Supplemental Materials on

Semantic Line Detection Using Mirror Attention and Comparative Ranking and Matching

Dongkwon Jin, Jun-Tae Lee, and Chang-Su Kim

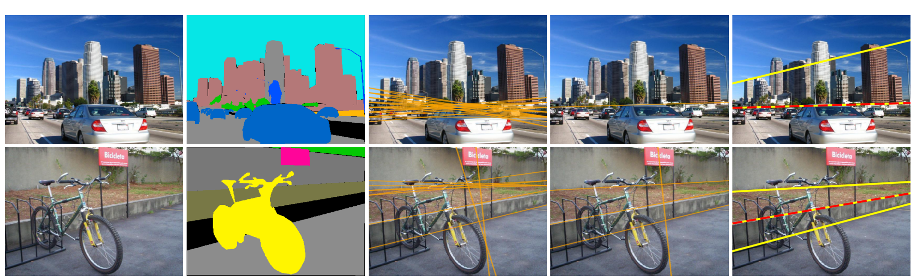
School of Electrical Engineering, Korea University, Seoul, Korea
{dongkwonjin,jtlee}@mcl.korea.ac.kr, changsukim@korea.ac.kr

A.1 SEL_Hard Dataset

A.1.1 Annotation process of SEL_Hard

We select 300 images and their segmentation labels from the ADE20K image segmentation dataset [55], as exemplified in Fig. A-1(a) and (b). For each selected image, we first generate an initial set of semantic lines using the segmentation labels automatically, and then refine the set manually.

First, we sample candidate lines as described in Section 3. For each candidate, we measure the edge score and the heterogeneity score. The edge score is the ratio of segmentation boundary pixels over all pixels on the candidate line. The heterogeneity score is the inverse of the inner product of the normalized segmentation label distributions in the two adjacent regions along the candidate line. The candidate line is initially declared as a semantic line, when both its edge and heterogeneity scores are higher than thresholds 0.2 and 2, respectively, as in Fig. A-1(c). Then, we eliminate the overlapping lines using the NMS scheme [28] with the edge scores. However, as shown in Fig. A-1(d), redundant lines are detected because of occluded regions or complex boundaries and some semantic lines (too implied or less obvious) are not detected. Therefore, we manually remove those redundant lines and add missing lines. Finally, in Fig. A-1(e), we choose the primary semantic line manually as well. Notice that SEL_Hard is constructed for testing semantic line detectors and is not used for training them. Fig. A-2 shows example images in SEL_Hard.



(a) (b) (c) (d) (e)

Fig. A-1: Annotation process of SEL_Hard

Fig. A-2: Examples images in the SEL_Hard dataset, where primary lines are depicted by dashed red ones and additional semantic lines are by yellow ones.

Table A-1: Ablation studies in terms of the mirror attention module and the comparative ranking and matching (R-Net and M-Net) on the SEL_Hard dataset.

		AUC_A	AUC_P	AUC_R
I.	D-Net(without attention)+NMS	74.65	80.62	76.56
II.	D-Net(with attention, but no flipped feature map)+NMS	76.17	79.03	75.58
III.	D-Net(with spatial-channel attention)+NMS	75.60	76.61	73.89
IV.	D-Net(with mirror attention)+NMS	76.60	82.28	77.35
V.	D-Net(with mirror attention)+R-Net+M-Net	80.68	87.19	77.69

A.1.2 Ablation studies on SEL_Hard

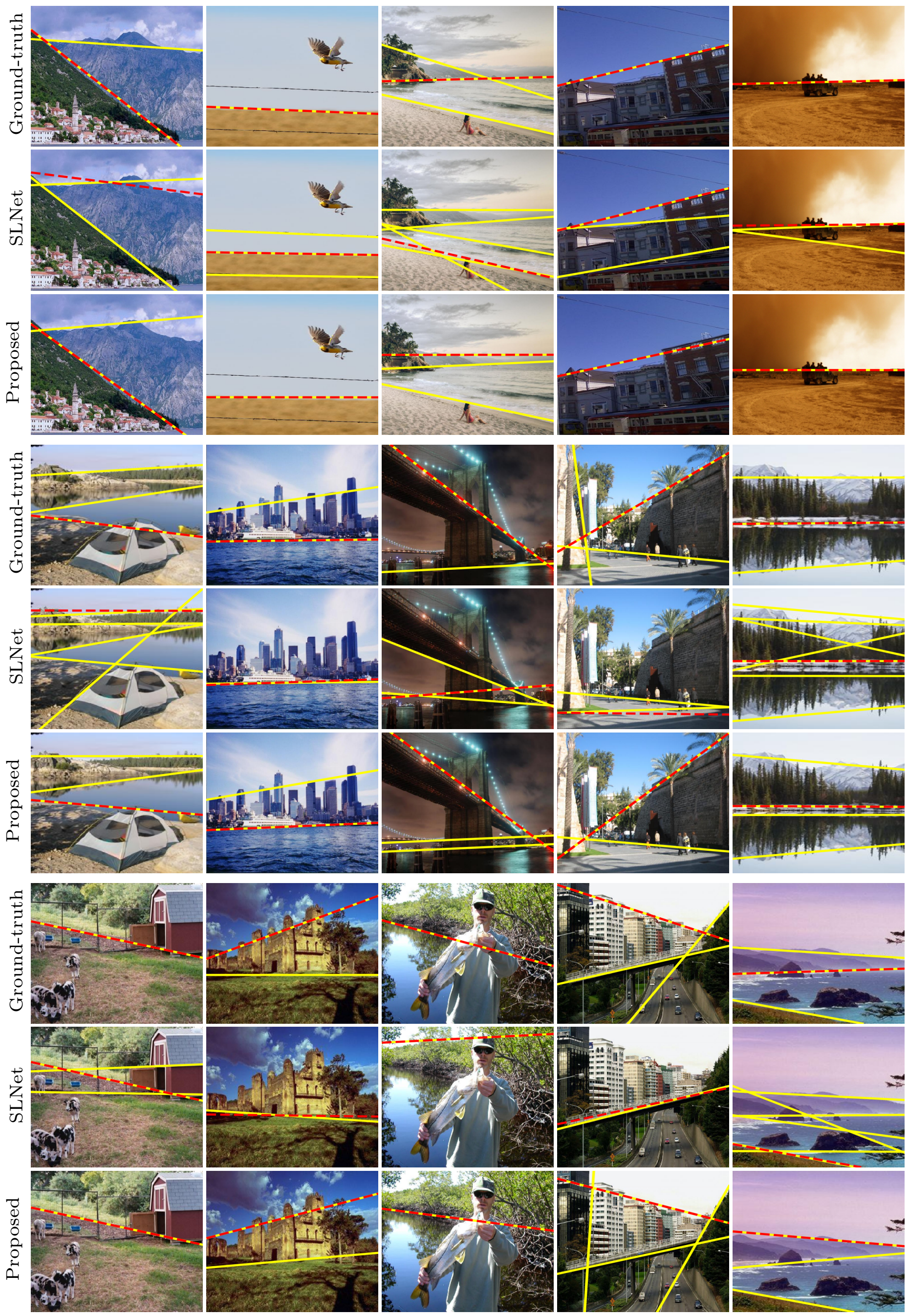
We conduct ablation studies to analyze the efficacy of the proposed D-Net, R-Net, and M-Net on the SEL_Hard dataset. Similar to Table 2 in the paper, Table A-1 compares the performances of several ablated models.

Efficacy of mirror attention model: As compared with no attention in method I, the two attention schemes in methods II and III improve the accuracy scores but degrade the precision and recall scores. In contrast, the proposed mirror attention model in method IV improves AUC_A, AUC_P, and AUC_R by 0.43, 3.25, and 1.77, respectively, compared with method II. It means that the mirroring of feature maps across semantic lines facilitates highlighting of informative regions, leading to the performance improvement.

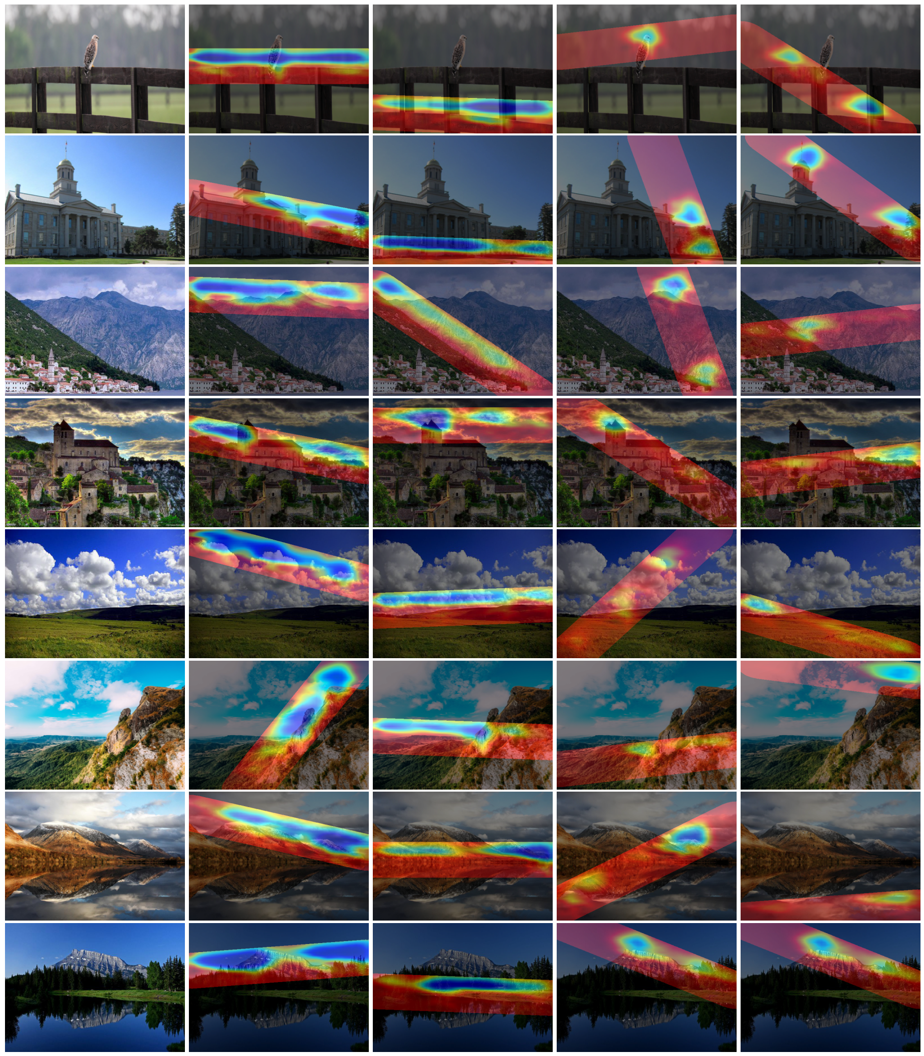
Efficacy of R-Net and M-Net: By comparing methods IV and V, we see that the proposed DRM algorithm improves all three scores using R-Net and M-Net, instead of the NMS scheme. Especially, in terms of AUC_A and AUC_P, DRM outperforms NMS by significant margins of 4.08 and 4.90, respectively. This means that the proposed comparative ranking and matching scheme selects reliable semantic lines and removes redundant ones effectively even on challenging scenes.

A.2 Semantic Line Detection Results

We compare more detection results. Primary (dashed red) and multiple (solid yellow) semantic lines are depicted.



A.3 Examples of Attention Masks

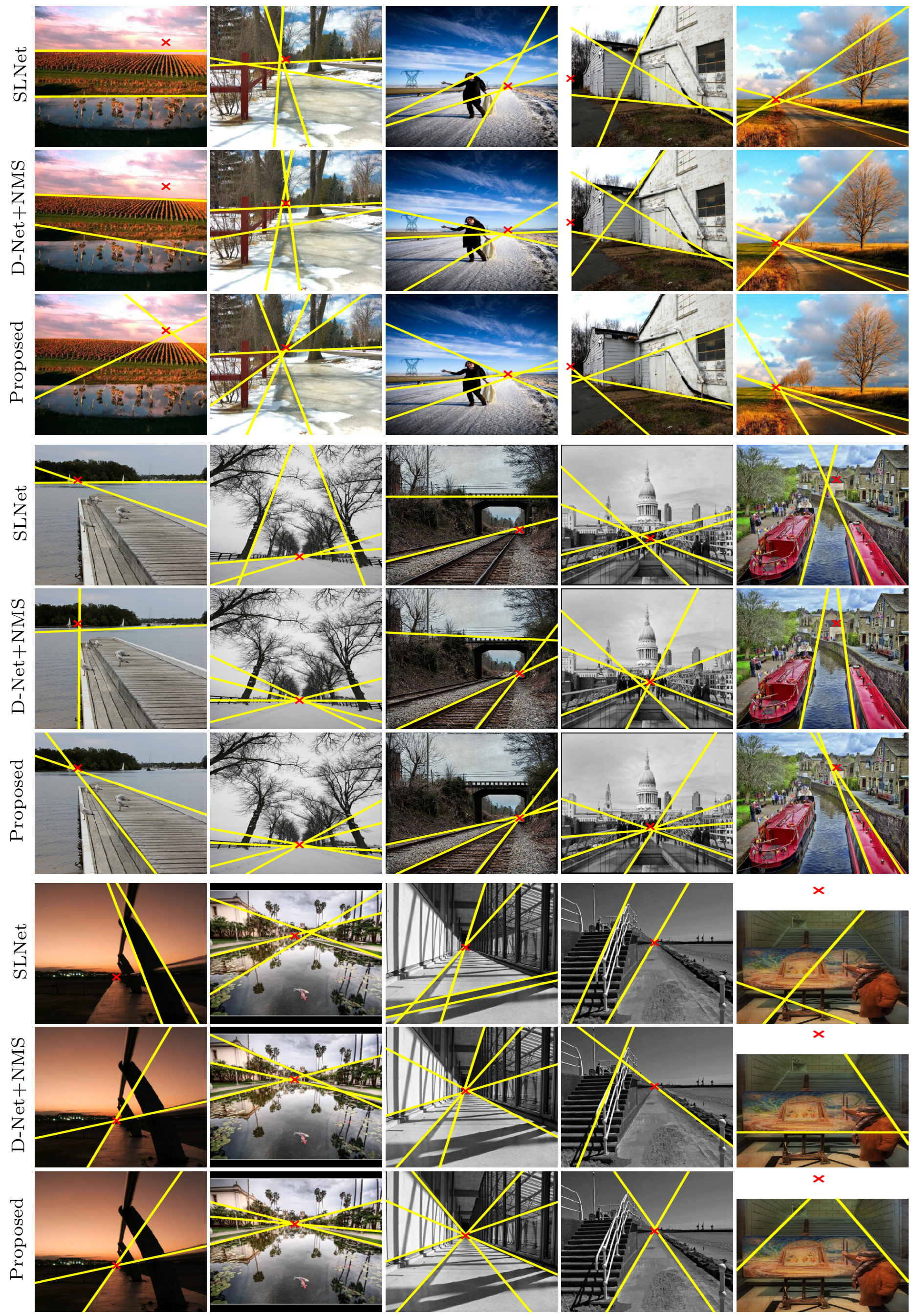


(a) (b) (c) (d) (e)

Fig. A-3: Attention masks for semantic line detection on the SEL dataset: (a) shows an input image. (b) and (c) are attention masks for semantic lines. (d) and (e) are those for negative lines. Note that the mirror attention module tends to assign small weights to one region and big weights to the other in the case of a semantic line. No such tendency is observed for negative lines. Attention masks are color coded: red and blue depict big and small values.

A.4 Dominant Parallel Line Detection

We provide more detection results on the AVA landscape dataset [58]. Dominant parallel lines are in yellow, while ground-truth vanishing points are depicted by red crosses.



A.5 Reflection Symmetry Axis Detection

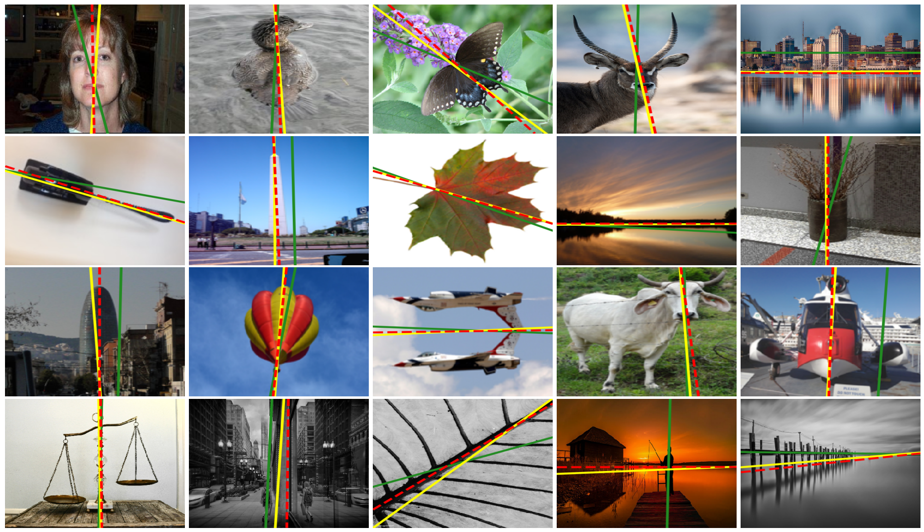
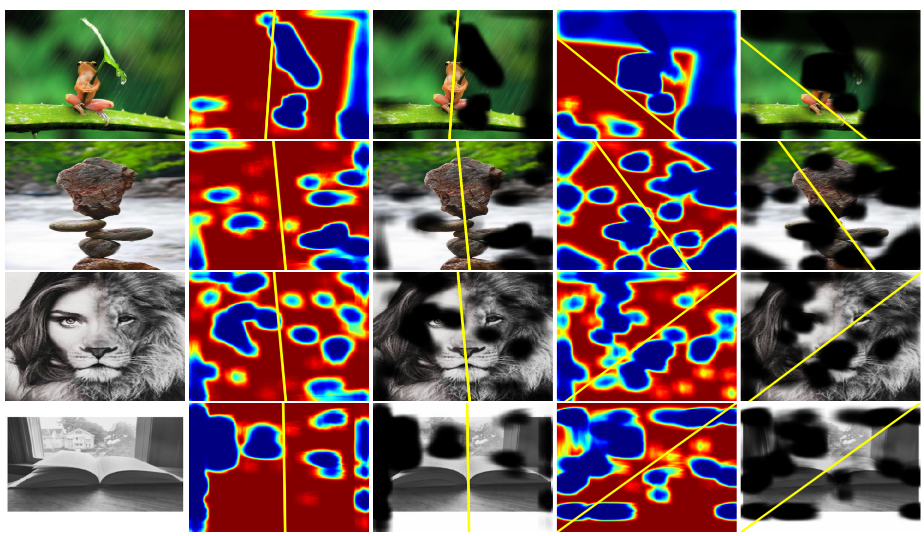


Fig. A-4: Detection results of symmetry axes. The ground-truth axes are in red, the detection results of [34] are in green, and those of the proposed algorithm are in yellow.



(a) (b) (c) (d) (e)

Fig. A-5: Attention masks for symmetry axis detection on the SYM_Hard dataset: (a) shows an input image. (b) and (d) are attention masks for two candidate axes. (c) and (e) are weighted images by the masks in (b) and (d), respectively. Note that the mirror attention module tends to assign small weights to asymmetric regions with respect to a candidate axis, while big weights to symmetric ones. Similar to Fig. A-3, attention masks are color coded.