Supplementary for CUNSB-RFIE: Context-aware Unpaired Neural Schrödinger Bridge in Retinal Fundus Image Enhancement

A. Experiment Details

Image Enhancement Experiment. The experiment was conducted on the synthetically degraded retinal fundus images, created by combining Light Transmission Disturbance, Image Blurring, and Retinal Artifact. The training and testing sets consisted of 3500 and 1891 high-quality retinal images from the EyeQ dataset, respectively. The model was trained for 130 epochs using the Adam optimizer, with an initial learning rate of 2×10^{-4} and β values set to 0.5 and 0.999, respectively. The learning rate was linearly decayed to 0 after running the first 80 epochs and the batch size was set to 8. The trained weights with best performance were then evaluated on complete DRIVE and IDRID datasets. SSIM and PSNR scores were calculated between the synthetically generated high-quality images and the corresponding ground-truth, where the low-quality counterparts were generated using the degrading algorithms described in [2].

Downstream Segmentation. Two Downstream segmentation tasks were conducted to demonstrate the ability to preserve intricate details in low-quality fundus images after enhancement, focusing on both high-frequency structure (i.e., lesion structure) and low-frequency information (i.e., blood vessel structure). Specifically, a vessel segmentation task was conducted on the DRIVE dataset and a diabetic lesion segmentation was performed on the IDRID dataset.

For vessel segmentation, we used the official split of the DRIVE dataset, dividing the training and testing sets equally. Performance was evaluated using the Area under ROC (AUC), the Area under Precision-Recall curve (PR), Sensitivity (SE), and Specificity (SP). For lesion segmentation, 54 subjects were used in the training set and 27 subjects in the testing set, with a focus on large lesion blocks such as Hard Exudates (EX) and Hemorrhages (HE). The evaluation metrics for this task were AUC, PR and the Jaccard index.

A vanilla UNet [1] model was trained from scratch for both segmentation tasks, respectively. For vessel segmentation, we used the Adam optimizer with cross-entropy loss as the objective function. The initial learning rate and batch size were set to 5×10^{-5} and 64, respectively. For lesion segmentation, the same optimizer was used. And the learning rate was set to 2×10^{-4} with a weight decay equal to 5×10^{-4}. The batch size was set to 8 the model was trained for 300 epochs.

B. Experiment Results

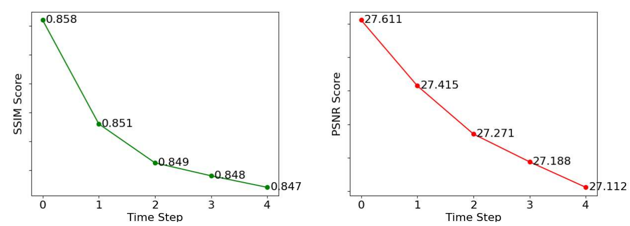

Figure 1. **Left**: SSIM scores decrease from 0.858 to 0.847 as time increases. **Right**: PSNR scores decrease from 27.611 to 27.122 as time increases.

From Fig. 1, we can find that the quality of synthetic images gradually drop as time step t_i increase, which indicate the smooth influnece shown in CUNSB-RFIE.

References

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. 1

[2] Z. Shen, H. Fu, J. Shen, and L. Shao. Modeling and Enhancing Low-Quality Retinal Fundus Images. *IEEE Trans Med Imaging*, 40(3):996–1006, 2021. 1