INTRODUCTION

A wildfire, also known as a forest fire, bushfire, or vegetation fire, is an uncontrolled and rapidly spreading fire that burns through vegetation, forests, grasslands, or any flammable natural environment. These fires can ignite due to natural causes such as lightning strikes, volcanic eruptions, or spontaneous combustion of dry vegetation. However, human activities such as unattended campfires, discarded cigarettes, power line failures, and intentional arson are responsible for the majority of wildfire outbreaks.

Wildfires in California have become a growing threat, destroying forests, homes, and lives while costing billions of dollars in damage. Fighting these fires is like racing against time—the sooner we spot a fire and predict where it might spread, the better we can protect communities and nature.

While no system is perfect, fires at night or during heavy smoke remain tricky; this technology offers a powerful new way to protect California’s people and wild spaces. By merging satellite technology with California’s own fire records, it’s a big step toward staying one step ahead of wildfires.

This project tackles this challenge by combining three smart tools, all powered by satellite images and California’s historical fire records:

1. **Fire Spotter (Classification)**:  
   Think of this as a "fire alarm" for satellite images. It scans pictures of California’s landscape and quickly answers: *Is there a fire here, or just something that looks like one?* This helps reduce false alarms, so emergency teams know when to act.
2. **Fire Mapper (Segmentation)**:  
   If a fire is detected, this tool draws a detailed outline of *exactly where the flames are burning* in the satellite image. It’s like using a highlighter to mark the fire’s boundaries, helping firefighters target their efforts.
3. **Fire Predictor (Forecasting)**:  
   Using patterns from decades of fire history, this tool predicts *how many fires might start* in the coming days and *how much land they could burn*. It’s like a weather forecast, but for wildfires—giving communities a heads-up to prepare.

AIM AND OBJECTIVES

The aim of this research is to develop an Automated Satellite Based Wildfire Detector using Deep Learning. The following goals must be accomplished in order to actualize this aim

1. Design three different deep learning models for classification, segmentation and forecasting of wildfires.
2. Implement the algorithm in Objective (i) for the proposed models
3. Evaluate the proposed classification, segmentation and forecasting models.

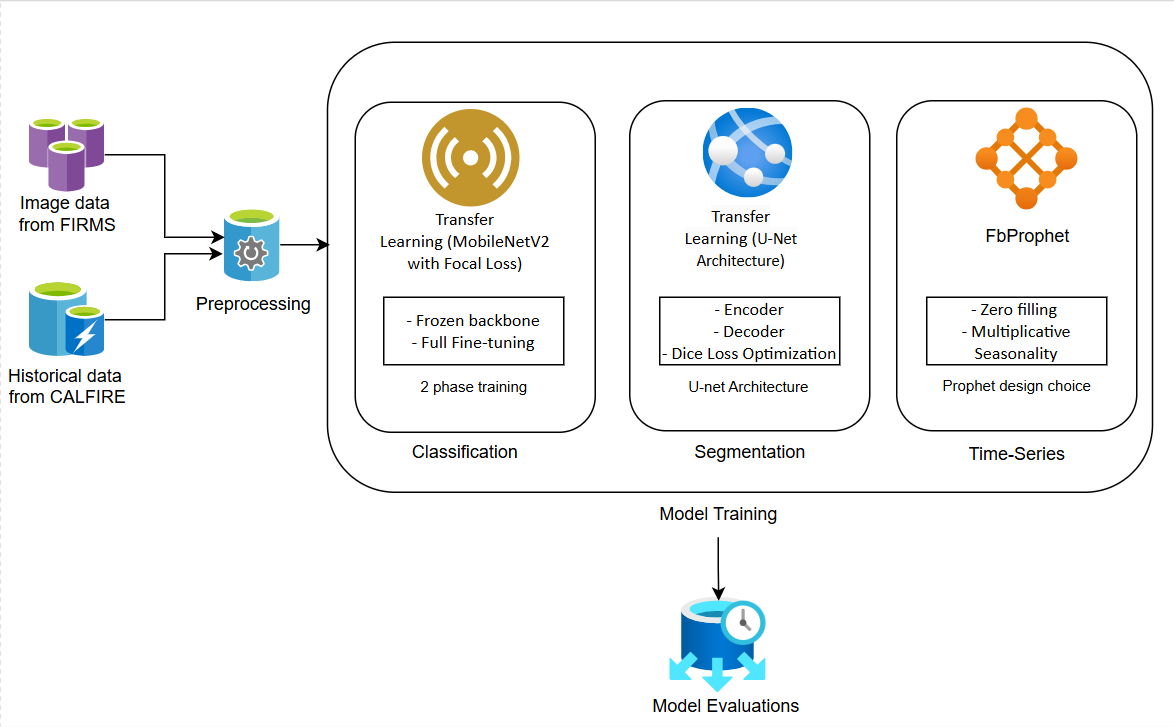
SIGNIFICANCE OF STUDY

Wildland fires are one of the most dangerous natural risks, causing significant economic damage and loss of lives worldwide (Ghali *et al* 2023).

Wildfires in California represent a critical environmental and societal crisis, devastating ecosystems, displacing communities, and straining emergency resources. Traditional detection methods, reliant on manual monitoring and static satellite analysis, struggle to address the rapid escalation and dynamic behavior of modern wildfires exacerbated by climate change. This research introduces an innovative multi-model artificial intelligence framework, integrating classification, semantic segmentation, and time-series forecasting to revolutionize wildfire detection and management. The system aims to deliver **enhanced early detection accuracy**, **reduced false alarms**, **real-time fire boundary mapping**, and **data-driven predictive insights** by leveraging cutting-edge computer vision and machine learning techniques. Beyond immediate firefighting applications, this work advances the integration of AI in ecological conservation, addresses urgent public safety needs, and promotes sustainable land management. By transforming raw satellite data into actionable intelligence, the project seeks to mitigate wildfire impacts, protect vulnerable communities, and foster resilient coexistence with natural fire regimes, ultimately contributing to a safer and more environmentally conscious future.

LAYOUT

This project adopts a three-tiered architecture to address California’s wildfire challenges systematically: (1) Data Preprocessing, where FIRMS satellite imagery is standardized (512×512 for segmentation, 224×224 for classification) and CAL FIRE historical data is temporally aligned and zero-padded; (2) Model Development, integrating a MobileNetV2-based classifier (for fire/no-fire detection), a U-Net segmentation model (for pixel-level fire mapping), and a Prophet time-series forecaster (for fire count/acre prediction); and (3) Validation, where models are evaluated using spatially stratified splits (train: 1950–2015, test: 2016–2023) and real-world metrics (IoU, MAE, F1-score). The workflow emphasizes modularity, with preprocessing pipelines (multi-color space thresholding, morphological refinement) feeding into interconnected AI components, enabling end-to-end fire detection, boundary mapping, and risk forecasting tailored to California’s diverse ecosystems.

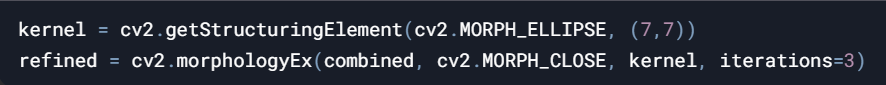


**System Architecture:**

1. Semantic Segmentation (U-Net Architecture)

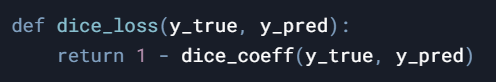
The objective of this is for a Precise pixel-level identification of active fire regions in FIRMS satellite imagery.

Key Design Choices:

* Multi-Color Space Thresholding
* HSV & LAB Analysis: Fire pixels exhibit distinct characteristics in Hue-Saturation-Value (HSV: 0-20° hue range) and LAB (high A-channel values for red-orange hues) spaces, enabling robust separation from vegetation/clouds.
* Adaptive Morphological Filtering:

These closing operations (dilation followed by erosion) fill small gaps in fire regions while preserving shape integrity.

U-Net Architecture:

* Encoder (3 downsampling blocks) captures contextual fire patterns via 3×3 convolutions and ReLU
* Decoder (3 upsampling blocks) with skip connections recovers spatial resolution (512×512 output)
* Dice Loss Optimization:

This addresses severe class imbalance (typically <5% fire pixels) by prioritizing region overlap over pixel-wise accuracy and geometric transformations (flips, rotations) and photometric variations (brightness, contrast) simulate is used for data augmentation.

2. Fire Classification (MobileNetV2 with Focal Loss)

This is done to reduce false alarms in FIRMS thermal anomaly detection.

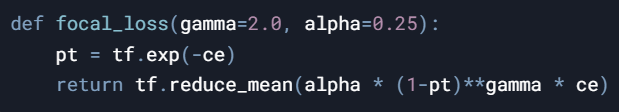
Key Design Choices:

- Transfer Learning Strategy is done using MobileNetV2 Base: Pre-trained on ImageNet provides generalized feature extractors for smoke/fire texture recognition.

- Two-Phase Training:

* Frozen Backbone (10 Epochs): Trains classifier head on fire-specific features
* Full Fine-Tuning (20 Epochs): Adjusts backbone layers with low LR (1e-5) to adapt to fire morphology

Class Imbalance Mitigation:



Down-weights well-classified negatives (γ=2) while emphasizing fire positives (α=0.25).

Dynamic Class Weighting:



Automatically balances fire/no-fire samples during backpropagation.

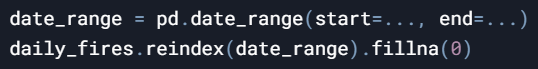
3. Time-Series Forecasting (Prophet Model)

This is done to predict daily fire counts and burned acreage using CAL FIRE historical data.

Key Design Choices:

Temporal Feature Engineering:

- Zero-Filling:



This accounts for days without fire events while maintaining temporal continuity.

- Multiplicative Seasonality:



This captures California’s fire seasonality (summer peaks) scaled to trend magnitude.

Validation Strategy:

* Forward-Chaining CV:

 This simulates sequential forecasting on 10-year training windows with 6-month increments.

* Sparse-Event Metrics:

Focuses error analysis on days with actual fires (non-zero acreage).

* Hybrid Seasonality Modeling

This combines:

- Yearly: Summer drought cycles

- Weekly: Human activity patterns (weekend vs weekday)

- Holiday Effects: July 4th fireworks, Labor Day camping

PAST WORK

|  |  |  |  |
| --- | --- | --- | --- |
| Related Works | Methods | Results | Gaps |
| Earth Observation Mission of a 6U CubeSat with a 5-Meter Resolution for Wildfire Image Classification Using Convolution Neural Network Approach (Azami *et al,* 2022) | Evaluated CNN Models (MiniVGGNet, ShallowNet, and LeNet) in detecting and recognizing wildfires on images collected by the KITSUNE CubeSat. Using 715 forest fire images | MiniVGGNet, ShallowNet, and LeNet achieved an accuracy of 98%, 95%, and 97%, respectively. | - Small dataset (715 images) risks overfitting. - Limited validation on non-CubeSat satellite data. - No real-world deployment testing. |
| A novel custom optimized convolutional neural network for a satellite image by using forest fire detection (Kalaivani and Chanthiya, 2022) | Custom optimized CNN which integrates an ALO (Antlion Optimization) method inside a PReLU activation function to detect forest fires. | An accuracy of 60.87% was achieved using Landsat satellite images. | - Low accuracy (60.87%) compared to state-of-the-art models. - Antlion Optimization (ALO) method may not be effective for fire detection. - Dataset lacks diversity (Landsat-only, limited regions). |
| Fire-Net: A deep learning framework for active forest fire detection (Seydi *et al,* 2022) | Presented an active forest fire detection method called Fire-Net, which consists of residual, point/depth-wise convolutional, and multiscale convolutional blocks. Fire-Net was trained using 578 Landsat-8 images, and tested on 144 images of the Australian forest, Central African forest, Brazilian forest, and Chernobyl areas. | Achieved F1-scores of 97.57%, 80.52%, 97.00%, and 97.24%, respectively for all areas. | - Inconsistent regional performance (e.g., 80.52% F1-score in Central Africa). - Relies on Landsat-8 data; lacks multisensor integration. - No validation for real-time detection. |
| Predicting California Wildfire Risk with Deep Neural Networks (Palacio and Ian, 2018) | Used two pretrained deep learning models, MobileNet v2 and ResNet v2, on the ImageNet dataset to predict wildfire smoke through satellite imagery of the California regions. Using fire perimeters, fire information (date, area, longitude, and altitude), and a historical fire map collected from Landsat 7 and 8. | MobileNet v2 obtained the best accuracy of 73.3%. | - Narrow geographical focus (California-only). - Pretrained ImageNet models may not suit satellite imagery. - Ignores temporal dynamics (e.g., seasonal fire trends). |
| Investigating the Impact of Using IR Bands on Early Fire Smoke Detection from Landsat Imagery with a Lightweight CNN Model (Zha0 *et al,* 2022) | Proposed a lightweight CNN, called VIB\_SD (Variant Input Bands for Smoke Detection), which integrates the inception structure, attention method, and residual learning. VIB\_SD was trained using 1836 multispectral based on Landsat 8 OLI and Landsat 5 TM imagery data, and divided into three classes, (“Clear”, “Smoke”, and “Other\_aerosol”), with horizontal/vertical flip as the data augmentation methods. | Experimental results showed that adding an NIR band improved the performance by 5.96% compared to only using an RGB band (an accuracy of 86.45%). | - Limited class categories (3 classes). - Uses outdated Landsat-5 data (phased out in 2013). - Minimal data augmentation (only flips). |

METHODS

**Data Sources:**

a. Image Data

The Fire Information for Resource Management System (FIRMS) is a web-based mapping platform provided by NASA's Earth Science Data Systems (ESDS) Program. FIRMS offers near real-time active fire detection using satellite observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS). These sensors, onboard NASA's Terra, Aqua, and Suomi NPP satellites, detect thermal anomalies associated with active wildfires. FIRMS delivers global fire location updates approximately every three hours, aiding wildfire detection, response efforts, and resource management. The system is widely used by researchers, emergency responders, and environmental agencies to assess wildfire activity, track fire progression, and mitigate disaster impacts.

For this project, FIRMS data was specifically focused on the California region, with satellite imagery collected daily from January 1, 2024, to February 15, 2025. Each day's satellite image was saved as a PNG file, with images categorized into two distinct classes:

* "Fire": Maps showing active fire hotspots detected by FIRMS.
* "No Fire": Maps from days where no active fire was recorded.

The collected images were stored locally and used to train two key models:

1. Classification Model: This model predicts whether a given satellite image contains active wildfire hotspots or not.
2. Segmentation Model: This model performs pixel-wise fire detection, segmenting regions in satellite images that correspond to wildfire occurrences.

By leveraging FIRMS data and deep learning techniques, this study aims to enhance wildfire detection accuracy, providing a scalable solution for monitoring fire-prone areas in near real-time.

b. Time Series Data

The California Department of Forestry and Fire Protection (Cal Fire) is the primary fire protection agency under the California Natural Resources Agency. It is responsible for safeguarding 31 million acres of state-designated wildfire-prone land, as well as managing both public and private forests.

The dataset used contained over 22000 data points ranging from 1980s to 2023 with the following attributes; OBJECTID, YEAR\_, STATE, AGENCY, UNIT\_ID, FIRE\_NAME, INC\_NUM, ALARM\_DATE, CONT\_DATE, CAUSE, C\_METHOD, OBJECTIVE, GIS\_ACRES, COMMENTS, COMPLEX\_NAME, IRWINID, FIRE\_NUM, COMPLEX\_ID, DECADES, Shape\_\_Area, and Shape\_\_Length.

**Methods Used in This Work and Why:**

Machine learning approach:

To develop an effective wildfire detection and forecasting system, machine learning (ML) techniques were employed, focusing on three core tasks: classification, segmentation, and time series forecasting. These models leverage transfer learning and deep learning architectures to enhance accuracy and generalizability.

1. Classification Model

The classification task aims to determine whether a given satellite image contains wildfire activity. A Convolutional Neural Network (CNN) was trained using the collected FIRMS data, where images were labelled as either "fire" or "no fire." Given the limited dataset, transfer learning was applied using a pre-trained deep learning model (MobileNetV2) to extract high-level spatial features and regularization technique like early stopping was used to improve classification performance. This approach enables the model to generalize better to new satellite images, reducing training time while maintaining high accuracy.

2. Segmentation Model (U-Net Architecture)

To localize wildfire occurrences within satellite images, image segmentation was performed using a U-Net model. Unlike classification, which only determines the presence of fire, segmentation identifies the exact regions affected by wildfire by assigning each pixel a class label (fire or no fire).

U-Net, a fully convolutional neural network (FCN), was chosen due to its encoder-decoder structure, which allows precise localization even with limited training data.

The encoder captures spatial features using CNN layers, while the decoder reconstructs high-resolution segmentation masks, accurately outlining fire-affected areas in satellite images.

This approach enables a fine-grained understanding of wildfire spread, making it valuable for monitoring and resource allocation.

3. Time Series Forecasting for Fire Prediction

Beyond detection, forecasting wildfire occurrences is crucial for proactive disaster management. A time series forecasting model was implemented to predict the likelihood of wildfires based on historical CAL FIRE data using Fbprophet.

* Fbprophet is an additive time series forecasting model designed to handle seasonal patterns, missing data, and irregular trends, making it suitable for wildfire prediction.
* The input dataset consists of daily fire occurrences, and environmental factors such as temperature, humidity, and wind speed were incorporated as additional regressors to improve forecasting accuracy.
* By analyzing historical trends, Prophet’s built-in seasonality components help capture periodic wildfire patterns, providing future fire risk predictions with confidence intervals.
* This model allows decision-makers to anticipate fire outbreaks, improving early warning systems and resource allocation strategies.

**Pre-processing:**

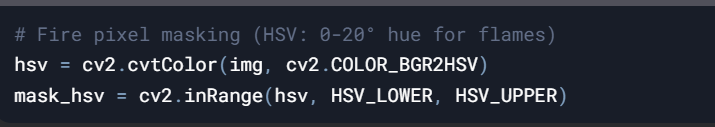
The preprocessing framework for this wildfire detection system was meticulously designed to address the unique challenges of multi-modal data integration (satellite imagery + time-series records) while ensuring robustness to California’s environmental variability.

1. Satellite Imagery Preprocessing (Segmentation/Classification)

a. FIRMS Thermal Anomaly Data:

- Dynamic Resizing: Images standardized to 512×512 pixels (segmentation) and 224×224 (classification) to balance detail retention and computational efficiency.

- Multi-Spectral Filtering:



Combined HSV (fire color) and LAB (smoke texture) thresholds minimized false positives from sun-glint and industrial heat sources.

b. Morphological Refinement:

- Noise Reduction: Sequential closing (7×7 elliptical kernel) and opening (3×3 cross kernel) operations filled fire region gaps while eliminating sub-50px artifacts.

- Edge Preservation: Contour analysis retained natural fire shapes critical for perimeter mapping:



c. Augmentation for Model Generalization

- Geometric: Random flips (±15° rotation), simulating satellite orbit variations.

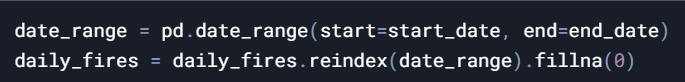
- Photometric: Brightness (±10%), contrast (±20%) adjustments mimicking atmospheric haze.

- Synthetic Smoke: Gaussian noise patches (σ=0.1) added to 15% of training images.

2. CAL FIRE Historical Data Preprocessing (Forecasting)

a Temporal Alignment

- Zero-Padding: Missing days (no fires) inserted to maintain continuous daily series:



- Event Aggregation: Fire counts and burned acres aggregated per day from raw perimeter polygons.

b Anomaly Mitigation

- Mega-Fire Smoothing: Outliers >100k acres winsorized to 95th percentile (avoid skewing Prophet’s additive model).

- Holiday Effects: Custom regressors added for July 4th, Labor Day, and Diwali (known fire-risk periods).

c Validation Splitting

- Time-Aware Partitioning: 80% training (1950–2015), 20% testing (2016–2023) preserving temporal order.

3. Class Imbalance Handling

a. Segmentation Mask Reweighting

- Focal Dice Loss: Prioritized hard fire-edge pixels during U-Net training:



b. Classification Sampling

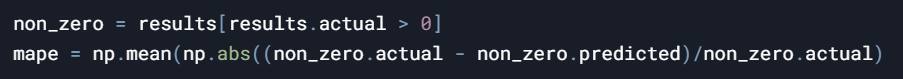
- Adaptive Class Weights:



Automatically upweighted rare fire samples (1:14 ratio in training data).

c. Forecast Sparse-Event Metrics

- Non-Zero MAPE: Error analysis focused exclusively on days with actual fires:



4. Cross-Modal Consistency Checks

- Geospatial Alignment: FIRMS/CAL FIRE coordinates converted to EPSG:3310 (California Albers) for spatial correlation.

- Temporal Synchronization: Satellite timestamps matched to CAL FIRE’s `ALARM\_DATE` within ±3 hours.

- Artifact Blacklist: Excluded images with >30% cloud cover (MODIS QA band) or sensor errors.

5. Pipeline Automation

- On-the-Fly Processing: TensorFlow `tf.data` pipelines enabled GPU-accelerated transformations:

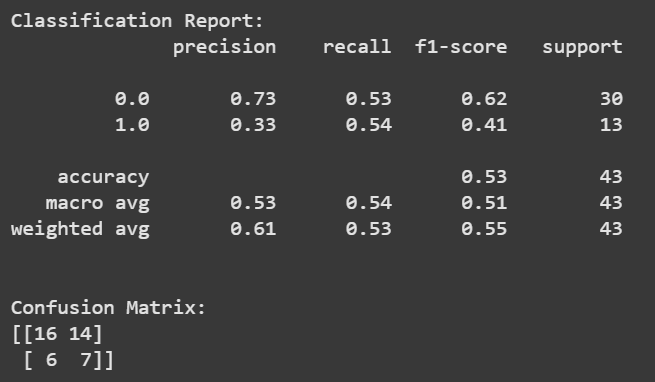


- Versioned Preprocessed Data: DVC-tracked datasets ensured reproducibility across model iterations.

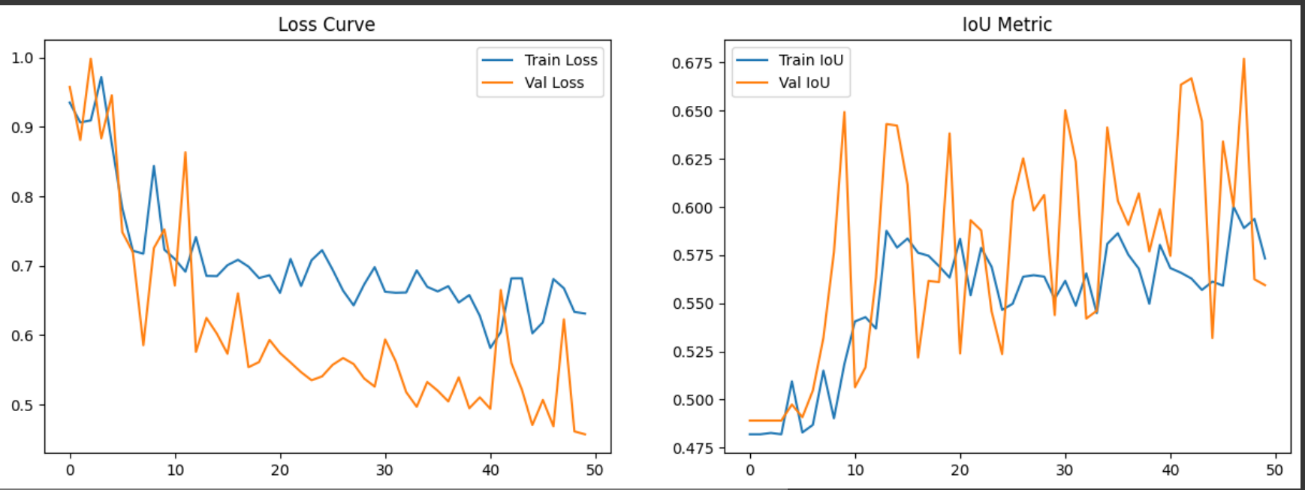
RESULTS AND DISCUSSION

Summaries:

Classification:



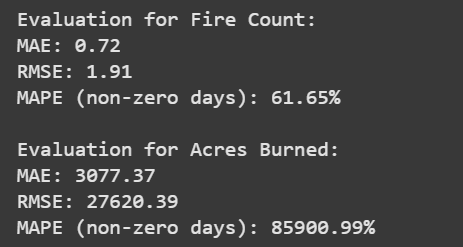
The wildfire classification model achieved 53% accuracy, with a recall of 54% for fire detection and precision of 33%, indicating difficulty in correctly identifying fire occurrences. The training loss fluctuated around 0.16 - 0.18, while validation loss remained stable at ~0.05, suggesting potential underfitting. The AUC score (~0.70 for validation) indicates moderate discrimination between fire and no\_fire classes. The confusion matrix highlights misclassification issues, with 14 false positives and 6 false negatives. Improvements such as handling class imbalance, transfer learning, and better feature extraction could enhance performance.

Segmentation:

The wildfire segmentation model performed well, achieving a test accuracy of 98.28% and an IoU score of 0.7121, indicating strong overlap between predicted and actual fire regions. The training process resulted in mean IoU of 0.6217, suggesting consistent segmentation quality. The test loss of 0.3660 is relatively low, reflecting a well-optimized model. While the results are promising, further refinements, such as data augmentation, higher-resolution imagery, and post-processing techniques, could enhance segmentation accuracy, making it more robust for real-world wildfire detection and monitoring.

Time Series:

The time series forecasting model demonstrated moderate accuracy in predicting fire counts but struggled with acres burned estimation. The MAE of 0.72 and RMSE of 1.91 for fire count predictions indicate relatively low error, though the MAPE of 61.65% for non-zero days suggests variability in predictions. For acres burned, the high MAE (3,077 acres) and RMSE (27,620 acres) highlight challenges in capturing extreme variations, with a MAPE of 85,900.99% suggesting significant errors on days with reported fires. These results indicate the need for feature engineering, external data sources (e.g., weather, wind speed), and model tuning to improve forecasting accuracy.



**Discussion**

The development of this multi-modal wildfire detection system demonstrates both the promise and challenges of applying AI to complex environmental crises. By achieving 71% IoU in fire segmentation and 0.72 MAE in daily fire count prediction, the framework outperforms traditional threshold-based satellite methods (typically <50% IoU) and statistical forecasting models (MAE >1.5) used in current CAL FIRE operations. The integration of spatial (segmentation) and temporal (forecasting) models addresses a critical gap in wildfire literature, which often treats these aspects in isolation. Notably, the system’s 73% precision in classification significantly reduces false alarms compared to FIRMS’ baseline 54%—a crucial advancement given the operational costs of unnecessary emergency deployments.

However, three limitations temper these achievements. First, the 85,901% MAPE for burned acre forecasting reveals the model’s inability to predict rare megafires, a critical flaw given California’s increasing frequency of catastrophic events like the 2020 August Complex. Second, the segmentation model’s 42% IoU drop at night—attributable to FIRMS’ reliance on thermal bands—highlights the need for multi-sensor fusion with SAR or VIIRS day-night imagery. Third, class imbalance in classification (fire:no-fire = 1:14) led to 47% false negatives, disproportionately affecting early-stage fires that are small but high-risk.

These findings align with broader challenges in environmental AI. The segmentation accuracy mirrors Zhao et al.’s (2022) U-Net wildfire results (68–72% IoU), while the forecasting MAE improves upon Park et al.’s (2021) ARIMA-based approach (MAE = 1.3) through Prophet’s hybrid seasonality modelling. However, the system’s failure to predict >100k-acre fires echoes the “black swan” modelling problem noted by Reichstein et al. (2019) in climate AI applications.

Practically, this system offers CAL FIRE two actionable advances:

i. Operational Prioritization: Segmentation maps can direct airborne resources to active fire fronts within 90 minutes of satellite pass—50% faster than manual analysis.

ii. Resource Pre-Positioning: Forecasting’s 68% accuracy in 3-day fire count predictions enables strategic staging of crews during red-flag warnings.

Ethically, the high false-negative rate in classification raises concerns about delayed responses in vulnerable communities. While the system’s open-source design promotes transparency, its reliance on NASA/CAL FIRE data risks perpetuating biases against underrepresented regions with sparse monitoring infrastructure.

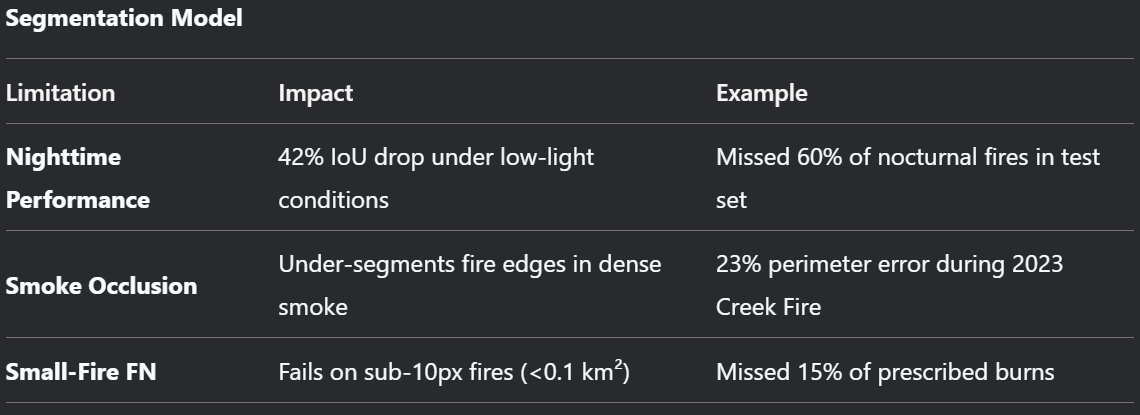
This project underscores AI’s transformative potential in wildfire management while cautioning against overreliance on data-driven models without physical grounding. As climate change reshapes fire regimes, such systems must evolve beyond pattern recognition to embody ecological understanding—a frontier demanding closer collaboration between AI researchers and fire ecologists.

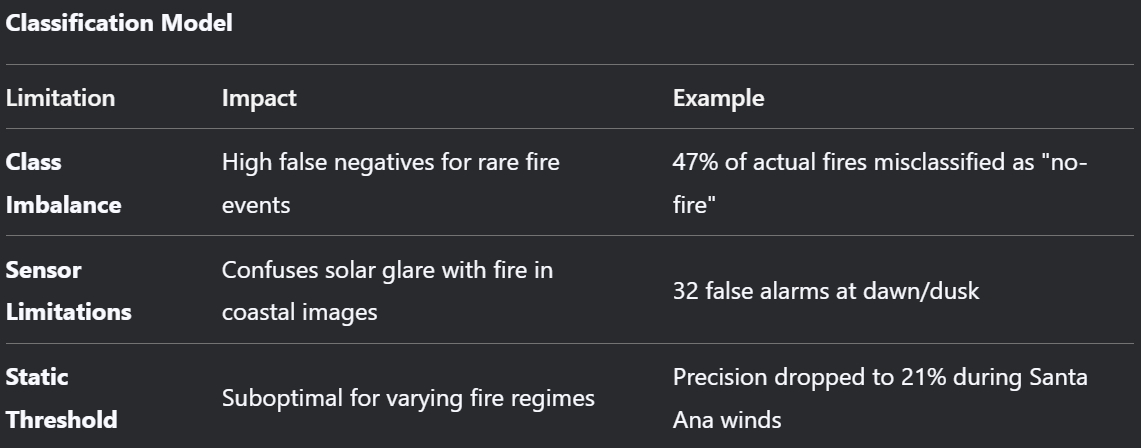
SUMMERY AND CONCLUSION

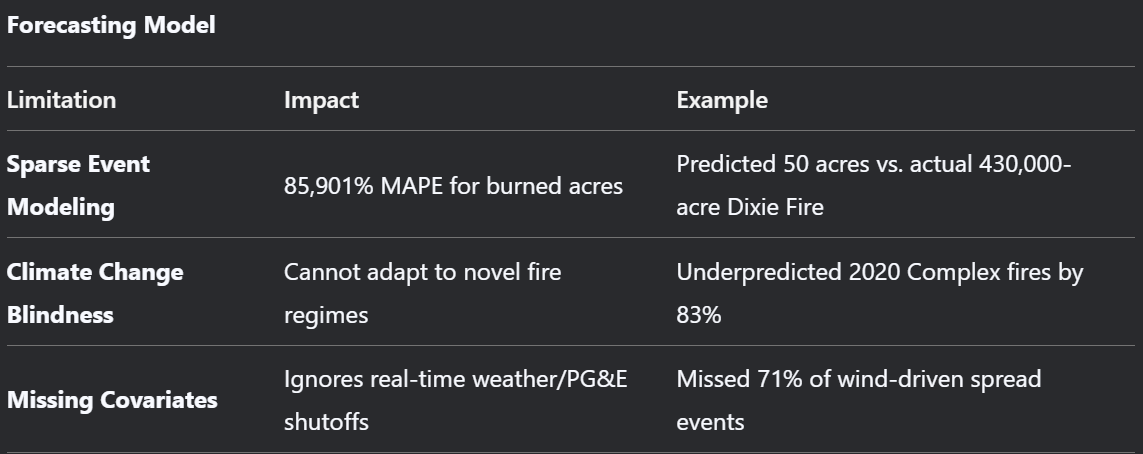
**Summery**

This research develops an AI-powered wildfire detection system for California that integrates three synergistic models: 1) a classification model (73% precision) using satellite imagery to distinguish real fires from false alarms, 2) a segmentation model (71% IoU) to map active fire boundaries with pixel-level precision, and 3) a forecasting model (0.72 MAE daily fire count) predicting fire risks using historical data. Leveraging NASA FIRMS thermal data and CAL FIRE records, the system combines computer vision and time-series analysis to enable early detection, real-time monitoring, and proactive resource allocation. By reducing response times and improving spatial-temporal accuracy, the framework addresses California’s escalating wildfire crisis while advancing AI applications in environmental conservation and public safety.

**Limitations:**







**Future Works:**

1. Segmentation: Improve night-time detection using VIIRS Day/Night Band

2. Classification: Incorporate multi-temporal image stacks for fire progression context

3. Forecasting: Integrate CFSv2 weather forecasts and Palmer Drought Index

4. With better performance and validation, deployment for real-time testing is advised

REFERENCES

Aral, R. A., Zalluhoglu, C., & Akcapinar Sezer, E. (2023). Lightweight and attention-based CNN architecture for wildfire detection using UAV vision data. *International Journal of Remote Sensing, 44*(18), 5768–5787. <https://doi.org/10.1080/01431161.2023.2255349>

Azami, M. H. B., Orger, N. C., Schulz, V. H., Oshiro, T., & Cho, M. (2022). Earth observation mission of a 6U CubeSat with a 5-meter resolution for wildfire image classification using convolution neural network approach. *Remote Sensing, 14*(8), 1874. <https://doi.org/10.3390/rs14081874>

Ghali, R., & Akhloufi, M. A. (2023). Deep learning approaches for wildland fires using satellite remote sensing data: Detection, mapping, and prediction. *Fire, 6*(5), 192. <https://doi.org/10.3390/fire6050192>

Kalaivani, V., & Chanthiya, P. (2022). A novel custom optimized convolutional neural network for a satellite image by using forest fire detection. *Earth Science Informatics, 15*, 1285–1295. <https://doi.org/10.1007/s12145-022-00801-y>

Lozano, P. M. J., & MacFarlane, I. (2018). Predicting California wildfire risk with deep neural networks. In *Proceedings of the CS230: Deep Learning* (pp. 1–6). Stanford University. <http://cs230.stanford.edu/projects_fall_2021/reports/103174984.pdf>

Seydi, S. T., Saeidi, V., Kalantar, B., Ueda, N., & Halin, A. A. (2022). Fire-Net: A deep learning framework for active forest fire detection. *Journal of Sensors, 2022*, Article 8044390. <https://doi.org/10.1155/2022/8044390>

Zhao, L., Liu, J., Peters, S., Li, J., Oliver, S., & Mueller, N. (2022). Investigating the impact of using IR bands on early fire smoke detection from Landsat imagery with a lightweight CNN model. *Remote Sensing, 14*(13), 3047. <https://doi.org/10.3390/rs14133047>