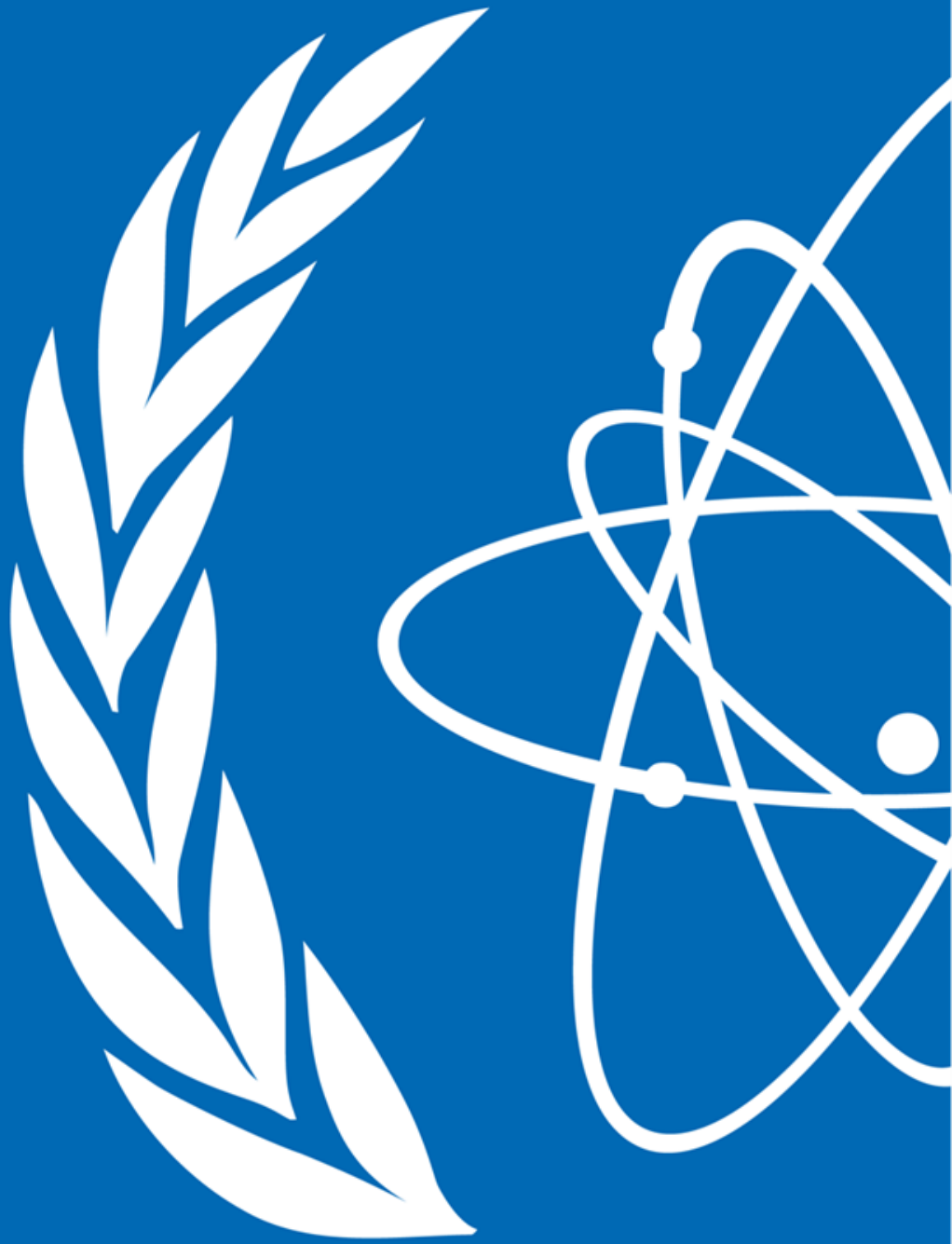
# AI for Mapping Buildings: Experience from IAEA

Matthew T Martinez
Image Scientist and Technical Expert
Division of Information Management
State Infrastructure Analysis



# The IAEA

- Created in 1957 in response to fears and expectations by the diverse use of nuclear technology
- The genesis was Eisenhower's "Atoms for Peace" address to the General Assembly of the United Nations on Dec 8, 1953.
- The US Ratification of Statute by President Eisenhower on July 29, 1957 marked the official birth of the IAEA.
- The IAEA objective is to promote and control the Atom:

"The Agency shall seek to accelerate and enlarge the contribution of atomic energy to peace, health and prosperity throughout the world. It shall ensure, so far as it is able, that assistance provided by it or at its request or under its supervision or control is not used in such a way as to further any military purpose."



# Safeguards are essential to international security

Safeguards are based on legally-binding international treaties and agreements (NPT, Safeguards agreements, Additional Protocol)

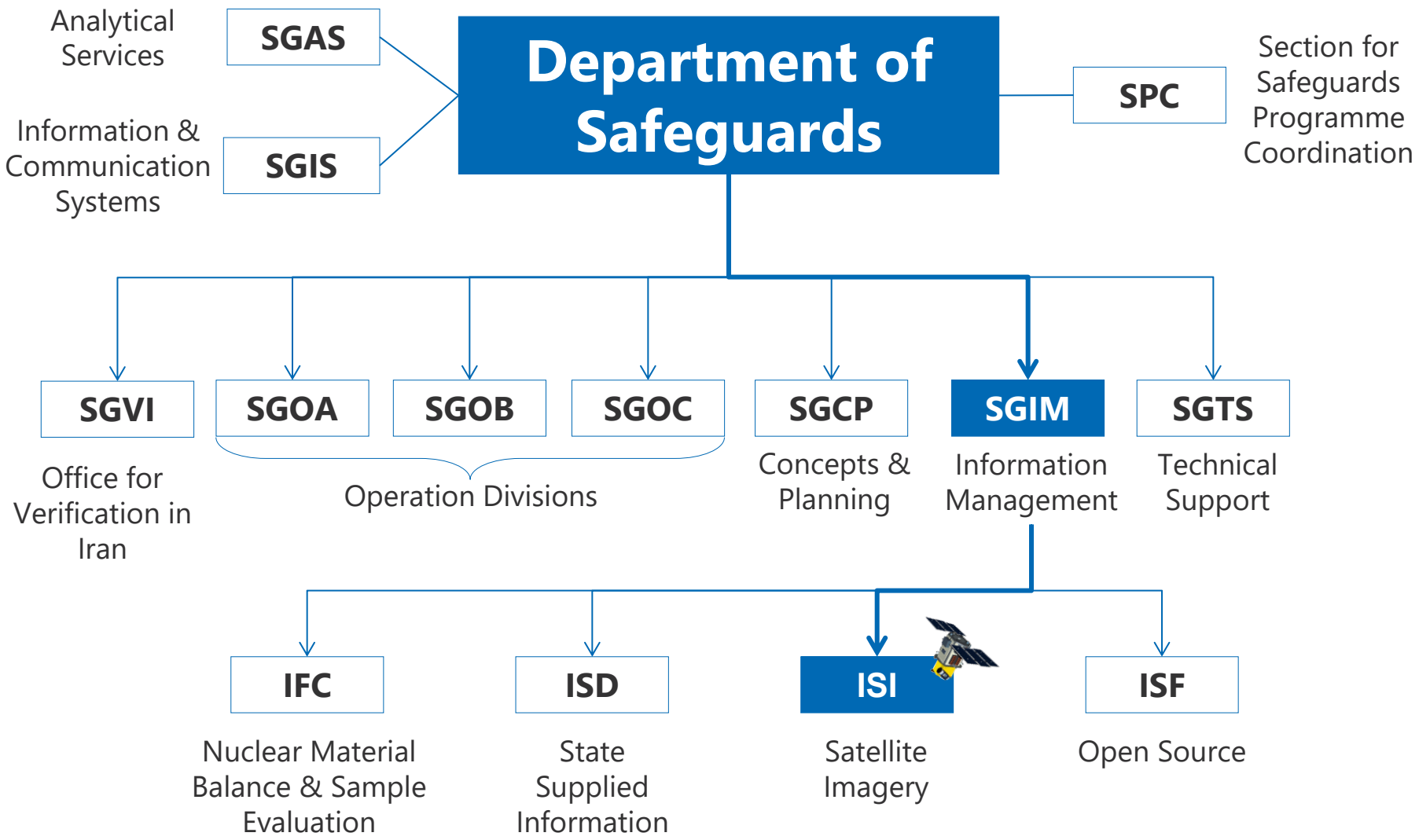
Diversion of nuclear material

misuse of material & facilities

undeclared nuclear material & activities



# IAEA – Department of Safeguards (SG)



# Division information Management

**collects, processes,** evaluates, **and analyzes**

**The IAEA collects and processes safeguards-relevant information about a State from three sources**

## ❶ State Declared Information

**Starting Point Essential**

- Facility design information; Nuclear material accountancy
- Voluntary reporting
- AP declarations

## ❷ Information From Field Activities

**In-Field Activities**

- Inspections
- Design information verification
- Complementary access

## ❸ Other Safeguards-relevant information

**Freely available through subscription**

- Open sources, incl.
  - S&T
  - Nuclear trade
  - Satellite imagery
- Third party

# VALIDATION

*correctness and completeness of declarations.*

# Use of Imagery for Safeguards Purposes

## Acquiring, processing, analyzing, exploiting satellite imagery and geospatial information to support nuclear verification activities.



Identify and Monitor Nuclear Fuel Cycle Sites & Activities



Verify States' Declarations



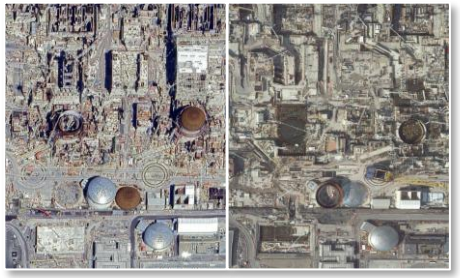
Support In-Field Activities



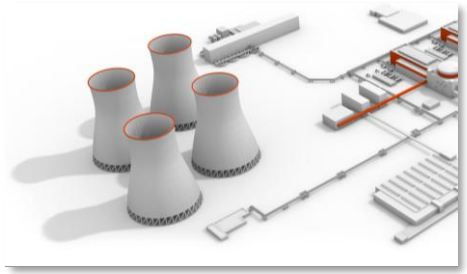
Investigate Alleged Undeclared Activities



Mineral exploration



Construction
Dismantlement



Gain insight into infrastructure layout



Monitor remote, difficult or denied sites of high interest
Search large geographic areas

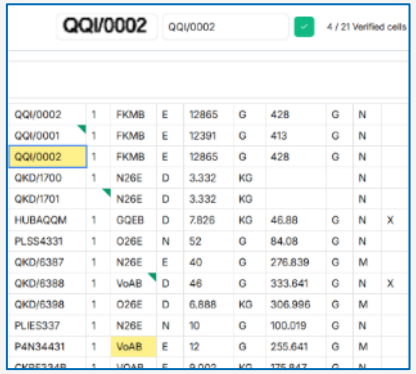


# SGIM Innovation Team

# AI in Safeguards Information Analysis

AI support for **efficient processing and validation** of safeguards-relevant information - leveraging **subject matter expertise** and **innovation** to support departmental strategic objectives

## State-declared information

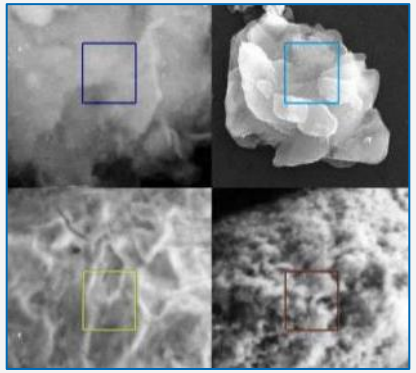



Significant reduction in time spent on manual data entry tasks



Extracting building footprints from site maps supporting AP review and consistency checks, as well as the **GEO** project

## Information from verification activities



Morphological comparison tool to support Environmental Sample Analysts

## Other SG-relevant information




Integrated in OSIS/OSIS-ST: prioritization, categorization, transcription, etc., supporting the **SG Lake** project



Detecting building footprints and changes from satellite imagery supporting analysts and consistency checks

# Digitizing Site Maps

## Business challenge

- Site maps provided by States are highly diverse and contain **complex geospatial information**
- Current digitization is very **manual and time-consuming**

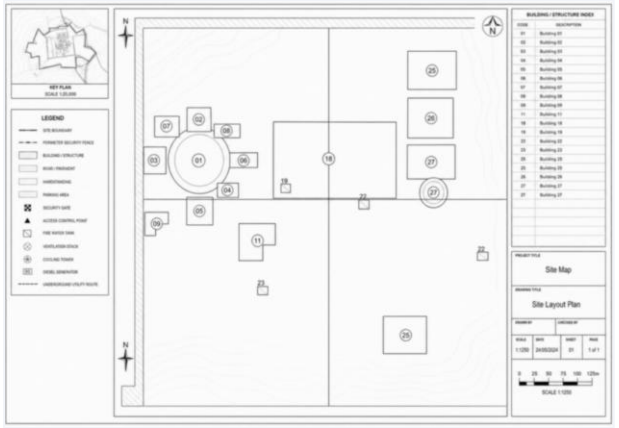
## Our approach

- **End-to-end AI pipeline** with several specialized models
- **Fine-tuning** on synthetic maps to address complex layouts

## Status

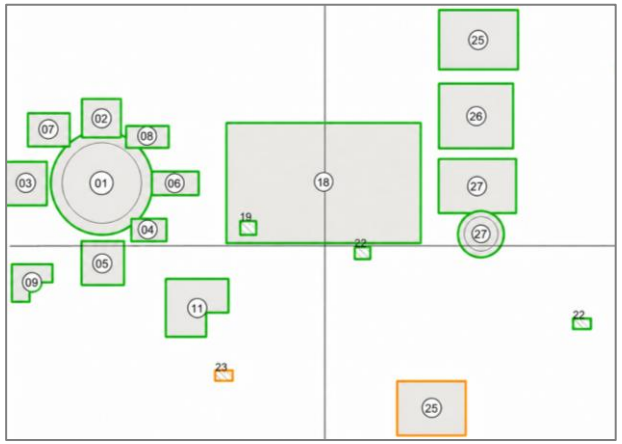
- Feasibility study completed, all **AI models** and **end-to-end pipeline** ready
- **Operational pilot** on selected States to define future large-scale implementation

## AI-Assisted Extraction and Review



### Input

**Site Map PDF**
Attachment in 2.a.(iii)



### Output

**Digitized site map**

Geo-referenced with highlighted changes

*Synthetic Map for Demonstration Purposes*

# Safeguards Infrastructure Recognition in Satellite Imagery

**Business need:** Extract building footprints, detect infrastructure changes, support consistency analysis

## Our approach

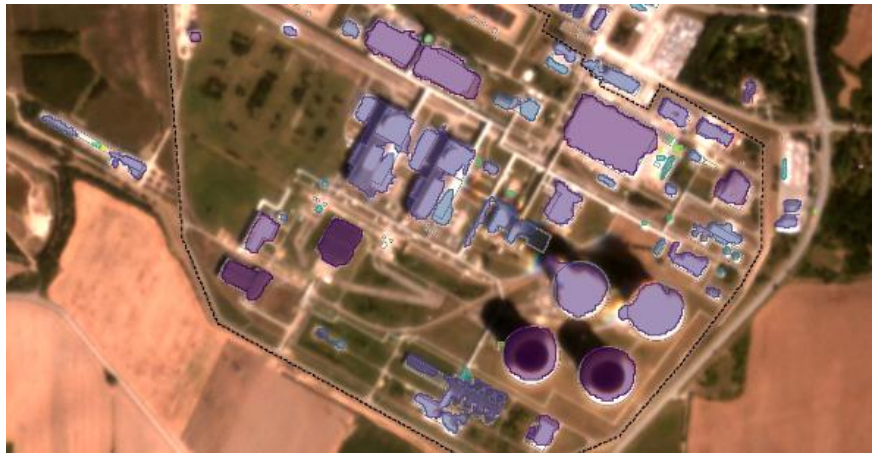
- Self-supervised pre-training on **high-resolution satellite imagery**
- **Fine-tuning** of models on **validated in-house datasets**

## Status

- Model available for detecting **nuclear-specific buildings**
- Model is being integrated to support use cases such as **change detection**, **prioritization**, **and consistency checks**

## Detection of nuclear-specific buildings





# $M^3$

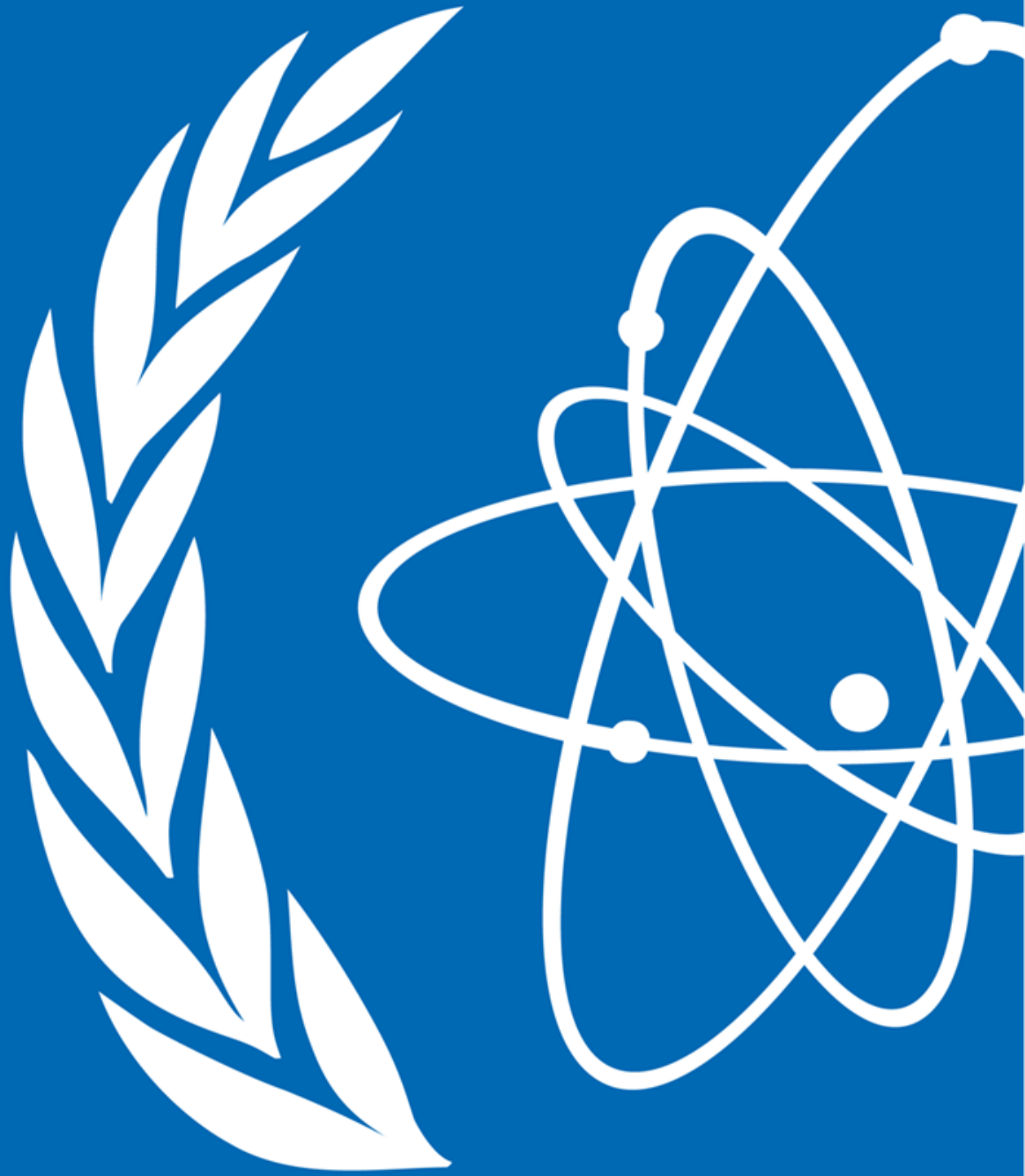
## A Multistage, Multitask, and Multispectral Framework for Building Footprint Segmentation in Industrial Zones

**Denis Groshev**

Technical University of Munich, Germany

**Matthew T. Martinez**

International Atomic Energy Agency, Vienna, Austria

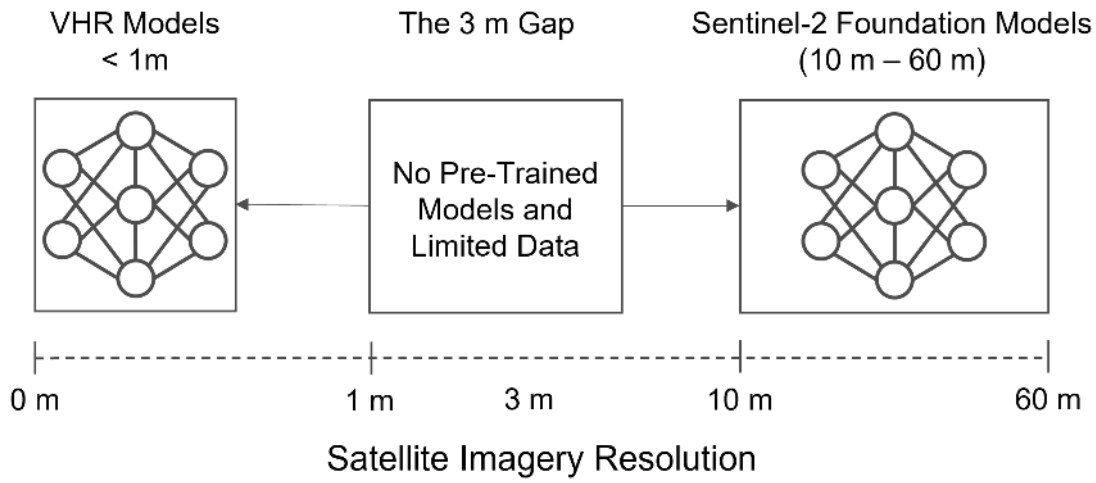


# Problem Statement

## The Resolution–Cost Trade-off

### 🔀 Transfer learning gap:

- Existing Sentinel-2  foundation models do not transfer well
- High-resolution models do not down scale effectively
- “Resolution gap” where 3 m is underserved




### 📊 Limited training data:

- Scarce availability of public data at 3m unlike Sentinel-2 (10m) or publicly available sub-meter.
- Requires augmentation strategies (down sampling of high-res, synthetic data, etc)

| | |
|---|---|
| Sub-meter: | 📸📸📸📸📸📸 |
| 3m: | 📸 |
| 10-60m: | 📸📸📸📸📸📸📸📸📸📸📸📸📸 |

Limited 3m Data 📉
→ Augmentation Strategies
→ Usable Training Data ✨

3 m resolution presents unique challenges that require tailored solutions for accurate models

# Our Approach

## M³: Three Ideas, One Framework



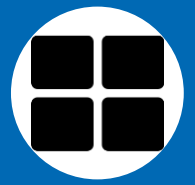
### Multistage

SSL pretraining on 1.8 m imagery → general building fine-tune at 3 m → industrial fine-tune at 3 m



### Multitask

Segmentation + edge detection + signed distance map heads provide dense, complementary geometric supervision



### Multispectral

Native 8-band input (vs. RGB-only priors) with zero-padding for legacy 4-band sensors

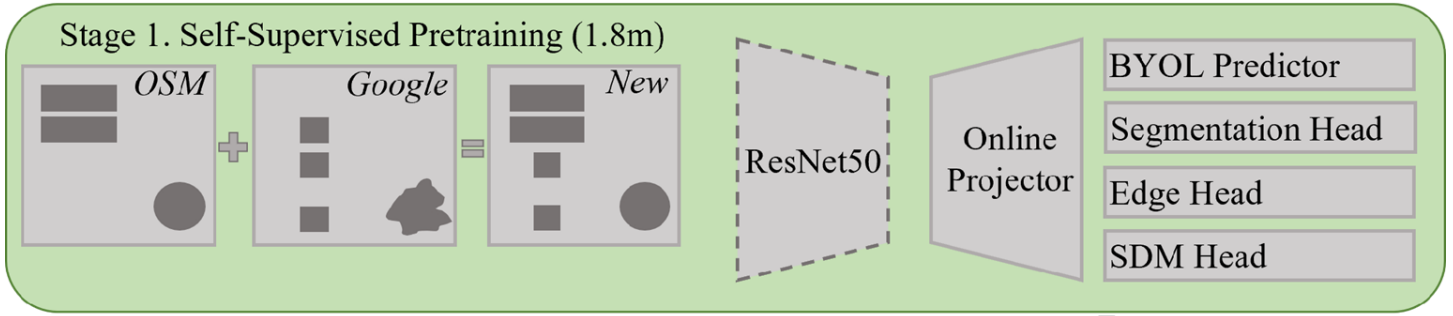
Open-source footprints are treated as geometric priors — not absolute ground truth.

# Methods

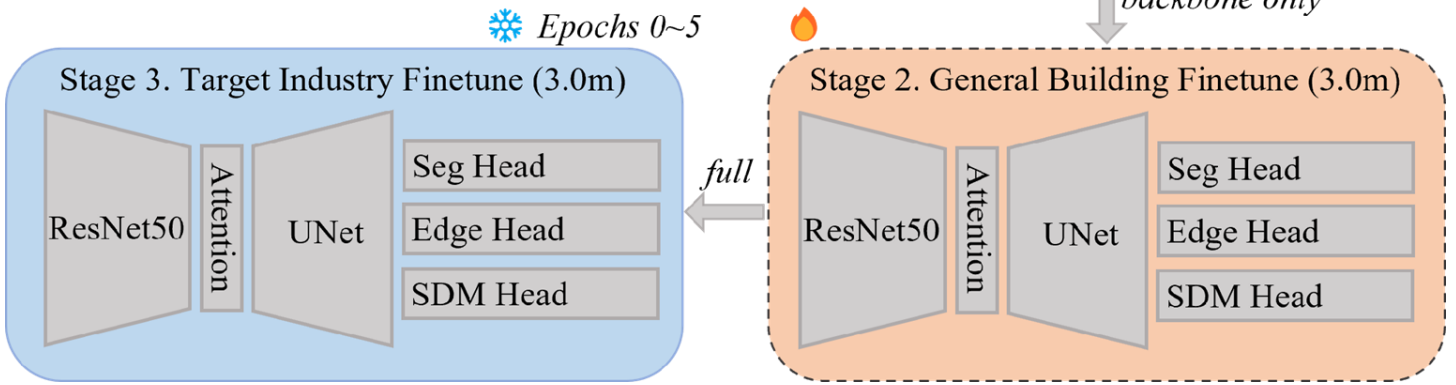
## Three-Stage Training Pipeline

📊 Multistage Training: Pretrain (1.8m) → General Buildings Finetune (3.0m) → Industrial Sites Finetune (3.0m)










*Imagery from Functional Map of the World

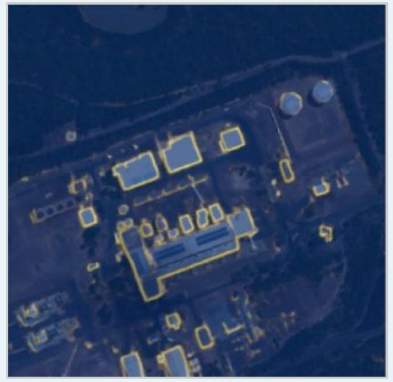
# Methods

## Geometric Priors as Auxiliary Supervision

### Edge — Morphological Gradient

$\text{Edge} = (M \oplus B) - (M \ominus B)$

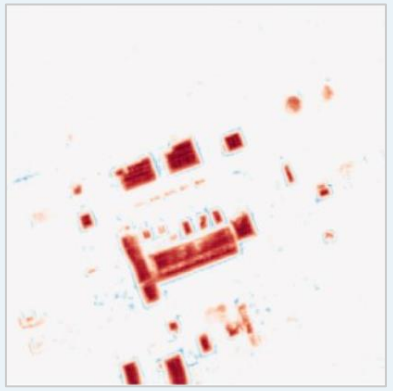
Dilation minus erosion of the binary mask M. Thickened boundaries reduce gradient sparsity from extreme class imbalance.



### Signed Distance Map (SDM)

Euclidean Distance Transform

Positive inside buildings, negative outside. A continuous signal that regularizes boundaries under noisy, shifted OSM supervision.



### Stage 1 — Pretraining Loss

$L = 0.65\, L_{BYOL} + 0.15\, L_{seg} + 0.10\, L_{edge} + 0.10\, L_{sdm}$

Asymmetric loss applies reduced penalty in unlabeled regions, avoiding punishment for valid structures missing from filtered labels.

### Stages 2 & 3 — Fine-tune Loss

$L_{ft} = 0.65\, L_{seg} + 0.20\, L_{edge} + 0.15\, L_{sdm}$

Weighting shifts toward segmentation once representation learning is complete — geometric heads now sharpen the final footprint.

# Methods

## Open-Source Label Fusion



OpenStreetMap (OSM)

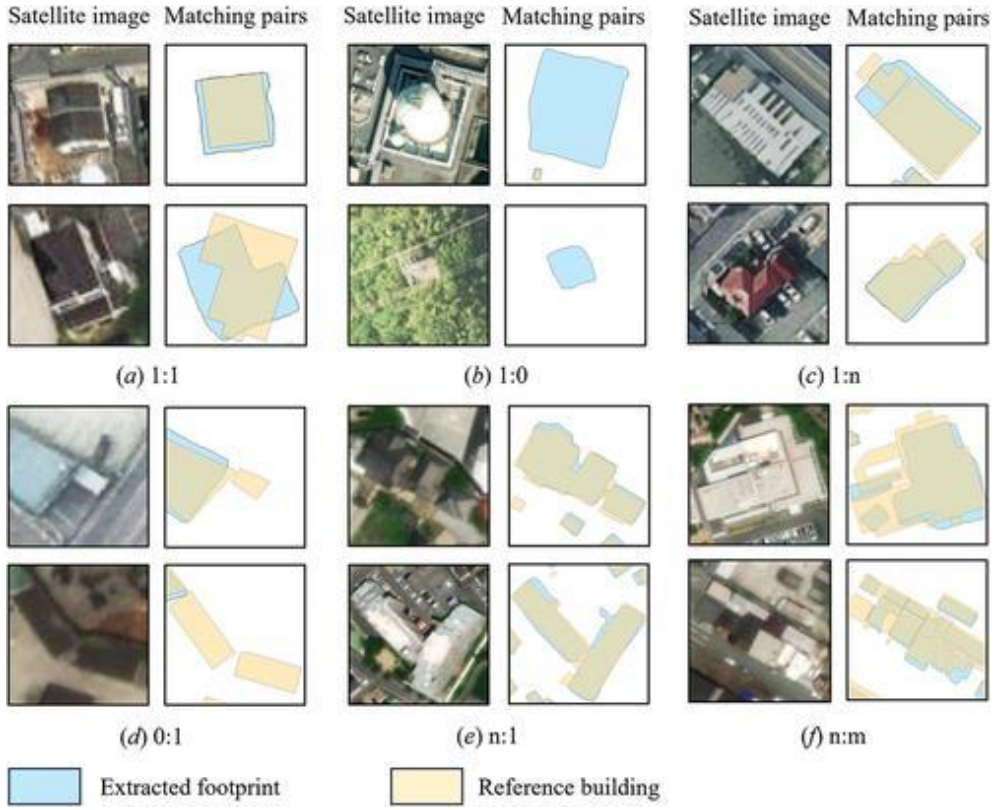


Google-Microsoft Open Buildings



Combined Labels

- Size Threshold - Structures under 6 m² discarded as sub-pixel artifacts
- Spatial conflict - Candidates with centroid inside an OSM polygon, or > 70% overlap, removed
- Redundancy suppression - Candidates geometrically similar to nearby OSM polygons suppressed in local clusters (≤ 15 buildings)



# Results

## Multistage + Multitask Ablation

| Pretrain | Adaptation | Fine-tune Weighting | IoU | F1 | SDM MAE |
|---|---|---|---|---|---|
| BYOL only | Pretrain → Target | Seg only | 0.542 | 0.551 | – |
| BYOL+Seg | Pretrain → Target | Seg only | 0.579 | 0.607 | – |
| BYOL+Seg+E+S | Pretrain → Target | 0.8 Seg, 0.2 Edge | 0.605 | 0.629 | – |
| BYOL+Seg+E+S | Pretrain → Target | 0.65/0.20/0.15 | 0.642 | 0.664 | 0.175 |
| BYOL+Seg+E+S | Pretrain → General 3m → Target | Seg only | 0.611 | 0.619 | – |
| BYOL+Seg+E+S | Pretrain → General 3m → Target | 0.8 Seg, 0.2 Edge | 0.658 | 0.672 | – |
| **BYOL+Seg+E+S** | **Pretrain → General 3m → Target** | **0.65/0.20/0.15** | **0.721** | **0.733** | **0.113** |

**0.642 → 0.721**

IoU gain from the intermediate 3 m general building adaptation stage

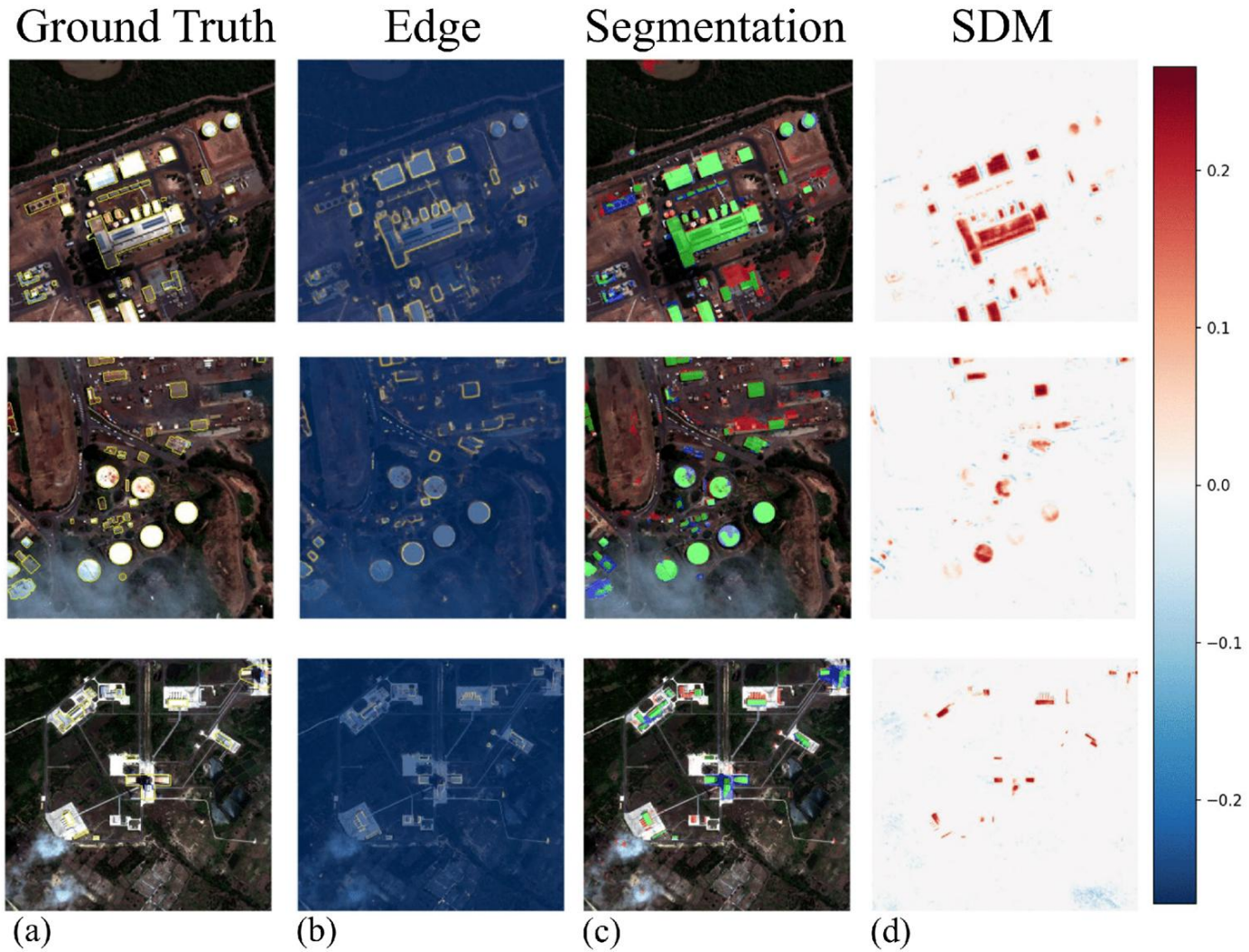
**0.175 → 0.113**

SDM MAE drop when the full multitask weighting is used

**F1 = 0.733**

Best backbone: auxiliary tasks during pretraining, after full adaptation

# Qualitative Segmentation Results




**Green** = correct **Red** = incorrect **Blue** = missing from label

Results:

- The model recovers buildings absent from the ground-truth labels (shown as “false positives”)
- Provides a viable path toward automated label refinement and anomaly detection, not just segmentation.

*Imagery from Functional Map of the World

# Conclusion & Future Work

## M³ Reduces Reliance on Expensive Labels

- Multistage adaptation + multitask geometric priors + multispectral input jointly close the 3m domain-resolution gap
- Treating OSM as a structural prior (not ground truth) makes training robust to systematic omission and misregistration
- Supplementing OSM with machine-derived footprints (Google/Microsoft) yields a 5% performance improvement
- Future work: student-teacher pseudolabeling to extend coverage into regions with sparse open-source data

**+5%**

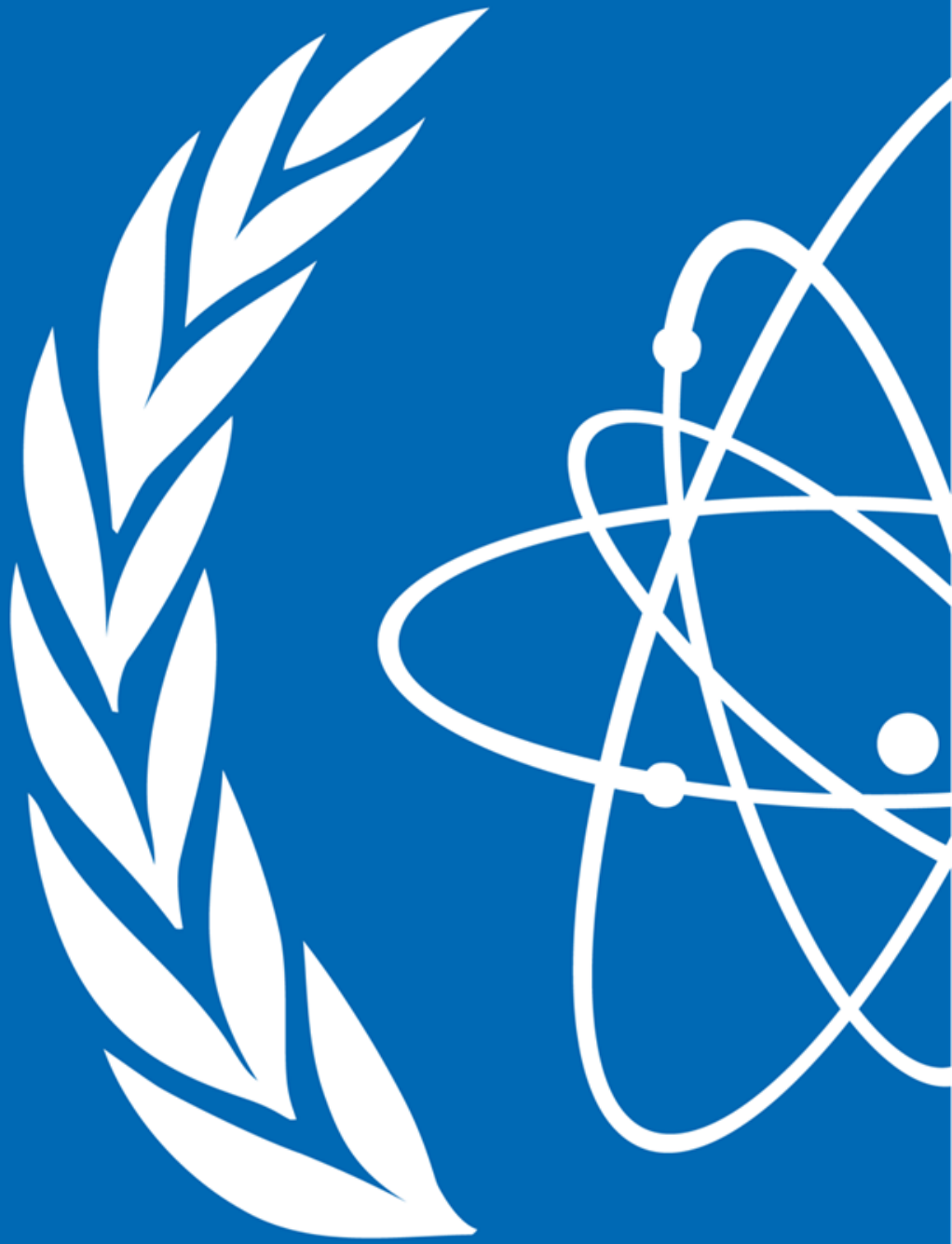
segmentation performance from fusing OSM with ML-derived footprints

**0.721 IoU**

full M³ configuration on industrial buildings

# Robust Change Detection of Nuclear Infrastructure using Satellite Imagery for Large-Scale Safeguards Verification

Matthew T Martinez
Image Scientist and Technical Expert
Division of Information Management
State Infrastructure Analysis Section



# Introduction

What changed?



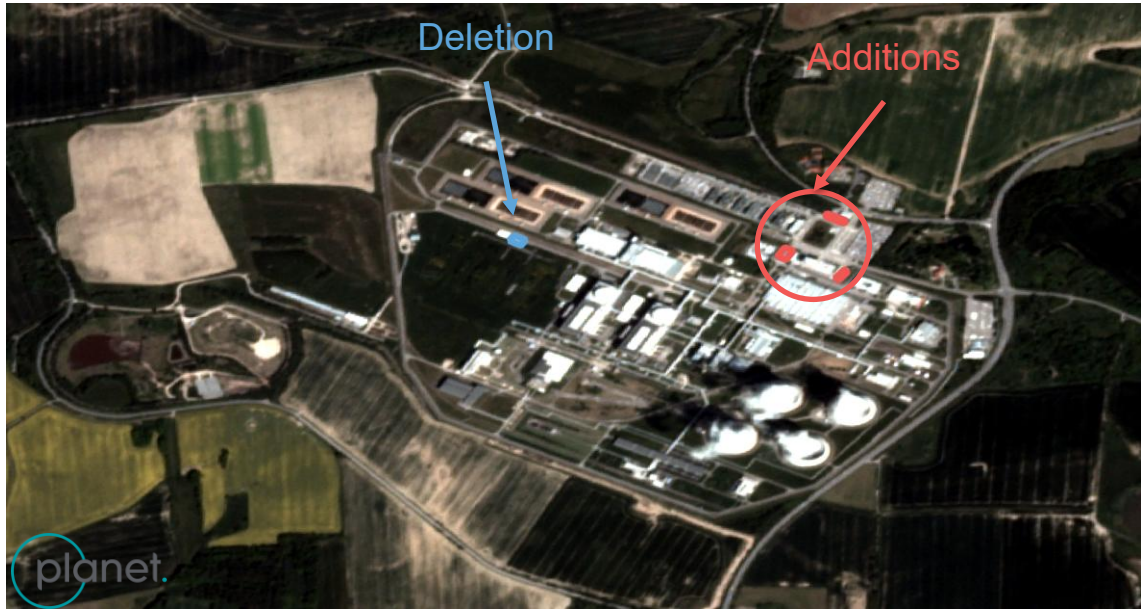

2021



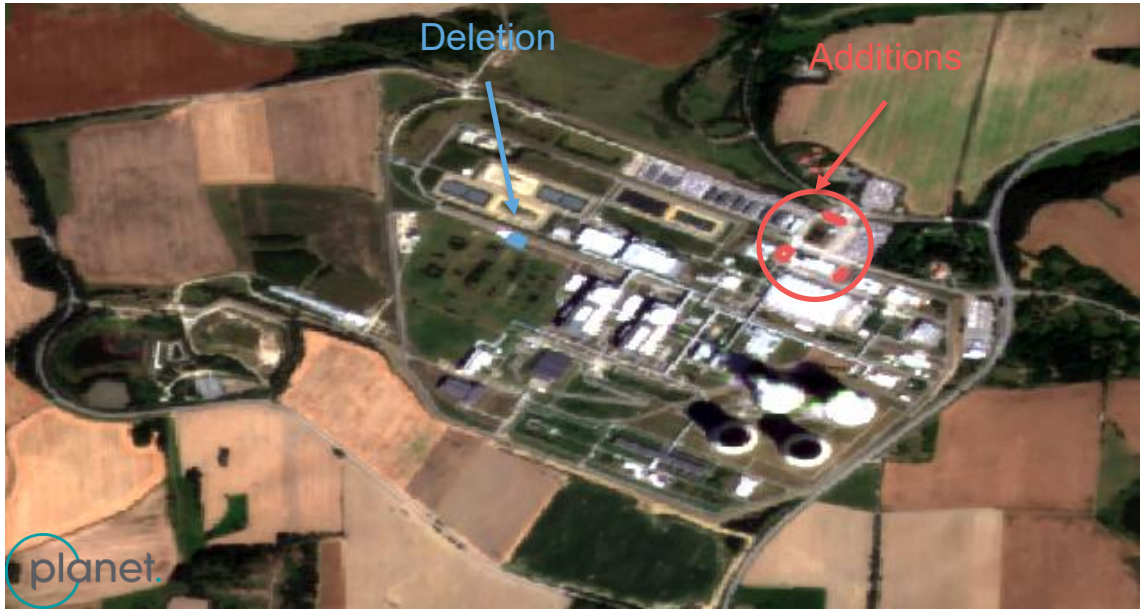

2022

# Introduction

What changed?




2021




2022

# Core Challenge

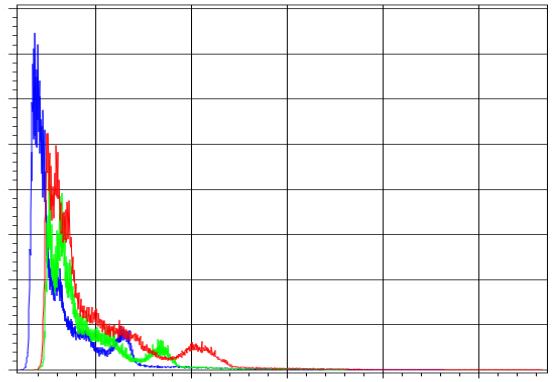
🌿 Interannual Phenological



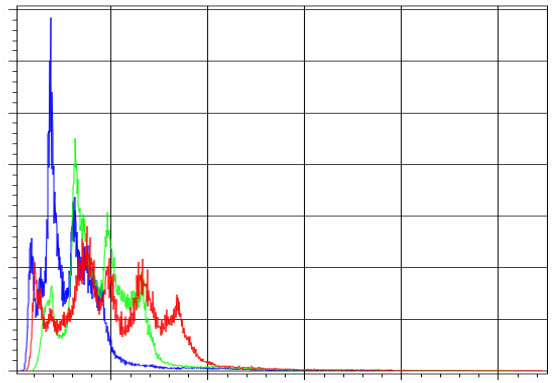




📷 / 📷 Sensor hardware generational differences



🌈 Spectral and band shift



# Solution

- 📐 **Proposed Solution:**
  - A sensor-agnostic, auditable change detection framework driven by geometric structural priors




🔄 Test Time Augmentations

Building Footprint Model

Ensemble Average

$P_{footprint}$

$P_{edge}$

SDM

Geometric Priors

$P_{footprint}$ Building footprint probability

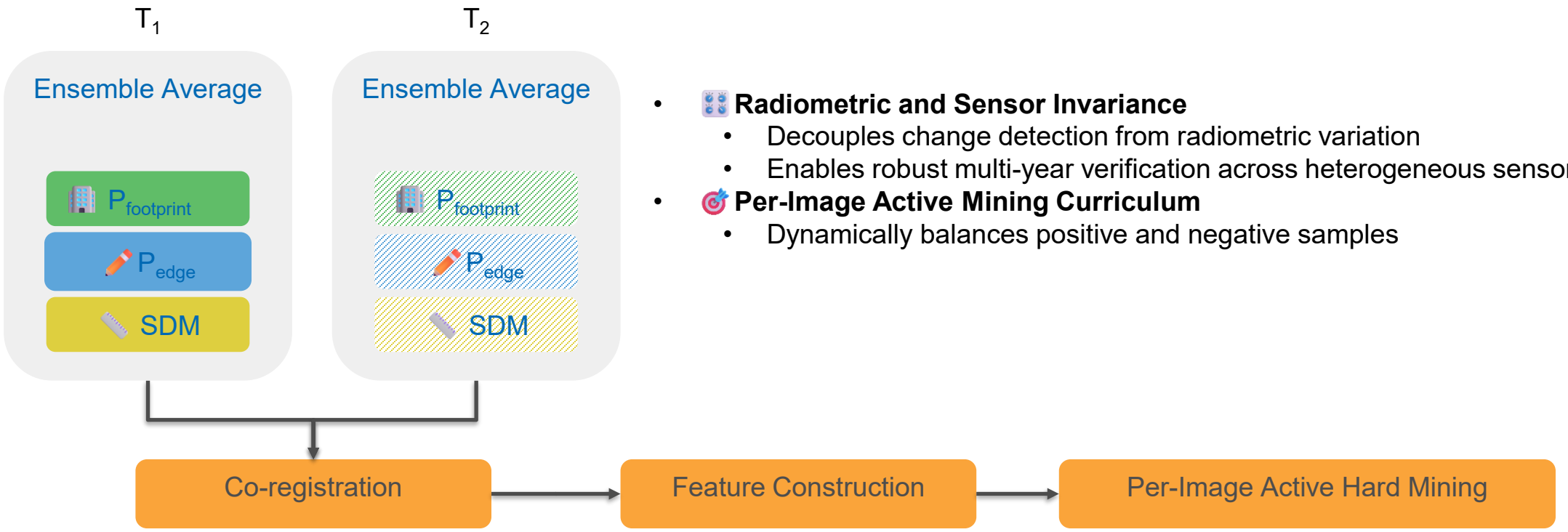
$P_{edge}$ Building edge probability

SDM Signed distance map ➔ orthogonal distance from every pixel to the nearest building boundary

# Solution

## 📐 Proposed Solution:

A sensor-agnostic, auditable change detection framework driven by geometric structural priors




- 🎛️ **Radiometric and Sensor Invariance**
  - Decouples change detection from radiometric variation
  - Enables robust multi-year verification across heterogeneous sensors
- 🎯 **Per-Image Active Mining Curriculum**
  - Dynamically balances positive and negative samples

# Infrastructure Changes Data

**5**
Consecutive Years
(2018 – 2022 )

**451**
Filtered
PlanetScope
Scenes

**52**
IAEA-Analyst
Annotated Scenes

**427**
Verified Changes

- Acquisition Filtering Constraints
- Seasonal Alignment: Jun – Sep (North) / Jan – Apr (South)
- Sun Geometry: Minimizes solar zenith & shadow distortions
- Atmospheric Screening: Excludes clouds, haze, and shadow artifacts

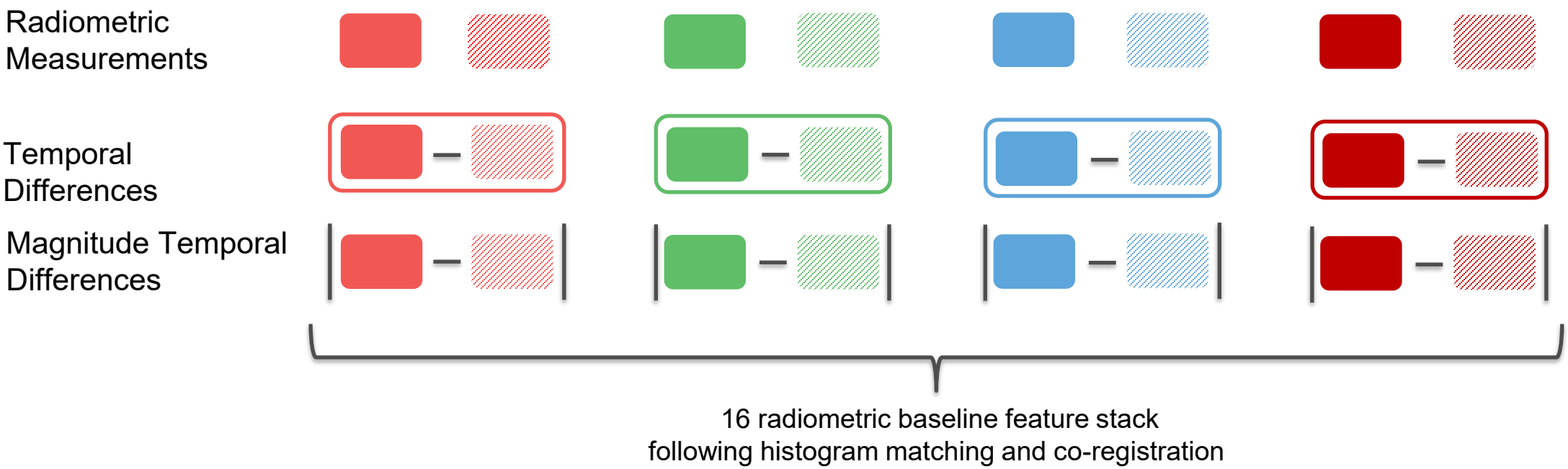
## Annotated Event Distribution

**+ 237**
Structural additions

**− 190**
Structural additions

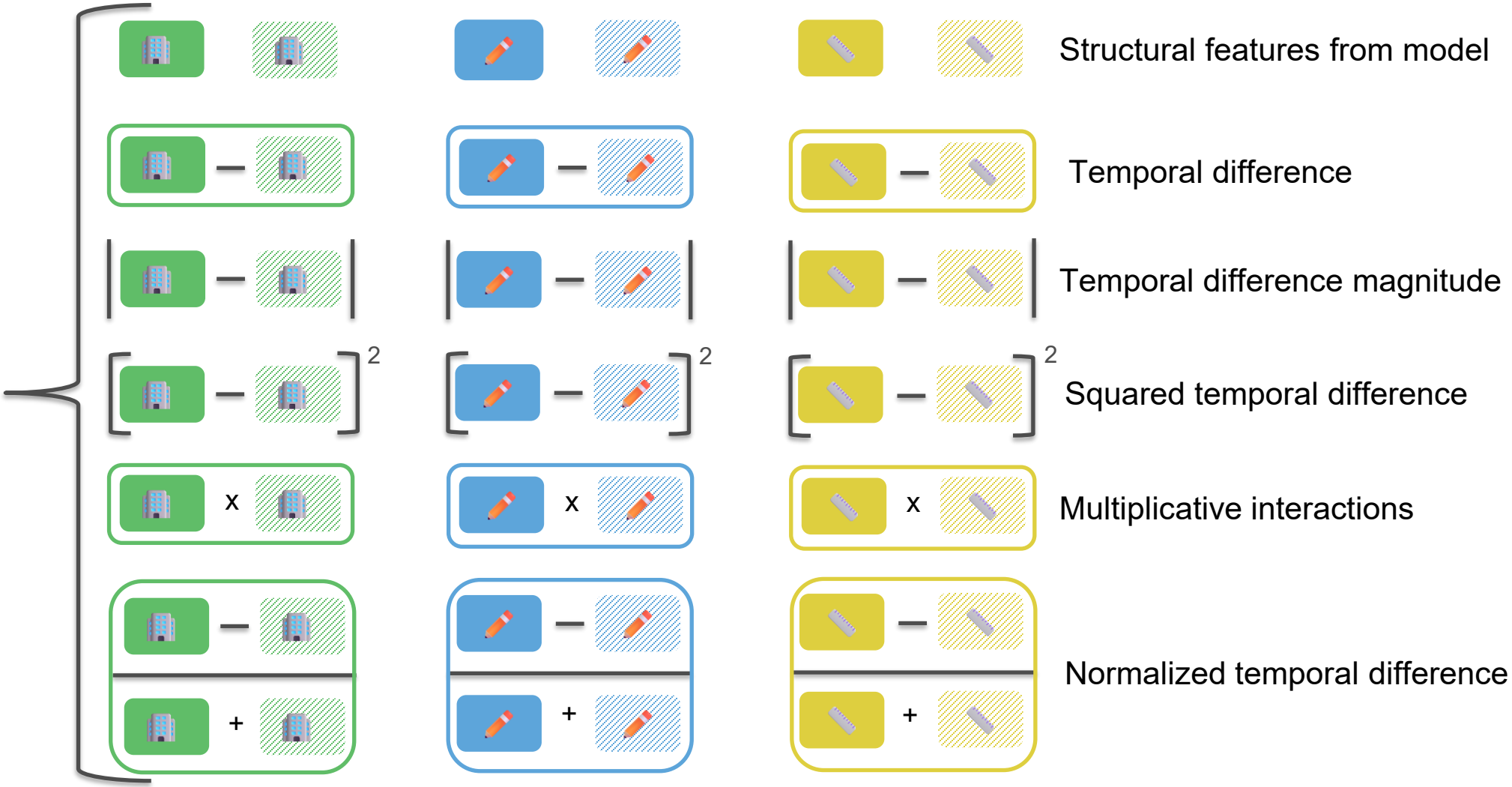
Changes represent < 0.5% total area
➔ extreme class imbalance

# Infrastructure Change Detection
# Part 1: Feature Extraction



# Infrastructure Change Detection
# Part 1: Feature Extraction

21 Structural Feature Stack Designed to isolate physical changes from environmental noise



# Infrastructure Change Detection
# Part 2: Feature Extraction

🔀 60/40 Hybrid Sampling Ratio
60% misclassification margin | 40% unfirm background sampling

📉 Temperature – Scaled Inverse Margin
Prioritizes ambiguous boundary pixels via scaled probability distribution

🔁 Persistent Selection History
Tracks pixel utilization across epochs to refine decision boundaries iteratively

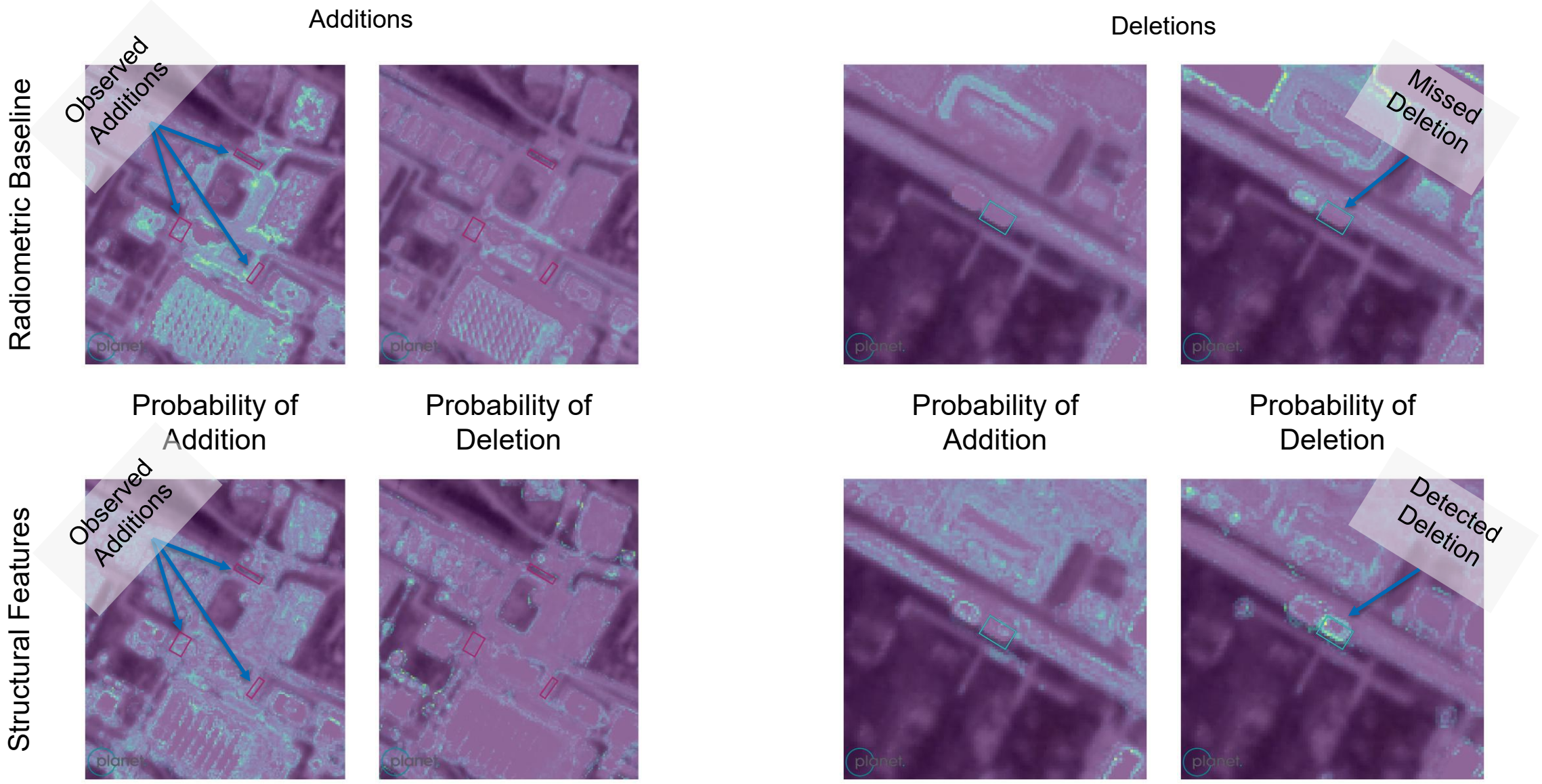
Per-Scene Budget

5,000

Negative pixels / Training Scene

Guarantees strict computational efficiency during per-scene active mining iterations.

# Results


Additions
Deletions
Radiometric Baseline
Observed Additions
Missed Deletion
Probability of Addition
Probability of Deletion
Probability of Addition
Probability of Deletion
Structural Features
Observed Additions
Detected Deletion

# Conclusion

### 🌐 Explainable & Auditable Framework
Maps raw imagery into explicit geometric structural features

### 🛡️ Environmental & Sensor Invariance
Abstracting scenes into structural representations decouple analysis from radiometric noise ensures compatibility across evolving hardware

### ⚖️ 16 Feature Radiometric Baseline
Effectively captures spectral shifts for building additions but struggles with complex structural deletions

### ⚖️ 21 Full Structural Stack
Provides complete spatial context to isolate and localize both additions and subtle deletion events

### 📈 Active Hard Negative Mining
Employs a 60/40 temperature-scaled margin distribution to continuously prioritize hard negatives near decision boundaries, directly solving extreme class imbalance.

### 🎯 Operational Impact
Replaces unfiltered pixel differences with confidence-ranked change alerts, delivering a prioritized, auditable review queue for image analysts.