UDC 004.8:629.7

IMPROVING THE QUALITY OF INSPECTION ROUTE PLANNING FOR AN FPV DRONE

Traktirnyi D. S.[1]; Zhulkovskyi O. O., Candidate of Technical Sciences, Prof.[2]
[1,2] Dniprovsky State Technical University
[1,2] Kamianske, Ukraine
[1] barmaglot888@gmail.com; [2] olalzh@ukr.net

Abstract*. This study investigates the impact of reinforcement learning on the quality of inspection route planning for an FPV drone. An attention-based neural network model was fine-tuned using the policy gradient method and compared against a greedy algorithm and supervised learning. Across 180 independent graphs, the average total priority increased by 0.739 points compared to the baseline neural network model.*
Keywords: *FPV drone; route planning; reinforcement learning; policy gradient; attention mechanism.*

Introduction. Inspecting industrial structures with an unmanned aerial vehicle (UAV) requires determining waypoint visit order based on importance, recording duration, travel time, and return requirements. Due to limited flight time, some waypoints must be excluded, making route quality dependent on sequential decisions.

Modern UAV mission planning includes combinatorial, heuristic, and intelligent methods [1, 2]. Neural network capabilities for such problems are summarized in [3], and attention-based routing models trained via complete route evaluation are proposed in [4]. Further reinforcement learning advancements appear in [5]. While supervised learning mimics pre-existing routes, adjusting parameters via rewards from complete routes directly optimizes final planning quality – defining the direction of this research.

The research aims to work is to evaluate the impact of fine-tuning an attention-based neural network model via the policy gradient method on the quality of FPV drone inspection route planning under constrained mission time.

The main part. The routing problem is modeled as an undirected weighted graph where vertices represent navigation positions, edges define feasible transition times, and waypoints carry priorities and inspection durations. The goal is to maximize total visited priority while accounting for mission duration, time margin, and return to start.

Three planning approaches are evaluated: a greedy heuristic (ratio of priority to time cost), an attention-based neural network trained via supervised learning, and the same model fine-tuned using reinforcement learning. All methods were tested on identical graphs with independent feasibility checks.

In the neural network planner, an encoder generates contextual features while a decoder sequentially selects the next waypoint. An action mask blocks repeated visits, unreachable nodes, and transitions preventing a safe return. Route termination uses a dedicated stop action.

Fine-tuning applied a policy gradient with a rollout baseline. The model stochastically builds complete routes, receiving reward $R = P - 0.1 \cdot T / B$ for feasible paths (where P is total priority, T is mission duration, and B is time budget). This weighting ensures gaining one priority point outweighs saving time.

Parameters were updated on the advantage over the baseline. Training spanned 10 epochs over 7500 synthetic graph episodes. Final evaluation covered 180 unseen graphs (60 each with 10, 40, and 120 waypoints).

The average total priority values of the routes obtained on the independent test graphs are presented in Table 1.

Table 1. Average total priority of inspection routes

Number of waypoints	Greedy method	Supervised learning	Reinforcement learning
10	16.350	16.383	16.283
40	68.850	69.383	69.733
120	198.000	199.350	201.317
All graphs	94.400	95.039	95.778

According to Table 1, the fine-tuned model achieved the highest average route priority among the three investigated methods. The largest improvement was obtained for 120 waypoints, a smaller gain was observed on graphs with 40 waypoints, and a slight decrease occurred for 10 waypoints. All generated routes complied with the time constraints, and the planning speed after fine-tuning remained close to the baseline.

Conclusions. This study investigated FPV drone inspection route planning as a time-constrained waypoint selection problem. To improve route quality, policy gradient fine-tuning was applied to an attention-based neural network model. The results were compared against a greedy algorithm and the baseline model trained with supervised learning. Across 180 independent test graphs, the fine-tuned model achieved the highest average total priority of 95.778 points, which is 0.739 points higher than the baseline neural network. The largest gain was achieved on 120-waypoint graphs, whereas no improvement was observed for the simplest problem instances. The obtained results demonstrate the potential of reinforcement learning for planning more complex inspection missions. Future research should focus on testing this approach on graphs of varying structures and larger scales.

Use of AI. During the preparation of the introduction of this abstract, ChatGPT (model GPT-6 Astra) was used to search and summarize information from scientific sources, and Codex (model GPT-6 Astra) was used as an auxiliary tool for software implementation, training, and verification of the neural network model.

REFERENCES

[1] Zhulkovskyi, O., Zhulkovska, I., Urbonas, R., Vokhmianin, H., Klimov, R., Kashcheev, M. (2026). Intelligent prediction of thermal regimes in engineering and technological systems. In: *2026 16th International Conference on Advanced Computer Information Technologies (ACIT)*, 242–246. https://doi.org/10.1109/ACIT69831.2026.11699964

[2] Zherebetskyi, V. (2026). Unmanned Aerial Vehicles Mission Planning: a Brief Overview. *Cybernetics and Computer Technologies*, (3), 65–79. https://doi.org/10.34229/2707-451X.26.3.6

[3] Wang, F., He, Q., & Li, S. (2024). Solving combinatorial optimization problems with deep neural network: A survey. *Tsinghua Science and Technology, 29*(5), 1266–1282. https://doi.org/10.26599/TST.2023.9010076

[4] Kool, W., van Hoof, H., & Welling, M. (2019). Attention, learn to solve routing problems! *arXiv:1803.08475 [stat.ML]*. https://doi.org/10.48550/arXiv.1803.08475

[5] Kwon, Y.-D., Choo, J., Kim, B., Yoon, I., Gwon, Y., & Min, S. (2020). POMO: Policy optimization with multiple optima for reinforcement learning. In: *34th Conference on Neural Information Processing Systems (NeurIPS 2020).* https://doi.org/10.48550/arXiv.2010.16011