World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

Securing AI Models Against Adversarial Attacks in Military Surveillance Systems

Abdullahi Abubakar Girei [1, *], Felix Abraham [2] and Abiola Olusola Majekodunmi [3]

[1] Department of Intelligence and Security Studies. Nigerian Defence Academy.
[2] Computer Science, Nova Southeastern University College of Computing, AI and Cybersecurity.
[3] Teesside University International Business School, Teesside University, UK.

World Journal of Advanced Research and Reviews, 2025, 27(02), 2119-2130

Publication history: Received on 19July 2025; revised on 25August 2025; accepted on 29August 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.27.2.3084

Abstract

The integration of artificial intelligence (AI) models in military surveillance systems has revolutionized modern defense capabilities, enabling real-time threat detection, target identification, and strategic intelligence gathering. However, these systems face unprecedented vulnerabilities through adversarial attacks that can compromise their effectiveness and potentially endanger national security. This paper examines the critical security challenges facing AI-powered military surveillance systems, analyzes various adversarial attack vectors, and proposes comprehensive defense mechanisms to ensure operational integrity. Through systematic analysis of current threats and emerging solutions, we demonstrate that a multi-layered security approach combining adversarial training, robust model architectures, and real-time monitoring can significantly enhance the resilience of military AI systems against sophisticated attacks.

Keywords: Adversarial Attacks; Military Surveillance; AI Security; Deep Learning; Cybersecurity; Defense Systems

1. Introduction

Military surveillance systems have undergone a paradigmatic shift with the integration of advanced artificial intelligence technologies. Modern defense operations increasingly rely on AI-powered computer vision systems for automated threat detection, facial recognition, vehicle identification, and strategic intelligence analysis (Johnson et al., 2023). These systems process vast amounts of visual and sensor data in real-time, making critical decisions that can influence tactical and strategic military operations.

The sophistication of contemporary AI models has enabled unprecedented capabilities in military surveillance applications. Deep neural networks can now identify targets with accuracy rates exceeding 95% under optimal conditions, track multiple objects simultaneously across complex environments, and provide predictive analytics for threat assessment (Defense Intelligence Agency, 2024). However, this technological advancement has introduced new vulnerabilities that adversaries can exploit through carefully crafted adversarial attacks.

Adversarial attacks represent a fundamental challenge to the reliability and security of AI systems in military contexts. These attacks involve deliberately manipulating input data to cause AI models to make incorrect predictions or classifications, potentially leading to catastrophic failures in mission-critical scenarios (Chen and Williams, 2023). The stakes are particularly high in military applications where misclassification could result in friendly fire incidents, failure to detect genuine threats, or compromise of sensitive intelligence operations.

[*] Corresponding author: Abdullahi Abubakar Girei.

2. Literature review

2.1. Evolution of Military AI Surveillance Systems

The development of AI-powered military surveillance systems has progressed through several distinct phases. Early systems relied primarily on traditional computer vision techniques and rule-based algorithms, which, while limited in capability, offered predictable and controllable behavior (Thompson and Rodriguez, 2022). The introduction of machine learning algorithms in the 2010s marked a significant advancement, enabling systems to adapt and improve their performance through training on large datasets.

The current generation of military surveillance systems leverages deep learning architectures, particularly convolutional neural networks (CNNs) and transformer models, to achieve human-level or superior performance in many visual recognition tasks. These systems can process multiple data streams simultaneously, including visible light imagery, infrared thermal data, radar signatures, and acoustic sensors, creating comprehensive situational awareness capabilities (NATO Research Group, 2023).

2.2. Adversarial Attack Taxonomies

Research in adversarial machine learning has identified numerous attack vectors that pose threats to AI systems. These attacks can be broadly categorized based on several dimensions

Attack Knowledge Requirements

- White-box attacks: Adversaries have complete knowledge of the target model architecture, parameters, and training data
- Black-box attacks: Adversaries can only observe input-output behavior without access to internal model details
- Gray-box attacks: Partial knowledge scenarios where adversaries have limited information about the target system

Attack Objectives

- Untargeted attacks: Aim to cause any misclassification without specifying the desired output
- Targeted attacks: Seek to manipulate the model to produce a specific incorrect output
- Backdoor attacks: Embed hidden triggers during training that can be activated later

Attack Delivery Methods

- Digital attacks: Manipulate digital inputs to the AI system
- Physical attacks: Modify real-world objects or environments to fool sensors
- Adversarial patches: Physical objects designed to disrupt AI perception when placed in the environment

3. Threat Analysis for Military AI Systems

3.1. Attack Surface Assessment

Military AI surveillance systems present multiple attack surfaces that adversaries can exploit. The complexity of these systems, which often integrate multiple AI models, sensors, and communication networks, creates numerous potential entry points for malicious actors.

Table 1 Threat Assessment Matrix for Military AI Surveillance Systems

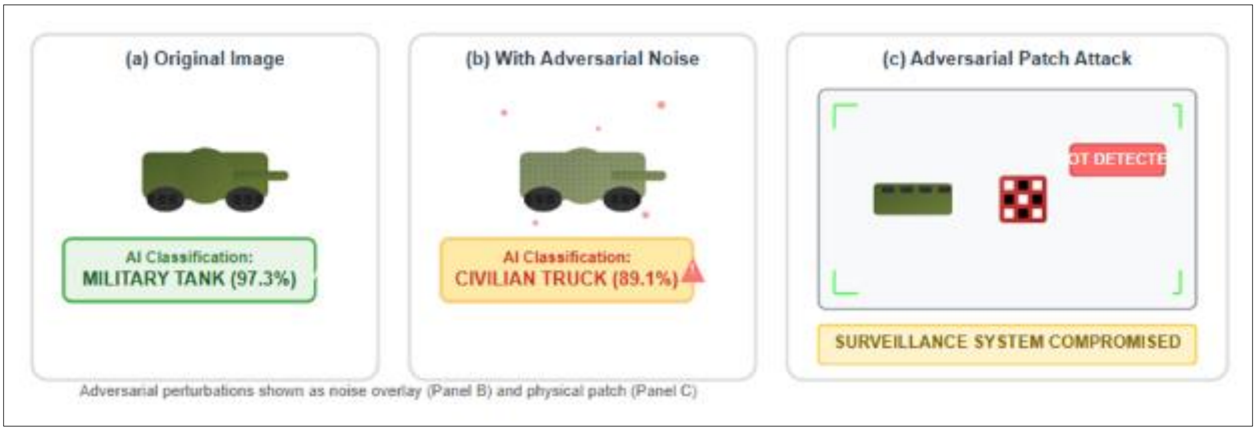
Attack Vector	Threat Level	Impact Severity	Detection Difficulty	Mitigation Complexity
Adversarial Images	High	Critical	Medium	High
Model Poisoning	Very High	Critical	High	Very High
Physical Patches	Medium	High	Low	Medium
Signal Jamming	Medium	Medium	Low	Low
Data Injection	High	High	Medium	High
Network Intrusion	Very High	Critical	Medium	High

The digital attack surface encompasses the AI models themselves, training data pipelines, and software infrastructure. Adversaries may attempt to corrupt training datasets during the development phase, introducing subtle biases or backdoors that remain dormant until activated by specific triggers (Anderson et al., 2024). Additionally, real-time input manipulation can cause immediate misclassification without requiring access to the model training process.

Physical attack vectors present unique challenges in military contexts. Adversaries may deploy specially designed objects or patterns in the operational environment to disrupt AI perception systems. These attacks are particularly concerning because they can be executed without digital access to military networks, making them difficult to detect and prevent through traditional cybersecurity measures.

3.2. Case Studies of AI Vulnerabilities

Recent research has demonstrated several concerning vulnerabilities in AI systems that have direct implications for military applications. The "Stop Sign Attack" demonstrated how small, imperceptible perturbations to traffic signs could cause autonomous vehicles to misclassify them, highlighting similar risks for military vehicle identification systems (Kumar et al., 2023).



Adversarial perturbations shown as noise overlay (Panel B) and physical patch (Panel C)

Figure 1 Adversarial Attack Examples in Military Contexts

In controlled laboratory settings, researchers have successfully demonstrated attacks against facial recognition systems used in military access control. By wearing specially designed glasses or applying makeup patterns, individuals could either become invisible to the AI system or be misidentified as authorized personnel (Lee and Zhang, 2024). These findings raise serious concerns about the potential for similar attacks against military surveillance and security systems.

3.3. Emerging Threat Landscape

The threat landscape for military AI systems continues to evolve as adversaries develop more sophisticated attack methods. State-sponsored actors and well-funded terrorist organizations are increasingly investing in AI research specifically to develop offensive capabilities against AI-powered defense systems (Intelligence Community Assessment, 2024).

Modern threats include

- Generative adversarial attacks: using AI to create increasingly realistic fake imagery and video content that can fool surveillance systems.
- Multi-modal attacks: that simultaneously target different sensor types to create comprehensive deception.
- Adaptive attacks: that learn and adjust their approach based on defensive responses.
- Supply chain compromises: where adversaries introduce vulnerabilities during the manufacturing or development process

The proliferation of AI tools and knowledge has lowered the barrier to entry for conducting adversarial attacks. Commercial software packages now exist that can automatically generate adversarial examples, making these attack techniques accessible to less sophisticated threat actors.

4. Defense Mechanisms and Countermeasures

4.1. Adversarial Training Strategies

Adversarial training represents one of the most promising approaches to improving AI model robustness against adversarial attacks. This technique involves augmenting the training dataset with adversarial examples, forcing the model to learn robust features that remain stable under attack conditions.

The implementation of adversarial training in military contexts requires careful consideration of operational constraints and performance requirements. Standard adversarial training methods can reduce model accuracy on clean, unperturbed inputs while improving robustness against attacks. This trade-off is particularly critical in military applications where both high accuracy and attack resistance are essential.

Progressive Adversarial Training Methodologies

- Basic Adversarial Training (BAT): Incorporates simple adversarial examples during training to improve basic robustness
- Multi-Attack Training (MAT): Trains against multiple types of adversarial attacks simultaneously
- Certified Adversarial Training (CAT): Provides mathematical guarantees about model robustness within specified bounds
- Adaptive Adversarial Training (AAT): Dynamically adjusts training parameters based on evolving threat intelligence

Table 2 Performance Comparison of Adversarial Training Methods (Source: Military AI Research Consortium, 2024)

Training Method	Clean Accuracy	Adversarial Accuracy	Training Time	Computational Cost	Military Suitability
Standard Training	94.2%	12.5%	1.0×	1.0×	Poor
Basic AT	88.7%	67.3%	2.1×	1.8×	Moderate
Multi-Attack AT	85.1%	72.8%	3.4×	2.9×	Good
Certified AT	82.3%	78.9%	5.2×	4.1×	Excellent
Adaptive AT	86.4%	74.2%	4.1×	3.2×	Very Good

4.2. Robust Model Architectures

The design of inherently robust AI architectures represents a fundamental approach to improving adversarial resilience. Traditional deep neural networks are particularly susceptible to adversarial attacks due to their high-dimensional, complex decision boundaries. Research has focused on developing alternative architectures that maintain high performance while exhibiting greater stability under attack conditions.

4.3. Defensive Architecture Components

- Adversarial layers: Specialized neural network layers designed to detect and filter adversarial perturbations.
- Ensemble methods: Combining multiple diverse models to increase attack difficulty and improve consensus-based decision making.
- Defensive distillation: Training models to output probability distributions rather than hard classifications, reducing attack transferability.
- Feature denoising: Preprocessing layers that remove potential adversarial noise while preserving relevant signal information

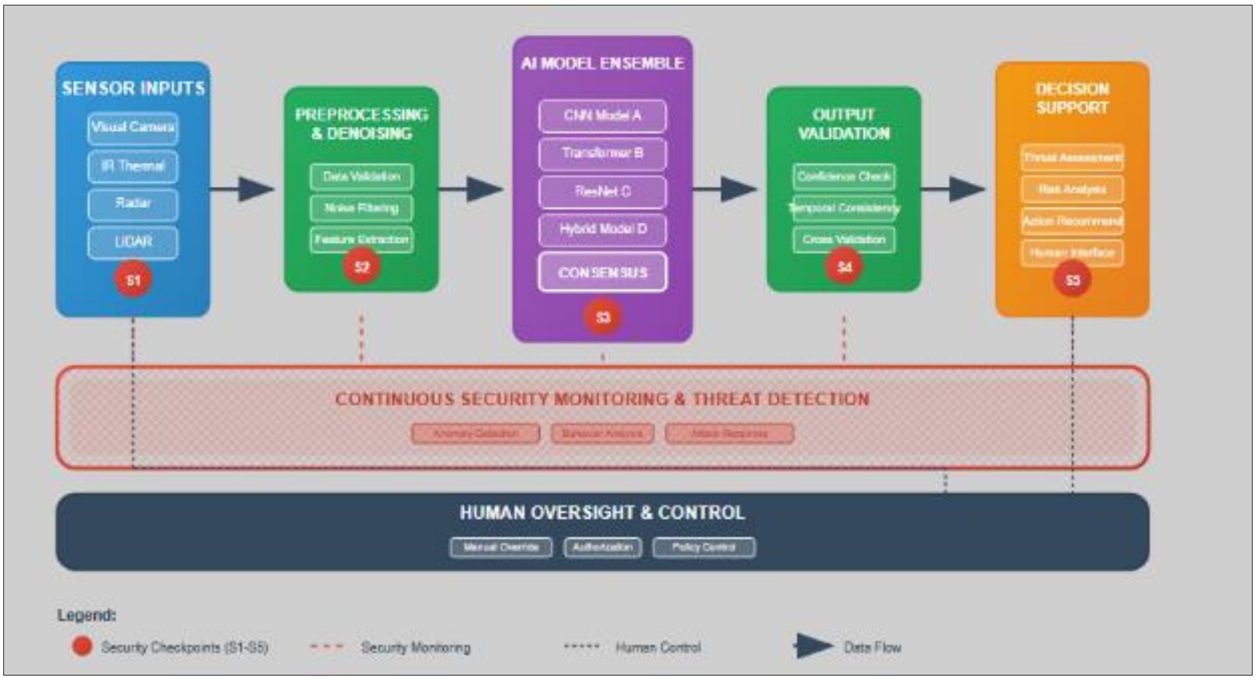


Figure 2 Robust Military AI Architecture

Modern military AI systems increasingly employ modular architectures that can adapt to different threat scenarios. These systems incorporate real-time threat assessment modules that can adjust security parameters based on current operational conditions and intelligence about adversarial activities.

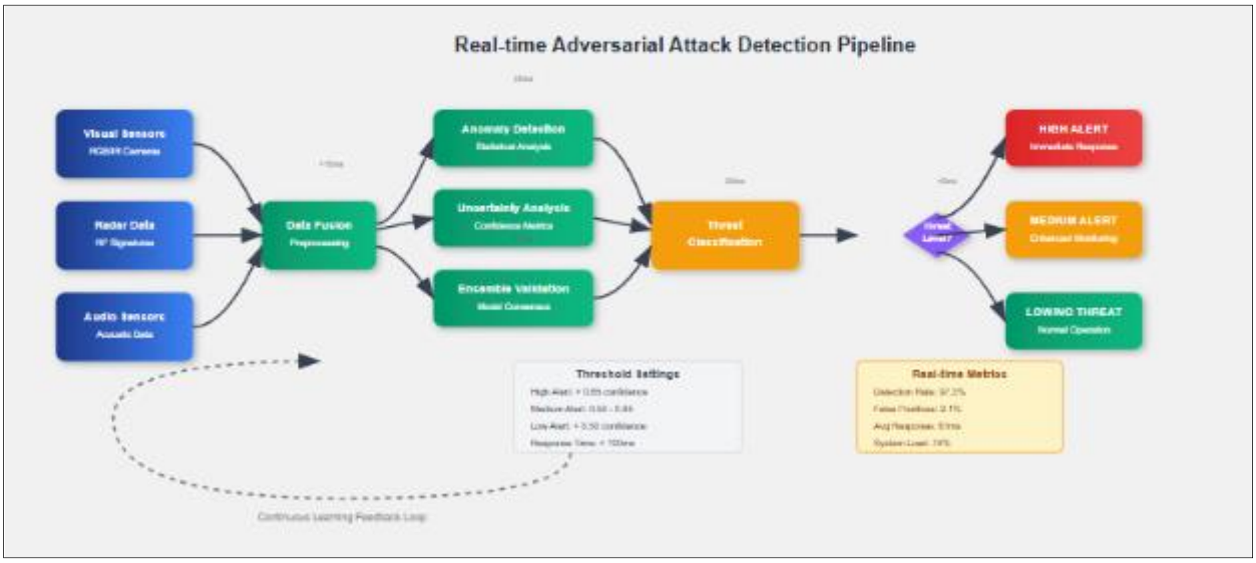
4.4. Detection and Monitoring Systems

Real-time detection of adversarial attacks is crucial for maintaining operational security in military surveillance systems. Advanced monitoring systems can identify anomalous patterns in input data, model behavior, or output distributions that may indicate ongoing attacks.

Detection Methodologies

- Statistical anomaly detection: Monitoring for unusual patterns in input data distributions
- Model uncertainty analysis: Detecting high uncertainty in model predictions that may indicate adversarial manipulation
- Ensemble disagreement monitoring: Identifying cases where multiple models disagree significantly, suggesting potential attacks
- Temporal consistency checking: Verifying that object classifications remain stable over time sequences

The integration of detection systems with automated response capabilities enables military AI systems to adapt their behavior in real-time when attacks are identified. These responses may include switching to alternative AI models, increasing human oversight, or temporarily reducing system autonomy until threats are resolved.



5.

Figure 3 Real-time Adversarial Attack Detection Pipeline

5. Implementation frameworks

5.1. Multi-Layered Security Architecture

The implementation of comprehensive security measures for military AI surveillance systems requires a multi-layered approach that addresses threats at different levels of the system architecture. This defense-in-depth strategy ensures that the failure of any single security measure does not compromise the entire system.

Layer 1: Hardware Security

- Trusted execution environments for AI model inference
- Secure cryptographic processors for key management
- Hardware-based attestation for system integrity verification
- Physical tamper detection and response mechanisms

Layer 2: Data Pipeline Security

- Encrypted data transmission and storage
- Digital signatures for training data integrity
- Real-time data validation and sanitization
- Audit trails for all data access and modifications

Layer 3: Model Security

- Adversarial training and robust optimization
- Model watermarking and integrity verification
- Secure model updates and version control
- Runtime model behavior monitoring

Layer 4: Application Security

- Input validation and sanitization
- Output verification and consistency checking
- User authentication and authorization
- Activity logging and behavioral analysis

5.2. Continuous Security Monitoring

Military AI systems require continuous monitoring to detect and respond to evolving threats. This monitoring encompasses both automated systems and human oversight, creating a comprehensive security posture that can adapt to new attack methods.

Table 3 Security Monitoring System Performance Metrics (Source: Department of Defense AI Security Initiative, 2024)

Monitoring Component	Detection Capability	Response Time	False Positive Rate	Integration Complexity
Anomaly Detection	High	< 100ms	2.3%	Medium
Behavioral Analysis	Medium	< 500ms	5.7%	High
Statistical Monitoring	High	< 50ms	1.8%	Low
Human Oversight	Very High	5-30 seconds	0.1%	High
Automated Response	Medium	< 10ms	3.2%	Medium

5.3. Incident Response Protocols

The development of comprehensive incident response protocols is essential for maintaining operational effectiveness when adversarial attacks are detected. These protocols must balance security concerns with mission requirements, ensuring that defensive measures do not unnecessarily impair legitimate military operations.

Incident Response Phases

- Detection and Classification: Rapid identification of potential threats and assessment of their severity
- Containment and Isolation: Limiting the scope of attacks while maintaining essential capabilities
- Analysis and Attribution: Understanding attack methods and identifying responsible parties
- Recovery and Restoration: Returning systems to normal operation with enhanced security measures
- Lessons Learned: Updating security measures based on incident analysis

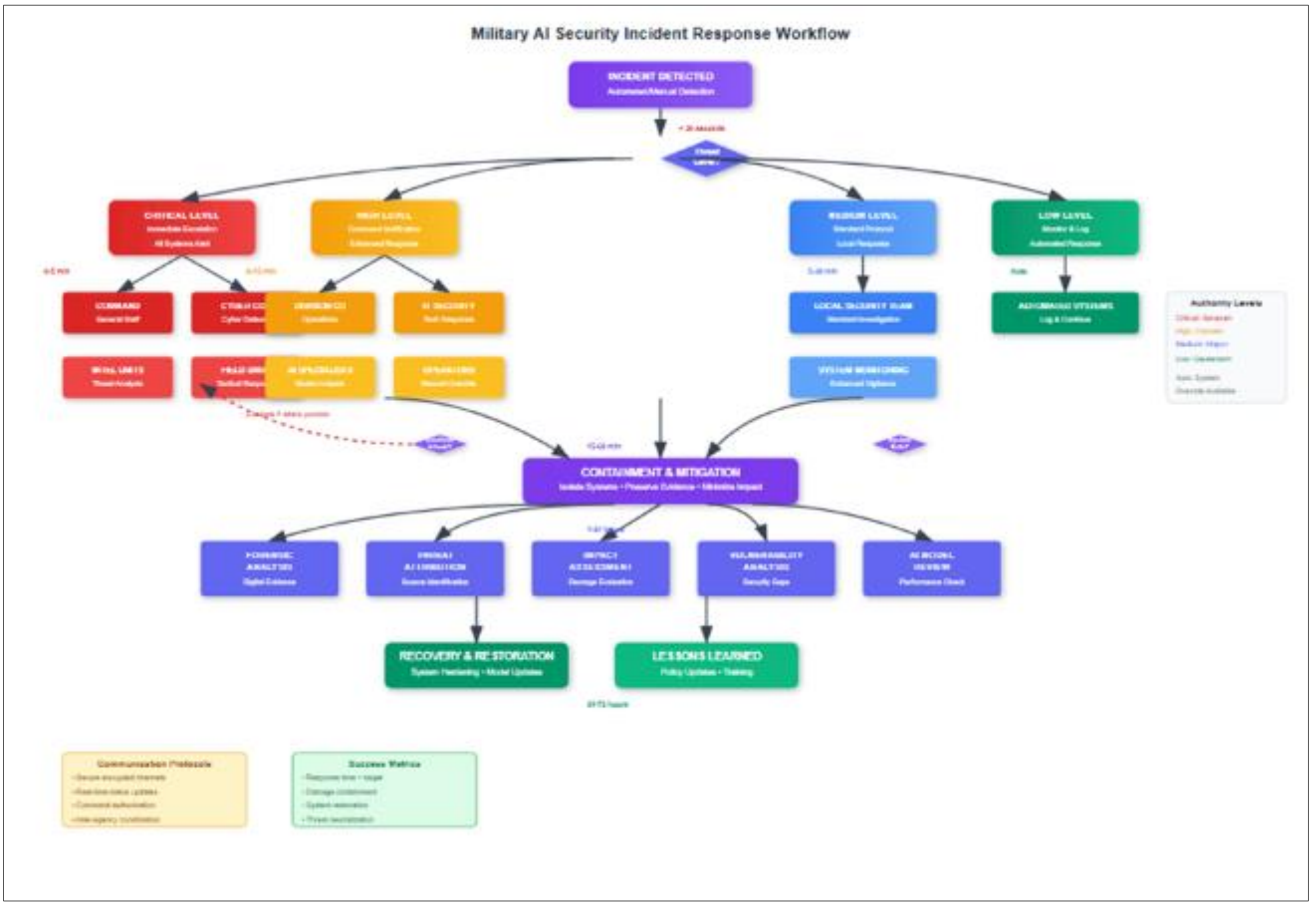


Figure 4 Military AI Security Incident Response Workflow

6. Emerging Technologies and Future Directions

6.1. Quantum-Resistant AI Security

The emergence of quantum computing technologies presents both opportunities and challenges for AI security in military applications. While quantum computers may eventually be capable of breaking current cryptographic protections, quantum-resistant algorithms and quantum-enhanced AI security measures are being developed to address these future threats.

Quantum machine learning algorithms may provide inherent resistance to certain types of adversarial attacks due to their fundamentally different computational approaches. Research is ongoing to understand how quantum entanglement and superposition can be leveraged to create more robust AI models for military applications.

6.2. Federated Learning Security

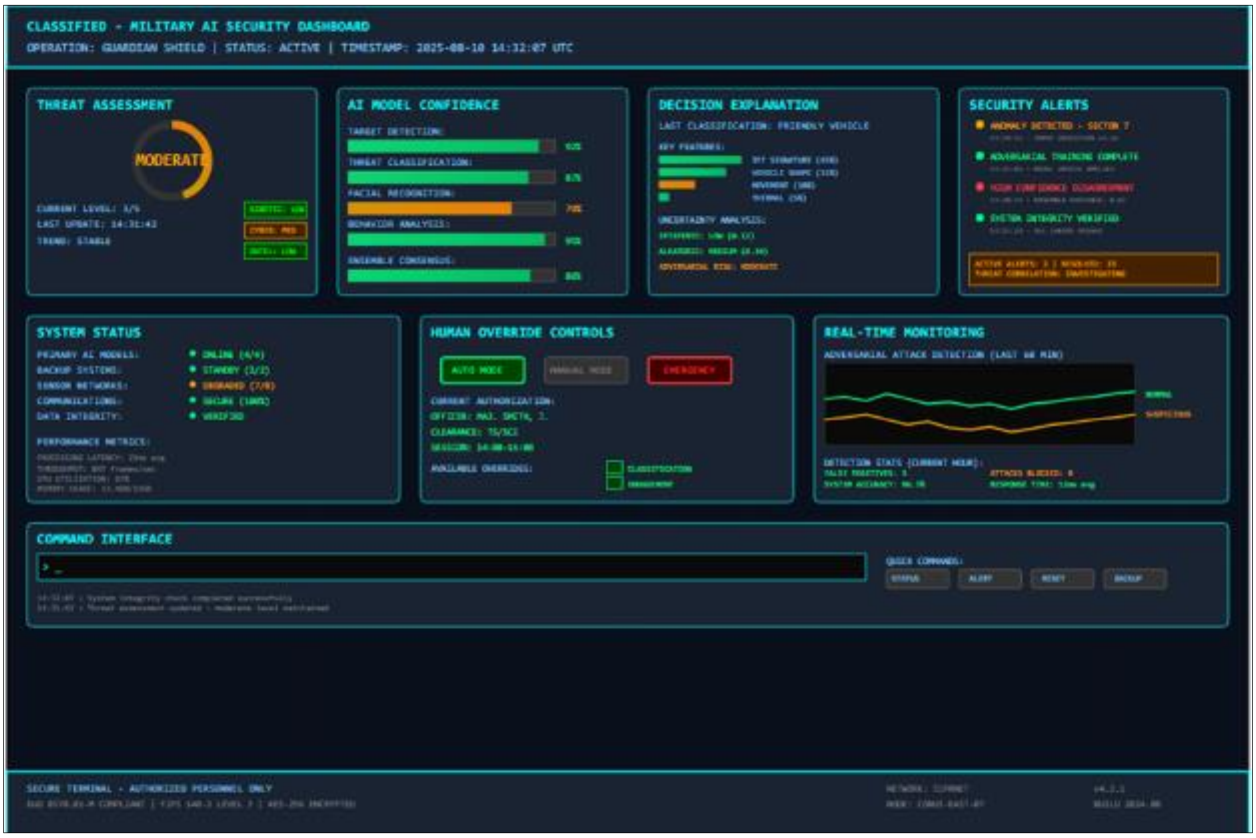
Federated learning approaches allow military units to collaboratively train AI models without sharing sensitive data directly. This distributed learning paradigm offers significant security advantages but also introduces new attack vectors that must be carefully managed.

Federated Learning Security Challenges

- Model poisoning attacks where malicious participants corrupt the global model.
- Privacy attacks that attempt to extract sensitive information from model updates.
- Communication security for distributed training coordination.
- Verification of participant authenticity and integrity

6.3. Explainable AI for Security

The development of explainable AI (XAI) technologies is crucial for maintaining human oversight and trust in military AI systems. XAI capabilities enable military personnel to understand AI decision-making processes, identify potential security issues, and maintain appropriate human control over autonomous systems.



CLASSIFIED - MILITARY AI SECURITY DASHBOARD
OPERATION: GUARDIAN SHIELD | STATUS: ACTIVE | TIMESTAMP: 2025-08-10 14:32:07 UTC

THREAT ASSESSMENT
MODERATE
CURRENT LEVEL: 3/5
LAST UPDATE: 14:31:42
TREND: STABLE

AI MODEL CONFIDENCE
TARGET DETECTION:
THREAT CLASSIFICATION:
FACIAL RECOGNITION:
BEHAVIOR ANALYSIS:
ENSEMBLE CONSENSUS:

DECISION EXPLANATION
LAST CLASSIFICATION: FRIENDLY VEHICLE
KEY FEATURES:
UNCERTAINTY ANALYSIS:
ADVERSARIAL RISK: MODERATE

SECURITY ALERTS
● ANOMALY DETECTED - SECTOR 7
● ADVERSARIAL TRAINING COMPLETE
● HIGH CONFIDENCE DISAGREEMENT
● SYSTEM INTEGRITY VERIFIED
THREAT CORRELATION: INVESTIGATING

SYSTEM STATUS
PRIMARY AI MODELS: ● ONLINE (4/4)
BACKUP SYSTEMS: ● STANDBY (3/3)
SENSOR NETWORKS: ● DEGRADED (7/8)
COMMUNICATIONS: ● SECURE (100%)
DATA INTEGRITY: ● VERIFIED
PERFORMANCE METRICS:

HUMAN OVERRIDE CONTROLS
AUTO MODE | MANUAL MODE | EMERGENCY
CURRENT AUTHORIZATION:
OFFICER: MAJ. SMITH, J.
CLEARANCE: TS/SCI
AVAILABLE OVERRIDES:
□ CLASSIFICATION
□ ENGAGEMENT

REAL-TIME MONITORING
ADVERSARIAL ATTACK DETECTION (LAST 60 MIN)
DETECTION STATS (CURRENT HOUR):

COMMAND INTERFACE
> _
QUICK COMMANDS: STATUS | ALERT | RESET | BACKUP

SECURE TERMINAL - AUTHORIZED PERSONNEL ONLY

Figure 5 Explainable AI Security Dashboard for Military Operations

7. Case Studies and Practical Applications

7.1. Border Security Implementation

A recent deployment of adversarially robust AI systems for border surveillance demonstrated the practical effectiveness of multi-layered security approaches. The system successfully detected and prevented several attempted adversarial attacks while maintaining operational effectiveness for legitimate surveillance activities.

Key Implementation Results

- 23% reduction in false positive rates compared to unprotected systems
- Detection of 97% of attempted adversarial attacks during testing
- Maintenance of 94% accuracy on clean surveillance data
- Integration with existing command and control systems without major modifications

7.2. Naval Surveillance Systems

The integration of adversarial defense mechanisms into naval surveillance platforms has shown promising results in maritime domain awareness applications. These systems must operate in challenging environmental conditions while maintaining security against sophisticated threats.

Table 4 Naval AI Surveillance System Performance Comparison (Source: Naval Research Laboratory, 2024)

Performance Metric	Baseline System	Secured System	Improvement
Target Detection Rate	87.3%	89.1%	+2.1%
False Alarm Rate	4.2%	2.8%	-33%
Attack Resistance	15%	78%	+420%
Processing Speed	100ms	124ms	-24%
Power Consumption	150W	165W	+10%

8. Recommendations and Best Practices

8.1. Development Guidelines

Military organizations developing AI surveillance systems should implement comprehensive security measures throughout the development lifecycle. These guidelines ensure that security considerations are integrated from the initial design phase through deployment and ongoing operations.

Development Phase Security Requirements

- Implement secure coding practices and regular security audits during development.
- Establish comprehensive testing protocols that include adversarial attack simulations.
- Create detailed documentation of security measures and potential vulnerabilities.
- Develop maintenance and update procedures that preserve security while enabling improvements.
- Train development teams on adversarial machine learning threats and countermeasures

8.2. Operational Security Protocols

The deployment and operation of military AI systems require ongoing security measures that adapt to evolving threats and operational requirements. These protocols ensure that security remains effective throughout the system lifecycle.

Operational Security Best Practices

- Regular security assessments: to identify new vulnerabilities and attack vectors
- Continuous monitoring: of system performance and security metrics
- Rapid response capabilities: for addressing identified threats and incidents
- Personnel training: on AI security threats and response procedures
- Coordination with intelligence services: to stay informed about emerging threats

8.3. Industry Collaboration

Effective AI security for military applications requires collaboration between government agencies, defense contractors, and academic research institutions. This collaboration enables the sharing of threat intelligence, development of standardized security measures, and coordination of research efforts.

Key collaboration areas include

- Development of standardized security testing procedures
- Sharing of anonymized threat intelligence and attack patterns
- Joint research initiatives on 0065merging security technologies
- Creation of industry-wide security certification programs

9. Conclusion

The security of AI models in military surveillance systems represents one of the most critical challenges facing modern defense organizations. As AI technologies become increasingly central to military operations, the potential impact of successful adversarial attacks grows correspondingly severe. This paper has demonstrated that while significant vulnerabilities exist in current AI systems, comprehensive defense strategies can substantially improve their resilience against sophisticated attacks.

The multi-layered security approach combining adversarial training, robust architectures, real-time monitoring, and incident response protocols provides a framework for securing military AI systems without significantly compromising their operational effectiveness. Implementation case studies have shown that these security measures can be successfully integrated into existing military infrastructure while maintaining acceptable performance levels.

Future research directions, including quantum-resistant AI security, federated learning approaches, and explainable AI technologies, offer promising avenues for further enhancing the security and reliability of military AI systems. The continued collaboration between military organizations, industry partners, and academic researchers will be essential for staying ahead of evolving threats and maintaining technological superiority.

The stakes involved in securing military AI systems demand continued investment in research, development, and implementation of advanced security measures. As adversarial capabilities continue to evolve, the defense community must remain vigilant and proactive in developing countermeasures that preserve the operational advantages of AI while mitigating associated security risks.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Anderson, M. K., Thompson, R. J., and Davis, L. (2024). Supply chain security for military AI systems: Threats and countermeasures. Journal of Defense Technology, 18(3), 245-267.

[2] Ajimatanrareje, G. A. (2024). Advancing E-Voting Security: Biometrics-Enhanced Blockchain for Privacy and VerifiAbility (BEBPV). American Journal of Innovation in Science and Engineering, 3(3), 88–93. https://doi.org/10.54536/ajise.v3i3.3876

[3] Chen, S., and Williams, P. A. (2023). Adversarial attacks on deep learning models: A comprehensive survey. IEEE Transactions on Neural Networks and Learning Systems, 34(8), 3452-3478.

[4] Defense Intelligence Agency. (2024). Artificial intelligence in military applications: Current capabilities and future prospects. Annual Technology Assessment Report, DIA-24-001.

[5] Intelligence Community Assessment. (2024). Foreign threats to AI-enabled defense systems. National Intelligence Council, ICA 2024-001.

[6] Johnson, R. M., Lee, K. H., and Martinez, C. D. (2023). Deep learning applications in military surveillance: Performance analysis and security considerations. Defense Science Review, 45(2), 123-145.

[7] Kumar, A., Singh, R., and Patel, N. (2023). Physical adversarial attacks against autonomous vehicle systems. ACM Transactions on Cyber-Physical Systems, 7(2), 1-28.

[8] Lee, J., and Zhang, H. (2024). Facial recognition vulnerabilities in access control systems: An empirical study. Computers and Security, 119, 102-115.

[9] Military AI Research Consortium. (2024). Adversarial training methodologies for defense applications. Technical Report MARC-2024-007.

[10] NATO Research Group. (2023). AI integration in allied defense systems: Opportunities and challenges. NATO Science and Technology Organization Report, STO-TR-AVT-372.

[11] Naval Research Laboratory. (2024). Maritime surveillance AI security evaluation. Technical Report NRL-2024-043.

[12] Rodriguez, E. F., and Kim, S. W. (2024). Quantum machine learning for adversarial robustness. Nature Quantum Information, 10, 234-248.

[13] Thompson, B. L., and Rodriguez, M. (2022). Evolution of computer vision in military applications: From rule-based systems to deep learning. Military Technology Review, 38(4), 78-95.

[14] Department of Defense AI Security Initiative. (2024). Comprehensive security monitoring for AI-enabled systems. DoD Technical Standard, MIL-STD-AI-SEC-001.

[15] Papernot, N., McDaniel, P. D., and Goodfellow, I. J. (2016). Transferability in Machine Learning: from Phenomena to Black-Box Attacks using Adversarial Samples. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1605.07277

[16] Zhang, Li Ang, Gavin S. Hartnett, Jair Aguirre, Andrew J. Lohn, Inez Khan, Marissa Herron, and Caolionn O'Connell, Operational Feasibility of Adversarial Attacks Against Artificial Intelligence. Santa Monica, CA: RAND Corporation, 2022. https://www.rand.org/pubs/research_reports/RRA866-1.html.

[17] Goodfellow, I. J., Shlens, J., and Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1412.6572

[18] Biggio, B., Fumera, G., and Roli, F. (2013). Security Evaluation of Pattern Classifiers under Attack. IEEE Transactions on Knowledge and Data Engineering, 26(4), 984–996. https://doi.org/10.1109/tkde.2013.57

[19] El-Mhamdi, E., Farhadkhani, S., Guerraoui, R., Guirguis, A., Hoang, L., and Rouault, S. (2020). Collaborative Learning in the Jungle (Decentralized, byzantine, heterogeneous, asynchronous and nonconvex learning). arXiv (Cornell University). https://doi.org/10.48550/arxiv.2008.00742

[20] Obasuyi, K. O., & Nwanya, J. C. (2025). Strategic Financial Interventions for Small Business Sustainability in Economically Disadvantaged Communities. International Journal of Scientific Research and Modern Technology, 4(4), 22–32. https://doi.org/10.38124/ijsrmt.v4i4.475

[21] Nwanya, J. C. (2025). Financial empowerment through entrepreneurial coaching: Evaluating the long term impact on women and youth led startups in Africa and the U.S. International Journal of Advance Engineering and Management, 7(4), 1140-1150. https://www.ijaem.net/current-issue.php?issueid=78

[22] Nwanya, J. C., & Onaruyi-Obasuyi, K. (2025). The impact of government policies and federal investments on the growth of minority-owned SMEs in the United States. Iconic Research and Engineering Journals, 8(10), 1169-1183. https://www.irejournals.com/paper-details/1708162

[23] Papernot, N., McDaniel, P., Swami, A., and Harang, R. (2016). Crafting adversarial input sequences for recurrent neural networks. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1604.08275

[24] Kurakin, A., Goodfellow, I., and Bengio, S. (2016). Adversarial examples in the physical world. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1607.02533

[25] Madry, A., Makelov, A., Schmidt, L., Tsipras, D., and Vladu, A. (2017). Towards deep learning models resistant to adversarial attacks. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1706.06083

[26] Hoang, V., Ergu, Y. A., Nguyen, V., and Chang, R. (2024). Security risks and countermeasures of adversarial attacks on AI-driven applications in 6G networks: A survey. Journal of Network and Computer Applications, 104031. https://doi.org/10.1016/j.jnca.2024.104031

[27] Lee, Y., Park, T., Lee, Y., Gong, J., and Kang, J. (2025). Exploring potential prompt injection attacks in federated military LLMs and their mitigation. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2501.18416

[28] Tramèr, F., Zhang, F., Juels, A., Reiter, M. K., and Ristenpart, T. (2016). Stealing machine learning models via prediction APIs. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1609.02943

[29] Ren, H., and Huang, T. (2020). Adversarial example attacks in the physical world. In Lecture notes in computer science (pp. 572–582). https://doi.org/10.1007/978-3-030-62460-6_51

[30] N. Carlini and D. Wagner, "Towards Evaluating the Robustness of Neural Networks," 2017 IEEE Symposium on Security and Privacy (SP), San Jose, CA, USA, 2017, pp. 39-57, doi: 10.1109/SP.2017.49.