Autonomous Drone Systems for Low-Cost Early Wildfire Detection and Tactical Response

Anish Kamatam

October 6, 2025

Abstract

Wildfires pose an escalating global threat, causing billions of dollars in damages annually and devastating ecosystems and communities. Existing detection methods, such as satellite imaging and ground-based sensors, often suffer from limited resolution, high costs, and delayed response times, leaving critical gaps in early intervention. In this work, we present an autonomous drone system that leverages onboard edge AI, thermal and RGB vision, LiDAR-based 3D modeling, and GPS mapping to provide real-time wildfire detection and risk assessment at a fraction of the cost of current solutions. Our design emphasizes affordability and scalability, achieving flight times exceeding 20 minutes on lightweight platforms. By integrating computer vision models with thermal-RGB sensor fusion, the system differentiates between true fire events and common false positives such as vehicles or animals. The addition of LiDAR enhances spatial awareness, enabling 3D terrain reconstruction and structural fire assessment, which supports tactical response planning. Field evaluations demonstrate accurate fire localization with low latency, enabling rapid alerting to first responders. We show that this approach improves detection accuracy while significantly reducing operational cost, providing a scalable solution for climate-resilient wildfire management.

1 Introduction

Wildfires have emerged as one of the most destructive natural disasters of the 21st century, fueled by rising global temperatures, prolonged droughts, and rapid urban expansion into fire-prone areas. In 2024, the United States experienced a sharp escalation in wildfire activity, with 64,897 reported wildfires burning 8.9 million acres, more than triple the 2.7 million acres recorded in 2023 and well above the ten-year average [1]. This surge far exceeded recent averages, with seven of ten U.S. regions experiencing above-normal fire activity. The destruction was likewise unprecedented, as over 4,500 structures were lost to wildfires in 2024, alongside untold ecological damage and economic impact. This escalating wildfire risk is driven by a dangerous confluence of factors, climate warming, prolonged drought, and decades of fuel accumulation, compounded by growing development in the wildland–urban interface [2]. Hotter, drier conditions and dense vegetation have created a landscape where small ignitions can rapidly explode into large conflagrations. There is an urgent need for improved wildfire detection and response systems to address this crisis.

Early detection is widely recognized as one of the most crucial determinants of wildfire outcomes[3]. Even a few minutes' delay can mean the difference between swift containment and an uncontrollable disaster. For instance, during California's 2020 Apple Fire (33,000 acres), camera footage revealed the blaze growing from ignition to a full "apocalyptic" inferno in under 40 minutes [3]. Such extreme fire behavior, increasingly common under climate-stressed conditions, leaves a narrow window for intervention. Faster detection enables a shift from reactive evacuation and structure defense to proactive attack on nascent fires[3]. Consequently, there is intense research interest in technologies that can spot wildfires at the earliest possible stage, before they rage out of control. Current state-of-the-art detection approaches, however, face significant limitations in terms of latency, coverage, and cost, which this work aims to overcome.

Satellite-based fire detection is a cornerstone of global wildfire monitoring, exemplified by systems like NASA's MODIS and VIIRS sensors aboard Earth observation satellites. These thermal imaging instruments continuously scan for hot spots and have provided near-real-time fire data for decades [4]. Satellites offer unmatched broad-area coverage, but their inherent constraints lead to important detection gaps. The spatial resolution of MODIS (1 km per pixel active fire product) and even the newer VIIRS ($\tilde{3}75$ m) means that small or incipient fires often go undetected until they grow larger [4][5]. Under typical conditions, MODIS can only reliably detect fires around 1000 m^2 in size (roughly 0.25 acre), with very best-case thresholds on the order of 100 m^2[5]. Moreover, closed forest canopies and heavy smoke can obscure thermal signatures, further reducing satellite detection rates for early-stage wildfires [4]. A recent performance study in Turkey found that MODIS identified as little as 0.6–16% of small (¡1 ha) fires, and VIIRS only 1.3–25%, when compared to ground-truthed fire records [4]. Temporal latency is another concern: polar-orbiting satellites like Terra/Aqua (MODIS) overpass a given location at most a few times per day, and even with NOAA's near-real-time processing, fire alerts typically come with a lag of up to 1–3 hours after ignition [4]. Geostationary platforms (e.g. GOES) improve refresh rates to minutes, but at the cost of coarse spatial resolution ($\tilde{2}$ km) that can still miss small ignitions. In summary, satellite systems remain invaluable for large-scale fire tracking, but their coverage comes at the expense of detection delay and a detection threshold too high for the smallest, most nascent fires.

On the other end of the spectrum, ground-based wildfire detection has traditionally relied on human observers, from fire lookout towers to 911 phone reports, which is a method inherently prone to delay and incomplete coverage[4]. A fire often must grow large enough to produce visible smoke or flame before it is noticed, by which time critical minutes or hours have passed. In recent years, more automated terrestrial systems have emerged. Notably, camera networks augmented with AI are now scanning high-risk landscapes for the first sign of smoke. For example, California's ALERTCalifornia program operates over 1,100 mountaintop cameras streaming live video to an AI system that flags potential smoke plumes for human verification[4]. This augmentation allows a few analysts to effectively monitor thousands of feeds, as the AI filters out false positives and highlights true anomalies in real time[4]. Commercial startups like

Figure 1: This image shows CV-enabled smoke detection by Pano AI, where computer vision models identify wildfire smoke plumes in real time using bounding boxes for localization.

Pano AI have similarly deployed tower-mounted 360° cameras with computer vision, achieving impressive reductions in detection times. In one 2023 incident in Washington state, a CV enabled camera system alerted fire agencies within minutes of ignition and helped get resources on scene 20–30 minutes faster than conventional means[4]. These examples demonstrate that ground-based CV models can significantly improve latency and provide rich visual situational awareness (smoke plume location, size, growth rate) to dispatchers. However, coverage and scalability remain limiting factors for terrestrial systems. Camera networks require substantial infrastructure (towers with power and communications), and they are typically focused on the wildland–urban interface and other accessible areas. Vast remote backcountry or wilderness areas may have no line-of-sight coverage. Even where cameras exist, complex terrain or weather (fog, darkness) can hinder visibility. In short, while ground AI cameras are a promising step forward, they cannot yet provide a low-cost early detection grid across all wildfire-prone landscapes.

In light of these limitations, attention is turning to autonomous Unmanned Aerial Vehicles (UAVs) as a platform for next-generation wildfire detection. UAVs combine the mobility of an aircraft with the potential for targeted, high-resolution sensing on demand. Unlike static cameras, drones can patrol vast areas and reposition dynamically to investigate triggers or high-risk zones. Compared to satellites, they fly at low altitudes (tens to hundreds of meters), yielding fine-grained imagery and thermal data and the ability to detect much smaller heat sources. Recent studies highlight that drones equipped with thermal and optical sensors offer a cost-effective alternative to satellite surveillance, especially for remote or inaccessible terrain [6]. By operating closer to the ground, UAVs can capture high-detail images with improved spatial and temporal resolution, essentially bridging the scale gap between orbital and ground observations. They can also venture into areas unsafe for humans, for example, to scout behind a ridge or at night, thus making fire monitoring safer and more comprehensive. Notably, thermal-infrared cameras onboard UAVs enable detection of heat signatures through smoke and in darkness, when optical cameras alone would fail[7]. This multi-modal sensing (thermal + RGB) means a drone can potentially spot a smoldering ember or hotspot before flames are even visible, addressing scenarios of early-stage or obscured fires. Furthermore, the addition of LiDAR sensing on UAVs opens a new dimension in wildfire monitoring. LiDAR can generate precise 3D maps of the terrain and vegetation in real time, which is invaluable for situational awareness. A UAV-mounted LiDAR provides immediate geometric context – for instance, mapping the local topography and fuel structure around a detected fire – which can help predict how the fire might spread and identify natural firebreaks or vulnerable structures[7]. This 3D modeling capability augments traditional 2D imagery with depth information, effectively enabling an on-the-fly "digital twin" of the environment that fire managers can use for decision support. Such rich data (thermal hot spots, visual imagery, and 3D terrain) traditionally required separate satellite passes and ground surveys to compile, but an agile drone can gather it in one flight.

2 Methodology

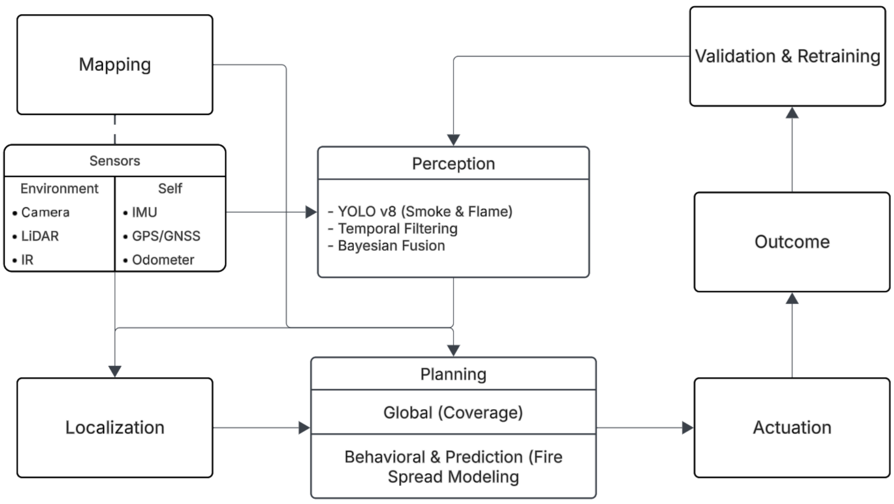

Figure 2: High-Level Software Architecture of AGNI X (2025). The system integrates multi-modal sensing, Bayesian perception, geo-spatial localization, predictive planning, and a validation-driven learning loop for autonomous wildfire detection.

2.1 Overview

AGNI X is a fully autonomous aerial intelligence system engineered for early wildfire detection, localized confirmation, and continuous model self-improvement. Its architecture draws inspiration from autonomous vehicle (AV) software stacks, adopting layered modules for perception, localization, planning, and control, but extends them with thermal–visual fusion, spatio-temporal filtering, and a data-driven validation loop tailored for unstructured wildfire environments (Figure 2). The methodology below details each subsystem and the mathematical principles unifying them into a closed-loop, self-learning platform.

2.2 Mapping & Contextual Data

A Mapping Database forms the foundation of AGNI X's situational awareness. It integrates:

- Digital Elevation Models (DEMs) for terrain contouring and line-of-sight estimation,
- No-Fly Zones (NFZs) for regulatory compliance,
- Vegetation and fuel maps for spread modeling, and
- Historical fire and wind datasets for risk weighting.

Mapping provides *a priori* context to both Localization and Global Planning. DEM height fields refine georeferencing by constraining projection rays from the camera onto terrain. Simultaneously, the planner consumes map layers to avoid restricted airspace and prioritize high-risk regions.

Formally, each grid cell g_i in the patrol map carries a weight:

$$w_i = f(v_i, s_i, r_i)$$

where:

- v_i encodes vegetation density,
- s_i represents terrain slope, and
- r_i denotes historical ignition rate.

Real-time LiDAR elevation maps are fused with preloaded DEM tiles to continuously refine terrain accuracy, ensuring centimeter-level correspondence between global coordinates and local features. The planner minimizes total weighted distance under flight constraints.

2.3 Sensor Suite

AGNI X employs a heterogeneous set of sensors that enable it to perceive both the external environment and its own motion state.

The environmental sensors include an RGB camera, a long-wave infrared (LWIR) thermal imager, and a LiDAR unit. Together, they provide complementary visual, thermal, and geometric cues essential for detecting active fire regions. The RGB camera identifies smoke and flame textures, while the thermal imager isolates heat plumes and hotspots through radiometric signatures. The onboard LiDAR contributes an additional three-dimensional context, generating local terrain meshes that enhance elevation awareness, slope estimation, and canopy structure understanding. This allows AGNI X to distinguish surface fires from elevated heat sources and maintain accurate line-of-sight projections over complex terrain.

The self-state sensor, comprising an inertial measurement unit (IMU), GPS/GNSS receiver, barometer, and odometer, continuously estimate the drone's orientation, position, and altitude. These measurements support stable flight control, precise geo-referencing, and drift correction during GPS degradation or occlusion.

All sensor modalities are hardware-time-synchronized, ensuring sub-30 ms latency across data streams. A multi-sensor ROS 2 pipeline manages acquisition, buffering, and timestamp alignment, guaranteeing that each perception frame integrates spatially and temporally consistent information. LiDAR point clouds also assist in obstacle mapping and height correction for geo-projection, enabling centimeter-level localization when fused with preloaded digital elevation maps (DEMs).

2.4 Localization

Localization estimates the drone's six-degree-of-freedom (6-DoF) state vector:

$$x_t = \begin{bmatrix} x & y & z & \phi & \theta & \psi \end{bmatrix}^T$$

using a visual–inertial Extended Kalman Filter (EKF). Inputs include IMU angular rates, GPS positions, and visual odometry from image features.

The prediction and update steps are defined as follows:

$$\hat{x}_{t|t-1} = f(\hat{x}_{t-1|t-1}, u_t)$$

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(z_t - h(\hat{x}_{t|t-1}))$$

The resulting geo-pose enables precise projection of detection pixels into world coordinates. During GPS loss, the fusion of inertial and visual cues maintains relative positional stability for several seconds, sufficient for local alerting and short-term navigation continuity.

2.5 Perception Layer

The perception layer translates multi-modal sensor data into actionable fire hypotheses in real time. It processes synchronized RGB, thermal, and LiDAR cues to identify smoke, flame, and heat sources under varying atmospheric conditions.

2.5.1 Detection

An RGB head based on YOLOv8-S identifies smoke and flame texture cues in visual imagery, while a complementary thermal head isolates high-temperature clusters in LWIR frames. Both models execute on an NVIDIA Jetson Orin Nano, an embedded GPU platform optimized for edge inference. It sustains real-time dual-stream (RGB + thermal) detection at 20–25 FPS within a 15 W power envelope, meeting AGNI X's onboard latency and energy constraints.. This configuration sustains 20–25 frames per second dual-stream inference, ensuring sub-second end-to-end latency between image capture and detection while remaining flight-ready for small-form-factor UAVs.

2.5.2 Temporal Filtering

Frame-to-frame jitter and flicker are suppressed using temporal filtering, which smooths detection confidence values over time. We apply an exponential moving average (EMA):

$$\hat{c}_t = \alpha c_t + (1 - \alpha)\hat{c}_{t-1} \tag{1}$$

where:

- c_t = instantaneous detection confidence at frame t,
- $\hat{c}_t$ = smoothed confidence after filtering,
- $\alpha \in [0, 1]$ = weighting factor controlling responsiveness (higher values react faster, lower values smooth more).

A detection is validated only if it persists above a confidence threshold for multiple consecutive frames, filtering out short-lived false positives from reflections or transient heat sources.

2.5.3 Bayesian Sensor Fusion

To integrate complementary cues from the visible and infrared spectra, AGNI X employs probabilistic late fusion. If p_{rgb} and p_{ir} represent the independent detection confidences from the RGB and thermal models, respectively, the fused fire likelihood is computed as:

$$p_{\text{fused}} = 1 - (1 - p_{\text{rgb}})(1 - p_{\text{ir}}) \tag{2}$$

where:

- p_{rgb} = probability of fire presence inferred from the RGB detector,
- p_{ir} = probability of fire presence inferred from the thermal detector,

- p_{fused} = joint fused probability assuming conditional independence.

This formulation emphasizes consensus across modalities: a confident signal in either RGB or IR increases the overall fused likelihood, while disagreements are moderated by multiplicative attenuation. The result is a robust, smoke- and lighting-invariant confidence score that feeds into the localization and planning layers.

2.6 Outcome Generation

Upon high-confidence verification, the system packages an *Outcome Event* containing the following metadata:

- Coordinates (latitude, longitude, altitude)
- Confidence score and modality
- Timestamp and detection ID
- RGB/IR thumbnails

Alerts are serialized into `GeoJSON` and `CAP` formats and uplinked via LTE/5G. Operators receive a live map ping with accompanying media evidence, enabling rapid human confirmation and situational awareness.

2.7 Validation and Retraining

All outcomes, true and false, enter a *Validation & Retraining* pipeline. Confirmed fire events reinforce positive samples, while false alarms enrich the negative dataset (e.g., fog, dust, sunlight).

A semi-supervised active learning loop ranks uncertain clips for annotation. Retrained YOLO and thermal heads are versioned in an MLOps registry; successful models are deployed over-the-air (OTA) during maintenance windows.

This loop ensures that AGNI X continuously adapts to new terrains, sensor noise profiles, and atmospheric variations, maintaining operational robustness across environments.

2.8 Continuous Feedback Loop

AGNI X embodies a closed, learning-centric autonomy cycle:

1. Mapping defines terrain context.
2. Sensors capture environmental and self-state data.
3. Perception fuses multi-modal evidence into actionable detections.
4. Localization & Planning translate detections into flight behaviors.
5. Control & Actuation execute maneuvers and generate outcomes.
6. Validation & Retraining analyze outcomes to refine models.

This persistent feedback enables lifelong learning, progressively reducing detection latency, false positives, and coverage gaps across successive deployments.

3 Conclusion

This paper presented the design methodology and system architecture for AGNI X, an autonomous aerial platform for real-time wildfire detection. By combining multi-modal sensing, probabilistic sensor fusion, and a closed-loop validation pipeline, AGNI X establishes a foundation for robust and adaptive fire monitoring in challenging environments. Although the platform is still under development, the proposed architecture demonstrates how lightweight UAVs equipped with embedded AI can bridge critical gaps left by satellites and ground-based systems.

Future work will focus on hardware integration and flight testing, including validation of perception accuracy, localization precision, and system endurance under real-world wildfire conditions. Additional directions include extending AGNI X into multi-drone swarms, integrating predictive fire spread models, and refining the continuous retraining pipeline with region-specific datasets. Ultimately, AGNI X aims to provide first responders with an affordable, scalable, and intelligent tool for rapid wildfire detection and response.

4 References

1. National Interagency Coordination Center, *Wildland Fire Summary and Statistics: Annual Report 2024*. Boise, ID: U.S. Department of the Interior, U.S. Department of Agriculture, and U.S. Department of Commerce, 2024.

2. Synolakis CE, Karagiannis GM. Wildfire risk management in the era of climate change. PNAS Nexus. 2024 May 7;3(5):pgae151. doi: 10.1093/pnasnexus/pgae151. PMID: 38715728; PMCID: PMC11075647.

3. "Artificial intelligence detects fires early, protecting people and infrastructure," *Civil Engineering Source*, American Society of Civil Engineers, Nov. 14, 2024. [Online]. Available: https://www.asce.org/publications-and-news/civil-engineering-source/article/2024/11/14/artificial-intelligence-detects-fires-early-protecting-people-infrastructure.

4. K. A. Coskuner, "Assessing the performance of MODIS and VIIRS active fire products in the monitoring of wildfires: a case study in Turkey," *iForest – Biogeosciences and Forestry*, vol. 15, no. 2, pp. 85–94, Mar. 2022.

5. NASA Earthdata, "FIRMS FAQ," *Earthdata*, [Online].

6. D. Meimetis, "An Architecture for Early Wildfire Detection and Spread Estimation," *SCIEpublish*, 2024.

7. M. R. Ahmed, A. J. L. Yang, H. Rahman, M. M. Rahman, and J. Park, "Drone-Based Early Wildfire Detection Using Multisensor Fusion and Edge Computing," *Fire*, vol. 7, no. 12, p. 443, Dec. 2024. doi: 10.3390/fire7120443. [Online].