Improving Single-Image Defocus Deblurring: How Dual-Pixel Images Help Through Multi-Task Learning
Supplementary Material

Abdullah Abuolaim Mahmoud Afifi Michael S. Brown
York University
{abuolaim,mafifi,mbrown}@eecs.yorku.ca

This supplementary material provides an ablation study (Sec. S1) to show the effectiveness of the multi-decoder stitching at the middle stage vs. other options (i.e., late-stage and no stitching) of the proposed multi-task dual-pixel network (MDP). Additional qualitative results are also provided as follows:

- Fig. S1 provides a qualitative comparison with other single-image defocus deblurring methods tested on the Canon DP dataset [1]. These methods are: the just noticeable defocus blur method (JNB) [4], the edge-based defocus blur estimation method (EBDB) [2], the deep defocus map estimation method (DMENet) [3], and the DPDNet (single) [1].
- Fig. S2 demonstrates qualitatively the deblurring generalization ability of our proposed MDP. In this experiment, MDP is trained only on the Canon data, but tested on images from other cameras.
- Figs. S3, S4, S5, and S6 show examples from our newly captured DLDP dataset along with results of high-quality reconstructed DP views and image motion of the synthesized eight views. The eight views are generated based on the description in Sec. 4.4 of the main paper, and visualized as image motion by alternating through eight views (i.e., NIMAT effect).
- Fig. S7 shows qualitatively how our proposed MDP is able to generalize for other cameras. In this experiment, we synthesize eight views from a single-input image captured by other cameras. The eight views are generated based on the description in Sec. 4.4 of the main paper, and visualized as image motion by alternating the eight views (i.e., NIMAT effect).

Note: This supplemental material contains animated figures for better visualization. However, the IEEE PDF eXpress validator does not allow the animation package. Therefore, we provide videos in the "videos" directory – located in the same path as this PDF file – for all the animated figures in the main paper as well as the supplementary material. We also provide in-PDF animated figures in our arXiv version [1].

Table S1: This table reports results on an ablation study performed to examine the effectiveness of the multi-decoder stitching design for defocus deblurring. Three variations of our proposed MDP based on three different stitching options: (1) no stitching, (2) late-stage stitching and (3) middle-stage stitching. Results reported are on the Canon DP deblurring dataset [1].

MDP variation	PSNR ↑	SSIM ↑	MAE ↓
MDP (no stitching)	25.03	0.757	0.042
MDP (late-stage stitching)	25.16	0.759	0.041
MDP (middle-stage stitching)	**25.35**	**0.763**	**0.040**

S1. Ablation study of multi-decoder stitching

In this section, we investigate the utility of having multiple weight sharing stages by introducing a variation of MDP network with different multi-decoder stitching options: (1) *no stitching* that makes the latent space $\mathcal{X}$ the only weight sharing stage, (2) *late-stage stitching* at the last block, and (3) the original proposed MDP with *middle-stage stitching*. We report the results of MDP variations in Table S1. The training procedure followed is the same as for all MDP variations as described in Sec. 4.1 of the main paper.

The results in Table S1 show that *middle-stage stitching* achieves the best results as it allows weight sharing at multiple stages compared with the *no stitching* variation. On the other hand, there is a noticeable drop in the deblurring performance when *late-stage stitching* is applied as the sharpness of the deblurring decoder (i.e., $\mathbf{Dec}_s$) is affected by the half-PSF blur present in feature maps of the synthesized DP views (i.e., $\mathbf{Dec}_l$ and $\mathbf{Dec}_r$) at this later stage.

[1] https://arxiv.org/pdf/2108.05251.pdf

(a) Input

(b) EBDB [2]

(c) DMENet [3]

(d) DPDNet (single) [1]

(e) Ours

(f) Ground truth

Figure S1: Additional qualitative comparisons with other single-image defocus deblurring methods on the test set of the Canon DP dataset [1]. We compare our results with the following single-image defocus deblurring methods: EBDB [2], DMENet [3], and DPDNet (single) [1]. Note that DPDNet was originally introduced to use DP images as input, but the authors in [1] also provided the same model trained on a single image, denoted as DPDNet (single). Our method produces the best quantitative and arguably best qualitative results.

(a) Input (b) EBDB [2] (c) DPDNet (single) [1] (d) Ours

Figure S2: Qualitative comparison with single-image defocus deblurring methods on other camera devices (top: Samsung S6, bottom: Flickr image [erasmusa CC BY-NC 2]. We compare our results with the following single-image defocus deblurring methods: EBDB [2] and DPDNet (single) [1]. Note that DPDNet was originally introduced to use DP images as input, but the authors in [1] also provided the same model trained on a single image, denoted as DPDNet (single). Our method generalizes well for unseen cameras during the training stage and produces arguably best qualitative results compared with other methods.

References

[1] Abdullah Abuolaim and Michael S Brown. Defocus deblurring using dual-pixel data. In *ECCV*, 2020.

[2] Ali Karaali and Claudio Rosito Jung. Edge-based defocus blur estimation with adaptive scale selection. *TIP*, 27(3):1126–1137, 2018.

[3] Junyong Lee, Sungkil Lee, Sunghyun Cho, and Seungyong Lee. Deep defocus map estimation using domain adaptation. In *CVPR*, 2019.

[4] Jianping Shi, Li Xu, and Jiaya Jia. Just noticeable defocus blur detection and estimation. In *CVPR*, 2015.

(a) Input

(b) Our generated NIMAT effect

(c) Our DP views

(d) GT DP views

Figure S3: An example from our DLDP dataset. (a) Input combined image $\mathbf{I}_c$. (b) Our novel NIMAT effect generated using the proposed MDP. (c) Animated results of our synthesized DP views. (d) Animated ground truth DP views. Our MDP is able to generate high-quality eight/DP views. **Note: (b), (c), and (d) are animated figures. The animation video is in the "videos" directory – located in the same path as this PDF file. We also provide in-PDF animated figures in our arXiv version: `https://arxiv.org/pdf/2108.05251.pdf`.**

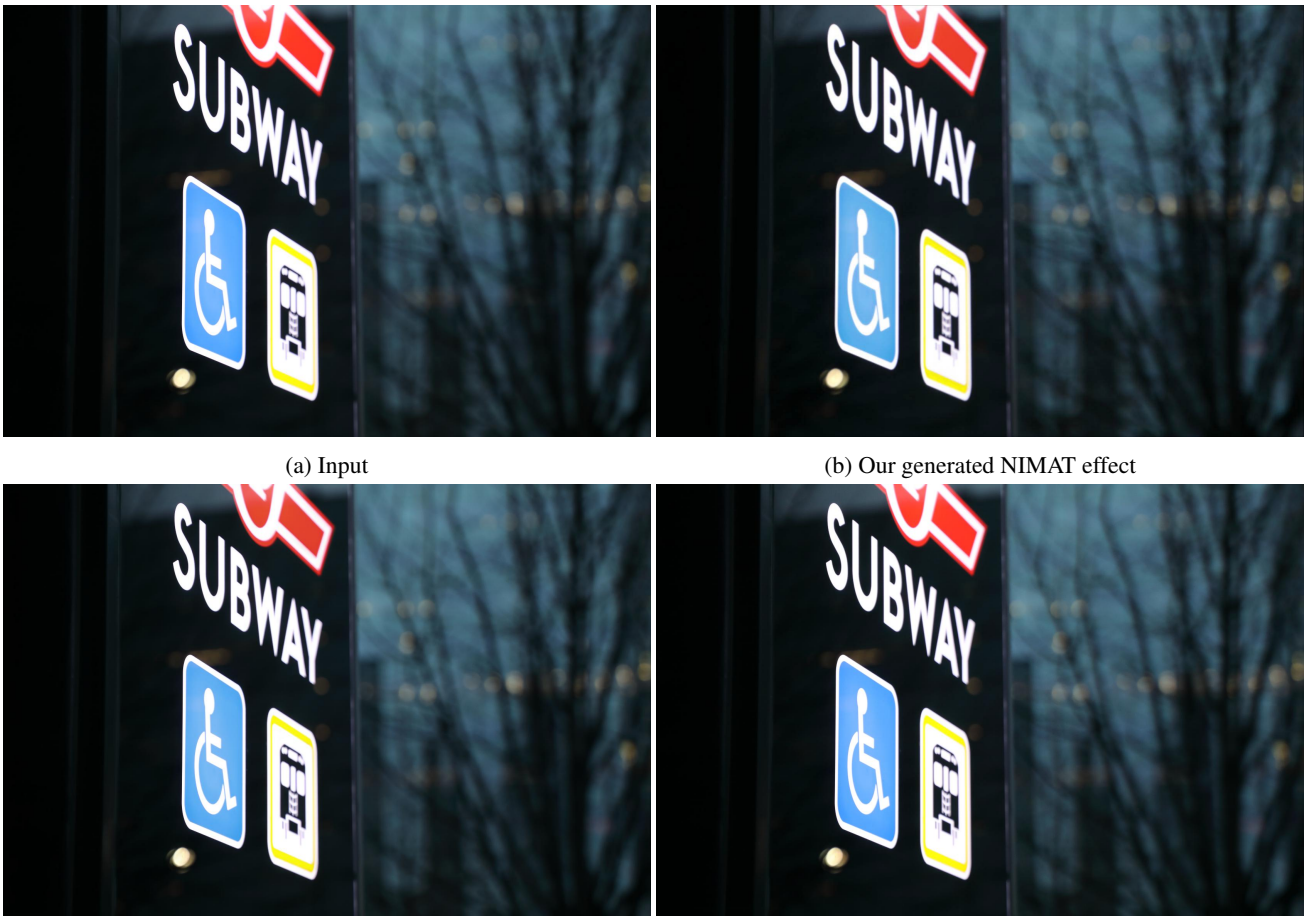
(a) Input

(b) Our generated NIMAT effect

(c) Our DP views

(d) GT DP views

Figure S4: An example from our DLDP dataset. (a) Input combined image $\mathbf{I}_c$. (b) Our novel NIMAT effect generated using the proposed MDP. (c) Animated results of our synthesized DP views. (d) Animated ground truth DP views. Our MDP is able to generate high-quality eight/DP views. **Note: (b), (c), and (d) are animated figures. The animation video is in the "videos" directory – located in the same path as this PDF file. We also provide in-PDF animated figures in our arXiv version: `https://arxiv.org/pdf/2108.05251.pdf`.**



(a) Input

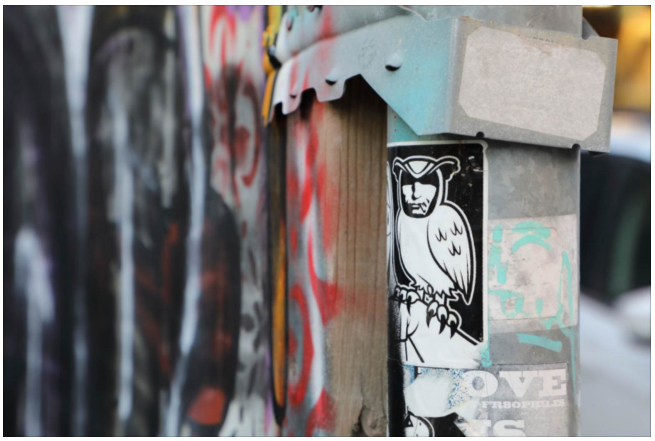
(b) Our generated NIMAT effect

(c) Our DP views

(d) GT DP views

Figure S5: An example from our DLDP dataset. (a) Input combined image $\mathbf{I}_c$. (b) Our novel NIMAT effect generated using the proposed MDP. (c) Animated results of our synthesized DP views. (d) Animated ground truth DP views. Our MDP is able to generate high-quality eight/DP views. **Note: (b), (c), and (d) are animated figures. The animation video is in the "videos" directory – located in the same path as this PDF file. We also provide in-PDF animated figures in our arXiv version: `https://arxiv.org/pdf/2108.05251.pdf`.**

(a) Input



(b) Our generated NIMAT effect

(c) Our DP views

(d) GT DP views

Figure S6: An example from our DLDP dataset. (a) Input combined image $\mathbf{I}_c$. (b) Our novel NIMAT effect generated using the proposed MDP. (c) Animated results of our synthesized DP views. (d) Animated ground truth DP views. Our MDP is able to generate high-quality eight/DP views. **Note: (b), (c), and (d) are animated figures. The animation video is in the "videos" directory – located in the same path as this PDF file. We also provide in-PDF animated figures in our arXiv version: `https://arxiv.org/pdf/2108.05251.pdf`.**

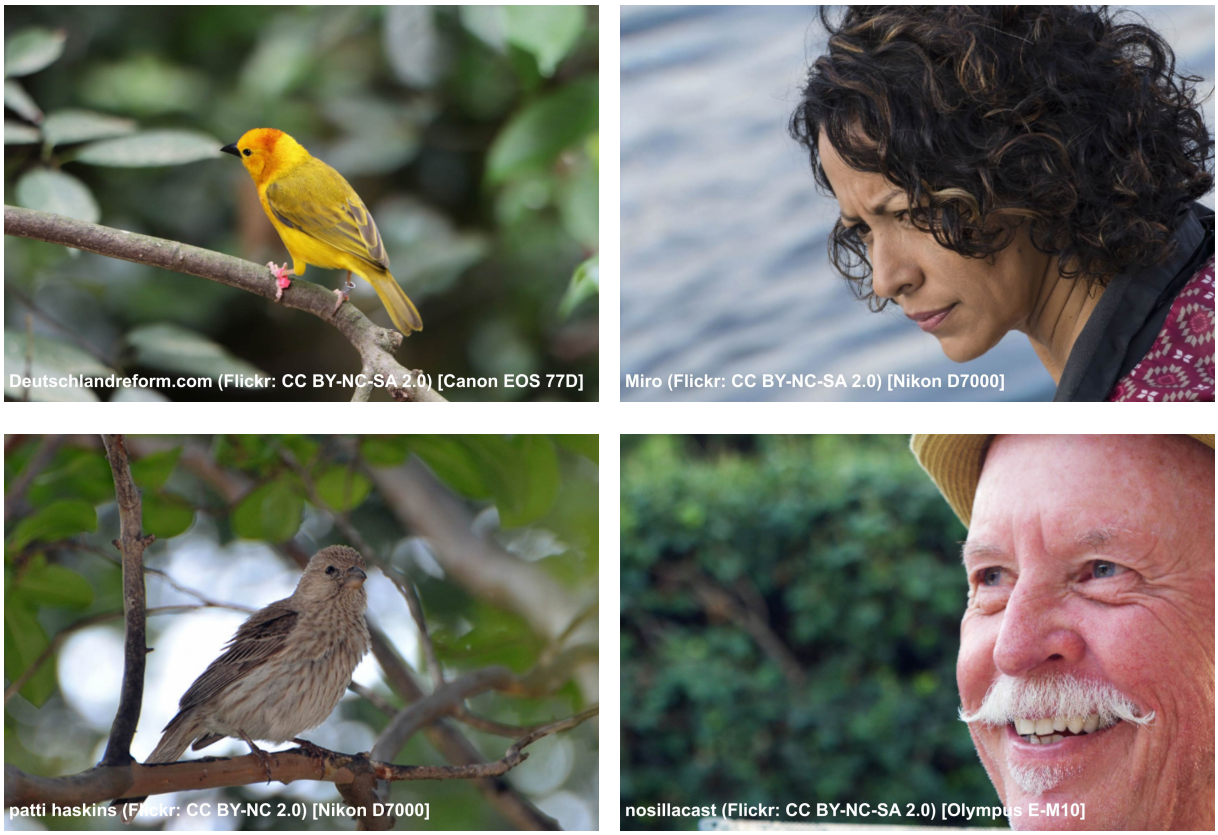


Figure S7: Our novel NIMAT effect. Multi-view synthesis results of our proposed MDP applied to other cameras than the Canon 5D DSLR camera (used for training). These results are synthesized from a single input image captured by new camera devices, in which they do not have the ground truth DP views. Our MDP produces high-quality eight views that can be used to create an aesthetically pleasing NIMAT effect. Furthermore, these results demonstrate a good generalization ability of our MDP as it can provide high-quality views from images that are captured by unseen camera device during the training stage. **Note: these images are animated figures. The animation video is in the "videos" directory – located in the same path as this PDF file. We also provide in-PDF animated figures in our arXiv version: `https://arxiv.org/pdf/2108.05251.pdf`.**