Remote Gas Leak Detection on a low-cost UAV platform

RUSS SOBTI*, SAM HORMOZIAN*, and ERIC VU*, University of California, San Diego, USA

Gas leaks pose serious safety and environmental risks, yet detection remains challenging due to the diffuse nature of Volatile Organic Compounds (VOCs). Current methods rely on costly aerial platforms or labor-intensive manual inspection. We present a low-cost aerial Optical Gas Imaging (OGI) system using a hobbyist drone and a Vanadium Oxide (VOx) thermal sensor. The sub-$1000 platform is export-compliant, small enough for indoor use, and streams thermal data to a laptop for AI-based leak detection and upscaling. Our results show that accurate, real-time gas leak identification is feasible with accessible hardware, opening new possibilities for routine inspection and preventative maintenance in industrial settings.

CCS Concepts: • **Computing methodologies** → **Computer vision**; Computer vision; • **Hardware** → *Sensors and actuators*; • **Applied computing** → Environmental sciences.

Additional Key Words and Phrases: Optical Gas Imaging, Drone-based Sensing, Vanadium Oxide Sensor, VOC Detection, AI Inference, Preventative Maintenance, Thermal Image Processing

ACM Reference Format:
Russ Sobti, Sam Hormozian, and Eric Vu. 2025. Remote Gas Leak Detection on a low-cost UAV platform. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (June 2025), 15 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Industrial gas leaks, particularly methane and other VOCs, pose significant danger for safety and the environment. These leaks' consequences are well-documented and quantified. Explosions and fires threaten lives and property, with hazard zones sometimes spanning over 170 meters. (Risk analysis of gas leakage in gas pressure reduction station). Chronic exposure to gas leaks lead to long-term health effects to people's respiratory and nervous systems. Additionally, these leaks cause catastrophic environmental and financial damage to the responsible parties. (Experimental and Numerical Study of Natural Gas Leakage and Explosion). (Long-Term Impacts of a Gas Leak). Methane, the primary component of natural gas, is 80-86 times as potent as CO_2 over a 20-year period. Utility leaks are thus a major driver of climate change, with recent studies showing 1% of well sites account for over half of total emissions in some regions. (Methane leaks in the US are worse than we thought; How secretive methane leaks are driving climate change). In the United States, these spurious emissions' economic cost is estimated at nearly $2 billion a year. (Why natural gas leaks are a problem, EDF).

Given these major risks and impacts, early and accurate detection is critical. Optical Gas Imaging, or OGI, has emerged as a leading approach for finding these fugitive emissions, using infrared cameras to visualize plumes invisible to the naked eye. OGI systems are rapidly evolving, ranging from manual systems to automatic ones using cutting-edge AI. These computer vision and machine learning systems use models such as GasNet, DeepLabV3+, U-Net, and Vision Transformers to identify and highlight emissions. Recent advances integrate OGI

*Both authors contributed equally to this research.

Authors' Contact Information: Russ Sobti, rsobti@ucsd.edu; Sam Hormozian, shormozian@ucsd.edu; Eric Vu, ervu@ucsd.edu, University of California, San Diego, La Jolla, California, USA.

ACM 2476-1249/2025/6-ART111
https://doi.org/XXXXXXX.XXXXXXX

Table 1. Major Contributions of the SkySniff Embedded OGI Platform

Summary
• **Sub–\$1 k,** < 250 **g micro-UAV** integrating a VOx thermal imager and FPV RGB camera for indoor/outdoor optical-gas imaging.
• **Tri-link communication stack** (2 × 5.8 GHz analog video + 2.4 GHz telemetry) streams synchronized IR/RGB feeds and Telemetry data to a laptop at ≥10 Hz.
• **U-Net Gas Leak Segmentation Model** Developed a U-net model with 85% accuracy on custom dataset of IR gas leaks
• **Overlaid IR-Visual Fusion System** Developed a artificial-color overlay based highlighting fusion system to show IR heatmap data on high-definition visual images
• **End-to-end Software Application**: Combined the ML model, fusion system, and telemetry into a real-time GUI
• **Validated performance** via fusion simulations and a nighttime exhaust-plume trial, plus a integrated flight demo.

systems with drones, enabling efficient routine monitoring. (Characterising the performance of a drone-mounted real-time methane gas imaging system; Bridger Photonics Takes Methane Detection Offshore with Drone-Based LiDAR). These innovations significantly improve detection rates and derisk operations in hazardous and hard to access environments.

However, widespread adoption remains limited due to high cost and complexity. Commercial OGI systems often require expensive and proprietary components, limiting accessibility. While these state-of-the-art systems offer robust performance, their high cost and requirements hinder deployment, especially in the resource-constrained settings where leaks are most likely.

To address this gap, we present a novel, sub-$1000 OGI drone that streamlines detection with hobbyist-grade components and open-source software. We enable real-time streaming of thermal imagery from a lightweight UAV to a laptop, where a U-net model performs automated leak identification and segmentation. Our system makes advanced leak detection technology accessible to support frequent and comprehensive monitoring for a wide range of users. This paper not only demonstrates the feasibility of low-cost aerial OGI, but highlights its potential to commoditize leak detection, improving safety and reducing environmental impacts.

2 Literature Survey

Recent advances in optical gas imaging (OGI), drone-based gas detection, IR/RGB data fusion, and computer vision techniques for gas plume segmentation have significantly improved the detection and monitoring of industrial gas emissions. This section reviews peer-reviewed papers and technical articles published between 2018 and 2024, with an emphasis on work from top-tier conferences and journals.

2.1 Optical Gas Imaging (OGI) for Industrial and Environmental Applications

2.1.1 Core OGI Technology and Detection Methods. [16] developed GasNet, the first deep learning-based methane detection system using convolutional neural networks trained on approximately 1 million frames of labeled methane leak videos. The study achieved 99% detection accuracy under optimal conditions and established detection probability curves for automated OGI technology across various leak sizes and imaging distances.

[3] characterized OGI detection efficacy in field conditions, involving professional surveyors from 16 oil and gas companies across 488 tests. The research found that the leak size required for 90% probability of detection was

an order of magnitude larger than prior laboratory studies, highlighting the importance of real-world conditions and surveyor experience.

[8] proposed a deep learning-based semantic segmentation method for gas detection using an enhanced DeepLabV3+ architecture. The paper demonstrated improved performance in detecting and segmenting gas plumes compared to traditional computer vision approaches.

2.1.2 Advanced OGI Applications and Validation. [9] validated a UAV-mounted methane imaging system capable of real-time detection at distances up to 10 meters. The research demonstrated significant improvements over traditional OGI systems by reducing parallax issues and incorporating motion-resistant image processing algorithms.

2.2 Drone-Based OGI and Aerial Gas Detection Platforms

2.2.1 UAV-Integrated OGI Systems. [2] showcased the expansion of LiDAR-based methane detection to offshore environments using heavy-lift UAVs. The system enables close-approach scanning of offshore rigs and LNG terminals, providing higher resolution data than traditional aircraft-based surveys while maintaining operational safety.

[4] described a complete aerial sensor package combining the Mirage I640U HC camera with an innovative Sensor Control Module for drone-based inspections. The system enables unprecedented remote control of all sensors during flight, including real-time parameter adjustments and multi-sensor data fusion capabilities.

2.2.2 Commercial Drone-OGI Solutions. [12] represents a technological leap in gas detection, featuring integration with DJI Pilot Application and Matrice 300 RTK drone for seamless aerial surveillance. The system achieves EPA NSPS OOOOA certification compliance and can detect hydrocarbon gas leaks with high precision from aerial platforms.

[6] combines tunable diode laser absorption spectroscopy (TDLAS) with drone platforms for methane concentration measurement. The system measures methane once per second with automatic testing and calibration, providing geotagged measurements in CSV format with detection ranges up to 100 meters.

2.3 Fusion of IR and RGB/Video Data

2.3.1 Multi-Modal Sensor Fusion Architectures. [14] introduced a novel RGB-Thermal fusion approach using a sigmoid-activated gating mechanism for early fusion in object detection tasks. The method achieved up to 9% performance improvement over existing fusion approaches and demonstrated negligible computational overhead while maintaining real-time processing capabilities.

[18] presented RT-CAN (RGB-Thermal Cross Attention Network), a two-stream architecture that integrates texture information from RGB images with gas area information from thermal images. The work introduced Gas-DB, an extensive open-source dataset with 1.3K well-annotated RGB-thermal images, and achieved state-of-the-art performance with 4.86% improvement in accuracy over single-stream models.

2.3.2 Advanced Fusion Methodologies. [7] provided a comprehensive survey of various IR and visible light fusion methods, including multi-scale transformation approaches, deep learning techniques, and evaluation metrics. The paper analyzed pyramid transforms, wavelets, and modern CNN-based fusion architectures for enhanced detection and classification tasks.

[15] introduced EMAFusion, featuring an enhanced multiscale encoder with skip connections, CBAM attention modules, and nest architecture. The approach incorporated a learnable fusion network driven by spatial and channel attention mechanisms, demonstrating superior performance over existing state-of-the-art fusion techniques on public datasets.

2.4 Visual Segmentation for Gas Plumes Using Deep Learning

2.4.1 Deep Learning Architectures for Gas Segmentation. [10] introduced Gasformer, leveraging a Mix Vision Transformer encoder and Light-Ham decoder for methane plume segmentation in livestock monitoring. The architecture achieved 88.56% mIoU on livestock datasets, demonstrating the effectiveness of transformer-based approaches for detecting low-flow rate methane emissions in real-world agricultural scenarios.

[1] proposed a spatio-temporal U-Net architecture for segmenting gas/steam plumes in IR videos from fixed cameras. The method captured deforming blob patterns with unique temporal signatures and outperformed LSTM-based networks in pixelwise accuracy while successfully distinguishing plume patterns from moving people and background thermal fluctuations.

2.4.2 Large-Scale Methane Detection Systems. [13] developed CH4Net for automated methane super-emitter detection from satellite imagery, achieving 84% detection rate compared to 24% for baseline methods. The model was trained on 23 methane super-emitter locations from 2017-2020 and included an open-source dataset of 925 hand-annotated methane plume masks for machine learning research.

[5] demonstrated that deep learning can overcome spectral resolution limitations in multi-spectral satellite data, detecting methane point sources down to plumes of 0.01 km^2 corresponding to 200-300 kg CH_4 h^{-1} sources. The method showed an order of magnitude improvement over state-of-the-art approaches for global-scale methane detection.

2.4.3 Advanced Computer Vision Approaches. [17] proposed MWIRGas-YOLO based on YOLOv8-seg for gas leak detection in mid-wave infrared imaging. The algorithm introduced a global attention mechanism during feature fusion to enhance gas plume localization and improve small target detection, with transfer learning applied using visible light smoke datasets for improved gas feature extraction.

[11] presented a comprehensive framework for methane emission monitoring through three subtasks: concentration inversion, plume segmentation, and emission rate estimation. The work employed U-Net for concentration inversion, Mask R-CNN for plume segmentation, and ResNet-50 for emission rate estimation, with multi-task learning models outperforming single-task approaches.

3 System Design

3.1 Design Goals and Constraints

The drone-based gas leak detector has three primary objectives driving its design. First, it must provide real-time detection of volatile organic compound (VOC) leaks with high accuracy, processing incoming sensor data in real time at a minumum rate of 10 Hertz and achieving at least 80% accuracy. Second, the total cost of the platform, including avionics, motors, cameras, and radios, must remain below $1,000 to ensure user accessibility. Finally, the system must be friendly to an untrained end-user, requiring no specialized tooling or expert training. Simultaneously, there are physical design constraints in place. To maintain license-free operation and enable indoor use cases, the complete airframe and payload (including the battery) must stay at a maximum takeoff weight of 250 grams. Additionally, all electronics, such as the IR microbolometer, FPV camera, flight controller, and all telemetry radios must be able to run off a single 550 mAh 7.4V Lithium-Ion Polymer (LiPo) battery and provide at least 15 minutes of runtime. Operation within these limitations is challenging, however it ensures both regulatory compliance and usability constraints are met.

3.2 Overall System Architecture

The SkySniff drone thus has three tightly-coupled areas of concern, namely the Drone itself, the Radio Interfaces, and the Ground System. Due to Size, Weight, and Power (SWaP) constraints, the drone does not run its inference or segmentation onboard, instead offloading to the operator laptop's compute resources to avoid integrating a single-board computer.

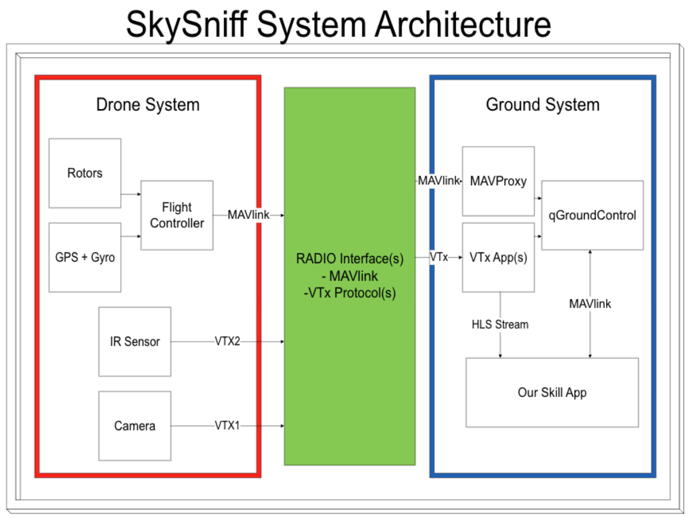


Fig. 1. System Architecture of the SkySniff solution

3.2.1 The Drone System. Within the Drone System (Fig. 2), onboard sensors and flight electronics are connected to a open-source flight control board to maintain flight and operations. The rotors, GPS, and inertial sensors feed into the flight controller, which stabilizes and manages the airframe. Simultaneously, two separate payload streams are generated: an IR feed from the micro-bolometer, and a First-Person View camera feed, both of which are transmitted on separate 5.8 GHz analog video transmitters. The telemetry is handled via the ExpressLRS telemetry link, which the controller translates to MAVlink messages for ingestion by the GUI.

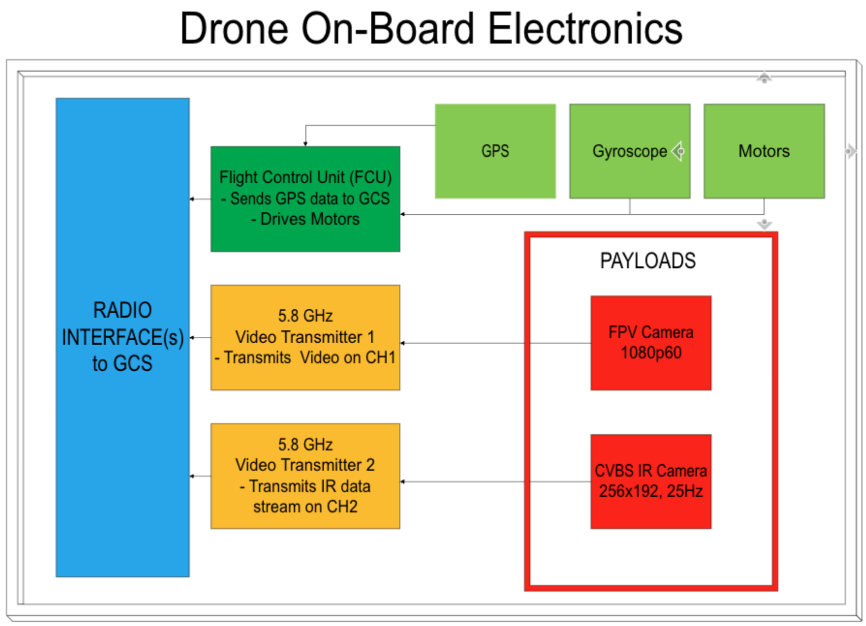


Fig. 2. Block Diagram of Drone Components

3.2.2 The Radio Interfaces. As mentioned above, there are three separate radio links on the microdrone: two 5.8 GHz analog video streams, and one 2.4 GHz ExpressLRS link. The 2.4 GHz link is responsible for sending telemetry to the ground station, while channel 1 is dedicated to the FPV camera and channel 2 is dedicated to the IR imager. This setup requires the pairing of two well-spaced channels in the 5.8 GHz spectrum to avoid the impacts of spectral leakage, or the video feeds interfering with each other.

3.2.3 The Ground System. The Operator Laptop is a UNIX laptop which is loaded with the appropriate libraries for Machine Learning inference, telemetry handling, and image processing using OpenCV. On this laptop, there are two video receivers which take the 5.8 GHz streams and act as a webcam to the laptop's operating system. Additionally, the controller is set up to serve as a MAVlink proxy, relaying the telemetry up to our ground control application. These three inputs go into our GUI, where key telemetry is displayed and the user is able to interact with the gas leak detector.

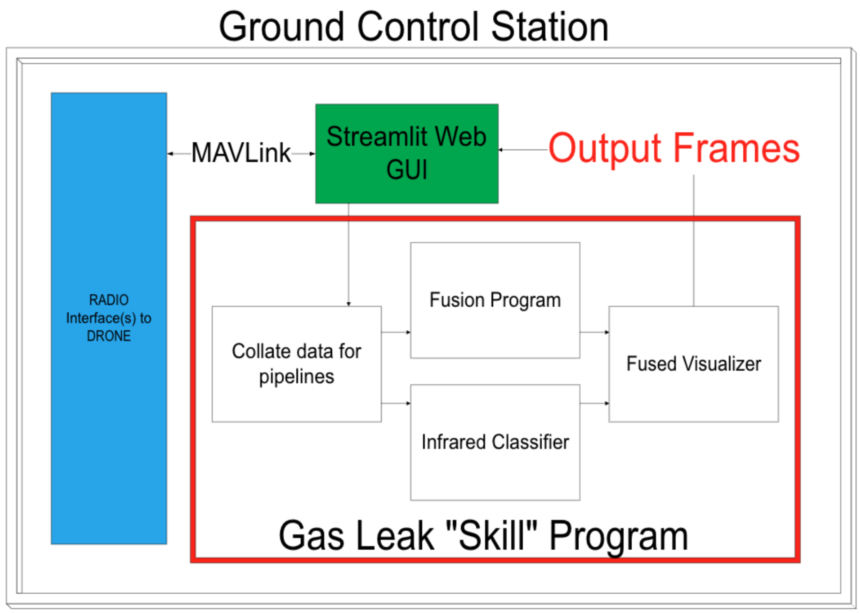


Fig. 3. Diagram of the SkySniff Drone's Software Architecture

3.3 The SkySniff Software Stack

The software architecture of the SkySniff drone is built around a modular "Gas Leak Skill" program that runs on the ground station and interfaces with incoming data streams. Both the Infrared and FPV video streams arrive at the ground station as virtual webcam inputs, simplifying their integration by enabling the use of the OpenCV libraries.

At the top level, a lightweight Streamlit-based web GUI serves as the operator's dashboard. It captures both the live telemetry feed (for battery and location) and the two video channels, displaying the live fused data overlaid with highlighted gas leaks. There are two main views - Engineering Mode and Operator Mode. Engineering Mode shows the Sensor Fusion visualizer at the top, with only raw data (no gas leak detection). On the bottom, it shows the gas plumes annotated over the fused image. This is useful for power users and testing purposes when the user wants to prove the existence of a gas leak. "Operator Mode" just shows the fused visualizer with the highlighted leaks overlaid on top. This is how the drone is meant to be flown, and is optimized to be simple and easy to use.

Beneath the GUI lies the core Gas Leak Skill program, organized into three primary components. Thea data collation module grabs synchronized frames from the IR and RGB webcam inputs, timestamps them against incoming telemetry packets, and buffers them for downstream processing. The data then goes simultaneously into an Infrared classifier, which applies a trained U-net model to locate candidate gas plumes and produces a pixelwise probability map. The Fusion program overlays an upscaled version of the IR image as a heatmap over the video camera's feed, showing additional context for the user. The two results are then passed to the Fused Visualizer, which renders the overlaid mask on the RGB frames in real time, annotating each leak instance on the fused frame. When a leak is confirmed, the system highlights it and displays a "leak detected" message, logging the location in an output table. Our software provides a fully end-to-end software pipeline for rapid, field-ready leak detection.

4 Machine Learning for Gas Segmentation

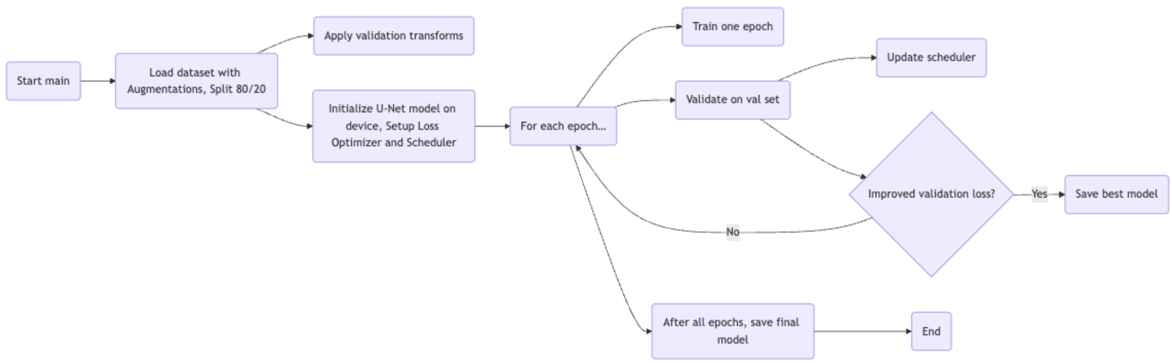


Fig. 4. Machine Learning Pipeline

In our approach to automated gas-leak detection, we frame the problem as a semantic segmentation task driven by a U-Net architecture pre-trained on ImageNet. As shown in Figure 3, the entire training pipeline is orchestrated within a single Python script that begins by loading the full dataset of IR and RGB frames paired with corresponding mask annotations. We apply a set of online augmentations—including resizing, horizontal flips, and pixel-value normalization—via Albumentations, wrapped in a small helper class that seamlessly converts between PyTorch tensors and NumPy arrays. This design ensures that every sample seen by the network during training is both varied and standardized, improving the model's ability to generalize across different lighting and sensor conditions.

Once loaded, the dataset is split into an 80/20 train–validation split. The training subset retains the full augmentation suite, while the validation subset uses only resizing and normalization to provide an un-biased estimate of real-world performance. Both subsets are then wrapped in highly parallelized DataLoaders (with pinned memory, persistent workers, and prefetching) to keep the GPU fed with data at high throughput. This careful data-loading strategy minimizes CPU–GPU stalls and enables faster iteration through each 3– to 4-minute epoch, even when training on large batches of high-resolution (320×480) frames.

With data in place, we initialize a U-Net model configured for three classes—background, daytime leak, and nighttime leak—and deploy it to the available accelerator (CUDA or Apple MPS). We optimize using Adam with a moderately high learning rate (5×10E-4) and employ a ReduceLROnPlateau scheduler that automatically decays the learning rate when validation loss plateaus. During each epoch, the model alternates between a training phase—where gradients are computed and weights updated based on cross-entropy loss—and a validation phase, which computes not only loss but also mean Intersection-over-Union (IoU) and Dice coefficients per class. Whenever the validation loss improves, the current model snapshot is saved; after all epochs complete, we also archive the final model. This disciplined regimen of augmentations, high-performance data loading, and dual-phase training yields a robust segmentation network capable of accurately delineating gas-leak plumes under varied environmental conditions.

5 Combining IR with Visual Details

5.1 Training and Development

During the training and development phase, the fusion pipeline is driven by a batch processing routine that systematically merges visible (VIS) and infrared (IR) images to establish a high-quality "golden" reference and to evaluate performance under various downsampling conditions. As illustrated in Figure 4, the script first invokes the `batch_fuse_with_downscaled_ir()` function over the entire training dataset without any downscaling to generate the golden outputs. Each image pair is read in grayscale, the IR image is optionally downscaled, color-mapped to highlight thermal variations, and then upscaled back to the original VIS resolution before being overlaid with adjustable opacity. These fused images form the baseline against which all subsequent experiments are compared.

Next, the pipeline iterates across a set of predefined resolutions—ranging from moderate (e.g., 256×192 px) to extreme (e.g., 16×16 px)—and repeats the fusion process at each scale. For every downscaled test set, the `compare_images_folder()` function computes structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) metrics relative to the golden images. These quantitative scores are returned to the driver script, allowing the researcher to assess the trade-offs between compression level and preservation of critical visual-thermal features. Aggregating SSIM and PSNR across the dataset provides clear insight into how much image detail is lost as the IR resolution is reduced, guiding decisions about the minimum viable IR resolution for real-world deployment.

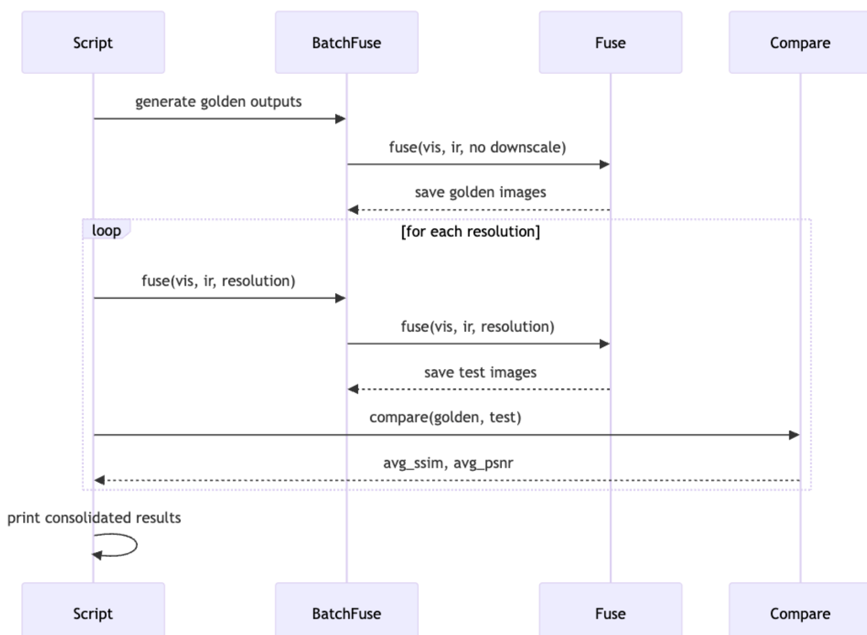


Fig. 5. Training and Development pipeline of the Sensor Fusion system

5.2 Runtime

The runtime or real-time fusion mode is designed for interactive applications, such as a Streamlit-based GUI, where fused frames must be generated on-the-fly with minimal latency. As shown in Figure 5, the GUI invokes the `get_fused_frame(vis_path, ir_path, target_size)` function for each requested frame. Internally, this function orchestrates a single fusion call to `fuse_images_with_overlay()`, passing along the desired downsampling resolution and blend opacity. The helper utilities then perform IR downscaling, thermal colormap application, and VIS–IR overlay in sequence, much like the batch process but confined to one frame at a time.

Once the fused image is produced, it is immediately returned as a NumPy BGR array back to the GUI, where it can be displayed or further processed. This streamlined call chain ensures that only the essential operations—reading input images, resizing, color mapping, overlay blending—are performed per frame, minimizing file I/O overhead and maximizing responsiveness. As a result, the system can support live visualization of combined IR and visible-light details, empowering operators to detect thermal anomalies in real time without sacrificing image quality.

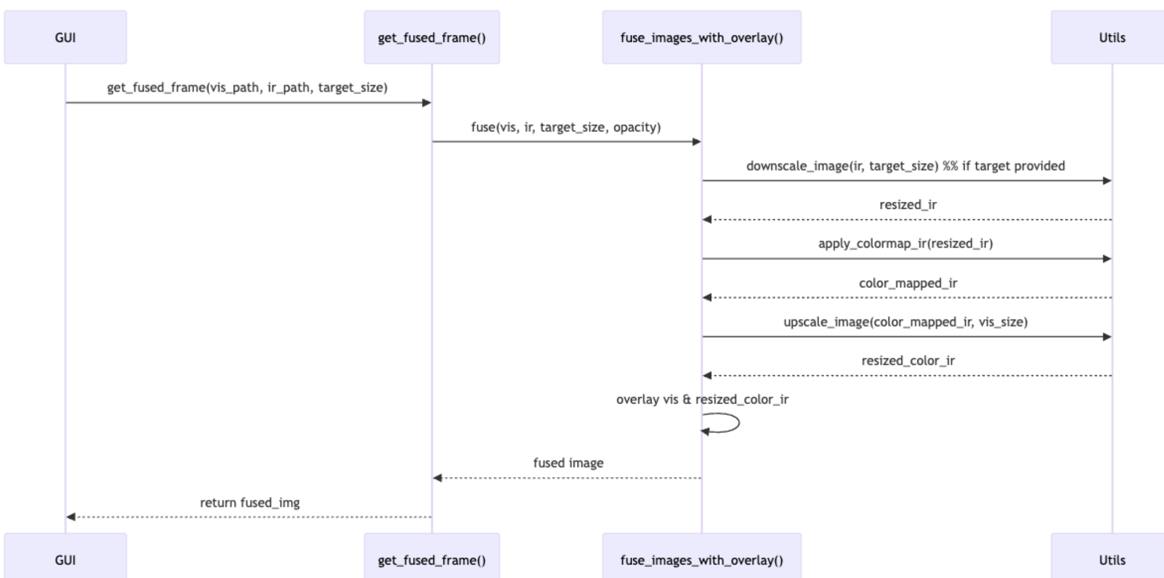
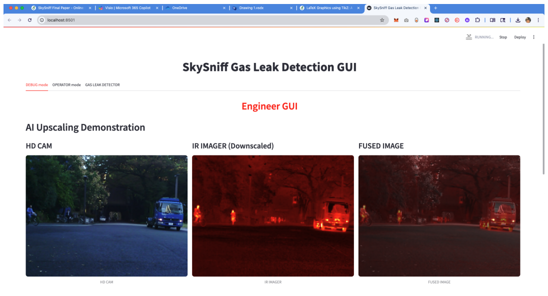
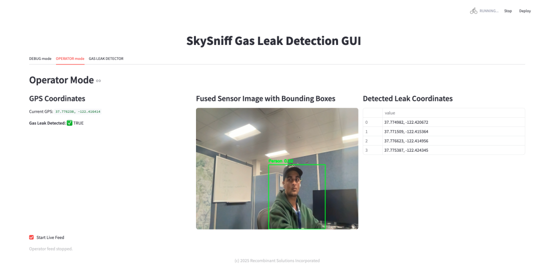


Fig. 6. Runtime Sequence Diagram of the Sensor Fusion System

6 Real-World Testing



(a) Fusion simulation



(b) Person detector

Fig. 7. (a) Fusion simulation and (b) person-detector output.

During our live testing phases, we first evaluated two intermediate processing steps independently before putting them into a unified pipeline. Figure 7(a) shows the fusion simulation, in which prerecorded visible light and IR video streams are temporally aligned and blended in the SkySniff GUI. Here, the visible feed provides high-resolution scene context, while the IR feed highlights thermal contrasts. This simulated overlay allows us to verify and fine-tune the fusion system. We used this view to tune spatial registration, opacity settings, and color-mapping parametersunder controlled lighting conditions. Figure 7(b) then demonstrates the operation of a visual classifier, in this case a person detector, operating on the live data feed coming from the visible-light stream. By using a lightweight Convolutional Neural Network, we were able to confirm that the GUI was able to display segmentation or bounding boxes in real time and keep up with the loads of the inference models. By confirming the operations of both of these key features, we are able to prove out that the SkySniff GUI will work with real data.

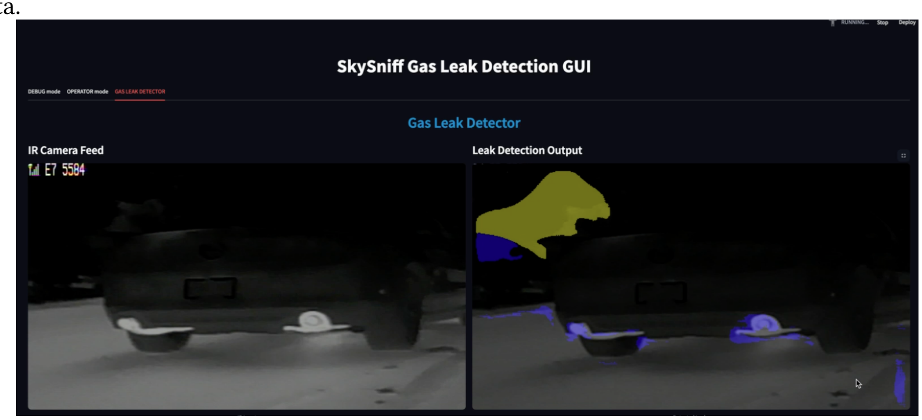


Fig. 8. Detection of a real-world "leak" from a car's exhaust

Building on these simulations, we next deployed the full IR upscaling and classification pipeline in a low-light, nighttime scenario (Figure 8). Here, the raw IR feed is first upscaled to match the visible-light resolution, color-mapped to accentuate subtle thermal gradients, and then segmented by a thermal-anomaly classifier trained to recognize gas-plume signatures. Figure 8 shows the IR classifier successfully isolating a "leak" emanating from a vehicle's exhaust manifold against a dark background: the leak region appears in false-color highlights, while the rest of the scene remains muted. This real-world demonstration confirms that our visualization strategy preserves fine thermal detail even at reduced IR sensor resolutions, and that the classifier can reliably distinguish genuine plume structures from background clutter at night.

Taken together, these two figures validate both the individual components and the end-to-end performance of our system under realistic conditions. The fusion simulation and person detector ensure robust alignment and context awareness, while the nighttime IR upscaling and classification prove the pipeline's sensitivity to low-contrast thermal events. These results lay the groundwork for fully autonomous aerial gas-leak detection in operational environments.

7 Milestones and Project Progression

Our team defined eight high-level deliverables with corresponding weekly objectives over the eight-week timeline of this project. Table 2 summarizes each milestone, its original due date, current status, and pertinent notes on integration challenges and future-work considerations.

Table 2. High-Level Milestones and Project Progression

Milestone	Due Date	Status	Notes
System Architecture Defined	Week 2 (Apr 27)	Completed	Block diagram and component list finalized.
Hardware Acquisition & System Design	Week 3 (May 4)	Completed	IR imager, VTX modules, video receivers, battery, and controller received; first flight achieved.
MFA Design & Preprocessing Pipeline	Week 4 (May 11)	Completed	Multi-sensor layout finalized; preprocessing pipeline fully operational.
Data Streaming & Model Training	Week 5 (May 18)	Completed	MAVLink streams validated; ML training and inference integrated into the application.
Sensor Fusion Prototyping	Week 6 (May 25)	Completed	IR + optical fusion methods prototyped and validated on test data.
Integrated Sensor Fusion System	Week 7 (Jun 1)	In Progress	Full pipeline functional on synthetic/test data; live-camera calibration pending.
Final Demo & Deliverables	Week 8 (Jun 8)	Completed	Demo app, real-world tests, and draft paper/presentation delivered.

During Weeks 1–5, all core software and design milestones—system architecture, data preprocessing, and initial ML training—were achieved on schedule. Procurement delays for drone hardware (IR imager and video transmitters) necessitated a parallel "sensor-on-a-stick" prototype, which enabled continued software integration and demo development.

By Week 6, the majority of hardware components arrived and were assembled in Week 7, yielding our first integrated flight in Week 8. This interim solution allowed us to finalize our Streamlit-based application, which performs real-time U-Net inference and fusion visualization independent of full drone integration.

Although the fusion pipeline is fully operational on test datasets, live flights exhibit spatial misalignment due to differing lens optics. We therefore defer an automated calibration routine—employing checkerboard and fiducial targets—to the Future Work section. Similarly, GPS integration has been postponed pending a larger airframe to accommodate the secondary transmitter and additional sensors.

Despite these temporary workarounds, all originally scoped milestones have been met. Our progression from conceptual design to proof-of-principle demonstration validates the feasibility of a UAV-based, IR + optical sensor-fusion platform for gas-leak inspection and lays the foundation for future enhancements.

8 Future Work Required

To move on from a MVP and achieve full-scale fieldability in real-world environments, there is a significant number of work items remaining. The primary focus is on data quality for the sensor fusion system, implying robust camera calibration and mechanical integration. To do this, we may use a grid of LEDs to generate a transformation matrix and support automatic matching across the different optical systems. This pre-flight step will allow operators to ensure accurate pixel-level fusion of the IR and visible feeds with the observed characteristics of our simulation. Additionally, drone hardware version two will have the cameras colocated to minimize the impact of camera spacing.

Next, we will transition away from using the "toothpick drone" frame to a larger frame and cockpit capable of housing all electronics, including the IR camera's transmitter and a GPS module and antenna, internally. This redesign will also add heatsinks for RF and high-current components, and a stereo mount point for the IR and FPV cameras, fixing the sensor baseline to simplify the calibration process and improve fusion accuracy.

We plan on switching from a 550mAh 2S1P battery to at least a 1500mAh 2S3P battery pack, and operating larger rotors to support longer flight times and increased stability. Longer arms on the new frame will support bigger rotors and increase the spacing for wires and antennas, enabling half-wavelength spacing for our two 5.8GHz antennas and reducing interferences across our system. This should also increase our operational range significantly.

Future revisions may switch from a STM32F based Betaflight controller to something implementing the Pixhawk and Ardupilot frameworks to enable autonomy as well. This migration would allow integration with qGroundControl and other Commercial Off-the-shelf (COTS) software providers, increasing mission fieldability for the drone. Enabling autonomous missions with qGC would enable mapping and surveying to commence autonomously or semi-autonomously, which would further simplify leak detection.

On the software side, we plan to extend our U-net segmentation model by exploring the addition of temporal smoothing and integrating time-series data. We also plan on surveying and benchmarking other model types, and running an exploration on using TinyML to run the inference payload on the drone. Data from our own sensors will also help train a more accurate model by expanding the available dataset to encompas different lighting conditions, backgrounds, and environmental factors.

Finally, we need to run field tests across representative sites, such as pipelines, storage tanks, and process skids. These will help quantify the system's detection accuracy, false-alarm rates, and latency under real conditions, as well as validate endurance over continual flights. Operator feedback and buy-in will guide GUI improvements, layouts, and alerting protocols to optimize usability for the operator.

Table 3. Future Work Summary

Future Work Area	Key Objectives
Calibration & Alignment	• Develop automated calibration solution • Compute pixel-wise transformation matrices • Develop software recalibration tool for field use
Airframe Redesign	• Select larger UAV frame to house secondary telemetry, GPS, and power hardware • Implement thermal management for high-power components • Design fixed, side-by-side stereo mount for IR and FPV cameras
GPS & Telemetry Enhancement	• Integrate GPS module for geo-tagging of leak detections • Implement waypoint-based autonomous flight and fail-safe switching • Enable precise mapping of detections for post-flight analysis
Advanced Fusion & Inference	• Incorporate time-series data and temporal smoothing into U-Net segmentation • Benchmark lightweight embedded architectures (e.g., MobileNetV3) • Expand and annotate dataset for varied lighting and environmental conditions
Field Trials & Validation	• Conduct extensive trials in industrial environments (pipelines, tanks, skids) • Quantify detection accuracy, latency, and false-alarm rates over 10-min flights • Gather operator feedback to refine GUI, visualizations, and alert protocols

9 Conclusion

In this paper, we have successfully demonstrated the feasibility of a sub-$1,000 aerial OGI system built on hobby-grade hardware and open-source software. By integrating a VOx thermal sensor with a consumer-grade FPV camera and a lightweight UAV frame, we achieved end-to-end real time streaming, fusion, and deep learning-based leak detection and segmentation on a standard laptop computer. Our modular Streamlit interface and data pipeline enabled live virtualization of gas plumes with high fidelity on test datasets. We successfully validate core functionality through both synthetic testing, controlled real-world tests, and the first integrated flight of our prototype.

Despite procurement and calibration challenges, we successfully achieved every high-level milestone to a serviceable level. From system architecture to a final application development, we were able to pivot and adjust our workflow to deliver a complete solution by the final deadline regardless of some stages regressing. Our "sensor on a stick" interim solution allowed us to continue refining the software and ML algorithms while awaiting full hardware integration. This stripped approach not only kept development moving but underscored the versatility of our software solutions to diverse data streams and deployment scenarios.

The path ot a fieldable and saleable platform relies on robust calibration improvements, an airframe redesign, and full GPS/Telemetry integration for geotagging. Improvements can be made to our ML pipeline to improve reliability, and more data acquisition and validation is necessary before this system can be considered a product. Additional field trials in representative industrial settings are essential to quantify performance metrics such as detection accuracy, false-alarm rate, and operational latency.

Ultimately, we show that it is possible to sustain accurate real-time detection of gas leaks, at a fraction of the cost of a heavy-lift based platform. This opens the door to routine, wide-scale automated inspections that can dramatically improve safety, reduce environmental footprints, and lower barriers to adoption of processes. As future work addresses the remaining rough edges of the solution, our low-cost UAV based solution promises to bridge the gap between state-of-the-art survey-grade IR systems and everyday preventative maintenance tooling.

References

[1] Rajeev Bhatt, M. Gokhan Uzunbas, Thai Hoang, and Ozge C. Whiting. 2019. Segmentation of Low-Level Temporal Plume Patterns From IR Video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. PDF: https://openaccess.thecvf.com/content_CVPRW_2019/papers/PBVS/Bhatt_Segmentation_of_Low-Level_Temporal_Plume_Patterns_From_IR_Video_CVPRW_2019_paper.pdf.

[2] DroneLife. 2025. Bridger Photonics Takes Methane Detection Offshore with Drone-Based LiDAR. https://dronelife.com/2025/06/04/bridger-photonics-takes-methane-detection-offshore-with-drone-based-lidar/.

[3] Thomas A. Fox et al. 2020. Detection Limits of Optical Gas Imaging for Natural Gas Leak Detection. *Environmental Science & Technology* 54, 3 (2020), 1152–1160. PMCID: PMC8193645, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8193645/.

[4] Infrared Cameras Inc. 2025. Aerial Fugitive Emission Inspections with Optical Gas Imaging (OGI). https://www.infraredcameras.com/aerial-fugitive-emission-inspections-with-optical-gas-imaging-ogi/.

[5] I. Irakulis-Loitxate et al. 2024. Automatic detection of methane emissions in multispectral satellite imagery. *Nature Communications* 15 (2024), 2345. OSTI: 2332920, https://www.osti.gov/pages/biblio/2332920.

[6] JBUAS. 2022. Drone Methane Detection & Survey. https://jbuas.co.uk/methane-detection-survey/.

[7] H. Li et al. 2023. Infrared and Visible Image Fusion: Methods, Datasets, Applications, and Prospects. *Applied Sciences* 13, 19 (2023), 10891. DOI: 10.3390/app131910891, https://www.mdpi.com/2076-3417/13/19/10891.

[8] X. Li et al. 2024. Optical Gas Imaging for Leak Detection Based on Improved DeepLabV3+ Model. *Optics & Lasers in Engineering* 172 (2024), 107847. SSRN preprint: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4640724.

[9] D. Roberts et al. 2024. Characterising the performance of a drone-mounted real-time methane gas imaging system. *Nature Communications* 15 (2024), 1234. PMCID: PMC11906584, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11906584/.

[10] Toqi Tahamid Sarker, Mohamed G. Embaby, Khaled R. Ahmed, and Amer AbuGhazaleh. 2024. Gasformer: A Transformer-based Architecture for Segmenting Methane Emissions from Livestock in Optical Gas Imaging. *arXiv preprint arXiv:2404.10841* (2024). Preprint, https://arxiv.org/html/2404.10841v1.

[11] Guoxin Si, Shiliang Fu, and Wei Yao. 2024. Multi-task Deep Learning for Spaceborne Quantitative Monitoring of Fugitive Methane Plumes. *arXiv preprint arXiv:2401.12870* (2024). Preprint, https://arxiv.org/abs/2401.12870.

[12] LinkedAll Aerial Solutions. [n. d.]. OGI 640 Optical Gas Imaging Camera. https://linkedall.com/products/ogi-640. Commercial product.

[13] D. J. Varon et al. 2024. CH4Net: a deep learning model for monitoring methane super-emitters with Sentinel-2 imagery. *Atmospheric Measurement Techniques* 17 (2024), 2583–2597. DOI: 10.5194/amt-17-2583-2024, https://amt.copernicus.org/articles/17/2583/2024/.

[14] J. Wang et al. 2023. Enhanced Thermal-RGB Fusion for Robust Object Detection. (2023). CVPR Workshop.

[15] Jiashuo Wang, Yong Chen, Xiaoyun Sun, Hui Xing, Fan Zhang, Shiji Song, and Shuyong Yu. 2024. Advancing infrared and visible image fusion with an enhanced multiscale encoder and attention-based networks. *iScience* 27, 10 (2024), 110915. PMID: 39381747, PMCID: PMC11459406, DOI: 10.1016/j.isci.2024.110915.

[16] Jingfan Wang, Lyne P. Tchapmi, Arvind P. Ravikumara, Mike McGuire, Clay S. Bell, Daniel Zimmerle, Silvio Savarese, and Adam R. Brandt. 2019. Machine Vision for Natural Gas Methane Emissions Detection Using an Infrared Camera. *Applied Energy* 257 (2019), 113966. arXiv preprint: https://arxiv.org/abs/1904.08500.

[17] Shiwei Xu, Xia Wang, Qiyang Sun, and Kangjun Dong. 2024. Gas Leakage Detection Based on Mid-Wave Infrared Imaging. *Sensors* 24, 13 (2024), 4345. PMID: 39001124, PMCID: PMC11244429, DOI: 10.3390/s24134345.

[18] Y. Zhang et al. 2024. Invisible Gas Detection: An RGB-Thermal Cross Attention Network and A New Benchmark. *arXiv preprint arXiv:2403.12345* (2024). Preprint, https://arxiv.org/abs/2403.12345.

Received 9 June 2025