Supplementary for Visible Watermark Removal with Dynamic Kernel and Semantic-aware Propagation

Xing Zhao
1033874657@sjtu.edu.cn
Li Niu*
ustcnewly@sjtu.edu.cn
Liqing Zhang
zhang-lq@cs.sjtu.edu.cn

MoE Key Lab of Artificial Intelligence
Shanghai Jiao Tong University
Shanghai, China

In the supplementary, we will first introduce the details of our used datasets in Section 1. Then, we will provide more examples of restored background images in Section 2. After that, we will discuss the limitations of our method in Section 3. Finally, we will show the efficiency comparison with baselines in Section 4.

1 Dataset Details

In this work, we conduct experiments on two watermark removal datasets: CLWD [6] and LOGO30K [1], which will be detailed next.

CLWD [6]: In CLWD, the training (*resp.*, test) set contains 60,000 (*resp.*, 10,000) images. The training set contains 160 different watermarks and the test set contains 40 different watermarks. In the training (*resp.*, test) set, the background images are randomly selected from the training (*resp.*, test) set of PASCAL VOC2012 [2]. The watermarks are gathered from public image websites. When overlaying the watermark on the background image, the rotation angle, size, position, and transparency of the watermark in each watermarked image are randomly decided. The transparency of the watermark is set within the scope of [0.3, 0.7].

LOGO30K [1]: LOGO30K contains 28352 training images and 4051 test images. The background images are selected from the VAL2014 subset of MSCOCO [5] dataset. The watermark collection includes more than 1k different and famous logos from the Internet. The transparency of the watermark are in the range of (0.35, 0.85). All background images and watermarks do not overlap in the training and test partitions.

2 More Visualization Results for Watermark Removal

In this section, we visualize the watermark-free image produced by our method and baselines (SLBR [4], SplitNet [1], BVMR [3], and WDNet [6]) on CLWD [6] dataset (see Figure 1)

*Corresponding author.
© 2022. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

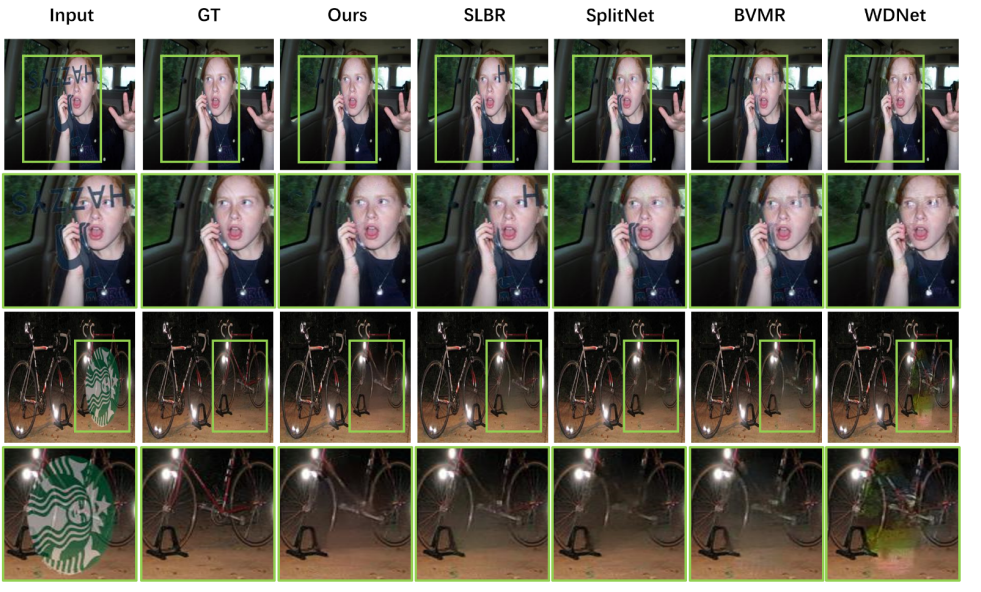

Figure 1: Visualization of watermark-free images produced by various methods on CLWD [6]. “Input" indicates the watermarked image and “GT" indicates the ground-truth background image. We zoom in the watermarked region in green bounding box for better observation.

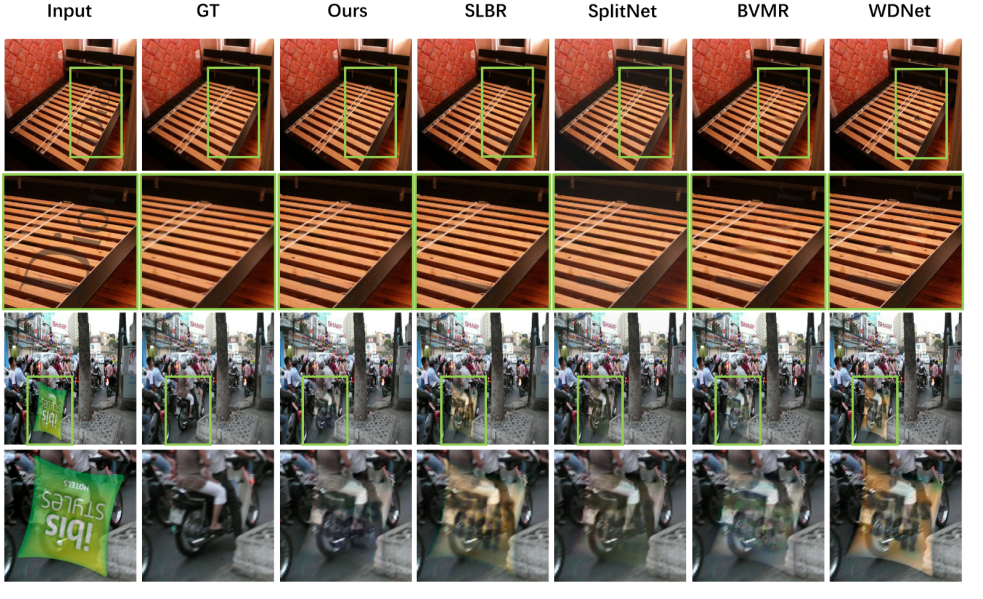

Figure 2: Visualization of watermark-free images produced by various methods on LOGO30K [1]. “Input" indicates the watermarked image and “GT" indicates the ground-truth background image. We zoom in the watermarked region in green bounding box for better observation.

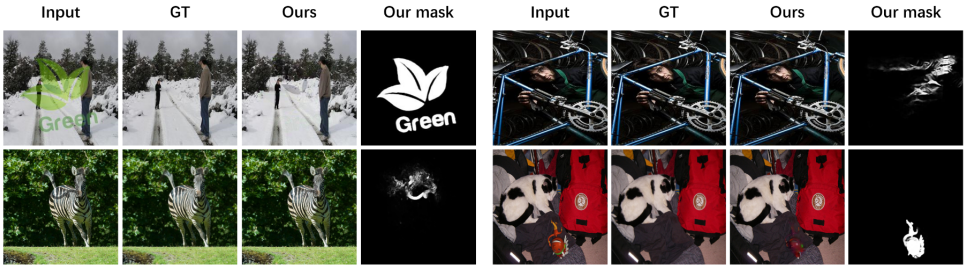

Figure 3: Example of failure cases of our approach on CLWD [6] (the first row) and LOGO30K [1] dataset (the second row). “Input" indicates the watermarked image and “GT" indicates the ground-truth background image.

and LOGO30K [1] dataset (Figure 2), from which we can observe that our method can more accurately restore the textures and patterns in the watermark areas. For instance, in the first example of Figure 1 and the first example of Figure 2, our model can remove the watermark text well, while the results of baselines have obvious residual text (*e.g.*, the letter “H" on the face and the letter “D" on the bed). In the second example of Figure 1 and the second example of Figure 2, the watermark is a color block, which is a more difficult type of watermark to deal with. Our model outperforms the baseline methods in terms of the removal effect, with lighter watermark traces and smaller chromatic aberration. In the results produced by baselines, there exists noticeable color cast, that is, the ground (*resp.*, the motorbike) is tinted with the color of watermark logo in Figure 1 (*resp.*, 2). In our results, the colors of restored ground and motorbike are closer to their original colors.

3 Discussion on Limitation

Although our method can generally localize the watermarked region and restore the background well, there still exist some failure cases, especially when the textures and patterns of the watermark or background are complicated. We show some examples of failure cases in Figure 3, where our method is less effective when the watermark pattern is more complex. In the top left example, our model can remove most of the watermark. However, for the watermarked region with dark color (the dark green leaves), although our model can predict the watermark mask accurately, it cannot completely remove the dark color and there are some remaining leaves. In the bottom right example, the watermark has 3D effect and our model lacks the ability to handle such watermarks. In the bottom left and top right examples, the watermark pattern is complex and the watermark color is similar to the background color, making it difficult to localize the watermark area and thus leading to poor watermark removal results.

4 Efficiency Comparison with Baselines

In Table 1, we compare with two baseline methods in terms of the number of training epochs, training time (hour), inference time (second), and the number of model parameters. All methods are trained and evaluated on Linux version 4.18.0-240.el8.x86_64, with 512GB

	#epoch	Train(h)	Test(s)	#param(M)
SplitNet [1]	50	63	0.022	32.61
SLBR [4]	60	56	0.021	21.39
HRNet+OCR [7]	—	—	0.047	70.56
Ours	50	79	0.071	88.67
Ours (w/o HRNet+OCR)	—	—	0.024	18.11

Table 1: The number of training epochs, training time (hour), inference time (second), and the number of model parameters of different methods on CLWD Dataset [6].

memory, Intel(R) Xeon(R) Platinum 8358 CPU @ 2.60GHz, and one A100 SXM4 40GB GPU. We also report the results of pretrained segmentation model HRNet+OCR used in our semantic-aware propagation module. As shown in Table 1, HRNet+OCR occupies a large proportion of inference time and model parameters, while the complexity of the remaining network is of the same order of magnitude as baselines. For efficiency, we will consider using more light-weighted segmentation model in the future works.

References

[1] Xiaodong Cun and Chi-Man Pun. Split then refine: stacked attention-guided resunets for blind single image visible watermark removal. In *AAAI*, 2021.

[2] Mark Everingham, SM Eslami, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes challenge: A retrospective. *International journal of computer vision*, 111(1):98–136, 2015.

[3] Amir Hertz, Sharon Fogel, Rana Hanocka, Raja Giryes, and Daniel Cohen-Or. Blind visual motif removal from a single image. In *CVPR*, 2019.

[4] Jing Liang, Li Niu, Fengjun Guo, Teng Long, and Liqing Zhang. Visible watermark removal via self-calibrated localization and background refinement. In *ACMMM*, 2021.

[5] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *ECCV*, 2014.

[6] Yang Liu, Zhen Zhu, and Xiang Bai. Wdnet: Watermark-decomposition network for visible watermark removal. In *WACV*, 2021.

[7] Yuhui Yuan, Xilin Chen, and Jingdong Wang. Object-contextual representations for semantic segmentation. 2020.