Physics-Based Generative Adversarial Models for Image Restoration and Beyond Supplemental Material

Jinshan Pan, Jiangxin Dong, Yang Liu, Jiawei Zhang, Jimmy Ren, Jinhui Tang, Yu-Wing Tai and Ming-Hsuan Yang

OVERVIEW

In this supplemental material, we first analyze the effect of the proposed physics-based generative adversarial models in Section 1. Then, we show that the proposed algorithm can be applied to dynamic scene deblurring, and super-resolving blurred images in Section 2-3, respectively. Finally, we show more visual comparisons in Section 4.

1 PERFORMANCE ANALYSIS

As mentioned in Section 6 of the manuscript, the proposed method without the physics model constraint reduces to the GAN model [1] with the loss function (9), which is also similar to pix2pix [2]. We also examine the effect of the proposed physicsbased GAN models on text image deblurring. In this supplemental material, we clarify the differences from the most related methods.

1.1 Additional Comparisons

BaseGAN. Our method without the physics model reduces to the GAN model with the loss function (9). However, without the physics model, this method cannot generate physically correct results where the colors and structures are not preserved well as shown in Figures 1(b) and 2(b).

CycleGAN. CycleGAN (which is designed for the problem when the paired images are not available) involves two discriminative networks and two generative networks. The cycle consistency loss is used to constrain the network. However, as this algorithm is trained on unpaired images, it may lead to the color distortion in the recovered images (see Figures 1(c) and 2(c)).

PCycleGAN. As mentioned in Section 6 of the manuscript, our algorithm involves two discriminators. It is similar to the CycleGAN algorithm. For fair comparisons, we train the CycleGAN algorithm on

the paired images (i.e., PCycleGAN). According to our analysis in Section 6 of the manuscript, training the CycleGAN algorithm on paired images usually leads to trivial solutions due to the cycle consistency loss (see Figures 1(d) and 2(d)).

Although the proposed method is able to recover clear text images, a natural question is whether the recovered character is the ground truth character. To demonstrate the effectiveness of the proposed method, we use the OCR accuracy [3] to evaluate the accuracy of each recovered character. Table 1 shows that the proposed method is able to generate clear text images thus facilitating OCR accuracy.

1.2 Effect of the Adversarial Loss

The discriminator \mathcal{D}_g and its corresponding adversarial loss are used to generate more realistic images. To demonstrate this effect, we compare with the proposed method without using the adversarial loss w.r.t. \mathcal{D}_g on the text image deblurring dataset. In addition, we ignore the discriminator \mathcal{D}_h and its corresponding adversarial loss in the proposed algorithm to demonstrate its effect on text image deblurring. We use the OCR accuracy [3] to measure the accuracy of the recovered characters. The results in Table 1 show that using the discriminators \mathcal{D}_g and \mathcal{D}_h with the corresponding adversarial losses can improve the OCR accuracy.

2 DYNAMIC SCENE DEBLURRING

As mentioned in Section 5.3 of the manuscript, the proposed method can be applied to dynamic scene deblurring. In this supplemental material, we evaluate our method on the dynamic scene deblurring dataset [4]. The proposed algorithm performs favorably against the the state-of-the-art methods as shown in Table 2. The results in Figures 3 and 4 show that the proposed method generates much clean images compared to the state-of-the-art algorithms.

y for every vset- he identifions of gysical neighbor and a vset-path y is always the tp. Wth by showing the the tablic are for meighbors. all do	s for exery vset- he identifiers of wsical neighbor and a vset-path y is always the ap. Wh by showing the the table are for neighbors. all do	ar avony with- indentifiers (eff (path iteal moighthan puths if a visit-path is always the 3.2 Showing the visit a rable are for wintur ighthors. aff dis C C C C C C C C	the influence of partial in the partial i	y for every vset- he identifiers of ysical neighbor and a vset-path y is always the up. by showing the vs. the table are for neighbors. 3 6 1 7 0 8 2 9 10 4	2 Wh et a thu et a thu et a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a thu a th
two major elements w ment:	heo major elements wi ment	o major elements withint	two major elements wi ment:	two major elements wi	
mation flow: through a t Situation	timetion Row: Decuph (itantion	t Sibration	mation flow: throug	h é
in I, root not in o(I). A In model checking [1] unavoidable when I ha offices be abstracted in others be abstracted in relation. The need for or, more generally, prop Simulation is not suffici In order to model part as "unknown" on the l	It is a start is off. A bit start the there is a first unserviced when I has about a first of the start of a start is a start of the second the second and the second the second and the second the second start of Sympletics is so to be at an and the second part is and the second the interview of the second start is a start of the second start is a start is a start of the second start is a start is a start of the second start is a start is a start of the second start is a start of the second is a start of the second start is a start of the second is a start of the second start is a start of the second is a start of the second start is a start of the second start of the second start is a start of the second	the structure of the second se	25 J. top both is of P. A. Bit snowld thereforg [1] innovalided when J has edited by when J has edited by when J has and investigation when and he selation. The word his of, name proceeding prop Sympletions is not infilted by order to reach part	In I , tsut not in $a(I)$. A In model checking [1] unavoidable when I has often be abstracted in and models were used in relation. The need for or, more generally, prop Simulation is not suffici In order to model part as "unknown" on the 1	

Fig. 1. Effectiveness of the proposed physics-based GAN models on image deblurring. The result by BaseGAN still contains blurred characters. The result generated by CycleGAN contains significant blur residual and the background contains some color distortions. The result generated by PCycleGAN still contains significant blur residual. All of these results are in line with our analysis. In contrast, our method generates a much clearer image with recognizable characters compared to BaseGAN.

3 SUPER-RESOLVING BLURRED IMAGES

We note that the recent work [9] focuses on superresolving blurred text and face images. We generate the training dataset according to [9] and compare the proposed method with [9]. The scaling factor for this problem is set to be 4. We randomly select 20 blurred face images for test where the test images and training images are not overlapped. Table 3 shows that the proposed method performs favorably against [9].

4 MORE EXPERIMENTAL RESULTS

In this section, we show more experimental results and compare our method with the state-of-the-art algorithms.



(a) Input

(b) BaseGAN

(c) CycleGAN

(e) Ours

Fig. 2. Effectiveness of the proposed physics-based GAN models on image dehazing. As discussed, the CycleGAN algorithm usually leads to the results with color distortions, e.g., the white blobs in the sky in (c). The PCycleGAN tends to generate trivial solutions. Thus, the results by this method still contain haze residual. BaseGAN does not preserve the main structures of the recovered image. In contrast, our method generates much clearer images.

TABLE 1

Quantitative evaluations of the proposed method using OCR accuracy on the text image deblurring dataset [3]. The proposed method is able to generate clear text images which are close to the ground truth text images.

Methods	CNN [3]	BaseGAN	CycleGAN	PCycleGAN	without \mathcal{D}_h	without \mathcal{D}_g	Ours
OCR accuracy	91.5%	93.0%	49.8%	56.3%	94.2%	91.8%	98.2%

TABLE 2

Quantitative evaluations with the state-of-the-art methods on the dynamic scene deblurring dataset by Nah et al. [4]. "Ours-o" denotes the proposed method which is trained on the uniform blur dataset. "Ours-n" denotes the proposed method without using physics model, which is trained on the dataset by Nah et al. "Ours-D" denotes the proposed method which is trained on the dataset by Nah et al.

Methods	Xu [5]	Pan [<mark>6</mark>]	BaseGAN	Nah [<mark>4</mark>]	DeblurGAN [7]	Ours-o	Ours-n	Ours-D
PSNR	20.30	23.50	26.68	29.23	28.70	29.12	28.45	30.02
SSIM	0.7407	0.8336	0.8862	0.9162	0.9023	0.9018	0.9012	0.9201



(a) Input



(e) Nah [4]



(f) DeblurGAN [5]





4

(c) Whyte [8]

(d) Pan [6]





(g) Ours

(h) GT

Fig. 3. The proposed method can be applied to dynamic scene deblurring and generates comparable deblurred results.



(e) Nah [**4**]

(f) DeblurGAN [5]

(g) Ours

(h) GT

Fig. 4. The proposed method can be applied to dynamic scene deblurring and generates comparable deblurred results.

TABLE 3

Quantitative evaluations of the proposed method on the problem of super-resolving blurred face images. The proposed method is able to generate clear high-resolution face images with higher PSNR and SSIM values.

Methods	Xu [9]	Ours
PSNR	21.75	22.16
SSIM	0.9245	0.9302

3 4 5 6	10110 101101 1011010	5	LOILES MAR	345	10110 10110 101101 101101	5456	10110 101101 1011010
8	101100110	*	Test Book Big	1.1	10110021	8	101100110
9	1011001101	1 Santa	I TRATECTORY	5	102100110	9	1011001101
10 difficult (to observe that	100		5	10-20-0-0	difficult	to observe that
(a	a) Input	(b)	Xu et al. [5]	(c) I	Pan et al. [<mark>10</mark>]	(d) 1	Pan et al. [<mark>6</mark>]
2	1031	3	10110	3	TOTI	3	1011
4	10110	4	10110	4	10110	4	10110
5	101101	5	101101	5	TOTFOT	5	101101
6	2011010	6	1011000	6	1011010	6	1011010
7	20110011	7	10110011	x	IOIPODII	7	10110011
S	201100110	8	101100110	s	101100110	8	101100110
9	2011001101	9	1011001101	9	1011001101	9	1011001101
10	10110011020	10	10110011090	10	10110011010	10	10110011010
difficult	to observe that	difficult	to observe that	40ffrendt	to ebserve ting	difficult	to observe that

(e) Nah et al. [4]

(f) CNN [3] (g) CycleGAN [11]

Fig. 5. More comparisons from text image deblurring dataset [3]. The proposed method generates images with much clearer characters.

ID 2. THE NOT IN HE23. A 编制:#公理解制: 按时用: (因此) #41.4.5。 · · · · In model checking (1) the advanted charactering In model checking (1) uncessidable when I ha consecutivities where I like uncervoldable where I has often be abstracted in and been adventised in a spectrum often be abstracted in . and models were used in and models work ward in and models were used in relation. The need for pediations. The model for relation. The need for en, meder generality, george or, more generally, prop or, more generally, prop Simulation is not onffici Reconfluctures in much stations Simulation is not suffici En order to model part his medicar his manufact game In order to mewled part "undracents" on the l as "manufaterment" and thus as "unknown" on the l 100 (a) Input (b) Xu et al. [5] (c) Pan et al. [10] (d) Pan et al. [6] HA I, DAUL DOL ID B(I). A ta I, Luck mek in et J]. A ALL NESTADA A LOUGE in i, tsut not in a(i). A In model checking [1] In model checking and I water which were a In model checking [1] upprovidable when I has unussoidable when I has unavoidable when I has at lost methods and works othen be abstracted in often be abstracted in often be abstracted in en brace were character time and models were used b and models were used in and models were used in ted by we will we used at relation. The seed for relation. The need for relation. The need for es cover toancage touta or, more gruerade, prop or, more generally, prop or, more generally, prop structures of every plant weither Simulation is not suffici Fundation is not suffici Simulation is not suffici Ex order See reashed pase! ha order to model par In order to model par In order to model part "vertisseers" on itse # as "unknown" on the ! as "unknown" on the 1 28 as "unknown" on the l - - · · · · - - -

(e) Nah et al. [4]

(f) CNN [3]

(g) CycleGAN [11]

(h) Ours

(h) Ours

Fig. 6. More comparisons from text image deblurring dataset [3]. The proposed method generates images with much clearer characters.



Fig. 7. Visual comparisons on face image deblurring. The proposed method generates much clearer face images.



PSNR/SSIM 33.04/0.9370



PSNR/SSIM 32.58/0.9599



PSNR/SSIM 25.98/0.9172



PSNR/SSIM 31.87/0.8215











PSNR/SSIM 30.22/0.8735



PSNR/SSIM 22.20/0.7418



PSNR/SSIM

32.92/0.8774



PSNR/SSIM

33.34/0.8881

PSNR/SSIM

33.60/0.9290

相相民

PSNR/SSIM

33.35/0.8881

PSNR/SSIM

33.82/0.9278



Fig. 8. The proposed algorithm can be applied to image super-resolution (×4) problem and generates comparable results.











8

Fig. 9. Comparisons of image dehazing results. The colors of the stone in (b)-(d) are over-estimated. There still exist haze residuals in (e)-(g). The proposed method generates much clearer images.



Fig. 10. Comparisons of image dehazing results. The result by the dark channel prior contains artifacts as shown in (b). The details of the result in (d) are not preserved well. The results in (e)-(g) still contain haze residuals. The proposed method generates much clearer images.



(e) Ren et al. [21]

(f) Cai et al. [22]

(g) PCycleGAN

(h) Ours

Fig. 11. Comparisons of image dehazing results. The result by the dark channel prior contains artifacts as shown in (b). There still exist haze residuals in (c)-(g). The proposed method generates much clearer images.



Fig. 12. Comparisons of image dehazing results. There still exist haze residuals in (b)-(g). The proposed method generates much clearer images.



(e) Chen et al. [20]

(f) Cai et al. [22]

(h) Ours

Fig. 13. Image dehazing results on the dataset [24]. The proposed method generates much clearer images.



Fig. 14. Image dehazing results on the dataset [24]. The proposed method generates much clearer images.





Fig. 15. Image dehazing on a real image. The proposed method generates a comparable result.



Fig. 16. Image dehazing on a real image. The proposed method generates a much clearer and brighter image.



Fig. 17. Image dehazing on a real image. The proposed method generates a much clearer and brighter image.



Fig. 18. Image dehazing on a real image. The proposed method generates a much clearer and brighter image, where the regions of the sky are better restored.

REFERENCES

- I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial nets," in *NIPS*, 2014, pp. 2672–2680. 1
- [2] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *CVPR*, 2017, pp. 1125–1134. 1, 12
- [3] M. Hradis, J. Kotera, P. Zemcík, and F. Sroubek, "Convolutional neural networks for direct text deblurring," in *BMVC*, 2015, pp. 6.1–6.13. 1, 3, 5
- pp. 6.1–6.13. 1, 3, 5
 [4] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *CVPR*, 2017, pp. 3883–3891. 1, 3, 4, 5
- [5] L. Xu, S. Zheng, and J. Jia, "Unnatural L₀ sparse representation for natural image deblurring," in *CVPR*, 2013, pp. 1107–1114. 3, 4, 5, 6
- [6] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior," in CVPR, 2016, pp. 1628– 1636. 3, 4, 5, 6
- [7] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "Deblurgan: Blind motion deblurring using conditional adversarial networks," in *CVPR*, 2018, pp. 8183–8192. 3
- [8] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *International Journal of Comput*er Vision, vol. 98, no. 2, pp. 168–186, 2012. 4
- [9] X. Xu, D. Sun, J. Pan, Y. Zhang, H. Pfister, and M.-H. Yang, "Learning to super-resolve blurry face and text images," in *ICCV*, 2017, pp. 251–260. 2, 4
 [10] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "L₀-regularized intensity
- [10] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "L₀-regularized intensity and gradient prior for deblurring text images and beyond," *IEEE TPAMI*, vol. 39, no. 2, pp. 342–355, 2017. 5, 6
- [11] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired imageto-image translation using cycle-consistent adversarial networks," in *ICCV*, 2017, pp. 2223–2232. 5
- [12] S. Cho and S. Lee, "Fast motion deblurring," in SIGGRAPH Asia, vol. 28, no. 5, 2009, p. 145. 6
- [13] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in ECCV, 2010, pp. 157–170. 6
- [14] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in CVPR, 2011, pp. 2657–2664.
- [15] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in ECCV, 2014, pp. 184–199. 7
- [16] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image superresolution using very deep convolutional networks," in CVPR, 2016, pp. 1646–1654. 7
- [17] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *CVPR*, 2017, pp. 4681–4690. 7
- [18] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in CVPR, 2009, pp. 1956–1963. 8, 9, 10, 11, 12
- [19] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in CVPR, 2016, pp. 1674–1682. 8, 9, 10, 11, 12
- [20] C. Chen, M. N. Do, and J. Wang, "Robust image and video dehazing with visual artifact suppression via gradient residual minimization," in ECCV, 2016, pp. 576–591. 8, 9, 10, 11, 12
- [21] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *ECCV*, 2016, pp. 154–169. 8, 9, 10
 [22] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An
- [22] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE TIP*, vol. 25, no. 11, pp. 5187–5198, 2016. 8, 9, 10, 11, 12
- [23] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *ICCV*, 2013, pp. 617–624. 10, 11, 12
- [24] http://www.vision.ee.ethz.ch/ntire18/. 10
- [25] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE TIP*, vol. 24, no. 11, pp. 3522–3533, 2015. 12