Human Condition Monitoring Using Unmanned Aerial System (UAS)

Dr. Adeel Khalid, Kennesaw State University

Adeel Khalid, Ph.D. Professor Industrial and Systems Engineering Office: 470-578-7241

Mr. Owais Ahmed, Kennesaw State University

As a doctoral researcher in the AERO Lab, I was responsible for the end-to-end development of the computer vision pipeline and the technical implementation of the project. I also played a key role in mentoring undergraduate students, equipping them with practical skills to develop and train computer vision models for real-world problem solving.

Human Condition Monitoring Using Unmanned Aerial System (UAS)

ABSTRACT

This research explores the use of Unmanned Aerial Systems (UASs) as a means to engage an undergraduate student in a hands-on research project within a university research laboratory. While many early-stage undergraduate students express interest in research, they often lack a clear understanding of what it entails. To bridge this gap, a structured mentorship model pairs an undergraduate student with a graduate researcher on a project that involves the integration of thermal imaging sensors mounted on a UAS platform for data collection and analysis.

The project develops a real-time thermal imaging system to assess human conditions using live UAS footage. The system distinguishes between mobile (upright) and immobile (non-upright) individuals by analyzing thermal video streams transmitted directly from the UAS to a ground station. A deep learning model, based on You Only Look Once version 12 (YOLOv12), is trained on a custom thermal dataset featuring a variety of human postures and environmental conditions. This solution offers a reliable and privacy-conscious approach for applications in safety monitoring, search-and-rescue missions, and emergency response scenarios.

This paper details the project, its technical contributions, and highlights student learning outcomes, mentorship practices, and opportunities for improvement.

Keywords: Artificial Intelligence (AI), Unmanned Aerial Systems (UAS), Machine Learning (ML), Thermal Imaging, Infrared Imaging, Object Detection, YOLO, Undergraduate Research, Engineering Education, Project-Based Learning

INTRODUCTION

The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) has significantly transformed the way autonomous systems perceive and interpret their surroundings. Within university research laboratories, student-driven projects have explored the integration of AI and ML with Unmanned Aerial Systems (UAS) equipped with thermal imaging sensors to address real-world problems. These efforts aim to improve environmental awareness, safety monitoring, and inspection efficiency through data-driven automation.

Unmanned Aerial System (UAS) technology has seen exponential growth in recent years due to its versatility, accessibility, and integration of advanced sensors and onboard computing. Their ability to autonomously perform complex missions such as surveillance, reconnaissance, and search-and-rescue (SAR) has revolutionized the aviation and robotics industries[1]. However, most systems still rely on human operators to monitor live video streams to identify critical events such as human falls or injuries.

With the growing advancement in AI and ML, these tasks can now be automated using deep learning models capable of real-time inference. Integrating these models with infrared imaging enhances usability in dark or occluded conditions while preserving human privacy. Thermal vision enables UASs to detect human presence and condition even in low-light or visually obstructed environments, which is especially valuable for emergency response and rescue operations.

Figure 1: Field testing setup showing the UAS and student team during dataset collection and model deployment.

Several recent studies have explored the use of AI-enabled UASs for object and human detection, identification, and position determination. Ramachandran et al.[2] provides a comprehensive survey of object detection and tracking methods using UASs, comparing deep learning and traditional image processing approaches and cataloging their relative strengths and failure modes—a taxonomy that informs our selection of detection architecture. Jain et al.[3] complements this survey by identifying the specific challenges of aerial image datasets, including scale variation, viewpoint distortion, and small object prevalence, which directly guided our data augmentation and training strategy. Lin et al.[4] developed a fall detection system using AI-based edge computing, achieving 96% accuracy and a processing speed of 11.5 frames per second (FPS). Their approach, YOLO-LW (a lightweight modification of YOLOv3-tiny), efficiently detected human falls and transmitted alerts over Wi-Fi without sharing video data. Our framework builds on Lin et al.'s edge-based alert mechanism by extending detection from RGB to thermal imagery and relocating inference from onboard edge hardware to a ground station with greater computational capacity, enabling use of the more expressive YOLOv12 architecture. Similarly, Chang et al.[5] implemented YOLOv4-tiny with OpenPose for human posture detection and facial recognition, enabling a UAS to autonomously detect falls and deliver real-time notifications to users. Our work extends this posture-detection paradigm by replacing RGB-based facial recognition—which raises privacy concerns—with a purely thermal pipeline where keypoint geometry alone determines condition state. Sonkar et al.[6] used a low-altitude long-endurance

(LALE) fixed wing vertical takeoff and landing (VTOL) UAS with a Raspberry Pi to transfer data from the aircraft to a ground station for annotation. The objects detected using this method were inanimate objects like plants, burned areas, helipads, and under construction areas.

In the context of SAR, UASs have been used extensively. Papyan et al.[7] utilized UAS-based systems for the detection of individuals, particularly by human scream and other distinct human distress signals. They elaborate on the unique challenges they faced with locating humans through aerial acoustics and highlighted one solution using AI to analyze sound frequencies. They note that deep learning-based networks, such as Convolutional Neural Networks (CNNs), can be trained to filter and identify noise and estimate Direction of Arrival (DOA). Weber et al.[8] implemented a UAS employed with Pixhawk autopilot, a modular electronics payload, and infrared and RGB sensors to enable both thermal-based rescue and agricultural imaging. They proved the human detection capability within real-world test zones and demonstrated a low-cost UAS system for civil applications. Bhandari et al.[9] developed an autonomous fixed-wing UAS system designed for SAR operations. Object recognition for target detection used OpenCV (C++) algorithms for color, shape, and edge segmentation, supported by a neural network (NN)-based feature extractor. Hayner et al.[10] also proposed the solution of a small UAS (sUAS) with computer vision algorithms for object detection in the context of real-time human detection. They combined a versatile suite of object detection algorithms and aerial data from both the visual and radiometric spectrum to have the ability to analyze and process real-time data faster than existing systems at greater or comparable accuracy. Harrington et al.[11] developed a UAS swarm modeling and simulation framework for monitoring migration flows and SAR operations during the Mediterranean refugee crisis, incorporating operational variables including altitude variation, wind effects, and multi-vehicle coordination. They found that UAS swarms drastically improve situational awareness, detection rates, and rescue coordination. Their framework also serves as a decision-support tool for optimizing resource allocation, swarm configuration, and future UAS acquisition strategies. While Harrington et al. account for a broader set of environmental variables (e.g., altitude variation, wind disturbance) than the present work, our study complements their macro-level analysis by providing the micro-level perception capability—thermal human detection and condition classification—that such swarm frameworks require at the individual agent level.

Pertaining to both YOLO and SAR, there have been a multitude of papers written. Boudjit et al.[12] presents research progress in the development of applications for the detection and identification of persons using the CNN-based YOLOv2 architecture deployed on a UAS camera. Their system detects both the position and posture of a person through deep-learning-based computer vision in real time; however, it operates exclusively on RGB imagery and does not consider thermal modalities, which limits performance in low-visibility conditions. Rizk et al.[13] demonstrated an AI-assisted UAS system using YOLOv3 deployed on a NVIDIA Jetson Xavier NX for real-time human detection. The model, trained with over 3000 labeled images, achieved 82% accuracy across diverse environments. Jayalath et al.[14] also designed a UAS combining YOLO and Fast R-CNN for autonomous identification and localization of humans in emergency missions, achieving 84.1% true positives and showing reliability improvements with higher-resolution imagery. Caputo et al.[15] demonstrated promising human detection performance using approaches based on the YOLOv5 detection algorithm, achieving strong results on RGB aerial imagery and validating that computer vision can effectively aid SAR operations with UASs. Critically, they identify the future use of thermal imaging as a key direction for improving detection accuracy in adverse conditions. Our work directly addresses this gap identified by Caputo et al. by implementing a thermal-only detection pipeline using the more recent YOLOv12 architecture, which incorporates attention-based feature extraction absent from earlier YOLO variants. McGee et al.[16] used the YOLOv3 algorithm as the neural network and thermal image acquisition to perform object detection and classification of people in need of SAR in outdoor environments. However, they were limited by their airframe (DJI Phantom) and payload capacity, as well as their training data, and recommended implementing different environments and angles to increase the robustness of the model. Our dataset of 2,426 images captured across four thermal color palettes directly addresses this data diversity limitation.

Although these systems demonstrate impressive results, most rely on onboard processing or static datasets, limiting flexibility and scalability. In addition, few studies utilize direct thermal data streaming from a UAS to an external ground station for inference. This approach allows the use of more powerful computing resources while maintaining real-time responsiveness. A general overview of this approach is shown in Figure 2.

Therefore, this study proposes a real-time human condition detection framework that decouples perception and reasoning. The YOLOv12 model[17,18] is fine-tuned solely for person detection and 2D pose estimation using thermal imagery collected from a UAS. The detected keypoints are then processed by a geometric posture analysis module, which interprets joint alignment and body orientation to determine whether an individual is “Mobile” (upright and capable of self-directed movement) or “Immobile” (non-upright, indicating a person who may be incapacitated, injured, or otherwise unable to move). This classification is intentionally agnostic to the cause of immobility; a person lying down for rest and an unconscious individual both present the same non-upright thermal signature. The distinction between these states is left to first responders who can use the system’s localization output for further assessment. The resulting system provides a privacy-preserving, interpretable, and efficient framework for real-time search-and-rescue and safety-monitoring missions using unmanned aerial systems.

This study also serves dual purposes: advancing technical capabilities while providing meaningful research experiences for undergraduate students.

EDUCATIONAL FRAMEWORK AND STUDENT INVOLVEMENT

Motivation for Undergraduate Research Engagement

Many early-stage undergraduate students express interest in research but often lack understanding of what research entails. Traditional coursework provides foundational knowledge, but hands-on research experience helps students develop critical thinking, problem-solving skills, and professional competencies that are difficult to cultivate in classroom settings alone. The literature on undergraduate research experiences (UREs) consistently demonstrates that participation in structured research activities enhances technical competency, increases self-efficacy in STEM disciplines, and strengthens career decision-making[19]. Hsieh[20] similarly reports that undergraduate engineering students who participate in mentored research programs show higher rates of graduate school enrollment and co-authored publications.

The project described in this paper was specifically designed to engage an undergraduate student with limited prior research experience while still producing a meaningful technical outcome. It is important to note that the project is structured and scaffolded: the student works within a well-defined problem boundary with a known methodological approach, rather than engaging in fully open-ended research where the problem formulation itself is part of the inquiry. This design choice is intentional—the goal is to provide an accessible entry point that introduces the *practices* of research (data collection, iterative experimentation, technical documentation) without the ambiguity that can overwhelm a novice researcher.

Several institutions have employed UAS platforms as vehicles for undergraduate research engagement in engineering. The University of Cincinnati's UAV MASTER Lab[21] integrates undergraduates into UAS design, analysis, and experimentation projects across aerospace and mechanical engineering disciplines. Similarly, drone-based project-based learning (PBL) curricula have been shown to promote engagement and self-efficacy among underrepresented students in STEM[22]. These precedents support the mentorship model adopted in the present work.

Project Structure and Mentorship Model

The project team consisted of one graduate research assistant (the lead author), who served as the primary technical developer and mentor, and one undergraduate student from mechanical engineering. The graduate student was responsible for the overall system architecture, model development, real-time deployment pipeline, and scholarly communication. The undergraduate student was progressively introduced to each phase of the project through a structured mentorship model.

Table 1 summarizes the task progression, the engineering concepts exercised at each phase, and the approximate time allocation.

In **Phase 1**, the undergraduate student was introduced to the project scope through a guided literature review. The graduate mentor assigned a set of foundational papers[2,4,15] and conducted one-on-one discussion sessions to build the student's understanding of the problem space. The student also received introductory training on UAS flight regulations (FAA Part 107) and safe operating procedures.

In **Phase 2**, the student performed hands-on data work: extracting individual frames from recorded thermal video and annotating bounding boxes using the Roboflow platform. The graduate mentor reviewed annotation quality and provided feedback on labeling consistency, teaching the student how data quality directly impacts downstream model performance.

In **Phase 3**, the graduate mentor introduced the fundamentals of computer vision and machine learning through structured one-on-one sessions. Topics covered included how convolutional neural networks extract spatial features, how the YOLO architecture frames detection as regression, and how training hyperparameters (batch size, learning rate, augmentation policy) affect convergence. The student then ran training experiments under supervision, learning to configure the Ultralytics API, monitor loss curves, and interpret precision-recall tradeoffs. These exercises were explicitly connected to concepts from the student's coursework in calculus

Table 1: Undergraduate student task progression: phases, engineering concepts, and time allocation.

Phase	Tasks	Concepts Exercised	Approx. Time
Phase 1: Foundations	Literature review, UAS operation basics, thermal camera theory, data capture	Sensor physics, academic writing, information synthesis	3–4 weeks
Phase 2: Data & Annotation	Frame extraction from thermal video, bounding box annotation via Roboflow, quality checking	Image representation, data curation, labeling standards	3–4 weeks
Phase 3: ML Fundamentals	Introduction to CNNs, YOLO architecture walkthrough, model training with Ultralytics API, loss curve interpretation	Gradient descent, optimization, evaluation metrics	4–5 weeks
Phase 4: Integration	Assisting with pipeline testing, documentation, conference paper preparation	Technical writing, system-level thinking, peer review	2–3 weeks

(gradient descent) and linear algebra (matrix operations).

In **Phase 4**, the student contributed to pipeline testing, project documentation, and preparation of the conference manuscript, gaining exposure to the scholarly communication process.

Learning Objectives

The project was designed to achieve several educational outcomes for the undergraduate participant:

(1) Develop practical skills in UAS operation, sensor integration, and data collection methodologies.

(2) Gain hands-on experience with machine learning workflows including data preparation, model training, and evaluation.

(3) Understand the iterative nature of research through hypothesis testing, experimentation, and refinement.

(4) Build professional skills including technical documentation, teamwork, and communication of results.

(5) Connect classroom concepts in physics, programming, and engineering to a real-world

application.

METHODOLOGY

System Overview

The proposed framework operates in two stages: first, a thermal person-detection module localizes individuals in the scene; second, a pose-based condition-estimation module analyzes the detected person to infer their physical state. An overview of this approach is shown in Figure 2.

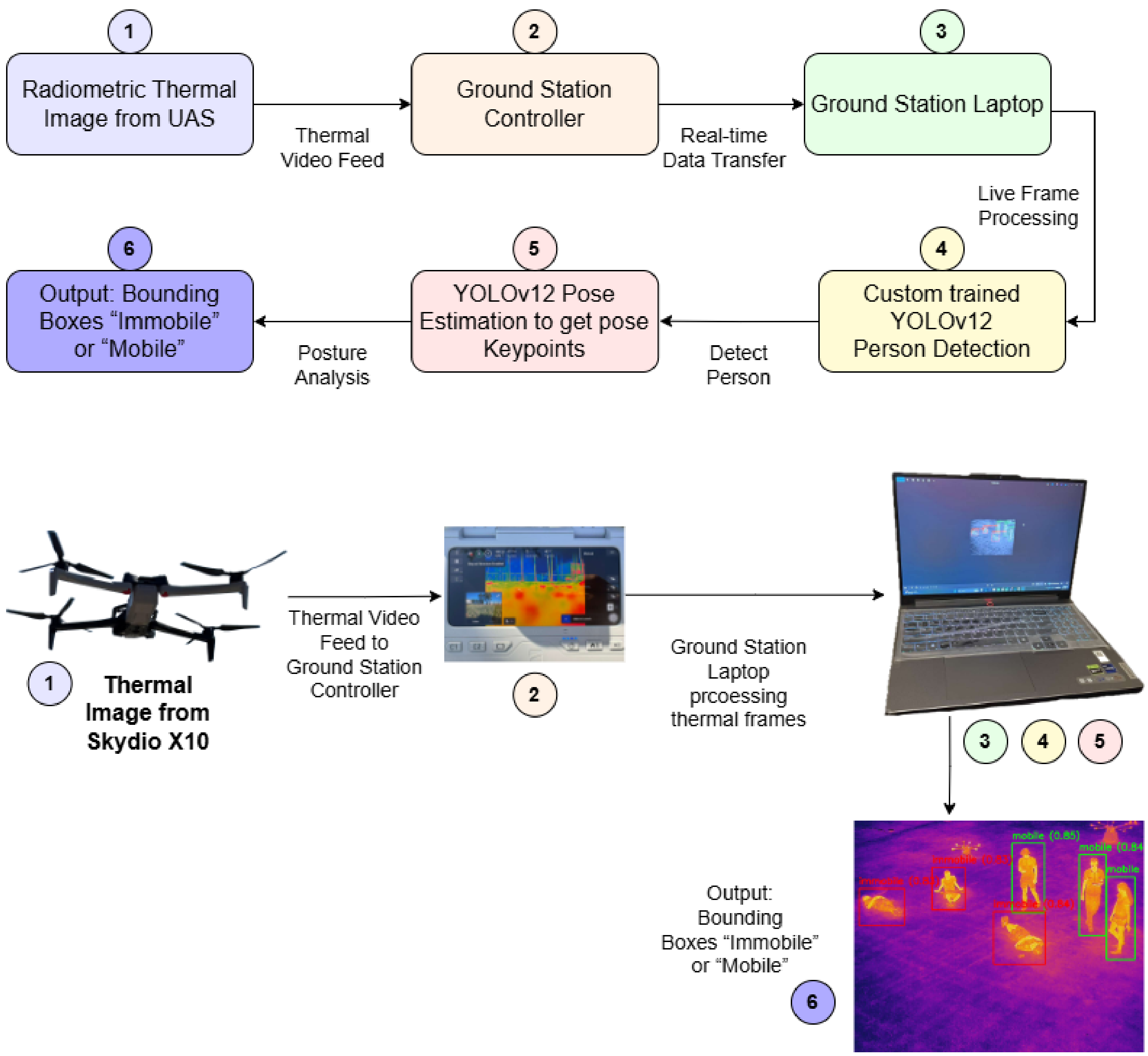


Figure 2: Overview of the proposed UAS thermal-vision pipeline showing live feed transmission from the UAS to the ground station running the deep learning model for human condition detection.

Deep Learning Model

The You Only Look Once (YOLO) family of object detectors, first introduced in 2016[23], frames detection as a single regression problem from image pixels to bounding box coordinates and class probabilities. This one-stage formulation trades the region proposal step used in two-stage detectors (e.g., Faster R-CNN) for substantially higher inference speed, a property critical for real-time aerial applications where frame throughput directly impacts operational utility. Successive YOLO variants have improved accuracy through architectural innovations: YOLOv3 introduced multi-scale prediction; YOLOv5 popularized anchor-free heads; and YOLOv8 further refined the decoupled head design. In the context of UAS-based human detection, prior work has deployed YOLOv2[12], YOLOv3[13,16], YOLOv4-tiny[5], and YOLOv5[15]—all on RGB imagery. The present work adopts YOLOv12[17,18], the most recent iteration at the time of this study, which introduces attention-centric feature extraction that is particularly beneficial for thermal imagery where object-background contrast differs from visible-spectrum scenes.

The YOLOv12 model architecture consists of three main components:

Backbone: Employs attention-enhanced feature extraction modules that emphasize regions of interest such as thermal human contours in aerial frames. Unlike convolution-only backbones in prior YOLO versions, the attention mechanism allows the network to selectively weight spatial regions with high thermal contrast.

Neck: Performs multi-scale feature fusion to improve recognition of individuals across varying altitudes and thermal contrasts. This is essential for aerial imagery where apparent subject size changes with flight altitude.

Detection Head: Handles classification, bounding box regression, and pose estimation independently while sharing the backbone representation.

Dataset Collection and Preparation

Students collected thermal imagery using a Skydio X10 UAS equipped with the VT300-L payload package, which includes a radiometric thermal camera with specifications including 640×512 pixel resolution and less than 30 mK thermal sensitivity. The UAS provides multiple thermal color palettes, allowing collection of imagery under various conditions.

Individual frames were extracted from recorded thermal videos and annotated by students using the Roboflow platform. Each frame contains one or more human subjects captured from an aerial viewpoint in thermal format, where individuals are represented by heat signatures rather than identifiable features. This ensures privacy preservation.

The final dataset comprised 2,426 thermal images captured across four infrared spectrums (Iron-Red, Grayscale, Rainbow, and White-Hot), covering variations in illumination, background temperature, subject orientation, and partial occlusion. The dataset was partitioned as shown in Table 2.

Hardware and Software Infrastructure

All model training and inference were performed on a ground station laptop (NVIDIA RTX 4090 GPU, 64 GB RAM, Intel i9-13900HX CPU) running Ubuntu 22.04 with Python 3.10, PyTorch 2.1, and the Ultralytics YOLOv12 framework. The thermal video feed from the Skydio X10's VT300-L payload was streamed in real time to the ground station controller over the UAS's native Wi-Fi link. Frame extraction and annotation were performed using Roboflow's web-based interface. This ground-station-based inference architecture, as opposed to onboard edge computing, allows use of high-performance GPUs that would exceed the payload and power constraints of the UAS platform, enabling deployment of larger and more accurate models without compromising airframe endurance.

Table 2: Thermal dataset split used for model training and evaluation.

Subset	**Images**	**Proportion**
Training	1,944	80%
Validation	316	13%
Test	166	7%

Training Process

The model was trained using stochastic gradient descent with 640×640 input images, batch size of 16, and an initial learning rate of 0.01. Early stopping terminated training at epoch 354 once validation metrics stabilized. Students participated in monitoring training progress and analyzing convergence behavior.

Two-Stage Detection Pipeline

The complete system separates person detection from posture analysis:

Stage 1 (Person Detection): Given an input thermal frame, the trained model predicts person bounding boxes after non-maximum suppression. Each box defines a region of interest for further analysis.

Stage 2 (Pose Estimation): Each detected region is analyzed to extract 2D body keypoints including shoulders, hips, knees, and ankles. A geometric decision rule then classifies each individual as "Mobile" (upright) or "Immobile" (non-upright) based on vertical ordering of joints and aspect ratio of the bounding box. The term "Immobile" is deliberately chosen over terms such as "unconscious" or "fallen" because the system cannot infer the cause of a non-upright posture from thermal keypoints alone; a resting individual and an incapacitated individual present equivalent thermal signatures.

RESULTS

Model Performance

The trained detector achieved strong performance on the held-out test set, as shown in Table 3.

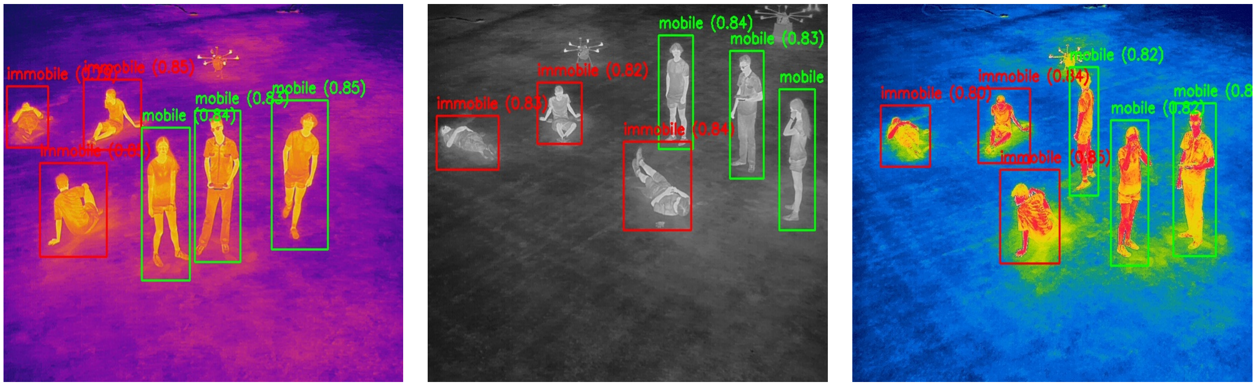
Table 3: Performance metrics on the held-out test split. Definitions are provided below for readers unfamiliar with object detection evaluation conventions.

Metric	Value	Interpretation
Precision	0.885	Of all bounding boxes the model predicted as "person," 88.5% were correct (i.e., 11.5% false positive rate).
Recall	0.815	Of all actual persons in the test images, the model successfully detected 81.5% (i.e., 18.5% were missed).
mAP@0.5	0.892	Mean Average Precision at 50% Intersection-over-Union (IoU) threshold—the standard detection accuracy metric.
mAP@0.5:0.95	0.607	Mean Average Precision averaged across IoU thresholds from 0.50 to 0.95, measuring spatial localization tightness.

The precision of 0.885 indicates that the detector produces relatively few false alarms, which is operationally important in SAR contexts where each false positive would divert responder attention. The recall of 0.815 means approximately one in five individuals may go undetected in a single pass; in practice, this is mitigated by the real-time nature of the system, which processes multiple overlapping frames as the UAS traverses the search area. The mAP@0.5 of 0.892 compares favorably with prior UAS-based human detection studies: Rizk et al.[13] reported 82% accuracy with YOLOv3 on RGB imagery, Jayalath et al.[14] achieved 84.1% true positives, and McGee et al.[16] reported lower performance with YOLOv3 on thermal data from a more constrained dataset. The gap between mAP@0.5 (0.892) and mAP@0.5:0.95 (0.607) indicates that while the model reliably detects persons, the predicted bounding boxes do not always tightly fit the target at stricter IoU thresholds—a common limitation in thermal imagery where object boundaries are inherently less defined than in visible-spectrum images due to thermal diffusion.

Field Validation

The complete two-stage pipeline was validated on imagery collected during additional flight sessions. These images were captured at different times of day, under varying ground temperatures, and across multiple thermal color palettes not explicitly represented during training. The detector reliably localized visible individuals with stable confidence values, and the pose-based classifier consistently produced correct condition assessments.

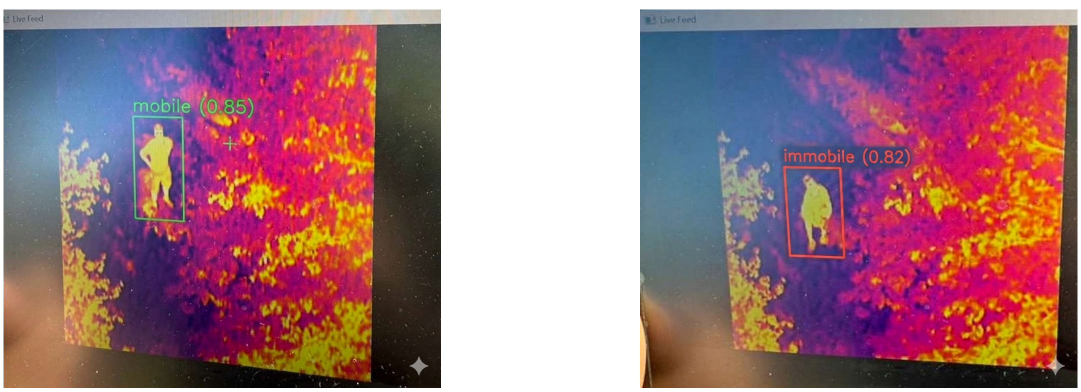


(a) Iron-Red palette (b) Grayscale palette (c) Rainbow palette

Figure 3: Two-stage inference on field imagery across different thermal palettes. Green boxes indicate mobile (upright) individuals; red boxes indicate immobile (non-upright) postures.

Real-Time Deployment

The framework was deployed on a live thermal video feed from the UAS. The system sustained throughput of 25–30 frames per second, confirming suitability for real-time aerial deployment. Field evaluations demonstrated robustness under challenging conditions, including successful identification of subjects partially obscured by vegetation.



(a) Detection of upright subject classified as "Mobile."

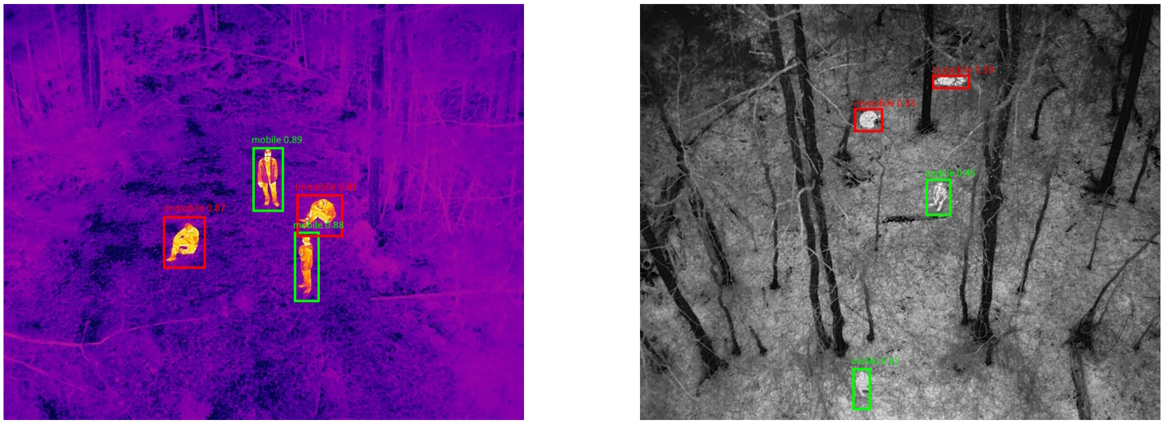
(b) Detection of partially occluded subject classified as "Immobile."

Figure 4: Real-time inference on live UAS thermal feed.

Stress Test: Dense Forest Canopy at Kennesaw Mountain

To evaluate the system under conditions representative of real-world SAR operations, a dedicated stress test was conducted at Kennesaw Mountain National Battlefield Park in Marietta, GA. This site features dense deciduous and mixed forest with heavy canopy cover, uneven terrain, and natural ground clutter—conditions that pose significant challenges for both thermal detection and UAS-based observation.

Four human subjects were positioned in natural, mixed postures (a combination of upright and non-upright) within the forest. The UAS was operated at altitudes of 40–50 m above ground level, positioning the aircraft above the tree canopy. This configuration simulates a realistic SAR scenario in which the UAS cannot descend below the canopy and must detect individuals through gaps in the foliage. Thermal video was captured across three color palettes: Iron-Red, Grayscale, and Rainbow.



(a) Iron-Red palette: detection in a small forest clearing.

(b) Grayscale palette: detection under heavy canopy occlusion.

Figure 5: Stress test results from Kennesaw Mountain dense forest environment. Green boxes indicate mobile (upright) individuals; red boxes indicate immobile (non-upright) postures. (a) All four subjects detected in a small clearing with confidence values of 0.87–0.89. (b) Subjects detected under dense tree cover with reduced confidence, demonstrating the system's ability to localize individuals through canopy gaps.

Figure 5a shows detection results in the Iron-Red palette within a small forest clearing. All four subjects were successfully detected, with two classified as mobile (confidence 0.89 and 0.88) and two as immobile (confidence 0.87 and 0.83). Confidence values remained high in this scenario because the clearing provided relatively unobstructed thermal signatures despite the surrounding tree canopy.

Figure 5b presents a more challenging case using the Grayscale palette, where subjects are positioned under dense canopy with significant tree trunk occlusion. The system still successfully localized individuals, though with notably reduced confidence values compared to the clearing scenario. This reduction is expected: partial occlusion by branches and foliage fragments the thermal silhouette, reducing the number of visible keypoints available for both detection and posture classification. Critically, no false positives were generated by warm tree trunks or sun-heated ground patches despite these objects producing thermal signatures comparable in intensity to human subjects. This robustness suggests that the model learned to discriminate based on spatial morphology rather than thermal intensity alone.

The Kennesaw Mountain stress test demonstrates that the proposed framework maintains operational utility in forested SAR environments—the precise conditions where thermal detection offers the greatest advantage over RGB-based systems, as foliage that visually conceals individuals is partially transparent to longwave infrared radiation. The observed reduction in confidence under heavy canopy quantifies a known limitation and provides a baseline for future work on occlusion-aware detection.

STUDENT LEARNING OUTCOMES AND ASSESSMENT

The undergraduate participant was a mechanical engineering student with no prior experience in machine learning, computer vision, or UAS-based research. Assessment of learning outcomes was conducted through direct observation by the graduate mentor, evaluation of project deliverables, and a structured end-of-project reflective interview.

Assessment Methods and Observed Evidence

Table 4 maps each learning objective to its assessment method and the observed evidence for the undergraduate student.

Observed Outcomes

The undergraduate student demonstrated measurable growth across the project timeline:

Technical skills: By the end of Phase 3, the student could independently configure training runs using the Ultralytics API, interpret precision-recall tradeoffs, and diagnose common training failures (e.g., overfitting from insufficient augmentation). At the start of the project, the student had no familiarity with Python-based ML frameworks or thermal imaging.

Research process understanding: The student developed an appreciation for the iterative nature of research through direct experience. A notable example occurred when the student discovered that inconsistent annotation quality in early batches led to degraded recall; this prompted a re-annotation effort and a subsequent improvement in model performance, providing a concrete lesson in how data quality propagates through the ML pipeline.

Professional development: The student improved technical communication abilities through documentation tasks and manuscript preparation. Contributing as a co-author on this conference paper provided exposure to the scholarly communication process, including peer review, revision, and technical writing conventions.

Career clarity: In the end-of-project interview, the student indicated that the experience clarified interest in pursuing graduate studies or industry roles involving robotics and AI, areas not previously considered prior to joining the project.

Student Feedback

Feedback was collected through a structured reflective interview at the conclusion of the project. Key responses included:

> *"Before this project, machine learning was just a buzzword to me. Actually training a model on data I helped collect and annotate made it click."* — Undergraduate student, Mechanical Engineering

Table 4: Mapping of assessment methods to learning objectives and observed evidence for the undergraduate participant.

Learning Objective	Assessment Method	Evidence / Outcome
UAS operation & data collection	Practical evaluation by graduate mentor	Student demonstrated independent thermal data capture by Phase 2; participated in all field sessions including Kennesaw Mountain stress test
ML workflow proficiency	Project deliverables (annotated dataset, training logs)	Annotated portion of 2,426-image dataset with consistent labeling quality; independently ran training experiments and interpreted loss curves by Phase 3
Iterative research process	Mentor observation, reflective interview	Student identified and corrected annotation errors that degraded model recall; demonstrated understanding of iterative refinement
Professional skills	Technical documentation, conference paper contribution	Contributed to literature review, documentation, and manuscript preparation; co-authored this conference paper
Connecting coursework to applications	Reflective interview	Student explicitly connected gradient descent (calculus) and matrix operations (linear algebra) to training behavior observed during experiments

> *"The one-on-one mentorship was the most valuable part. Having someone walk me through why the model was failing on certain thermal palettes, and then showing me how to fix it with augmentation, taught me more about debugging than any class project."* — Undergraduate student, Mechanical Engineering

When asked about areas for improvement, the student suggested that having a written onboarding guide or reference manual at the start of the project would have reduced the initial ramp-up time. The student also noted that more exposure to the mathematical foundations of attention mechanisms would have deepened understanding of the YOLOv12 architecture. These observations inform the recommendations for future projects discussed below.

DISCUSSION

Balancing Technical and Educational Goals

A key challenge in undergraduate research projects is balancing technical rigor with educational accessibility. This project addressed this challenge by structuring tasks in progressive phases while maintaining high standards for overall project outcomes.

It is worth acknowledging that the scaffolded structure of this project—with a clearly defined problem, a known methodological approach, and bounded scope—differs from the open-ended inquiry characteristic of independent research. The undergraduate student working within this framework did not experience the ambiguity of problem formulation or the possibility of fundamentally negative results that define authentic research practice. However, we contend that for an early-stage undergraduate with no prior ML experience, this scaffolding is pedagogically appropriate: it introduces the *mechanics* of research (systematic data collection, iterative experimentation, peer review of results) while managing cognitive load. As the student progressed through phases, increasingly complex tasks were introduced, providing a graduated transition toward independent technical decision-making. This approach aligns with the structured-to-open progression advocated in the course-based undergraduate research experience (CURE) literature[19].

Data annotation, while straightforward, was essential for project success and gave the student a meaningful contribution from the earliest stages. More complex tasks such as model training and loss curve interpretation were introduced progressively as the student demonstrated readiness.

Connections Between Technologies

The system pipeline illustrated in Figure 2 establishes explicit connections between the technologies and methods used in this work. At the hardware level, the Skydio X10 UAS with its VT300-L thermal payload provides the raw radiometric data stream. This stream is transmitted via the UAS controller to the ground station laptop, where the software stack (Python, PyTorch, Ultralytics YOLOv12) processes frames in real time. The pipeline separates concerns into two distinct stages: a learned component (YOLOv12 person detection and keypoint estimation, trained on our 2,426-image thermal dataset) and a rule-based component (geometric posture analysis using joint positions). This separation is a deliberate design choice—the learned component handles the perceptually difficult task of localizing humans in thermal clutter, while the rule-based component provides interpretable, deterministic condition classification without requiring additional labeled training data for each posture class. The ground-station architecture, rather than onboard edge computing used by Lin et al.[4] and Rizk et al.[13], enables deployment of larger models (YOLOv12 vs. YOLOv3-tiny) at the cost of requiring a reliable communication link between the UAS and ground station.

Framework for Future Projects

The mentorship structure developed for this project provides a template for engaging undergraduates in research:

(1) Define projects with clear technical goals and multiple entry points for students at different

skill levels.

(2) Create explicit connections between project tasks and underlying engineering principles.

(3) Implement regular check-ins and feedback mechanisms to monitor student progress and address difficulties.

(4) Document procedures and best practices to facilitate onboarding of new team members.

(5) Recognize student contributions through co-authorship, presentations, and other professional development opportunities.

Limitations

While the mentorship framework achieved its goals for this project, several limitations should be noted:

This study reports on a single undergraduate participant, which limits generalizability. The observed outcomes represent one case study rather than a statistically powered sample. Future implementations should track outcomes across multiple students and cohorts to validate the mentorship model.

Assessment of learning outcomes relied on mentor observation, project deliverables, and a reflective interview rather than controlled pre/post evaluations with validated instruments (e.g., the SURE survey[19]). Future iterations will incorporate pre-project baseline assessments to enable measurement of learning gains rather than post-only perceptions.

The specific focus on computer vision and UAS technology, while engaging, represents a bounded problem space that may not generalize to all research domains. The student should be made aware that open-ended research typically involves greater ambiguity in problem definition and methodology selection than the scaffolded structure used here.

CONCLUSION

This work presented a thermal human-condition detection framework designed for aerial search-and-rescue operations while simultaneously serving as a platform for undergraduate research engagement. The technical system achieved strong performance with precision of 0.885, recall of 0.815, and mAP@0.5 of 0.892, comparing favorably with prior UAS-based detection systems that operate on RGB imagery. The mobile/immobile classification, based on geometric analysis of pose keypoints, provides an interpretable and privacy-preserving condition assessment suitable for real-world SAR deployment. A dedicated stress test at Kennesaw Mountain confirmed operational viability under dense forest canopy with zero false positives from environmental thermal sources.

The project also demonstrated a one-on-one mentorship model for engaging an undergraduate mechanical engineering student with no prior ML experience. Through a phased task progression—from literature review and data annotation through model training and manuscript preparation—the student developed functional competency in UAS operation, computer vision fundamentals, and the iterative practices of technical research. The student's reflective feedback highlighted the value of connecting classroom mathematics to applied ML and the benefit of direct graduate-student mentorship.

The mentorship framework developed through this project offers a reproducible model for engaging early-stage undergraduates in scaffolded research activities. By providing appropriately structured tasks, connecting hands-on activities to underlying principles, and offering one-on-one guidance, research laboratories can introduce undergraduates to technical research practices while maintaining project quality. We acknowledge that this scaffolded approach differs from open-ended independent research and serves as a preparatory stage rather than a substitute.

Future work will expand the dataset to additional environments and seasons, continue refinement of the detection system with multi-class posture categorization, and implement validated pre/post assessment instruments (e.g., the SURE survey [19]) across a larger cohort of undergraduate participants to enable rigorous measurement of student learning gains.

ACKNOWLEDGMENTS

The authors acknowledge the use of the Roboflow platform for dataset preparation and annotation, and the Ultralytics framework for model training and evaluation.

AI-based writing assistance tools were used for spell checking, grammatical error correction, and formatting during the preparation of this manuscript. All technical content, results, and interpretations remain the sole responsibility of the authors.

References

[1] S. Waharte. Supporting search and rescue operations with UAVs. In *2010 International Conference on Emerging Security Technologies*, pages 142–147, Canterbury, 2010.

[2] A. Ramachandran. A review on object detection in unmanned aerial vehicle surveillance. *International Journal of Cognitive Computing in Engineering*, 2021.

[3] A. Jain. AI-enabled object detection in UAVs: Challenges, design choices, and research directions. *IEEE Network*, 35(10):129–135, 2021.

[4] B.-S. Lin. Fall detection system with artificial intelligence-based edge computing. *IEEE Access*, 10:4328–4339, 2022.

[5] K.-J. Chang. Flying watchdog: A drone with edge AIoT for residential safety and fall detection by face and posture recognition. In *IEEE APWCS 2022*, Hualien, 2022.

[6] S. Sonkar. Real-time object detection and recognition using fixed-wing LALE VTOL UAV. *IEEE Sensor Journal*, 22(21):20738–20747, 2022.

[7] N. Papyan. AI-based drone assisted human rescue in disaster environments: Challenges and opportunities. *Pattern Recognition and Applications*, 34:169–186, 2024.

[8] A. Weber. A dual-use unmanned aerial system for precision agriculture and search and rescue applications. In *AIAA SciTech Forum*, San Diego, CA, USA, 2016.

[9] S. Bhandari. Search and rescue using unmanned aerial vehicle. In *AIAA SciTech Forum*, Kissimmee, FL, USA, 2015.

[10] C. Hayner. Real-time human detection with integration of visual and thermal data from high altitude sUAS. In *AIAA SciTech Forum*, 2021.

[11] T. Harrington. UAV swarms for migration flow monitoring and search and rescue mission support. In *AIAA SciTech Forum*, Kissimmee, FL, USA, 2018.

[12] K. Boudjit. Human detection based on deep learning YOLO-v2 for real-time UAV applications. *Journal of Experimental and Theoretical Artificial Intelligence*, 43(3):527–544, 2021.

[13] M. Rizk. Toward AI-assisted UAV for human detection in search and rescue missions. In *International Conference on Decision Aid Sciences and Application*, Bahrain, 2021.

[14] J. Jayalath. Drone-based autonomous human identification for search and rescue missions in real-time. In *10th International Conference for Information and Automation for Sustainability*, Sri Lanka, 2021.

[15] S. Caputo. Human detection in drone images using YOLO for search-and-rescue operations. In *Advances in Artificial Intelligence*, volume 13196. Springer, Cham, 2021.

[16] J. McGee. Unmanned aerial vehicle and artificial intelligence for thermal target detection in search and rescue applications. In *International Conference on Unmanned Aircraft Systems (ICUAS)*, Greece, 2020.

[17] Y. Tian, Q. Ye, and D. Doermann. YOLOv12: Attention-centric real-time object detectors. *arXiv preprint arXiv:2502.12524*, 2025.

[18] Y. Tian, Q. Ye, and D. Doermann. YOLOv12: Attention-centric real-time object detectors. Software Release, 2025. Available: https://github.com/sunsmarterjie/yolov12, Accessed: October 2025.

[19] D. Lopatto. Undergraduate research experiences support science career decisions and active learning. *CBE—Life Sciences Education*, 6(4):297–306, 2007.

[20] S. Hsieh. Research experiences for undergraduate engineering students. In *2013 ASEE Annual Conference & Exposition*, Atlanta, GA, 2013.

[21] University of Cincinnati. Unmanned aerial vehicle multi-agent systems technology research (UAV MASTER) lab. https://www.ceas.uc.edu/research/centers-labs/uav-master-lab.html, 2024. Accessed: March 2025.

[22] R. Goode, T. Dunlap, and B. Nejmeh. Aerial drone: An effective tool to teach information technology and cybersecurity through project based learning to minority high school students in the U.S. *Education and Information Technologies*, 26:6413–6451, 2021.

[23] J. Redmon. You only look once: Unified, real-time object detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2016.