IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 1, December 2025

Impact Factor: 7.67

Drone-Based AI System for Wildfire Monitoring and Risk Prediction

Rakh Om[1], Patil Yash[2], Padole Omkar[3], Prof. P. V. Gaikwad[4]

Student, Department of Computer Engineering[123]

Professor, Dept, of Computer Engineering[4]

Adsul Technical Campus, Chas, Ahilyanagar, Maharashtra, India

Abstract: *Wildfires pose a significant threat to ecosystems, human lives, and infrastructure worldwide. Traditional wildfire detection and risk assessment methods often suffer from limitations such as delayed detection and low confidence in certain regions. In this paper, we propose a novel computational system based on Machine Learning for wildfire risk assessment using data collected by drones. The system can integrate various sensors to capture spatiotemporal data on environmental factors such as temperature, humidity, and vegetation. By leveraging high-resolution data collected through autonomous drone missions, our system enhances wildfire risk estimation and enables proactive mission planning. Although the system is mainly designed to address wildfire monitoring using drone-collected data, it can be easily adapted to other environmental monitoring applications and other sources of data. We demonstrate the effectiveness of our approach through a comprehensive evaluation and validation process in both simulated and real world environments. Our work contributes to advancing wildfire monitoring capabilities, improving early detection, and mitigating the impact of wildfires on communities and the environment*

Keywords: Aerial drones, artificial intelligence, environmental monitoring, machine learning, risk assessment, spatiotemporal data, wildfire detection, wildfire risk estimation

I. INTRODUCTION

Wildfires are one of the most devastating natural disasters, causing significant economic losses, ecological damage, and severe health effects. Throughout 2023, the Copernicus Atmosphere Monitoring Service (CAMS1) has monitored wildfires worldwide. Using its Global Fire Assimilation System (GFAS), CAMS provided accurate data on wildfire intensity andcarbonemissions[1].Manyregionsexperienced record-breaking wildfire activity that year. For example, Canada saw the highest wildfire carbon emissions in the CAMS record, which dates back to 2003. Greece also experienced the largest wildfire in the European Union to date. According to CAMS estimates, in 2023, global wildfires generated approximately 2,170 megatons of carbon emissions. According to a report by the UN Environment Programme (UNEP2) and GRID-Arendal,3 changes in climate and land make wildfires more frequent and intense. They expect a global increase of extreme fires of up to 14% by 2030, 30% by the end of 2050, and 50% by the end of the century. The report, Spreading like Wildfire: The Rising Threat of Extraordinary Landscape Fires [2], finds an elevated risk in previously unaffected regions, such as the Arctic. Impacts of wildfires are not limited to the environment. Theyalsohavesignificanteconomic,ecological,andphysical and mental health effects [3]. These impacts extend for days, weeks, and even years after the flames subside. In Brazil, the Pantanal region has been particularly affected by wildfires. With fast climate changes, the region has been experiencing a significant increase in temperature [4], [5]. Also a decrease in precipitation leads to a decrease in soil moisture and an increase in the number of days without precipitation [6], [7]. This situation has led to a drop in the water mass in the soil during the drought season, which has been increasing the risk of wildfires. Consequently, in 2020, wildfire occurrence has reached record highs of over 40,000 km2. The extent of area affected represents a 376% increase compared to the annual average of the area burnt in the last two decades, doubling the value achieved in 2019 [8]. The extent of wildfires in that region has been linked to the occurrence of heatwaves [9] that are rising together with extreme hot conditions [4]. Future climate

DOI: 10.48175/IJARSCT-30114

ISSN 2581-9429 IJARSCT

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

changes are expected to provide even more favourable conditions for wildfires to occur [10], [11]. At the sametime, advances in agricultural use of the Pantanal region boosted an increase in the practice of burning areas to prepare the soil for later use [5]. A prediction scenario for 2030 shows that if measures for environmental protection and combating f ires are not adopted, this year, 20% of the Pantanal areas will be for agricultural and pasture use [5]. Despite the problems caused by wildfires in Brazil, initiatives are being implemented to foster more responsible and sustainable fire management and control. For instance, the action plan for integrated fire management in the Pantanal biome [12] aims to improve fire management's efficiency and economic viability. Such initiatives aim to integrate traditional, local and ecological knowledge and existing scientific knowledge to develop a collective strategy within the socio-ecological context. Communities are encouraged to collaborate in creating a cooperative network for forest f ire prevention, prioritizing areas for financial investment, management, and protection [8]. These initiatives are aligned with international thinking and efforts, as highlighted in the UNEP and GRID-Arendal report [2], which calls for a significant shift in government spending on wildfires—from reaction and response towards prevention and preparedness. There is also an urgent need for an improved understanding of the benefits of wildfire prevention that can be achieved through enhancedmonitoring activities. Currently, Brazil also has the INPE's Wildfire Program Portal,4 which aggregates data on precipitation, temperature, humidity, topography, and vegetation to calculate fire risk. However, due to the system's wide coverage and data collection granularity, this approach tends to be susceptible to late fire detection and low confidence in certain regions. While high-resolution fixed sensors are limited to small areas and lack mobility flexibility, an alternative approach and possible solution involves using autonomous drone missions to collect high-resolution data in regions of interest. The choice of the region of interest can be based on the system's risk estimation or other criteria, such as few historical data or low confidence in the risk estimation for that particular area can define it. This approach can be the building block of a very proactive prevention and monitoring system, capable of, based on its predictions, deciding the best locations for new missions to take a better look at high-risk regions or to collect more data in regions previously undersampled. A. OBJECTIVE Our objective is to develop a computational system based on Machine Learning (ML) for calculating wildfire risk, including a system for capturing, managing, annotating, and fusing data from the various sensors embedded in Unmanned Aerial Vehicles (UAVs), particularly drones. The system will enable both wildfire risk estimation and mission planning management for data collection via UAVs in areas of interest. This approach aligns with the three stages of wildfire man agement: pre-fire, active-fire, and post-fire tasks. According to Boroujeni et al. [13], these three stages are described as: 1) Pre-fire: activities such as vision-based vegetation fuel measurement to assess and mitigate fire risk before it occurs; 2) Active-fire: tasks like fire growth modeling to predict and manage the behavior of an ongoing fire; and 3) Post-fire: efforts including evacuation planning and assessing the aftermath of the fire. By focusing on wildfire risk estimation and mission planning, our system primarily addresses the pre-fire stage through predictive modeling and data collection. However, it also supports the active-fire stage by providing real-time data to assist in fire growth modeling and response efforts. The specific objectives are: • Develop technology for wildfire prevention and pattern recognition and missions related to wildfire combat; and • Develop a flexible data management technology for civilian or military applications, facilitating using Arti f icial Intelligence (AI) in scenarios not covered in this proposal. Achieving these objectives will result in significant advancementsinwildfireriskmanagementandresponse.The following contributions highlight the key outcomes of this initiative: B. CONTRIBUTIONS • Development of a novel computational system based on ML for wildfire risk calculation, integrating data from various sensors embedded in UAVs; • Creation of a comprehensive system for capturing, man aging, annotating, and fusing sensor data, facilitating efficient wildfire risk estimation and mission planning; and • Advancement of environmental monitoring capabilities, enhancing early wildfire detection and response strate gies. This paper introduces a novel system that advances the state-of-the-art by seamlessly integrating multisensor data from UAVs with sophisticated AI algorithms for enhanced wildfire detection and risk assessment. Specifically, our system addresses the limitations of existing approaches by offering a robust framework for data fusion, employing advanced ML models, and providing an intuitive user interface for real-time decision support.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

II. BACKGROUND AND RELATED WORKS

Increasing in frequency, extent, and magnitude of damage, wildfiresrepresent oneofthemostprevalentnaturaldisasters. Projections antecipate a continuing escalation of this trend, especially in response to anthropogenic activity [14]. The capacity to enhance wildfire detection systems and calculate fire risk autonomously and automatically with the aid of UAVs, and potentially manned aircraft, meets the demand of the Ministry of Defense, led by the Air Force. For example, the Aeronautics' Science, Technology, and Innovation Plan PCA 11-217/2018, foresees the study and development of ''Unmanned Aerial Vehicles for defense missions and for civil defense operations, public security, and environmental monitoring''. This project proposes developing a system based on Artificial Intelligence (AI) and Machine Learning (ML) to model fire risk using data collected by UAVs. The Unmanned Aircraft System (UAS) was designed to fully managethesystemsandoperationsofUnmannedAerial Vehicles (UAVs). Recent advancements in AI and UAS have significantly contributed to wildfire management. This section discusses the state-of-the-art research in the context of pre-fire, active f ire, and post-fire management stages.

A. WILDFIRE PREDICTION MODELS

Understanding and predicting wildfire risk is a multifaceted problem that has attracted significant research attention. Several studies have focused on various aspects of wildfire risk prediction, utilizing different methodologies and data types. Traditional statistical models and ML techniques have been extensively used in wildfire prediction. Boroujeni et al. [13] provide a comprehensive survey on the statistical and MLmodels used for wildfire prediction. Their work empha sizes the need for robust models capable of handling complex interactions between variables. This survey is a foundational background for understanding wildfire prediction models' progression and current state and is recommended for further reading. Rubí and Gondim [15] compare various ML models for predicting wildfire occurrence risk in the Brazilian Federal District region. They found that fire risk can be predicted with an accuracy of 0.99. The study also highlighted that the models were most sensitive to Normalized Difference Vegetation Index (NDVI), atmospheric pressure, and relative humidity. By incorporating a comprehensive dataset enriched with Brazilian governmental open data, they achieve high prediction accuracy, demonstrating the potential of Machine Learning in wildfire risk assessment. Kabir et al. [16] focus on wildfire prediction in the United States using time series forecasting models. They employ the Neural Basis Expansion Analysis for Time Series (N-BEATS) model to forecast the total area burned by wildfires weekly and monthly. Their model demonstrates improved performance over other state-of-the-art models, showcasing the potential of advanced time series forecasting techniques in wildfire prediction. Our proposed system builds on recent advancements by focusing on Machine Learning-based wildfire risk calcu lation. It integrates data from multiple sensors on drones, including a variation of the Visible band of NDVI (vNDVI) [17], humidity, and temperature. The vNDVI was selected based on research by Costa et al. [17], which demonstrated that this variation of the Visible band of NDVI provides a rapid and low-cost tool for assessing vegetation information. Using low-cost RGB cameras and a genetic algorithm, this method estimates NDVI values from RGB maps with an overall Mean Percentage Error (MPE) of 6.89 percent and a Mean Absolute Error (MAE) of 0.052. Moreover, our system is prepared to incorporate any other type of information that can be collected from the UAV sensors. This integration facilitates efficient mission planning and real-time data analysis, enhancing wildfire risk estimation, improving early detection, and supporting comprehensive wildfire management strategies.

B. INDEPENDENT VARIABLES AND DATA PROCESSING

Xu et al. [18] in their review article discuss the wide range of independent variables that are crucial for wildfire prediction. These variables include climate and meteo rological conditions, socio-economic factors, terrain and hydrological features, and historical wildfire records. The authors highlight various data processing techniques to handle these variables' diverse formats and resolutions. They also emphasize the importance of addressing collinearity and evaluating the significance of different variables. This review provides valuable insights into the preprocessing and significance evaluation of independent variables, making it an essential reference for further understanding and development. Rubí and Gondim [15] emphasize the role of variables like NDVI, atmospheric pressure, and relative humidity in predicting wildfire risk. They demonstrate how the careful selection and processing of these variables can lead to high prediction accuracy in ML models. NDVI and vNDVI are crucial tools in wildfire prediction research using drones due to their ability to monitor vegetation health accurately [17]. NDVImeasures the difference between near-infrared and red light reflection, with higher values indicating healthier vege tation [19]. vNDVI, incorporating visible and near-infrared spectral bands, enhances sensitivity to stressed or sparse vegetation. Even, this makes it more sensitive to changes in visible reflectance caused by fire [17]. These indices enable the early detection of fire-prone areas by identifying dry or stressed vegetation susceptible to wildfires [20]. Integrating vegetation indices data with ML algorithms can identify patterns and correlations with historical wildfire occurrences, enhancing early warning systems and proactive measures. The spatiotemporal analysis capabilities of these veg etation indices allow for tracking changes in vegetation health over time and across regions, aiding in predicting seasonal wildfire risks and assessing the impact of climate change [21]. Using drones for data collection is cost-effective and flexible compared to traditional methods, enabling more frequent and widespread monitoring. The adaptability of suggested vegetation indices to various environments ensures their applicability in diverse regions, enhancing the generalizability of wildfire prediction models. Additionally, integrating NDVI and vNDVI data with other environmental data, such as temperature and humidity, enriches datasets and improves the predictive power of Machine Learn ing models, leading to more comprehensive wildfire risk assessments. Furthermore, [22] modeled the probability of wildfires in the Mediterranean basin in Spain, using as predictive variables the housing density in the region, distance to populated places, land use, distance and density of roads, and the interface between different land uses (for instance, forest agriculture). Reference [23] also studied the sensitivity of wildfiretogeographic,climatic, andanthropomorphicfactors and found that precipitation levels and terrain inclinations are essential variables for the phenomenon. Accordingly, [24] found that wildfires in forests tend to increase in occurrence with more extensive terrain inclination, except for inclinations larger than 25 degrees. Our system aligns with these recommendations by incor porating various independent variables collected by UAVs. It processes these variables to ensure high data quality and relevance, addressing collinearity and evaluating their importance to improve the prediction accuracy.

C. DEEP LEARNING ADVANCEMENTS

Recent advancementsindeeplearning haveshownpromising results in wildfire risk prediction. Xu et al. [18] mainly discuss the application of deep learning models and their performance evaluation metrics. They highlight the need for more effective time series forecasting algorithms and the utilization of three-dimensional data, including ground and trunk fuel. This comprehensive review of deep learning tech niques and evaluation metrics is essential for understanding these models' current capabilities and limitations. Shamta and Demir [25] focus on developing a deep learning-based surveillance system for early forest fire detection using UAVs. They examine the performance of 139868 various deep learning models for identifying forest fires, including YOLOv8 and YOLOv5. Their system employs NVIDIA Jetson Nano for real-time detection and includes a ground station interface for data display and intervention planning. Their results show high accuracy in detecting forest fires, demonstrating the effectiveness of deep learning algorithms in enhancing UAV-based wildfire surveillance. Kabir et al. [16] demonstrate using the N-BEATS model for time series forecasting of wildfire occurrences in the United States, showing its superiority over other state-of-the art models. Ourproposedsolutionincorporatesadvanceddeep-learning models to enhance the accuracy of wildfire risk prediction. We also use YOLO to detect stationary and moving objects, like persons or cars. By leveraging state-of-the-art algorithms and performance metrics, our system aims to improve the detection and classification of wildfires, providing timely and reliable information for intervention planning.

D. MODEL EVALUATION AND PERFORMANCE METRICS

Evaluating the performance of wildfire prediction models is critical. Xu et al. [18] address the limitations of current eval uation metrics and suggest improvements. They recommend extracting more accurate historical firepoint data and using enhanced model evaluation metrics to improve prediction accuracy. The focus on evaluation metrics is crucial for developing more reliable and effective wildfire prediction models. Rubí and Gondim [15] highlight the importance of feature sensitivity analysis in improving model accuracy. Their study demonstrates how the careful evaluation of model performance can lead to significant improvements in prediction accuracy. Our system incorporates advanced evaluation metrics to ensure the reliability and accuracy of wildfire risk predictions. We aim to provide more accurate and actionable information for wildfire management by utilizing enhanced model evaluation techniques.

E. INTEGRATION AND ANALYSIS

The integration of diverse data sources and the analysis of their combined effect on wildfire prediction remains challenging. Xu et al. [18] provide insights into preprocessing methods and the importance of considering data from differ ent magnitudes and resolutions. Their review emphasizes the need for effective data integration and analysis techniques to improve wildfire prediction accuracy. Kanand et al. [26] discuss the integration of UAS, multiple sensor technologies, and 5G mobile networks for wildfire detection and disaster monitoring. Their Saxony, Germany project utilizes VIS and Thermal IR cameras to detect smoke or hotspots, combining multi-observation data for automated analysis and localization. This approach showcases the potential of sensor fusion technologies and advanced network capabilities in improving the accuracy and efficiency of wildfire detection systems Rashid et al. [27] present CompDrone, a framework that combines computational wildfire modeling with social media-driven drone sensing (SDS) to monitor wildfire propagation. CompDrone addresses challenges such as the limited availability of social signals in remote regions and the need for effective drone dispatch strategies. Their evaluation demonstrates that CompDrone outperforms existing methods in predicting wildfire spread, highlighting the potential of integrating computational models and SDS for reliable wildfire monitoring. Our proposed system addresses these challenges by integrating data from multiple sensors present on UAVs and facilitating efficient mission planning and real-time data analysis.

F. OVERVIEW

These works collectively demonstrate the potential of lever aging AI, UAVs, and sensor fusion technologies in various stages of wildfire management. Our system aims to enhance wildfire risk estimation, improve early detection, and support comprehensive wildfire management strategies by leveraging advanced data integra tion and analysis techniques. In summary, the background context provided by Borou jeni et al. [13] and the comprehensive review by Xu et al. [18] offer a valuable foundation for understanding the current state of wildfire risk prediction research. These works are recommended for further reading to gain deeper insights into the methodologies and challenges in this field. The next section will introduce our proposed solution, highlighting its unique contributions and improvements over existing methods.

III. PROPOSED FRAMEWORK

Our proposed framework for wildfire risk prediction and management leverages advanced ML algorithms and autonomous drone technology to enhance the monitoring and management of wildfire-prone areas. The framework is designed to be robust, flexible, and scalable, capable of integrating data from various sensors and providing insights to assist in wildfire management. The overall conceptual components of our framework encompass several key elements, each representing a critical aspect of the system. These components include data acqui sition and sensor integration, data management and fusion, API and communication, AI and ML, operator console, and embedded hardware and edge computing. Together, these elements form a comprehensive solution for wildfire risk prediction and management. The detailed architecture for the system, including the relationships and data flow among subsystems, is depicted in Figure 1. Data comes from sensors embedded in drones with the associated metadata — such as geolocation, sensor type, and timestamp. Then, data is processed to be stored in a database. A data transformation process is applied to the raw data to generate a dataset in a standard format — we chose the GeoJSON format for this purpose. An inference module performs two main tasks: extract features from the raster data — effectively, create vector data — and predict the value of the target variable in regions of interest. The front-end module, which we call the operator console, is responsible for displaying not only the inference results but also the confidence level of the prediction. The operator console is also responsible for displaying the drone's location and planning the next collection of data. The system is designed with scalability in mind, enabling the integration of additional drones and sensors as necessary. The framework is built

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

around several key components, each fulfilling distinct roles within the system. Below, we provide a brief overview of these subsystems:

A. DATA ACQUISITION AND SENSOR INTEGRATION

The data acquisition subsystem collects high-resolution data relevant to wildfire conditions. This includes visible spectrum imagery, thermal imagery, humidity, temperature, and other environmental data. These data can be collected from real-world sensors or simulated sources, ensuring comprehensive coverage and flexibility in data gathering. Therefore, data acquisition pertains more to identifying the relevant environmental data to be collected than the sensors' definition. Some environmental data, such as the vNDVI index [17], are derived from the raw data collected by the sensors. The data acquisition subsystem is responsible for processing this raw data to generate the derived data. This subsystem also ensures that the data collected is correctly annotated with metadata, such as geolocation, timestamp, sensor type, and other relevant information. This metadata is crucial for data management, fusion, and transformation processes. Section IV provides a detailed discussion of the data acquisition subsystem, including the selection of sensors and data collection strategies.

B. DATA MANAGEMENT AND FUSION

The data management and fusion subsystem is responsible for storing, organizing, and processing the data collected by the sensors. This subsystem ensures that the data is correctly stored in a database, allowing for efficient retrieval and processing. It also handles data fusion from multiple sensors, ensuring that the data is properly integrated and synchronized. Considering the spatio-temporal nature of the data col lected, there are two general types of data: raster data and vector data. Raster data represents dense data matrices, such as images or temperature maps. Vector data, on the other hand, represents a single numerical value associated to a specific geographic location, such as the humidity level or the presence of an object. Section V provides a detailed discussion of the data management and fusion subsystem, including the selection of data storage technologies and data processing strategies.

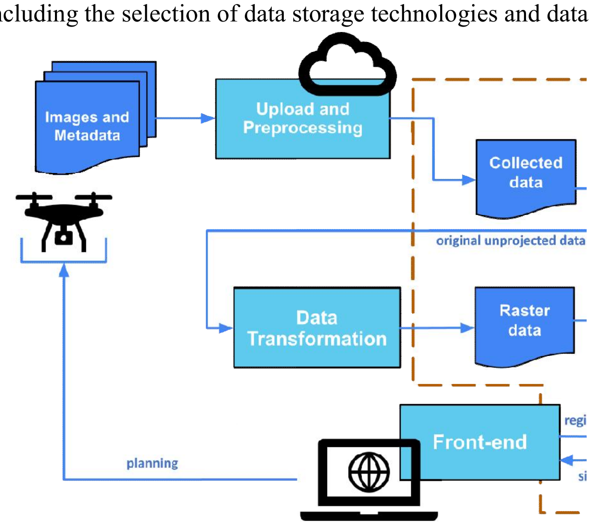


FIGURE 1. Framework architecture showing relationships and data flow among subsystems. Data flow from drone sensors to storage. Data transformation ensures data with geolocation is in the proper format. The inference module extracts features and predicts target variables, while the operator console displays results, confidence levels, drone locations, and plans for future data collection

C. API AND COMMUNICATION

The API and communication subsystem provides the nec essary interfaces for data exchange. It ensures seamless communication between different system components, facil itating efficient data transfer, which allows data interop erability between drones, sensors and other parts of the system. To implement the API and communication subsystem, we consider using standard protocols and formats, such as RESTful APIs and JSON. These technologies ensure the systemis flexible andscalable, allowing easy integration with other systems and components. Moreover, by providing a standardized interface, the system becomes more robust and easier to maintain and extend in the future. Section VI provides a detailed discussion of the API and communication subsystem, including the selection of communication protocols and data formats.

D. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

This subsystem manages the ML models used for data transformation and inference. It includes a selection of appropriate AI algorithms, training, validation procedures, and the deployment of models.
Amongthetechniquesusedinthissubsystem,wehighlight the use of deep learning and computer vision algorithms for feature extraction in raster data. Since environmental monitoring consists of tracking certain variables over time and space, using computer vision to process raster data is essential for the system to extract the features of interest from the raw data. For instance, one might be interested in observingthenumberoftreesinaspecificregion.Inthiscase, the system must be able to detect the trees in the image and 139870 count them — effectively converting the raster data in vector data. Once data is transformed into vector data, the AI and ML subsystem is also responsible for predicting the value of the target variable in regions of interest not directly observed. This prediction is needed since data collection is sparse both in time and space. Finally, it is essential to employ ML methods that can provide a measure of confidence in estimations. This builds trust in the system and enables proactive decisions of where new data collection missions should be undertaken with more significant gain. Section VII provides a detailed discussion of the AI and MLsubsystem,includingtheselectionofalgorithmsweused.

E. OPERATOR CONSOLE

The operator console provides a user-friendly interface for viewing the system's predictions and outputs. It is critical for decision-making because it enables the visualization of wildfire risk prediction on a map. With the visual and numerical estimation of target variables related to wildfire risk, the operator can make informed decisions about the necessity of taking action in a specific area to prevent or contain fire events. The confidence level of the prediction is also displayed, providing the operator with a measure of the reliability of the information. Using the operator console, the operator can also plan the next data collection mission, ensuring that the system collects data most efficiently. This component involves developing features like AI and ML API consumption, data visualization, and map representation. This can be accomplished through building a comprehensive user interface, so the operator can easily access andvisualize wildfire prediction estimates. The design focuses on usability and efficiency to support operators in effective decision-making for wildfire prevention. Section VIII provides a detailed discussion of the operator console subsystem, including the selection of data visualiza tion technologies and user interface design.

F. EMBEDDED HARDWARE AND EDGE COMPUTING

This subsystem involves augmenting the data-gathering capabilities of the DJI Matrice M30T drone (i.e., the one acquired for this project) with additional sensors and embedded hardware. The selection and integration of these sensors are crucial to ensure efficient data collection and synchronization. The focus is on embedding the sensors into the system to optimize its capabilities for providing the necessary data for wildfire monitoring while ensuring seamless integration with other system components. In this context, edge computing enables efficient data processing closer to the data source. Section IX provides a detailed discussion of the embedded hardware and edge computing subsystem, including the selection of sensors and edge computing technologies.

G. SIMULATION AND REAL-WORLD TESTING

This subsystem is dedicated to validating and verifying the system through simulations based on case studies of interest. It ensures that the system performs as expected under controlled conditions. Real-world testing involves planning, preparing, and executing functional project demonstrations in actual environments. This step ensures that the system

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

can handle real-world challenges and conditions effectively. This comprehensive framework ensures that all aspects of wildfire risk prediction and management are covered, from data acquisition and processing to AI model deployment and real-world validation. Section X provides a detailed discussion of the simulation and real-world testing subsystem, including the selection of case studies and testing strategies.

IV. DATA ACQUISITION AND SENSOR INTEGERTION

Data acquisition for this study involved collecting informa tion from both real-world sensors and simulated sources, as a way of ensuring comprehensive coverage and flexibility. An essential component of our data acquisition strategy was the use of a DJI Drone M30T, which conducted a total of 37 flights, accumulating 3 hours and 34 minutes of flight time. The drone was employed in the capture of real-time data across various locations and environmental conditions. Figure 2 shows a summary of data collection missions performed. For these flights, all 4 cameras of the drone were utilized, capturing images in visible light with wide aperture, standard visible, visible with analog zoom, and infrared spectra. The synchronization of imaging with telemetry data (including geographical positioning, time, etc.) required the development of a publicly available software tool, accessible at GitHub.5 This tool also includes functionalities for generating GeoTIFF files and performing data filtering, essential for interoperability with the data management system developed in this project. Most flights were conducted in São José dos Campos/SP, Brazil, at the vicinities of the ‘‘Parque de Inovaço Tecnológ ica’’6 characterized by significant variations in land use and land cover (see Figure 3). Data capture during the drier seasons also facilitated variations in vegetation indices.

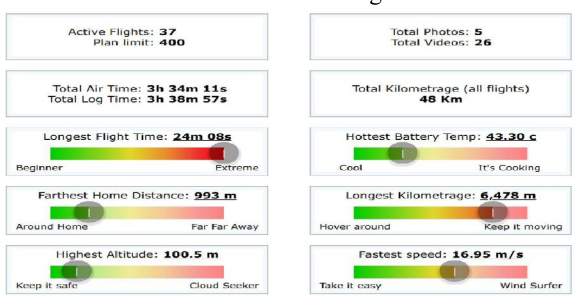


FIGURE 2. Summary of DJI M30T drone flight missions for the tests in this project.

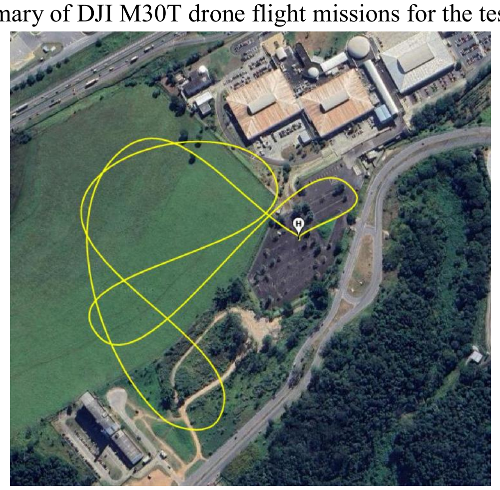


FIGURE 3. Example of flight conducted from the Pit parking area. The drone flew over the Parque de Inovaço Tecnológica (PIT) in São José dos Campos/SP, Brazil. Different types of vegetation and buildings can be observed.

The embedded system developed also permitted the capture of temperatureandhumiditydataduringflights.After all variables of interest are synchronized both in time and space, they are stored in a database for further processing. See Section V for more details, including our definition of raster andvector. Asummaryofthedatacollectedandderived during the drone flights is presented in Table 1.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

TABLE 1. Summary of real-world data collected during drone flights and derived from other sensors. Each piece of raster or vector data is associated with a geolocation and timestamp.

Variable	Source
Visible light images	DJI M30T
Infrared images	DJI M30T
Temperature	Embedded system
Humidity	Embedded system
VARI	Derived from images
vNDVI	Derived from images
People detection	Derived from images (AI/ML)

Wehighlight that vegetation indices are mathematical cal culations applied to images to estimate vegetation properties, such asbiomass,chlorophyll content, and canopycover.They are essential tools for monitoring vegetation health, land cover changes, and agricultural productivity. In our system, we employed two such indices: VARI (Vis ible Atmospherically Resistant Index) and vNDVI (Visible band Normalized Difference Vegetation Index), which are derived from the RGB imagery captured by the drone. The people detection data is also derived from the images, using AI/ML algorithms to identify and locate people in the images. More details on the AI/ML algorithms used in this project are presented in Section VII. Furthermore, alongside the acquisition of real-world data, we used realistic simulations to enhance our capability of capturing data in adiversesetofsituations andcontextsandin high quantities. The data-augmentation capabilities provided by the simulations played a crucial role in our ability to train our ML models, and in testing and validating our system's capabilities under a diverse set of conditions. Detailed discussions regarding the simulation method ologies and their corresponding results are presented in section X.

V. DATA MANAGEMENT

In the realm of spatiotemporal data analytics, fusion, and effective data management present significant challenges due to the diversity of file formats, including proprietary and closed formats from sensors vendors, and the mul tiplicity of data sources such as sensors, databases, and external platforms. Additionally, the data is typically large, heterogeneous, collected in a distributed manner, and may be ingested via streaming. These characteristics make the process complex and demanding. To overcome these challenges, our approach aims to provide a unified view of the data while preparing it for seamless ingestion by ML pipelines. We propose a robust data management architecture that segments data processing and enrichment into well-defined stages, each designed to enhance data quality, standardization and scalability. The architecture illustrated in fig. 4 comprises three independent persistence layers: raw, raster and vector data that can be implemented using appropriate persistence solutions based on their specific use cases.

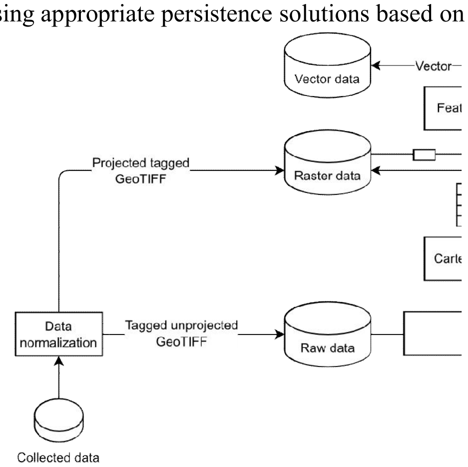


FIGURE 4. Data ingestion architecture with three stages: raw, raster and vector persistence layers.

Given the virtual impossibility of ingesting all existing formats andencodings, westandardized the ingestion process to accept only GeoTIFF data, delegating this task of data normalization to thedatasources.Thisdecisionwasdrivenby the fact that GeoTIFF is a widely used format for geospatial data, supported by most tools and libraries in the geospatial

DOI: 10.48175/IJARSCT-30114

ISSN 2581-9429 IJARSCT

ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

domain. Standardizing on a single format also significantly reduces the complexity of the fusion process and minimizes the number of potential bugs in the production code by limiting the range of exceptions that need to be handled. In our architecture, raw data is not the collected initially data but rather data that has already been converted into a GeoTIFF format. The delegated task Data normalization in f ig. 4 shows that this transformation results in either raw or raster data, which is preferred for ingestion. Raw data is defined as (1) data in the GeoTIFF format, (2) tagged with the metadata CAPTURE_TIMESTAMP, an ISO8601formatted datetime string indicating the time the data was captured, and (3) unprojected raster data, i.e., data in the original spatial reference system of the data source. Raster data, in contrast to raw data, is (1) projected into a Cartesian coordinate system, (2) a single tile of a raw data, (3) has the coordinate reference centered on each tile, and (4) is also tagged with CAPTURE_TIMESTAMP. As depicted in fig. 4, the ingestion of raw data auto matically converts it into raster data through the ingestion pipeline during the Cartesian grid tiling task. This task not only projects raw data into the Cartesian coordinate system but also perform a tiling normalization, resulting in multiple raster data tiles of fixed predefined size. In contrast, ingestion of raster data bypasses the Cartesian grid tiling task and directly stores data into the raster layer of persistence, assuming that the tiling normalization has already been performed. The primary reason for tiling normalization is to enhance performance. For instance, retrieving specific areas involves reading only the tiles related to the requested area instead of a large region. It also facilitates parallel processing of data through domain decomposition, and indexing can be applied to provide a scalable solution to handle many tiles. The normalization of the tile size is motivated to streamline its use in ML pipelines. Maintaining a fixed size for all tiles allows us to easily feed the data into ML models without requiring additional pre-processing. Although we do not strictly enforce a fixed tile size, we strongly recommend it for performance reasons and understand that different tile sizes in the raster layer indicate a lower level of data curation quality. Over time, as the quality of sensors improves and areas of interest result in varying tile sizes, continuous data renormalization is necessary to maintain data quality. Vector data represents the final layer of persistence in our architecture. It is defined as an actual number representing the measurement of a feature of interest, such as temperature, humidity, or wind speed, intended for consumption by a ML model. Each data point is annotated with a tag name, a string denoting the feature's name, and it includes the same spatiotemporal information as the center of the tile from which it originates. This spatiotemporal information comprises longitude, latitude, altitude and timestamp. During the ingestion pipeline, raster data is automatically transformed into vector data by the Feature extraction task. This task is scheduled after the persistence of raster data. It is designed to run an external feature extraction pipeline, resulting in an actual number representing the value of the feature at the center of the tile. In other words, each vector data point represents information aggregation from the perspective of the feature at the center of the tile. Thus, the relationship among the data types is as follows: one raw data input results in multiple raster data tiles, and one raster data tile may generate one or more vector data points, each with a different feature tag. Both raw and raster data can bemultichanneldata, containing multiple bands, and the Feature extraction task must be aware of this to properly generate the vector data. However, only a single type of raw and raster data is expected in this architecture.

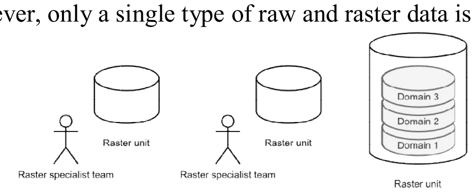


FIGURE 5. Logical organization of raster units.

Tomanagedifferent types of raster data, one must replicate our architecture for each type of data, as shown in fig. 5. We define a raster unit as a logical unit comprising an instance of our data architecture dedicated to managing a single type of raster data, and it is associated with a team of specialists responsible for curating this data. Each raster unit is responsible for managing the raw, raster, and vector data, including the definition of the tile size, the feature tags, and the feature extraction pipeline. In the example shown in fig. 5, we illustrate three raster units representing three distinct raster data types. This example demonstrates that a team can benefit from modern database resources by transparently partitioning the database into three segments representing a spatial division, potentially assigning different teams for the duration of each partition domain. This architecture is intended to provide a highly scalable and flexible solution for

DOI: 10.48175/IJARSCT-30114

managing different raster data types. It mirrors real-world scenarios where different teams of specialists typically manage distinct data types. However, a significant limitation of our data architecture is that it restricts the generation of vector data to a single type of raster data. In other words, features derived from multiple raster data types are not supported. We decided to restrict the generation of vector data to a single type of raster data primarily for performance reasons. If really necessary, the raster specialist team can incorporate all the required information for the feature extraction into a single raster data file by leveraging the multichannel capacity of the GeoTIFF format. However, we understand that a better approach to dispose of the data for ML models as vector data and let this one to learn the optimal way to combine the information in the database. Additionally, for engineering composite features of inter est, our data architecture can be utilized as a component in a more complex data pipeline. For example, one could create an additional vector database that combines information from different vector or raster datasets.

A. METADATA AND PERSISTENCE TECHNOLOGIES

Raw data typically involves large geospatial data files that can efficiently persist in a distributed file system such as HDFS or S3. Since our solution is agnostic to geospatial features from the database, we prefer to use a scalable NoSQL solution. However, for small databases or teams interested in traditional geospatial analytics, a traditional Relational Database Management System (RDBMS) such as PostgreSQL with PostGIS extension can be used. Raster data comprises two components: the raster data itself and its spatiotemporal metadata. During the persistence of raster data, the tile's corners and the data capture timestampareextracted to composethemetadata.Weregister the following metadata: top-left, top-right, bottom-left, and bottom-right of the latitude and longitude of the corners of the tile, the latitude and longitude of the center of the tile, and the timestamp of the data capturing. Note that our definition of tile is not necessarily a rectangular area. We register the following metadata: all the possible corners of the tile, derived from the combinations of top-down, left-right, and high-low of the coordinates' longitude, latitude, and altitude; the center of the tile using the exact coordinates; and the timestamp of the data capture. It is important to note that our definition of a tile is not necessarily a rectangular area. In fact, the managementofraster data is entirely performed using the metadata, with the raster image itself being used only for ML models. Therefore, the metadata can be implemented in a traditional RDBMS. Since queries for regions are performed over the metadata, indexes must be created to provide a scalable solution. For large databases, we suggest using a distributed database such as Cassandra, HBase, or MongoDB. For databases with high-frequent measurements of the same spatial area, we recommend using a time-series database like TimescaleDB, which offers transparent partitioning by time. Vector data is similar to raster data but contains a single actual numberinsteadofarasterimage.Itincludesatagname andjust the coordinates of the center of the tile. Therefore, the same storage solution used for raster data can be applied to vector data.

VI. API AND COMMUNICATION

Effective API design and communication protocols are critical to ensuring seamless interaction between the various components of our proposed system. A robust API design enable us to hide implementation details and provide a standardized interface for data exchange. In the system, there are two major paths of data exchange: the first is the communication with the data management system, which is responsible for storing and managing the data of the system; the second is the communication between the front-end and the inference engine, which is responsible for processing the data and generating the predictions of the variables of interest.

A. DATA STORAGE AND RETRIEVAL

In our architecture, each raster unit is implemented as a stateless microservice to ensure high availability and scalability. By adopting a stateless design, we can manage each service independently, enabling better load balancing and failover strategies. This design choice is crucial for maintaining continuous operation and reducing downtime. We employ a RESTful interface for seamless integra tion between services, utilizing the Flask framework. The RESTful approach provides a standardized and easy way for services to communicate, ensuring compatibility and use across different components of the system. Flask was chosen due to its simplicity and widespread use, which facilitates rapid development and maintenance. To optimize performance within a raster unit, it is recommended that its persistence services for the three data layers (raw, raster, and vector) be located on the same machine or in close proximity. This configuration reduces data transfer latency and improves the efficiency of data processing tasks, as the data does not need to traverse a network. Such a setup is particularly beneficial when handling large geospatial datasets, where minimizing delay is critical. Our architecture also specifies that the transformation of raster data to vector data shouldbeexecutedlocallywithinthe raster unit whenever possible. This design choice stems from the specialization of each raster unit in curating a specific type of data. By localizing these transformations, the team responsible for each raster unit can leverage their expertise to fine-tune and optimize the processes. Alternatively, for f lexibility reasons, the inference engine can read raster data and add vector data directly within the raster unit.

Thecombinedcollectionofvectordatafromallraster units is made available to a centralized inference engine, which integrates these data for final processing and visualization. This centralized inference engine is responsible for consoli dating the transformed data and making it accessible to the front-end module for user interaction and visualization. Overall, this architecture promotes a modular and efficient approach to data management and communication, where each raster unit operates independently yet cohesively within the broader system, enabling effective data storage, retrieval, and analysis.

B. USER INTERFACE

Thecommunicationbetweentheoperatorconsole(front-end) and the inference engine is simpler than the communication with the data management system in terms of data exchange. Awebinterfacerunningontheuser'sbrowserusesaRESTful API to request certain variable of interest, a region and a time span. The inference engine then uses these parameters to retrieve the necessary data from the data management system andperformtheanalysis. The console receives the results and displays them to the user. A point of attention is the authentication of the user. Differently from the communication between the inference engine and the data management system, many users can access the front-end over time. Therefore, an authentication layer is necessary to ensure that only authorized users can access the inference system and to manage the permissions of each user. The proposed system is designed with a modular and generic architecture, making it adaptable to complement existing solutions and systems. This architectural decision includes avoiding a fixed authentication method, thereby offering flexibility to orchestrate authentication at the infrastructure level or within the hosting system. This approach facilitates integration with diverse technological environments, enabling interoperability and extending the system's utility across different applications and use cases.

VII. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

AI techniques have been employed in various aspects of this project, including raster to vector data transformation and spatial regression tasks. Both deep learning and classic Machine Learning (ML) techniques were utilized to address these tasks. The selection and optimization of these models are derived from the context in which they are applied. A. RASTER TO VECTOR DATA TRANSFORMATION For raster to vector data transformation, two AI techniques were utilized. Part of the data collected by drones are raster images, in our experiments, top-view images in the visual and near-infrared spectrum. These images must be processed to extract augmented relevant information, such as vegetation indices and the presence of people. Firstly, to extract vegetation indices, indices such as VARI and NDVI can be directly calculated from the images. However, in some cases, the images may not contain the necessary bands to calculate these indices. In these cases, vegetation index regression was performed using techniques inspired by [17], where a new visible band index (vNDVI) was proposed. Secondly, the presence or proximity of people in the area is also an important aspect of fire risk [22]. To detect people in the images, a YOLO detection model [28] was fine-tuned on a dataset containing top-view images of people. YOLO is a real-time object detection system that can detect multiple objects in a single image using a combination of clustering, deep learning, and regression methods. Fine-tuning involves training from a pre-trained and typically generic model for a more specific task, in this case, person detection from a top-view or aerial angle. Fine-tuning is usually performed with smaller datasets, for less time, and with a lower learning rate. For our fine-tuning process we selected a dataset con taining top-view images of people, and augmented it using generated data from our foto-realistic simulations. Our base model was the YOLOv4 [28] model pre-trained in the COCO dataset. The specifics of the fine-tuning, such as dataset augmentation and hyperparameter adjustments, are determined by the implementer based on their needs and constraints. In our case, the fine-tuning resulted in a model achieved a mean-average precision (mAP) of 90%, with an intersection over union (IoU) threshold of 0.5, being very suitable for our task. Finally, other variables of interest can be extracted from the images, such as the presence of vehicles, buildings, and other objects. Image segmentation and further classification of terrain types can also be performed using AI techniques [29], [30]. B. SPATIAL REGRESSION To feed the operator console with relevant information, the system needs to provide two types of data: the predictions of the variable of interest and the uncertainty of these predictions. The data is shown in an overlay on the map, and the operator can use it to make decisions about the fire risk in the area and to plan new data collection missions. Temperature, humidity, vegetation indices and the presence of people are the variables of interest supported by the current version of the system. Spatial regression tasks were addressed using a method developed specifically for the project:

• K-NearestNeighbors(KNN)withBootstrapping:We employ KNN combined with bootstrapping for spatial regression. KNN estimates values at unvisited points in space based on nearby measurements. Bootstrapping, which involves re-sampling with replacement, helps in estimating the distribution of predictions and assessing uncertainity.

• Adherence to Tobler's Law: The method adheres to Tobler's First Law of Geography [31], ensuring spatial coherence in predictions by considering that nearby locations are more related than distant ones. This particular task is similar to the problem of spatial interpolation, where the goal is to estimate the value of a variable at an unmeasured location based on the values of the variable at nearby locations. Kriging interpolation is a common method used in spatial statistics for this purpose [32]. Other spatial regression methods are also available, such as geographically weighted regression (GWR) [33] and spatial autoregressive models [34]. These methods can be used to model spatially varying relationships between variables. The developed spatial regression method, KNN with bootstrapping, is non-parametric and does not require prior training, only accesstoknownmeasurementdata.Thisformat has the advantage of being easily deployed in varied locations without the need of adaptation or further training. Additionally, the repeated estimation of the variable of interest using bootstrapping is especially useful because it not only provides the average value ofthevariable wearelooking for, but also the probability distribution of the prediction. The variance of this distribution is a good indicator of how sure or unsure the model is regarding that prediction, and we use it as a measure of uncertainty of the model's inference at each point in space. While specific implementation details such as model selection and optimization can be adjusted based on the project's requirements, the methods outlined here provide a foundation anddemonstrateeffectiveapproachesforthetasks described.

VIII. OPERATOR CONSOLE

The Console is a lightweight web application developed based on the Web Server Gateway Interface (WSGI) specification. It is composed of both a back-end and a front end component. The front-end part, henceforth called User Interface (UI), is considered a key component of this project because front-end applications allow end users to perform manyhardtasks byabstracting complex concepts and turning them into intuitive operations. In this project, the operator utilizes the Console to perform the following actions through the user interface:

• Map visualization: Displaying a map with countries, cities, neighborhoods, and other spatial and geographic elements; • Parameters definition: Parameter selection to filter data for the desired prediction, such as start and end dates, and the variable of interest, such as temperature, humidity, pressure, etc.;

• Initial and final coordinates selection: Coordinates definition to set the region of interest for the prediction and data retrieval;

• AI data sending and retrieval: Sending and getting data from the back-end through a RESTful request to the AI module of the solution;

• Prediction visualization: Projection, on a map, of pre dicted values and confidence levels for each data point within the region of interest previously selected by the operator.

Map visualization is built using a Python wrapper for the Javascript library Leaflet.js, which allows data visualization on OpenStreetMap maps without giving up Python data manipulation capabilities. The same logic applies to data prediction visualization: Javascript and third-party libraries render the data points retrieved from the AI module by the back end of the Console. While parameter selection is accomplished with regular HTML forms and fields, coordinates

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

selection is also built using Javascript to capture the input from the user's click dynamically. The only action involving back-end operations is sending and retrieving data from the AI module. This action is processed on the server side because the amount of data that the AI module returns after querying the database may be too large for the client side to handle in the front end, which could result in application interruption, poor performance, or a bad experience for the end user.

IX. EMBEDDED HARDWARE

In addition to the existing sensors on the drone (i.e., visible spectrum and thermal images), we developed and integrated a module with temperature and humidity sensors. Figure 6 presents a schematic of device communications, such as sensing, control, and recording modules.

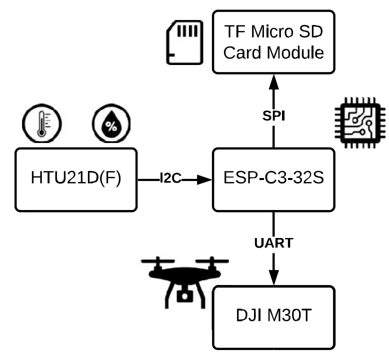


FIGURE 6. Communication scheme of devices, such as sensing, control and recording modules.

This module uses the HTU21D(F) sensor, a dedicated plug-and-play digital humidity sensor with TE Connectivity temperature output. Its compact size (i.e., 18×16×11.5 mm) and its lightweight design (i.e., 1.4 grams) make it ideal for embedded applications. The sensor has a measurement time of 50msandprovidescalibrated and linearized signals in I2C digital format. The direct interface with the microcontroller is carried out through the digital humidity and temperature output module, and the resolution of both relative humidity andtemperaturecanbeadjustedbycommand7(from8/12bits to 12/14 bits). Some specifications of the sensor module are presented in Table 2.

Characteristics	Ranges
Supply voltage	3.3 to 5
Operational humidity range	0 to 10(
Humidity signal accuracy (5% to 95%RH)	$\pm$ 2
Operational temperature range	-40 to 12
Temperature signal accuracy (-30 to 90^{o}C)	$\pm$ 1

TABLE 2. Specifications and operation mode of the HTU21D(F) sensor module

A. INTEGRATED SYSTEM FOR DATA ACQUISITION (ISA DATA)

The integrated data acquisition system, named ISA-Data, uses an ESP-C3-32S microprocessor board with USB, Wi Fi, and Bluetooth connectivity. This microprocessor board is knownforitslow-powerconsumption,highperformance,and robust features, making it outstanding for on-board aircraft applications.8 The ESP-C3-32S features: • 32-bit RISC-V microcontroller: operating at up to 160 MHz. The microcontroller provides sufficient processing power to handle the demands of acquiring and processing data from temperature and humidity sensors; • 400 KB RAM memory: RAM memory ensures suffi cient storage for the data collected by the sensors and for running the data acquisition software; • 4 MBFlash Memory: Flash memory stores embedded system firmware and configuration data; and • USB, Wi-Fi, and Bluetooth Connectivity: these inter faces allow to communicate with other aircraft systems, transfer data to a computer on the ground, or connect to mobile devices for remote monitoring. We prioritized the ESP-C3-32S because it combines low power consumption, high performance, and robust features. To capture temperature and humidity signals, we use the HTU21D(F) module, which operates at voltage levels between 3.3 and 5V. The temperature measurement range is −40to125◦C,andtherelative humidity varies between 0 and 100%.

B. DATA LOGGING AND SYNCHRONIZATION

After processing the sensor data, the results are recorded in log format, synchronized with the aircraft's image captures. To do this, we use a micro SD card read/write module9 (TF Micro SD Card Module).

C. DIRECT COMMUNICATION BETWEEN SYSTEMS

Taking advantage of the USB connection on the micropro cessor board, weconductdirectcommunicationtestsbetween 7https://learn.adafruit.com/adafruit-htu21d-f-temperature-humidity sensor 8https://www.waveshare.com/wiki/ESP-C3-32S-Kit 9https://learn.adafruit.com/adafruit-micro-sd-breakout-board-card tutorial the aircraft system and the onboard temperature and humidity data acquisition system.10 For communication between the different components of the system, we use three main protocols: 1) Serial Peripheral Interface (SPI): ideal for fast and synchronized communication between the microcon troller and high-speed peripherals, such as sensors and displays; 2) Inter-Integrated Circuit (I2C): low-cost, low-power protocol, ideal for connecting multiple devices to a single bus, such as temperature and humidity sensors; and 3) UniversalAsynchronousReceiverTransmitter(UART): used for asynchronous serial communication between two devices, such as the microcontroller and a computer for debugging or sending data. Choosing the appropriate protocol depends on several factors, such as data transfer speed, number of connected devices, power consumption, and distance between devices. Regarding temperature and humidity sensors, I2C is a good option due to its simplicity, low-power consumption, and ability to connect multiple sensors on a single bus [35].

X. SIMULATION AND REAL-WORLD TESTING

The simulation of scenarios involving autonomous aircraft has been evolving alongside the expansion of their civil and military use. However, demonstrating the feasibility of these proposals may depend on realistic simulations that help identify potential yet unclear obstacles. Developing new solutions without these simulation tools is a practically unfeasible task given the complexity of the scenarios [36]. The simulation module is a critical component of our frameworkfordisasterriskprediction,enablingthevalidation andverificationofthesystemundercontrolledconditionsthat closely mimic real-world scenarios. This module leverages the Unreal Engine 5 for creating photorealistic 3D environ ments and integrates with Microsoft's AirSim to simulate UAVoperations with high fidelity. The creation of 3D scenarios within Unreal Engine 5 is a meticulous process that involves manually designing the environment to reflect real-world conditions as closely as possible. Despite being a manual process, the scenarios are informed by actual data collected from UAV flights and various sensors, ensuring that weather conditions, object placements, and other environmental factors are accurately represented. The advanced rendering capabilities of Unreal Engine 5 allow for the creation of highly detailed and immersive virtual worlds, which are essential for realistic simulation. AirSimopen-sourceplatform[37],developedbyMicrosoft, is integrated into our simulation framework to provide realistic visual and physical simulations of UAV operations. This integration involves setting up the virtual environment in Unreal Engine 5 and configuring AirSim to interact seamlessly with this environment. AirSim supports a variety of UAV models and simulates complex flight dynamics, altitude, gravity, GPS, wind, and physical interactions, which are critical for creating a realistic and functional simulation. To demonstrate and validate the simulated environment, we defined two primary tasks for the simulated UAVs: (i) monitoring an area of interest via a predefined path, and (ii) detecting individuals within the area of interest. These tasks are designed to be scenario-agnostic, meaning they can be replicated across different simulated environments as they are developed. Initially, simulated data were collected in custom-built scenarios to test and refine the simulation module and data flow processes. These tasks provide a robust foundation for evaluating the system's capabilities and performance in different conditions. For the person detection scenario, we used the flight mission specified with the real drone obtained for the project as inspiration. Figure 7 shows in (a) the real images captured with the drone and contrasts with (b) showing images of the scenario built in Unreal and a view from the simulated drone. The monitoring task involves the UAV following a predefined path within the simulated environment, collect ing data through its onboard sensors. This data includes high-resolution imagery, environmental variables such as temperature and humidity, and object detection information. The goal is to replicate real-world UAV operations as closely as possible, ensuring that the data collected in the simulation is representative of actual flight data. This approach allows for thorough testing and validation of the system's data acquisition and processing capabilities.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

The detection task focuses on the UAV autonomously identifying individuals or objects within the area of interest. This capability is crucial for applications such as search and rescue or security surveillance. The UAV utilizes advanced AI algorithms and sensor fusion techniques to detect and analyze objects within the environment. This task tests the system's ability to process and interpret complex data, providing valuable insights into its operational effectiveness and potential areas for improvement. The simulation module plays a pivotal role in our disaster risk prediction framework, providing a realistic and controlled environment for testing and validating the system. By integrating Unreal Engine 5 and AirSim, we have created a sophisticated and immersive simulation platform that enhances the accuracy and reliability of UAV operations in disaster management. This comprehensive approach ensures that all aspects of data acquisition, processing, and AI model deployment are thoroughly tested and validated, ultimately contributing to more effective disaster risk prediction and management.

XI. ANALYSIS

Thesystemwasdesignedtosupportandenhancethemonitor ing of potential wildfire areas. Therefore, we have developed a system for collecting, managing, annotating, and fusing multisensor data, which can be easily integrated with other

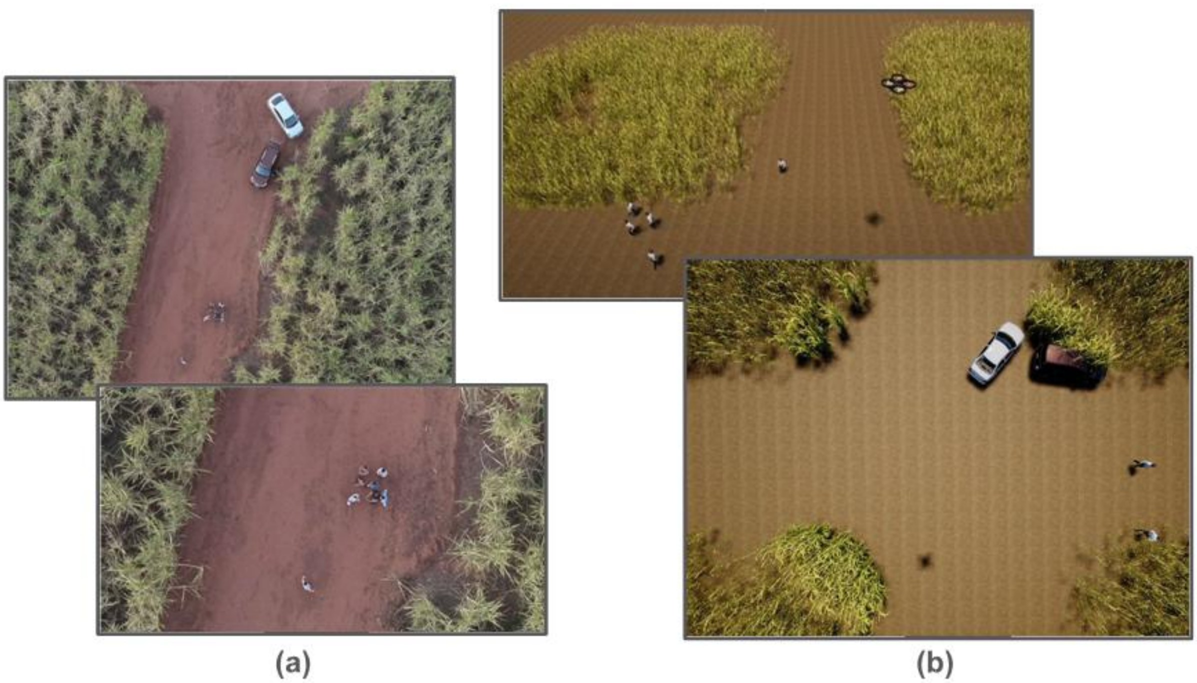


FIGURE 7. (a) Real-world drone image, (b) simulated scenario

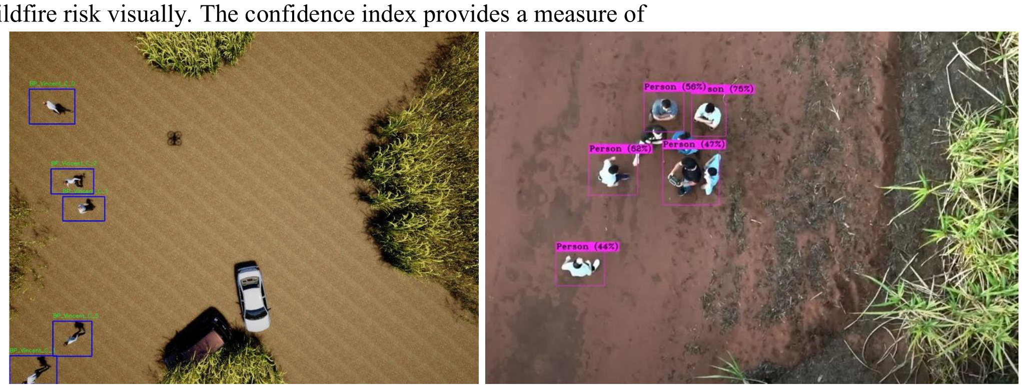
related data systems such as satellite information, vegetation maps, or other structured or unstructured data. The system's f lexibility facilitates the use of AI in various scenarios, extending to both civilian and military applications.

A. PERSON DETECTION

To analyze the simulated environment, two tasks were defined to be performed via a simulated drone: 1. Monitoring of the area of interest through a predefined path; 2. Detection of persons in the area of interest. In the first task, the monitoring followed a quadrangular path starting from the four points of the area of interest. The simulated drone navigated in a straight line between these points. For the second task, the simulated drone positioned itself at the center of the four points of the area of interest. The defined tasks are agnostic and independent of the scenario, allowing them to be conducted again as other scenarios are included. The simulated environment allows for the capture of images from the drone's simulated camera, including panoramic perspective images, and the creation of datasets with ground truth associations for each object of interest, in this case, persons. A camera was specially positioned for the simulation to generate data from an overhead or aerial angle (perpendicular to the ground), analogous to what is found in the real drone. Figure 8(a) illustrates the ground truth for person detec tion in a simulated scenario. This includes the accurate labeling of individuals within the simulated environment, which is crucial for training and validating the detection model. Figure 8(b) shows the results of the specialized YOLOv3 applied to real-world data captured by the M30T drone. Thefine-tuned model exhibits effective detection capabilities in actual conditions, suggesting the system's prospective performance in practical applications. Person detection consolidates the index of persons in the area ofinterest, serving as ariskmetric.Urbanareaswithhigh movement have a greater risk. This system's flexibility and efficiency in detecting individuals in different environments underscore its potential for enhancing situational awareness and risk assessment.

B. SYSTEM IN ACTION

The system's user interface allows operators to interact seamlessly with the predictions of the variable of interest, demonstrating the effective integration of data flow from the back-end to the front-end. In Figure 9, the operator selects a region of interest on the mapinterface. Theinterfaceprovidesfiltersforspecifyingthe date range and selecting specific metrics such as the vNDVI. This selection initiates a RESTful request to the back-end, which processes the query through the AI module to retrieve relevant data, filling the grid with the predictions. Figure 10 showcases the resulting visualizations for the selected area. The map displays the confidence index and risk estimation, allowing the operator to assess wildfire risk visually. The confidence index provides a measure of



(a) (b)

FIGURE 8. (a) Person detection ground truth in simulated scenario (b) Person detection in real-world image with fine-tuned YOLO model.

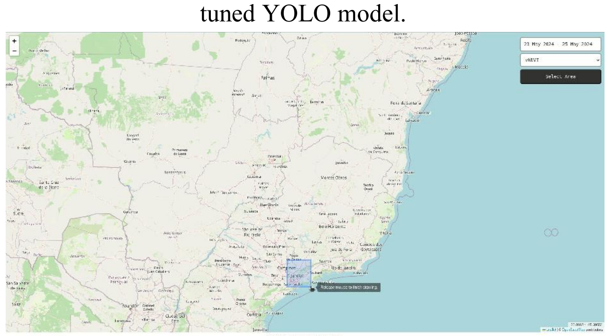


FIGURE 9. Selecting area of interest with filters for a date range and vNDVI as the variable of interest.

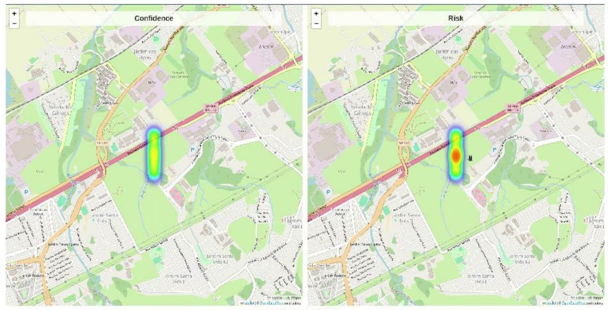


FIGURE 10. Visualizations for the confidence index and risk estimation for a chosen area.

the reliability of the predictions, aiding in decision-making processes. The interface abstracts complex data processing and AI analysis, providing the operator with visual tools for wildfire risk assessment. This interaction suggests the system's feasi bility and practicality in real-world applications, promoting efficient data flow and accessible information.

C. IMPACTS

Thedeploymentofourproposedsystemisanticipatedtoyield substantial benefits across economic, environmental, social, and technological domains.

Economically, the integration of unmanned flights can significantly reduce operational costs associated with wild f ire monitoring and management. The automation of data collection and analysis processes not only lowers the need for manual labor but also enhances the efficiency of surveillance operations. Furthermore, the indirect economic benefits stemming from improved wildfire detection and response capabilities can mitigate the costs related to property damage and loss of resources. Environmentally, the system holds promise for early detec tion of wildfires, which is crucial in preventing large-scale forest destruction. By aiding in the timely identification and combat of illegal wildfires, the system contributes to the preservation of ecosystems. Additionally, its application can extend to other environmental monitoring challenges, such as detecting illegal deforestation, thereby promoting broader ecological conservation efforts. From a social perspective, the system enhances the safety, life, and health of residents and passersby in monitored areas. Early detection and accurate risk assessment can lead to quicker evacuations and more efficient resource deployment, reducing the potential for human casualties. Furthermore, the project fosters the training and development of highly specialized human resources, enriching the skill set required for advanced environmental monitoring and AI applications. Technologically, our system represents a good use of AI for environmental monitoring. The integration of spa tiotemporal data from various sensors and the application of sophisticated AI algorithms enhance the precision and reliability of predictions. The system's architecture allows for continuous updates and improvements, ensuring that the latest advancements in AI can be incorporated. This adaptability paves the way for future innovations and broader applications, extending the utility of the system beyond the scope of wildfire management.

D. LESSONS LEARNED

Throughout the development and implementation phases, several key lessons were learned. We demonstrated that the system functions as intended, with effective communication between modules. The data management system is agnostic to the nature of the sensors, making it extensible to enhance the civil sector's capacity in utilizing AI for scenarios not covered in this proposal. Even with sparse data for certain regions, the use of AI methods allowed us to interpolate missing data points and provide comprehensive data coverage. Simulation proved promising for representing vegetation indices in simulated scenarios with different types of vegetation. Realistic simulations can also aid in the training of MLmodels. The data representation and storage model effectively handled real mission data from drone flights as well as datasets from simulated flights. The system showed its capability by correlating data from various sensors (humidity, temperature) with aerial images of the regions of interest, and detecting the presence of people in these areas. As technological contributions is worth to mention: • Software modulefor storage, management, and retrieval of data captured by drones or other multisensor data sources. • Software library for training AI models to estimate wildfire risk. • Operator console (Human-Machine Interface) for deci sion support based on the interpretation of predictions generated by AI models.

E. CONSOLIDATED INSIGHTS AND IMPLICATIONS

Despite the successes, we encountered several challenges and limitations that provide valuable insights for future work. Data transformation is currently performed server-side, necessitating the transportation of data for processing. Edge processing could alleviate this, but further investigation is needed into its implications. Simulationscenarioswerecraftedbasedonhistoricalflight mission data or expected real-flight characteristics. Automat ing the construction of these scenarios could enhance efficiency. Investigating the creation of digital twins for these scenarios might provide more robust and parallelizable solutions. Wedid not simulate multiple drones in a single execution. Implementing this with more parallel processing could significantly improve efficiency. Additionally, employing digital twins for drones, specialized in various flight tasks, warrants further exploration. Introducing a swarm of drones, or even just two, poses implications for data transformation and communication. This could increase complexity and requires careful consid eration and further study.

1) FUTURE DIRECTIONS

Investigating the feasibility and benefits of edge processing for data transformation to reduce latency and increase processing efficiency. Also, real-time communication with the drones might be beneficial in time sensity missions and decision-making. Developing automated tools for scenario construction, potentially leveraging digital twin technology, to create more dynamic and realistic simulation environments. Exploring the implementation of multiple drone simula tions in parallel, including the development of digital twins specialized for different flight tasks to enhance mission planning and execution. Assessing the impact of drone swarms on data processing and communication infrastructure to better understand and manage the increased complexity. Moreover, the usage of the uncertainty of the predictions to guide the drone's path and optimize the data collection process is a promising research direction.

XII. CONCLUSION

The proposed system introduced in our paper leverages AI and drone technology to enhance wildfire monitoring and risk assessment. By integrating data from various sensors and applying sophisticated AI algorithms, the system provides accurate predictions and visualizations for wildfire risk through a selection of variables of interest, such as vNDVI and urban movement. The system's user interface allows operators to interact seamlessly with the predictions of the variable of interest, demonstrating the effective integration of data flow from the back-end to the front-end. The system's deployment is anticipated to yield substantial benefits across economic, environmental, social, and technological domains. The system's architecture allows for continuous updates and improvements, ensuring that the latest advancements in AI and sensors can be incorporated. This adaptability paves the way for future innovations and broader applications, extending the utility of the system beyond the scope of wildfire management.

REFERENCES

[1] Copernicus Atmosphere Monitoring Service. (Jan. 2024). 2023: A Year of Intense Global Wildfire Activity. Accessed: Jul. 24, 2024. [Online]. Available: https://atmosphere.copernicus.eu/2023-year-intense global-wildfire-activity

[2] United Nations Environment Programme and GRID-Arendal. (Feb. 2022). Spreading Like Wildfire: The Rising Threat of Extraordinary Landscape Fires. Accessed: Jul. 21, 2024. [Online]. Available: https://www.unep.org/resources/report/spreading-wildfire-rising-threat extraordinary-landscape-fires

[3] E.Grant and J. D. Runkle, ''Long-term health effects of wildfire exposure: A scoping review,'' J. Climate Change Health, vol. 6, May 2022, Art. no. 100110.

[4] R. Libonati, J. L. Geirinhas, P. S. Silva, A. Russo, J. A. Rodrigues, L. B. C. Belém, J. Nogueira, F. O. Roque, C. C. DaCamara, A. M. B. Nunes, J. A. Marengo, and R. M. Trigo, ''Assessing the role of compound drought and heatwave

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

events on unprecedented 2020 wildfires in the Pantanal,'' Environ. Res. Lett., vol. 17, no. 1, Jan. 2022, Art. no. 015005.
[5] J. F. Marques, M. B. Alves, C. F. Silveira, A. A. E. Silva, T. A. Silva, V. J. D. Santos, and M. L. Calijuri, ''Fires dynamics in the Pantanal: Impacts of anthropogenic activities and climate change,'' J. Environ. Manage., vol. 299, Dec. 2021, Art. no. 113586.
[6] W. L. Lázaro, E. S. Oliveira-Júnior, C. J. D. Silva, S. K. I. Castrillon, and C. C. Muniz, ''Climate change reflected in one of the largest wetlands in the world: An overview of the northern Pantanal water regime,'' Acta Limnologica Brasiliensia, vol. 32, Sep. 2020, Art. no. e104.
[7] J. L. Geirinhas, A. C. Russo, R. Libonati, D. G. Miralles, A. M. Ramos, L. Gimeno, and R. M. Trigo, ''Combined large-scale tropical and subtropical forcing on the severe 2019–2022 drought in South America,'' NPJ Climate Atmos. Sci., vol. 6, no. 1, p. 185, Nov. 2023.
[8] L. C. Garcia, J. K. Szabo, F. de Oliveira Roque, A. de Matos Martins Pereira, C. Nunes da Cunha, G. A. Damasceno-Júnior, R. G. Morato, W. M. Tomas, R. Libonati, and D. B. Ribeiro, ''Record-breaking wildfires in the world's largest continuous tropical wetland: Integrative fire management is urgently needed for both biodiversity and humans,'' J. Environ. Manage., vol. 293, Sep. 2021, Art. no. 112870.
[9] P. S. Silva, J. L. Geirinhas, R. Lapere, W. Laura, D. Cassain, A. Alegría, and J. Campbell, ''Heatwaves and fire in pantanal: Historical and future perspectives from CORDEX-CORE,'' J. Environ. Manage., vol. 323, Dec. 2022, Art. no. 116193.
[10] J. Marengo, L. Alves, and R. Torres, ''Regional climate change scenarios in the Brazilian Pantanal watershed,'' Climate Res., vol. 68, nos. 2–3, pp. 201–213, May 2016.
[11] A.F.S.Ribeiro, P. M. Brando, L. Santos, L. Rattis, M. Hirschi, M. Hauser, S. I. Seneviratne, and J. Zscheischler, ''A compound event-oriented framework to tropical fire risk assessment in a changing climate,'' Environ. Res. Lett., vol. 17, no. 6, Jun. 2022, Art. no. 065015.
[12] (2023). Plano De Ação Para O Manejo Integrado Do Fogo No Bioma Pantanal. Accessed: Jul. 21, 2024. [Online]. Available: https://www.gov.br/ibama/pt-br/assuntos/fiscalizacao-e-protecao ambiental/incendios-florestais/arquivos/20230620PlanoacaoPantanal.pdf
[13] S.P.H.Boroujeni, A. Razi, S. Khoshdel, F. Afghah, J. L. Coen, L. O'Neill, P. Fule, A. Watts, N.-M.-T. Kokolakis, and K. G. Vamvoudakis, ''A comprehensive survey of research towards AI-enabled unmanned aerial systems in pre-, active-, and post-wildfire management,'' Inf. Fusion, vol. 108, Aug. 2024, Art. no. 102369.
[14] S. C. P. Coogan, F.-N. Robinne, P. Jain, and M. D. Flannigan, ''Scientists' warning on wildfire—A Canadian perspective,'' Can. J. Forest Res., vol. 49, no. 9, pp. 1015–1023, Sep. 2019.
[15] J.N.S.RubíandP.R.L.Gondim,''Aperformancecomparisonofmachine learning models for wildfire occurrence risk prediction in the Brazilian federal district region,'' Environ. Syst. Decis., vol. 44, no. 2, pp. 351–368, Jun. 2024.
[16] M. K. Kabir, K. K. Ghosh, M. F. U. Islam, and J. Uddin, ''Wildfire prediction in the United States using time series forecasting models,'' Ann. Emerg. Technol. Comput., vol. 8, no. 2, pp. 32–42, Apr. 2024.
[17] L. Costa, L. Nunes, and Y. Ampatzidis, ''A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms,'' Comput. Electron. Agricult., vol. 172, May 2020, Art. no. 105334.
[18] Z. Xu, J. Li, S. Cheng, X. Rui, Y. Zhao, H. He, and L. Xu, ''Wildfire risk prediction: A review,'' 2024, arXiv:2405.01607.
[19] J. W. Rouse Jr., R. H. Haas, D. Deering, J. Schell, and J. C. Harlan, ''Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation,'' NASA/GSFCT, Washington, DC, USA, Final Rep., 1974. [Online]. Available: https://ntrs.nasa.gov/ api/citations/19750020419/downloads/19750020419.pdf
[20] E. R. Hunt, P. C. Doraiswamy, J. E. Mcmurtrey, C. S. T. Daughtry, E. M. Perry, and B. Akhmedov, ''A visible band index for remote sensing leaf chlorophyll content at the canopy scale,'' Int. J. Appl. Earth Observ. Geoinf., vol. 21, pp. 103–112, Apr. 2013.
[21] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, ''Overview of the radiometric and biophysical performance of the MODIS vegetation indices,'' Remote Sens. Environ., vol. 83, nos. 1–2, pp. 195–213, Nov. 2002.

[22] L. Vilar, S. Herrera, E. Tafur-García, M. Yebra, J. Martínez-Vega, P. Echavarría, and M. P. Martín, ''Modelling wildfire occurrence at regional scale from land use/cover and climate change scenarios,'' Environ. Model. Softw., vol. 145, Nov. 2021, Art. no. 105200.

[23] A. D. Syphard, T. Sheehan, H. Rustigian-Romsos, and K. Ferschweiler, ''Mapping future fire probability under climate change: Does vegetation matter?'' PLoS ONE, vol. 13, no. 8, Aug. 2018, Art. no. e0201680.

[24] J. Zhai, Z. Ning, R. Dahal, and S. Yang, ''Wildfire susceptibility of land use and topographic features in the western United States: Implications for the landscape management,'' Forests, vol. 14, no. 4, p. 807, Apr. 2023.

[25] I. Shamta and B. E. Demir, ''Development of a deep learning-based surveillance system for forest fire detection and monitoring using UAV,'' PLoS ONE, vol. 19, no. 3, Mar. 2024, Art. no. e0299058.

[26] T. Kanand, G. Kemper, R. König, and H. Kemper, ''Wildfire detection and disaster monitoring system using uas and sensor fusion technologies,'' Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci., vol. XLIII-B3, pp. 1671–1675, Aug. 2020.

[27] M. T. Rashid, Y. Zhang, D. Zhang, and D. Wang, ''CompDrone: Towards integrated computational model and social drone based wildfire monitoring,'' in Proc. 16th Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS), May 2020, pp. 43–50.

[28] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ''YOLOv4: Optimal speed and accuracy of object detection,'' 2020, arXiv:2004.10934.

[29] A. M. G. Barros and J. M. C. Pereira, ''Wildfire selectivity for land cover type: Does size matter?'' PLoS ONE, vol. 9, no. 1, Jan. 2014, Art. no. e84760.

[30] A. A. Deshmukh, S. D. B. Sonar, R. V. Ingole, R. Agrawal, C. Dhule, and N. C. Morris, ''Satellite image segmentation for forest fire risk detection using Gaussian mixture models,'' in Proc. 2nd Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC), May 2023, pp. 806–811.

[31] H. J. Miller, ''Tobler's first law and spatial analysis,'' Ann. Assoc. Amer. Geographers, vol. 94, no. 2, pp. 284–289, Jun. 2004.

[32] N. Cressie, Statistics for Spatial Data. Hoboken, NJ, USA: Wiley, 2015.

[33] A.Fotheringham,C.Brunsdon,andM.Charlton,GeographicallyWeighted Regression: The Analysis of Spatially Varying Relationships. Hoboken, NJ, USA: Wiley, 2003. [Online]. Available: https://books.google.com.br/books?id=9DZgV1vXOuMC

[34] D. A. Griffith, ''Spatially autoregressive models,'' in International Encyclopedia of Human Geography, R. Kitchin and N. Thrift, Eds., Amsterdam, The Netherlands: Elsevier, 2009, pp. 396–402.

[35] R. Kamal, Microcontrollers: Architecture, Programming, Interfacing and System Design. Chennai, India: Pearson, 2011.

[36] A. Mairaj, A. I. Baba, and A. Y. Javaid, ''Application specific drone simulators: Recent advances and challenges,'' Simul. Model. Pract. Theory, vol. 94, pp. 100–117, Jul. 2019.

[37] S. Shah, D. Dey, C. Lovett, and A. Kapoor, ''AirSim: High-fidelity visual and physical simulation for autonomous vehicles,'' in Field and Service Robotics, M. Hutter and R. Siegwart, Eds., Cham, Switzerland: Springer, 2018, pp. 621–635.