Session 1: Semantic Segmentation with U-Net

Advanced Deep Learning for Pixel-Level Analysis

Stylianos Kotsopoulos
EU-Philippines CoPhil Programme

Session Overview

Duration: 1.5 hours
Type: Theory + Discussion
Goal: Understand semantic segmentation and U-Net for EO

You will learn: - What semantic segmentation is and why it matters for EO - U-Net architecture: encoder, decoder, skip connections - Losses for segmentation: CE, Dice, IoU - EO applications: floods, land cover, roads, buildings

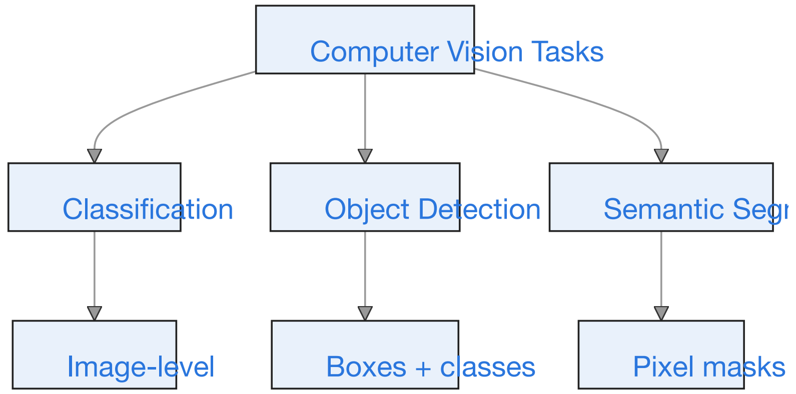
Prerequisites: - Day 2 Session 3 (CNNs) - Basic Python/Colab

Resources: - Hands-on lab in Session 2 (flood mapping)

Concept of Semantic Segmentation

Pixel-wise classification vs other CV tasks

- Classification → single label per image
- Object detection → boxes + labels per instance
- Semantic segmentation → label for every pixel



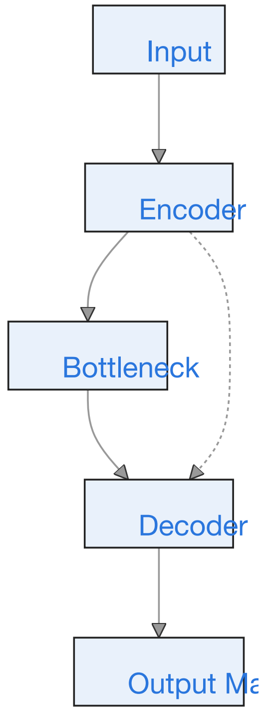
Why segmentation for EO?

- Precise boundaries (flood edges, building footprints)
- Accurate area calculations (km^2 flooded)
- Enables fine-grained change detection

U-Net Architecture

High-level overview

- Encoder (contracting path): features ↑, resolution ↓
- Bottleneck: global context
- Decoder (expansive path): upsample + refine
- Skip connections: fuse detail from encoder into decoder



Skip connections: the key

- Preserve high-resolution details lost during pooling
- Concatenate encoder features with upsampled decoder features
- Sharper boundaries, better small structures (roads, levees)

Encoder details

- Blocks: 2× Conv(3×3, ReLU) → MaxPool(2×2)
- Spatial dims halve, channels double
- Multi-scale feature hierarchy (edges → textures → semantics)

Decoder details

- Upsample via TransposedConv or Bilinear + Conv
- Concatenate with corresponding encoder features (same size)
- 1×1 Conv at end to produce class logits (per pixel)

Loss Functions for Segmentation

Imbalance is common in EO

- Flood pixels << background pixels
- Buildings/roads are sparse

Options

- Pixel-wise Cross-Entropy (with class weights)
- Dice loss (overlap-focused)
- IoU/Jaccard loss
- Combined loss (BCE + Dice) → robust baseline

Practical recommendation

Use Combined (BCE + Dice) for flood/building tasks with imbalance.

EO Applications

Flood Mapping (DRR)

- Sentinel-1 SAR (VV/VH) → binary flood masks
- U-Net excels at delineating flood boundaries in dark backscatter regions

Land Cover Mapping (NRM)

- Multi-class masks from Sentinel-2 bands/indices
- Precise boundaries between forest / agriculture / urban

Roads & Buildings (Urban)

- Thin linear features and compact footprints benefit from skips

Implementation Notes

- Input sizes: 256×256 or 512×512 chips
- Normalization: SAR (dB scaling), Optical (0–1)
- Augmentation: flips/rotations; ensure image & mask transformed identically
- Metrics: Dice, IoU, Precision/Recall (minority class focus)

Summary & Prep for Lab

Key takeaways

- U-Net = encoder–decoder + skip connections → precise masks
- Overlap-focused losses improve minority-class performance
- Perfect fit for EO segmentation tasks (floods, land cover, roads)

Next (Session 2)

- Hands-on: Train U-Net flood mapper on Sentinel-1 SAR
- Evaluate with IoU/Dice; export masks for GIS