Universität Zürich[UZH]

Department of Geography

Drone Technology for Early Detection of Smoke in Forest Areas

GEO 511 Master's Thesis

Author: Luca Capitanio, 18-754-721

Supervised by: Dr. Yves Bühler (buehler@slf.ch), Dr. Céline Portenier

Faculty representative: Prof. Dr. Alexander Damm

25.04.2025

Acknowledgements

I would like to express my genuine gratitude to Dr. Céline Portenier and Dr. Yves Bühler for their constant support throughout the development of this thesis. As my supervisors, they guided me through both the theoretical aspects and the practical execution of this research. Their expertise, encouragement, and patience were indispensable, and I am deeply thankful of the time and effort they dedicated to supporting me at every stage of this project.

I would also like to extend my sincere thanks to Prof. Dr. Damm-Reiser, who provided important feedback and played a key role in the final revision of this thesis. His critical insights and attention to detail were precious in refining and improving the quality of this work. As the faculty representative, his support in ensuring the academic rigor of the thesis was crucial, and I am really glad for his contributions.

I would like to express my sincere gratitude to all the volunteers who generously contributed their time and energy to participate in the data collection campaign in Davos. Their dedication, particularly in assisting with the ignition, maintenance, and safe extinguishing of the controlled fires, was essential to the success of this research. Without their help, this project would not have been possible.

I must thank my family and my girlfriend. They have always been by my side, providing solid support in every personal, academic, and professional decision I have made. Their love and encouragement have been a constant source of strength, and I am grateful for their presence in my life.

Finally, I want to thank all the people who have been part of my university journey and shared unforgettable moments with me. Whether through collaborative work, conversations, or simply sharing experiences, these memories have enriched this period of my life, making it all the further meaningful. I am truly thankful for their presence and support along the way.

Thank you all for your contributions and for being part of this journey.

Abstract

Wildfires are becoming increasingly frequent and destructive, posing significant risks to ecosystems, human health, and infrastructure. Timely and accurate detection of wildfire smoke is crucial for early intervention, which can mitigate the devastating effects of these fires. This research explores the use of Unmanned Aerial Vehicles (UAVs) equipped with optical sensors and deep learning models, specifically YOLOv8, to detect wildfire smoke in challenging environments. The primary objective of this study is to evaluate whether optical sensors, when paired with YOLOv8, can provide a viable and cost-effective alternative to traditional thermal sensors for early smoke detection, particularly in areas with dense vegetation, rugged terrain, and varying environmental conditions.

Field experiments were conducted in the Büelenwald region of Switzerland, where controlled wildfire scenarios were simulated to generate diverse smoke conditions. The YOLOv8 model, known for its real-time object detection capabilities, was trained and tested on this specialized dataset to assess its performance in detecting smoke at early stages. The study focuses on optimizing detection accuracy through various techniques, including multi-rotation image analysis, area thresholding, and fine-tuning of the detection algorithm.

The findings suggest that optical sensors equipped with YOLOv8 can reliably detect smoke, offering significant advantages in terms of operational feasibility, speed, and cost-effectiveness compared to traditional thermal sensors. However, the study also highlights the limitations of optical sensors under certain conditions, such as low visibility and adverse weather. By examining the influence of environmental factors like weather, terrain, and vegetation on detection performance, the research provides valuable insights into improving the reliability of smoke detection systems.

This research contributes to the development of advanced, real-time wildfire monitoring solutions, demonstrating the potential of UAVs, deep learning, and optical sensors in enhancing wildfire detection and management. The results offer a pathway for future applications in wildfire monitoring, providing scalable and adaptive detection strategies that can help mitigate the growing risks posed by wildfires.

Table of Contents

Abbreviations

AI	Artificial Intelligence
AP	Average Precision
BA	Balanced Accuracy
CNN	Convolutional Neural Network
DFL	Distribution Focal Loss
FN	False Negative
FP	False Positive
GAN	Generative Adversarial Network
GOES-16	Geostationary Operational Environmental Satellite 16
GPU	Graphics Processing Unit
GPS	Global Positioning System
GRU	Gated Recurrent Unit
IoT	Internet of Things
IoU	Intersection over Unit
IR	Infrared
LSTM	Long Short-Term Memory
LWIR	Long Wavelength Infrared
MAP	Mean Average Precision
MODIS	Moderate Resolution Imaging Spectroradiometer
MTG-I1	Meteosat Third Generation Imager
MWIR	Middle Wavelength Infrared
PAHs	Polycyclic Aromatic Hydrocarbons
PM2.5	Particulate Matter
PTSD	Post-Traumatic Stress Disorder
RGB	Red Green Blue
RGB-D	Red Green Blue - Depth
SAOP	Situation Assessment and Observation Planning

SLF	Swiss Federal Institute for Snow and Avalanche Research
SSD	Single Shot MultiBox Detector
TN	True Negative
TP	True Positive
UAV	Unmanned Aerial Vehicle
US	United States
UZH	University of Zurich
RNN	Recurrent Neural Network
VIIRS	Visible Infrared Imaging Radiometer Suite
WUI	Wildland-Urban Interface
YOLO	You Only Look Once

List of Figures

List of Tables

1. Introduction

1.1 The Growing Threat of Wildfires

Wildfires are a complex and dynamic natural phenomenon, playing a critical role in shaping ecosystems by influencing vegetation patterns, nutrient cycling, and biodiversity (Akhloufi et al., 2021). Historically, wildfires have contributed to the resilience of many ecosystems, where periodic fires naturally clear underbrush, maintain soil fertility and open spaces for new vegetation growth (Benzerkri et al., 2020). However, the increasing frequency, size and intensity of wildfires are shifting this natural balance, resulting in extensive ecological and human impacts (Benzerkri et al., 2020). Factors such as climate change, land-use changes, and fire suppression practices have amplified the wildfire risk, causing fires to deviate from traditional, ecologically beneficial patterns toward more destructive outcomes (Barmpoutis & Stathaki, 2020; Benzerki et al., 2020).

One of the most significant drivers of intensified wildfire activity is climate change (Aslan et al., 2019). Rising global temperatures have led to extended fire seasons, particularly in fire-prone regions such as California (Boroujeni et al., 2024), Australia (Chuvieco et al., 2020), and the Mediterranean (Benzekri et al., 2020). In recent decades, these regions have experienced record-breaking fire seasons, with fires burning millions of hectares and causing unprecedented economic losses (Dimitropoulos et al., 2016). California's extreme wildfire activity in 2017 and 2018 resulted in $40 billion in damages and exposed millions to harmful air quality conditions, raising serious public health concerns (Boroujeni et al., 2024). Similarly, the 2025 wildfires, particularly those in Los Angeles and surrounding areas, caused devastating economic losses estimated around $250 billions and forced the evacuation of over 180'000 residents (Qiu et al., 2025). In southern Europe, intense fire seasons are increasingly common, with areas like Portugal and Greece experiencing some of the most destructive fires on record due to hotter, drier summers and changes in precipitation patterns (Benzekri et al., 2020). These shifts contribute to drier vegetation, a critical factor that makes landscapes more susceptible to ignition and rapid fire spread (Bouguettaya et al., 2022).

Figure 1.1: Aerial view of a neighborhood of Santa Rosa, California, after the wildfires of 2017 [Photo by Josh Fields on pexels.com]

Wildfires also have deep and lasting consequences on soil health and ecosystem services, especially in forested and mountainous regions (Barmpoutis & Stathaki, 2020). The temperature and duration of fires substantially impact the post-fire recovery of soils, often leading to altered physical and chemical properties that affect soil fertility, water retention, and erosion rates (Barmpoutis et al., 2020). Intense fires can destroy organic matter in the soil, which is crucial for nutrient cycling and expose topsoil layers to erosion, further complicating the natural recovery process (Barmpoutis et al., 2020). In regions like the Alps, where slopes are steep and ecosystems unstable, such post-fire erosion can be especially damaging, creating downstream effects on freshwater systems due to ash and sediment runoff (Burke et al., 2021). Aquatic ecosystems are particularly vulnerable to these changes, as runoff can alter water chemistry, increase turbidity, and harm freshwater species, underscoring the connection of terrestrial and aquatic impacts in fire-affected regions (Akhloufi et al., 2021).

Human factors, such as urban expansion into the wildland-urban interface (WUI), the zone where developed areas intermingle with natural landscapes, increase fire hazards and complicate management efforts (Barmpoutis & Stathaki, 2020). In the United States (US), the number of homes in high-risk fire zones has significantly grown, placing more structures and people in danger while driving up control costs as firefighting efforts prioritize private property protection (Barmpoutis et al., 2020). The expansion of the WUI has further contributed to the

rising costs on wildfire response, now reaching billions annually in fire-prone regions like the western US (Barmpoutis & Stathaki, 2020).

The social and health impacts of wildfires are significant and far-reaching (Chuvieco et al., 2020). Exposure to wildfire smoke, which contains harmful pollutants like polycyclic aromatic hydrocarbons (PAHs) and particulate matter (PM2.5), poses substantial health risks (Bailon-Ruiz et al., 2022). These pollutants can aggravate respiratory conditions, contribute to cardiovascular issues and over time, increase cancer risk (Bailon-Ruiz et al., 2022). The smoke's spread, reaching well beyond the immediate fire zones, has become a pressing environmental and health issue, with recent fires in North America affecting air quality thousands of kilometers away (Chen et al., 2018). This exposure has also been linked to mental health repercussions, with survivors often experiencing post-traumatic stress disorder (PTSD) and eco-anxiety due to the traumatic experience of fire-related displacement or property loss (Aslan et al., 2019).

Despite the historical low wildfire activity in some regions, such as Switzerland's Alpine areas, changing climate and land-use patterns are increasing fire propensity, highlighting the need for region-specific adaptation strategies (Wastl et al., 2013). Alpine regions are experiencing increased fire risks due to drier winters and summers, driven by climate warming and altered precipitation regimes (Wastl et al., 2013). Even in regions like the Swiss Alps, specific areas like Valais and Ticino face heightened fire risk due to local topography and climate (Schumacher et al., 2006). For example, Ticino, located south of the Alps, experiences a Mediterranean-influenced climate with dry winters, making it susceptible to wildfires, especially under strong wind conditions (Schumacher et al., 2006). In contrast, the inner Alpine valleys of Valais, sheltered from precipitation by surrounding mountains, also face heightened risk due to drying trends and land abandonment that leave vegetation unmanaged and prone to fire spread (Zumbrunnen et al., 2011). The CH2018 scenarios indicates that near-surface air temperature will rise in the summer months by 2.5°C to 4.5°C in comparison to today, showing up to a quarter less rainfall and making so the topic of wildfire monitoring of crucial relevance for the future of our region (Fischer et al., 2019).

As climate change continues to intensify, future projections indicate that extreme fire seasons may become increasingly common worldwide, making effective fire management and mitigation critical (Chuvieco et al., 2020). However, current fire management practices face challenges due to the complexity of wildfire dynamics and regional differences in climate and

land use (Crowley et al., 2023). Additionally, socioeconomic factors such as population density and local economic activity, can variably impact fire risk, influencing factors like forest management, land use, and infrastructure investment (Chuvieco et al., 2020).

1.2 Conventional Monitoring

Conventional wildfire monitoring methods play a crucial role in mitigating the extensive impacts of wildfires on ecosystems, economies, and public safety. As wildfires contribute to significant carbon emissions and threaten biodiversity, the need for effective detection and response systems has grown, emphasizing the role of monitoring technologies in fire management strategies (Okoro et al., 2024). Several approaches, including satellite remote sensing, airborne monitoring, ground patrols, and ground-based sensor networks, have proven crucial in this field, each offering distinct advantages and facing individual limitations.

Satellite-based remote sensing has become fundamental in wildfire management, providing a continuous overview of large geographic areas (Okoro et al., 2024). Since the early 1970s, when Landsat satellite data became available, satellites have enabled multitemporal and multispectral data analysis for pre-fire vegetation assessment, active fire detection, and post-fire recovery mapping (Chuvieco et al., 2020). Systems such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS), equipped with optical and thermal sensors, allow scientists and fire managers to monitor vegetation conditions, detect hotspots, and track fire spread (Kang et al., 2023). However, polar-orbiting satellites like MODIS and VIIRS revisit specific areas only once or twice daily, resulting in significant temporal delays in data collection (Kang et al., 2023). This limits their utility in wildfire events in near real-time, particularly during peak fire activity periods, which can delay timely decision-making and response efforts (Kang et al., 2023).

To address these temporal limitations, geostationary satellites like the Geostationary Operational Environmental Satellite 16 (GOES-16) and the Meteosat Third Generation Imager (MTG-I1) provide near real-time updates every five to ten minutes, making them highly effective for monitoring active fires and detecting new ignitions across wide landscapes (Xu et. Al, 2021). However, the lower spatial resolution of geostationary satellites, typically around 2 kilometers, reduces their sensitivity to smaller fires or fires within dense vegetation, potentially delaying initial detection (Bushnaq et al., 2021). Atmospheric interference further complicates satellite monitoring; cloud cover, humidity, and dust can obscure fire signals, while

sun-heated surfaces may trigger false alarms (Kang et al., 2023). These factors collectively limit the reliability of satellite data in regions with complex weather patterns, requiring satellite-based monitoring to be supplemented by other technologies for a more accurate and responsive fire management approach.

Manned aircraft equipped with high-resolution thermal and optical cameras offer a flexible, high-resolution method for assessing fire characteristics, spread, and intensity in real time (Bailon et al., 2022). This capability is particularly useful for targeted monitoring in high-risk areas or when satellite data is insufficient. Airborne monitoring enables fire managers to gather precise data to inform immediate containment strategies, thus improving fire spread models (Thangavel et. al, 2022). However, the high operational costs associated with fuel, maintenance and specialized equipment, as well as the need for trained personnel, often restrict the deployment of aircraft to critical fire events (Crowley et al., 2023). Furthermore, environmental factors, such as flight range, duration and adverse weather conditions, limit the frequency and geographical reach of airborne monitoring (Crowley et al., 2023). These constraints, combined with safety risks during extreme fire conditions, make airborne monitoring an important yet resource-intensive part of wildfire management (Crowley et al., 2023).

Ground patrols provide on-the-ground information about fire behavior and environmental conditions, often verifying satellite or airborne data (Syaufina et al., 2022). These patrols, usually conducted by trained personnel with mobile applications and Global Positioning System (GPS) devices, deliver localized data on fire intensity, spread, and fuel moisture (Syaufina et al., 2022). This real-time, high-precision data is crucial in assessing conditions that remote sensors may not capture, including local wind speed, humidity, and terrain features that influence fire dynamics (Syaufina et al., 2022). However, the possibility of ground patrols is limited by the time required to traverse rugged or extensive terrain, resulting in portions of fire-prone areas potentially remaining unmonitored, especially during peak fire seasons when personnel may be distributed across multiple areas (Syaufina et al., 2022). Logistical challenges and safety risks in active fire zones further hinder the coverage and efficiency of ground patrols, underscoring the need for complementary monitoring methods to provide broader spatial coverage (Crowley et al., 2023).

High-resolution cameras, often set up on watchtowers, are progressively used for ground-based wildfire detection. These setups usually combine both optical and thermal imaging and are frequently paired with sensors to improve detection accuracy (Okoro et al., 2024). Combined,

these sensors can successfully identify early signs of fire, such as smoke and temperature increases, especially in close-proximity situations like buildings or small forested regions (Okoro et al., 2024). However, these systems are constrained by limited range and visibility, requiring careful positioning and many installations to monitor extensive forest areas effectively (Okoro et al., 2024). The costs of installing and maintaining such sensor networks can be quite high, especially for large or remote areas, making them less feasible as the primary monitoring approach for wide landscapes. (Bouguettaya et al., 2023)

Each conventional wildfire monitoring method offers strengths and limitations. Satellite systems offer large-scale monitoring, while airborne observations provide targeted, high-resolution data, and ground patrols deliver localized, real-time insights. Fixed terrestrial cameras add an additional layer of monitoring but require considerable investment to achieve widespread coverage. Together, these technologies contribute to a multi-tiered wildfire management approach that, despite constraints, enhances fire prevention, detection, and response capabilities, supporting broader efforts to moderate the extensive damage caused by wildfires.

1.3 UAVs Monitoring

1.3.1 Drones for Wildfire Detection

The use of Unmanned Aerial Vehicles (UAVs) in wildfire detection and management has expanded significantly in recent years due to their versatility, cost-effectiveness, and ability to collect real-time data in otherwise inaccessible areas (Okoro et al., 2024). UAVs have become priceless for developing early fire detection, monitoring fire progression, and assessing post-fire impacts, all of which are essential for mitigating the devastating social, economic, and environmental effects of wildfires (Akhloufi et al., 2021).

One of the primary reasons UAVs have gained traction in wildfire management is their ability to operate in real time, providing continuous, high-resolution imagery and data on active fires (Bushnaq et al., 2021). Unlike satellites, which are limited by their orbit schedules and spatial resolution, UAVs can be deployed over specific zones of interest, capturing data on fire size, perimeter, and progression (Bailon et al., 2022). UAVs can be equipped with various sensors, including optical and infrared cameras, allowing them to capture thermal signatures and visible light data, which are crucial for assessing fire severity and predicting its spread (Boroujeni et al., 2024). The operational flexibility of UAVs is another advantage. They can be deployed by

a single operator and navigate complex terrains, reducing human exposure to hazardous conditions (Viseras et al., 2021). Furthermore, UAVs are less expensive to operate than helicopters or airplanes, making them an attractive option for budget-constrained fire management agencies (Akhloufi et al., 2021).

The development of autonomous UAV fleets has further boosted their utility in wildfire management (Viseras et al., 2021). Coordinated UAVs working in swarms can cover large areas, enabling continuous monitoring across a fire's span and supporting real-time data integration for extensive fire mapping (Okoro et al., 2024). Advances in machine learning and deep learning have increased the capacity of these fleets to operate autonomously and coordinate effectively (Okoro et al., 2024). These algorithms enable UAVs to navigate dynamically changing fire conditions, optimize their flight paths, and efficiently collect data with minimal human oversight (Okoro et al., 2024). The Situation Assessment and Observation Planning (SAOP) system, for instance, exemplifies how multiple UAVs can work in unison to generate fire maps and make predictive assessments of fire spread (Bailon et al,. 2022).

Moreover, UAVs have proven valuable for pre-fire risk assessments and post-fire damage evaluations, complementing their active fire monitoring role. In pre-fire stages, UAVs help assess vegetation dryness, fuel load, and other environmental variables that influence fire risk, enabling more accurate predictions of fire-prone areas (Boroujeni et al., 2024). Post-fire, UAVs can survey the impacted areas to document ecological damage, soil erosion, and changes in vegetation, which informs recovery works and risk assessment for future fires (Akhloufi et al., 2021).

Despite their benefits, UAVs face technical and operational challenges. Flight duration and battery life remain limiting factors, especially for small UAVs tasked with extended missions (Boroujeni et al., 2024). Advances in battery technology and energy-efficient flight algorithms are actively addressing these limitations (Viseras et al., 2021). The need for real-time processing also places demand on UAVs' computational capabilities, often requiring on-board data processing or dependence on ground-based systems for processing large volumes of data (Okoro et. al, 2024). In addition, effective deployment of UAVs requires robust communication networks, particularly in remote areas where wildfire management is needed most. Integrating UAVs with Internet of Things (IoT) networks can improve their reach and functionality, enabling UAVs to serve as aerial data relays for ground-based sensors in difficult-to-access areas (Bushnaq et al., 2021).

UAVs have transformed wildfire management by enabling rapid, targeted, and flexible monitoring solutions (Bailon et al., 2022). Their ability to operate autonomously and deliver real-time data develops fire detection and monitoring, ultimately supporting more efficient and informed wildfire management (Viseras et al., 2021). As these technologies continue to evolve, the integration of UAVs with Artificial Intelligence (AI)-driven algorithms and IoT infrastructure promises to further boost the efficiency in wildfire management strategies, offering the opportunity for more adaptive and responsive fire monitoring systems (Bushnaq et al., 2021).

1.3.2 Thermal vs Optical Cameras

Thermal cameras have become an effective tool in various fields for detecting and analyzing heat patterns that are invisible to the human eye (Bouguettaya et al., 2022). By acquiring the infrared (IR) radiation emitted from objects, thermal cameras generate images based on heat signatures rather than visible light. These cameras, often operating within the Middle Wavelength InfraRed (MWIR, 3’000 to 5’000 nm) and Long Wavelength InfraRed (LWIR, 8’000 to 14’000 nm) spectral ranges, have been proved useful in applications like firefighting, wildlife monitoring, search and rescue and industrial inspections (Bouguettaya et al., 2022). Thermal cameras offer several benefits that make them suitable for wildfire detection, especially when mounted on UAVs (Shamsoshoara et al., 2021). Unlike optical cameras, which rely on visible light, thermal cameras detect infrared radiation, allowing them to capture heat signatures regardless of lighting conditions. This capability enables UAVs to detect fires in low-visibility scenarios, such as darkness or smoke-filled environments, making them a powerful tool for continuous wildfire monitoring (Sousa et al., 2019). Additionally, thermal cameras help distinguish fires from other visible elements that may mimic flames, such as sunlight reflections or red-colored objects (Sousa et al., 2019). UAV-mounted thermal cameras can scan large areas quickly and relay high-resolution thermal imagery, allowing firefighters to detect hotspots, track fire movement, and assess surrounding vegetation for potential risks. In recent studies, researchers like Shamsoshoara et al. (2021) have demonstrated the effectiveness of UAV-based thermal cameras in detecting early-stage fire outbreaks, which could significantly enhance early warning systems and facilitate rapid responses.

Despite these advantages, thermal cameras also have limitations when employed for wildfire monitoring. One significant issue is that dense tree canopies can obstruct the view of ground-

level heat sources, making it challenging to detect fires in forested areas. This limitation means that fires underneath thick foliage may not reach the thermal sensors on UAVs, potentially delaying detection until the flames breach the canopy or expand to exposed areas. Another challenge is distinguishing wildfire heat signatures from other naturally heated objects in the environment. During summer months, materials such as rocks or metal objects absorb considerable solar radiation, reaching temperatures high enough to produce misleading thermal signatures. These false positives can confuse fire detection, leading to inefficient deployment of firefighting resources.

Optical cameras play a central role in wildfire monitoring by capturing images within the visible spectrum (400 to 700 nm). These cameras, also known as RGB cameras, are favored for their affordability, high spatial resolution, ease of use, and lightweight design, making them well-suited for small UAV platforms and low-cost applications in forestry (Chen et al., 2018). UAVs equipped with optical cameras can acquire high-resolution imagery, enabling the detection of smoke and flames in the early stages of a wildfire, mostly under favorable visibility conditions. These abilities make optical cameras valuable for a range of ecological monitoring tasks, including initial fire and smoke identification from the air (Barmpoutis et al., 2020). Optical cameras offer multiple benefits that have popularized their use in wildfire detection. Their high spatial resolution allows them to capture fine details, which can be critical for identifying early signs of wildfire spread. This detail is particularly advantageous for smaller UAVs engaged in forest monitoring and ecological studies (Barmpoutis et al., 2020). Advances in optical camera technology have further enhanced their functionality. For example, CMOS 360° cameras mounted on UAVs, as explored by researchers like Barmpoutis & Stathaki (2020), provide an expansive field of view, overcoming the limited range of traditional optical sensors and reducing the need for multiple cameras to achieve comprehensive area coverage. Additionally, RGB-D cameras, combining visible spectrum imaging with depth sensing, add another layer of data, allowing UAVs to estimate the height and extent of flames, as demonstrated in wildfire studies by Novac et al. (2020). Their capability to capture fine details makes them particularly effective in spotting early fire indicators, such as smoke plumes emerging from beneath tree canopies, where thermal imaging might fail to detect ground-level hotspots.

Despite these advantages, optical cameras on UAVs face considerable limitations in wildfire monitoring. Their dependence on visible light restricts their effectiveness in low-visibility scenarios, such as nighttime or dense smoke conditions, making them inadequate for

continuous monitoring in challenging environments. Detecting fires in densely forested areas is also problematic, as heavy canopy cover can obstruct views of flames and smoke, preventing UAV-mounted optical cameras from accurately identifying ground-level fire activity. Environmental factors such as sunlight angles, cloud cover, and shadows can further complicate image analysis, creating false positives or obscuring actual fire signatures (Govil et al., 2020). Optical cameras alone are not sufficient for comprehensive wildfire detection, especially in low-visibility circumstances. In contrast, thermal cameras excel in detecting active flames and heat signatures regardless of lighting conditions. Consequently, an integrated approach combining both thermal and optical sensors is needed for a more robust and reliable wildfire monitoring system. By leveraging the strengths of each technology, UAV-based wildfire detection can achieve greater accuracy, earlier detection, and improved situational awareness for firefighters and emergency responders.

1.3.3 AI-Driven Detection

Computer vision, which provides machines with the ability to visually interpret environments, has become a keystone in the application of UAVs for early wildfire detection (Novac et al., 2020). Two primary methods dominate computer vision approaches to wildfire detection: traditional machine learning and deep learning. Traditional machine learning techniques focus on extracting key features manually, relying on color transformations and predefined characteristics (Badmouths et al., 2020). However, these methods require extensive expert input for feature selection and often do not perform optimally in complex wildfire environments, such as dense forests, where flames or smoke may be obscured by vegetation or affected by background noise (Jadon et al., 2019).

In contrast, deep learning techniques use powerful neural networks to automate feature extraction, an innovation made possible by progresses in Convolutional Neural Network (CNN) architectures and enhanced hardware like Graphics Processing Units (GPUs), along with software platforms like TensorFlow and PyTorch (Badmouths et al., 2020). Today, deep learning models can address the complex patterns in composite images and are widely applied in image processing tasks across diverse fields, from self-driving vehicles (Totakura et al., 2021) to plant disease identification (Saleem et al., 2019). In wildfire detection, deep learning has excelled by overcoming traditional methods' limitations, processing complex variations in visual data such as shifting lighting conditions, occlusions from dense tree canopies, and

varying fire appearances, ultimately enabling UAVs to detect fires with greater precision (Lee et al., 2017). Deep learning algorithms have transformed UAV-based wildfire detection due to their ability to adapt to variations in the visual characteristics of flames and smoke (Zhao et al., 2018). In wildfire scenarios, CNNs are particularly well-suited for the challenges of aerial imagery due to their capacity to capture detailed spatial features even in complex, chaotic scenes. For instance, CNNs have proven effective for detecting both flames and smoke, which are critical visual cues in identifying early fire stages (Bouguettaya et al., 2022). While some studies prioritize flame detection for its direct association with fire (Goyal et al., 2020), others focus on smoke, which can indicate fire even in dense, forested environments where flames may be hidden (Zhang et al., 2016). Recent research suggests that models capable of detecting both flame and smoke simultaneously are better suited for real-world applications, as this dual focus helps address detection gaps that arise when relying on only one feature (Hossain et al., 2020).

Wildfire detection via deep learning can be divided into three core methods: image classification, object detection, and semantic segmentation, each offering specific strengths in different wildfire monitoring contexts:

Image classification models categorize input images into predefined classes, such as fire versus non-fire. CNNs are the preferred architecture due to their capability to handle high-dimensional data, accurately extracting meaningful features from 2D images (Novac et al., 2020). Lee et al. (2017) used AlexNet, GoogLeNet, and VGG-Net on UAV imagery, achieving accuracy rates of 94.8% to 99%, underscoring CNNs' effectiveness in distinguishing fire from non-fire images. Although highly accurate, image classification can struggle with smaller fire features, which may limit its application in early detection scenarios where fire starts as small spots (Novac et al., 2020).

Object detection algorithms go beyond classification by identifying the location of fire features within images, drawing bounding boxes around flames or smoke (Solovyev et al., 2021). This spatial data is crucial for situational awareness, enabling responders to assess fire spread more accurately. Object detection methods are split into two categories: two-stage models, such as R-CNN (Region-based CNN), which segment images in multiple steps, and single-stage models, like You Only Look Once (YOLO) and Single Shot MultiBox Detector (SSD), which process images in one pass, optimizing for speed (Yadav, 2020). YOLO, has gained popularity for wildfire detection due to its real-time capabilities; YOLOv2, for instance, demonstrated

98.3% accuracy and 99.14% F1-score in a study on smoke detection from UAV imagery (Yadav, 2020). These results highlight YOLO's efficacy for rapid-response applications, allowing responders to access critical fire data without delay.

For tasks requiring pixel-level precision, semantic segmentation models classify each pixel in an image, which allows for detailed mapping of fire, smoke, and vegetation (Bouguettaya et al., 2022). Unlike object detection that uses bounding boxes, semantic segmentation assigns each pixel to a class, making it more suitable for complex wildfire scenes where fire contours are essential. DeepLab and U-Net are commonly used architectures in segmentation tasks, with DeepLabV3+ showing success in separating fire and smoke pixels from background vegetation in UAV imagery (Barmpoutis et al., 2020). In addition to CNNs, advanced architectures like Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) have introduced new dimensions to wildfire detection (Benzekri et al., 2020). RNNs, particularly Long Short-Term Memory (LSTM) networks, manage temporal dependencies in video data, adding a memory component useful for tracking fire progression across frames (Bouguettaya et al., 2022). Benzekri et al. (2020) applied RNN models to sensor data from forest monitoring networks, achieving up to 99.89% accuracy with Gated Recurrent Unit (GRU) and LSTM models, which underscores the potential of RNNs in sequential data analysis for active fire monitoring.

GANs, on the other hand, have become crucial for data augmentation in deep learning, creating synthetic wildfire images to supplement limited datasets (Aslan et al., 2019). These synthetic images aid in training robust models that can generalize better across diverse wildfire scenarios. Aslan et al. (2019) employed GANs to generate realistic fire images for training, showing the potential of this technique to reinforce model performance in low-data environments (Aslan et al., 2019). As wildfires are unpredictable and data collection can be challenging, GANs provide a practical solution for enhancing dataset size and diversity.

Despite the advances, deep learning approaches face challenges, particularly in complex wildfire scenarios. UAV imagery often encounters issues such as variations in smoke and flame appearance, occlusions from trees or structures, and changing light conditions throughout the day (Badmouths et al., 2020). Furthermore, segmentation models require extensive computational resources for training and inference, making real-time application challenging in remote or resource-limited settings (Bouguettaya et al., 2022). The choice of architecture, training dataset, and feature extraction techniques must be carefully considered to optimize

model performance while minimizing processing time and computational load (Zhang et al., 2016).

Deep learning-based computer vision has expanded the capabilities of UAVs in wildfire management, with each method offering specific advantages. While CNN-based models are prevalent for accurate classification and detection, advanced architectures like RNNs and GANs provide new paths for handling dynamic data and increasing model robustness. By combining these methods within UAV frameworks, researchers and practitioners can address the urgent need for early wildfire detection and response in a way that is both precise and adaptive to complex, real-world conditions.

1.4 Research Objectives

This research aims to investigate whether optical sensors, when paired with the YOLOv8 deep learning model, can serve as a solid complementary source to thermal sensors for early smoke detection in wildfire scenarios. While numerous smoke detection algorithms already exist, including those based on thermal sensing, optical imagery, and even multispectral data, this study will evaluate and improve upon these methods by testing their performance in challenging field conditions.

The idea for this study stems from a suggestion by Schutz & Rettung Zürich, who, considering the limitations of thermal sensors in the region due to environmental factors such as dense vegetation, rugged terrain, and varying weather conditions, proposed evaluating the effectiveness of optical sensors as an alternative. This study aims to assess the efficiency of optical sensors, particularly in environments where thermal sensors face challenges. To this end, UAV-based flights have been organized to collect data, which will then be employed in this research to explore the potential of optical sensors.

Specifically, this research explores whether the combination of optical sensors and YOLOv8 can overcome the limitations of thermal sensors, especially in environments characterized by dense vegetation, rugged terrain, and varying environmental conditions. YOLOv8 will be employed to assess its potential in detecting smoke at the earliest possible stages of a wildfire, with the goal of minimizing fire spread and reducing its devastating impacts on the environment and human populations.

In addition to evaluating the optical sensor-YOLOv8 combination, this study will consider the influence of environmental factors such as weather conditions, terrain type, and vegetation density on smoke detection performance. These factors often complicate the detection process, and their interaction with both the sensor and model could significantly affect detection timeliness and accuracy. Understanding how these elements influence the system will provide insights into strategies for improving detection reliability in real-world wildfire scenarios.

Moreover, this research will refine existing smoke detection methods by optimizing smoke detection approaches to enhance reliability and reduce false positives. Given the challenges in smoke detection, particularly in dynamic and unpredictable natural environments, this study will apply a series of advanced image processing techniques, such as multi-rotation analysis, area thresholding, and algorithm fine-tuning, to improve the accuracy and specificity of smoke detection. These enhancements aim to ensure that the detection system is not only accurate but also robust enough to handle the complexities of wildfire scenarios.

The research will explore the following key questions:

R1*: Can optical sensors equipped on UAVs, when paired with YOLOv8, provide a reliable and efficient alternative to thermal sensors for early wildfire smoke detection in complex, vegetation-dense environments? What are the strengths and limitations of these methods?*

R2*: What methods, technologies, or approaches can be employed to optimize the quality of smoke detection in complex wildfire scenarios, enhancing reliability and reducing false positives?*

R3: *In what real-world scenarios can the combined use of optical sensors and YOLOv8-based detection systems be most effectively applied, and what potential benefits could this approach offer in terms of early detection, cost-efficiency, and operational feasibility?*

This investigation will provide valuable insights into the feasibility of optical sensors as a complementary tool for wildfire detection. By combining optical sensing technology with advanced machine learning models, this study will contribute to improve the scalability and cost-effectiveness of wildfire detection systems, potentially leading to more efficient and responsive wildfire monitoring. The results will help shape more adaptive wildfire management strategies, which are essential for mitigating the risks posed by increasingly frequent and severe wildfire events worldwide.

2. Methods

2.1 Data Collection

2.1.1 Test Site and Experimental Setup

The data collection for this study has been conducted in the Büelenwald, in the Dischma Valley, located in the Davos area of Switzerland highlighted in Figure 2.1. This region is part of a testing zone supervised by the Swiss Federal Institute for Snow and Avalanche Research (SLF) and spans elevations between 1550 and 2150 meters above sea level. The presence of challenging mountainous terrain and dense vegetation typical of high-altitude forests makes it an ideal location for simulating wildfire scenarios. The site is selected for its relevance to real-world regions prone to wildfires, as its diverse topography and vegetation types closely resemble areas vulnerable to forest fires globally. The high-altitude setting also introduces atmospheric factors, such as variable wind currents and light scattering, which can influence smoke dispersal and visibility, providing valuable data for refining smoke detection algorithms under complex conditions.

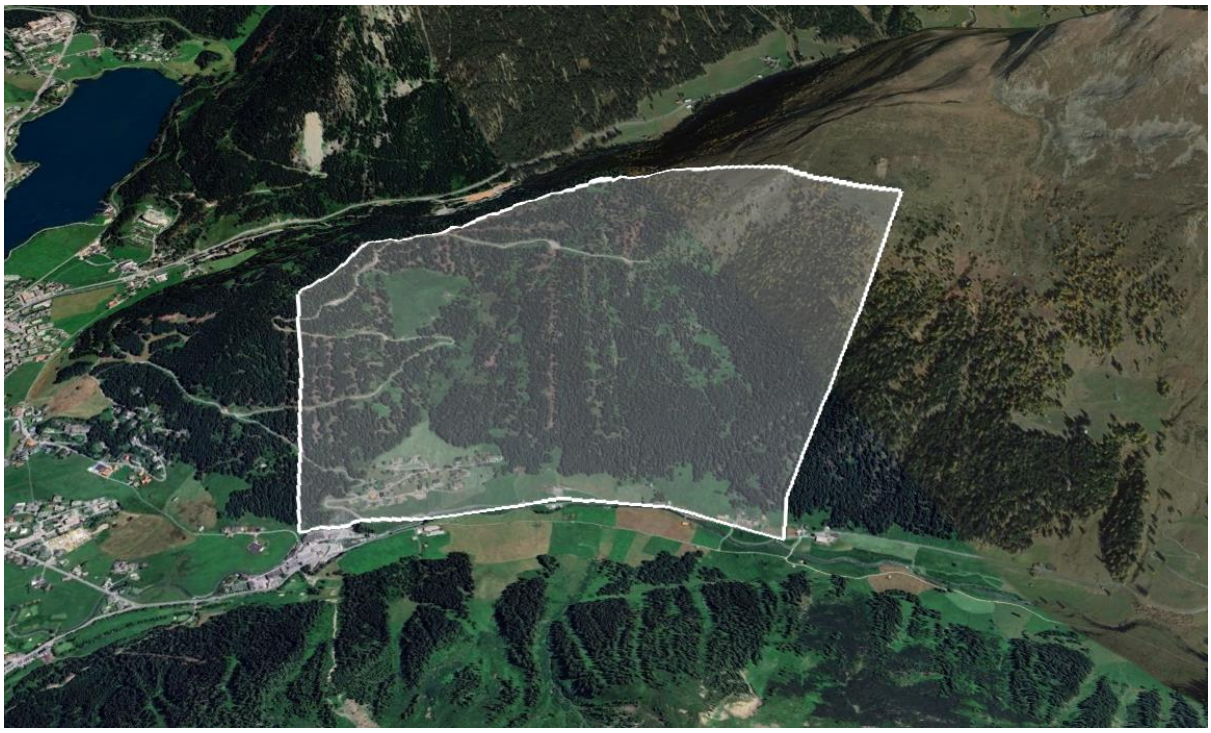

Figure 2.1: The SLF Test Site Büelenwald (highlighted in white), where the data collection campaign was conducted [Background image from Google Earth, © 2025 Airbus. Image reproduced in accordance with Google Earth's terms of use]

Three flights were conducted over 2 days, with carefully designed experimental setups that involved controlled fire scenarios to simulate different wildfire smoke conditions. These fire configurations produced a range of smoke densities, from low-density plumes to more concentrated smoke clouds. This variation was important for gathering diverse data, allowing for thorough evaluation of the algorithm's performance under different smoke intensities. By simulating realistic wildfire scenarios and creating repeatable conditions, the experiment aimed to test the model's ability to detect smoke in various behaviors, setting the stage for both initial assessments and future refinements of the smoke detection system.

The Wingtra One Gen II drone visible in Figure 2.2, equipped with a Sony RX1R II camera, was used for high-resolution data collection, capturing 42-megapixel images ideal for documenting environmental features and smoke dispersion. The camera's high spatial resolution enabled precise identification of smoke plumes, even in areas with obstructing elements like trees and dense foliage. This detailed imagery was needed for analyzing smoke movement and dispersion, particularly in complex terrains with limited visibility. Additionally, the consistency and quality of the images ensured reliable data for algorithm development and testing, supporting the fine-tuning of smoke detection under varying conditions, including fluctuating visibility, lighting, and atmospheric effects. Such precision is crucial for real-world wildfire monitoring, where accurate, real-time data is essential.

Figure 2.2: The Wingtra One Gen II drone employed for the data collection campaign

2.1.2 Drone Flights and Products

The first flight, conducted on the 7th of June 2024, served as an introductory field experiment to evaluate the initial conditions for smoke detection. During this flight, three small fires were ignited at the locations represented in Figure 2.3 across the Büelenwald area. These controlled fires, monitored and extinguished by three individuals, were intentionally spaced to produce distinct smoke plumes within the coverage area, allowing the drone to capture varying smoke patterns in a single flight. While the smoke was intentionally kept at controllable volumes for clarity in initial algorithm testing, the amount produced was limited, and consequently, its visibility was minimal in the collected data. Despite the limited effectiveness for evaluating the algorithm, this first flight provided crucial insights, revealing the need for adjustments to optimize experimental conditions for better smoke production.

This initial session resulted in the acquisition of 1 orthomosaic and 533 orthophotos, documenting smoke dispersal patterns, fire locations, and surrounding vegetation. While these images were not directly used for algorithm evaluation, they placed the foundation for improving the flight setup for future tests.

Figure 2.3: Fire locations applied on the orthomosaic of the covered area obtained during Flight 1

Developing on the insights from the first flight, the second flight, conducted on the 27th of August 2024, adopted a more refined approach. The area for the drone overflight represented in Figure 2.4 was reduced compared to the first flight, allowing for a higher density of data. This flight saw three fires ignited, with two individuals assigned to each fire to increase the volume and continuity of smoke. The additional workforce facilitated the smoke management, ensuring substantial and continuous smoke plumes that could be captured from various angles and altitudes. Focusing on a smaller area allowed for two separate flights on the same day, effectively doubling the volume of data captured. The second flight generated 1 orthomosaic, 544 orthophotos, and 544 raw JPG images.

Later on, the same day, the third flight reflected the conditions of the second flight, ensuring additional data capture in the same controlled area. The purpose was to provide an extended dataset, ensuring that the results were comparable across multiple instances, and to analyze the algorithm's performance under repeated, consistent conditions. As in the second flight, the third flight produced 1 orthomosaic, 544 orthophotos, and 544 raw JPG images, reinforcing the dataset's redundancy and consistency. This extended dataset allowed for a deeper analysis of the smoke detection model’s stability and performance when exposed to similar conditions across both data collection instances.

Figure 2.4: Fire locations applied on the orthomosaic of the covered area obtained during Flight 2 and Flight 3

The inclusion of raw JPG images as a product during Flight 2 and Flight 3 was intentional, as the orthophotos and orthomosaics had limitations in capturing smoke in the desired detail for the specific analysis goals. The orthophotos, due to the applied geometric correction, disrupted the structural continuity of the smoke plume, making it difficult to assess the smoke's extent accurately. The orthomosaic, which involves the merging of multiple captures, often led to the cancellation or distortion of the smoke plume, creating noise that hindered clear identification. These raw JPG images provide a more layered and varied view of the smoke patterns, contributing to a more detailed and precise assessment of the dispersal. Additionally, raw JPG images are the most likely format to be used in real-time detection scenarios, making them particularly suitable for rapid analysis and decision-making in wildfire monitoring.

2.1.3 Operational Considerations

To guarantee the safety and efficiency of the experiments, protocols were implemented throughout each phase of data collection. These protocols were designed not only to safeguard participants but also to maintain the integrity and validity of the data. Local authorities and residents were informed of the planned flights and controlled fires in advance, ensuring that any concerns about smoke sightings were addressed. This active communication helped minimize potential disturbances and confusion among the public, inducing an environment of transparency and cooperation.

Each flight was supervised by Dr. Yves Bühler, Group Leader in Alpine Environment and Natural Hazards at the Swiss Federal Institute for Snow and Avalanche Research (SLF). Dr. Bühler’s knowledge in environmental safety and hazard management ensured that the experiments followed the safety standards. His supervision was fundamental for maintaining high levels of data integrity and environmental conformity throughout the testing process. Moreover, Dr. Bühler’s involvement allowed for real-time adjustments to the flight plans and data collection methods, particularly with regard to managing environmental factors such as wind, temperature and humidity, which could influence the visibility of smoke and the quality of the data collected. The successful execution of the controlled fires was made possible thanks to the assistance of Dr. Portenier, as well as volunteers from both SLF and the University of Zurich (UZH). Their involvement in the ignition, monitoring, and safe extinguishing of the fires ensured that the experimental conditions were carefully controlled, producing the necessary smoke volumes for data collection. These operational considerations underscore the critical importance of guaranteeing safety and maintaining controlled conditions when

collecting quality data. The research team's collaboration with local authorities and the contributions of Dr. Portenier and the volunteers were very important in conducting the experiments smoothly, without causing undue alarm to the community. This careful planning and coordination reflect the study's commitment to responsible and ethical data collection practices, ensuring that the testing procedures aligned with both scientific and social responsibility standards.

The strategy employed in the data collection resulted in a comprehensive dataset. This is essential for developing and refining an advanced UAV-based smoke detection algorithm, capable of detecting smoke in complex, real-world conditions. By conducting flights under a range of managed variables, the team was able to create a dataset that will serve as the foundation for assessing the algorithm's ability to perform reliably in a real world scenario. The consistency of conditions across flights, coupled with the controlled variability in smoke production, will provide a solid basis for evaluating the model's performance.

2.2 Data Analysis

2.2.1 YOLOv8 for Wildfire Detection

YOLO is a real-time object detection framework that offers an ideal balance between high detection accuracy and rapid processing speed, making it one of the most widely adopted models in the field of computer vision (Jiang et al., 2024). Introduced by Redmon (2016), YOLO revolutionized object detection by implementing a single-pass detection method, enabling the analysis of entire images in one forward pass of a CNN. This single-pass approach differentiates YOLO from traditional models that rely on region proposals or multi-stage processing, which can be slower and more computationally demanding (Redmon, 2016). In YOLO, an image is divided into a grid, with each cell responsible for predicting bounding boxes and class probabilities for detected objects in real time (Terven et al., 2023). This grid-based structure allows YOLO to accomplish notable speed, making it ideal for applications requiring real-time detection, such as autonomous driving, surveillance, and wildfire monitoring (Saydirasulovich et al., 2023). One of YOLO's key advantages is its ability to rapidly detect objects without compromising on accuracy. This balance between speed and precision has been progressively refined in subsequent versions, with YOLOv3 incorporating multi-scale prediction layers to improve detection performance, especially for smaller objects at varying scales (Casas et al., 2024). The evolution of YOLO continued with YOLOv4 and

YOLOv5, each introducing significant architectural developments, novel loss functions, and evolved data augmentation techniques (Casas et al., 2024). These improvements have reinforced both the accuracy and flexibility of YOLO across diverse contexts and hardware environments. For example, YOLOv4 integrated data augmentation and bag-of-freebies techniques to boost performance without increasing inference time, making the model particularly suitable for environments with limited processing power (Jiang et al., 2022). YOLOv5 further increased the model's usability by offering multiple sizes, allowing users to adapt YOLO to specific hardware resources, from compact edge devices to high-capacity GPUs (Diwan et al., 2023). One of the most notable strengths of YOLO is its adaptability across different hardware environments and its capacity to be fine-tuned on domain-specific datasets (Jiang et al., 2024). Through transfer learning, YOLO can be customized to specialized tasks, such as smoke detection in wildfire monitoring. This malleability enables YOLO to achieve high accuracy even in challenging conditions by learning to recognize tricky features that might be overlooked by more general models (Jiang et al., 2024). As a result, YOLO not only improves detection accuracy but also reduces false positives (Terven et al., 2023). Its success stems from its speed, accuracy, and scalability, keeping it at the forefront of computer vision (Terven et al., 2023). With each iteration improving on the last, YOLO remains a key tool for fast, reliable object recognition across diverse applications (Redmon, 2016).

To enhance wildfire detection capabilities, a pretrained version of YOLOv8 was employed in this study, specifically adapted and fine-tuned on a specialized dataset to improve its ability to detect smoke and fire. The model was trained by Tleuliyev (2023) using the D-Fire dataset exposed in Table 2.1, which contains around 21,000 labelled images with bounding box annotations for both smoke and fire. Although the D-Fire dataset includes both smoke and fire labels, our study focused exclusively on detecting smoke, as it is typically the first visible sign of a potential wildfire.

Table 2.1: Composition of the D-Fire Dataset employed for the training of the YOLOv8 algorithm in the study conducted by Tleuliyev (2023)

Number of images		Number of bounding boxes	
Category	**# Images**	**Class**	**# Bounding boxes**
Only fire	1'164	Fire	14'692
Only smoke	5'867	Smoke	11'865
Fire and smoke	4'658		
None	9'838		

Prior to the training, a custom configuration file was created to optimize key model parameters specifically for smoke detection. Hyperparameters, such as learning rate, batch size, and the number of epochs, were carefully selected to optimize the model's architecture for precise smoke and fire detection. Parameters like confidence threshold and Intersection over Union (IoU) were also chosen to enhance sensitivity and precision, which are crucial for minimizing false positives. The training was conducted over approximately 130 epochs, ensuring the model achieved convergence at an optimal level of accuracy. This extent provided enough time for the model to stabilize, preventing overfitting while enabling it to learn a strong representation of smoke features. Monitoring logs and convergence graphs during the training process confirmed that the model was on track to meet its target accuracy, as showed by the training metrics in Figure 2.5.

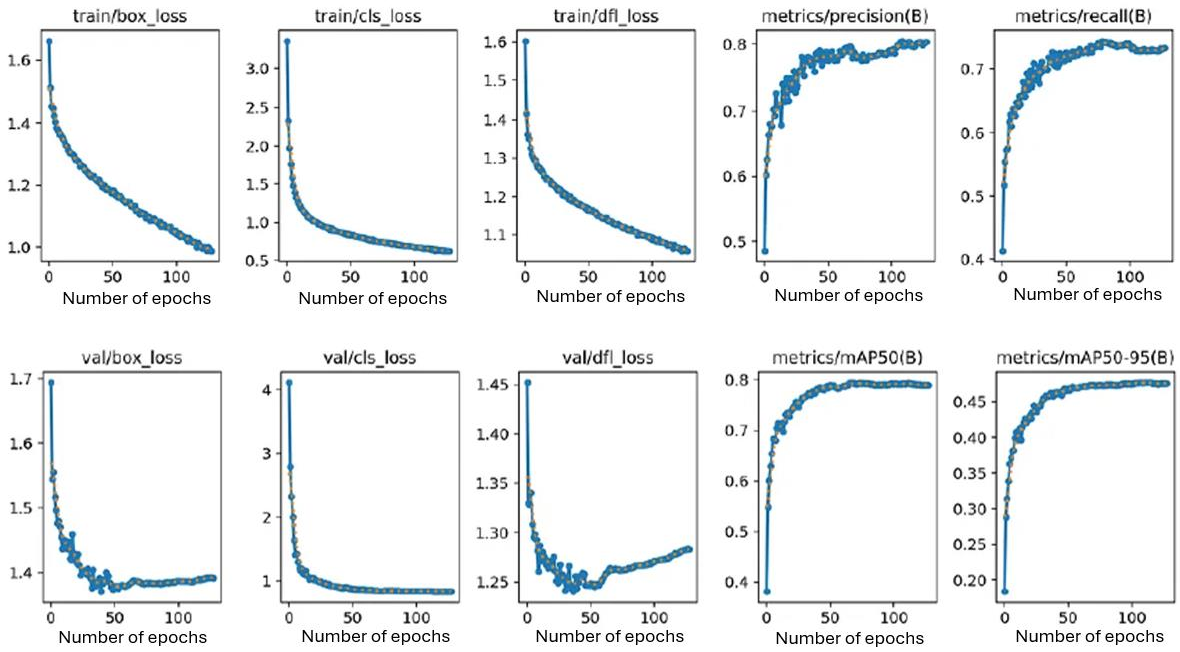

Figure 2.5: Evolution of multiple evaluation metrics as a function of the number of training epochs on the D-Fire dataset obtained during the study conducted by Tleuliyev (2023)

After training, the model's accuracy was carefully evaluated using metrics such as Precision, Recall, F1-score, and Mean Average Precision (mAP) at an IoU threshold of 0.50 (mAP50). These metrics, described in section 2.3, were helpful in verifying the model's ability to consistently detect smoke while minimizing false positives. Notably, the nano version of YOLOv8 achieved an mAP50 of 0.79 on the test set, demonstrating its ability to accurately detect smoke even in challenging environmental contexts. The efficiency of the nano model makes it particularly suitable for UAV deployment, where computational resources are limited but high detection accuracy remains essential, making it the best choice for our study case.

This trained YOLOv8 model, specifically focused on smoke and fire detection, serves as the foundation for the wildfire detection methodology implemented in this study and will be the starting point to evaluate and trying to improve its performance in a real case wildfire detection scenario. This targeted approach will try to maximize the model's accuracy and efficiency for early intervention in wildfire management.

2.2.2 Simple Image Analysis

An initial analysis was performed using the raw JPG images collected during Flights 2 and 3. The main objective was to evaluate the performance of the pretrained YOLOv8 model across different image resolutions, as resolution can significantly impact the accuracy of smoke detection. Three resolution levels were selected for the analysis: the original full-resolution images at 7952x5304 pixels, segmented images at 2000x2000 pixels, and further segmented images at 1000x1000 pixels.

To create the lower-resolution images, a custom segmentation process was developed using Python. This process involved dividing each full-resolution image into square sections of 2000x2000 and 1000x1000 pixels. This segmentation pictured in Figure 2.6 allowed for the isolated analysis of smaller regions within the larger image, potentially highlighting smoke more effectively by focusing on specific areas and minimizing irrelevant background details. However, due to the nature of the segmentation, exceptions were made for sections at the image boundaries, specifically the rightmost column and bottom row, where the segmented sections were smaller and irregular in size.

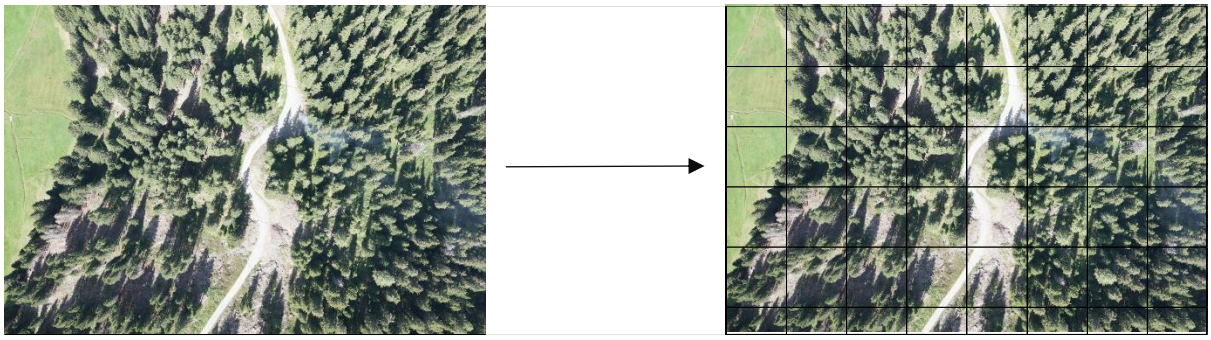

Figure 2.6: Segmentation example from a full resolution image (left) to 1000x1000 pixels segments (right)

The pretrained YOLOv8 model was used for this first analysis whose process is depicted in Figure 2.7. As part of YOLOv8's standard preprocessing pipeline, all images were automatically resized to 640x640 pixels while being input into the model.

Along with testing different image resolutions, the analysis also involved adjusting the confidence threshold across a range from 0.1 to 0.9. The confidence threshold is a key parameter in YOLOv8 that determines the minimum level of certainty required for an object to be classified as detected. Lower thresholds make the model more tolerant in identifying potential smoke regions, which can be useful for detecting faint or partially obscured smoke. On the other hand, higher thresholds demand greater certainty, reducing false positives but possibly leading to missed detections, especially if the smoke is less pronounced. By systematically adjusting the confidence threshold, the analysis aimed to identify the optimal setting for each resolution, balancing detection accuracy with the trade-offs between false positives and false negatives. This inclusive approach sought not only to determine the best confidence level for reliable detection but also to assess whether image segmentation and resolution adjustments could simplify the model's task by isolating the most relevant portions of the image. Ultimately, the goal was to identify the ideal combination of image resolution and confidence threshold that maximized YOLOv8's performance in smoke detection within dense, forested environments. This layered analysis, varying both resolution and confidence threshold, should provide valuable insights into how segmentation and resolution scaling affect the effectiveness of smoke detection.

Figure 2.7: Example of a 1000x1000 pixels segment (left) and the detection result (right) after being processed through the YOLOv8 pretrained model

2.2.3 Enhanced Image Analysis

Building on the results from the initial analysis, two additional methods were introduced to improve detection accuracy: a multi-rotation process and an area threshold filter. Each method was tested individually to assess its impact on detection performance, and then both methods were combined to evaluate their joint efficiency. These methods were applied to the resolution and confidence level identified as optimal during the initial Simple Image Analysis.

The first enhancement involves applying a multi-rotation process to each image before analysis, as showed in Figure 2.8. Since smoke detection can be influenced by the orientation of the image, each image was rotated into four distinct orientations (0°, 90°, 180°, and 270°) to determine whether examining multiple perspectives improves the model's ability to detect smoke. Each rotated version was then analyzed using the pretrained YOLOv8 model.

During the analysis, each rotated image was evaluated individually, and the model recorded the number of orientations in which smoke was detected. This counting process provided insights into the consistency of smoke detection across different orientations. If multiple orientations detect smoke, it suggests stronger confidence in the detection's validity. In contrast, if only one or none of the rotations yield detections, it may suggest that the initial detection was weak or a false positive. This approach provides a more comprehensive analysis and helps determine whether multi-rotation improves the model's reliability in identifying smoke.

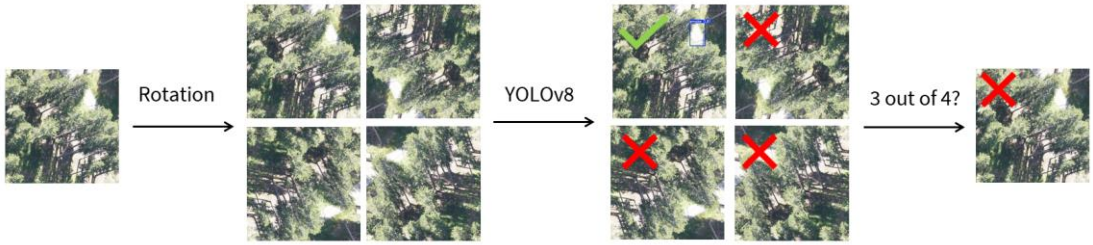

Figure 2.8: Example of a 1000x1000 pixels segment (left) that gets first rotated in the 4 possible orientations (middle-left), all of them are processed to the pretrained YOLOv8 model and signed with or without detection (middle-right), finally based on how many orientations showed a positive detection the image is classified as containing or not containing smoke (right)

The second enhancement involves applying an area threshold filter to the detections, setting a minimum size for each detected object to be considered valid. This thresholding method is particularly useful for filtering out smaller, irrelevant objects that might be misclassified as smoke due to their size and shape. By focusing only on detections that meet the predefined minimum area, this approach aims to increase accuracy by eliminating false positives often caused by small objects or background elements that resemble smoke.

The area threshold was first tested independently of the multi-rotation method to assess its isolated impact on detection accuracy. Smoke clouds typically occupy a significant portion of the image; therefore, setting a minimum area requirement allows the model to discard detections that are too small to represent actual smoke as showed in Figure 2.9. This is particularly helpful in complex environments where natural objects like rocks, shadows, or foliage might resemble smoke-like features at a smaller scale. By establishing a minimum area, the model focuses on larger shapes that are more likely to represent real smoke clouds, thus reducing noise and improving accuracy.

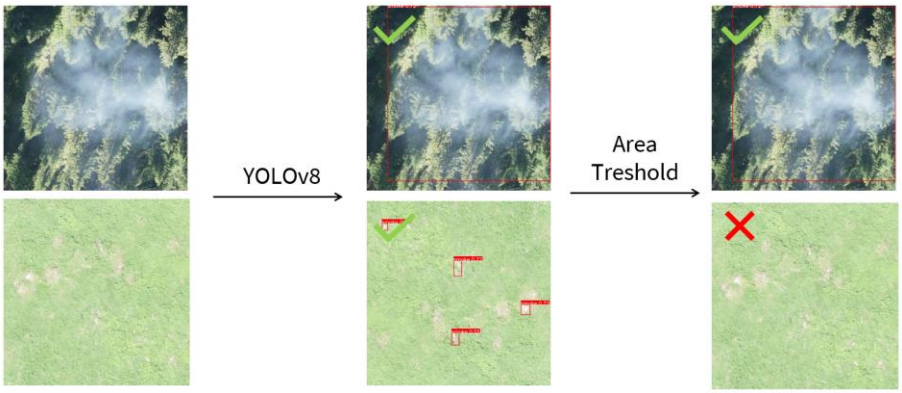

Figure 2.9: Example of two 1000x1000 pixels segments (left) that show detection of smoke after being processed through the pretrained YOLOv8 model (middle) and then kept (above) or discarded (below) based on the area of the detection

After evaluating each enhancement method independently, the area threshold was combined with the multi-rotation process. For each rotated orientation, the area threshold was applied after the YOLOv8 model processed the image, ensuring that only those detections meeting both the minimum area requirement and the orientation criteria were retained. By combining these methods, the process leverages both orientation diversity and size filtering to enhance detection reliability.

This combined approach showed in Figure 2.10 ensures that only valid detections, based on both orientation and area, are counted, ultimately improving the accuracy and consistency of smoke identification. The area threshold serves as a filtering tool that complements the multi-rotation method, refining the detection results by removing detections that do not meet the necessary criteria. Together, these two methods offer a potential strategy for improving YOLOv8's smoke detection capabilities.

Figure 2.10: Example of a 1000x1000 pixels segment (left) that is first rotated in the 4 possible orientations (middle-left), then processed through the pretrained YOLOv8 model, and if enough orientations show smoke detection (middle-right) the area threshold is applied and the detection is either kept or discarded (right)

2.2.4 Algorithm Tuning Process

To improve the YOLOv8 model's ability to detect smoke in complex forest environments, three fine-tuning phases were implemented. Each phase was designed to progressively enhance the model's ability to generalize smoke detection and minimize false positives. These fine-tuning phases were conducted separately for datasets derived from two distinct training flights, ensuring that each model was adapted to its respective environmental conditions and perspectives.

The first phase of the tuning process focused on refining the model's foundational detection capabilities using the original datasets from Flight 2 and Flight 3 exposed in Table 2.2, supplemented with targeted data augmentations. The augmentation process involved selecting

smoke images and applying rotation in four possible directions to simulate different UAV perspectives. For the non-smoke images, no rotation was applied, as their diversity already provided enough variation for the training process. The goal of this phase was to enhance the model's initial ability to detect smoke while minimizing the false positives caused by specific environmental features. The parameters used for the tuning process involved 50 epochs, early stopping patience of 10 epochs and 320x320 resolution for the input images, which were resized to balance computational efficiency with the need for sufficient detail. After this initial fine-tuning, the two separate models were evaluated. The model tuned with Flight 2 data was tested using images from Flight 3, and vice versa. This approach ensured that each model was not evaluated on the dataset it was tuned on, providing a realistic assessment of its performance on data from a different flight.

Table 2.2: Datasets composition employed during the first Tuning Phase

First Tuning Datasets				
	Flight 2		Flight 3	
	Smoke	No Smoke	Smoke	No Smoke
Training	209	354	198	309
Validation	60	96	63	77
Testing	23	40	16	41
Total	**292**	**490**	**277**	**427**

In the second phase, the original model was tuned and tested in a similar way to the first phase. However, this time, the model was tuned on an extended dataset, where the smoke images were augmented by flipping (both vertically and horizontally), effectively tripling the number of images containing smoke. The non-smoke images were augmented by rotating them in four orientations, which quadrupled the size of this dataset portion as visible in Table 2.3. This expansion aimed to provide a broader set of variations to further refine the model's performance.

To increase the quality of the fine-tuning process, several parameters were adjusted. The number of epochs was increased to 100, with a patience parameter of 20 epochs for early stopping to ensure that the model had ample time to converge effectively. The input resolution for images remained at 320x320 pixels to maintain uniformity and balance training speed with model accuracy. These changes aimed to improve the model's ability to generalize, reducing both false positives and false negatives, while ensuring that the model could detect smoke across specific scenarios and environmental conditions. After tuning, the models were again

evaluated on a separate flight dataset, the model trained on Flight 2 data was tested on Flight 3, and vice versa.

Table 2.3: Datasets composition employed during the second Tuning Phase

Extended Tuning Datasets				
	Flight 2		Flight 3	
	Smoke	No Smoke	Smoke	No Smoke
Training	636	1406	492	1153
Validation	169	398	174	327
Testing	71	156	73	131
Total	**876**	**1960**	**831**	**1708**

The third phase shifted the approach by combining the extended datasets from both Flight 2 and Flight 3, resulting in a single Combined Tuning Dataset. This dataset represented in Table 2.4 included all the augmented images from both datasets of the second phase, providing a richer and more diverse training set for improved fine-tuning. The training was conducted using the same parameters as in the second phase to maintain consistency.

Because both Flight 2 and Flight 3 were used in the combined dataset, the model could no longer be tested on an independent flight dataset. Instead, the final algorithm was evaluated on the "Flame 2" dataset created with images of smoke and non-smoke scenes that were captured with the use of UAVs (Hopkins et al., 2023). This external dataset was specifically chosen to assess how well the model could generalize to new scenarios within the UAV-specific context.

To evaluate the impact of the tuning, the performance of the final model was compared with that of the original YOLOv8 model. This comparison allowed for a detailed assessment of whether the tuning process improved the model's ability to detect smoke in real-world different wildfire detection scenarios, specifically from the UAV perspective.

Table 2.4: Datasets composition employed during the third Tuning Phase

Combined Tuning Dataset		
	Flight 2 + Flight 3	
	Smoke	No Smoke
Training	1128	2559
Validation	343	725
Testing	144	287
Total	**1707**	**3668**

These three phases formed a systematic approach to develop the YOLOv8 model's reliability for UAV-based smoke detection. The workflow of the three distinct phases is represented in Figure 2.11.

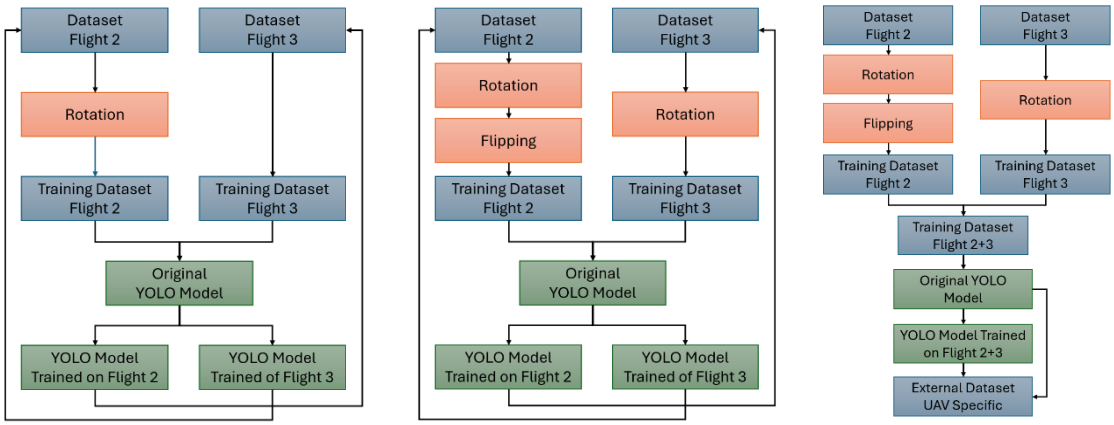

Figure 2.11: Tuning and Evaluation Workflow for the first Tuning Phase (left), the second Tuning Phase (middle) and the third Tuning Phase (right)

2.3 Evaluation metrics

To assess the performance of each model in detecting smoke, a range of evaluation metrics were calculated. These metrics provide insights into the model's accuracy, precision, sensitivity, specificity, and other key performance indicators, helping to determine its reliability and effectiveness.

1. Analyzed: This represents the total number of images analyzed, which serves as the baseline for calculating various metrics.

2. No Smoke / Smoke: These columns indicate the number of images classified as containing no smoke versus those containing smoke, providing an overview of the dataset's composition during evaluation.

3. True Positives (TP): The number of images correctly identified as containing smoke. This metric reflects the model's ability to accurately detect smoke when it is present.

4. False Positives (FP): The number of images incorrectly classified as containing smoke when there is none. False positives indicate instances where the model falsely detects smoke, impacting its precision.

5. True Negatives (TN): The number of images correctly identified as not containing smoke. True negatives represent cases where the model accurately recognizes the absence of smoke.

6. False Negatives (FN): The number of images that contain smoke but were incorrectly classified as not containing it. False negatives reflect instances where the model fails to detect smoke, impacting its sensitivity.

7. Accuracy: This metric is calculated as the ratio of correctly classified images (both true positives and true negatives) to the total number of images (Casas et al, 2023). Accuracy provides an overall measure of the model's performance:

$$Accuracy = \frac{TP + TN}{Analized}$$ (Casas et al, 2023)

8. Precision: Precision, also known as Positive Predictive Value, is the ratio of true positives to the sum of true positives and false positives (Casas et al, 2023). It indicates how reliable the model's positive smoke detections are:

$$Precision = \frac{TP}{TP + FP}$$ (Casas et al, 2023)

9. Sensitivity (Recall): Also called Recall or True Positive Rate, sensitivity measures the model's ability to correctly detect all instances of smoke. It is the ratio of true positives to the sum of true positives and false negatives (Swift et al., 2020):

$$Sensitivity = \frac{TP}{TP + FN}$$ (Swift et al., 2020)

10. Specificity: Specificity, or True Negative Rate, is the ratio of true negatives to the sum of true negatives and false positives (Swift et al., 2020). It measures the model's ability to correctly identify non-smoke images:

$$Specificity = \frac{TN}{TN + FP}$$ (Swift et al., 2020)

11. Balanced Accuracy (BA): Balanced Accuracy is the average of sensitivity and specificity, providing a balanced view of the model's performance in detecting smoke and non-smoke cases, especially useful in datasets with imbalanced classes (Brodersen, 2010):

$$BA = \frac{Sensitivity + Specificity}{2}$$ (Brodersen, 2010)

12. F1 Score: The F1 Score is the harmonic mean of precision and sensitivity (Zhao & Li, 2020). It is a useful metric when there is an uneven class distribution and helps to balance false positives and false negatives (Zhao & Li, 2020):

$$F1 = 2x\frac{PrecisionxSensitivity}{Precision + Sensitivity}$$ (Zhao & Li, 2020)

13. F2 Score: Similar to the F1 Score, the F2 Score also combines precision and sensitivity, but it gives more weight to sensitivity (Prasetiyo et al., 2021). This is useful when it is more important to capture all instances of smoke, even at the cost of some false positives:

$$F2 = \frac{5xPrecisionxSensitivity}{4xPrecision + Sensitivity}$$ (Prasetiyo et al., 2021)

14. Box Loss: Box loss is a measure of how well the predicted bounding boxes match the ground truth bounding boxes (Wang et al., 2022). It typically involves calculating the difference between the predicted and actual coordinates (center, width, and height) of

the bounding box (Wang et al., 2022). A lower box loss means more accurate localization of the detected object.

15. Class Loss: Class loss is used to assess how well the model classifies the objects within the detected bounding boxes. It calculates the error between the predicted class probabilities and the actual class labels for the detected object (Wu et al., 2022). This is typically measured using cross-entropy loss for classification tasks (Wu et al., 2022).

16. DFL Loss (Distance-Intersection over Union Loss): DFL loss is an advanced loss function used to improve bounding box localization by evaluating the distance between the predicted and ground truth bounding boxes based on their overlap (IoU) (Li et al., 2021). It encourages better localization of the detected objects and is particularly useful in complex detection tasks (Li et al., 2021).

17. mAP50 (mean Average Precision at IoU threshold 50%): mAP50 is the average precision computed at a fixed Intersection over Union (IoU) threshold of 50% (Zhu et al., 2020). This metric evaluates how well the model can detect objects by calculating precision and recall at the 50% IoU threshold, where a detection is considered correct if the predicted bounding box overlaps the ground truth by at least 50% (Zhu et al., 2020).

18. mAP50-95 (mean Average Precision at IoU thresholds from 50% to 95%): mAP50-95 is the mean of average precision scores calculated at multiple IoU thresholds, ranging from 50% to 95%, usually in increments of 5% (Reis et al., 2023). This metric provides a more comprehensive evaluation of the model's performance across different levels of overlap between predicted and actual bounding boxes, highlighting the model's ability to detect objects with higher precision at varying thresholds (Reis et al., 2023).

These metrics collectively provide a complete evaluation of each model's performance, offering insights into both its strengths and weaknesses in smoke detection. By examining these

metrics, we can understand the trade-offs involved in tuning for sensitivity against precision and assess how well the model balances these metrics across different scenarios.

3. Results

3.1 Image Analysis

3.1.1 Simple Image Analysis

The results of the Simple Image Analysis shown in Figure 3.1 display the values of various evaluation metrics in relation to the three different resolutions employed during the analysis of Flight 2 and Flight 3.

Using lower resolutions than the original images significantly increases the number of false positives, which rise from a few dozen to over a thousand, as well as true positives, which increase from just a few to several dozen in both flights. This trend is clearly influenced by the greater number of images analyzed resulting from the segmentation process, and thus requires the evaluation of normalized parameters. The accuracy value is slightly lower during the analysis of full-resolution images compared to lower resolution segments, while specificity shows little variation, consistently ranging between 0.92 and 0.96. Precision results are inconsistent between the two flights: in Flight 2, recision is higher when analyzing 2000x2000 resolution images, while in Flight 3, precision is higher when analyzing full-resolution images. However, the increase in sensitivity is clear, as both flights show much higher values for segmented images, with sensitivity rising from values below 0.1 to above 0.4 in the case of the 1000x1000 resolution images in Flight 3. The same trend is observed for F1 and F2 scores, which are significantly lower for full-resolution images and higher when using the 2000x2000 resolution segments.

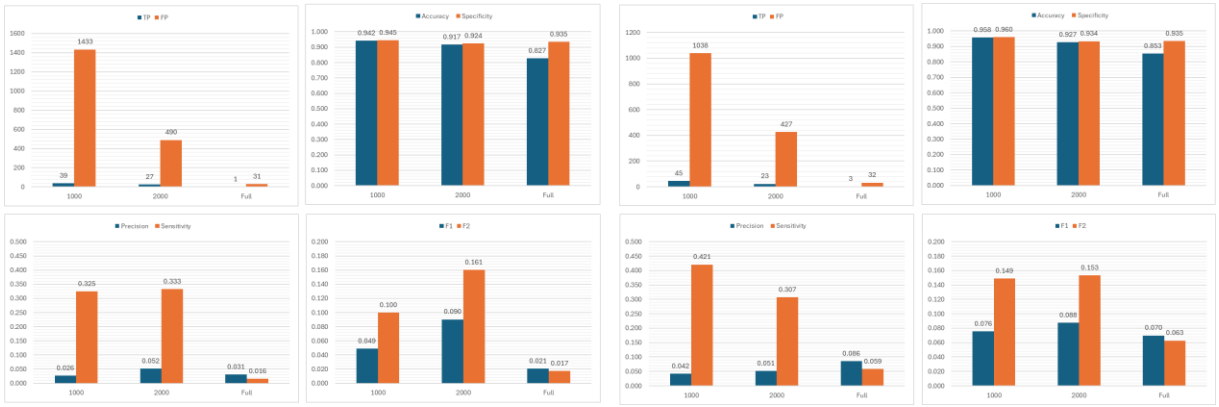

Figure 3.1: Values of diverse evaluation metrics across different resolutions for Flight 2 (left) and Flight 3 (right) during the Simple Image Analysis

The graphics in Figure 3.2 integrate the confidence level, along with resolution, as a parameter for determining F2. It is evident that the trend is similar for both flights. Using a low confidence level leads to higher F2 values, which consistently decrease as the confidence level increases. However, using excessively low confidence levels, such as 0.1, results in a drastic drop in F2 values for both flights. In most cases, segments with a resolution of 2000x2000 provide the best F2 results at the same confidence level when compared to other resolutions. The peak performance for both flights is achieved when using a confidence level of 0.2, applied to the 2000x2000 resolution segments. These results make the use of these two parameters a starting point for attempting to improve the current F2 value during the subsequent Enhanced Image Analysis.

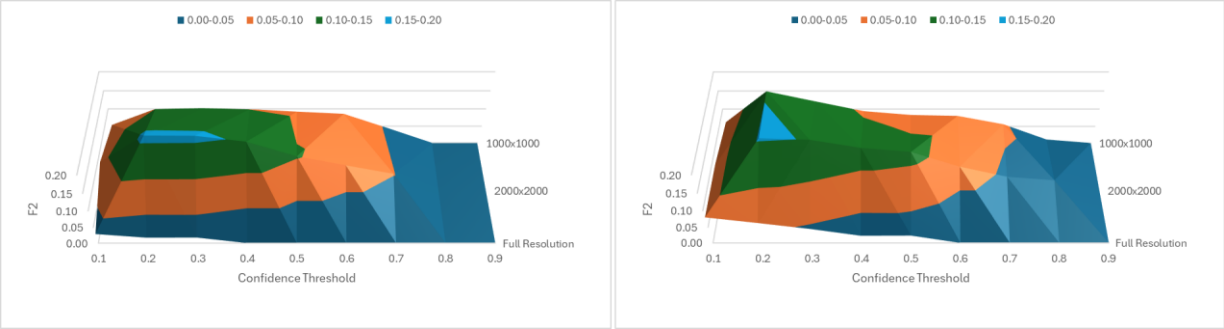

Figure 3.2: F2 Score for varying resolution and confidence level on Flight 2 (left) and Flight 3 (right) during the Simple Image Analysis

3.1.2 Enhanced Image Analysis

Figure 3.3 shows the F2 value as a function of the minimum number of image rotations analyzed that contain smoke detections. In the case of Flight 2, using a threshold of 1 results in a lower F2 value compared to the original one, while thresholds of 2, 3, and 4 increase it. The threshold of 3 provides the best results, peaking at around 0.25, nearly double the initial F2 value. For Flight 3, similar results are observed. With thresholds of 1 or 2, the F2 values are lower than the original one, but they increase with thresholds of 3 or 4. Once again, the peak is reached with a threshold of 3, which shows around a 50% increase in the F2 value compared to the original. Therefore, 3 seems to be the ideal minimum number of rotations presenting detections required to achieve the best performance in our specific case study.

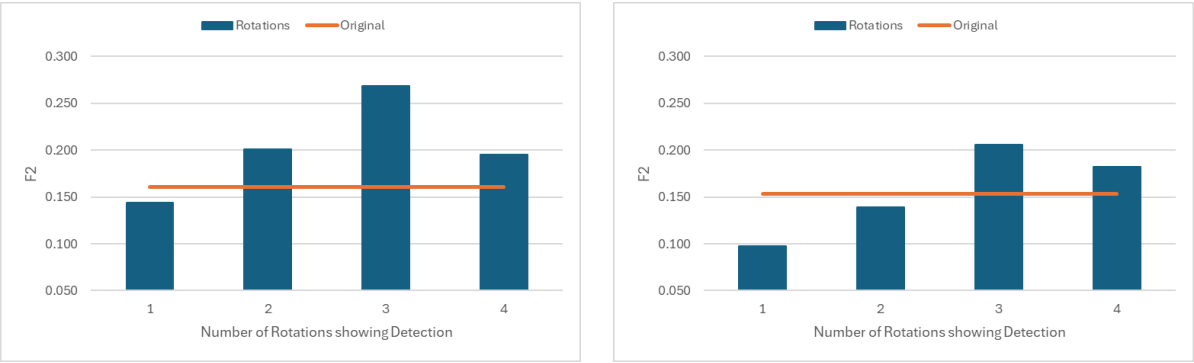

Figure 3.3: F2 Score for number of rotations showing detections on Flight 2 (left) and Flight 3 (right) during the Enhanced Image Analysis

Figure 3.4 shows the different F2 values obtained in both flights as a function of the area threshold applied. In the case of Flight 2, the F2 value increases steadily until reaching 500'000 pixels, where it drops sharply and then remains at values similar to or lower than the original. For Flight 3, the trend is similar, with the F2 value increasing as the area threshold rises, peaking at 500'000 pixels. After this point, the trend reverses, leading to a continuous decrease in F2. These results suggest that a minimum area threshold of approximately 400'000-500'000 pixels is the ideal value for achieving the best performance in our specific case study.

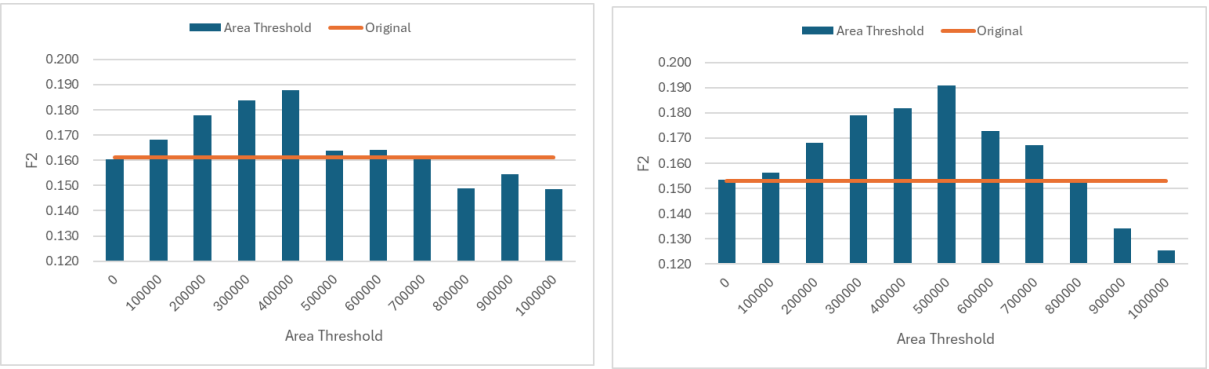

Figure 3.4: F2 Score for different area thresholds on Flight 2 (left) and Flight 3 (right) during the Enhanced Image Analysis

Figure 3.5 compares the results of various metrics using the original model, multi-rotation, area threshold, and the combined method. It is evident that each method leads to a significant reduction in the number of false positives in both flights, dropping from over 400 to values ranging from 63 with the combined method in Flight 3 to 116 with the multi-rotation and combined methods in Flight 2. The number of true positives always decreases compared to the

original method, with the loss of correct detection of a number of images ranging from 1 to 5 units, depending on the method used.

The accuracy and specificity values show only a slight increase compared to those obtained during the original analysis. Sensitivity decreases slightly with all methods, but precision improves significantly, especially with the multi-rotation and combined methods, increasing from an F2 value of around 0.05 to over 0.2 with the combined method in Flight 3. The F1 and F2 scores improve compared to the original with all methods, achieving the best performance with the combined method in both flights.

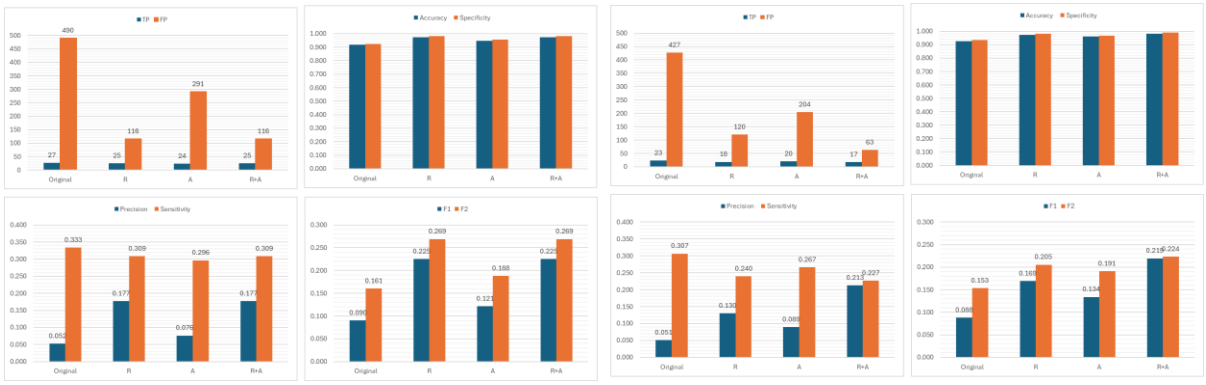

Figure 3.5: Values of diverse metrics across different enhancing methods for Flight 2 (left) and Flight 3 (right)

Figure 3.6 graphically represents the F2 value as a function of the area threshold used and the method applied. Using the area threshold method alone does not lead to significant performance improvements. This changes when the combined method is used, where the area threshold and multi-rotation are combined, resulting in a notable increase in F2. In the case of Flight 2, a clear peak is observed with an area threshold of 0 pixels, indicating that the simple application of the multi-rotation method yields the best results. F2 remains higher compared to the original method and the area threshold alone, until a threshold of around 700’000 pixels. A similar trend is observed in Flight 3, where again, the combination of the area threshold up to 300’000 pixels and the multi-rotation method yields the highest F2 values.

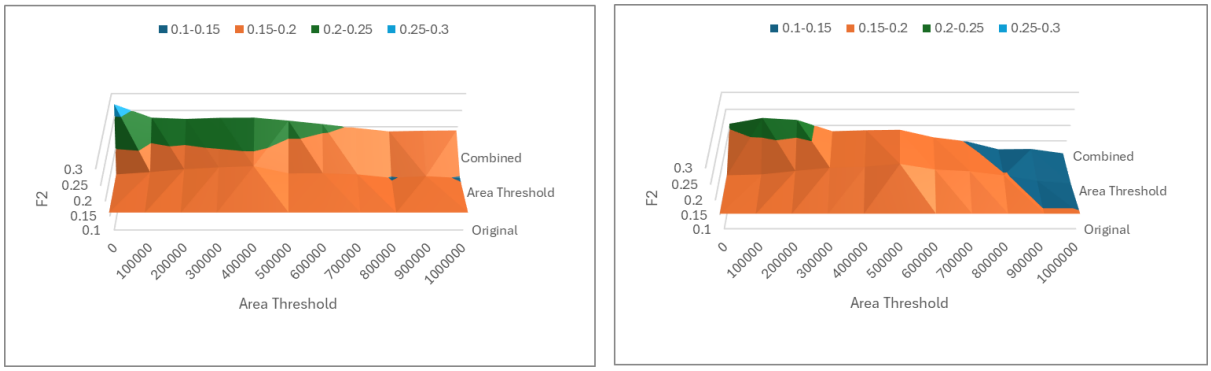

Figure 3.6: F2 Score for Combined Methodology compared to Original and Area Threshold on Flight 2 (left) and Flight 3 (right) during the Enhanced Image Analysis

3.2 Algorithms Tuning Process

The values obtained during the first tuning phase are shown in Figure 3.7 for Flight 2 and Figure 3.8 for Flight 3. The top row presents values for the training sub dataset, while the bottom row shows values for the evaluation sub dataset, allowing us to assess whether performance improvements are generalizable to the entire dataset. In both flights, for the first three columns on the left indicating the loss metrics, we observe a constant decrease in values, while in the two columns on the right, there is an increase as the number of epochs grows. This indicates a tuning process that is enhancing the quality of the detection, increasing the mAP50-95 values from around 0.1 at the beginning to over 0.3 in Flight 3 and more than 0.4 in Flight 4.

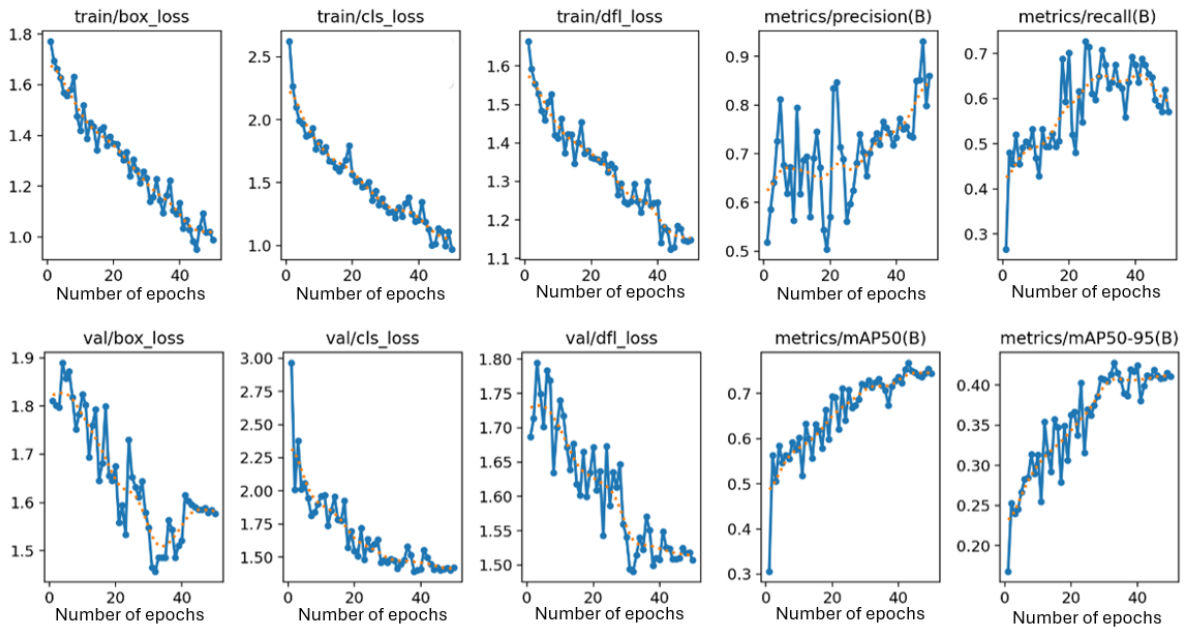

Figure 3.7: Training metrics obtained during the first Tuning Phase on Flight 2

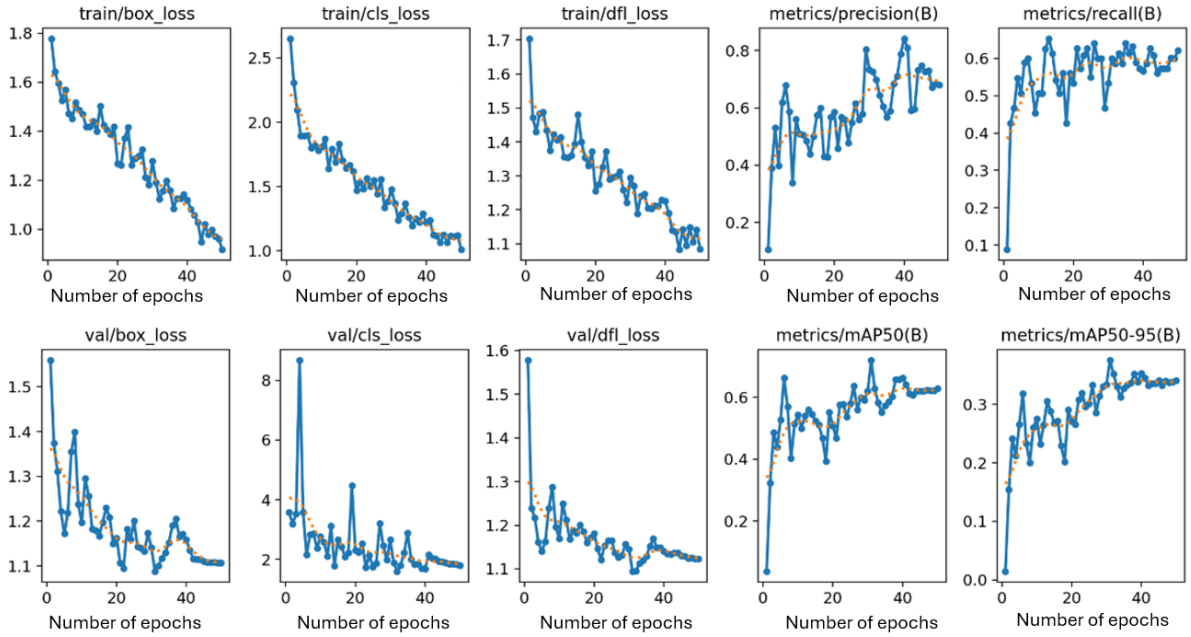

Figure 3.8: Training metrics obtained during the first Tuning Phase on Flight 3

The same trend becomes even more evident during the second tuning phase with the use of the extended training datasets, shown for Flight 2 in Figure 3.9 and for Flight 3 in Figure 3.10. The increased number of epochs for this training phase allows for a clearer view of the correct tuning process. A strong and consistent decrease is observed in the loss metrics in the 3 columns on the left. Additionally, as in the previous tuning phase, we have an increase in the values in the two rightmost columns, with the mAP50-95 value rising from around 0.2 at the start of the process to nearly 0.5 at the end.

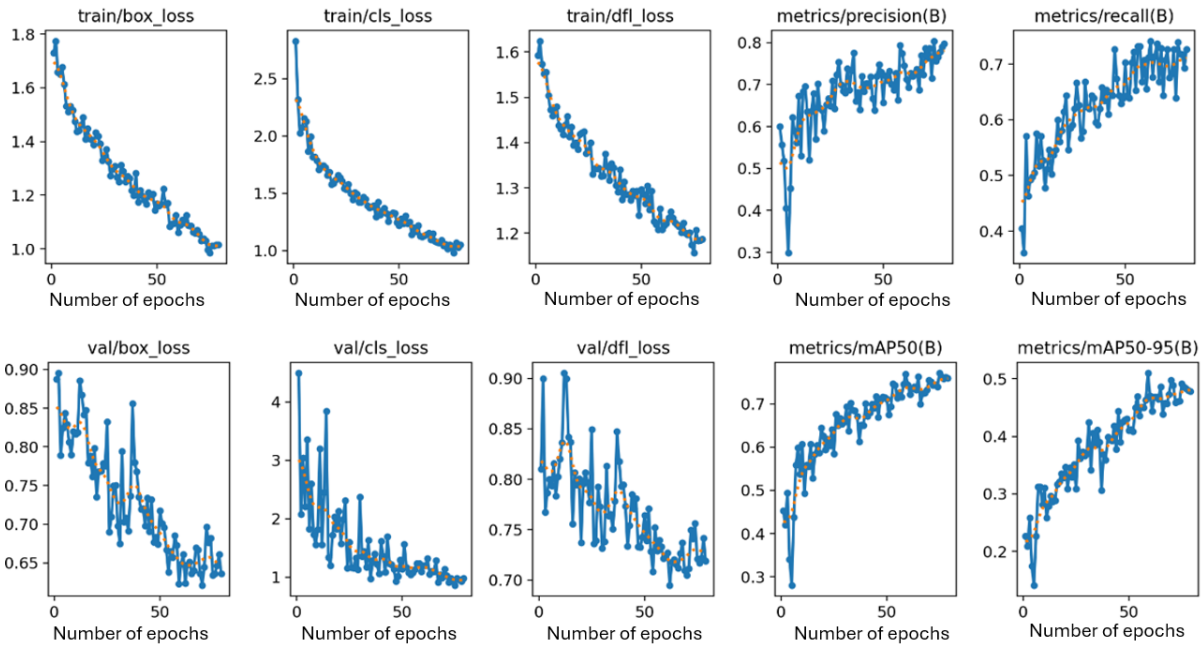

Figure 3.9: Training metrics obtained during the second Tuning Phase on Flight 2

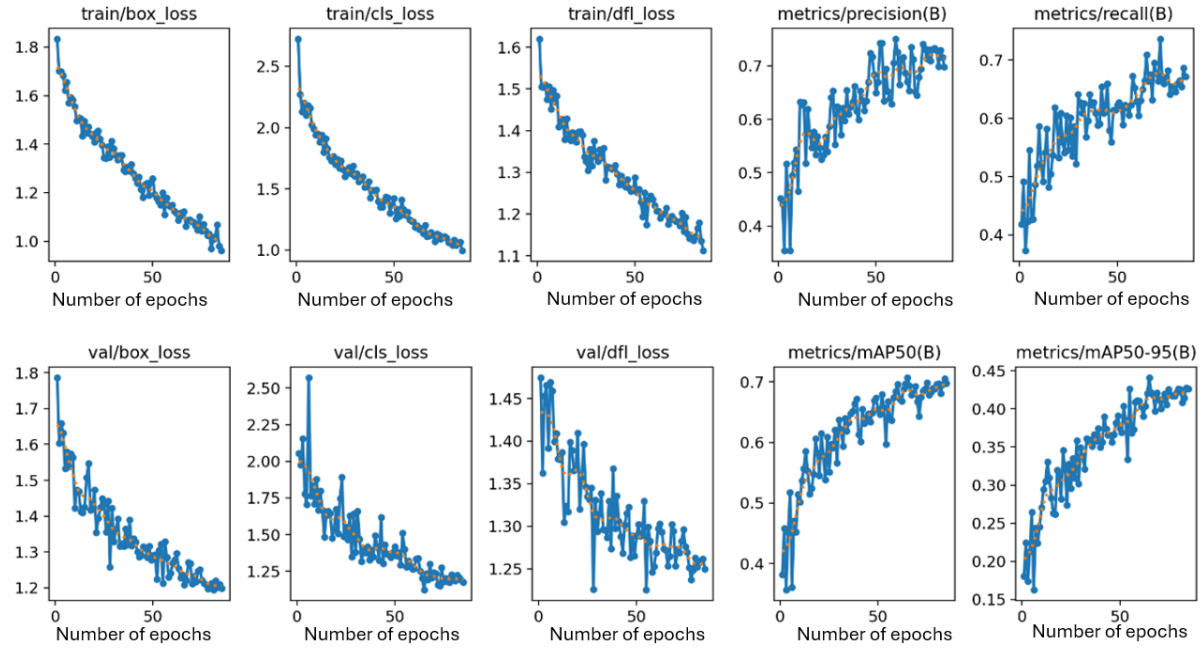

Figure 3.10: Training metrics obtained during the second Tuning Phase on Flight 3

The same trend is observed during the third tuning phase, performed with the combined datasets, as shown in Figure 3.11. This time, the tuning process lasts for a total of 100 epochs, allowing for better convergence of the metric values toward the desired results. Once again, we observe a significant decrease in the loss metrics in the three leftmost columns and an increase in the metrics in the two rightmost columns, where the mAP50-95 value rises from the initial 0.2 to nearly 0.5, indicating a much more accurate classification.

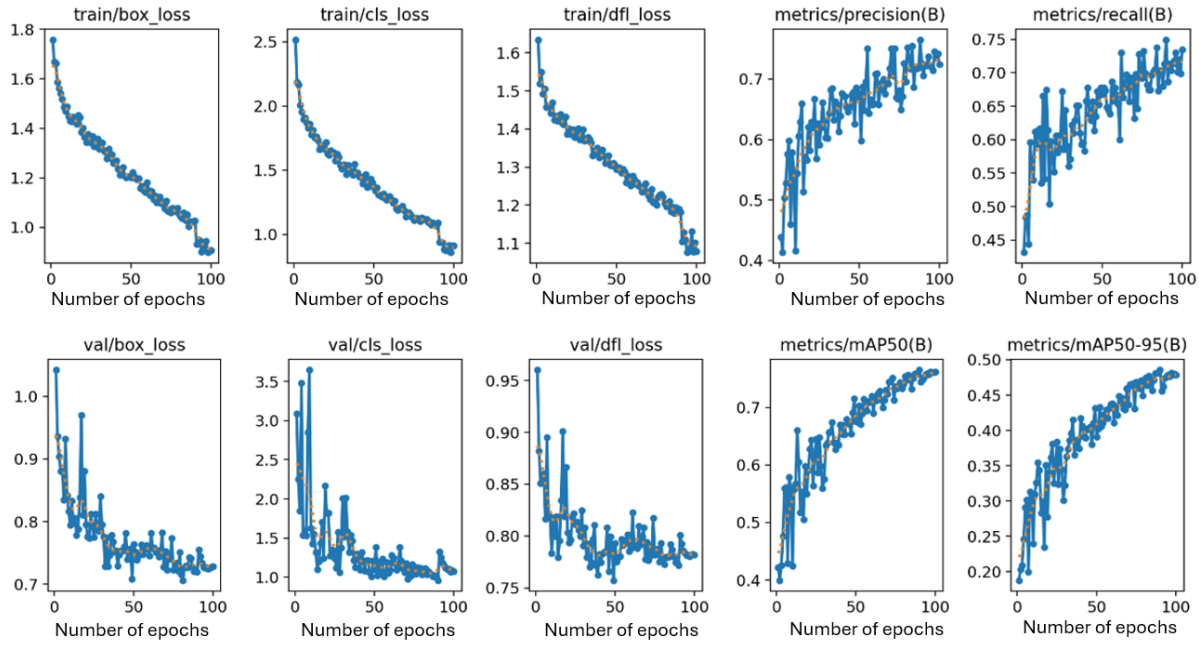

Figure 3.11: Training metrics obtained during the third Tuning Phase on Flight 2 and Flight 3

3.3 Algorithms Tuning Evaluation

3.3.1 At 2000x2000 Resolution

The results shown in Figure 3.12 compare the outcomes obtained from the original model and those of the models developed during the first (T1) and second (T2) tuning phases.

The number of false positives decreases significantly, dropping from over 400 with the original model to a few dozen with the T1 and T2 models. The number of true positives also improves, increasing from 27 initially to 39 with T2 in Flight 2, and from 23 to 40 with T2 in Flight 3. Accuracy and specificity also show substantial improvement, rising from initial values near 0.9 to values approaching 1 with T1 and T2. A noticeable further increase is observed in precision, which rises from around 0.05 in the original model to nearly 0.85 with T2 in both flights, as well as in sensitivity, which increases from 0.3 initially to around 0.4 with T1 and 0.5 with T2. The best results obtained by the tuned models across various metrics are reflected in the F1 and F2 scores, with F2 rising from an initial 0.33 to 0.53 with T2 in Flight 2, and from 0.15 to 0.58 with T2 in Flight 3. T1 and T2 clearly outperform the original model across all evaluation metrics, with T2 providing better results than T1.

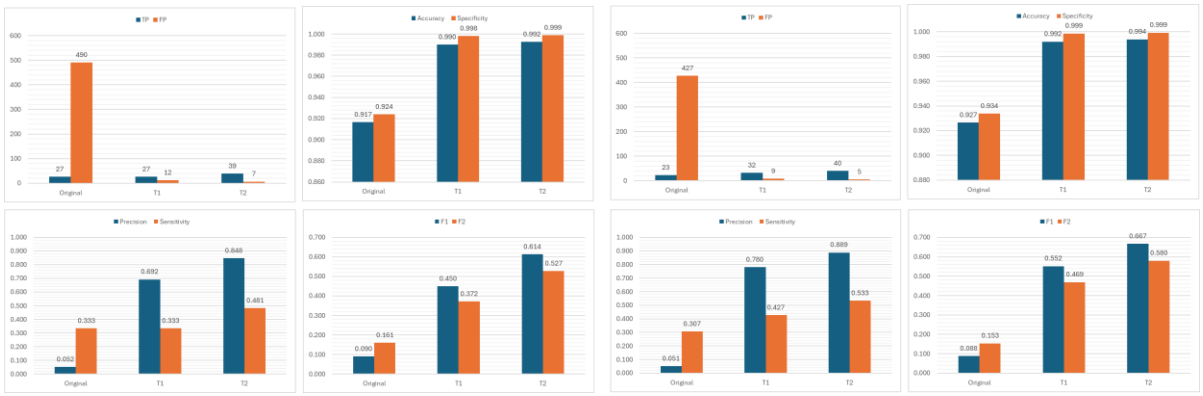

Figure 3.12: Values of diverse metrics using the original model, the first tuned model and the second tuned model for Flight 2 (left) and Flight 3 (right) after the first and second Tuning Phases

The graphical representation in Figure 3.13 shows the F2 values as a function of the confidence level and the model used. The two flights exhibit very similar results, with higher F2 values observed when using low confidence levels, ranging from 0.1 to 0.3, and values that significantly decrease as the confidence level increases. Additionally, for the same confidence level, it is evident that T1 always provides higher results compared to the original model, and T2, in turn, offers better results than T1.

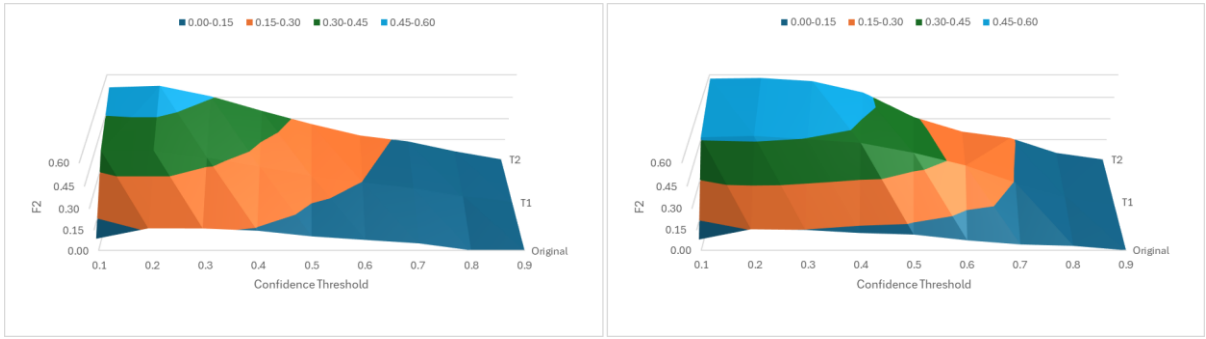

Figure 3.13: Comparison of F2 Score for Original Model, Tuning 1 and Tuning 2 on Flight 2 (left) and Flight 3 (right) at 2000x2000 pixel resolution after the first and second Tuning Phases

Figure 3.14 shows the F2 values when the multi-rotation method is applied to the T2 model. In this case, the F2 value obtained using this method does not show a significant increase when compared to the original model. For Flight 2, using thresholds of 1 or 2 results in a slight increase in F2, while thresholds of 3 or 4 lead to values lower than the original. For Flight 3, once again, with a threshold of 1, we observe a slightly higher F2 value compared to the original model, but values decrease when using a minimum of 2, 3, or 4 orientations presenting detections.

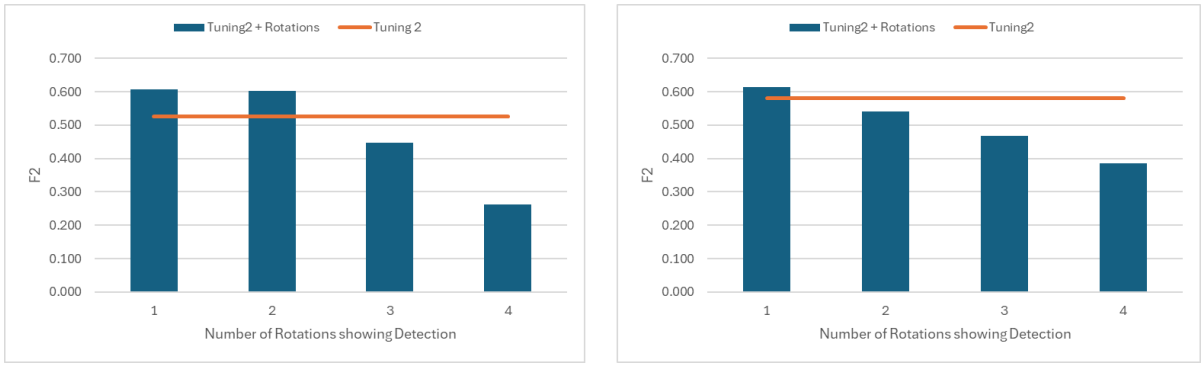

Figure 3.14: Comparison of F2 Score for Tuning 2 and Tuning 2 + Rotation on Flight 2 (left) and Flight 3 (right)

3.3.2 Full Resolution

Figure 3.15 shows the metrics obtained after analyzing the full-resolution images using the original model, T1, and T2. As with the 2000x2000 resolution, the T1 and T2 models result in a significant reduction in false positives in both flights, decreasing from several dozen to zero in the case of T2. True positives also increase substantially, rising from a few to 8 in Flight 2 and 16 in Flight 3. Accuracy and sensitivity slightly improve compared to the original model when using T1 and T2, reaching values very close to 1. Precision also improves dramatically, increasing from 0.03 in the original model to 1.00 for T2 in Flight 2, and from 0.09 to 1.00 in Flight 3. The best detection and classification capabilities are also reflected in the F1 and F2 values, with F2 rising from 0.02 in the original model to 0.15 for T2 in Flight 2, and from 0.06 in the original model to 0.36 for T2 in Flight 3.

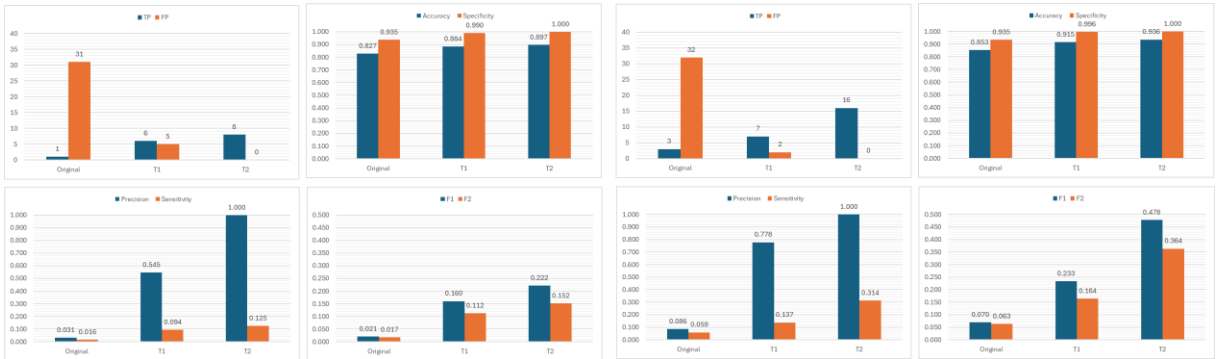

Figure 3.15: Values of diverse metrics using the original model, the first tuned model and the second tuned model for Flight 2 (left) and Flight 3 (right)

The graphical representation in Figure 3.16 shows the F2 value as a function of the confidence level and the model used for the analysis. In this case, the results obtained in the two flights

differ, with Flight 3 generally showing higher F2 values. As in the case of the 2000x2000 pixel segments, lower confidence levels also yield better results for full-resolution images. The highest F2 values are achieved with a confidence level between 0.1 and 0.3, while higher confidence levels lead to a decrease in performance. Once again, the tuned models offer better results compared to the original model, with T1 outperforming the original model and T2 further improving and significantly increasing the results obtained by T1.

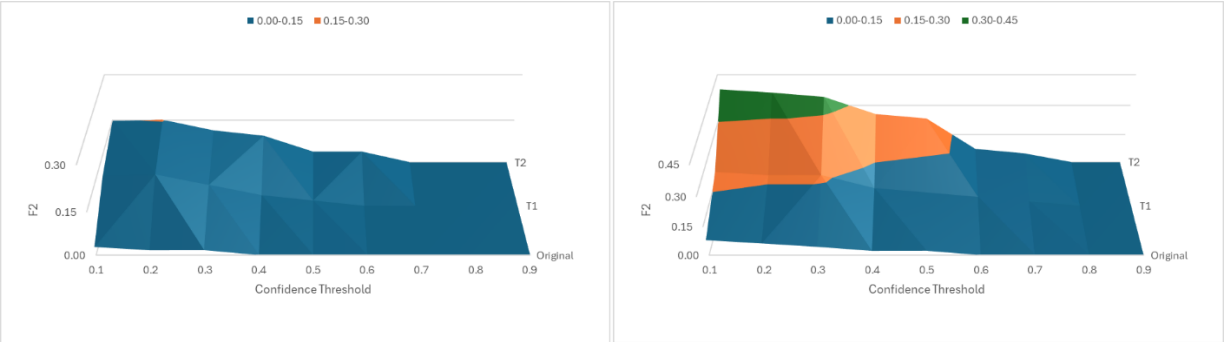

Figure 3.16: Comparison of F2 Score for Original Model, Tuning 1 and Tuning 2 on Flight 2 (left) and Flight 3 (right) at full resolution

The improved detection capabilities of the T2 algorithm compared to the original algorithm can be seen in Figure 3.17. Although in the second frame the smoke is identified with a lower confidence level by the T2 algorithm compared to the original, using the tuned algorithm we are also able to correctly identify the smoke in the first and third images.

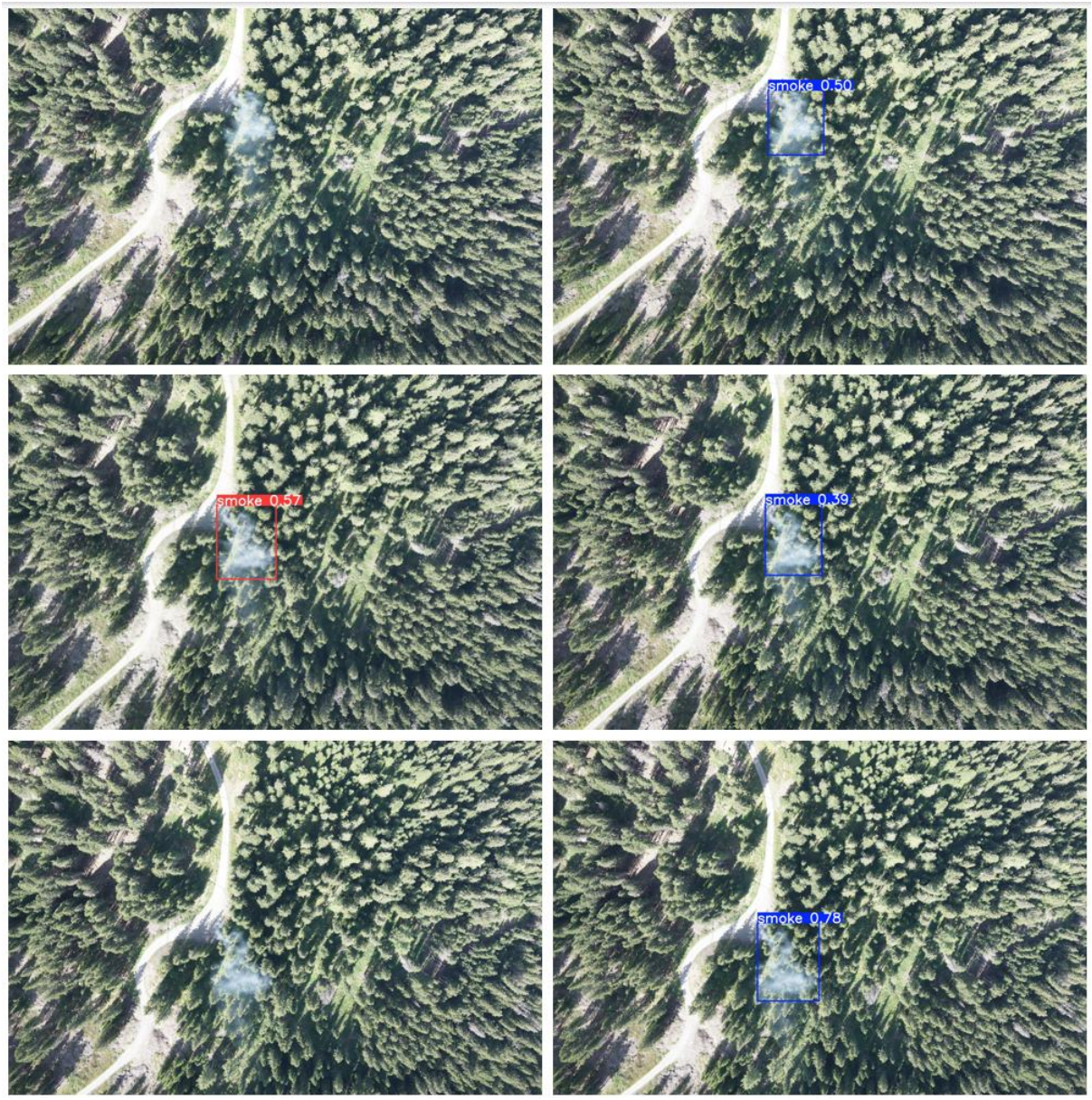

Figure 3.17: Comparison of the results of the detection process on 3 consecutive Flight 3 images with the original algorithm (left) and the tuned T2 algorithm (right)

The same improvement in performance can also be seen in Figure 3.18, where again the results comparing the detection of the original algorithm and the tuned T2 algorithm are displayed. In this case, the original algorithm was not able to correctly identify smoke in any of the 3 different images. In contrast, the tuned T2 algorithm is able to identify smoke correctly in all 3 images considered.

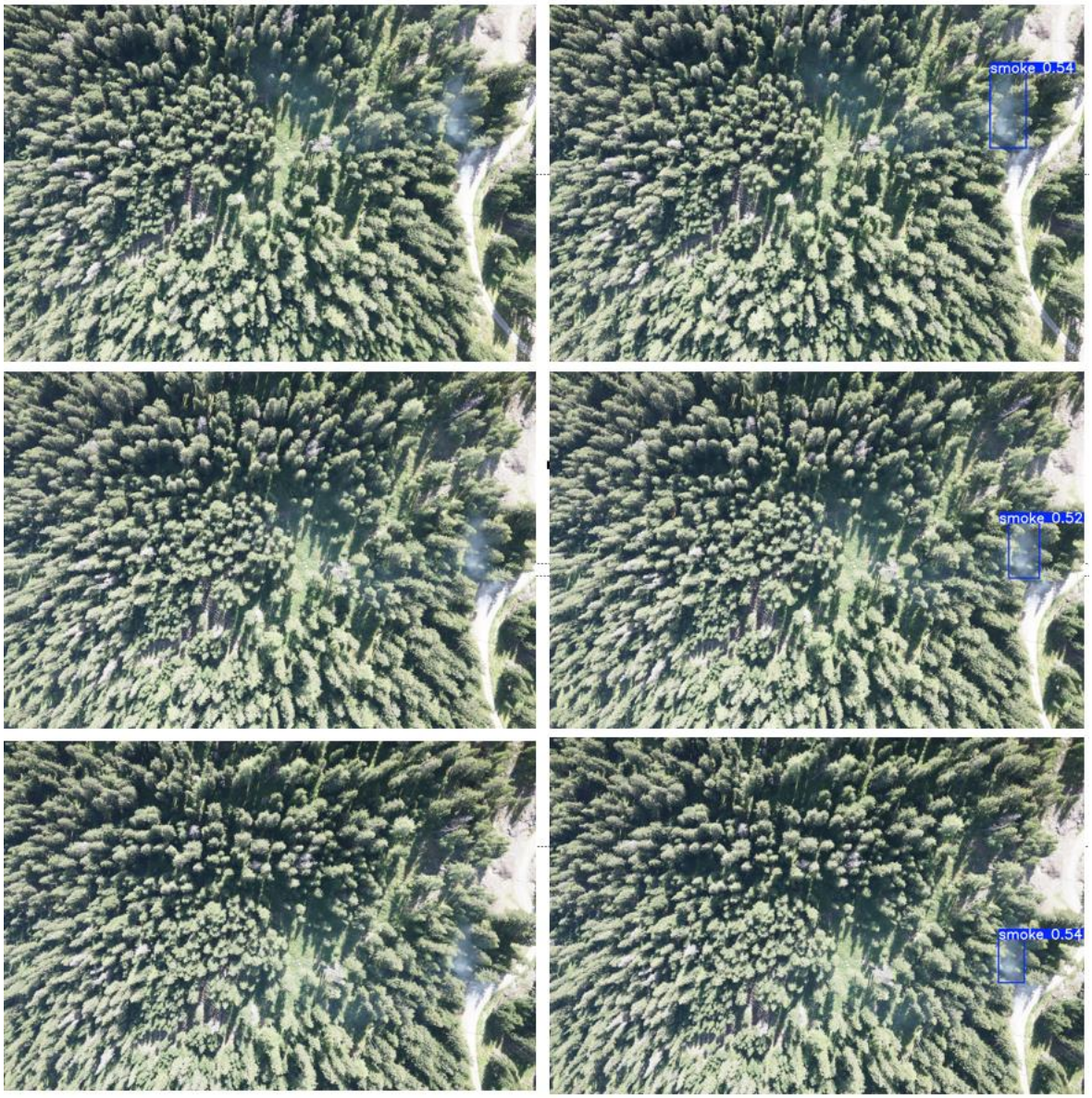

Figure 3.18: Comparison of the results of the detection process on 3 others consecutive Flight 3 images with the original algorithm (left) and the tuned T2 algorithm (right)

3.3.3 Comparison on external Dataset

Figure 3.19 shows the difference in evaluation metrics by subtracting the results obtained from the model of the third tuning phase (T3) from those of the original model, providing an overview of the differences between the two models. First, it is clear that no single model consistently outperforms the other across all confidence levels. The original model achieved better results primarily with lower confidence levels such as 0.1 and 0.2, as well as higher levels between 0.6 and 0.8, where the difference in F2 values becomes significant, ranging from -0.065 to -0.077. On the other hand, the T3 model performs better with confidence levels

of 0.4 and 0.5, as well as at 0.9, where a much larger number of images are correctly classified as positive.

Δ	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
> 0.1	-14	94	-94	14	-0.020	-0.019	-0.003	-0.219	-0.111	-0.011	-0.006
> 0.2	-41	29	-29	41	-0.013	-0.006	-0.008	-0.067	-0.038	-0.007	-0.008
> 0.3	-59	8	-8	59	-0.013	-0.002	-0.012	-0.019	-0.015	-0.007	-0.010
> 0.4	-27	-5	5	27	-0.004	0.001	-0.006	0.012	0.003	-0.002	-0.004
> 0.5	-26	-4	4	26	-0.004	0.001	-0.005	0.009	0.002	-0.002	-0.004
> 0.6	-142	-1	1	142	-0.027	0.000	-0.029	0.002	-0.013	-0.016	-0.024
> 0.7	-361	0	0	361	-0.068	0.000	-0.074	0.000	-0.037	-0.048	-0.065
> 0.8	-382	0	0	382	-0.072	0.000	-0.078	0.000	-0.039	-0.070	-0.077
> 0.9	572	0	0	-572	0.108	0.000	0.117	0.000	0.059	0.195	0.140

Figure 3.19: Difference in evaluation metrics as a function of confidence level obtained by subtracting those of the original algorithm from those obtained from the tuned algorithm T3, where blue indicates better performance and red worse.

4. Discussion

4.1 Image Analysis Review

4.1.1 Simple Image Analysis

The segmentation process used during the analysis allowed YOLOv8 to improve its performance. Although the number of false positives increased, the higher number of true positives led to a greater sensitivity and consequently a better F2 score. In our specific case, the 1000x1000 resolution achieved better results compared to the original resolution, but the high number of false positives caused by the loss of image context made the detection process less ideal. The 2000x2000 resolution, on the other hand, provided the best results, maintaining a low number of false positives while identifying smoke clouds that were too small to be properly detected in the original resolution images. A similar approach was explored by Ozge Unel et al. (2019), who addressed the challenge of small object detection in remote sensing imagery by segmenting high-resolution images into smaller patches. This strategy enabled the detection system to focus on localized regions where small objects, such as pedestrians or vehicles, were more likely to appear (Ozge Unel et al., 2019). By processing these smaller image patches rather than the entire high-resolution image at once, their model demonstrated a notable improvement in detection accuracy (Ozge Unel et al., 2019). This image tiling technique significantly enhanced the model's ability to identify small objects that could have otherwise been missed in a full-resolution analysis (Ozge Unel et al., 2019). This approach aligns with the segmentation strategy used in this study, reinforcing the effectiveness of dividing images into smaller sections to improve detection performance in complex environments.

The choice of confidence level also plays a critical role in the model’s performance. As the data suggests, lower confidence levels increase the model’s sensitivity by detecting a larger portion of the smoke signals, including those that are faint or partially obscured. However, this increase in sensitivity comes with a significant rise in false positives, where non-smoke features are incorrectly identified as smoke. On the other hand, raising the confidence threshold reduces false positives and improves precision. At higher confidence thresholds, precision can reach its peak, indicating that all detections are accurate. However, this comes at the expense of

sensitivity, as the model becomes more conservative and misses several weak or faint smoke signals, which are often the earliest indicators of a potential wildfire.

This trade-off between precision and sensitivity has been widely acknowledged in object detection research. In a study by Mukhiddinov et al. (2023), a similar effect was observed when varying the confidence thresholds in a wildfire smoke detection system using UAV imagery. They found that lowering the threshold increased sensitivity, enabling the detection of more faint smoke signals, but also led to a significant increase in false positives (Mukhiddinov et al., 2023). On the other hand, raising the threshold improved precision, reducing false positives but missing early-stage smoke that could be critical for timely wildfire intervention (Mukhiddinov et al., 2023). These findings align with our observations, suggesting that the optimal balance between sensitivity and precision is crucial for effective detection systems, particularly in real-time applications. Further supporting this, Gonçalves et al. (2023) evaluated YOLO-based models for detecting smoke and wildfires in both ground and aerial images, highlighting the trade-off between sensitivity and precision as confidence levels were adjusted. Their study confirmed that lower confidence thresholds result in more true positives but at the cost of increased false positives, which can overwhelm the model (Gonçalves et al., 2023). Conversely, higher thresholds were shown to reduce false positives and improve precision, but this made the system less sensitive to early, subtle signs of smoke, a finding that parallels our results, particularly with the drop in sensitivity at higher confidence thresholds (Gonçalves et al., 2023). Additionally, a broader study by Hao et al. (2022), while not focused specifically on smoke detection, examined the impact of image quality and the distance of objects on object detection performance. Their findings validated our results by showing that adjusting the confidence threshold influences detection accuracy, where lower thresholds increase sensitivity and allow for the detection of smaller objects but at the cost of false positives (Hao et al., 2022). This suggests that for wildfire detection, especially in dynamic environments, sensitivity may need to be prioritized, at least in the early stages of detection, to catch subtle smoke signals before they escalate.

4.1.2 Enhanced Image Analysis

The three primary methods, simple analysis, multi-rotation, and combined techniques demonstrated distinct advantages and trade-offs in enhancing the YOLOv8 model's smoke detection capabilities. Each approach played a fundamental role in addressing specific

challenges encountered in complex forested environments, where smoke visibility can be highly variable due to environmental factors.

The multi-rotation method enhances the smoke detection process by addressing the challenge of false positives and improving the overall consistency of detections. By rotating each image through multiple orientations as shown in Figure 4.2, the model can analyze smoke from different angles and assess detection consistency. When detections appear across multiple rotations, the likelihood of them being true positives increases. Conversely, if a detection only appears in one rotation or a limited number of perspectives, it becomes more likely to be a false positive. This approach ensures that only detections that are consistent across various angles are considered valid, thus reducing the number of false alarms. The multi-rotation method not only filters out inconsistent false positives but can also improves the detection of previously misclassified false negatives. Some smoke patterns, which may not be clearly visible from one perspective, can become more distinguishable from different angles, allowing the model to correctly identify more instances of smoke. By requiring detections to be consistent across multiple rotations, the method strengthens the detection framework, ensuring that valid smoke instances are reliably identified while minimizing errors.

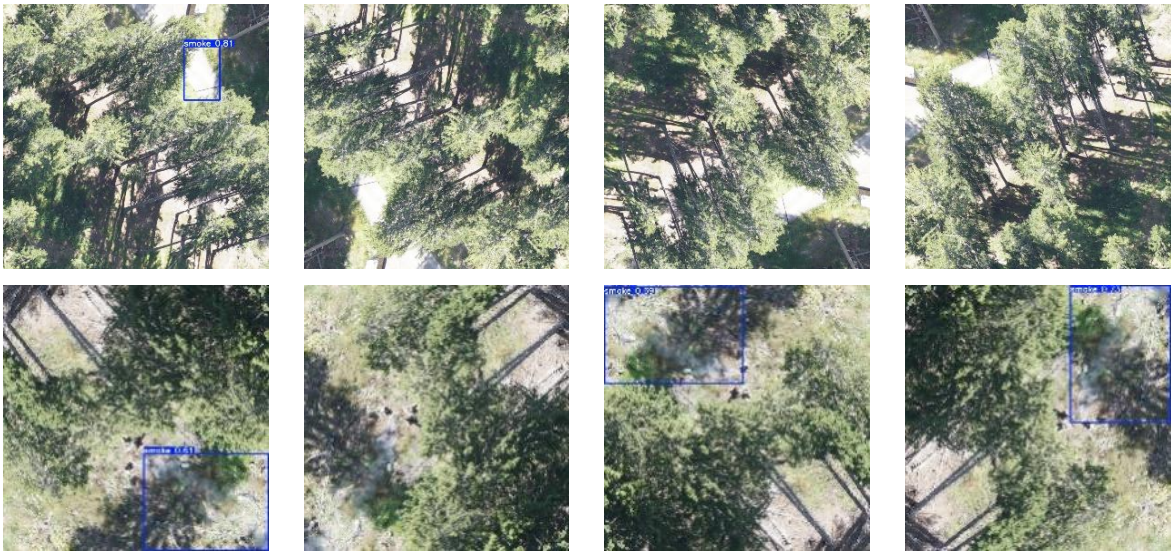

Figure 4.1:Example of a false positive showing detection only in one of the four rotations (above) and a true positive showing detection in three out of four (below)

This multi-rotation approach yields similarities to the methods used in recent studies on object detection. For example, in Li et al. (2025), the authors utilized multi-scale and rotated feature extraction to detect rotated targets in remote sensing images. Their method addressed the challenge of detecting objects from various angles, improving accuracy by analyzing them

from different orientations (Li et al., 2025), much like how multi-rotation in our study improves the detection of smoke from various perspectives. Similarly, Fu et al. (2020) employed a rotation-aware CNN for small object detection in remote sensing images. The network in their approach integrated multi-scale features and rotation-invariant mechanisms, which helped improve robustness when objects appeared at different orientations (Fu et al., 2020). Their method focused on handling rotations, similar to the goal of multi-rotation in our smoke detection approach. Both studies, while focused on different domains, emphasize the importance of considering multiple orientations to increase detection accuracy (Li et al., 2025; Fu et al., 2020) However, the multi-rotation method comes with a trade-off: while it significantly improves precision by eliminating false positives, it reduces sensitivity. As the model demands consistency across more rotations, subtle or partial smoke signals that are not visible in all orientations are excluded, leading to missed detections. This reduction in sensitivity is a critical factor to consider, as it may result in overlooking smaller smoke signals.

Area thresholding takes a different approach to refining the detection process by applying a minimum size requirement for detected smoke instances. In our case, a significant number of false positives are caused by small grey-colored objects that exhibit a structural shape similar to a rising smoke columns. By implementing a minimum area thresholding method, we can effectively prevent a substantial portion of these false detections. These false positives are primarily generated by objects such as rocks in open fields, dry tree tops, or light distortions caused by shadows on asphalt surfaces as showed in Figure 4.2. Introducing a minimum area filter is a logical approach, as all these objects tend to have a smaller area compared to the natural dispersion of a smoke cloud in the air. This ensures that the model prioritizes larger, more realistic smoke formations, significantly improving detection reliability.

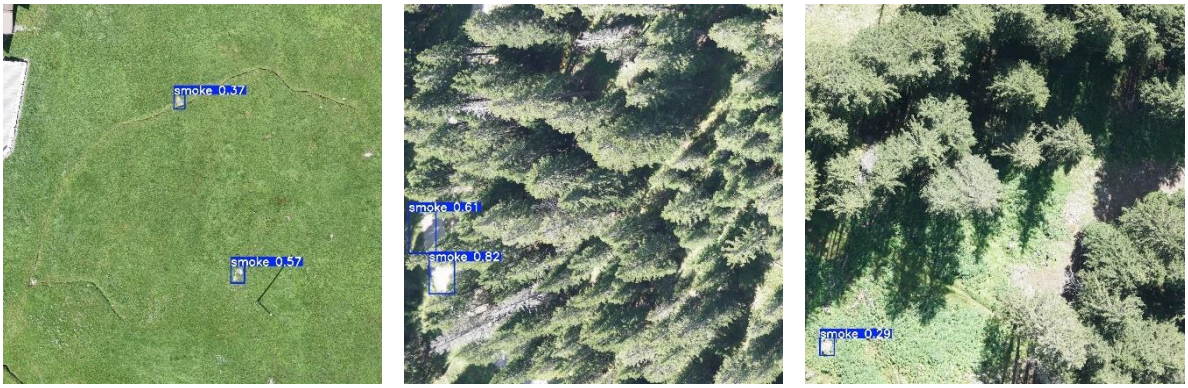

Figure 4.2: Examples of false positives smaller area (above) compared to true positives bigger area (below)

A similar approach is seen in the Fast R-CNN framework, where a similar filtering technique is employed to exclude regions of interest that fall outside a specified size range. This is done to prevent the model from focusing on irrelevant small objects that could lead to false positives. By excluding these non-relevant regions, the model is able to concentrate on the most important areas of the image, much like the application of the minimum area filter in our method to improve smoke detection accuracy. Additionally, Girshick et al. (2015) in their work on Fast R-CNN demonstrated that applying such filters significantly improved the accuracy of object detection by reducing noise and focusing on the most likely candidates. Likewise, the study by Liu et al. (2022) also utilizes a minimum size threshold for excluding small, irrelevant objects in object detection tasks. They introduced a multi-scale feature extraction method that filtered out small objects, enhancing the overall performance of the detection system. This concept is similarly applied in our work, where the minimum area thresholding is used to filter out small objects, improving the reliability of detecting larger, more relevant smoke plumes.

The combination of multi-rotation and area thresholding provides the best overall performance, leveraging the strengths of both techniques. Multi-rotation strengthens detection reliability by ensuring consistency across different perspectives, while area thresholding eliminates small, irrelevant objects that could otherwise be misclassified as smoke. This integrated approach enhances the model's ability to distinguish true smoke instances while minimizing false positives more effectively than either technique alone. By applying both filtering mechanisms, the model achieves improved precision without excessively sacrificing recall. This combination results in a model that is both more stable and more reliable for operational wildfire detection, as it effectively reduces the impact of false positives while maintaining high specificity. However, this combined approach comes at the cost of reduced sensitivity, similar to the multi-rotation method alone. The stricter filtering criteria introduced by the area

thresholding limit the detection of smaller or partially obscured smoke signals. While the combined method significantly improves precision and specificity, the reduction in sensitivity emphasizes the ongoing challenge of balancing the detection of faint or partially obscured smoke with the need to minimize false positives.

The improvements provided by segmentation, multi-rotation, and area thresholding demonstrate that segmentation alone is not sufficient to achieve optimal performance in smoke detection. While increasing resolution can enhance precision and specificity, it does not completely resolve the challenges of false positives or unreliable detections. Each method contributed unique strengths to the smoke detection process, with the simple analysis method serving as an effective baseline. The choice of approach should depend on the operational context, balancing the need for precision with the need to detect all possible instances of smoke.

4.2 Tuning Process Review

The tuning process aimed to refine the YOLOv8 model's smoke detection capabilities by adapting it to the complex challenges of wildfire monitoring. This process consisted of three distinct phases, each designed to optimize different aspects of the model's performance. Initially, the model demonstrated high precision but struggled with a large number of false positives. The high number of false positives reduced precision and highlighted the difficulty in distinguishing smoke from other environmental features, such as shadows, vegetation, and terrain. Sensitivity was also relatively low, which meant that the model failed to detect a significant number of smoke instances.

Following the first tuning phase, precision improved, and the number of false positives decreased, suggesting that the model was becoming better at identifying true instances of smoke. The second tuning phase further improved precision, reducing false positives, and specificity reached near-perfect levels. This suggests that this phase helped the model become better at excluding irrelevant features such as roads, rocks, and vegetation. The F2 score showed significant improvement, reflecting the model's better capability to detect true positives without increasing false positives. This is a positive outcome because high specificity helps reduce false alarms. In terms of F1 and F2 scores, both improved throughout the tuning phases. The second phase exhibited the highest F1 score, showing that the model was increasingly capable of balancing precision and sensitivity. The F2 score showed even more substantial improvements, with the second phase achieving the highest values. These findings

are consistent with Yang et al. (2024), who improved the YOLOv8 model for wildfire detection by incorporating deformable convolutions and CoordAtt mechanisms, achieving significant precision improvements without compromising recall.

Interestingly, when the multi-rotation method was applied to the tuned models, the results were less favorable than when applied to the original model. The original model showed a more significant improvement in performance, especially regarding increased precision as more rotations were applied. In contrast, the fine-tuned models showed less improvement with the multi-rotation method, suggesting that the multi-rotation approach may be more effective when applied to the original YOLOv8 model compared to its use on a finely-tuned version. The fine-tuned model, having already been adapted to specific dataset characteristics, may have become too selective in its detection, with the multi-rotation method adding complexity without providing significant benefits. While the multi-rotation method remains valuable, its effectiveness should be evaluated based on the current state of the model. This finding aligns with Saydirasulovich et al. (2023), who also fine-tuned YOLOv8 for wildfire smoke detection and noted that further modifications to the model, did not always improve performance significantly.

The analysis of image resolutions and confidence levels revealed that the model's performance was influenced by these factors. At lower resolutions, the model tended to detect more smoke instances but at the cost of a higher number of false positives, which lowered precision. Higher resolutions reduced false positives but also resulted in fewer true detections, as the model became more selective. The 2000x2000 resolution provided the best balance between detecting true smoke instances and minimizing false positives. The model performed better at lower confidence levels, as these settings allowed it to capture more instances of smoke, though they led to a higher number of false positives. Increasing the confidence level beyond this range improved precision but decreased sensitivity, leading to fewer detected instances of smoke. A comparison of Flight 2 and Flight 3 showed a consistent pattern, with Flight 3 generally yielding higher F2 scores at lower confidence thresholds. This suggests that the dataset from Flight 3 contained clearer smoke patterns, which made detections more reliable. Both flights confirmed that the 2000x2000 resolution, combined with moderate confidence levels, was the optimal choice for achieving the highest F2 scores. Similar conclusions were reached by Ramos et al. (2024), who found that hyperparameter optimization of YOLOv8, particularly in adjusting confidence thresholds, had a significant impact on precision and recall, underscoring the importance of fine-tuning these parameters based on the specific dataset and environment.

The third phase of tuning and its testing on an external database, which had characteristics different from those of the datasets obtained from the two flights in the Büelenwald region, highlighted the limitations of the tuning process. The original model performed better than the models produced from our datasets. In most cases, smoke was correctly detected, with a relatively low number of false positives. The model obtained from the third tuning phase was able to classify a considerable number of smoke-containing images with a higher confidence level, allowing the use of a higher confidence level in practical applications. However, it also revealed the shortcomings of the tuning process, producing a significantly higher number of false positives compared to the original model. These challenges are consistent with findings from Rothmeier et al. (2024), who explored the fine-tuning of object detection models for adverse weather conditions. Their study demonstrated that, although the model showed significant improvements under specific conditions, the performance dramatically decreased when applied to general conditions outside the fine-tuned scenario (Rothmeier et al., 2024). This overfitting issue, where the model excels on the specialized dataset but struggles with more generalized data, mirrors the difficulties observed in our smoke detection model after fine-tuning. The increased number of false positives in the practical application phase highlights a similar risk of overfitting, where the model becomes too tailored to the training data and loses its ability to generalize effectively, thus affecting the robustness of its predictions in real-world scenarios.

It is important to note that a large portion of these false positives were caused by a few features that repeatedly triggered false detections. One example is the sky, which, in the case of the original algorithm, did not cause significant problems, but this was not the case with our final tuned model. These false positives were triggered by the similarity between images captured by UAVs with a horizontal perspective of the ground, where a grey sky is depicted in the upper part of the image and green vegetation below. This composition closely resembled the images used in the tuning process, where often, in 2000x2000 resolution images, a portion of the image depicted smoke opposing the green vegetation underneath. However, in our case, the images were captured using a vertical perspective of the terrain studied. Other elements that occasionally caused false positives included rocks resembling columns of smoke and puddles of water. In these cases, the problem was observed in both the original model and the fine-tuned one.

Figure 4.3: A false positive detected in the "Flame 2" dataset (Hopkins et al., 2023) by the tuned algorithm caused by the grey sky opposed to the green vegetation (left) and a segment of image used for the third phase of the tuning process that sees grey smoke opposed to green vegetation (right)

4.3 Optimizing Early Wildfire Detection with UAVs and YOLOv8

The combined use of optical sensors and deep learning models, such as YOLOv8, for wildfire detection represents an innovative and effective approach compared to traditional systems. The use of UAVs equipped with optical sensors offers significant advantages over optical satellites, particularly in terms of real-time data collection, higher spatial resolution, and operational flexibility. However, with recent advancements in optical satellite technology, high-resolution optical satellites have significantly improved large-scale detection, providing broad coverage and continuous monitoring (Li et al., 2019). In this context, it is valuable to explore the differences and synergies between these two systems and how UAV-based systems with YOLOv8 compare to the state-of-the-art in wildfire detection.

UAVs equipped with optical sensors offer numerous operational advantages over optical satellites, especially in wildfire detection. One key differentiator is the ability of UAVs to capture real-time data. While optical satellites are limited by their revisit times, often only once or twice a day, UAVs can be deployed on-demand to fly over a specific wildfire or a high-risk area, delivering immediate data (Kang et al., 2023). This ability to provide real-time feedback is crucial for fire management teams, enabling faster decision-making and resource allocation to mitigate wildfire damage (Okoro et al., 2024). Furthermore, UAVs offer superior spatial

resolution, particularly in complex terrains such as dense forests or mountainous regions, where optical satellites may struggle to detect small smoke plumes or subtle environmental changes (Akhloufi et al., 2021). With high-resolution cameras, UAVs can capture more detailed imagery, which is critical for detecting early signs of wildfire, such as faint smoke or minor temperature variations. Unlike satellites, which may not be able to identify subtle indicators of fire in thick vegetation, UAVs can provide precise and localized data, improving the reliability of early detection (Boroujeni et al., 2024). Operational flexibility is another advantage of UAVs. Unlike satellites, which follow predetermined orbits, UAVs can be dispatched at any time to monitor a wildfire's progression. This makes UAVs particularly suitable for tracking fires across irregular and rugged terrains, where other monitoring systems may not be able to provide accurate or timely data. Additionally, UAVs are more cost-effective than satellite systems, which require substantial infrastructure and operational costs (Bailon et al., 2022). UAVs can be deployed at a fraction of the cost, enabling scalable and widespread monitoring, especially in remote areas or regions where traditional systems may be limited (Chen et al., 2022).

Despite their advantages, UAVs do have limitations compared to optical satellites. One of the primary challenges is their limited flight time due to battery constraints. UAVs are typically restricted to shorter flight durations, meaning their coverage is limited to smaller areas. In contrast, satellites can monitor vast areas continuously without the need for recharging, making them more suitable for long-term, large-scale wildfire detection (Xu et al., 2021). Another limitation of UAVs is their susceptibility to weather conditions. Strong winds, rain, or other adverse weather conditions can interfere with UAV operations, potentially compromising their effectiveness during critical wildfire events. Satellites, particularly those in geostationary orbit, are generally unaffected by such conditions, providing more consistent coverage regardless of local weather (Bushnaq et al., 2021). The development of high-resolution optical satellites has significantly enhanced their role in wildfire detection. Modern satellites can now capture images with resolutions as fine as 30 cm, providing highly detailed imagery of large areas, including fire perimeters, smoke plumes, and land-use changes (Xu et al., 2021). These high-resolution satellites are crucial for monitoring the spread of wildfires in remote or hard-to-reach areas, where ground-based sensors or UAVs may not be able to provide effective coverage. Despite these advancements, optical satellites still face limitations, particularly in terms of temporal resolution. The revisit time of satellites is often too slow for early wildfire detection, which is critical for minimizing damage. Wildfires can develop rapidly, and the ability to detect

them in their early stages is essential for effective intervention (Kang et al., 2023). While high-resolution satellites offer detailed imagery, they may not provide the same level of immediacy or accuracy as UAVs, which can monitor fire behavior in real time. Additionally, while high-resolution satellites improve detection accuracy, they are often limited by atmospheric interference. Clouds, dust, and smoke can obscure satellite imagery, making it difficult to monitor fires during certain weather conditions (Bushnaq et al., 2021). UAVs, however, can be deployed immediately over a fire, avoiding the delays caused by cloud cover and other atmospheric factors.

The integration of YOLOv8 with UAVs marks a significant advancement in wildfire detection. YOLOv8 is a deep learning-based model known for its real-time object detection capabilities, which makes it highly suitable for dynamic environments like wildfire zones. The model can process images in real time, identifying key features such as smoke or flames, which is critical for early detection and rapid response (Jiang et al., 2022). YOLOv8's ability to classify and localize objects with high precision makes it particularly effective for detecting early signs of wildfires, providing immediate alerts to fire management teams (Jiang et al., 2024). Compared to traditional satellite systems, YOLOv8-equipped UAVs offer faster and more precise detection, minimizing false positives and false negatives. Unlike optical satellites, which often require manual interpretation of imagery, YOLOv8 can automatically detect and classify wildfire-related features, improving efficiency and accuracy (Terven et al., 2023). The model can be trained to detect both smoke and fire, enabling UAVs to capture detailed, real-time data of both the initiation and progression of wildfires (Zhang et al., 2016). YOLOv8’s integration with UAVs allows for high-resolution data collection, enabling real-time fire detection while also minimizing operational costs. This combination of cutting-edge AI technology and cost-effective drone platforms presents a scalable solution for wildfire management, particularly in regions where satellite monitoring might be cost-prohibitive or insufficient (Boroujeni et al., 2024).

4.4 Methodological limitations and Operational Considerations

This study presents several limitations related to the data collection process, environmental variables, and the model's application to real-world wildfire scenarios. The primary dataset used to train and test YOLOv8 was derived from controlled fire scenarios. While these controlled environments provided valuable insights into the model’s performance, they may

not fully replicate the dynamic, unpredictable conditions of natural wildfires. The controlled fires had limited variations in smoke intensity and patterns, which may have restricted the model's ability to generalize to more complex, real-world wildfire scenarios.

Environmental variability, such as changes in weather conditions, terrain complexity, and vegetation density, posed challenges to the model's performance. Dense forests, mountainous terrain, and varying atmospheric conditions can affect the UAV's ability to capture clear smoke signals. While the study accounted for diverse conditions through controlled flights, real-world scenarios often present greater unpredictability, which could further limit the model's effectiveness. For example, forest canopies can obscure smoke, and variations in lighting, such as glare or shadows, can complicate detection, introducing noise that may interfere with accurate smoke identification. Elements such as clouds and fog produced by the typical evaporation of forests can also constitute important disturbing elements in mountain contexts that must be prevented through specific and targeted tuning processes.

Another key limitation relates to the dataset itself. The "D-Fire" dataset, used for training YOLOv8. Thus dataset is relatively small and specific, containing images from limited conditions that may not represent the full range of smoke appearances, intensities, and dispersions encountered in wildfires. This limits the model's ability to generalize to a wider variety of wildfire scenarios. Additionally, while YOLOv8 excels in real-time detection, its performance is still constrained by the quality of input data, such as image resolution, lighting, and the presence of visual obstructions. In real-world settings, smoke may be faint or obscured by trees, buildings, or other objects, which could affect detection accuracy.

The image resolution and segmentation process also presented challenges. While segmenting full-resolution images into smaller regions improved detection by focusing on smaller areas, it also reduced spatial context, particularly for larger smoke plumes that span multiple segments. This limitation may have caused the model to miss broader features of smoke, which could be critical for identifying larger or more significant wildfires. Additionally, the trade-off between image resolution and computational demands remains a challenge. Lower resolution improves real-time processing but sacrifices fine detail, whereas higher resolution provides more detail but increases computational load, complicating fast decision-making in emergency scenarios.

The generalizability of the findings is another important consideration. While this study demonstrated promising results in the Büelenwald region, environmental factors such as climate, terrain, and vegetation vary significantly across wildfire-prone areas globally. These

factors, along with local fire behavior, could impact the model's performance in other regions. The model's reliance on a relatively specific dataset also limits its ability to generalize to unforeseen wildfire scenarios.

Operational and practical limitations must also be addressed. UAVs are constrained by battery life, limiting their flight duration and coverage area in a single mission. Additionally, the complexities of operating UAVs in challenging terrains or adverse weather conditions could hinder their widespread adoption for wildfire monitoring. Effective communication systems, robust flight planning, and coordination with firefighting teams will be essential for integrating UAVs into wildfire monitoring systems. Moreover, the real-time data processing capabilities of UAVs need further refinement to ensure immediate, actionable insights during wildfire events. External computational resources or onboard processing may become bottlenecks in situations requiring rapid decision-making.

During the analysis, several additional limitations became apparent. One key limitation is the variability in pixel representation, which is influenced by changes in drone altitude and the lack of geometric corrections in the images. In practice, a fixed pixel threshold for area-based filtering may correspond to different physical ground sizes depending on the altitude at which the drone operates. This variability can introduce discrepancies, as a threshold that is appropriate at a lower altitude might miss smaller smoke plumes, while at higher altitudes, it could erroneously include irrelevant features such as small shadows or vegetation clusters. As such, the effectiveness of area thresholding becomes highly dependent on the altitude and geometry of the images, requiring careful calibration for different flight conditions.

Another limitation arises from the reliance on RGB imagery alone for smoke detection. While RGB images provide valuable visual data, they can also introduce ambiguities, especially when smoke is faint or partially obscured by other environmental factors. Smoke signals can be visually similar to mist, fog, or cloud cover, making it challenging to differentiate between these phenomena solely based on color information. This issue is particularly prevalent in dense forest environments, where the background can be highly cluttered and visually complex. As a result, the model may struggle to accurately distinguish between smoke and other natural features, leading to both false positives and missed detections.

The environmental diversity across datasets also poses a challenge. While fine-tuning on datasets from different flight conditions improved the model's performance, it also highlights the variability in terrain, lighting, and atmospheric conditions. These factors can greatly

influence the visibility of smoke, as lighting changes can alter the color of smoke or even obscure it completely. Thus, the model's effectiveness in detecting smoke in diverse environmental conditions may require more extensive training on varied datasets that reflect the broad range of real-world scenarios.

Lastly, the need for real-time processing in real-world applications adds computational challenges, particularly when the model has to handle high-resolution images or large areas during flight. Future UAV-based systems need to address these challenges by exploring lighter versions of the model or techniques like model pruning or quantization, which reduce computational load without sacrificing detection accuracy.

While this study demonstrated significant improvements in smoke detection through fine-tuning and advanced methods, there are still several challenges to address. The need for more training datasets, the ambiguity introduced by relying solely on RGB imagery, and the potential for enhanced detection through multi-spectral and thermal sensors suggest that addressing these limitations will be crucial for the further development and deployment of UAV-based smoke detection systems.

5. Conclusion and Outlook

5.1 Restatement of the Research Problem

The increasing frequency and severity of wildfires make early detection and rapid response critical for minimizing their impact on ecosystems, public health, and infrastructure. Wildfires threaten biodiversity, air quality, and human lives, highlighting the need for more effective detection solutions. Traditional methods like satellite remote sensing, airborne monitoring, ground patrols, and fixed cameras have limitations, including resolution, real-time capabilities, and adaptability to environmental changes. Satellites face delays due to their orbit, airborne methods are costly and weather-dependent, and ground-based methods are resource-intensive and limited by terrain. This study focused on UAVs equipped with optical sensors, which offer flexibility, cost-effectiveness, and real-time data collection, addressing many traditional system limitations. It assessed whether UAV-mounted optical sensors, paired with deep learning models like YOLOv8, could provide a more efficient and complementary means of early smoke detection. The research also examined how environmental factors such as terrain, vegetation, and weather affect detection accuracy and timeliness. Understanding these influences is critical for optimizing detection in various environments. The goal was to evaluate if optical sensors could serve as a cost-effective alternative to thermal sensors, explore how environmental conditions impact their reliability and consider how they can be integrated into current surveillance methods.

5.2 Summary of key findings

5.2.1 Optical Sensors as a Viable Alternative to Thermal Sensors

While thermal sensors have been the standard for detecting flames due to their ability to identify heat signatures, this study showed that optical sensors, paired with deep learning models like YOLOv8, offer a viable and cost-effective complementary alternative for early smoke detection, though they still face generalization challenges. Thermal sensors are effective for identifying flames or hotspots but lose effectiveness in the early stages of a wildfire, when smoke is present but flames are not visible. Additionally, environmental factors like temperature changes and reflective surfaces can cause false positives or delays in detection. In

contrast, optical sensors detect smoke in the visible spectrum, making them ideal for early wildfire response, as they can identify smoke before flames appear. Unlike thermal sensors, they do not rely on temperature differences, and they are lightweight and cost-effective, which makes them well-suited for UAV-based monitoring systems that prioritize low weight and cost. UAVs equipped with optical sensors can efficiently cover large areas and provide real-time visual data for continuous monitoring, especially in regions where thermal sensors are too costly or impractical. When combined with models like YOLOv8, optical sensors enable rapid and accurate smoke detection, even in complex terrains or dense vegetation.

Although optical sensors have limitations, such as reduced effectiveness in low-visibility conditions like heavy smoke or fog, integrating them with thermal sensors provides a more comprehensive solution. This study confirms that optical sensors, when paired with machine learning models, offer a reliable and cost-effective alternative or complement to thermal sensors for early smoke detection, improving response times and supplementing traditional methods. Their cost-effectiveness, real-time data capabilities, and ability to detect early smoke make them a promising option for large-scale, UAV-based wildfire detection. With further optimization, optical sensors could play a crucial role in enhancing the efficiency and scalability of wildfire detection, particularly in regions with limited access to traditional thermal sensing technologies.

5.2.2 YOLOv8 Performance, Image Resolution, and Enhancement Techniques

A key finding of the study was the impact of image resolution on detection accuracy. While full-resolution images provide high detail, the resizing process for YOLOv8 can result in the loss of critical features needed to detect faint smoke patterns. To address this, image segmentation techniques were used, dividing full-resolution images into smaller sections of 2000x2000 and 1000x1000 pixels. This allowed the model to focus on smaller, more relevant regions, improving accuracy and sensitivity. The study also explored enhancement techniques like multi-rotation and area threshold filtering to optimize YOLOv8's performance. Multi-rotation involved rotating images at four orientations, allowing the model to detect smoke from different perspectives, especially useful when smoke appeared differently depending on the UAV's orientation. Analyzing the images on these varied perspectives significantly improved its robustness in real-world scenarios, ensuring consistent smoke detection from multiple angles. Area threshold filtering ensured that only objects above a certain size were considered

valid detections, eliminating irrelevant objects or background noise that could be misidentified as smoke. By setting a minimum size threshold for smoke plumes, this technique reduced false positives and increased accuracy. The combination of multi-rotation and area threshold filtering proved highly effective. Multi-rotation accounted for changing smoke orientations, while area threshold filtering ensured only relevant detections were considered. Together, these enhancements significantly improved the model's ability to detect smoke, especially in challenging environments where smoke direction and scale varied.

5.2.3 Fine-tuning and the Challenges of Generalizing Smoke Detection

One of the key findings of this study was the significant improvement in YOLOv8's performance through fine-tuning, tailored to the unique characteristics of wildfire datasets from UAV flights. The fine-tuning process was conducted in three phases. In the first phase, YOLOv8 was adapted to specific environmental conditions like lighting, terrain complexity, and smoke density using datasets from Flights 2 and 3. Smoke images were augmented with rotations at four orientations to help the model distinguish smoke from other features. In the second phase, the training dataset was expanded using image mirroring, tripling the smoke dataset and exposing the model to diverse conditions. Key hyperparameters such as learning rate and batch size were adjusted, with training lasting 50 epochs in the first phase and 100 in the second, allowing the model to focus on fine details through a reduced learning rate. The first two phases of fine-tuning demonstrated the potential of adapting YOLOv8 to specific conditions, like those in the Büelenwald region, reducing false positives and improving detection accuracy. Tailoring the model to regional characteristics made it more reliable in detecting smoke.

Despite these advances, the third phase revealed challenges in generalizing the model across diverse environments. Smoke varies in density, color, and form based on fire intensity, weather, and landscape, and factors like UAV altitude, distance, and angle can alter detection. Environmental elements such as water, rocks, or varying lighting can also resemble smoke, leading to false positives. This third phase highlighted the need for a more robust, universal model. More diverse and extensive training datasets are crucial for improving the general performance of the model. While fine-tuning worked well for Büelenwald, region-specific models are not feasible for all areas. Therefore, the most important conclusion is that significantly increasing training data, covering a broader range of environmental conditions,

will enhance the general performance of wildfire detection models. This would create a more robust model that could be adapted for specific applications and perform reliably across diverse environments.

5.2.4 Environmental Influences and Detection Trade-offs

This study highlighted the significant role environmental factors play in the performance of YOLOv8 for wildfire smoke detection. Terrain, vegetation density, weather, and lighting conditions were found to heavily influence the model's ability to reliably detect smoke. Dense vegetation and complex terrain, like forests or mountainous regions, can obstruct smoke detection by optical sensors, making it hard to see smoke plumes. Lighting conditions, such as glare, shadows, or low light, can further complicate detection, as varying light angles throughout the day may obscure smoke. Atmospheric factors like humidity and wind also impact detection, with humidity influencing smoke dispersion and wind causing smoke to shift, making it difficult to track consistently. These dynamic conditions require continual refinement of the detection algorithms to adapt to changing environments, ensuring accurate smoke detection.

Another key finding was the trade-off between sensitivity and false positives. With a low confidence threshold, YOLOv8 had higher sensitivity, detecting faint or partially obscured smoke, but this led to more false positives, with features like clouds or shadows being misidentified as smoke. Increasing the threshold reduced false positives but also missed detections, decreasing true positives. This emphasizes the need for a balance between sensitivity and specificity, crucial in wildfire detection, where both false negatives and positives must be minimized.

The findings suggest that careful calibration of detection thresholds is essential for reliable early smoke detection while minimizing false alarms. Furthermore, incorporating more training data that represents a variety of environmental conditions would significantly improve model performance. Adjusting and fine-tuning the model with diverse training data is crucial for optimizing its performance in real-world wildfire scenarios. By continually refining the model with data reflecting changing environments, smoke detection accuracy under diverse conditions can be enhanced.

5.3 Insight and Future Perspectives

This study significantly contributes to wildfire detection by evaluating and improving UAV-based systems using optical sensors integrated with deep learning models. Traditional methods, such as thermal imaging and satellites, are limited by their real-time responsiveness and adaptability, as well as their inability to provide detailed data in areas difficult to reach. In contrast, combining optical sensors with models like YOLOv8 offers a scalable, cost-effective, and flexible solution for real-time wildfire monitoring. These systems can be quickly deployed in underserved or hard-to-reach regions, crucial for global wildfire management.

YOLOv8's ability to process images in real-time and detect smoke accurately, even in complex environments like varying terrain and lighting, marks a significant advancement in deep learning for environmental monitoring. This real-time detection enables rapid intervention, potentially saving lives, protecting ecosystems, and minimizing wildfire impacts. UAVs with optical sensors provide proactive monitoring, allowing fire management teams to act sooner than with traditional methods, improving resource allocation and reducing wildfire spread. This research also highlights the intersection of machine learning and remote sensing in environmental monitoring. YOLOv8's use of data augmentation techniques like multi-rotation and area threshold filtering improves performance in dynamic environments. Fine-tuning with domain-specific datasets focused on smoke detection further enhances model accuracy. The integration of high-quality, diverse training data equips the model to generalize across various wildfire scenarios, showing potential for broader applications such as deforestation monitoring and wildlife conservation.

The environmental and societal benefits of early wildfire detection are significant. It helps reduce biodiversity loss, soil degradation, and harmful pollutant release by enabling faster response times. Early detection can limit fire intensity and duration, protecting ecosystems and reducing greenhouse gas emissions. From a societal perspective, it minimizes health risks by reducing exposure to harmful wildfire smoke, which contains pollutants that affect respiratory health, particularly in vulnerable populations. Faster responses also improve evacuation plans and reduce firefighting costs.

As wildfires increase due to climate change, governments must adopt more scalable, effective monitoring technologies. UAV-based systems offer a flexible solution, and this research provides a detailed analysis of integrating UAVs into operational smoke detection. Policymakers could incorporate UAVs into national or regional wildfire monitoring,

establishing guidelines for their safe deployment. This would encourage adoption, equipping fire management agencies with an essential tool for combating severe wildfires. Furthermore, integrating UAV data into broader wildfire management systems, such as GIS, can enhance fire risk predictions, evacuation plans, and air quality monitoring. UAVs could also support long-term monitoring of fire-affected areas, aiding ecosystem recovery and biodiversity restoration. Post-fire UAV imagery can track vegetation changes, soil erosion, and regrowth, supporting environmental recovery strategies.

5.4 Final Remarks

This research highlights the potential of UAV-based optical sensors combined with deep learning models like YOLOv8 for enhancing early wildfire detection. Detecting smoke before flames become visible is crucial for minimizing wildfires' impact on ecosystems, human populations, and infrastructure. UAVs offer a cost-effective, scalable, and flexible solution, addressing limitations in traditional detection systems like resolution, temporal coverage, and adaptability. YOLOv8, when fine-tuned, proved effective in detecting smoke in diverse environments, with techniques such as data augmentation and resolution segmentation improving sensitivity and reducing false positives.

However, challenges remain, especially with environmental factors like terrain and weather, which affect detection accuracy in real-world wildfire scenarios. Despite these challenges, the study lays a strong foundation for future work, focusing on integrating multi-sensor systems, refining smoke detection algorithms, and testing UAVs in live wildfire events. It also emphasizes the need for collaboration between machine learning, remote sensing, and wildfire management experts to develop comprehensive solutions.

As wildfires become more frequent and severe, the development of advanced technologies for timely, actionable insights is crucial for firefighting teams and emergency responders. Persistent advancements in UAV technology and deep learning algorithms promise to improve wildfire detection's speed, accuracy, and scalability. Integrating these technologies into existing management frameworks could significantly enhance wildfire detection, response, and mitigation efforts. Ultimately, this study contributes to both scientific understanding and real-world applications, positioning UAV-based optical sensors and YOLOv8 as key tools in global wildfire management, helping save lives, protect ecosystems, and reduce the social and economic costs of wildfires.

6 Literature

Akhloufi, M. A., Couturier, A., & Castro, N. A. (2021). Unmanned aerial vehicles for wildland fires: Sensing, perception, cooperation and assistance. Drones, 5(1), 15.

Aslan, S., Güdükbay, U., Töreyin, B. U., & Cetin, A. E. (2019, May). Early wildfire smoke detection based on motion-based geometric image transformation and deep convolutional generative adversarial networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8315-8319). IEEE.

Bailon-Ruiz, R., Bit-Monnot, A., & Lacroix, S. (2022). Real-time wildfire monitoring with a fleet of UAVs. Robotics and Autonomous Systems, 152, 104071.

Barmpoutis, P., Papaioannou, P., Dimitropoulos, K., & Grammalidis, N. (2020). A review on early forest fire detection systems using optical remote sensing. Sensors, 20(22), 6442.

Barmpoutis, P., Stathaki, T., Dimitropoulos, K., & Grammalidis, N. (2020). Early fire detection based on aerial 360-degree sensors, deep convolution neural networks and exploitation of fire dynamic textures. Remote Sensing, 12(19), 3177.

Barmpoutis, P., & Stathaki, T. (2020). A novel framework for early fire detection using terrestrial and aerial 360-degree images. In Advanced Concepts for Intelligent Vision Systems: 20th International Conference, ACIVS 2020, Auckland, New Zealand, February 10–14, 2020, Proceedings 20 (pp. 63-74). Springer International Publishing.

Benzekri, W., El Moussati, A., Moussaoui, O., & Berrajaa, M. (2020). Early forest fire detection system using wireless sensor network and deep learning. International Journal of Advanced Computer Science and Applications, 11(5).

Boroujeni, S. P. H., Razi, A., Khoshdel, S., Afghah, F., Coen, J. L., O'Neill, L., ... & Vamvoudakis, K. G. (2024). A comprehensive survey of research towards AI-enabled unmanned aerial systems in pre-, active-, and post-wildfire management. Information Fusion, 102369.

Bouguettaya, A., Zarzour, H., Taberkit, A. M., & Kechida, A. (2022). A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. Signal Processing, 190, 108309.

Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010, August). The balanced accuracy and its posterior distribution. In 2010 20th international conference on pattern recognition (pp. 3121-3124). IEEE.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. Proceedings of the National Academy of Sciences, 118(2), e2011048118.

Bushnaq, O. M., Chaaban, A., & Al-Naffouri, T. Y. (2021). The role of UAV-IoT networks in future wildfire detection. IEEE Internet of Things Journal, 8(23), 16984-16999.

Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverría, F. (2023). Assessing the effectiveness of YOLO architectures for smoke and wildfire detection. IEEe Access, 11, 96554-96583.

Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverria, F. (2024). Yolov5 vs. yolov8: Performance benchmarking in wildfire and smoke detection scenarios. Journal of Image and Graphics, 12(2), 127-136.

Çetin, A. E., Dimitropoulos, K., Gouverneur, B., Grammalidis, N., Günay, O., Habiboğlu, Y. H., ... & Verstockt, S. (2013). Video fire detection–review. Digital Signal Processing, 23(6), 1827-1843.

Chen, Y., Zhang, Y., Xin, J., Yi, Y., Liu, D., & Liu, H. (2018, July). A UAV-based forest fire detection algorithm using convolutional neural network. In 2018 37th Chinese Control Conference (CCC) (pp. 10305-10310). IEEE.

Chen, X., Hopkins, B., Wang, H., O'Neill, L., Afghah, F., Razi, A., ... & Watts, A. (2022). Wildland fire detection and monitoring using a drone-collected rgb/ir image dataset. IEEE Access, 10, 121301-121317.

Chuvieco, E., Aguado, I., Salas, J., García, M., Yebra, M., & Oliva, P. (2020). Satellite remote sensing contributions to wildland fire science and management. Current Forestry Reports, 6, 81-96.

Crowley, M. A., Stockdale, C. A., Johnston, J. M., Wulder, M. A., Liu, T., McCarty, J. L., ... & White, J. C. (2023). Towards a whole-system framework for wildfire monitoring using Earth observations. Global change biology, 29(6), 1423-1436.

de Venancio, P. V. A., Lisboa, A. C., & Barbosa, A. V. (2022). An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices. Neural Computing and Applications, 34(18), 15349-15368.

Diwan, T., Anirudh, G., & Tembhurne, J. V. (2023). Object detection using YOLO: Challenges, architectural successors, datasets and applications. multimedia Tools and Applications, 82(6), 9243-9275.

Dimitropoulos, K., Barmpoutis, P., & Grammalidis, N. (2016). Higher order linear dynamical systems for smoke detection in video surveillance applications. IEEE Transactions on Circuits and Systems for Video Technology, 27(5), 1143-1154.

Dupuy, J. L., Fargeon, H., Martin-StPaul, N., Pimont, F., Ruffault, J., Guijarro, M., ... & Fernandes, P. (2020). Climate change impact on future wildfire danger and activity in southern Europe: a review. Annals of Forest Science, 77, 1-24.

Fischer, A., Kotlarski, S., Strassmann, K., Schär, C., Croci-Maspoli, M., Knutti, R., & Schwierz, C. (2019, January). The new Swiss climate change scenarios CH2018. In Geophysical Research Abstracts (Vol. 21).

Fu, K., Chang, Z., Zhang, Y., Xu, G., Zhang, K., & Sun, X. (2020). Rotation-aware and multi-scale convolutional neural network for object detection in remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing, 161, 294-308.

Gomez Isaza, D. F., Cramp, R. L., & Franklin, C. E. (2022). Fire and rain: A systematic review of the impacts of wildfire and associated runoff on aquatic fauna. Global Change Biology, 28(8), 2578-2595.

Ghetu, C. C., Rohlman, D., Smith, B. W., Scott, R. P., Adams, K. A., Hoffman, P. D., & Anderson, K. A. (2022). Wildfire impact on indoor and outdoor PAH air quality. Environmental science & technology, 56(14), 10042-10052.

Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

Goyal, S., Shagill, M., Kaur, A., Vohra, H., & Singh, A. (2020). A yolo based technique for early forest fire detection. Int. J. Innov. Technol. Explor. Eng, 9, 1357-1362.

Gonçalves, L. A. O., Ghali, R., & Akhloufi, M. A. (2024). YOLO-Based models for smoke and Wildfire Detection in Ground and aerial images. Fire, 7(4), 140.

Goss, M., Swain, D. L., Abatzoglou, J. T., Sarhadi, A., Kolden, C. A., Williams, A. P., & Diffenbaugh, N. S. (2020). Climate change is increasing the likelihood of extreme autumn wildfire conditions across California. Environmental Research Letters, 15(9), 094016.

Govil, K., Welch, M. L., Ball, J. T., & Pennypacker, C. R. (2020). Preliminary results from a wildfire detection system using deep learning on remote camera images. Remote Sensing, 12(1), 166.

Hao, Y., Pei, H., Lyu, Y., Yuan, Z., Rizzo, J. R., Wang, Y., & Fang, Y. (2023, October). Understanding the impact of image quality and distance of objects to object detection performance. In 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 11436-11442). IEEE.

Hasan, M. W., Shanto, S., Nayeema, J., Rahman, R., Helaly, T., Rahman, Z., & Mehedi, S. T. (2024). An Explainable AI-Based Modified YOLOv8 Model for Efficient Fire Detection. Mathematics, 12(19), 1-21.

Hopkins, B., O'Neill, L., Afghah, F., Razi, A., Rowell, E., Watts, A., ... & Coen, J. (2023). Flame 2: Fire detection and modeling: Aerial multi-spectral image dataset. IEEE DataPort.

Hossain, F. A., Zhang, Y. M., & Tonima, M. A. (2020). Forest fire flame and smoke detection from UAV-captured images using fire-specific color features and multi-color space local binary pattern. Journal of Unmanned Vehicle Systems, 8(4), 285-309.

Hrelja, I., Šestak, I., & Bogunović, I. (2020). Wildfire impacts on soil physical and chemical properties-a short review of recent studies. Agriculturae Conspectus Scientificus, 85(4), 293-301.

Jadon, A., Omama, M., Varshney, A., Ansari, M. S., & Sharma, R. (2019). FireNet: a specialized lightweight fire & smoke detection model for real-time IoT applications. arXiv preprint arXiv:1905.11922.

Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. Procedia computer science, 199, 1066-1073.

Kang, Y., Sung, T., & Im, J. (2023). Toward an adaptable deep-learning model for satellite-based wildfire monitoring with consideration of environmental conditions. Remote Sensing of Environment, 298, 113814.

Lee, W., Kim, S., Lee, Y. T., Lee, H. W., & Choi, M. (2017, January). Deep neural networks for wild fire detection with unmanned aerial vehicle. In 2017 IEEE international conference on consumer electronics (ICCE) (pp. 252-253). IEEE.

Li, C., Du, Y., Zhang, X., & Wu, P. (2025). YOLOGX: an improved forest fire detection algorithm based on YOLOv8. Frontiers in Environmental Science, 12, 1486212.

Li, X., Li, Z., Feng, R., Luo, S., Zhang, C., Jiang, M., & Shen, H. (2019). Generating high-quality and high-resolution seamless satellite imagery for large-scale urban regions. Remote Sensing, 12(1), 81.

Li, X., Wang, W., Hu, X., Li, J., Tang, J., & Yang, J. (2021). Generalized focal loss v2: Learning reliable localization quality estimation for dense object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11632-11641).

Liu, J., Yang, D., & Hu, F. (2022). Multiscale object detection in remote sensing images combined with multi-receptive-field features and relation-connected attention. Remote Sensing, 14(2), 427.

Lv, K. (2024). CCi-YOLOv8n: Enhanced Fire Detection with CARAFE and Context-Guided Modules. arXiv preprint arXiv:2411.11011.

Mansoor, S., Farooq, I., Kachroo, M. M., Mahmoud, A. E. D., Fawzy, M., Popescu, S. M., ... & Ahmad, P. (2022). Elevation in wildfire frequencies with respect to the climate change. Journal of Environmental management, 301, 113769.

Mukhiddinov, M., Abdusalomov, A. B., & Cho, J. (2022). A wildfire smoke detection system using unmanned aerial vehicle images based on the optimized YOLOv5. Sensors, 22(23), 9384.

Novac, I., Geipel, K. R., de Domingo Gil, J. E., de Paula, L. G., Hyttel, K., & Chrysostomou, D. (2020, January). A framework for wildfire inspection using deep convolutional neural networks. In 2020 IEEE/SICE International Symposium on System Integration (SII) (pp. 867-872). IEEE.

Okoro, S. C., Lopez, A., & Unuriode, A. (2024). A Synergistic Approach to Wildfire Prevention and Management Using AI, ML, and 5G Technology in the United States. arXiv preprint arXiv:2403.14657.

Ozge Unel, F., Ozkalayci, B. O., & Cigla, C. (2019). The power of tiling for small object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops.

Prasetiyo, B., Muslim, M. A., & Baroroh, N. (2021, June). Evaluation performance recall and F2 score of credit card fraud detection unbalanced dataset using SMOTE oversampling technique. In Journal of physics: conference series (Vol. 1918, No. 4, p. 042002). IOP Publishing.

Qiu, M., Chen, D., Kelp, M., Li, J., Huang, G., & Yazdi, M. D. (2025). The rising threats of wildland-urban interface fires in the era of climate change: The Los Angeles 2025 fires. The Innovation.

Ramos, L., Casas, E., Bendek, E., Romero, C., & Rivas-Echeverría, F. (2024). Hyperparameter optimization of YOLOv8 for smoke and wildfire detection: Implications for agricultural and environmental safety. Artificial Intelligence in Agriculture, 12, 109-126.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Reis, D., Kupec, J., Hong, J., & Daoudi, A. (2023). Real-time flying object detection with YOLOv8. arXiv preprint arXiv:2305.09972.

Rothmeier, T., Huber, W., & Knoll, A. C. (2024). Time to shine: Fine-tuning object detection models with synthetic adverse weather images. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 4447-4456).

Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. Plants, 8(11), 468.

Saydirasulovich, S. N., Mukhiddinov, M., Djuraev, O., Abdusalomov, A., & Cho, Y. I. (2023). An improved wildfire smoke detection based on YOLOv8 and UAV images. Sensors, 23(20), 8374.

Schumacher, S. A. B. I. N. E., & Bugmann, H. (2006). The relative importance of climatic effects, wildfires and management for future forest landscape dynamics in the Swiss Alps. Global Change Biology, 12(8), 1435-1450.

Shamsoshoara, A., Afghah, F., Razi, A., Zheng, L., Fulé, P. Z., & Blasch, E. (2021). Aerial imagery pile burn detection using deep learning: The FLAME dataset. Computer Networks, 193, 108001.

Syaufina, L., Rahmawan, H., Trisminingsih, R., Ammarullah, A. N., Sitanggang, I. S., & Ardiansyah, F. (2022, September). Development of Hotspots Ground Check Module in the Mobile Application for Forest and Land Fires Prevention Patrol. In 2022 International Conference on Electrical and Information Technology (IEIT) (pp. 333-337). IEEE.

Solovyev, R., Wang, W., & Gabruseva, T. (2021). Weighted boxes fusion: Ensembling boxes from different object detection models. Image and Vision Computing, 107, 104117.

Sousa, M. J., Moutinho, A., & Almeida, M. (2019). Classification of potential fire outbreaks: A fuzzy modeling approach based on thermal images. Expert systems with applications, 129, 216-232.

Swift, A., Heale, R., & Twycross, A. (2020). What are sensitivity and specificity?. Evidence-Based Nursing, 23(1), 2-4.

Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. Machine learning and knowledge extraction, 5(4), 1680-1716.

Thangavel, K., Spiller, D., Sabatini, R., Marzocca, P., & Esposito, M. (2022). Near real-time wildfire management using distributed satellite system. IEEE Geoscience and Remote Sensing Letters, 20, 1-5.

Tleuliyev, A. (2023). Stay ahead of the flames: A comprehensive guide to wildfire prevention with YOLOv8. Institute of Smart Systems and AI. Medium.

To, P., Eboreime, E., & Agyapong, V. I. (2021). The impact of wildfires on mental health: a scoping review. Behavioral Sciences, 11(9), 126.

Totakura, V., Vuribindi, B. R., & Reddy, E. M. (2021). Improved safety of self-driving car using voice recognition through CNN. In IOP Conference Series: Materials Science and Engineering (Vol. 1022, No. 1, p. 012079). IOP Publishing.

Viseras, A., Meissner, M., & Marchal, J. (2021). Wildfire front monitoring with multiple uavs using deep q-learning. IEEE Access.

Wang, Q., Ma, Y., Zhao, K., & Tian, Y. (2022). A comprehensive survey of loss functions in machine learning. Annals of Data Science, 9(2), 187-212.

Wastl, C., Schunk, C., Lüpke, M., Cocca, G., Conedera, M., Valese, E., & Menzel, A. (2013). Large-scale weather types, forest fire danger, and wildfire occurrence in the Alps. Agricultural and Forest Meteorology, 168, 15-25.

Wu, S., Yang, J., Wang, X., & Li, X. (2022). Iou-balanced loss functions for single-stage object detection. Pattern Recognition Letters, 156, 96-103.

Xu, S. G., Kong, S., & Asgharzadeh, Z. (2021, July). Wildfire detection using streaming satellite imagery. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS (pp. 2899-2902). IEEE.

Yadav, R. (2020). Deep learning based fire recognition for wildfire drone automation. Can. Sci. Fair J, 3(2), 1-8.

Yang, Z., Shao, Y., Wei, Y., & Li, J. (2024). Precision-boosted forest fire target detection via enhanced YOLOv8 model. Applied Sciences, 14(6), 2413.

Zhang, Q., Xu, J., Xu, L., & Guo, H. (2016, January). Deep convolutional neural networks for forest fire detection. In 2016 International forum on management, Education and information technology application (pp. 568-575). Atlantis Press.

Zhao, Y., Ma, J., Li, X., & Zhang, J. (2018). Saliency detection and deep learning-based wildfire identification in UAV imagery. Sensors, 18(3), 712.

Zhao, L., & Li, S. (2020). Object detection algorithm based on improved YOLOv3. Electronics, 9(3), 537.

Zhu, H., Wei, H., Li, B., Yuan, X., & Kehtarnavaz, N. (2020). A review of video object detection: Datasets, metrics and methods. Applied Sciences, 10(21), 7834.

Zumbrunnen, T., Pezzatti, G. B., Menéndez, P., Bugmann, H., Bürgi, M., & Conedera, M. (2011). Weather and human impacts on forest fires: 100 years of fire history in two climatic regions of Switzerland. Forest Ecology and Management, 261(12), 2188-2199.

7. Appendices

Table 7.1: Simple Image Analysis results on Flight 2 for different resolutions and confidence levels

Confidence	Resolution	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
0.1	1000	26112	25992	120	44	3547	22445	76	0.861	0.012	0.367	0.864	0.615	0.024	0.054
0.2	1000	26112	25992	120	39	1433	24559	81	0.942	0.026	0.325	0.945	0.635	0.049	0.100
0.3	1000	26112	25992	120	35	1199	24793	85	0.951	0.028	0.292	0.954	0.623	0.052	0.102
0.4	1000	26112	25992	120	27	851	25141	93	0.964	0.031	0.225	0.967	0.596	0.054	0.099
0.5	1000	26112	25992	120	20	584	25408	100	0.974	0.033	0.167	0.978	0.572	0.055	0.092
0.6	1000	26112	25992	120	15	376	25616	105	0.982	0.038	0.125	0.986	0.555	0.059	0.086
0.7	1000	26112	25992	120	6	202	25790	114	0.988	0.029	0.050	0.992	0.521	0.037	0.044
0.8	1000	26112	25992	120	0	41	25951	120	0.994	0.000	0.000	0.998	0.499	#DIV/0!	#DIV/0!
0.9	1000	26112	25992	120	0	0	25992	120	0.995	#DIV/0!	0.000	1.000	0.500	#DIV/0!	#DIV/0!
0.1	2000	6528	6447	81	30	1402	5045	51	0.777	0.021	0.370	0.783	0.576	0.040	0.085
0.2	2000	6528	6447	81	27	490	5957	54	0.917	0.052	0.333	0.924	0.629	0.090	0.161
0.3	2000	6528	6447	81	24	403	6044	57	0.930	0.056	0.296	0.937	0.617	0.094	0.160
0.4	2000	6528	6447	81	18	284	6163	63	0.947	0.060	0.222	0.956	0.589	0.094	0.144
0.5	2000	6528	6447	81	11	200	6247	70	0.959	0.052	0.136	0.969	0.552	0.075	0.103
0.6	2000	6528	6447	81	7	125	6322	74	0.970	0.053	0.086	0.981	0.534	0.066	0.077
0.7	2000	6528	6447	81	4	60	6387	77	0.979	0.063	0.049	0.991	0.520	0.055	0.052
0.8	2000	6528	6447	81	0	11	6436	81	0.986	0.000	0.000	0.998	0.499		
0.9	2000	6528	6447	81	0	0	6447	81	0.988		0.000	1.000	0.500		
0.1	Full	544	480	64	2	87	393	62	0.726	0.022	0.031	0.819	0.425	0.026	0.029
0.2	Full	544	480	64	1	31	449	63	0.827	0.031	0.016	0.935	0.476	0.021	0.017
0.3	Full	544	480	64	1	24	456	63	0.840	0.040	0.016	0.950	0.483	0.022	0.018
0.4	Full	544	480	64	0	14	466	64	0.857	0.000	0.000	0.971	0.485		
0.5	Full	544	480	64	0	7	473	64	0.869	0.000	0.000	0.985	0.493		
0.6	Full	544	480	64	0	2	478	64	0.879	0.000	0.000	0.996	0.498		
0.7	Full	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
0.8	Full	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
0.9	Full	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		

Table 7.2: Simple Image Analysis results on Flight 3 for different resolutions and confidence levels

Confidence	Resolution	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
0.1	1000	26112	26005	107	51	3614	22391	56	0.859	0.014	0.477	0.861	0.669	0.027	0.062
0.2	1000	26112	26005	107	45	1038	24967	62	0.958	0.042	0.421	0.960	0.690	0.076	0.149
0.3	1000	26112	26005	107	37	1062	24943	70	0.957	0.034	0.346	0.959	0.652	0.061	0.121
0.4	1000	26112	26005	107	23	771	25234	84	0.967	0.029	0.215	0.970	0.593	0.051	0.094
0.5	1000	26112	26005	107	16	516	25489	91	0.977	0.030	0.150	0.980	0.565	0.050	0.083
0.6	1000	26112	26005	107	13	365	25640	94	0.982	0.034	0.121	0.986	0.554	0.054	0.081
0.7	1000	26112	26005	107	7	185	25820	100	0.989	0.036	0.065	0.993	0.529	0.047	0.056
0.8	1000	26112	26005	107	1	38	25967	106	0.994	0.026	0.009	0.999	0.504	0.014	0.011
0.9	1000	26112	26005	107	0	0	26005	107	0.996		0.000	1.000	0.500		
0.1	2000	6528	6453	75	27	1375	5078	48	0.782	0.019	0.360	0.787	0.573	0.037	0.079
0.2	2000	6528	6453	75	23	427	6026	52	0.927	0.051	0.307	0.934	0.620	0.088	0.153
0.3	2000	6528	6453	75	20	352	6101	55	0.938	0.054	0.267	0.945	0.606	0.089	0.149
0.4	2000	6528	6453	75	14	239	6214	61	0.954	0.055	0.187	0.963	0.575	0.085	0.127
0.5	2000	6528	6453	75	11	166	6287	64	0.965	0.062	0.147	0.974	0.560	0.087	0.115
0.6	2000	6528	6453	75	6	105	6348	69	0.973	0.054	0.080	0.984	0.532	0.065	0.073
0.7	2000	6528	6453	75	3	47	6406	72	0.982	0.060	0.040	0.993	0.516	0.048	0.043
0.8	2000	6528	6453	75	2	10	6443	73	0.987	0.167	0.027	0.998	0.513	0.046	0.032
0.9	2000	6528	6453	75	0	0	6453	75	0.989		0.000	1.000	0.500		
0.1	Full	544	493	51	5	104	389	46	0.724	0.046	0.098	0.789	0.444	0.063	0.080
0.2	Full	544	493	51	3	32	461	48	0.853	0.086	0.059	0.935	0.497	0.070	0.063
0.3	Full	544	493	51	2	18	475	49	0.877	0.100	0.039	0.963	0.501	0.056	0.045
0.4	Full	544	493	51	1	12	481	50	0.886	0.077	0.020	0.976	0.498	0.031	0.023
0.5	Full	544	493	51	1	7	486	50	0.895	0.125	0.020	0.986	0.503	0.034	0.024
0.6	Full	544	493	51	0	3	490	51	0.901	0.000	0.000	0.994	0.497		
0.7	Full	544	493	51	0	3	490	51	0.901	0.000	0.000	0.994	0.497		
0.8	Full	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
0.9	Full	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		

Table 7.3: Enhanced Analysis Results on Flight 2 for the three different methods on segments of 2000x2000 resolution and confidence level 0.2

Rotations	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
>0	6528	6447	81	50	1361	5086	31	0.787	0.035	0.617	0.789	0.703	0.067	0.144
>1	6528	6447	81	31	416	6031	50	0.929	0.069	0.383	0.935	0.659	0.117	0.201
>2	6528	6447	81	25	116	6331	56	0.974	0.177	0.309	0.982	0.645	0.225	0.269
>3	6528	6447	81	14	21	6426	67	0.987	0.400	0.173	0.997	0.585	0.241	0.195
Area	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
0	6528	6447	81	27	490	5957	54	0.917	0.052	0.333	0.924	0.629	0.090	0.161
100000	6528	6447	81	27	452	5995	54	0.922	0.056	0.333	0.930	0.632	0.096	0.168
200000	6528	6447	81	26	381	6066	55	0.933	0.064	0.321	0.941	0.631	0.107	0.178
300000	6528	6447	81	25	331	6116	56	0.941	0.070	0.309	0.949	0.629	0.114	0.184
400000	6528	6447	81	24	291	6156	57	0.947	0.076	0.296	0.955	0.626	0.121	0.188
500000	6528	6447	81	20	266	6181	61	0.950	0.070	0.247	0.959	0.603	0.109	0.164
600000	6528	6447	81	19	236	6211	62	0.954	0.075	0.235	0.963	0.599	0.113	0.164
700000	6528	6447	81	18	219	6228	63	0.957	0.076	0.222	0.966	0.594	0.113	0.160
800000	6528	6447	81	16	197	6250	65	0.960	0.075	0.198	0.969	0.583	0.109	0.149
900000	6528	6447	81	16	178	6269	65	0.963	0.082	0.198	0.972	0.585	0.116	0.154
1000000	6528	6447	81	15	166	6281	66	0.964	0.083	0.185	0.974	0.580	0.115	0.149
R+A	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
0	6528	6447	81	25	116	6331	56	0.974	0.177	0.309	0.982	0.645	0.225	0.269
100000	6528	6447	81	19	73	6374	62	0.979	0.207	0.235	0.989	0.612	0.220	0.228
200000	6528	6447	81	18	59	6388	63	0.981	0.234	0.222	0.991	0.607	0.228	0.224
300000	6528	6447	81	18	53	6394	63	0.982	0.254	0.222	0.992	0.607	0.237	0.228
400000	6528	6447	81	18	52	6395	63	0.982	0.257	0.222	0.992	0.607	0.238	0.228
500000	6528	6447	81	17	47	6400	64	0.983	0.266	0.210	0.993	0.601	0.234	0.219
600000	6528	6447	81	16	44	6403	65	0.983	0.267	0.198	0.993	0.595	0.227	0.208
700000	6528	6447	81	15	43	6404	66	0.983	0.259	0.185	0.993	0.589	0.216	0.196
800000	6528	6447	81	14	40	6407	67	0.984	0.259	0.173	0.994	0.583	0.207	0.185
900000	6528	6447	81	14	36	6411	67	0.984	0.280	0.173	0.994	0.584	0.214	0.187
1000000	6528	6447	81	14	34	6413	67	0.985	0.292	0.173	0.995	0.584	0.217	0.188

Table 7.4: Enhanced Analysis Results on Flight 3 for the three different methods on segments of 2000x2000 resolution and confidence level 0.2

Rotations	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
>0	6528	6453	75	35	1454	4999	40	0.771	0.024	0.467	0.775	0.621	0.045	0.098
>1	6528	6453	75	22	471	5982	53	0.920	0.045	0.293	0.927	0.610	0.077	0.139
>2	6528	6453	75	18	120	6333	57	0.973	0.130	0.240	0.981	0.611	0.169	0.205
>3	6528	6453	75	12	17	6436	63	0.988	0.414	0.160	0.997	0.579	0.231	0.182
Area	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
0	6528	6453	75	23	427	6026	52	0.927	0.051	0.307	0.934	0.620	0.088	0.153
100000	6528	6453	75	21	351	6102	54	0.938	0.056	0.280	0.946	0.613	0.094	0.156
200000	6528	6453	75	21	304	6149	54	0.945	0.065	0.280	0.953	0.616	0.105	0.168
300000	6528	6453	75	21	266	6187	54	0.951	0.073	0.280	0.959	0.619	0.116	0.179
400000	6528	6453	75	20	230	6223	55	0.956	0.080	0.267	0.964	0.616	0.123	0.182
500000	6528	6453	75	20	204	6249	55	0.960	0.089	0.267	0.968	0.618	0.134	0.191
600000	6528	6453	75	17	175	6278	58	0.964	0.089	0.227	0.973	0.600	0.127	0.173
700000	6528	6453	75	16	163	6290	59	0.966	0.089	0.213	0.975	0.594	0.126	0.167
800000	6528	6453	75	14	143	6310	61	0.969	0.089	0.187	0.978	0.582	0.121	0.153
900000	6528	6453	75	12	135	6318	63	0.970	0.082	0.160	0.979	0.570	0.108	0.134
1000000	6528	6453	75	11	128	6325	64	0.971	0.079	0.147	0.980	0.563	0.103	0.125
R+A	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
0	6528	6453	75	18	120	6333	57	0.973	0.130	0.240	0.981	0.611	0.169	0.205
100000	6528	6453	75	17	63	6390	58	0.981	0.213	0.227	0.990	0.608	0.219	0.224
200000	6528	6453	75	16	52	6401	59	0.983	0.235	0.213	0.992	0.603	0.224	0.217
300000	6528	6453	75	13	45	6408	62	0.984	0.224	0.173	0.993	0.583	0.195	0.182
400000	6528	6453	75	13	39	6414	62	0.985	0.250	0.173	0.994	0.584	0.205	0.185
500000	6528	6453	75	13	36	6417	62	0.985	0.265	0.173	0.994	0.584	0.210	0.186
600000	6528	6453	75	11	29	6424	64	0.986	0.275	0.147	0.996	0.571	0.191	0.162
700000	6528	6453	75	10	27	6426	65	0.986	0.270	0.133	0.996	0.565	0.179	0.148
800000	6528	6453	75	8	21	6432	67	0.987	0.276	0.107	0.997	0.552	0.154	0.122
900000	6528	6453	75	8	18	6435	67	0.987	0.308	0.107	0.997	0.552	0.158	0.123
1000000	6528	6453	75	7	18	6435	68	0.987	0.280	0.093	0.997	0.545	0.140	0.108

Table 7.5: Tuned Models Performance on Flight 2 at resolution 2000x2000 for varying confidence levels

Original	Confidence	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
	0.1	6528	6447	81	30	1402	5045	51	0.777	0.021	0.370	0.783	0.576	0.040	0.085
	0.2	6528	6447	81	27	490	5957	54	0.917	0.052	0.333	0.924	0.629	0.090	0.161
	0.3	6528	6447	81	24	403	6044	57	0.930	0.056	0.296	0.937	0.617	0.094	0.160
	0.4	6528	6447	81	18	284	6163	63	0.947	0.060	0.222	0.956	0.589	0.094	0.144
	0.5	6528	6447	81	11	200	6247	70	0.959	0.052	0.136	0.969	0.552	0.075	0.103
	0.6	6528	6447	81	7	125	6322	74	0.970	0.053	0.086	0.981	0.534	0.066	0.077
	0.7	6528	6447	81	4	60	6387	77	0.979	0.063	0.049	0.991	0.520	0.055	0.052
	0.8	6528	6447	81	0	11	6436	81	0.986	0.000	0.000	0.998	0.499		
	0.9	6528	6447	81	0	0	6447	81	0.988		0.000	1.000	0.500		
T1	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	0.1	6528	6447	81	29	37	6410	52	0.986	0.439	0.358	0.994	0.676	0.395	0.372
	0.2	6528	6447	81	27	12	6435	54	0.990	0.692	0.333	0.998	0.666	0.450	0.372
	0.3	6528	6447	81	23	10	6437	58	0.990	0.697	0.284	0.998	0.641	0.404	0.322
	0.4	6528	6447	81	18	6	6441	63	0.989	0.750	0.222	0.999	0.611	0.343	0.259
	0.5	6528	6447	81	13	4	6443	68	0.989	0.765	0.160	0.999	0.580	0.265	0.191
	0.6	6528	6447	81	10	1	6446	71	0.989	0.909	0.123	1.000	0.562	0.217	0.149
	0.7	6528	6447	81	7	0	6447	74	0.989	1.000	0.086	1.000	0.543	0.159	0.106
	0.8	6528	6447	81	3	0	6447	78	0.988	1.000	0.037	1.000	0.519	0.071	0.046
	0.9	6528	6447	81	0	0	6447	81	0.988		0.000	1.000	0.500		
T2	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	0.1	6528	6447	81	40	23	6424	41	0.990	0.635	0.494	0.996	0.745	0.556	0.517
	0.2	6528	6447	81	39	7	6440	42	0.992	0.848	0.481	0.999	0.740	0.614	0.527
	0.3	6528	6447	81	33	5	6442	48	0.992	0.868	0.407	0.999	0.703	0.555	0.456
	0.4	6528	6447	81	25	3	6444	56	0.991	0.893	0.309	1.000	0.654	0.459	0.355
	0.5	6528	6447	81	18	2	6445	63	0.990	0.900	0.222	1.000	0.611	0.356	0.262
	0.6	6528	6447	81	12	0	6447	69	0.989	1.000	0.148	1.000	0.574	0.258	0.179
	0.7	6528	6447	81	9	0	6447	72	0.989	1.000	0.111	1.000	0.556	0.200	0.135
	0.8	6528	6447	81	4	0	6447	77	0.988	1.000	0.049	1.000	0.525	0.094	0.061
	0.9	6528	6447	81	0	0	6447	81	0.988		0.000	1.000	0.500		
T2 + R	**Nr. Rotations**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	>0	6528	6447	81	54	67	6380	27	0.986	0.446	0.667	0.990	0.828	0.535	0.607
	>1	6528	6447	81	46	12	6435	35	0.993	0.793	0.568	0.998	0.783	0.662	0.602
	>2	6528	6447	81	32	1	6446	49	0.992	0.970	0.395	1.000	0.697	0.561	0.448
	>3	6528	6447	81	18	0	6447	63	0.990	1.000	0.222	1.000	0.611	0.364	0.263

Table 7.6: Tuned Models Performance on Flight 3 at resolution 2000x2000 for varying confidence levels

Original	Confidence	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
	0.1	6528	6453	75	27	1375	5078	48	0.782	0.019	0.360	0.787	0.573	0.037	0.079
	0.2	6528	6453	75	23	427	6026	52	0.927	0.051	0.307	0.934	0.620	0.088	0.153
	0.3	6528	6453	75	20	352	6101	55	0.938	0.054	0.267	0.945	0.606	0.089	0.149
	0.4	6528	6453	75	14	239	6214	61	0.954	0.055	0.187	0.963	0.575	0.085	0.127
	0.5	6528	6453	75	11	166	6287	64	0.965	0.062	0.147	0.974	0.560	0.087	0.115
	0.6	6528	6453	75	6	105	6348	69	0.973	0.054	0.080	0.984	0.532	0.065	0.073
	0.7	6528	6453	75	3	47	6406	72	0.982	0.060	0.040	0.993	0.516	0.048	0.043
	0.8	6528	6453	75	2	10	6443	73	0.987	0.167	0.027	0.998	0.513	0.046	0.032
	0.9	6528	6453	75	0	0	6453	75	0.989		0.000	1.000	0.500		
T1	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	0.1	6528	6453	75	34	35	6418	41	0.988	0.493	0.453	0.995	0.724	0.472	0.461
	0.2	6528	6453	75	32	9	6444	43	0.992	0.780	0.427	0.999	0.713	0.552	0.469
	0.3	6528	6453	75	30	5	6448	45	0.992	0.857	0.400	0.999	0.700	0.545	0.448
	0.4	6528	6453	75	28	3	6450	47	0.992	0.903	0.373	1.000	0.686	0.528	0.423
	0.5	6528	6453	75	23	0	6453	52	0.992	1.000	0.307	1.000	0.653	0.469	0.356
	0.6	6528	6453	75	17	0	6453	58	0.991	1.000	0.227	1.000	0.613	0.370	0.268
	0.7	6528	6453	75	9	0	6453	66	0.990	1.000	0.120	1.000	0.560	0.214	0.146
	0.8	6528	6453	75	3	0	6453	72	0.989	1.000	0.040	1.000	0.520	0.077	0.050
	0.9	6528	6453	75	0	0	6453	75	0.989		0.000	1.000	0.500		
T2	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	0.1	6528	6453	75	42	23	6430	33	0.991	0.646	0.560	0.996	0.778	0.600	0.575
	0.2	6528	6453	75	40	5	6448	35	0.994	0.889	0.533	0.999	0.766	0.667	0.580
	0.3	6528	6453	75	38	2	6451	37	0.994	0.950	0.507	1.000	0.753	0.661	0.559
	0.4	6528	6453	75	32	0	6453	43	0.993	1.000	0.427	1.000	0.713	0.598	0.482
	0.5	6528	6453	75	20	0	6453	55	0.992	1.000	0.267	1.000	0.633	0.421	0.313
	0.6	6528	6453	75	13	0	6453	62	0.991	1.000	0.173	1.000	0.587	0.295	0.208
	0.7	6528	6453	75	10	0	6453	65	0.990	1.000	0.133	1.000	0.567	0.235	0.161
	0.8	6528	6453	75	3	0	6453	72	0.989	1.000	0.040	1.000	0.520	0.077	0.050
	0.9	6528	6453	75	0	0	6453	75	0.989		0.000	1.000	0.500		
T2 + R	**Nr. Rotations**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	>0	6528	6453	75	48	43	6410	27	0.989	0.527	0.640	0.993	0.817	0.578	0.614
	>1	6528	6453	75	37	5	6448	38	0.993	0.881	0.493	0.999	0.746	0.632	0.541
	>2	6528	6453	75	31	0	6453	44	0.993	1.000	0.413	1.000	0.707	0.585	0.468
	>3	6528	6453	75	25	0	6453	50	0.992	1.000	0.333	1.000	0.667	0.500	0.385

Table 7.7: Tuned Models Performance on Flight 2 at full resolution for varying confidence levels

Original	Confidence	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
	> 0.1	544	480	64	2	87	393	62	0.726	0.022	0.031	0.819	0.425	0.026	0.029
	> 0.2	544	480	64	1	31	449	63	0.827	0.031	0.016	0.935	0.476	0.021	0.017
	> 0.3	544	480	64	1	24	456	63	0.840	0.040	0.016	0.950	0.483	0.022	0.018
	> 0.4	544	480	64	0	14	466	64	0.857	0.000	0.000	0.971	0.485		
	> 0.5	544	480	64	0	7	473	64	0.869	0.000	0.000	0.985	0.493		
	> 0.6	544	480	64	0	2	478	64	0.879	0.000	0.000	0.996	0.498		
	> 0.7	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.8	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.9	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
T1	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	> 0.1	544	480	64	6	15	465	58	0.866	0.286	0.094	0.969	0.531	0.141	0.108
	> 0.2	544	480	64	6	5	475	58	0.884	0.545	0.094	0.990	0.542	0.160	0.112
	> 0.3	544	480	64	4	5	475	60	0.881	0.444	0.063	0.990	0.526	0.110	0.075
	> 0.4	544	480	64	2	1	479	62	0.884	0.667	0.031	0.998	0.515	0.060	0.039
	> 0.5	544	480	64	1	0	480	63	0.884	1.000	0.016	1.000	0.508	0.031	0.019
	> 0.6	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.7	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.8	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.9	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
T2	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	> 0.1	544	480	64	8	5	475	56	0.888	0.615	0.125	0.990	0.557	0.208	0.149
	> 0.2	544	480	64	8	0	480	56	0.897	1.000	0.125	1.000	0.563	0.222	0.152
	> 0.3	544	480	64	6	0	480	58	0.893	1.000	0.094	1.000	0.547	0.171	0.115
	> 0.4	544	480	64	5	0	480	59	0.892	1.000	0.078	1.000	0.539	0.145	0.096
	> 0.5	544	480	64	2	0	480	62	0.886	1.000	0.031	1.000	0.516	0.061	0.039
	> 0.6	544	480	64	2	0	480	62	0.886	1.000	0.031	1.000	0.516	0.061	0.039
	> 0.7	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.8	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		
	> 0.9	544	480	64	0	0	480	64	0.882		0.000	1.000	0.500		

Table 7.8: Tuned Models Performance on Flight 3 at full resolution for varying confidence levels

Original	Confidence	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
	> 0.1	544	493	51	5	104	389	46	0.724	0.046	0.098	0.789	0.444	0.063	0.080
	> 0.2	544	493	51	3	32	461	48	0.853	0.086	0.059	0.935	0.497	0.070	0.063
	> 0.3	544	493	51	2	18	475	49	0.877	0.100	0.039	0.963	0.501	0.056	0.045
	> 0.4	544	493	51	1	12	481	50	0.886	0.077	0.020	0.976	0.498	0.031	0.023
	> 0.5	544	493	51	1	7	486	50	0.895	0.125	0.020	0.986	0.503	0.034	0.024
	> 0.6	544	493	51	0	3	490	51	0.901	0.000	0.000	0.994	0.497		
	> 0.7	544	493	51	0	3	490	51	0.901	0.000	0.000	0.994	0.497		
	> 0.8	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
	> 0.9	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
T1	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	> 0.1	544	493	51	8	9	484	43	0.904	0.471	0.157	0.982	0.569	0.235	0.181
	> 0.2	544	493	51	7	2	491	44	0.915	0.778	0.137	0.996	0.567	0.233	0.164
	> 0.3	544	493	51	7	2	491	44	0.915	0.778	0.137	0.996	0.567	0.233	0.164
	> 0.4	544	493	51	4	1	492	47	0.912	0.800	0.078	0.998	0.538	0.143	0.096
	> 0.5	544	493	51	3	0	493	48	0.912	1.000	0.059	1.000	0.529	0.111	0.072
	> 0.6	544	493	51	2	0	493	49	0.910	1.000	0.039	1.000	0.520	0.075	0.049
	> 0.7	544	493	51	1	0	493	50	0.908	1.000	0.020	1.000	0.510	0.038	0.024
	> 0.8	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
	> 0.9	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
T2	**Confidence**	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
	> 0.1	544	493	51	17	3	490	34	0.932	0.850	0.333	0.994	0.664	0.479	0.379
	> 0.2	544	493	51	16	0	493	35	0.936	1.000	0.314	1.000	0.657	0.478	0.364
	> 0.3	544	493	51	15	0	493	36	0.934	1.000	0.294	1.000	0.647	0.455	0.342
	> 0.4	544	493	51	11	0	493	40	0.926	1.000	0.216	1.000	0.608	0.355	0.256
	> 0.5	544	493	51	10	0	493	41	0.925	1.000	0.196	1.000	0.598	0.328	0.234
	> 0.6	544	493	51	3	0	493	48	0.912	1.000	0.059	1.000	0.529	0.111	0.072
	> 0.7	544	493	51	2	0	493	49	0.910	1.000	0.039	1.000	0.520	0.075	0.049
	> 0.8	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		
	> 0.9	544	493	51	0	0	493	51	0.906		0.000	1.000	0.500		

Figure 7.9: Performance of the Original YOLOv8 model (above), Tuned Model (middle) and their difference (below) in the evaluation metrics on the external dataset where red indicated worst and green better performances

Original	Analized	No Smoke	Smoke	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	BA	F1	F2
> 0.1	5313	430	4883	4812	48	382	71	0.978	0.990	0.985	0.888	0.937	0.988	0.986
> 0.2	5313	430	4883	4787	33	397	96	0.976	0.993	0.980	0.923	0.952	0.987	0.983
> 0.3	5313	430	4883	4746	21	409	137	0.970	0.996	0.972	0.951	0.962	0.984	0.977
> 0.4	5313	430	4883	4650	12	418	233	0.954	0.997	0.952	0.972	0.962	0.974	0.961
> 0.5	5313	430	4883	4504	6	424	379	0.928	0.999	0.922	0.986	0.954	0.959	0.937
> 0.6	5313	430	4883	4336	2	428	547	0.897	1.000	0.888	0.995	0.942	0.940	0.908
> 0.7	5313	430	4883	3857	0	430	1026	0.807	1.000	0.790	1.000	0.895	0.883	0.825
> 0.8	5313	430	4883	2624	0	430	2259	0.575	1.000	0.537	1.000	0.769	0.699	0.592
> 0.9	5313	430	4883	185	0	430	4698	0.116	1.000	0.038	1.000	0.519	0.073	0.047
Tuned	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
> 0.1	5313	430	4883	4798	142	288	85	0.957	0.971	0.983	0.670	0.826	0.977	0.980
> 0.2	5313	430	4883	4746	62	368	137	0.963	0.987	0.972	0.856	0.914	0.979	0.975
> 0.3	5313	430	4883	4687	29	401	196	0.958	0.994	0.960	0.933	0.946	0.977	0.966
> 0.4	5313	430	4883	4623	7	423	260	0.950	0.998	0.947	0.984	0.965	0.972	0.957
> 0.5	5313	430	4883	4478	2	428	405	0.923	1.000	0.917	0.995	0.956	0.957	0.932
> 0.6	5313	430	4883	4194	1	429	689	0.870	1.000	0.859	0.998	0.928	0.924	0.884
> 0.7	5313	430	4883	3496	0	430	1387	0.739	1.000	0.716	1.000	0.858	0.834	0.759
> 0.8	5313	430	4883	2242	0	430	2641	0.503	1.000	0.459	1.000	0.730	0.629	0.515
> 0.9	5313	430	4883	757	0	430	4126	0.223	1.000	0.155	1.000	0.578	0.268	0.187
Δ	**Analized**	**No Smoke**	**Smoke**	**TP**	**FP**	**TN**	**FN**	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**BA**	**F1**	**F2**
> 0.1	0	0	0	-14	94	-94	14	-0.020	-0.019	-0.003	-0.219	-0.111	-0.011	-0.006
> 0.2	0	0	0	-41	29	-29	41	-0.013	-0.006	-0.008	-0.067	-0.038	-0.007	-0.008
> 0.3	0	0	0	-59	8	-8	59	-0.013	-0.002	-0.012	-0.019	-0.015	-0.007	-0.010
> 0.4	0	0	0	-27	-5	5	27	-0.004	0.001	-0.006	0.012	0.003	-0.002	-0.004
> 0.5	0	0	0	-26	-4	4	26	-0.004	0.001	-0.005	0.009	0.002	-0.002	-0.004
> 0.6	0	0	0	-142	-1	1	142	-0.027	0.000	-0.029	0.002	-0.013	-0.016	-0.024
> 0.7	0	0	0	-361	0	0	361	-0.068	0.000	-0.074	0.000	-0.037	-0.048	-0.065
> 0.8	0	0	0	-382	0	0	382	-0.072	0.000	-0.078	0.000	-0.039	-0.070	-0.077
> 0.9	0	0	0	572	0	0	-572	0.108	0.000	0.117	0.000	0.059	0.195	0.140

The files used and elaborated during this research include:

- The raw JPG images obtained during the execution of Flight 2 and Flight 3 in the Büelenwald region.

- A portion of the Python codes used to execute the various processes implemented throughout the research.

- The original Pytorch file for YOLOv8 developed in the study by Tleuliyev (2019).

- The Pytorch files for YOLOv8 developed during the various fine-tuning phases conducted in this study.

These files are available at the following repository link:
https://drive.google.com/drive/folders/1lCLej7aYVJOIleax2upmywNHU2g8Szhb?usp=drive_link

Form for the disclosure of AI

☒ I have read the information sheet on the Master's thesis.

☒ Yes, I have used AI for this work, namely: Deepl, ChatGPT-4

I have used AI for the following purposes:

- I used Deepl for writing and translating the text from Italian to English in order to achieve better speed and spontaneity in drafting the work.

- I used ChatGPT-4 as a tool for the development, refinement, and adaptation of the Python codes used for the various purposes pursued in the thesis.

- I used ChatGPT-4 as a tool for the revision of grammatical errors or typographical mistakes within the text.

With my signature I confirm that the above information is true.

Date

25 April, 2025

Signature

Luca Giorgio Capitanio

Personal Declaration

I hereby declare that the material contained in this thesis is my own original work. Any quotation or paraphrase in this thesis from the published or unpublished work of another individual or institution has been duly acknowledged. I have not submitted this thesis, or any part of it, previously to any institution for assessment purposes. Furthermore, I confirm not to have used any AI tool for this work other the ones stated in the appendix.

Date

April 25, 2025

Location

Zürich

Signature

Luca Giorgio Capitanio