

Article

Not peer-reviewed version

Towards Near-Real-Time Wildfire Monitoring: A Deep Learning Application Using GOES Observations

Mukul Badhan , [Majid Bavandpour](#) , Kasra Shamsaei , [Dani Or](#) , [George Bebis](#) , [Neil P. Lareau](#) , [Qunying Huang](#) , [Hamed Ebrahimian](#) *

Posted Date: 4 March 2026

doi: 10.20944/preprints202603.0342.v1

Keywords: wildfire; remote sensing; geostationary operational environmental satellite (GOES); visible infrared imaging radiometer suite (VIIRS); deep learning; operational monitoring; super resolution



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Towards Near-Real-Time Wildfire Monitoring: A Deep Learning Application Using GOES Observations

Mukul Badhan ¹, Majid Bavandpour ², Kasra Shamsaei ⁵, Dani Or ², George Bebis ¹, Neil P. Lareau ³, Quanying Huang ⁴ and Hamed Ebrahimian ^{2,*}

¹ Computer Science and Engineering Department, University of Nevada, Reno, NV, USA

² Civil and Environmental Engineering Department, University of Nevada, Reno, NV, USA

³ Physics Department, University of Nevada, Reno, NV, USA

⁴ Geography Department, University of Wisconsin-Madison, WI, USA

⁵ Cloudfire Inc.

* Correspondence: hebrahimian@unr.edu

Abstract

Monitoring the progression of large wildfires in near-real-time is essential for active-fire situational awareness and emergency response management. Current satellite-based wildfire monitoring systems face a trade-off between temporal and spatial resolution: geostationary satellites such as GOES offer frequent (~5 minutes) but coarse observations (~2 km), while low earth orbit (LEO) instruments such as VIIRS provide fine spatial detail (~375 m) with limited temporal coverage (twice per day). To bridge this gap, this study introduces a deep learning (DL) approach that enables near real-time, high-resolution wildfire monitoring using GOES data. The proposed approach consists of two main steps: a segmentation step to distinguish active fire regions from background areas and a regression step to estimate the active fire pixels brightness temperature (BT) across a region of interest. The output of these steps is combined to generate a high-resolution fire location and BT maps. To train the DL model, multi-spectral GOES inputs are paired with VIIRS-derived fire observations from several wildfires across the United States. Spatial consistency between GOES and VIIRS data is achieved through parallax correction, reprojection, resampling, and per-image normalization. Ablation studies are performed to demonstrate the impact of different assumptions (e.g., background values in the VIIRS ground truth) and strategies (e.g., loss functions) throughout the development process. The results show that the proposed DL approach effectively enhances GOES imagery, improving both BT estimation and fire boundary localization. Overall, the proposed method offers a practical and scalable solution for wildfire boundary detection and thermal mapping using existing satellite systems.

Keywords: wildfire; remote sensing; geostationary operational environmental satellite (GOES); visible infrared imaging radiometer suite (VIIRS); deep learning; operational monitoring; super resolution

1. Introduction

Wildfires have emerged as one of the most destructive and fast-growing natural hazards in the U.S., with far-reaching impacts on communities, ecosystems, public health, and economy. In 2023, the U.S. recorded 56,580 wildfires that burned 2.7 million acres, destroyed at least 4,318 structures, and caused insured losses exceeding \$1.6 billion—figures that rise substantially when accounting for uninsured assets, as well as indirect and long-term losses (Jones Ben, n.d.; *National Interagency Coordination Center Wildland Fire Summary and Statistics Annual Report 2023*, n.d.). An important part of wildfire management strategy, active-fire response and decision making necessitates wildfire

monitoring solutions that are fast, accurate, cost-effective, and spatially detailed. In operational contexts, detecting fire presence alone is insufficient; what's required is the delivery of georeferenced maps that precisely locate active fires and quantify their intensity, ideally in near-real-time. Rapid fire spread makes real-time monitoring critical for situational awareness, decision making, and emergency response.

Traditional wildfire monitoring systems—such as ground-based cameras, airborne sensors, and human reports—provide localized insights but suffer from information latency, high deployment costs, coverage gaps, and logistical challenges, especially when monitoring large or inaccessible areas. In contrast, satellite remote sensing offers a scalable and consistent means of observing vast landscapes. Among available satellites, polar-orbiting platforms such as Visible Infrared Imaging Radiometer Suite (VIIRS) and Moderate Resolution Imaging Spectroradiometer (MODIS) provide global coverage with spatial resolutions that are relevant to the wildfire management problems (i.e., 375 m and 1 km respectively); however, their temporal frequency is limited to just a few overpasses per day, which is often insufficient for tracking the rapid progression of wildfires. Geostationary satellites offer a compelling alternative. While they observe only a portion of the Earth's surface, they provide continuous coverage over large regions at higher temporal resolution. The Geostationary Operational Environmental Satellite-R (GOES-R) series (GOES-16, 17, and 18), for example, continuously monitor the Americas and can capture the evolution of wildfires at 5-minute intervals. Despite its coarser spatial resolution (~2 km infrared channels), the temporal granularity of GOES imagery makes it a valuable resource for near-real-time wildfire monitoring.

Traditional approaches to wildfire detection using GOES imagery are primarily threshold-based. One of the most widely used methods is the Wildfire Automated Biomass Burning Algorithm (WF-ABBA) (Koltunov et al., 2012), which has played a foundational role in operational fire detection. WF-ABBA is a dynamic contextual thresholding algorithm that utilizes the 3.9 μm mid-infrared and 10.7 μm thermal infrared bands, along with the visible band when available. It also incorporates the 12 μm split-window band to help distinguish hot targets from opaque cloud cover. Despite its success, WF-ABBA suffers from limitations, including a high false-alarm rate during daytime, coarse spatial resolution, and an inability to resolve small or low-intensity fires. To improve reliability, alternative thresholding-based frameworks such as the GOES Early Fire Detection (GOES-EFD) system (Koltunov et al., n.d.) have been introduced, yet these methods remain constrained by predefined rules and sensitivity to atmospheric conditions.

To overcome the limitations of thresholding-based detection method, recent studies have explored deep learning (DL) models for automatic fire segmentation using satellite imagery. Toan (2019) proposed one of the earliest DL studies to leverage GOES imagery for wildfire detection. This work is notable for being the first to aim toward a real-time streaming platform for wildfire monitoring based on pixel-wise classification. The authors utilized all 16 spectral bands from GOES-R, treating each multiband image as a 3D input volume. The proposed model is a simple deep convolutional neural network (CNNs) composed of 3D convolutional layers followed by fully connected layers, trained to classify each pixel as either fire or non-fire. Building upon these advances, Zhao and Ban (2022) employed a temporal modeling strategy that applies Gated Recurrent Units (GRUs) to GOES-R time series imagery, the network captures dynamic fire behavior over time. The input to the model was preprocessed using dedicated spectral indices, including a normalized difference between the Middle Infrared (MIR) and Thermal Infrared (TIR) bands to enhance fire signals, and a custom smoke/cloud mask based on the 12 μm band (Band 15). Using VIIRS active fire points as supervisory labels, the network was trained to segment burning areas. The results showed improved early detection capability and spatial localization compared to the GOES-R, while also reducing false positive detection in mid-latitude regions. These developments demonstrate a growing shift toward data-driven and temporally-aware (i.e., considering past snapshots) fire detection systems using GOES data. However, despite these advances, past studies (e.g., (Koltunov et al., 2012; Zhao and Ban, 2022)) have either combined subsets of GOES bands using predefined equations and thresholds, which can lead to loss of information, or utilized all 16 channels as input,

which is computationally expensive and may not fully exploit their distinct spectral properties. Additionally, past studies (e.g., (Malik et al., 2021; Toan, 2019)) are restricted to limited spatial domains, which hinders generalizability and operational scalability. Moreover, many models (e.g., (Badhan et al., 2024; Saleh et al., 2024; Toan, 2019)) rely on polar-orbiting VIIRS/MODIS data for supervision without accounting for the spatial and temporal disparities between GOES and VIIRS observation. These limitations motivate the need for an improved development that can operate in near-real-time, leverage optimal spectral bands, and generalize across diverse fire-prone regions without requiring temporal sequences.

In a recent work (Badhan et al., 2024), a DL framework was developed to enhance the spatial resolution of GOES-17 wildfire imagery. The model was trained using contemporaneous GOES-VIIRS observation pairs. Its objective was to produce VIIRS-like resolution at GOES's high temporal frequency, enabling more accurate estimation of fire location and brightness temperature (BT). Although the potential of this approach was demonstrated, it tends to underestimate BT values compared to the VIIRS ground truth. To address this critical challenge, here we propose a revised and more general framework for near real-time wildfire monitoring across the continental U.S. A key novelty of this work is a two-step DL pipeline, which is designed to address the underestimation of the BT values at fire locations. In this approach, segmentation is first employed to isolate fire-affected regions, followed by regression to estimate pixel-level BT. The outputs of these two steps are then combined to enable more accurate delineation of fire boundaries and mapping of BT values. The two-step pipeline allows two separate models to specialize in their respective objectives of spatial localization and BT estimation, leading to more accurate results than the single-step approach in the prior work (Badhan et al., 2024), as will be presented later in the paper.

Furthermore, several other technical advancements are introduced over the previous work (Badhan et al., 2024) and existing literature to improve the performance of the wildfire monitoring framework. These include leveraging a broader set of GOES spectral channels to correct for environmental effects, refining preprocessing through parallax correction and normalization, and improving the data pre-processing procedure. We also systematically examine the effects of modeling choices, such as background value selection and loss function design, through a comprehensive ablation study summarized in appendix. Collectively, these advancements address the key limitations of the earlier model and enable higher-fidelity BT estimation, more accurate fire boundary delineation, and a scalable framework for continuous wildfire monitoring at GOES's 5-minute observation interval.

The remainder of the paper is organized as follows. Section 2 describes the satellite data sources, including VIIRS and GOES, and outlines the preprocessing steps for dataset construction and alignment. Section 3 details the proposed two-step segmentation-regression approach, including the DL architectures, loss functions, evaluation metrics, and ablation study setup. Section 4 presents the training and testing results, including both quantitative and visual evaluations, and discusses the impact of different background values and loss functions through ablation studies. The performance of the new framework is compared with the previous work to demonstrate the improvement in both fire boundary and BT estimation. Finally, Section 5 concludes the paper with a summary of key findings and suggestions for future work.

2. Data Sources and Preprocessing

In this study, a wildfire dataset is developed to support training the supervised DL for active fire monitoring. The dataset is constructed from open-source satellite imagery and products, including low spatial resolution GOES imagery used as the model input and the VIIRS Active Fire Product serving as the ground truth. This work advances the prior efforts (Badhan et al., 2024) by improving data quality, diversifying data through expanding spatial and temporal fire event coverage, incorporating additional spectral features, and implementing corrections to reduce spatial and temporal misalignments. Data sources, area of study, and the dataset construction and corrections are presented in this section.

2.1. Products and Data Sources

Two primary satellite products are used to construct the dataset: VIIRS I-band Active Fire Product (Schroeder et al., 2025) and GOES ABI Level 1B Radiance product (ABI-L1B-Rad) (Hager and Lemieux, n.d.). VIIRS I-band Active Fire Product provides near-real-time active fire detections. VIIRS is onboard the Suomi NPP and NOAA-20 satellites, both operating in low earth polar orbits with approximately 12-hour revisit time. Each satellite provides two 12-hours interval observations per day (one during daytime and one at night for most locations). The VIIRS I-band Active Fire Product detects fire locations using an algorithm described in the Algorithm Theoretical Basis Documents (ATBD) (Schroeder et al., 2014; Ti, 2016; *Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m Active Fire Detection and Characterization Algorithm Theoretical Basis Document 1.0*, 2016), which integrates multiple spectral bands to identify thermal anomalies. The VIIRS fire detection algorithm identifies active fires through a sequence of fixed threshold and contextual tests. The algorithm uses BT values from channels I4 (3.55–3.93 μm) and I5 (10.5–12.4 μm), reflectance from I1 (0.60–0.68 μm), I2 (0.846–0.885 μm), and I3 (1.58–1.64 μm), along with geometric parameters such as solar and view zenith angles and relative azimuth.

The algorithm begins by masking optically thick clouds, and most water bodies using BT value and reflectance thresholds from I1, I2, and I3. To reduce false positives, additional exclusion tests are applied to identify and flag pixels affected by sun glint or bright surfaces (e.g., urban areas or sands) using geometric parameters and reflectance-based criteria. Fire candidate pixels are initially flagged based on fixed thermal thresholds in the I4 and I5 infrared channels capturing strong thermal anomalies typical of active fires. For each candidate pixel, a dynamic background is estimated from surrounding valid pixels, which are not obscured by clouds, are not already flagged as fire, are not marked as low quality, and share the same surface classification (land or water) as the candidate. This background is calculated as the median I4 value within a 501×501 pixels window. Contextual fire detection tests then refine the classification by comparing the candidate's thermal signal to the background using a dynamically expanding window (11×11 to 31×31 pixels), computing both absolute (e.g., I4 BT value) and relative (e.g., I4–I5) BT differences. If insufficient valid pixels are available, the pixel is labeled "unclassified". Candidates that pass contextual tests are assigned a confidence level: low, nominal, or high, based on the strength of their anomaly, viewing geometry, and potential contamination. Secondary tests further identify residual low-confidence fires and screen out remaining false positives, particularly in sun glint zones and regions affected by the south Atlantic magnetic anomaly. The resulting detections are provided in the CSV format and include fire coordinates, confidence levels, BT values from the I4 and I5 channels, and Fire Radiative Power (FRP). This product serves as the ground truth in this study.

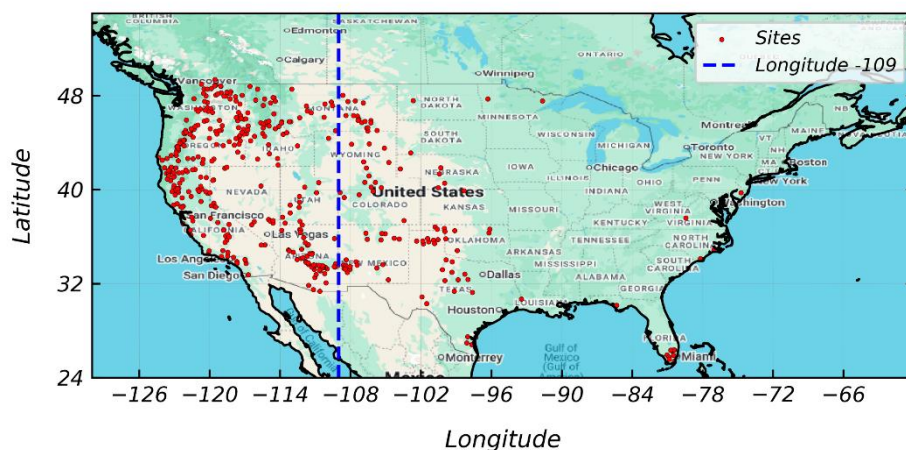
GOES ABI-L1B-Rad, the second satellite product, provides radiances measurements from the Advanced Baseline Imager (ABI). The radiance values are converted to BT values in this study. The ABI sensor captures data across 16 spectral bands, spanning from the visible spectrum (0.47 μm) to the infrared (13.3 μm), some of which are suitable for thermal anomaly detection. This data is provided in the NetCDF format and sourced from the GOES-East and GOES-West satellites. Together, the GOES satellites offer broad coverage across the western hemisphere and enable continuous multispectral observation with temporal resolutions of 5 minutes. While these observations are more frequent than those from VIIRS, they have lower spatial resolution due to the higher orbital altitude of GOES. GOES-16 operates as GOES-East; GOES-18, which replaced GOES-17 on January 2023, operates as GOES-West ("Earth from Orbit: NOAA's GOES-18 is now GOES West," n.d.). To ensure uninterrupted spatial and temporal coverage, the data from GOES-16, GOES-17, and GOES-18 are included by selecting the appropriate satellite based on each wildfire's location and timestamp.

We considered three spectral bands from the GOES product as input to the DL model towards improving the accuracy of active fire prediction. Specifically, Bands 7 (3.80–4.00 μm), 14 (10.8–11.6 μm), and 15 (11.8–12.8 μm) are chosen to leverage complementary thermal characteristics across the middle and thermal infrared ranges. Band 7 (Middle Infrared, MIR) is sensitive to temperature

associated with active fires and is commonly used in operational fire detection products (Schmidt, n.d.). Band 14 and Band 15 (Thermal Infrared, TIR) provide BT values that help remove cloud coverage (Barducci et al., 2004). Zhao and Ban (2022) introduced a multi-spectral fire detection approach that combines GOES Bands 7, 14, and 15 using a set of equations to identify active fires through the normalized difference between Bands 7 and 14, followed by cloud removal using Band 15. Their approach reduces the multi-band input to a single-channel representation, which is then used as input to a DL model for fire segmentation. While this approach produces a single-channel input for segmentation that captures the general location of fire activity, the outputs often do not align precisely with ground truth fire perimeters. In the proposed DL framework in this study, all three GOES Bands 7, 14, and 15 are provided as individual inputs to the model, allowing more intricate and adaptive relationships to be learned for detecting fire locations rather than relying on a predefined combination of bands. By leveraging multispectral context, the model is anticipated to achieve higher accuracy in detecting and mapping fires compared to a single-band method.

2.2. Dataset Construction and Preprocessing

To support high-quality supervised learning and improve generalization across diverse landscapes and fire conditions, this study selects wildfire events spanning the continental U.S. (CONUS). The selected events cover a wide range of geographic regions and various times of year, supporting generalized model development. The wildfire events are selected from the Wildland Fire Incident Locations dataset provided by the National Interagency Fire Center (National Interagency Fire Center (NIFC), n.d.) database, a public platform that tracks major fire incidents. To ensure the fires are clearly visible in satellite images, only fires that burned more than 10,000 acres are included. This threshold provides a balance between the number of events and the spatial prominence of fire signals. Using this criterion, 208 wildfire events from 2019 to 2024 are selected. Figure 1.a depicts the spatial distribution of these fire events, and Figure 1.b summarizes their temporal distribution by month. Together, these figures highlight the broad geographic and temporal coverage of the dataset. In Figure 1a, a dashed blue line at longitude -109° marks the boundary (as defined in this study) between GOES-East and GOES-West coverage zones (Schmit et al., 2017).



(a)

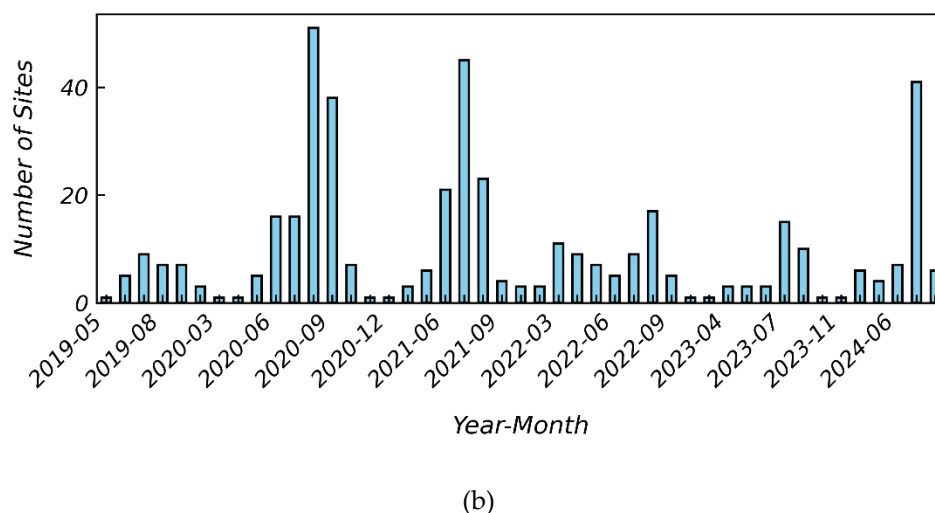


Figure 1. (a) Geographic distribution of wildfire sites used in this study, mapped across the continental U.S. Red dots indicate wildfire locations. The dashed blue line at longitude -109° separates regions observed by GOES-West (left) and GOES-East (right). (b) Histogram of wildfire occurrences by year and month from 2019 to 2024, illustrating the temporal distribution of fire events in the dataset.

To construct the dataset, a region of interest (ROI) is defined for each wildfire event. For simplicity and consistency, we fix the ROI at 1.2 degrees in both latitude and longitude and center it on the wildfire site's central coordinates. Within each ROI, data from GOES and VIIRS are extracted and paired to form a snapshot, which is defined as a contemporaneous GOES-VIIRS image pair captured at a specific timestamp. To ensure spatial alignment, GOES and VIIRS are reprojected to a common geographic coordinate system using an appropriate EPSG code (OGP, 2012), a standardized identifier that specifies a particular coordinate reference system (e.g., UTM zone based on the event's longitude). GOES is resampled to match the 375-meter spatial resolution of VIIRS. VIIRS active fire (hotspot) points are resampled into a 2D BT map, and nearest-neighbor interpolation (Badhan et al., 2024) is applied to fill gaps between detected fire pixels. Since the VIIRS product includes only detected active fire locations, a constant background BT value is assigned to all remaining pixels within the ROI. This value is selected to preserve the bimodal distribution between background and fire pixels and to maintain physical relevance. Additional details on the selection and justification of this background value are provided in the Section 3. A saturation condition in VIIRS, where I4 pixels at the core of intense fires either reach the nominal maximum of 367 K or fold to artificially low values (~ 208 K), occurs due to the extreme thermal intensity of active fires and the sensor's dynamic range. This saturation condition is corrected using the procedure described in (Badhan et al., 2024), to ensure physically consistent VIIRS inputs. After applying the correction procedure described in (Badhan et al., 2024), the effective VIIRS range becomes 283–367 K. In comparison, GOES exhibits broader ranges of 209–413 K (Band 7), 202–342 K (Band 14), and 198–342 K (Band 15). These differences reflect sensor-specific characteristics, including orbital configuration, spectral response, and spatial resolution, as well as the fact that GOES covers both fire and background. It should be noted that during alignment with GOES, some VIIRS hotspots are not detected at the exact pairing timestamp but instead appear within short intervals of 2 to 10 minutes. To correct this temporal inconsistency, all VIIRS hotspots within a 10-minute window are aggregated, and the latest timestamp is assigned prior to resampling into the resulting BT maps.

Lastly, the parallax effect (Pestana and Lundquist, 2022) in GOES ABI imagery is corrected to improve geolocation accuracy and enable reliable alignment with higher-resolution VIIRS reference. The parallax effects arise due to the satellite's off-nadir viewing geometry, which causes elevated terrain features, such as mountains, to appear spatially displaced from their true locations. The dislocation grows with both elevation and viewing angle. In mountainous regions, this can lead to

geolocation errors that can exceed 5 km. The consequence of this parallax is a misalignment between GOES imagery and ground truth datasets of VIIRS. Standard ABI products project each pixel from the ABI Fixed Grid (defined by scan and elevation angles) onto the Earth coordinates. This is done using a fixed satellite position and a smooth Earth ellipsoid model (GRS80)(Moritz, 2000). The GRS80 model approximates Earth as an oblate spheroid, with a slightly larger radius at the equator than at the poles. However, the projection ignores actual land elevation and assumes all terrain lies on the reference ellipsoid. This simplification causes parallax effect.

To correct the parallax effect, a geometric orthorectification correction method presented by Pestana and Lundquist (2022) is applied. Considering terrain offsets from the GRS80 ellipsoid, the GOES-R projection process is modified using a digital elevation model (DEM) to create orthorectified ABI images. Figure 2 illustrates the orthorectification method, which computes the intersection between line-of-sight (r_s) vectors, extending from the satellite sensor through the ABI Fixed Grid image plane, and the Earth's surface, modeled using DEM-based elevations. The satellite is positioned at a constant distance H from the Earth's center, directly above the sub-satellite point, which is the point on the Earth's surface located directly beneath the satellite, at longitude λ_0 . For each DEM grid cell, defined by geodetic coordinates (latitude ϕ , longitude λ , and surface elevation Z), the geocentric latitude (ϕ_c) is defined as the angle between the equatorial plane and the line connecting the Earth's center to the point on the surface. It differs from the geodetic latitude, which is measured relative to the ellipsoid normal. The geocentric latitude can be computed as

$$\phi_c = \text{atan} \left(\frac{r_{pol}^2}{r_{eq}^2} \cdot \tan(\phi) \right) \quad (1)$$

where r_{eq} is the equatorial radius of earth (semi-major axis) and is approximately equal to 6.378×10^3 km and r_{pol} is the polar radius of the earth (semi-minor axis) and is approximately equal to 6.356×10^3 km. Next, the Earth-centered distance (r_c) to the surface point, is computed as

$$r_c = r_{pol} \sqrt{\frac{1}{1 - e^2 \cos^2(\phi_c)} + Z} \quad (2)$$

where, e is eccentricity of the reference ellipsoid, which quantifies how much the Earth's shape deviates from a perfect sphere (e.g., 0.0818 for GRS80) and Z is surface elevation above the ellipsoid in km.

The surface point is then expressed in the satellite-centered cartesian coordinate system, where S_x , S_y , and S_z are the distance from the satellite location as origin to the surface point and are calculated as

$$\begin{aligned} S_x &= H - r_c \cos(\phi_c) \cos(\lambda - \lambda_0) \\ S_y &= -r_c \cos(\phi_c) \sin(\lambda - \lambda_0) \\ S_z &= r_c \sin(\phi_c) \end{aligned} \quad (3)$$

where H is approximately equal to 42.164×10^3 km for GOES, and λ_0 is -75° for GOES-East and -137° for GOES-West. Finally, the coordinates are projected onto the ABI Fixed Grid as scan angle (x) and elevation angle (y), using:

$$\begin{aligned} x &= \text{asin} \left(\frac{-S_y}{\sqrt{S_x^2 + S_y^2 + S_z^2}} \right) \\ y &= \text{atan} \left(\frac{S_z}{S_x} \right) \end{aligned} \quad (4)$$

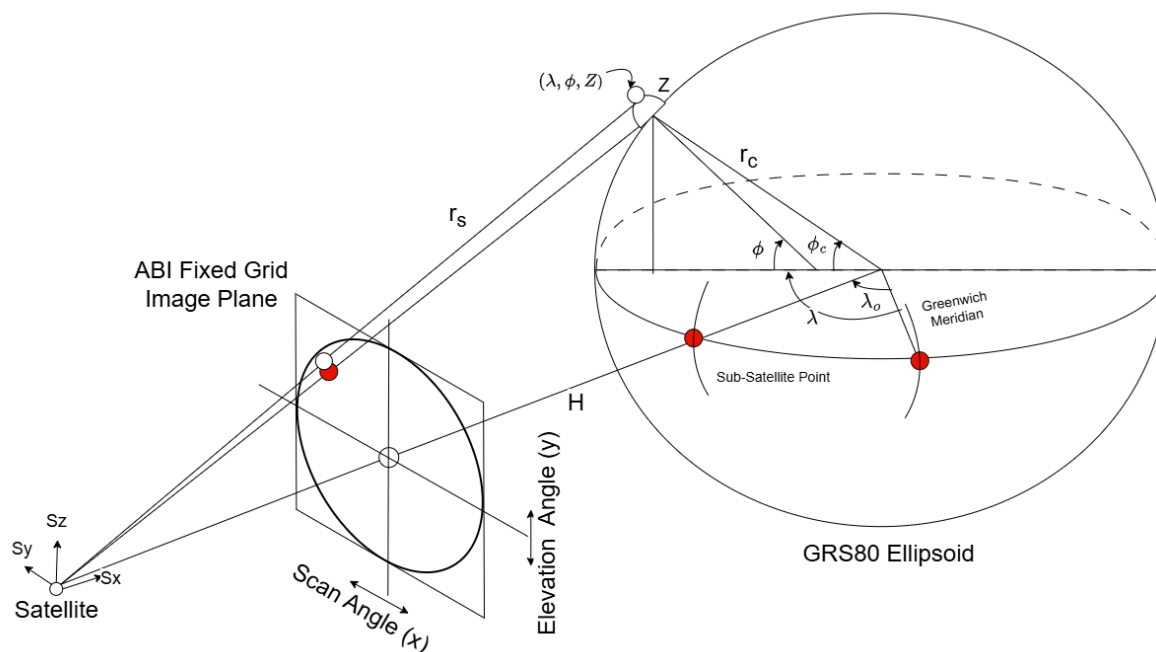


Figure 2. The diagram illustrates the transformation from geodetic coordinates (λ, ϕ, Z) on the GRS80 reference ellipsoid to the ABI fixed grid image plane. The satellite is positioned at a constant distance H from Earth's center, directly above the sub-satellite point at longitude λ_0 .

This transformation creates a direct link between terrain-aware surface locations and the satellite's viewing geometry, enabling correction of the apparent location of each pixel in GOES imagery. It is applied to each fire event in our dataset to obtain accurate surface locations by accounting for both elevation and satellite viewing angle. To prevent losing edge pixels that may shift during correction, a 0.2-degree buffer is added around the ROI before applying the transformation. After correction, the buffer is removed to restore the original ROI and maintain alignment with VIIRS data. Figure A-1 is provided in Appendix A to illustrate the effects of parallax correction.

3. Proposed Deep Learning for Wildfire Remote-sensing Enhancement Approach (DL-WREN)

To enhance the resolution of GOES imagery for active fire identification and BT estimation, a two-step approach including segmentation (Jadon, 2020) and regression (Wang et al., 2021) is developed using VIIRS data as ground truth. This two-step approach addresses a key challenge identified in the single-step approach in the prior work (Badhan et al., 2024): when background suppression and BT estimation are performed simultaneously, the resulting BT values at fire locations are often underestimated. In the proposed approach, the GOES input is processed independently through segmentation and regression models. The segmentation process detects active fire pixels and filters out background noise, producing a binary fire mask. The regression process predicts BT values across the image, independently of the segmentation output. The final BT prediction is obtained by applying a pixel-wise multiplication between the fire mask and the regression output, ensuring that BT values are retained only at predicted fire locations. This decoupling allows each model to specialize in its respective objective, i.e., spatial localization in segmentation and BT estimation in regression, leading to more accurate results than the single-step approach in the prior work (Badhan et al., 2024). The proposed approach differs from single-stage methods by enabling modular optimization and flexibility in model design. It allows the integration of task-specific architectures, normalization schemes, and loss functions tailored to each sub-task.

Segmentation is used to classify pixels within the ROI as either active fire or background by learning spatial and spectral patterns from GOES imagery. A DL-based segmentation approach is used instead of simple thresholding because high BT values alone are not uniquely indicative of fire. Elevated BT values can occur in valley regions where lower elevations exhibit warmer surface temperatures than surrounding higher terrain, and in some cases these background temperatures can even exceed the BT values of small or cooler fires. This behavior is observed in our dataset and makes threshold-based methods unreliable. The DL-based segmentation approach can leverage contextual cues to better distinguish fire pixels from confounding background signals. It should be noted that the segmentation in this study differs from typical segmentation setups, where ground truth is directly derived from the input data. Here, VIIRS observations, collected from a different sensor, serve as approximate references rather than perfectly aligned labels. This cross-sensor discrepancy introduces uncertainty into the training process and should be considered when evaluating model performance.

To estimate accurate BT values in active fire regions, a separate regression step is introduced to address the imbalance caused by the dominance of background pixels, which comprise approximately 98% of the ROI on average. In the prior work (Badhan et al., 2024), the model was found to consistently underestimate BT values in fire regions due to the disproportionate influence of background pixels in both inputs and ground truth. This imbalance shifts the optimization toward minimizing errors in non-fire areas, resulting in suppressed BT values even where active fire is present. To mitigate the effects of this background-driven bias and to stabilize the mapping between GOES and VIIRS BT fields, two supporting strategies are introduced: z-score data normalization and background-value adjustment in the VIIRS ground truth. These strategies do not resolve the imbalance itself, but they help prevent the regression model from being pulled toward low background values and instead encourage the predicted BT in fire regions to approach their true VIIRS fire-region values.

Per-image z-score normalization is applied to both GOES inputs and VIIRS ground truth to enhance generalization across daytime and nighttime conditions and to ensure physically meaningful BT predictions. Each image is transformed by subtracting its mean BT and dividing by its standard deviation, thereby standardizing the data to a mean of zero and a standard deviation of one, as described in Eq. (5).

$$x_i^{norm} = \frac{x_i - \mu}{\sigma} \quad (5)$$

where x_i is the original BT value at i^{th} pixel location, μ is the mean BT values across the image, and σ is the standard deviation of BT values across the image.

This normalization process preserves internal contrast between fire and background while removing absolute differences in BT associated with time of day or scene-wide variability. For example, a nighttime image with generally lower observed BT values and a daytime image with higher BT values will both be rescaled so that their internal contrasts, such as between fire and background, remain comparable. In the previous work (Badhan et al., 2024), global min-max normalization was used, which limited the model's ability to generalize across varying conditions. In contrast, per-image z-score normalization allows the model to focus on relative BT patterns (e.g., identifying fire regions), rather than being biased by global shifts in the ambient BT value. This normalization approach is motivated by two key observations: (1) both GOES Band-7 and VIIRS fire regions exhibit a clear day-night bimodality in mean BT values, with systematically higher temperatures during the day and lower at night, and (2) contemporaneous GOES and VIIRS observations show a linear correlation in their mean BT values, as illustrated in Figure A-2 in Appendix A using data from the Dixie Fire between July 14 and August 15. These patterns guided the use of per-image z-score normalization, since the per image mean and standard deviation used for normalization are not available at inference time. To enable proper de-normalization during prediction, the regression model is trained to estimate these per-image statistics directly from the GOES input, as described in Section 3.1.

To prevent suppression of BT values in the predicted active fire regions, a value of 240 K is assigned to the background pixels in the VIIRS ground truth for training. In the previous work (Badhan et al., 2024), background pixels were set to 0 K, which introduced a mismatch with the GOES inputs, where background BT values are typically much higher. This discrepancy, combined with the numerical dominance of background pixels, caused the model to favor lower overall predictions, especially in fire regions. The choice of 240 K as a background value helps to reduce this bias by providing a more realistic scenario, encouraging the model to produce higher BT outputs where appropriate. However, if the background value is set too high (e.g., close to fire BTs), the per-image z-score normalization compresses the contrast between fire and background, leading the model to learn unimodal predictions without clearly separating fire regions. Although the ground truth retains bimodality, the model's predictions do not, resulting to poor fire-background discrimination during inference. Therefore, the value of 240 K represents a practical compromise: it is high enough to avoid suppressing predicted fire-region BT values, yet low enough to preserve separation between fire and background in the model's output. The rationale for this choice is further supported by an ablation study on varying background value, which is presented in Appendix C.

In summary, the proposed two-step approach addresses critical limitations of prior single-step methods by decoupling fire localization and BT estimation into separate segmentation and regression stages. This design enables targeted strategies, such as per-image z-score normalization and background value adjustment to overcome the dominance of background/non-fire pixels and improve prediction accuracy. The segmentation model specializes in spatial localization of fire pixels, while the regression model focuses on predicting accurate BT values. Detailed descriptions of the model architecture, loss functions, and evaluation metrics are provided in the following subsections.

3.1. DL Architecture and Model Elements

The segmentation and regression models in this study are based on the U-Net architecture (Weng and Zhu, 2021), an encoder–decoder framework originally designed for biomedical image segmentation. The U-Net used in this study features a symmetric structure, in which the encoder compresses spatial information while capturing abstract features, and the decoder restores resolution while refining predictions. A key characteristic of U-Net is the use of skip connections that pass high-resolution feature maps from the encoder to the decoder, enabling better localization. The architecture is constructed using common building blocks. 2D convolutional layers (Conv2d) (Uchida et al., 2018) extract spatial features using learnable 3×3 filters while Rectified Linear Units (ReLU) (Agarap, 2019) introduce non-linearity by zeroing out negative activations. Batch Normalization (BatchNorm2d) (Ioffe and Szegedy, 2015) normalizes the output of convolutional layers across each mini-batch to stabilize and accelerate training. Max pooling (MaxPool2d) (Zhao and Zhang, 2024) reduces spatial dimensions by selecting the maximum value in each 2×2 window, enabling spatial down-sampling while preserving important features. In the decoder, transposed convolution (TransposeConv) (Sahito et al., 2023) is used to increase spatial resolution as it learns to up-sample feature maps by applying learned kernels that reverse the effects of down-sampling. Adaptive average pooling (AdaptiveAvgPool2d) (Yang et al., 2024) reduces a feature map to a fixed spatial size by averaging over spatial regions, enabling a consistent output regardless of the input dimensions. To streamline the design, a modular unit called DoubleConv block is used throughout the encoder and decoder. Each DoubleConv block consists of two sequential 3×3 Conv2d layers, each followed by ReLU and BatchNorm2d. These blocks extract and refine spatial features and form the backbone of each stage in the U-Net.

The segmentation model uses a deep U-Net variant tailored for pixel-wise binary classification. It maps $128\times 128\times 3$ GOES input images to a 128×128 binary fire mask. The encoder consists of four stages, each built with a DoubleConv block followed by MaxPool2d, progressively halving the spatial resolution and doubling the feature depth until reaching a $1024\times 8\times 8$ representation at the bottleneck. The decoder mirrors this structure: TransposeConv layers up-sample the features and skip connections concatenate encoder outputs to recover spatial detail. Each decoder stage includes a

DoubleConv block to refine the combined features. A final 1×1 Conv2d projects the 64-channel decoder output to a single channel, and a sigmoid activation produces a binary mask indicating active fire regions. Batch normalization is used throughout the network to ensure stable feature learning. The complete architecture of the segmentation model is detailed in Table B-1, which outlines the layers, their configurations, and output dimensions at each stage of the network.

The regression model adopts a shallower U-Net variant designed to estimate three outputs per image: a normalized BT map ($1 \times 128 \times 128$), the mean (μ), and the standard deviation (σ) of BT values. These per-image statistics allow for recovery of BT values by z-score de-normalization during inference. The encoder consists of two DoubleConv blocks, each followed by MaxPool2d, reducing the resolution while increasing feature richness. At the bottleneck, the feature depth is expanded to 256. The decoder performs up-sampling via TransposeConv and recovers spatial detail using skip connections. We have removed BatchNorm2d from the last layer. Unlike the segmentation model, BatchNorm2d is removed from the final decoder stage since our experiments have shown that it enforces batch-wise statistics that often conflict with regression output. A 1×1 Conv2d layer outputs the normalized BT map. Separately, an AdaptiveAvgPool2d compresses the decoder's final feature map to $64 \times 1 \times 1$, which is passed through two fully connected layers: one for μ and another for σ . The predicted σ is passed through an exponential function to ensure it remains positive. The complete architecture is described in Table B-2 of Appendix B, which lists each layer and its function in the regression pipeline.

3.2. Loss Functions

For the segmentation task, Binary Cross-Entropy (BCE) (Mao et al., 2023) loss, as defined in Eq. (6), is used to train the model to predict a binary fire mask. BCE is commonly used in binary segmentation problems and derived from the Bernoulli distribution. It minimizes the difference between predicted probabilities and ground truth labels.

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_{b,i} \log \hat{p}_i + (1 - y_{b,i}) \log(1 - \hat{p}_i)] \quad (6)$$

where $y_{b,i} \in \{0,1\}$ represents the binarize label of the i^{th} pixel in the VIIRS ground truth image (1: active fire and 0: no fire), $\hat{p}_i \in [0,1]$ represents the model-predicted probability of fire at that pixel, and N represents the total number of pixels across the image. The loss function is calculated after flattening the two-dimension image into one-dimension array.

For the regression task, a statistically supervised, region-weighted, root means square error loss ($L_{SS-RWRMSE}$) is used to estimate pixel-wise BT values. Here, "statistically supervised" denotes additional supervision on the mean and standard deviation, helping align the overall BT distribution between prediction and ground truth. This loss not only penalizes raw prediction error but also enforces agreement in spatial statistics (mean and standard deviation) between predicted and ground truth BT values. In addition, the ranges of predicted BT values are expected to remain consistent with VIIRS observations, which saturate at 367 K. The objective is to enhance GOES imagery such that its outputs closely mimic the dynamic range and fidelity of VIIRS. This loss encourages accurate prediction in fire regions while maintaining consistency with BT statistics. The total loss contains three terms:

$$L_{SS-RWRMSE} = L_{RWRMSE} + L_{MRMSE} + L_{SDRMSE} \quad (7)$$

where, L_{MRMSE} measures the root mean square error between the ground truth mean (μ) and the predicted mean ($\hat{\mu}$) and L_{SDRMSE} measures the root mean square error between the ground truth standard deviation (σ) and the predicted standard deviation ($\hat{\sigma}$).

$$L_{MRMSE} = \sqrt{\frac{\sum_{i=1}^B (\mu_i - \hat{\mu}_i)^2}{B}}, \quad L_{SDRMSE} = \sqrt{\frac{\sum_{i=1}^B (\sigma_i - \hat{\sigma}_i)^2}{B}} \quad (8)$$

where B represents the batch size (number of images used in each batch). The Region-Weighted RMSE loss (L_{RWRMSE}), addresses the class imbalance between fire and background/non-fire regions by assigning different weights to their respective errors. Specifically, higher weights can be given to fire pixels to emphasize their importance, while lower weights are assigned to background pixels to prevent them from dominating the loss:

$$L_{RWRMSE} = W_F \cdot L_{FRMSE} + W_B \cdot L_{BRMSE} \quad (9)$$

where W_F and W_B are hyperparameters to be tuned for the fire and background/non-fire regions, respectively. These weights allow the model to prioritize errors in fire-affected areas, which are critical for wildfire monitoring, while still maintaining stability in the surrounding background. Physically, this formulation reflects the fact that small but intense fire regions have disproportionate importance compared to the much larger background. The sensitivity of the loss function to W_F and W_B is further analyzed through the ablation studies, where their impact on prediction accuracy is systematically examined.

The two loss components, L_{FRMSE} (fire root mean square error loss) and L_{BRMSE} (background root mean square error loss), are defined as:

$$L_{FRMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2 \cdot I_i}{\sum_{i=1}^N I_i}} \quad I_i = \begin{cases} 1, & y_i \neq 0 \\ 0, & y_i = 0 \end{cases} \quad (10)$$

$$L_{BRMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2 \cdot (1 - I_i)}{\sum_{i=1}^N (1 - I_i)}}$$

where y_i is the ground truth BT of the i^{th} pixel, \hat{y} is the predicted BT value at that pixel, N is the total number of pixels, and I_i is a binary fire mask indicating whether a pixel belongs to the fire region (nonzero y_i). Physically, L_{FRMSE} quantifies the prediction error only within fire regions, ensuring that localized BT within fire region is captured accurately. In contrast, L_{BRMSE} measures the prediction error in the much larger background, stabilizing the overall reconstruction and preventing spurious noise. Together, these terms ensure that the contribution from sparse but critical fire pixels is not overwhelmed by the background.

When no distinction is made between fire and background regions (i.e., without assigning separate weights to each), the regression objective simplifies to a global RMSE (L_{GRMSE}), where all pixels contribute equally. In this case, the loss no longer separates fire- and background-specific errors but instead evaluates overall prediction accuracy across the entire image.

$$L_{GRMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (11)$$

This global formulation can be seen as a baseline version of the loss, where the image is treated as a single unit without differentiating between fire and background regions.

3.3. Evaluation Metrics

The performance of the segmentation and regression models is evaluated using a set of metrics designed to assess both the accuracy of the predicted fire locations and the BT values at those locations. For segmentation, evaluation metrics are used to quantify how well the predicted shape of active fire regions matches the corresponding regions in the ground truth. For regression, the metrics evaluate the accuracy of BT values within active fire regions and the model's ability to suppress background BT values.

Intersection over Union (IOU) measures the agreement between the prediction and ground truth by quantifying the degree of overlap of the fire area between ground truth and model prediction. The

E_{IOU} values range from 0 to 1, where 0 represents no overlap between the VIIRS ground truth and the model-predicted fire regions, 1 indicates perfect overlap (i.e., no error), and intermediate values correspond to partial disagreement in fire extent and shape. Prior to evaluation, the predicted probability map is binarized using Otsu's thresholding (Xu et al., 2011), an adaptive method that selects an optimal threshold value based on the distribution of predicted intensities. After binarization, all pixels above the threshold are set to 1 (fire), and those below are set to 0 (no fire). The E_{IOU} metric is then computed as

$$E_{IOU} = \frac{\sum_{i=1}^N y_{b,i} \hat{y}_{b,i}}{\sum_{i=1}^N [y_{b,i} + \hat{y}_{b,i} - y_{b,i} \hat{y}_{b,i}]} \quad (12)$$

where $\hat{y}_{b,i}$ and $y_{b,i}$ represent the predicted and ground truth binarized presence (1) or absence (0) of fire at the i^{th} pixel location, respectively.

The fire and background root mean square error (E_{FRMSE} and E_{BRMSE}) quantify the prediction error of BT values in distinct spatial regions based on the ground truth. The term E_{FRMSE} measures the prediction error in the active fire regions (i.e., where ground truth has non-zero BT values), and E_{BRMSE} measures the prediction error in background regions (i.e., where ground truth has zero BT values). The E_{FRMSE} and E_{BRMSE} values start from 0, where 0 represents zero error in the predicted BT values (i.e., perfect prediction) while larger values correspond to more error in the predicted BT values. Both metrics are expressed in K and are computed on denormalized outputs, allowing the error magnitudes to retain physical interpretability.

These metrics are calculated using following equations.

$$E_{FRMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2 I_i}{\sum_{i=1}^N I_i}} \quad I_i = \begin{cases} 1, & y_i \neq 0 \\ 0, & y_i = 0 \end{cases} \quad (13)$$

$$E_{BRMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2 (1 - I_i)}{\sum_{i=1}^N (1 - I_i)}}$$

It should be noted that although these metrics are conceptually similar to the loss components defined in Eq. (10), they are used here solely for evaluation rather than optimization. Physically, E_{FRMSE} captures how accurately the model reconstructs BT values within active fire regions, where temperatures are high and spatially localized. In contrast, E_{BRMSE} measures prediction stability in the non-fire background, where BT values should remain near ambient levels.

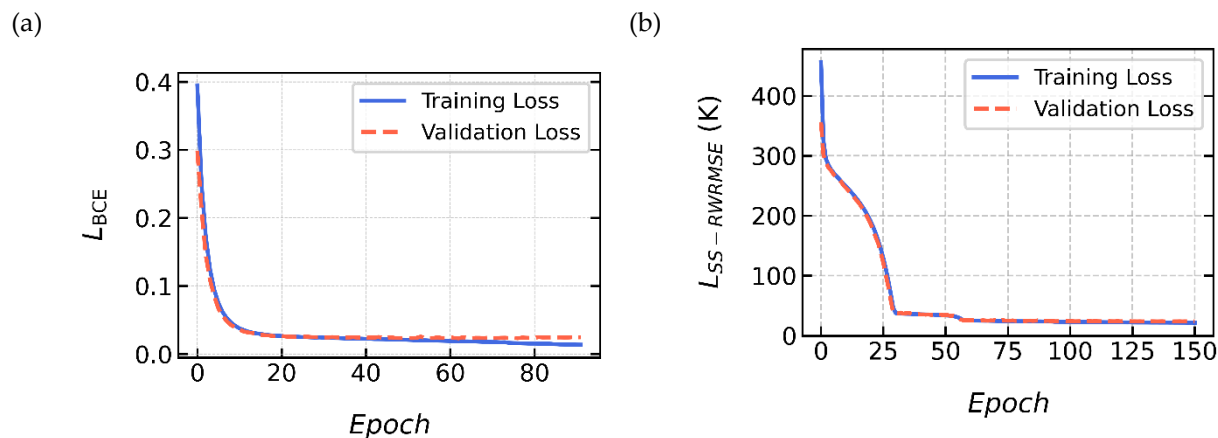
4. Results

4.1. Training

All models are implemented using PyTorch (v2.0.1) with CUDA 11.7 and cuDNN 8.5 support. Training is performed on a workstation equipped with NVIDIA RTX A6000 GPUs (each with 48 GB of VRAM), running CUDA version 12.0. Separate hyperparameters are used for segmentation and regression tasks, tuned individually to optimize their respective performance. The regression model is trained for 150 epochs with a batch size of 32 and a learning rate of 3×10^{-5} , requiring approximately one hour. For the segmentation task, a learning rate of 8×10^{-5} is used, with identical batch size and epoch count. The segmentation model completes training in approximately 55 minutes. The Adam optimizer (Kingma and Ba, 2017) is used for both tasks due to its efficient handling of sparse gradients and adaptive learning capabilities. To enhance generalization and training stability, the learning rate is reduced by a factor of 0.5 if the validation loss plateaus, with a minimum improvement threshold of 1×10^{-5} . Early stopping is employed to prevent overfitting, halting training if the validation loss does not improve for 30 consecutive epochs.

Each sample in the dataset is derived from a ROI corresponding to a wildfire event snapshot. These ROIs are divided into fixed-size patches of 128×128 pixels (approximately 48×48 km) using a sliding-window technique with no overlap. To ensure complete spatial coverage, windows at the edges of each ROI that are smaller than 128×128 are adjusted by extending into the preceding region, allowing full inclusion without missing any portion of the scene. Additionally, a filtering step is applied to exclude samples with fires that cannot be reliably detected by GOES, as their inclusion has been found to degrade DL model performance. Specifically, a sample is discarded if the fire observed by VIIRS has an active fire coverage of less than 0.37% of the total area (approximately 506.35 km²) and a total fire FRP, which quantifies the rate of radiant heat energy emitted by the fire, below 600 MW. Whereas the 10,000-acre threshold described in Section 2.2 is applied at the event selection stage to ensure the inclusion of large and clearly visible wildfires, this filtering step is applied at the snapshot level to remove weak or uninformative samples within those events. The dataset consists of 17,061 reference samples collected from a diverse set of wildfire events across the CONUS. A predefined 64/16/20 split ratio is applied, and samples are randomly assigned to the training, validation, and testing subsets, resulting in 10,918, 2,730, and 3,413 samples, respectively. To confirm that the random assignment does not bias the results, multiple shuffling seeds are tested while maintaining the same split ratio. All seeds produced similar model performance, indicating that the dataset's diversity makes the results robust to the specific random split.

To increase diversity and reduce overfitting, basic data augmentation techniques are applied during training. These include random horizontal and vertical flipping of all input channels and corresponding ground truth. The learning curves for both regression and segmentation tasks are shown in Figure 3. These plots demonstrate smooth convergence of the training and validation losses, with no signs of overfitting, underscoring the effectiveness of the training strategy and regularization techniques.



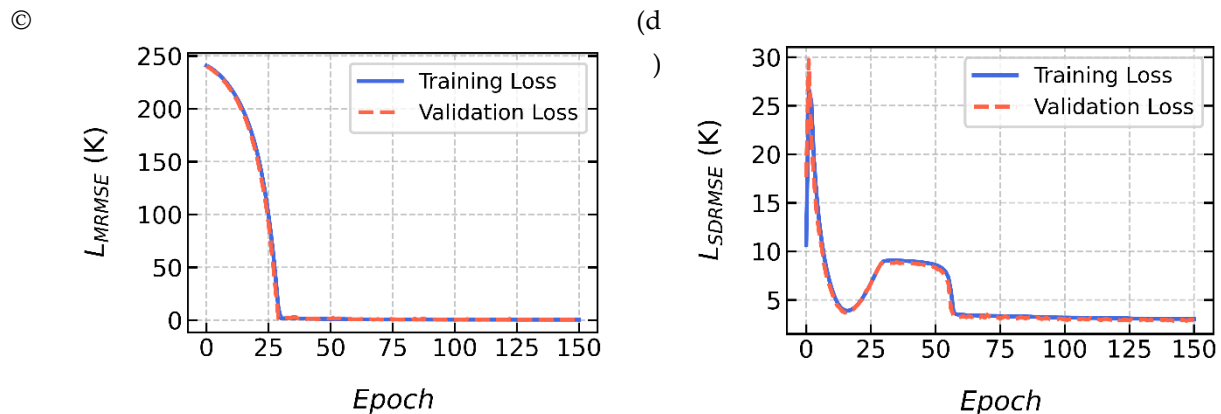
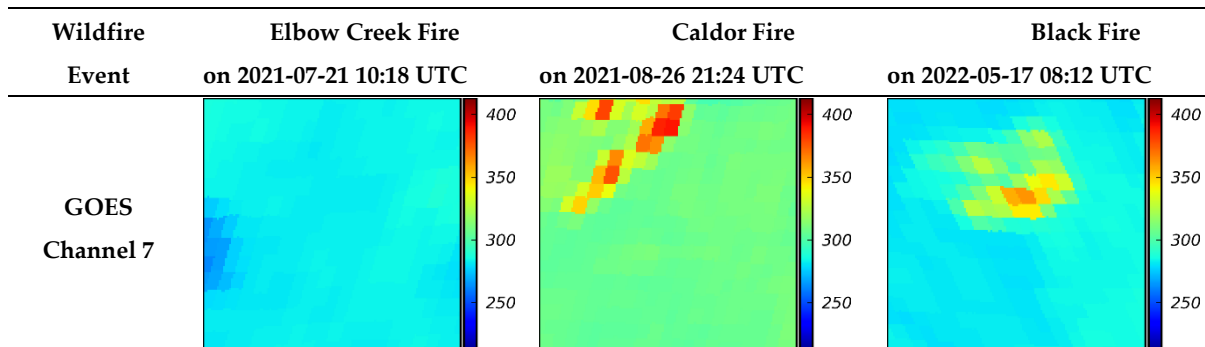


Figure 3. Training curves (i.e., Loss vs. Epoch) of (a) L_{BCE} for segmentation step, (b) $L_{SS-RWRMSE}$ loss for regression step, (c) L_{MRMSE} loss in mean values of ground truth and prediction, (d) L_{SDRMSE} loss in standard deviation value of ground truth and prediction.

4.2. Testing of the proposed DL-WREN

This section presents the results of the proposed two-step approach described in Section 3.1. The proposed approach first employs a U-Net segmentation model trained using BCE loss to localize fire regions. It then employs a subsequent U-Net regression model trained with the $L_{SS-RWRMSE}$ loss function (see Eq. (7)), which combines the RMSE of 2D BT predictions with the RMSE between the predicted and ground truth global mean and standard deviation. The 2D prediction component uses the weighted RMSE loss (Eq. (9)) emphasizes fire region BT value accuracy, using weights of $W_F = 0.75$ for active fire pixels and $W_B = 0.25$ for background pixels. The final BT prediction is obtained by element-wise multiplication of the segmentation and regression outputs, suppressing background noise while enhancing fire-specific details.

Figure 4 presents visual results for three representative wildfire events among the test set, Elbow Creek (2021-07-21), Caldor (2021-08-26), and Black Fire (2022-05-17), are selected to illustrate poor IOU (less than 0.33), average IOU (more than or equal 0.33 less than 0.66), and good IOU (more than or equal to 0.66), respectively. For each case, GOES input channels (7, 14, and 15) are shown alongside the segmentation prediction, regression prediction, the final element-wise product, and the VIIRS ground truth. It should be noted that the grey areas represent background pixels, which are excluded from the BT range and therefore not shown in the color bar. Otsu's thresholding is applied to suppress small spurious values. The lower limit of the color scale is set to 260 K, instead of the non-saturated VIIRS minimum of 283 K, to enable visualization of under-predicted values that would otherwise be clipped. Evaluation metrics including E_{IOU} for fire region segmentation, fire-pixel RMSE (E_{FRMSE}), and background RMSE (E_{BRMSE}) as introduced in Section 3.3 are also listed in the figure for each case.



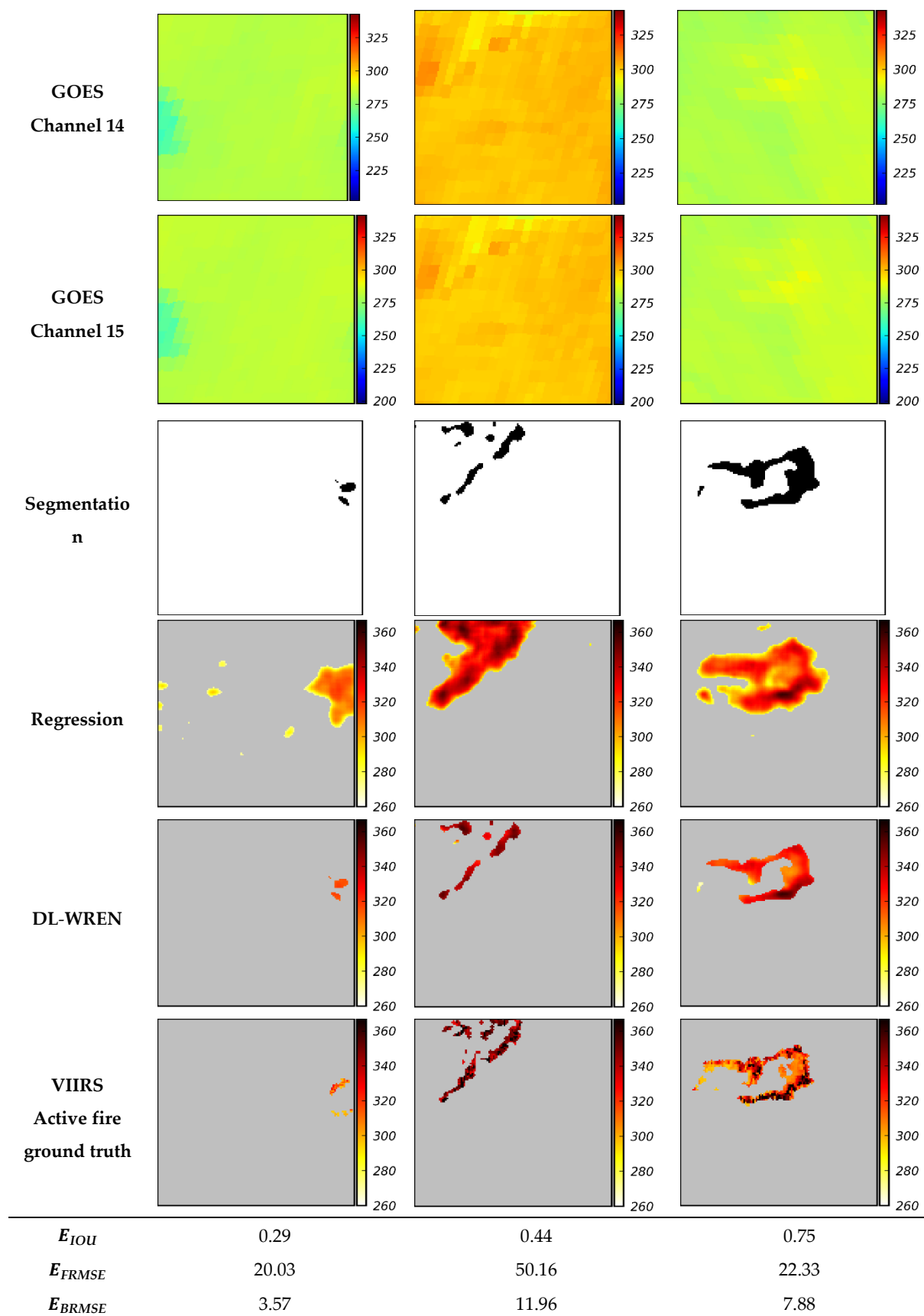


Figure 4. Visual comparisons for three wildfire events with varying IOU levels (low, medium, high). The rows display, from top to bottom, GOES input channels (7, 14, and 15), segmentation output, regression output, DL-WREN (i.e., element-wise product of segmentation output and regression output), and VIIRS active fire ground truth. The three final rows report IOU, fire area RMSE (in Kelvin), and background area RMSE (in Kelvin) scores

for each case. Brightness temperature values are color-coded in Kelvin (K), with color bars provided for temperature scale interpretation.

As shown in Figure 4, the two-step DL-WREN framework improves both spatial localization and BT estimation. The segmentation step isolates fire regions and suppresses false activations, while the regression step provides per-pixel BT predictions reasonably close to the ground truth through region-weighted optimization (Eq. (7)). Together, these components reduce background interference and enhance detection across diverse wildfire scenarios.

Table 1 compares the superior performance of the DL-WREN framework with a previous model developed by the authors (Badhan et al., 2024), which is referred to as the baseline model herein. Both the DL-WREN and baseline model have been applied to the test dataset and the resulting evaluation metrics are calculated and reported in Table 1. As can be seen, the proposed DL-WREN framework substantially outperforms the baseline model. DL-WREN improves spatial localization accuracy from $E_{IOU} = 0.24$ to 0.40, representing a 66.7 % relative increase over the test dataset. In terms of BT accuracy, DL-WREN reduces the fire-region RMSE (E_{FRMSE}) from 187.2 K to 37.6 K, corresponding to a 79.9% reduction, while the background RMSE (E_{BRMSE}) decreases from 31.3 K to 5.9 K, yielding an 81.2% reduction. These gains demonstrate that the two-step design effectively mitigates the trade-off observed in single-stage models by decoupling spatial boundary refinement from BT value estimation. The design choice is justified through ablation studies (Appendix C), which guide the selection of optimal configurations for segmentation and regression. Overall, the two-step approach effectively resolves the trade-off observed in single-stage models by separately leveraging spatial cues for boundary refinement and BT values cues for accurate BT estimation.

Table 1. Average evaluation of proposed approach and baseline.

	E_{IOU}	E_{FRMSE} (K)	E_{BRMSE} (K)
Baseline	0.24	187.2	31.3
DL-WREN	0.40	37.6	5.9

5. Conclusions

This study aimed to improve wildfire monitoring through deep learning (DL) by enhancing both fire location detection and BT estimation from GOES satellite imagery. Building upon our previous work, the main focus was to provide BT predictions within their physical range while improving spatial accuracy. By combining segmentation and regression into a two-step framework, fire regions and fire intensity could be estimated at 5-minute intervals, enabling practical near-real-time wildfire monitoring. Several methodological innovations distinguish this work from earlier efforts. A comprehensive dataset was constructed spanning 2019–2024 across the entire continental United States (CONUS), allowing the models to generalize across diverse geographic and climatic conditions. Preprocessing was improved through terrain-based parallax correction of GOES observations, which substantially reduced geolocation errors relative to VIIRS reference data. Following a study of spectral characteristics, three GOES bands, Band 7 (3.80–4.00 μm), Band 14 (10.8–11.6 μm), and Band 15 (11.8–12.8 μm), were selected as inputs. Separate architectures were identified for segmentation and regression, representing a two-step approach. For segmentation, a U-Net architecture was employed, while regression was carried out with a shallower U-Net variant featuring modifications such as the removal of batch normalization in later layers and additional output branches to predict mean and standard deviation alongside per-image z-score normalized values. These predicted statistics were subsequently used to de-normalize the per-image z-score normalized prediction outputs, enabling accurate brightness temperature estimation in physical units. To improve regression accuracy, a weighted RMSE loss was employed. Fire-region pixels were given higher weights than background pixels to counter the strong pixel imbalance, ensuring that the

model learned to predict brightness temperatures in fire regions without being dominated by background signals. The framework generates fire predictions by multiplying binarized segmentation outputs with BT estimates from the regression model, effectively combining spatial localization with physical fire intensity.

Evaluation demonstrated that the proposed framework delivers predictions that are reasonably close to ground truth. For fire-region detection, the optimized segmentation design achieved an IOU of 0.40, while regression achieved a fire-region RMSE of 38.94 K. Visual comparisons with ground-truth observations further validated the quantitative findings, confirming that the framework reliably captures both fire extent and brightness temperature dynamics. Overall, this study shows that DL, when coupled with geostationary satellite observations and preprocessing steps such as terrain-based parallax correction, spectral band selection, and normalization selection, provides a viable pathway toward near-real-time wildfire monitoring. The ability to deliver fire location and BT estimates at 5-minute intervals across CONUS represents a substantial advancement for fire monitoring. The proposed framework can support firefighters, emergency managers, researchers, and policymakers by enabling refined fire progression information. Future work can incorporate temporal dependencies using past GOES observations, as well as contextual factors such as wind direction and fuel type, with the aim of further improving predictive accuracy and reliability for operational deployment.

Author Contributions: **Mukul Badhan:** Conceptualization, Methodology, Software, Validation, Formal Analysis, Visualization, Writing - Original Draft. **Majid Bavandpour:** Methodology, Formal Analysis, Writing - Review & Editing. **Kasra Shamsaei:** Methodology, Formal Analysis, Writing - Review & Editing. **Dani Or:** Methodology, Formal Analysis, Supervision, Writing - Review & Editing. **George Bebis:** Methodology, Formal Analysis, Supervision, Writing - Review & Editing. **Neil P. Lareau:** Conceptualization, Methodology, Writing - Review & Editing. **Qunying Huang:** Writing - Review & Editing. **Hamed Ebrahimian:** Conceptualization, Methodology, Formal Analysis, Supervision, Project Administration, Funding Acquisition, Writing - Review & Editing.

Funding: This work has been supported through the National Science Foundation's Leading Engineering for America's Prosperity, Health, and Infrastructure (LEAP-HI) program by grant number CMMI-1953333, as well as U.S. Army Engineer Research and Development Center (ERDC) contracts W912HZ24F0414 and W912HZ25CA008. The opinions and perspectives expressed in this study are those of the authors and do not necessarily reflect the views of the sponsors.

Appendix

GOES-VIIRS Alignment Analysis

This appendix provides additional visual results supporting the data preprocessing steps. It includes examples of GOES-VIIRS spatial misalignment before and after orthorectification, and an analysis of day-night variability and cross-sensor correlation in BT values over the active fire regions.

Figure A-1 illustrates examples of the misalignment of GOES-17's band 7 and VIIRS active fire product, its corrected misalignment using orthorectification, the misalignment of GOES-16's band 7 and VIIRS, and its corrected misalignment using orthorectification. GOES-17 exhibits a systematic eastward shift, while GOES-16 shows a westward shift relative to the VIIRS ground truth. After applying orthorectification, these misalignments are significantly reduced.

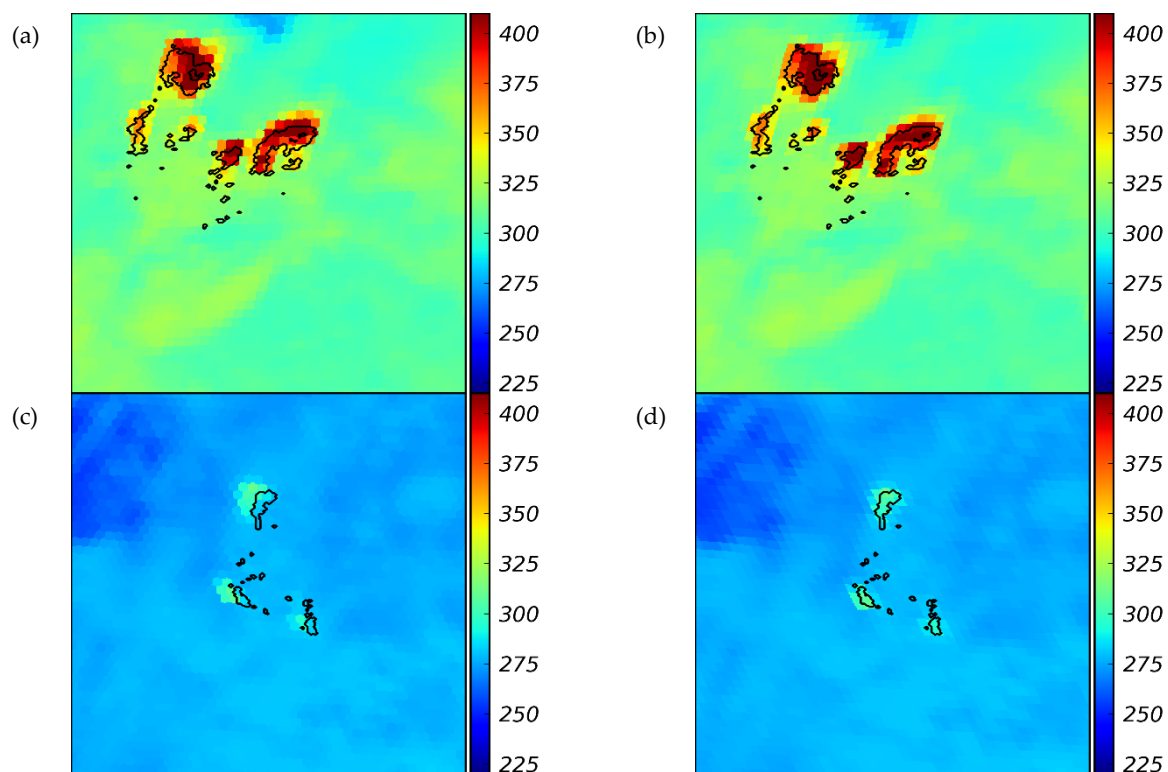


Figure A1. Effect of orthorectification on GOES and VIIRS alignment. For Dixie Fire on August 05, 2021, 21:12 UTC (longitude: -121, latitude: 40): (a) Original GOES-West imagery overlaid with VIIRS boundaries, showing misalignment, (b) Orthorectified GOES-West imagery, demonstrating improved alignment with VIIRS. For Hermits Peak Fire May 04, 2022, 08:54 UTC (longitude: -105, latitude: 35): (c) Original GOES-East imagery overlaid with VIIRS boundaries, showing misalignment, (d) Orthorectified GOES-East imagery, demonstrating improved alignment. The results highlight how orthorectification reduces spatial displacement, improving consistency between GOES and VIIRS fire observations. In these figures, the BT observations are from GOES, while the fire boundaries extracted from VIIRS are shown with black lines.

Figure A-2 provides quantitative evidence supporting the normalization strategy adopted in this study. Panels (a) and (b) show that both GOES Band-7 and VIIRS active-fire BT values exhibit a pronounced day-night bimodal distribution, with systematically higher mean BT during daytime and lower values at night when computed over the same wildfire snapshot. Panel (c) further demonstrates a strong linear relationship between contemporaneous GOES and VIIRS mean BT values over VIIRS-detected fire pixels, indicating that although the absolute BT values differ between sensors and across time of day, their relative BT values remain consistent. Together, these results indicate that absolute BT magnitudes are strongly influenced by diurnal variability, while relative contrasts within each image are more stable and informative for fire characterization. These observations motivated the use of per-image z-score normalization, which removes global BT shifts while preserving internal thermal structure. The analysis in Figure A-2 thus provides empirical justification for the normalization approach described in the main text and underlines the model design in which per-image mean and standard deviation are estimated directly from GOES inputs to enable physically consistent de-normalization at inference time.

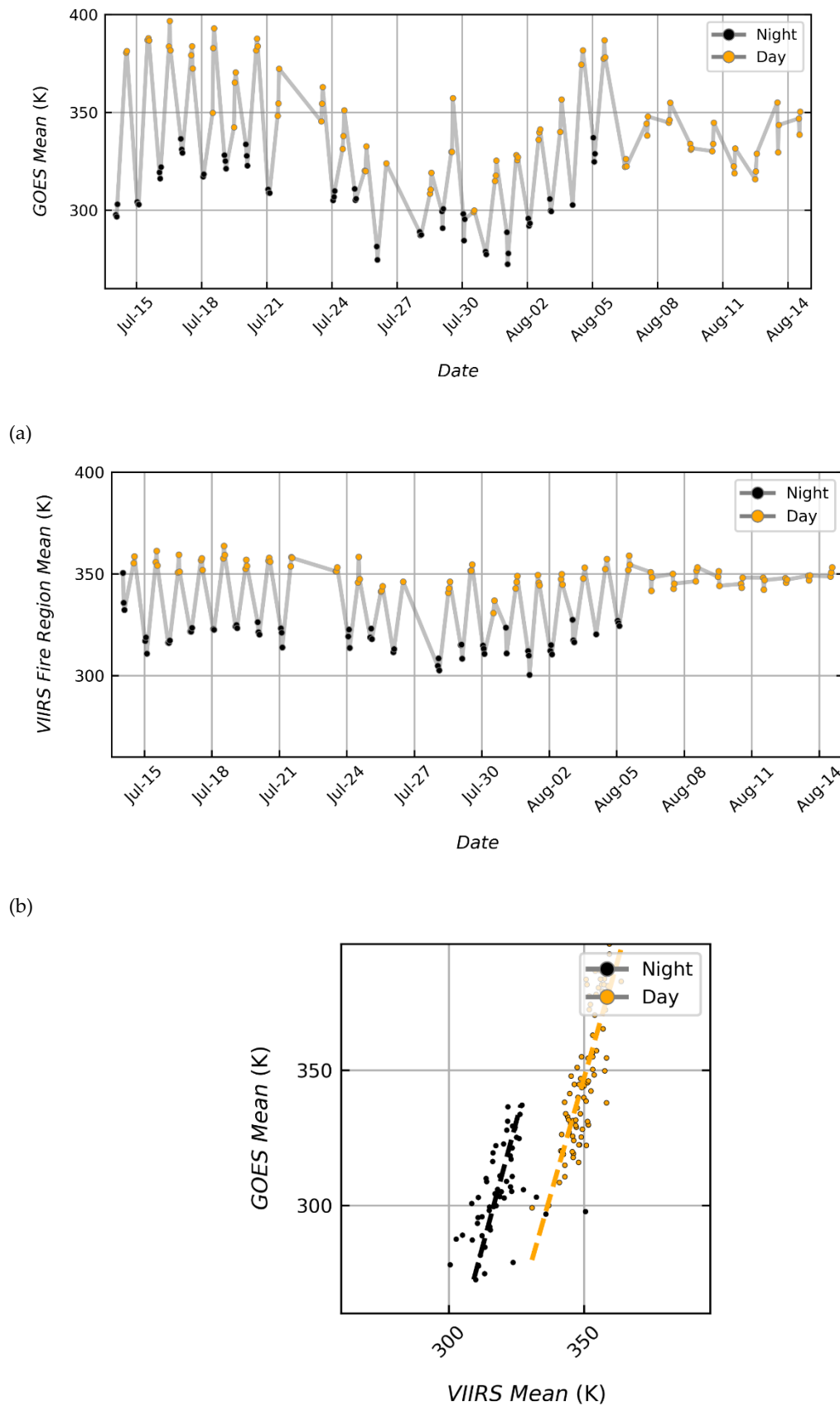


Figure A2. Day–night variability and cross-sensor correlation in BT expressed in K for the Dixie Fire from July 14 to August 15: (a) Mean GOES Band 7 BT values computed only over regions corresponding to VIIRS-detected active fire pixels, showing a distinct day–night bimodal pattern, (b) Mean VIIRS active fire pixels BT values over time, similarly exhibiting clear day–night separation, and (c) Scatter plot of contemporaneous GOES and VIIRS

mean BT values over VIIRS active-fire region, demonstrating a strong linear relationship. Black dots represent nighttime observations while yellow dots represent daytime observations.

Segmentation and Regression Model Architecture used in DL-WREN

This appendix provides a detailed breakdown of the DL architecture used in DL-WREN for fire segmentation and regression. Table B-1 describes the U-Net variant used to generate pixel-wise fire masks, and Table B-2 describes the U-Net based regression model that predicts z-score-normalized BT along with per-image statistical parameters (μ and σ) needed to reconstruct full-scale temperature values.

Table B1. Architecture of the U-Net variant used for segmentation. The model receives a 3-channel GOES input image and outputs a 1-channel binary mask representing the probability of fire presence for each pixel. “Conv2d (3→64)” denotes a 2D convolutional layer with 3 input and 64 output channels, while “×2” indicates two consecutive convolutional layers within a DoubleConv block. “TransposeConv” refers to a transposed convolution (also called deconvolution) used for up-sampling in the decoder path. Batch normalization (BatchNorm2d) is applied after each convolution unless otherwise specified. The final layer uses a sigmoid activation to produce pixel-wise probabilities between 0 and 1, later thresholded to generate binary fire masks.

Stage	Layer Type	Output Shape	Kernel Size	Stride	Padding	Activation	Normalization	Details
Input	–	$3 \times 128 \times 128$	–	–	–	–	–	3 channel GOES input
Encoder 1	Conv2d (3→64) ×2	$64 \times 128 \times 128$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	$64 \times 64 \times 64$	2×2	2	0	–	–	Down-sampling
Encoder 2	Conv2d (64→128) ×2	$128 \times 64 \times 64$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	$128 \times 32 \times 32$	2×2	2	0	–	–	Down-sampling
Encoder 3	Conv2d (128→256) ×2	$256 \times 32 \times 32$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	$256 \times 16 \times 16$	2×2	2	0	–	–	Down-sampling
Encoder 4	Conv2d (256→512) ×2	$512 \times 16 \times 16$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	$512 \times 8 \times 8$	2×2	2	0	–	–	Down-sampling
Bottleneck	Conv2d (512→1024) ×2	$1024 \times 8 \times 8$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Expands feature depth
Decoder 1	TransposeConv (1024→512)	$512 \times 16 \times 16$	2×2	2	0	–	–	Up-sampling
	Conv2d (1024→512) ×2	$512 \times 16 \times 16$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Skip from Encoder 4
Decoder 2	TransposeConv (512→256)	$256 \times 32 \times 32$	2×2	2	0	–	–	Up-sampling
	Conv2d (512→256) ×2	$256 \times 32 \times 32$	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Skip from Encoder 3
Decoder 3	TransposeConv (256→128)	$128 \times 64 \times 64$	2×2	2	0	–	–	Up-sampling

	Conv2d (256→128) ×2	128 × 64 × 64	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Skip from Encoder 2
Decoder 4	TransposeConv (128→64)	64 × 128 × 128	2×2	2	0	–	–	Up- sampling
	Conv2d (128→64) ×2	64 × 128 × 128	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Skip from Encoder 1
Output	Conv2d (64→1)	1 × 128 × 128	1×1	1	0	Sigmoid	–	Final prediction (binary mask)

Table B2. Architecture of the U-Net variant used for regression. The model takes a 3-channel GOES input and predicts a z-score-normalized brightness temperature (BT) map, along with the mean (μ) and standard deviation (σ) of BT for each input image. These parameters allow reconstruction of the full-scale temperature in Kelvin (K) during evaluation. “Conv2d (3→64)” indicates a 2D convolutional layer with 3 input and 64 output channels, while “×2” denotes two successive convolutional operations within the same block (commonly referred to as a DoubleConv block). “FC” refers to a fully connected (linear) layer. Two parallel fully connected branches are used: one predicts the per-image mean (μ) and the other predicts the standard deviation (σ), with an exponential activation (“Exp.”) ensuring σ remains positive. Batch normalization (BatchNorm2d) follows each convolution unless otherwise noted.

Stage	Layer Type	Output Shape	Kernel Size	Stride	Padding	Activation	Batch Normalization	Details
Input	–	3×128×128	–	–	–	–	–	3 channel GOES input
Encoder 1	Conv2d (3→64) ×2	64×128×128	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	64×64×64	2×2	2	0	–	–	Downsample
Encoder 2	Conv2d (64→128) ×2	128×64×64	3×3	1	1	ReLU ×2	BatchNorm2d ×2	DoubleConv block
	MaxPool2d	128×32×32	2×2	2	0	–	–	Downsample
Bottleneck	Conv2d (128→256) ×2	256×32×32	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Expands feature depth
Decoder 1	ConvTranspose2d (256→128)	128×64×64	2×2	2	0	–	–	Upsample
	ConvTranspose2d (256→128)	128×64×64	3×3	1	1	ReLU ×2	BatchNorm2d ×2	Skip from Encoder 2
Decoder 2	ConvTranspose2d (128→64)	64×128×128	2×2	2	0	–	–	Upsample
	Conv2d (128→64) ×2	64×128×128	3×3	1	1	ReLU ×2	BatchNorm2d ×1	Skip from Encoder 1
Final Output	Conv2d (64→1)	1×128×128	1×1	1	0	–	–	Final z-score normalized BT prediction
Global Pool	AdaptiveAvgPool2d	64×1×1	–	–	–	–	–	Spatial pooling

								over full feature map
FC μ	Linear (64→1)	1	-	-	-	-	-	Predicts per-image mean
FC σ	Linear (64→1)	1	-	-	-	Exp.	-	Predicts per-image std, enforced positive

Ablation Studies

Appendix C presents the ablation studies conducted to evaluate the impact of key design choices in the proposed approach, focusing on both the regression and segmentation tasks of the model. These studies quantify how specific modeling decisions, such as loss function design and background BT value in ground truth, affect the accuracy of fire detection and BT estimation.

For the segmentation task, the study addresses the severe class imbalance caused by the small number of active fire pixels relative to the background. To address this, six loss functions are considered in the ablation study, each chosen for its ability to handle data sparsity, class imbalance, or limited spatial overlap. In addition to the BCE loss explained before in Eq. (6), other considered loss functions include the Focal Loss, Jaccard Loss, Tversky Loss, Focal Tversky Loss, and Combo Loss (Jadon, 2020). The goal of this evaluation is to determine which loss formulation provides more accurate localization of fire pixels. Since each loss has a different formulation, the relevant hyperparameters (α , β and γ) are tuned separately for each method to ensure fair comparison. It should be noted that although the same symbols (α , β and γ) are used throughout for clarity and consistency with prior literature, their specific meaning varies across loss functions. The rationale and mathematical formulation of each loss function are discussed in detail below.

Focal Loss (Lin et al., n.d.) integrates two key mechanisms into the BCE loss to better address the challenges of fire segmentation: a class weighting mechanism and a focus on hard examples mechanism. These mechanisms are controlled by the hyperparameters α and γ , respectively. The class weighting mechanism, governed by α (ranging from 0 to 1), adjusts the relative importance of fire (foreground) versus background pixels. A higher α increases the penalty for misclassifying fire pixels, which are scarce relative to the vast number of background pixels, thereby directly addressing class imbalance. The focus on hard examples mechanism, controlled by γ (typically between 1 and 3), reduces the contribution of well-classified pixels (those where \hat{p}_i is close to the true label) by down-weighting their loss using a modulating factor $(1 - \hat{p}_i)^\gamma$ or $(\hat{p}_i)^\gamma$. This shifts the training focus toward hard-to-classify pixels, such as low-intensity or small fires, which are often misclassified due to their similarity to background. Compared to standard BCE, which treats all pixels equally, Focal Loss emphasizes learning from rare and hard-to-classify examples, making it particularly effective for tasks with high class imbalance and ambiguous boundaries. The Focal Loss is defined as:

$$L_{\text{Focal}} = -\frac{1}{N} \sum_{i=1}^N [\alpha y_{b,i} (1 - \hat{p}_i)^\gamma \log \hat{p}_i + (1 - \alpha) (1 - y_{b,i}) (\hat{p}_i)^\gamma \log (1 - \hat{p}_i)] \quad (\text{C-1})$$

The Jaccard Loss, as defined in the following equation, is a region-based loss function that encourages better overlap between predicted and ground truth binary masks. By directly optimizing the IOU score, it is particularly effective for segmentation tasks that require precise spatial alignment.

$$L_{\text{Jaccard}} = 1 - \frac{\sum_{i=1}^N y_{b,i} \hat{p}_i}{\sum_{i=1}^N [y_{b,i} + \hat{p}_i - y_{b,i} \hat{p}_i]} \quad (\text{C-2})$$

In the above equation, the numerator $\sum_{i=1}^N y_{b,i} \cdot \hat{p}_i$ represents the intersection between the predicted fire region and the ground truth, i.e., the number of pixels correctly identified as fire. The denominator includes the union of the predicted and ground truth regions. Specifically, it sums all

pixels identified as fire by either the ground truth or the model (or both), then subtracts the overlapping pixels to avoid double-counting. Physically, this means the model is rewarded when it accurately predicts fire pixels and penalized when it predicts too many false positives or misses fire pixels.

Tversky Loss (Salehi et al., 2017), as defined in Eq. (C-3), is a region-based loss function and extends the Jaccard loss by introducing adjustable weights to balance the importance of false positives (FP) and false negatives (FN). This flexibility is important in fire segmentation tasks, where missing a fire pixel (FN) may be more critical than incorrectly predicting one (FP). The Tversky Loss can be calculated as follows:

$$L_{Tversky} = 1 - \frac{\sum_{i=1}^N y_{b,i} \hat{p}_i}{\sum_{i=1}^N [y_{b,i} \hat{p}_i + \beta(1 - y_{b,i}) \hat{p}_i + (1 - \beta)y_{b,i}(1 - \hat{p}_i)]} \quad (C-3)$$

where, β (ranging from 0 to 1) is a hyperparameter that controls the trade-off between penalizing false positives and false negatives. The second term in the denominator increases when the model predicts fire where there is no fire (FP), while the third term increases when it fails to predict fire where it exists (FN). By adjusting β , the loss function can be biased toward desired behavior. For example, setting $\beta > 0.5$ emphasizes reducing false positives, promoting conservative predictions, whereas $\beta < 0.5$ favors reducing false negatives, encouraging the model to detect more fire pixels even at the risk of overprediction. Physically, this loss encourages the model to learn according to the relative severity of the two types of errors. In fire detection, where missing an active fire can have serious consequences, Tversky Loss enables the network to prioritize fire predictions over overall accuracy, making it well-suited for highly imbalanced segmentation problems.

Focal Tversky Loss (Abraham and Khan, 2019) builds on Tversky Loss by adding a focusing parameter γ , which acts as a spotlight, intensifying the penalty for incorrect or uncertain predictions while dimming the influence of easy, confidently correct ones. By raising the Tversky Loss to the power of $\gamma > 1$, the loss becomes more sensitive to pixels where the model struggles such as small, faint, or partially detected fires forcing the network to pay more attention to these cases during training. The Focal Tversky Loss is defined as follows.

$$L_{Focal_Tversky} = (L_{Tversky})^\gamma \quad (C-4)$$

Finally, Combo Loss (Taghanaki et al., 2019), is formulated as a weighted sum of modified binary cross-entropy (MBCE) (Jadon, 2020) and Dice loss to combine their complementary strengths (Eq. (C-7)). The MBCE component (Eq. (C-5)) introduces a weighting factor β to address class imbalance by adjusting the relative importance of foreground/fire and background/non-fire terms. This loss provides a smooth gradient behavior, meaning it changes gradually with respect to the model's predictions, which helps the model train steadily and avoid unstable jumps. Dice Loss (Eq. (C-6)), a special case of the Tversky loss with $\beta = 0.5$, enhances the learning signal from overlapping regions by doubling their weight in both numerator and denominator. This property makes Dice Loss particularly effective for segmenting small or sparse regions, such as active-fire pixels. This gives Dice a high sensitivity to spatial overlap, encouraging precise localization. By combining MBCE's smooth gradient behavior with Dice Loss's sensitivity to spatial overlap, Combo Loss balances pixel-wise classification and region-level accuracy. The hyperparameter α controls the contribution of each component.

$$L_{MBCE} = -\frac{1}{N} \sum_{i=1}^N \left[\beta (y_{b,i} - \log(\hat{p}_i)) + (1 - \beta)(1 - y_{b,i}) \log(1 - \hat{p}_i) \right] \quad (C-5)$$

$$L_{Dice} = 1 - \frac{\sum_{i=1}^N 2y_{b,i} \cdot \hat{p}_i}{\sum_{i=1}^N [y_{b,i} + \hat{p}_i]} \quad (C-6)$$

$$L_{Combo} = \alpha L_{MBCE} + (1 - \alpha) L_{Dice} \quad (C-7)$$

For the regression task, two important factors are examined for their influence on BT prediction accuracy as in the regression ablation study. First, the effect of assigning different constant values to

background/non-fire pixels in the VIIRS ground truth is examined to reduce the underestimation of BT in fire regions caused by mismatched background intensities. Several background values are tested during experimentation, with 0 K, 220 K, 240 K, and 260 K reported in the next appendix. Second, the effect of spatially selective weighting of fire regions in the loss function is examined. For this ablation study, model is trained with the $L_{SS-RWRMSE}$ (Eq. (9)), varying the fire-region weight W_F among 0.25, 0.5, and 0.75 to emphasize learning from the sparse but critical fire regions. Quantitative and qualitative results from these evaluations are presented in the following section to guide the final model design.

Ablation Study Results

Appendix D provides the results of the ablation studies. It reports performance values, shows visual comparisons, and explains how each tested option-loss functions, background settings, and spatial weighting affects the model performance. These results are used to justify the final configuration of the model.

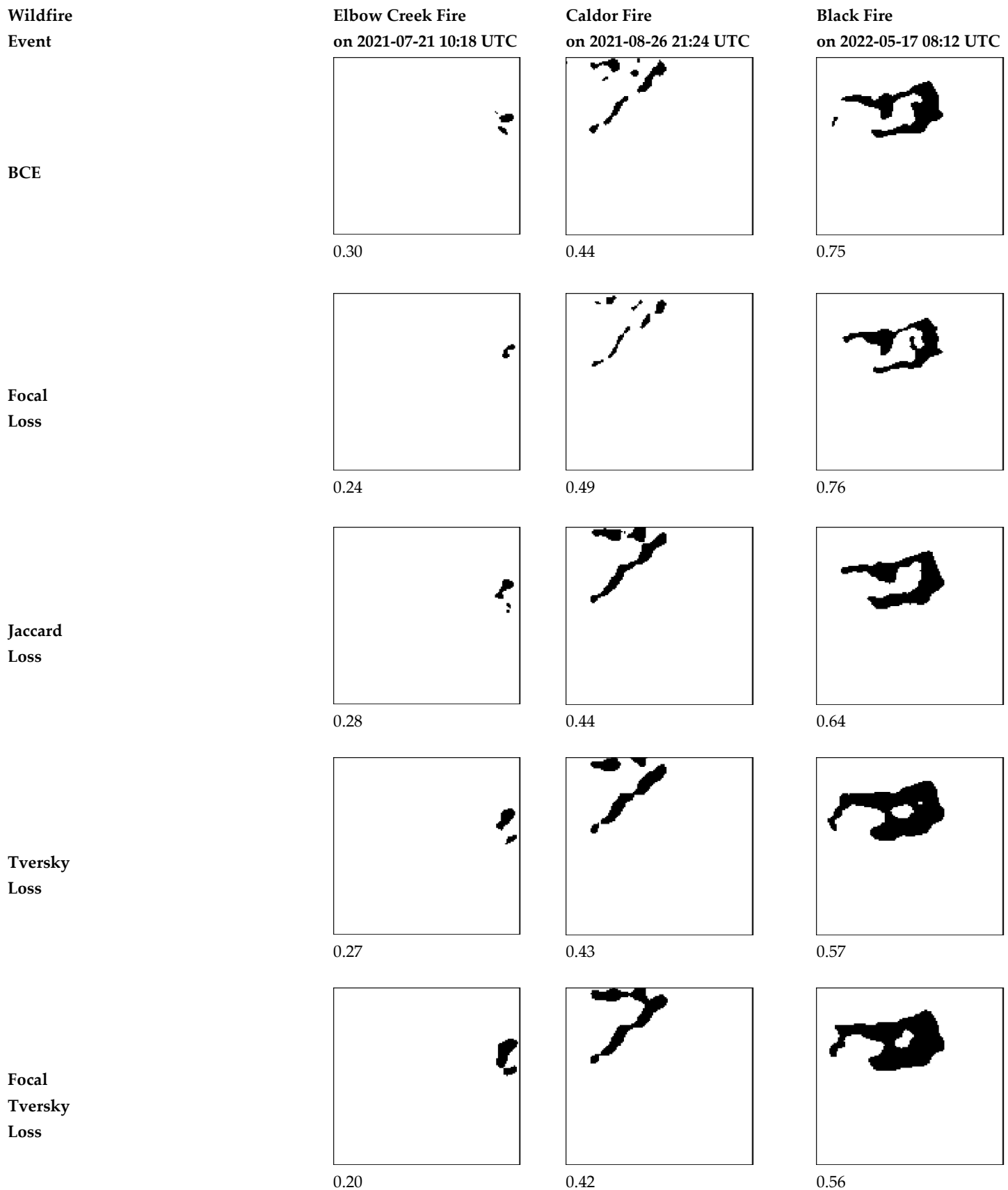
Results for different loss functions evaluated for fire region segmentation are presented in Table D-1, with IOU used as the performance metric. For each loss function, the relevant hyperparameters (α , β and γ) are tuned separately, since their roles differ depending on the formulation (e.g., class weighting in Focal Loss versus imbalance control in Tversky-based losses).

Table D1. Evaluation of loss functions for fire region segmentation.

Loss Function	Alpha	Beta	Gamma	E_{IOU}
BCE	N/A	N/A	N/A	0.40
Focal Loss	0.25	N/A	1	0.40
Jaccard Loss	N/A	N/A	N/A	0.37
Tversky Loss	N/A	0.75	N/A	0.38
Focal Tversky Loss	N/A	0.75	2	0.38
Combo Loss	0.25	0.25	N/A	0.38

Segmentation predictions from each loss function along with binarized VIIRS ground truth are visualized in Figure D-1 for the three representative wildfires illustrated previously, with the corresponding IOU scores shown underneath the predictions. BCE and Focal Loss produce the most spatially consistent and compact fire regions, resembling the VIIRS active fire ground truth. In contrast, Jaccard Loss slightly underestimates the fire extent, while Tversky and Focal Tversky Losses tend to overestimate it by expanding the fire boundaries beyond the true active regions, resulting in lower IOU values. These overpredictions are particularly visible in the Black Fire case, where the fire shape appears overly thickened compared to the VIIRS. Combo Loss exhibits similar behavior, showing fragmented and spatially inconsistent predictions in smaller events like Elbow Creek. Although BCE is a generic pixel-wise loss and does not explicitly account for class imbalance, it performs surprisingly well for this task. This likely stems from the inconsistent relationship between GOES and VIIRS fire region sizes across events. For instance, in the Elbow Creek and Black Fire cases, GOES captures a smaller, more diffused core of the fire compared to the VIIRS reference, whereas in the Caldor Fire, the GOES-derived region appears broader. Loss functions such as Tversky and Focal Tversky, which emphasize imbalance correction, tend to overcompensate for these variations, expanding predictions beyond the true active fire extent and reducing IOU. Combo Loss avoids major over- or underprediction, but it consistently misses finer structural details. Even in larger fires, the thinner extensions of the fire are not fully captured, leading to lower IOU values. Both BCE and Focal Loss achieve comparable IOU scores (0.40), but BCE is ultimately selected for its simplicity and stable performance across diverse fire conditions. Overall, BCE yields the most balanced and physically

consistent segmentation behavior, making it the preferred configuration for the final segmentation model.



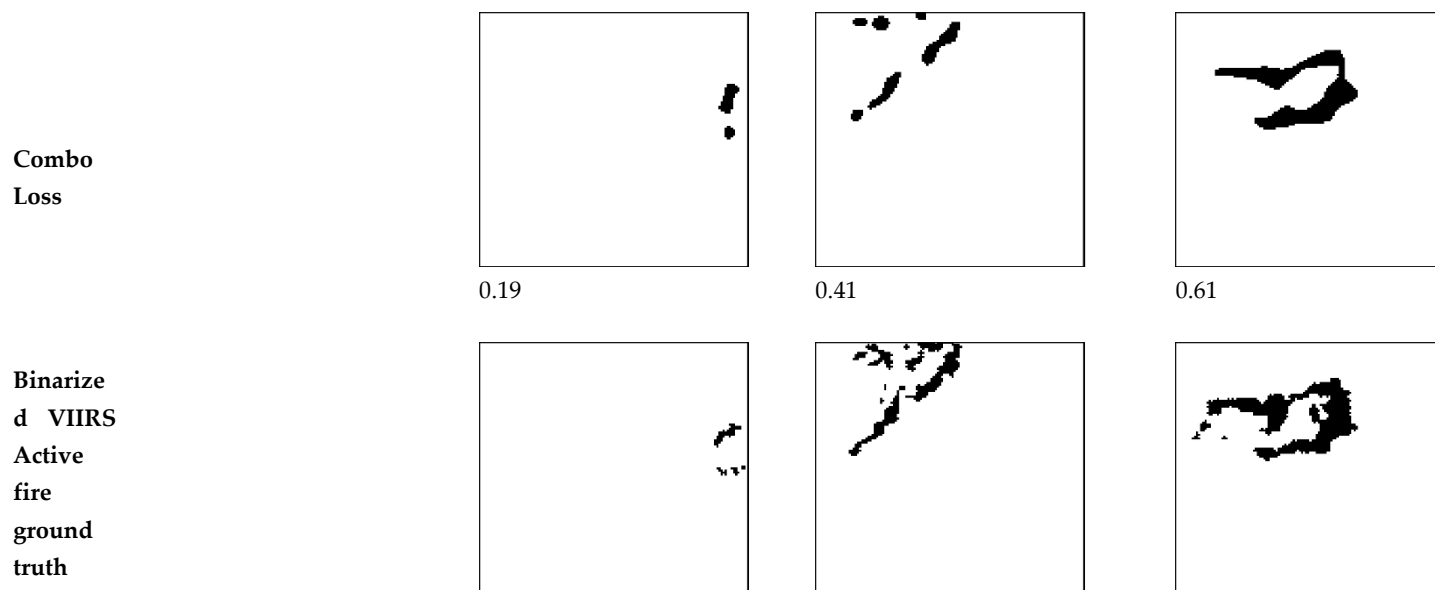


Figure D1. Segmentation predictions using six different loss functions and their binarized VIIRS active fire ground truth for the Elbow Creek, Caldor, and Black Fire events.

The effect of assigning different constant background values (0 K, 220 K, 240 K, and 260 K) to non-fire pixels in the VIIRS ground truth is examined to address the underestimation in the predicted BT. As summarized in Table D-2, setting the background to 0 K results in substantial errors, with a fire region RMSE (E_{FRMSE}) of 153.72 K and a background RMSE (E_{BRMSE}) of 21.6 K. In contrast, using more physically realistic values reduces prediction error: at 240 K, the E_{FRMSE} decreases to 57.51 K and the E_{BRMSE} to 6.08 K, while the IOU remains moderate (0.33). Although increasing the background to 260 K further lowers RMSE ($E_{FRMSE} = 52.14$ K, $E_{BRMSE} = 5.47$ K), this comes at the cost of reduced contrast in small fire regions. To evaluate these background settings fairly, background is removed from the predictions using Otsu's thresholding (Xu et al., 2011). Visual comparisons for the same three representative events are shown in Figure D-2. Predictions are generated from models trained with background values of 0 K, 240 K, and 260 K, and displayed under both full-range and clipped (260–367 K) visualization. With a 0 K background, predictions often fall below the physically plausible BT range, obscuring non-fire structure: clipping removes much of the predicted signal, especially for Elbow Creek. By contrast, a 240 K background yields outputs within the expected range, preserving spatial coherence and capturing both large and small fires. A 260 K background slightly improves RMSE but reduces visual contrast, again most evident in the Elbow Creek Fire. Overall, these results indicate that assigning a 240 K background achieves the best trade-off: it produces physically consistent BT estimates, maintains structure in non-fire regions, and reliably differentiates fires of varying sizes.

Table D2. Evaluation with different background values.

Background Value	E_{IOU}	E_{FRMSE} (K)	E_{BRMSE} (K)
0	0.35	153.72	21.6
220	0.33	64.33	7.07
240	0.33	57.51	6.08
260	0.32	52.14	5.47

Elbow Creek Fire

Caldor Fire

Black Fire

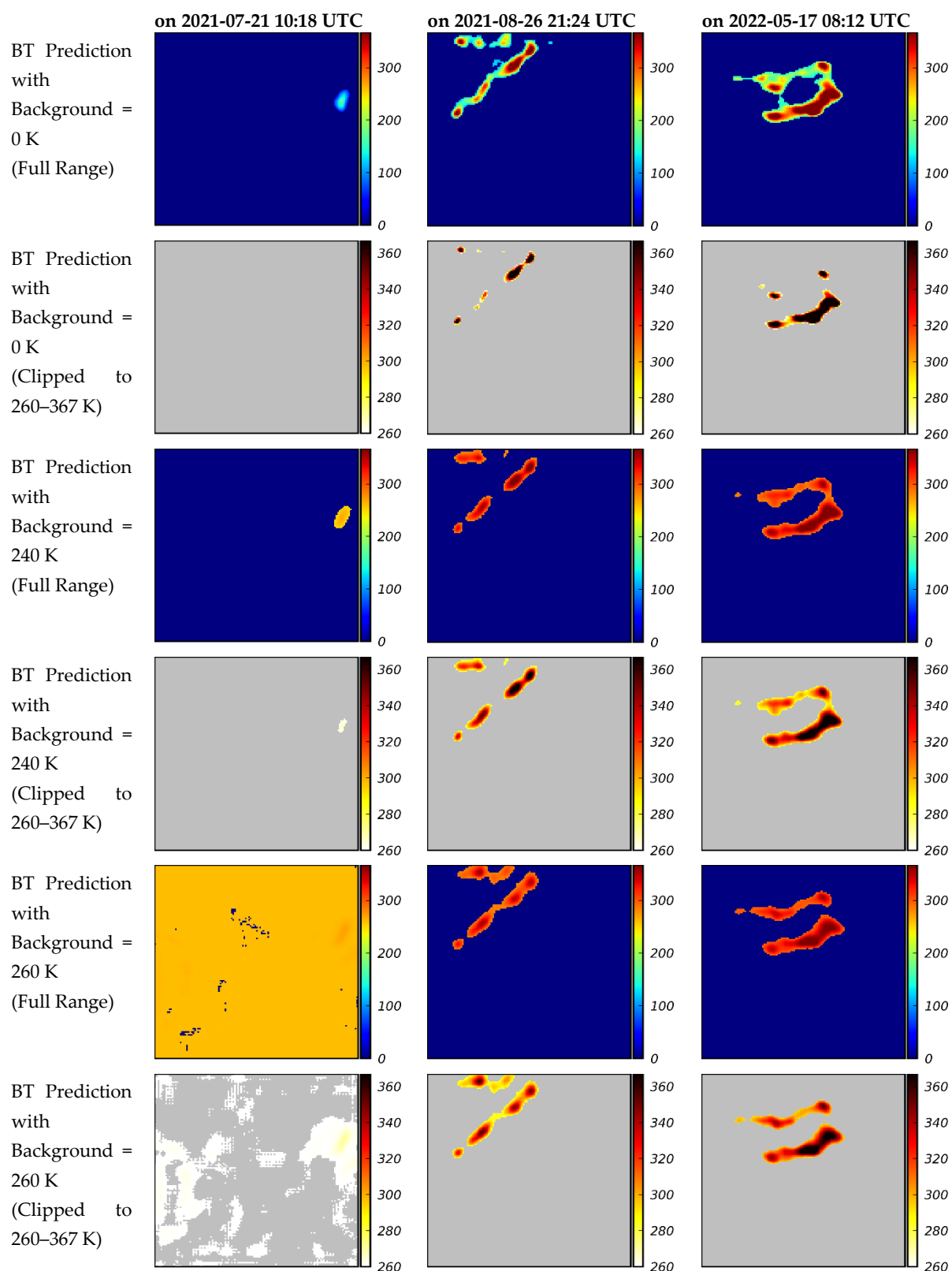


Figure D2. Predicted BT maps from the regression model under different background settings. Columns show Elbow Creek (2021-07-21), Caldor (2021-08-26), and Black Fire (2022-05-17). Predictions are shown for models trained with background values of 0 K, 240 K, and 260 K, visualized under both full prediction range and the physically plausible BT range (260–367 K). Increasing the background value reduces unrealistically low BT predictions in non-fire regions and improves overall RMSE. However, excessively high background values (e.g., 260 K) begin to reduce contrast in smaller fire regions, as seen in the Elbow Creek Fire. A background value of

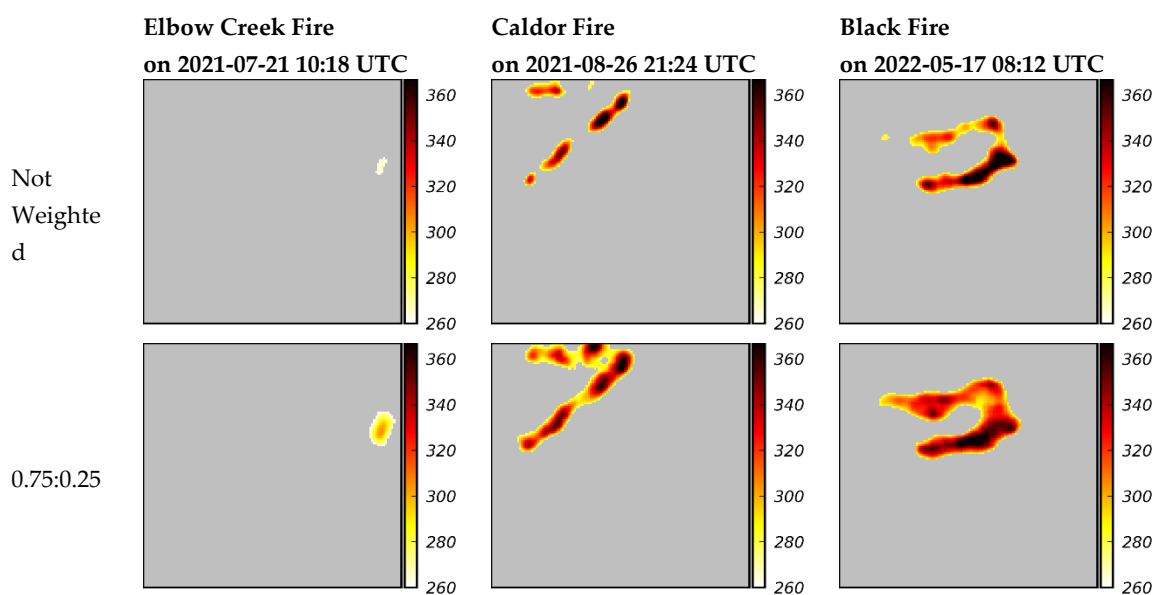
240 K provides a balance between accurate BT estimation and clear differentiation of both large and small fire regions.

To further improve regression accuracy in fire regions, the spatial weighting term introduced in Eq. (9) of the total loss formulation (Eq. (7)) is evaluated with different background-to-fire weighting ratios. Three configurations are evaluated here: (1) background:fire = 0.75:0.25, (2) 0.50:0.50, and (3) 0.25:0.75 where higher weights assigned to fire pixels (W_F) increase their relative contribution to the loss. As shown in Table D-3, a steady decrease in fire region RMSE (E_{FRMSE}) is observed—from 57.23 K in the unweighted case to 41.02 K, 24.53 K, and 18.49 K—demonstrating that better accuracy can be achieved when greater importance is placed on sparse but high-impact fire pixels during training. However, this improvement comes with a trade-off: IOU values steadily decrease, indicating a loss in spatial overlap with the ground truth; and background RMSE (E_{BRMSE}) increases as more emphasis is placed on fire regions. These trends highlight the challenge of balancing precise fire region estimation with overall segmentation quality.

Table D3. Regression performance with z-score normalization and fire region weighting ($L_{SS-RWRMSE}$).

W_B	W_F	Background Value	E_{IOU}	E_{FRMSE} (K)	E_{BRMSE} (K)
Not Weighted	Not Weighted	240	0.33	57.23	6.36
0.75	0.25	240	0.30	41.02	5.77
0.50	0.50	240	0.19	24.53	9.00
0.25	0.75	240	0.11	18.49	13.05

Visual comparisons in the following figure reinforce these findings, with BT maps from three wildfire events showing progressively sharper fire regions and higher predicted BTs as the fire region weight increases. While the most aggressive weighting (0.75) produces the lowest E_{FRMSE} , it also introduces artifacts and reduces IOU.



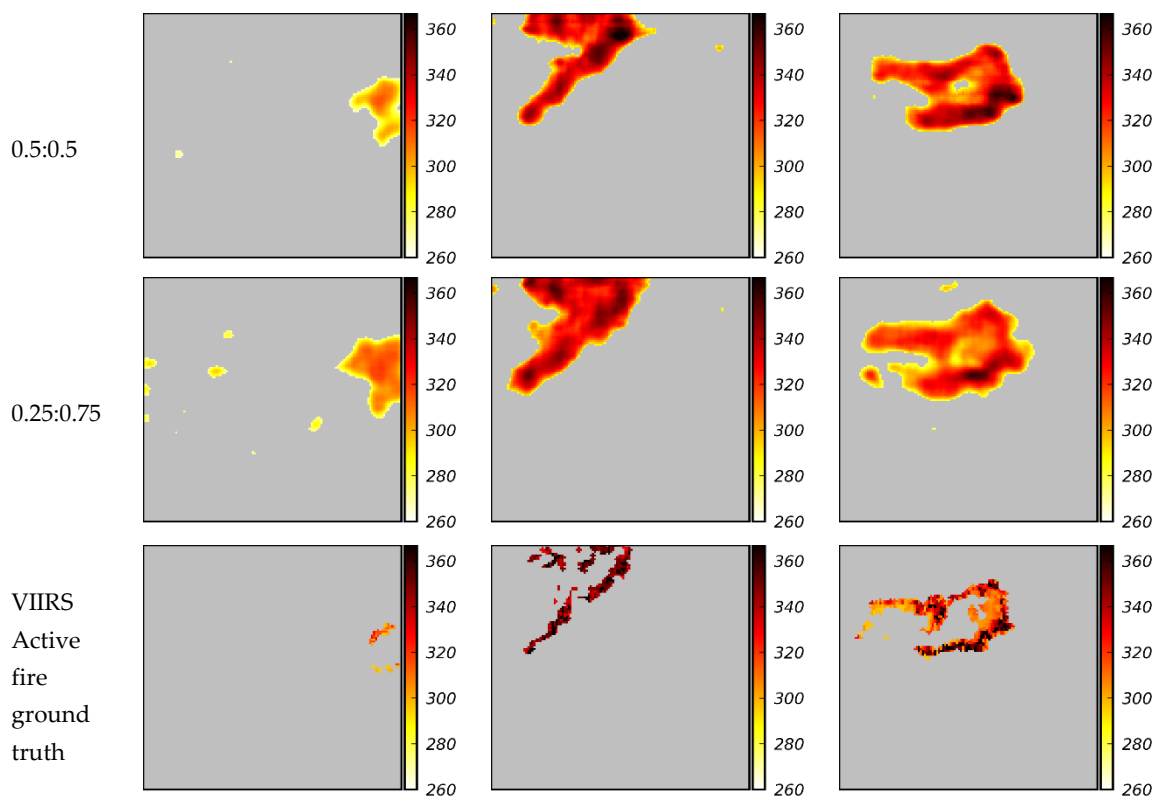


Figure D3. Predicted BT maps from the z-score normalized regression model trained with different fire region weights (background:fire) (Not Weighted, 0.75:0.25, 0.5:0.5, and 0.25:0.75) along with VIIRS ground truth across the three representative wildfire events.

Collectively, these ablation results clarify the complementary strengths and weaknesses of the segmentation and regression modules. The segmentation model trained with BCE loss provides the most spatially consistent fire localization ($E_{IOU} = 0.40$). The regression model optimized with fire-region weighting ($W_F = 0.75$ for active fire and $W_B = 0.25$) produces physically accurate brightness temperatures ($E_{FRMSE} = 18.49$) but tends to overestimate fire extent, lowering spatial overlap ($E_{IOU} = 0.11$). This contrast underscores a key challenge: emphasizing spatial precision often reduces BT accuracy, and vice versa.

The proposed DL-WREN approach addresses this imbalance by multiplying the regression output with the segmentation mask. The segmentation stage suppresses the regression model's boundary overestimation, while the regression stage provides accurate BT values within the detected fire regions. Some fire pixels may be slightly reduced due to imperfect segmentation, but the overall effect is a more stable and physically meaningful reconstruction. When combined, the DL-WREN achieves an E_{IOU} of **0.40**, an E_{FRMSE} of **37.6 K**, and an E_{BRMSE} of **5.9 K** representing the best overall trade-off between spatial accuracy and BT fidelity. Although the combined errors are higher than those of the regression-only model, this is expected because the segmentation mask can remove a small portion of true fire pixels along with background. However, the two-step framework maintains strong BT predictions while enforcing precise localization, producing the most reliable and physically meaningful representation of active fires across diverse wildfire scenarios.

Wildfire Event List

This appendix summarizes the wildfire events used in this study in Table E-1. It provides geographic coordinates and active time ranges for each fire included in the dataset.

Table E1. Wildfire event details.

Site	Latitude	Longitude	Start time	End time
Tucker	41.73	-121.24	7/30/2019	7/30/2019
Walker	40.05	-120.67	9/6/2019	9/16/2019
Taboose	37.03	-118.34	9/8/2019	9/17/2019
Red Bank	40.12	-122.64	9/6/2019	9/11/2019
Saddle Ridge	34.33	-118.48	10/11/2019	10/28/2019
Kincade	38.79	-122.78	10/24/2019	10/30/2019
Maria	34.3	-119	11/1/2019	11/1/2019
Bighorn Fire	32.53	-111.03	6/7/2020	7/1/2020
Mangum Fire	36.61	-112.34	6/11/2020	6/27/2020
Bush Fire	33.63	-111.56	6/14/2020	6/26/2020
North Complex Fire	39.69	-120.12	8/15/2020	11/17/2020
SCU Lightning Complex	37.35	-121.44	8/16/2020	8/31/2020
CZU Lightning Complex	37.1	-122.28	8/16/2020	8/23/2020
White River Fire	44.74	-121.64	8/18/2020	8/29/2020
August Complex	39.87	-122.97	8/18/2020	9/22/2020
Christie Mountain	49.36	-119.54	8/19/2020	8/22/2020
Palmer Fire	48.83	-119.56	8/19/2020	8/22/2020
LNU Lightning Complex	38.59	-122.24	8/18/2020	9/29/2020
Santiam Fire	44.82	-122.19	9/1/2020	9/18/2020
Beachie Creek Fire	44.75	-122.14	9/2/2020	9/13/2020
Creek Fire	37.2	-119.3	9/5/2020	9/9/2020
Cold Springs Fire	48.85	-119.57	9/7/2020	9/9/2020
Holiday Farm Fire	44.15	-122.45	9/8/2020	9/17/2020
Slater Fire	41.77	-123.38	9/8/2020	9/29/2020
Glass Fire	38.56	-122.5	9/27/2020	10/2/2020
Blue Ridge Fire	33.88	-117.68	10/26/2020	10/27/2020
Silverado Fire	33.74	-117.66	10/26/2020	10/26/2020
Bond Fire	33.74	-117.67	12/3/2020	12/3/2020
Joseph Canyon	45.99	-117.08	6/5/2021	7/14/2021
Telegraph Fire	33.21	-111.09	6/4/2021	6/19/2021
Pinnacle Fire	32.87	-110.2	6/11/2021	6/30/2021
Dixie	40	-121	7/14/2021	8/14/2021
Backbone Fire	34.34	-111.68	6/17/2021	7/8/2021
S-503	45.09	-121.48	6/19/2021	8/16/2021
Rafael Fire	34.94	-112.16	6/18/2021	6/29/2021
Lava	41.46	-122.33	6/26/2021	9/2/2021
Tennant	41.66	-122.04	6/28/2021	7/3/2021
Wrentham Market	45.49	-121.01	6/29/2021	6/29/2021
Salt	40.85	-122.34	6/30/2021	7/9/2021
Beckwourth Complex	36.57	-118.81	8/17/2021	9/21/2021
Tamarack	38.63	-119.86	7/17/2021	9/25/2021
Jack	43.32	-122.69	7/6/2021	10/28/2021
Bootleg	42.62	-121.42	7/6/2021	7/30/2021
Lick Creek Fire	46.26	-117.42	7/7/2021	8/13/2021
Grandview	44.47	-121.4	7/12/2021	7/13/2021
Elbow Creek	45.87	-117.62	7/15/2021	9/9/2021
Middle Fork Complex	43.87	-122.41	8/3/2021	10/28/2021
Rough Patch Complex	43.51	-122.68	8/3/2021	10/28/2021
River Complex	41.14	-123.02	8/1/2021	9/26/2021

Monument	40.75	-123.33	7/30/2021	9/26/2021
McCash	41.56	-123.4	8/2/2021	9/26/2021
Antelope	41.52	-121.92	8/3/2021	9/9/2021
Bull Complex	44.88	-122.01	8/12/2021	10/19/2021
McFarland	40.35	-123.03	7/30/2021	9/15/2021
Black Butte	44.09	-118.33	8/4/2021	8/17/2021
Devil's Knob Complex	41.91	-123.27	8/15/2021	9/26/2021
Fox Complex	42.21	-120.6	8/14/2021	8/15/2021
French	35.69	-118.55	8/19/2021	10/5/2021
Cougar Peak	42.28	-120.61	9/8/2021	9/18/2021
Windy	36.05	-118.63	9/12/2021	10/7/2021
KNP Complex	36.57	-118.81	9/12/2021	10/17/2021
Alisal	34.52	-120.13	10/12/2021	10/13/2021
Bertha Swamp Road	30.15	-85.33	3/5/2022	3/8/2022
62 Fire	34.8	-98.91	3/14/2022	3/18/2022
Kidd	32.33	-98.83	3/18/2022	3/18/2022
Big L	32.36	-98	3/20/2022	3/20/2022
Hayfield South	26.81	-97.66	3/24/2022	3/26/2022
Bell 2101	31.26	-97.62	3/27/2022	3/28/2022
Canadian River Bottom	35.75	-100.54	3/29/2022	4/1/2022
Washita River Fire	35.82	-99.89	4/1/2022	4/1/2022
Borrega	27.47	-98.02	3/31/2022	3/31/2022
L 30	25.89	-80.49	3/31/2022	4/2/2022
Beaver River Fire	36.83	-100.79	4/6/2022	4/7/2022
Hermits Peak	35.72	-105.4	4/11/2022	6/16/2022
Cooks Peak	36.24	-105.04	4/18/2022	6/16/2022
Tunnel	35.3	-111.59	4/19/2022	6/17/2022
Cerro Pelado	35.77	-106.58	4/23/2022	5/12/2022
Road 702	40.16	-100.53	4/23/2022	4/23/2022
Little Highline	35.4	-101.79	4/23/2022	4/24/2022
Bear Trap	33.85	-107.53	5/4/2022	5/31/2022
Smoke Stack Lightning	33.19	-100.17	5/3/2022	5/3/2022
Wildcat	33.82	-111.79	7/14/2022	7/15/2022
L39	26.36	-80.4	5/5/2022	5/7/2022
San Rafael	31.35	-110.61	5/8/2022	5/8/2022
Black (Arizona)	33.2	-108.06	5/14/2022	6/17/2022
Coconut	33.86	-99.34	5/18/2022	5/19/2022
Mesquite Heat	32.27	-99.96	5/18/2022	5/19/2022
Contreras	31.85	-111.57	6/14/2022	6/21/2022
Pipeline	35.27	-111.67	6/12/2022	6/17/2022
Mullica River Fire	39.73	-74.72	6/20/2022	6/20/2022
Dempsey	32.8	-98.25	6/23/2022	6/25/2022
Willow Creek	44.16	-117.39	6/29/2022	6/29/2022
Halfway Hill	38.91	-112.34	7/8/2022	7/12/2022
Moose	45.37	-114.09	7/18/2022	10/21/2022
Casino	43.13	-102.87	7/19/2022	7/19/2022
Bray	43.05	-114.92	7/19/2022	7/19/2022
Oak	37.55	-119.92	7/22/2022	7/28/2022
217 Fire	36.51	-99.12	7/26/2022	7/26/2022
McKinney	41.83	-122.9	7/30/2022	9/11/2022
Elmo	47.82	-114.51	7/30/2022	8/18/2022

Carter Canyon	41.76	-103.76	7/31/2022	7/31/2022
Vantage Highway	46.97	-120.18	8/2/2022	8/4/2022
Miller Road	45.13	-121.34	8/2/2022	8/2/2022
Cedar Creek	43.73	-122.17	8/5/2022	10/20/2022
Campbell	40.91	-123.6	8/6/2022	9/2/2022
Four Corners	44.54	-116.17	8/16/2022	10/10/2022
Ross Fork	43.81	-114.98	8/14/2022	9/12/2022
Rum Creek	42.64	-123.63	8/27/2022	9/5/2022
Patrol Point	45.43	-114.94	9/1/2022	10/21/2022
Parks	48.98	-120.7	8/30/2022	10/20/2022
Sturgill	45.28	-117.53	8/30/2022	9/12/2022
Russell Mountain	48.79	-116.57	8/30/2022	9/14/2022
Trail Ridge	45.77	-113.87	8/27/2022	9/28/2022
Nebo	45.13	-117.11	8/30/2022	9/29/2022
Williams Creek	45.69	-115.65	8/31/2022	10/19/2022
Double Creek	45.43	-116.74	8/30/2022	10/10/2022
Mountain	41.43	-122.64	9/2/2022	9/11/2022
Fairview	33.72	-116.89	9/6/2022	9/8/2022
Mosquito	39.01	-120.74	9/7/2022	9/17/2022
Eden 2	42.65	-114.21	9/8/2022	9/9/2022
Bolt Creek	47.73	-121.35	9/10/2022	10/20/2022
Bovee	41.86	-100.31	10/2/2022	10/2/2022
Hwy 123	36.63	-96.13	4/3/2023	4/8/2023
Zink	36.31	-96.2	4/6/2023	4/8/2023
Great Lakes	34.86	-77.05	4/21/2023	6/7/2023
Sandy	25.93	-81.03	5/3/2023	5/12/2023
Pass	33.41	-108.29	5/25/2023	7/31/2023
Wilbur	34.55	-111.47	6/7/2023	6/20/2023
Ridge	35.91	-111.99	6/7/2023	7/28/2023
Hat Rock	45.91	-119.15	6/13/2023	6/13/2023
Pulp Road	34.11	-78.2	6/16/2023	6/17/2023
Pilot	34.87	-113.29	7/5/2023	7/8/2023
Beehive	31.5	-111.16	7/2/2023	7/2/2023
Divide	33.54	-108.39	7/19/2023	8/7/2023
Flat	42.52	-124.04	7/16/2023	9/23/2023
Hayden	44.7	-113.74	7/20/2023	8/1/2023
Newell Road	45.83	-120.46	7/22/2023	7/25/2023
Pasture	33.57	-108.6	7/24/2023	8/7/2023
Elkhorn	45.5	-115.32	7/30/2023	8/4/2023
York	35.29	-115.29	7/28/2023	7/30/2023
Prior	33.29	-108.39	7/28/2023	8/7/2023
Eagle Bluff	48.88	-119.48	7/29/2023	7/31/2023
Niarada	47.83	-114.59	7/31/2023	8/20/2023
Middle Ridge	47.52	-114.41	7/31/2023	8/4/2023
Lookout	44.22	-122.15	8/8/2023	9/18/2023
Kelly	41.86	-123.84	8/15/2023	9/23/2023
Elliot	41.6	-123.52	8/15/2023	9/23/2023
Mosquito	41.37	-123.67	8/16/2023	9/23/2023
Pearch	41.3	-123.49	8/17/2023	9/23/2023
Oregon	48.03	-117.23	8/18/2023	8/20/2023
Gray	47.54	-117.74	8/18/2023	8/20/2023

River Road East	47.38	-114.81	8/18/2023	8/20/2023
Tiger Island	30.67	-93.44	8/23/2023	8/27/2023
Anvil	42.75	-124.34	9/14/2023	9/23/2023
Still	34.58	-111.25	9/19/2023	10/19/2023
Matts Creek	37.59	-79.43	11/16/2023	11/20/2023
Bettys Way	41.34	-100.45	2/26/2024	3/19/2024
Smokehouse Creek	35.85	-101.43	2/27/2024	2/28/2024
Windy Deuce	35.63	-101.85	2/27/2024	2/28/2024
Grape Vine Creek	35.41	-100.83	2/27/2024	2/28/2024
Catesby Fire	36.44	-99.96	2/28/2024	2/28/2024
Indios	36.29	-106.7	5/23/2024	5/29/2024
Pioneer	48.18	-120.53	6/11/2024	8/4/2024
Antone	34.03	-108.36	6/11/2024	6/13/2024
Post	34.8	-118.88	6/16/2024	6/17/2024
Sites	39.31	-122.34	6/18/2024	6/19/2024
South Fork	33.34	-105.77	6/17/2024	6/19/2024
Little Valley	43.86	-117.5	6/26/2024	6/27/2024
Shelly	41.47	-123.06	7/3/2024	7/9/2024
Silver King	38.48	-112.37	7/5/2024	7/9/2024
Lake	34.79	-120.05	7/6/2024	7/9/2024
Wilder	41.9	-118.55	7/7/2024	7/7/2024
Deer Springs	37.3	-112.31	7/7/2024	7/9/2024
Horse Gulch	46.69	-111.7	7/17/2024	7/17/2024
Falls	43.85	-119.43	7/16/2024	8/8/2024
Cow Valley	44.36	-117.76	7/18/2024	8/7/2024
McGhee	45.27	-106.53	7/16/2024	7/18/2024
Deadman	45.22	-106.62	7/16/2024	7/17/2024
Trout	36	-118.32	7/16/2024	8/1/2024
Black (New Mexico)	33.61	-111.14	8/5/2024	8/5/2024
Lone Rock	45.17	-119.96	7/16/2024	7/23/2024
Cougar Creek	46.04	-117.32	7/17/2024	8/3/2024
Boneyard	44.96	-119.48	7/18/2024	8/7/2024
Durkee	44.55	-117.48	7/18/2024	8/7/2024
Lane 1	43.65	-122.79	7/19/2024	8/8/2024
Swawilla I	47.95	-118.78	7/20/2024	8/3/2024
Jack Canyon Mutual Aid	45.45	-118.93	7/19/2024	7/19/2024
Snake	45.06	-119.01	7/19/2024	8/2/2024
Monkey Creek	45.01	-119.19	7/18/2024	7/23/2024
Adam Mountain	43.62	-122.73	7/19/2024	8/7/2024
Courtrock	44.73	-119.38	7/21/2024	8/8/2024
Big Horn	45.9	-120.24	7/23/2024	7/23/2024
Crazy Creek	44.36	-120.03	7/22/2024	8/7/2024
Telephone	43.76	-118.86	7/22/2024	8/7/2024
Retreat	46.67	-120.99	7/23/2024	8/8/2024
Coyote	44.73	-117.09	7/23/2024	8/7/2024
Thompson	44.62	-117.4	7/23/2024	8/7/2024
Stockade Canyon	40.78	-119.74	7/29/2024	7/29/2024
Park	39.82	-121.8	7/29/2024	8/8/2024
Borel	35.52	-118.68	7/29/2024	8/1/2024
Limepoint	45.08	-116.75	7/29/2024	8/8/2024
Sand Stone	33.76	-111.59	8/5/2024	8/5/2024

Lower Granite	46.57	-117.55	7/29/2024	8/2/2024
Pleasant Valley	42.31	-104.73	7/31/2024	8/2/2024
Wildcat Creek	43.54	-105.11	8/2/2024	8/4/2024
Town Gulch	44.88	-117.25	8/7/2024	8/8/2024
Paddock	44.16	-116.51	8/6/2024	8/8/2024
Warner Peak	42.43	-119.81	8/7/2024	8/8/2024

References

1. Abraham, N., Khan, N., 2019. ISBI 2019 : 2019 IEEE International Symposium on Biomedical Imaging : April 8-11, 2019, Hilton Molino Stucky, Venice, Italy. IEEE.
2. Agarap, A.F., 2019. Deep Learning using Rectified Linear Units (ReLU).
3. Badhan, M., Shamsaei, K., Ebrahimian, H., Bebis, G., Lareau, N.P., Rowell, E., 2024. Deep Learning Approach to Improve Spatial Resolution of GOES-17 Wildfire Boundaries Using VIIRS Satellite Data. *Remote Sens. (Basel)*. 16. <https://doi.org/10.3390/rs16040715>
4. Barducci, A., Guzzi, D., Marcoionni, P., Pippi, I., 2004. Comparison of fire temperature retrieved from SWIR and TIR hyperspectral data. *Infrared Phys. Technol.* 46, 1–9. <https://doi.org/10.1016/j.infrared.2004.03.001>
5. Earth from Orbit: NOAA's GOES-18 is now GOES West [WWW Document], n.d. . <https://www.nesdis.noaa.gov/news/earth-orbit-noaas-goes-18-now-goes-west>.
6. Hager, L., Lemieux, P., n.d. Data Stewardship Maturity Report for NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 1b Radiances. <https://doi.org/10.25923/wm7w-cn62>
7. Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.
8. Jadon, S., 2020. A survey of loss functions for semantic segmentation, in: 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2020. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/CIBCB48159.2020.9277638>
9. Jones Ben, n.d. Wildfires in the US: Alarming Statistics and Trends [WWW Document]. <https://www.dryad.net/post/wildfires-usa-statistics>.
10. Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization.
11. Koltunov, A., Ustin, S.L., Prins, E.M., 2012. On timeliness and accuracy of wildfire detection by the GOES WF-ABBA algorithm over California during the 2006 fire season. *Remote Sens. Environ.* 127, 194–209. <https://doi.org/10.1016/j.rse.2012.09.001>
12. Koltunov, A., Ustin, S.L., Quayle, B., Schwind, B., n.d. GOES early fire detection (GOES-EFD) system prototype.
13. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., n.d. Focal Loss for Dense Object Detection.
14. Malik, A., Jalin, N., Rani, S., Singhal, P., Jain, S., Gao, J., 2021. Wildfire Risk Prediction and Detection using Machine Learning in San Diego, California, in: Proceedings - 2021 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People, and Smart City Innovations, SmartWorld/ScalCom/UIC/ATC/IoP/SCI 2021. Institute of Electrical and Electronics Engineers Inc., pp. 622–629. <https://doi.org/10.1109/SWC50871.2021.00092>
15. Mao, A., Mohri, M., Zhong, Y., 2023. Cross-Entropy Loss Functions: Theoretical Analysis and Applications.
16. Moritz, H., 2000. Geodetic reference system 1980. *J. Geod.* 74, 128–133.
17. National Interagency Coordination Center Wildland Fire Summary and Statistics Annual Report 2023, n.d.
18. National Interagency Fire Center (NIFC), n.d. Wildland Fire Incident Locations [Dataset] [WWW Document]. <https://data-nifc.opendata.arcgis.com/datasets/nifc::wildland-fire-incident-locations/about>.
19. Pestana, S., Lundquist, J.D., 2022. Evaluating GOES-16 ABI surface brightness temperature observation biases over the central Sierra Nevada of California. *Remote Sens. Environ.* 281. <https://doi.org/10.1016/j.rse.2022.113221>
20. Sahito, Faisal, Zhiwen, P., Sahito, Fahad, Ahmed, J., 2023. Transpose convolution based model for super-resolution image reconstruction. *Applied Intelligence* 53, 10574–10584. <https://doi.org/10.1007/s10489-022-03745-4>

21. Saleh, A., Zulkifley, M.A., Harun, H.H., Gaudreault, F., Davison, I., Spraggon, M., 2024. Forest fire surveillance systems: A review of deep learning methods. *Heliyon*. <https://doi.org/10.1016/j.heliyon.2023.e23127>
22. Salehi, S.S.M., Erdogmus, D., Gholipour, A., 2017. Tversky loss function for image segmentation using 3D fully convolutional deep networks.
23. Schmidt, C.C., n.d. NOAA NESDIS CENTER for SATELLITE APPLICATIONS and RESEARCH GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Fire / Hot Spot Characterization.
24. Schmit, T.J., Griffith, P., Gunshor, M.M., Daniels, J.M., Goodman, S.J., Lehair, W.J., 2017. A closer look at the ABI on the goes-r series. *Bull. Am. Meteorol. Soc.* 98, 681–698. <https://doi.org/10.1175/BAMS-D-15-00230.1>
25. Schroeder, W., Giglio, L., Hall, J., 2025. Collection 2 Visible Infrared Imaging Radiometer Suite (VIIRS) 375-m Active Fire Product User's Guide Version 1.2.
26. Schroeder, W., Oliva, P., Giglio, L., Csiszar, I.A., 2014. The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment. *Remote Sens. Environ.* 143, 85–96. <https://doi.org/10.1016/j.rse.2013.12.008>
27. Taghanaki, S.A., Zheng, Y., Kevin Zhou, S., Georgescu, B., Sharma, P., Xu, D., Comaniciu, D., Hamarneh, G., 2019. Combo loss: Handling input and output imbalance in multi-organ segmentation. *Computerized Medical Imaging and Graphics* 75, 24–33. <https://doi.org/10.1016/j.compmedimag.2019.04.005>
28. Ti, 2016. Algorithm Theoretical Basis Document For NOAA NDE VIIRS Active Fire Compiled by the SPSRB Common Standards Working Group.
29. Toan, N.T., 2019. A deep learning approach for early wildfire detection from hyperspectral satellite images, in: 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA). IEEE, p. 48. <https://doi.org/10.1109/RITAPP.2019.8932740>
30. Uchida, K., Tanaka, M., Okutomi, M., 2018. Coupled convolution layer for convolutional neural network. *Neural Networks* 105, 197–205. <https://doi.org/10.1016/j.neunet.2018.05.002>
31. Using the EPSG Geodetic Parameter Dataset Revision history, 2004.
32. Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m Active Fire Detection and Characterization Algorithm Theoretical Basis Document 1.0, 2016.
33. Wang, Z., Chen, J., Hoi, S.C.H., 2021. Deep Learning for Image Super-Resolution: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* <https://doi.org/10.1109/TPAMI.2020.2982166>
34. Weng, W., Zhu, X., 2021. UNet: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access* 9, 16591–16603. <https://doi.org/10.1109/ACCESS.2021.3053408>
35. Xu, X., Xu, S., Jin, L., Song, E., 2011. Characteristic analysis of Otsu threshold and its applications. *Pattern Recognit. Lett.* 32, 956–961. <https://doi.org/10.1016/j.patrec.2011.01.021>
36. Yang, J., Chen, F., Das, R.K., Zhu, Z., Zhang, S., 2024. Adaptive-avg-pooling based attention vision transformer for face anti-spoofing, in: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 3875–3879. <https://doi.org/10.1109/ICASSP48485.2024.10446940>
37. Zhao, L., Zhang, Z., 2024. A improved pooling method for convolutional neural networks. *Sci. Rep.* 14. <https://doi.org/10.1038/s41598-024-51258-6>
38. Zhao, Y., Ban, Y., 2022. GOES-R Time Series for Early Detection of Wildfires with Deep GRU-Network. *Remote Sens. (Basel)*. 14. <https://doi.org/10.3390/rs14174347>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.