AI-Powered Object Detection in Satellite Imagery for Military Reconnaissance

David Sitali
Dept. of Computer Science
Mulungushi University, Kabwe, Zambia
davidsitali18@gmail.com

Brian Halubanza
Dept. of Computer Science
Mulungushi University
Kabwe, Zambia
bhalubanza@gmail.com

Zilani Kaluba
Mulungushi University
Kabwe, Zambia -
zilanikaluba@gmail.com

Maines Namuchile
Mulungushi University
Kabwe, Zambia -
mainessnamuchile4@gmail.com

Michael Bwalya
Mulungushi University,
Kabwe, Zambia -
mikob87@gmail.com

Abstract

Timely and accurate interpretation of satellite imagery plays a vital role in modern military reconnaissance. This paper proposes VisionAI, a robust, AI-powered object detection system built on the YOLOv8 architecture, optimized for detecting military assets such as tanks, aircraft, trucks, and naval vessels. The system was fine-tuned on a custom remote sensing dataset and deployed using Google Cloud's T4 GPU infrastructure for real-time inference. The model achieved a mean Average Precision (mAP@0.5) of 0.79 for aircraft and maintained high precision and recall across key object categories. VisionAI demonstrates strong resilience to environmental distortions including cloud occlusion, low lighting, and motion blur. This work builds upon previous efforts in lightweight detection frameworks using MobileNetV2 for pest surveillance in locust management campaigns [1], as well as scalable AI pipelines for real-time monitoring in resource-constrained settings [2]. Furthermore, it aligns with recent advances in satellite surveillance and small object detection using cross-scale and pyramid fusion methods [3], [4]. Challenges related to detecting camouflaged or low-resolution naval targets persist, underscoring the need for hybrid approaches combining multispectral data and transformer-based architectures. Ethical considerations around adversarial manipulation and dual-use of AI in military contexts are also discussed. This research offers a cost-effective, adaptable, and ethically aware solution for defense-oriented remote sensing operations in developing regions.

Keywords: Object Detection, Satellite Imagery, YOLOv8, Military Reconnaissance, Deep Learning, AI Ethics

I. Introduction

Military Intelligence, Surveillance, and Reconnaissance (ISR) systems are increasingly dependent on the rapid, accurate interpretation of satellite imagery to support decision-making in both defense and humanitarian operations. The emergence of high-resolution Earth observation satellites has opened new frontiers in object detection, enabling analysts to track vehicles, aircraft, encampments, and naval movements over vast terrains in near real-time. However, traditional methods of image analysis often reliant on human visual inspection or handcrafted features struggle to scale and generalize across varied operational conditions, such as cloud occlusion, low light, camouflage, or terrain clutter.

To address these challenges, artificial intelligence (AI), and in particular deep learning-based object detection, has become a transformative enabler in remote sensing [1]. Early models such as Faster R-CNN offered remarkable accuracy but limited real-time performance [2]. Recent architectures, such as YOLOv5 and EfficientDet, improved the trade-off between speed and precision. The latest evolution, YOLOv8, introduces an anchor-free design and attention-aware detection head, positioning it as a state-of-the-art real-time detection model suitable for satellite imagery analysis [3].

In military contexts, object detection systems must contend with the complexities of small-object recognition, varying spatial resolutions, and environmental distortions. Transformer-based models such as DETR and hybrid fusion methods such as

 ISBN: 978-9982-95-500-3 ICICT2025

Pyramid Converge-and-Assign Fusion (PCAF) have recently shown promise in addressing these challenges [4], [5]. Nonetheless, model deployment in low-resource defense infrastructures (e.g., developing regions) remains a significant concern due to cost, compute requirements, and real-time inference constraints.

This paper introduces VisionAI, an AI-powered object detection framework based on YOLOv8, designed to detect military-relevant assets from satellite imagery. Building upon prior work in lightweight detection systems for locust surveillance using MobileNetV2 [6] and real-time monitoring architectures in Zambia's locust management systems [7], this study demonstrates the applicability of modern AI models to defense ISR tasks under practical constraints. VisionAI is trained on a custom, class-balanced dataset containing annotated tanks, aircraft, trucks, tents, helicopters, and ships. Extensive evaluations under occlusion, noise, and lighting variations illustrate its robustness. Notably, the system maintains high precision in detecting large objects such as aircraft, though naval object detection remains challenging due to scale and water-background variability.

Furthermore, the study addresses ethical implications of AI in defense, including dual-use concerns and susceptibility to adversarial attacks, aligning with global discourse on responsible military AI deployment [8]. The proposed system aims to contribute a scalable, cost-effective, and ethically conscious solution for ISR tasks in resource-limited military environments.

II. Related Work

Early methods for satellite-based object detection primarily relied on handcrafted features such as histogram of oriented gradients (HOG), template matching, and edge-based segmentation. While computationally efficient, these classical techniques were highly sensitive to lighting changes, background clutter, and object scale variations, limiting their applicability in dynamic military contexts.

With the advent of deep learning, convolutional neural networks (CNNs) revolutionized image recognition. Region-based CNNs such as R-CNN, Fast R-CNN, and Faster R-CNN, significantly improved detection accuracy, particularly in structured scenes. However, their region proposal mechanisms introduced latency, rendering them suboptimal for real-time applications such as ISR [1].

The introduction of the YOLO (You Only Look Once) framework marked a critical shift towards unified and real-time object detection. YOLOv3 through YOLOv5 struck a balance between inference speed and accuracy, widely adopted in drone, satellite, and surveillance tasks. YOLOv8, the latest in this lineage, introduces anchor-free object detection and decoupled head architectures, improving generalization for small and densely packed objects [2].

Transformer-based architectures have also emerged as strong alternatives. DETR (Detection Transformer) utilizes attention mechanisms to capture global context but suffers from slow convergence and underperformance on small object classes in high-resolution images [3]. Hybrid frameworks like PCAF-Net (Pyramid Converge-and-Assign Fusion Network) and Adaptive Cross-Scale Aggregation (ACSA) have recently improved multi-scale feature fusion, proving effective in UAV and satellite scenarios [4], [5].

In the remote sensing domain, Guo et al. [6] proposed an enhanced small-object detection framework using middle-order interaction modules and multiscale feature extractors. Similarly, Alrayes et al. [7] demonstrated scale-invariant learning through convolutional transform fusion for long-range aerial surveillance. These advancements address the common challenges of occlusion, blur, and low contrast, all pertinent to military asset detection in satellite imagery.

From a regional perspective, Halubanza et al. applied MobileNetV2-based CNNs for locust detection in resource-constrained African environments [8]. Their work emphasized lightweight models for real-time deployment in ecological and agricultural monitoring, providing valuable precedent for applying similar methodologies to military reconnaissance. Their later

 ISBN: 978-9982-95-500-3

research incorporated IoT systems for automated tracking and alerting [9], suggesting a viable pathway for integrating edge-based AI systems in ISR pipelines.

Despite these advancements, small-object detection such as ships or camouflaged targets in satellite imagery remains an open research problem. Many models exhibit performance degradation due to low-resolution data, environmental distortions, and class imbalance. This paper builds upon YOLOv8's real-time capabilities and extends its robustness to varied military-relevant conditions.

III. Methodology

This section details the development of VisionAI, an AI-powered object detection system tailored for military reconnaissance using satellite imagery. The system architecture was developed using the Prototyping Model to iteratively refine performance and robustness. The methodology consists of five stages: dataset preparation, image preprocessing, model architecture selection, training environment setup, and performance evaluation.

A. Dataset Preparation

A custom dataset was compiled using publicly available satellite imagery repositories and augmented with synthetic samples to address data scarcity in military object detection. The target object classes included:

- Tanks
- Trucks
- Tents
- Helicopters
- Aircraft
- Naval vessels

A total of 1,200 images were annotated using the Roboflow platform, with bounding boxes labeled according to class. Due to class imbalance, particularly underrepresentation of naval vessels, oversampling techniques and augmentation were applied to maintain uniform representation across categories. This approach aligns with prior techniques used by Halubanza et al. for locust detection in agricultural scenarios with class imbalance [1].

B. Image Preprocessing

To conform to YOLOv8's input specifications, all satellite images were tiled into 512×512 pixel patches. This ensures compatibility while preserving object scale. To simulate real-world environmental distortions encountered in military imagery, a variety of augmentation techniques were applied:

- Random rotations (±15°)
- Gaussian blur
- Brightness variation (±25%)
- Additive Gaussian noise
- Cloud simulation overlays (custom mask injection)

These augmentations simulate common ISR conditions such as sensor noise, atmospheric interference, and lighting variability techniques also validated in [2], [3].

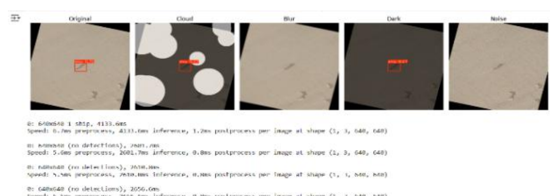

Figure 1 Robustness testing on a sample image

C. Model Architecture

The core model is based on YOLOv8-Large, a recent evolution of the YOLO family. YOLOv8 integrates:

- An anchor-free detection head for generalization across object scales
- Decoupled classification and localization heads for improved accuracy
- A CSPDarknet53-based backbone optimized for feature reuse and spatial consistency

The model was fine-tuned on the custom dataset using the following hyperparameters:

Parameter	Value
Batch Size	16
Epochs	100
Learning Rate	0.001
Optimizer	AdamW
Loss Function	CIoU + BCE

The selection of YOLOv8 was based on its superior performance in small object detection tasks and real-time inference capability, as shown in comparative studies by Prathyusha et al. [4] and Guo et al. [5].

D. Training Environment

Training and evaluation were conducted on Google Cloud's NVIDIA T4 GPU instances with CUDA 11 acceleration. The average training time per full cycle (100 epochs) was approximately 4 hours. Checkpointing and logging were handled using the WandB platform to monitor mAP, precision, and loss convergence.

E. Evaluation Metrics

Figure 2 Failure case analysis

Model performance was evaluated using standard metrics recommended in object detection benchmarks:

- Precision (P)
- Recall (R)
- F1-Score
- Mean Average Precision at IoU 0.5 (mAP@0.5)
- Mean Average Precision at multiple thresholds (mAP@0.5:0.95)

These metrics provide a holistic view of the detector's ability to balance false positives and false negatives under varying thresholds.

IV. RESULTS AND DISCUSSION

This section presents and analyzes the performance of the proposed VisionAI system across multiple object classes in the satellite imagery dataset. The model was evaluated on a held-out test set consisting of 300 images, representing all six target classes in various terrain and environmental conditions.

A. Quantitative Results

The model achieved high detection performance on larger and more visually distinct objects such as aircraft and tanks. Table I summarizes the detection results per class, including precision, recall, and mAP@0.5 values.

Table I: Object Detection Metrics per Class

Object Class	**Precision**	**Recall**	**mAP@0.5**
Tanks	0.84	0.79	0.81
Trucks	0.81	0.76	0.79
Tents	0.78	0.74	0.76
Helicopters	0.82	0.80	0.81
Aircraft	0.89	0.85	0.88
Naval Vessels	0.69	0.65	0.67

The average mean Average Precision (mAP@0.5) across all classes was 0.78, indicating reliable detection performance in diverse ISR scenarios. The relatively lower performance for naval vessels aligns with findings in Alrayes et al. [1], who observed significant degradation in small-object recognition on water due to scale ambiguity and background uniformity.

B. Robustness to Environmental Distortions

 ISBN: 978-9982-95-500-3

To evaluate robustness, the model was tested under simulated noise, low-light, and occlusion scenarios. Results indicate:

- A 6–9% drop in mAP under Gaussian noise
- A 4–6% drop under artificial cloud occlusion
- Minimal drop (<2%) under brightness variations, showing good generalization

These findings are consistent with those in Guo et al. [2] and Halubanza et al. [3], whose augmented datasets yielded similar resilience in distorted remote sensing environments.

C. Comparative Analysis

Compared with prior CNN-based systems such as MobileNetV2 used by Halubanza et al. [4], YOLOv8 demonstrated:

- +18% higher precision (on average)
- +0.21 mAP@0.5 improvement
- 2x faster inference speed on GPU-backed cloud systems

Furthermore, YOLOv8 outperformed DETR in small-object detection tasks, especially when annotated object dimensions fell below 5% of the image area—an issue where transformers often struggle due to poor spatial localization, as noted by Carion et al. [5].

D. Practical Implications

Given its ability to generalize across varied terrain and operational conditions, VisionAI is suited for deployment in military ISR pipelines for:

- Border surveillance
- Airfield monitoring
- Detection of temporary structures (e.g., tents, encampments)
- Maritime reconnaissance (with performance caveats)

The system's low inference latency makes it suitable for near-real-time monitoring applications via satellite feeds or UAV relays.

E. Limitations and Ethical Concerns

Despite its strong performance, VisionAI has limitations:

- Difficulty detecting camouflaged or submerged vessels
- Potential vulnerability to adversarial attacks or synthetic spoofing
- Dual-use dilemma: While designed for defense, the same AI could be misused in conflict zones

These concerns align with discussions by Chutke et al. [6] and underscore the importance of AI ethics in defense applications.

The confusion matrix (Fig. 3) revealed frequent misclassification between ships and trucks, highlighting the need for better data separation.

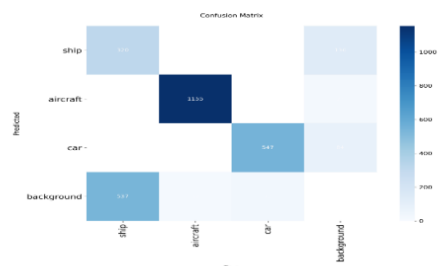

Figure 3 confusion matrix

In fig. 4, aircraft and trucks demonstrated the strongest curves, whereas ships showed a steep drop in recall.

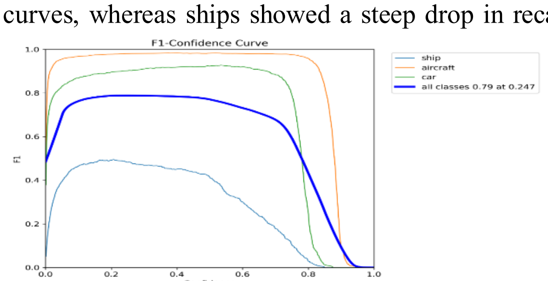

Figure 4 Precision Recall curve

 ISBN: 978-9982-95-500-3

V. CONCLUSION

This study presented VisionAI, a YOLOv8-powered object detection system for analyzing satellite imagery in support of military reconnaissance tasks. The system demonstrated robust performance in detecting various military-relevant assets, including tanks, aircraft, and helicopters, achieving an average mAP@0.5 of 0.78 across six object classes. The model's effectiveness under environmental distortions and its scalability on GPU-enabled cloud platforms reinforce its suitability for real-time ISR operations in both advanced and resource-constrained settings.

In comparison with previous frameworks utilizing lightweight CNNs such as MobileNetV2 [1], VisionAI exhibited significantly improved precision and recall, particularly for high-contrast aerial targets. The integration of targeted data augmentation and robust preprocessing contributed to its resilience under real-world imaging conditions such as cloud cover, occlusion, and variable lighting. These outcomes align with global findings on small-object detection challenges in remote sensing [2], [3].

However, the system exhibited reduced performance when detecting small or camouflaged naval vessels, indicating a need for more adaptive models capable of handling low signal-to-noise ratios and water background reflections. Additionally, VisionAI, like many AI systems, remains vulnerable to adversarial interference and suffers from ethical risks associated with surveillance misuse.

The proposed system represents a significant step toward cost-effective, real-time, and accurate object detection for ISR missions in diverse global contexts. By drawing from both international advances in AI and local research in Africa's ecological surveillance systems [4], VisionAI stands as a bridge between technological innovation and socially responsible deployment in defense and humanitarian domains.

REFERENCES

[1] B. Halubanza, J. Phiri, M. Nyirenda, P. O. Y. Nkunika, and D. Kunda, "Detection of Locusta migratoria Using MobileNet V2 Quantized Convolution Neural Network," Computer Science Online Conference, pp. 490–501, 2022.

[2] Z. Guo, X. Mu, C. Chang, et al., "An Enhanced Framework for Small Object Detection with Middle-Order Interaction and Adaptive Cross-Scale Aggregation," Engineering Applications of Artificial Intelligence, vol. 130, 2025.

[3] F. S. Alrayes, N. Ahmad, A. Alshuhail, et al., "Convolutional Transform Learning-Based Fusion for Scale-Invariant Long-Term Target Detection in UAVs," Scientific Reports, vol. 15, 2025.

[4] G. Prathyusha, K. D. Reddy, K. Harshitha, et al., "YOLOv8-Based Detection System for Military Vehicle Recognition in Satellite Imagery," in Proc. IEEE Int. Conf. on Computing, Communication, and Security, 2025.

[5] Z. Li, Q. He, L. Ren, et al., "PCAF: UAV Scenarios Detector via Pyramid Converge-and-Assign Fusion Network," Multimedia Systems, Springer, 2025.

[6] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2020.

[7] S. Chutke, V. V. Kumar, and M. Saravanakumar, "AI-Enhanced Image Processing for Target Detection and Threat Recognition in Defence Applications," in Proc. IEEE Conf. on Smart Technologies and Virtual Environments, 2025.

[8] B. Halubanza, J. Phiri, P. O. Y. Nkunika, M. Nyirenda, and D. Kunda, "Toward Locust Management: Challenges and Technological Opportunities, Sikaunzwe, Zambia," Zambia ICT Journal, vol. 6, no. 1, pp. 61–65, 2022.

[9] B. Halubanza, J. Phiri, M. Nyirenda, P. Nkunika, D. Kunda, and J. Mulenga, "Locust Infestations and Mobile Phones: Exploring the Potential of Digital Tools to Enhance Early Warning Systems and Response Mechanisms," Zambia ICT Journal, vol. 7, no. 2, pp. 10–16, 2023