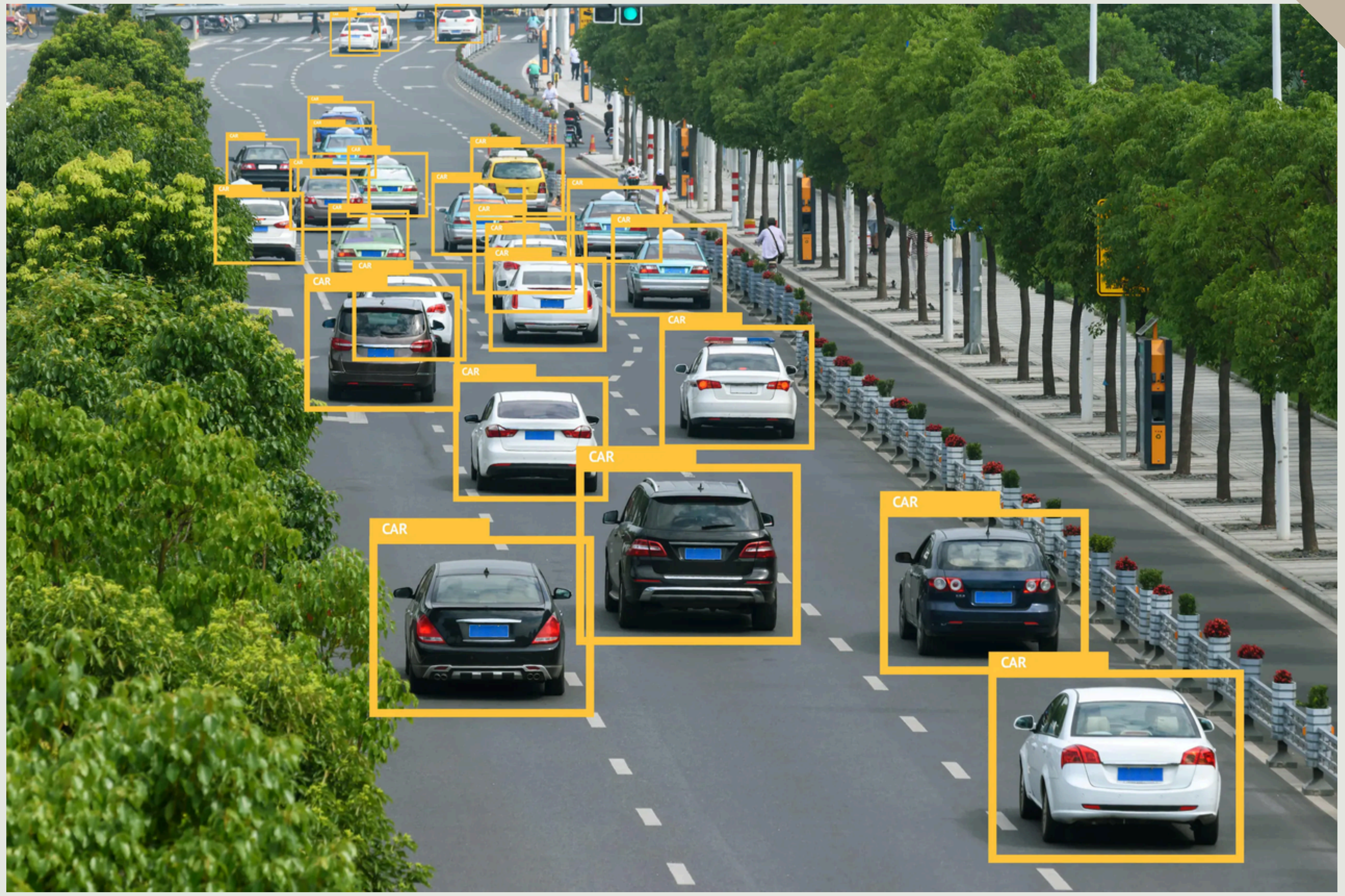
METHODOLOGY

Developed a robust, autonomous vehicle pipeline integrating deep learning models, classical planning algorithms, and explainable AI (XAI).

- Perception Module:
  - Object Detection
  - Road Segmentation
- Planning Module:
  - Pathfinding
  - Collision Avoidance
- Explainability Module:
  - Object Detection Interpretation
  - Road Segmentation Interpretation

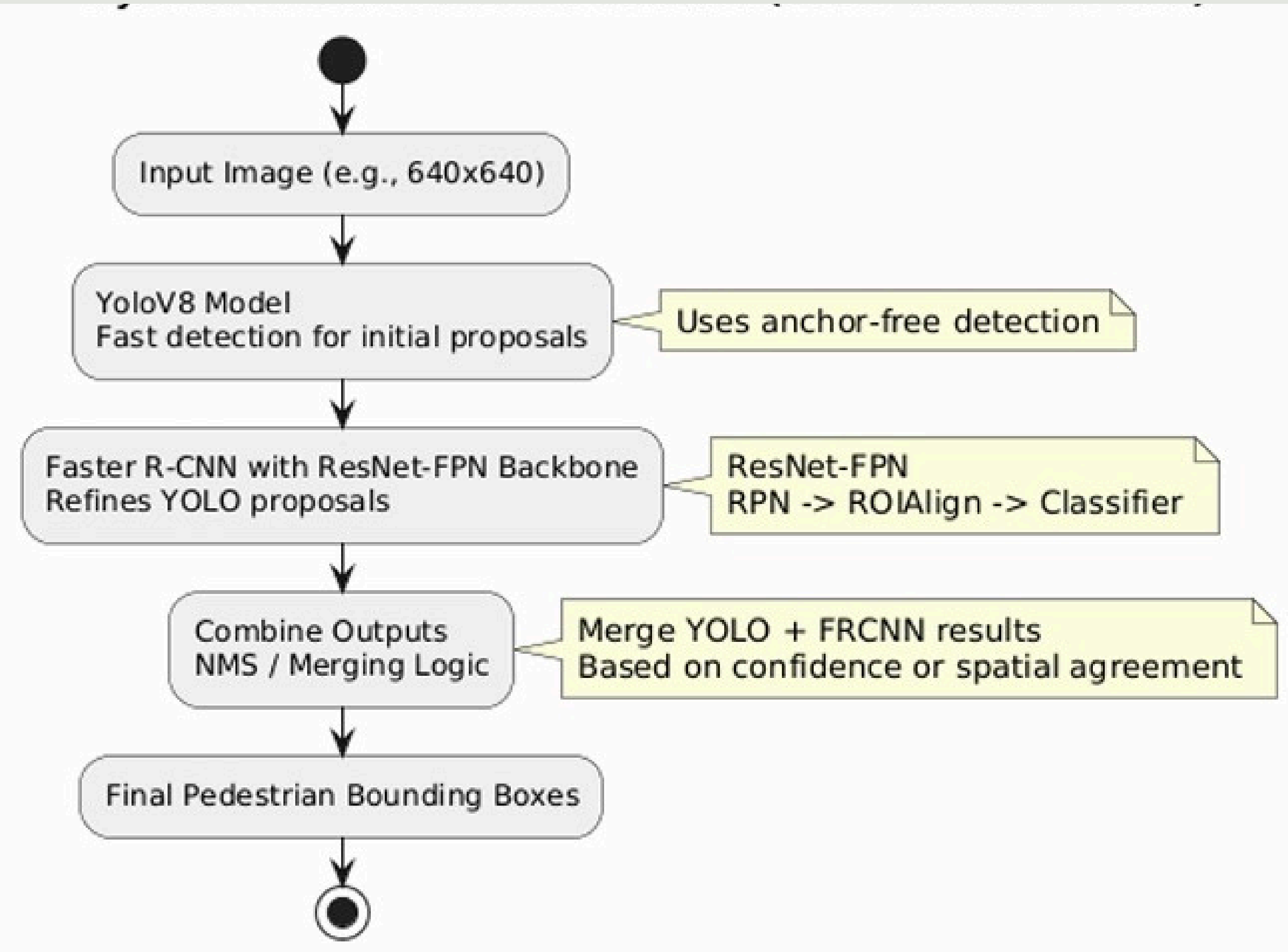
Object Detection

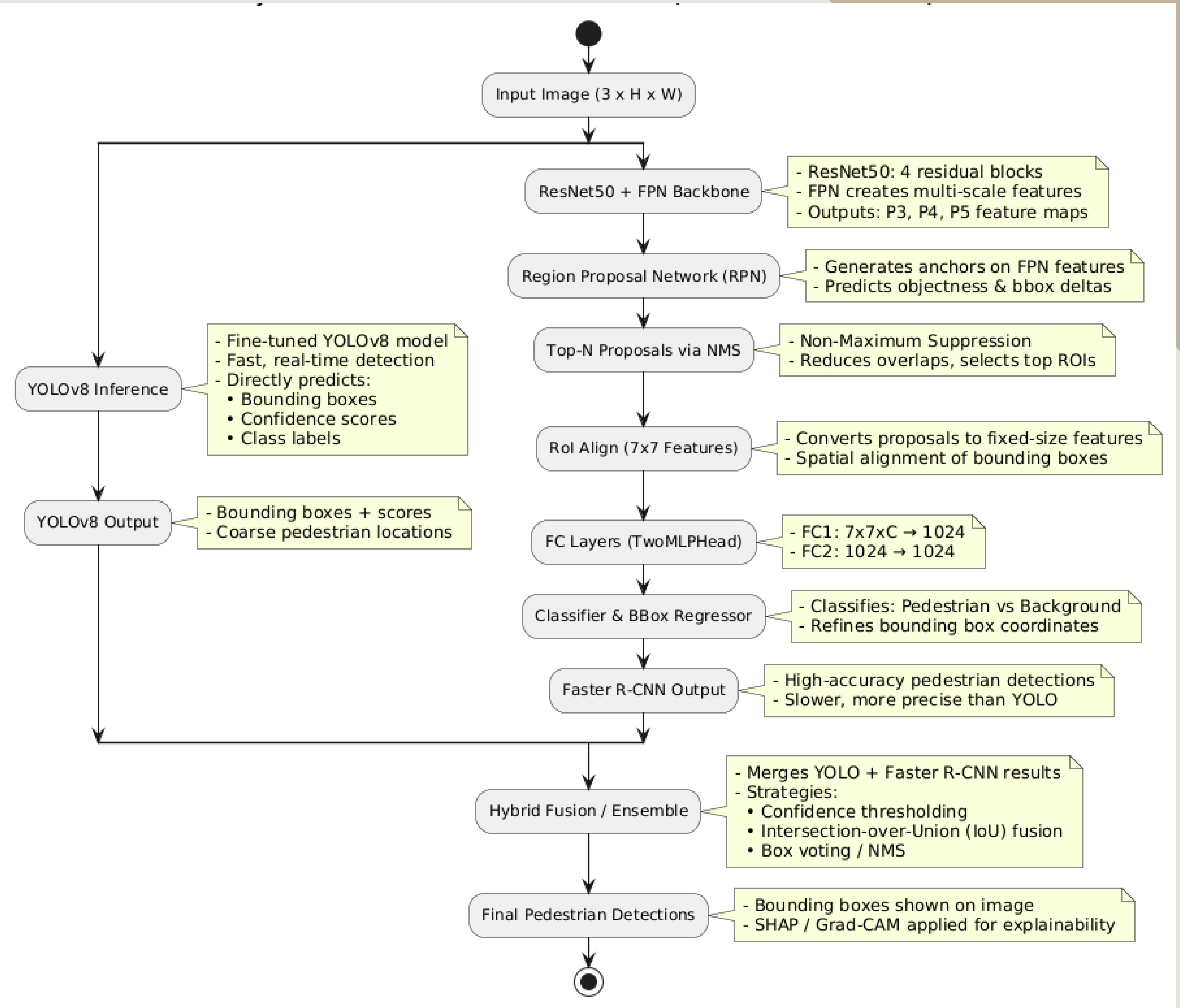
- Combined YOLOv10 (for speed) and Faster R-CNN (for accuracy).
- Trained on KITTI dataset.
- Used IoU threshold of 0.5 in Non-Maximum Suppression.
- Explainability via SHAP to identify image regions influencing predictions.



Object Detection:

Model Architecture



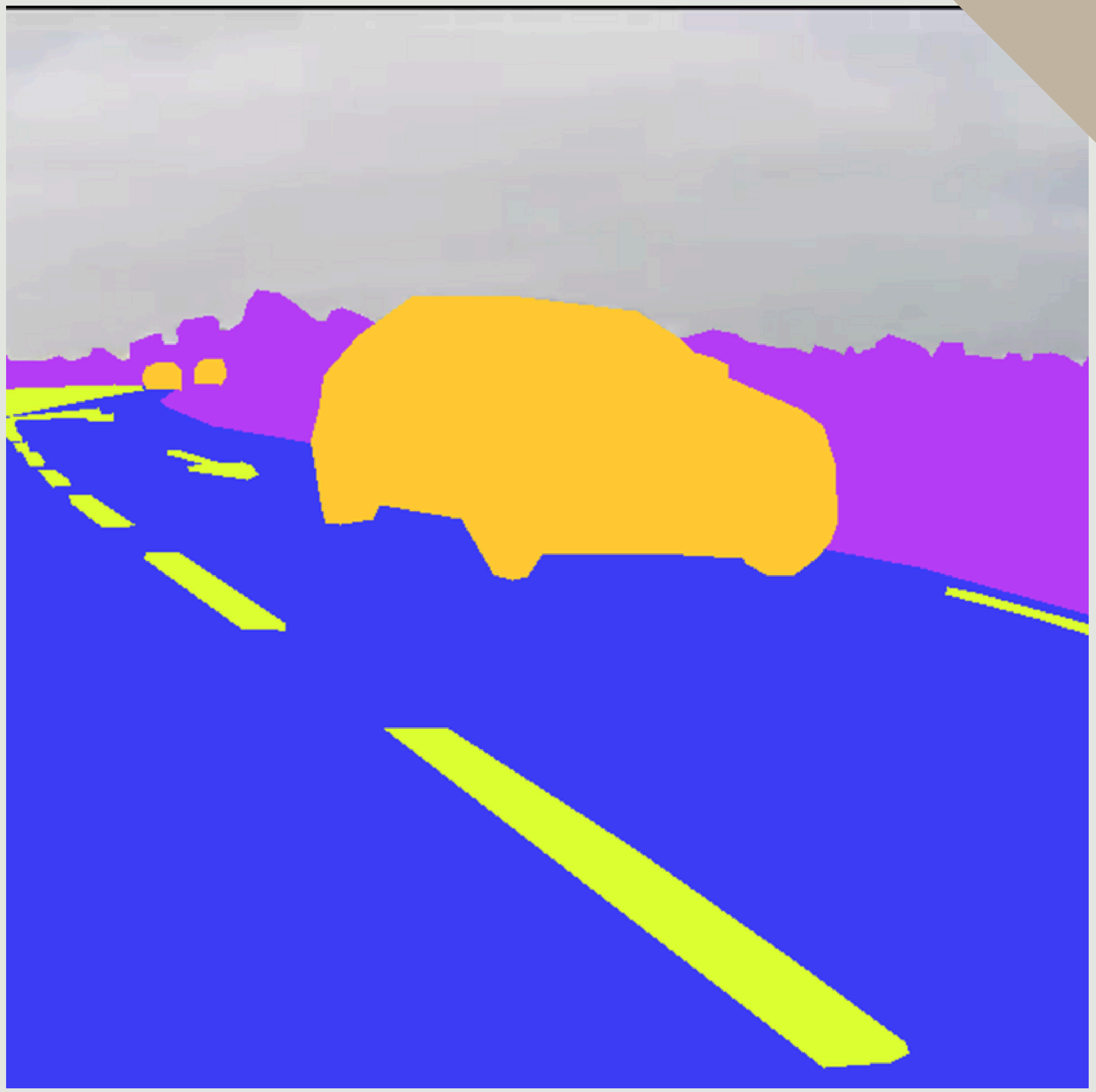


Object Detection: Performance Comparision

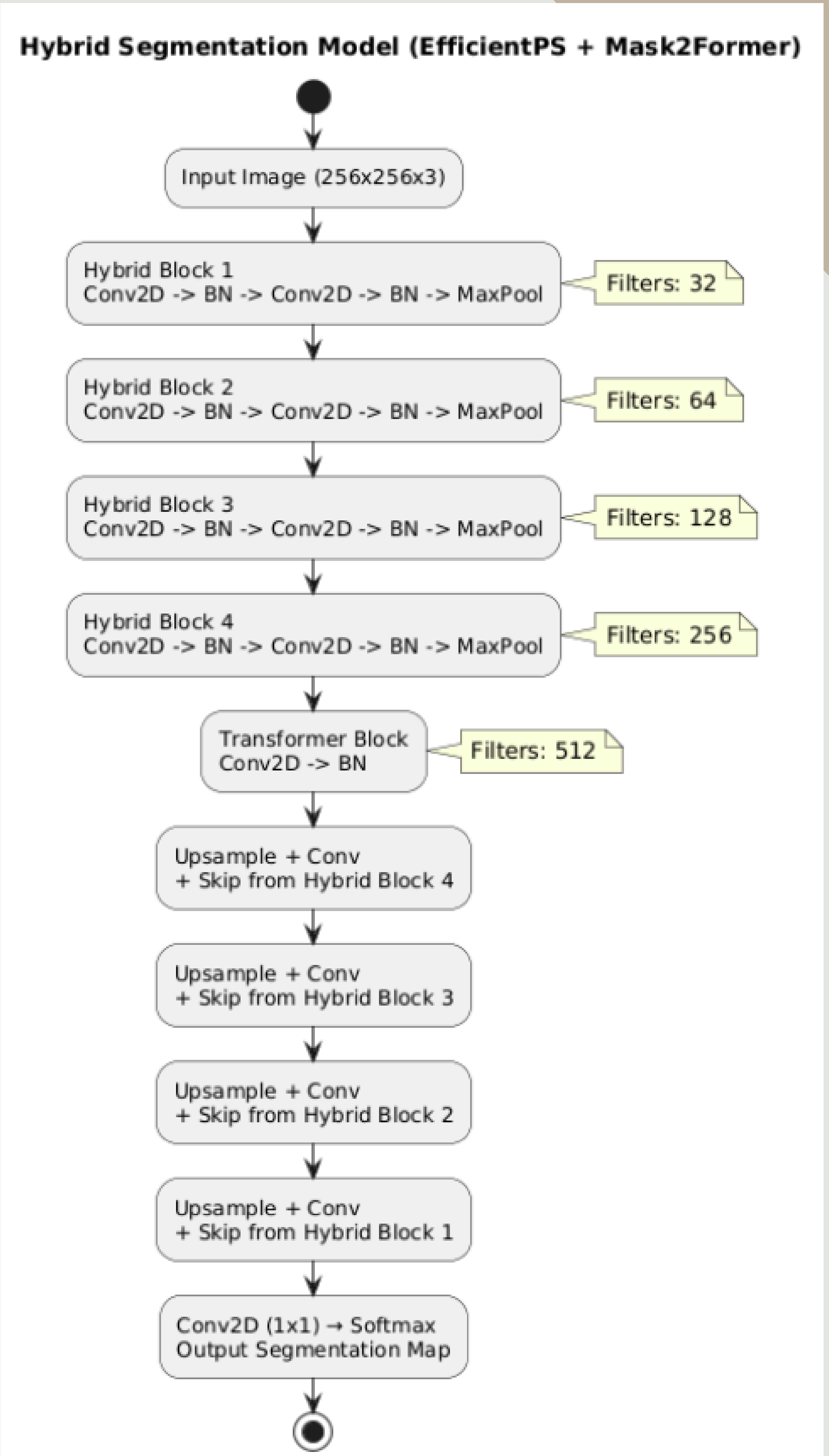
Model	mAP@0.5	Pedestrian mAP	Inference Time (ms)	FPS
Hybrid YOLOv10 + Faster R-CNN	0.85–0.92	0.87–0.94	20–30	33–50
YOLOv8	0.75–0.82	0.78–0.85	10–15	66–100
SSD	0.65–0.75	0.68–0.78	15–25	40–66
Standalone Faster R-CNN	0.80–0.88	0.82–0.90	50–60	16–20

Road Segmentation

- Used EfficientPS for multi-scale context.
- Added Mask2Former for transformer-based mask refinement.
- Trained on Cityscapes dataset with 13-class segmentation.
- Converted output to binary drivable mask.
- Used Grad-CAM for interpretability of segmentation layers.



Road Segmentation: Model Architecture

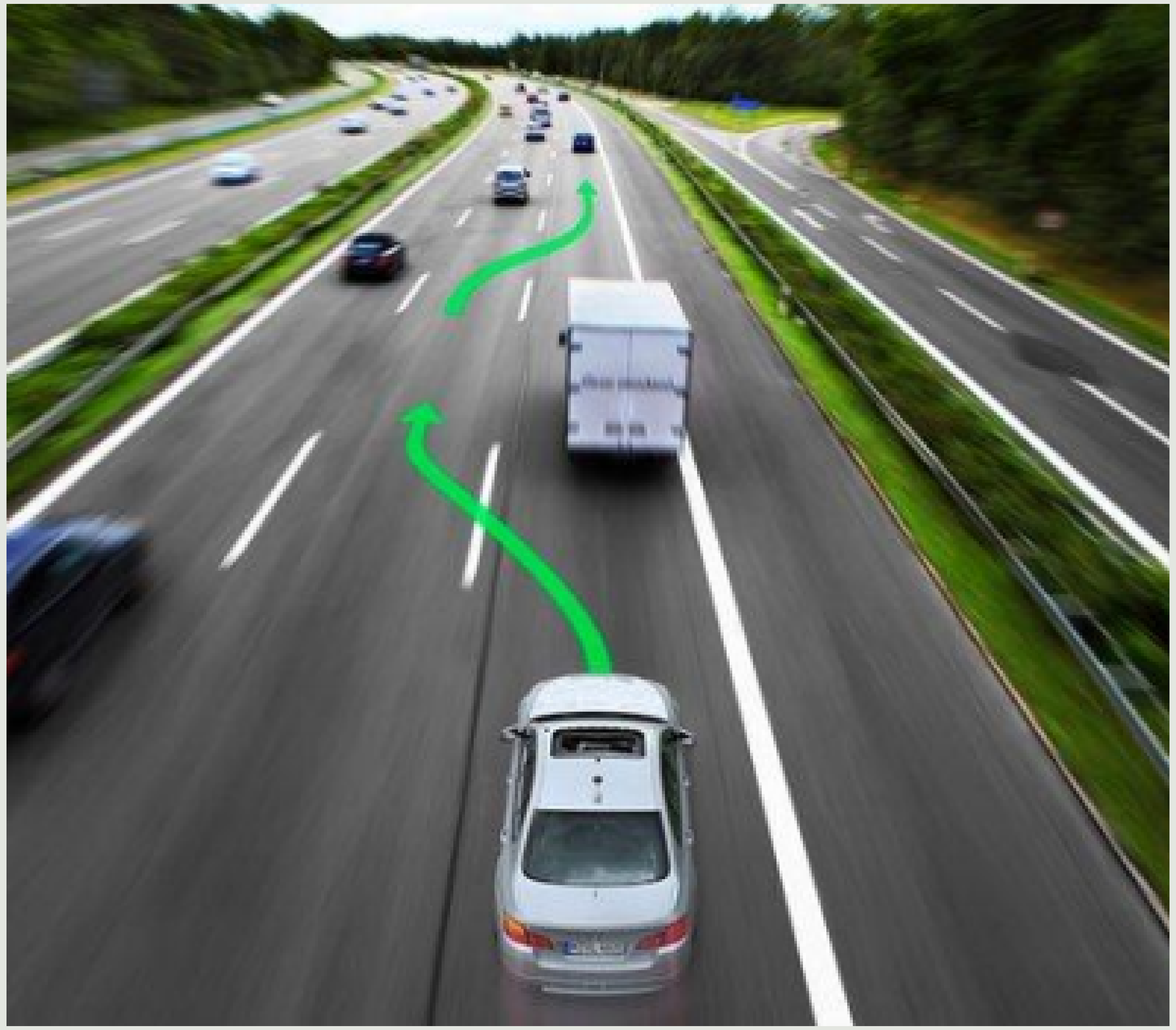


Road Segmentation: Model performance analysis

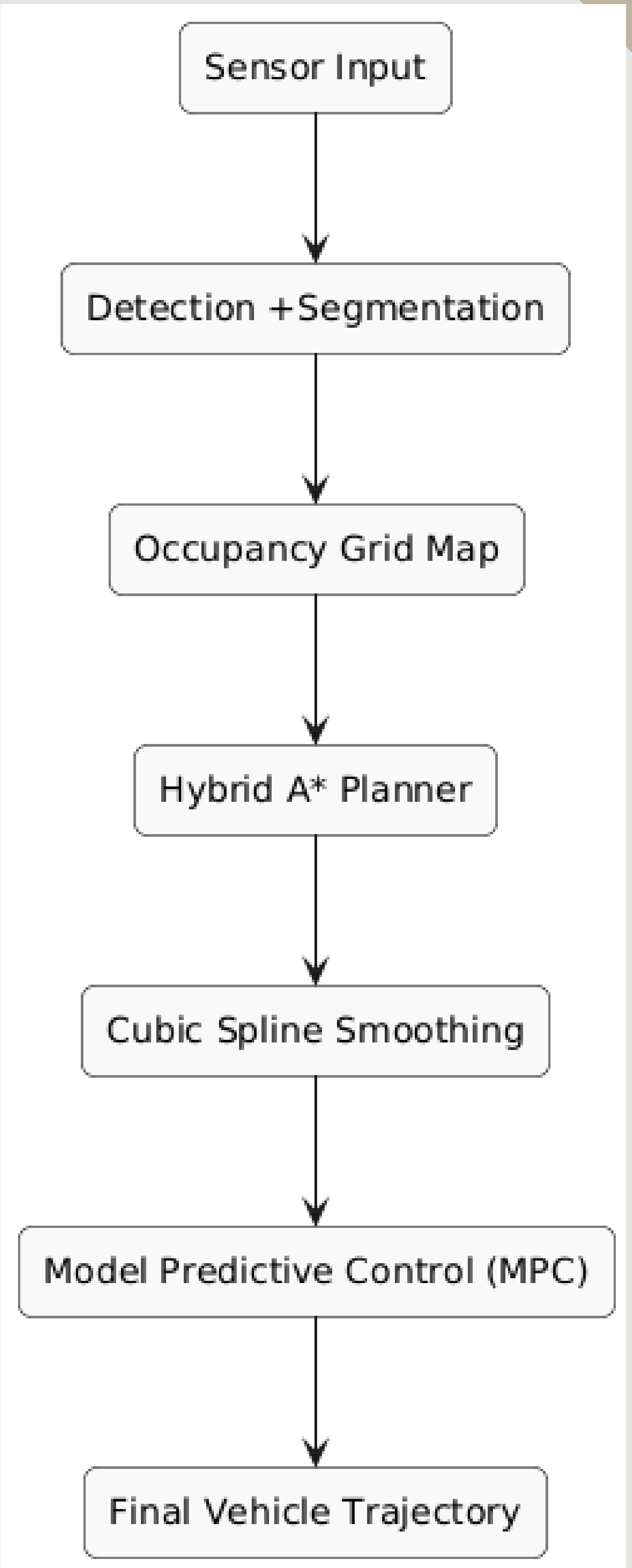
Model	Train Accuracy	Val Accuracy	Train Loss	Val Loss
Hybrid	82.86%	**80.05%**	0.5714	**0.6930**
U-Net	**80.86%**	78.48%	**0.6543**	0.7485
DeepLabV3+	78.71%	74.18%	0.6895	0.8927

Pathfinding Algorithm

- Input: Drivable mask + Object bounding boxes.
- A* Search Algorithm finds the shortest path on a road segmentation mask using a cost+heuristic strategy (Euclidean distance).
- Dynamic Goal Detection selects the safest and widest road segment at the top of the frame as the navigation target.
- Cubic Spline Smoothing generates a smooth, drivable path from discrete A* waypoints, ensuring real-world usability.
- Off-road Correction intelligently adjusts points back to the road if deviations occur, enhancing safety and reliability.



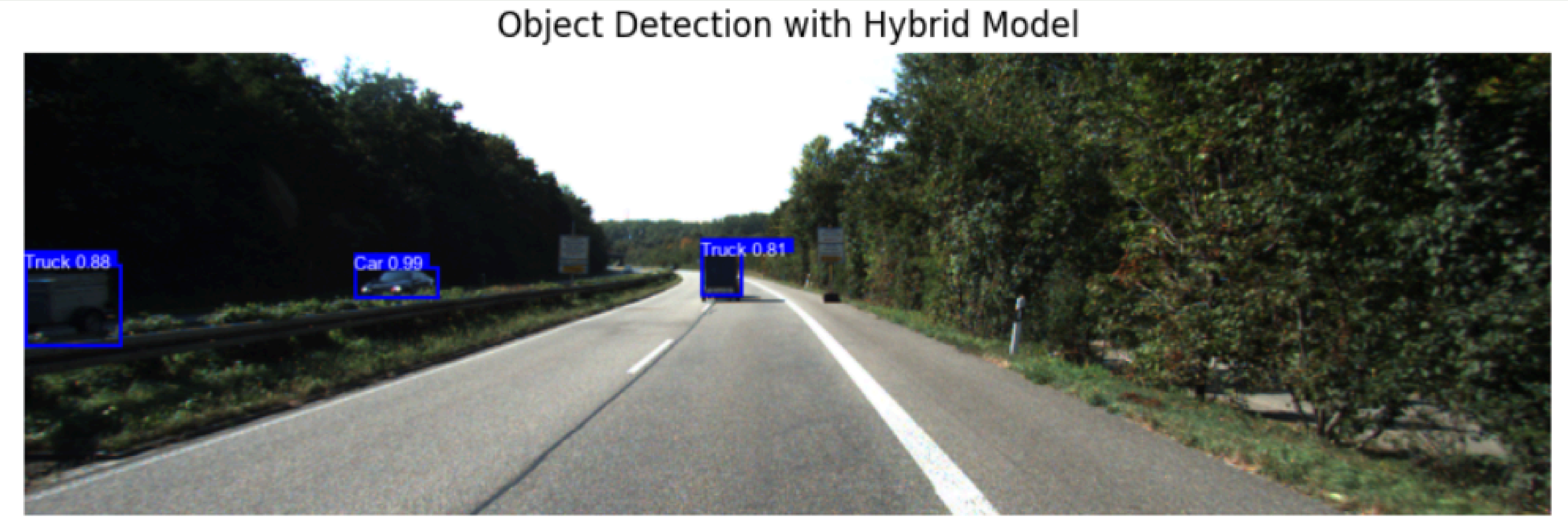
Pathfinding Algorithm Explained



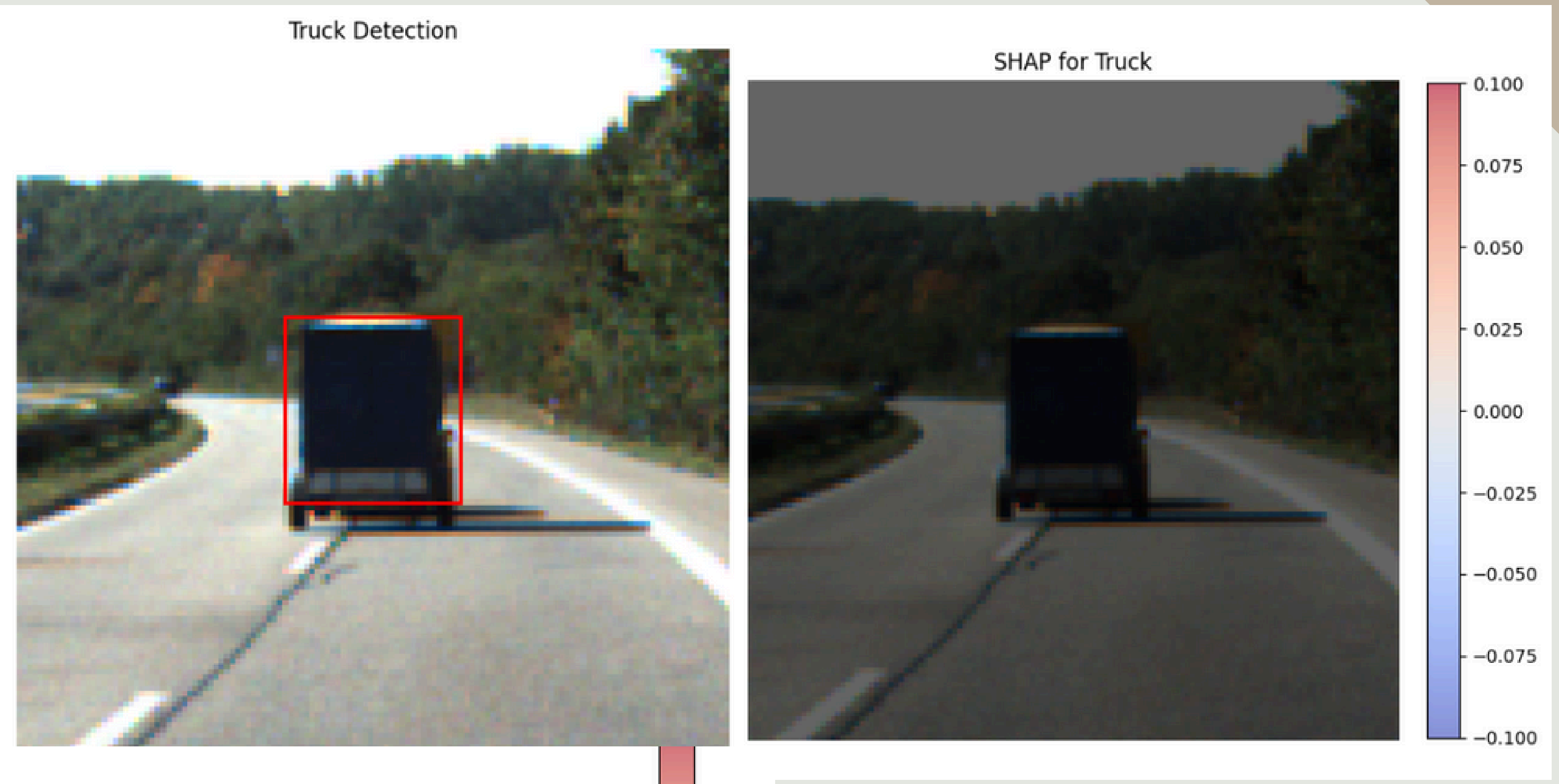
Overall Model Importence

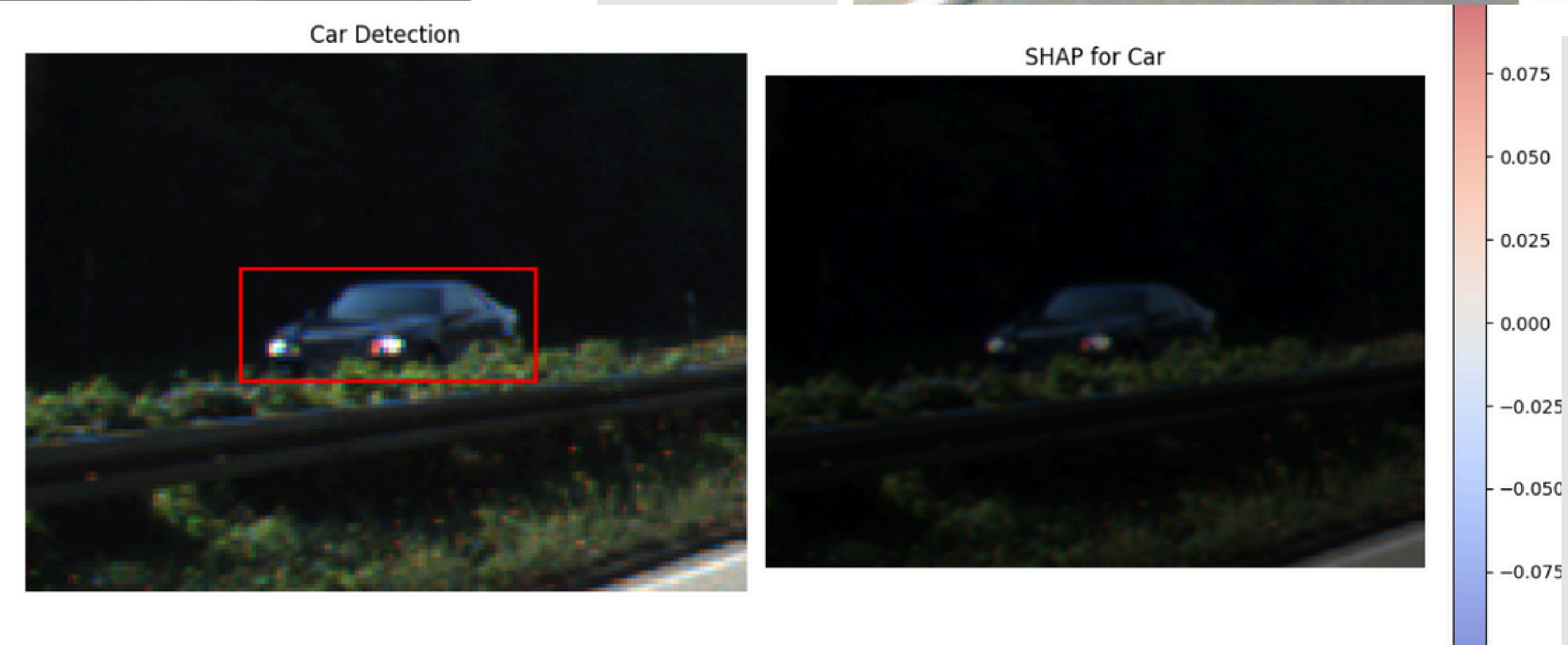
Component	Model	Dataset	Type	Notes
Object Detection	YOLOv10 + Faster R-CNN (ResNet50-FPN)	KITTI	Fine-tuned (YOLO) + Faster R-CNN (Custom head)	Fast inference with accurate bounding
Road Segmentation	Hybrid (EfficientPS + Mask2Former)	Cityscapes	From Scratch	Multiscale fusion with Transformer decoder
Pathfinding	A* Algorithm with Cubic Spline Smoothing	-	Custom	Ensures safe path
Explainable AI (XAI)	SHAP, Grad-CAM	Model Outputs	Post-hoc	Interpretability
Output	Safe Path with Overlay	-	Real-time	Integrated visualization

RESULT: OBJECT DETECTION

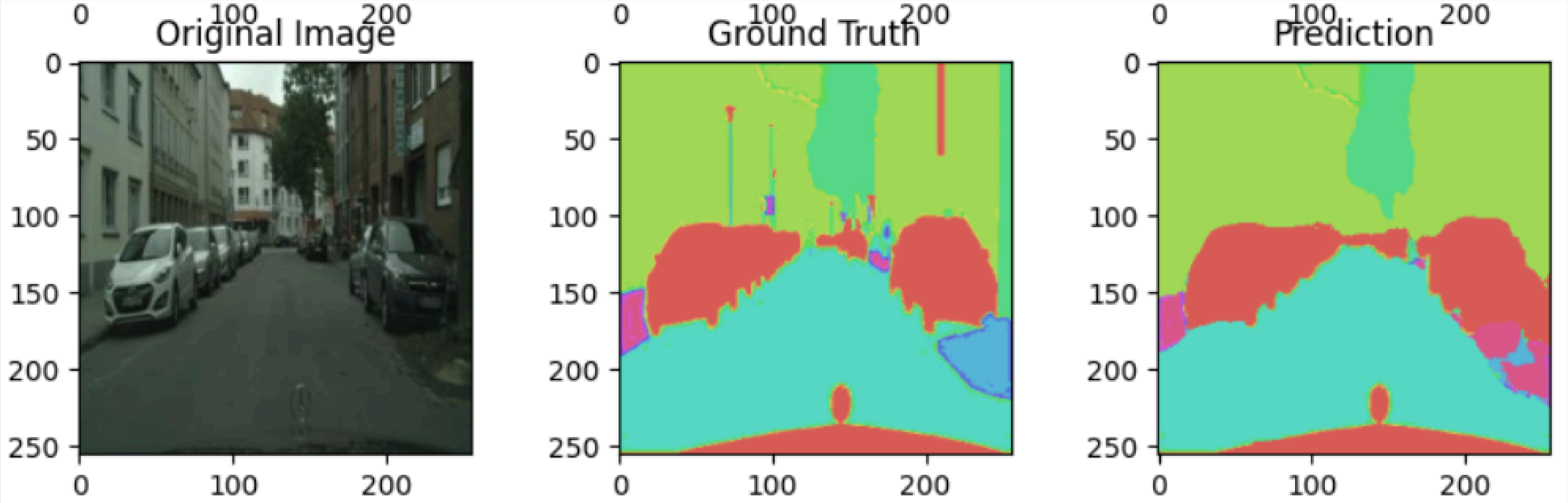


Result: SHAP XAI for Object detection

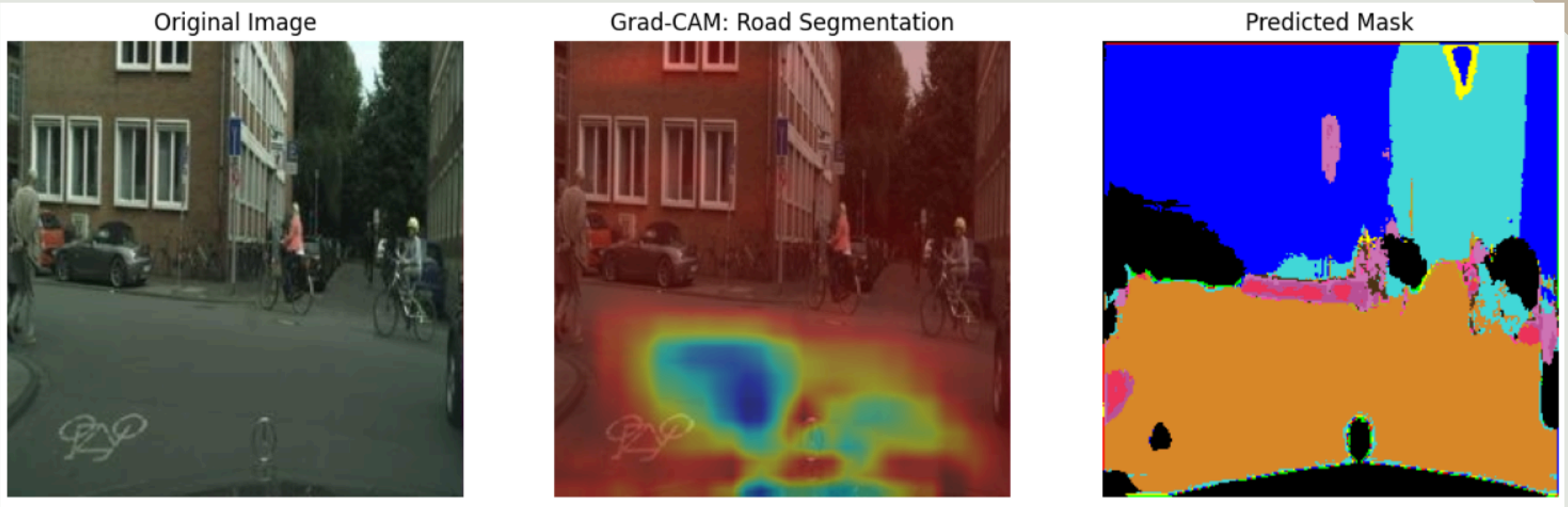




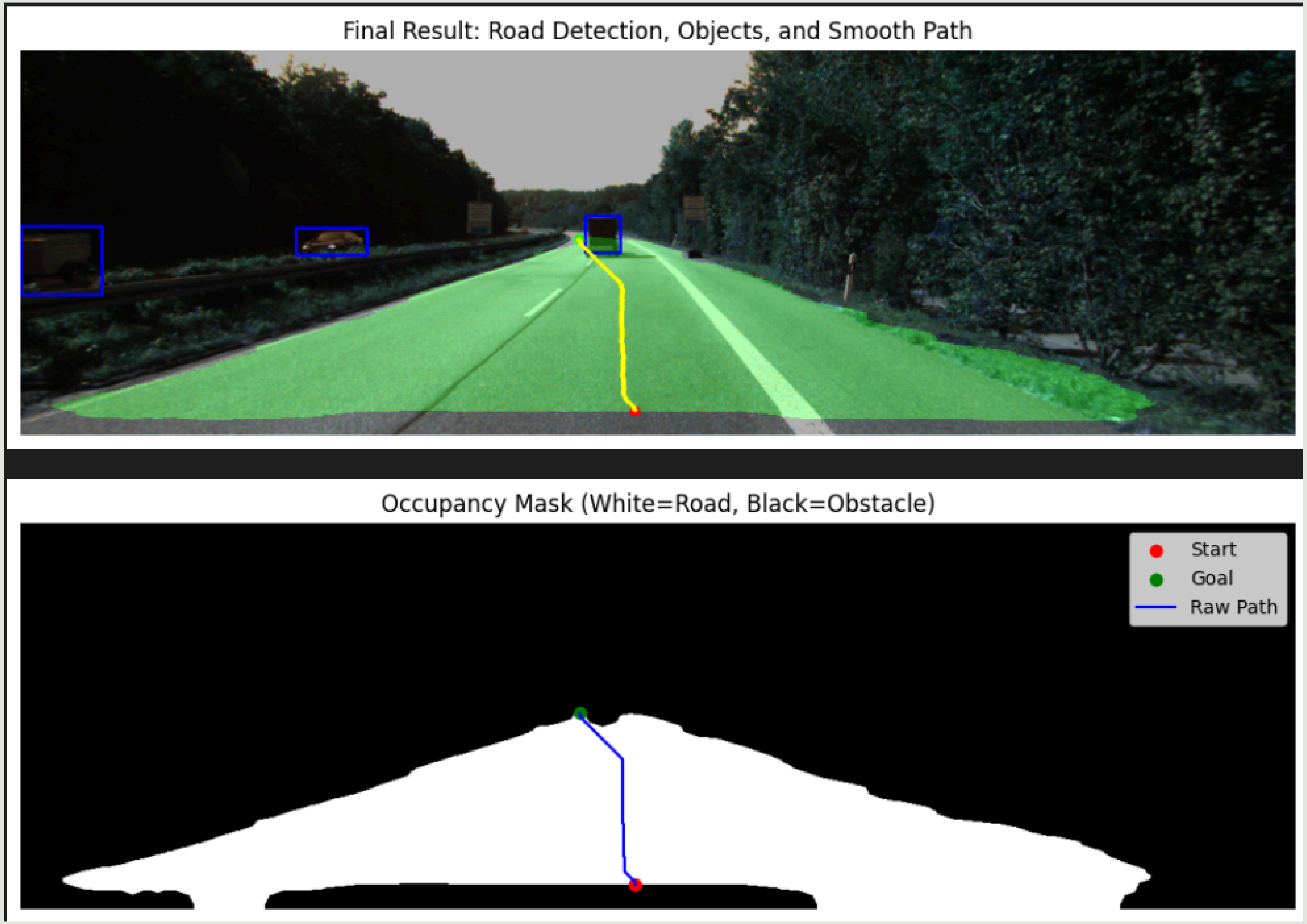
RESULT: ROAD SEGMENTATION



Result: Grad-cam XAI for Road segmentation



RESULT: PATH FINDING

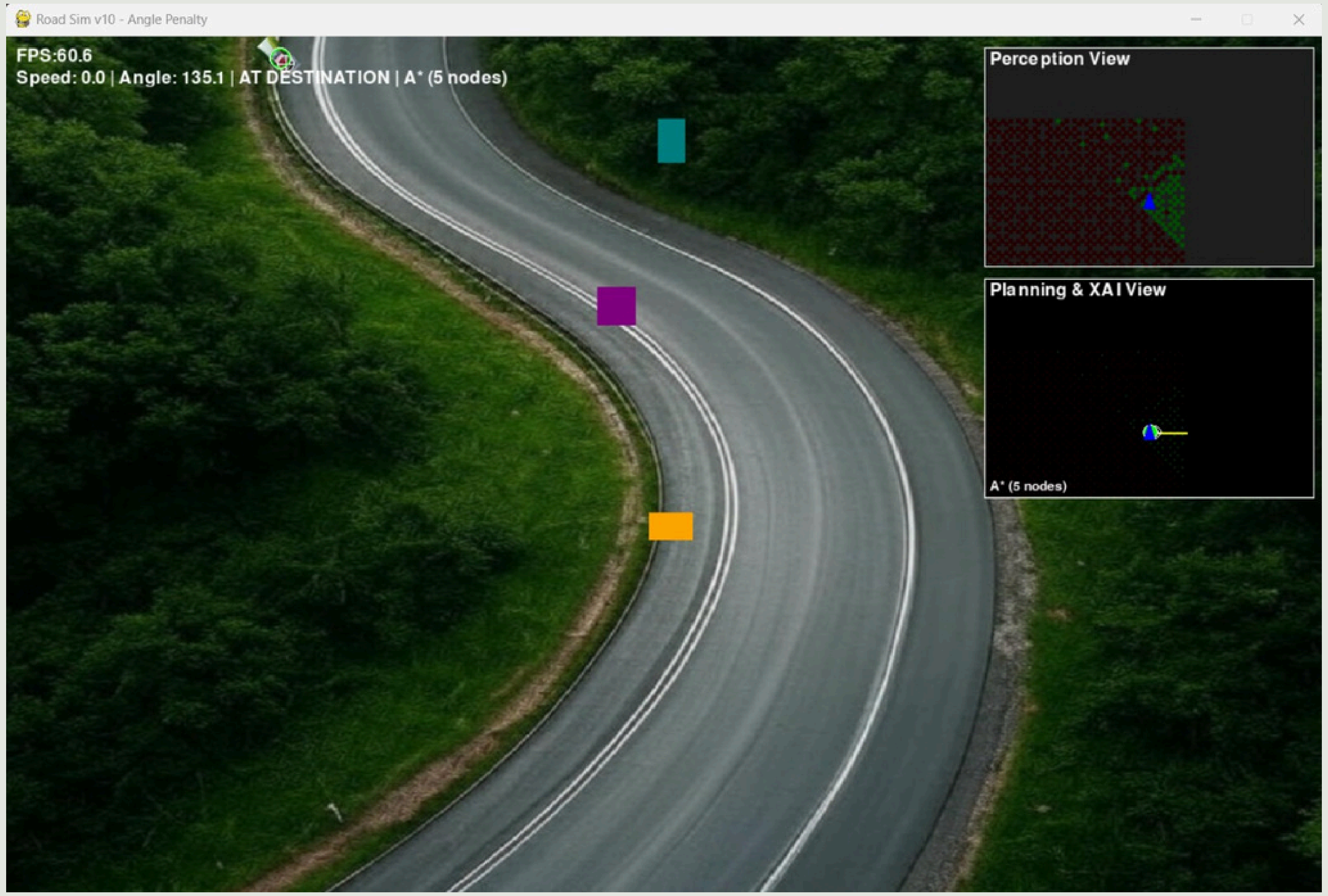


SIMULATION

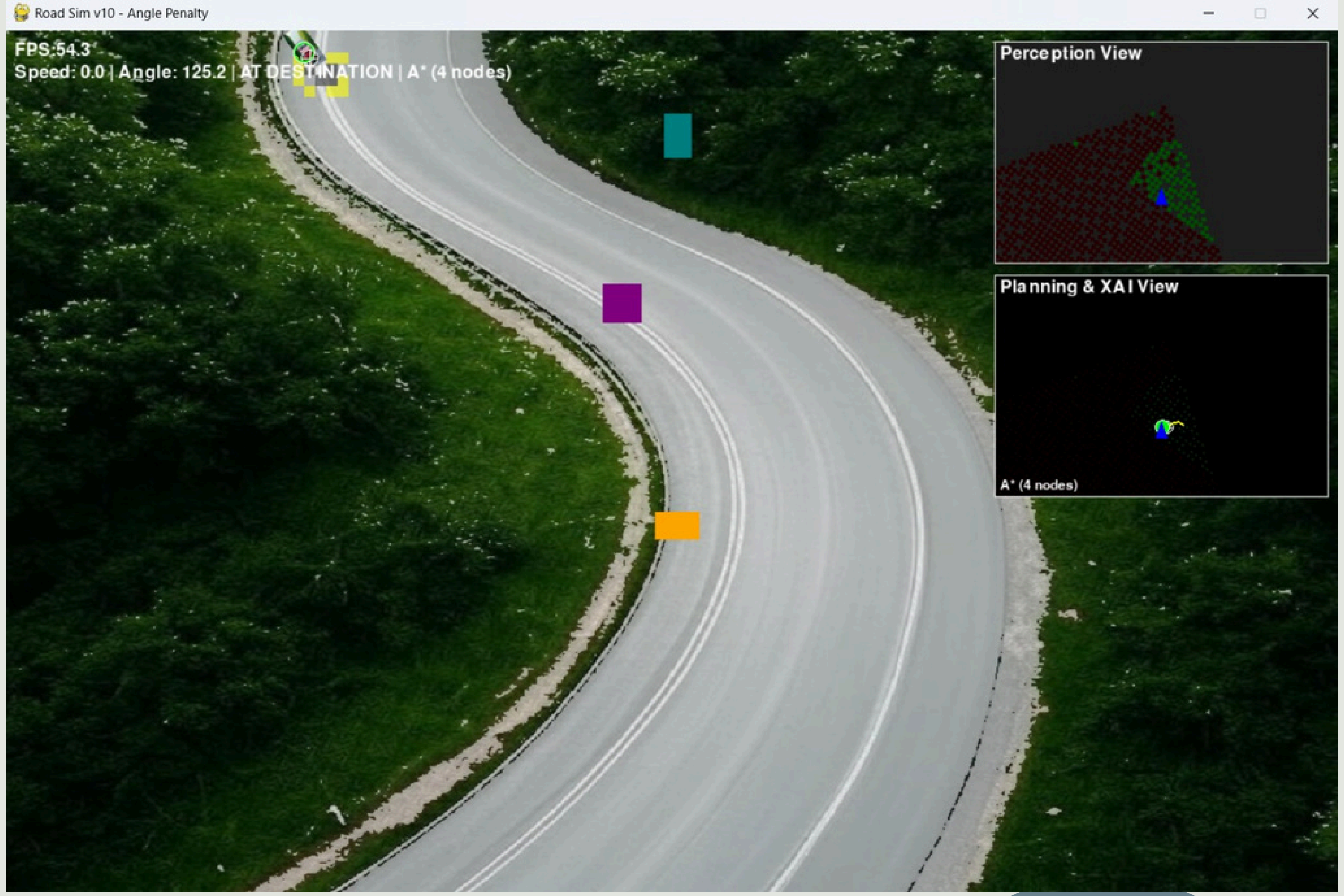
We designed a simple yet effective simulation that mimics real-world autonomous driving scenarios.

- A* Algorithm is used for Path Planning.
- Road background image and Car image are taken for simulation.
- Boxes are treated as obstacles/objects.

SIMULATION



Road Sim v10 - Angle Penalty
FPS:60.6
Speed: 0.0 | Angle: 135.1 | AT DESTINATION | A* (5 nodes)
Perception View
Planning & XAI View
A* (5 nodes)



Road Sim v10 - Angle Penalty
FPS:54.3
Speed: 0.0 | Angle: 125.2 | AT DESTINATION | A* (4 nodes)
Perception View
Planning & XAI View
A* (4 nodes)

CONCLUSION

This project proposes a modular pipeline for autonomous vehicles by integrating object detection, road segmentation, and pathfinding with explainable AI. YOLOv10 and Faster R-CNN ensure accurate Object detection, while EfficientPS and Mask2Former handle road segmentation.A* Search Algorithm-based planner enables safe navigation, and XAI tools like SHAP and Grad-CAM enhance transparency and trust in decision-making.

REFERENCES

- Explainable Artificial Intelligence for Autonomous Driving: A Comprehensive Overview and Field Guide for Future Research Directions"
- "Dynamic Lane Segmentation for Autonomous Vehicles Using Neural Networks"
- "On-Road Trajectory Planning of Connected and Automated Vehicles in Complex Traffic Settings: A Hierarchical Framework of Trajectory Refinement"
- "Deep Learning-Based object Detection in Autonomous Vehicles: Substantial Issues and Challenges"

Thank You