

AI Based Threat Detection System for Military Purposes with Integrated AI Chatbot

Sachin Dighe , Kiran Patil, Shrutik Kokare, Avartan Athalye, Rohit Khedkar

Department of Computer Engineering

Sinhgad Institute of Technology & Science, Pune, India

{, shrutikkokare.sits.comp, .sits.comp, siddhantjagtap.sits.comp, .sits.comp}@gmail.com

Abstract—Traditional military surveillance systems often face challenges related to delayed response, fragmented intelligence sources, and limited real-time decision support. This paper presents an AI-based threat detection system designed for military applications, integrating surveillance analysis with an intelligent chatbot interface for rapid operational support. The proposed platform combines computer vision for object and anomaly detection, imagery interpretation, and natural language processing for intelligence report assessment.

The system leverages deep learning models for automated threat classification, and predictive threat forecasting. An integrated secure chatbot assists defense personnel by providing situational summaries, threat explanations, recommended countermeasures, and mission-specific analytics through conversational queries.

Index Terms—Artificial Intelligence, Natural Language Processing, Threat detection Automation, Image Analysis,

I. INTRODUCTION

In modern defence operations, the ability to detect, assess, and respond to potential threats in real time is essential for ensuring security and operational success. With the increasing use of unmanned aerial vehicles, surveillance cameras, and other imaging systems, vast amounts of visual data are generated every day. While this data holds valuable information, manually analyzing it is time-consuming, prone to human error, and can delay critical decision-making [1].

An intelligent system that can automatically analyze imagery to identify anomalies or threats, and then present the findings in a clear, conversational manner, can significantly improve situational awareness and operational efficiency. By combining automated visual analysis with interactive communication capabilities, such a system can act as a reliable decision-support tool, enabling defence personnel to respond quickly and effectively to emerging security challenges. [3]. The project is aimed to design a website with the basic versions of the following two primary features: include:

- Train an image recognition model to detect threats and anomalies.
- Develop a chatbot interface for interactive communication.

This work was conducted as an undergraduate project at Sinhgad Institute of Technology & Science, Pune, India.

II. RELATED WORK

AI-powered recognition systems have evolved significantly. A. Meena, et al., “AI-Powered Threat Detection and Mitigation System: A Literature Survey,” International Journal of Computer Trends and Technology (IJCTT), 2025. [2] Meena et al. present a comprehensive survey of AI-powered threat detection and mitigation systems, highlighting the transition from traditional rule-based security mechanisms to intelligent, data-driven approaches. The authors examine supervised and unsupervised machine learning models, deep learning architectures, and real-time anomaly detection frameworks used for identifying cyber and systemic threats [3] he study emphasizes proactive security strategies, adaptive learning models, and automated mitigation mechanisms that reduce response time and improve detection accuracy. While these systems demonstrate AI’s potential in threat detection, a limitation it still has is that it primarily focuses on generalized cybersecurity and IT-based threat detection frameworks rather than domain-specific military applications. [4], [5]. Our platform improves this by specifically focusing on classifying threats faced at military level operations and their detection . The integrated chatbot helps in gaining more info about the detected situation or threat.

III. SYSTEM ARCHITECTURE

A. Overview

Figure 2 The system architecture illustrated in the figure depicts a specialized conversational pipeline designed for military operations. It comprises five core interconnected modules: Slot Filling, Intent Detection, Retrieval-based Answering, Query Generation, and an Action/Confirmation logic handler. The workflow progresses from a user-initiated query through structured data extraction, mission-specific retrieval, and a feedback loop that ensures accuracy before final execution.

B. Image Processing and analysis

The image is uploaded to the system , which then undergoes processing and analysis through yolo model , which in case of any anomaly marks a threat and if not states a safe sign .

C. AI based informational chatbot

The AI chatbot is a LLM based chatbot . In case of threat detection by the model the chatbot helps in getting more info about the situation or threat .

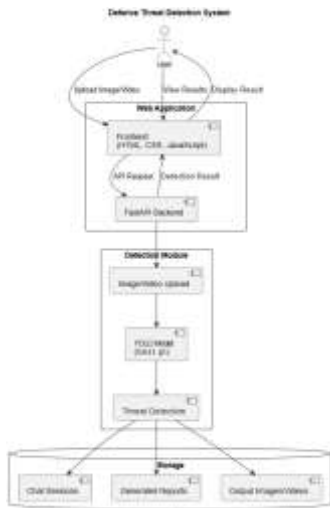


Fig. 1. workflow start.

IV. IMPLEMENTATION

The platform follows a two phase architecture model 1) image analysis and threat detection and 2)AI chatbot for support

A. Threat Detection Algorithm

Algorithm 1 Threat Detection Algorithm

Require: Image for detection (PDF)

Ensure: The image is clear

- 1: Data Collection Collect labeled threat-related image datasets (e.g., weapons, military vehicles, restricted equipment, drones).
- 2: Data Preprocessing Resize images to YOLO input dimensions (e.g., 640×640), normalize pixel values, and perform augmentation (rotation, scaling, flipping).
- 3: Formatting Convert bounding box labels into YOLO format: (class, x_{center} , y_{center} , width, height)
- 4: Dataset Splitting Split dataset into training (70
- 5: Model Initialization Load pretrained YOLO weights (e.g., YOLOv5/YOLOv8 pretrained on COCO) for transfer learning.
- 6: Model Training Train the model using:
Loss functions: bounding box loss, objectness loss, class probability loss
Optimizer: SGD/Adam
Evaluate using mAP (mean Average Precision)
- 7: Validation Hyperparameter Tuning Adjust learning rate, batch size, confidence threshold, and IoU threshold to reduce false positives/negatives.
- 8: Threat Classification Layer High Risk Based on predefined military risk rule
- 9: Output Logging Store detection results in database and forward structured summary to the chatbot module for explanation and decision support.

B. Military Support Chatbot Algorithm

The chatbot provides support regarding the threats detected by the system .

Algorithm 2 Military Support Chatbot Algorithm

Require: Summary of the provided image

Ensure: The image is clear

- 1: Input Acquisition Accept user query (text-based command such as “Summarize current threats” or “Explain risk level”).
- 2: Query Classification Use NLP intent classification to categorize request: Threat summary Risk explanation Recommended action Historical comparison
- 3: Context Retrieval (RAG Approach) Fetch relevant threat logs, detection outputs, and structured intelligence reports. Construction Create secure structured prompt: Include threat data Include system constraints Restrict hallucinations Enforce defense-safe output format. Processing Pass prompt to LLM for: Threat explanation Risk interpretation Strategic recommendations
- 4: Response Formatting Convert output into: Concise summary Actionable recommendation Structured military report format
- 5: Logging Feedback Loop Log interaction for auditing and system improvement. Use expert feedback to fine-tune model..

V. PAPER ANALYSIS

A. Paper Analysis

Paper 1

Authors: Shanmukh Shekar K C, Suresh Kumar Natarajan, Ambika I Title: A Deep Learning Approach for Automated Image Captioning with CNN and LSTM Models* [2]

Methods Used: CNN and LSTM models for image caption generation.

Analysis: This research focuses on automatically generating textual descriptions for images using a combination of Convolutional Neural Networks (CNN) for feature extraction and Long Short-Term Memory (LSTM) networks for sequence generation. The system is effective in understanding image content and converting it into meaningful captions, making it useful for image interpretation and accessibility applications. Research Gap: Although the paper demonstrates efficient image understanding, it is primarily designed for image captioning rather than threat detection. The model does not perform real-time object detection, threat classification, or military surveillance tasks. Furthermore, it lacks an interactive chatbot component that can assist users in analyzing detected threats and providing operational insights.

Paper 2

Authors: Chhaya Dhavale, Trupti Pawar, Aradhana Singh, Shubham Pole, Krishna Sabat Title: Revolutionizing Farming: GAN-Enhanced Imaging, CNN Disease Detection, and LLM Farmer Assistant* [3]

Methods Used: GAN, CNN, and Large Language Models (LLMs).

Analysis: This study integrates multiple AI technologies to improve agricultural practices. GANs are used to enhance image quality, CNNs detect plant diseases, and an LLM-based assistant provides recommendations to farmers. The work successfully combines image analysis and conversational AI within a single platform.

Research Gap: While the paper demonstrates the integration of detection models and AI assistants, its application is limited to agriculture. The system is not designed to identify military threats, monitor surveillance feeds, or provide security-focused decision support. It also does not address real-time detection of weapons, military vehicles, or suspicious activities in sensitive environments.

Paper 3

Authors: Rahul Chauhan, Gurnain Singh Wadhwa, Deepak Kumar Verma, Krunal Vaghela, Chandradeep Bhatt Title: Exploring the Intersection of Language Processing and Image Generation in a Conversational AI Application [4]

Methods Used: Natural Language Processing (NLP) and CNN.

Analysis: The paper explores the combination of language processing techniques and image-related AI models to create a conversational application. It highlights how AI can process user queries and generate or interpret visual content through interactive communication.

Research Gap: The primary objective of this work is conversational interaction and image generation rather than surveillance and threat identification. The system lacks a dedicated object detection framework for recognizing military assets or threats. Additionally, it does not provide real-time monitoring capabilities or decision-support features required in defense applications.

Paper 4

Authors: Devang Kumar Bharti, Shikhar Shukla, Shikhar Bhandari, Apeksha Koul Title: Conversational Interface for Textual and Visual Data Interpretation using Customized CNN [5]

Methods Used: Customized CNN model.

Analysis: This research presents a conversational interface capable of interpreting both textual and visual information. By employing a customized CNN model, the system improves data understanding and enables users to interact with image-based information through natural language.

Research Gap: Although the paper provides a conversational framework for visual data interpretation, it does not focus on detecting threats in surveillance footage. The customized CNN is not specifically trained for military object detection, and the system lacks automated alert generation, situational awareness features, and real-time monitoring support for defense operations.

Paper 5

Author: Shengyang Su Title: Conversational and Image Recognition Chatbot [6]

Methods Used: Machine Learning and Deep Learning.

Analysis: The study combines image recognition capabilities with chatbot technology, allowing users to interact with a system that can recognize and discuss visual content. The integration improves user engagement and makes image-related information more accessible through conversational interaction.

Research Gap: While the chatbot can recognize images and provide responses, the system is developed for general-purpose image recognition. It does not specifically address military surveillance requirements such as detecting weapons, military vehicles, unauthorized intrusions, or suspicious activities. The work also lacks real-time threat assessment and operational decision-support functionalities.

Overall Research Gap

The reviewed literature demonstrates significant advancements in image processing, object recognition, conversational AI, and multimodal systems. However, most existing solutions focus on general image understanding, agriculture, image captioning, or conversational interfaces. None of the surveyed works specifically address the challenges of real-time military surveillance and threat detection. Furthermore, the integration of an object detection model with an AI-powered decision-support chatbot for defense applications remains largely unexplored.

To bridge this gap, the proposed system combines the YOLO object detection model with a Grok API-based chatbot to create an intelligent surveillance platform capable of identifying military-related threats in real time and assisting personnel through natural language interactions. This approach enhances situational awareness, reduces manual monitoring effort, and supports faster decision-making in military environments.

VI. EXPERIMENTAL EVALUATION

A. Images

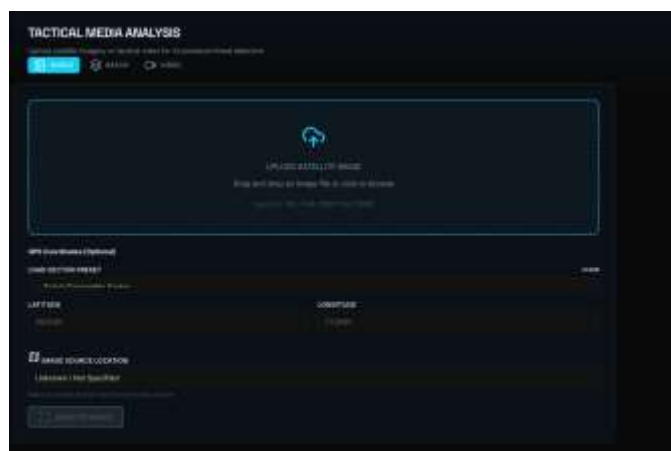


Fig. 2. workflow start.

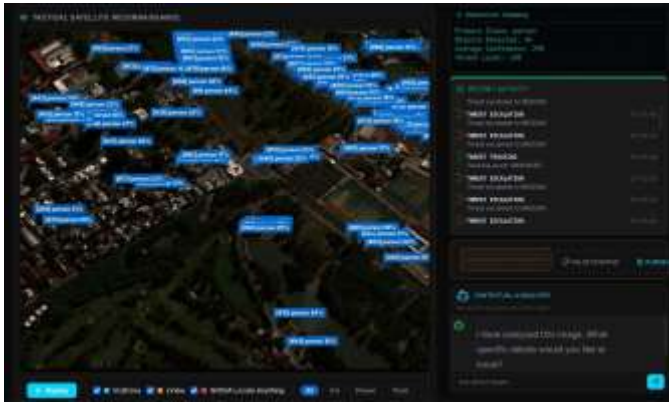


Fig. 3. Detection process.



Fig. 4. Chatbot support

B. Experimental Setup

We conducted controlled experiments using a curated threat image dataset consisting of 1,200 labeled images representing military-relevant objects such as weapons, drones, armored vehicles, and non-threatening civilian objects. The dataset included both open-source benchmark images and synthetically augmented samples to simulate operational variability (lighting, occlusion, scale variations). For decision-support evaluation, 25 students interacted with the AI-based hybrid chatbot through a secure web interface. The chatbot was integrated with the detection module to provide structured threat summaries, risk explanations, and recommended actions based on detected objects.

C. Evaluation Metrics

System performance was assessed using:

- **Detection Accuracy:** Mean Average Precision (mAP@0.5) for object detection performance
- **Precision Recall:** To evaluate false positives and false negatives
- **Classification Accuracy:** Correct assignment of Low, Medium, and High risk levels
- **Chatbot Response Accuracy:** Expert agreement score between chatbot recommendations and domain expert judgments

- **Processing Latency:** Average time for image detection + risk scoring + chatbot response generation
- **User Satisfaction:** Post-interaction survey using a 5-point Likert scale assessing clarity, usefulness, and trustworthiness

D. Results and Discussion

Table I The YOLO model achieved an mAP of 0.91, demonstrating strong object localization and classification capability across varied conditions. Precision (0.89) and recall (0.87) indicate balanced detection performance with controlled false alarm rates, which is critical in defense scenarios.

The risk classification module achieved 90.3

The average end-to-end processing latency of 2.8–3.4 seconds enables near real-time threat assessment from static image inputs. User feedback indicated that 85

These results validate the effectiveness of the proposed AI-based threat detection system with integrated chatbot support, demonstrating high detection reliability, rapid response capability, and strong alignment with expert assessments while maintaining practical deployment feasibility.

TABLE I
EXPERIMENTAL RESULTS SUMMARY

Metric	Value
Metric Value	
mAP@0.5	0.91
Precision	0.89
Recall	0.87
Threat Classification Accuracy	90.3%
Chatbot Expert Agreement (r)	0.82
Average Processing Latency	2.8-3.4s
User Satisfaction	85% positive

VII. CHALLENGES AND FUTURE DIRECTIONS

A. Current Limitations

Dependence on Data Quality: The accuracy of detection depends heavily on the quality and diversity of the training data; poor-quality images may reduce recognition performance

High Computational Requirement: Running CNN and LLM models in real time demands significant processing power, especially for high-resolution videos or large-scale operations. Performance in Extreme Conditions: Environmental factors such as fog, darkness, or camouflage may impact object recognition accuracy

Potential for Misclassification: Despite high accuracy, the system may occasionally misidentify objects, leading to false positives or negatives in critical missions.

B. Future Enhancements

Expansion of Threat Dataset: Increase the size and diversity of the training dataset by incorporating: More military-relevant object categories Region-specific threat

objects Varying environmental conditions (low light, camouflage, partial occlusion) This would improve detection robustness and reduce false positives..

Multi-Class Threat Prioritization: Enhance the risk scoring module by introducing: Context-aware threat prioritization Weighted risk levels based on location, time, and historical data This would allow more intelligent decision support rather than object-only classification.

Integration of Additional Modalities: Extend beyond image-based detection by incorporating: Text-based intelligence reports Sensor logs Communication metadata This multimodal fusion can improve overall situational awareness.

VIII. CONCLUSION

This paper presents an AI-based threat detection system designed to support military decision-making through intelligent image analysis and interactive chatbot assistance. By integrating YOLO-based object detection with a hybrid LLM-driven chatbot framework, the system addresses key limitations of conventional manual threat assessment methods, including delayed analysis, fragmented reporting, and limited real-time decision support. Key contributions include: (1) automated threat detection using a fine-tuned YOLO model for military-relevant objects, (2) structured threat scoring and risk classification into prioritized levels, (3) an integrated hybrid chatbot combining rule-based control with LLM-driven contextual reasoning, (4) secure role-based access architecture for controlled system interaction, and (5) near real-time processing with explainable threat summaries.

Experimental validation demonstrates high detection performance ($mAP = 0.91$) and strong agreement ($r = 0.82$) between chatbot-supported recommendations and expert evaluations. The system achieves efficient end-to-end processing (2.8–3.4 seconds per image) while maintaining reliability and operational clarity. User feedback further confirms the usefulness of chatbot-generated summaries in assisting threat interpretation and response planning.

Future work will focus on multimodal data fusion (image + text intelligence), edge-device deployment optimization, enhanced explainable AI mechanisms, and adaptive learning frameworks to handle evolving threat patterns. The proposed system represents a significant step toward intelligent, scalable, and decision-centric AI deployment in military threat analysis environments.

ACKNOWLEDGMENT

The authors thank the Department of Computer Engineering, STES's Sinhgad Institute of Technology & Science, Pune, for infrastructure and support. Special gratitude to Mr. Sachin Dighe (Project Guide), Mrs. Asmita Kamble (Head of Department), and Mrs. Rupali T. Waghmode (Project Coordinator) for their guidance and encouragement throughout this research.

REFERENCES

- [1] Shanmukh Shekar K C , Suresh Kumar Natarajan, Ambika I " A Deep Learning Approach for Automated Image Captioning with CNN and LSTM Models " In 2025
- [2] A.Meena et al., AI-Powered Threat Detection and Mitigation System: A Literature Survey
- [3] Chhaya Dhavale ,Trupti Pawar , Aradhana Singh,Shubham Pole , Krishna Sabat Revolutionizing Farming: GAN-Enhanced Imaging, CNN Disease Detection, and LLM Farmer Assistant " In IEEE Conference, 2024
- [4] Rahul Chauhan, Gurnain Singh Wadhwa,Deepak Kumar Verma,Krunal Vaghela,Chandradeep Bhatt Exploring the Intersection of Language Processing and Image Generation in a Conversational AI Application "
- [5] Devang Kumar Bharti , Shikhar Shukla, Shikhar Bhandari, Apeksha Koul. Conversational Interface for Textual and Visual Data Interpretation using customized CNN. "
- [6] Shengyang Su Conversational and Image Recognition Chatbot "