Comparative Analysis of Multi-Source Satellite Imagery in North Jakarta, Indonesia

Tanish Verma tanish_v@mfs.iitr.ac.in
Sargam Goyal sargam_g@mfs.iitr.ac.in
Priyani Nagle priyani_n@mfs.iitr.ac.in
Shradhdha Agrawal agrawal_ss@mfs.iitr.ac.in

Comparative Analysis of Multi-Source Satellite Imagery in North Jakarta, Indonesia

This project conducts a detailed comparative study of optical and radar satellite imagery over a 5x5 km area in North Jakarta, Indonesia. Using PlanetScope high-resolution optical data and Sentinel-1 SAR radar data, we analyze spectral characteristics, perform edge detection, and apply deep learning for object detection. The study highlights the strengths and limitations of each data source in capturing urban, vegetation, and water features.

Data Acquisition and Preprocessing

PlanetScope Scene

Acquired a 4-band ortho_analytic_4b image dated December 20, 2024, with 5–10% cloud cover. Digital Number (DN) values were converted to reflectance using metadata scale factors via Dockerized tools.

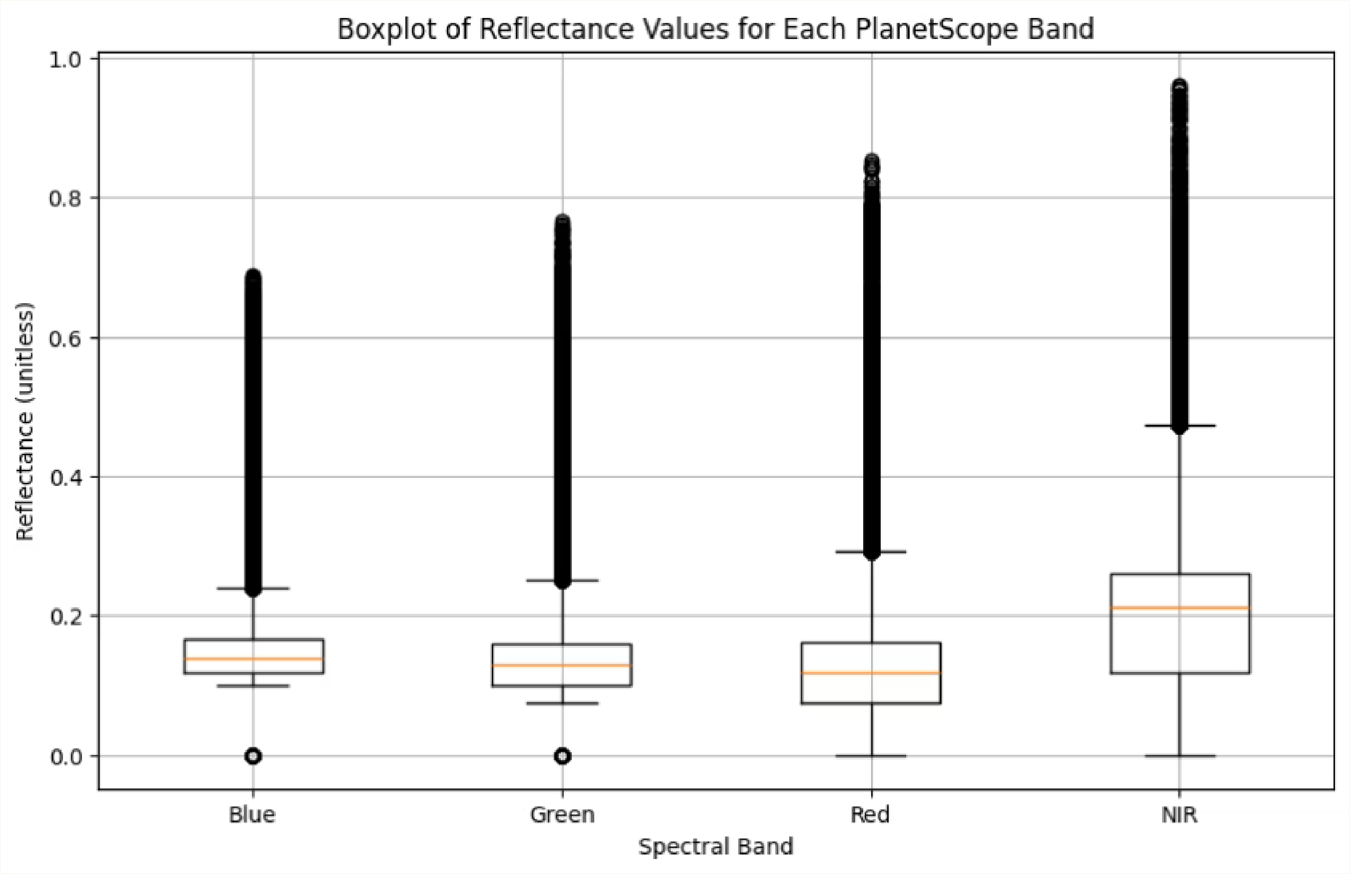
This ensured radiometric standardization for meaningful spectral spectral analysis.

Sentinel-1 SAR Image

Obtained Interferometric Wide (IW) GRD data from ASF dated dated December 18, 2024. Preprocessing in SNAP included orbit included orbit file application, thermal noise removal, radiometric calibration to sigma0, speckle filtering, terrain terrain correction, and conversion to dB scale.

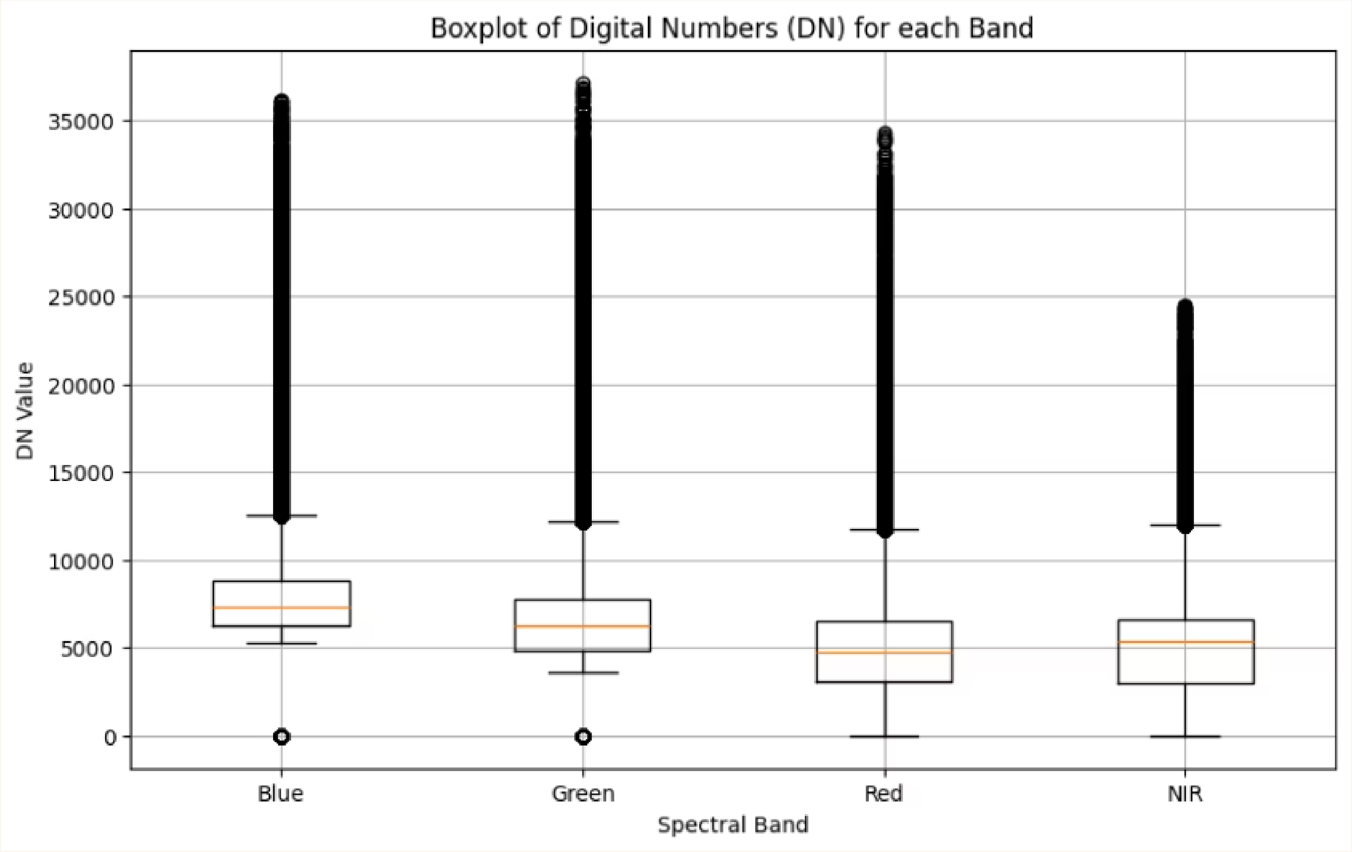
The final GeoTIFF was cropped to match the AOI for analysis. analysis.

Statistical Image Analysis: Boxplots of PlanetScope Data



Reflectance v/s Spectral Band

DN values



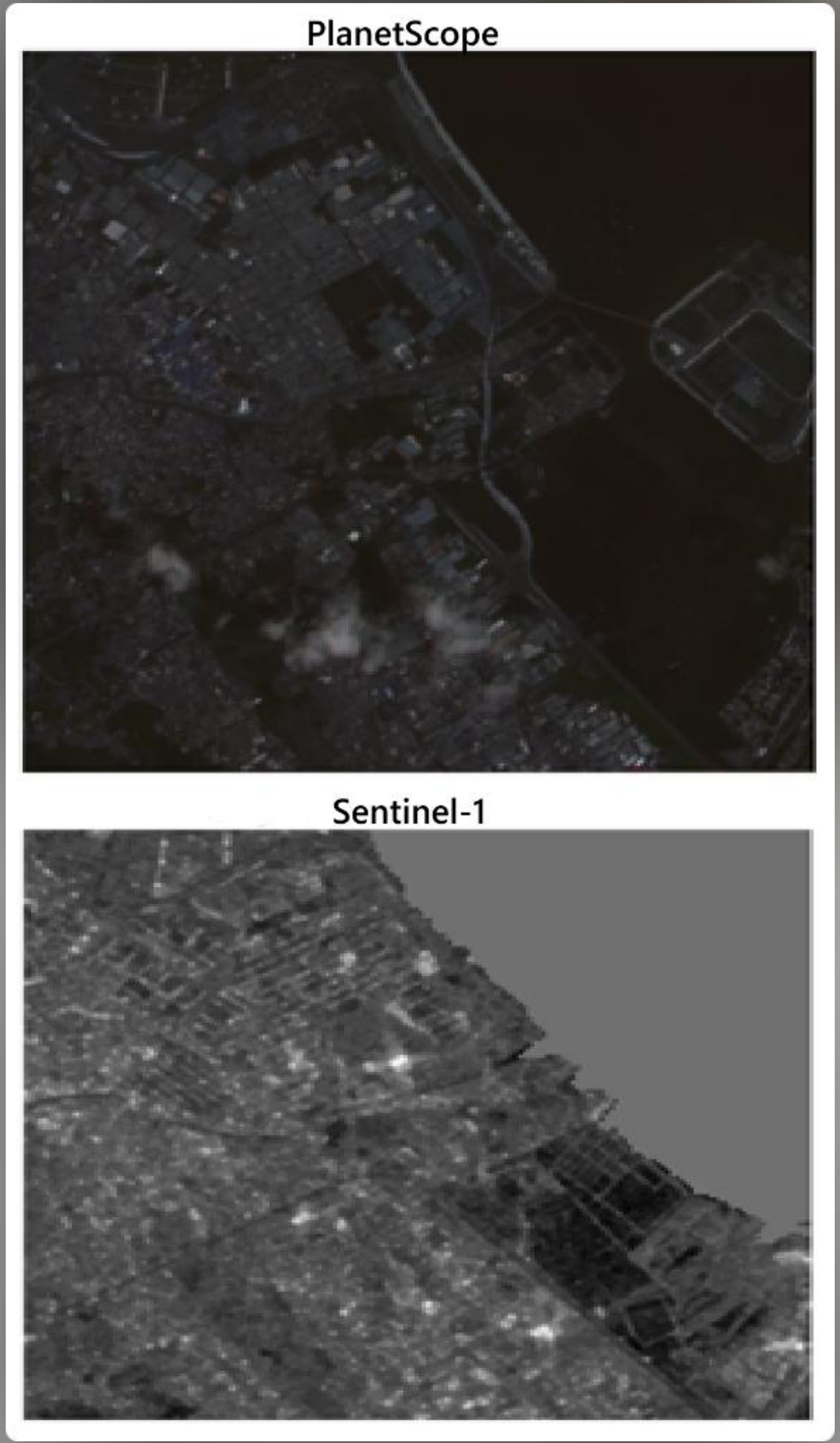
DN Values v/s Spectral Band

DN values show higher medians in Red and NIR bands, reflecting vegetation and urban features. Reflectance conversion normalized values (0–1), highlighting urban and vegetation distinctions.

Statistical Image Analysis: Boxplots of Sentinel-1 Data

Sentinel-1 Boxplots

Sigma0_VV band had the broadest dynamic range with strong backscatter from urban areas. VH band showed lower medians. Log-scaled VV_dB compressed extremes, aiding subtle surface differentiation.



Comparative Observations: PlanetScope vs Sentinel-1 Imagery

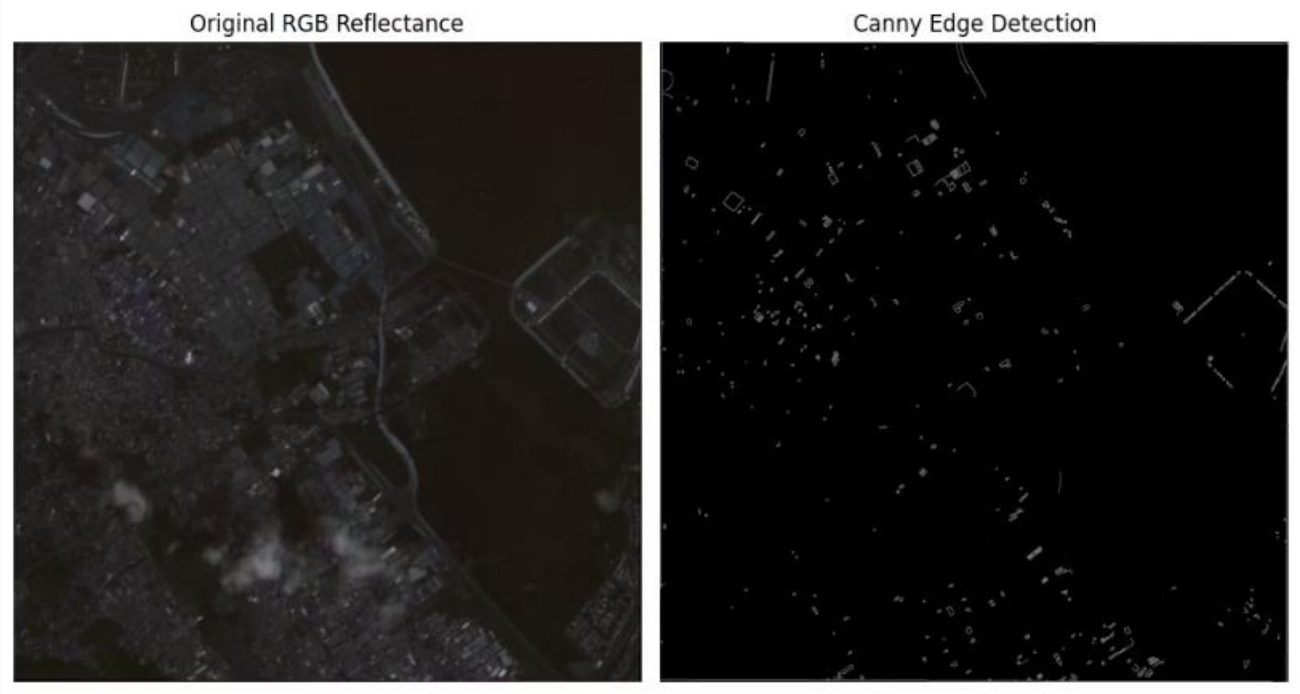
Feature Type	PlanetScope (Optical)	Sentinel-1 (Radar)
Urban Areas	High spatial clarity, visible structures	Strong backscatter, lower spatial detail
Vegetation	Strong NIR response, detailed detailed spectral info	Moderate textural differentiation
Water Bodies	Easily isolatable via low reflectance	Very dark due to low radar return
Overall Detail	High spectral and spatial resolution	Moderate, primarily textural
Weather Impact	Cloud-sensitive	All-weather capable

PlanetScope excels in visual detail and spectral richness, while Sentinel-1 provides all-weather imaging and structural insights, making their integration valuable.

Feature Detection Using Canny Edge Extraction over:

I. PlanetScope

- Detected edges were **sharp and continuous**, especially around:
  - Rooftops and building perimeters
  - Roadways and field boundaries
  - Coastlines and water-land interfaces
- The high **spatial and spectral resolution** enabled detection of fine details and small-scale features.

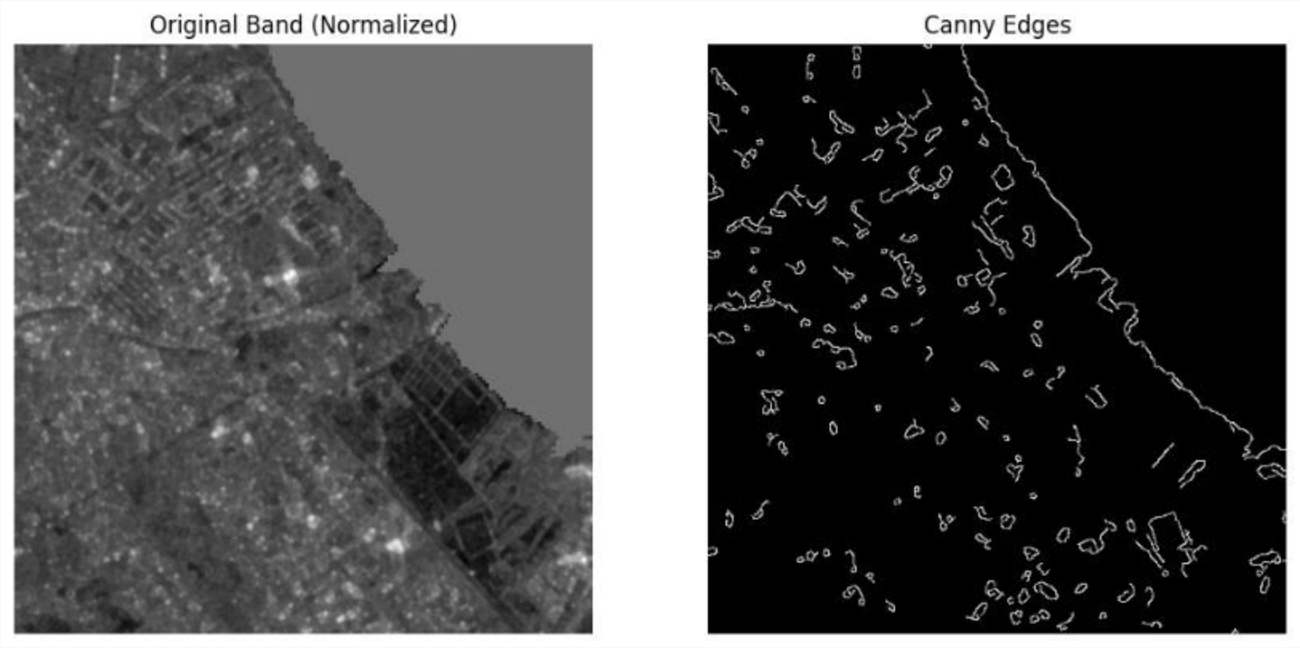


Feature Detection Using Canny Edge Extraction over: over:

II. Sentinel-I Imagery

- Edge detection highlighted key structures, though results results were influenced by the **inherent radar texture**. **texture**
- Clear edges were found in:
  - **Urban areas** with strong radar returns (e.g., rooftops, ports)
  - **Coastal boundaries**, where intensity shifted rapidly

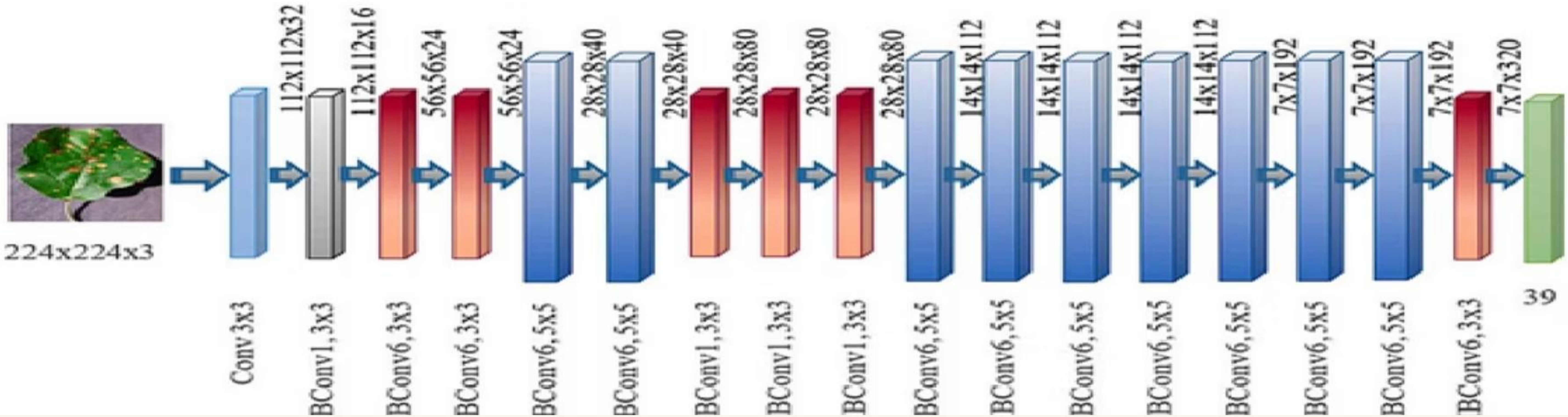
- Finer elements like small roads or vegetation boundaries were less distinct compared to PlanetScope.



Summary

Image Source	Feature Detection Quality	Notable Detected Features
PlanetScope	High (precise and dense)	Buildings, roads, coastline, vegetation
Sentinel-1	Moderate (textural edges)	Urban blocks, coastlines, large structures

The application of Canny edge detection proved effective for structural interpretation. In SAR imagery, the method particularly enhanced boundary visibility in areas with **strong radar contrast**, validating its utility for tasks like **urban segmentation** or **change detection** in future workflows



EfficientNet: Scalable CNN Models for Object Detection

1 Baseline and Scaling

EfficientNet-B0 is the baseline CNN with balanced depth, width, and resolution. Compound scaling uniformly increases these dimensions from B0 to B7 for improved accuracy and efficiency.

2 Model Variants

Models range from B0 (5.3M parameters) to B7 (66M parameters), with increasing input resolution and accuracy (~77% to ~84%).

3 Applications

Used in diverse domains including satellite satellite imagery analysis, medical imaging, imaging, and real-time systems, balancing balancing performance and computational computational cost.

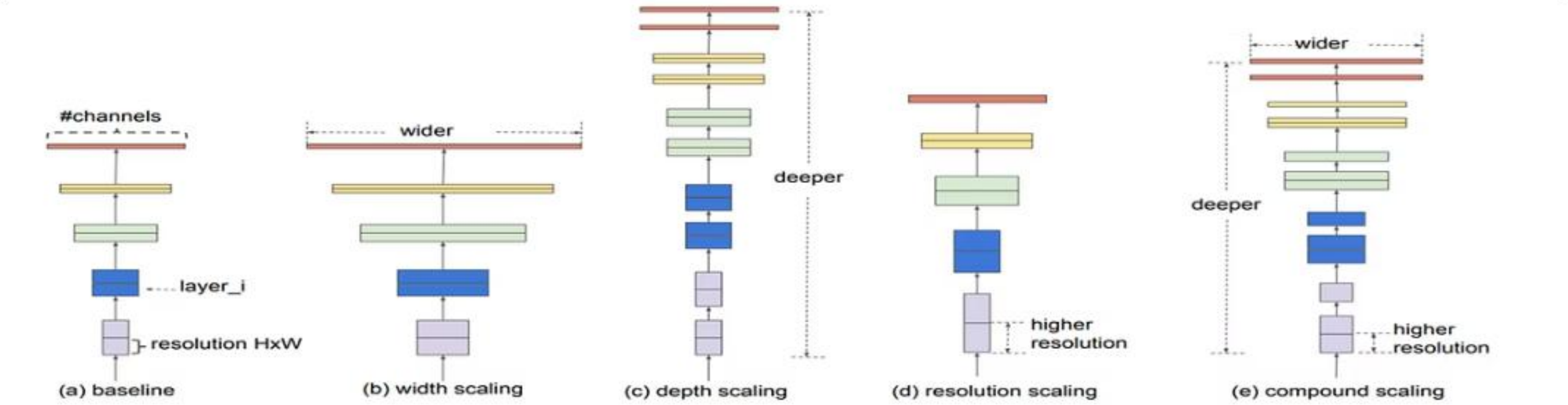


Figure 2. **Model Scaling.** (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

Key Features of EfficientNet:

I. EfficientNet Architecture:

- Uses **neural architecture search (NAS)** to design a mobile-sized baseline model (EfficientNet-B0).
- Scales B0 to B7 using the compound method, achieving state-of-the-art accuracy with fewer parameters and FLOPs.
- Key building block: **MBConv** (mobile inverted bottleneck) with squeeze-and-excitation optimization.

II. Compound Scaling Method:

- Compound scaling was introduced to address the inefficiency of conventional single-dimension scaling (depth, width, or resolution alone), by **uniformly balancing all three dimensions** using a principled **dimensions** using a principled approach using a compound coefficient ϕ.
- Unlike traditional methods, , compound scaling **improved accuracy while reducing computational costs**, enabling models like EfficientNet to achieve state-of-the-art performance with fewer parameters and faster inference.

$$
\begin{aligned}
\text{depth: } & d = \alpha^{\phi} \\
\text{width: } & w = \beta^{\phi} \\
\text{resolution: } & r = \gamma^{\phi} \\
\text{s.t. } & \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\
& \alpha \geq 1, \beta \geq 1, \gamma \geq 1
\end{aligned} \qquad (3)
$$

Model Optimization

Although our EfficientNet-based model achieved a high validation accuracy of approximately **98%** without extensive tuning, we include this section to outline the standard optimization pipeline that would typically be followed in deep learning projects. This demonstrates methodological awareness and best practices expected in academic and applied research, even when performance is already near-optimal.

Model optimization often begins with tuning the following key hyperparameters:

- **Learning Rate:**
- Initial learning rate values (e.g., 1e-3, 1e-4) are tested and adjusted dynamically using learning rate schedulers such as **ReduceLROnPlateau** or **Cosine Annealing**.
- **Batch Size:**
- Commonly ranges between **8 to 64**, depending on memory constraints. Smaller batch sizes can stabilize learning but increase training time.
- **Optimizer Selection:**
- Optimizers like **Adam**, **SGD with momentum**, or **AdamW** are evaluated based on convergence speed and generalization.
- **Epochs and Early Stopping:**
- Typically, training is conducted for 20–50 epochs, with **early stopping** used to halt training once validation loss plateaus, preventing overfitting.

In our case, default settings provided optimal results early on, so extensive tuning was not required.

Model Evaluation

To assess performance beyond raw accuracy, a set of evaluation metrics and tools are commonly used:

Metric	Purpose
Accuracy	Overall classification correctness
Precision	Focus on false positive minimization
Recall	Focus on false negative minimization
F1-Score	Harmonic mean of precision and recall
Loss Curve	Indicates model convergence and potential overfitting
Confusion Matrix	Highlights class-wise performance and errors

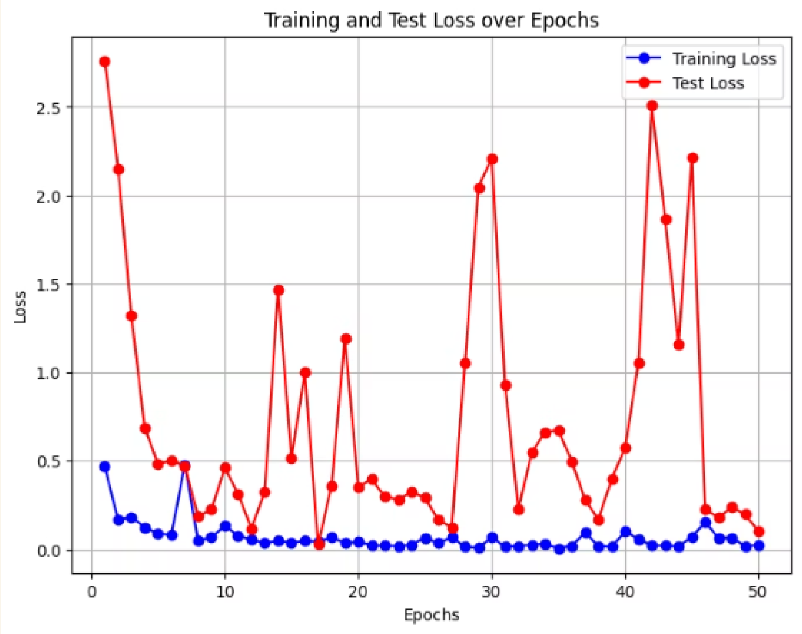
These tools ensure that optimization efforts are informed by quantitative trends and not just overall accuracy.

Evaluation Results

Class	Precision	Recall	F1-Score	Support
Class 0	1.00	0.96	0.98	28
Class 1	0.92	1.00	0.96	12
Accuracy			**0.97**	**40**
Macro Avg	0.96	0.98	0.97	40
Weighted Avg	0.98	0.97	0.98	40

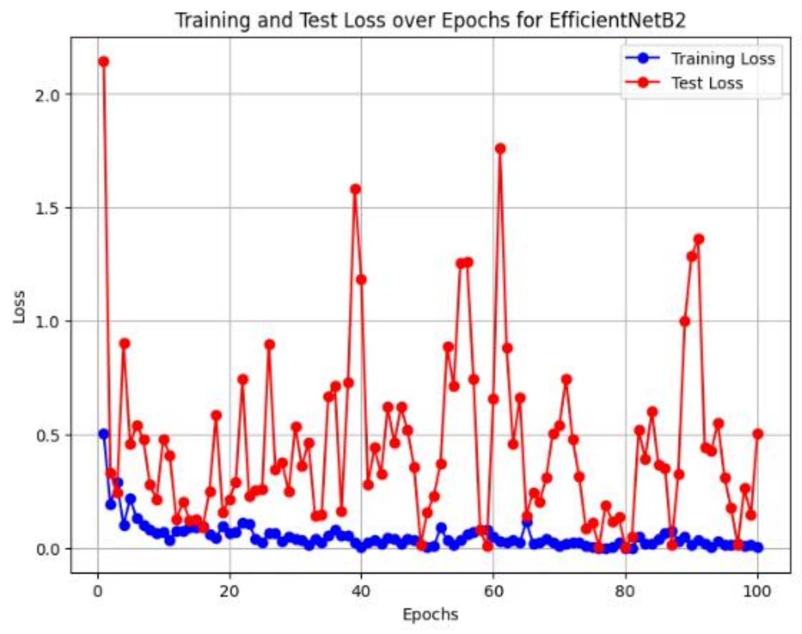
These tools ensure that optimization efforts are informed by quantitative trends and not just overall accuracy.

Evaluation Visualization

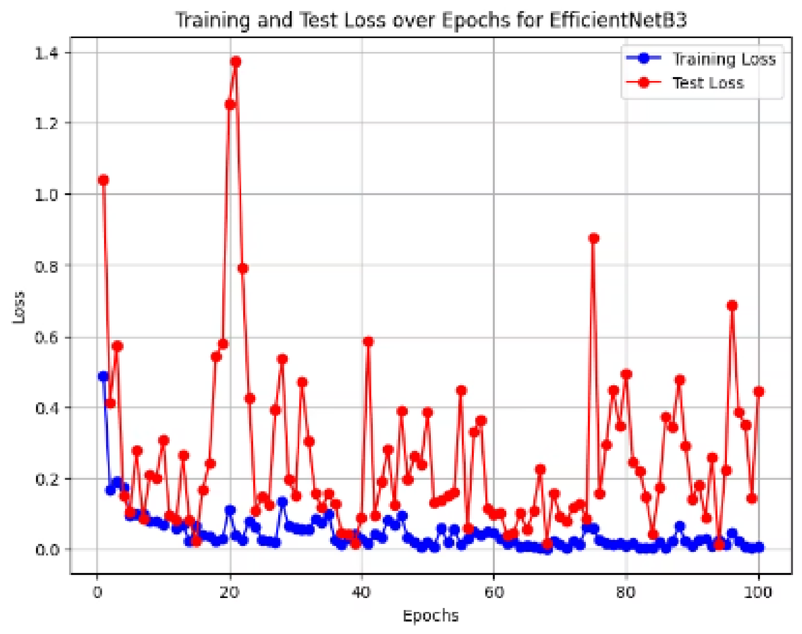


EfficientNet-B0

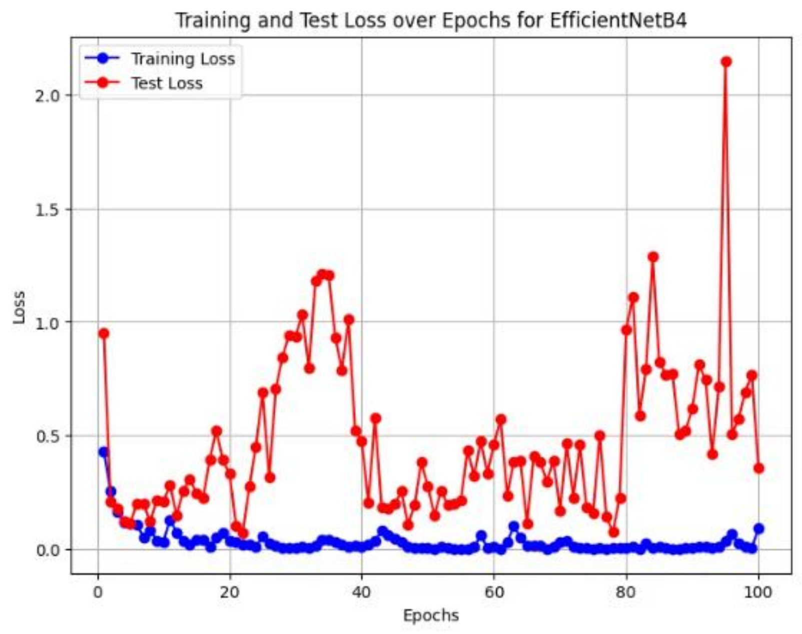
EfficientNet-B1



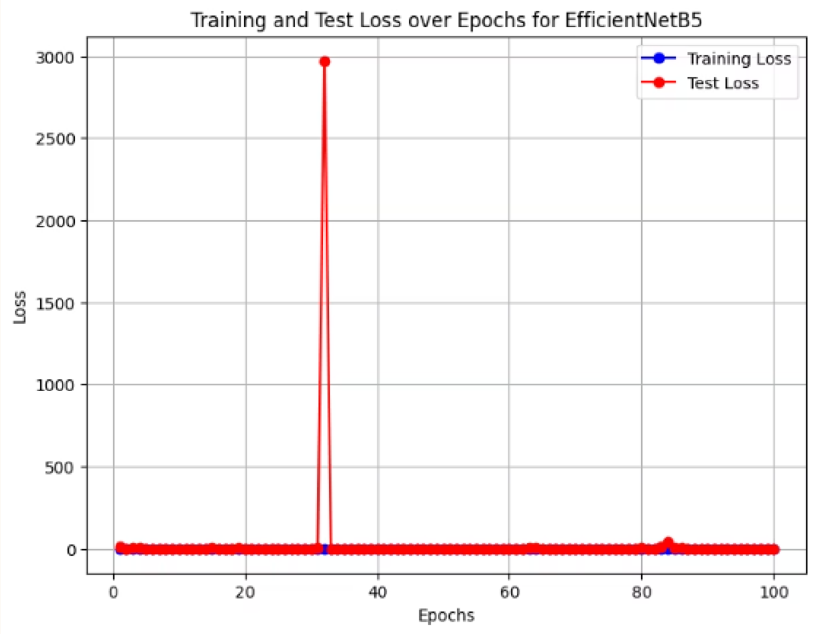
EfficientNet-B2



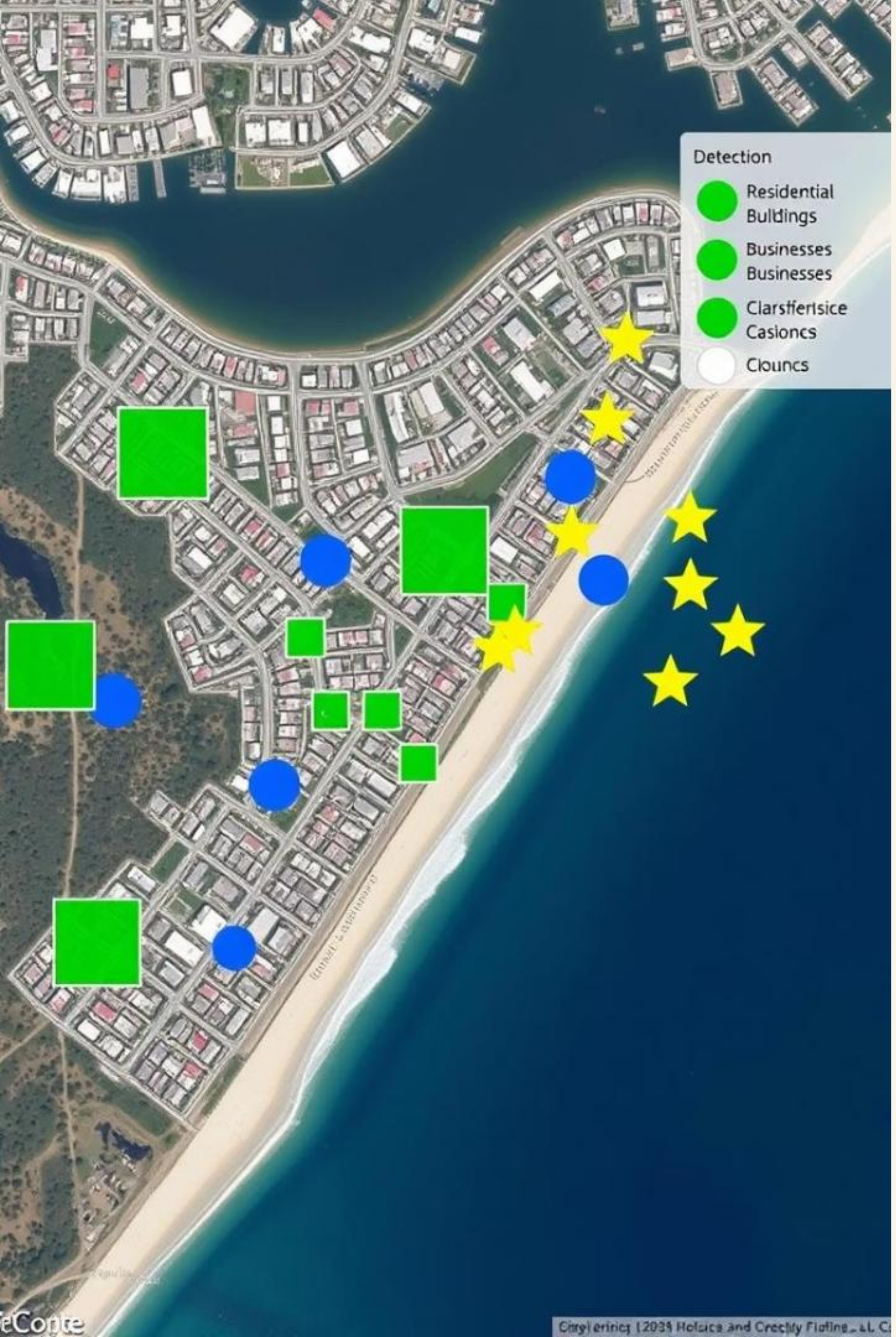
EfficientNet-B3



EfficientNet-B4



EfficientNet-B5



Conclusions and Future Directions

1

Key Findings

PlanetScope provides superior spatial and spectral detail, while Sentinel-1 offers all-weather imaging and structural insights. Canny edge detection effectively extracts features in both datasets.

2

Challenges

Temporal mismatch between datasets and speckle noise in SAR (Synthetic Aperture Radar) imagery posed difficulties. EfficientNet training was straightforward with high baseline accuracy.