Statement of Purpose

Yanjie Zhang

Introduction

The vision of smart cities and the advent of artificial general intelligence (AGI) are driving forces behind the rapid advancements in autonomous vehicle technology (Campisi et al., 2021, Louati et al., 2024). In a smart city, autonomous vehicles are expected to navigate seamlessly, ensuring safety and efficiency through interaction with their environment and other vehicles in real time (Sadaf et al., 2023). The development of AGI further amplifies this vision by enabling machines to perform tasks with human-like intelligence, paving the way for transformative applications in urban mobility.

At the core of these innovations lies the need for reliable and up-to-date road information, which is crucial for the safe and efficient operation of autonomous vehicles. High-definition (HD) maps (Shimada et al., 2015; Jiang et al., 2019; Liu et al., 2020), enriched with detailed road features, have become essential components in this ecosystem. However, creating and maintaining these maps require significant resources and continuous updates to reflect the dynamic urban environment. Currently, the utilization of crowdsourced data presents a promising solution to this challenge (Guo et al., 2024; Zhang, Zhang, & Liu, 2021). By leveraging the widespread availability of consumer-grade sensors and cameras, it is possible to collect vast amounts of data from ordinary vehicles and public transportation. This data can then be used to detect and classify critical road features, ensuring that autonomous vehicles have access to accurate and current information.

My research focuses on developing advanced models for detecting and classifying road features from crowdsourced data, such as extracting valuable information from social sensing or VGI. This work involves creating algorithms that can efficiently mine and integrate diverse data sources to enhance the autonomous decision-making capabilities of vehicles. By effectively leveraging crowdsourced data, my aim is to contribute to the advancement of smarter, safer, and more efficient autonomous driving systems within the domain of urban science.

I believe that Dr. Yingwei Yan's work on VGI, data analysis, data quality, data applications in the social sciences and data privacy aligns closely with my research interests in developing models for extracting valuable road information from crowdsourced data. His expertise in handling large, unstructured geospatial datasets and improving data quality in social sensing resonates with my aim to enhance the autonomous decision-making capabilities of vehicles using VGI. Dr. Yan's research will be invaluable in guiding my work on leveraging crowdsourced data for smarter urban mobility solutions.

Methodology and Conceptual Framing

My research will be framed within the emerging field of social sensing, focusing on enhancing the quality and utility of VGI data for urban science applications. The methodological approach will involve a combination of data quality assessment techniques, machine learning algorithms, and geostatistical analysis to develop frameworks for evaluating and improving VGI data, to support autonomous driving systems by providing reliable environmental information.

The increasing reliance on crowdsourced data for autonomous driving has led to significant advancements in developing models for detecting and classifying road features from imagery and sensor data. Key studies and identified gaps in the field are outlined below.

Zhou et al. (2023) proposed a method for extracting lane information using LMPNet, a network based on the Feature Pyramid Network (FPN) for lane detection from road image sequences. The approach involves projecting lane markings from perspective space into a three-dimensional (3D) space to calculate world coordinates, followed by clustering and fitting using an improved DBSCAN algorithm and B-spline curves. This study achieved an 82.6% accuracy for lane marking within 1 meter, demonstrating the effectiveness of using consumer-grade cameras and low-precision positioning equipment for lane detection.

Another critical development is presented by Bu et al. (2023), who utilized monocular cameras on buses to detect crosswalk changes. Their system employs image analysis techniques to monitor and update crosswalk information, achieving an 86% accuracy in crosswalk change detection on offline datasets and 66% on online datasets. This approach highlights the viability of using public vehicles for continuous data collection and the importance of addressing localization and observability challenges.

In the work by Dabeer et al. (2017), an end-to-end system for generating 3D maps using consumer-grade equipment was developed. The system achieved a mean absolute accuracy of less than 20 cm at any sign corner with only 25 journeys, demonstrating the feasibility of using crowdsourced data for precise detection and classification of road features.

The study by Kim et al. (2021) proposed a system for updating road feature information using crowdsourced data. They employed a combination of clustering algorithms and machine learning models to process data from various sources, achieving significant improvements in the accuracy and reliability of road feature detection and classification.

Qin et al. (2023) focused on traffic flow-based mapping in complex urban environments. Their study utilized vehicle trajectories to generate maps, addressing challenges such as data heterogeneity and urban density. They introduced advanced machine learning techniques to enhance the detection and classification of road features, which is crucial for real-time applications in autonomous driving.

According to the aforementioned research results, three main gaps have been identified:

1. Data integration: Many existing methods focus on homogeneous data sources. There is a need for advanced data fusion techniques to integrate heterogeneous data, particularly for combining Street View Imagery (SVI) to extract road element information. Research in this area, aimed at supporting autonomous driving decision-making, remains in its infancy;
2. Real-Time Update: Ensuring real-time capabilities in detecting and classifying road features remains a challenge due to the processing time and data volume involved;
3. Accuracy and Reliability: While significant progress has been made, further improvements are needed in the accuracy and reliability of crowdsourced data, particularly in diverse and complex environments.

Building on these advancements, my research will focus on developing models for detecting and classifying urban features from crowdsourced data such as SVI and integrating this information with other data sources to enhance the decision-making capabilities of autonomous vehicles. This involves:

1. Data collection and preprocessing. Collecting and preprocessing SVI and other potential available related data;
2. Machine learning models. Developing models for detecting and classifying road features such as lanes, traffic signs, and crosswalks from SVI;
3. Data fusion techniques. Employing advanced techniques to integrate information from diverse sources;

4. Real-time updates. Designing algorithms for rapid detection and classification of urban features;
5. Evaluation. Conducting comprehensive evaluations in various environments to assess the proposed method's accuracy and reliability.

By addressing the identified gaps and leveraging the potential of crowdsourced data, this research aims to extract key urban elements from crowdsourced data to provide prior information for autonomous vehicles. This will facilitate the construction of digital cities and support the safer operation of autonomous vehicles in open urban environments.

This enriched literature review provides a comprehensive overview of the current state of research and highlights the critical gaps that my proposed research aims to address.

Therefore, the research questions are summarized as follows:

1. How can valuable information be effectively extracted from crowdsourced SVI to enhance the detection and classification of road features for autonomous vehicles?
2. What machine learning models and data fusion techniques can be developed to integrate diverse data sources, such as SVI and vehicle sensor data, for real-time updates?
3. How can the proposed methods improve accuracy, reliability, and scalability across various driving environments?

In order to study the above questions, I may need to use some of the methods mentioned.

1. Data collection and preprocessing: Collecting crowdsourced data, including SVI, camera images, from multiple sources such as public open source platforms like Cityscapes (Cordts et al., 2016) and Mapillary Vistas (Neuhold et al., 2017), public transportation vehicles, etc. Ensure consistency across data sets by normalizing image resolutions, synchronizing timestamps, and removing noise through filtering techniques. Employ data augmentation methods to increase the diversity of training data for machine learning models.
2. Machine learning models for detection and classification: Develop and train deep learning models, such as convolutional neural networks (CNNs) based on architectures like LMPNet (Zhou et al., 2023), to detect urban elements from SVI; Utilize techniques like perspective transformation to project urban elements into 3D space. Implement object detection models, leveraging frameworks like YOLO or Faster R-CNN (Joiya, 2022), to identify and classify traffic signs. Employ transfer learning to adapt pre-trained models for this specific task; Use image segmentation techniques to detect traffic elements in imagery dataset. Even explore the use of recurrent neural networks (RNNs) for temporal consistency in elements detection across video frames.
3. Data fusion techniques: Develop a data fusion framework to combine information from SVI and other potential data. Employ Bayesian networks or Kalman filters to merge data from different sources, ensuring robustness and reducing uncertainty; Use clustering algorithms such as DBSCAN or special clustering to group detected features and redundancy (Zhou et al., 2023; Kim et al., 2021). Apply these techniques to integrate data from diverse crowdsourced data.
4. Real-time update: Design algorithms capable of processing incoming data in real time, updating the urban feature database dynamically (Guo et al., 2024). Implement parallel processing techniques to handle large volumes of data efficiently; Set up a backend server to aggregate and process data, ensuring low-latency communication with autonomous vehicles. Utilize cloud-based platforms for scalability and reliability.

5. Evaluation: Define performance metrics to evaluate the accuracy, reliability, and scalability of the proposed method. Metrics may include detection accuracy, false positive and false negative rates, and processing latency; Conduct experiments in various city environments, such as urban, suburban, and highway settings, to assess the method's performance. Use ground truth data for validation; Perform case studies to demonstrate the method's effectiveness in real world scenarios. Analyze the impact of different data sources and fusion techniques on overall method performance.
6. Expect outcomes: Improved models for detecting and classifying urban elements features from crowdsourced data, leading to more accurate and reliable data for autonomous vehicles; A robust framework for integrating diverse data sources, ensuring comprehensive information; Algorithms that enable real-time updates, enhancing the responsiveness of detecting features. A scalable system capable of handling large volumes of data from various sources, ensuring consistent performance in different environments.

Past Education and Experiences

I believe I am a strong candidate for this program due to my solid academic background, including an outstanding undergraduate thesis in transportation engineering from Nanjing Agricultural University and comprehensive research experience in a sub-project of a National Key R&D Program of China at Tongji University.

During my master's degree, I developed strong programming and project management skills. For coding, I developed an HD map information management database using PostgreSQL and designed the entire process of information interaction between connected vehicles and HD maps using MATLAB. I also integrated this with the RoadRunner simulator for visualization. For project management, I completed a group standard project, managing the entire process from initiation to publication. This project involved coordinating 26 HD map-related units in China, with the assistance of the Chinese Society for Geodesy Photogrammetry and Cartography, and participating more than 10 seminars over 14 months.

Relevant Information

During my master's degree at Tongji University, I focused on high-definition maps, studying the organization and management of map data and proposing an approach for information interaction between autonomous vehicles and HD maps. Throughout this research, I have published one journal paper (in Chinese) (Zhang et al., 2024) in Geomatics and Information Science of Wuhan University and a conference paper (Zhang & Huang, 2023) presented as an oral report at the ISPRS 2023 Geospatial Week.

Next Steps and Conclusion

Looking ahead, I aim to contribute to the evolving field of social sensing by developing new methodologies to improve data quality and ensuring data privacy in VGI systems, as well as extracting valuable road environment information for use in vehicle-road collaboration to assist autonomous driving. I believe that my proposed research aligns with the broader goals of the Department of Geography at NUS, particularly in the areas of spatial data science and social sensing technologies.

I am eager to bring my skills and research to NUS, where I hope to contribute meaningfully to the field of urban sensing and autonomous vehicle collaboration. I am excited about the opportunity to collaborate with faculty members at NUS, particularly in projects focused on enhancing VGI applications in the social sciences.

Through this research, I hope to contribute to the department's efforts to advance the understanding and practical use of crowd-sourced geographic data for the benefit of society.

References

Bu, T., Mertz, C., Dolan, J. (2023). Toward Map Updates with Crosswalk Change Detection Using a Monocular Bus Camera. *2023 IEEE Intelligent Vehicles Symposium (IV)*. pp. 1–8.

Campisi, T., Severino, A., Al-Rashid, M.A., Pau, G. (2021). The Development of the Smart Cities in the Connected and Autonomous Vehicles (CAVs) Era: From Mobility Patterns to Scaling in Cities. *Infrastructures* 6, 100.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 3213–3223.

Dabeer, O., Ding, W., Gowaiker, R., Grzechnik, S.K., Lakshman, M.J., Lee, S., Reitmayr, G., Sharma, A., Somasundaram, K., Sukhavasi, R.T., Wu, X. (2017). An end-to-end system for crowdsourced 3D maps for autonomous vehicles: The mapping component. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 634–641.

Guo, Y., Zhou, J., Li, X., Tang, Y., Lv, Z. (2024). A Review of Crowdsourcing Update Methods for High-Definition Maps. *IJGI* 13, 104.

Jiang, K., Yang, D., Liu, C., Zhang, T., Xiao, Z. (2019). A Flexible Multi-Layer Map Model Designed for Lane-Level Route Planning in Autonomous Vehicles. *Engineering* 5, 305–318.

Joiya. (2022). Object detection: YOLO vs FASTER R-CNN. *Modernization in Engineering Technology and Science*.

Kim, K., Cho, S., Chung, W. (2021). HD Map Update for Autonomous Driving With Crowdsourced Data. *IEEE Robot*. Autom. Lett. 6, 1895–1901.

Liu, J., Zhan, J., Guo, C., Lei, T., Li, Y. (2020). Data logic structure and key technologies on intelligent high-precision map. *Journal of Geodesy and Geoinformation Science*, 3(3), 1-17.

Louati, A., Louati, H., Kariri, E., Neifar, W., Hassan, M.K., Khairi, M.H.H., Farahat, M.A., El-Hoseny, H.M. (2024). Sustainable Smart Cities through Multi-Agent Reinforcement Learning-Based Cooperative Autonomous Vehicles. *Sustainability* 16, 1779.

Neuhold, G., Ollmann, T., Bulo, S.R., Kontschieder, P. (2017). The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. *2017 IEEE International Conference on Computer Vision (ICCV)*. pp. 5000–5009.

Qin, T., Huang, H., Wang, Z., Chen, T., Ding, W. (2023). Traffic Flow-Based Crowdsourced Mapping in Complex Urban Scenario. *IEEE Robot*. Autom. Lett. 8, 5077–5083.

Sadaf, M., Iqbal, Z., Javed, A.R., Saba, I., Krichen, M., Majeed, S., Raza, A. (2023). Connected and Automated Vehicles: Infrastructure, Applications, Security, Critical Challenges, and Future Aspects. *Technologies* 11, 117.

Shimada, H., Yamaguchi, A., Takada, H., Sato, K. (2015). Implementation and Evaluation of Local Dynamic Map in Safety Driving Systems. *JTTs* 05, 102–112.

Zhang, P., Zhang, M., Liu, J. (2021). Real-Time HD Map Change Detection for Crowdsourcing Update Based on Mid-to-High-End Sensors. *Sensors* 21, 2477.

Zhang, Y.J., Huang, W. (2023). AN APPROACH OF HIGH DEFINITION MAP INFORMATION INTERACTION. *Int. Arch. Photogramm. Remote Sens*. Spatial Inf. Sci. XLVIII-1/W2-2023, 861–866.

Zhang, Y.J., Huang W., Liu, X.T., Zhang, F.Y., Wu, H.B., Ying,S., Liu, C. (2024). An Approach for High Definition Map Information Interaction for Autonomous Driving. *Geomatics and Information Science of Wuhan University*. 49(4): 662-671.

Zhou, J., Guo, Y., Bian, Y., Huang, Y., Li, B. (2023). Lane Information Extraction for High Definition Maps Using Crowdsourced Data. *IEEE Trans. Intell. Transport. Syst.* 24, 7780–7790.