

For the ablation study, FRB and AGB in each AFR module have been designed with lightweight convolution, and cannot be modified into smaller structure. To answer the concern that performance improvement may come from more parameters than the baseline, we try to modify the baseline to a model with similar parameters as the AFR-Net. Specifically, for models in Table 1 of the paper, we modify the baseline as the combination of two encoders (one as replacement of AGB1), one FRB, and one decoder. The modified model (denoted as BaseR) shows nearly the same complexity as AFR-Net with one refinement module. However, the PSNR/SSIM on RS-Data is 26.35/0.907, on RD-Data is 29.11/0.908, and the results are much worse than AFR-B as described in Table 1 of the paper. Furthermore, the same strategy is used to modify the baseline model to be the same complexity as AFR-Net with different numbers of AFR modules as in Table 2, and the results are similar that the modified baseline models are outperformed by the AFR-Net variants by a large margin. Such results indicate that the performance improvement is not from the more parameters but from the novel attention guided refinement design. The above results and analysis will be added to the paper in the camera ready version.