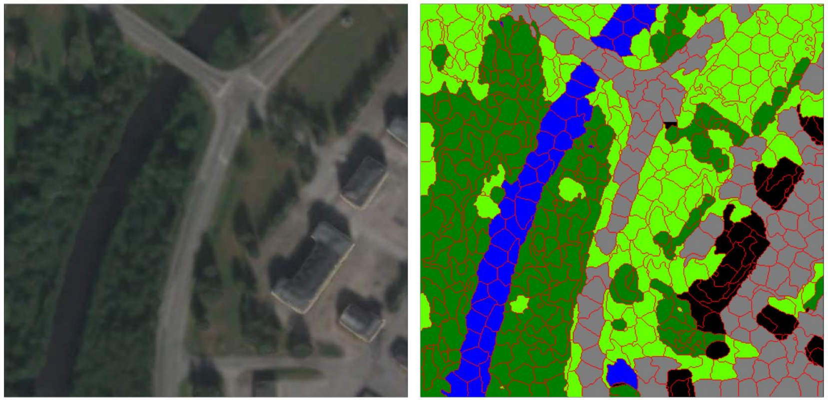
Satellite Image Segmentation

Advanced Image Processing - 22/01/2026

Colin Manyri



Project Overview and Objectives

<u>Project Objectives</u>

- Implement and compare multiple image segmentation methods. (thresholding, histograms, CNN)

- Challenges of optimizing computational complexity (low-power computer).
- Focus on methodology over absolute performance metrics

I. Dataset Presentation and Feature Extraction
II. Multiples image segmentations techniques
III. Inference, Post-Processing, and Predictions
IV. Results and Performances
V. Conclusion and Perspectives

Presentation of the Dataset

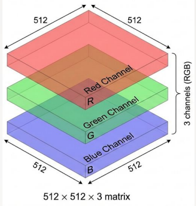
Land Cover from Aerial Imagery - LandCover.ai

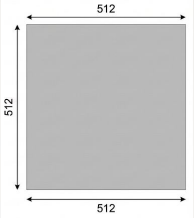
A dataset for mapping of buildings, woodlands, water and roads from AIs.

Dataset structure

https://landcover.ai.linuxpolska.com/

- **Images**: Around 10,000 satellite images, with a resolution of 512 × 512, stored in *.jpg* format.
- **Masks (Labels)**: Each image is associated with a **monochannel** segmentation mask, used as ground truth and stored in *.png* format.



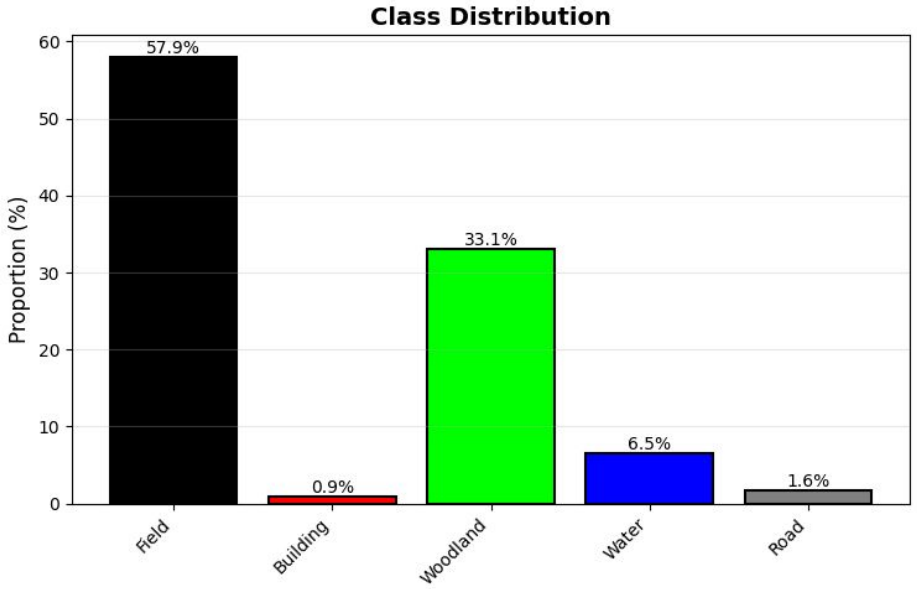


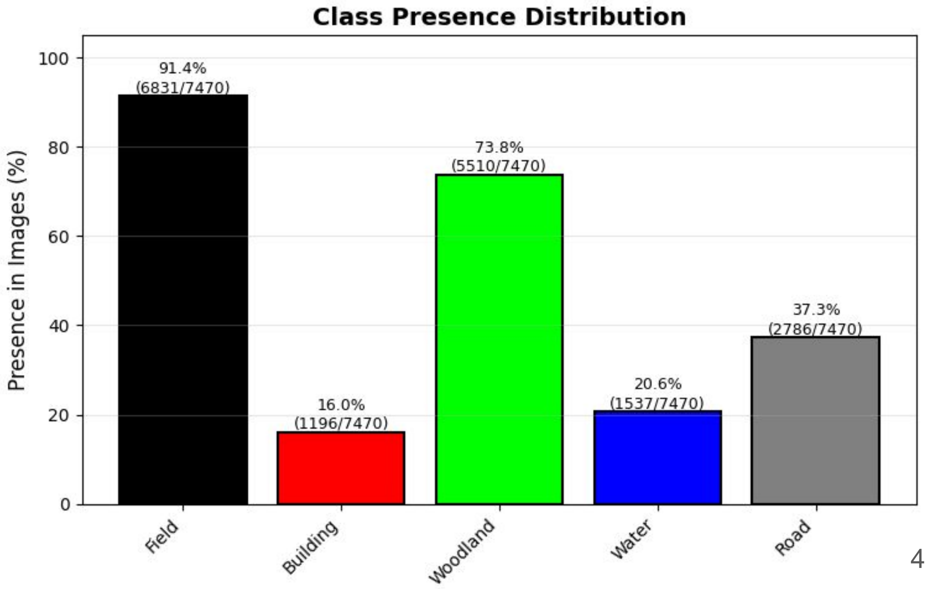
- **Data Splits**: The dataset is divided into ***Train, Val and Test*** subsets for model development and evaluation.
- **Classes**: The segmentation task includes 5 semantic classes: *Field, Woodland, Building, Water, and Road.*

About the classes

Class imbalance: lots of *Field*, few *Buildings* and *Roads*.

- **Accuracy is misleading:** high accuracy ≠ good classification.
- **Focus on minority classes:** monitor recall and consider alternative metrics (F1, IoU).
- **Use data augmentation or class weighting** to help underrepresented classes.





Feature choice

The current observation

- Heterogeneous dataset
- Images vary in intensity, texture, etc…
- **Hard to know a priori which features are most discriminative.**

The best solution

- Extracting a broad set of features (color, texture, geometry, multi-scale context) ensures all possible discriminative cues are captured.
- Dimensionality reduction (PCA) or feature selection can be applied later.

Robustness for probabilistic models : Ensures that all relevant patterns can be captured by histogram-based or learning-based models.

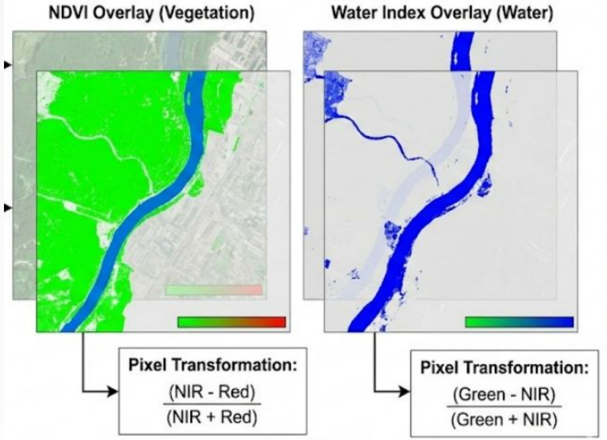
Color / Spectral
RGB, HSV

Intensity / Context:
Grayscale, Multi-scale Blur

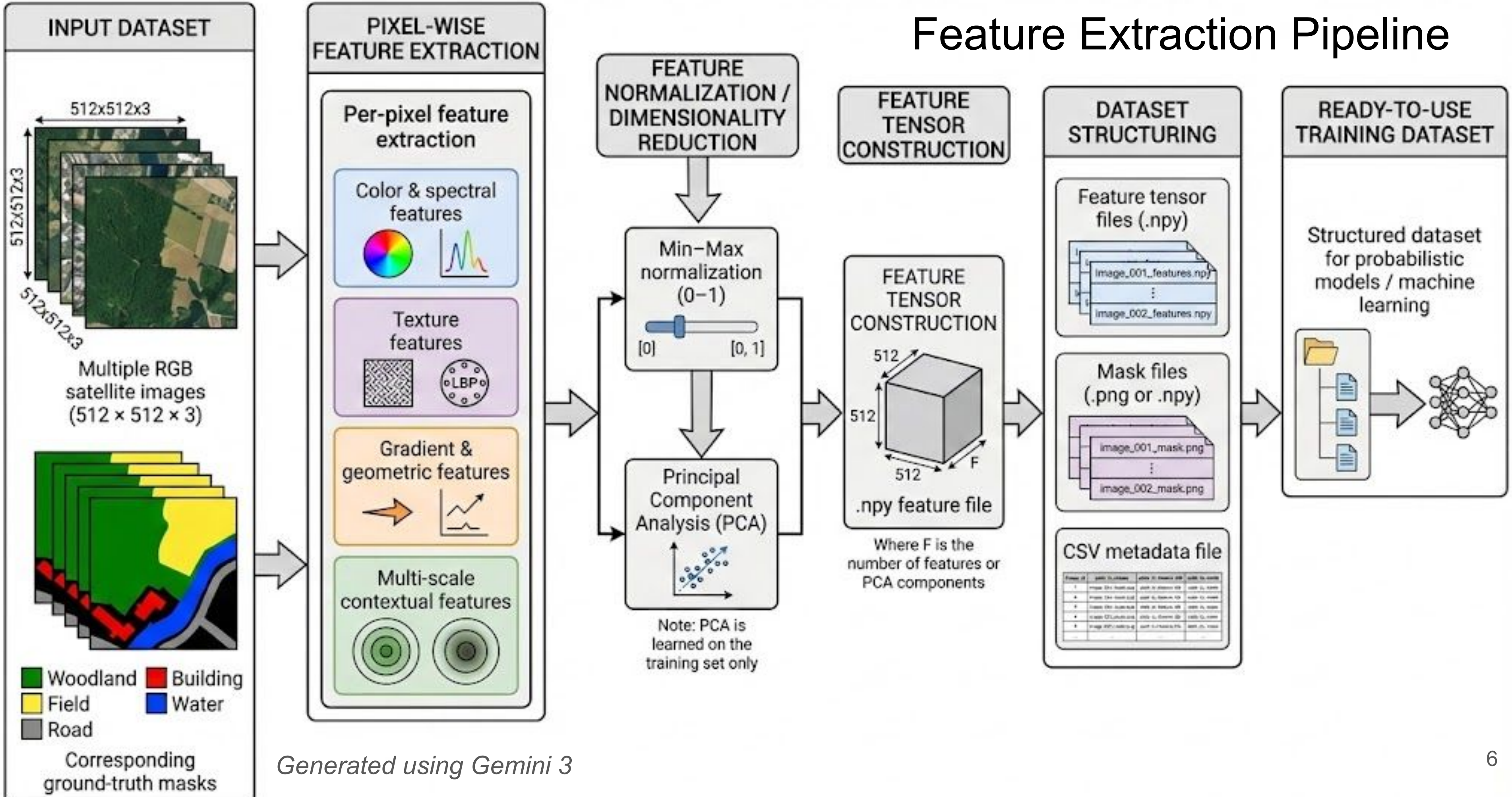
Gradient / Geometry
Gradient, Anisotropy

Texture
Variance, Entropy, LBP

Indices
NDVI, Water

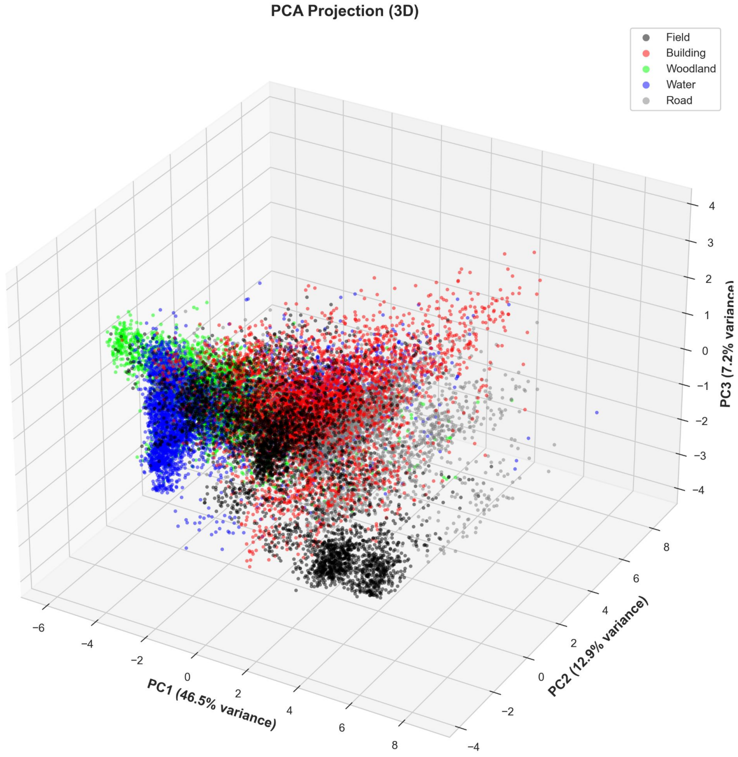


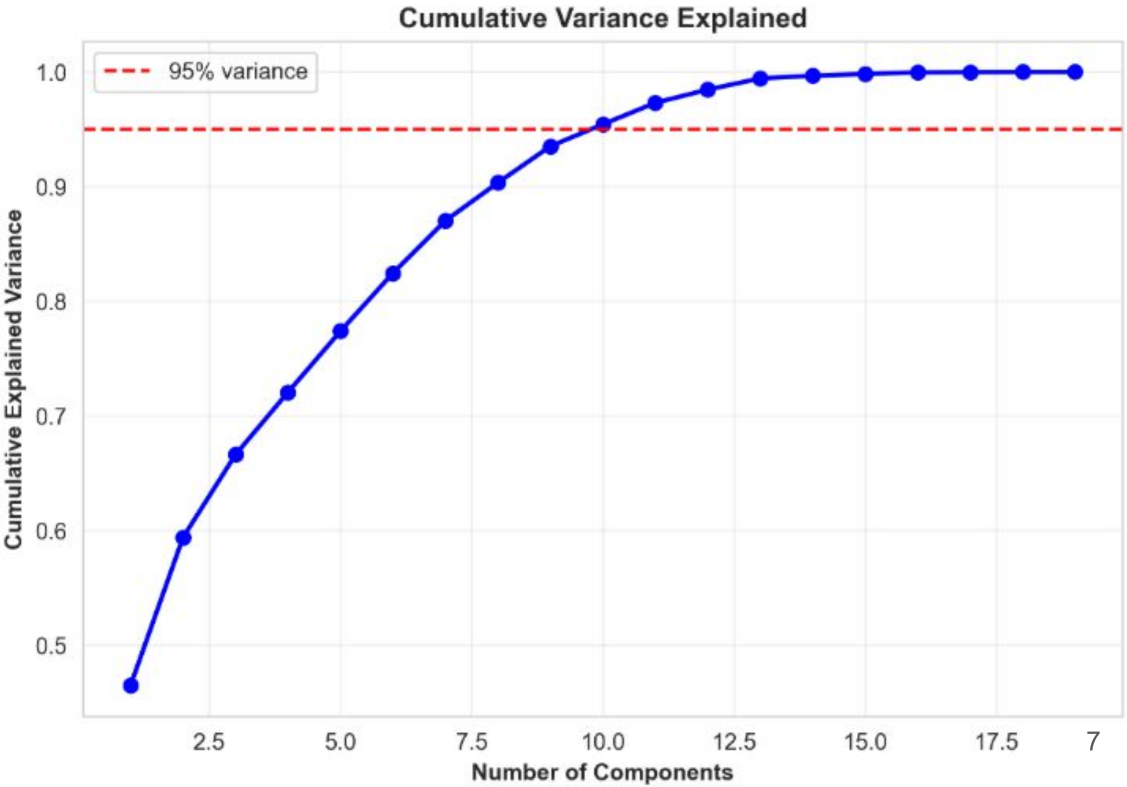
Feature Extraction Pipeline



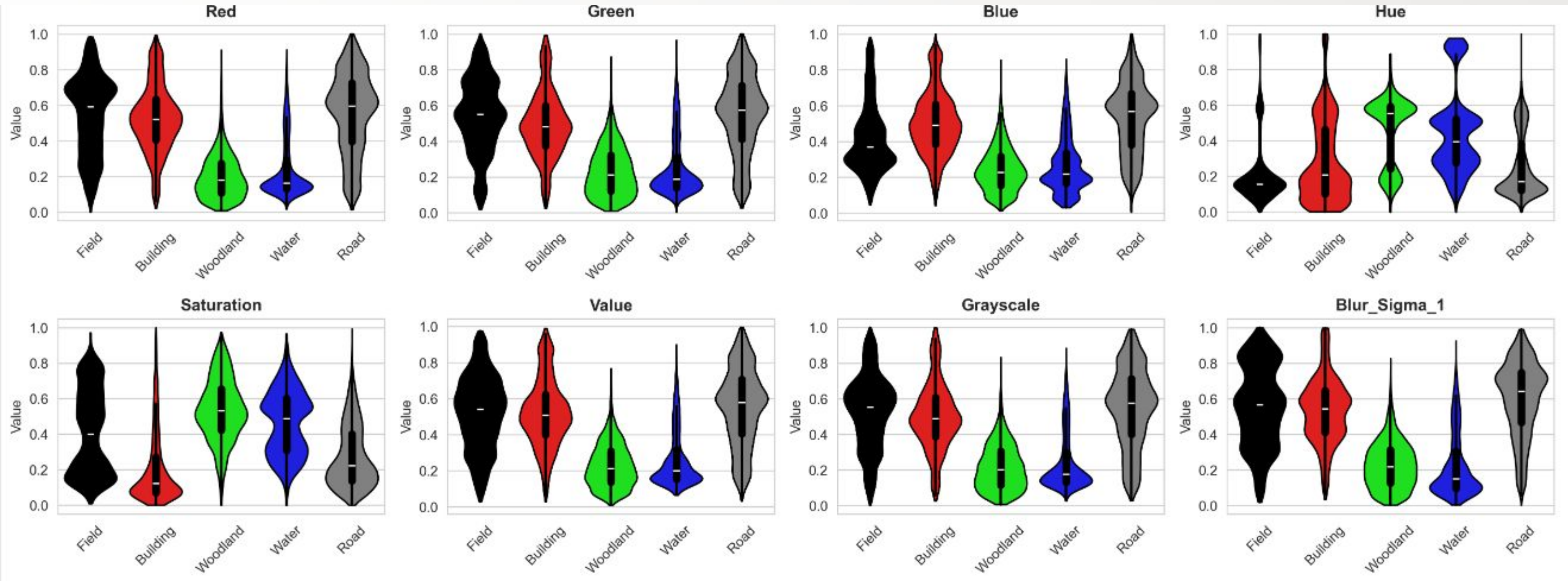
Generated using Gemini 3

Selected Feature analysis





Selected Feature analysis



Histogram-Based Probabilistic Segmentation

- Each pixel is represented by a **feature vector** x∈ℝ^F
- For each class c, a **feature distribution** is learned using **histograms**
- Segmentation is performed by assigning each pixel to the most probable class:

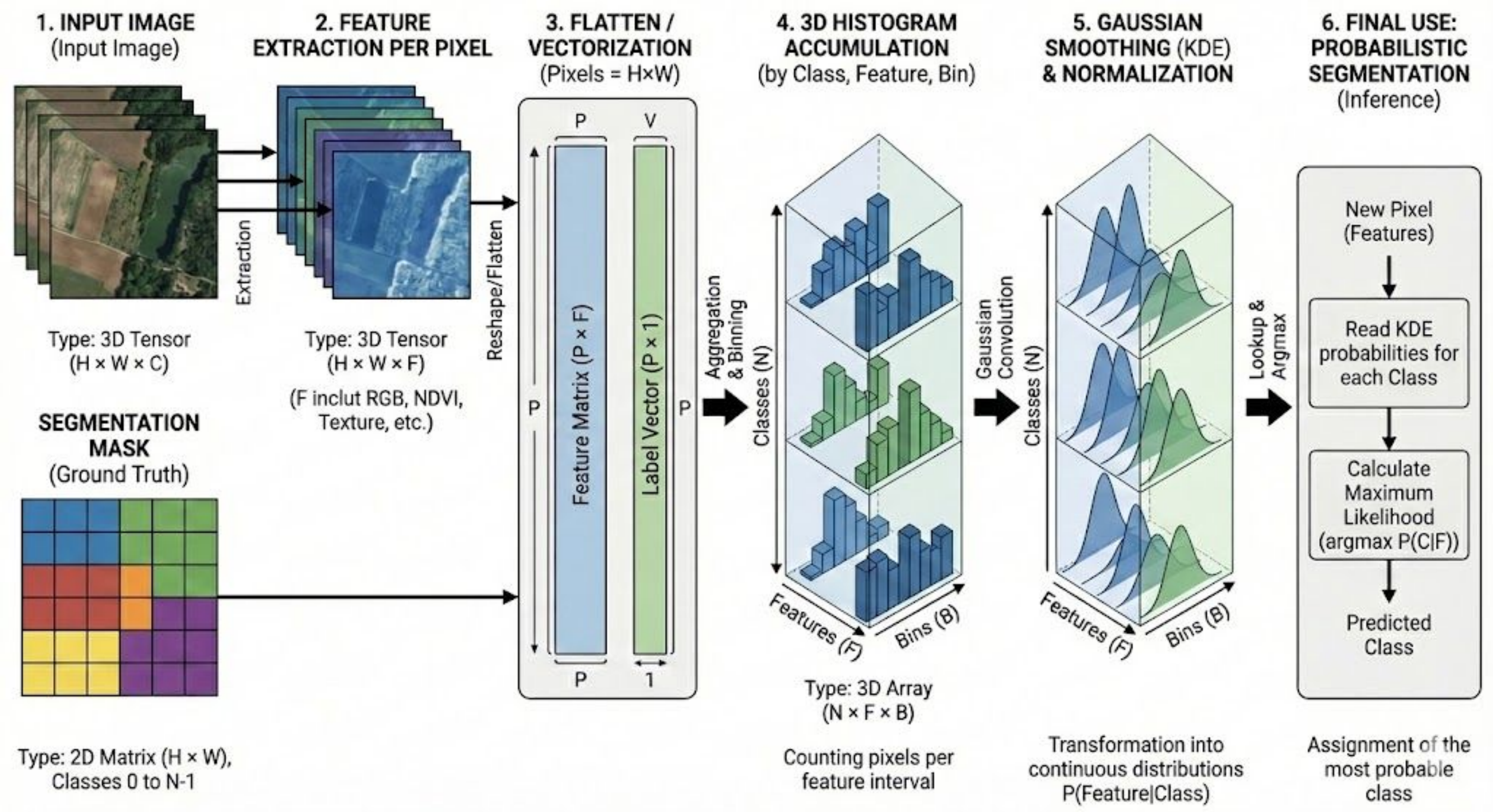
$$\hat{c} = \arg\max_{c} \; P(\mathbf{x} \mid c)$$

- Feature space is discretized into bins
- For each class ccc, a histogram estimates the density:

$$P(\mathbf{x} \mid c) \approx \frac{N_{c,\mathrm{bin}(\mathbf{x})}}{N_c}$$

Aspect	Multi-thresholding	Histogram-based Segmentation
Decision type	Hard, binary decision	Soft, probabilistic decision
Feature usage	Single or few features	Multi-dimensional feature vectors
Decision boundary	Fixed thresholds	Data-driven decision surfaces
Handling uncertainty	Not modeled	Explicitly modeled
Class overlap	Poorly handled	Naturally handled
Robustness to noise	Low	Higher

Histogram-Based Probabilistic Segmentation

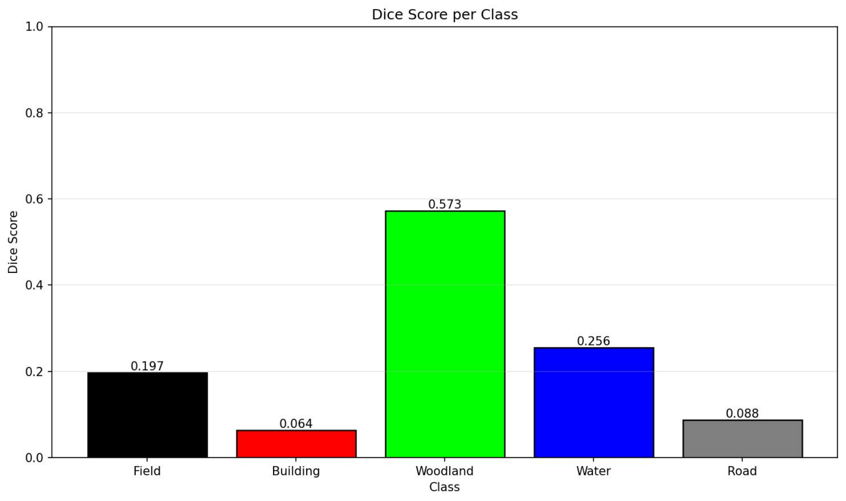


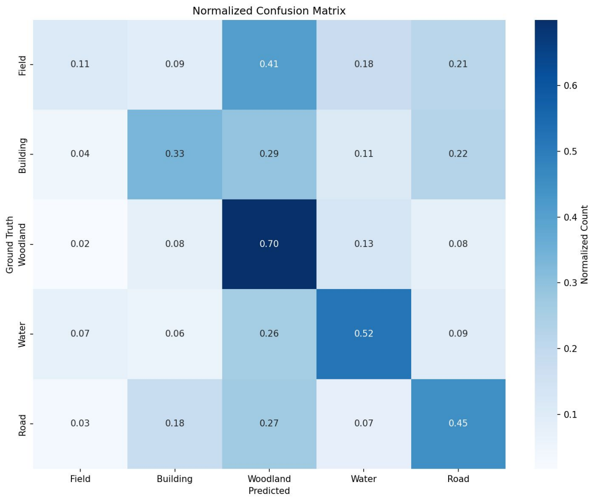
Generated using Gemini 3

Results

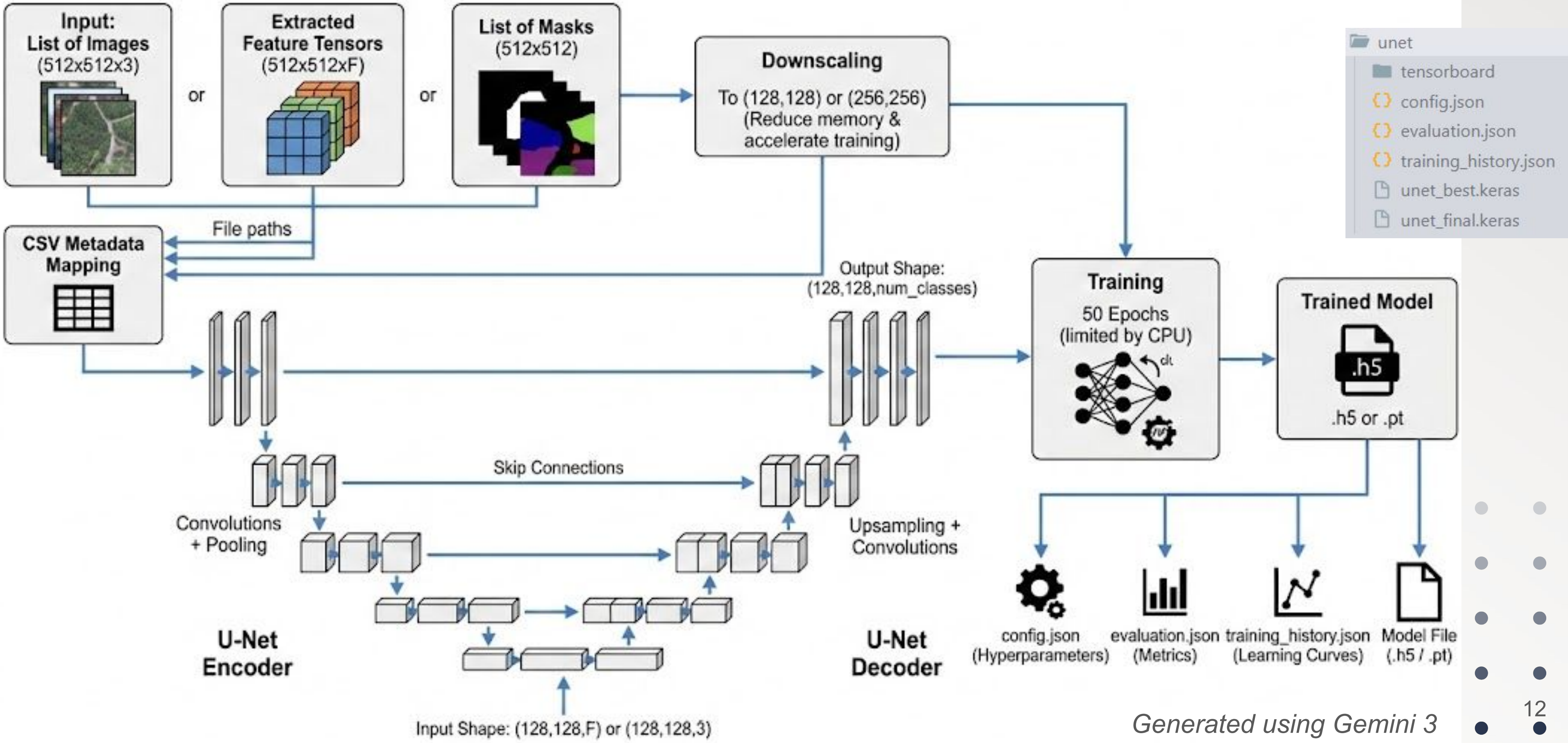
Histogram

Very bad results for now



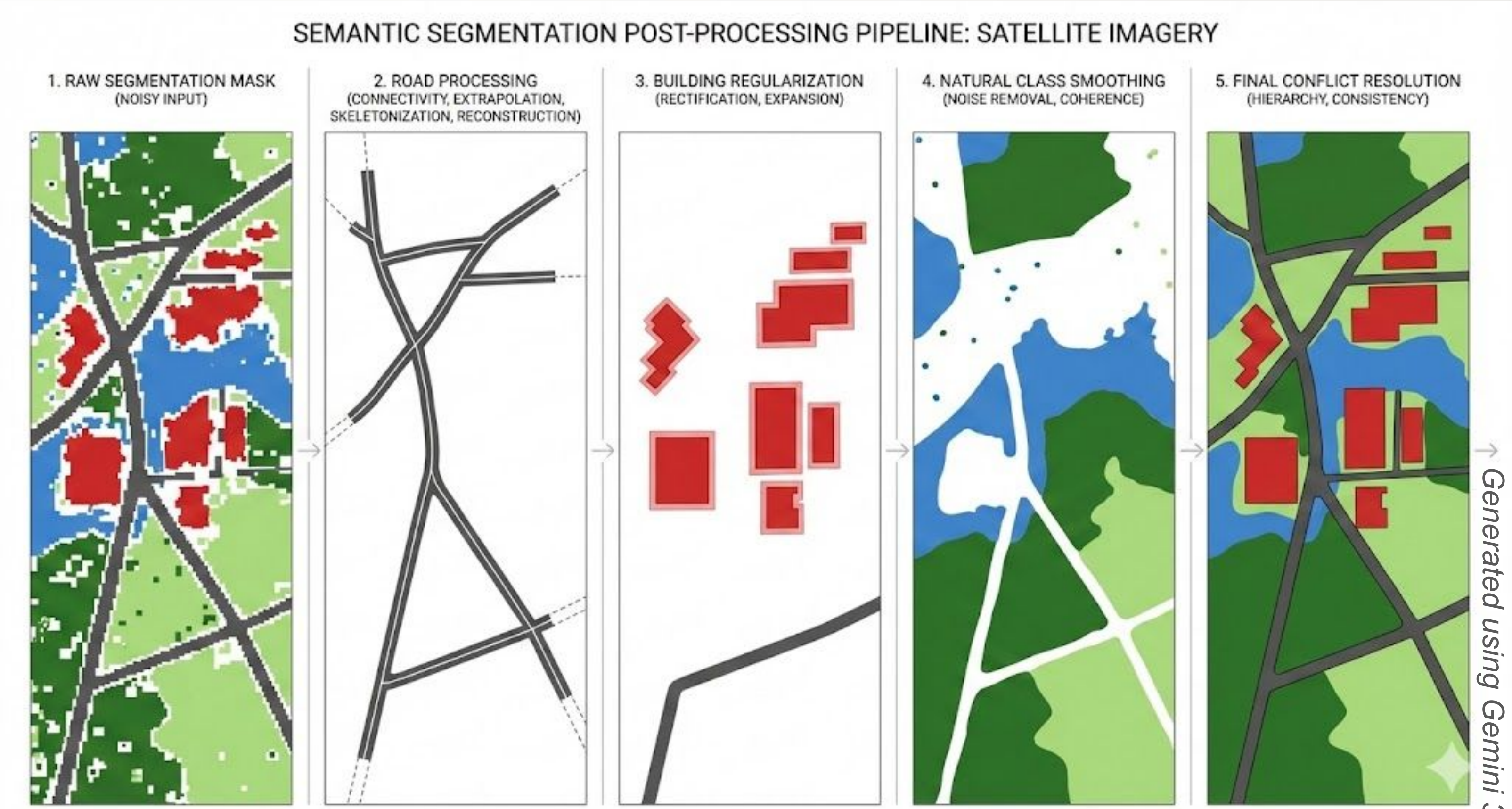


Another Technique: CNN Pipeline using U-Net



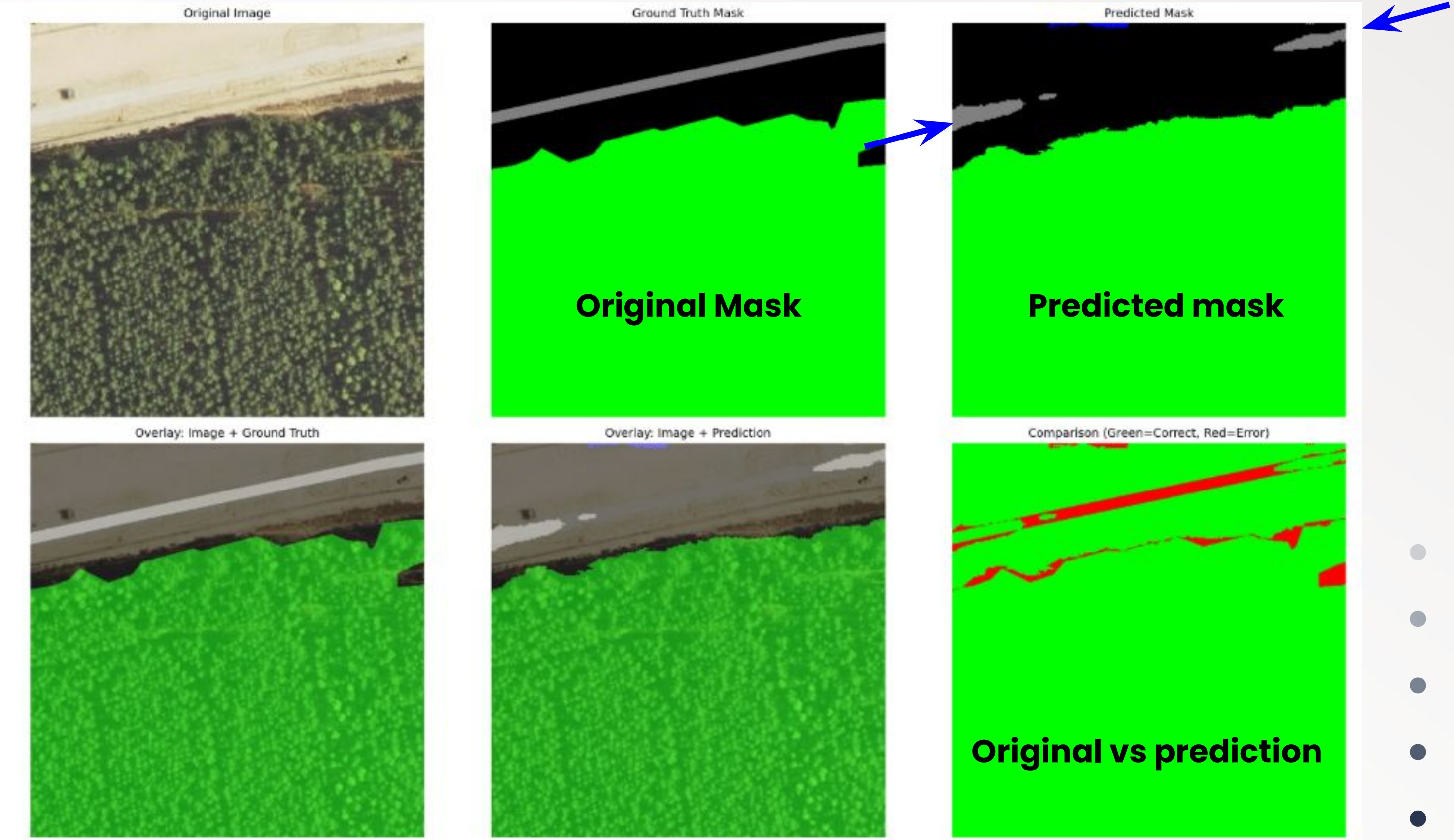
Generated using Gemini 3

Post-Processing : *Main Ideas*

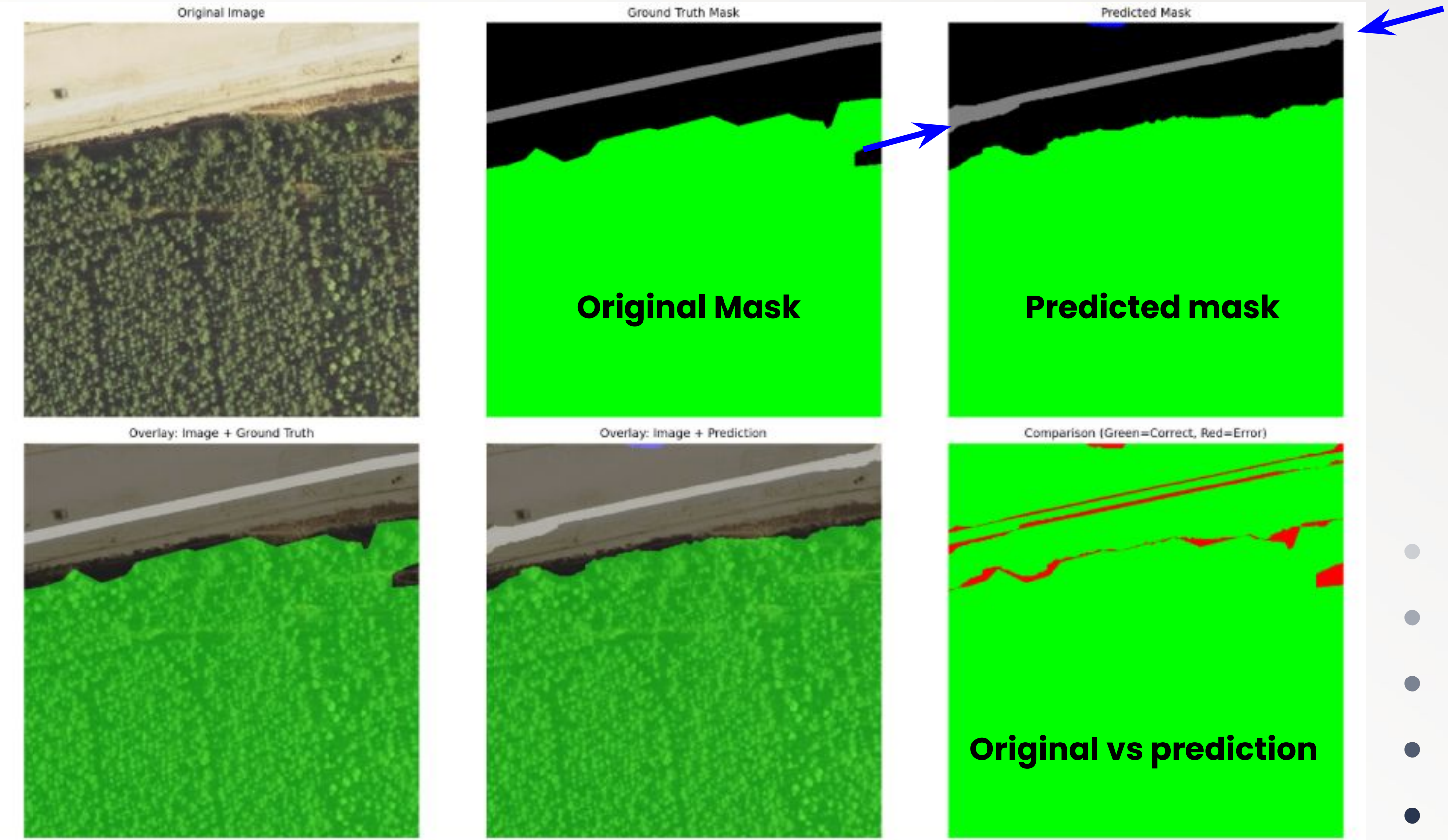


Generated using Gemini 3

Post-Processing Example : *Before*



Post-Processing Example : *After*



Choice of Evaluation Metrics

Dice Coefficient

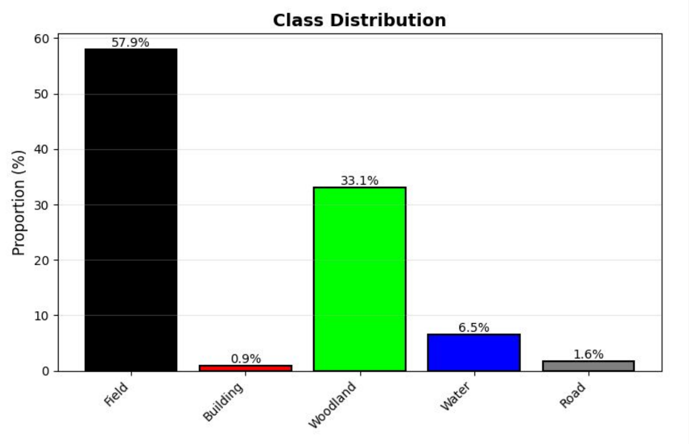
- Measures **spatial overlap** between prediction and ground truth
- Sensitive to **false negatives and false positives**, unlike accuracy
- Ideal for **imbalanced classes**

$$\text{Dice} = \frac{2TP}{2TP + FP + FN}$$

- *TP (True Positives)*
- *FP (False Positives)*
- *FN (False Negatives)*

Why not Accuracy

- Large class imbalance in satellite images
- accuracy can be **misleadingly high**



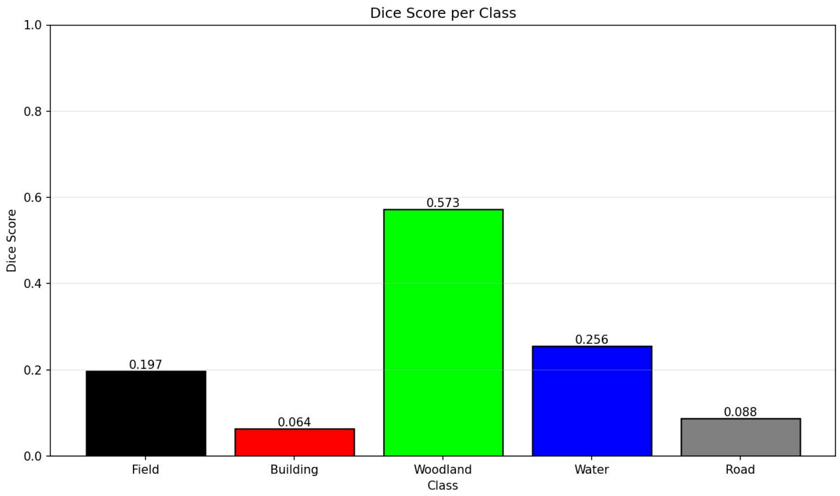
Class-specific strategy

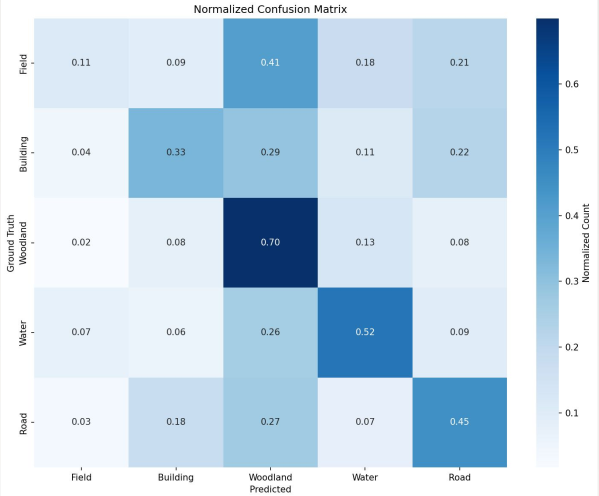
- **Roads and Buildings**: prioritize **high recall**
- Slight over-prediction (FP) is acceptable

Results

Histogram

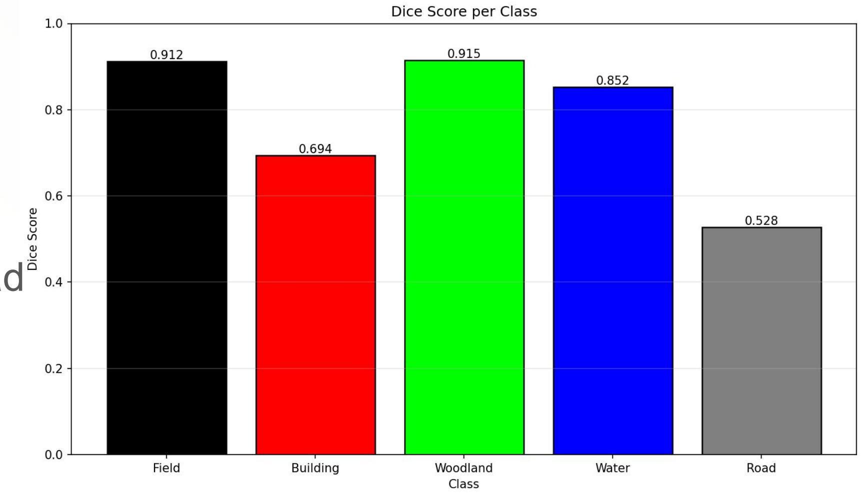
Very bad results
for now

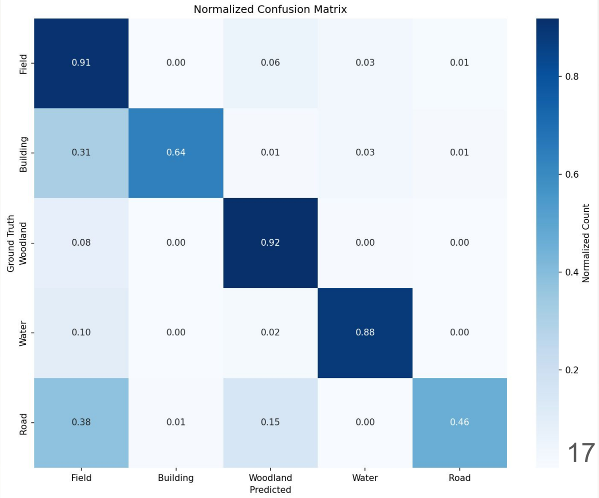




Unet

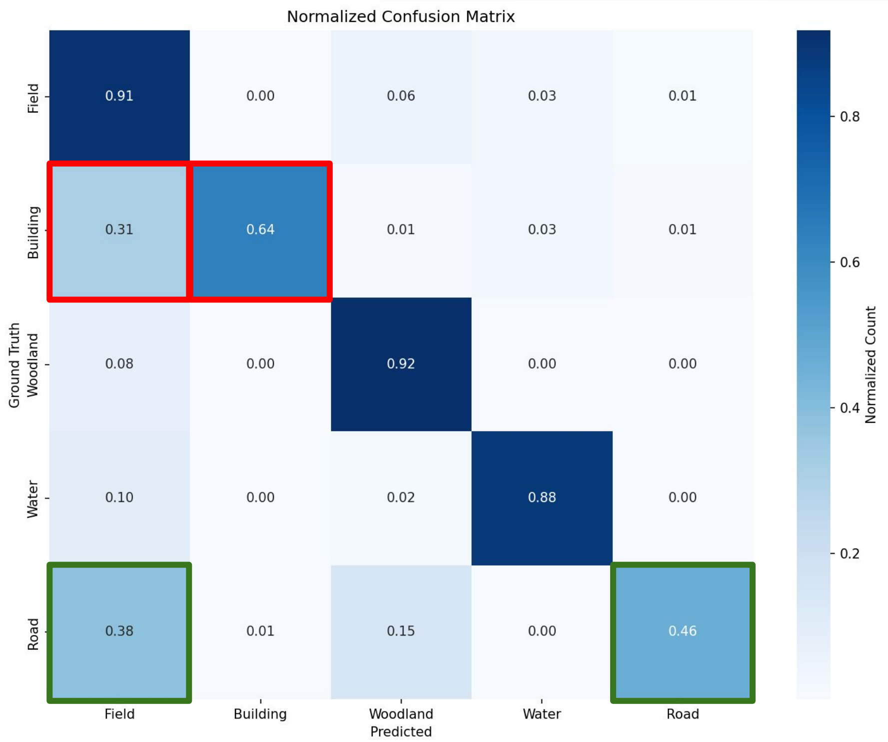
Good results
Moderate Dice
on Building and Road



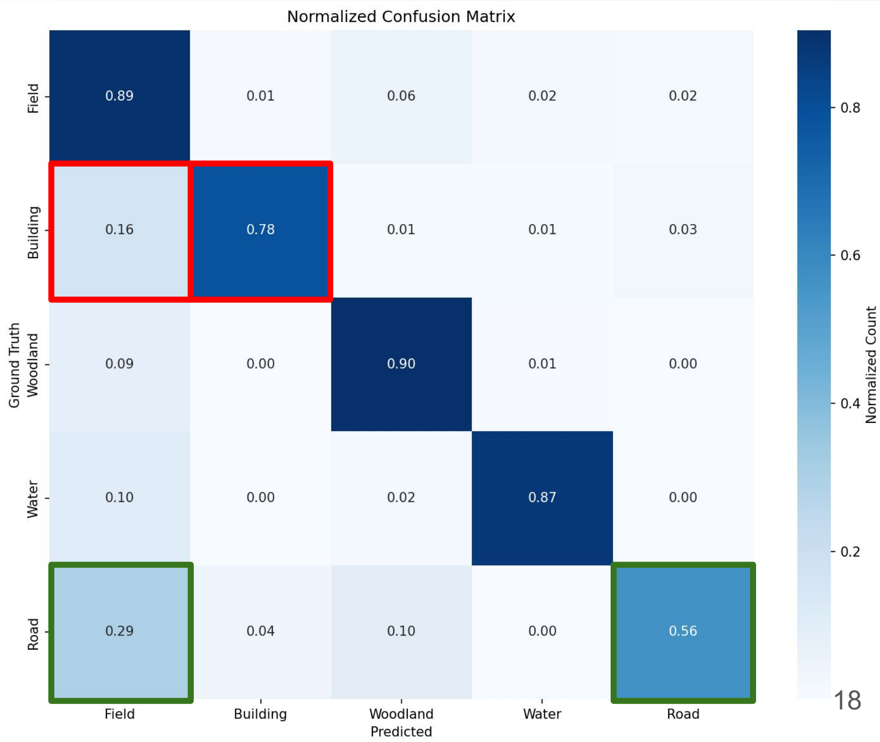


Post-Processing Impact

Before Post-Processing

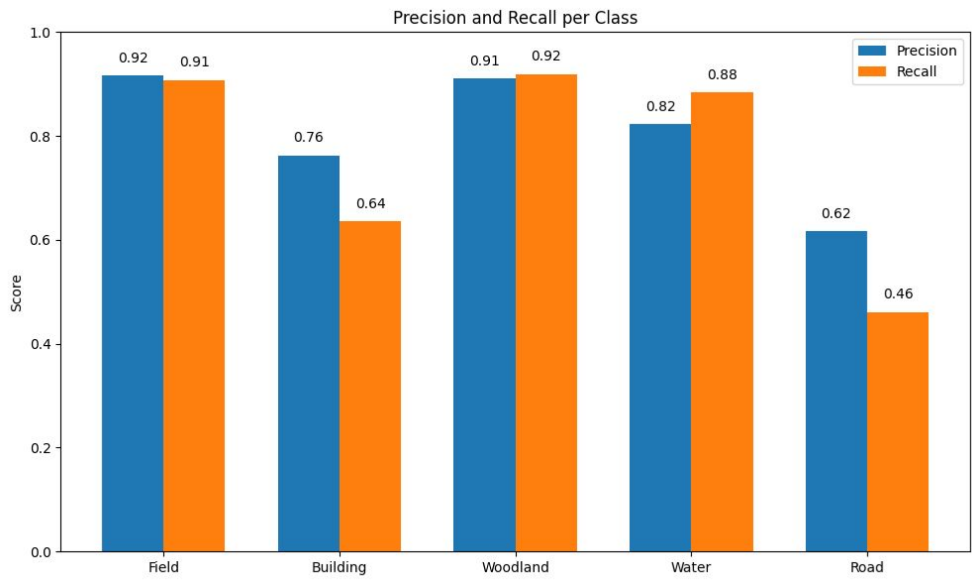


After Post-Processing

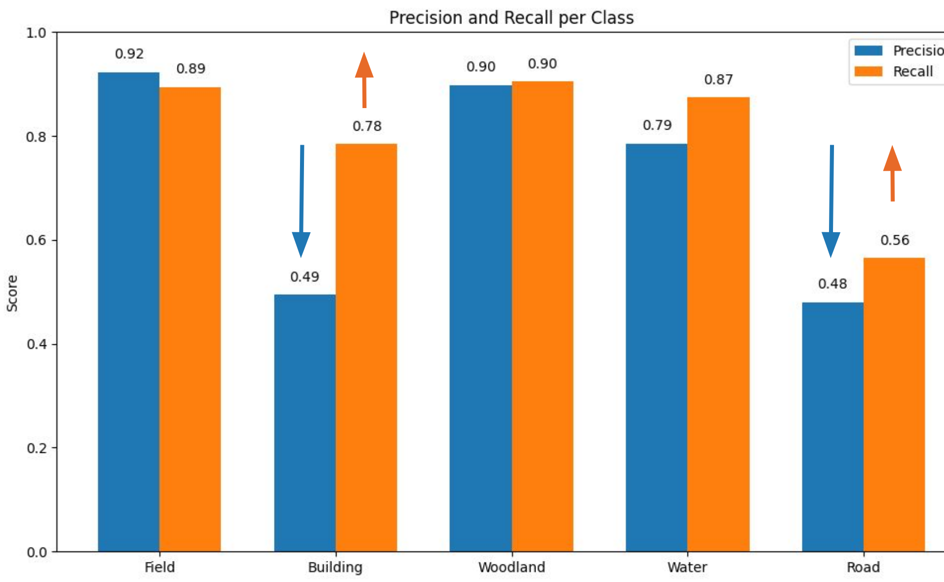


Post-Processing Impact

Before Post-Processing



After Post-Processing



Conclusion : Improvement Perspectives

Learning & Actions Taken

- **Built a pipeline** for feature extraction, preprocessing, training, and evaluation
- **Achieved good U-Net segmentation results**, while handling class imbalance and prioritizing recall for small structures
- Learned the importance of selecting relevant features for effective segmentation

Perspectives / Future Work

- **Improve computation efficiency**, code robustness
- **Explore better histogram-based methods** and additional or more relevant features
- **Improve Post-Processing pipeline** for better performances

References :

First image: https://www.mdpi.com/2072-4292/8/4/329
Dataset Official Link: https://landcover.ai.linuxpolska.com/
Dataset Kaggle Link: https://www.kaggle.com/datasets/aletbm/land-cover-from-aerial-imagery-landcover-ai

GitHub Repository

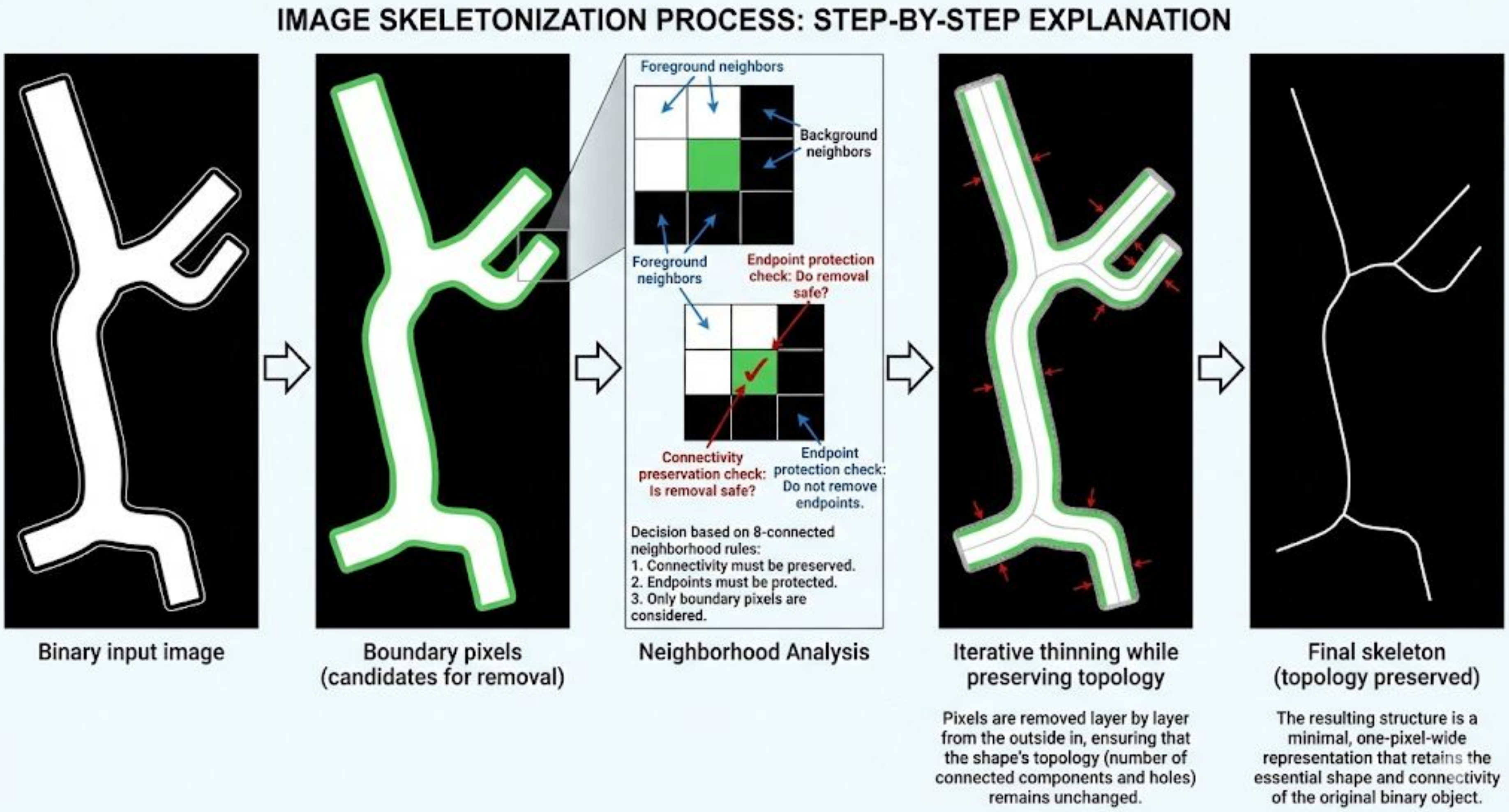


IMAGE SKELETONIZATION PROCESS: STEP-BY-STEP EXPLANATION
Foreground neighbors
Background neighbors
Foreground neighbors
Endpoint protection check: Do removal safe?
Connectivity preservation check: Is removal safe?
Endpoint protection check: Do not remove endpoints.
Decision based on 8-connected neighborhood rules:
1. Connectivity must be preserved.
2. Endpoints must be protected.
3. Only boundary pixels are considered.
Binary input image
Boundary pixels (candidates for removal)
Neighborhood Analysis
Iterative thinning while preserving topology
Final skeleton (topology preserved)
Pixels are removed layer by layer from the outside in, ensuring that the shape's topology (number of connected components and holes) remains unchanged.
The resulting structure is a minimal, one-pixel-wide representation that retains the essential shape and connectivity of the original binary object.