

Received 21 October 2022, accepted 4 November 2022, date of publication 17 November 2022, date of current version 23 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3222805

RESEARCH ARTICLE

Wildland Fire Detection and Monitoring Using a Drone-Collected RGB/IR Image Dataset

XIWEN CHEN¹, (Member, IEEE), BRYCE HOPKINS², (Member, IEEE),
HAO WANG¹, (Member, IEEE), LEO O'NEILL³,
FATEMEH AFGHAH², (Senior Member, IEEE),
ABOLFAZL RAZI¹, (Senior Member, IEEE), PETER FULÉ³, JANICE COEN⁴,
ERIC ROWELL⁵, AND ADAM WATTS⁶

¹School of Computing, Clemson University, Clemson, SC 29634, USA

²Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634, USA

³School of Forestry, Northern Arizona University, Flagstaff, AZ 84001, USA

⁴National Center for Atmospheric Research, Boulder, CO 80301, USA

⁵Desert Research Institute, Reno, NV 89512, USA

⁶U.S. Forest Service, Pacific Wildland Fire Science Laboratory, Seattle, WA 98103, USA

Corresponding author: Fatemeh Afghah (fafghah@clemson.edu)

This work was supported in part by the Air Force Office of Scientific Research under Award FA9550-20-1-0090; and in part by the National Science Foundation under Grant CNS-2232048, Grant CNS-2204445, Grant CNS-2038741, and Grant CNS-2038759.

ABSTRACT Current forest monitoring technologies including satellite remote sensing, manned/piloted aircraft, and observation towers leave uncertainties about a wildfire's extent, behavior, and conditions in the fire's near environment, particularly during its early growth. Rapid mapping and real-time fire monitoring can inform in-time intervention or management solutions to maximize beneficial fire outcomes. Drone systems' unique features of 3D mobility, low flight altitude, and fast and easy deployment make them a valuable tool for early detection and assessment of wildland fires, especially in remote forests that are not easily accessible by ground vehicles. In addition, the lack of abundant, well-annotated aerial datasets – in part due to unmanned aerial vehicles' (UAVs') flight restrictions during prescribed burns and wildfires – has limited research advances in reliable data-driven fire detection and modeling techniques. While existing wildland fire datasets often include either color or thermal fire images, here we present (1) a multi-modal UAV-collected dataset of dual-feed side-by-side videos including both RGB and thermal images of a prescribed fire in an open canopy pine forest in Northern Arizona and (2) a deep learning-based methodology for detecting fire and smoke pixels at accuracy much higher than the usual single-channel video feeds. The collected images are labeled to “fire” or “no-fire” frames by two human experts using side-by-side RGB and thermal images to determine the label. To provide context to the main dataset's aerial imagery, the included supplementary dataset provides a georeferenced pre-burn point cloud, an RGB orthomosaic, weather information, a burn plan, and other burn information. By using and expanding on this guide dataset, research can develop new data-driven fire detection, fire segmentation, and fire modeling techniques.

INDEX TERMS Data-driven fire detection, prescribed fire, fire modeling, fire data, unmanned aerial vehicle (UAV), deep learning.

I. INTRODUCTION

Despite marked improvements in wildfire response, rapid public reporting, shrinking intervals between coarse geostationary satellite observations (as little as 30 seconds in

4 km midlatitude pixels), and the distribution of visible spectrum smoke-observing cameras (e.g. Alert Wildfire [1]) that have markedly improved awareness that a wildland fire ignition has occurred (“wildfire detection”), uncertainty about their specific location and extent (“wildfire mapping”) and their environment remains. In addition to the coarse spatial resolution of high frequency observations, which limits

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval¹.

mapping, satellites and fire lookouts may be unable to detect low intensity fires, fire in its earliest stages, or through clouds, vegetation, or heavy smoke. While higher resolution polar orbiting platforms can provide more detailed delineation of a fire, these revisit a location less frequently, which can result in delayed observation by these sensors in the initial attack stages of a wildfire. Both fire science and management still have a pressing need to rapidly identify (through observation and interpretation of data) and map a fire's specific location, extent, and environment. Broader unmet needs among stakeholders, the public, and various aspects of scientific research include expanded, more available, higher temporal frequency observations of changing fire extent, intensity, and status in a wider range of atmospheric conditions (as flight restrictions due to time of day or weather impede intelligence gathering of some systems) than are currently available.

Various platforms with varying radiometric instrumentation and data analysis algorithms, resolution, and observational footprint have been proposed and tested to meet these needs. Examples include the moderate resolution geostationary satellites such as GOES 17 [2] and on demand fire mapping aircraft (e.g. the Colorado Multi-Mission Aircraft [3], the USDA Forest Service National Infrared Operations (NIROPS) [4], and FireGuard ([5])). Concurrently, wildland fire science has been developing models and datasets across a range of fire sizes, intensities, and type (e.g. prescribed burns and wildfires). Such datasets are also desired for model improvement and applications (e.g. initialization and validation of computational fluid dynamics-based (CFD) fire models, machine learning-type modeling, data assimilation or steering, observing phenomena, and distinguishing mechanisms of fire spread). However, the specific needed observation characteristics (resolution, temporal frequency, latency, spectral characteristics) are rarely simultaneously met and remain an active area of interdisciplinary research between engineering and wildland fire science.

Unmanned aerial systems (UAS) can offer fast and high-resolution mapping of wide areas. The unique features of drone systems in terms of 3D mobility, flexibility, low flight altitude (below the clouds), and fast and easy deployment make them a valuable tool for early detection and assessment of wildfires. Several UAS-based fire monitoring operations using both large [6] and small UASs [7], [8], [9], [10] have been deployed. Authors in [11] studied the solution to deploy low-cost hardware UAV platform with efficient onboard image processing algorithms.

With the recent advances in nanotechnologies-based semiconductors, the new generations of TPUs and GPUs can provide significant onboard computation capabilities for edge devices including drones. Therefore, TPU/GPU-equipped UAVs can perform on-the-fly detection of wildfires as well as onboard fire spread modeling. Emerging communication technologies such as LoRA and nano satellites as well as the advances in integrating cellular-connected drones in 5G/6G supports reliable utilization of drones in remote wild-fire management operations [12], [13]. Data-driven onboard

fire detection and modeling models rely on the availability of large-scale and multi-modal training datasets. However, due to safety concerns and regulation and flight restrictions related to UAVs' operation, in particular, during prescribed and wildfires, there are only few datasets of aerial fire images available. This fact leaves a lingering need to build up credibility in UAV reliability and safety in disaster relief operations, developing operation guidelines and collecting diverse datasets.

As technology has enabled more sophisticated measurements as well as their integration, a new value in coordinated datasets has been recognized. Efforts to enrich data collected over fires with observation metadata, contemporaneous measurements made by other platforms or observers near or above the fire line, and weather information are aimed at improving the quality of inputs to fire detection and modeling systems. While challenging due to the logistical and technical difficulties of merging measurements across disciplinary areas and matching sensor output with system input requirements, the wildland fire science and modeling communities have proposed ways to prioritize and achieve measurement coordination [14], [15], [16]. The present work is one example of a responsive effort to coordinate fire measurements, focused machine-learning inputs and other significant information into a coordinated fire library.

New generations of fire detection technology, modeling systems, and their underlying scientific understanding show remarkable advances, but discontinuities remain between instruments' outputs and present modeling systems' needs. Conversely, in some instances the conceptualization of measurements is initially not met with adequate sensors to provide needed measurements, although the response of some sensor development and deployment efforts to partially bridge these gaps has been encouraging [17], [18]. Finally, novel uses of sensors to gather valuable information relevant to fire science offer additional frontier-expanding examples [19] that are likely to increase with the proliferation of interdisciplinary teams converging on wildland fire applications.

Technological proliferation also will yield substantial benefits for measurements of fire ignition and behavior, and one example is the combination of small unmanned aircraft systems [20] and so-called "co-witnessed" Electro-Optical/Infrared (EO/IR) imaging payloads offered by several manufacturers. Simultaneous recording of visible-spectrum and longwave infrared provides a relatively high-resolution still or video image for purposes such as direct image georectification, orthomosaicking, and spectral or textural imagery analysis, while the infrared image component permits the tracking of the fire line, embers, and significant terrain features (as well as personnel) when smoke would obscure these features in visible-spectrum imagery. The combined imagery permits the rough position and movement of fire to be observed under most combustion, weather, and visibility conditions, providing relatively uninterrupted input to machine-learning or geospatial models that focus on spatial resolutions appropriate for the relatively small spatial

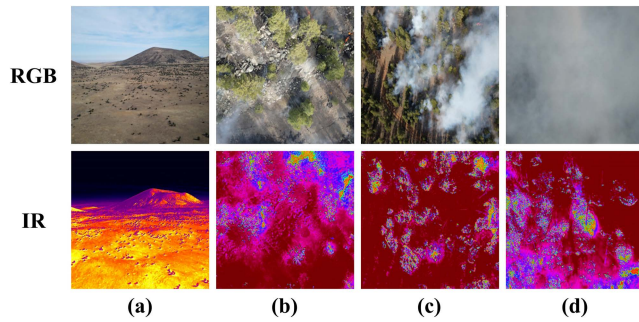


FIGURE 1. Handpicked frame pairs from FLAME 2 Dataset [21]. (a) No flame with No smoke, (b) flame with no smoke, (c)(d) flame with smoke.

coverage typically captured from low-altitude UAS (i.e., scenes or landscape units of a few ha in size or smaller).

We summarized our contributions as follow:

- We present a unique dataset of side-by-side RGB/IR imagery collected during a prescribed fire near Flagstaff, Arizona in 2021 [21]. The images are jointly labeled by two human experts with fire/no-fire and smoke/no-smoke¹ labels.
- In addition to aerial images, data on weather information, and georeferenced pre-burn point cloud data points are included in the dataset.
- We implement the DL-based benchmarks with RGB-thermal fusion for fast fire classification and demonstrate the comparative analysis.
- We also develop a hybrid method by cascading the DL-based classifier and a fire detection localization for frames labeled as fire by processing dual-feed imagery.

It is expected that the methods developed to produce the FLAME2 dataset and others to follow can facilitate fire detection and modeling, as well as fire management.

II. BACKGROUND AND RELATED WORKS

Wildfire researchers have assembled diverse datasets to both detect wildland fires (Section II-B) and initialize and/or validate a wide range of kinematic and dynamic fire behavior models applied across a wide range of fire types, sizes, and environments (Section II-C). Some prescribed fire experiments and observations made in each are listed in [22].

Observations may be collected in-situ from the ground or tower, or remotely from aircraft. In hindsight, data collection – guided by simpler needs of legacy models – often appears incomplete, overlooks a needed metric, or has inadequate characteristics (in spatial or temporal frequency) to evaluate more than gross fire characteristics and lacks the data density to constrain data-driven models. For unplanned landscape-scale fires encompassing larger temporal and spatial scales, observations from already in operation, satellite-based sensors have been more useful (e.g., [23]).

¹The smoke/no-smoke label refers to frames where the experts visually estimate that smoke fills greater than 50% of the RGB frame. See Fig 3 for more information on the labeling process.

However, satellite-based wildfire detection and mapping (such as datasets available through NASA's Fire Information for Resource Management System (FIRMS)) [24] that has the spatial resolution to resolve fire lines lack frequent updates (thus, 12-hourly observations do not clarify during which part of the day the growth occurred), whereas sensors that provide high frequency updates lack spatial granularity. A particular challenge for operational applications' on-the-fly model steering is that standard communication of observations can have latency of hours.

A. DRONE-COLLECTED FIRE DATASETS

UAV-based wildfire detection enables high precision and, when coupled with communication, real-time measurements. In order to train neural network (NN)-based UAV fire detection systems, there is a need for datasets of aerial imagery of wildfires and prescribed burns, optimally at high revisit rates. Table 1 provides a review of existing fire datasets. As shown in this table, while there are a plethora of publicly available fire detection-based datasets online, there are few datasets that provide aerial imagery of wildfires/prescribed burns [7]. More importantly, the existing UAV-collected fire image datasets often only include the color or thermal images. To our knowledge, there is no dataset of aerial images which provide side-by-side thermal and RGB images. Our proposed FLAME2 [21] aims to fill this gap by providing a collection of dual RGB/IR drone videos and images taken during a prescribed burn in northern Arizona in November 2021 (refer to Section III, Data Collection, for more information). Another key feature of this dataset is the potential input for fire spread and high resolution burn intensity/severity modeling. The key difference between FLAME2 and our previously proposed FLAME dataset [7] is using dual-channel side-by-side cameras in FLAME 2, while FLAME provides RGB/IR images though not side-by-side. While most existing datasets only include image samples, the FLAME2 supplemental dataset includes geo-referenced point clouds of the prescribed burn area, weather and burn conditions, and pre-burn aerial RGB videos, all of which serve to expand the dataset's potential applications beyond fire detection and masking. For example, previous work to fuse thermal and color images into high resolution, accurate 3-dimensional point clouds could use images collected for this project to create a burn intensity model or pre- and post- fire georeferenced temperature mapping [25], [26].

B. DEEP LEARNING-BASED WILDFIRE DETECTION

A series of convolutional neural network (CNN) models demonstrate an outstanding performance on classification, which is the most upstream task. These models further can be fine-tuned to a wide variety of downstream tasks such as object detection, semantic segmentation, and instance segmentation. Fire patterns can also be recognized by CNN-based networks, since the wildfire images are still natural images that have rich local covariance. The low-level features such as contours and edges of the fire can be extracted well

TABLE 1. Comparison of similar fire image datasets.

Dataset Name	Collection Method	RGB/IR	Perspective	Fire Area	Labeling Type	Num of Images	Year	Supplemental Data
DataCluster Labs' Fire and Smoke Dataset [27]	Cellphone videos	RGB	Terrestrial	Mixed ^a	Fire or Smoke/No Fire or smoke	7000+ frames	2021	N
AIDER ^b [28] [29]	Search Engines	RGB	Aerial	Mixed	Fire/No Fire	1000 images	2020	N
Dataset for Forest Fire Detection [30]	Search Engines	RGB	Mixed ^c	Rural	Fire/No Fire	1900 images	2020	N
Fire Detection by Dhruvil Shah [31]	Search Engines	RGB	Terrestrial	Mainly Urban	Fire/No Fire	3225 images	2020	N
FireNet [32]	Search Engines	RGB	Terrestrial	Mixed	Bounding Boxes	502 images	2019	N
Fire Detection From closed-circuit television (CCTV) [33]	CCTV	RGB	Terrestrial	Mixed	Fire/no Fire Smoke/no Smoke	864 frames	2019	N
Furg Fire Dataset [34] [35]	Search Engines	RGB	Terrestrial	Urban	Bounding Boxes	365,702 frames	2018	N
CAIR's Fire Detection Image Dataset ^d [36]	Search Engines	RGB	Terrestrial	Mainly Urban	Fire/No Fire	651 images	2017	N
Mivia's Fire Detection Dataset [37] [38]	CCTV	RGB	Terrestrial	Mixed	Fire/No Fire	62,690 frames	2014	N
FLAME ^e [7]	Drone Footage	Mixed ^f	Aerial	Rural Pile Burns	Fire/No Fire and Masking	47,992 frames	2020	Minimal ^g
FLAME 2^h [21]	Drone Footage	side-by-side Dual RGB/IR	Aerial	Rural Prescribed Burns	Fire/No Fire Smoke/No Smoke	53,451 frames	2022	Yes ⁱ

^a Fire Area "Mixed" refers to including at least two of the following categories: rural, urban, or other fires. "Rural" includes forest/wildland fires, "Urban" includes fires located near buildings or indoors, Other includes any other location (ex: dark undecipherable background, solid colored background, computer generated background).

^b AIDER - Aerial Image Dataset for Emergency Response Applications.

^c Perspective "Mixed" refers to having some aerial imagery and some terrestrial imagery.

^d CAIR - Centre for Artificial Intelligence Research.

^e FLAME - Fire Luminosity Airborne-based Machine learning Evaluation.

^f RGB/IR "Mixed" refers to having both RGB and IR though not side-by-side.

^g Burn area provided as well as human weather estimates.

^h FLAME 2 - Fire detection and modelLing: Aerial Multi-spectral image dataset.

ⁱ See Table 4.

by CNNs with translational invariant. At deeper layers, high-level (semantic) features will be learned from these low-level features from the beginning layers. Then, classification can be performed based on these extracted high-level features.

Wildland fire shapes are spatially and temporally dynamic and depend on factors such as wind, terrain, and fuel including state, amount, physical properties, and arrangement. Consequently, annotating new task data and training a model from scratch is often required when the model needs to be deployed on a new data/task. However, in most scenarios, these two operations are expensive and time consuming, and we expect to learn a new model with prior knowledge transferred from the related domain to accelerate and enhance the new model. This approach is named *Transfer Learning*. We often name the data from the related domain as source data, and the data from the current task as target data. The taxonomy of Transfer Learning can be defined based on the attributes of source data and target data: i) Homogeneous (feature spaces are related, data and labels have the same attributes), ii) Heterogeneous (feature spaces are not directly related or the attributes of

data/labels are not same). In wildfire detection, homogeneous Transfer Learning is used in many papers [39], [40]. The basic idea is fine-tuning the state-of-art CNN models (often very large), which are pre-trained on some large natural image datasets, such as ImageNet-1K (~1.28M images with 1000 classes). These pre-trained models are proven to have enough capacity to extract low-level and high-level representations from the imagery signals. After simply fine-tuning the classifier by the last fully-connected layers, the model can perform classification on the wildfire dataset. It is noteworthy that only parameters in the fine-tuned layers (which are few) require training, which can greatly improve efficiency with maintaining the strong ability of the large networks. For example, a naive MobileNetV2 [41] has 3.4M parameters, but in the fine-tuned model for binary classification only has 2.5K parameters are trainable. There are a large amount of studies that take advantage of Transfer Learning. For instance, [40] applies pre-trained Inception-v3 [42] while authors in [39] and [43] fine-tuned MobileNetV2 for fire/no fire binary classification.

Besides classification for fire images, some studies turn their attention to higher-level vision tasks. One paradigm would be real-time detection and localization of the fire/smoke area, where output is the bounding box. Many well-known object detection frameworks offer the opportunity to detect fire in real-time and diverse environments for surveillance videos. For example, You-only-look-once-(YOLO) series [44], [45] are employed in multiple works [46], [47], [48], in which single-shot multi-box detector (SSD) series [49] are applied in [50] and [51]. It is noteworthy that both recent versions of YOLO and SSD are one-stage detection methods but address the trade-off between accuracy and speed by applying well-designed efficient components. Moreover, some works are interested in fire area segmentation, which aims to extract the fire regions at pixel-level from the background. For instance, [52] and [7] take advantage of U-net, which is one of the most famous segmentation frameworks. The works [53] and [54] design a network that learns from both optical and thermal domains and performs the flame segmentation. These two works also refer to a topic known as *Multi-modality Learning*. They are interested in detecting fire by aggregating optical images and thermal infrared images since robust informativeness can be offered by the thermal images under bad visibility such as thick smoke, occlusion by clouds, high altitudes, other natural barriers, or due to camera resolution and imaging imperfectness. Specifically, the thermal infrared cameras can capture thermal radiation of flame according to the spectral variability in the range between mid-wavelength infrared and long-wavelength infrared regardless of optical occlusion [55].

Appropriate *Feature Fusion* at different levels is the key component to empower the additional achievement of the DL-based detection by the Multi-modality learning. For instance, authors in [53] use an encoder-decoder structure. Specifically, the features from two spectral ranges are combined at different levels in the encoder, and then these representations will be concatenated with the flowing feature maps of the decoder at different levels. Authors in [54] are interested in remote sensing data. They perform fusion at the beginning and extract the low-level features of the fused signal. Then they apply two streams to learn the different scales of these features and integrate them at the end of the network. The main challenge of using fusion models is the need for dual camera system which might not be present in most existing datasets.

As discussed earlier, one issue would be naive DL-based model requires heavy computation on training, which can be addressed by fine-tuning the pre-trained models [39], [43] to a low-complexity architecture. In addition to this issue, another is the lack of generalization, since the data domain shift is caused by many underlying characteristics such as camera angle, image scale, terrain, etc. This issue often results in catastrophic forgetting, meaning that if a model trains on an old fire detection task and then trains on a new fire detection task, the model would perform poorly on the old task (i.e.,

the model forgets the knowledge of the old task). Except training all data simultaneously, incremental/continuous learning provides a potential solution by guiding the model update based on the domain knowledge of two tasks. There are two approaches to achieving continuous learning. One branch of CL with privacy-preserving is known as the *regularization method*, which only imposes constraints on the weights regardless of any old data. For example, authors in [56] use Fisher information as their penalty criterion. The other branch is *replay*, which tries to reserve the old task knowledge by maintaining some representative data from the old task when training the model on a new task. For example, [57] takes advantage of knowledge distillation, which first chooses the important samples of the old tasks. Then, it employs a cross-entropy loss between the old model and the new model to ensure they have a similar mapping function. To the authors' knowledge, this paradigm should attract the attention of the community, but only limited works focus on it. This may be because of the lack of a standardized benchmark.

C. DATA-DRIVEN FIRE SPREAD MODELING

In this section, we first introduce the datasets used for data-driven fire modeling. Satellite active fire detection has characteristics that are desirable for data driven modeling, being routinely and globally collected, potentially providing the data density needed for data-driven and machine learning approaches. For example, broad properties about fire activity can be learned from products that compile fire activity over a day. Daily fire growth datasets include [58], [59]. A third is the "Next Day Wildfire Spread" dataset. The latter includes a total of 18,445 samples collected during 2012 to 2020 across the United States [60]. The dataset represents fire areas as 'fire masks' during consecutive days to show the fire spread. Besides the fire data, the dataset includes information on wind direction and speed, elevation, minimum and maximum temperatures, humidity, precipitation, drought index, normalized difference vegetation index. Another dataset is the "WildfireDB" which includes 17,820,834 datapoints of daily fire occurrences from 2012 to 2018. Each datapoint includes a polygon cell on fire in a time step of one day). The dataset also contains weather information, canopy and vegetation base density, height, cover, elevation and slope measured in 2016. Other datasets, like Monitoring Trends in Burn Severity (MTBS), [61], and [62] contain final fire perimeters. Historic wildland fire perimeters are delivered through [63].

Methods and datasets such as the FLAME2 example presented here are used to advance understanding of fire behavior and models developed to replicate it. Historically, a key predicted outcome of models is rate of spread of a landscape-scale fire line, the interface between burning and not-yet-ignited areas, particularly, the flaming front, which is the current leading, actively burning segment of the fire. Examples of specific use cases, based on the current state of understanding, follow:

- 1) A dataset can be used to provide regular updates in the evolving fire extent. While the FLAME2 dataset covers a small area, larger segments of fires could be mapped with coordinated multi-UAS activity. Currently, estimates of fire spread are based on temporally coarser satellite or aircraft revisits and an average expansion rate is calculated assuming the fire grows perpendicularly outward at each point to the current perimeter. When the time between observation frames is short, much more instantaneous estimates of rate of spread can be made, capturing more dynamic and more subtle aspects of fire spread such as transient bursts and slipping fire growth, where at points along a protrusion, fire propagation is outward, while winds briefly carry the fire sideways downwind. Thus, such datasets (e.g. [64], [65], [66]) allow better models to test better representation of transient aspects of fire growth (e.g. [67]).
- 2) High temporal and spatial resolution datasets can be used to improve fire models with detailed representation of fire physics, for example, revealing fire dynamics ([68]) and propagation mechanisms ([69], [70]) underlying rate of spread within wildland fires. Spatial heterogeneity in fuel properties can greatly influence (or interrupt) fire spread in small and/or low intensity fires. Spotting (the lofting and landing of burning embers, generating new fire ignitions outside the existing fire line) can be an important propagation mechanism but whether spots impact fire spread or are overrun is indeterminate. Mapping long distance spots is also important for forecasting, because where an individual spot will land is deterministically unpredictable.
- 3) Newer data-driven and machine learning modeling approaches (e.g. [10], [71], [72], [73], [74]) provide an alternative to ever more detailed physical models in predicting fire spread but often require many large datasets. These approaches are aided by high spatial and temporal resolution data and, compared to other data gathering platforms, UAV-based mapping is becoming (relatively) more technically easy.

III. DATA COLLECTION

The broadcast burn was located on the Kaibab National Forest, Arizona, USA, adjacent to the South Rim of Grand Canyon National Park. Imagery was collected November 19, 2021. The unit is composed of pinyon-juniper woodland and ponderosa pine forest. Objectives of the burn were to reduce surface fuel loading to $<2.2 \text{ kg/m}^2$ in pinyon-juniper and $0.67\text{--}2.2 \text{ kg/m}^2$ in ponderosa dominant areas, maintain 7 logs and 3–5 snags per hectare, and minimize large diameter gambel oak (*Quercus gambelii* Nutt.) and ponderosa pine (*Pinus ponderosa* Laws.) fire-related mortality. Hand ignitions with drip torches were used to burn the perimeter of the unit. Once perimeter burning was complete, a helicopter with plastic sphere dispenser capabilities ignited the interior of the unit. Fire intensity in the study area was primarily surface fire with 0.25–0.75 m flame lengths occasionally reaching

5–10 m when consuming large logs or ladder fuels. Spread rate was estimated at 300–600 m/hr. Severity was relatively low, burning litter and coarse woody debris with minimal spread in open areas, occasionally consuming the grass component when fuels were contiguous. In the pinyon-juniper fuel type, tree density and fuel loading were low, consisting of needle cast under trees and coarse woody debris. This created a spatially complex mosaic of burned area. In the ponderosa fuel type, understory fuel loading was higher and more homogeneous, enabling surface fire to spread between tree cover. Fire severity was also relatively low.

UAS-based 4K RGB and thermal infrared data were collected using a DJI Mavic 2 Enterprise Advanced system. These UAS are primed to collect simultaneous RGB and thermal imagery through two sensors that are integrated into a single platform. This system produces dual coincident imagery collection, though the RGB and thermal imagery differ slightly in resolution and domain. The thermal sensor on the Mavic Enterprise Advanced uses an uncooled vanadium oxide (VOx) microbolometer sensor that captures a 640×512 pixel array. The sensor can characterize temperatures from -40°C to 150°C for high gain image capture and from -40°C to 550°C for low gain image capture. Spectral ranges for the thermal imager are 8–14 μm . The 1/2" CMOS Visual Sensor captures passive RGB imagery with a maximum image size of 8000×6000 pixels. Video resolution is sampled at either 3840×2160 pixels at 30 frames per second (FPS) or 1920×1080 pixels at 30 FPS. For both systems, GPS location information are only carried through individual image capture that are exported in the DJI R-JPEG format. MP4 video formats used to record video carry no meta data related to the GPS location of the UAS at the time of data capture.

A DJI Matrice 200 drone with a Zenmuse X4S camera was used to collect pre-burn imagery of the site for photogrammetry. The X4S has a 20 megapixel, CMOS 1" sensor. Two videos were recorded at 3840×2160 pixels, 30 FPS, and 122 m and 61 m above ground, respectively for a total of 10 minutes.

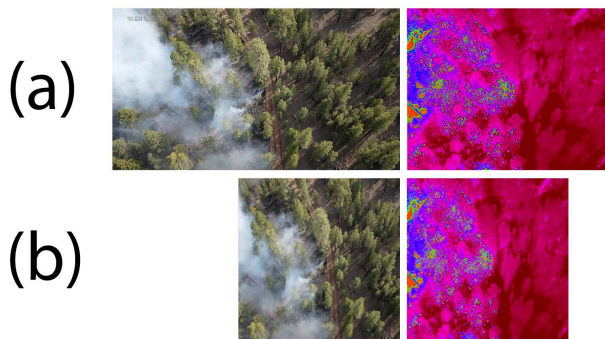
Several sampling challenges arose from measuring fire characteristics over an active prescribed fire. Primary considerations while working within a United States Forest Service active incident were maintaining active air to ground communications while aerial ignitions were underway. Due to the helicopter-based ignition for the prescribed fire, UAS air operations had to wait for ignition operations to be completed and the helicopter to leave the fire's air space, limiting our sampling to the perimeter of the fire that was ignited last. The dual imagery collected with the Mavic Enterprise Advanced was collected at 60 m limiting our focal view of the fire.

IV. DATASET DESCRIPTION

The FLAME2 Dataset collected during this prescribed burn consists of 53,451 dual RGB/IR frame pairs extracted from seven different RGB/IR video pairs. A reduced resolution set of frame pairs is included to aid in image processing.

TABLE 2. Description of the prescribed burn.

Date	11-19-2021
Coordinates (imagery collection area)	35.8794, -111.8561
Spot weather forecast	Temp: 14 – 15 °C, Relative Humidity: 18 – 20%, Wind: Southwest 4.5 m/s gusting to 9 m/s, Haines Index: 3
Burn objectives	Reduce surface fuel loading, maintain coarse woody debris, and minimize oak and ponderosa mortality
Outcome	Low severity, no torching, consumption of surface fuels, patchy fire spread estimated at 300-600 m/hr
Flame Length	0.25-0.75 m surface fires, reaching 5-10 m for large logs or ladder fuels

**FIGURE 2.** Depicts results of FOV corrections to Frame Pair 39340 from FLAME 2 [21] (randomly picked). The 254p × 254p pair have an approximate 1:1 correlation of the size and positions of elements in the frame.

A detailed breakdown of the main dataset's contents can be seen in Table 3.

A. IMAGE FRAME PREPARATION AND LABELING

The 254p frame pairs were created for the purpose of testing the dataset with various CNN's (as seen in Table 7). To improve CNN performance, the 254p RGB and IR frames were cropped and scaled such that they possess approximately the same field of view and level of zoom, with the aim of having a 1:1 correlation between RGB and IR frames. A sample frame pair exhibiting the 1:1 correlation can be seen below in Fig 2.

To achieve the approximate 1:1 correlation of the 254p frame pairs that is seen in Fig 2, the left and right sides of the 1.78:1 aspect ratio RGB videos were cropped to match the 1.25:1 aspect ratio IR videos. For the 4k RGB videos, 570 pixels were cropped from both the left and the right side of the frame and for the 2k RGB videos, 285 pixels were cropped from both sides. Next, both the RGB and thermal videos were scaled down to 254p × 254p, the desired input resolution for the CNN model. Finally, the frames were extracted from the 254p videos. This entire process of cropping, scaling, and extracting frames was completed in a single action using Adobe Media Encoder.

After converting the video pairs into frames, we noticed that each pair's IR video had slightly more frames than the pair's RGB video. The difference was approximately 20 additional IR frames for every 10,000 frames of IR/RGB video (seen in the slightly different video durations in Table 3). As a result, higher numbered frames from each video pair would possess a significant amount of frame drift. For example RGB frame 10,000 might match IR frame 10,020. After matching

several frame pairs from each video pair, we noticed that the frame drift was not consistently increasing as video time increased. Many video pairs appeared to have extra IR frames at the start or end of the recording in addition to increasing frame shift as video time increases. Thus, we expect that the frame shift originates from the combination of slightly desynchronized IR/RGB camera start/stop recording times and a slight difference in IR/RGB camera frame rates. To correct the frame drift, each RGB/IR video pair's resulting set of frame pairs were visually searched for frame matches. With frame matches, the approximate frame drift at various points in each video pair's frames could be measured. Based on the frame matches, an appropriate number of IR frames were removed from the beginning and end of each video pair's IR frames. Then, IR frames were removed linearly (usually 1 every 500 or 600 frames) to equalize the number of IR frames and RGB frames per video pair. After applying the above procedure to each video pair's frames, the frame pairs were again searched visually for frame matches. The largest observed frame drift after the above frame drift correction procedure was 5 frames, significantly improved from the initial maximum frame drift of 18 frames.

The images have been labeled with two binary labels, agreed on by two human experts. The Fire/No Fire label indicates whether or not the expert observes fire to be present in either the RGB or IR frame. The Smoke/No Smoke label indicates whether the experts visually estimate that smoke fills at least 50% of the RGB frame. The 50% threshold was chosen as a human expert can visually estimate 50% more quickly and accurately than other thresholds. Given the ambiguous definition of the smoke/no-smoke label mixed with inconsistent transparency of smoke, users should consider the label indicates a high smoke presence in the frame. Note that the two labels are not mutually exclusive. The dataset was labeled by two individuals independently, with finalized labels requiring agreement by both experts. The labeling procedure can be found below in Fig 3.

By labeling based on dual RGB/IR data, labels are assigned more accurately than would be for purely RGB data. The IR frames are unimpeded by smoke, enabling accurate Fire/No Fire labeling regardless of how much smoke is obscuring the RGB frame. Fig 4 depicts several example frame pairs where the dual RGB/IR setup improves labeling accuracy and provides more information than a sole RGB camera would provide.

A breakdown of dataset's total number of frames with each label can be seen below in Table 5.

TABLE 3. FLAME 2 dataset contents [21].

Data Name	Data Type	Camera	File Type	Resolution (p x p)	Duration	FPS	File Size	Labeled
1) Video Pair 1	RGB Video	M2EA RGB Cam	.mp4	3840 x 2160	291 seconds	30	3.50 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	291 seconds	30	322 MB	N
2) Video Pair 2	RGB Video	M2EA RGB Cam	.mp4	3840 x 2160	183 seconds	30	2.20 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	183 seconds	30	199 MB	N
3) Video Pair 3	RGB Video	M2EA RGB Cam	.mp4	1920 x 1080	404 seconds	30	1.69 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	404 seconds	30	446 MB	N
4) Video Pair 4	RGB Video	M2EA RGB Cam	.mp4	1920 x 1080	301 seconds	30	1.26 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	301 seconds	30	333 MB	N
5) Video Pair 5	RGB Video	M2EA RGB Cam	.mp4	1920 x 1080	267 seconds	30	1.12 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	268 seconds	30	297 MB	N
6) Video Pair 6	RGB Video	M2EA RGB Cam	.mp4	1920 x 1080	185 seconds	30	800 MB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	185 seconds	30	206 MB	N
7) Video Pair 7	RGB Video	M2EA RGB Cam	.mp4	1920 x 1080	239 seconds	30	1.00 GB	N
	IR Video	M2EA IR Cam	.mp4	640 x 512	240 seconds	30	266 MB	N
8) Original Resolution Frame Pairs	RGB Frames	M2EA RGB Cam	.jpg	3840 x 2160	13,700 frames	-	58.3 GB	Y
	RGB Frames	M2EA RGB Cam	.jpg	1920 x 1080	39,751 frames	-	46.6 GB	Y
	IR Frames	M2EA IR Cam	.jpg	640 x 512	53,451 frames	-	21.6 GB	Y
9) Scaled Resolution Frame Pairs	RGB Frames	M2EA RGB Cam	.jpg	254 x 254	53,451 frames	-	3.04 GB	Y
	IR Frames	M2EA IR Cam	.jpg	254 x 254	53,451 frames	-	5.30 GB	Y
10) Frame Pair Labels	text file	-	.txt	-	-	-	543 B	-

TABLE 4. FLAME 2 supplementary dataset contents [21].

Data Name	File Type	File Size	Description
11) README	.txt	3.1 KB	Details supplementary dataset contents and additional resources
12) README	.pdf	51 KB	Alternate README w/ hyperlinks and improved readability
13) Fire Plan	.pdf	2.6 MB	Incident Action Plan (IAP) for the Blue Stem prescribed burn
14) RGB Pointcloud	.laz	108 MB	Georeferenced RGB pointcloud of preburn area
15) RGB Orthomosaic	.jpg	5.4 MB	Image of RGB orthomosaic produced w/ preburn videos
16) Weather Forecast	.jpg	6.4 MB	Picture of national weather forecast during prescribed burn

Data Name	Data Type	Camera	File Type	Resolution (p x p)	Duration	FPS	File Size	Labeled
17) Preburn Video 1	Preburn RGB Video	Zenmuse X4S	.mp4	3840 x 2160	291 seconds	30	2.66 GB	Preburn
18) Preburn Video 2	Preburn RGB Video	Zenmuse X4S	.mp4	3840 x 2160	183 seconds	30	1.56 GB	Preburn

TABLE 5. FLAME 2 Dataset [21] frame pair label breakdown.

Label	Number of Image Pairs
Fire, Smoke	25,434
Fire, No Smoke	14,317
No Fire, No Smoke	13,700
No Fire, Smoke	0
Total Number	53,451

B. SUPPLEMENTAL DATASET

We prepared a supplemental dataset for public access and future analyses, including a pre-burn georeferenced point cloud and orthomosaic (Table 4) [21]. In addition to the photogrammetric products, we included a weather forecast, burn plan, and external resources such as Light Detection And Ranging (LiDAR) point clouds and Digital Elevation Models (DEMs). The dense point cloud and orthomosaic image of the site were generated using the pre-burn imagery collected with the DJI M200 and analyzed with Agisoft MetashapePro [75]. Videos were parsed into images every 60 frames, resulting

in 303 photos with sufficient overlap for image alignment. The dense point cloud was generated using structure-from-motion methods [76]. Settings included sequential image pre-selection, the “high” accuracy setting for image alignment, and “high” quality point cloud. To showcase of the point cloud’s potential uses, Fig. 5 depicts a top-down 2D plot of the normalized preburn point cloud, effectively mapping out canopy heights. The orthomosaic was created from a digital elevation model sourced from the point cloud with the “hole filling” setting enabled. Lastly, the point cloud and orthomosaic were georeferenced using the coordinates of natural features visible in the images and satellite imagery. A mean error of 0.705 m and a resolution of 7.6 cm/px were attained.

V. CASE STUDIES USING FLAME2 DATASET

In this section, we describe two DL-based models for fire classification and flame detection using the FLAME2 dataset as sample use cases. The dataset can be also utilized for prescribed fire modeling using the data provided in FLAME2 supplementary dataset.

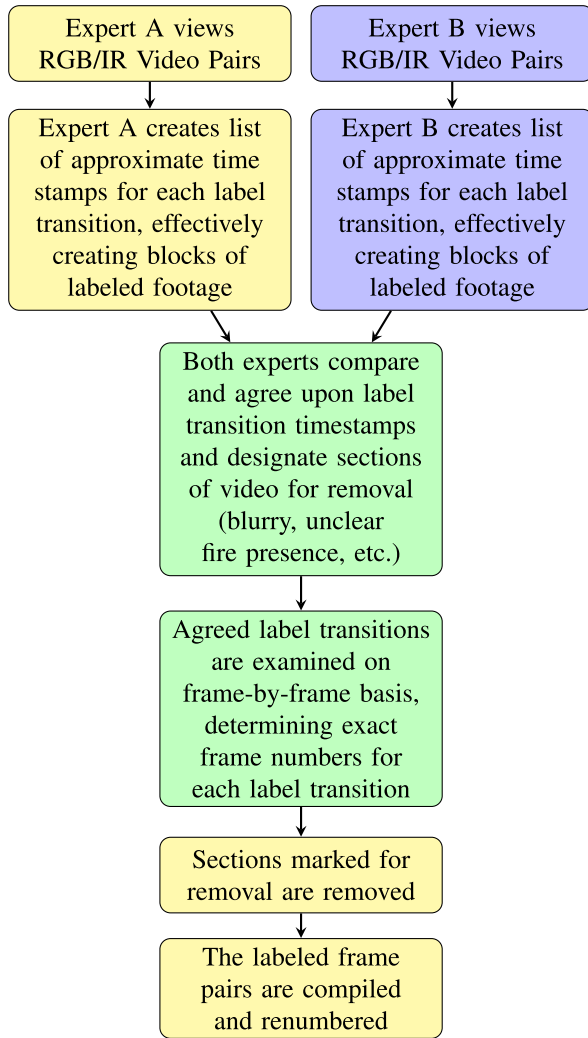


FIGURE 3. Frame pair labeling procedure. Expert A's actions are in yellow, Expert B's in blue, and joint actions are in green.

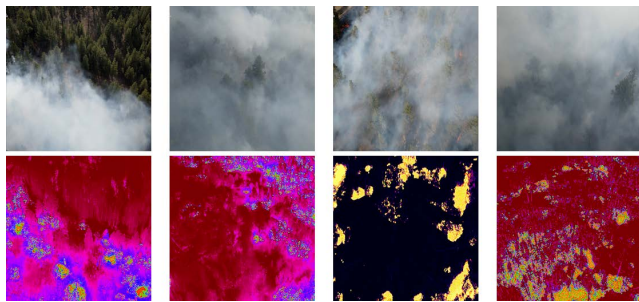


FIGURE 4. Depicts several hand picked frame pairs from FLAME 2 [21] where the IR frame provides information that is obscured by smoke in the RGB frame.

A. DL-BASED WILDFIRE CLASSIFICATION

Unlike the methods mentioned in Section II-B which classify wildfire to only two conditions, we aimed to categorize pixels in our FLAME2 dataset into four types: Flame with Smoke (YY), Flame with No Smoke (YN), Smoke with No Flame (NY), and No Flame No Smoke (NN). (Note that the class

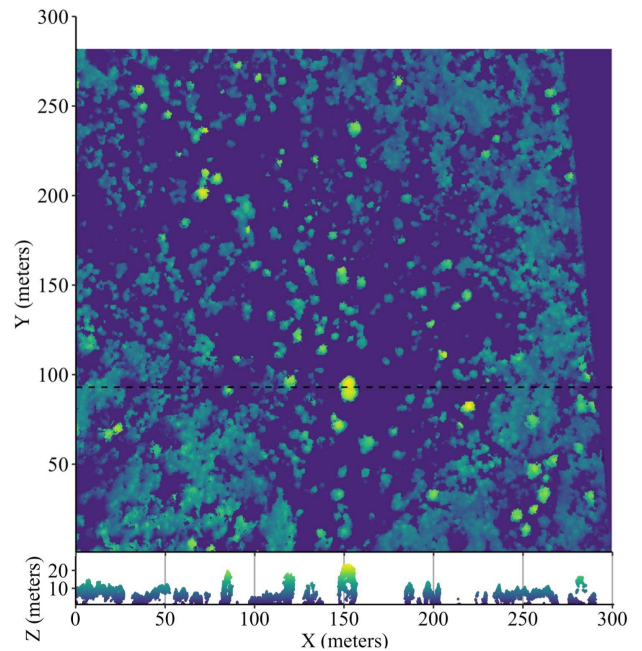


FIGURE 5. Nadir view (top) and cross-section (bottom) of the point cloud with relative height values represented in color. Yellow-green dots are tall tree crowns. The structure-from-motion derived point cloud and orthomosaic can be accessed in the supplementary dataset [21].

'Smoke with No Flame' does not exist in this dataset, but we retained this type for future research. However, this class has a practical meaning since smoke with no fires can easily be confused with fog and clouds.)

First, we applied some widely used machine learning and deep learning classification models (i.e. "benchmarks"), including Logistic Regression, LetNet(1989) [77], Vgg(2014) [78], MobileNet(2017) [79], and ResNet(2016) [80], to our dataset, as well as our method ("Flame" [7], described below and shown in Fig. 6). Note that Vgg, MobileNet, and ResNet are pre-trained (i.e., each model has been trained on some other datasets that have some prior knowledge to understand the features of images, but the high-level features may not be enough for the new task of fire detection). Then, we need only modify some layers of the pre-trained models (their "classifiers") and train the classifiers on our dataset; then, those models can have good accuracy on our dataset. The number of parameters of each model is shown in Table 6. Note that only the trainable parameters are required to update in the training phase, which hugely accelerates the training progress. The Flame net, taking advantage of spreadable convolutional layers, only has less than 1000 trainable parameters.

The input of the model can be only RGB images ($f(RGB; \theta)$) or IR images ($f(IR; \theta)$) or we can benefit from inputting both RGB and IR pairs ($f(RGB, IR; \theta)$), where $f(\cdot; \theta)$ denotes the network mapping function parameterized by θ . When inputting RGB-IR pairs, we consider two simple approaches to perform feature fusion (shown in Fig 7), named *Early Fusion* and *Late Fusion*. Specifically, in Early Fusion, we simply concatenate the paired images and modify the

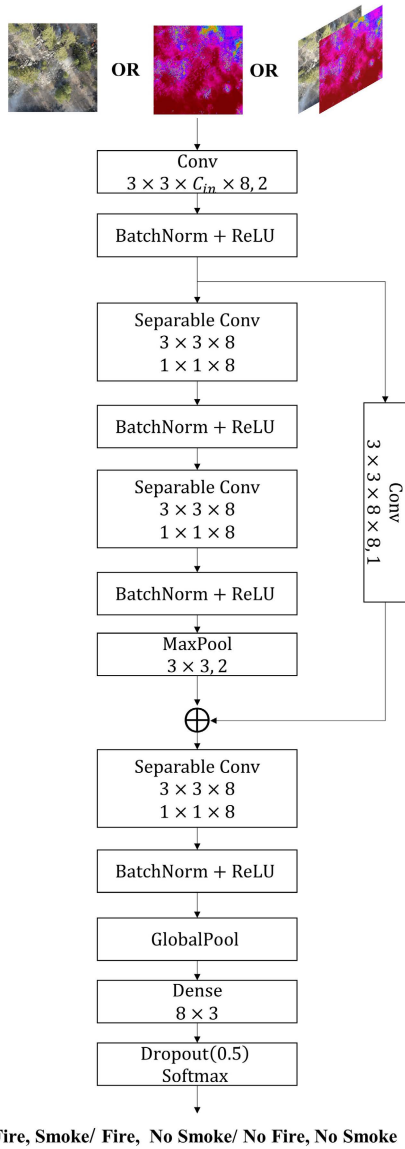


FIGURE 6. Architecture of the Flame network introduced in [7]. For the regular convolutional layer, the form of parameters is $k \times k \times C_{in} \times C_{out}$, stride. For separable convolutional layers, the form of parameters is $k \times k \times C_{in}$ and $k \times k \times C_{out}$. For the max pooling layer, the form of parameters is $k \times k$, stride. For the dense layer, the form of parameters is $C_{in} \times C_{out} \cdot k$ denotes the kernel size, C_{in} and C_{out} denote the number of input channels and output channels, respectively.

number of channels of the input of the first layer from 3 to 6. In Late Fusion, either RGB or IR will be fed to two streams with the same architecture (i.e. the left stream learns from the RGB domain while the right stream learn from the IR domain). We then concatenate the extracted features from each stream and feed the fused feature to a fully-connected layer to perform classification. Thus, subsequent layers can learn the high-level representation from both RGB and IR domains.

B. IMAGE PROCESSING-BASED FLAME DETECTION

In addition to the frame-by-frame fire classification, image processing-based methods for flame detection and

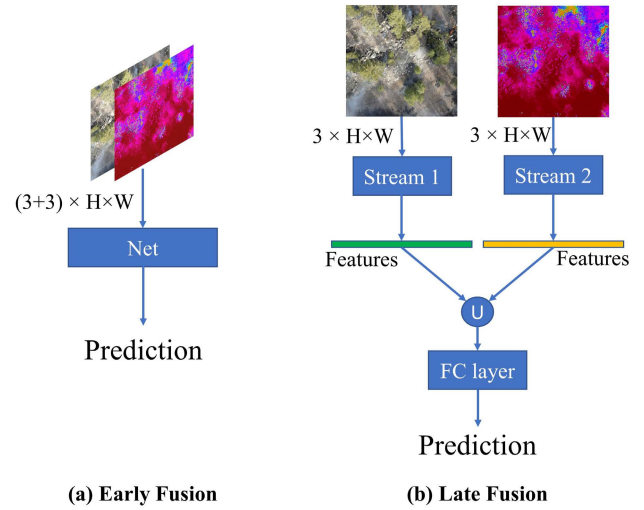


FIGURE 7. Two approaches for feature fusion.

TABLE 6. Number of parameters of each model.

Model	Mode	Total Para.	Trainable para.
Logistic	RGB	5.81E+05	5.81E+05
Logistic	IR	5.81E+05	5.81E+05
Logistic	Early Fusion	1.16E+06	1.16E+06
LeNet5	RGB	3.17E+07	3.17E+07
LeNet5	IR	3.17E+07	3.17E+07
LeNet5	Early Fusion	3.17E+07	3.17E+07
LeNet5	Late Fusion	6.33E+07	0.00E+00
Flame	RGB	7.95E+02	7.95E+02
Flame	IR	7.95E+02	7.95E+02
Flame	Early Fusion	1.01E+03	1.01E+03
Flame	Late Fusion	2.50E+03	2.50E+03
VGG16	RGB	1.34E+08	1.20E+08
VGG16	IR	1.34E+08	1.20E+08
MobileNetV2	RGB	2.23E+06	3.84E+03
MobileNetV2	IR	2.23E+06	3.84E+03
MobileNetV2	Late Fusion	7.73E+06	3.29E+06
ResNet18	RGB	1.12E+07	1.54E+03
ResNet18	IR	1.12E+07	1.54E+03
ResNet18	Late Fusion	2.26E+07	2.64E+05

geo-localization can be applied to the FLAME2 dataset. The flame zone task can provide valuable information to the first responders in the areas impacted by the wildfire in order to steer the fire management strategy.

Many works study ML-based flame detection strategies. Although, these methods achieve high classification rates in specific areas, the results are not easily transferable to other environments, primarily because of the limited number of available labeled wildfire flame and smoke datasets, as mentioned earlier in section II. B. Some researchers approach this problem by building their own datasets, then labeling the dataset manually or based on some machine learning techniques.

In [81], researchers applied an YOLOv3-based object detection method to detect flame area and smoke zones in their created RGB-image dataset. In [82], the authors manually labeled the FUEGO wildfire detection dataset, then using an Inception v3-based model detected smoke and flame on these RGB images by drawing bounding boxes around the flame and smoke areas. Reference [83] implemented another

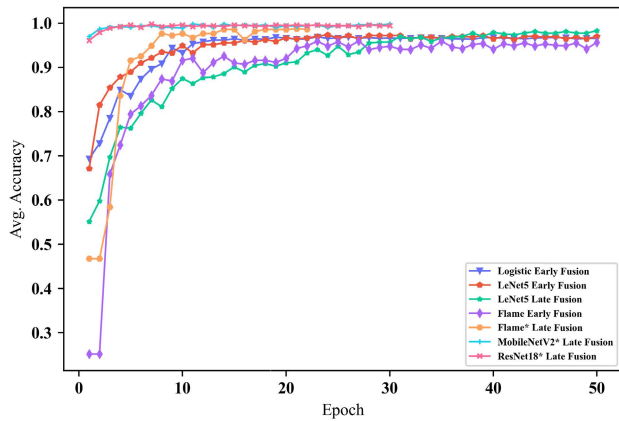


FIGURE 8. Averaged accuracy via time on test set of different models with input of RGB-thermal pairs. The models with the sign ‘*’ denote the fine-tuned pre-trained models, while the remaining models denote the models training from scratch.

neural network model for flame detection, manually labeling the wildfire images from the internet. They used a trained MobileNet V2-based model to divide each image into blocks according to the possibility of flame and smoke. Likewise, [7] provided the idea of using Xception-based model to discriminate the wildfire and non-wildfire images, as well as using Subject Matter Expert (SME) effect to mark the fire pixels to generate fire segmentation masks.

However, the data collection and labeling methodology of these studies are not sustainable. On one hand, the manual labeling is labor costly and time consuming. On the other hand, machine learning-based auto-labeling lacks the required accuracy and can trigger false alarms. For instance, a trained Inception-v3 model will reach a general accuracy of 93.6% in the Portuguese firefighters portal database (RGB). However, the model may have a high chance (over 95%) to misclassify the sunset images as flame, meaning that the model is suitable only under some weather and illumination conditions [40].

While the aforementioned works only focus on RGB images, the FLAME2 dataset includes both RGB and IR images captured with a coaxial infrared-camera. Noting this fact, we divided the images into flame zones and background based on pixel-wise representation values since the pixel intensity of IR images represent the fire intensity.

A Maximally Stable Extremal Regions (MSER) method was applied to detect the image blob features and then generate bounding boxes for detected features. As in many other object detection applications, the MSER method usually generates many partially overlapping bounding boxes for the same object. One simple solution to avoid many unnecessary calls and achieve more precise localization of the fire flames would be using Non-Maximum Suppression (NMS) method to eliminate the overlapping bounding boxes in favor of the strongest one. The confidence of flames in a single box is calculated as below:

$$C_{boxi} = \frac{\sum I_{pixel}}{W_{boxi} \times H_{boxi} \times 255}, \quad (1)$$

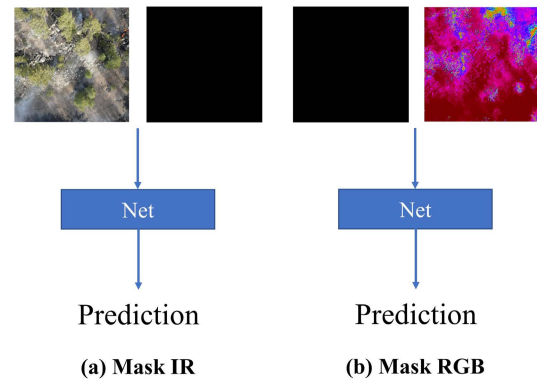


FIGURE 9. Scenario of ablation analysis by masking one of the dual-channel.

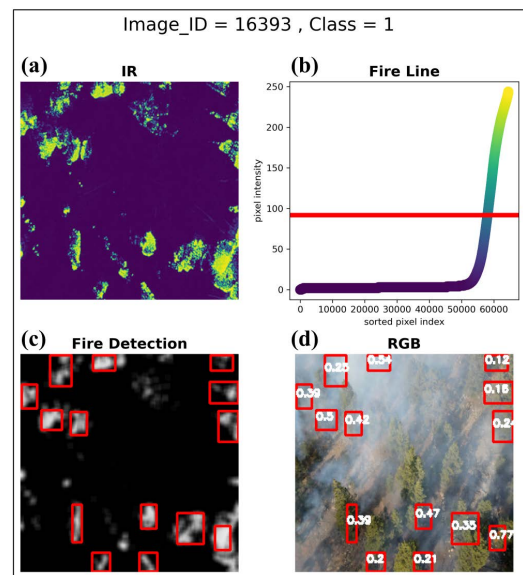


FIGURE 10. Flame detection based on pixel segmentation and MSER-NMS, (a) is the IR image in gray-scale, (b) is the segmentation based on pixel intensity, (c) is the result of pixel segmentation and MSER-NMS, and (d) is the RGB image with bounding boxes, the box location is copied from (c).

where I denotes the intensity of an individual pixel in a single box, W and H denote the width and the height of the correspond box, respectively. The confidence score C_{boxi} represents the fire density and intensity at each bounding box, higher is better.

By fine-tuning the suppression processes of non-maximum parameters and the threshold of pixel intensity (fire-line for segmentation) on IR images, the algorithm identifies areas with higher confidence to contain flames.

VI. EXPERIMENTAL RESULTS

A. FIRE CLASSIFICATION

To perform the per-frame fire classification, for each experiment, we only used 1% randomly sampled data and split it into 80% to train and 20% to test. This operation can accelerate the testing process and allow us to evaluate more models in

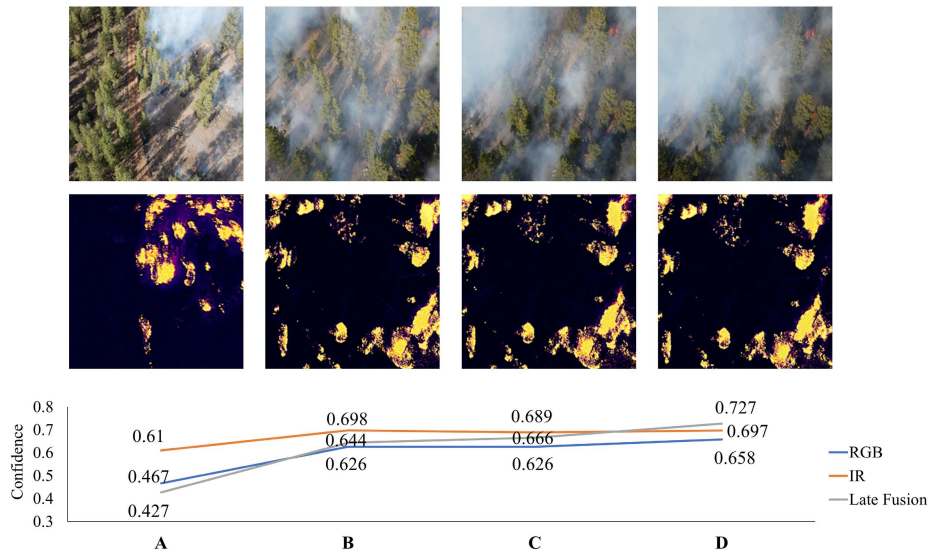


FIGURE 11. Detection confidence for sequential frames by pre-trained ResNet18 with RGB, IR, and Late Fusion, respectively. All of them are detected as YN class. The confidence is calculated as the softmax probability of the predicted class.

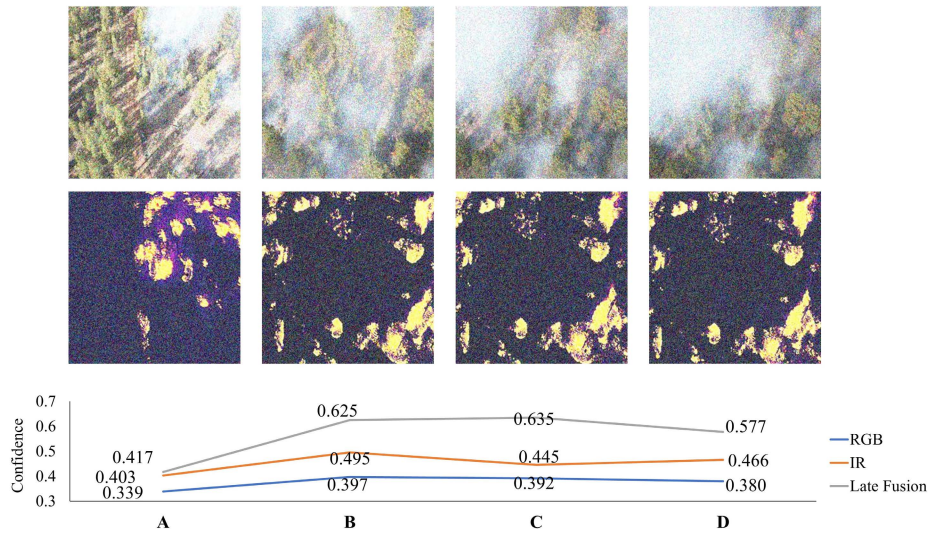


FIGURE 12. Detection confidence under Gaussian noise ($\sigma = 0.5$) for sequential frames by pre-trained ResNet18 with RGB, IR, and Late Fusion, respectively. All of them are detected as YN class. The confidence is calculated as the softmax probability of the predicted class.

a short time. Also, it is helpful in a sense that the consecutive frames in a video are very similar. With no sampling, very close frames can appear in both training and test datasets providing misleading high accuracy. The results are more reliable with sampling. We perform ten experiments for each model. We use the ADAM [84] optimizer with $1e-3$ learning rate for the Flame network, and $1e-4$ for the other models. The batch size is set to 64, considering the *label smoothing* with probability 0.2 is used in the training phase. To achieve a fair comparison, we train the models that learn from scratch with 50 epochs, and the pre-trained models with 30 epochs. Besides focusing on the accuracy of the classification, we are

more interested in the macro-level metrics, such as macro F1 score, macro recall, and macro precision. This is because in reality, the wildfire frequency is often low, and the performance of the model cannot be simply reflected by the classification accuracy. For example, one may have only one wildfire sample in a total of 1000 images. If the model labels all samples as “no fire”, then the accuracy would be 99.9%. For each class, the recall (R), precision (P), and F1 score are calculated by:

$$R = \frac{TP}{TP + FN} \quad (2)$$

TABLE 7. Comparison of averaged performance of different models with different mode. The models with the sign '*' denote the fine-tuned pre-trained models.

Model	Mode	F1 Score	Precision	Recall	Accuracy
Logistic	RGB	89.57	90.99	89.48	90.37
Logistic	IR	92.61	92.94	92.43	92.43
Logistic	Early Fusion	96.71	96.92	96.65	96.54
LeNet5	RGB	95.39	95.86	95.12	95.33
LeNet5	IR	92.3	92.19	92.79	92.15
LeNet5	Early Fusion	97.16	97.35	97.1	97.01
LeNet5	Late Fusion	98.45	98.35	98.58	98.32
Flame	RGB	94.53	95.18	94.38	94.86
Flame	IR	86.81	87.47	86.91	85.79
Flame	Early Fusion	95.67	96.17	95.26	95.70
Flame	Late Fusion	98.63	98.76	98.52	98.60
VGG16*	RGB	99.92	99.9	99.93	99.91
VGG16*	IR	97.35	97.57	97.26	97.29
MobileNetV2*	RGB	99.36	99.33	99.42	99.35
MobileNetV2*	IR	97.51	97.65	97.43	97.38
MobileNetV2*	Late Fusion	99.82	99.78	99.87	99.81
ResNet18*	RGB	98.46	98.57	98.37	98.32
ResNet18*	IR	96.54	96.97	96.27	96.26
ResNet18*	Late Fusion	99.5	99.46	99.56	99.44

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$F1 = 2 \times \frac{P \times R}{P + R}, \quad (4)$$

where TP denotes True Positive, TN denotes True Negative, FP denotes False Positive, and FN denotes False Negative. The marco level metrics take the average of the calculated class-wise metrics, regardless of the number of samples in each class.

The results are shown in Table 7 and Fig. 8. We found the pre-trained models can fast converge to high accuracy in early epochs (some as soon as the first epoch) and generally outperform our customized models. This is reasonable because they are pre-trained on the large dataset such as ImageNet dataset, meaning that they have a richer experience in understanding different features of the fire imagery, while our customized model has only been trained using a few hundred images.

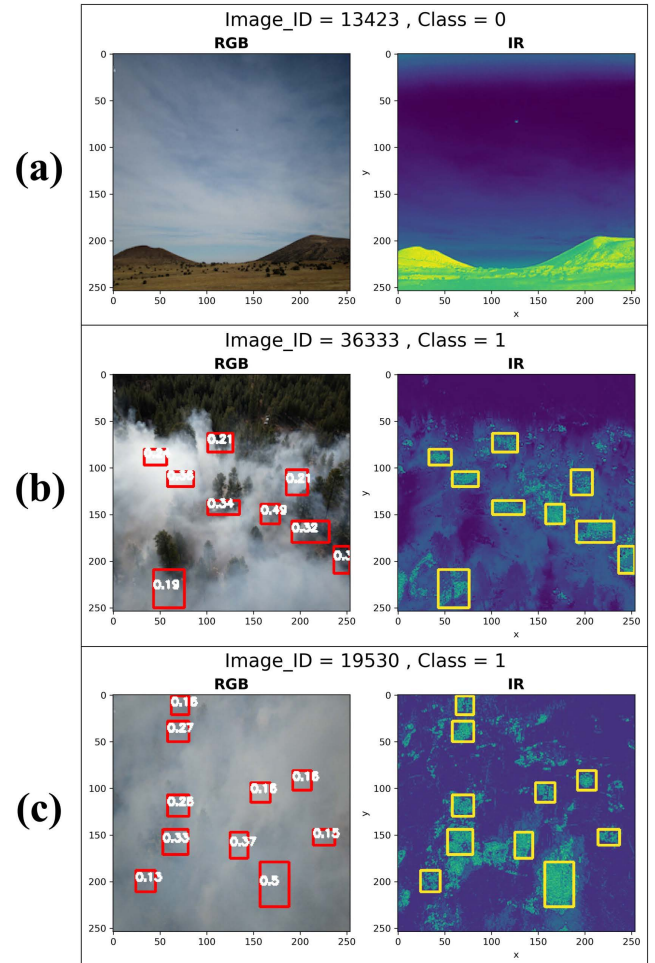
We also show the ablation analysis for the fusion models. We approximate the individual contribution from one domain by masking one domain input (the scenario is shown in Fig. 8). The results are shown in Table 8, which indicate the information from different domains contributes reasonably to the fusion model.

More examples of detection confidence for sequential frames by ResNet18 with different modes are shown in Figs. 11 and 12. Fig. 11 demonstrates the detection confidence for clean frames. The trends of the curves from the three modes are consistent. Fig. 12 presents the detection confidence for frames with Gaussian noise. The late fusion model exhibits stronger robustness against single-channel models.

It is noteworthy that the high accuracy of ResNet18 is achieved within one dataset, hence it is not directly transferrable to other datasets. The achieved accuracies of such studies are highly dependent upon testing conditions such as the size of dataset, weather conditions, resolution and length of videos, camera resolution and altitude, etc. Compared to

TABLE 8. Ablation analysis for models with dual-channel input.

Model	Mode	Mask IR	Mask RGB	No Mask
Logistic	Early Fusion	62.28	42.64	96.71
LeNet5	Early Fusion	81.13	42.14	97.16
LeNet5	Late Fusion	86.13	50.9	98.45
Flame	Early Fusion	56.53	60.6	95.67
Flame	Late Fusion	80.21	71.06	98.63
MobileNetV2	Late Fusion	88.5	82.6	99.82
ResNet18	Late Fusion	93.87	89	99.5

**FIGURE 13.** Flame detection in different scenarios. (a) No flame, no smoke, (b)(c) Flame with smoke.

other models, the advantage of the proposed Flame net is its fewest parameters (shown in Table 6) and much shorter training time.

B. FLAME DETECTION

We tested the performance of our flame detection system equipped with the MSER-NMS operation in several scenarios and evaluated its accuracy and robustness. Fig.10 shows the result of flame detection, where the flame detection's accuracy is not affected by smoke. Thus, our proposed framework is simple, stable, computationally efficient, and labor-free. As more results shown in Fig.13, the fire's smoke did not prevent

our system from detecting flames and the system successfully drew out the flame area in each frame. We can implement a hierarchical method by first classifying frames into fire versus no-fire, and then running the detection algorithm only on fire frames, to be more computationally efficient.

In order to verify the suitability of the proposed detection method in real-time scenarios, we examine the processing speed of the proposed method with real wildfire images. We observe that the system achieves an average of 83 FPS on 39,700 flame images with $256p \times 256p$ resolution, a frame much higher than the usual 30 FPS rate of most UAV cameras. The program has been tested on a PC with Intel i7-11700 CPU and 16GB memory. Since the computation time depends mainly on the performance of the CPU and RAM, we believe that it has the potential of being deployed on edge devices such as UAVs and mobile sites, toward making real-time aerial monitoring ubiquitous.

VII. CONCLUSION

In this work, we present FLAME2, a dual-feed prescribed fire imaging dataset collected by unmanned aerial systems in a ponderosa pine forest using side-by-side visual and thermal camera feeds. In addition, we present and apply our new deep learning-based analysis methods to the dataset to accurately label and segment, frame by frame, pixels with fire and/or smoke present. The accuracy, measured against delineation by two human experts, was above 94%. The outcomes of the work are twofold. First, the work is a step toward developing methods for real time wildfire detection using low cost, (semi-)autonomous sensing systems. Second, the observational procedure, as well as the FLAME2 and supplemental datasets, can be used as a template by other researchers to improve data-driven fire modeling techniques. Future studies could investigate fire intensity modeling and perimeter mapping using data from FLAME2. Advances in operationally ready aerial fire detection may use similar algorithms tested above for real-time detection of small wildfires or spot fires of large wildfires. The transferability and explainability of DL models refer to Domain Adaption, Incremental Learning, and Explainable-AI is also planned to be addressed by studying multiple related datasets in the future.

ACKNOWLEDGMENT

The authors would like to thank USDA Forest Service and Kaibab National Forest for providing access to the prescribed fire.

REFERENCES

- [1] (2018). *ALERT Wildfire*. [Online]. Available: <https://www.alertwildfire.org/>
- [2] *Geostationary Operational Environmental Satellites (GOES)-R Series*. Accessed: Jul. 15, 2022. [Online]. Available: <https://www.goes-r.gov/multimedia/dataAndImageryVideosGoes-17.html>
- [3] *Multi-Mission Aircraft (MMA)*. Accessed: Jul. 15, 2022. [Online]. Available: <https://dfpc.colorado.gov/MMA>
- [4] *USDA Forest Service National Infrared Operations (NIROPS)*. Accessed: Jul. 15, 2022. [Online]. Available: <https://fsapps.nwcg.gov/nirops/>
- [5] C. T. Lopez, "DOD extends 'firefly' related 'FireGuard' support to extinguish wildfires," U.S. Dept. Defense, Sep. 2021. Accessed: Jul. 15, 2022. [Online]. Available: <https://www.defense.gov/News/News-Stories/Article/Article/2768197/dod-extends-firefly-related-fireguard-support-to-extinguish-wildfires/>
- [6] V. G. Ambrosia, S. Wegener, T. Zajkowski, D. V. Sullivan, S. Buechel, F. Enomoto, B. Lobitz, S. Johan, J. Brass, and E. Hinkley, "The ikhona unmanned airborne system (UAS) western states fire imaging missions: From concept to reality (2006–2010)," *Geocarto Int.*, vol. 26, no. 2, pp. 85–101, Apr. 2011.
- [7] A. Shamsoshoara, F. Afghah, A. Razi, L. Zheng, P. Z. Fulé, and E. Blasch, "Aerial imagery pile burn detection using deep learning: The FLAME dataset," *Comput. Netw.*, vol. 193, Jul. 2021, Art. no. 108001.
- [8] L. Merino, F. Caballero, J. R. M. de Dios, I. Maza, and A. Ollero, "Automatic forest fire monitoring and measurement using unmanned aerial vehicles," in *Proc. 6th Int. Congr. Forest Fire Res.* Citeseer, 2010.
- [9] F. Afghah, A. Razi, J. Chakareski, and J. Ashdown, "Wildfire monitoring in remote areas using autonomous unmanned aerial vehicles," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2019, pp. 835–840.
- [10] Q. Huang, A. Razi, F. Afghah, and P. Fule, "Wildfire spread modeling with aerial image processing," in *Proc. IEEE 21st Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Aug. 2020, pp. 335–340.
- [11] M. Basso, D. Stocchero, R. V. B. Henriques, A. L. Vian, C. Bredemeier, A. A. Konzen, and E. Pignaton de Freitas, "Proposal for an embedded system architecture using a GNDVI algorithm to support UAV-based agrochemical spraying," *Sensors*, vol. 19, no. 24, p. 5397, Dec. 2019.
- [12] M. Gharib, S. Nandadapu, and F. Afghah, "An exhaustive study of using commercial LTE network for UAV communication in rural areas," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.
- [13] S. H. Alsamhi, F. A. Almalki, F. Afghah, A. Hawbani, A. V. Shvetsov, B. Lee, and H. Song, "Drones' edge intelligence over smart environments in B5G: Blockchain and federated learning synergy," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 295–312, Mar. 2022.
- [14] Y. Liu, A. Kochanski, K. R. Baker, W. Mell, R. Linn, R. Paugam, and J. Mandel, "Fire behaviour and smoke modelling: Model improvement and measurement needs for next-generation smoke research and forecasting systems," *Int. J. Wildland Fire*, vol. 28, pp. 570–588, 2019.
- [15] W. E. McNamara and D. Mell, "An approach to integrated data management for three-dimensional, time-dependent fire behaviour model evaluation," *Int. J. Wildland Fire*, vol. 30, no. 12, pp. 911–920, 2021.
- [16] S. Prichard, N. Larkin, R. Ottmar, N. French, K. Baker, T. Brown, C. Clements, M. Dickinson, A. Hudak, A. Kochanski, R. Linn, Y. Liu, B. Potter, W. Mell, D. Tanzer, S. Urbanski, and A. Watts, "The fire and smoke model evaluation experiment—A plan for integrated, large fire-atmosphere field campaigns," *Atmosphere*, vol. 10, no. 2, p. 66, Feb. 2019.
- [17] K. Nelson, J. Boehmle, A. Khlystov, H. Moosmüller, V. Samburova, C. Bhattarai, E. Wilcox, and A. Watts, "A multipollutant smoke emissions sensing and sampling instrument package for unmanned aircraft systems: Development and testing," *Fire*, vol. 2, no. 2, p. 32, Jun. 2019. [Online]. Available: <https://www.mdpi.com/2571-6255/2/2/32>
- [18] J. Aurell, B. Gullett, A. Holder, F. Kiro, W. Mitchell, A. Watts, and R. Ottmar, "Wildland fire emission sampling at fishlake national forest, Utah using an unmanned aircraft system," *Atmos. Environ.*, vol. 247, Feb. 2021, Art. no. 118193.
- [19] N. Barber, E. Alvarado, V. R. Kane, W. E. Mell, and L. M. Moskal, "Estimating fuel moisture in grasslands using UAV-mounted infrared and visible light sensors," *Sensors*, vol. 21, no. 19, p. 6350, Sep. 2021.
- [20] A. C. Watts, V. G. Ambrosia, and E. A. Hinkley, "Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use," *Remote Sens.*, vol. 4, no. 6, pp. 1671–1692, 2012. [Online]. Available: <http://libproxy.clemson.edu/login?url=https://www.proquest.com/scholarly-journals/unmanned-aircraft-systems-remote-sensing/docview/1537378479/se-2>
- [21] B. Hopkins, L. O'Neill, F. Afghah, A. Razi, A. Watts, P. Fule, and J. Coen, "FLAME 2: Fire detection and modeling: Aerial Multi-spectral image dataset," *IEEE DataPort*, 2022. [Online]. Available: <https://iee-dataport.org/open-access/flame-2-fire-detection-and-modeling-aerial-multi-spectral-image-dataset>, doi: 10.21227/swyw-6j78.
- [22] J. Coen, M. Cruz, D. Rosales-Giron, and K. Speer, "Coupled fire-atmosphere model evaluation and challenges," in *Wildland Fire Dynamics: Fire Effects and Behavior From a Fluid Dynamics Perspective*, K. Speer and S. Goodrick, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2022. [Online]. Available: <https://www.cambridge.org/core/books/wildland-fire-dynamics/ABB400C822D978A3F4BF02861F5BA5D3>

- [23] J. L. Coen and W. Schroeder, "Use of spatially refined remote sensing fire detection data to initialize and evaluate coupled weather-wildfire growth model simulations," *Geophys. Res. Lett.*, vol. 40, pp. 5536–5541, Oct. 2013.
- [24] *Fire Information for Resource Management System (FIRMS)*. Accessed: Jul. 15, 2022. [Online]. Available: <https://firms.modaps.eosdis.nasa.gov/>
- [25] Y. Hou, M. Chen, R. Volk, and L. Soibelman, "Investigation on performance of RGB point cloud and thermal information data fusion for 3D building thermal map modeling using aerial images under different experimental conditions," *J. Building Eng.*, vol. 45, Jan. 2022, Art. no. 103380.
- [26] C. J. Moran, V. Hoff, R. A. Parsons, L. P. Queen, and C. A. Seielstad, "Mapping fine-scale crown scorch in 3D with remotely piloted aircraft systems," *Fire*, vol. 5, no. 3, p. 59, Apr. 2022.
- [27] (2021). DataCluster Labs. *Fire and Smoke Dataset*. [Online]. Available: <https://www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset>
- [28] C. Kyrkou and T. Theodoridis, "EmergencyNet: Efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1687–1699, 2020.
- [29] C. Kyrkou and T. Theodoridis, "Deep-Learning-Based aerial image classification for emergency response applications using unmanned aerial vehicles," 2019, *arXiv:1906.08716*.
- [30] A. Khan and B. Hassan. (Aug. 2020). *Dataset for Forest Fire Detection*. [Online]. Available: <https://data.mendeley.com/datasets/gjmr63rz2r/1>
- [31] D. Shah. (Jul. 2020). *Fire Detection*. [Online]. Available: <https://github.com/jackfrost1411/fire-detection>
- [32] M. Olafenwa. (Aug. 2019). *FireNET*. [Online]. Available: <https://github.com/OlafenwaMoses/FireNET>
- [33] R. Pande. (2019). *Fire Detection from CCTV*. [Online]. Available: <https://www.kaggle.com/datasets/ritupande/fire-detection-from-cctv>
- [34] G. Samarth, N. Bhowmik, and T. Breckon, "Experimental exploration of compact convolutional neural network architectures for non-temporal real-time fire detection," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2019, pp. 653–658.
- [35] A. J. Dunnings and T. P. Breckon, "Experimentally defined convolutional neural network architecture variants for non-temporal real-time fire detection," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1558–1562.
- [36] J. Sharma and M. Goodwin. (2017). *Fire Detection Image Dataset*. [Online]. Available: <https://github.com/cair/Fire-Detection-Image-Dataset>
- [37] P. Foggia, A. Saggese, and M. Vento, "Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1545–1556, Sep. 2015.
- [38] R. D. Lascio, A. Greco, A. Saggese, and M. Vento, "Improving fire detection reliability by a combination of videoanalytics," in *Proc. Int. Conf. Image Anal. Recognit. (ICIAR)*, 2014, pp. 477–484.
- [39] H. Wu, H. Li, A. Shamsoshoara, A. Razi, and F. Afghah, "Transfer learning for wildfire identification in UAV imagery," in *Proc. 54th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2020, pp. 1–6.
- [40] M. J. Sousa, A. Moutinho, and M. Almeida, "Wildfire detection using transfer learning on augmented datasets," *Expert Syst. Appl.*, vol. 142, Mar. 2020, Art. no. 112975.
- [41] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
- [42] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
- [43] A. Jadon, A. Varshney, and M. S. Ansari, "Low-complexity high-performance deep learning model for real-time low-cost embedded fire detection systems," *Proc. Comput. Sci.*, vol. 171, pp. 418–426, Jan. 2020.
- [44] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 7263–7271.
- [45] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
- [46] S. Saponara, A. Elhanashi, and A. Gagliardi, "Real-time video fire/smoke detection based on CNN in antifire surveillance systems," *J. Real-Time Image Process.*, vol. 18, no. 3, pp. 889–900, Jun. 2021.
- [47] M. Park and B. C. Ko, "Two-step real-time night-time fire detection in an urban environment using static ELASTIC-YOLOv3 and temporal fire-tube," *Sensors*, vol. 20, no. 8, p. 2202, Apr. 2020.
- [48] P. Li and W. Zhao, "Image fire detection algorithms based on convolutional neural networks," *Case Stud. Thermal Eng.*, vol. 19, Jun. 2020, Art. no. 100625.
- [49] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer*, 2016, pp. 21–37.
- [50] S. Wu and L. Zhang, "Using popular object detection methods for real time forest fire detection," in *Proc. 11th Int. Symp. Comput. Intell. Design (ISCID)*, Dec. 2018, pp. 280–284.
- [51] A. Nguyen, H. Nguyen, V. Tran, H. X. Pham, and J. Pestana, "A visual real-time fire detection using single shot multibox detector for uav-based fire surveillance," in *Proc. IEEE 8th Int. Conf. Commun. Electron. (ICCE)*, Jan. 2021, pp. 338–343.
- [52] V. S. Bochkov and L. Y. Kataeva, "WUUNet: Advanced fully convolutional neural network for multiclass fire segmentation," *Symmetry*, vol. 13, no. 1, p. 98, Jan. 2021.
- [53] S. Guo, B. Hu, and R. Huang, "Real-time flame segmentation based on RGB-thermal fusion," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2021, pp. 1435–1440.
- [54] S. T. Seydi, V. Saeidi, B. Kalantar, N. Ueda, and A. A. Halin, "Fire-Net: A deep learning framework for active forest fire detection," *J. Sensors*, vol. 2022, pp. 1–14, Feb. 2022.
- [55] A. Bouguettaya, H. Zarzour, A. M. Taberkit, and A. Kechida, "A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms," *Signal Process.*, vol. 190, Jan. 2022, Art. no. 108309.
- [56] J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 13, pp. 3521–3526, 2017.
- [57] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2935–2947, Dec. 2017.
- [58] T. Artés, D. Oom, D. de Rigo, T. H. Durrant, P. Maiani, G. Libertà, and J. San-Miguel-Ayanz, "A global wildfire dataset for the analysis of fire regimes and fire behaviour," *Scientific Data*, vol. 6, no. 1, pp. 1–11, Dec. 2019.
- [59] S. A. Parks, C. Miller, M.-A. Parisien, L. M. Holsinger, S. Z. Dobrowski, and J. Abat-Zoglou, "Wildland fire deficit and surplus in the western United States, 1984–2012," *Geocarto Int. J.*, vol. 6, pp. 1–13, Dec. 2015.
- [60] Attribution 4.0 International (CC BY 4.0). (2020). *Next Day Wildfire Spread—A Data Set to Predict Wildfire Spreading From Remote-Sensing Data*. [Online]. Available: <https://www.kaggle.com/fantineh/next-day-wildfire-spread>
- [61] J. Eidenshink, B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. Howard, "A Project for Monitoring Trends in Burn Severity," *Fire Ecol.*, vol. 3, no. 1, pp. 3–21, Jun. 2007.
- [62] J. Balch, L. St Denis, A. Mahood, N. Mietkiewicz, T. Williams, J. McGlinchy, and M. Cook, "FIRED (fire events delineation): An open, flexible algorithm and database of U.S. fire events derived from the MODIS burned area product (2001–2019)," *Remote Sens.*, vol. 12, p. 3498, Oct. 2015.
- [63] (2022). Attribution 4.0 International (CC BY 4.0). *National Interagency Historic Wildland Fires*. [Online]. Available: https://data-niic.opendata.arcgis.com/search?tags=Category%2Chistoric_w%ildlandfire_opendata
- [64] G. M. Schag, D. A. Stow, P. J. Riggan, R. G. Tissell, and J. L. Coen, "Examining landscape-scale fuel and terrain controls of wildfire spread rates using repetitive airborne thermal infrared (ATIR) imagery," *Fire*, vol. 4, no. 1, p. 6, Feb. 2021.
- [65] D. A. Stow, P. J. Riggan, E. J. Storey, and L. L. Coulter, "Measuring fire spread rates from repeat pass airborne thermal infrared imagery," *Remote Sens. Lett.*, vol. 5, no. 9, pp. 803–812, Sep. 2014.
- [66] D. Stow, P. Riggan, W. Schag, G. Brewer, R. Tissell, J. Coen, and E. Storey, "Measuring fire spread rates from repeat pass airborne thermal infrared imagery," *Remote Sens. Lett.*, vol. 40, pp. 4876–4897, Sep. 2020.
- [67] J. L. Coen and W. Schroeder, "The high park fire: Coupled weather-wildland fire model simulation of a windstorm-driven wildfire in Colorado's front range," *J. Geophys. Res., Atmos.*, vol. 120, no. 1, pp. 131–146, Jan. 2015.
- [68] T. L. Clark, L. Radke, J. Coen, and D. Middleton, "Analysis of small-scale convective dynamics in a crown fire using infrared video camera imagery," *J. Appl. Meteorol.*, vol. 38, no. 10, pp. 1401–1420, Oct. 1999.
- [69] L. F. Radke, T. L. Clark, J. L. Coen, C. A. Walther, R. N. Lockwood, P. J. Riggan, J. A. Brass, and R. G. Higgins, "The Wildfire experiment (WIFE): Observations with airborne remote sensors," *Can. J. Remote Sens.*, vol. 26, no. 5, pp. 406–417, Oct. 2000.
- [70] J. Coen, S. Mahalingam, and J. Daily, "Infrared imagery of crown-fire dynamics during FROSTFIRE," *J. Appl. Meteorol.*, vol. 43, no. 9, pp. 1241–1259, Sep. 2004.

- [71] D. Radke, A. Hessler, and D. Ellsworth, "FireCast: Leveraging deep learning to predict wildfire spread," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 1–7. [Online]. Available: <https://www.ijcai.org/Proceedings/2019/0636.pdf>
- [72] T. I. Zohdi, "A machine-learning framework for rapid adaptive digital-twin based fire-propagation simulation in complex environments," *Comput. Methods Appl. Mech. Eng.*, vol. 363, May 2020, Art. no. 112907.
- [73] M. C. Rochoux, S. Ricci, D. Lucor, B. Cuenot, and A. Trouvé, "Towards predictive data-driven simulations of wildfire spread—Part I: Reduced-cost ensemble Kalman filter based on a polynomial chaos surrogate model for parameter estimation," *Natural Hazards Earth Syst. Sci.*, vol. 14, no. 11, pp. 2951–2973, Nov. 2014.
- [74] M. C. Rochoux, C. Emery, S. Ricci, B. Cuenot, and A. Trouvé, "Towards predictive data-driven simulations of wildfire spread—Part II: Ensemble Kalman filter for the state estimation of a front-tracking simulator of wildfire spread," *Natural Hazards Earth Syst. Sci.*, vol. 15, no. 8, pp. 1721–1739, Aug. 2015.
- [75] (2021). *Metashape*. [Online]. Available: <http://www.agisoft.com/downloads/installer/>
- [76] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds, "'Structure-from-motion' photogrammetry: A low-cost, effective tool for geoscience applications," *Geomorphology*, vol. 179, pp. 300–314, Dec. 2012.
- [77] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and D. and L. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
- [78] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [79] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
- [80] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [81] F. M. A. Hossain, Y. M. Zhang, and M. A. Tonima, "Forest fire flame and smoke detection from UAV-captured images using fire-specific color features and multi-color space local binary pattern," *J. Unmanned Vehicle Syst.*, vol. 8, no. 4, pp. 285–309, Dec. 2020.
- [82] K. Govil, M. L. Welch, J. T. Ball, and C. R. Pennypacker, "Preliminary results from a wildfire detection system using deep learning on remote camera images," *Remote Sens.*, vol. 12, no. 1, p. 166, Jan. 2020.
- [83] H. Pan, D. Badawi, and A. E. Cetin, "Computationally efficient wildfire detection method using a deep convolutional network pruned via Fourier analysis," *Sensors*, vol. 20, no. 10, p. 2891, May 2020.
- [84] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.



XIWEN CHEN (Member, IEEE) received the B.S. degree in electrical engineering from Northern Arizona University, in 2019. He is currently pursuing the Ph.D. degree with the School of Computing, Clemson University. His current research interests include computer vision, image processing, information theory, driving safety analysis, image-based security protocols, and information retrieval in complex networks.



BRYCE HOPKINS (Member, IEEE) is currently pursuing the B.S. degree in electrical engineering with Clemson University, where he will pursue the M.S. degree in electrical engineering. While pursuing the bachelor's degree, he is working as an Under Graduate Research Assistant under Dr. Fatemeh Afghah with the Intelligent Systems and Wireless Networking (IS-WiN) Laboratory. His current research interests include multiagent systems, computer vision, and autonomous UAV decision making.



HAO WANG (Member, IEEE) received the B.S. degree in electrical engineering from Northern Arizona University, in 2018. He is currently pursuing the Ph.D. degree with the School of Computing, Clemson University. He has research experience in cybersecurity and environmental science. His current research interests include smart transportation, building energy management, and low-level vision.



LEO O'NEILL is currently pursuing the M.S. degree with the School of Forestry, Northern Arizona University. He spent four years as a Seasonal Wildland Firefighter for the U.S. Forest Service prior to graduate school. His research interests include the remote sensing of forest, fuel measurements, and the utility of machine learning in this field.



FATEMEH AFGHAH (Senior Member, IEEE) is currently an Associate Professor with the Department of Electrical and Computer Engineering, Clemson University, where she is also the Director of the Intelligent Systems and Wireless Networking (IS-WiN) Laboratory. Prior to joining Clemson University, she was an Associate Professor with the School of Informatics, Computing and Cyber Systems, Northern Arizona University (NAU), Flagstaff, AZ, USA, from 2015 to 2021.

She has coauthored more than 100 technical papers. Her research interests include wireless communication networks, decision making in multiagent systems, radio spectrum management, hardware-based security, and artificial intelligence in healthcare. She was a recipient of several awards, including the NSF CRII Award, in 2017, the Air Force Office of Scientific Research Young Investigator Award, in 2019, and the National Science Foundation (NSF) CAREER Award, in 2020. She serves as an Editor for several journals, including *Ad hoc Networks* journal and *Computer Networks* journal.



ABOLFAZL RAZI (Senior Member, IEEE) received the B.S. degree in electrical engineering from Sharif University, in 1998, the M.S. degree in electrical engineering from Tehran Polytechnic, in 2001, and the Ph.D. degree in electrical engineering, University of Maine, in 2013. He is currently an Associate Professor of computer science with Clemson University. He held two postdoctoral positions in the field of machine learning and predictive modeling at Duke University,

from 2013 to 2014, and Case Western Reserve University, from 2014 to 2015. His current research interests include smart connected vehicles, driving safety analysis, wireless networking, the Internet of Things, and predictive modeling. He was a recipient of several competitive awards including the Best Research of MCI, in 2008, the Best Graduate Research Assistant of the Year Award from the College of Engineering, University of Maine, in 2011, and the Best Paper Award from the IEEE/CANEUS Fly By Wireless Workshop, in 2011.



PETER FULÉ is currently a Regents' Professor with the School of Forestry, Northern Arizona University. His research interests include intersection of forests, wildfire, climate and people, using multiple research techniques including tree rings and simulation modeling to assess tree growth, and forest fires over many centuries.



ERIC ROWELL is currently an Assistant Research Professor at the Desert Research Institute. His research interests include intersection of technology and fire for improved understanding of fire effects, processes, and the role that fuels play in multiple-use wildland and prescribed fire. Working across a range of landscapes, he employs wide array of fine-scale remote sensing and statistical modeling to investigate roles of fire and plant interactions and the role of fuels inputs for computational fluid dynamics-based fire behavior models.



JANICE COEN received the B.S. degree in engineering physics from Grove City College, and the M.S. and Ph.D. degrees from the Department of Geophysical Sciences, University of Chicago. She holds positions of Project Scientist with the Mesoscale and Microscale Meteorology Laboratory, National Center for Atmospheric Research, and a Senior Research Scientist at the University of San Francisco. She studies fire behavior and its interaction with weather using coupled weather-fire CFD models and flow analysis of high speed IR fire imagery. Her recent work investigated the mechanisms leading to extreme wildfire events, fine-scale wind extrema that lead to ignitions by the electric grid, and integration of coupled models with satellite active fire detection data to forecast landscape-scale wildfires.



ADAM WATTS received the Ph.D. degree. He leads the Fire and Environmental Research Applications Team, U.S. Forest Service's Pacific Wildland Fire Science Laboratory. His backgrounds in fire ecology, atmospheric science, and development of unmanned platforms and their payloads, he uses interdisciplinary teams to approach difficult challenges in wildland fire management and implications for ecosystems and climate. His other current research projects include the fire and smoke model evaluation experiment (FASMEE) and international collaborations to find appropriate technologies and approaches for regionally and globally important fire-related challenges. His past experiences have included work as an alligator biologist, Peace Corps agroforestry volunteer, and graduate work at the University of Florida.

...