Machine Learning-Based Detection of Cyberattacks Against UAVs

Alex Burnside and Abdulrahman Takiddin

1. Introduction

Cyber-physical systems (CPS) are systems that combine software, communication networks, and sensors to control physical devices. Examples include smart power grids, self-driving cars, and autonomous drones.

Disruptions to CPS caused by system faults or cyberattacks can quickly lead to physical consequences. In unmanned aerial vehicles (UAVs), commonly known as drones, corrupted data can destabilize flight, redirect navigation, or cause a complete loss of control.

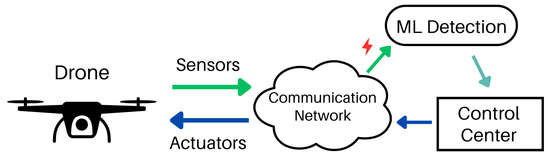

Figure 1. Simplified UAV cyber-physical feedback loop

Machine learning (ML), a form of AI, is widely used to detect abnormal behavior in CPS [2]. However, before advanced approaches can be trusted, their performance must be evaluated against simpler baseline methods. This project establishes these reproducible baselines for UAV attack detection.

2. Methods

- Drone flight data including normal operation and four simulated attack types (false data injection, denial of service, evil twin, and replay) was used [1].
- Data was partitioned into cyber (communication/control data), physical (onboard sensor data), and combined measurement sets.
- Five machine learning models were trained and tuned to detect compromised behavior, including two classical models and three neural network-based models [2].
- To simulate real-world data uncertainty, measurements were randomly altered by controlled percent-level deviations during testing (often called noise) [3].
- Performance was evaluated using detection accuracy and false alarm rate.

Figure 2. Simulated cyber-attack injection into communication layer

3. Results

Under ideal (noise-free) conditions, all models achieved **high detection accuracy**, ranging from 94.5% to 100%. False alarm rates (FAR) remained low, peaking at 7.2% depending on feature configuration and model choice.

As seen in Fig. 3, models trained on physical measurements **outperformed** those trained on cyber-only inputs, with Random Forest (RF) achieving the highest overall accuracy. However, combining cyber and physical measurements **slightly reduced** ideal accuracy.

These results suggest that physical measurements provide the **strongest signal** for attack detection under ideal conditions.

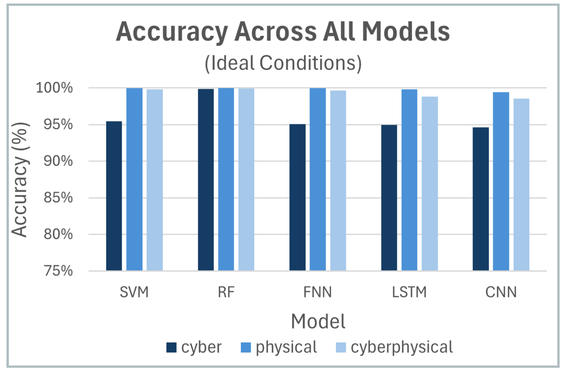

Figure 3. Detection Accuracy Across All Models With No Noise

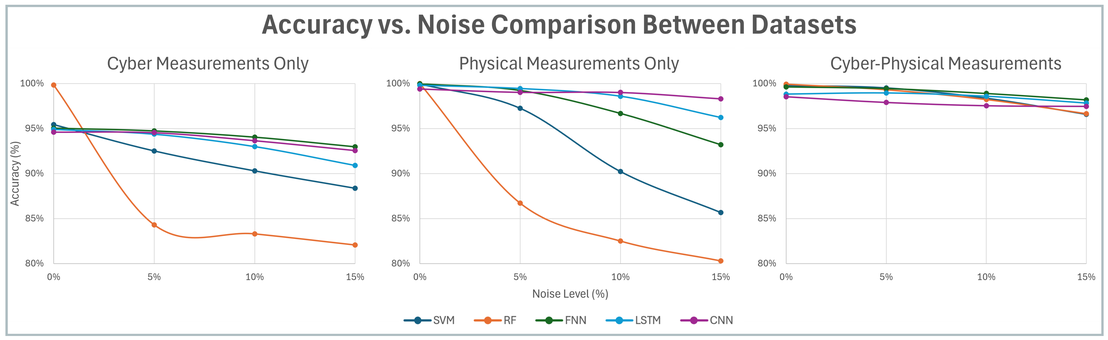

Figure 4. Detection Accuracy vs. Measurement Noise Across Datasets

When measurement noise was introduced (Fig. 4), performance **degraded unevenly** across datasets. Cyber-only and physical-only models showed substantial accuracy drops, with worst-case performance falling to 78%. In contrast, cyber-physical models maintained **consistently higher accuracy** across all noise levels.

False Alarm Rate vs. Noise Between Datasets

Model	Cyber Measurements Only				Physical Measurements Only				Cyber-Physical Measurements			
	None	Mild	Moderate	Severe	None	Mild	Moderate	Severe	None	Mild	Moderate	Severe
SVM	2%	15%	26%	35%	0%	1%	3%	5%	0%	0%	2%	4%
RF	0%	51%	53%	56%	0%	38%	50%	55%	0%	1%	3%	5%
FNN	0%	2%	6%	10%	0%	1%	3%	4%	0%	1%	1%	2%
LSTM	8%	8%	17%	21%	0%	1%	2%	9%	1%	1%	1%	3%
CNN	8%	9%	11%	17%	1%	1%	1%	3%	1%	1%	2%	2%

FAR (%) 0% – 60%

Figure 5. False Alarm Rate vs Measurement Noise Heatmap

Notably, neural network-based models were generally **more robust** to noise than classical approaches, which degraded more rapidly despite strong ideal performance. This pattern was less pronounced in false alarm behavior, although RF exhibited consistently high false alarm rates under noise (Fig. 5).

4. Conclusion

This study establishes reproducible baseline benchmarks for UAV attack detection across cyber, physical, and combined measurement sets. While physical-only models achieved the highest accuracy under ideal conditions, models trained on both cyber and physical measurements demonstrated a **greater robustness** to measurement noise.

These findings reveal a trade-off between peak detection accuracy and resilience to measurement uncertainty, and establish reproducible benchmarks for evaluating advanced UAV detection models.

5. Discussion

As shown in the results, performance under ideal conditions does not fully reflect real-world reliability. When measurement noise was introduced, models trained on cyber-only or physical-only inputs degraded substantially, whereas cyber-physical models remained more stable.

These findings are consistent with prior cyber-physical intrusion detection research showing that fusing cyber and physical measurements enhances robustness by providing complementary signals [4].

Limitations: Attacks and noise were simulated due to lack of publicly available data, and evaluation was not cross-validated against similar UAV test sets.

References

[1] S. C. Hassler et al., "Cyber-Physical Intrusion Detection System for UAVs," IEEE Trans. Intell. Transp. Syst., 2024.

[2] J. Richeson et al., "Ensemble Learning-Based Intrusion Detection for Aerial Base Stations," IEEE ICC, 2025.

[3] M. Karahan et al., "Nonlinear Modeling and Robust Control of a Quadrotor UAV Under White Gaussian Noise," ISMSIT, 2021.

[4] S. R. Fahim et al., "GNN-Based Detection of False Data Injection Attacks on Voltage Stability," IEEE Open Access J. Power Energy, 2025.