

## A APPENDIX

This section presents additional examples of CenterlineNet in situations involving occlusions and challenges common to aerial and satellite imagery. The main paper showed skeletonized results – but we wanted to show that our pre-thinned results are already very thin. We provide both skeletonized and non-skeletonized visualizations to demonstrate how our patch-reciprocal-intersection loss function handles predictions under an added occlusion. In each of the images, a random rectangle is replaced by the mean color of the satellite image.

### A.1 CENTERLINE NET WITH PATCH-RECIPROCAL-INTERSECTION

Fig. 7 (a full-page figure) shows centerlines extracted by our model for numerous examples. Our quantitative evaluation first skeletonizes and then matches pixels to ground-truth, but that does not capture the fact that our predictions are already very thin. In these figures, no bipartite matching is done, we simply compare the ground-truth the thresholded prediction (with a threshold of 50%). Green pixels are true positives for predictions, while red pixels are false positives. We direct the reader’s attention to these observations:

- **Occlusion fill.** The network bridges many occluded gaps, though with lower certainty, visible as slightly brighter blobs in the background pixels. This is most visible in row 2 image set 1, and to some extent in row 3 image set 2
- **Split-lane awareness.** On divided highways such as row 5 set 1 marked as B with red arrow, the model outputs two parallel centerlines. It hallucinates plausible two-lane roads in the occluded area in row 6 image set 1.
- **Plausible hallucination.** When detail is completely hidden, the network invents reasonable geometry (e.g., row 6 image set 2); we regard this as a feature rather than a flaw.
- **Minor structures.** Driveways and access roads appear inconsistently, for example row 6 image set 2 in the bottom right quadrant, mirroring the uneven annotation quality in OpenStreetMap.

## B ADDITIONAL EXPERIMENTAL RESULTS

Table 5 compares model variants trained with and without occlusion. Included metrics Precision, Recall, and F1 scores. Models trained with occlusion show clear improvement, especially in complex junction detection, highlighting the benefit of occlusion-aware training.

## C FAILURE AND ABNORMAL CASES

Fig. 8 illustrates selected examples where CenterlineNet encounters difficulties in accurately extracting road centerlines. We observe the following patterns:

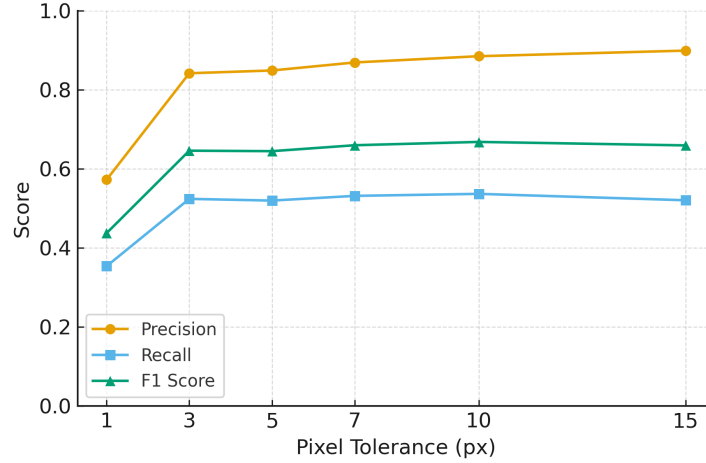
1. Occlusion near the edge of an image is more damaging - presumably, because there is not as much context.
2. Occasionally, roads are missing in our validation set, but are found by the model. These are not true failures.
3. Excessive shadows or forests that occlude the road cause failures. There is a limit to how much occlusion we can handle. This is similar to occlusion near the edges - there is not enough context to predict the occluded region.

## D NOTATION AND HYPER-PARAMETERS

Table 4 summarizes the variables and hyperparameters used in our loss functions and best saved model.

## E PIXELWISE TOLERANCE ABLATION STUDY

Fig. 6 shows various pixel spatial tolerances used to determine the practicality of pixel spatial tolerance selection. Real-world applications may require understanding models across a range of tolerances. The figure reports CenterlineNet’s precision, recall, and F1 scores as a function of the pixel matching threshold based on Centerline1M dataset. These results demonstrate how spatial tolerance parameters impact both network extraction and connectivity evaluation under varying annotation registration conditions. Sharp incline in scores from 1 to 3 pixel tolerance but not a significant impact in score increase from 3 to 15 pixels.



**Figure 6:** Precision, recall, and F1 scores of CenterlineNet as a function of pixel-matching tolerance.

**Table 4:** Complete Training and Evaluation Implementation Details. All CenterlineNet variants are implemented using the DeepLabUNetPrecise backbone with 512×512 inputs. Patch, Reciprocal, and Intersection correspond to the three components of our proposed loss.

Component	Specification
<b>Dataset and Preprocessing</b>	
Dataset	USGS–OSM satellite imagery crops (512×512)
Resolution	1 meter/pixel
Training Samples	512×512 pixel patches extracted from large aerial scenes
Normalization	ImageNet mean/std: (0.485, 0.456, 0.406) / (0.229, 0.224, 0.225)
<b>Network Architecture</b>	
Backbones	UNet, DeepLabUNetPrecise, DeepLabv3+, ResNet50+FPN
Input Size	512×512×3
Output Channels	1 (binary centerline mask)
<b>Training Hyperparameters</b>	
Optimizer	Adam
Learning Rate	5e-5
Weight Decay	1e-4
Batch Size	8–16 depending on model and GPU memory
Epochs	100 with early stopping
Loss Function	BCE (roads) + Patch Alignment + Reciprocal + Intersection Loss
<b>Patch Alignment Loss Parameters</b>	
Patch Size	11×11 pixels
Search Radius	5 pixels
$\alpha$ (False Positive Weight)	5.0
$\beta$ (Singleton Weight)	0.5
<b>Hardware and Runtime</b>	
Hardware Used	NVIDIA TITAN RTX (24GB) and GTX 1080Ti (11GB)
Framework	PyTorch 1.x
CUDA Version	11.x
Training Time	4–8 hours per model (100 epochs, depending on variant)
GPU Memory Usage	8–15GB depending on loss components



**Figure 7:** Qualitative results of occluded predictions using CenterlineNet





**Figure 8:** Failure prediction cases

**Table 5:** Bipartite Matching Metrics Grouped by Method and Occlusion.

We **bold** the best score in each row, and in case of a tie both are bolded. The second-best are underlined.

Valence & Metric	UNet		Patch		Patch-Reciprocal		Patch-Recip-Intersect	
	Non-Occ	Occ	Non-Occ	Occ	Non-Occ	Occ	Non-Occ	Occ
<b>Bipartite Matching Metrics</b>								
Overall Precision	0.51	0.51	0.29	<b>0.82</b>	0.27	0.67	0.28	<u>0.81</u>
Overall Recall	<u>0.50</u>	0.26	0.47	0.49	0.45	<b>0.54</b>	0.43	0.49
Overall F1 Score	0.51	0.37	0.36	<b>0.61</b>	0.34	0.60	0.34	<b>0.61</b>
<b>End (Valence 1)</b>								
Precision	0.02	0.00	0.01	<b>0.18</b>	0.00	0.04	0.00	<b>0.18</b>
Recall	0.07	0.00	0.02	<u>0.23</u>	0.01	0.15	0.02	<b>0.24</b>
F1 Score	<u>0.03</u>	0.00	0.01	<u>0.20</u>	0.01	0.07	0.01	<b>0.21</b>
<b>Straight (Valence 2)</b>								
Precision	0.39	0.06	0.21	<b>0.65</b>	0.21	0.42	0.20	<b>0.65</b>
Recall	0.43	0.13	0.38	0.45	0.36	<b>0.49</b>	0.36	<u>0.46</u>
F1 Score	0.41	0.09	0.28	<b>0.53</b>	0.26	0.46	0.26	<b>0.53</b>
<b>T (Valence 3)</b>								
Precision	0.04	0.00	0.01	<u>0.25</u>	0.01	0.07	0.01	<b>0.26</b>
Recall	0.23	<b>0.32</b>	0.22	0.21	0.20	<u>0.24</u>	0.20	0.22
F1 Score	0.07	0.00	0.03	<u>0.23</u>	0.02	0.11	0.02	<b>0.24</b>
<b>Intersection (Valence 4)</b>								
Precision	0.12	0.01	0.03	<b>0.46</b>	0.03	0.15	0.04	<u>0.42</u>
Recall	0.24	<b>0.34</b>	0.23	0.25	0.22	<u>0.28</u>	0.23	0.25
F1 Score	0.16	0.01	0.05	<b>0.32</b>	0.05	0.20	0.07	<u>0.31</u>
<b>Complex (Valence 5+)</b>								
Precision	0.45	0.33	0.39	<u>0.80</u>	0.35	0.53	0.33	<b>0.82</b>
Recall	0.87	<b>0.96</b>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>	<u>0.90</u>
F1 Score	0.59	0.50	0.54	<u>0.85</u>	0.51	0.67	0.49	<b>0.86</b>