

AirNavigation: Let UAV Navigation Tell Its Own Story

Jianguo Jiang^{1*}, Zequan Wang^{1*}, Liang Yao¹, Shengxiang Xu², Fan Liu^{1†}

¹Hohai University

²Southeast University

{remain11, zequan, yaoliang}@hhu.edu.cn, xushx@seu.edu.cn, fanliu@hhu.edu.cn

Abstract

Testing autonomous navigation algorithms of Unmanned Aerial Vehicles (UAVs) in real-world scenarios often entails significant safety risks. In this paper, we aim to build a flexible yet user-friendly UAV autonomous navigation simulator. Ideally, it should closely emulate real-world environments, support diverse UAV models and algorithms, and provide a flexible evaluation framework. Existing frameworks fail to satisfy all three requirements simultaneously. To this end, we present AirNavigation, an integrated simulation platform designed to support the end-to-end workflow of UAV navigation research. Specifically, our system leverages Unreal Engine to simulate highly realistic environments and diverse UAV models. It further facilitates semi-automated scene generation and multi-modal synthetic training data production. To lower the barrier of adoption, we develop a suite of user-friendly interfaces to enable seamless integration of diverse navigation algorithms. Moreover, we introduce a novel evaluation system powered by large language models to deliver personalized and fine-grained performance analysis.

Code — <https://github.com/1e12Leon/AirNavigation>

1 Introduction

With advances in UAV technology (Qian et al. 2022; Srivastava and Prakash 2023) and AI (LeCun, Bengio, and Hinton 2015; Sharifani and Amini 2023), UAVs are now widely applied in domains such as agriculture and rescue (Srivastava and Prakash 2023; Su et al. 2023). Autonomous navigation, a core task of UAV embodied intelligence, has attracted extensive research in recent years (Chang et al. 2023; Arafat, Alam, and Moh 2023). Although testing navigation algorithms in real-world scenarios provides practical performance evaluation, uncontrollable environmental factors and potential algorithmic flaws often lead to safety risks.

Therefore, many scholars employ simulators to evaluate the performance of UAV navigation algorithms before real-world deployments (Dosovitskiy et al. 2017a; Giovagnola et al. 2023). These simulators can be broadly divided into two categories. Device-specific simulators (e.g., DJI Flight

Simulator (Casado and Bermúdez 2020) or VR-based UAV simulators (Albeaino et al. 2022)) only support particular UAV models and are often coupled with specific hardware configurations. Others focus on specific tasks or algorithms (e.g., AerialVLN (Liu et al. 2023) and OpenFly (Gao et al. 2025) for vision-and-language navigation). While these simulators work efficiently within their field, none of them provide a general, device-agnostic, and end-to-end framework that unifies the design, testing, and evaluation of diverse UAV navigation algorithms. A systematic comparison of representative UAV platforms is presented in Table 1.

To this end, we present **AirNavigation**, an integrated simulation platform designed to support the end-to-end workflow of UAV navigation research. As illustrated in Fig. 1, our system incorporates four key modules: **1) Scenario Construction**: built on Unreal Engine to simulate highly realistic environments and diverse UAV models, with support for semi-automated scene generation; **2) Data Generation**: enabling collection and annotation of multi-modal data, facilitating the development and training of navigation algorithms (Huo et al. 2023); **3) Navigation Interface**: providing user-friendly APIs that allow seamless integration of diverse navigation algorithms and allow natural-language API calls; **4) Performance Evaluation**: introducing a novel LLM-driven evaluation system to deliver more flexible and fine-grained performance analysis.

2 AirNavigation

2.1 Easy Scenario Construction

Compared to existing autonomous navigation simulators like CARLA (Dosovitskiy et al. 2017b), our system offers a more intuitive and user-friendly solution by providing rich scenario options. In particular, it includes 15 built-in environment templates, covering urban and highway scenes that meet the environmental requirements for UAV flight operations. We further adopt blueprint-based terrain modules that users can operate without knowledge of Unreal Engine, and integrate the CityBLD (Studios 2024) framework, which allows users to easily design urban environments of different scales and styles through a graphical interface. Additionally, our system includes dynamic environment simulation capabilities, simulating weather phenomena such as fog, rain, and snow, as well as realistic traffic flows.

*These authors contributed equally.

†Corresponding Author.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

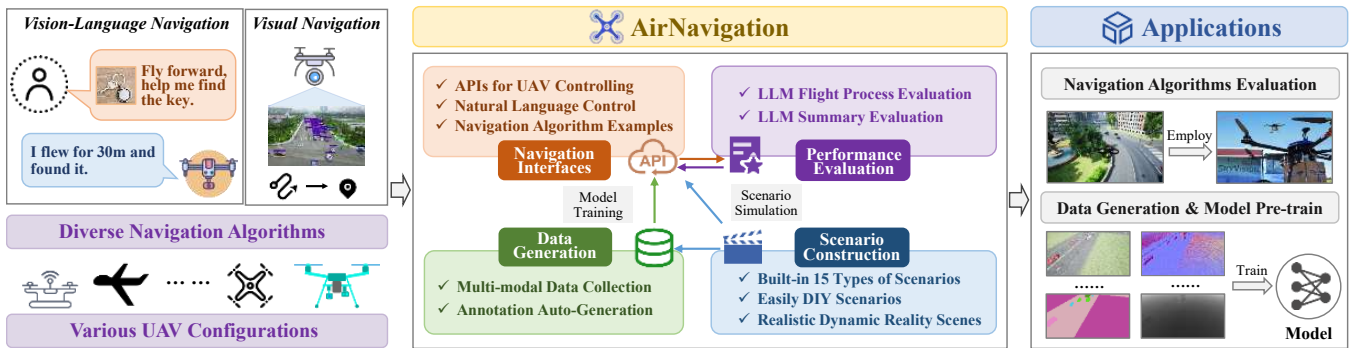


Figure 1: The overview of our AirNavigation. It integrates four main functionalities: easy scenario construction, automated multi-modal data generation, rich navigation interfaces, and intelligent performance evaluation.

Platform	Engine	Sensors	Map	Algorithms	LLM Eval.
Webots	ODE	RGB, LiDAR, IMU, GPS	Custom	RL, SLAM	✗
CARLA	UE	RGB, LiDAR, IMU, GNSS, Depth	Custom, built-in	Perception	✗
FlightGoggles	Gazebo-based	RGB, IMU	Hybrid	SLAM, OA	✗
CityNav	WebGL	RGB, IMU, GPS	Outdoor	VLN	✗
OpenFly	UE, Google Earth, GTAV, 3DGS	RGB, IMU, Depth, GPS	Outdoor, Custom	RL, VLN, OA	✗
EmbodiedCity	UE	RGB, LiDAR, IMU, GPS	Custom	VLN, Navigation, Planning	✗
Ours	UE	RGB, LiDAR, IMU, Seg, Depth...	Custom, built-in	LLM, CV, RL, SLAM...	✓

Table 1: Comparison of Different UAV Simulation Platforms. ‘RL’: Reinforcement Learning, ‘CV’: Computer Vision, ‘OA’: Obstacle Avoidance. Map types: Custom (platform-specific), Outdoor (realistic environments), Hybrid (multi-source).

2.2 Automatic Data Collection

Our AirNavigation supports the generation of multiple data types, including RGB images, depth, normalized disparity, segmentation, surface normal, infrared and IMU parameter data (e.g., altitude, orientation, velocity and so on). Through AirNavigation’s natural language control system, multiple UAVs can be simultaneously controlled for automated data collection. With three UAVs on predefined routes, the collection efficiency can reach 3×12 valid sets per minute. In addition, we develop an automatic annotation algorithm that utilizes information from (Yao et al. 2025a) segmentation and depth images to automatically generate bounding-box labels, whose effectiveness has been validated through experiments. Leveraging these automated pipelines, we construct the multi-modal image–annotation part of the UEMM-Air dataset (Yao et al. 2025b).

2.3 Rich Navigation Interfaces

Our AirNavigation is a device-agnostic simulation platform for navigation algorithms. We integrate complex control logic from AirSim (Shah et al. 2017) and provide comprehensive APIs for UAV control tasks. These APIs cover basic UAV control, map control, environment control, data collection, and auxiliary evaluation, and can be used by any type of UAV navigation algorithm, ensuring system flexibility and broad applicability (Xu et al. 2025). For better understanding, we also provide examples of API usage for navigation algorithms, including visual navigation and vision-language navigation examples. Users could deploy their algorithms according to the two examples. In addition, API documenta-

tions are available in the code repository for reference.

2.4 Intelligent Performance

To enhance the flexibility and depth of algorithm evaluation, we integrate excellent capabilities for both real-time evaluation and comprehensive post-flight summaries. This integration enables dynamic, adaptable evaluation frameworks. Specifically, we leverage Large Language Models (LLMs) within our system, utilizing prompt engineering to create tailored role and skill libraries. By feeding real-time UAV data at regular intervals, including numerical data (e.g., position, velocity, acceleration) as well as visual data, we enable continuous evaluation and recommendation, empowering people to refine their responses in real time.

2.5 Dialog-Based UAV Navigation

Our system supports dialog-based navigation, enabling seamless human-UAV collaboration through natural language commands (e.g., “Find the distant red truck”). By combining LLM-driven intent parsing with object-tracking algorithms, this capability enables the UAV to autonomously navigate complex environments. In addition, it supports object detection, observation validation, and result confirmation through interactive messaging mechanisms. Furthermore, the performance of this algorithm across different weather and scenarios has been validated through experiments, with further details available in the repository documentation. Through this dialog-based UAV navigation capability, we ensure intuitive, context-aware interactions, fostering efficient communication between the user and the UAV.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62372155, in part by Science Fund for Excellent Young Scholars of Jiangsu Province No. SBK20250300739, in part by the Research Funds of Jiangsu Hydraulic Research Institute under Grant 2025z065.

References

- Albeaino, G.; Eiris, R.; Gheisari, M.; and Issa, R. R. 2022. DroneSim: A VR-based flight training simulator for drone-mediated building inspections. *Construction Innovation*, 22(4): 831–848.
- Arafat, M. Y.; Alam, M. M.; and Moh, S. 2023. Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges. *Drones*, 7(2): 89.
- Casado, R.; and Bermúdez, A. 2020. A simulation framework for developing autonomous drone navigation systems. *Electronics*, 10(1): 7.
- Chang, Y.; Cheng, Y.; Manzoor, U.; and Murray, J. 2023. A review of UAV autonomous navigation in GPS-denied environments. *Robotics and Autonomous Systems*, 104533.
- Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; and Koltun, V. 2017a. CARLA: An open urban driving simulator. In *Conference on robot learning*, 1–16. PMLR.
- Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A. M.; and Koltun, V. 2017b. CARLA: An Open Urban Driving Simulator. In *Proceedings of the 1st Annual Conference on Robot Learning (CoRL)*, 1–16.
- Gao, Y.; Li, C.; You, Z.; Liu, J.; Li, Z.; Chen, P.; Chen, Q.; Tang, Z.; Wang, L.; Yang, P.; Tang, Y.; Tang, Y.; Liang, S.; Zhu, S.; Xiong, Z.; Su, Y.; Ye, X.; Li, J.; Ding, Y.; Wang, D.; Wang, Z.; Zhao, B.; and Li, X. 2025. OpenFly: A Comprehensive Platform for Aerial Vision-Language Navigation. arXiv:2502.18041.
- Giovagnola, J.; Megías, J. B. M.; Fernández, M. M.; Cuéllar, M. P.; and Santos, D. P. M. 2023. AirLoop: A simulation framework for testing of UAV services. *IEEE Access*, 11: 23309–23325.
- Huo, J.; Sun, Q.; Jiang, B.; Lin, H.; and Fu, Y. 2023. GeoVLN: Learning Geometry-Enhanced Visual Representation with Slot Attention for Vision-and-Language Navigation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)*, 23212–23221. IEEE Computer Society.
- LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *nature*, 521(7553): 436–444.
- Liu, S.; Zhang, H.; Qi, Y.; Wang, P.; Zhang, Y.; and Wu, Q. 2023. AerialVLN: Vision-and-Language Navigation for UAVs. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 15384–15394.
- Qian, C.; Wu, H.; Zhang, Q.; Yang, L.; and Jiang, Q. 2022. Design and Implementation of UAV Formation Cooperative System. In *ICAUS*.
- Shah, S.; Dey, D.; Lovett, C.; and Kapoor, A. 2017. AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. In *Field and Service Robotics*.
- Sharifani, K.; and Amini, M. 2023. Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, 10(07): 3897–3904.
- Srivastava, A.; and Prakash, J. 2023. Techniques, answers, and real-world UAV implementations for precision farming. *WPC*.
- Studios, W. 2024. CityBLD: Procedural City Creation Toolkit for Unreal Engine 5. <https://www.worldbld.com>.
- Su, J.; Zhu, X.; Li, S.; and Chen, W.-H. 2023. AI meets UAVs: A survey on AI empowered UAV perception systems for precision agriculture. *Neurocomputing*.
- Xu, S.; Zhang, J.; Di, S.; Luo, Y.; Yao, L.; Liu, H.; Zhu, J.; Liu, F.; and Zhang, M.-L. 2025. RobustFlow: Towards Robust Agentic Workflow Generation. *arXiv preprint arXiv:2509.21834*.
- Yao, L.; Liu, F.; Chen, D.; Zhang, C.; Wang, Y.; Chen, Z.; Xu, W.; Di, S.; and Zheng, Y. 2025a. RemoteSAM: Towards Segment Anything for Earth Observation. arXiv:2505.18022.
- Yao, L.; Liu, F.; Xu, S.; Zhang, C.; Di, S.; Ma, X.; Jiang, J.; Wang, Z.; and Zhou, J. 2025b. UEMM-Air: Enable UAVs to Undertake More Multi-modal Tasks. In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25)*, 12792–12798. Association for Computing Machinery.