Grounding Line Detection in DInSAR Images using Deep Learning

Davide Gattolin (885797)

Image & Video Understanding Project

Ca’ Foscari University of Venice

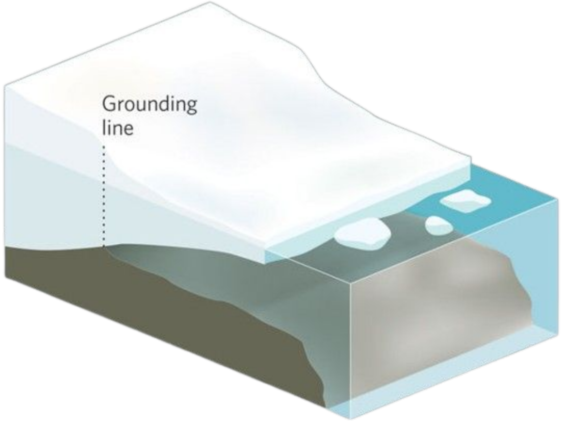
Why Grounding Lines Matter

They have an impact on:

Ice sheet stability & Ocean-ice interactions

-

Ice melting is strictly correlated to climate change



From Geophysical Task to Computer Vision Problem

Input:

- DInSAR satellite images (Phase + magnitude channels)
- Hand-made binary-mask as label

Output:

- Binary grounding line segmentation mask

This task is classified as:

- Semantic segmentation

Previous Approaches

Mohajerani et al. (2021)

- First CNN-based GL detector (U-Net)
- Encoder-decoder CNN with ASPP
- Weighted BCE loss

Ross et al. (2024)

- Extension to X-band SAR
- Higher spatial resolution

Goals of the project

Study effect of patch size on segmentation performance

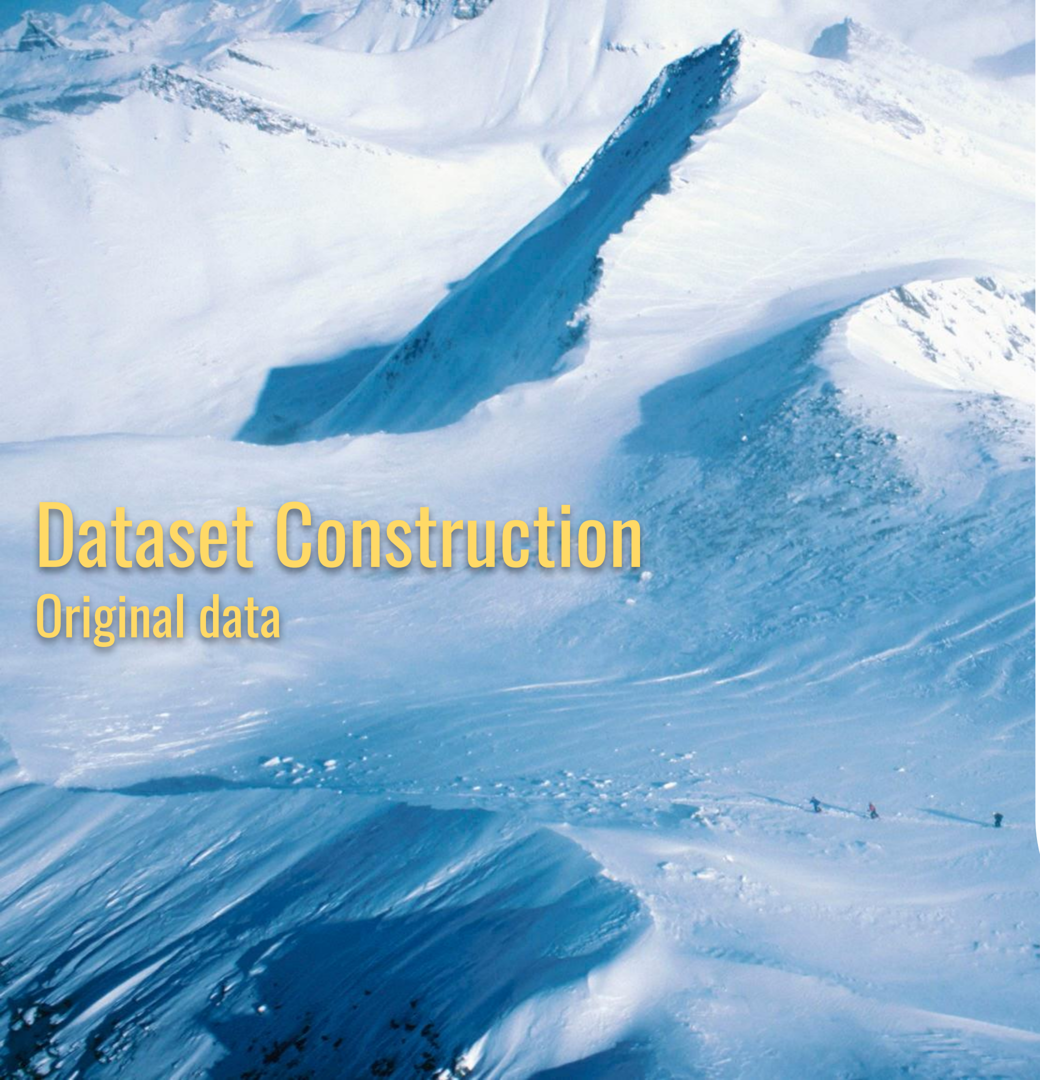
-

Design compound loss for severe class imbalance

-

Compare advanced U-Net-based architectures:

- U-Net
- Attention U-Net
- U-Net++



Dataset Construction

Original data

Data Source:

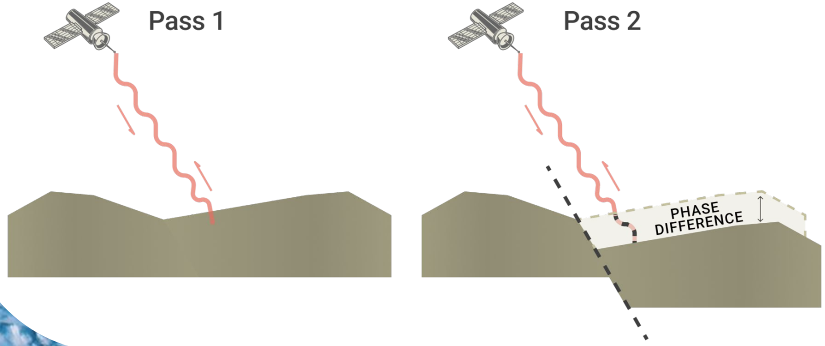
- Sentinel-1 DInSAR imagery from the Getz Ice Shelf region

Preprocessing, Complex signal decomposed into:

- Phase & Magnitude

Ground Truth Creation

- Grounding line geometries rasterized using the “All touched” rasterization to reduce label sparsity



Dataset Construction

Patch Sampling Strategy

Two datasets created:

- 128 × 128 patches
- 256 × 256 patches

Sampling design:

- Ensure spatial diversity
- Coverage map to reduce redundancy
- Mix of:
  - Patches containing GL
  - GL-less patches

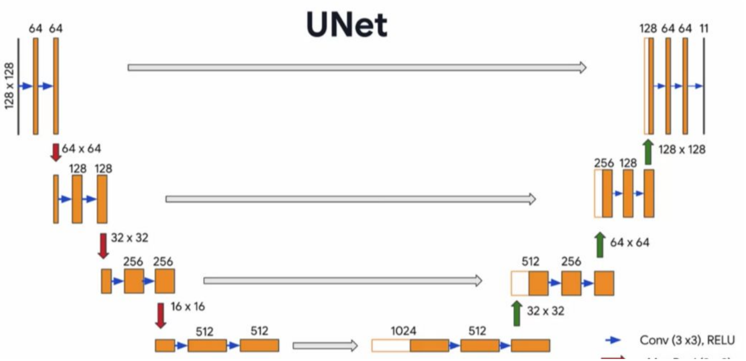
The NN architectures:
U-Net

U-Net:

- Encoder–decoder convolutional architecture
- Symmetric structure with skip connections
- Encoder extracts hierarchical spatial features
- Decoder progressively reconstructs spatial resolution

Why U-Net:

- Good balance between model complexity and generalization
- Provides reference performance for comparison with enhanced architectures
- Proven effectiveness in segmentation of thin structures



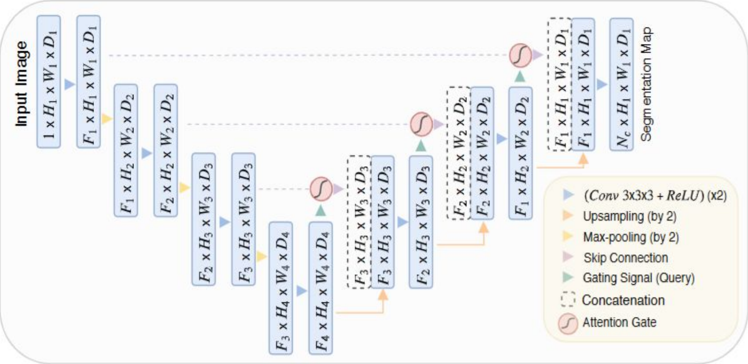
The NN architectures:
Attentional U-Net

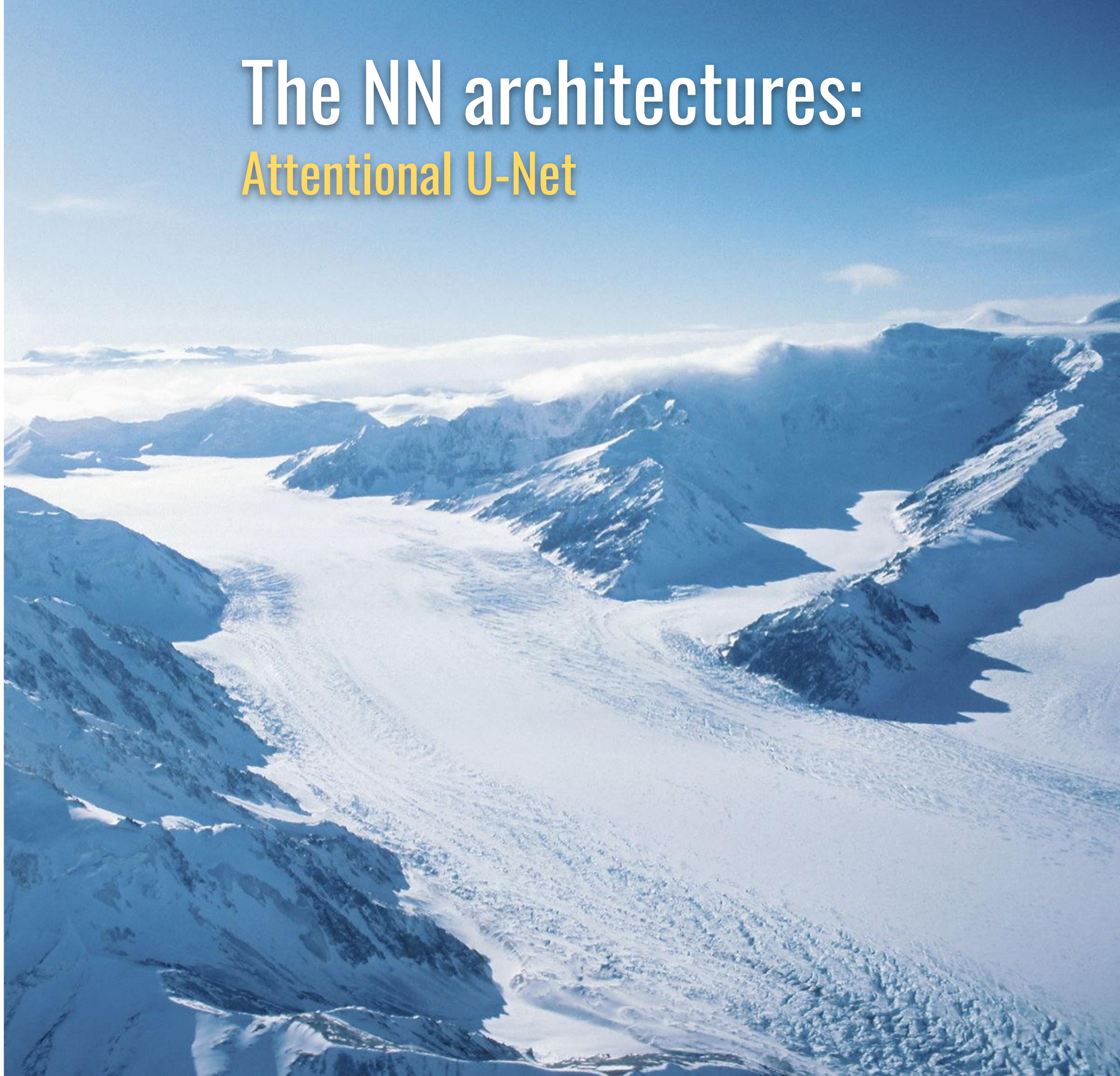
Attention U-Net

- Attention modules inserted into skip connections
- Feature maps are spatially weighted based on relevance to the target class
- Suppresses irrelevant background features before feature fusion

Why Att U-Net:

- Grounding lines are sparse and embedded in long-distance patterns
- Attention helps focus learning on relevant spatial patterns
- Improves detection of weak and noisy signal regions
- Expected to reduce background interference





The NN architectures:
U-Net++

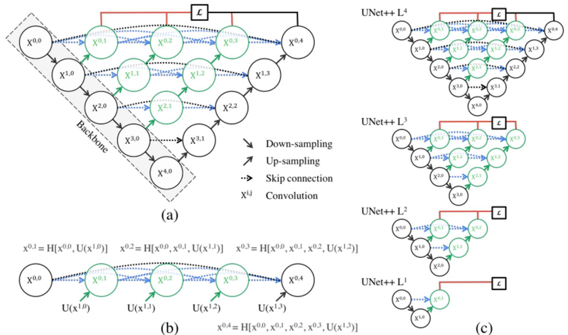
U-Net++:

- Dense skip connections reduce semantic gap between encoder and decoder features
- Multi-scale feature aggregation improves spatial detail reconstruction
- Deep supervision introduces auxiliary outputs at multiple decoder depths, enhancing gradient flow and stabilizes training

Why U-Net++:

- Multi-scale fusion improves localization of thin elongated structures
- Deep supervision encourages discriminative feature learning
- Higher representational capacity allows better modeling of complex spatial relationships

UNet++: A Nested U-Net Architecture



Loss Function

Given the peculiarities of the GLs structure a custom loss was designed, composed of:

Binary Cross Entropy (BCE)

- Provides stable pixel-wise probabilistic supervision
- Encourages global convergence during early training
- Penalizes misclassification uniformly across the image

Focal Tversky (FTL):

- Explicit control of False Negatives vs False Positives
- Increased sensitivity to small foreground objects
- Focal modulation emphasizes hard-to-classify pixels

The two different losses are combined as follows:

$$\mathcal{L} = \lambda \ \mathrm{BCE} + (1-\lambda) \ \mathrm{FTL}$$

Parameters are selected via validation ablation study.

Training Setup

Optimizer: Adam

-

Learning rate: 1e−4

-

Early stopping on validation loss (max of 500 epochs)

-

Patch-based training

-

Hyperparameter sweeps on:

- FTL parameters
- BCE weighting
- Deep supervision weights

Evaluation Metrics

Given the GL's structure and sparsity the choice of robust metrics for evaluating it is very important.

Overlap-based metrics

- Precision
- Recall
- Dice coefficient
- Tolerant IoU

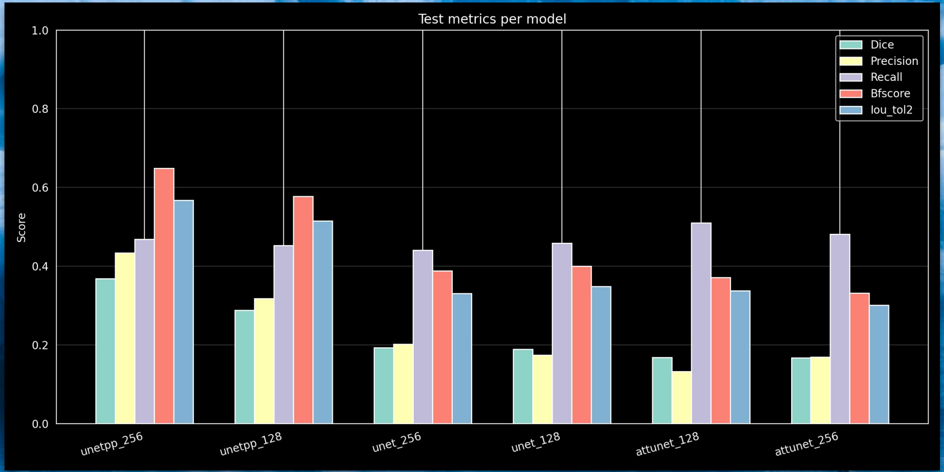
Boundary-based metrics

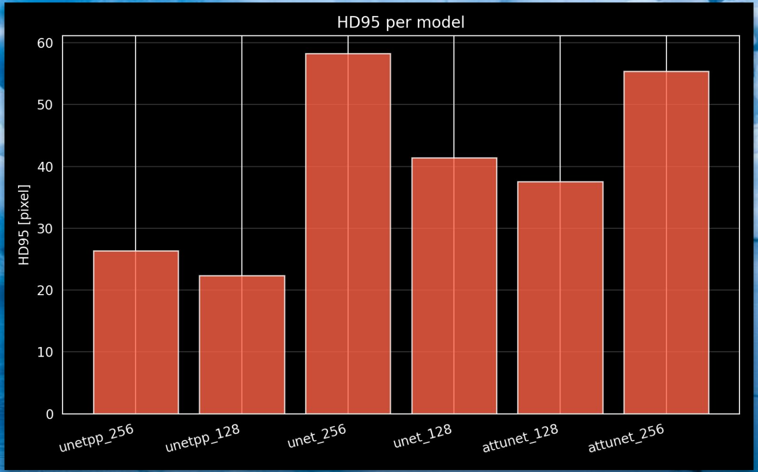
- Boundary F-score
- HD95 (Hausdorff Distance 95%)

Decision threshold calibrated via validation Dice maximization.

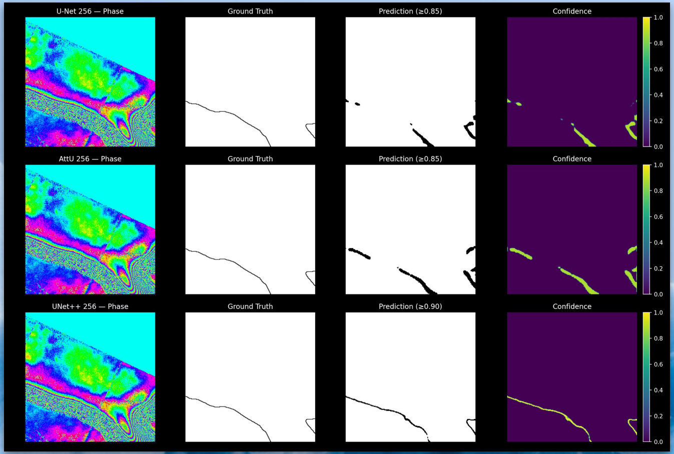
Results

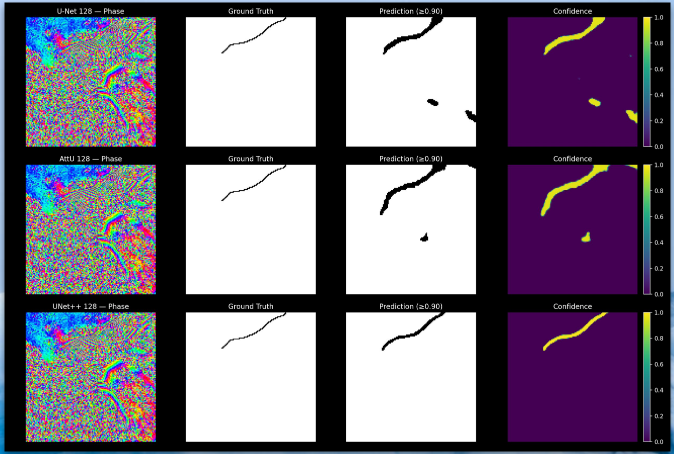
Model	Dice	Precision	Recall	BFScore	IoU_{tol2}	HD95	Threshold
UNet 128	0.19	0.17	0.46	0.40	0.35	41.37	0.90
UNet 256	0.19	0.20	0.44	0.39	0.33	58.25	0.85
AttUNet 128	0.17	0.13	**0.51**	0.37	0.34	37.53	0.90
AttUNet 256	0.17	0.17	0.48	0.33	0.30	55.41	0.85
UNet++ 128	0.29	0.32	0.45	0.58	0.52	**22.37**	0.90
UNet++ 256	**0.37**	**0.43**	0.47	**0.65**	**0.57**	26.38	0.90

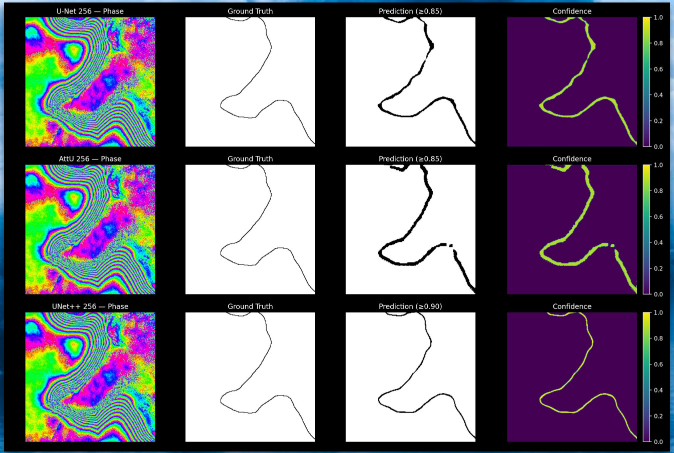


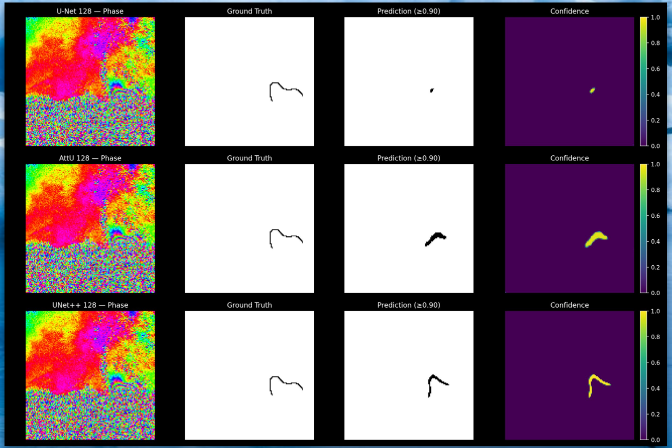


Results

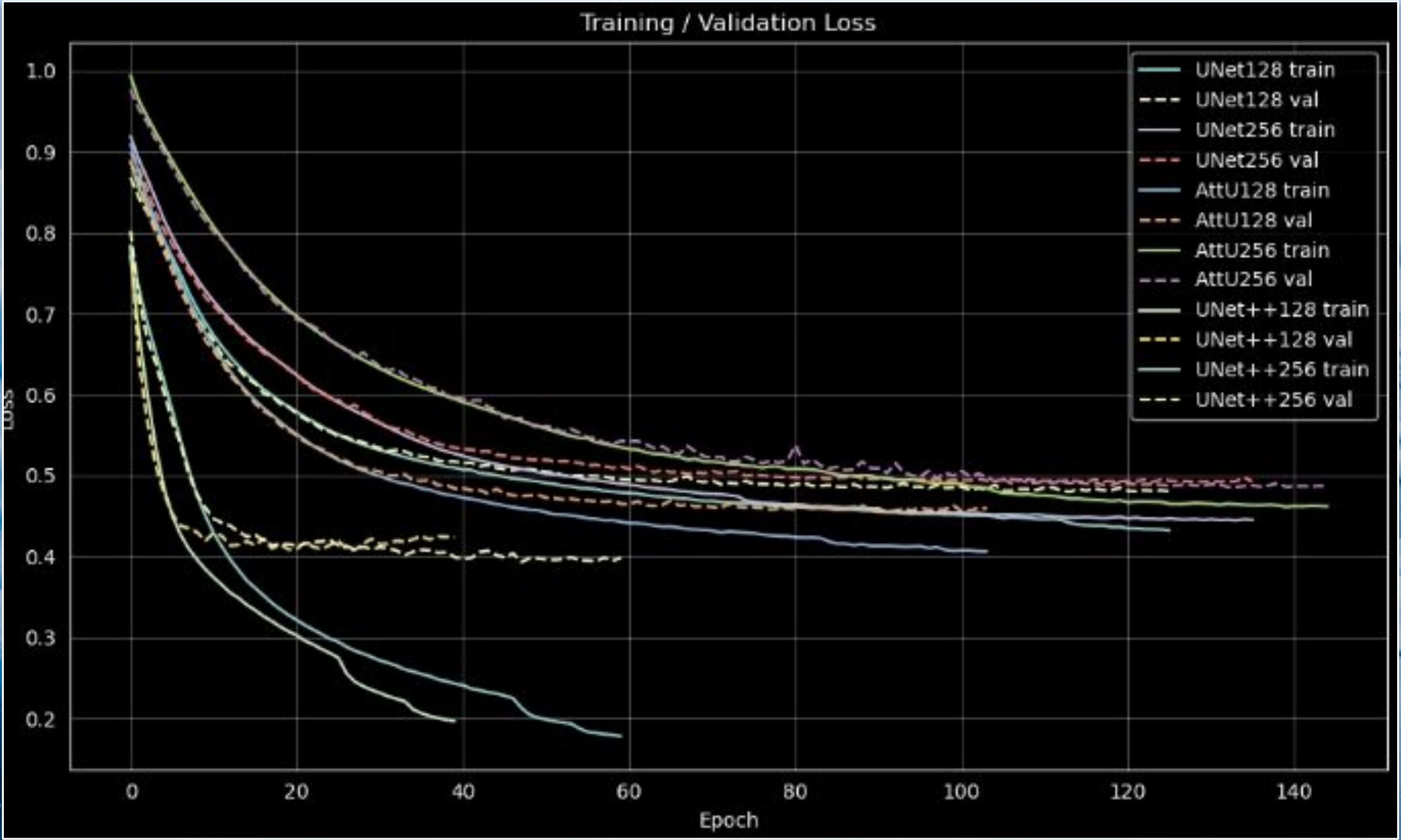

U-Net 256 — Phase
Ground Truth
Prediction (≥0.85)
Confidence
AttU 256 — Phase
Ground Truth
Prediction (≥0.85)
Confidence
UNet++ 256 — Phase
Ground Truth
Prediction (≥0.90)
Confidence


U-Net 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence
AttU 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence
UNet++ 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence


U-Net 256 — Phase
Ground Truth
Prediction (≥0.85)
Confidence
AttU 256 — Phase
Ground Truth
Prediction (≥0.85)
Confidence
UNet++ 256 — Phase
Ground Truth
Prediction (≥0.90)
Confidence


U-Net 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence
AttU 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence
UNet++ 128 — Phase
Ground Truth
Prediction (≥0.90)
Confidence

Results



Training / Validation Loss
Loss
Epoch
0
20
40
60
80
100
120
140
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
UNet128 train
UNet128 val
UNet256 train
UNet256 val
AttU128 train
AttU128 val
AttU256 train
AttU256 val
UNet++128 train
UNet++128 val
UNet++256 train
UNet++256 val

Results - Interpretation

U-Net:

- Prone to hallucinated detections in heterogeneous glaciological textures
- Limited responsiveness to extremely thin and low-contrast grounding line structures
- Difficulty preserving long-range spatial continuity

Attention U-Net:

- Enhanced grounding line detection through adaptive spatial feature selection
- Improved sensitivity to weak interferometric signals
- Tendency toward mild over-segmentation due to aggressive foreground enhancement

U-Net++:

- Superior geometric reconstruction of grounding line morphology
- Effective multi-scale feature integration via nested skip pathways
- Improved boundary precision and structural continuity
- Best overall trade-off between detection accuracy and spatial consistency

Results - Interpretation

Training vs validation loss trends show:

U-Net++:

- highest training performance
- stronger overfitting

Attention U-Net:

- best balance between capacity and generalization

U-Net:

- most stable but lowest accuracy

Future Work

Expand dataset diversity (Anderson’s or Ross’ datasets) and change the patches selection strategy

Investigate phase-only relevance

Introduce label smoothing strategies

Explore transformer-based segmentation models