Supplementary Material: Domain-adaptive Video Deblurring via Test-time Blurring

Jin-Ting He*,[1], Fu-Jen Tsai*,[2], Jia-Hao Wu[1], Yan-Tsung Peng[3], Chung-Chi Tsai[4], Chia-Wen Lin[2], and Yen-Yu Lin[1]

[1] National Yang Ming Chiao Tung University, Taiwan
jinting.cs12@nycu.edu.tw, jiahao.11@nycu.edu.tw, lin@cs.nycu.edu.tw
[2] National Tsing Hua University, Taiwan
fjtsai@gapp.nthu.edu.tw, cwlin@ee.nthu.edu.tw
[3] National Chengchi University, Taiwan
ytpeng@cs.nccu.edu.tw
[4] Qualcomm Technologies, Inc., San Diego
chuntsai@qti.qualcomm.com

1 Algorithm of Selecting Relatively Sharp Patches

In the Relatively-Sharp Detection Module (RSDM), we adopt an adaptive sharpness threshold $\eta^{(i)}$ to ensure the number of relatively sharp patches extracted from the video $V^{(i)}$ to reach a specified percentage of the number of total frames in $V^{(i)}$. More precisely, we set $\eta^{(i)}$ to select the top $r\%$ relatively sharp patches to serve as pseudo-sharp images. Here, we describe how to determine the adaptive sharpness threshold $\eta^{(i)}$ in Algorithm 1 to select the top $r\%$ relatively sharp patches. First, the target number of selected patches is calculated as $N = r\% \times |V^{(i)}|$, where $|V^{(i)}|$ is the number of total frames in $V^{(i)}$, and the initial threshold is initialized as $\eta^{(i)} = 0.5$. Next, we choose relatively sharp patches based on the set threshold $\eta^{(i)}$ via RSDM (described in Algorithm 2). If we select more patches than the specified limit N, we decrease $\eta^{(i)}$ by 0.01 to choose fewer patches in the next iteration. On the other hand, if the number of the patches is less than N, we increase $\eta^{(i)}$ by 0.01 to obtain more patches next round. With the updated $\eta^{(i)}$, we iterate the above steps until the top $r\%$ relatively sharp patches in $V^{(i)}$ are collected, where r is set to 20 in our work. In the end, we obtain a set of relatively sharp patches, denoted as $S^{(i)}$ in Algorithm 1, for reblurring.

2 Computational cost of our method

Table 1 presents the computational cost of each component of our method for adapting the deblurring model, ESTRNN. The total runtime for processing $3,000$ images from the BSD testing dataset is 945 seconds, or 0.315 seconds per image on average, on an NVIDIA 3090 GPU. Additionally, our adaptation is conducted offline, allowing the adapted models to operate during inference without incurring additional overhead.

Algorithm 1 Selecting Top $r\%$ Relatively Sharp Patches

$V^{(i)}$: A blurred video
$N \leftarrow r\% \times |(V^{(i)}|$: Target number of selected patches
$\eta^{(i)} \leftarrow 0.5$: Adaptive sharpness threshold
$P \leftarrow N$: Number of selected patches in the previous iteration
while True **do**
  $S^{(i)} \leftarrow$ **RSDM**$(V^{(i)}, \eta^{(i)})$: **RSDM** returns the selected relatively sharp patches
  $N^{(i)} \leftarrow |S^{(i)}|$: Number of selected relatively sharp patches
  if $N == N^{(i)}$ **then**
    return $S^{(i)}$
  else if $N > N^{(i)}$ **then**
    $\eta^{(i)} = \eta^{(i)} + 0.01$
    if $N < P$ **then**
      return $S^{(i)}$
    end if
  else if $N < N^{(i)}$ **then**
    $\eta^{(i)} = \eta^{(i)} - 0.01$
    if $N > P$ **then**
      return $S^{(i)}$
    end if
  end if
  $P \leftarrow N^{(i)}$
end while

Algorithm 2 Extracting the relatively sharp patches based on $\eta^{(i)}$ (**RSDM**)

$\eta^{(i)}$: Adaptive sharpness threshold
$V^{(i)}$: Blurred video
$M^{(i)} \leftarrow BME(V^{(i)})$: Blur magnitude map estimated by Blur Magnitude Estimator
$Mask^{(i)} \leftarrow Binarize(M^{(i)}, \eta^{(i)})$: Blur magnitude map is binarized based on $\eta^{(i)}$
$S^{(i)} \leftarrow []$: Selected relatively sharp patches
for $t \leftarrow 1$ **to** $|Mask^{(i)}|$ **do**
  $s \leftarrow$ regionprop$(Mask_t^{(i)})$: t-th sharp patch cropped by regionprop library [1]
  if shape$(s) \geq (256, 256)$ **then**
    Append s to $S^{(i)}$
  end if
end for
return $S^{(i)}$

Table 1: Runtime for each component of our method when adapting ESTRNN on the BSD testing dataset (3000 frames). Note that the fine-tuning time will very across differents deblrred models

RSDM	DBCGM	Reblurring	Fine-tuning	Total
291s	177s	171s	306s	945s

Table 2: Ablation studies on different strategies to set a domain-specific blur magnitude in the Magnitude Adaptation Process.

Method	Baseline	Method 1	Method 2	Ours
PSNR	25.57	28.70	28.74	**29.44**

3 Effects of Fusion Strategy on the Magnitude Adaptation Process

In Table 2, we compare various strategies for setting a domain-specific blur magnitude in the Magnitude Adaptation Process of the DBCGM. First, we average the blur magnitudes across all the considered frames $\{V_{t-2}^{(i)}, ..., V_{t+2}^{(i)}\}$) to obtain the domain-specific blur magnitude (Method 1), namely

$$\tilde{M}_t^{(i)} = \text{Avg}(M_{t-2}^{(i)}, M_{t-1}^{(i)}, M_t^{(i)}, M_{t+1}^{(i)}, M_{t+2}^{(i)}). \tag{1}$$

Second, we average the blur magnitudes excluding the center frame $V_t^{(i)}$ (Method 2), *i.e.*,

$$\tilde{M}_t^{(i)} = \text{Avg}(M_{t-2}^{(i)}, M_{t-1}^{(i)}, M_{t+1}^{(i)}, M_{t+2}^{(i)}). \tag{2}$$

Lastly, the proposed method (denoted as Ours) for determining the blur magnitude is depicted as follows

$$\tilde{M}_t^{(i)} = \text{Norm}(M_t^{(i)}) \cdot \text{Avg}(M_{t-2}^{(i)}, M_{t-1}^{(i)}, M_{t+1}^{(i)}, M_{t+2}^{(i)}). \tag{3}$$

The experimental results show that all the methods improve performance, where the proposed method achieves the best results.

4 Domain-adaptive Reblurred Results on Real-world Datasets

We demonstrate reblurred results obtained using ID-Blau [8] based on the domain-specific blur conditions generated by our Domain-adaptive Blur Condition Generation Module (DBCGM) on several datasets, including BSD-1ms8ms [9] in Figure 1, BSD-2ms16ms [9] in Figure 2, BSD-3ms24ms [9] in Figure 3, RBVD [2] in Figure 4, and RealBlur [6] in Figure 5. These results show that we can generate blurred images consistent with blur patterns present in test videos.

5 Domain-adaptive Deblurred Results on Real-world Datasets

We provide additional qualitative results for deblurring models using our domain adaptation scheme, denoted as Ours, compared to those without using the scheme, denoted as Baseline, in Figures 6 $\sim$ 25. These deblurring models are trained on the GOPRO [4] dataset, and tested on the RealBlur [6], RBVD [2], and three versions of BSD [9] datasets. We demonstrate the qualitative comparisons based on four state-of-the-art video deblurring models, including DSTNet [5] in Figures 6 $\sim$ 10, ESTRNN [9] in Figures 11 $\sim$ 15, MMP-RNN [7] in Figures 16 $\sim$ 20, and Shift-Net [3] in Figures 21 $\sim$ 25.

6 Limitations

6.1 Limitation for single-image deblurring

Our proposed method requires consecutive frames from a video to accurately compile blur conditions. This requirement inherently limits the applicability of our approach to scenarios where video data is available, and consecutive frames can be processed. Consequently, our method is not suitable for single-image deblurring models, which aim to restore sharpness in individual frames without the context provided by adjacent frames.

Fig. 1: Qualitative reblurred results of our scheme on BSD-1ms8ms [9].

Fig. 2: Qualitative reblurred results of our scheme on BSD-2ms16ms [9].

Fig. 3: Qualitative reblurred results of our scheme on BSD-3ms24ms [9].

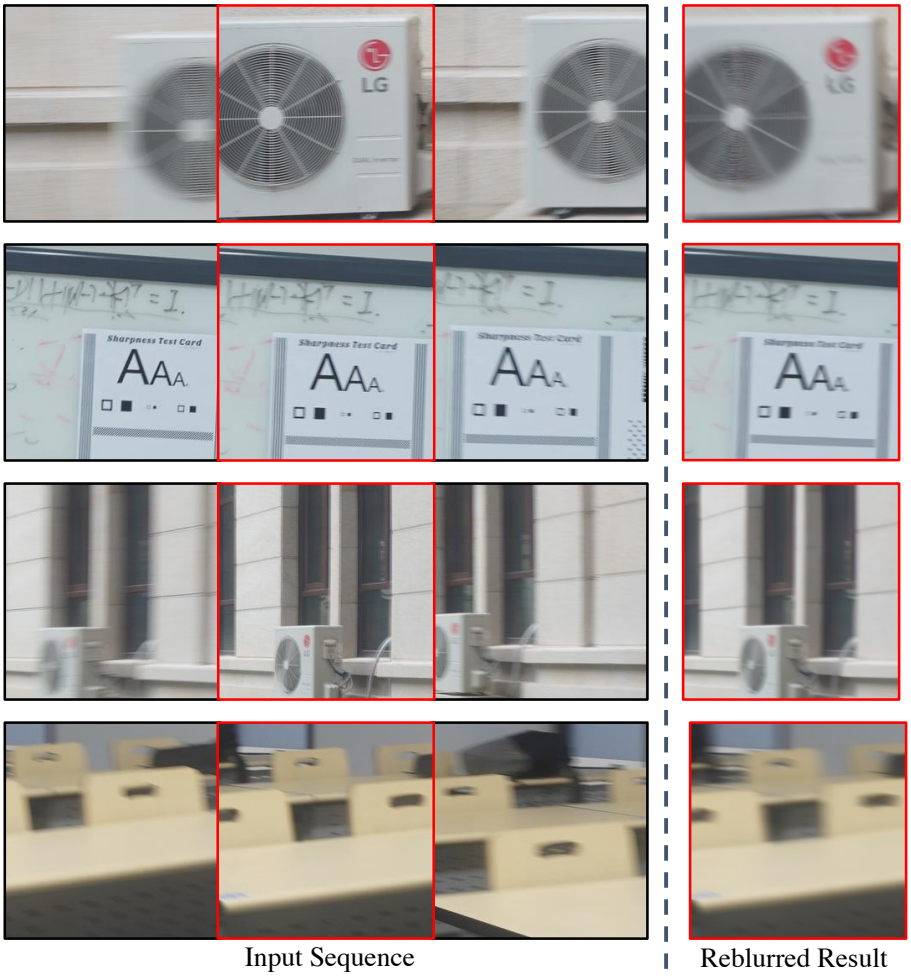

Fig. 4: Qualitative reblurred results of our scheme on RBVD [2].

Fig. 5: Qualitative reblurred results of our scheme on RealBlur [6].

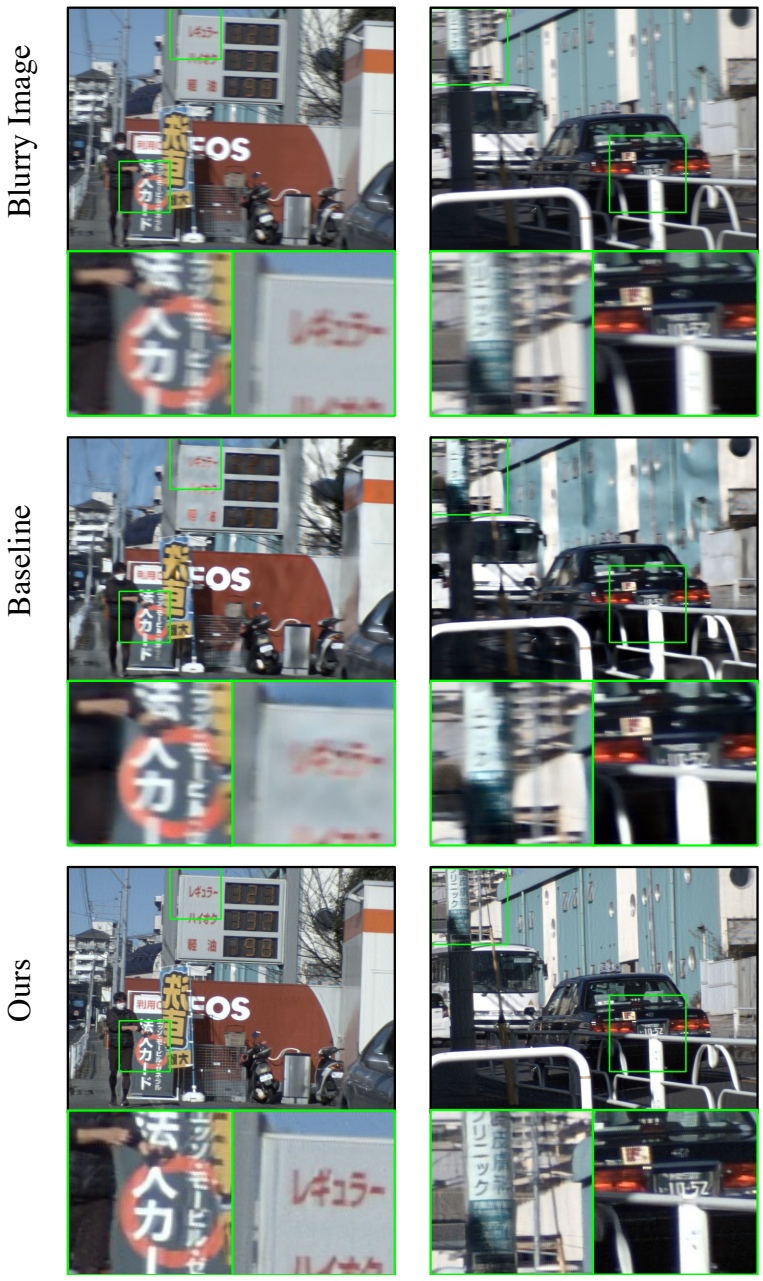

Fig. 6: Qualitative results of DSTNet [5] on the BSD-1ms8ms [9] dataset.

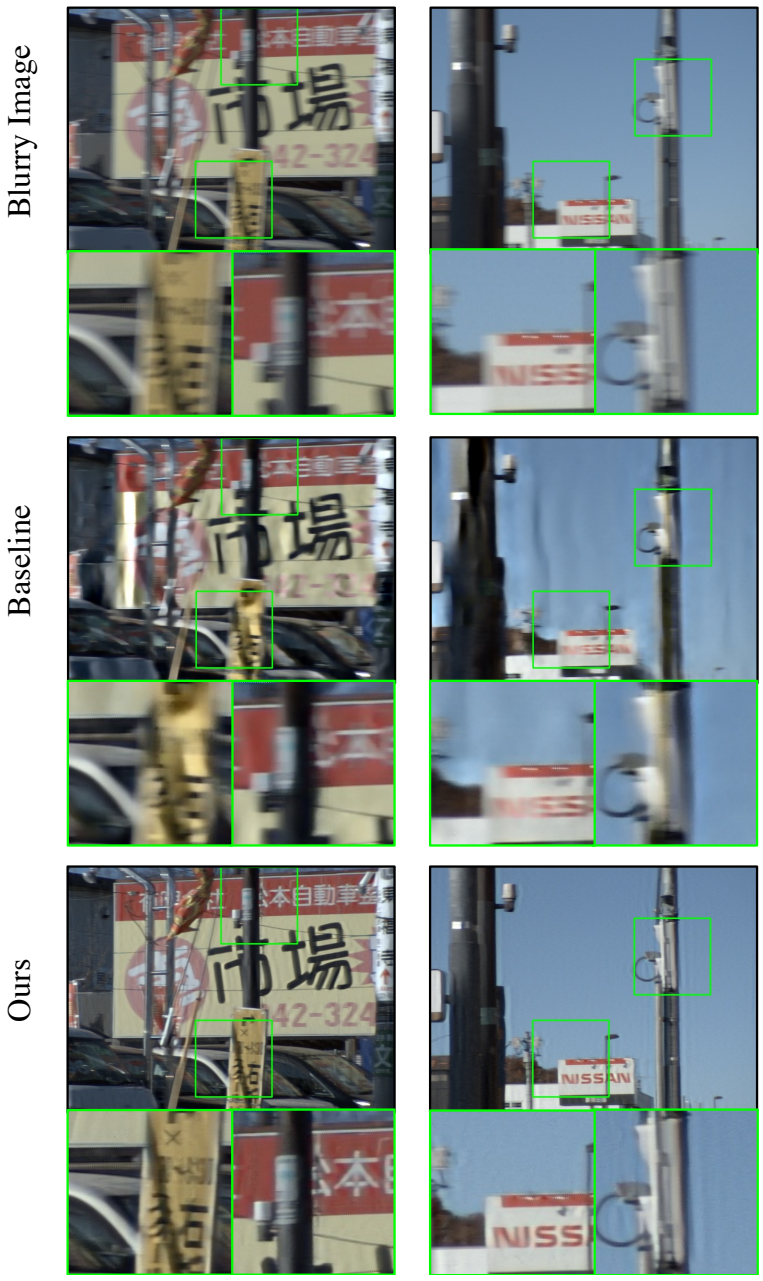

Fig. 7: Qualitative results of DSTNet [5] on the BSD-2ms16ms [9] dataset.

Fig. 8: Qualitative results of DSTNet [5] on the BSD-3ms24ms [9] dataset.

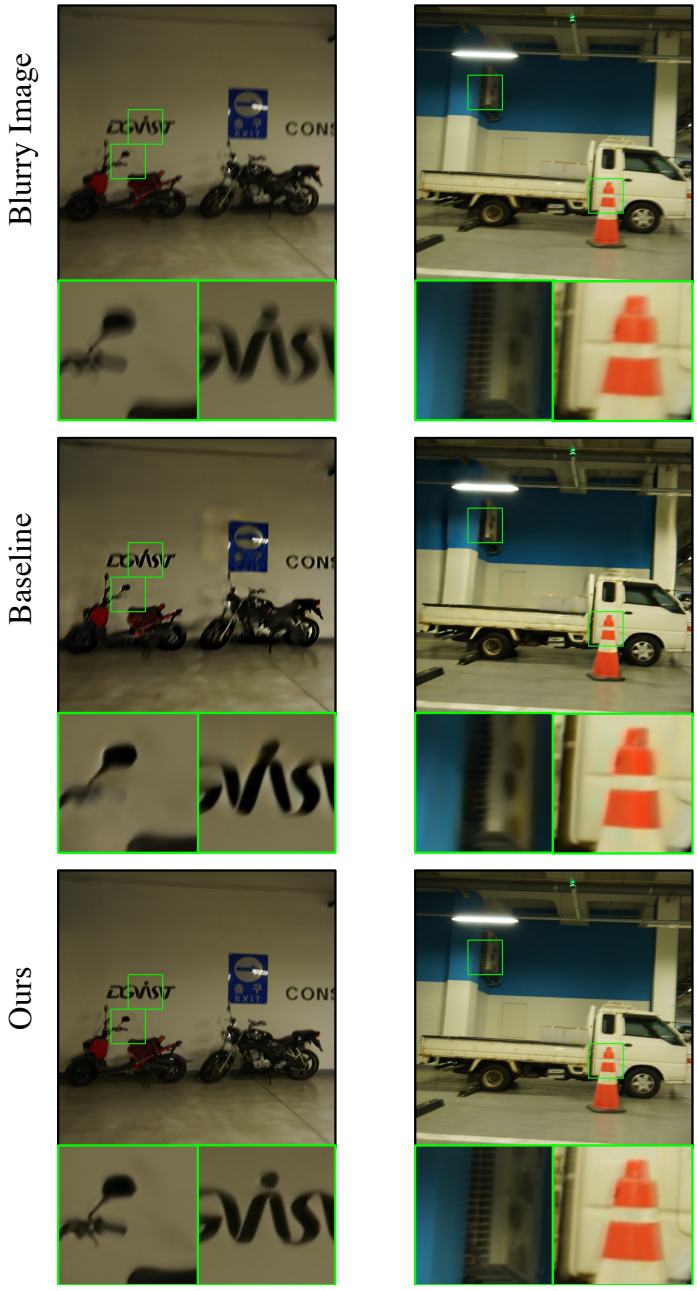

Fig. 9: Qualitative results of DSTNet [5] on the RealBlur [6] dataset.

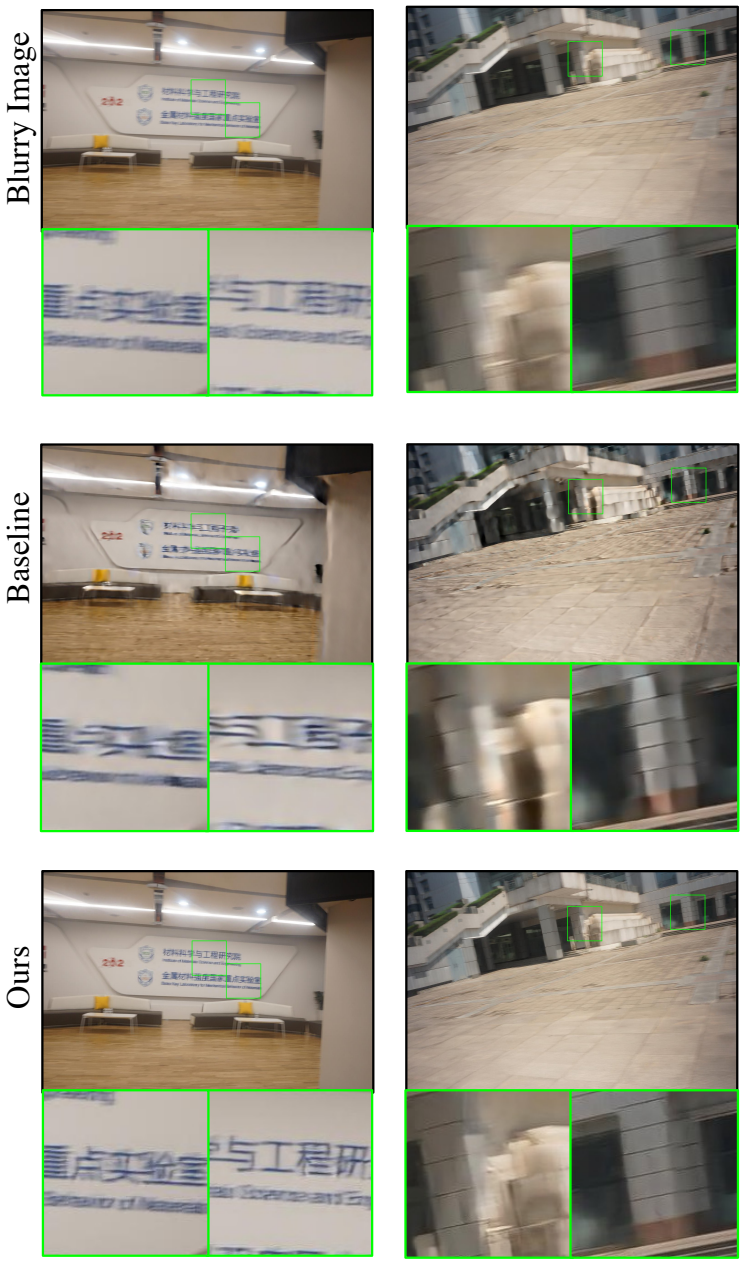

Fig. 10: Qualitative results of DSTNet [5] on the RBVD [2] dataset.

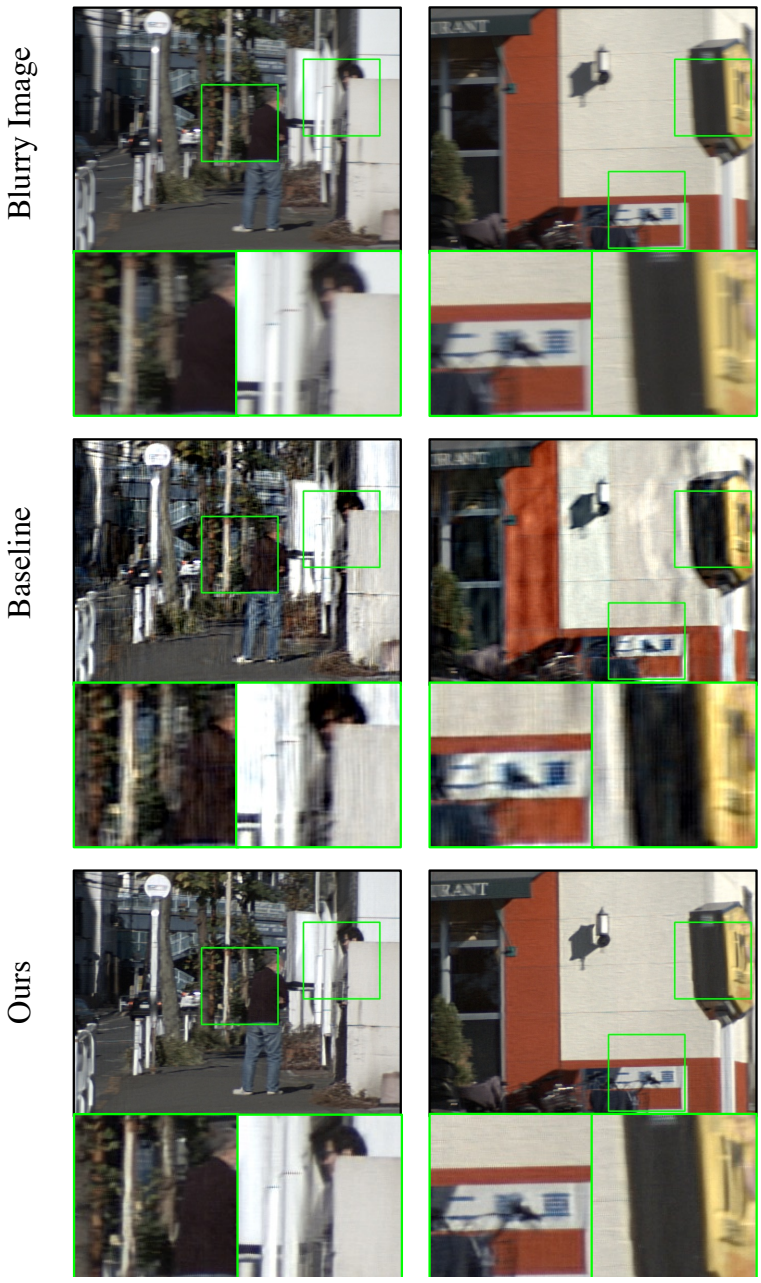

Fig. 11: Qualitative results of ESTRNN [9] on the BSD-1ms8ms [9] dataset.

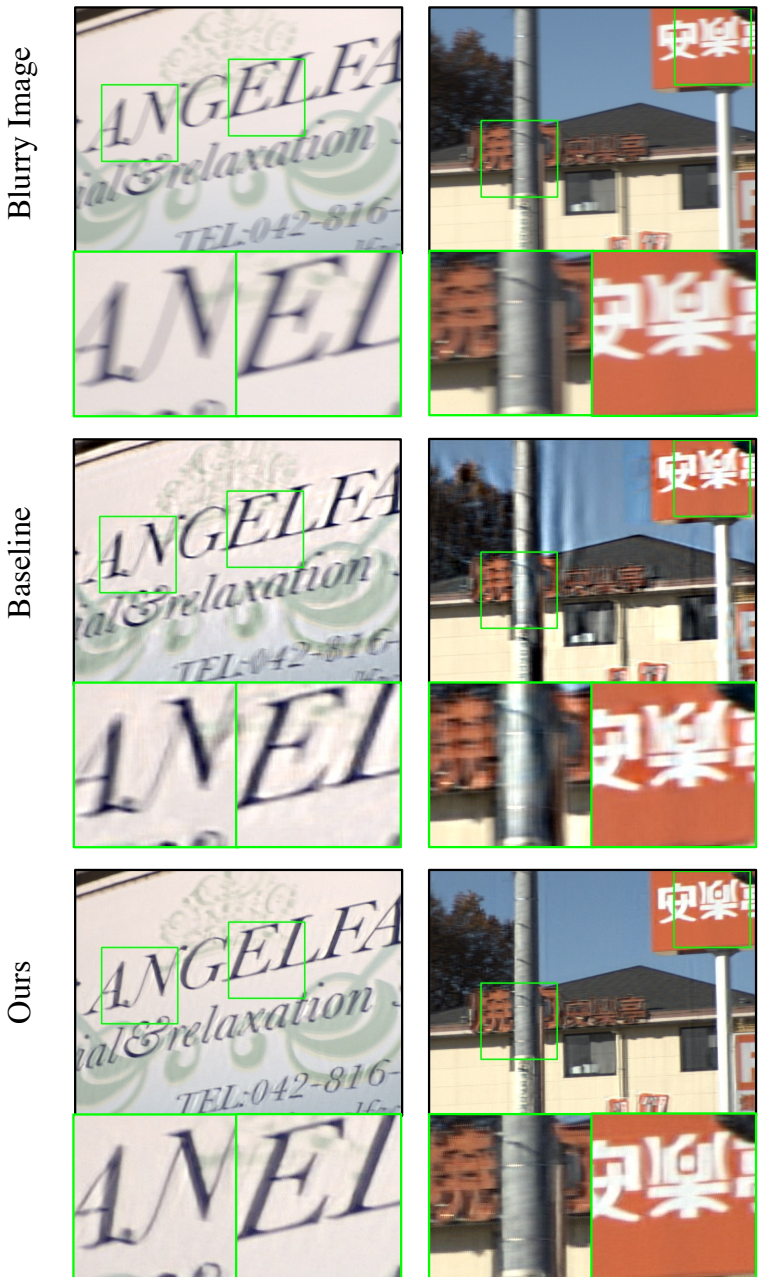

Fig. 12: Qualitative results of ESTRNN [9] on the BSD-2ms16ms [9] dataset.

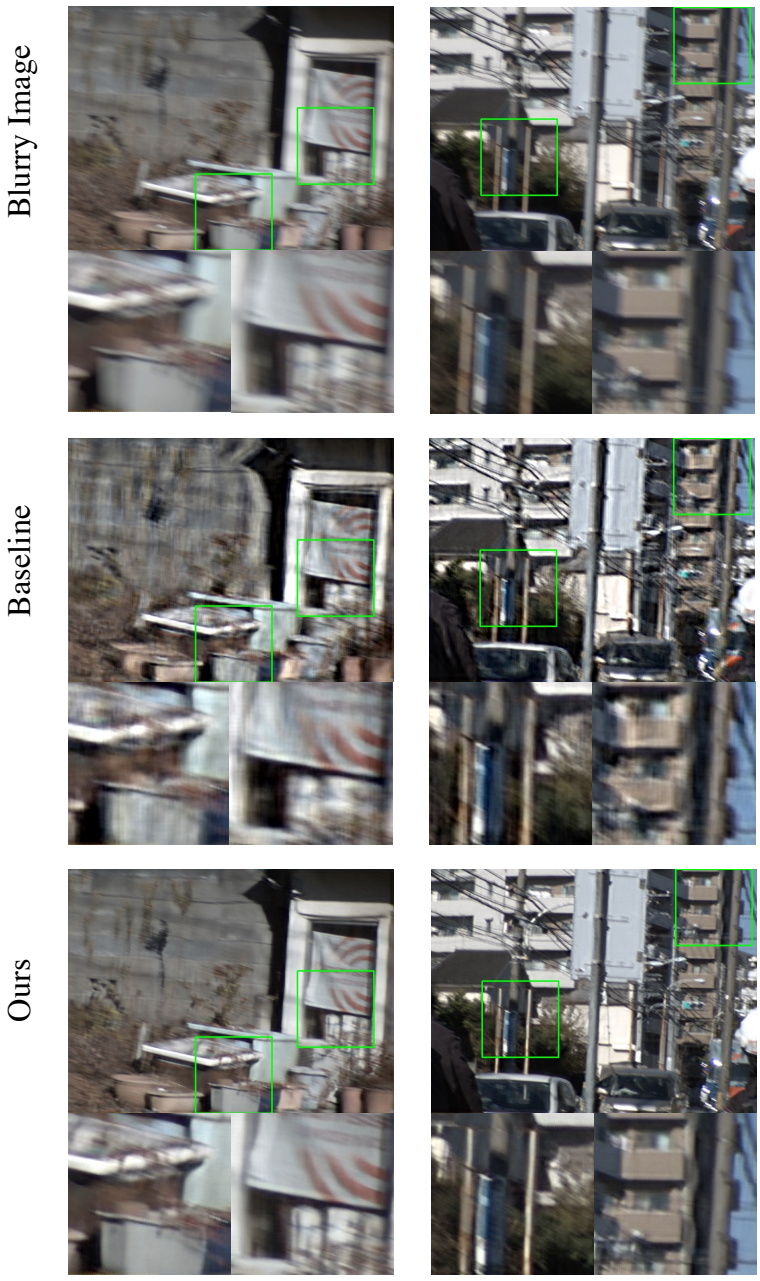

Fig. 13: Qualitative results of ESTRNN [9] on the BSD-3ms24ms [9] dataset.

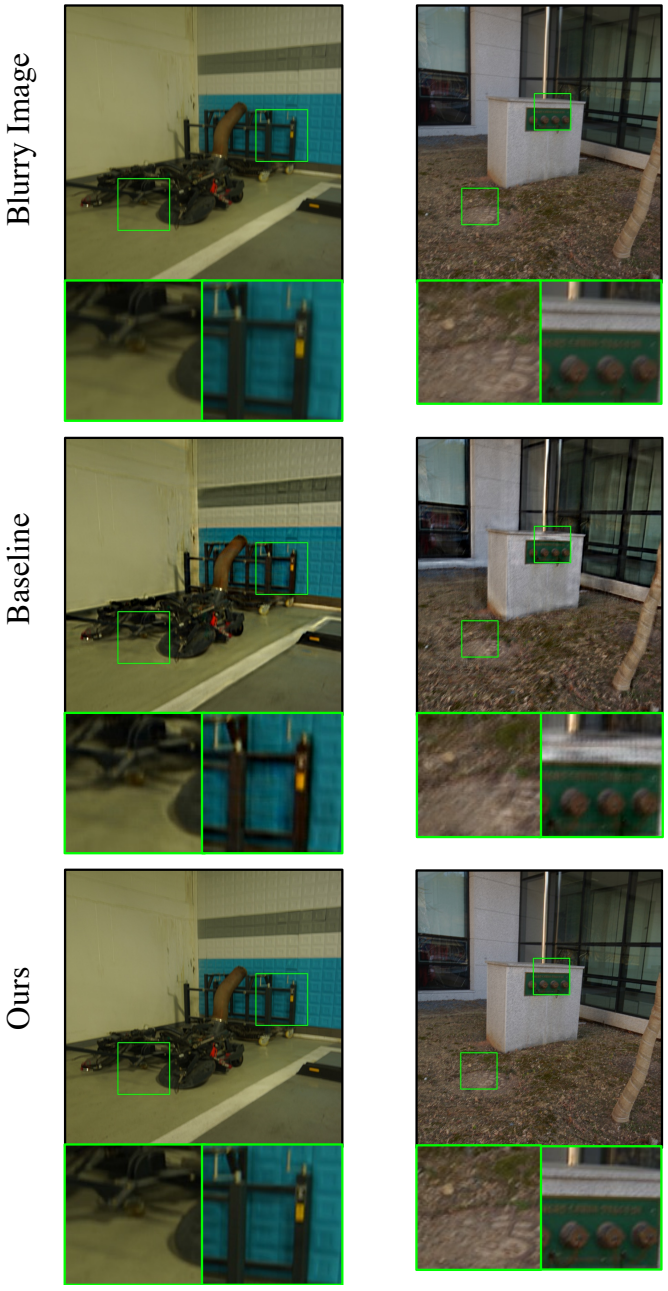

Fig. 14: Qualitative results of ESTRNN [9] on the RealBlur [6] dataset.

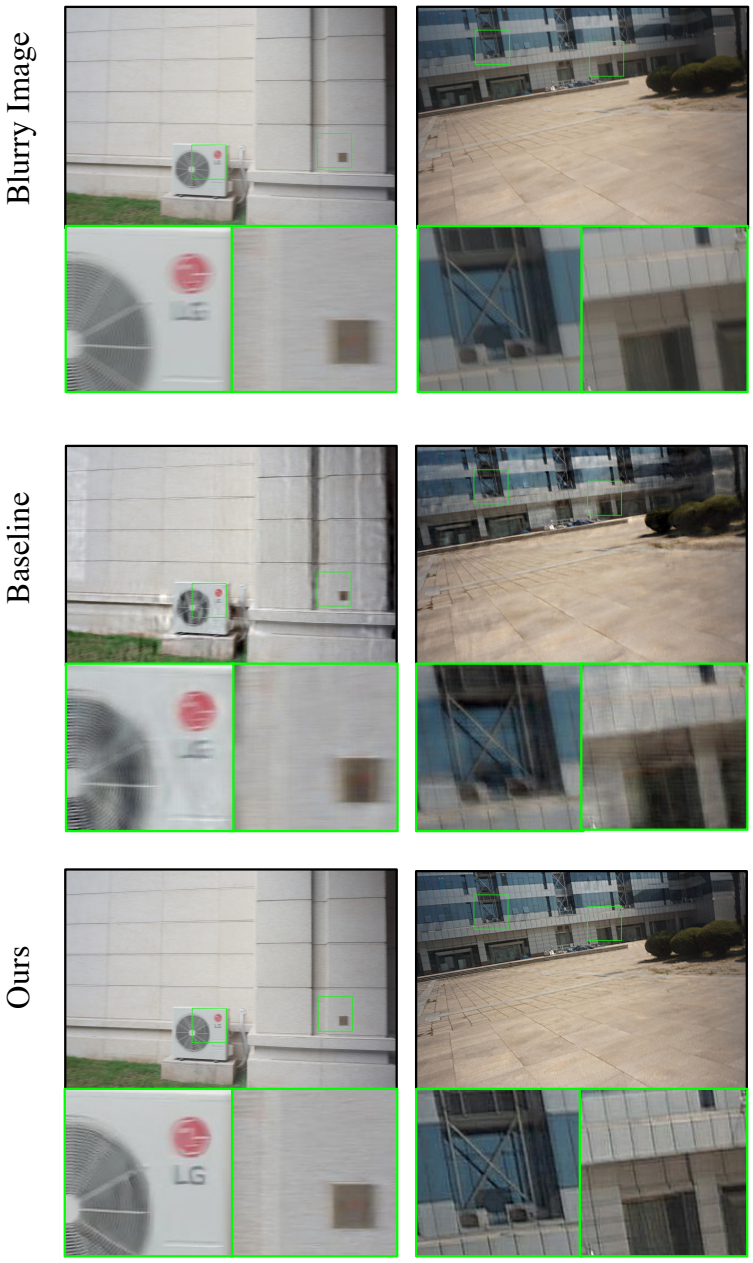

Fig. 15: Qualitative results of ESTRNN [9] on the RBVD [2] dataset.

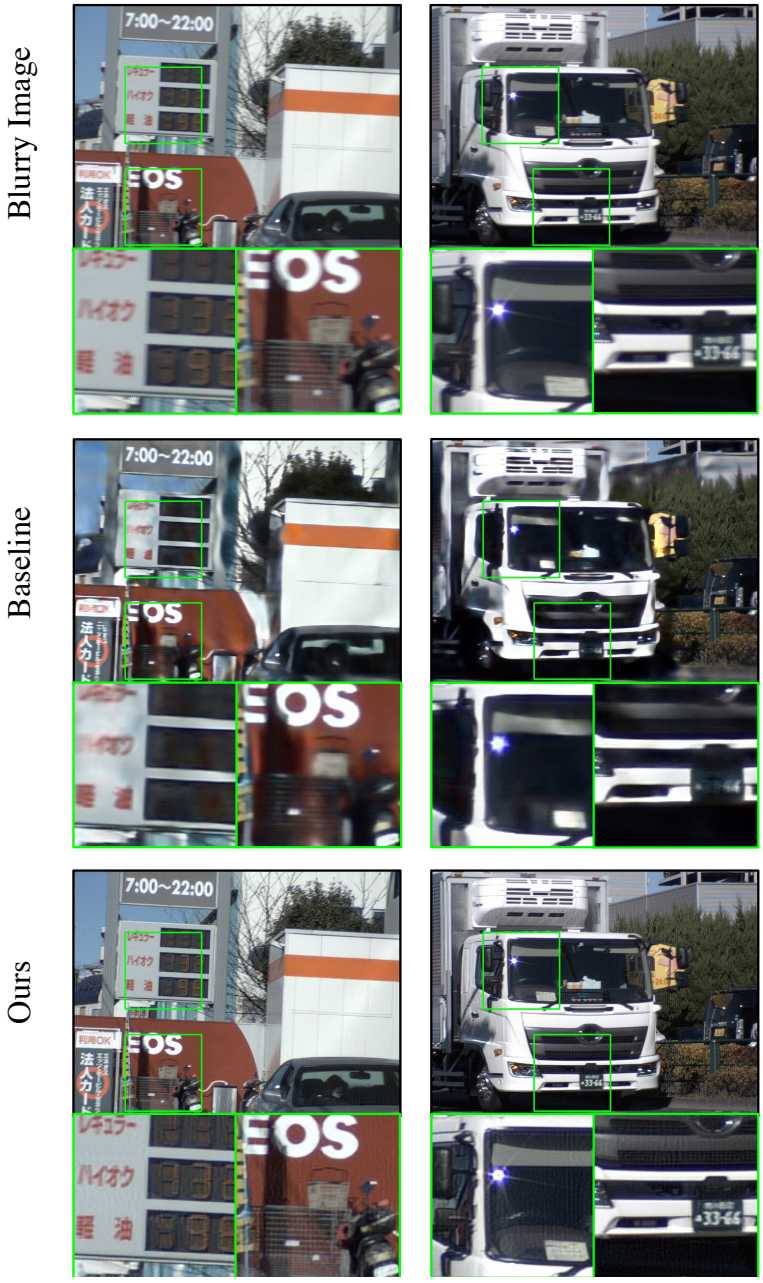

Fig. 16: Qualitative results of MMP-RNN [7] on the BSD-1ms8ms [9] dataset.

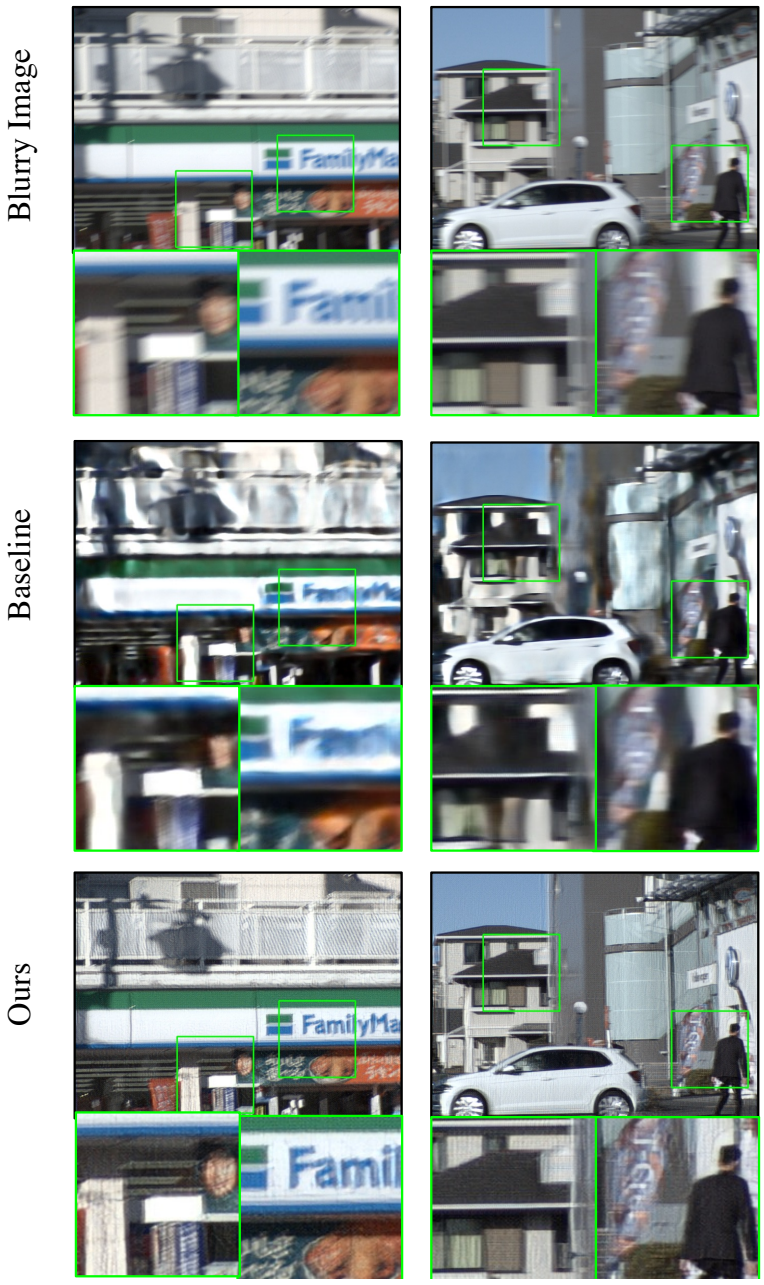

Fig. 17: Qualitative results of MMP-RNN [7] on the BSD-2ms16ms [9] dataset.

Fig. 18: Qualitative results of MMP-RNN [7] on the BSD-3ms24ms [9] dataset.

Fig. 19: Qualitative results of MMP-RNN [7] on the RealBlur [6] dataset.

Fig. 20: Qualitative results of MMP-RNN [7] on the RBVD [2] dataset.

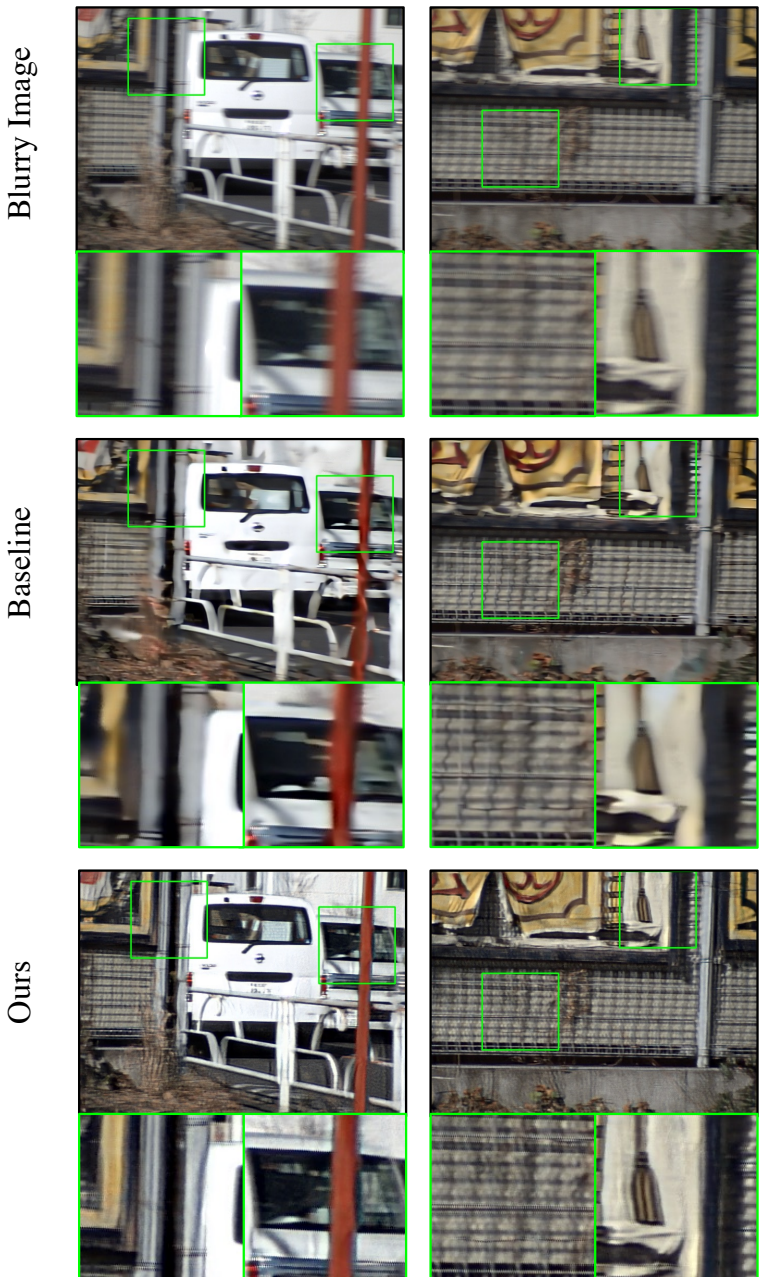

Fig. 21: Qualitative results of Shift-Net [3] on the BSD-1ms8ms [9] dataset.

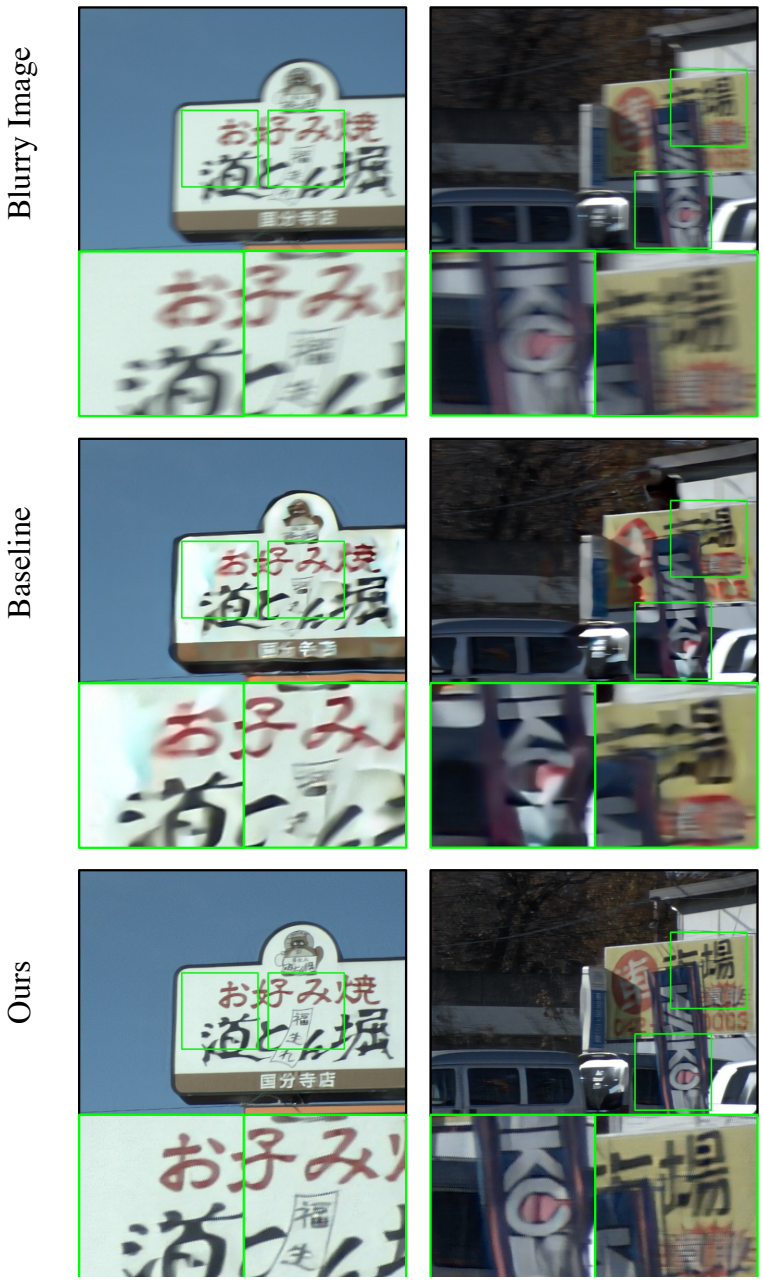

Fig. 22: Qualitative results of Shift-Net [3] on the BSD-2ms16ms [9] dataset.

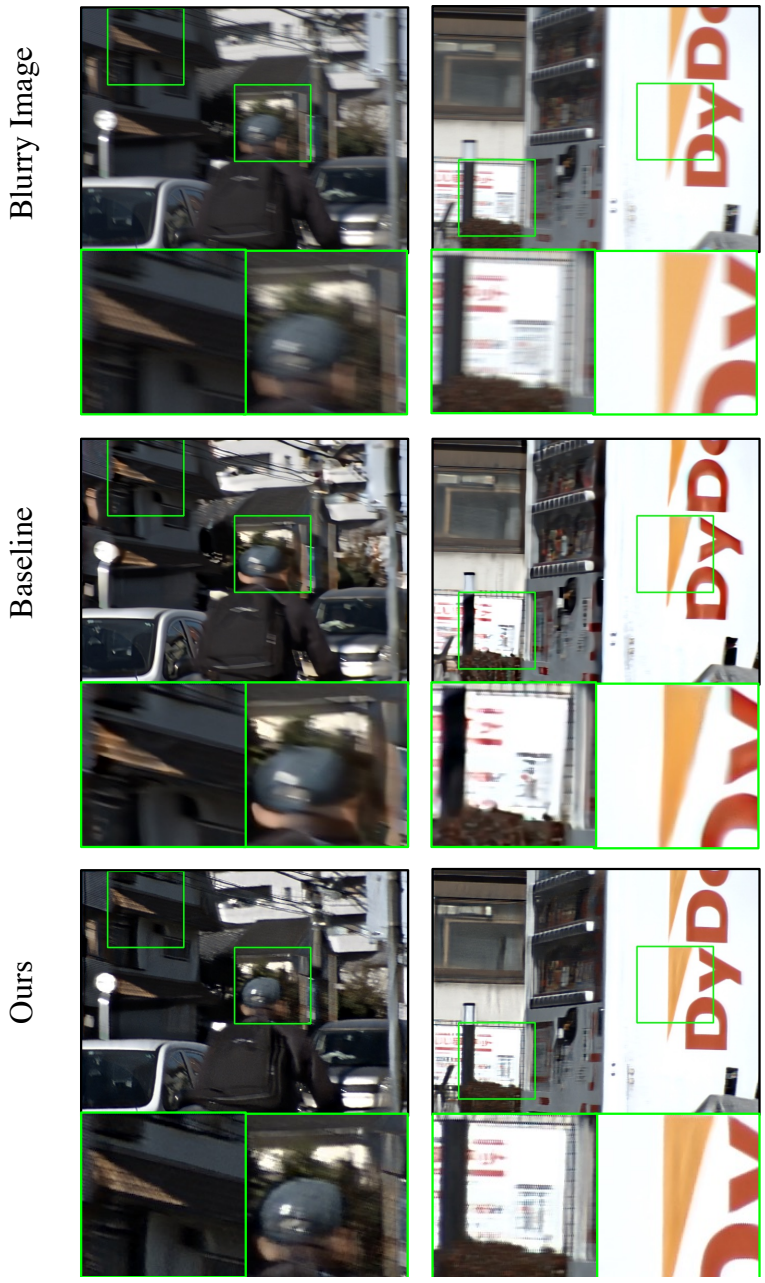

Fig. 23: Qualitative results of Shift-Net [3] on the BSD-3ms24ms [9] dataset.

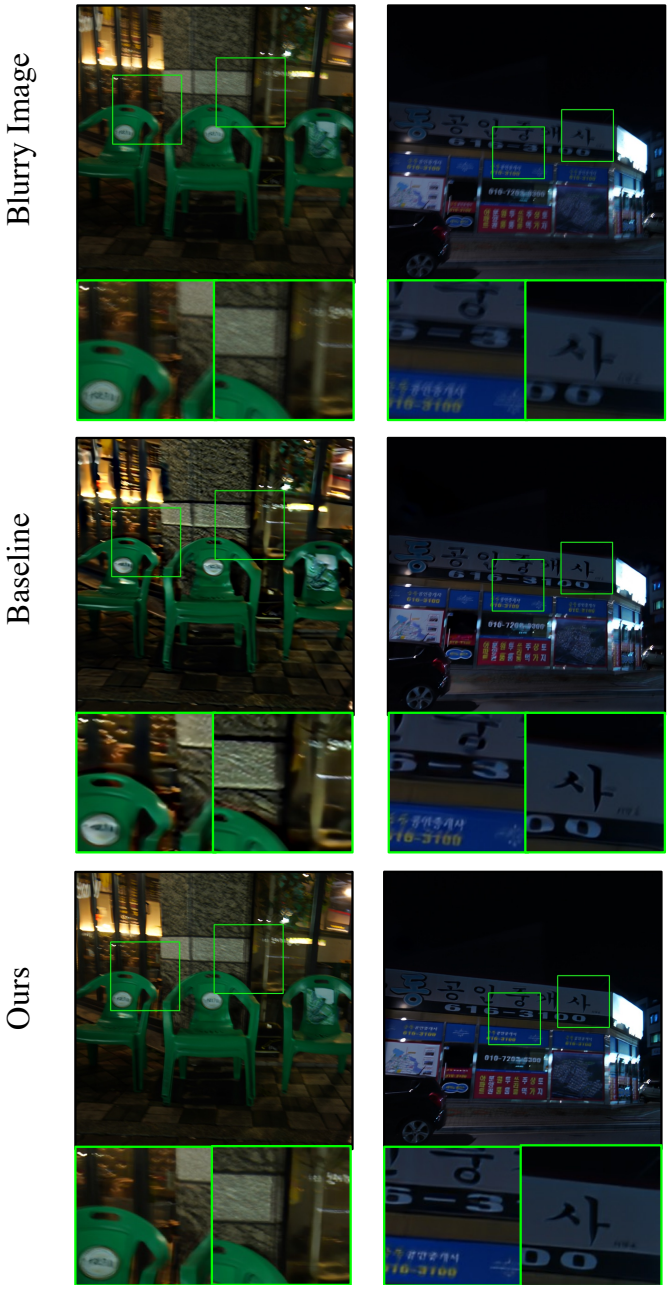

Fig. 24: Qualitative results of Shift-Net [3] on the RealBlur [6] dataset.

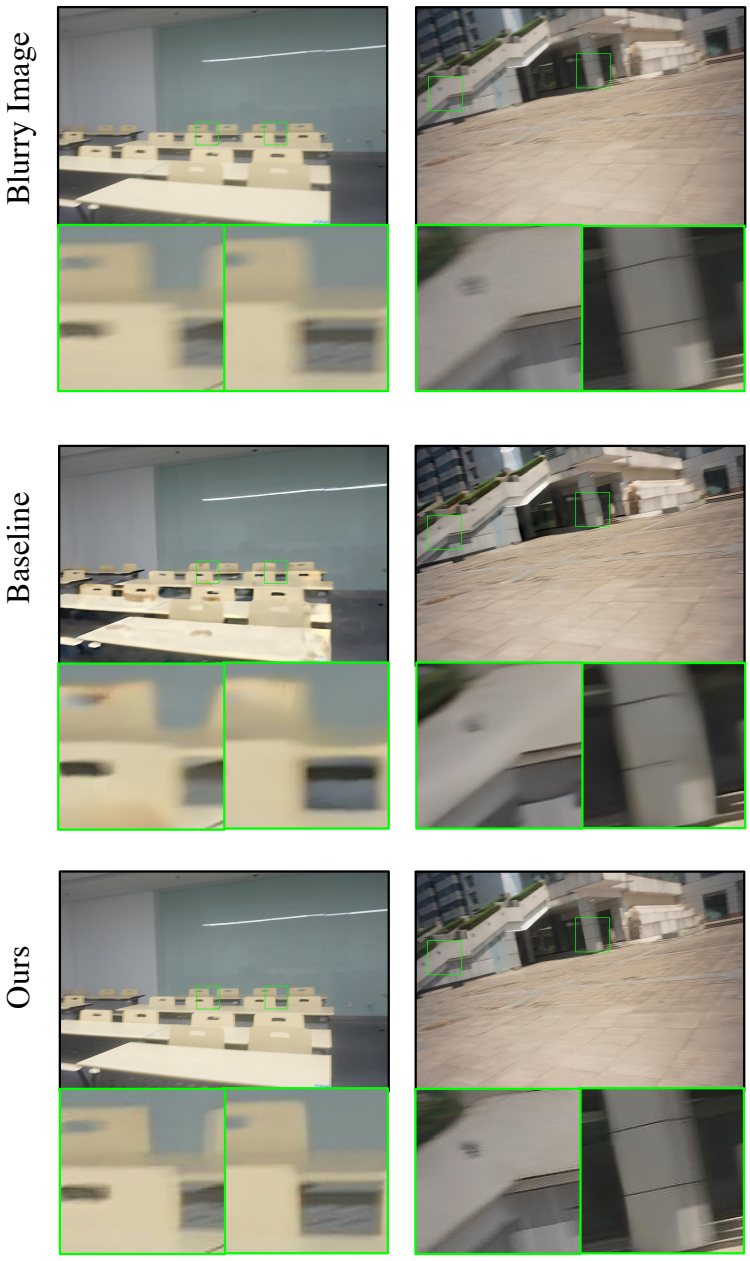

Fig. 25: Qualitative results of Shift-Net [3] on the RBVD [2] dataset.

References

1. https://scikit-image.org/docs/stable/auto_examples/segmentation/plot_regionprops.html
2. Chao, Z., Hang, D., Jinshan, P., Boyang, L., Yuhao, H., Lean, F., Fei, W.: Deep recurrent neural network with multi-scale bi-directional propagation for video deblurring. In: AAAI (2022)
3. Li, D., Shi, X., Zhang, Y., Cheung, K.C., See, S., Wang, X., Qin, H., Li, H.: A simple baseline for video restoration with grouped spatial-temporal shift. In: CVPR (2023)
4. Nah, S., Kim, T.H., Lee, K.M.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: CVPR (2017)
5. Pan, J., Xu, B., Dong, J., Ge, J., Tang, J.: Deep discriminative spatial and temporal network for efficient video deblurring. In: CVPR (2023)
6. Rim, J., Lee, H., Won, J., Cho, S.: Real-world blur dataset for learning and benchmarking deblurring algorithms. In: ECCV (2020)
7. Wang, Y., Lu, Y., Gao, Y., Wang, L., Zhong, Z., Zheng, Y., Yamashita, A.: Efficient video deblurring guided by motion magnitude. In: ECCV (2022)
8. Wu, J.H., Tsai, F.J., Peng, Y.T., Tsai, C.C., Lin, C.W., Lin, Y.Y.: Id-blau: Image deblurring by implicit diffusion-based reblurring augmentation. In: CVPR (2024)
9. Zhong, Z., Gao, Y., Zheng, Y., Zheng, B., Sato, I.: Real-world video deblurring: A benchmark dataset and an efficient recurrent neural network. IJCV (2022)