Multi-Task Learning for Fire Ignition Maps

Parastoo Hashemi : s339438
Yousef Fayyaz: s341999

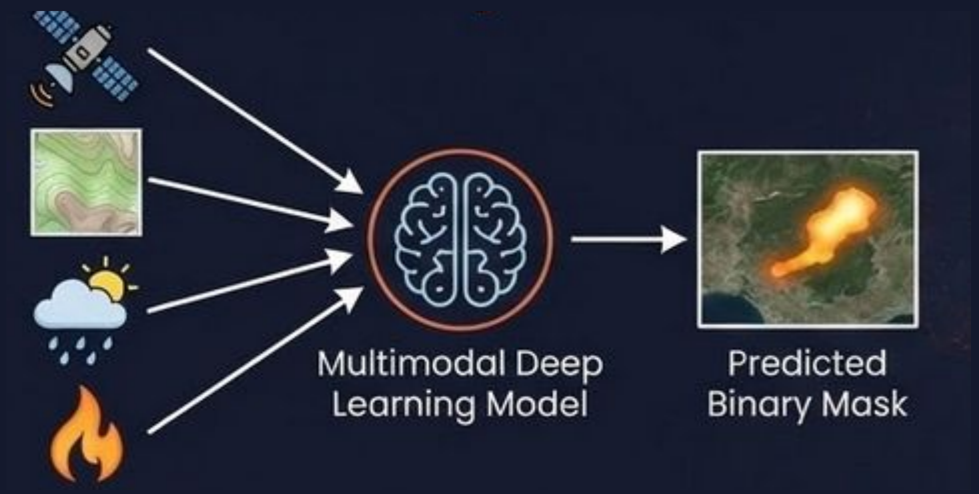
Project Objective

Core Objective

To develop a multimodal deep learning model capable of **Predicting the final burned area extent right at the moment of ignition,** using only **pre-fire multimodal data that** fused into a multimodal model

Input: Pre-fire environmental data (Satellite, Topography, Weather) + Ignition Point.

Output: Predicted binary mask of the final burned area.



The Paradigm Shift

Current State (Reactive): Remote sensing is primarily used for post-event damage assessment.

Target State (Proactive): To estimate the spread of fire immediately to allocate resources immediately during the "Golden Hour".

Value Proposition

Environmental Impact

Proactive predictions reduce environmental losses by enabling faster response before uncontrolled spread.

Fewer hectares burned → lower CO_2 emissions, preserved ecosystems, and reduced long-term regeneration costs.

Response Efficiency

Early burned-area forecasts help Civil Protection allocate firefighting teams, helicopters, and ground units efficiently.

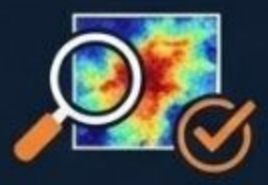
Better situational awareness during the first critical hours improves containment and reduces operational risk.

Innovation & Scalability

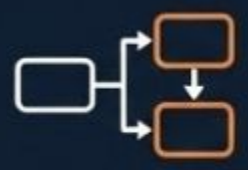
The multimodal AI system generalizes across regions where Sentinel-2, Landsat, DEM, ERA5, and infrastructure layers are available.

The framework can be replicated globally, providing a reusable template for data-driven wildfire risk management.

Research Questions



Can pre-fire multimodal data predict final burned area with sufficient accuracy to support early emergency decisions?



Why are standard metrics (accuracy) insufficient for spatial wildfire decision-making?

which input data modalities contribute most to predictive performance?

LEARNING PROBLEM DEFINITION: BINARY SEMANTIC SEGMENTATION

INPUT: MULTIMODAL DATA CUBE

OUTPUT: DENSE SPATIAL MASK

Ground Truth/Target

Predicted Segmentation Mask

TASK FORMULATION: PIXEL-LEVEL CLASSIFICATION

Each pixel is strictly classified as "Burned" (1) or "Not Burned" (0).

DENSE SPATIAL OUTPUT

Generates a full spatial map, not a single value. Focus is on the entire spatial extent.

LEARNING OBJECTIVE: SPATIAL COHERENCE

Simple Accuracy (Poor Geometry)

Geometry-Aware (High Coherence)

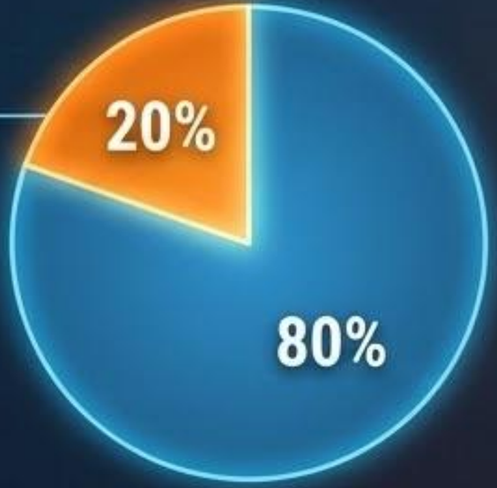
Emphastzes capturing the precise shape and boundary, not just pixel counts.

SCIENTIFIC INTENT

Positioning the work as **dense prediction** (not patch-level) and **geometry-aware segmentation** (not detection).

DATASET PREPARATION & CHALLENGES

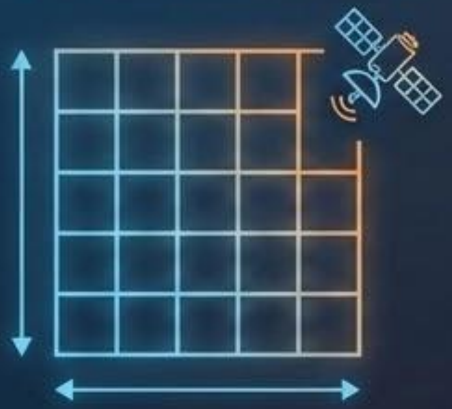
DATASET SPLIT (80/20)



TRAINING SET (872 Fires)
Diverse geographical locations (Piedmont)

VALIDATION SET (218 Fires)
Independent evaluation, maintaining diversity

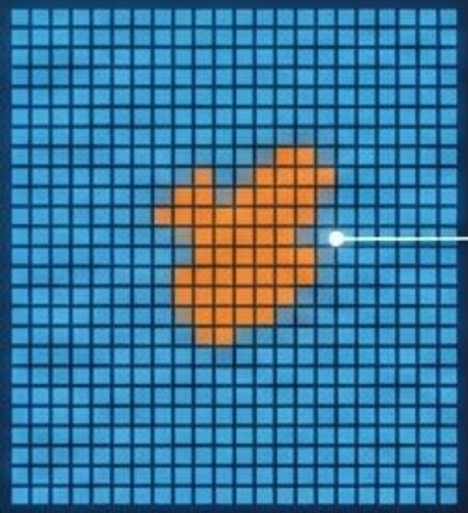
INPUT CONFIGURATION



256x256
PATCHES

All inputs processed as 256×256 patches for sufficient spatial context and computational feasibility.

DATASET CHALLENGE: EXTREME CLASS IMBALANCE



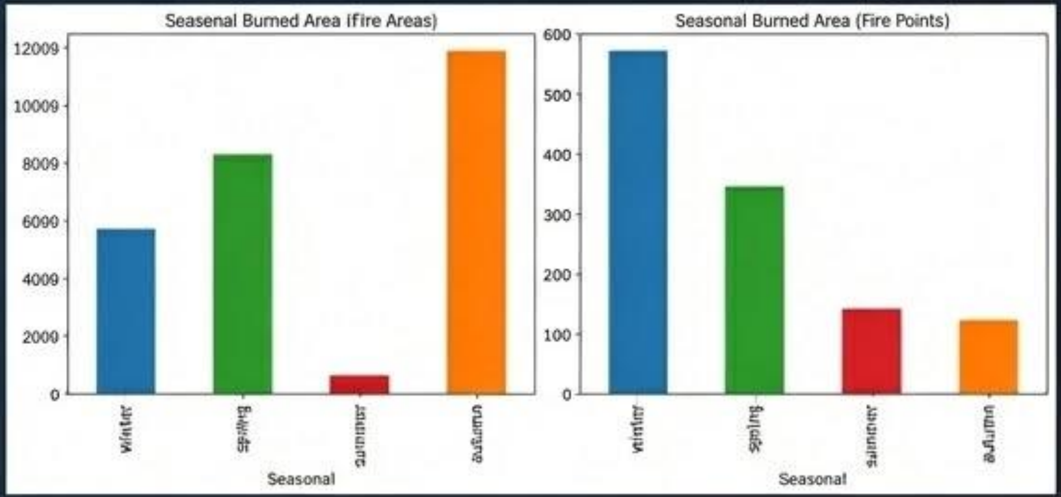
Burned pixels are rare (e.g., <1%). Standard models bias towards the majority background class.

SOLUTION: MULTITASK LEARNING STRATEGY
Jointly optimizing burned area segmentation to address imbalance.

Dataset Characteristics & Challenges

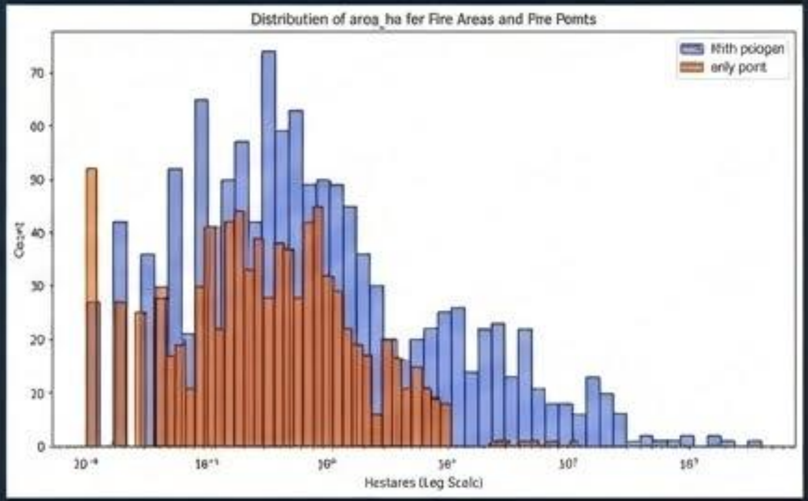
Scientific Context: Seasonal Patterns & Imbalance

Seasonal Disparities: Frequency vs. Severity



- The number of fires doesn't match how much land gets burned.
- Winter has the most fire starts, but they usually stay small.
- Autumn has fewer fires, but they grow much bigger and burn more land.
- This means we need to predict how *bad* a fire will be, not just if it will start.

Size Distribution & Dataset Imbalance



- Our data is very unbalanced, with a 'long tail' of rare events.
- Most fires are tiny (less than 1 hectare, like a small field).
- The "Polygon" data (blue) shows the rare but huge fires (>100 ha) that cause the most damage.
- This imbalance makes it hard to teach a computer model to predict the big, important fires.

Loss Design, Optimization Objective, and Primary Metric: IoU

Binary Cross-Entropy (BCE)

Captures pixel-wise classification accuracy.

Dice Loss

Promotes regional overlap, addressing of global shape.

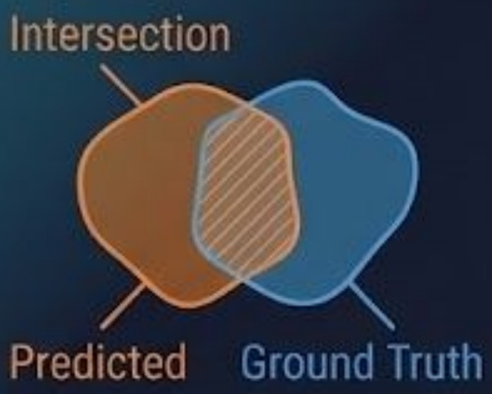
Combined Loss

Balances local precision with spatial coherence.

Alignment: The objective function is directly optimized for the evaluation metric (IoU).

SCIENTIFIC INTENT

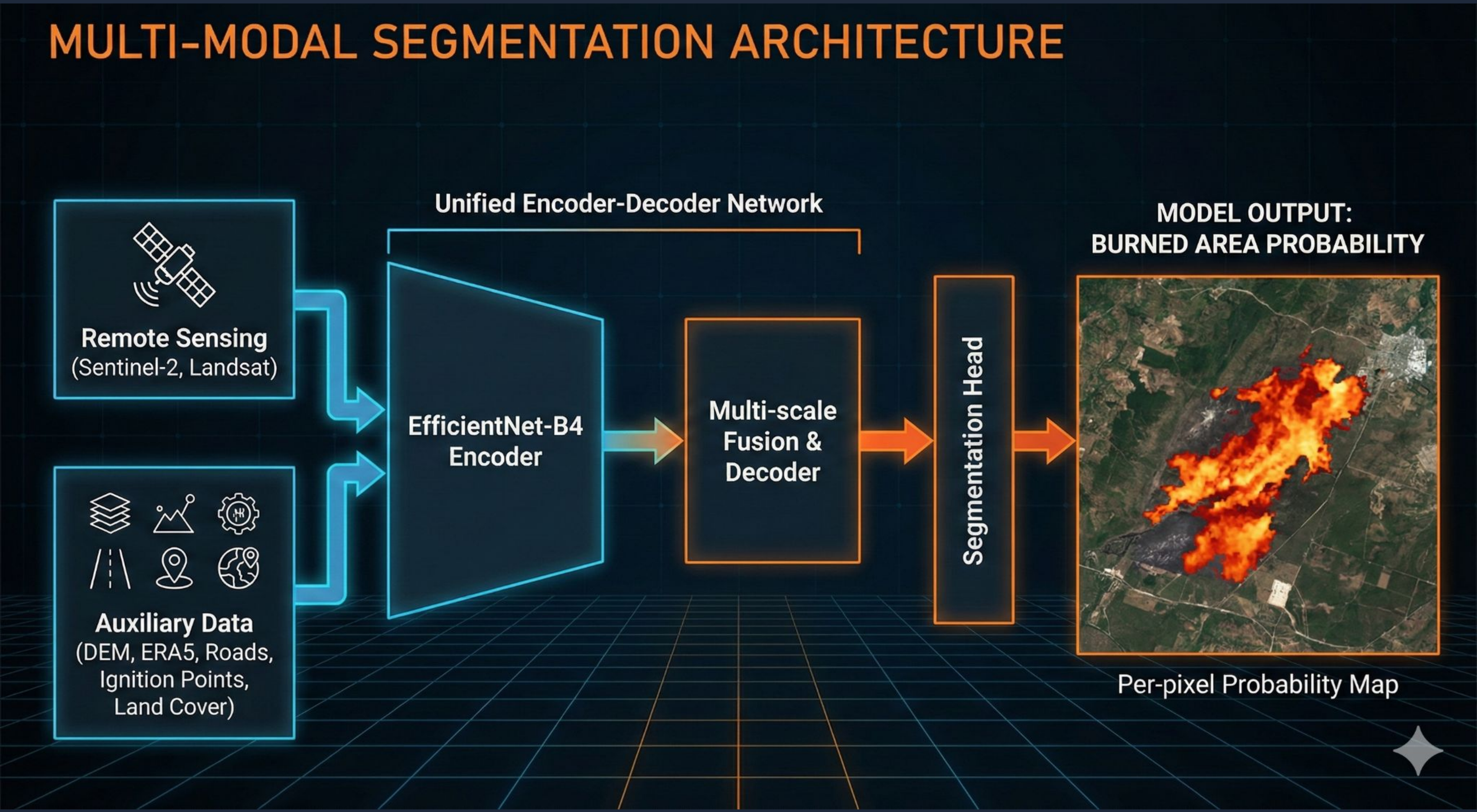
- Awareness of class imbalance.
- Awareness of the metric–loss mismatch.

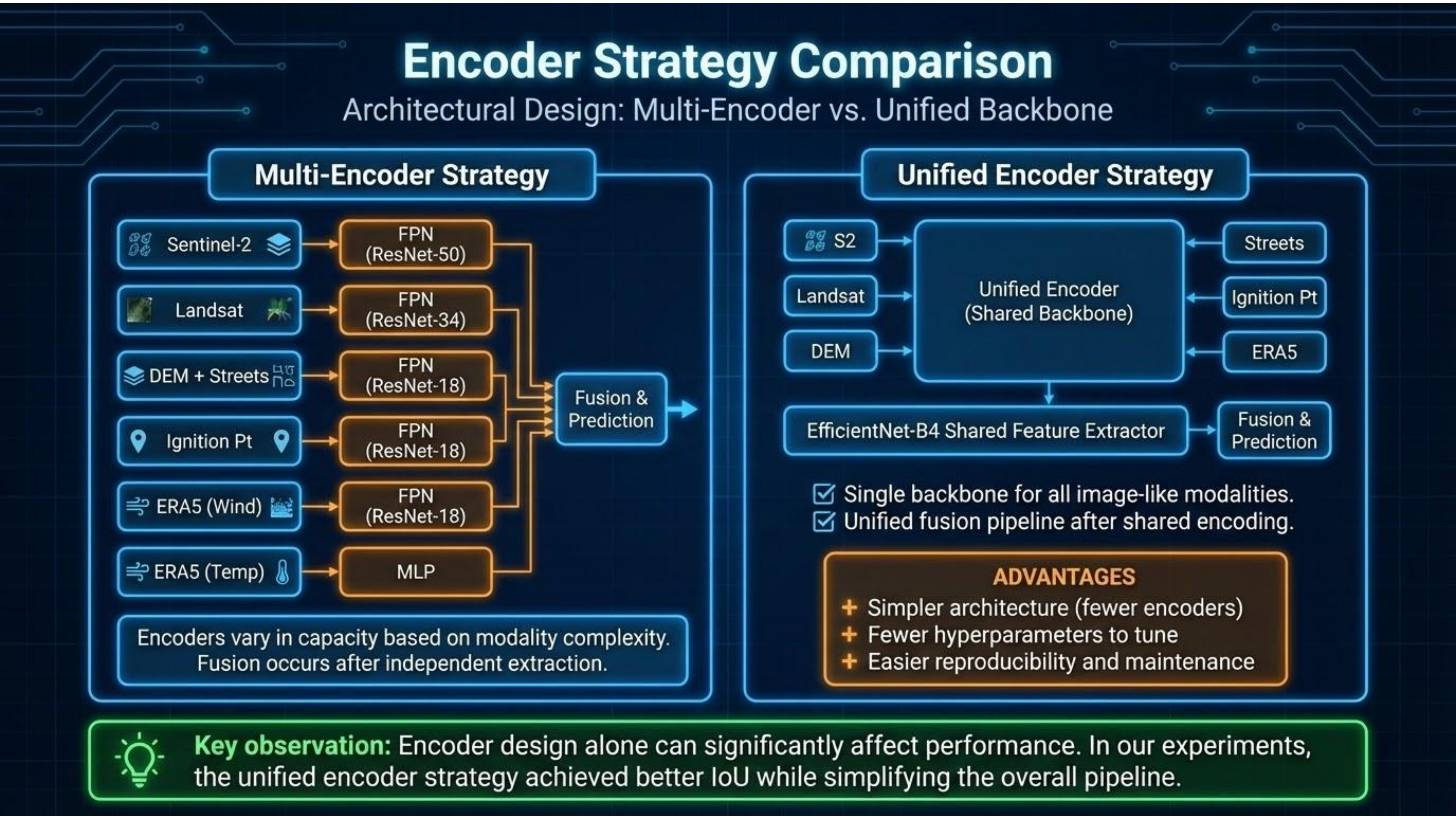


$$\text{IoU} = \frac{\text{Intersection}}{\text{Union}}$$

PRIMARY METRIC: IoU

Intersection over Union drives the evaluation.


MULTI-MODAL SEGMENTATION ARCHITECTURE
Unified Encoder-Decoder Network
Remote Sensing
(Sentinel-2, Landsat)
Auxiliary Data
(DEM, ERA5, Roads, Ignition Points, Land Cover)
EfficientNet-B4 Encoder
Multi-scale Fusion & Decoder
Segmentation Head
MODEL OUTPUT:
BURNED AREA PROBABILITY
Per-pixel Probability Map


Encoder Strategy Comparison
Architectural Design: Multi-Encoder vs. Unified Backbone
Multi-Encoder Strategy
Sentinel-2
FPN (ResNet-50)
Landsat
FPN (ResNet-34)
DEM + Streets
FPN (ResNet-18)
Ignition Pt
FPN (ResNet-18)
ERA5 (Wind)
FPN (ResNet-18)
ERA5 (Temp)
MLP
Fusion & Prediction
Encoders vary in capacity based on modality complexity.
Fusion occurs after independent extraction.
Unified Encoder Strategy
S2
Landsat
DEM
Streets
Ignition Pt
ERA5
Unified Encoder (Shared Backbone)
EfficientNet-B4 Shared Feature Extractor
Fusion & Prediction
Single backbone for all image-like modalities.
Unified fusion pipeline after shared encoding.
ADVANTAGES
+ Simpler architecture (fewer encoders)
+ Fewer hyperparameters to tune
+ Easier reproducibility and maintenance
Key observation: Encoder design alone can significantly affect performance. In our experiments, the unified encoder strategy achieved better IoU while simplifying the overall pipeline.

RQ1 — Can Pre-Fire Multimodal Data Support Early Emergency Decisions?

Pre-fire multimodal data enables meaningful early estimation of final burned area, even before fire spread is observed.

Quantitative Evidence

+70%
IoU improvement via decision threshold optimization

High precision regime reduces false alarms (critical for emergency response)

Operational Interpretation

- Goal is early containment planning, not perfect boundaries
- Conservative predictions prioritize high-confidence burned cores
- Large fires (highest impact events) are detected with strong spatial coherence
- Model supports resource prioritization, not final damage assessment

LIMITATIONS

- Small or rapidly evolving fires remain challenging
- Model prioritizes precision over recall
- Intended for early-stage guidance only

RQ3 – Which modalities contribute most?

METHOD

We estimate modality contribution using inference-time ablation: run the same trained multimodal model on the same validation subset, but **zero-out one modality at a time.** The performance drop (ΔIoU, ΔF1) indicates contribution.

Conclusion: The modality causing the largest drop in IoU/F1 is the most influential. **Landsat** and **ERA5** show the highest contribution.

Landsat & ERA5 (High) → **Other** (Medium) → **Ignition** (Low)

Modality Importance: Performance Drop (Higher is Better)

Baseline vs Multinodal: Quantitative Gain

EXPERIMENTAL SETUP

Parameter	Configuration
Model	Baseline vs Fusion
Encoder	ResNet34 vs EffNet-B4
Input	Sentinel vs Multimodal
Epochs	40 vs 120

Table 1. *Controlled setup differences including encoder architecture, input modalities, and training schedule duration.*

PERFORMANCE METRICS

IOU IMPROVEMENT

+0.251 `0.152 → 0.403`

F1 IMPROVEMENT

+0.311 `0.263 → 0.574`

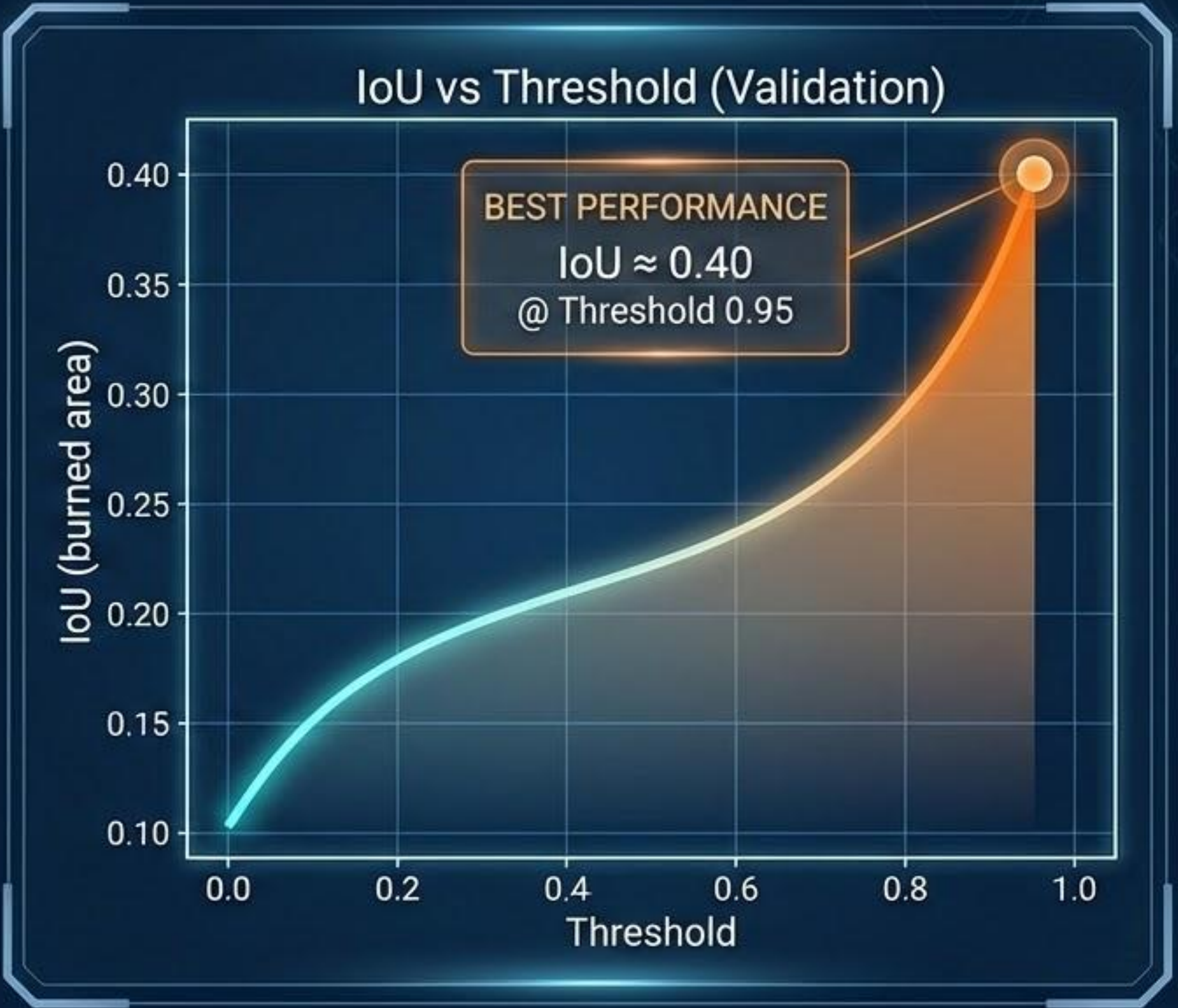
Insight: Main gain comes from recall (0.212 → 0.572), indicating significantly fewer missed burned pixels in early detection scenarios.

Methodology: Baseline evaluated on quick CPU subset (~120 val samples); thresholds {0.50, 0.75, 0.90, 0.95}. Multimodal data sourced from

EFFECT OF DECISION THRESHOLD ON IOU

- Default threshold (0.5) significantly underestimates performance.
- IoU increases monotonically with the threshold.
- **Best Performance:** Validation IoU ≈ 0.40 at threshold 0.95.

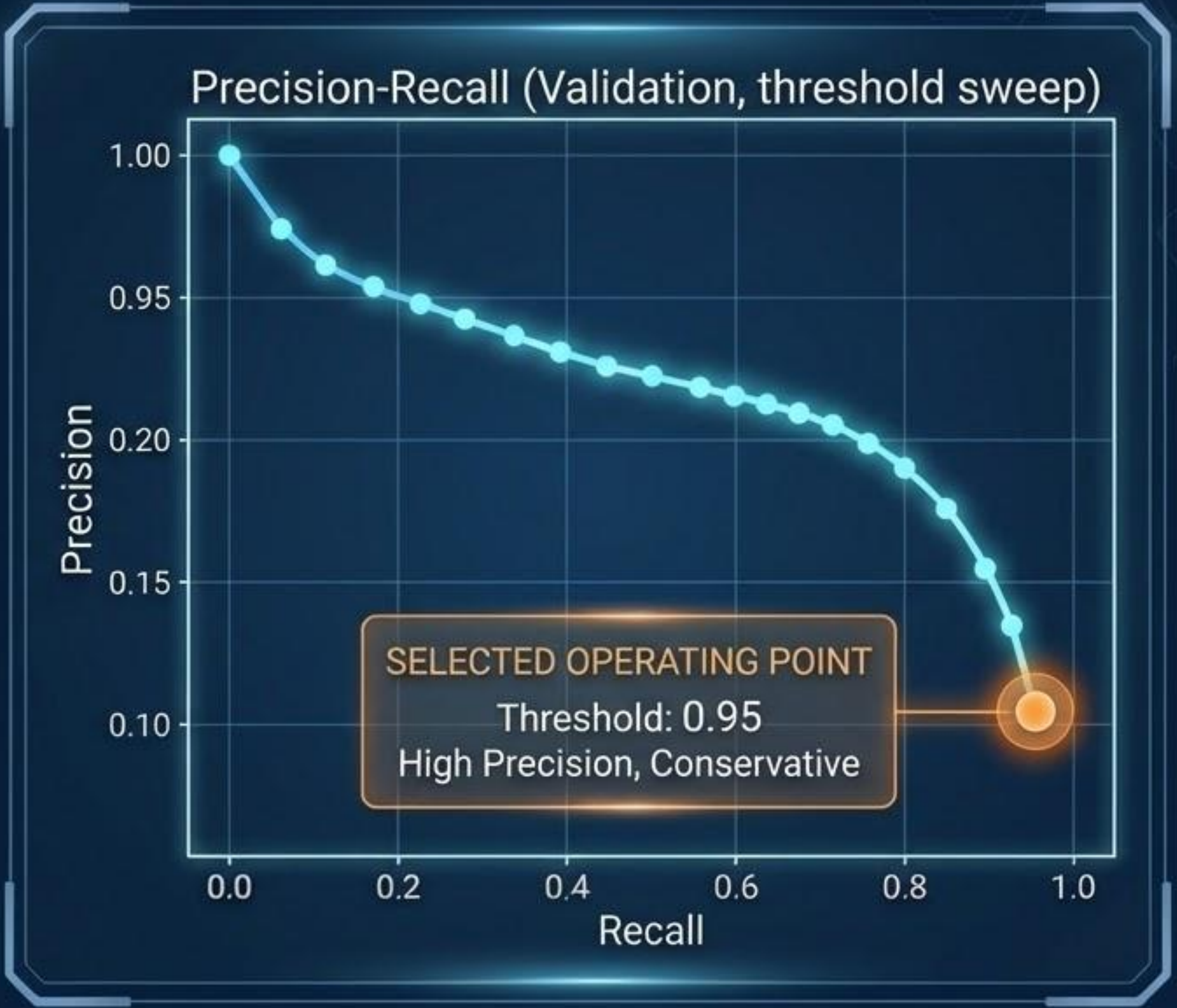
Threshold calibration improves IoU by **~70%** without retraining.



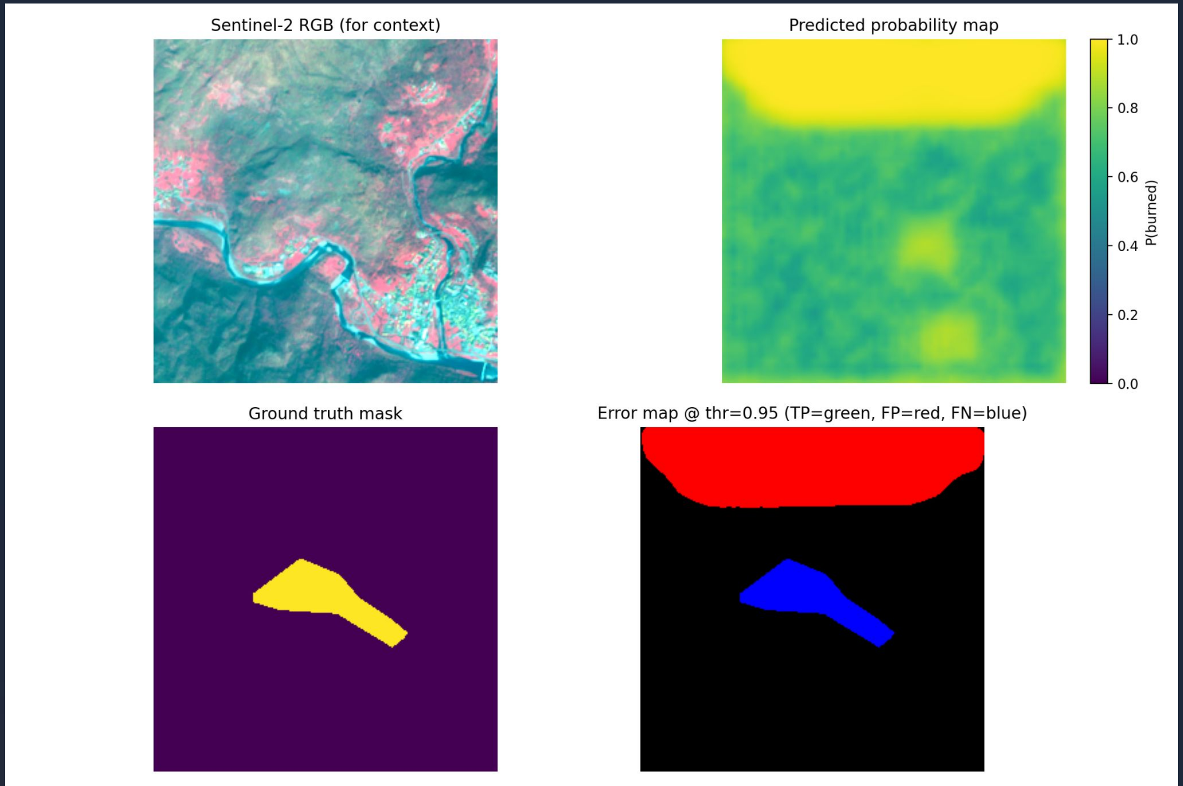
PRECISION–RECALL BEHAVIOUR

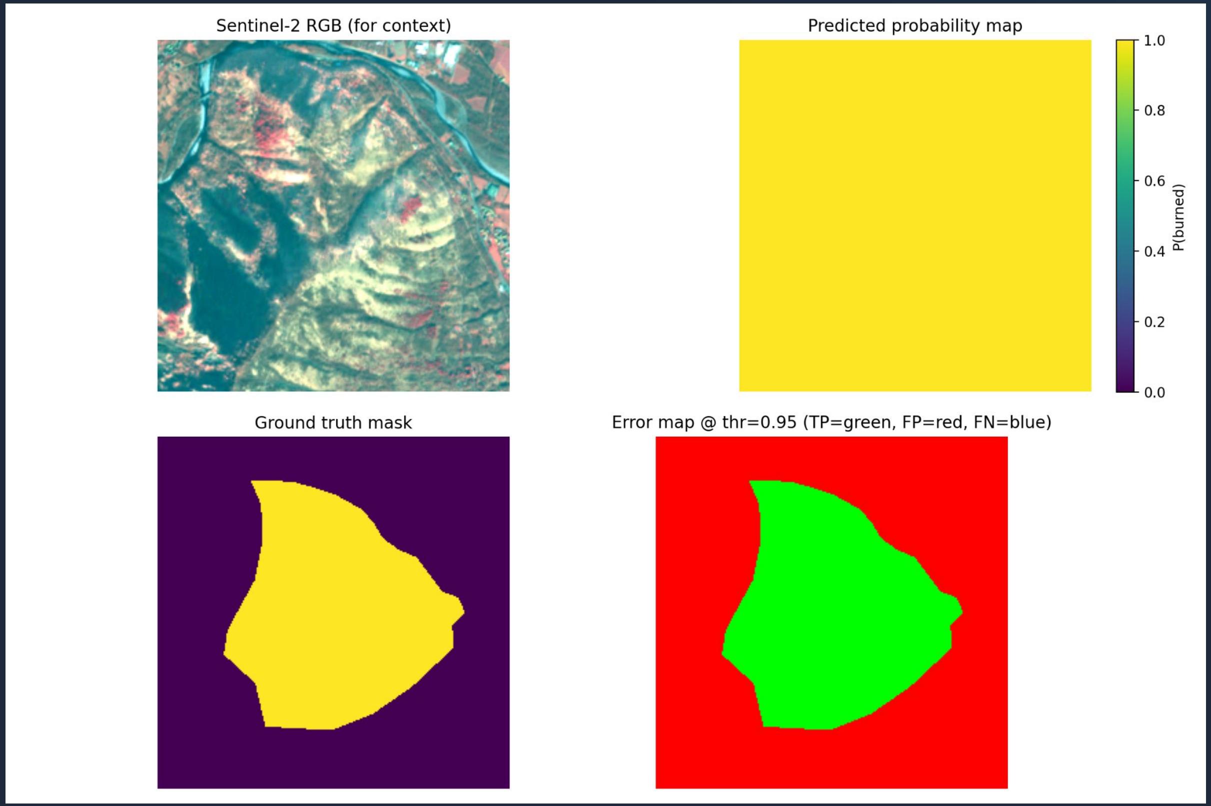
- **Low thresholds:** High recall but many false positives.
- **High thresholds:** Conservative predictions (high precision).

Decision: Selected operating point (0.95) prioritizes spatial accuracy to minimize false alarms.



QUALITATIVE EVALUATION





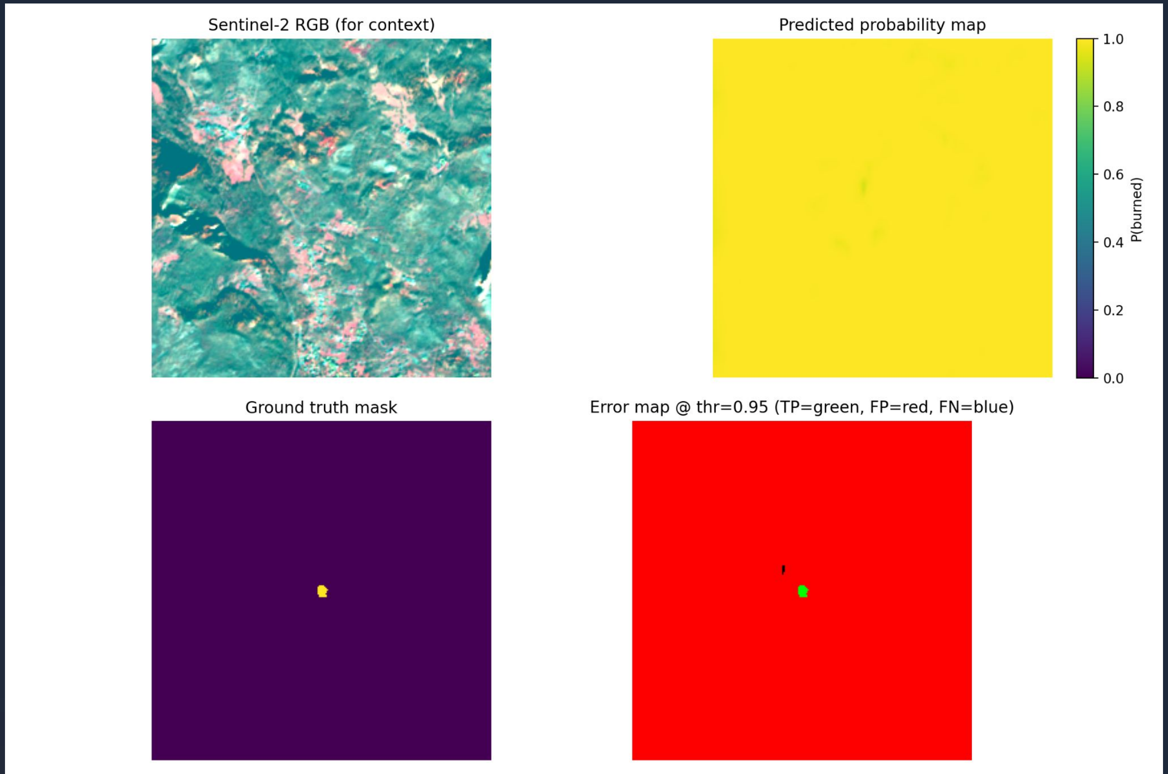
SMALL / DIFFICULT FIRE

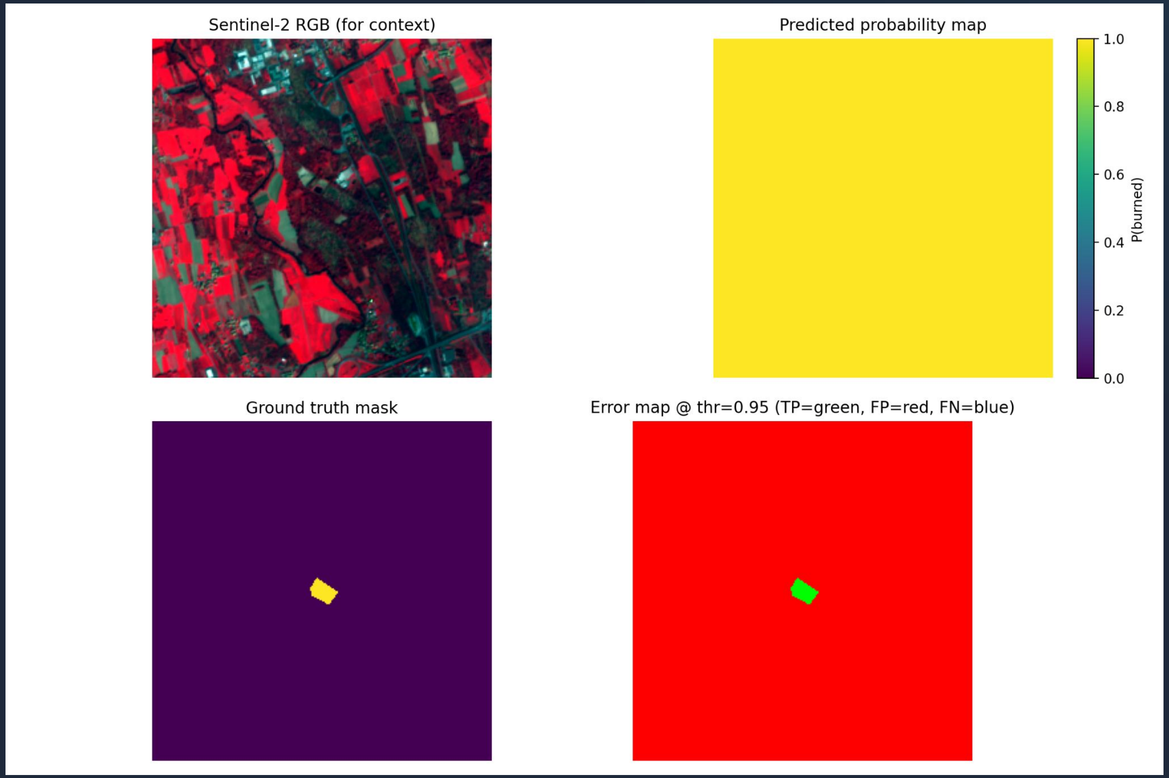
Model shows conservative behavior, avoiding noise.

LARGE / WELL-DETECTED FIRE

Large fires are detected with high spatial coherence.

QUALITATIVE EVALUATION





SMALL / DIFFICULT FIRE

Model shows conservative behavior, avoiding noise.

SMALL / DIFFICULT FIRE

small fires are detected with normal spatial coherence.

CONCLUSION

Signal Utility

Pre-fire multimodal data provides useful signals for predicting final burned areas.

Data Integration

Integrating multiple data sources improves segmentation performance compared to a baseline approach.

Encoder Architecture

Encoder design plays a key role: a unified high-capacity encoder achieved better IoU than alternative encoder strategies.

Design Balance

Careful architectural choices can improve accuracy while keeping the model pipeline simple and reproducible.

> "Overall, this project shows that **effective multimodal integration** combined with a **well-chosen encoder strategy** can significantly enhance burned-area prediction performance."

THANKS FOR YOUR ATTENTION!

ANY QUESTIONS?

Disclaimer: Our models are good, but they can't predict *your* questions.