1. Introduction

In this material, we present more experiment results and qualitative comparisons to show the effectiveness of our CODE and discuss the broader impact. From the visual results, one could see that CODE recovers image structures better and image details finer, thus obtaining clearer restorations.

2. Ablation for Aggregation Factors

In addition to the analysis experiments presented in the main body of the paper, here we show more results about $r_c$ and $r_s$. The experimental settings are the same as other analysis experiments, which are conducted on Set12 with noise level of 50. As shown in Tabs. 1 and 2, our method is not sensitive to $r_c$ and $r_s$. Increasing $r_c$ or decreasing $r_s$ only leads to slight performance drop.

<table>
<thead>
<tr>
<th>$r_c$</th>
<th>#Params</th>
<th>FLOPs</th>
<th>PSNR</th>
<th>SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>18.67M</td>
<td>25.84G</td>
<td>27.92</td>
<td>0.8077</td>
</tr>
<tr>
<td>4</td>
<td>12.18M</td>
<td>22.44G</td>
<td>27.93</td>
<td>0.8083</td>
</tr>
<tr>
<td>8</td>
<td>8.94M</td>
<td>20.75G</td>
<td>27.90</td>
<td>0.8074</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$r_s$</th>
<th>#Params</th>
<th>FLOPs</th>
<th>PSNR</th>
<th>SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>[16, 8, 4, 2]</td>
<td>9.51M</td>
<td>22.10G</td>
<td>27.91</td>
<td>0.8072</td>
</tr>
<tr>
<td>[32, 16, 8, 4]</td>
<td>12.18M</td>
<td>22.44G</td>
<td>27.93</td>
<td>0.8083</td>
</tr>
<tr>
<td>[64, 32, 16, 4]</td>
<td>17.40M</td>
<td>23.11G</td>
<td>27.92</td>
<td>0.8079</td>
</tr>
</tbody>
</table>

3. Qualitative Results on Image Denoising

In addition to the qualitative results presented in the main body of the paper, we show more results on grayscale and color image denoising at noise level 15, 25 and 50. As shown in Figs. 1 to 5, CODE demonstrates better visual results on grayscale image denoising compared with the other methods. To be specific, DnCNN [11] and FFDNet [12] leave noises and artifacts in the images, DRUNet [10] and SwinIR [4] obtain oversmoothed results and lose the details in the door and window. However, our CODE fully recovers the structures of the door and the window while preserving the textures better. Through careful observation, one can find that our visual results are the closest ones to GTs. Figs. 8 to 10 and Figs. 11 to 13 respectively show the qualitative results of color image denoising on Kodak24 [1] and McMaster [13]. From the figures, one can observe that DnCNN and FFDNet have residual noises and distortions, DRUNet and SwinIR smooth out the image details, thus obtaining the blurry images. In contrast, our CODE adequately removes the noises without introducing artifacts, while achieving better restoration on details and textures.

4. Qualitative Results on Motion Deblurring

In this section, we show the qualitative results on motion deblurring. Fig. 14 and Fig. 15 show the visual results on GoPro [5] test set and HIDE [6] dataset, respectively. From the figures, one can observe that DGAN [2] and DGANv2 [3] have residual black streaks and artifacts, SRN [8], DMPHN [9], and SAPHNet [7] obtain blurry results with distortions. By contrast, our CODE recovers better structures and more details with much less residual blur in the restored images. Such an observation is more obvious in Fig. 14, where CODE obtains almost the same result as GT.

5. Qualitative Results on JPEG Compression Artifact Reduction

Figs. 16 to 19 and Figs. 20 to 23 show the qualitative results on JPEG compression artifact reduction at compression level 10, 20, 30, and 40. From the figures, one can observe that DnCNN has residual artifacts, DRUNet, SwinIR, and CODE all obtain visually pleasant results with fine textures and structures, in which our CODE is much more ef-
ficient than the other two methods. Specifically, CODE has \( \sim 37\% \) parameters and \( \sim 31\% \) FLOPs of DRUNet, and only \( \sim 6\% \) FLOPs of SwinIR. In other words, our CODE can provide similar results with much less computation and/or memory costs.

6. Broader Impact

In this section, we discuss the impact of our CODE in a broader vision. Generally, CODE is much more efficient than other competing image restoration Transformers, and thus has more opportunities to be deployed on mobile devices. However, CODE is a general neural network and might be trained with customized data and used for unauthorized purposes, such as watermark removal, which might prejudice the rights of others. Moreover, the training and testing of the model consume a lot of electricity, which causes carbon emissions.

References

Figure 1. Qualitative comparisons of grayscale image denoising (noise level 15) on the Set12 dataset.

Figure 2. Qualitative comparisons of grayscale image denoising (noise level 25) on the Set12 dataset.

Figure 3. Qualitative comparisons of grayscale image denoising (noise level 15) on the BSD68 dataset.

Figure 4. Qualitative comparisons of grayscale image denoising (noise level 25) on the BSD68 dataset.
Figure 5. Qualitative comparisons of **grayscale image denoising** (noise level 50) on the BSD68 dataset.

Figure 6. Qualitative comparisons of **color image denoising** (noise level 15) on the CBSD68 dataset.

Figure 7. Qualitative comparisons of **color image denoising** (noise level 25) on the CBSD68 dataset.

Figure 8. Qualitative comparisons of **color image denoising** (noise level 15) on the Kodak24 dataset.

Figure 9. Qualitative comparisons of **color image denoising** (noise level 25) on the Kodak24 dataset.
Figure 10. Qualitative comparisons of **color image denoising** (noise level 50) on the Kodak24 dataset.

Figure 11. Qualitative comparisons of **color image denoising** (noise level 15) on the McMaster dataset.

Figure 12. Qualitative comparisons of **color image denoising** (noise level 25) on the McMaster dataset.

Figure 13. Qualitative comparisons of **color image denoising** (noise level 50) on the McMaster dataset.

Figure 14. Qualitative comparisons of **motion deblurring** on the GoPro test set.

Figure 15. Qualitative comparisons of **motion deblurring** on the HIDE dataset.
Figure 16. Qualitative comparisons of **JPEG compression artifact reduction** (compression level $q = 10$) on the Classic5 dataset.

Figure 17. Qualitative comparisons of **JPEG compression artifact reduction** (compression level $q = 20$) on the Classic5 dataset.

Figure 18. Qualitative comparisons of **JPEG compression artifact reduction** (compression level $q = 30$) on the Classic5 dataset.

Figure 19. Qualitative comparisons of **JPEG compression artifact reduction** (compression level $q = 40$) on the Classic5 dataset.

Figure 20. Qualitative comparisons of **JPEG compression artifact reduction** (compression level $q = 10$) on the LIVE1 dataset.
Figure 21. Qualitative comparisons of JPEG compression artifact reduction (compression level $q = 20$) on the LIVE1 dataset.

Figure 22. Qualitative comparisons of JPEG compression artifact reduction (compression level $q = 30$) on the LIVE1 dataset.

Figure 23. Qualitative comparisons of JPEG compression artifact reduction (compression level $q = 40$) on the LIVE1 dataset.