AI Student Symposium 2026

CSC 525

Team 35

Automated Bridge Surface Crack Detection using UAV Imagery and Deep Learning Segmentation

Kunj Patel, Ishaan Patel | Instructor: Dr. Edgar Lobaton | Neural Networks

Motivation

Bridge inspection today relies on inspectors physically climbing or hanging off structures to identify cracks. This process faces three critical pain points:

SAFETY — Inspectors work at height, in traffic, and on difficult terrain exposing them to hazardous environments.

COST & TIME — A single inspection can run tens of thousands of dollars, bottlenecking infrastructure projects.

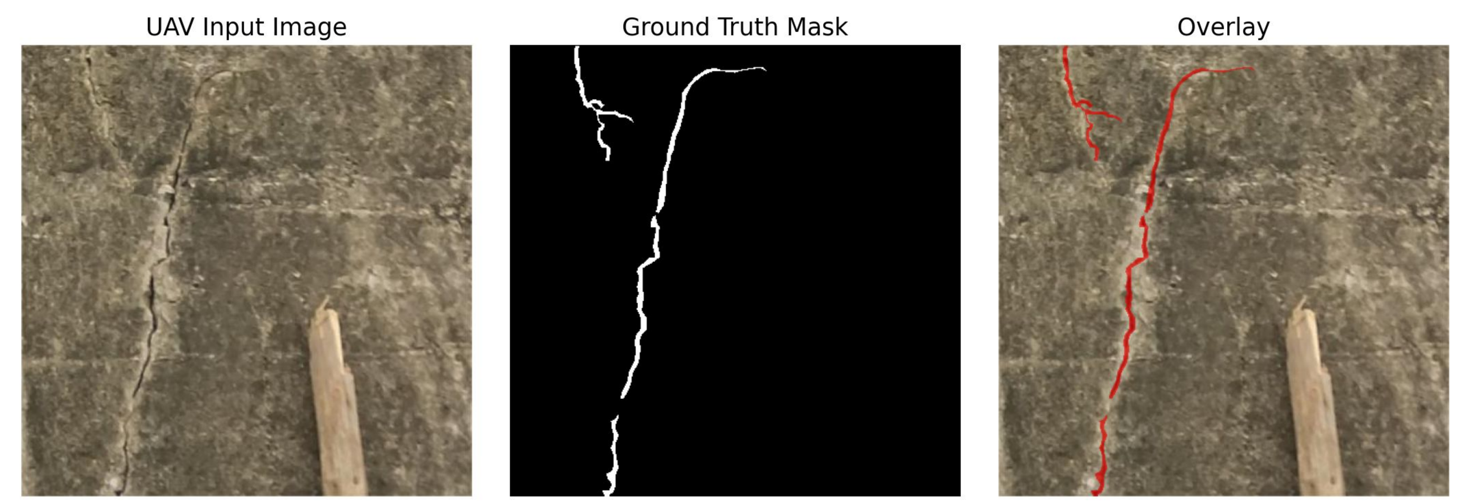
CONSISTENCY — Visual inspection is subjective; two engineers may disagree about the same crack.

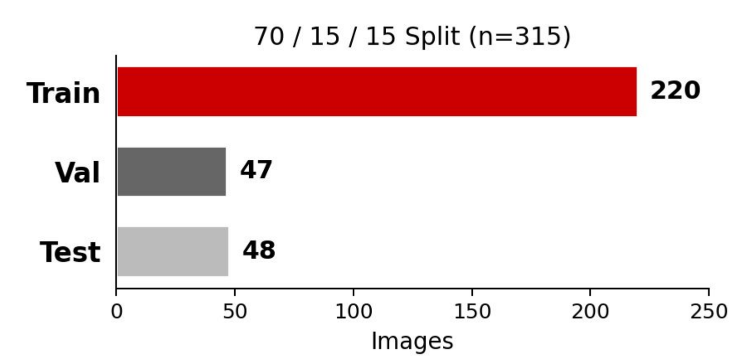
OUR SOLUTION

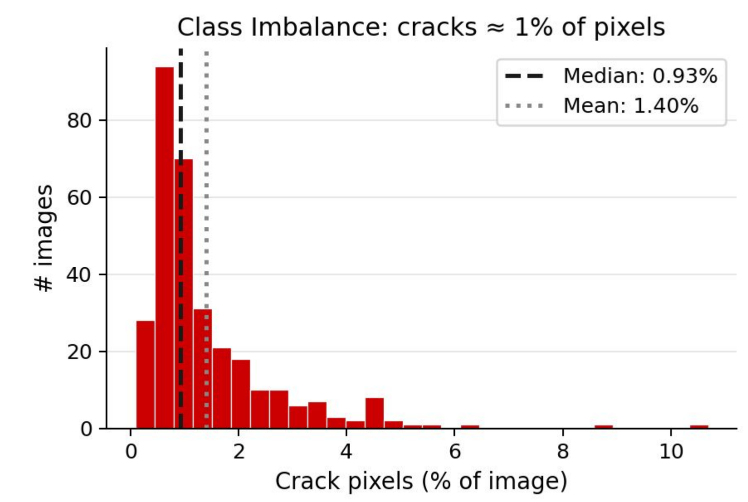
A UAV-drone system that automatically detects cracks at the pixel level. Output is a full spatial mask showing exactly where cracks exist giving engineers a prioritized repair map to augment human decision making instead of replacing it.

Data & EDA

UAV-based crack dataset (Kaggle): 315 image/mask pairs at 448×448 with pixel-level binary masks.

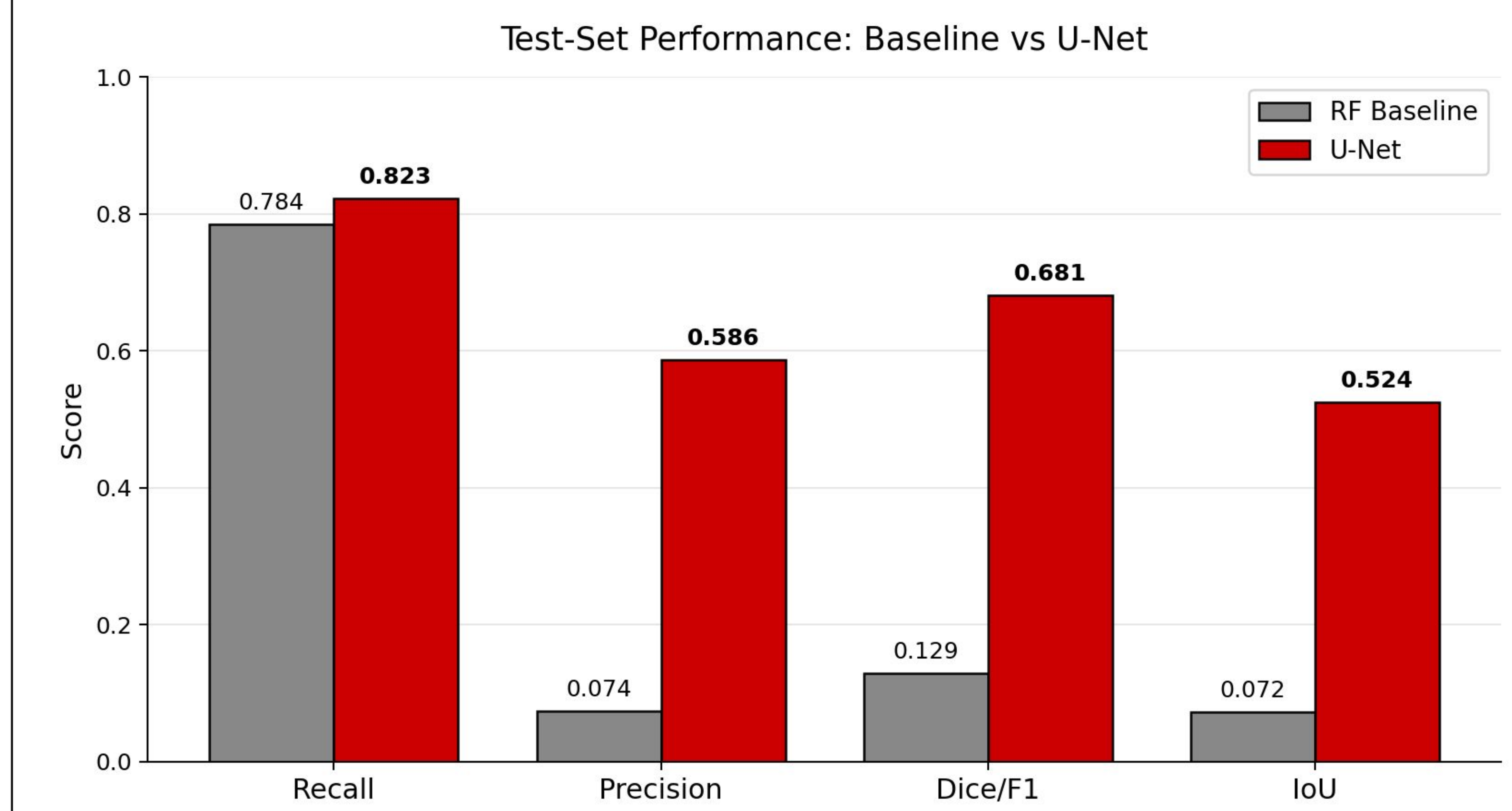






Takeaway: Severe class imbalance (median ~0.93% crack pixels) motivates our choice of Dice-based loss and recall/IoU as primary metrics.

Results | Quantitative



+8× Precision vs. baseline

+5× Dice / F1 improvement

82.3% Recall on test set

58.6% Precision on test set

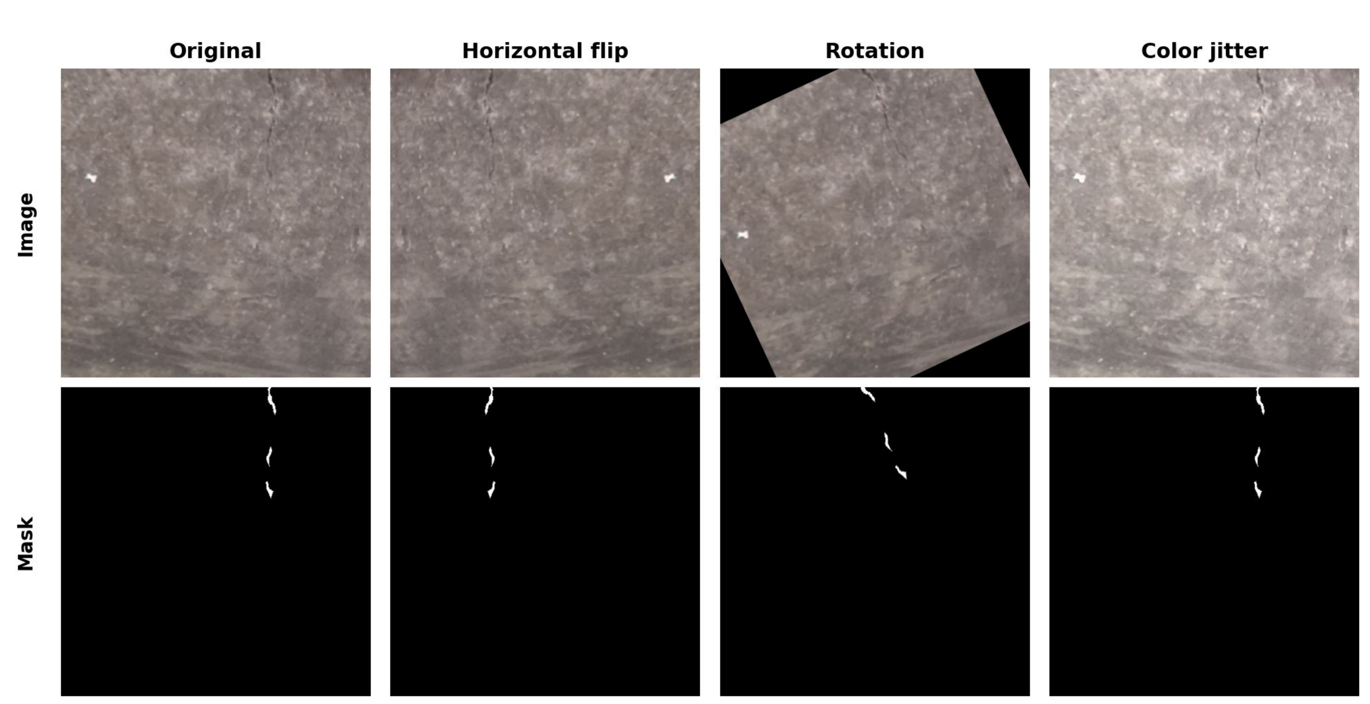
Approach

BASELINE | Random Forest

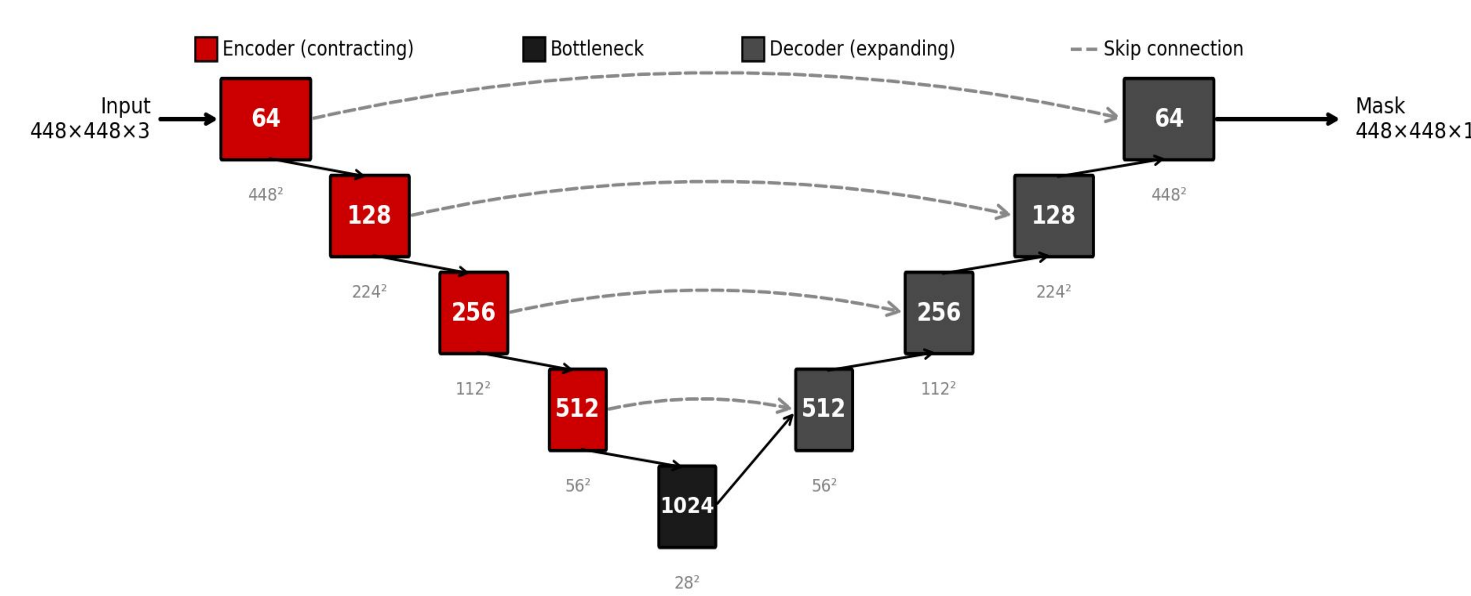
- Tile each image into 16×16 patches.
- Extract handcrafted features: Local Binary Patterns (texture), Histogram of Oriented Gradients (edges), and intensity statistics.
- Train class-balanced Random Forest (n_estimators=100, max_depth=20).
- Reconstruct full 448×448 mask by labeling each patch independently.

PROPOSED | U-Net Segmentation

- Fully convolutional encoder–decoder. Trained end-to-end at 448×448.
- Combined loss = 0.5·BCE + 0.5·Dice. Adam optimizer, lr=1e-4, 25 epochs.
- Spatial + photometric augmentations: flips, rotation, color jitter.
- 70/15/15 train/val/test split (220/47/48 images).



Network Architecture



4-level U-Net (~31M parameters). Skip connections preserve fine edge detail across scales.

Loss Function

$$\mathcal{L} = 0.5 \cdot \mathrm{BCE} + 0.5 \cdot (1 - \mathrm{Dice})$$

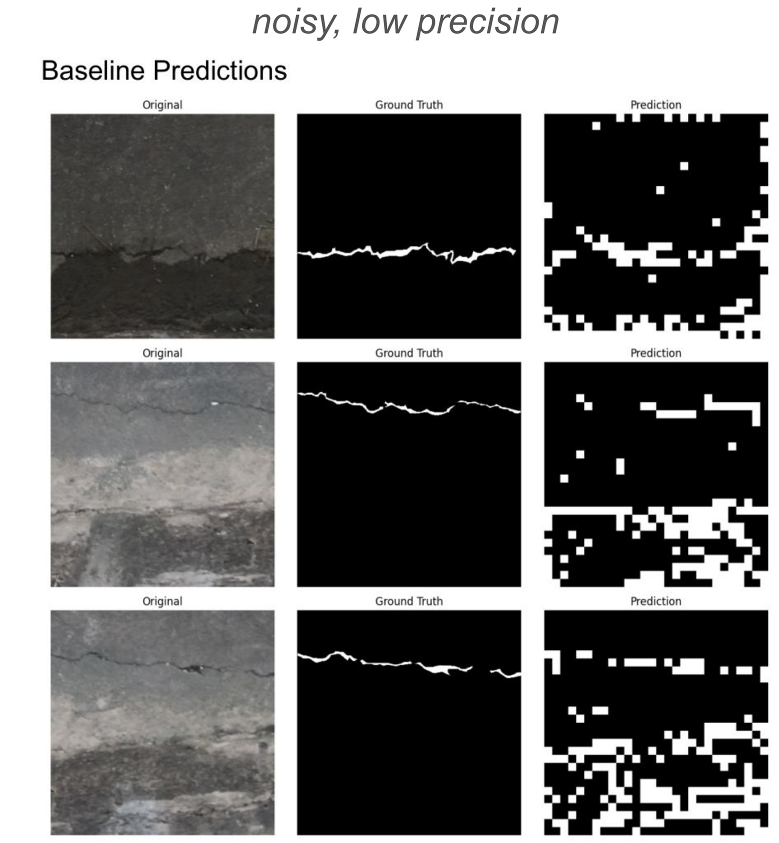
BCE
Pixel-level cross-entropy.
Keeps gradients stable.

Dice
Directly optimizes overlap.
Handles severe class imbalance.

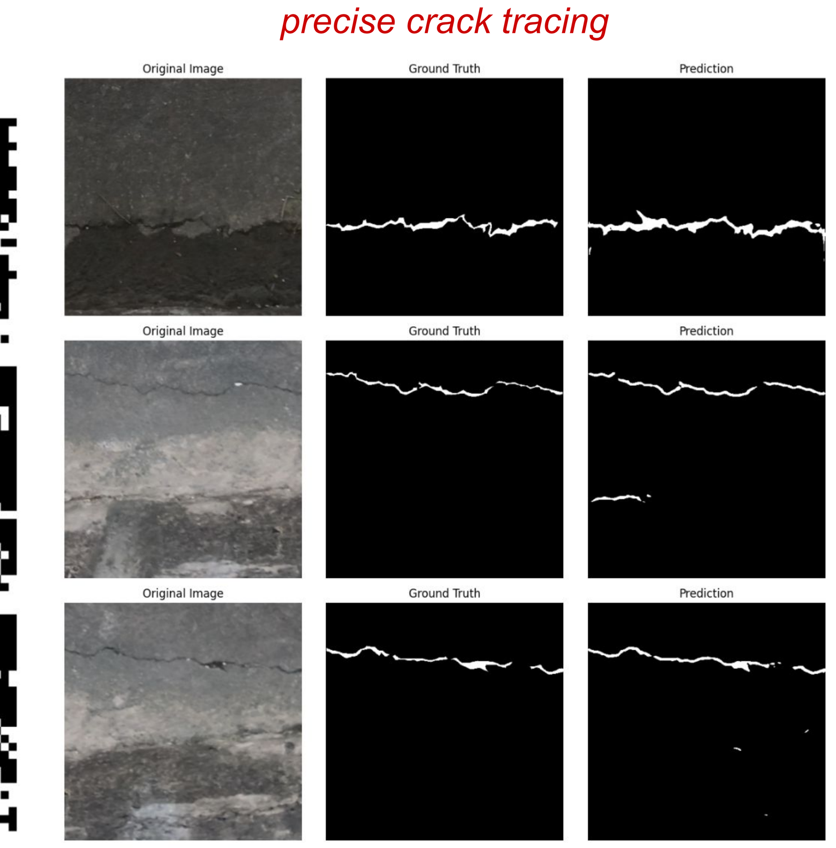
Results | Qualitative

RF Baseline
noisy, low precision

Baseline Predictions



U-Net
precise crack tracing



Why U-Net wins: The Random Forest baseline classifies 16×16 patches independently, so it has no awareness of surrounding image structure as predictions fragment into disconnected blocks. The U-Net learns spatial context end-to-end, and its skip connections preserve both fine edge detail and global crack continuity, producing clean, traceable masks.

References

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. MICCAI.
[2] Ziya07 (2024). UAV-based Crack Detection Dataset. Kaggle.
[3] Milletari, F., Navab, N., & Ahmadi, S. (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. 3DV.