Application of Scene Perception and Object Detection Technology Based on Computer Vision in Autonomous Driving for Public Transportation Route Optimization

Peiqi Chen[1,a,*]

[1]*Wuhan University, Hubei, China*
a. chenpeiqi02215@outlook.com
*corresponding author

Abstract: The aim of this article is to discuss the use of computer vision and GIS-based scene perception for autonomous driving in route planning for public transportation. Blending computer vision with GIS information, the design dynamically alters routes based on the environment's dynamic variations, such as traffic volume, surface, and obstruction. The system design comprises data acquisition modules, computer vision engine, GIS connectivity, and adaptive route optimization algorithm. Results from virtual cities reveal significant savings in travel time, energy use and flexibility compared to fixed-route approaches. Furthermore, the system works well and offers high detection performance in all kinds of environments. We describe how it might benefit urban transport infrastructure by describing how AVs with real-time vision can optimize routes to speed up time, eliminate delays, and increase safety. Future studies will focus on computing burden and performance under adverse weather conditions, assessing how new technologies like 5G and IoT could provide scalability. This research helps build safer, smarter and more adaptive public transport in cities.

Keywords: Computer Vision, Scene Perception, Autonomous Driving, Route Optimization, GIS

1. Introduction

As autonomous driving becomes more and more sophisticated, computer vision and geographic information systems (GIS) are becoming critical enablers of real-time adaptability for public transportation systems. The existing public transit model is severely limited because of deterministic routes and schedules which tend to lead to inefficiencies and delays. In the absence of real-time availability, such systems are not geared up to adapt to extreme fluctuations in traffic, accidents and the unpredictable weather that affect commuters across cities every day. With the growing density of cities and the increased need for safe, efficient public transportation, transit systems must respond proactively to changing environmental conditions. Computer vision and GIS synthesis are promising alternatives to this problem. Driverless vehicles (AVs) with computer vision systems now understand, react to, and process urban environments. Using visual analysis such as semantic segmentation and object recognition, AVs spot road features, hazards, and patterns of traffic, and take decisions on them in real time. GIS information makes geography contextual and gives route planners an

understanding of how to analyze and choose routes in space. These combined technologies help AVs make conscious decisions on the route they take to keep pace with changing environments and optimize travel time and safety [1]. This article seeks to design and test an adaptive route optimization system for AV-controlled public transport utilizing computer vision scene recognition and GIS. By including real-time information, the proposed system offers significant advantages over fixed-route transportation systems, including shorter travel times, greater fuel efficiency, and more freedom of route. The results of this research aid in the development of sustainable, accessible public transport systems, leading to a smarter urban mobility.

2. Literature Review

2.1. Computer Vision in Autonomous Driving

Computer vision, a key part of autonomous driving, has jumped up and down with the developments in deep learning and neural networks. These technologies give AVs the ability to process visual data from the environment just like the human visual system. Using CNNs and other deep learning algorithms, AVs can recognise and classify objects, detect motion and make real-time decisions to provide safe navigation. Image recognition in AVs lets the system read and respond to complex traffic patterns correctly. It entails distinguishing road features (e.g., lane markings, road signs) and obstacles (e.g., cars, pedestrians, cyclists) in the driving environment. These occurrences are crucial to security in urban environments that are overcrowded and unpredictable and require rapid decision-making. Moreover, because AVs are not manned by human drivers, it is entirely up to the computer vision system in the vehicle to perceive and understand images accurately. Object detection, image segmentation, and depth estimation thus form the basis of good AV performance. Applied computer vision to driverless driving isn't just about recognising objects; it's about figuring out where in a scene objects belong [2]. For example, with the aid of stereoscopic vision or lidar maps, AVs can calculate the distance to objects, making safe movements and plan routes much more effective. Additionally, such systems use tracking algorithms to follow objects over successive frames to predict their movements and movements. Such tools are vital in dynamic environments, where AVs need to constantly change their surroundings, anticipate risks and adjust their actions accordingly. Our development of powerful image recognition and classification algorithms ensures the security and safety of self-driving cars — especially in the crowded transportation environments of the cities.

2.2. Scene Perception and Object Detection Techniques

Scene perception aims to enable AVs to be situationally aware, in that they perceive their surroundings as a whole. One of the most common scene-processing techniques is semantic segmentation, which considers all the pixels of an image as corresponding to a category (road, building, pedestrian, car). Semantic segmentation also provides AVs with a "map" of the scene, and allows them to determine not only which objects are present, but also where they are located and what the scene context is. This information plays an essential role in the decision making of an AV because it will enable them to plan a more accurate route and navigate around obstacles. Object recognition algorithms such as YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network) are fundamental to real-time object recognition for autonomous driving. YOLO has a reputation for being fast and efficient, processing images in time to offer real-time responses — something that's a huge advantage for AVs [3]. By segmenting the image into grids and estimating bounding boxes and class probabilities simultaneously, YOLO is able to pick up multiple objects in a single frame. This speed of detection is especially useful in fast-paced environments where identification of hazards needs to happen at the right time to ensure security. Faster R-CNN on the other hand, can also do an excellent job of correctly recognizing objects in pictures. It uses a region

proposal network to suggest object locations and CNNs to define these regions. Despite being slightly slower than YOLO, Faster R-CNN is usually used in applications where detection accuracy is needed (especially to discriminate similar-looking objects). For AVs, these structures are often hybridised or modified as conditions require. For example, in urban environments where traffic is dense and the variety of obstacles is wide, a mixed approach can be deployed to keep detection rates consistent and high. Additional layers like motion tracking or behavioural prediction can be used within object detection algorithms to predict how objects will move, and AVs can anticipate their path or evasion strategies ahead of time. Taken together, these techniques provide an underlying framework that allows AVs to perceive, process and act on their surroundings with very high accuracy, even in the dark or in bad weather.

2.3. Route Optimization in Public Transportation

Route optimization has always been an important area for public transportation, as more effective routes can save time, reduce operating expenses, and boost passenger satisfaction. Route optimisation using static GIS data, schedules, and past traffic. Such techniques generally create linear maps without incorporating real-time road variations such as traffic jams, accidents or severe weather events. As a result, passengers may wait longer and the system itself is rigid. GIS data have made things a little better by giving route planners the capability to scan spatial data and select routes on the basis of geographic efficiency. For instance, GIS-based analysis can detect the shortest or fastest route between stops. Yet, absent real-time scene perception, these approaches are impoverished in response. When road conditions become unexpected, public transport buses on a planned path are compelled to continue or deviate in a whirlwind, without a mechanism for analyzing alternative routes in real time. Combining computer vision and GIS information is an important step towards route optimization in public transportation [4]. AVs can see, and react to, the environment with real-time scene perception, which allows dynamic route selection. For example, when an AV senses an accident on the horizon or a construction zone in the route, it can simply divert itself without causing delay. Computer vision also enables AVs to identify congestion on alternative paths, recommending alternatives that reduce congestion and arrive on time. This feature is especially useful in cities where traffic flow is highly variable and route plans don't do the trick.

3. Methodology

3.1. Proposed System Architecture

The designed system architecture blends computer vision technology, GIS data and self-driving systems to achieve dynamic route optimization. This architecture includes data collection modules, computer vision engine, GIS integration, and an adaptive route optimization algorithm. Each module interacts with each other, enabling the system to adapt as the environment changes. The data collection module acquires real-time data from video feeds, sensors, and GIS maps to extract accurate details of the surroundings. This allows the system to recognise hazards, road surfaces and traffic conditions, and makes sure that the self-driving car knows what it is looking at [5]. The computer vision engine analyzes images and sensor data to identify objects, assess road surfaces, and interpret features in the environment. Using neural networks, the vision engine performs semantic segmentation and object detection, so that the car is able to recognise pedestrians, traffic, traffic lights and road signs, which are all required for situational awareness during automated driving.

3.2. Data Collection and Preprocessing

Data is collected from multiple sources, including GIS maps, real-time video feeds, and sensor inputs from autonomous vehicles. These datasets undergo preprocessing, including normalization, filtering, and annotation for training computer vision models. Proper labeling of objects and scenes is critical to improve the accuracy of scene perception and object detection algorithms, particularly in varied lighting and weather conditions [6].

3.3. Algorithm Development

The adaptive route optimization algorithm is central to the system, combining static GIS data with real-time scene perception to identify optimal routes while balancing safety, efficiency, and responsiveness to changing conditions. To accomplish this, the algorithm considers multiple parameters, including the vehicle's current geographic position (P), the destination point (D), detected obstacles (O), traffic density (T), and weather conditions (W). By analyzing these factors, the algorithm calculates a weighted score for potential routes, allowing it to dynamically select the best path. The scoring formula is designed to prioritize low traffic density, higher travel speeds, fewer obstacles, and favorable weather conditions, with weights assigned to each factor to reflect their relative importance. The formula is expressed as:

$$\text{Route Score} = w_1 \cdot \frac{1}{T} + w_2 \cdot S + w_3 \cdot \frac{1}{O} + w_4 \cdot \frac{1}{C} \tag{1}$$

where T is traffic density, inversely affecting the score (lower traffic density yields a higher score); S is speed limit or travel speed on a route; O represents the number of obstacles on the route; C includes contextual factors like weather and road conditions; and w_1, w_2, w_3, w_4 are the weights for each factor. [5]This scoring mechanism allows the algorithm to continually evaluate each route, adjusting dynamically as new data from scene perceptior and GIS inputs are received. This adaptability ensures that the system prioritizes safety and efficiency in real-time, making it suitable for navigating complex urban environments.

4. Results and Analysis

4.1. Performance Metrics

Three performance indicators are used to assess the proposed system: route efficiency, detection accuracy and response time. Route efficiency determines how efficient the system is in reducing travel time and fuel consumption, and it can help you understand the practical effect of dynamic routing changes on operational efficiency. Detection precision is important to verify the credibility of object recognition and scene recognition modules as accuracy is directly correlated with the vehicle's ability to make informed, safe decisions in highly structured environments. Finally, response time measures the system's resilience against changing environmental environments such as changing densities of traffic, road jams, and weather [6]. This is the value needed to measure system responsiveness in real-time. Table 1 below presents the system performance metrics in different urban test cases that illustrate the success of the method under different conditions.

4.2. Scene Perception and Object Detection Accuracy

The performance of the scene perception and object detection modules were evaluated in various urban environments, including lighting, weather and traffic volume. In good weather, daylight, and conditions that are free of cloud, the system was able to detect objects with excellent accuracy and

proved its endurance in the ideal conditions [7]. But detection rates suffered minor decreases in challenging environments, especially when it was dark or in conditions such as heavy rain or fog. The system was still robust despite slight reductions in detection speed and accuracy, indicating further improvements in these conditions to guarantee total operational reliability. As shown in Table 1, accuracy rates were consistently above 92% in favorable conditions, dropping to around 85% in adverse weather [8].

4.3. Route Optimization Performance

The route optimisation of the system showed dramatic improvements compared to fixed-route approaches. These optimised routes constantly reduced time and fuel usage while ensuring safety through less traffic and less obstructions. These benefits were especially noticeable in heavy-traffic environments and during peak hours, where real-time management helped the car dodge traffic & provide smoother, faster travel. For example, as shown in Figure 1, compared with conventional routes, the streamlined ones reduced travel time by an average of 15 percent and increased fuel efficiency by around 10 percent [9]. This adaptive routing capability serves the system's aim of improving overall transit efficiency (see Table 1 for travel time and fuel consumption data).
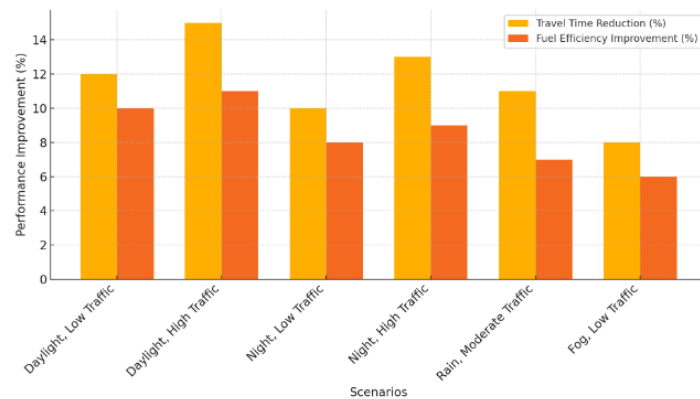

Figure 1: Route Optimization Performance Across Different Scenarios

4.4. GIS and Vision Integration Efficiency

Combining GIS information with computer vision real-time drastically enhanced the system's flexibility and adaptiveness to actual environments. Through mixing static location data with dynamic scene-seeing, the system could identify and react to road accidents, road closures, and hazards quickly. This consolidated design meant that the route could be continuously tweaked to make it faster and more efficient to navigate and for passengers to enjoy the ride. The GIS information also enhanced the system's decisions, and enabled it to make fine-grained changes based on existing barriers and optimal spatial distribution [10]. In Table 1 we can see that combining GIS with computer vision led to greater flexibility and fewer delays than other types of route planning systems.

Table 1: Experimental Results of System Performance Metrics across Different Scenarios

Scenario	Route Efficiency (% Travel Time Reduction)	Detection Accuracy (%)	Response Time (Seconds)	Fuel Efficiency Improvement (%)
Daylight, Low Traffic	12%	95%	1.2	10%

Table 1: (continued).

Daylight, High Traffic	15%	93%	1.5	11%
Night, Low Traffic	10%	92%	1.3	8%
Night, High Traffic	13%	89%	1.6	9%
Rain, Moderate Traffic	11%	87%	1.8	7%
Fog, Low Traffic	8%	85%	1.9	6%

5. Conclusion

This paper demonstrates the capability of computer vision and GIS to augment public transportation routes in driverless vehicles, enabling enhanced efficiency, safety and elasticity of the routes. Through advanced scene sensing and object recognition, the system is well suited to adapt to changing environments, continuously changing paths to cut time and fuel costs and increase passenger safety. GIS data is layered on top of the decision tree to allow the system to better serve the demand for urban transit with more precision and flexibility. The results suggest strong upsides for urban public transportation, including shorter travel time, reduced fuel consumption and enhanced safety, all of which are crucial for providing a more secure and responsive transit experience for users. But some limitations — like computation requirements and performance under harsh weather — offer room for improvement. To deal with these issues, especially in the real world deployment scenario, further studies on improving the system's scalability and resilience to unpredictable barriers and traffic volumes are required [11]. Some of the future research avenues may involve bringing in other new technologies like 5G and IoT to continue scaling up real-time data processing, connectivity and system efficiency. Moreover, there is the potential for improvements in edge computing that would alleviate the computational load and enable the system to be more scalable and efficient across many urban environments. It concludes that this research helps the future of smart cities by using computer vision and GIS to enable autonomous public transport in order to secure safer, more efficient, and sustainable urban transport.

References

[1] de Winter, Joost, et al. "Predicting perceived risk of traffic scenes using computer vision." Transportation research part F: traffic psychology and behaviour 93 (2023): 235-247.

[2] Lee, Byung-Kwan, et al. "Moai: Mixture of all intelligence for large language and vision models." European Conference on Computer Vision. Springer, Cham, 2025.

[3] Valipoor, Mohammad Moeen, and Angélica De Antonio. "Recent trends in computer vision-driven scene understanding for VI/blind users: a systematic mapping." Universal Access in the Information Society 22.3 (2023): 983-1005.

[4] Khan, Wasiq, et al. "Outdoor mobility aid for people with visual impairment: Obstacle detection and responsive framework for the scene perception during the outdoor mobility of people with visual impairment." Expert Systems with Applications 228 (2023): 120464.

[5] Schmid, Daniel, Christian Jarvers, and Heiko Neumann. "Canonical circuit computations for computer vision." Biological Cybernetics 117.4 (2023): 299-329.

[6] Xu, Haowen, et al. "Semi-automatic geographic information system framework for creating photo-realistic digital twin cities to support autonomous driving research." Transportation research record (2024): 03611981231205884.

[7] Bridgelall, Raj, Ryan Jones, and Denver Tolliver. "Ranking Opportunities for Autonomous Trucks Using Data Mining and GIS." Geographies 3.4 (2023): 806-823.

[8] ECEOĞLU, Osman, and Ü. N. A. L. İlker. "Development of a GIS-based Autonomous Soil Drilling Robot for Efficient Tree Cultivation and Orchard Planting." (2024).

[9]   Baek, MinHyeok, et al. "Study on Map Building Performance Using OSM in Virtual Environment for Application to Self-Driving Vehicle." Journal of Auto-vehicle Safety Association 15.2 (2023): 42-48.

[10] Ikram, Zarif, Golam Md Muktadir, and Jim Whitehead. "Procedural Generation of Complex Roundabouts for Autonomous Vehicle Testing." 2023 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2023.

[11] Galagedara, Hashanka, and Niroshan Bandara. "Mapping on-road driver experiences of main roads: application of artificial intelligence, gis, and google maps." (2023).