Aerial Road Segmentation in the Presence of Topological Label Noise

Corentin Henry
German Aerospace Center (DLR)
Oberpfaffenhofen, Germany
corentin.henry@dlr.de

Friedrich Fraundorfer
Institute of Computer Graphics and Vision
Graz University of Technology (TUG)
Graz, Austria

Eleonora Vig
Amazon
Berlin, Germany

High-resolution aerial images are expensive to annotate
- Large scale or high definition: from 1m/pixel to 5cm/pixel
- Many objects: cars, trees, parking places
- Small details: lane-markings, walls, danger-areas
- Frequent occlusion: vegetation, buildings, shadows

Roads are complex objects
- Types diversity: streets, highways, dirt paths
- Topology complexity: shape, width, connectivity
- Similarity to other objects: parking lots, sidewalks, bikeways

Large-scale datasets trade-off quality for quantity
- With inconsistencies from OpenStreetMap [1]
- With inaccurate polylines [2]
- With incomplete ground truths [3]

Examples of label noise and its effect on predictions
- (a-f) Triplets of RGB, confusion and probability maps of our baseline model
  - RGB: 1m/pixel satellite images from [1]
  - Confusion Maps: true positives, false negatives, false positives
  - Probability Maps: pixel-wise detection of roads in 0-100%

Labeling Roads in Aerial Images

Noise-Aware Loss Functions

Predictions are often more consistent and better localized than labels
- Injecting predictions as ground truth in losses using bootstrapping [7]

\[
BCE(y, p) = - \sum_{k} y_k \log(p_k)
\]

\[
SoftBootBCE(y, p) = - \sum_{k} [\beta y_k + (1 - \beta) p_k] \log(p_k)
\]

Reinforcing Noise-Resilience

Introduction of synthetic noise as data augmentation during training
- Improving the resistance towards label errors
- Using different noise types with gradual amplitudes:

Measuring Road Segmentation Quality

Pixel-wise accuracy
- IoU, F1-DICE, Precision, Recall

Topological accuracy
- Completeness, Correctness, Quality [8]

Experiments on Road Datasets

Results on noisy datasets: Massachusetts Roads [1] and DeepGlobe [3]

Qualitative prediction improvements on MA (top) and DG (bottom)

The training labels are not entirely trustworthy

Extracting Roads in Aerial Images

Fully-convolutional neural networks (FCNNs)
- Fine-grained segmentation: need for dedicated architectures
- U-Nets [4] are state-of-the-art
- Shallow architectures used for fast training in challenges
- Deep architectures used to leverage large-scale data


Diagram of the Dense-U-Net-121 architecture

References

[2] SpaceNet on Amazon Web Services (AWS). Datasets. The SpaceNet Catalog, 2018