

Explainable Artificial Intelligence framework for Robust Object Detection in Tactical Unmanned Vehicles (XAI-UAV)

Under the Guidance of

Dr. Deepasikha Mishra

Review - 1



VIT-AP
UNIVERSITY

Proposal Report

Capstone Project Proposal Report (Individual Report)

Guide Approval (initials/date):	<i>Deepasikha Mishra</i>	26/8/25
---------------------------------	--------------------------	---------

CAP4001– Capstone Project Proposal Report

Student Name	Konakalla Rishitha		
Student Register Number	22BCE7331		
Programme	B.Tech, Computer Science and Engineering - AIML		
Semester/Year	FallSem 2025-26		
Guide(s)	Dr. Deepasikha Mishra		
Project Title	XAI-UAV: Explainable Artificial Intelligence Framework for Robust Object Detection in Tactical Unmanned Aerial Vehicles		
Team Composition:			
Reg. No	Name	Major	Specialization
22BCE7331	Konakalla Rishitha	CSE	AIML
22BCE9807	Kana Hakshay Reddy	CSE	AIML
22BCE7558	Appala Pranav Sai	CSE	AIML

Motivation

- In high-stakes UAV applications like surveillance & defense, AI must not only be fast, it must be explainable
- Current object detection models used in drones (e.g., YOLO) are accurate, but behave like black boxes
- Lack of transparency reduces trust, limits adoption, and poses risks in human-AI decision-making
- We are motivated to bridge this gap by building a real-time, interpretable AI system for UAVs
- Our solution merges model optimization (pruning) with visual explanations (Grad-CAM) — ensuring both performance and accountability

Objectives

- Identify and address the lack of explainability in UAV-based object detection systems
- Optimize deep learning models (YOLOv5/YOLOv8) for real-time inference on edge devices (e.g., Jetson, FPGA)
- Apply pruning to reduce model size while preserving performance and interpretability
- Integrate explainability tools (Grad-CAM, Saliency Maps) into the UAV inference pipeline
- Build a ground station interface that visualizes detections alongside their explanations
- Evaluate detection accuracy (mAP, F1)

Hypothesis

- We hypothesize pruning can produce a lightweight YOLO model that is both accurate and interpretable
- We propose that integrating XAI methods (Grad-CAM, Saliency Maps) with the deployed model will enable real-time visual explanations onboard UAVs
- We believe this combination will allow UAV operators to understand and trust AI decisions without compromising performance
- Our pipeline will maintain detection accuracy while also delivering explanation fidelity and stability, bridging the gap between efficiency and transparency

Problem Survey

State of the Art

- UAV object detection: *VisDrone*, *UAVDT* benchmarks; YOLOv5/YOLOv8 widely adopted for aerial vision.
- Edge AI deployments: *Jetson Nano/Orin* enable onboard inference but limited by resource constraints.
- XAI methods (Grad-CAM, Saliency Maps) are established in medical and NLP domains, rarely extended to UAV detection.

Limitations Identified

- Existing UAV research optimizes models for speed but overlooks interpretability.
- No standardized inclusion of fidelity/stability metrics alongside accuracy.
- Edge pruning is often magnitude-based, risks accuracy drop and ignores explainability.

Premise of Proposal

- Introduce pruning: interpretable model compression.
- Integrate real-time explainability engine in UAV object detection pipeline.
- Deliver trustworthy, operator-validated UAV AI that balances accuracy, efficiency, and transparency.

Subject Knowledge

Object Detection Models

- YOLO family: single-shot, real-time detection
- Trained on UAV-specific datasets (VisDrone, UAVDT)
- Balancing accuracy (mAP) vs. efficiency (FPS, latency) is critical for UAV hardware

Optimization

- **Pruning:**
 - Uses BN values to rank neuron importance
 - Retains interpretability while reducing parameters
- **Deployment Workflow:**
 - PyTorch (training) → ONNX (intermediate format) → TensorRT (.engine model)
 - Runs efficiently on Jetson/FPGA hardware

Subject Knowledge

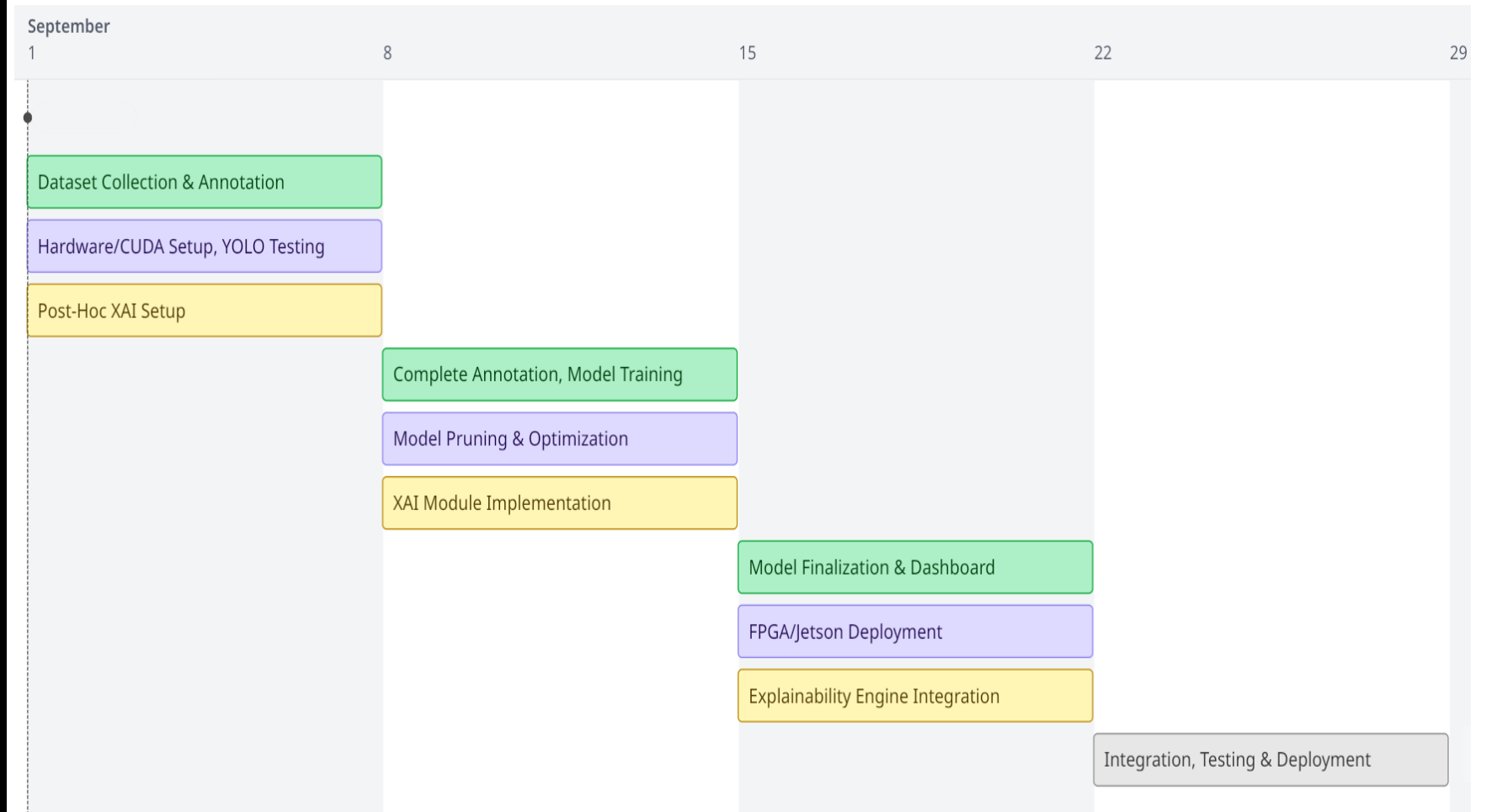
Evaluation Framework

- **Accuracy:** mAP, F1-score
- **Efficiency:** Latency per frame

Explainability Methods

- **Grad-CAM:** Highlights *which regions in feature maps* drive detections
- **Saliency Maps:** Pixel-level sensitivity showing *why detection confidence changes*
- **SHAP:** Quantifies *numerical importance* of image regions to predictions

Time Plan



Thank you!

Team members:

Konakalla Rishitha 22BCE7331

Kana Hakshay Reddy 22BCE9807

Appala Pranav Sai 22BCE7558