Road segmentation made simple: A practical comparison of segmentation models and post-processing techniques

Júlia Alves Porto, Apostolos Ziakopoulos and George Yannis

National Technical University of Athens - NTUA

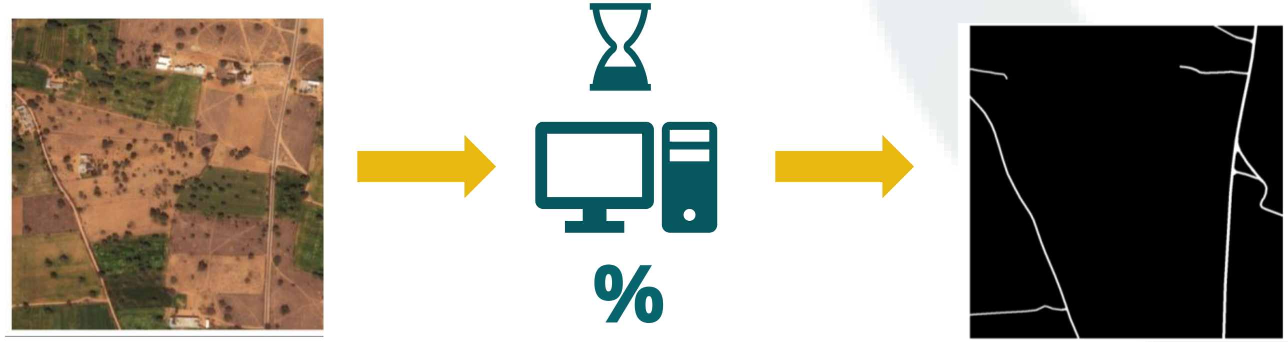
European Commission

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101119590.

IVORY

Introduction

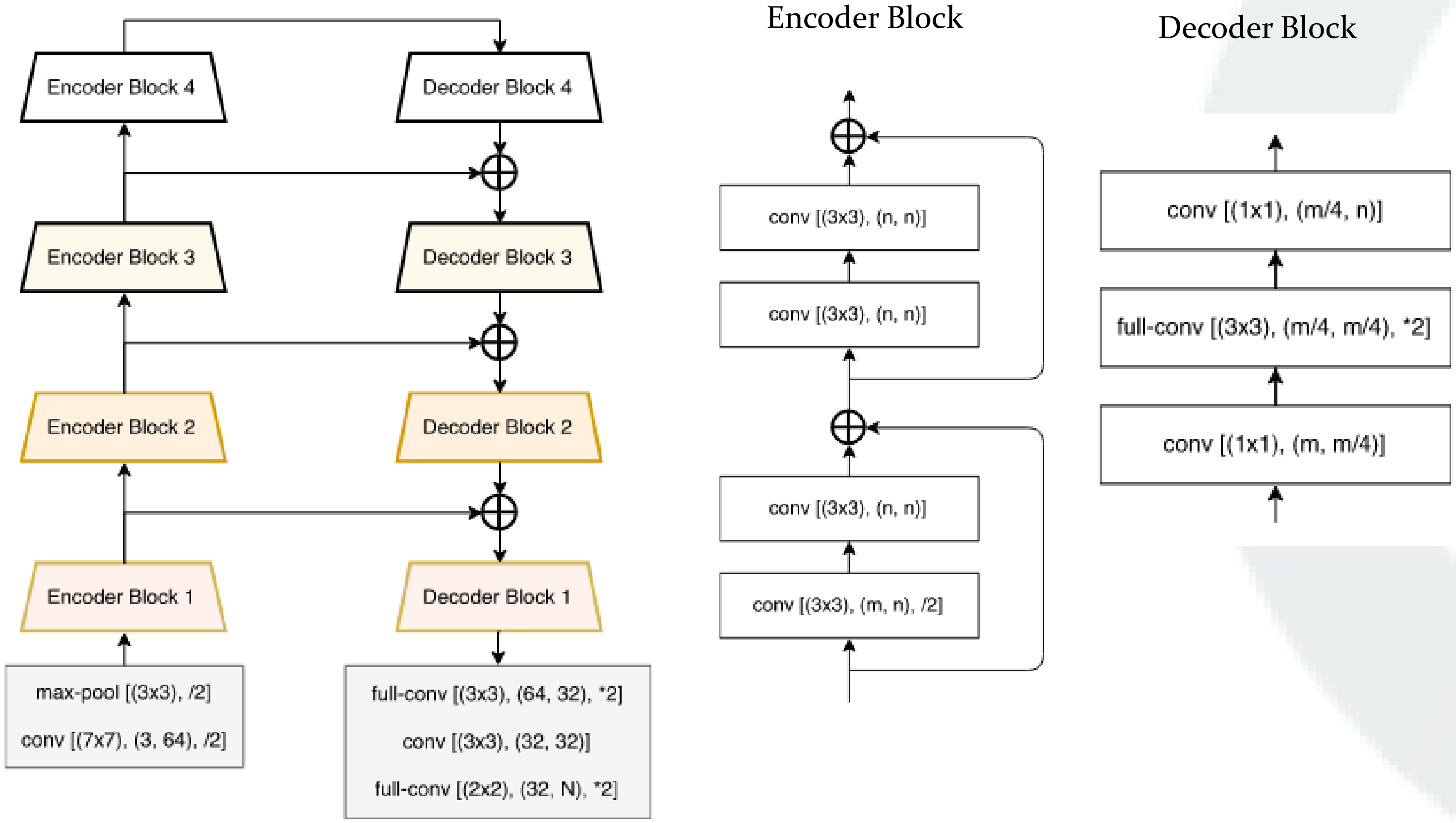
- Applications of Road Segmentation: mapping remote areas, updating existing maps, analyzing road geometry
- Popular models for road segmentation rely on graph convolutional network, deep neural network etc.
- Research objective: provide a practical comparison of open-access and high performing models and post-processing techniques cited in the literature



European Commission

Selected models

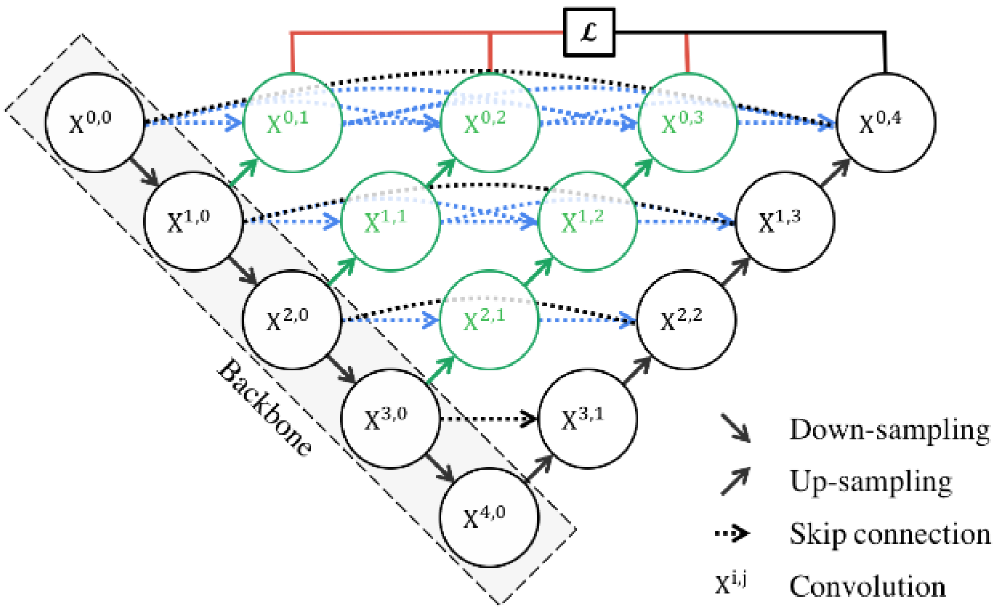
- LinkNet (Chaurasia & Culurciello, 2017)



- Hyperparameters:
  - 4 encoder-decoder blocks
  - Use Batchnorm normalization
  - Base number of channels: [1024, 512]
  - ResNet-18 as encoder
  - No pre-trained weights

Selected models

- Unet++ (Zhou et al., 2018)



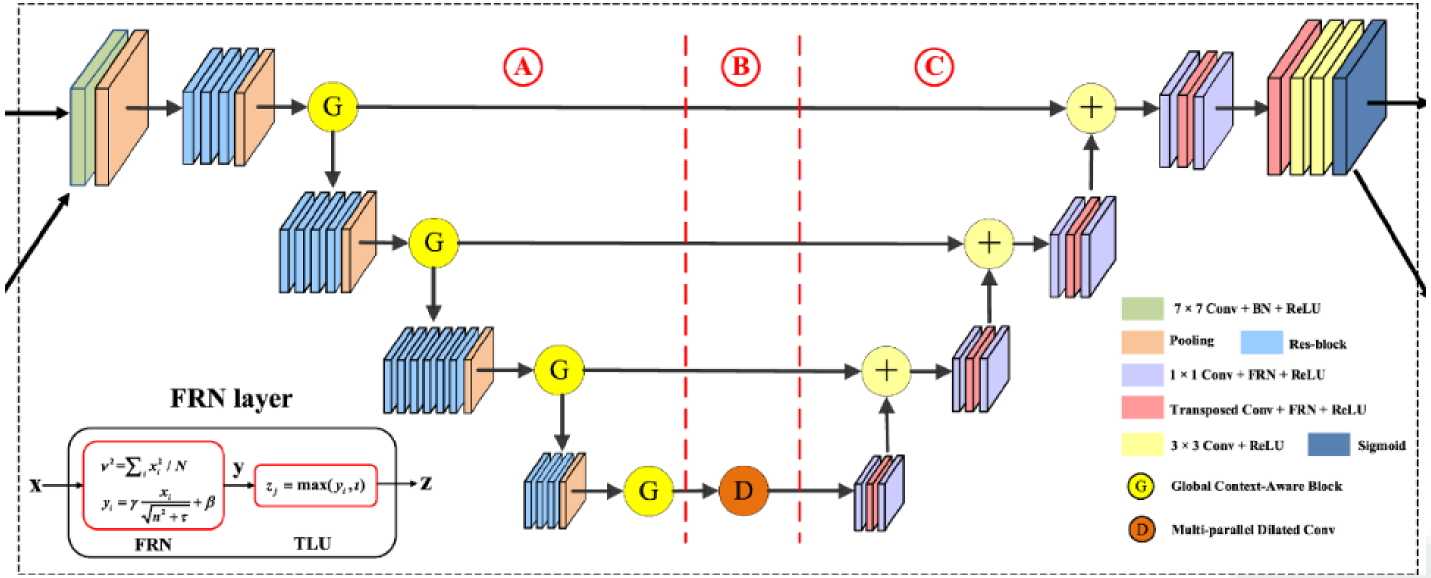
- Hyperparameters:
  - 5 depth layers
  - Use Batchnorm normalization
  - Base number of channels: [1024, 512]
  - ResNet-18 as encoder
  - No pre-trained weights

European Commission

IVORY

Selected models

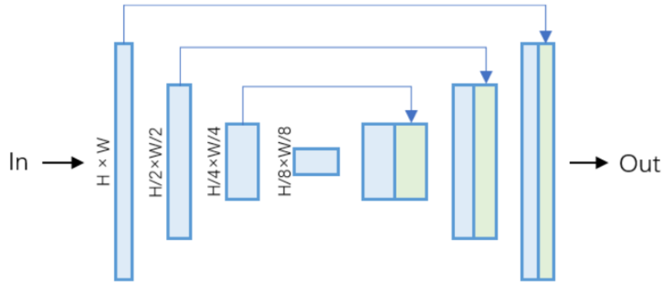
- GCB-Net (Zhu et al., 2021)

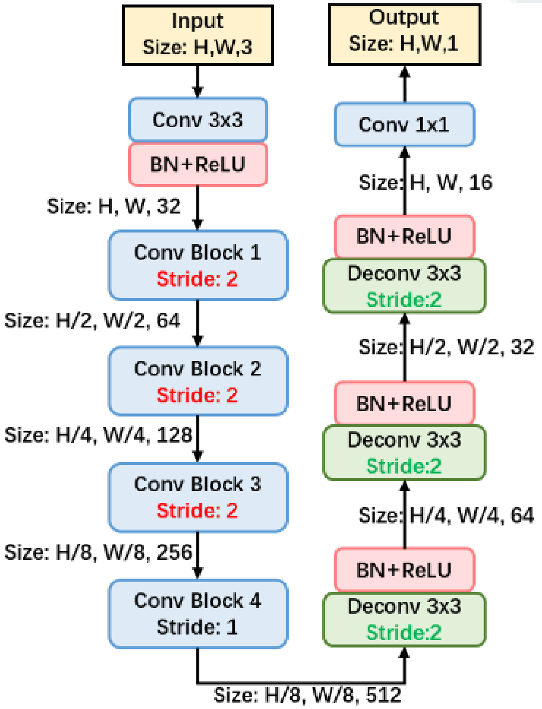


- Hyperparameters:
  - Base number of channels: [1024, 512]
  - Standard kernel size: [7, 3]

Selected models

- DiResSeg (Ding and Bruzzone, 2021)





- Hyperparameters:
  - Base number of channels: [1024, 512]
  - Standard kernel size: [7, 3]

Training and metric parameters

- DeepGlobeChallenge Dataset (Demir et al., 2018)
- Adam optimizer with 0.001 learning rate
- Augmentation: 90° flipping, blurring, brightness contrast
- Jaccard Loss: $1 - \frac{|P \cap T|}{|P \cup T|} = 1 - \frac{\sum(P \cdot T)}{\sum(P + T - P \cdot T) + \varepsilon}$
- F1-Score, IoU, Accuracy and Precision
- Early stop for 5 consecutive epochs with no improvement on IoU, 24 hours or 100 epochs
- Training on GPU NVIDIA GeForce RTX2080
- Dataset pre-processing: None, Resize, Crop

IoU

$$\frac{2 \times TP}{2 \times TP + FP + FN}$$

Precision

$$\frac{TP}{TP + FP}$$

Recall

$$\frac{TP}{TP + FN}$$

F1

$$2 \times \frac{Precision \cdot Recall}{Precision + Recall}$$

European Commission

IVORY

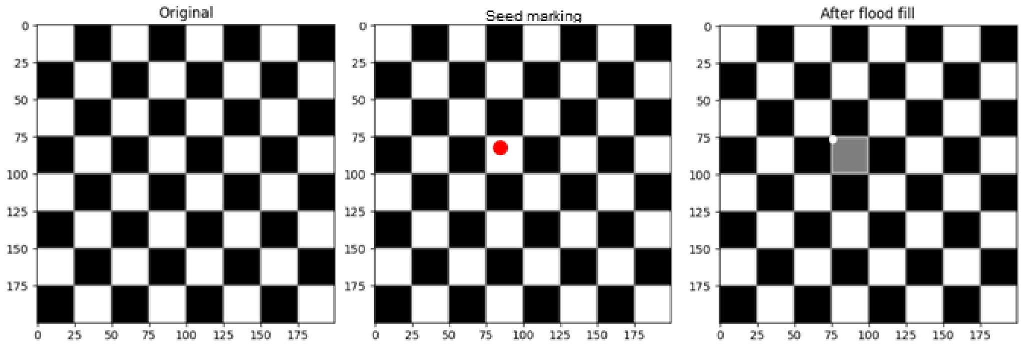
Results

Method	Pixel Size	Model	Kernel Size	Training Dataset Size	Training Time (min)	Total epochs	End of training	IoU	Precision	Recall	F1 Score	Flexible IoU (3 pixels)
None	1024	GCB	7	4604	-	-	24h Time Out	-	-	-	-	-
Resize	512	GCB	7	4604	962	26	No improvement	0.3063	0.3372	0.8001	0.4553	0.8194
Crop	512	GCB	7	6506	1313	27	No improvement	0.3984	0.4163	0.9100	0.5551	0.9154
None	**1024**	**GCB**	**3**	**4604**	**709**	**36**	**No improvement**	**0.4474**	**0.4631**	**0.9332**	**0.6045**	**0.9395**
Resize	512	GCB	3	4604	140	23	No improvement	0.3401	0.3912	0.7484	0.4883	0.7640
Crop	512	GCB	3	15048	1599	10	24h Time Out	0.4175	0.4402	0.9020	0.5734	0.9104
None	1024	DiResSeg	3	4604	206	21	No improvement	0.4585	0.6419	0.6335	0.6129	0.6558
Resize	512	DiResSeg	3	4604	62	11	No improvement	0.2613	0.5346	0.3487	0.3913	0.3769
Crop	512	DiResSeg	3	15093	288	34	No improvement	0.4798	0.6600	0.6519	0.6341	0.6628
None	**1024**	**DiResSeg**	**7**	**4604**	**1114**	**48**	**No improvement**	**0.5736**	**0.7404**	**0.7529**	**0.7160**	**0.7406**
Resize	512	DiResSeg	7	4604	1560	10	24h Time Out	0.3124	0.6541	0.3813	0.4511	0.4087
Crop	512	DiResSeg	7	15085	1626	8	24h Time Out	0.4751	0.5627	0.7709	0.6297	0.7809
None	1024	LinkNet	4	4604	216	14	No improvement	0.4084	0.4557	0.8342	0.5654	0.8323
Resize	512	LinkNet	4	4604	126	23	No improvement	0.4514	0.6667	0.6039	0.6010	0.6149
Crop	**512**	**LinkNet**	**4**	**15119**	**258**	**28**	**No improvement**	**0.5895**	**0.7563**	**0.7417**	**0.7285**	**0.7792**
None	1024	U-Net++	3	4604	-	1	24h Time Out	0.1514	0.7003	0.1727	0.2300	0.5271
Resize	512	U-Net++	3	4604	455	28	No improvement	0.3598	0.7100	0.4297	0.4955	0.4637
Crop	**512**	**U-Net++**	**3**	**15057**	**1917**	**2**	**24h Time Out**	**0.4167**	**0.5061**	**0.7369**	**0.5716**	**0.7437**

European Commission

Post-processing techniques

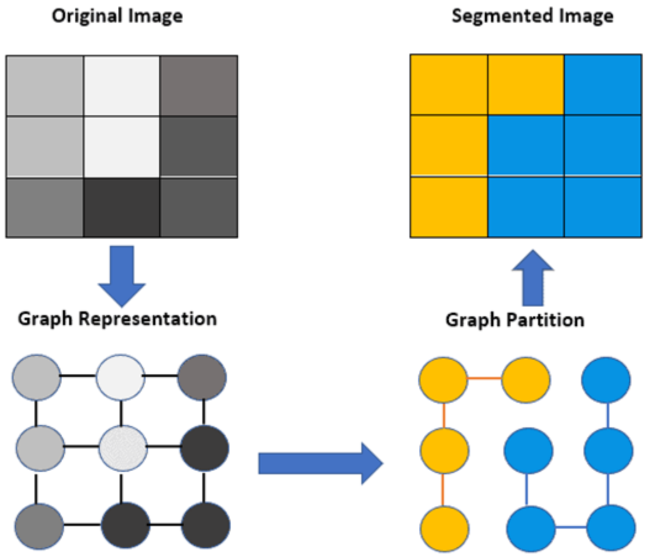
Region growth



- Gaussian smoothening $\sigma = 2$
- 2 random road pixels per quadrant, if any
- Threshold = 0.3
- Flood fill from Scikit-image library

Post-processing techniques

Graph-based segmentation

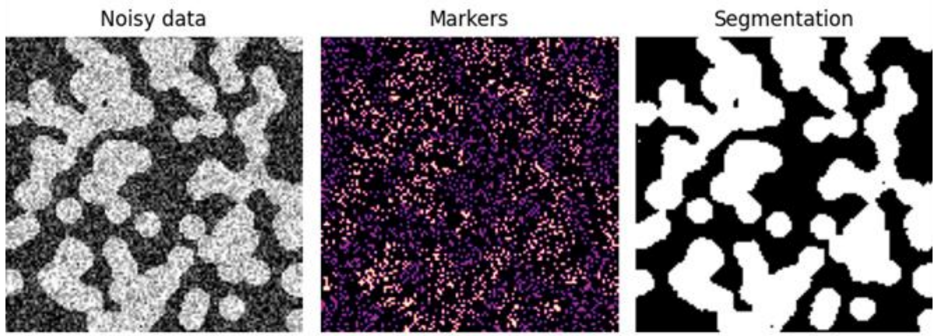


- Guided by Canny edge detection
- Using NetworkX and Scikit-image
- Build a graph from the binary image based on 4-connected neighbor pixels
- Component size threshold
- Intersection with Canny-edge pixels threshold
- Morphological closing with a 5x5 kernel

European Commission

Post-processing techniques

Energy-based segmentation



- Conditional Random Fields (Krähenbühl and Koltun, 2011)
- Markers: < 0.1 or > 0.9
- Walker guide: 80% of most prominent edges from original input and predicted probability map (Farid and Simoncelli, 2004)
- Random walker from Scikit-image

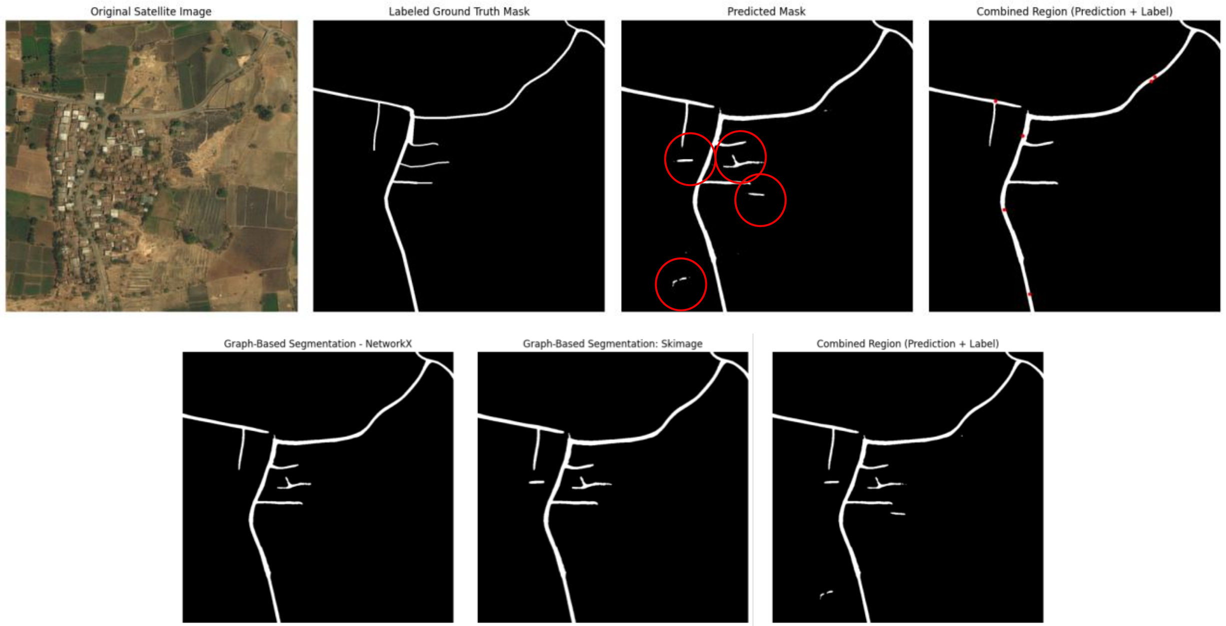
Results

Segmentation Model	Post-Processing Technique	Main Library	IoU	Precision	Recall	F1
LinkNet	-	-	0.5895	0.7563	0.7417	0.7285
LinkNet	Region growth	Skimage	0.5622	0.8086	0.6882	0.6530
LinkNet	**Graph-Based Segmentation**	**Networkx**	**0.6027**	**0.8366**	**0.7108**	**0.6933**
LinkNet	Graph-Based Segmentation	Skimage	0.4049	0.7127	0.5471	0.5071
LinkNet	Energy based	Skimage	0.5895	0.8401	0.6926	0.6811
DiResSeg	-	-	0.5736	0.7404	0.7524	0.7160
DiResSeg	Region growth	Skimage	0.5607	0.7330	0.7524	0.6484
DiResSeg	**Graph-Based Segmentation**	**NetworkX**	**0.6355**	**0.8074**	**0.7706**	**0.7266**
DiResSeg	Graph-Based Segmentation	Skimage	0.4159	0.6804	0.5770	0.5199
DiResSeg	Energy based	Skimage	0.6115	0.8155	0.7316	0.7056
GCB	-	-	0.4474	0.4631	0.9332	0.6045
GCB	**Region growth**	**Skimage**	**0.4535**	**0.4917**	**0.9054**	**0.5626**
GCB	Graph-Based Segmentation	Networkx	0.4479	0.4714	0.9358	0.5605
GCB	Graph-Based Segmentation	Skimage	0.2102	0.2988	0.4841	0.2852
GCB	Conditional Random Fields	Skimage	0.4295	0.4509	0.9374	0.5425
U-Net++	-	-	0.4167	0.5061	0.7369	0.5716
U-Net++	**Region growth**	**Skimage**	**0.4235**	**0.5752**	**0.7208**	**0.5228**
U-Net++	Graph-Based Segmentation	NetworkX	0.4169	0.5106	0.7927	0.5241
U-Net++	Graph-Based Segmentation	Skimage	0.2556	0.4034	0.5480	0.3471
U-Net++	Energy based	Skimage	0.3937	0.4779	0.7982	0.5023

European Commission

IVORY

Results



- Reduction of noise horizontally between post-processing techniques
- Unpaved roads captured by the model while not present in the training labels were not excluded by post-processing

European Commission

IVORY

Discussion and Conclusion

- Region Growth reduces noise, at the expense of removing some True Positive disconnected from the main road
- Graph-based segmentation can be very helpful for road segmentation given the graph-like structure of a road, as suggested by the growing use of Graph Neural Networks for road segmentation (Lian et al., 2022)
- Limitations: only one dataset, no pre-trained weights, strict early termination criteria, few ablation tests for models architecture

Road segmentation made simple: A practical comparison of segmentation models and post-processing techniques

Júlia Alves Porto, Apostolos Ziakopoulos and George Yannis

National Technical University of Athens - NTUA

European Commission

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101119590.

IVORY