Urban Traffic Noise Analysis with The Integration of Vision-Language Model and 3D WebGIS

Yueying Zhang[a], Zhiyong Wang[a*], Lina Huang[b,c], Runze Huang[a]

[a]School of Civil Engineering and Transportation, South China University of Technology, 381 Wushan Road, Guangzhou, 510641, Guangdong, China - ctzhangyueying@mail.scut.edu.cn, zwang1984@scut.edu.cn (*corresponding author), ctmike-huang02@mail.scut.edu.cn

[b]School of Resource and Environmental Science, Wuhan University, 129 Luoyu Road, Wuhan, 430079, Hubei, China - linahuang@whu.edu.cn

[c]Key Laboratory of Digital Mapping and Land Information Application, Ministry of Natural Resources, 129 Luoyu Road, Wuhan 430079, Hubei, China - linahuang@whu.edu.cn

Keywords: Traffic Noise, Grounding DINO, Noise Map, Three Dimension Visualization, Noise modeling.

Abstract

In the context of rapid urbanization, traffic noise pollution has emerged as a critical environmental issue. This study proposes an innovative three-dimensional dynamic noise mapping method addressing existing technical challenges in noise modeling and visualization. The key innovations include: Utilizing the Grounding DINO large-scale vision-language model to automatically extract traffic flow information from video surveillance data, significantly improving data acquisition efficiency and accuracy. Developing a Web-based three-dimensional visualization system using the Cesium platform, supporting interactive dynamic noise distribution display and innovatively introducing an audio feedback mechanism. The research method combines deep learning with spatiotemporal correlation analysis to effectively capture noise source parameters. The noise model adopts the CNOSSOS-EU standard, considering multiple factors including geometric divergence attenuation, atmospheric absorption, ground effects, and building reflection and diffraction. Using Jingxiu and Lianchi Districts in Baoding City, Hebei Province as a case study, the research validates the method's effectiveness. The three-dimensional visualization results demonstrate the approach's superior ability to reflect the physical characteristics of real-world acoustic environments, providing crucial technical support for urban planning and noise control decision-making. Key innovations include improved noise distribution accuracy, dynamic visualization capabilities, and the introduction of interactive audio feedback, offering a novel technical approach to urban noise assessment.

1. Introduction

With the rapid urbanization and exponential growth of urban populations globally, traffic noise pollution has emerged as one of the most pressing environmental challenges facing modern cities. Addressing this issue requires accurate identification and management of various noise pollution sources to enhance residents' quality of life (Poslončec-Petrić et al., 2016). According to the World Health Organization (WHO), environmental factors account for approximately 17% of total health impacts in urban areas, with traffic noise identified as a major and increasingly significant factor (Hamad et al., 2017) access to quiet spaces for relaxation and recovery. Long-term exposure to urban traffic noise significantly impacts human health, causing sleep disturbances, cognitive impairment in children and the elderly, cardiovascular problems, and mental health issues including stress and anxiety (Tobollik et al., 2019, Stansfeld et al., 2021, Radun et al., 2022, Seidler et al., 2019, Crombie et al., 2011).

The cumulative evidence from these studies underscores the urgent need for comprehensive urban planning strategies that effectively address traffic noise pollution. The multifaceted health impacts of noise exposure, ranging from physiological to psychological effects, highlight the importance of implementing robust noise management policies and developing innovative solutions for noise reduction in urban environments.

Accurate noise mapping and visualization systems have emerged as critical tools for the systematic assessment and management of urban noise pollution, particularly in contemporary urban environments (Beran et al., 2022, Wing et al., 2006). These systems play a pivotal role in understanding noise distribution patterns, identifying problem areas, and developing effective mitigation strategies.

While significant advances have been made in noise modeling techniques, modern urban environments present unprecedented complexities that continue to impede progress in this field, especially in cities characterized by high-density development and clusters of high-rise buildings.

The first challenge lies in modeling noise propagation where traffic noise exhibits intricate three-dimensional characteristics through vertical propagation patterns and complex multiple reflection effects between buildings (Xie et al., 2015, Cai et al., 2018). This complexity stems from the need to account for multiple variables simultaneously, including building geometries, material properties, atmospheric conditions, and the dynamic nature of noise sources, creating complex sound wave behaviors that are difficult to capture accurately in computational models.

The second critical challenge concerns the acquisition of essential input data, particularly accurate traffic flow information. This challenge is especially pronounced in areas where road segments lack dedicated traffic monitoring sensors, creating significant data gaps in noise modeling efforts. The absence of reliable traffic flow data represents a fundamental bottleneck in achieving accurate noise predictions. While recent technological advances, particularly in video analysis using deep learning methods such as You Only Look Once

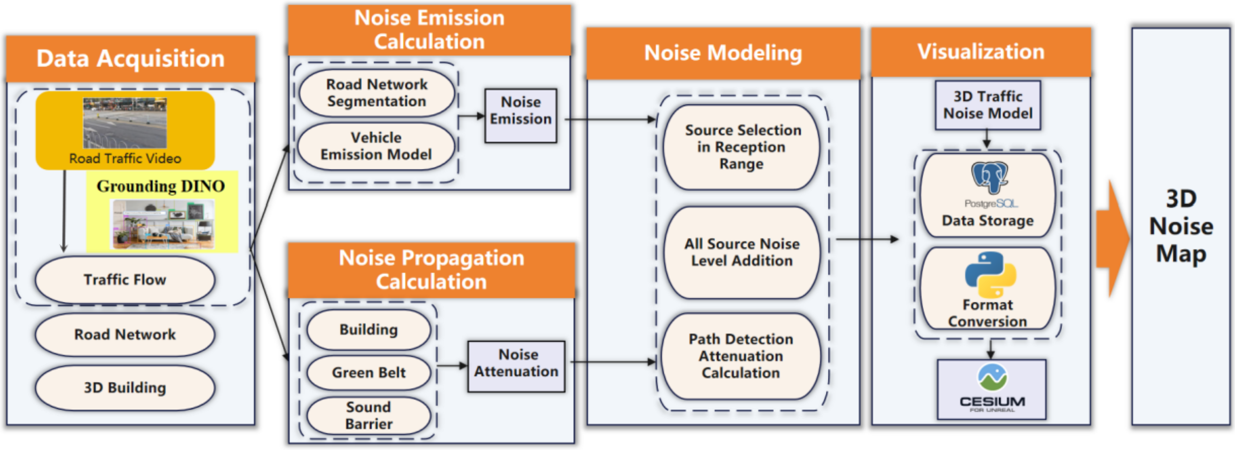

Figure 1. Overview of Our Proposed Approach

models (YOLO) (Jia et al., 2021, Zhang et al., 2022), have introduced novel approaches for traffic data collection, these solutions still face considerable limitations. One notable constraint is the accuracy of vehicle type recognition, which is crucial for precise noise modeling as different vehicle types generate distinct noise signatures. The inability to consistently and accurately differentiate between vehicle types can lead to significant uncertainties in noise prediction models.

The third major challenge relates to the limitations of existing visualization methods in meeting the demanding requirements of modern urban noise assessment (Cai et al., 2015). Current visualization techniques often struggle to effectively handle large-scale, multi-dimensional representations of noise distribution patterns. These limitations manifest in several ways: first of all, in the ability to simultaneously display multiple layers of noise information across different spatial and temporal scales; secondly, in the capacity to represent dynamic changes in noise patterns over time; and thirdly, in the capability to maintain high levels of detail while processing and displaying vast amounts of data. The challenge is further compounded by the need to balance computational efficiency with visualization quality, particularly when dealing with real-time or near-real-time data updates; Last but not least, the visualization challenge extends beyond technical limitations, as current systems struggle to provide intuitive interfaces for navigating complex three-dimensional noise data, particularly in meeting the diverse needs of stakeholders from urban planners to the general public who require different levels of detail and interaction methods.

In response to these challenges, this paper proposes an innovative three-dimensional dynamic noise mapping methodology that integrates advanced deep learning capabilities with sophisticated three-dimensional visualization technologies. This comprehensive approach improves the accuracy of urban noise distribution representation, providing important technical support for urban noise assessment and control.

2. Methodology

2.1 System Overview

This research presents a method for dynamic mapping of urban traffic noise based on deep learning and three-dimensional visualization, as illustrated in Figure 1. The proposed approach offers two major technical contributions in urban noise assessment and visualization.

The first contribution lies in traffic flow data acquisition through the implementation of the Grounding DINO large-scale vision-language model (Liu et al., 2023, Ren et al., 2024). This state-of-the-art model leverages powerful semantic understanding abilities to automatically extract detailed traffic flow information from existing video surveillance infrastructure.

The second contribution involves a sophisticated Web-based three-dimensional visualization system built on the Cesium platform. A particularly innovative aspect is the introduction of interactive sound simulation capabilities that correspond to visualized noise levels, while real-time audio feedback enhances users' understanding of noise impact. Additionally, spatial audio rendering accurately represents the noise distribution in three-dimensional space.

The integration of these components creates a comprehensive system that enables accurate noise analysis while supporting dynamic decision-making processes. The system's interactive nature enhances user understanding of noise-related issues through both visual and auditory channels.

2.2 Traffic Flow Data Acquisition

The system utilizes the Grounding DINO large-scale vision-language model to process urban surveillance video sequences. This model employs a transformer-based architecture that effectively bridges visual and textual information, allowing for precise vehicle detection and classification across different environmental conditions. The traffic flow analysis framework incorporates the following features:

(1) Traffic Flow Analysis: The system implements sophisticated algorithms for vehicle trajectory prediction and tracking.

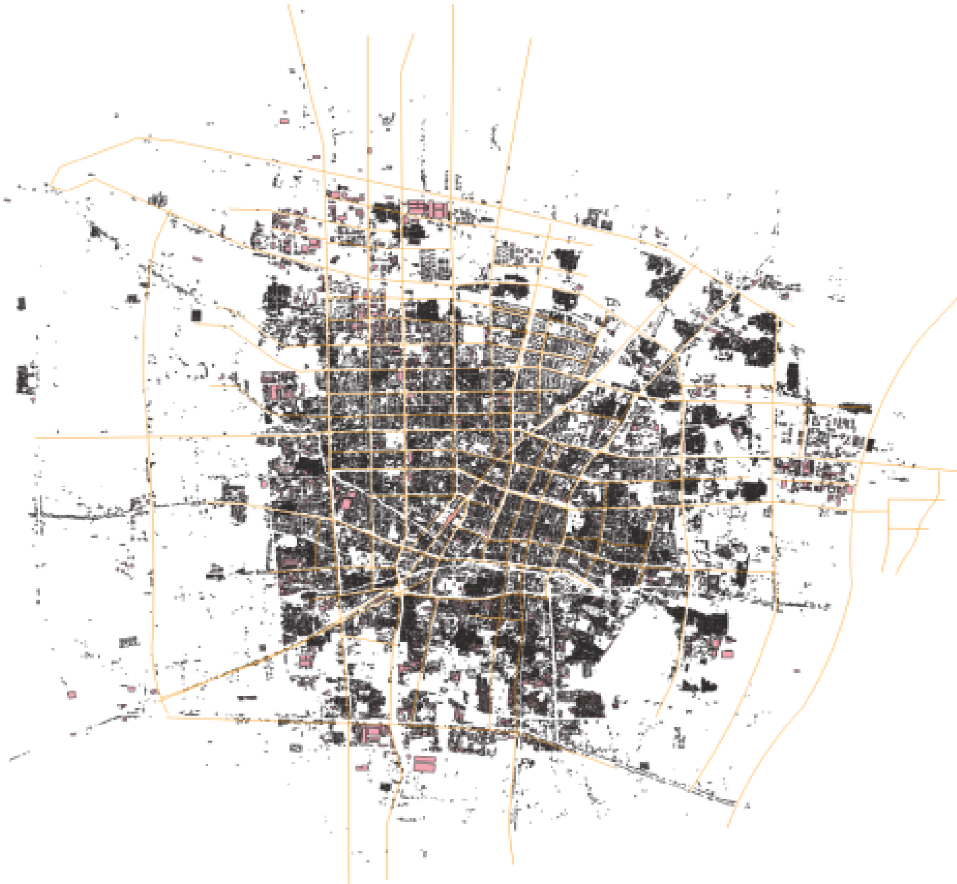

Figure 2. 2D Road and Building Map of the Study Area

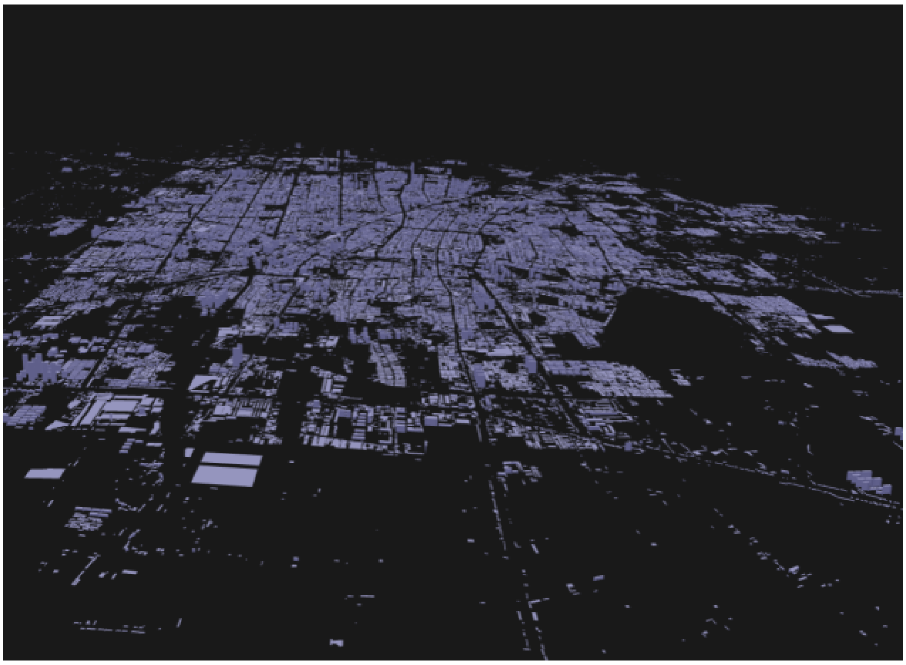

Figure 3. 3D Building Map of the Study Area

(2) Vehicle Classification: The system distinguishes between different vehicle categories with high precision, crucial for accurate noise emission calculation.

This comprehensive approach to data collection and integration provides a robust foundation for subsequent noise modeling and visualization processes.

2.3 Noise Modeling

The noise modeling framework adopts a grid-based spatial discretization approach to accurately represent sound pressure level data, which proves particularly effective in complex urban environments, especially in areas characterized by high-rise buildings. The system incorporates the CNOSSOS-EU standard noise prediction model (Shilton et al., 2015), establishing a comprehensive computational framework that addresses multiple environmental factors. In terms of geometric divergence attenuation, the framework implements spreading laws for both point and line sources while accounting for distance-dependent energy dissipation in three-dimensional space. The system thoroughly considers atmospheric absorption by modeling temperature and humidity effects, incorporating local meteorological conditions, and applying standardized absorption calculations.

Ground effects are modeled through a sophisticated approach that accounts for surface reflection and absorption across various surface types, while simultaneously considering terrain elevation changes and integrating ground impedance characteristics. The framework also implements detailed building reflection and diffraction calculations, including specular reflections, edge diffraction, building facade properties, and multiple reflection paths in urban canyons. This comprehensive modeling approach demonstrates particular effectiveness when applied to complex urban geometries with multiple building reflections, areas with significant terrain variations, high-density urban environments with multiple noise sources, and scenarios involving complex interference patterns.

2.4 3D Visualization System

The visualization system employs a frontend-backend separation architecture to optimize performance and maintainability. The backend, implemented in Python, manages essential functions including data processing and optimization, conversion to GLB format, noise reduction algorithms, and spatial interpolation calculations. Built on the Cesium three-dimensional visualization engine, the frontend visualization system implements comprehensive features for dynamic spatiotemporal evolution display, including real-time playback of noise distribution evolution, multiple time scale support, and flexible observation period selection.

The system provides sophisticated interactive navigation and control capabilities, featuring free viewpoint rotation, dynamic scaling capabilities, smooth translation controls, multi-level opacity adjustment, and real-time noise intensity section views. A distinctive audio feedback mechanism enhances the user experience by incorporating dynamic volume and pitch adjustment based on spatial position, simulation of different frequency noise characteristics, real-time acoustic feedback, and spatial audio rendering. The system's interactive design emphasizes user experience through intuitive interface design, responsive operations, clear visual feedback, integrated auditory feedback, and smooth performance with large-scale data. Through this integrated approach, the system serves as a comprehensive tool for urban noise assessment and control, supporting various stakeholders including urban planners, environmental departments, residents, and researchers in their efforts to understand and manage urban noise environments.

3. Evaluation

We conducted a comprehensive evaluation of our approach through a detailed case study implementation in Jingxiu District and Lianchi District of Baoding City, Hebei Province, China (see Figure 2 and Figure 3). These districts were specifically chosen for their diverse urban characteristics, featuring varying building densities, road networks, and land use patterns, making them ideal test cases for validating our model's versatility.

The three-dimensional noise maps (Figure 4 and Figure 5) were generated using high-resolution spatial data, incorporating detailed building geometries, terrain variations, and traffic flow patterns.

Compared to traditional two-dimensional methods, the three-dimensional visualization capabilities demonstrated advantages including clear representation of vertical noise propagation patterns along building facades, accurate depiction of noise reflection and diffraction effects in urban canyons, intuitive visualization of noise level variations at different heights, and enhanced understanding of the impact of building geometry on noise distribution.

The model exhibited particularly strong performance in complex urban scenarios, including accurate representation of noise propagation around high-rise buildings, precise modeling of multiple reflection paths in dense urban areas, effective visualization of noise patterns in areas with significant terrain variations, and realistic representation of traffic noise propagation along major transportation corridors.

These comprehensive evaluation results strongly indicate that our approach offers significant advantages for urban noise assessment and management. The system's ability to accurately model and visualize complex noise propagation patterns, combined with its efficient computational performance, makes it a valuable tool for urban planning and noise control decision-making processes. The intuitive three-dimensional visualization significantly enhances stakeholders' understanding of noise distribution patterns, enabling more informed decisions in urban development and environmental protection initiatives.

4. Conclusion

This research presents a comprehensive approach to urban traffic noise mapping by integrating advanced deep learning and three-dimensional visualization technologies. The proposed method addresses critical challenges in noise pollution assessment, offering significant improvements over traditional approaches. The key contributions of this study include: (1) Data Collection Innovation: Using Grounding DINO vision-language model to extract traffic flow data from surveillance videos, improving collection efficiency and accuracy. (2) Noise Model Optimization: Combining CNOSSOS-EU standard with the grid-based approach to simulate urban noise distribution more precisely by considering geometric divergence, atmospheric absorption, and building reflections. (3) Visualization Enhancement: Developing a 3D web visualization system based on Cesium platform to display noise distribution intuitively, making urban noise more understandable and perceptible.

The case study in Baoding City demonstrated the method's effectiveness in capturing spatiotemporal noise patterns and providing detailed insights into urban acoustic environments.

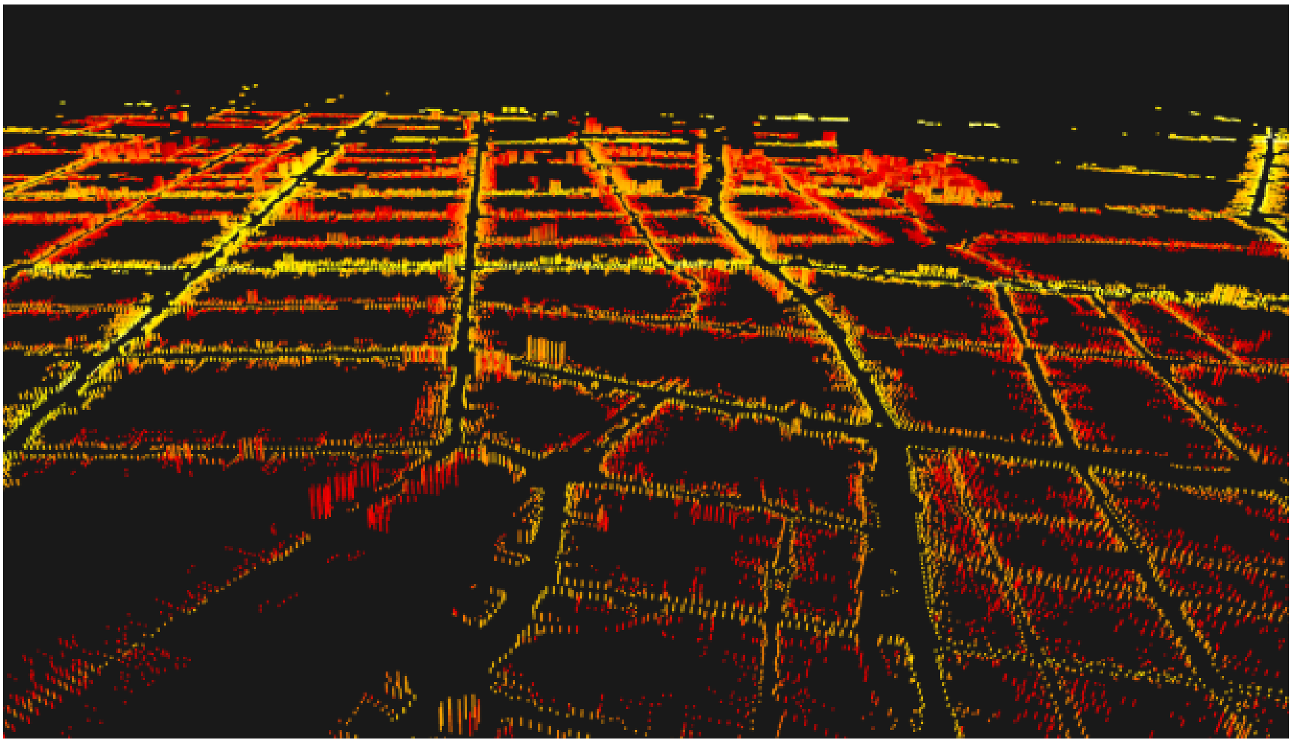

Figure 4. Local Noise Detail Map of the Study Area

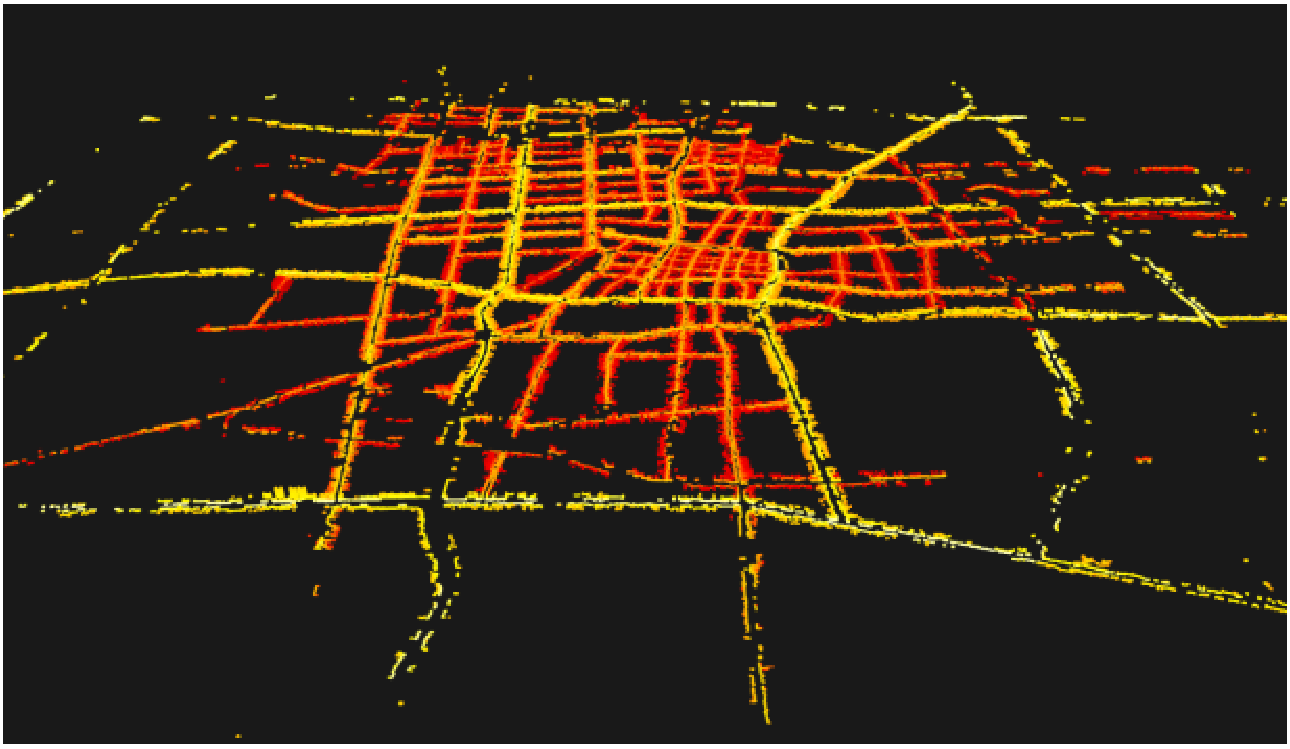

Figure 5. Global Noise Distribution Map of the Study Area

Future research directions could include expanding the model's applicability to more diverse urban environments, improving vehicle type recognition accuracy, developing more sophisticated deep learning algorithms for noise prediction, and integrating real-time data processing capabilities.

Acknowledgments

This work was supported by the "FundamentalResearch Funds for the Central Universities"(NO.2023ZYGXZR056), and "Open Research Fund Program of Key Laboratory of Digital Mapping and Land Information Application, Ministry of Natural Resources"(NO.ZRZYBWD202402).

References

Beran, D., Jedlička, K., Kumar, K., Popelka, S., Stoter, J., 2022. The third dimension in noise visualization–a design of new methods for continuous phenomenon visualization. *The Cartographic Journal*, 59(1), 1–17.

Cai, M., Yao, Y., Wang, H., 2018. Urban traffic noise maps under 3D complex building environments on a supercomputer. *Journal of Advanced Transportation*, 2018, 2814261.

Cai, M., Zou, J., Xie, J., Ma, X., 2015. Road traffic noise mapping in Guangzhou using GIS and GPS. *Applied Acoustics*, 87, 94–102.

Crombie, R., Clark, C., Stansfeld, S. A., 2011. Environmental noise exposure, early biological risk and mental health in nine to ten year old children: a cross-sectional field study. *Environmental Health*, 10(1), 39.

Hamad, K., Khalil, M. A., Shanableh, A., 2017. Modeling roadway traffic noise in a hot climate using artificial neural networks. *Transportation Research Part D: Transport and Environment*, 53, 161–177.

Jia, W., Xu, S., Liang, Z., Zhao, Y., Min, H., Li, S., Yu, Y., 2021. Real-time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector. *IET Image Processing*, 15(14), 3623–3637.

Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., Zhang, L., 2023. Grounding DINO: Marrying DINO with Grounded Pre-training for Open-Set Object Detection. *arXiv preprint*.

Poslončec-Petrić, V., Vuković, V., Frangeš, S., Bačić, Ž., 2016. Voluntary noise mapping for smart city. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 131–137.

Radun, J., Maula, H., Saarinen, P., Keränen, J., Hongisto, V., 2022. Health effects of wind turbine noise and road traffic noise on people living near wind turbines. *Renewable and Sustainable Energy Reviews*, 157, 112040.

Ren, T., Jiang, Q., Liu, S., Zeng, Z., Liu, W., Gao, H., Zhang, L., 2024. Grounding DINO 1.5: Advance the "Edge" of Open-Set Object Detection. *arXiv preprint*.

Seidler, A., Hegewald, J., Seidler, A. L., Schubert, M., Wagner, M., Dröge, P., Haufe, E., Schmitt, J., Swart, E., Zeeb, H., 2019. Is the whole more than the sum of its parts? Health effects of different types of traffic noise combined. *International Journal of Environmental Research and Public Health*, 16(9), 1665.

Shilton, S. J., Anfosso Lédée, F., Van Leeuwen, H., 2015. Conversion of existing road source data to use cnossos-eu. *Proc. of Euronoise*.

Stansfeld, S., Clark, C., Smuk, M., Gallacher, J., Babisch, W., 2021. Road traffic noise, noise sensitivity, noise annoyance, psychological and physical health and mortality. *Environmental Health*, 20(1), 1–15.

Tobollik, M., Hintzsche, M., Wothge, J., Myck, T., Plass, D., 2019. Burden of disease due to traffic noise in Germany. *International Journal of Environmental Research and Public Health*, 16(13), 2304.

Wing, L. C., Kwan, L. C., Kwong, T. M., 2006. Visualization of complex noise environment by virtual reality technologies. *Environment Protection Department (EPD): Hong Kong, China*, 76.

Xie, J., Ma, X., Cai, M., 2015. Road traffic noise mapping in Guangzhou using GIS and GPS. *Applied Acoustics*, 87, 94–102.

Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., Guo, X., 2022. Real-time vehicle detection based on improved yolo v5. *Sustainability*, 14(19), 12274.