Supplementary Material for CenterlineNet: Weakly Supervised Road Centerline Extraction Using Patch Alignment for Topologically Accurate Networks

Anonymous Submission

1 Additional Examples

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 satelite image.

1.1 CenterlineNet with Patch-Reciprocal-Intersection

Fig. 2 (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.

2 Additional Experimental Results

Table 2 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.

3 Failure and Abnormal Cases

Fig. 3 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.

4 Notation and Hyper-parameters

Table 1 summarizes the variables and hyperparameters used in our loss functions and best saved model.

5 Pixelwise Tolerance Ablation Study

Fig. 1 shows various pixel spatial tolerances used to to determine the practically 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 base 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.

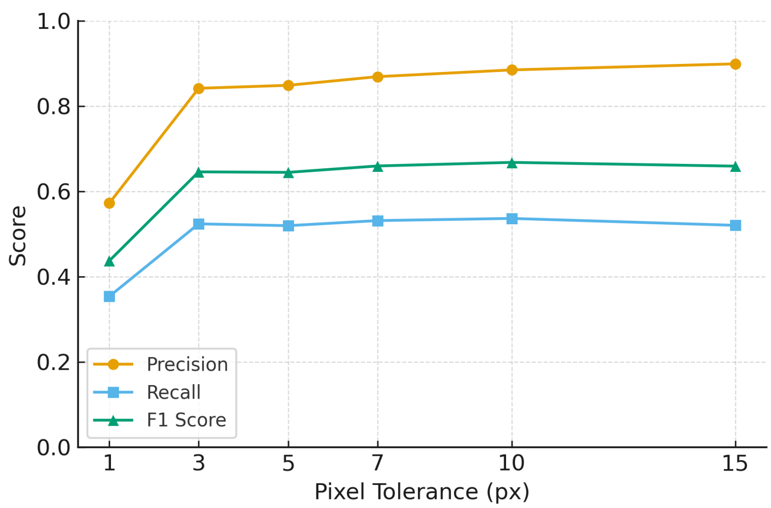

Figure 1: Precision, recall, and F1 scores of CenterlineNet as a function of pixel-matching tolerance.

Table 1: Summary of variables and hyperparameters in the loss functions.

Symbol / Parameter	Description	Default
$\mathbf{v}(\mathbf{x}) = (dx, dy)$	Displacement vector from pixel $\mathbf{x}$ to nearest ground-truth centerline	–
$\widehat{\mathbf{p}}(\mathbf{x})$	Predicted patch centered at $\mathbf{x}$	11×11
$\mathbf{p}(\mathbf{x} + \mathbf{v})$	Ground-truth patch at mapped location	11×11
$\mathcal{M}$	Label-derived road pixels within $d_{\max}$ of ground truth	$d_{\max} = 5$ px
$w_{\mathbf{x}}$	Softmax-in-group weight for pixel $\mathbf{x}$	$\tau = 0.1$
α	Weight for false-positive penalty	5.0
β	Weight for singleton penalty	0.5
r	Intersection search radius	10 px
Patch size K	Elements per patch (patch_size^2)	$K = 121$

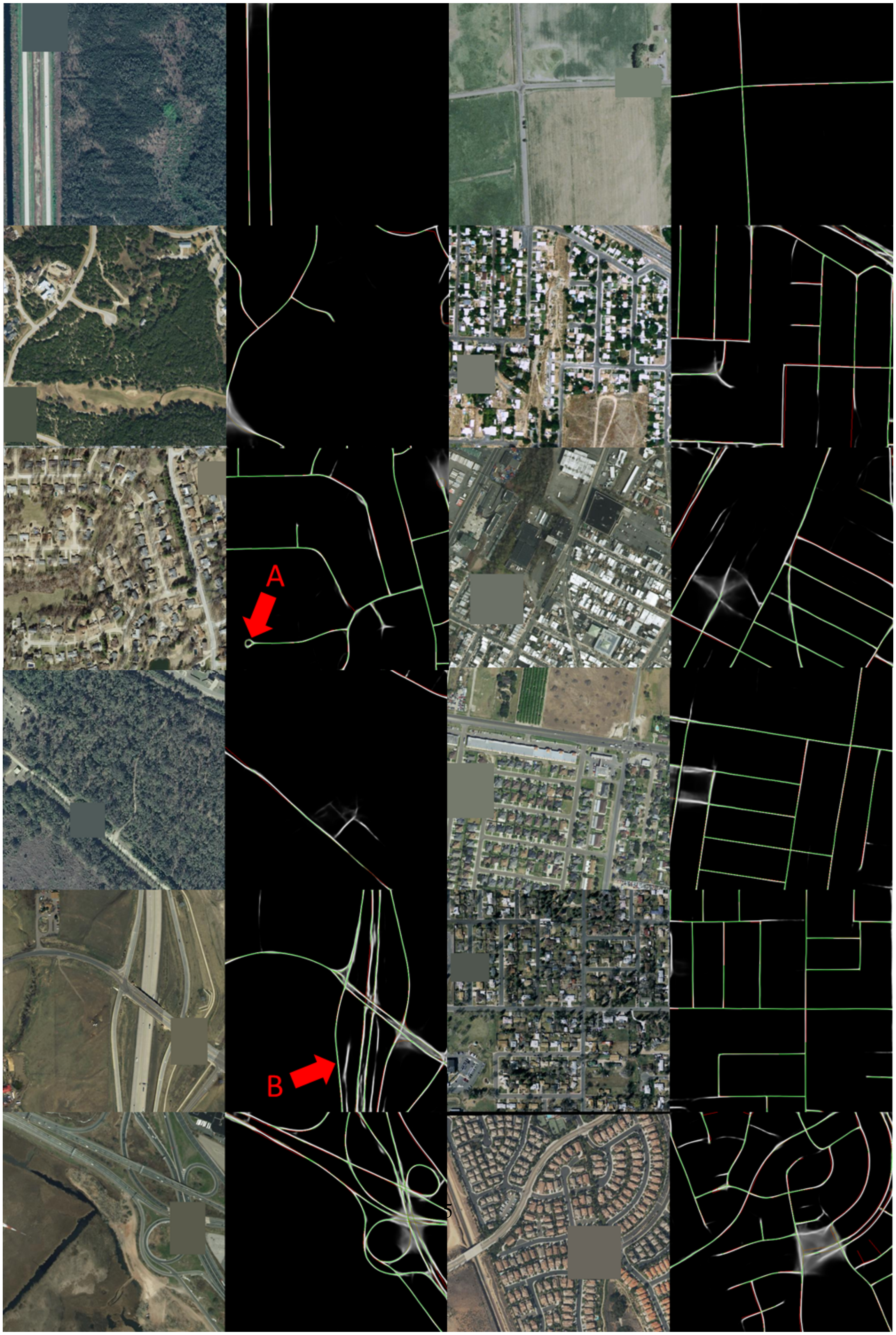

Figure 2: Qualitative results of occluded predictions using CenterlineNet

Figure 3: Failure prediction cases

Table 2: 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 <u>underlined</u>.

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