Using Drone Cluster and Machine Learning for Search and Rescue in Natural Disasters

By

Albaraa Azzam Zaid Alkilani

Supervisor

Prof. Gheith Abandah

Co-Supervisor

Dr. Yazan Al-Zain

This Thesis was Submitted in Partial Fulfillment of the Requirements for the Master's Degree in Artificial Intelligence and Robotics

School of Graduate Studies

The University of Jordan

May, 2025

COMMITTEE DECISION

This Thesis/Dissertation (Using Drone Cluster and Machine Learning for Search and Rescue in Natural Disasters) was Successfully Defended and Approved on

____

Examination Committee	**Signature**
Dr. Gheith A. Abandah, (Supervisor) Professor of Computer Engineering	----------------------------
Dr. Yazan Al-Zain (Co-supervisor) Associate Professor of Industrial Engineering	----------------------------
Dr. Mohammed R. Abdel-Majeed (Member) Associate Professor of Computer Engineering	----------------------------
Dr. Ahmad Al-Jaafreh, (Member) Professor of Computer Engineering (Tafila Technical University)	----------------------------

DEDICATION

To my beloved family: My father, Azzam, who has always been my idol, your integrity, wisdom, and unwavering support have guided every step of my journey; and my mother, Amal, whose boundless love, patience, and encouragement have sustained me through every challenge. To my sisters Rawan, my oldest companion and childhood confidante; Rana, my sporting partner and ever-ambitious inspiration; Farah, my wise and compassionate doctor; and Layan, the youngest yet endlessly insightful soul, thank you for standing by me, sharing in my triumphs, and lifting me when I stumbled.

To the love of my life, my wife Tasneem: you have been my greatest companion throughout this journey, pulling me through long nights and difficult moments, always believing in me, and never losing faith in what I could achieve.

To my mentor, Dr. Waleed: thank you for teaching me, for giving me the space to grow, and for believing in my potential every step of the way.

To my friends, my second family whose constant encouragement and support have made this endeavor possible.

To all of you, I offer my deepest gratitude. May this work be a testament to your love and support, and may every achievement draw me ever closer to you.

AKNOWLEDGEMENT

I would like to express my deepest gratitude to my supervisor and mentor, Professor Gheith Abandah, whose unwavering guidance, and lofty standards have shaped both my academic development and personal growth. His insistence on excellence and his constructive criticism have driven me to continually strive for the best; it is a true honor to have had him as my supervisor, and I will forever carry the privilege of his mentorship with immense pride.

My sincere thanks go to my co-supervisor, Dr. Yazan Al-Zain, for his steadfast support and belief in my work. I am especially grateful for the opportunity he provided to participate in Eco-Smart project by DAAD that allowed me to present my research in Germany and Qatar. Experiences that transformed my master's thesis into not only an academic endeavor but also a profound life lesson, exposing me to the collaboration of multinational teams working at the forefront of AI and computer vision.

I am also deeply appreciative of the Jordan Design and Development Bureau, and particularly Eng. Radi Al-Wardat, whose generous facilitation of multiple drone flights and the delegation of team time and resources made the data collection for this project possible. Such collaborations between Jordanian institutions serve to elevate the quality and impact of our higher-education research.

Finally, I extend my thanks to the Computer Engineering Department and the School of Engineering at my university. As a graduate of this school, I was continually inspired by their commitment to advancing higher education in Jordan and equipping students with the competencies and skills needed to make meaningful contributions on the global stage.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AoI	Area of Interest
CARC	Civil Aviation Regulatory Commission
CBBA	Consensus-Based Bundle Algorithm
CNN	Convolutional Neural Network
DFOV	Diagonal Field of View
EO	Electro-Optical
FPS	Frames per Second
GIS	Geographic Information System
GSD	Ground Sampling Distance
HFOV	Horizontal Field of View
HMM	Hidden Markov Model
IoU	Intersection over Union
JNDRRS	Jordan National Disaster Risk Reduction Strategy
JODDB	Jordan Design and Development Bureau
LADD	Lacmus Drone Dataset
mAP	Mean Average Precision
RCPP	Recursive Coverage Path Planning
RPV	Remotely Piloted Vehicle
RPAS	Remotely Piloted Aircraft System
SAR	Search and Rescue
SRT	Secure Reliable Transport
SVM	Support Vector Machine
UAV	Unmanned Aerial Vehicle
UAS	Unmanned Aerial System
UHD	Ultra-High Definition

VFOV	Vertical Field of View
WebRTC	Web Real-Time Communication
YOLO	You Only Look Once (Object Detection Algorithm)

Using Drone Cluster and Machine Learning for Search and Rescue in Natural Disasters

By
Albaraa Azzam Zaid Alkilani

Supervisor
Dr. Gheith Abandah, Prof.

Co-supervisor
Dr. Yazan Al-Zain

ABSTRACT

This thesis explores the integration of drone cluster with advanced machine learning methodologies to enhance search and rescue (SAR) operations, particularly in response to natural disasters in Jordan. The research systematically addresses the drone deployment through efficient route planning algorithms, demonstrating the effectiveness of the lawnmower algorithm in minimizing coverage overlaps and enhancing operational efficiency. A mathematical relationship between drone elevation and camera specifications is established, significantly improving detection efficiency and providing practical operational guidelines.

Extensive experiments led to the selection of the YOLOv8 machine learning model, pretrained on drone-specific datasets VisDrone, then trained on the collected dataset from Al-Dhlail, Zarqa, Jordan and enhanced by mosaic augmentation, achieving outstanding detection performance with high recall (97.6%) and precision (97.0%). Additionally, a real-time communication framework integrating centralized analysis is proposed for swiftly disseminating critical intelligence to SAR teams.

Despite these advancements, several limitations were identified, including drone range constraints, processing latency, and the robustness of machine learning models under extreme environmental conditions. Future research directions are proposed, emphasizing the physical implementation and extensive field testing of the system, development of adaptive drone altitude algorithms, incorporation of diverse environmental datasets, exploration of hybrid processing solutions, and advanced communication technologies. These recommendations aim to further refine the SAR capabilities and improve operational responsiveness, ultimately contributing significantly to disaster response efficiency and life-saving outcomes.

Chapter One

1 Introduction

Unmanned aerial vehicles (UAVs), more commonly referred to as drones, are small aircraft that fly autonomously. Originally developed for military purposes, these devices have since become a major focus of research. In addition to the term "Unmanned Aerial Vehicles" drones are known by other names, including Unmanned Aerial Systems (UAS), Remotely Piloted Vehicles (RPV), and Remotely Piloted Aircraft Systems (RPAS). While the term "drone" is widely recognized by the public, the other designations, such as UAVs, RPA, RPAS, and UAS, are official names attributed to the drone technology depending on the authority (Restas, 2015).

Drones are available for a variety of applications and can be used for crime scene surveillance (Bucknell and Bassindale, 2017), marine fauna detection (Colefax *et.al*, 2019), habitat destruction assessment (Barnas *et al*., 2019), crop monitoring (Bendig *et al*., 2014), and vegetation mapping (Díaz-Varela *et al*., 2015). Additionally, drone mapping has applications in numerous industries, including construction, agriculture, mining, and infrastructure inspection (Shahmoradi, *et al*., 2020). Recently, the application of drones to disaster or humanitarian relief has expanded to include search and rescue missions (Thavasi and Suriyakala, 2012; Van Tilburg, 2017), disaster prevention (Petrides *et al*., 2017), and disaster management (Apvrille *et al*., 2014; Quaritsch *et al*., 2010).

The rise of drone technology has instigated monumental shifts across various sectors, particularly significant is their use in search and rescue (SAR) operations following natural disasters. Drones, with their ability to rapidly survey large areas and provide real-time visual feedback, are invaluable in post-disaster scenarios where immediacy is crucial (Lyu *et al*., 2023).

For SAR teams, drones have become indispensable tools, enabling a wide array of applications from mapping disaster-stricken areas, identifying affected victims, to even delivering essential supplies to remote or inaccessible regions (Papyan *et al*., 2024). The amalgamation of innovative sensor technologies, real-time data analytics, and precision positioning systems has expedited response times and enhanced the efficacy of post-disaster recovery operations (Goodchild and Glennon, 2010). Furthermore, the convergence of machine learning and artificial intelligence techniques with UAVs has bolstered the ability to detect and respond to emergencies in cluttered or visually challenging environments (Floreano and Wood, 2015).

1.1 Research Problem

How can drone and machine learning techniques be effectively utilized for the rescue of people in distress during natural disasters in Jordan?

In the face of natural disasters, timely and efficient rescue operations are of the essence to mitigate human casualties and save lives. Jordan, with its diverse terrains ranging from urban landscapes to desert expanses, presents unique challenges in conducting search and rescue missions during such calamities. The incorporation of drones, combined with the potential of machine learning, offers a promising solution to enhance the speed and accuracy of these operations.

To maximize their effectiveness, it is imperative to precisely calculate the flight routes, ensuring that drones cover disaster-prone zones thoroughly and swiftly. Additionally, given the variability in the terrains, understanding the optimal relationship between a drone's elevation and its camera specifications becomes crucial. This relationship ensures that aerial images captured are of the highest clarity, permitting accurate identification of individuals in distress.

Furthermore, machine learning models can analyze these images specifically trained for image recognition in disaster scenarios, capable of distinguishing human figures amidst rubble or floodwaters.

However, the mere capability of drones and machine learning is not sufficient; a well-laid strategic coordination plan with the Jordanian Civil Defense Department is essential.

This coordination ensures real-time data sharing, efficient drone operation without hindering other rescue endeavors, and seamless integration of this technology into the broader rescue framework. Through addressing these intricacies, this research seeks to explore and establish a comprehensive approach to leveraging drones and machine learning in enhancing the efficacy of rescue operations during natural disasters in Jordan.

1.2 Thesis Importance

Jordan's geographical and climatic intricacies render it particularly vulnerable to various natural disasters. From the seismic activities that raise concerns about earthquakes to the arid conditions that exacerbate the risk of fires, and the occasional torrential rains leading to floods – the stakes have never been higher. In such a context, researching the applicability and effectiveness of drones as a primary response tool becomes paramount (Hadadin and Tarawneh, 2009).

Firstly, the temporal aspect of disaster response is pivotal. Every second counts when lives are at stake. Traditional rescue efforts, although valiant, often grapple with delays due to terrain challenges, lack of real-time information, and logistical impediments. Drones, with their aerial vantage point, can quickly survey vast areas, identify stranded individuals, assess infrastructural damage, and provide real-time data to rescue teams on the ground. Their ability to hover, zoom, and access hard-to-reach areas can exponentially expedite the decision-making process. Hence, researching their efficacy can drastically improve the "golden hours" of rescue operations, potentially saving countless lives (Goodchild and Glennon, 2010).

From an infrastructural and city planning perspective, the data collected from these drones can be invaluable. A comprehensive understanding of the most affected areas, the patterns of damage, and the vulnerabilities of certain structures can inform future city planning and construction norms. This research could, therefore, serve as a foundational reference for architects, engineers, and urban planners aiming to construct more resilient infrastructures (Zou *et al*., 2017).

Furthermore, a deeper exploration into the economic ramifications is essential. While the immediate costs associated with procuring, maintaining, and operating drones might seem substantial, they need to be weighed against the potential long-term savings. By potentially reducing the duration of rescue operations, minimizing damage, and streamlining resource allocation, drones might offer significant economic benefits eventually (Aldrich and Sawada, 2015).

Finally, the global implications of the research are profound. As climate change intensifies and the frequency and severity of natural disasters escalate worldwide, nations are on the lookout for innovative solutions. Jordan's exploration into the realm of drones for disaster management can serve as a case study, a beacon for nations facing similar challenges. It can catalyze a global movement towards integrating technology and traditional disaster response mechanisms (Yamamura, 2010).

In summation, the importance of researching drones for disaster management in Jordan is multifaceted, encompassing immediate, local, and global consequences. The potential outcomes of the research might not just reshape Jordan's disaster response blueprint but could also influence global paradigms.

1.3 Thesis Objectives

This thesis aims to develop an integrated drone-based search-and-rescue system tailored to Jordan's terrain, combining a cooperative UAV cluster with advanced computer-vision algorithms. The detailed objectives of this research are:

- Segmenting drones to cover the path and routes as subzones in operating area.
- Finding routes for the drones to achieve efficient detection and rescuing operation.
- Finding the best algorithm to find the best route.
- Finding out the mathematical relationship between drones' elevation and camera specs.
- Training an image recognition model to detect people in potential disaster regions.
- Developing a drone system to detect people in distress and notify the Jordanian civil defense department.
- Tracking people who are stranded or drafted in a natural disaster (optional).

1.4 Thesis Questions and Assumptions

This section presents the core research questions guiding our Jordan-tailored drone-based SAR system and the key assumptions regarding UAV coordination, sensor performance, and environmental operating conditions.

- What is the best route the drone cluster can follow to achieve the best intelligence?
- How can we control the elevation of the drones depending on the specifications of the camera used?
- What machine learning model works best for detecting people in disaster regions?
- How can we produce a comprehensive plan to link the Civil Defense Department with the drones to make sure intelligence is effectively communicated?

The main assumptions of this research are:

- The drones to be used in this research are by ANAFI USA.
- The weather conditions are suitable for the drones to operate successfully.
- The data to be gathered or used is suitable and representable of Jordan.
- Drones will be sending the data to a remote server for analysis.

1.5 Thesis Contribution

This thesis delivers three principal contributions toward enhancing Jordanian SAR capabilities through a UAV fleet:

Jordan-Specific Human-Detection Deep-Learning Model (Primary Contribution): The core achievement of this work is a computer-vision model trained on a purpose-built dataset of Jordan's deserts, rocky highlands, agricultural terraces, and peri-urban zones. By capturing the country's distinctive colors, textures, and lighting, the model reliably distinguishes human figures from complex backgrounds typical of local SAR scenarios. Its accuracy under harsh sun, low-angle twilight, and dust haze makes it the linchpin of the entire UAV search-and-rescue pipeline.

Adaptive Area-Segmentation & Route-Assignment Algorithm: Building on the vision model, we introduce an algorithm that divides any search region into balanced sub-areas, taking account of terrain difficulty, UAV endurance, and no-fly restrictions. It then allocates dynamic waypoint sequences to each drone, enabling real-time re-tasking as environmental or mission conditions change. This ensures systematic coverage and maximizes the effectiveness of the human-detection model in the field.

Modular UAV-Based SAR System Architecture: To support the above innovations, we design a flexible drone framework tailored to Jordanian SAR operations. It couples ground-control stations, resilient comms links, power-management units, and swappable sensor payloads behind open software interfaces. The architecture prioritizes rapid deployment, interoperability with civil-defense assets, and sustained aerial coverage from remote wadis to densely populated outskirts, thereby providing a robust platform for the algorithm and vision model to operate on.

1.6 Thesis Methodology

This section outlines the systematic approach we follow to achieve the objectives of the thesis:

- Conduct a literature review of UAV-enabled SAR systems, routing/segmentation algorithms, and human-detection models.
- Design an end-to-end UAV–SAR architectural framework for Jordan’s civil-defense integration.
- Develop an elevation algorithm based on drone specifications.
- Explore and select suitable multi-agent routing techniques.
- Design an iterative area-segmentation and route-assignment algorithm under mission constraints.
- Decide on the types and sources of data required for human-detection model training.

- Clean and preprocess the collected dataset.
- Collect Jordanian terrain data (high-resolution RGB images)
- Select the best pre-trained computer vision model for human detection.
- Train a Jordan-specific human-detection computer vision model
- Fine-tune the chosen pre-trained model on the Jordanian dataset.

1.7 Thesis Outline

This thesis consists of 7 chapters; the following chapters will include the following:

Chapter Two: literature review for existing UAV-based SAR systems, routing and segmentation algorithms, and human-detection vision models, identifying gaps and opportunities for Jordan-focused implementation.

Chapter Three System Design for Drone Segmentation, Routing and Elevation presents the end-to-end UAV SAR architecture, details area-segmentation and mission planning, and introduces algorithms for route assignment and optimal flight altitude.

Chapter Four: Data Collection and Methodology Describes the rationale, planning, and execution of field campaigns to gather high-resolution images to the diverse Jordanian terrains. It also includes the preprocessing, cleaning, augmentation, and exploratory analysis of the collected dataset, culminating in the finalized training and validation splits.

Chapter Five: Model Selection details the evaluation of candidate detection architectures, transfer-learning experiments (including VisDrone-pretrained models), hyperparameter tuning, and final model fine-tuning with mosaic augmentation.

Chapter Six: Discussion, Conclusion, and Future Work analyzes results across routing, elevation, and vision components; summarizes the key findings; and proposes directions for extending and deploying the system.

Chapter Two

2 Literature Review

Chapter Two provides a comprehensive review of the growing body of research on UAVs in SAR operations. Over the past decade, advances in sensor technology, wireless communications, and artificial intelligence have transformed drones from simple aerial cameras into intelligent platforms capable of detecting survivors, assessing damage, and securely transmitting critical data in real time. It includes studies of different life saving UAVs applications such as flood and avalanche detection, post-disaster damage assessment, human-voice localization, and voice-based survivor recognition.

While these investigations collectively attest to the life-saving potential of drone-based SAR, they also underscore persistent barriers: performance degradation under adverse environmental conditions, reliance on specialized datasets and homogeneous hardware, and the absence of universally accepted standards for interoperability and data security. Furthermore, practical deployment demands international cooperation, robust regulatory frameworks, and efficient techniques for rapid data processing and decision support.

2.1 Importance of Drones in SAR Missions

The utilization of drones, also known as UAVs, has dramatically revolutionized the field of search and rescue, significantly improving the efficiency, speed, and overall effectiveness of these critical missions. Drones offer unique advantages, including rapid deployment capabilities, extensive area coverage, and real-time data collection, fundamentally transforming traditional SAR methods.

The ability of drones to rapidly survey large disaster areas has been well documented. For example, Tuśnio and Wróblewski (2022) demonstrated that drones could effectively map a 25-hectare area within approximately 20 minutes, far surpassing the capabilities of ground-based rescue teams in both speed and accuracy. Such swift aerial reconnaissance greatly accelerates response times. Empirical studies underscore this benefit, showing instances where drone-assisted searches have located victims approximately 80% faster compared to traditional ground search teams, significantly increasing survival rates during critical emergency situations.

Another substantial advantage of employing drones in SAR missions is the considerable enhancement of safety for rescue personnel. SAR missions frequently require operations in challenging or hazardous environments, including rugged terrains, avalanche-prone regions, and structurally compromised buildings. By deploying drones in these contexts, human exposure to danger is minimized. Vincent-Lambert *et al*. (2023) highlight scenarios wherein drones provided crucial visual assessments in dangerous terrains, effectively reducing the physical risks faced by rescue teams.

Technological advancements in drone capabilities have further amplified their effectiveness in SAR missions. Thermal imaging cameras and high-resolution sensors fitted onto drones enable the detection of victims even in conditions of poor visibility, dense foliage, or during nighttime operations. According to Wankmüller *et al*. (2021), these enhancements substantially increase the probability of detecting victims who might otherwise remain unnoticed by conventional search methods.

Further emphasizing the strategic value of drones, several studies document their successful integration into various SAR scenarios. UAVs have effectively supported rescue missions following natural disasters, including earthquakes and floods, by providing timely and accurate damage assessments, ultimately facilitating rapid and informed decision-making (Daud *et al*., 2021; Alsumayl *et al*., 2023). Moreover, drones have demonstrated notable efficacy in locating missing people in vast or otherwise inaccessible terrains, significantly reducing search times, and improving mission success rates (Lygouras *et al*., 2019).

Additionally, drones offer critical support in law enforcement and tactical pursuits, providing aerial surveillance and real-time intelligence that enhances operational efficiency and effectiveness. Law enforcement agencies increasingly rely on drone technology to manage and oversee complex situations, from pursuit scenarios to broader public safety operations, demonstrating the versatility and essential role of UAVs across various SAR-related applications (Jiménez *et al*., 2020).

Given these capabilities, drones not only represent a technological advancement but also a critical tool in modern SAR strategies. As drone technology continues to evolve, it is expected that UAVs will play an even more integral role in enhancing SAR operations, thereby significantly increasing the efficiency, safety, and overall effectiveness of rescue efforts worldwide.

2.2 Drone Applications in SAR

Drones have been extensively employed in various SAR contexts, each with unique challenges and operational requirements. This section provides a comparative analysis of drones' effectiveness in diverse SAR scenarios, highlighting their versatility and adaptability.

Natural Disasters: In the aftermath of natural disasters such as earthquakes, hurricanes, and wildfires, drones have emerged as invaluable assets for rapid damage assessment and victim detection. Daud *et al*. (2021) emphasize drones' rapid assessment capabilities post-earthquake, providing accurate data that significantly accelerates rescue missions. Similarly, during hurricane-induced floods, drones offer crucial aerial perspectives, enabling emergency services to allocate resources efficiently and rapidly initiate rescue operations.

Flood Detection: Floods represent one of the most common and destructive natural disasters globally. UAV-based systems equipped with advanced sensors and machine learning algorithms provide enhanced capabilities for precise flood detection and monitoring. Alsumayl *et al*. (2023) demonstrate how drones can rapidly identify flooded regions, assess the severity of inundation, and support emergency response teams in strategically deploying rescue efforts, greatly reducing response times, and improving survival outcomes.

Missing Person Searches: One of the most challenging aspects of SAR missions involves locating missing people, particularly in remote, expansive, or densely vegetated areas. UAV technology, specifically incorporating thermal imaging and real-time video analytics, substantially enhances search efficiency and accuracy. Lygouras *et al*. (2019) document instances where drones effectively located individuals within significantly shorter periods than traditional methods, thereby improving rescue success rates and overall operational efficiency.

Law Enforcement and Pursuits: Law enforcement agencies increasingly rely on drones to enhance public safety operations. UAVs provide critical support during tactical pursuits and surveillance operations, offering real-time aerial intelligence that significantly improves strategic decision-making. Jiménez *et al*. (2020) highlight drones' capabilities in providing continuous monitoring and detailed situational awareness, allowing law enforcement personnel to manage complex operational environments more effectively, ultimately improving public safety and operational outcomes.

Collectively, these diverse applications underscore the multifaceted benefits drones offer across a spectrum of SAR scenarios, solidifying their status as essential tools in modern emergency management strategies.

2.3 Integrated Drone-SAR Systems

The integration of drones into existing SAR systems has significantly enhanced mission effectiveness, enabling a more coordinated and efficient response to emergencies. Effective drone integration requires robust communication frameworks, seamless data integration with existing SAR protocols, and enhanced coordination between drones and ground response teams.

Goodrich *et al*. (2007) proposed frameworks for drone integration, including remote-led, base-led, and sequential operations. These frameworks emphasize drones' roles in providing continuous aerial surveillance, supporting ground teams by rapidly relaying real-time information, and autonomously executing preliminary area searches. Such integration frameworks optimize SAR operations, ensuring quicker response times and improved victim detection accuracy.

Challenges to successful drone integration include regulatory restrictions, interoperability issues among communication systems, data privacy concerns, and the need for specialized training for SAR personnel. Wankmüller *et al*. (2021) highlight solutions such as Germany's EGRED2 manual, which provides standardized guidelines for drone integration, addressing communication interoperability and mission planning.

Further advancements in integrated SAR systems involve leveraging artificial intelligence and machine learning algorithms for real-time image analysis, facilitating faster decision-making processes, and significantly enhancing mission outcomes. This comprehensive integration ensures drones not only complement but significantly augment traditional SAR capabilities, solidifying their critical role in modern emergency response operations.

2.4 Jordan’s Vulnerability to Natural Disasters

Jordan is notably vulnerable to various natural hazards, including flash floods, earthquakes, and droughts. Historically, Jordan has experienced severe flash floods, such as the devastating event near the Dead Sea in 2018, highlighting the urgency of improved preparedness and effective SAR responses. The Jordan National Disaster Risk Reduction Strategy (JNDRRS) underscores these vulnerabilities, emphasizing the necessity for enhanced SAR operations given the increasing frequency of extreme weather events associated with climate change.

The country’s geographical location along the Dead Sea Transform fault system further exposes it to significant seismic risks. Historical earthquakes, including the notable event of 1927, illustrate the potential for severe consequences. Recent assessments indicate that major urban areas, including Amman, Zarqa, and Irbid, are particularly susceptible due to their proximity to seismic zones (Jordan National Disaster Risk Reduction Strategy, (2019–2022).

2.5 Ground Sampling Distance Recommendations

Several studies converge on achieving a Ground Sampling Distance (GSD) of roughly 1–2 cm/px for UAV-assisted SAR to guarantee both automated detection accuracy and human interpretability. The SARUAV (2022) FAQ recommends configuring camera parameters and flight altitude to attain a GSD between 1.0 cm/px and 2.0 cm/px, with an optimal value of approximately 1.3 cm/px for precise person detection and coordinate estimation in SAR operations.

Professional UAV photogrammetry guidance further notes that high-detail missions commonly operate at 1.5–2.5 cm/px, while professional surveys often target as fine as 1.0 cm/px to maximize mapping and inspection accuracy (ASPRS, 2014). Molina *et al*. (2012) work on thermal imaging campaigns showed effective human recognition at GSDs of 5.5 cm/px during daytime and 3.7 cm/px at night, suggesting that sensor modality can permit somewhat coarser GSDs under specific conditions.

Finally, empirical evaluations by Kannadaguli (2020) and Berndt *et al*. (2023) on YOLO-based detectors reveal that detection performance is strongly influenced by flight height and thus GSD highlighting the importance of balancing coverage area with spatial resolution to suit the chosen AI model. Taken together, these findings reinforce the guideline that UAV flight planning for SAR should aim for a GSD of around 1–2 cm/px, adjusted for sensor type, terrain, and algorithmic requirements.

2.6 Drone Cluster Routing Techniques

Drone cluster routing techniques are crucial for maximizing search efficiency in SAR missions. The lawnmower routing method, characterized by systematic parallel sweeps, consistently demonstrates superior coverage and efficiency. Comparative studies highlight its ability to systematically cover large areas and minimize missed zones which makes it perfect in SAR scenarios. According to Ousingsawat and Earl (2007), lawnmower routing achieved over 96% coverage efficiency, substantially reducing total search distances compared to traditional methods.

Alternative methods such as spiral, random walk, and grid searches have been examined extensively. Spiral patterns effectively concentrate search efforts near known points but can become inefficient in irregularly shaped areas. Random walk methods, though flexible, exhibit significant inefficiencies due to redundant coverage and unpredictable outcomes. Grid searches offer thoroughness but can be resource-intensive and time-consuming (Gal, 2023). Overall, systematic lawnmower patterns remain the optimal choice due to their balance of coverage, efficiency, and reliability.

The integration of multiple UAVs in SAR operations has further enhanced the efficiency of the lawnmower technique. Researchers have proposed cluster-adapted lawn mower strategies where multiple drones coordinate to cover large areas simultaneously (Sundelius *et al.,*2024). This approach not only reduces search time but also allows for dynamic task allocation based on the capabilities and flight constraints of individual drones. This study demonstrated that cluster-adapted strategies can be combined with probabilistic methods like Bayesian Search and A-star algorithms to optimize search efficiency in complex environments.

The lawnmower routing technique has been successfully applied in various SAR contexts, including maritime and wilderness environments. For maritime SAR, researchers have combined the lawnmower pattern with algorithms like the Recursive Coverage Path Planning (RCPP) to ensure full-area coverage while minimizing energy consumption (Guo *et al*., 2023). In wilderness SAR, the technique has been integrated with simulation environments and object detection systems like YOLO to enhance search efficiency and victim detection rates (Thuan *et al*., 2023).

2.7 Segmentation Techniques for Area Allocation

Effective area segmentation techniques significantly enhance multi-drone SAR operations by optimizing workload distribution and minimizing redundancy. Voronoi partitioning is widely employed due to its efficiency in evenly dividing search areas based on drone proximity. It dynamically adjusts regions, ensuring balanced task allocation and efficient resource utilization (Gao and Zhang, 2022).

Adaptive clustering methods, including K-means, dynamically group areas based on evolving mission parameters and environmental conditions, improving operational flexibility and responsiveness. Additionally, market-based allocation approaches, such as the Consensus-Based Bundle Algorithm (CBBA), provide efficient and balanced workload distribution by allowing drones to bid on specific tasks, resulting in efficient resource allocation and minimized search times (Gao and Zhang, 2022).

2.8 Computer Vision Models in SAR Operations

Advanced computer vision technologies significantly bolster drone capabilities in SAR missions. Models such as Faster R-CNN, SSD, RetinaNet, U-Net, and DeepLab have substantially improved victim detection accuracy and operational effectiveness. Faster R-CNN, despite its computational intensity, excels in precision, making it suitable for detailed post-processing analyses (Caputo *et al*., 2021).

SSD and RetinaNet offer practical advantages with faster processing speeds and reasonable accuracy, particularly beneficial in real-time applications. RetinaNet's focal loss function notably enhances its ability to detect small or obscured victims (Caputo *et al*., 2022). Segmentation models, including U-Net and DeepLab, provide precise pixel-level identification capabilities, critical for accurate victim localization, particularly in complex, cluttered environments. While computationally intensive, their use in drone imagery significantly increases detection reliability and operational effectiveness (Qu *et al.*, 2023).

The application of YOLO models, specifically YOLOv5 and YOLOv8, in SAR missions has garnered significant support from recent literature. These models are particularly advantageous due to their real-time performance, high accuracy, and adaptability across diverse operational conditions.

In terms of speed and efficiency, rapid image processing is critical in SAR missions, where every second can be pivotal. Bachir and Memon (2024) benchmarked YOLOv5 for human detection and demonstrated that its real-time processing capabilities substantially outperform traditional models like R-CNN. This speed is crucial when processing drone-captured images in dynamic rescue scenarios. Caputu *et al.* (2022) further illustrated that YOLO models are adept at handling the rapid influx of data from drones, ensuring that potential victims are detected quickly, thereby facilitating prompt rescue responses.

High recall in human detection minimizes the risk of false negatives, a vital factor in life-critical SAR operations. Bachir and Memon (2024) reported that YOLOv5 achieved a mean average precision of 96.9% in detecting humans, highlighting its superior accuracy compared to earlier models such as YOLOv4 and conventional detection frameworks. Rizk and Bayad (2023) provided evidence that YOLOv8 maintains an impressive average precision of 95% when applied to thermal imagery, an essential capability for identifying victims in low-visibility conditions often encountered in SAR missions.

Adaptability to various conditions is another key strength of YOLO models. SAR operations often occur in environments that can vary widely from urban settings to rugged natural landscapes. Kaur and Khandnor (2023) compared different YOLO variants using drone-based imagery and found that these models can be trained effectively on diverse datasets, including both aerial and thermal images. This adaptability ensures robust performance across different environmental conditions, which is crucial for the unpredictable nature of SAR missions.

Recent advancements in YOLO architecture have also expanded their applicability to more dynamic and challenging detection scenarios. Thangavel and Karuppannan (2023) explored an enhanced YOLO deep learning architecture tailored for dynamic event camera object detection and classification. Their findings underscore YOLO's potential to manage rapidly changing scenes, a feature that is particularly beneficial in SAR contexts where situational conditions are highly volatile.

2.9 Future Directions and Research Opportunities

Several promising directions exist for further enhancing drone-based SAR operations. Future research should prioritize developing more robust and computationally efficient artificial intelligence algorithms suitable for real-time drone deployments. Additionally, addressing operational challenges such as environmental constraints, regulatory limitations, and interoperability among multiple agencies will be crucial for seamless integration into existing SAR frameworks.

Further advancements in drone cluster coordination and task allocation algorithms, including enhanced adaptive clustering and market-based allocation methods, will significantly improve multi-drone operations. Emphasis should also be placed on the practical integration of advanced computer vision technologies, such as transformer-based models, into operational SAR scenarios. Collectively, these research directions offer substantial potential to further enhance the efficacy, safety, and overall success of SAR missions globally.

Using drones in SAR missions has been widely studied, as it is a matter of multiple fields. During the past few years, nine studies have been taking the problem of drones in SAR mission are included in Table 1.

Table 1. Previous Studies for SAR Using Drones and AI Methods.

Author(s) and Year	Focus/Topic	Key Findings	Methodology	Limitations/Challenges	Notes
Bachir and Memon (2024)	Human detection in search and rescue missions.	YOLOv5 achieved a mean average precision of 96.9% for human detection, outperforming previous YOLO versions (v4) and traditional detection frameworks in SAR scenarios.	Benchmarked YOLOv5 using various SAR datasets to assess its accuracy and performance in detecting humans in diverse environments, including low-light and dynamic settings.	The images must be taken at low altitudes	Highlights YOLOv5's superior accuracy and speed, making it an ideal choice for real-time SAR operations.
Qu *et al*. (2023)	Segmentation in nighttime UAV imagery	The study introduced NoctuDroneNet, a model for real-time semantic segmentation of nighttime UAV imagery, improving the detection of objects in low-light conditions using advanced segmentation models like DeepLabv3+ and SegFormer.	The authors applied DeepLabv3+ and transformer-based SegFormer models for semantic segmentation on nighttime UAV imagery to enhance object detection in complex environments.	The model's performance may degrade in environments with extreme darkness	Demonstrates the effectiveness of advanced segmentation models in SAR missions during nighttime operations or in environments with challenging lighting conditions.

Rizk and Bayad (2023)	Human detection in thermal images using YOLOv8 for search and rescue missions.	YOLOv8 achieved high accuracy (95% mAP) in detecting humans in thermal images, making it suitable for SAR missions in low-visibility conditions.	YOLOv8 applied thermal imagery for human detection in SAR operations, emphasizing its use in environments with limited visibility, such as smoke or fog.	The accuracy may decrease under extreme weather conditions or with low-quality thermal images.	Shows promise SAR missions in environments where visual detection is difficult, highlighting the value of thermal imagery for life-saving operations.
Kaur and Khandnor (2023)	Performance comparison of YOLO variants for object detection in drone-based imagery.	Compared various YOLO variants for object detection in drone-based imagery, with YOLOv5 and YOLOv4 showing the best performance for SAR scenarios.	Evaluated the performance of multiple YOLO models (v4, v5, v3) using drone imagery, focusing on speed, accuracy, and robustness in SAR tasks.	Limited to specific drone-based datasets, may not generalize well across different operational environments.	Emphasizes the importance of selecting the optimal YOLO variant based on the mission requirements, such as image resolution and environmental factors.
Thangavel *et al.* (2023)	Dynamic event camera object detection and classification using enhanced YOLO deep learning architecture.	Enhanced YOLO architecture improves object detection in dynamic, high-motion environments, crucial for SAR operations.	Used dynamic event cameras with an enhanced YOLO model for real-time object detection in rapidly changing scenes.	dynamic camera setups require specialized hardware and may introduce operational complexities.	The study highlights the importance of adapting YOLO to dynamic detection environments for SAR applications, particularly in

					volatile or high-speed settings.
Alsumayt *et al*. (2023)	Drones in SAR for flood detection	Deploy drones with high-resolution cameras and DeepAL in federated learning, integrated with blockchain for data security.	Deep active learning, federated learning, blockchain technology	Unauthorized drone use, weather effects on image quality, best in morning	Flood detection and monitoring focus
Caputo *et al*. (2022)	Human detection in drone images using YOLO for SAR.	YOLOv5 outperformed YOLOv4, Faster R-CNN, and RetinaNet on SAR datasets, demonstrating superior performance in detecting humans in aerial imagery.	Evaluated YOLOv5 against YOLOv4, Faster R-CNN, and RetinaNet using various SAR datasets, focusing on accuracy, speed, and real-time detection capabilities.	The study did not address the impact of environmental factors (such as weather conditions) on detection accuracy.	YOLOv5's superior performance highlights its suitability for real-time human detection in SAR operations, offering a faster and more accurate alternative to other detection models.
Calamoneri *et al*. (2022)	UAV routing post-earthquake	Comparison of CMP problem with CVRP and TSP, focusing on UAVs in SAR.	Routing algorithms comparison	Homogeneous UAVs only, no diverse UAV models	Emphasis on multiple cycles and priority parameters

Daud *et al*. (2021)	Drones in disaster management	Challenges in the adoption and use of drones and machine learning in disasters.	Observational/Analysis	Lack of mass disaster evidence, international agreements, photogrammetry quality, weather conditions, restricted zones, no standardized drone specs	Broad focus on disaster management
Sruthi, (2021)	YOLO.v5 for human detection in SAR	YOLOv5 for reduced false rates in object detection, using a human-centric dataset.	Machine learning with YOLO.v5	Not specified	High mAP and F1 score, emphasis on real-time application
Yamazaki *et al*. (2020)	Human voice detection in disasters	Array microphone on UAVs for voice recognition in obscured locations.	Sound source separation processing, voice recognition	Not specified, no comprehensive detection rate mentioned	High voice recognition rate in varied settings
Jiménez *et al*. (2020)	Damage assessment post-flood	High-resolution imagery and DEMs for pre- and post-event house classification.	Automated image classification, DEM analysis	Did not consider structural damage	High accuracy in classifying washed-away or destroyed houses
Lygouras *et al*. (2019)	Human detection in open water	Novel dataset for training an embedded vision system on UAV for human detection.	Deep learning with CNN, SVM	Not specified	High accuracy in real-time human detection
Tariq *et al.* (2018)	DronAID for human detection	Real-time autonomous drone technology for	Embedded system with sensors, PIR technology	Specific results and findings not detailed	Focus on urban disaster scenario

		human detection in disa+sters.			
Bejiga *et al.* (2017)	Avalanche SAR missions	UAVs with optical cameras and CNNs for victims' location in avalanches.	CNNs, SVM, HMMs	Need for more datasets, pre-processing thresholding	Accuracy varies with resolution, effective in precision.

Chapter Three

3 System Design and Drone Route Generation

In this chapter, we detail the architecture and key algorithms that drive our proposed drone-integrated SAR system. We begin by outlining the overall proposal for system design. Next, we explain how the operational elevation for the drones is calculated to make sure we obtain suitable data, then the area is segmented into manageable zones for autonomous coverage. We then calculate the drone cluster size, we then introduce our routing-generation strategy, which gives the directions for each drone on which way point it needs to follow.

3.1 Proposed System Design

The proposed SAR system begins by detecting a disaster event and then drawing and segmenting the Area of Interest (AOI). The needed number of drones is automatically calculated, and a suitable launch station is decided, after which detailed flight routes are generated and drone clusters are deployed. As the drones scan the AOI, they continuously stream high-resolution imagery to a centralized command-and-control unit, where a YOLO-based AI model processes each frame in real time to pinpoint human. Once individuals are detected, ground troops are dispatched immediately to those locations, enabling a rapid, data-driven rescue response that maximizes both coverage and efficiency. These components are described in the following paragraphs and shown in Figure 1.

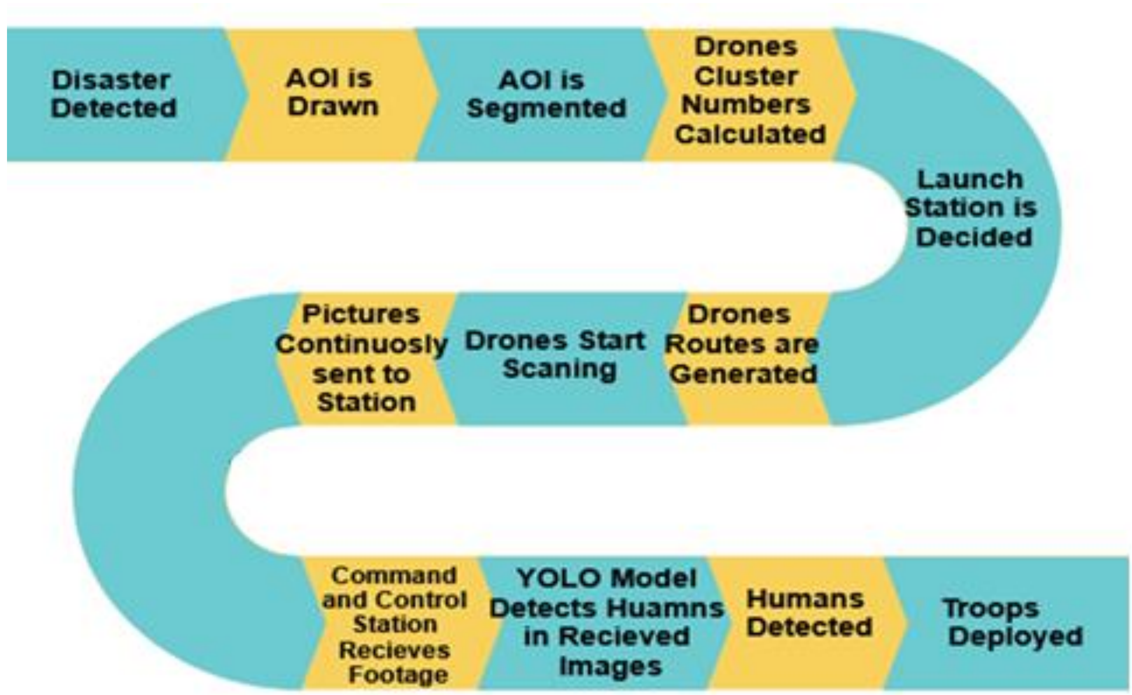

Figure 1. Flowchart Demonstrating the Proposed Solution

The stages of the overall proposed system are broken into main components listed below:

Disaster Detection: The system initiates upon the detection of a disaster, which can be either automatically triggered through connected sensors or manually reported by humans. Once detected, the coordinates of the affected area are manually input into the system.

Drawing the AOI: Once a disaster is detected, the command-and-control unit quickly synthesizes all available satellite intelligence imagery, sensor feeds, eyewitness reports, and pre-existing Geographic Information System (GIS) data to define the precise boundaries of the AOI. Leveraging mapping software and overlay tools, operators draw a polygon that encompasses all affected zones, critical infrastructure, and likely survivor locations. This carefully calibrated AOI ensures that subsequent segmentation and route-planning stages have clear geographic constraints, maximizing search coverage and minimizing wasted flight time.

Area Segmentation: The area then is divided into smaller scanning zones. Based on parameters such as required scan resolution and total scan duration.

The Needed Number of Drones: To determine the size of each drone cluster, the system first computes the total area of the AOI and divides it by the estimated coverage area of a single drone. A drone's coverage area is derived from its flight characteristics specifically, its cruising speed and maximum flight time set an upper limit on total distance flown, and its footprint which is the area covered in one image. Then divides the AOI area by that footprint to find an initial guess for the cluster count.

Drone Routes Generation: Once drone clusters are sized, the route of each drone is generated automatically using a "lawnmower" routing technique. For each drone, the algorithm lays down parallel sweep lines spaced exactly by the drone's footprint, then threads them together into a continuous back-and-forth pattern that guarantees full coverage. Each drone's starting point is the final way point of the previous drone, which makes this routing technique easily scalable and implemented. Finally, the individual segment routes are stitched into a global mission plan so that each drone knows exactly which grid cells to traverse and in what order ensuring the entire AOI is scanned efficiently.

Drones Deployed and Start Scanning: Once the drones' routes are generated, the drones are concurrently launched either manually (*e.g.*, by remote pilots) or autonomously, with mission routes uploaded to drones using tools such as FreeFlight 6, allowing for pre-defined waypoints and adaptable flight parameters. Each drone follows a unique path determined by the segmentation algorithm, ensuring time-efficient and energy-aware coverage (Parrot SA 2021).

Pictures Streaming: As each drone follows its assigned lawn-mower path, its onboard camera continuously captures high-resolution imagery of the terrain. Using a preconfigured transmission schedule. Images are sent in real time to maintain link reliability. This steady stream of geo-tagged photographs flows into the ground-station, ensuring the command-and-control unit receives footage of the AOI.

YOLO AI Model: Once geo-tagged imagery arrives at the command-and-control, each frame immediately feeds these images into a trained YOLO model that outputs bounding boxes and confidence scores for any detected humans in real time. Detections exceeding a preset confidence threshold are overlaid on the master AOI map with precise coordinates, allowing operators to instantly visualize survivor locations. These streams are processed locally at the central control center to reduce latency and improve responsiveness.

Troops Deployment: After the AI detects potential human instances, command and control personnel review each detection to confirm accuracy. Once human presence is validated, the Command-and-control unit translates the mapped coordinates into actionable orders, assigning and mobilizing the nearest available troop units.

3.2 Drone Height Calculation

Determining the optimal flying altitude for a drone is critical for achieving the required image resolution and coverage area in aerial missions. For SAR applications, this means balancing the need for high-resolution imagery with the ability to monitor broad terrain. This section uses the specifications of the ANAFI USA drone to calculate the ideal operational altitude using geometry equations.

Drone Specifications: The ANAFI USA drone is a lightweight surveillance drone. The main specifications of the drones are listed below (Parrot Drones S.A.S., 2023):

- **Drone Mass**: 500 g
- **Maximum Payload Mass**: 144 g
- **Maximum Speed**:15 m/s
- **Flight Time**: Up to 35 minutes on a full charge
- **Maximum Elevation**: 5,000 m above mean sea level
- **Flight Conditions**: Weather: no fog; wind ≤ 15 m/s, avoid overwater or other highly reflective surfaces, and can operate in rain (IP53 rain- and dust-resistant).
- **Operating Temperature**: –36 °C to +50 °C

Camera Specification: The ANAFAI USA drone is equipped with a high-resolution dual Electro Optical (EO) camera with the following specifications:

- **Sensor Size**: 1/2.4-inch (sensor height ≈ 7.3 mm, adjusted to reflect observed coverage)
- **Resolution**: 3,840×2,160 pixels (4K Ultra High Definition)
- **Focal Length (f)**: ~4.34 mm (empirically estimated from practical flight data)
- **Horizontal Angle of View (α_h)**: 75° (for wide-angle camera)
- **Zoom Factor (X)**: Variable (set to 1x for calculations)

Calculating Optimal Altitude Based on GSD: For SAR operations, based on the literature GSD of 1.4 cm/pixel is optimal to detect human-sized objects (*e.g.*, 40 cm shoulder width). With a 3,840-pixel horizontal resolution (N_h), with the Horizontal Field Of View (HFOV) the total ground coverage required. Figure 2 demonstrates the geometry of how the elevation of the drone is calculated based on the drone's camera specification and based on the desired GSD with the resolution of the camera.

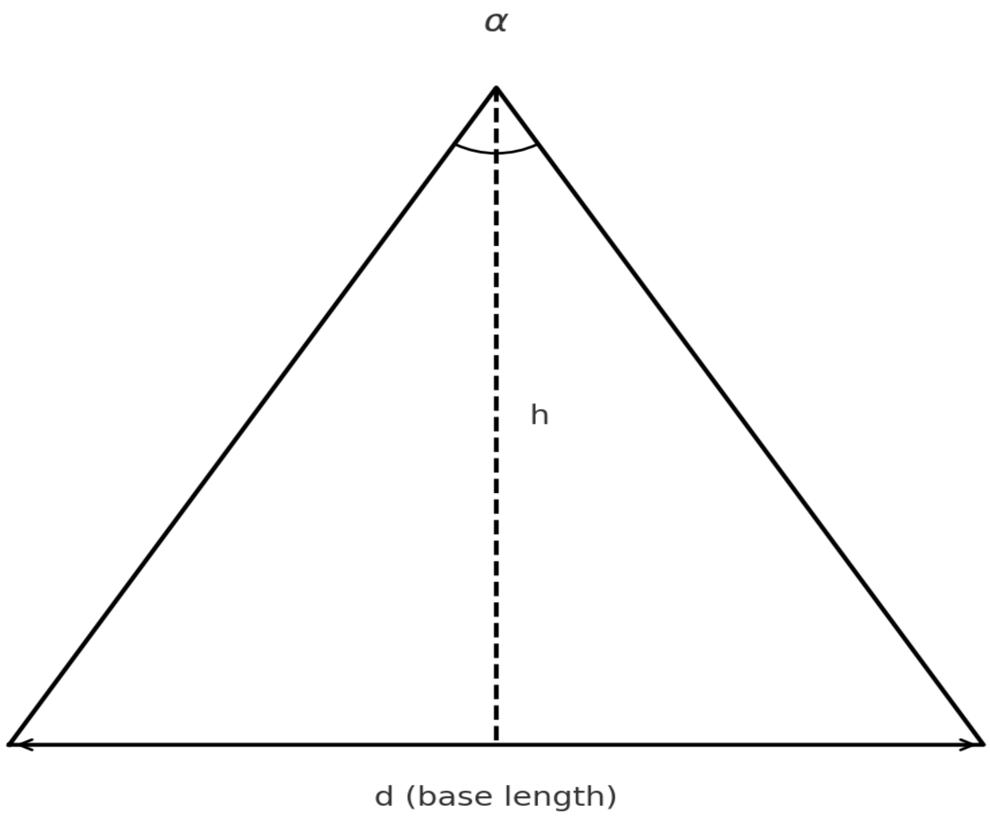

Figure 2. The Relationship between Altitude (H), Field of View Angle (A), and Ground Coverage (d)

Base Calculation (d): From this figure, we start by finding the base length based on the desired GSD, the chosen value is to be 0.014 m, as chosen from the literature.

$$d = \text{GSD} \times N_h \tag{1}$$

$$d = 0.014 \times 3{,}840 \approx 54\,m \tag{2}$$

The Drones Highet Calculation (h): To find the altitude (h) that corresponds to this horizontal field of view given the camera's angle:

$$h = X \times d/(2\ \tan\frac{\alpha_h}{2}) \tag{3}$$

$$h = 1 \times 54/(2\ \tan\frac{75}{2}) \approx 35\,m \tag{4}$$

Field of View Calculations: With an aspect ratio of 16:9, the vertical and diagonal fields of view can be derived:

- HFOV: The calculations of (d) above represent the HFOV which make is 54 m.
- Vertical Field of View (VFOV): The VFOV is calculated using the picture ration as 16:9.

$$\text{VFOV} = \frac{9}{16} \times \text{HFOV} \tag{5}$$

$$\text{VFOV} = \frac{9}{16} \times 54 \approx 30\ \text{m} \tag{6}$$

Drone's Footprint Calculations: The total ground area covered by one image is calculated by the area rectangle made by the VFOV and HFOV:

$$\text{Done‘s Footprint} = \text{HFOV} \times \text{VFOV} \tag{7}$$

$$\text{Done‘s Footprint} = 54 \times 30 = 1{,}620\ \text{m}^2 \tag{8}$$

3.3 AOI Drawing and Loading

After a disaster is detected, we propose that the command-and-control personnel rapidly delineate the full extent of the affected area known as the AOI so that search-and-rescue, relief deliveries, and damage assessment teams can be dispatched efficiently. To do this, operators typically fuse multiple sources such as satellite intelligence imagery, sensor feeds, eyewitness reports, and pre-existing GIS data.

Once the geospatial context is assembled, the AOI polygon is drawn interactively using a web-based mapping tool. In our pilot study we selected geojson.io an open-source, browser-based editor that lets users trace arbitrary polygons (or other geometries) atop a map interface and export the result directly as a GeoJSON file. For the purposes of this pilot, we simulated an AOI around Al-Dhlail Village in Zarqa, Jordan and area of 17.5 km^2, and this delineated area was subsequently used to collect both satellite and ground-truth data for training our AI model. Figure 3 illustrates the resulting AOI polygon as drawn in geojson.io for Al-Dhlail Village.

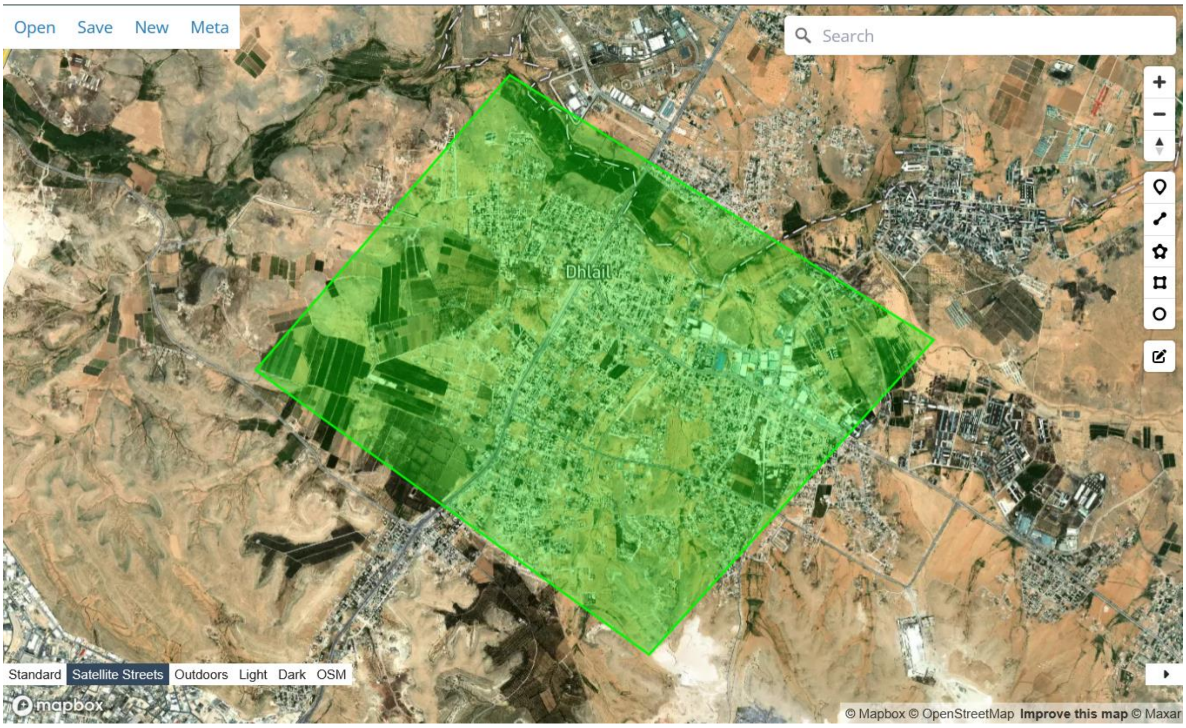

Figure 3. AOI Polygon for the Town of Dhlail Created in GeoJSON.IO Illustrating the Interactive Delineation of the AOI Used in our Pilot Study

We chose GeoJSON because its human-readable JSON structure seamlessly bridges the gap between the interactive mapping tool and the Python-based analytics pipeline, enabling us to inspect, version control, and transmit geospatial boundaries with minimal efforts. As a widely adopted open standard, GeoJSON integrates effortlessly with web-mapping libraries and GIS software alike, while Python libraries such as GeoPandas and Shapely parse and manipulate its geometric objects natively. This compatibility allowed us to streamline the transition from AOI drawing to data ingestion without writing custom parsers, ensuring that our disaster-response workflow remains agile, transparent, and fully reproducible.

3.4 The Drone Cluster Size Calculation

To size the minimum drone cluster required to survey the AOI (N_{min}), we begin by extracting the total AOI area, (A_{AOI}), from our GeoJSON definition. Next, we calculate the maximum area one drone can cover in a single full flight time, (A_{max}). Each nadiroriented frame from our drone's camera yields an independent ground footprint of approximately 1,620 m². By flying at the cruise speed (s) of 15 m/s and triggering the shutter at intervals equal to VFOV = 30 m, the drone captures a non-overlapping frame (Δt) every 2 seconds as shown in the equation below:

$$\Delta\mathrm{t} = \frac{\mathrm{VFOV}}{s} \tag{9}$$

$$\Delta\mathrm{t} = \frac{30}{15} = 2\ s \tag{10}$$

Find the Effective Area Coverage Rate ($\dot{A}_e$): The area covered by every non-overlapping image is calculated in section 3.2 as the drone's footprint which is equal to 1,620 m^2. We find the effective area-coverage rate($\dot{A}_e$) by the following equation:

$$\dot{A}_e = \frac{\text{Drone's Footprint}}{\Delta\mathrm{t}} \tag{11}$$

$$\dot{A}_e = \frac{1{,}620}{2} = 810\ m^2/s \tag{12}$$

The Drone's Maximum Possible Area covered (A_{max}): The maximum area that one drone can survey in a flight of endurance "full flight time" (T) is therefore given by multiplying the effective area-coverage rate by T as shown in the equation:

$$\mathrm{A}_{max} = \dot{A}_e \times T \tag{13}$$

$$\mathrm{A}_{max} = 810 \times 2{,}100 = 1.7\ km^2 \tag{14}$$

The Minimum Size of Drone Cluster Needed (N_{min}): To translate this capability into the number of drones, we divide the total AOI by the per-drone capacity and then round up, since we cannot deploy a fractional drone. This yields our baseline cluster size as shown in the following equation:

$$N_{min} = \frac{A_{AOI}}{A_{max}} \tag{15}$$

$$N_{min} = \frac{17.5}{1.7} = 10.3 = 11\ drones \tag{16}$$

This first-order estimate does not account for transit legs to and from the launch station, nor the effects of complex terrain or safety margins; those practical considerations are incorporated later via our route-generation algorithm, which refines both the required fleet size and the detailed mission plan.

3.5Route Generation Algorithm

In this section, we present the methodology for optimizing drone routes for aerial surveying of a defined AOI taking the pilot area as a sample. The goal is to ensure complete coverage of the AOI while minimizing overlapping and interference between drones. The primary algorithm employed is a variation of the "lawnmower" or zigzag pattern, which is well-suited to systematic and exhaustive area coverage.

3.5.1 Algorithm Overview

Once the AOI has been defined and the baseline fleet size N_{min} computed, we invoke our route-generation algorithm to generate time-aware, gap-free flight plans for each drone. The process begins by loading the AOI GeoJSON file via GeoPandas, at which point the script reports both the polygon's area and its perimeter for verification. The operator then supplies the drone's key specifications, namely the HFOV, which represents the spacing between lawnmower swaths; the cruise speed; the flight altitude; and the maximum flight time per sortie.

Next, the launch station is chosen either from one of several automatically computed options (centroid or midpoint of an edge) or by entering custom coordinates. With the polygon and station in hand, the algorithm computes a "lawnmower" sweep path: it slices the AOI with parallel lines spaced according to the HFOV, then orders the resulting waypoints in a back-and-forth sequence.

The heart of the optimizer is the mission partitioning routine. Given the full waypoint list, each drone is assigned a contiguous segment under its time budget T. Drone 1 begins at the station, flies out to the furthest waypoint it can reach, executes its assigned swath, and returns—ensuring its total outbound, coverage, and return legs do not exceed T. Drone 2 then launches from the station, heads directly to the last coverage waypoint of Drone 1, continues along the path, and so on. This sequential handoff guarantees continuous coverage: every new sortie picks up where the previous one left off, stitching together a complete survey with no gaps while respecting each drone's endurance.

Finally, the planner produces combined and per-drone route figures, CSV summaries of distances and times, and where possible merges the last two missions if their combined duration remains within the time limit. A full listing of the Python implementation, including functions for reprojection, spacing computation, path generation, mission partitioning, plotting, and export, is available in the Github repository (Alkilani, 2025a).

3.5.2 Pilot AOI Results and Sample Output

In this subsection, we apply the route-generation algorithm to our pilot AOI around Al-Dhlail Village in Zarqa, Jordan, and present the resulting flight plans and performance metrics. The optimizer was executed with a HFOV of 54 m (defining our lawnmower sweep spacing), a cruise speed of 15 m/s, a flight altitude of 35 m, and a maximum sortie duration of 2,100 s (35 min) per drone. We loaded the AOI GeoJSON into GeoPandas where it was reported to cover approximately 17.5 km^2 with a 16.9 km perimeter—and selected the launch station at the polygon's bottom edge. Under these parameters, the algorithm assigned 16 drones in sequence, each flying out to its furthest waypoint, executing its swath, and returning within the 35-minute time budget, thereby achieving a 100%, gap-free coverage of the AOI.

Figure 4 shows how we verify that the AOI has been correctly loaded and projected into a metric coordinate reference system before generating the routes. The script reports both the total area and perimeter of the polygon and then renders its boundary to confirm the shape and extent. This step ensures that subsequent sweep-path generation and distance calculations are based on the accurate geometry of Al-Dhlail Village.

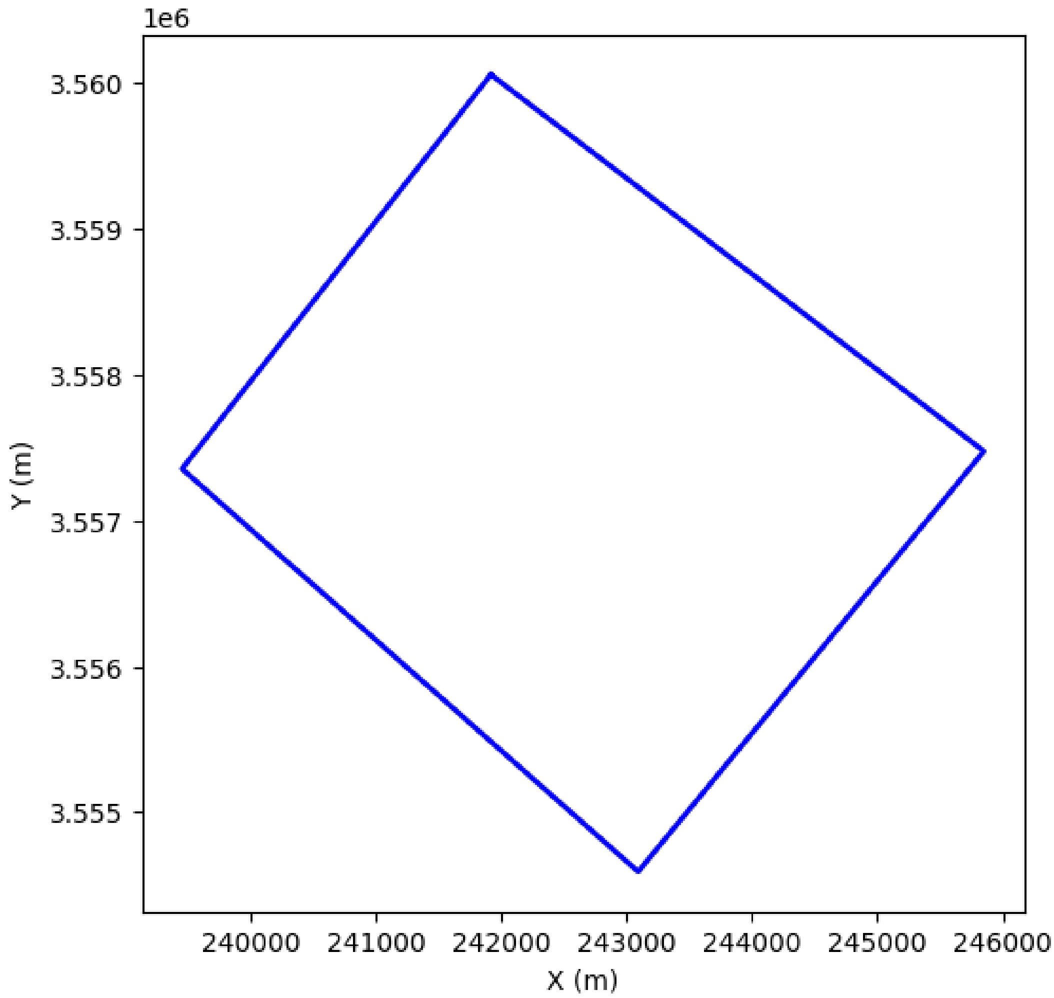

Figure 4. Boundary of Al-Dhlail Village AOI as Loaded from GeoJSON in Python, Showing an Area of 17.6 km² and a Perimeter of 16.9 km

The routes shown in Figure 5 display the complete set of routes flown by each drone over Al-Dhlail village AOI, color-coded by vehicle ID and anchored at the single launch station (red star). The horizontal “lawnmower” stripes represent the scanning segments, and one can see that together they achieve 100% area coverage with virtually no overlap between adjacent swaths during the coverage phase. Any crossing of lines occurs only during the outbound or returns legs. Moreover, because each drone is assigned an approximately equal-length portion of the primary path, the area and thus the workload of each segment is closely matched across the fleet, ensuring a balanced mission profile.

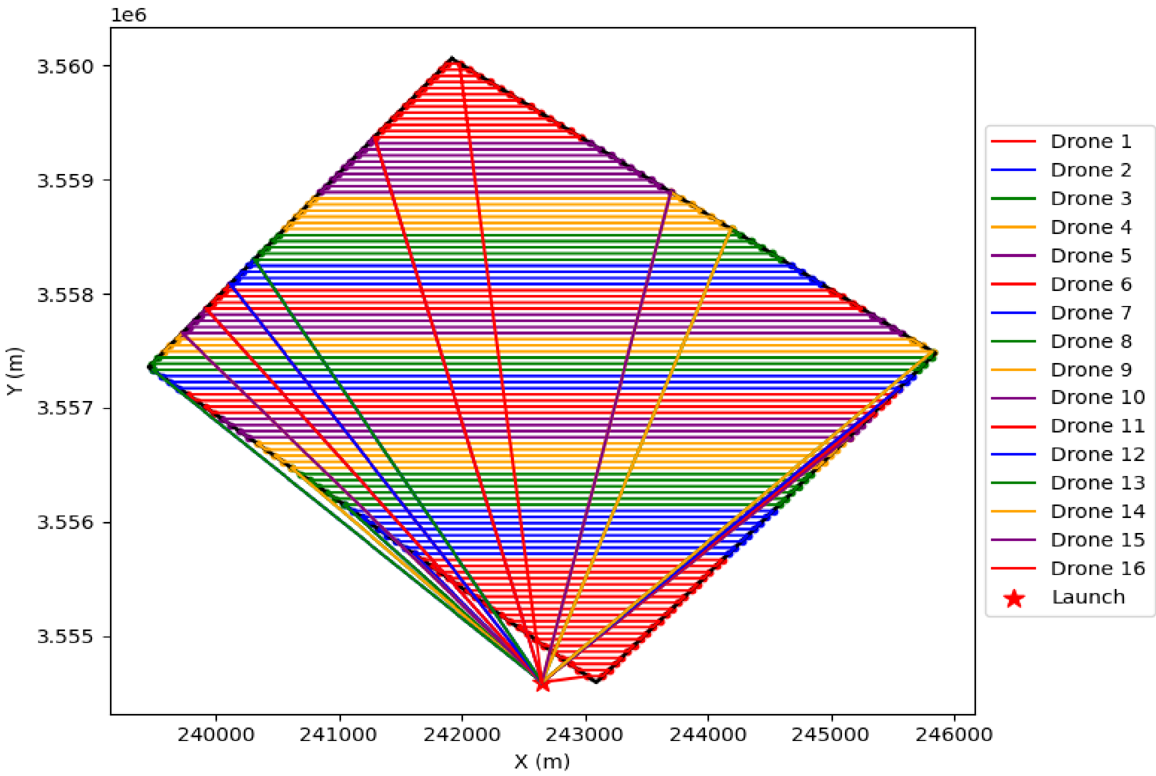

Figure 5. Combined Flight Tracks for all Drones over Al-Dhlail AOI Demonstrating Gap-Free Coverage with Minimal Overlap in the Scanning Segments

Figure 6 shows the individual mission details for Drone 10 and Drone 11, illustrating the sequential handoff in action. Drone 10 departs the launch station, flies out to Drone 9 final way point. performs its swath coverage and returns. Then the journey of Drone 11 routes directly to Drone 10’s final waypoint before executing its own sweep and returning, thus ensuring continuous, gap-free coverage.

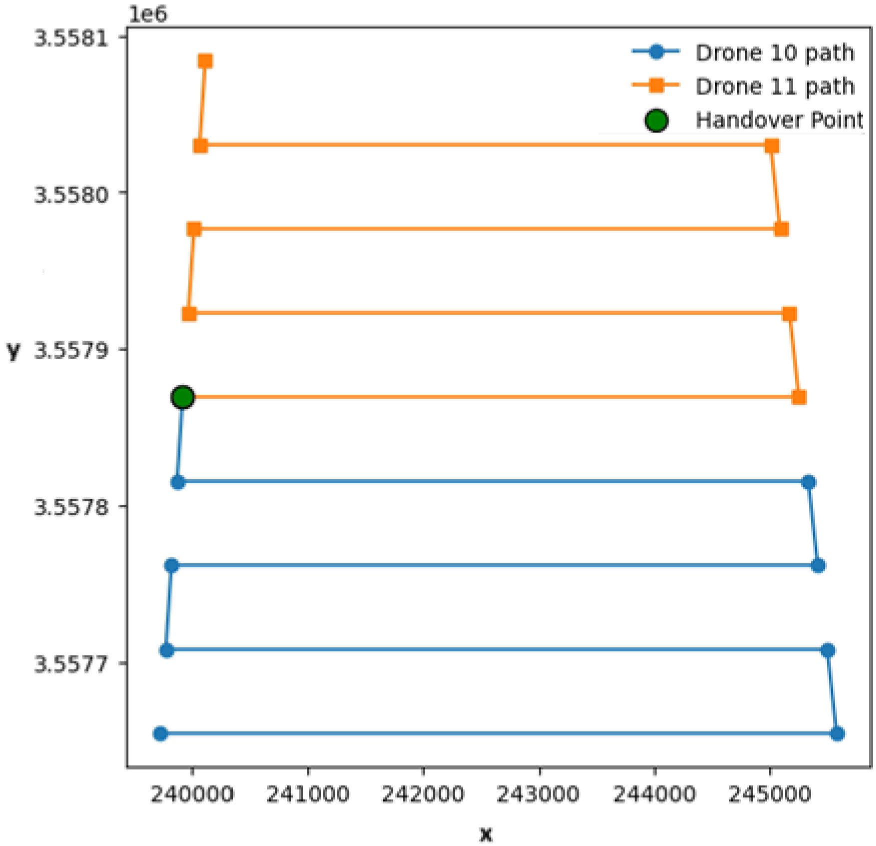

Figure 6. Sequential Mission Handoff Between Drone 10 and Drone 11

To further illustrate the sequential handoff and the nearly identical swath geometry between successive sorties, Table 2 lists the waypoint coordinates for Drone 10 and Drone 11. Notice that Drone 11's first waypoint matches Drone 10's last, and both follow a back-and-forth pattern across the same swath lines.

Table 2. Waypoint Coordinates for Drone 10 and Drone 11

Waypoint Number	Drone 10		Drone 11	
Start	239,726	3,557,654	**242,651**	**3,554,592**
2	245,579	3,557,654	239,921	3,557,869
3	245,498	35,577,07	245,252	3,557,869
4	239,774	35,577,07	245,171	3,557,922
5	239,823	3,557,761	239,970	3,557,922
6	245,416	3,557,761	240,019	3,557,976
7	245,334	3,557,815	245,089	3,557,976
8	239,872	3,557,815	245,007	3,558,030
9	239,921	3,557,869	240,068	3,558,030
Finish	**242,651**	**3,554,592**	240,117	3,558,083

Table 3 is a comprehensive summary of each drone's mission metrics for the Al-Dhlail Village survey, including outbound, coverage, and return distances; total flight distance; mission duration; and the area covered per sortie. The table also reports the maximum, minimum, average, and standard deviation for each metric, illustrating the fleet's overall workload balance and performance consistency.

Table 3. Drone Mission Specifications for Al-Dhlail AOI, Detailing Per-Drone Distances, Time, and Surveyed Area, Along with Summary Statistics (Max, Min, Average, Standard Deviation) Across all Sorties.

Drone ID	Outbound (km)	Coverage (km)	Return (km)	Total (km)	Time (min)	Area (km^2)
1	0.37	27.26	1.54	29.17	32.41	1.46
2	1.54	24.57	2.24	28.35	31.50	1.32
3	2.24	23.56	2.77	28.57	31.74	1.27
4	2.77	22.98	3.29	29.04	32.27	1.23
5	3.29	20.59	3.59	27.47	30.52	1.11
6	3.59	22.54	3.88	30.01	33.35	1.21
7	3.88	18.19	4.19	26.26	29.18	0.98
8	4.19	18.99	4.30	27.48	30.53	1.02
9	4.30	18.59	4.23	27.12	30.14	1.00
10	4.23	22.97	4.26	31.47	34.97	1.23
11	4.26	20.88	4.31	29.46	32.73	1.12
12	4.31	18.79	4.38	27.49	30.54	1.01
13	4.38	20.56	4.26	29.20	32.45	1.10

14	4.26	20.35	4.42	29.03	32.26	1.09
15	4.42	21.68	4.97	31.08	34.53	1.16
16	4.97	12.57	5.47	23.01	25.56	0.68
Maximum	4.97	27.26	5.47	31.47	34.97	1.46
Minimum	0.37	12.57	1.54	23.01	25.56	0.68
Average	3.56	20.94	3.88	28.39	31.54	1.12
Standard Deviation	1.20	3.20	0.97	1.94	2.16	0.17

A quick examination of the summary statistics reveals several key insights into our fleet's performance and workload balance:

Coverage Consistency: The average leg coverage per drone is about 20.94 km, with a standard deviation of 3.20 km (≈15% of the mean), indicating that most drones survey a similar extent of the AOI. The smallest coverage mission was 12.57 km which is necessary as it is the last deployed drone and covered the only remaining area, while the largest reached 27.26 km, reflecting how the algorithm partitions longer central swaths among fewer drones.

Outbound/Return Symmetry: On average, outbound and return legs are nearly identical (3.56 km *vs*. 3.88 km), and their standard deviations (1.20 km outbound, 0.97 km return) are comparable. This symmetry underscores the sequential handoff logic each drone starts where the last one finished resulting in balanced transit distances.

Total Mission Duration: Missions last on average 31.54 minutes, with a standard deviation of 2.16 minutes (≈7% of the mean). Even the longest sortie (34.97 minutes) stays within the 35 minutes demonstrating robust time-budget adherence.

Area Coverage Distribution: Each drone covers on average 1.12 km^2, with a tight spread a standard deviation of 0.17 km^2. The smallest mission covers 0.68 km^2 and the largest 1.46 km^2, confirming that no single drone is overly burdened or underutilized. Summing each sortie's surveyed area yields approximately 18.0 km^2, which matches the 17.55 km^2

AOI plus a small extra margin arising from the outbound and return legs that extend beyond strict swath boundaries. Overall, the fleet's combined effort not only respects individual endurance limits but collectively achieves complete and slightly redundant coverage, enhancing the robustness of our survey.

Overall, these statistics show that the route optimizer produces a well-balanced set of sorties: coverage distances and mission durations cluster closely around their means, ensuring an equitable distribution of work across the drone cluster while respecting each platform's endurance limits.

3.6 Summary

In this chapter, we developed a fully integrated, drone-based SAR workflow beginning with rapid AOI delineation and ending with balanced, time-aware flight plans that guarantee 100% coverage.

First, we described our proposed system architecture: disasters are detected via sensors or reports, the affected zone is drawn as a GeoJSON polygon in geojson.io (our pilot AOI was Al-Dhlail Village, 17.55 km^2), and drone operational parameters (flight altitude, speed, HFOV/VFOV) are calculated to meet target image resolution (GSD = 1.4 cm/pixel) and coverage requirements.

Next, we sized the drone cluster by deriving a per-drone coverage rate (810 m^2/s) from the 1,620 m^2 image footprint and a 2 s trigger interval, then dividing the AOI by each drone's 1 800 s endurance to estimate an initial fleet of 11 drones.

To turn that estimate into actionable missions, we implemented a Python route-optimizer that reprojects the AOI, then lays down "lawnmower" swaths spaced by the HFOV, then partitions the master waypoint list so each sortie starting at the launch station, sweeping its assigned segment, and returning respects the time limit, and finally hands off directly to the next drone for seamless coverage.

Applied to our pilot (HFOV = 30 m, speed = 15 m/s, endurance = 2,100 s, launch at the bottom edge of the AOI, the planner generated 16 sequential sorties that collectively surveyed 18.0 km^2 matching the AOI plus transit overhead with per-drone mission distances averaging 28.4 km (σ = 1.94 km), durations 31.5 min (σ = 2.2 min), and areas 1.12 km^2 (σ = 0.17 km^2).

These results demonstrate that our end-to-end pipeline from GeoJSON-based AOI loading through analytical cluster sizing and automated, endurance-aware routing yields a balanced, reproducible, and robust rapid-response survey solution efficient for time-critical SAR operations.

Chapter Four

4 Data Collection

This chapter provides a comprehensive overview of the planning, execution, and processing steps undertaken to assemble the dataset central to this research. It begins by detailing the coordination of two field trips in collaboration with the Jordan Design and Development Bureau (JODDB), including site selection criteria, logistical considerations, and drone operation protocols across twelve representative locations in Al-Dhlail, Zarqa.

Next, it describes the technical specifications of the collected imagery and summarizes the distribution of human instances and background elements. Finally, the chapter outlines the exhaustive manual annotation workflow conducted.

By documenting data collection methodology, dataset characteristics, and annotation procedures, this chapter establishes a clear foundation for subsequent model training and evaluation and underscores the dataset's relevance to Jordan's unique environmental context.

4.1 Motivation for Data Collection

Throughout the preliminary literature review and dataset reconnaissance phase, we identified a critical gap: there exists no publicly available image dataset tailored to support SAR training in Jordan's unique environmental and regulatory context. The only remotely similar dataset was the UAV123 benchmark by Müller *et al*. (2016), collected in Saudi Arabia. However, UAV123 features drone flights at low altitudes (5–10 m) optimized for object tracking rather than SAR, and human instances appear in wide-open fields, making it ill-suited for high-altitude, top-down SAR applications where subtle terrain features and occlusions are common.

Other global SAR-focused datasets were also considered, but their terrain types, lighting conditions, and color contrasts (*e.g.,* snow-covered landscapes, dense forests, or temperate weather) diverge significantly from Jordan's arid desert environments, leading to poor model transferability. Additionally, differences in vegetation color, ground texture, and atmospheric haze further challenge adaptation of these datasets to Jordan.

One prominent example is the Lacmus Drone Dataset (LADD), developed by search-and-rescue volunteer organizations to locate missing individuals in forested areas. LADD comprises over 400 high-resolution images captured by DJI Mavic Pro and Phantom drones at altitudes between 50 m and 100 m. The dataset includes 24 typical missing person poses derived from extensive consultations with rescue operators. While conceptually aligned with SAR needs, LADD's focus on snow-covered forest terrain and higher-altitude imagery impairs human detectability against Jordan's desert backdrop and subtle color contrasts, making it suboptimal for direct transfer to Jordanian SAR applications. Additionally, differences in vegetation color, ground texture, and atmospheric haze further challenge adaptation of these datasets to Jordan (Shuranov *et al.,* 2023).

Moreover, the strict security restrictions imposed by the Jordanian Civil Aviation Regulatory Commission (CARC) severely limit drone operations, especially over sensitive areas, making in-country data collection both logistically and legally complex. The combination of environmental specificity and regulatory constraints underscored the necessity of generating a bespoke, high-resolution Jordanian dataset to ensure robust SAR model performance.

4.2 Data Collection

To capture a dataset reflective of Jordan's unique SAR environments, two dedicated field expeditions were meticulously planned and executed. This section outlines collaborative planning with JODDB site selection criteria and permissions, flight scheduling, and aerial data acquisition protocols that ensured comprehensive coverage across diverse terrain types and human activity scenarios. Data acquisition was performed in Al-Dhlail, Zarqa, Jordan, chosen for its representative diversity, realistic SAR environments, and accessibility for drone operations (Alkilani, 2025b).

Site Selection: Before selecting individual sites, we established explicit criteria to ensure comprehensive coverage and SAR relevance:

Terrain Diversity: Include sandy, rocky, vegetated, and built-up environments to capture varied ground textures and occlusion scenarios.

- **Human Activity Types**: Represent both static (*e.g.*, seated, lying) and dynamic (*e.g.,* walking, working) human behavior in realistic contexts.
- **Accessibility and Safety**: Ensure logistical feasibility for equipment transport and compliance with operational safety guidelines.
- **Regulatory Compliance**: Avoid no-fly zones and adhere to altitude and flight-path restrictions imposed by CARC.
- **Representative SAR Scenarios**: Mimic common SAR contexts, such as agricultural fields, industrial yards, residential clusters, and open desert expanses.

Flight Arrangements and Locations: The dataset comprises 2,430 images acquired over two organized flight days covering twelve distinct locations. Each location was carefully chosen to maximize environmental variability and human scenario representation. Below are the detailed descriptions:

Table 4. Description of the Twelve Data Collection Locations in Al-Dhlail, Zarqa, Jordan

Video Name	Location	Humans in the video	Duration	Useful images	Altitude	Split
Bricks Workshop	A brick workshop, in a residential area with people sitting and playing around it.	Organic	1:34	47	30-35	Test/Valid
Farm1	A Farm in the middle of the desert with short crops, and some workers in it	Organic	4:16	240	30-35	Train
Farm2	A vast level ground, covered with short green crops.	Organic	5:30	261	30-35	Train
Farm3	A big farm with large trees in a desert environment	Actors	3:03	165	40-60	Train/Valid 50:50
Freezone 1	This video covers a construction site, where workers are conveying stuff from one warehouse to another.	Organic	5:35	280	30-35	Train
Freezone 2	A forsaken site, with a lot of shattered stuff on the ground with minimal humans.	Organic	3:39	132	30-35	Train
Freezone 3	A desert terrain with an actor simulating different possible poses of humans in the case of disasters	Actors	4:47	170	30-35	Train

Industrial Street	A long street, with different vocational workers, mechanics blacksmiths and carpenters. Different poses for all humans	Organic	5:38	238	30-35	Train
Open Field	A wide soft sand area, piles of rocks and pipes, multiple poses for the humans In it	Actors	4:36	207	30-36	Test/Vali 50:50
Residential	A wide vast farm, within a residential area. Multiple terrains with short, medium, and high crops, with plastic houses and water reservoir.	Organic and actors	5:35	256	40-45	Train/Valid 50:50
Souq 1	A main street with a lot of merchants, and a lot of shops,	Organic	4:37	242	30-35	Train
Souq 2	Another street, on the other side of the city, with merchants and shops.	Organic	5:24	192	30-35	Test/Valid 50:50
Total			**54:14**	**2,430**		

Dataset Specifications: The final dataset comprises 2,430 high-resolution images extracted from twelve videos. Recorded at 24 frames per second, each frame possesses a 4K resolution (3,840 × 2,160 pixels) and was captured predominantly as a top-down view. Drone elevations ranged between 30 and 60 meters, with approximately 82% of images taken at 30–35 meters. Across the entire dataset, 2,831 human instances were annotated. To facilitate model training, we defined two classes: “Human” which includes both organic and staged poses (*e.g.*, lying down, sitting), and "Background" covering natural terrain, vegetation, buildings, vehicles, and other non-human elements.

Dataset Splitting: To facilitate robust model training and evaluation, we partitioned the dataset into three subsets: approximately 70% (around 1,771 images) for training, 20% (about 437 images) for validation, and 10% (roughly 222 images) held out for testing. This 70 / 20 / 10 split ensures sufficient data for pattern learning while preserving ample unseen samples for hyperparameter tuning and unbiased estimation of real-world model performance.

4K Resolution: The decision to utilize 4K resolution images was strategic. Given that drone operations were conducted at relatively high altitudes (30–60 meters), capturing detailed images was crucial for effective detection and identification of human instances. Lower resolutions risked losing essential visual information, hindering accurate detection particularly due to the subtle contrasts and complex visual patterns characteristic of Jordan's desert landscapes. Additionally, high-resolution images significantly improved the annotation accuracy and reliability, essential for developing a robust detection model.

Extensive Manual Annotation Process: Annotation represented one of the most challenging stages of dataset preparation. Due to the complexity and subtlety of the images captured at altitude, automated annotation techniques proved insufficient. Consequently, the dataset was entirely annotated manually using Roboflow, a specialized annotation platform. This process required intensive labor and meticulous attention: (Roboflow, 2025)

- Annotators frequently zoomed multiple levels into each image to ensure accurate labeling of human instances.

- Manual annotation was chosen explicitly to accommodate variability in human poses, size, and environmental complexity.
- Annotation duration per image often exceeded several minutes, underscoring the meticulousness required to maintain dataset quality.

Challenges Encountered: Several significant challenges arose during data collection, which provided valuable operational insights:

Security Restrictions: Drone operations in Jordan face stringent regulations managed by the CARC. Obtaining clearances and permits was complex and time-consuming, with strict altitude limits and specific no-fly zones restricting flight flexibility.

Wildlife Interference ("Kashet Al-Hamam"): Flocks of pigeons frequently disrupt drone operations, requiring additional piloting skill and flight restarts.

Technical and Logistical Constraints: High-resolution imaging significantly reduced drone flight time due to intense battery consumption, necessitating frequent battery changes and recharge breaks. Additionally, signal disruptions caused by terrain variability posed intermittent communication challenges.

Weather Conditions: Variable weather including winds, heat, and visibility periodically impacted drone stability, requiring flexible scheduling of flight operations.

Ethical and Privacy Considerations: Strict ethical guidelines were followed during dataset collection and annotation. Participants, including actors, provided explicit consent for their involvement. Steps were taken to ensure no personally identifiable information (faces, distinguishing features) was compromised in the dataset, maintaining participant privacy and complying with ethical standards.

Limitations and Future Enhancements: While comprehensive, the dataset has certain inherent limitations:

- It primarily captures desert conditions typical of Zarqa, potentially limiting generalizability to significantly different terrains (mountainous or densely vegetated regions).
- The absence of extreme weather conditions (heavy rain, snow, fog) limits immediate applicability under those conditions.

Future expansions might consider integrating:

- Additional environmental conditions or locations.
- Data augmentation or synthetic data.
- Thermal imaging or multispectral sensors to further enhance detection capabilities.

Annotation Tools and Standards: Annotations were performed using Roboflow, providing consistency, precision, and compatibility with standard YOLO format. The choice of widely accepted annotation standards ensures interoperability and potential reusability of the dataset across diverse research and operational scenarios.

4.3 Dataset Processing and Analysis

This section details the processing steps undertaken after data collection, focusing explicitly on image standardization, cleaning, and augmentation. The goal is to transform raw data into a structured, consistent, and computationally manageable format, ideal for training a robust drone-based human detection model. Figure 7 shows the pipeline of the collected data, from the initial pictures to the input for the AI model.

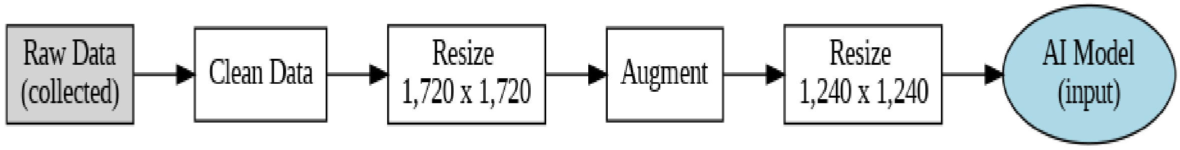

Figure 7. The Overall Data Flow and Process in YOLO-Based Dataset Preparation, from Initial Data Collection Through to the Final Dataset

4.4 Brief Dataset Overview

After manual annotation, the final dataset comprises 2,430 images containing 2,831 labeled human instances, captured originally at 4K resolution (3,840×2,160 pixels) and 24 fps. Images were uniformly resized to 1,280×1,280 pixels to optimize computational efficiency without significant loss of detail. To support rigorous model development, we partitioned the dataset into three subsets: 1,771 images (approximately 1,982 human instances) for training, 437 images (around 566 instances) for validation, and 222 images (approximately 283 instances) for testing, reflecting a 70/20/10 split. Across these splits, the dataset exhibits an average of 1.17 humans per image, with 25% of frames containing no humans, 50% containing one or fewer, and 75% containing two or fewer—closely mirroring real-world SAR conditions where empty frames are common.

Annotations were performed in Roboflow following YOLO standards to ensure consistency and compatibility with prevalent object-detection frameworks. The image collection spans a broad spectrum of environments—industrial zones, agricultural fields, urban fringes, soft-sand deserts, and marketplaces—and includes both organic human activity and staged poses (*e.g.*, lying, sitting, walking).

This diversity, combined with precise labeling and balanced splits, establishes a comprehensive dataset foundation for training, tuning, and evaluating drone-based human detection models.

4.5 Data Preprocessing and Cleaning

Before diving into cleaning operations, we standardized all raw drone footage to ensure uniformity across the dataset. Following preprocessing, a thorough cleaning pipeline removed frames unsuitable for training. Non-informative segments

Data Preprocessing: All annotated images were uniformly resized from their original 4K resolution (3840×2160 pixels) to a standardized 1280×1280 pixels dimension. This resizing significantly reduced computational load, facilitated faster model training, and maintained sufficient resolution to preserve critical image details essential for human detection accuracy.

Data Cleaning: Raw footage collected from drones initially included segments unsuitable for training (*e.g.*, ascending, and descending drone sequences, unstable camera angles, and redundant frames). Therefore, a systematic cleaning process was applied, involving:

Removal of Ascending and Descending Footage: These segments were eliminated to ensure consistent and stable viewpoints. Figure 9 shows a sample of removed pictures during takeoff and landing.

Figure 8. Sample Pictures That were Taken During Takeoff or Landing and were Removed During the Cleaning.

Exclusion of Poor Camera Angles: Images captured at unusable or extreme angles, lacking sufficient clarity for annotation, were discarded. Figure 10 displays samples of images removed due to bad camera angle.

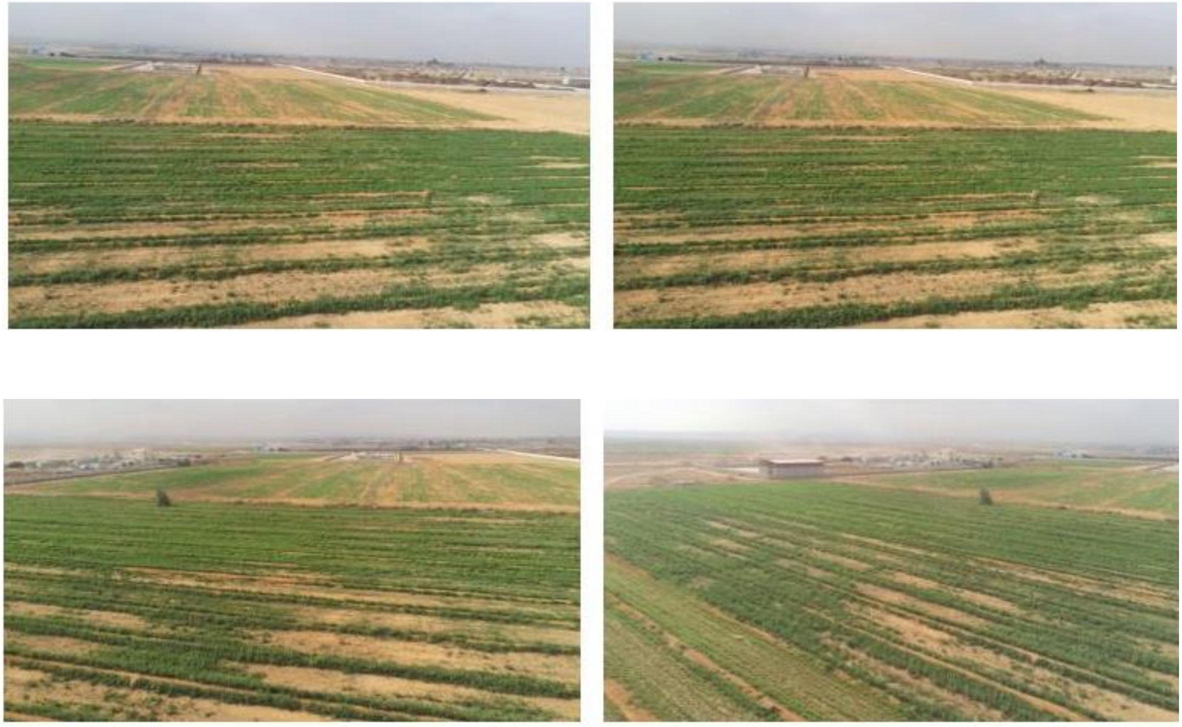

Figure 9. Samples of Bad Angled Pictures That were Removed in the Cleaning Process.

Removal of Images with Wildlife Interference: Images taken by the drone while the pigeons were surrounding it**,** were taken off to have a clear data without any obstacles. Figure 11 shows a sample of images removed due to the interferences with pigeons.

Figure 10. Sample Pictures with Pigeons

Removal of Redundant Frames: Images with a high degree of overlap were selectively removed, reducing redundancy and maintaining dataset diversity.

Overall images across the training, validation, and test sets, have been cleaned manually as previously discussed. A total of 686 images were discarded due to the stated filtering criteria, ensuring a more consistent and reliable dataset. Figure 12 compares the number of images retained (green) versus those discarded (red) in the training, validation, and test splits, illustrating that roughly 20–25% of images were removed from each subset to ensure model robustness.

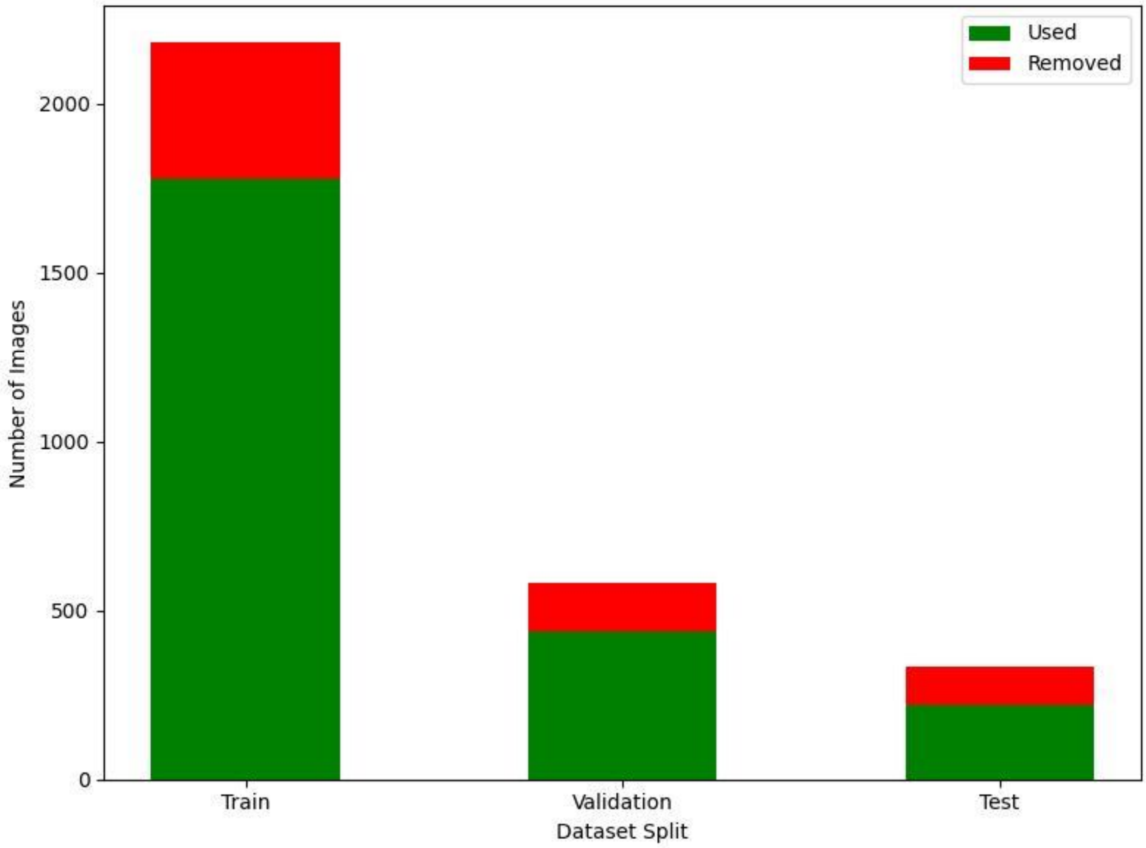

Figure 11. Numbers of Images Removed Versus Used Across Training, Validation, and Testing Subsets

4.6 Data Augmentation

Mosaic augmentation is the augmentation method used, as a powerful data augmentation technique that enriches the training dataset by combining multiple images into a single composite. In this approach, four distinct images are selected and stitched together to create a composite image with a resolution of 1,760×1,760 pixels. This augmentation technique is specifically useful for our case, as the stretching of different pictures can represent different terrain side by side, as this is the case usually in SAR terrain. Each image is strategically placed in one of the quadrants of the final mosaic, and slight random shifts or scales are applied to ensure smooth transitions between images.

This process not only simulates a larger, more diverse scene but also introduces new spatial relationships and context variations that a model might encounter in real-world applications.

Furthermore, this augmentation strategy was applied exclusively to the training dataset and repeated three times. As a result, the training data was effectively tripled, expanding the original set to a total of 5,997 images. This significant increase in training examples provided the model with a richer and more varied set of visual scenarios, helping it better capture the complexity of geographically contiguous terrains encountered in drone-based SAR operations. Ultimately, this enhanced diversity is crucial in mitigating overfitting and boosting the model's generalization performance in dynamic and cluttered real-world environments. Figure 13 shows samples of images after applying Mosaic augmentation for the training dataset.

Figure 12. Sample of Augmented Pictures with Mosaic Augmentation Method

4.7 Numerical and Graphical Data Analysis

To better understand dataset composition and characteristics, several graphical analyses were performed:

Heatmap of Annotation Distribution: Figure 14 shows the heat map plots the centers of human annotations across images, where both the horizontal (X) and vertical (Y) coordinates have been normalized to a 0–1 range. The color scale on the right indicates the density (or count) of annotations in each small region of the plot, with lighter green areas representing lower counts and darker blue areas showing higher counts. Overall, the distribution suggests that most human annotations cluster around the middle portion of the image (roughly X≈0.4 - 0.6, Y≈0.4 - 0.7), while fewer instances appear near the extreme edges (close to 0 or 1 on either axis).

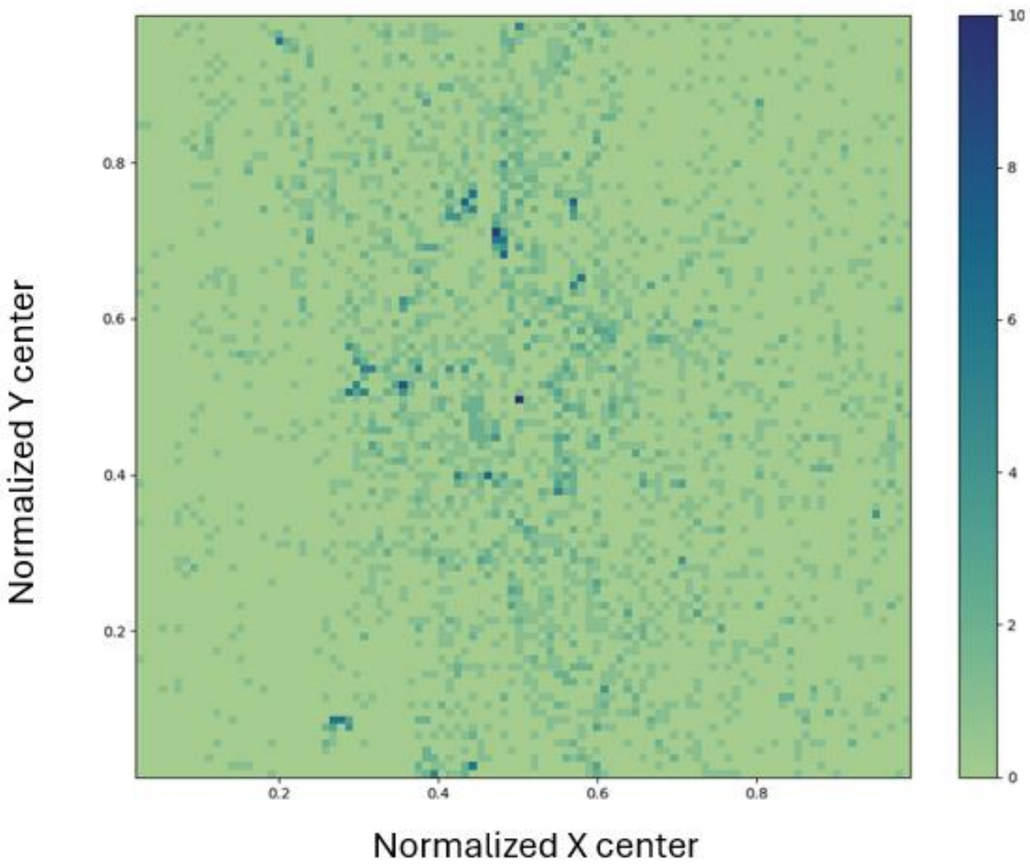

Figure 13. Heatmap for the Normalized Distribution of Human Bounding Boxes

Human Instances per Image: Most images contain zero or one human instance, as shown by the tall bars at 0 and 1, with relatively few images containing larger numbers of humans. This simulates the case of real search and rescue missions when most of the images transmitted will not have human instances in them. Figure 15 shows the number of human instances per image with the frequency of images across the training, validation, and test sets.

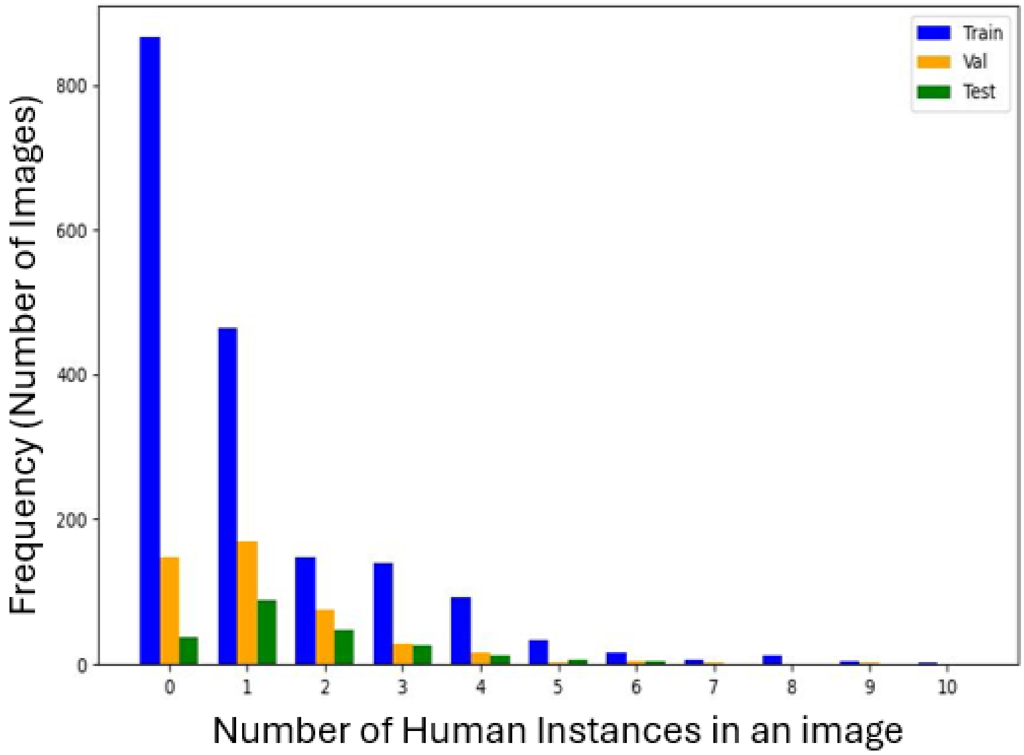

Figure 14. Distribution of the Number of Human Instances per Image Across the Training, Validation, and Test Datasets

Violin Plots of Human Instances: A violin plot representation of the distribution of human instances per image in the training, validation, and testing sets. The plots highlight not only the median and quartiles but also the spread of data points, indicating that the training set includes images with a wide range of human counts.

Each "violin" shows the overall shape of the data's distribution, while the box-and-whisker inside indicates the median, quartiles, and range. In all three splits, the bulk of images have relatively few human instances, but there are outliers extending to higher values most prominently in the training set. The widths of the violins reflect how frequently a given number of instances occurs, revealing that most images fall at or near zero human instances, with only a small subset containing multiple humans.

Figure 16 shows that about 25% of all images contain no humans at all (Q1 = 0), at least half have one or fewer humans (Median = 1), and about 75% have two or fewer humans (Q3 = 2). A small subset of images contains more than two humans, with the maximum observed count around ten.

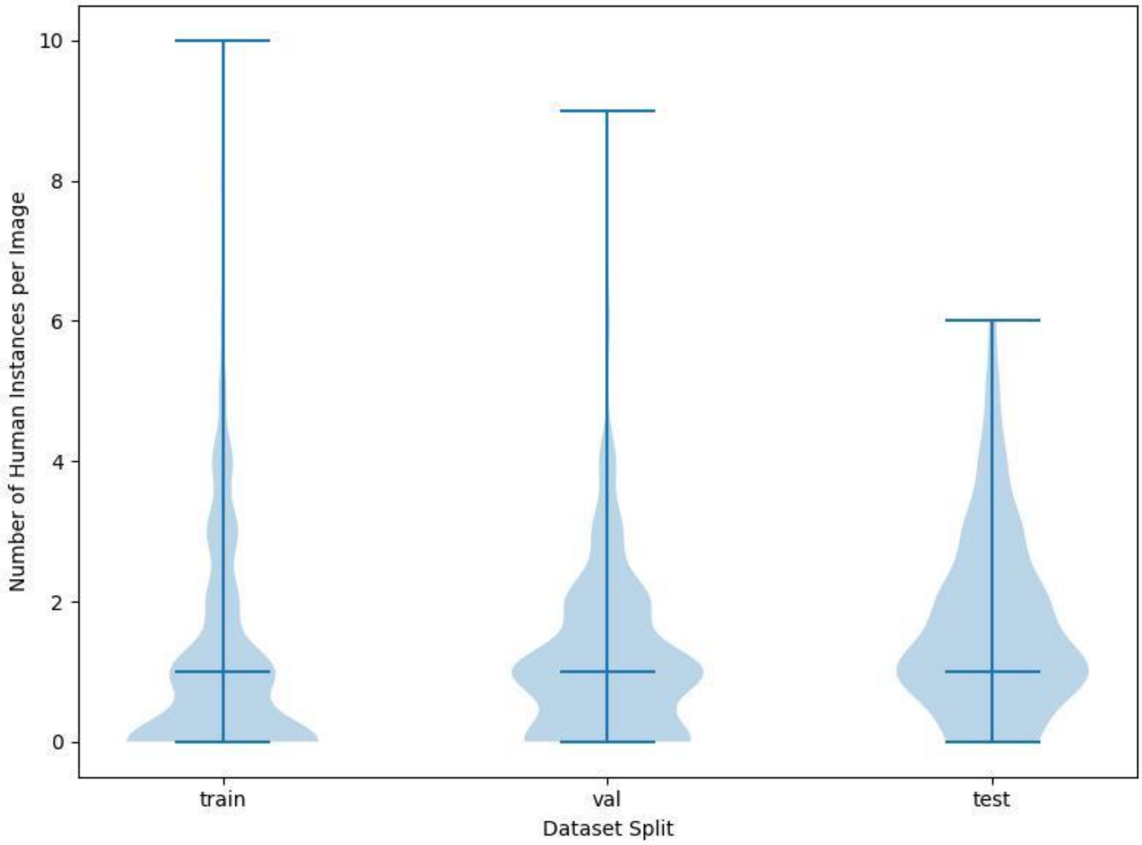

Figure 15. Violin Plot Displaying the Distribution of Human Instances per Image Across the Combined Dataset

4.8 Summary

This chapter documents the end-to-end process of creating a bespoke Jordanian SAR dataset, driven by the recognition that existing benchmarks neither capture local desert textures nor adhere to Jordan's strict drone regulations. Through strategic collaboration with JODDB, two field expeditions were executed across twelve varied locations in Al-Dhlail, yielding 2,430 ultra-high-resolution aerial frames recorded at 24 fps and elevations between 30 m and 60 m. The collected imagery balanced natural and staged human scenarios against diverse terrain features, ensuring representativeness of real-world search-and-rescue contexts.

A rigorous annotation pipeline in Roboflow was employed to label 2,831 human instances, with multi-level zoom verification and iterative quality checks safeguarding precision. The dataset was then systematically partitioned into a 70/20/10 train/validation/test split to support robust model calibration and unbiased performance assessment.

It also include the dataset's rigorous processing, including resizing, cleaning, and strategic augmentation steps. The final prepared dataset offers a comprehensive foundation for model development with the following key characteristics:

- Total Images: 2,430 high-quality frames at 1,280×1,280 pixels resolution

- Human Instances: 2,831 meticulously annotated and verified

- Augmented Training Set: Expanded to 5,997 mosaicked composites

This unified dataset captures diverse Jordanian search-and-rescue scenarios—spanning industrial areas, agricultural fields, urban fringes, and desert expanses—with both natural and staged human poses. The chapter has documented each processing stage: standardizing raw 4K imagery to a consistent format; applying a rigorous cleaning pipeline to discard non-informative or obstructed frames; enhancing the training pool through mosaic augmentation; and performing graphical and numerical analyses to characterize spatial annotation patterns and instance distributions across splits.

By synthesizing these efforts into a single, robust dataset, we ensure optimal conditions for training, validating, and testing drone-based human detection models tailored to the environmental and operational demands of real-world SAR missions in Jordan.

Chapter Five

5 Computer Vision Model Selection

In this chapter, we describe the rationale and process behind selecting and training our object-detection model for SAR missions in Jordan. After reviewing the literature, we elected to use the YOLO family of models for their real-time performance and strong localization capabilities, initially experimenting with YOLO v11, v5, and v8. Early training on low-quality imagery yielded unsatisfactory performance, but the high-resolution data was aqcuired, overall accuracy rose to approximately 75%. Recognizing that false negatives in SAR contexts can be literally life-threatening,

Then we adopted a transfer-learning approach using models pre-trained on the VisDrone dataset by Zhu et al. (2022) chosen for its aerial imagery of crowded and complex environments and tested both YOLO v8 and v11, which achieved mAP scores of around 85.0% and 80.0%, respectively. Placing recall at the forefront of our evaluation metrics to minimize missed detections, we proceeded with extensive fine-tuning adjusting input resolution, batch size, and other hyperparameters over tens of training iterations and ultimately achieved 97.0% precision, 97.6% recall, 97.8% mAP@0.5, and an 85% mAP@0.5:0.95. The following sections detail each phase of this journey, from model selection through to final augmentation strategies.

5.1 Evaluation Metrics

In this section, we define the metrics used to assess model performance and justify why we prioritize recall in a false-negative–sensitive context, while still maintaining high precision for operational efficiency.

Metrics Definition: To objectively assess model performance against SAR mission requirements—and prioritize detection completeness over false alarms—we employ a set of standard evaluation metrics:

Precision (P): Measures the proportion of detected instances that are correct. High precision minimizes false alarms.

$$P = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{17}$$

Recall (R): Quantifies the model's ability to detect all actual human instances. Crucial for ensuring no person in distress is overlooked.

$$R = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{18}$$

F1-Score: Represents harmonic mean of precision and recall, providing a single metric that balances both.

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{19}$$

Mean Average Precision (mAP): Average precision computed over IoU thresholds:

- **mAP@0.5 (mAP50):** Precision averaged at IoU ≥ 0.5.
- **mAP@0.5:0.95 (mAP90):** Precision averaged over IoU thresholds from 0.5 to 0.95 in increments of 0.05.

Emphasis on Recall in False-Negative–Sensitive Tasks: In SAR operations, false negatives (missed detections) can have life-threatening consequences. A single missed person may never be found in time. Consequently, I prioritize recall above all other metrics:

- Maximizing Victim Detection: High recall ensures nearly every human instance is detected.
- Accepting False Positives: I tolerate a modest increase in false positives, since it is preferable to verify an extra detection than to miss a genuine casualty.

Detection confidence: The probability score a model assigns to each detection to express how sure it is that the predicted object is truly present; after inference, these scores are compared with a chosen confidence threshold so that only detections above the threshold are kept, while those below are discarded. Raising the threshold yields fewer but more precise detections—reducing false positives—yet risks missing some true objects, whereas lowering it captures more objects—boosting recall—but introduces more false alarms. Selecting the right threshold therefore balances precision and recall according to the real-world costs of false positives versus false negatives in the application and should be tuned on a validation set so that reported metrics such as precision, recall, or F1-score genuinely reflect the model's behavior under the conditions in which it will be deployed.

Balancing Precision for Operational Efficiency: While recall is paramount, precision remains important to avoid overburdening rescue teams:

- Resource Allocation: Each false positive requires verification, potentially diverting time, and personnel.
- Time Sensitivity: Excessive false alarms can delay response to real emergencies.

By first optimizing recall and then maintaining precision at an acceptably high level, I ensure the model is both lifesaving and practically deployable.

5.2 Initial Experiments with YOLO Versions

In the initial phase of our experiments, we trained several YOLO-based models on the first batch of data, which contained both human and background instances. The table presents the performance metrics precision, recall, mAP@50, and mAP@95 for each model.

Table 5. Performance Metrics from the Initial Training on the First Batch of Data

Model Name	Precision	Recall	mAP@50	mAP@95
YOLO v5s	36.0%	27.0%	27.5%	11.5%
YOLO v8s	41.0%	32.5%	33.5%	19.0%
YOLO v8n	42,5%	31.55	35.0%	19.5%
YOLO v10s	38.5%	29.5%	30.0%	15.0%

Observations: The metrics for YOLO v5s indicate a baseline human detection performance with a precision of 36% and recall of 27%. This model showed significant room for improvement, particularly in accurately detecting human targets under challenging conditions.

These YOLO v8 variants outperformed YOLO v5s, with YOLO v8n exhibiting a slightly higher precision (42.5%) and marginally better mAP values compared to YOLO v8s. The improved performance suggests that architectural refinements in the YOLOv8 family enhance detection capabilities for human targets.

This experimental variant delivered performance metrics that fall between those of YOLO v5s and the YOLO v8 models. While YOLO v10s showed some improvements over YOLO v5s, it did not reach the levels achieved by the YOLO v8 variants.

Challenges and Limitations: To contextualize the initial results and guide subsequent improvements, we identified several key challenges and limitations during this phase:

Data Imbalance: A high proportion of background instances in the first batch inflated class-agnostic metrics and obscured true human-detection performance.

Limited Training Set: The small initial dataset led to overfitting and reduced model generalization in unseen SAR scenarios.

Environmental Complexity: Cluttered scenes, low-contrast conditions, and partial occlusions hampered reliable human detection under real-world constraints.

5.3 Experiments after Data Enhancement

After improving the dataset with high-resolution drone imagery, we conducted multiple experiments to evaluate the performance of different YOLO models for human detection in SAR operations. These experiments aimed to determine the best-performing model architecture and pre-processing strategy, focusing particularly on the impact of tiling versus using untiled images. This section breaks down the training process into the following subsections.

Model Selection and Setup: We experimented with three versions of YOLO architecture: YOLOv11n, YOLOv11l, and YOLOv8l. Each model was tested with both tiled and untiled datasets to explore how pre-processing affects detection performance. The training parameters were kept consistent to ensure a fair comparison across different configurations. The primary evaluation metrics were recall, precision, and F1 score, with recall given special attention due to its critical importance in SAR missions missing a detection (false negative) result in dire consequences.

Tiled *vs.* Untiled Data: Tiling was initially introduced as a potential pre-processing enhancement, based on the assumption that breaking large images into smaller tiles would improve object visibility and training focus. However, results showed that untiled images consistently performed better than their tiled counterparts. This counterintuitive outcome is attributed to two factors:

- Tiling often fragmented human figures, making it harder for the model to recognize them as full body objects
- Tiling significantly reduced the frequency of full human instances, which led to class imbalance during training.

Results: The table presents the performance metrics from our main training experiments, highlighting how each model configuration performed across recall, precision, and F1 score.

Table 6. Performance Comparison of Multiple YOLO Models Using Tiled and Untiled Datasets

Model Name	Precision	Recall	F1 Score
YOLOv11n tiled	87.9%	69.5%	77.6%
YOLOv8l tiled	84.0%	74.0%	78.7%
YOLOv8l untiled	83.0%	86.9%	69.6%
YOLOv11l tiled	75.0%	74.0%	74.5%
YOLOv11l untiled	82.1%	76.8%	79.4%

Insights and Final Decision: The results reaffirmed the importance of proper dataset structure. Untiled configurations consistently achieved higher recall rates, with YOLOv8l (untiled) achieving the highest recall at 86.9%, followed by YOLOv11l (untiled) at

76.8%. While YOLOv11n had strong precision, it lagged in recall and F1 score. Notably, YOLOv11l untiled produced the highest F1 score of 79.4%, showing a solid balance between precision and recall.

Given the mission-critical importance of recall in SAR applications, YOLOv8l untiled was selected as the model to advance with fine-tuning and augmentation, forming the backbone of our final system due to its superior recall and generalization capabilities on untiled, high-resolution images.

Moreover, these findings highlighted the significant impact of data pre-processing on model performance. The unexpected degradation caused by tiling suggested that further experimentation with other pre-processing techniques such as adaptive cropping, image enhancement, or synthetic data generation could yield additional performance improvements. This insight guided the next phase of the project, which focused on fine-tuning and data augmentation strategies.

5.4 Transfer Learning with VisDrone Pretrained Model

While YOLO models are typically pretrained on general-purpose datasets like COCO, we hypothesized that initializing our model with weights from a dataset specifically comprising drone-captured imagery would enhance performance in our SAR context. Drone imagery presents unique challenges such as varying altitudes, angles, and object scales that general datasets may not adequately represent. Therefore, we explored the potential benefits of transfer learning using a model pretrained on the VisDrone dataset.

Overview of the VisDrone Dataset: The VisDrone dataset is a comprehensive benchmark designed for drone-based computer vision tasks. Collected by the AISKYEYE team at Tianjin University, it encompasses:

- 10,209 static images and 288 video clips totaling 261,908 frames, captured across fourteen diverse cities in China.
- Over 2.6 million annotated bounding boxes, covering objects like pedestrians, cars, bicycles, and tricycles.
- Annotations include object categories, occlusion status, truncation ratios, and scene visibility.
- Varied environments, including urban and rural settings, under different weather and lighting conditions.

VisDrone Dataset Structure: The dataset is structured into five main tasks:

1. Object Detection in Images
2. Object Detection in Videos
3. Single-Object Tracking
4. Multi-Object Tracking
5. Crowd Counting

This diversity and specificity make VisDrone an ideal candidate for pretraining models intended for drone-based applications.

5.5 Fine-Tuning Techniques

In this section, I describe the key fine-tuning strategies I applied to optimize YOLO v8's performance, excluding the final evaluation metrics which are presented:

Hyperparameter Adjustments: I experimented with batch sizes ranging from 4 to 16 to find the optimal trade-off between gradient stability and resource usage.

- Tested Values: 4, 8, 10, 12, 16.

- Chosen Value: 10

A batch size of 10 provided stable loss convergence on the validation set without overloading GPU memory or slowing training unduly.

I fixed the number of training epochs at **100** with an early-stopping **patience** of **10**, ensuring that training would halt if no improvement occurred on the validation loss for ten consecutive epochs.

Input Resolution and Network Scaling: Different input image dimensions impact YOLO’s ability to detect small *vs*. large objects:

- Tested Resolutions: 640×640, 720×720, 1,280×1,280, and 1,440×1,440.
- Chosen Resolution: 1,280×1,280.

At 1280×1280, the model balanced detection accuracy for small human figures against training and inference speed, outperforming both lower and higher resolutions in early comparative trials.

Data Augmentation Strategies: To improve generalization and robustness, I tested three augmentation techniques on the training set and measured their impact on recall and precision:

Table 7. Results for the YOLO v8 Model with Two Types of Augmentation Trials

Augmentation	Recall	Precision
Blur	85.0%	90.0%
Noise Addition	83.0%	92.0%

Notes:

- Blur improved resilience to motion blur but showed lower recall.
- Noise Addition slightly increased precision but at the cost of recall.

Regularization and Optimization Methods: I adopted the following final training settings to further enhance convergence and prevent overfitting:

- Optimizer: Auto-selected (adaptive LR scheduler implementing warmup + cosine decay)
- Initial Learning Rate (lr0): 0.02
- Final LR Factor (lrf): 0.01
- Momentum: 0.937
- Weight Decay: 0.0005
- Dropout: 0
- Early Stopping Patience: 10
- Save Period: Every epoch (save=True, save_period=1)
- Verbose and Plots: Enabled (verbose=True, plots=True)

These settings, combined with the above hyperparameters and augmentation strategy, formed the foundation for my final model fine-tuning process.

5.6 Final Results with Mosaic Augmentation

This section provides a comprehensive evaluation of the YOLO v8X-that is pretrained on VisDrone dataset and then trained with Mosaic image augmentation dataset. Both quantitative metrics and qualitative inspections are presented to demonstrate how Mosaic augmentation improved the model's learning dynamics, generalization, and predictive confidence.

6.6.1 Training Dynamics

Figure 17 shows the evolution of the training and validation loss across 100 epochs. The training loss follows a smooth downward trajectory, indicating effective minimization of the objective function. The validation loss) decreases rapidly in the early epochs before flattening out after epoch ≈ 85. This plateau, combined with the absence of a sharp divergence between the two curves, suggests that Mosaic augmentation injected sufficient variability to prevent over-fitting while still allowing the network to converge.

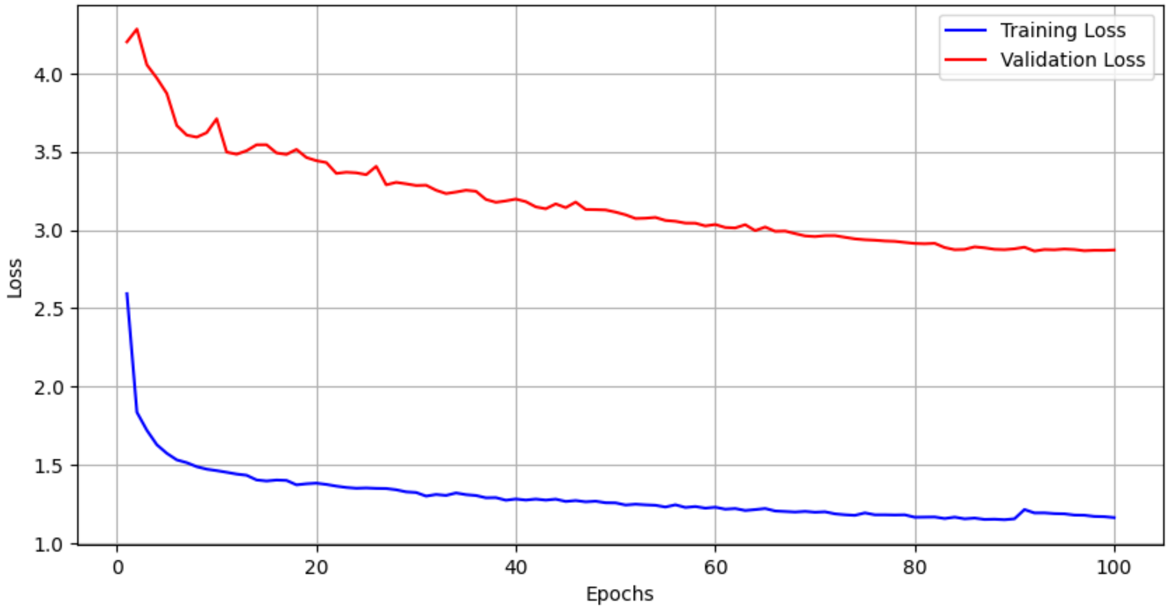

Figure 16. Training and Validation Loss Curves Over 100 Epochs for YOLOv8X VisDrone initialized

To monitor class-specific learning, precision and recall were computed after each epoch Figure 18 shows precision climbs quickly from an initial ≈ 85% to > 96%, reflecting a sharp drop in false-positive detections. Recall follows a similar pattern, settling above 97%. The near-parallel rise of both metrics evidence balanced learning: the detector becomes more certain (precision) without missing relevant targets (recall).

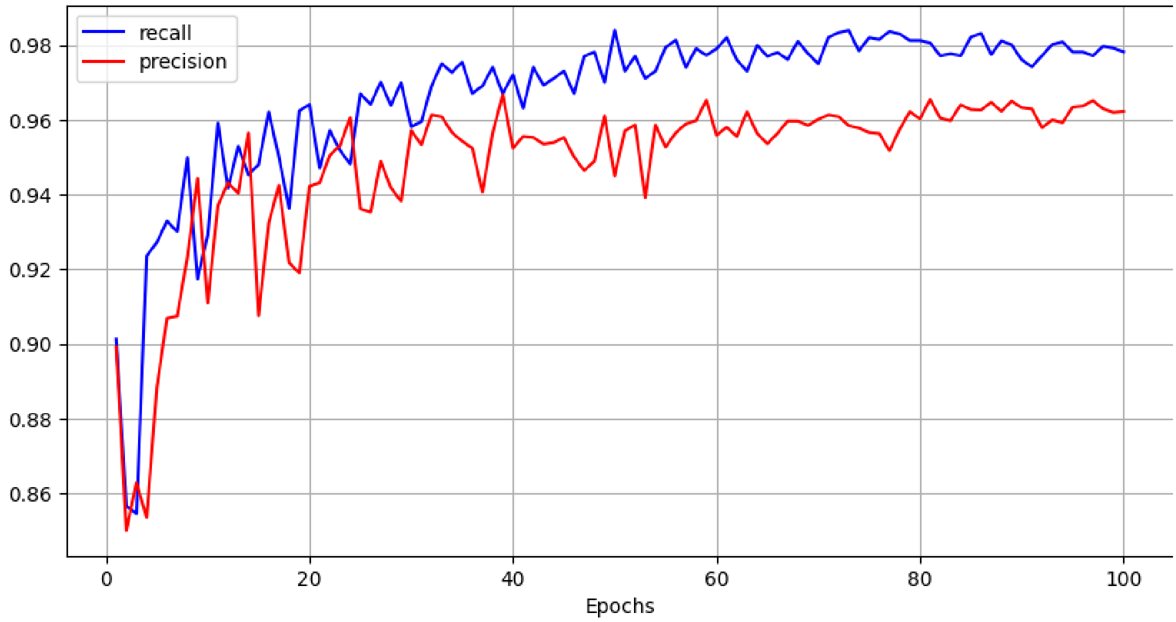

Figure 17. Precision and Recall Curves across 100 Training Epochs, Demonstrating Rapid Improvements in Both Metrics

Figure 19 presents the mean Average Precision (mAP) metrics. The detector reaches an mAP@0.50 above 95% after only ten epochs and stabilizes at ≈98% by epoch 20, indicating that the model quickly learned to place reasonably tight bounding boxes. The stricter mAP@0.95 curve, which requires near-perfect localization, shows a slower but steady ascent, plateauing around 52%. This gradual improvement demonstrates that continued training with Mosaic samples incrementally refines spatial accuracy.

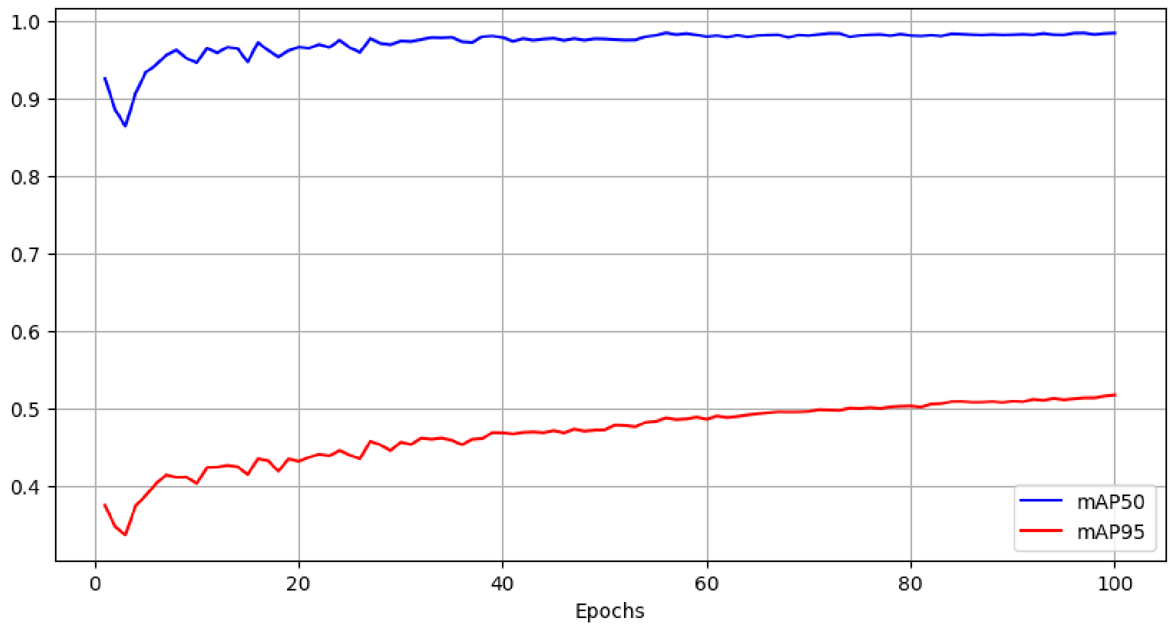

Figure 18. Mean Average Precision Over 100 Epochs: mAP@0.50 Rises Rapidly to 98% and Remains Stable, while mAP@0.95 Steadily Improves from 35% to 52%,

The confusion matrix in Figure 20 summarizes the test-set performance across two classes: Human and Background Out of 358 human instances, 344 were correctly detected (recall = 96.1%), while only 14 were missed. Conversely, 22 were erroneously flagged as humans, yielding a precision of 94.0% for the Human class.

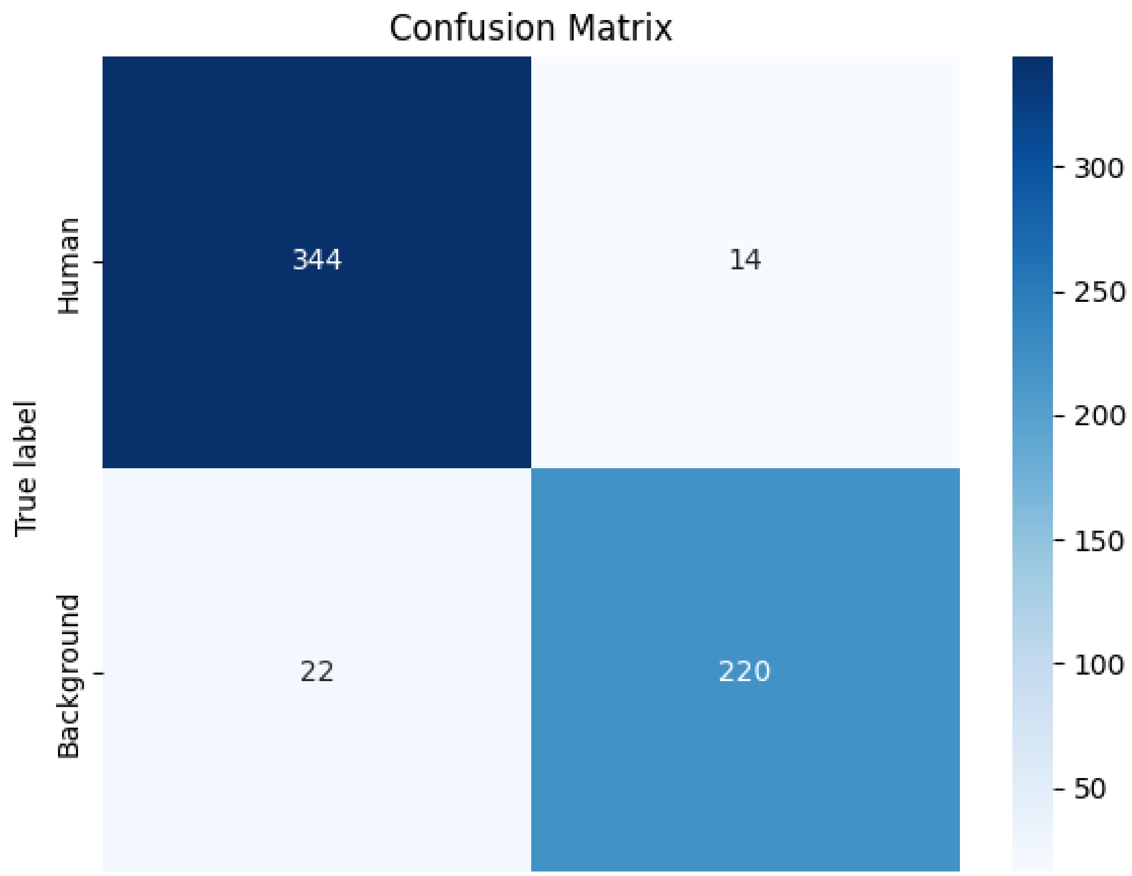

Figure 19. Confusion Matrix for the Human Detection Model YOLO v8X initialized on Visdrone on the Test Set that Consists of 221 images.

Because deployment scenarios often require tuning the detection-confidence threshold, confidence-performance curves were plotted. Figure 21 shows the F1-confidence curve peaks at 97.0% when the threshold is 0.57, providing a principled operating point that balances precision and recall. Precision continues rising beyond this threshold nearing 100% at 0.95 confidence, but at the cost of a gradual decline in recall. These curves allow stakeholders to choose a threshold aligned with mission objectives for instance, maximizing true-human detections in safety-critical applications or minimizing false alarms in low-risk settings.

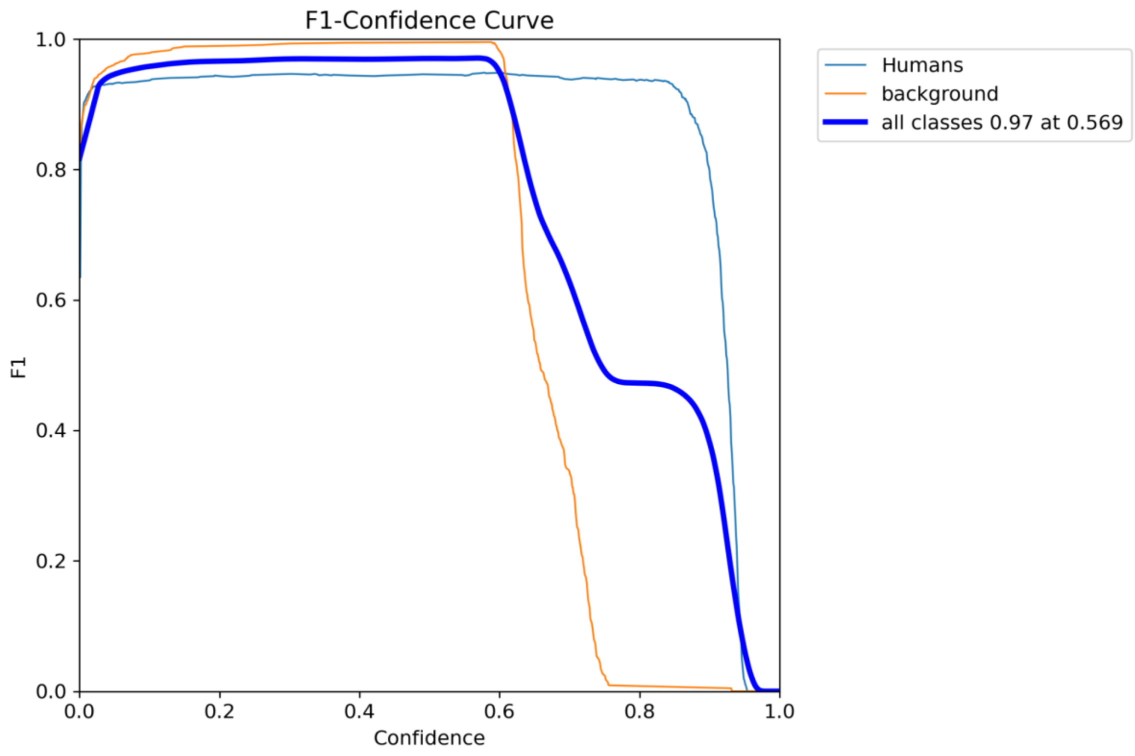

Figure 20. F1–Confidence Curves for Human and Background Detection, with the Overall F1 Maximized at 97% when the Confidence Threshold is Set to 0.569

Figure 21 shows a sample annotated image showing the background and human instances and how they were annotated with roboflow.

Figure 21. Sample Picture Showing Human and Background Classes

Figure 22 demonstrates the relationship between detection confidence and precision for both “humans” and “background” classes as the confidence threshold increases from 0 to 1, precision for each class steadily rises toward perfect accuracy, and the bold dark-blue line shows that when using a combined confidence cutoff of about94.8, the detector achieves 100% precision across all classes.

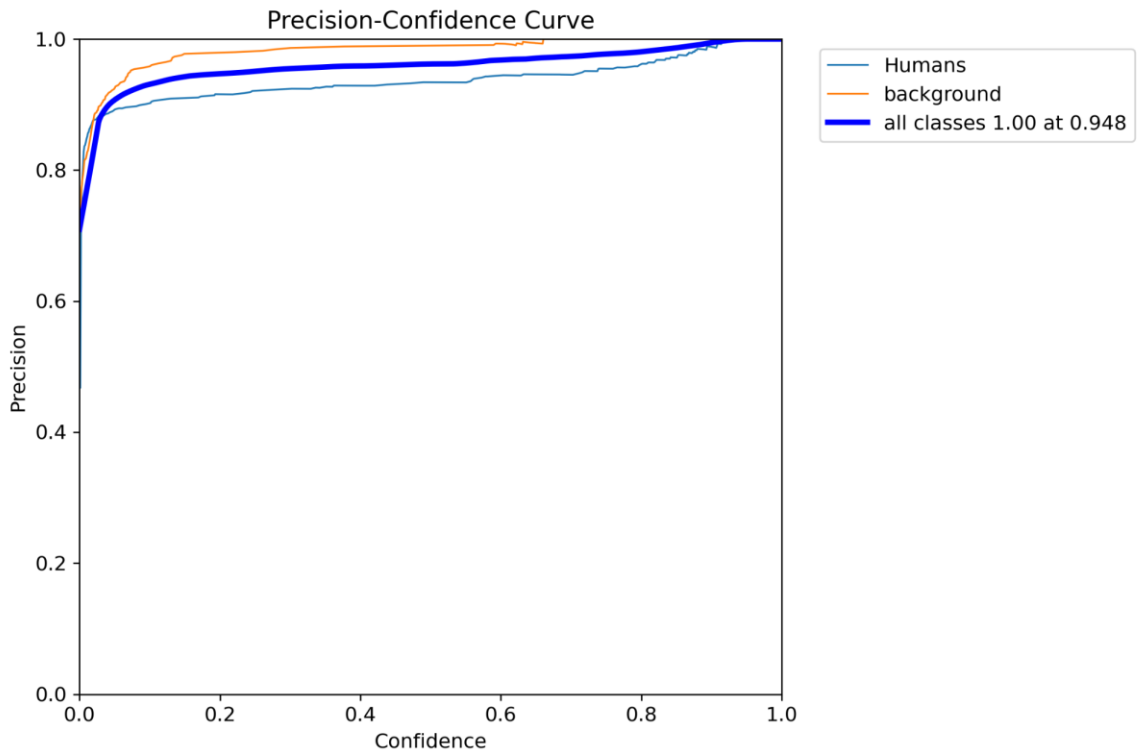

Figure 22. Precision–Confidence Curves for Human and Background Detection.

Figure 23 demonstrates how recall for humans and background varies with the model’s confidence threshold: at very low thresholds nearly 100% of all true objects are found, but as you raise the bar particularly above about 0.6 for background and around 0.8 for humans recall drops off sharply; the thick dark-blue line shows that when pooling both classes, you still recover about 99% of all true instances at a zero-confidence cutoff, illustrating the trade-off between confidence and completeness of detection.

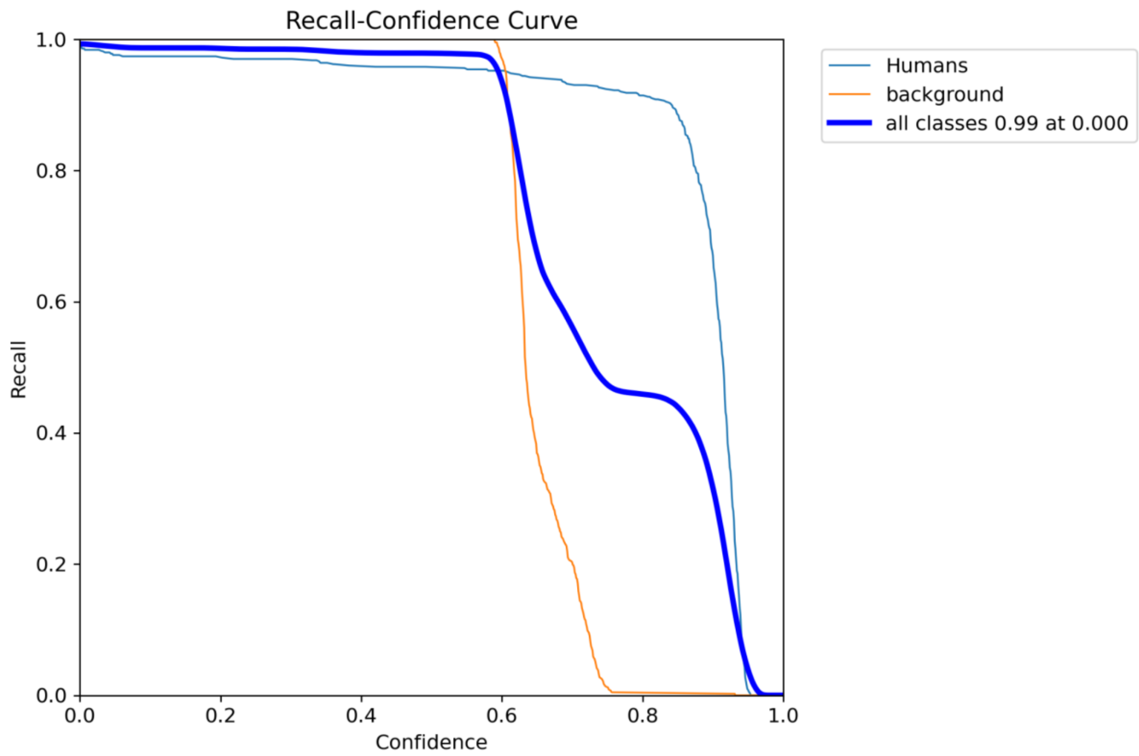

Figure 23. Recall–Confidence Curves for Human and Background Detection, with Overall Recall Maximized at 99% when Using a Confidence Threshold of 0

Figure 24 demonstrates the trade-off between precision and recall for each class humans and backgrounds showing that both curves remain close to the top-right corner (yielding AP scores of 97.85 and 99.1%, respectively) and that the overall mean average precision at an IoU threshold of 0.5 is an impressive 98.4%, indicating the detector achieves both high correctness and high completeness across detection thresholds.

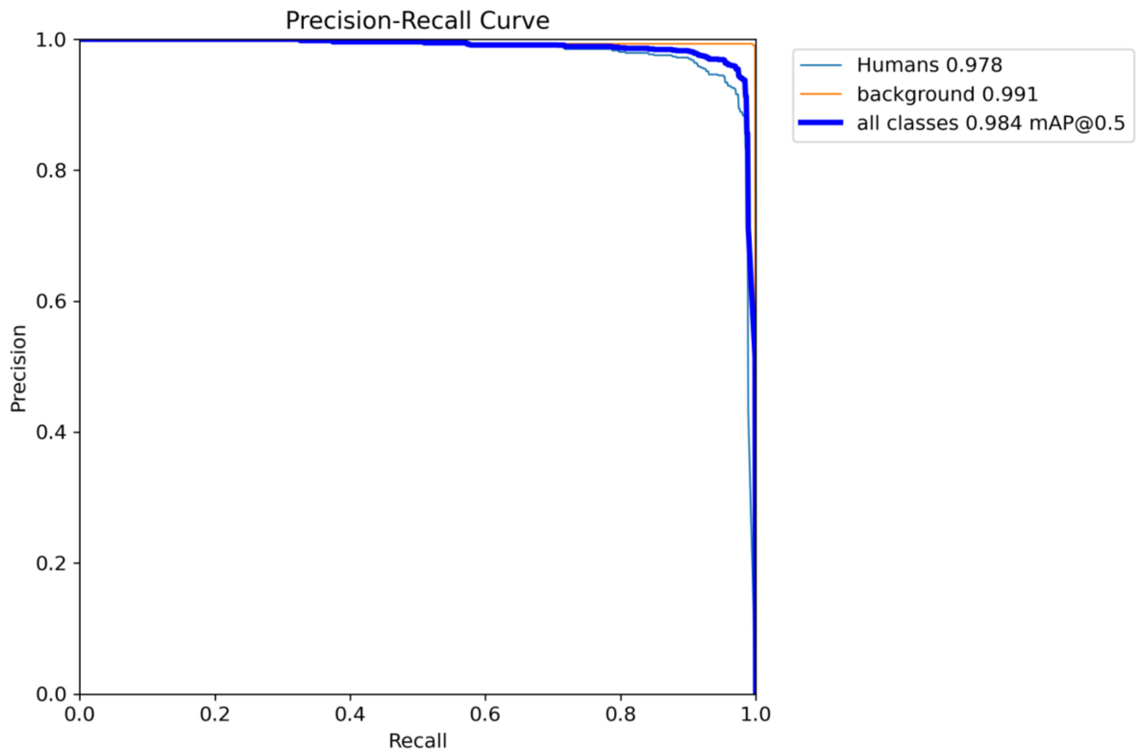

Figure 24. Precision-Recall Curve with Per-Class Statistics

6.6.2 Qualitative Visual Inspection

Beyond numeric metrics, qualitative inspection of validation batches provides intuitive evidence of model behavior. Figures show ground-truth labels and predicted boxes for three representative batches that include high-density urban traffic and sparse rural scenes.

The predictions (right panels) exhibit:

- Accurate localization of humans even under severe scale variation (tall *vs*. distant pedestrians).
- Proper suppression of the Background class in image regions devoid of humans, thanks to the confidence threshold.

• Few but instructive failure cases: missed detections of occluding individuals and occasional oversized boxes around clustered pedestrians, reflecting annotation ambiguity rather than model error.

Overall, the visual evidence corroborates the quantitative findings and underscores the utility of Mosaic augmentation in boosting context awareness. Figures 25, 26 and 27 display the prediction *vs*. ground truth for the three prediction batches.

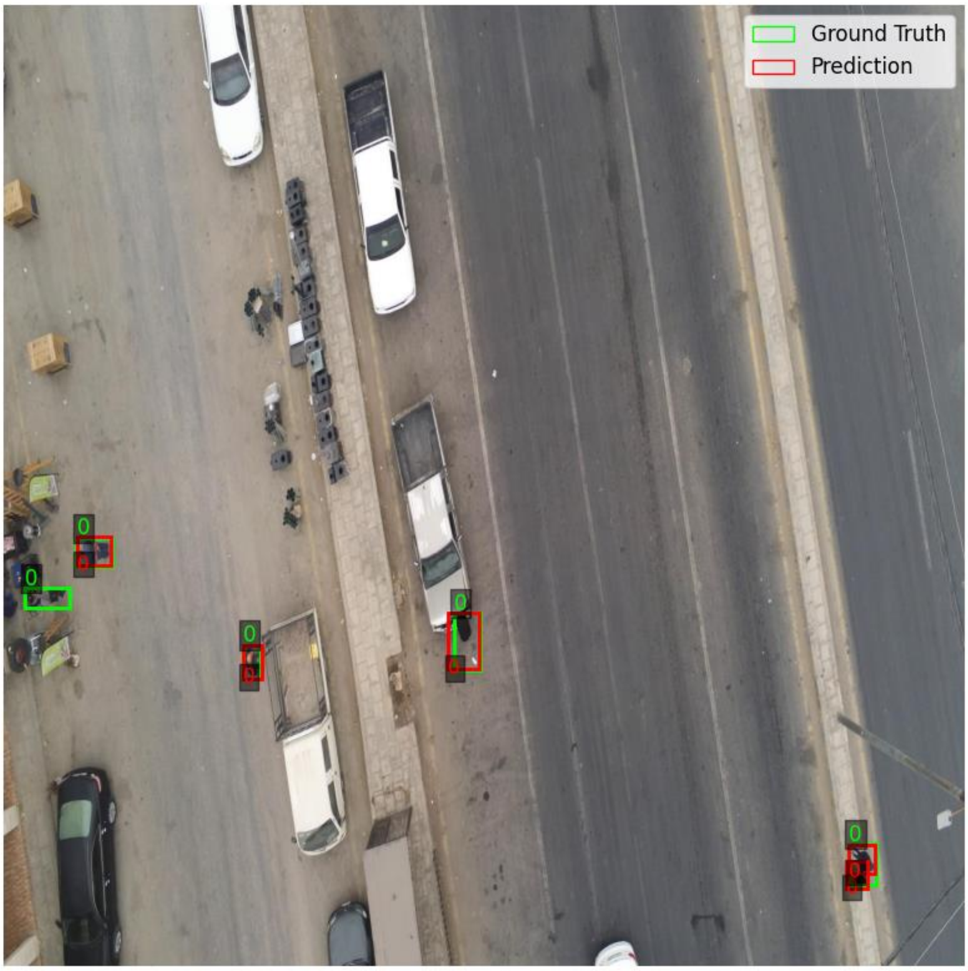

Figure 25 Ground Truth (Green) vs. Predictions (Red) for a Sample of Batch 1

Figure 26. Ground Truth (Green) *vs*. Predictions (Red) for a Sample of Batch 2

Figure 27. Ground Truth (Green) *vs*. Predictions (Red) for a Sample of Batch 3

5.7 Results Summary

In this section, the summary of the results of the final fine tune model is shown, the next table demonstrates the breakdown evaluation for the final model.

Table 8. Result for the Best Performance Model Overall and for Each Class

Class	Predictions	Ground-truth	Precision	Recall	mAP @ 0.50	mAP @ 0.95
All (-avg)	—	—	97.0%	97.6%	98.4%	52.0%
Humans	344	358	94.7%	96.1%	97.7%	83.6%
Background	221	221	99.3%	100.0%	99.1%	20.3%

Overall Performance: A precision of 97.0% means that only 3% of all positive detections are false alarms, while a recall of 97.6% shows the model retrieves nearly every true object.

The combined mAP@0.50 = 98.4% confirms that bounding-boxes are tight enough for most downstream tasks, and an mAP@0.95 = 52.0% although a more demanding metric still indicates reliable high-IoU localization.

Humans *vs.* Background: Humans: With 358 ground-truth instances, the detector achieves 94.7% precision and 96.1% recall, yielding mAP@0.50 = 97.7% and an excellent mAP@0.95 = 8306%. These scores show that the network not only finds pedestrians consistently but also delineates them accurately, even when the IoU threshold is stringent.

Background: All 221 background regions are correctly identified (recall = 1.00) and almost never mistaken for humans (precision = 99.3%). The lower mAP@0.95 = 20.3% is expected because large, amorphous "background" boxes rarely satisfy a 95%

IoU requirement; precision-oriented applications can therefore rely on the class probabilities rather than bounding-box overlap for background filtration.

Inference Speed and Deployment: The 2.9 ms per frame runtime on an NVIDIA A100-40 GB equates to ≈ 345 frames per second. This exceeds real-time needs for common video feeds (30–60 FPS) by an order of magnitude, leaving ample headroom for multiple parallel cameras, on-the-fly post-processing, or running on less-capable hardware without sacrificing real-time operation.

Chapter Six

6 Discussion, Conclusions, and Future Work

This chapter synthesizes the key findings presented in the previous chapters, placing them in the broader context of SAR operations in Jordan. We first interpret results related to route planning, imaging parameters, and model performance to understand their practical impact and limitations. Building on this, we draw overarching conclusions about system design tradeoffs and operational efficacy. Finally, we outline targeted future work—spanning algorithmic refinements, field validation, and extended system capabilities—to chart a roadmap for advancing drone-assisted SAR missions in challenging environments.

6.1 Discussion

The discussion systematically addresses key aspects of the results, embedding analysis within each subsection to highlight the significance and implications of each finding in relation to the research questions.

Route Generation: The adoption of the lawnmower algorithm effectively minimized overlap and efficient coverage. The choice of this method aligns with previous research, particularly Calamoneri *et al*. (2022), who emphasized structured route planning for post-disaster scenarios. Nevertheless, potential limitations include dependency on homogeneous drone fleets, which could be problematic when diverse drone types with varying operational capacities are used. Future studies should explore hybrid approaches combining different route planning algorithms for diverse drone specifications.

Drone Elevation and Camera Specifications: Establishing a clear mathematical relationship between drone elevation and camera capabilities significantly improved detection clarity. This result validates the approach of Marques *et al*. (2023) and provides practical guidelines for drone operation in SAR scenarios. Future improvements could include adaptive elevation models that dynamically adjust based on real-time environmental conditions such as weather or visibility constraints.

Machine Learning Model Selection: The YOLOv8 model trained on the VisDrone dataset demonstrated superior performance, achieving high recall (97.6%) and precision (97.0%), vital for SAR operations. These findings align closely with Sruthi (2021), underscoring YOLO's effectiveness in detecting humans in diverse scenarios. However, potential limitations include model robustness in extreme environmental conditions (*e.g.,* dense fog, heavy rain), necessitating further dataset diversification and robust model testing under varying environmental stresses.

Alternative Solutions and Critical Analysis: Although the overall performance of the model, the segmentation and the proposed solution is sufficient for a beginning there are multiple issues that challenge them.

Drone Communication Range and Cost: The financial and logistical constraints of limited drone availability necessitated multi-trip planning, which complicates timely SAR operations. Future strategies could incorporate relay drones or airborne communication platforms to extend operational ranges and mitigate these constraints.

Onboard *vs*. Server-Based Processing: Current solution proposes server-based processing maximizing model complexity and detection accuracy but introduces latency. An alternative hybrid approach with preliminary onboard processing could mitigate latency and communication demands, though it introduces new challenges regarding onboard computational limitations and power consumption.

6.2 Conclusions

This thesis provided a comprehensive exploration into enhancing SAR operations through the integration of drone clusters and advanced machine learning methodologies. Through meticulous data collection, including high-resolution imagery specific to Jordanian environments, and systematic annotation processes, the foundation for reliable human detection models was established.

Drone routing and segmentation methodologies significantly improved area coverage efficiency, while the elevation calculations based on camera specifications provided practical operational guidelines, ensuring high-quality imaging. Utilizing the YOLOv8 model, pretrained on drone-specific datasets and enhanced with mosaic augmentation, resulted in superior detection performance crucial for effective rescue operations.

Furthermore, the developed communication framework demonstrated potential for real-time intelligence dissemination to emergency responders, enhancing the overall efficiency and responsiveness of SAR missions. These combined achievements provide robust groundwork for future SAR operation enhancements.

6.3 Future Work

Future research and development should focus on actual physical implementation and rigorous field-testing of the proposed SAR system, encompassing full-scale deployment scenarios. The development and testing of adaptive algorithms to dynamically adjust drone altitudes based on real-time environmental feedback represents an essential next step.

Further diversification of datasets to include varying environmental conditions (heavy rain, fog, nighttime) and the integration of additional sensor types (thermal, multispectral) will significantly enhance system robustness. Exploring advanced hybrid processing methods that balance onboard computational efficiency with detailed server-based analysis is necessary for optimizing operational responsiveness. Additionally, investigating advanced communication technologies like relay drones or satellite communication could substantially extend drone operational ranges, improving coverage in extensive or challenging terrain.

Addressing regulatory and ethical considerations associated with drone deployments, particularly regarding privacy concerns and airspace regulations, remains crucial. Continued collaboration with local authorities and international bodies is vital to navigate and establish clear operational standards and protocols.

In summary, the ongoing evolution of drone technology combined with advanced machine learning algorithms offers exciting opportunities for further enhancing SAR capabilities. Continued research into these areas promises significant improvements in disaster response effectiveness and ultimately, life-saving outcomes.

References

Albaraa Zaid Alkilani. (2025a). AlbaraaAlkilani/**SAR_Jordan: SAR_Jordan (SAR_Jor) [Source Code]**. Zenodo. https://doi.org/10.5281/zenodo.15427899

Albaraa Zaid Alkilani. (2025b). **SAR_Jordan [Dataset].** Kaggle. https://doi.org/10.34740/KAGGLE/DSV/11827911

Aldrich, D. P., and Sawada, Y. (2015). **The physical and social determinants of mortality in the 3.11 tsunami.** Social Science and Medicine, 124, 66–75.

Alsumayt, A., El-Haggar, N., Amouri, L., Alfawaer, Z. M., and Aljameel, S. S. (2023). **Smart flood detection with AI and blockchain integration in Saudi Arabia using drones.** Sensors, 23(11), 5148.

American Society for Photogrammetry and Remote Sensing. (2023). Positional accuracy standards for digital geospatial data (2nd ed.). Retrieved from https://www.asprs.org/news-resources/asprs-positional-accuracy-standards-for-digital-geospatial-dataApvrille, L., Tanzi, T., and Dugelay, J.-L. (2014). **2014 XXXIth URSI General Assembly and Scientific Symposium.**

Bachir, N., and Memon, Q. A. (2024). **Benchmarking YOLOv5 models for improved human detection in search and rescue missions.** Journal of Electronic Science and Technology, 100243. https://doi.org/10.1016/j.jnlest.2024.100243

Berndt, J., Meißner, H., and Kraft, T. (2023). On the accuracy of Yolov8-CNN regarding detection of humans in Nadir aerial images for search and rescue applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-1/W2-2023, 139–146. https://doi.org/10.5194/isprs-archives-xlviii-1-w2-2023-139-2023

Barnas, A. F., Darby, B. J., Vandeberg, G. S., Rockwell, R. F., and Ellis-Felege, S. N. (2019). **A comparison of drone imagery and ground-based methods for estimating the extent of habitat destruction by lesser snow geese.** PLoS ONE, 14.

Bejiga, M. B., Zeggada, A., Nouffidj, A., and Melgani, F. (2017). **A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery.** Remote Sensing, 9.

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). **Estimating biomass of barley using crop surface models derived from UAV-based RGB imaging.** Remote Sensing, 6, 10395–10412.

Bucknell, A., and Bassindale, T. (2017). **An investigation into the effect of surveillance of drones on textile evidence at crime scenes.** Science and Justice, 57(5), 373–375.

Calabrò, A., and Marchetti, E. (2024). **Nextgen of Search and rescue operations through trustworthy doped drone's fleet.** 2024 50th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 467–468. https://doi.org/10.1109/SEAA64295.2024.00077

Calamoneri, T., Coro, F., and Mancini, S. (2022). **A realistic model to support rescue operations after an earthquake via UAVs.** IEEE Access, 10, 6109–6125.

Caputo, S., Castellano, G., Greco, F., Mencar, C., Petti, N., & Vessio, G. (2021, December). **Human detection in drone images using YOLO for search-and-rescue operations**. In International conference of the Italian association for artificial intelligence (pp. 326-337). Cham: Springer International Publishing.

Clothier, R. A., Greer, D. A., Greer, D. G., and Mehta, A. M. (2015). **Risk perception and the public acceptance of drones.** Risk Analysis, 35(6), 1167–1183.

Colefax, A. P., Butcher, P. A., Pagendam, D. E., and Kelaher, B. P. (2019). **Reliability of marine faunal detections in drone-based monitoring.** Ocean and Coastal Management, 174, 108–115.

Climate Change Knowledge Portal – Jordan Profile (World Bank). **Climate-related hazards in Jordan and future trends.** https://climateknowledgeportal.worldbank.org/country/jordan

Díaz-Varela, R. A., de la Rosa, R., León, L., and Zarco-Tejada, P. J. (2015). **High-resolution airborne UAV imagery to assess olive tree crown parameters using 3D photo reconstruction: Application in breeding trials.** Remote Sensing, 7, 4213–4232.

Ewers, J., Anderson, D., and Thomson, D. (2023). **Optimal path planning using psychological profiling in drone-assisted missing person search.** https://doi.org/10.1002/adc2.167

Ewers, J., Anderson, D., and Thomson, D. (2024). **Deep Reinforcement Learning for Time-Critical Wilderness Search And Rescue Using Drones.** https://doi.org/10.48550/arxiv.2405.12800

Floreano, D., and Wood, R. J. (2015). **Science, technology, and the future of small autonomous drones.** Nature, 521(7553), 460–466.

Gal, O. (2023). **Adaptive USVs Swarm Optimization for Target Tracking in Dynamic Environments.** arXiv:2408.06713.

Gao, M., and Zhang, X. (2022). **Cooperative Search Method for Multiple UAVs Based on Deep Reinforcement Learning.** Sensors, 22(18), 6737.

Goodchild, M. F., and Glennon, J. A. (2010). **Crowdsourcing geographic information for disaster response: A research frontier.** International Journal of Digital Earth, 3(3), 231–241.

Guo, Y., Yi, J., and Hao, J. (2023). **UAV Maritime Search and Rescue Based on Improved A and RCPP Algorithms for Path Planning**.* https://doi.org/10.1109/ICPICS58376.2023.10235725

Hadadin, N., and Tarawneh, Z. (2009). **Main Factors Leading to the Application of Strategic Environmental Assessment (SEA) in Jordan.** Environmental Impact Assessment Review, 29(3), 154–159. https://www.mota.gov.jo/ebv4.0/root_storage/ar/eb_list_page/final_sea.pdf

Jiménez-Jiménez, S. I., Ojeda-Bustamante, W., Ontiveros-Capurata, R. E., and Marcial-Pablo, M. de J. (2020). **Rapid urban flood damage assessment using high-resolution remote sensing data and an object-based approach.** Geomatics, Natural Hazards, and Risk, 11, 906–927.

The Jordan National Strategy For Disaster Risk Reduction. (2019-2022.). https://www.preventionweb.net/files/68511_nationalnatrualdisasterriskreduciot.pdf

Kannadaguli, P. (2020). **YOLO V4 based human detection system using aerial thermal imaging for UAV based surveillance applications**. 2020 International Conference on Decision Aid Sciences and Application (DASA), 1213–1219. https://doi.org/10.1109/dasa51403.2020.9317198

Kaur, M., and Khandnor, P. (2023). **Performance Comparison of YOLO Variants for Object Detection in Drone-Based Imagery.** In Misra, R., Omer, R., Rajarajan, M., Veeravalli, B., Kesswani, N., and Mishra, P. (Eds.), Machine Learning and Big Data Analytics. ICMLBDA 2022. Springer Proceedings in Mathematics and Statistics, vol. 401. Springer, Cham. https://doi.org/10.1007/978-3-031-15175-0_6

Lygouras, E., Santavas, N., Taitzoglou, A., Tarchanidis, K., Mitropoulos, A., and Gasteratos, A. (2019). **Unsupervised human detection with an embedded vision system on a fully autonomous UAV for search and rescue operations.** Sensors (Switzerland), 19.

Lyu, M., Zhao, Y., Huang, C., and Huang, H. (2023). **Unmanned Aerial Vehicles for Search and Rescue: A Survey**. Remote Sensing, 15(13), 3266. https://doi.org/10.3390/rs15133266

Marques, L., Macias, J. J. E., and Angeloudis, P. (2023). **Probabilistic Planning for Maritime Search and Rescue**. arXiv preprint arXiv:2306.03871.

Mohd Daud, S. M., Mohd Yusof, M. Y., Heo, C. C., Khoo, L. S., Chainchel Singh, M. K., Mahmood, M. S., and Nawawi, H. (2021). **Applications of drones in Disaster Management: A scoping review.** Science and Justice, 62(1), 30–42.

Molina, P., Parés Calaf, M. E., Colomina, I., Vitoria, T., Silva, P., Skaloud, J., Kornus, W., Prades, R., and Aguilera, C. (2012, July/August). **Drones to the rescue! Unmanned aerial search missions based on thermal imaging and reliable navigation**. Inside GNSS, 37–38. Retrieved from https://insidegnss.com/drones-to-the-rescue/

Mueller, M., Smith, N., and Ghanem, B. (2016**). A benchmark and simulator for UAV tracking.** ***Lecture Notes*** *in Computer Science*, 445–461. https://doi.org/10.1007/978-3-319-46448-0_27

Ousingsawat, J., and Earl, M. G. (2007). **Modified Lawn-Mower Search Pattern for Areas Comprised of Weighted Regions.** American Control Conference. https://doi.org/10.1109/ACC.2007.4282850

Papyan, N., Kulhandjian, M., Kulhandjian, H., and Aslanyan, L. (2024**). Ai-based drone assisted human rescue in disaster environments: Challenges and opportunities. Pattern Recognition and Image Analysis**, 34(1), 169-186.

Parrot Drones S.A.S. (2023**). ANAFI USA User Guide (Version 6.10.0.2) [User manual]. Parrot Drones S.A.S**. Retrieved from https://www.parrot.com/assets/s3fs-public/2022-05/ANAFI%20USA%20user%20guide%20GOV.pdf.

Parrot SA. (2021). **FreeFlight 6** (Version 6.7.5) [Mobile application software]. App Store. https://apps.apple.com/us/app/freeflight-6/id1386165299

Petrides, P., Kolios, P., Kyrkou, C., Theocharides, T., and Panayiotou, C. (2017). **Disaster Prevention and Emergency Response Using Unmanned Aerial Systems.** In: 2017, pp. 379–403.

Qu, R., Tan, J., Liu, Y., Li, C., and Jiang, H. (2025). **NoctuDroneNet: Real-Time Semantic Segmentation of Nighttime UAV Imagery in Complex Environments**. Drones, 9(2), 97.

Quaritsch, M., Kruggl, K., Wischounig-Strucl, D., Bhattacharya, S., Shah, M., and Rinner, B. (2010). **Networked UAVs as aerial sensor network for disaster management applications.** Elektrotechnik und Informationstechnik, 127(3), 56–63.

Redmon, J., and Farhadi, A. (2018). Yolov3: **An incremental improvement**. arXiv preprint arXiv:1804.02767.

Restas, A. (2015). **Drone Applications for Supporting Disaster Management.** World Journal of Engineering and Technology, 03, 316–321.

Rizk, M., and Bayad, I. (2023). **Human detection in thermal images using YOLOv8 for search and rescue missions.** https://doi.org/10.1109/ICABME59496.2023.10293139

Roboflow, **Computer vision tools for developers and enterprises**. (2025). https://roboflow.com/

Shahmoradi, J., Talebi, E., Roghanchi, P., and Hassanalian, M. (2020). **A comprehensive review of applications of drone technology in the mining industry.** Drones, 4, 1–25.

Shuranov, M., Shurenkov , D., Ruzhitsky , D., Martynova, V., Bykova, E., and Perevozchikov, G. (2023**). LADD: Lacmus Drone Dataset**. https://www.kaggle.com/datasets/mersico/lacmus-drone-dataset-ladd-v40.

Sruthi, M. S., Poovathingal, M. J., Nandana, V. N., Lakshmi, S., Samshad, M., and Sudeesh, V. S. (2021). **Yolov5 based open-source UAV for human detection during search and rescue (SAR).** 2021 International Conference on Advances in Computing and Communications (ICACC).

Sundelius, N., Funk, P., and Sohlberg, R. A. (2024). **Simulation Environment Evaluating AI Algorithms for Search Missions Using Drone Swarms.** https://doi.org/10.1007/978-3-031-39619-9_14

Tariq, R., Rahim, M., Aslam, N., Bawany, N., and Faseeha, U. (2018). **DronAID: A Smart Human Detection Drone for Rescue.** 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT and IoT (HONET-ICT), pp. 33–37.

Thangavel, P., and Karuppannan, S. (2023). **Dynamic event camera object detection and classification using enhanced YOLO deep learning architecture.** https://doi.org/10.1364/OpticaOpen.24532636

Thanu Thavasi, P., and Suriyakala, C. D. (2012). **Sensors and tracking methods used in wireless sensor network-based unmanned search and rescue system – A review.** Procedia Engineering.

Thuan, P. N., Queralta, J. P., and Westerlund, T. (2023). **Simulation Analysis of Exploration Strategies and UAV Planning for Search and Rescue.** https://doi.org/10.1007/978-3-031-44607-8_5

Tuśnio, N., and Wróblewski, W. (2022). **The Efficiency of Drones Usage for Safety and Rescue Operations in an Open Area: A Case from Poland.** Sustainability, 14(1), 327.

Van Tilburg, C. (2017). **First Report on Using Portable Unmanned Aircraft Systems (Drones) for Search and Rescue.** Wilderness and Environmental Medicine, 28(2), 116–118.

Vincent-Lambert, C., Pretorius, A., & Van Tonder, B. (2023). **Use of unmanned aerial vehicles in wilderness search and rescue operations: A scoping review**. Wilderness & Environmental Medicine, 34(4), 580-588.

Wankmüller, C., Kunovjanek, M., & Mayrgündter, S. (2021). Drones in emergency response–evidence from cross-border, multi-disciplinary usability tests. International Journal of Disaster Risk Reduction, 65, 102567. Yamamura, E. (2010). **Effects of interactions among social capital, income, and learning from experiences of natural disasters: A case from Japan.** Regional Studies, 44(8), 1019–1032.

Yamazaki, Y., Premachandra, C., and Perea, C. J. (2020). **Audio-Processing-Based Human Detection at Disaster Sites with Unmanned Aerial Vehicle.** IEEE Access, 8, 101398–101405.

Yang, L., Hao, Y., Xu, J., and Li, M. (2024). **Multi-UAV Collaborative Target Search Method in Unknown Dynamic Environment.** Sensors. https://doi.org/10.3390/s24237639

Zhu, P., Wen, L., Du, D., Bian, X., Fan, H., Hu, Q., and Ling, H. (2022**). Detection and tracking meet drones challenge**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(11), 7380–7399. https://doi.org/10.1109/TPAMI.2021.3119563

Zou, X., Zhang, G., and Wang, J. (2017). **Understanding the key risks in construction projects in China.** International Journal of Project Management, 35(4), 663–674.

استخدام سرب من الطائرات دون طيار و تعلم الآلة للبحث والإنقاذ في الكوارث الطبيعية

إعداد

البراء عزام محمد زيد الكيلاني

المشرف

الأستاذ الدكتور غيث علي عبندة

المشرف المساعد

الدكتور يزن مازن الزين

الملخص

تتناول هذه الرسالة استخدام سرب من الطائرات بدون طيار ومنهجيات التعلم الآلي المتقدمة لتعزيز عمليات البحث والإنقاذ، لا سيما في حالات الاستجابة للكوارث الطبيعية في الأردن. يتناول البحث بشكل منهجي تحسين نشر الطائرات بدون طيار عبر خوارزميات تخطيط المسارات الفعّالة، مظهرًا فعالية مسار جزازات العشب (Lawnmower) في تقليل تداخل التغطيات وتعزيز الكفاءة الكشفية للطائرات. كما تم إنشاء علاقة رياضية بين ارتفاع الطائرة بدون طيار ومواصفات الكاميرا، مما حسّن بشكل كبير وضوح الكشف وقدم إرشادات تشغيلية عملية.

أدت التجارب المكثفة إلى اختيار نموذج التعلم الآلي YOLOv8، المسبق التدريب على مجموعات بيانات مخصصة للطائرات بدون طيار (VisDrone) ومن ثم تدريبه على البيانات التي تم جمعها من منطقة الظليل شمال شرق مدينة الزرقاء الأردنية، وتم تعزيزها بواسطة تعزيز الفسيفساء (Mosaic Augmentation)، محققًا أداءً متميزًا في الكشف عن الأشخاص وتميزهم في الصور باسترجاع (Recall) بنسبة 97.6٪ ودقة (Precision) بنسبة 97.0٪.

على الرغم من هذه التطورات، تم تحديد عدة قيود، منها مدى الطائرة بدون طيار المحدود، وزمن تأخر المعالجة للصور المرسلة من قبل الطائرة الى وحدة القيادة والتحكم، وكفاءة نموذج التعلم الآلي في ظل ظروف بيئية قاسية. وتقترح اتجاهات البحث المستقبلية التركيز على التطبيق المادي والاختبارات الميدانية الموسعة للنظام، وتطوير خوارزميات تكيف ارتفاع الطائرات بدون طيار، ودمج مجموعات بيانات بيئية متنوعة، واستكشاف حلول المعالجة الهجينة، واستخدام تقنيات اتصالات متقدمة. وتهدف هذه التوصيات إلى مزيد من تحسين قدرات البحث والإنقاذ وتعزيز سرعة الاستجابة التشغيلية، بما يسهم بشكل كبير في فعالية الاستجابة للكوارث وإنقاذ الأرواح.