Fine-Tuning Large Language Models for Autonomous Tactical Deconfliction of Small Unmanned Aerial Systems

Iman Sharifi, Alex Zongo, and Peng Wei
George Washington University

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

CVPR JUNE 3-7, 2026 DENVER COLORADO

Background and Objective

Dense urban drone traffic requires **real-time tactical conflict resolution under safety-critical constraints**. Traditional rule-based and optimization approaches often struggle with scalability, adaptability, and interpretability in heterogeneous multi-agent environments.

This work investigates whether fine-tuned Large Language Models (LLMs) can act as cooperative tactical deconfliction policies for small unmanned aerial systems (sUASs).

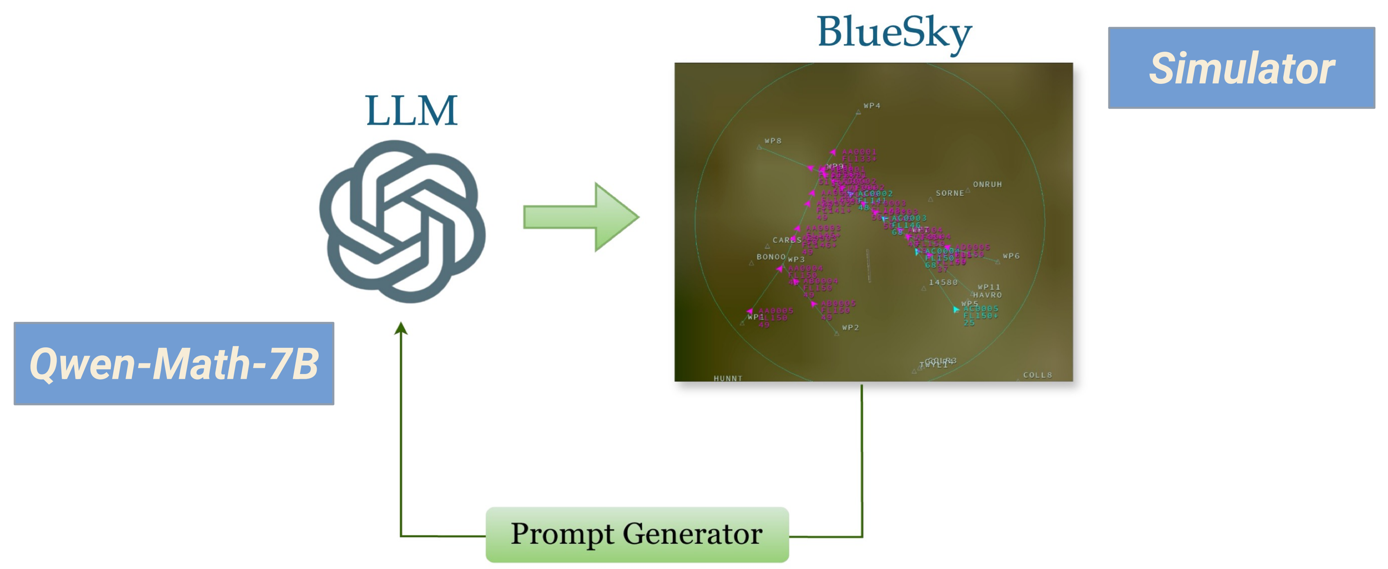
Research Questions:

- ❖ Can pre-trained LLMs manage domain-specific safety-critical tasks such as cooperative tactical deconfliction for small unmanned aerial systems (sUASs)?
- ❖ Do they need fine-tuning? If yes, how to fine-tune them without real-world data?
- ❖ How do LLMs handle multiple agents in a centralized manner?

Key Points:

- ❑ Dense low-altitude urban airspace
- ❑ Cooperative multi-agent deconfliction
- ❑ Partial observability and heterogeneous agents
- ❑ Human-aligned tactical reasoning using rule-based methods
- ❑ Fine-tuning instead of zero-shot prompting

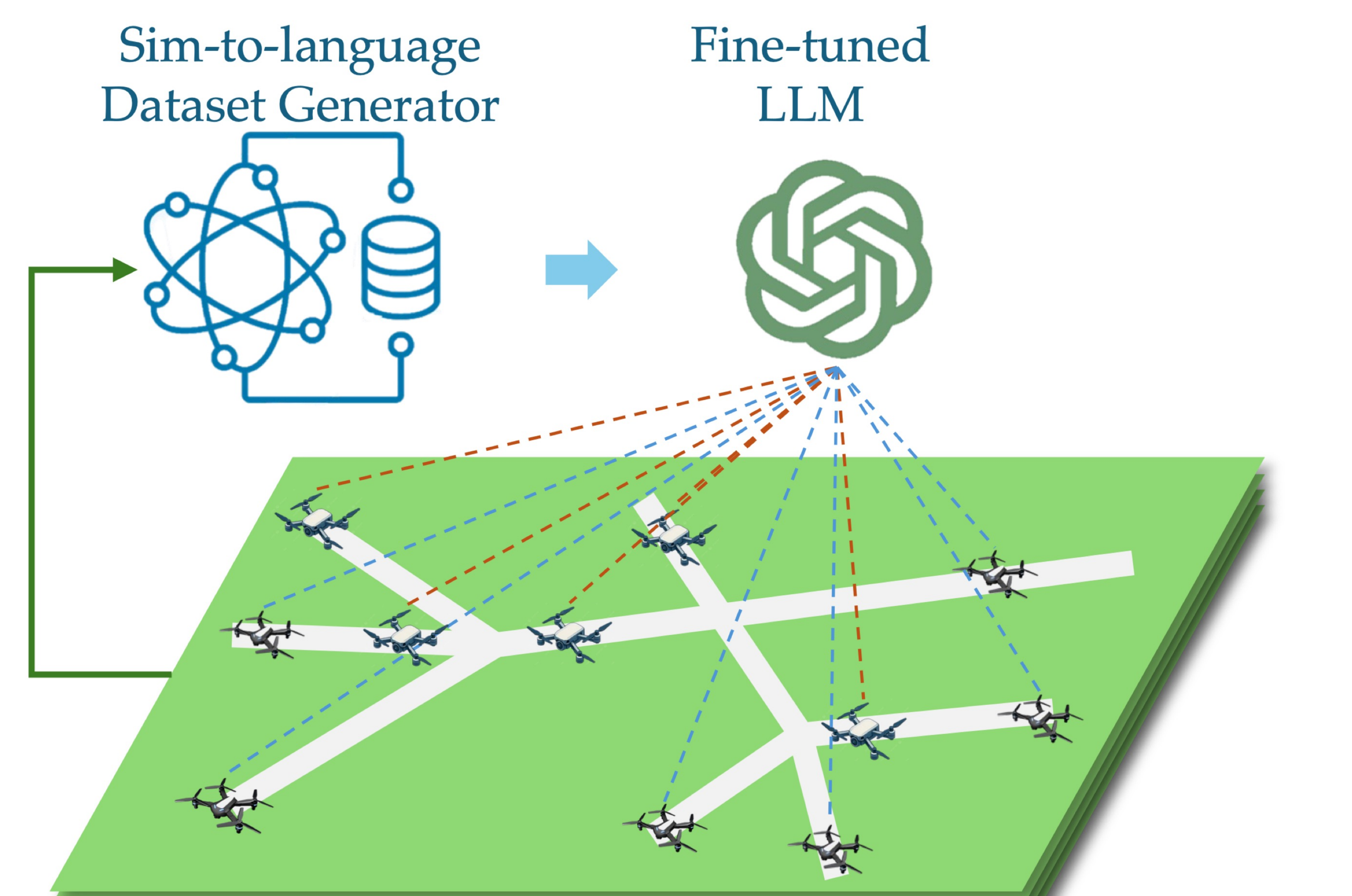
Pretrained (Base) LLM Evaluation



Initial Findings:

- The Base model acts conservatively, congesting the airspace.
- It occasionally generates unsafe actions.
- The inference time is at least 3 seconds for a single agent, for a variable-length response (with a maximum of 1024 output tokens).
- For shorter responses (e.g., 50 tokens), it generates unreasonable answers, although the inference time is reduced.

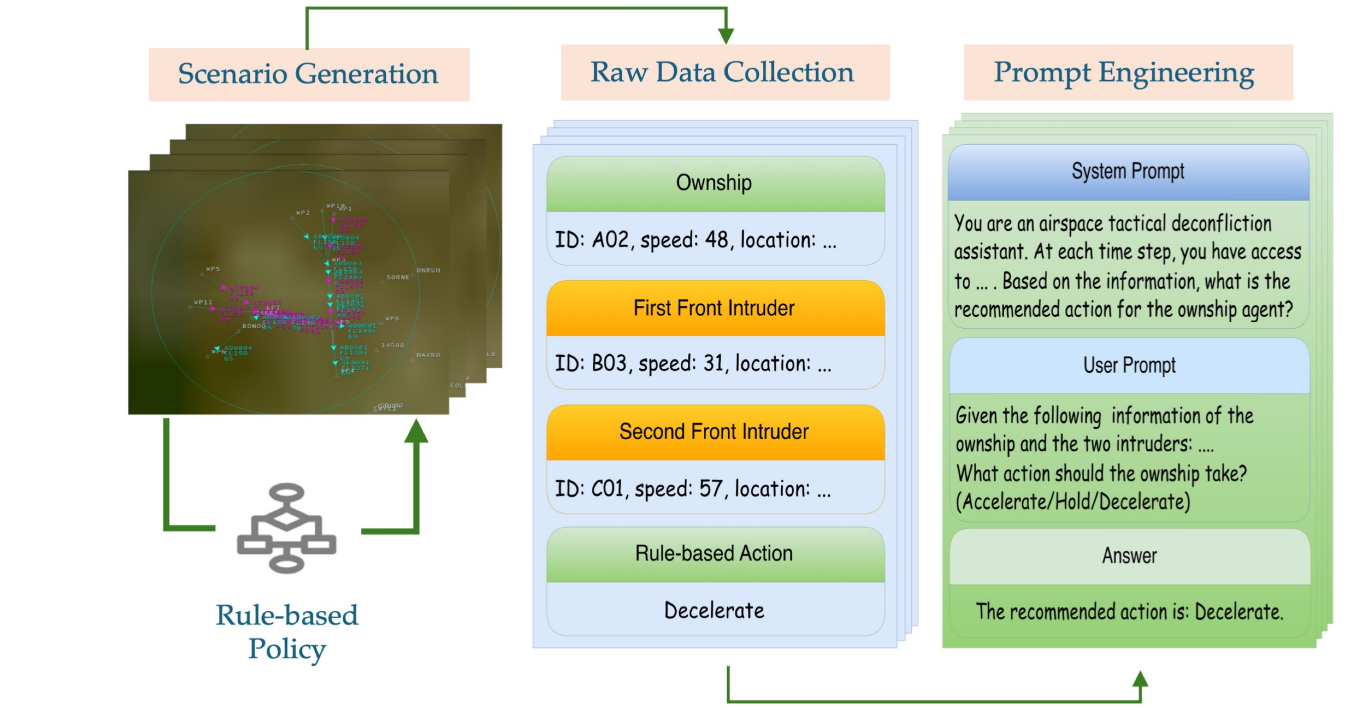
Proposed Approach



Main Contributions:

- We propose a **simulation-to-language** dataset generation pipeline.
- We fine-tune LLMs on the extracted dataset using two strategies:
  - **Supervised fine-tuning (SFT)** with low-rank adaptation (LoRA)
  - **Group-relative policy optimization (GRPO)**

Simulation-to-Language Pipeline



The pipeline **embeds human tactical reasoning** through interpretable rule-based supervision.

Simulation Details:

- ❑ BlueSky multi-agent simulator
- ❑ 20–30 sUASs per scenario
- ❑ 15–25 minute delivery operations in Frisco City, Dallas, Texas
- ❑ Heterogeneous drone configurations inspired by Google Wing and Amazon Prime Air drones
- ❑ Strong rule-based human-aligned labels, with ~38k samples
- ❑ Prompt engineering for structured reasoning

Experimental Results on the Test Dataset and Real-time Simulations

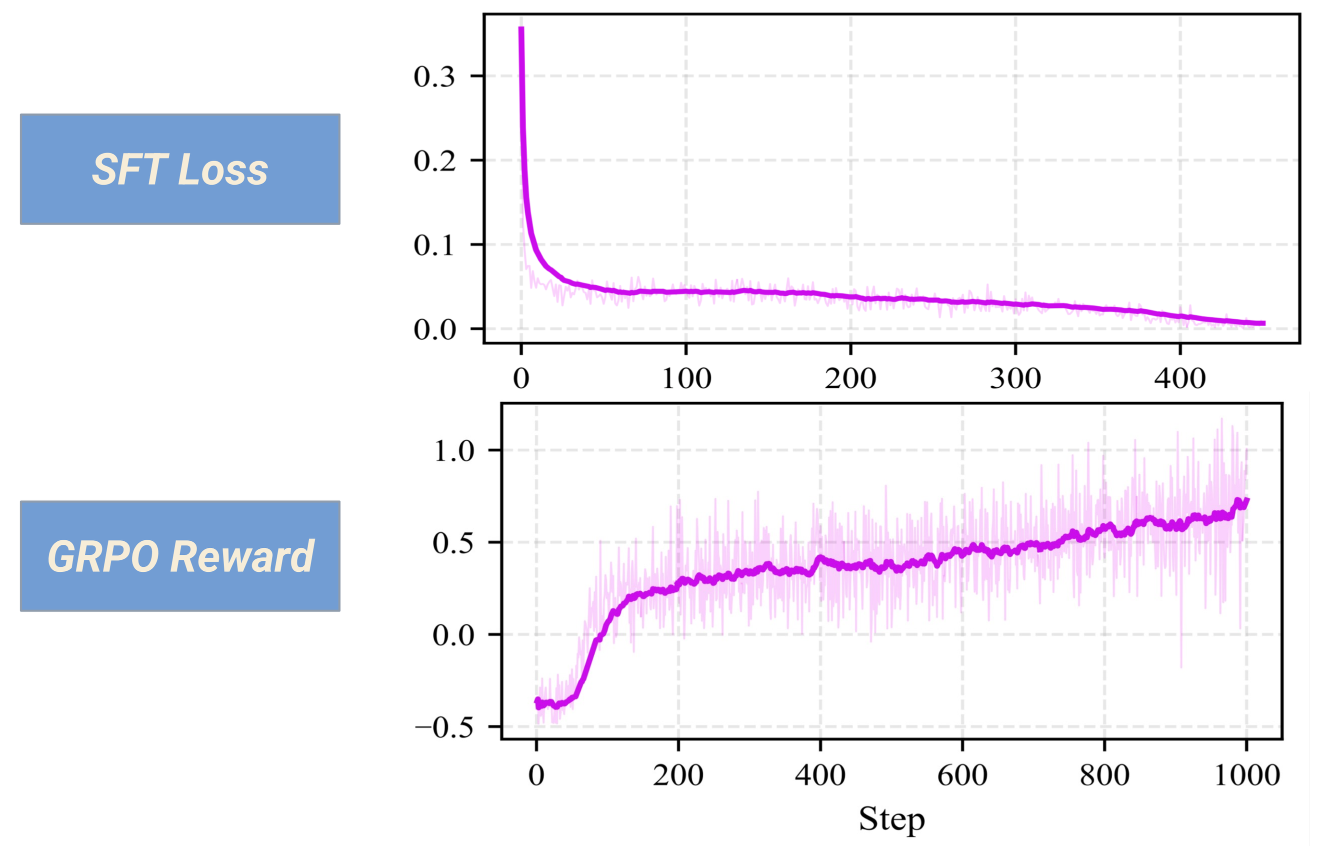
Test Dataset Results: Prediction Accuracy

Model	Accuracy	Precision	Recall	F1-score
Base	27	**75**	20	31
SFT	**88**	**75**	**66**	**69**
GRPO	53	**75**	40	50

BlueSky Simulation Results: Near mid-air collision (NMAC) rate for different combinations of Rule-based (R) and 10 LLM (L) agents interactions, and the operation time

	Base					SFT					GRPO				
Scen.	All	L–L	L–R	SR	Time	All	L–L	L–R	SR	Time	All	L–L	L–R	SR	Time
A	3.5±1.1	2.7±0.5	0.8±1.0	0.12±0.38	3.7±4.1	**1.0±0.8**	**0.7±0.7**	**0.3±0.5**	**0.77±0.29**	**5.7±0.6**	1.7±0.5	1.1±0.7	0.6±0.5	0.57±0.31	5.2±0.1
B	3.4±1.1	1.8±1.4	1.6±0.9	0.20±0.56	3.3±7.0	**1.9±1.2**	**0.9±0.7**	**1.0±0.8**	**0.62±0.36**	**8.1±0.9**	3.0±0.7	1.3±0.4	1.7±0.5	0.27±0.22	6.9±0.3
C	4.0±0.9	2.5±1.2	1.5±1.2	0.05±0.58	1.6±5.0	**1.9±0.7**	0.8±0.6	**1.1±0.9**	**0.52±0.37**	**7.5±0.7**	2.3±0.8	**0.6±0.5**	1.7±0.7	0.42±0.29	6.6±0.1

Training Performance



Conclusions

Key Takeaways:

- ❑ The proposed simulation-to-language pipeline facilitates fine-tuning LLMs on human-aligned tactical deconfliction strategies for small unmanned aerial systems.
- ❑ SFT provides the most reliable and robust performance for safety-critical deployment.
- ❑ GRPO also outperformed the Base model but did not achieve performance comparable to SFT.
- ❑ Both fine-tuning methods significantly reduced the number of near mid-air collisions (NMACs).

Challenges and Bottlenecks:

- ❖ High inference latency in centralized decision-making for multi-agent settings
- ❖ High computational cost and memory limitations, especially for GRPO, which requires generating multiple responses in each training iteration
- ❖ Lack of safety guarantees in safety-critical applications

Future Research:

Real-world data augmentation, safety module augmentation, comparison with rule-based and reinforcement learning methods

Contact

Iman Sharifi, Alex Zongo, and Peng Wei
George Washington University
Email: {i.sharifi, a.zongo, pwei}@gwu.edu
Website: https://web.seas.gwu.edu/pwei/

References

1. **Iman Sharifi, Alex Zongo, and Peng Wei. Fine-tuning Large Language Models for Cooperative Tactical Deconfliction for Small Unmanned Aerial Systems. In arXiv preprint arXiv:2603.28561, 2026.**
2. Justas Andriuskevicius and Junzi Sun. Automatic Control with Human-Like Reasoning: Exploring Language Model Embodied Air Traffic Agents. In 14th SESAR Innovation Days, SIDS 2024.
3. Alibaba Group AI Team. Qwen-Math: Mathematical Reasoning Models from Alibaba Cloud AI. Technical report, Alibaba Group, 2024.
4. Mark Brittain and Peng Wei. Autonomous Separation Assurance in An High-density En Route Sector: A Deep Multi-Agent Reinforcement Learning Approach. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pages 3256–3262, 2019.
5. Edward J. Hu, Yelong Shen, Philips Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-Rank Adaptation of Large Language Models. In International Conference on Learning Representations (ICLR), 2022.
6. Federico Errica, Davide Sanvito, Giuseppe Siracusano, and Roberto Bifulco. What Did I Do Wrong? Quantifying LLMs' Sensitivity and Consistency to Prompt Engineering. In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pages 1543–1558, 2025.
7. Dewi Gould, George De Ath, Ben Carvell, and Nick Pepper. AirTrafficGen: Configurable Air Traffic Scenario Generation with Large Language Models. ArXiv, abs/2508.02269, 2025.
8. Jacco Hoekstra and Joost Ellerbroek. BlueSky ATC Simulator Project: an Open Data and Open Source Approach. In International Conference on Research in Air Transportation, 2016.
9. Fabio Suim Chagas, Neno Ruseno, and Aurilla Aurelie Arntzen Bechina. Artificial Intelligence Approaches for UAV Deconfliction: A Comparative Review and Framework Proposal. In Automation, 6(4), 2025.
10. Kathleen Ge and William Coupe. Aviation-Specific Large Language Model Fine-Tuning and LLM-as-a-Judge Evaluation. In AIAA AVIATION FORUM and ASCEND 2025, page 3712, 2025.