

RESEARCH

Open Access



A multi-task approach to face deblurring

Ziyi Shen^{1,2}, Tingfa Xu^{1,2*}, Jizhou Zhang^{1,2}, Jie Guo^{1,2} and Shenwang Jiang^{1,2}

Abstract

Image deblurring is a foundational problem with numerous application, and the face deblurring subject is one of the most interesting branches. We propose a convolutional neural network (CNN)-based architecture that embraces multi-scale deep features. In this paper, we address the deblurring problems with transfer learning via a multi-task embedding network; the proposed method is effective at restoring more implicit and explicit structures from the blur images. In addition, by introducing perceptual features in the deblurring process and adopting a generative adversarial network, we develop a new method to deblur the face images with reservation of more facial features and details. Extensive experiments compared with state-of-the-art deblurring algorithms demonstrate the effectiveness of the proposed approach.

Keywords: Convolutional neural network, Face deblurring, Multi-task learning, Transfer learning

1 Introduction

The highly challenging task of estimating a clear image from its degraded blur image is referred to recover the sharp contents and textures. The formation process of image blur is usually formulated as

$$B = I * K + n \quad (1)$$

Where B and I indicate the blurred image and latent sharp image, K is the blur kernel, and n is the additive noise. $*$ denotes the convolution operator.

Image deblurring is an ill-posed problem in computer vision. There has been a remarkable process in the direction with solving the blur kernel and the latent image alternately. The success of state-of-the-art deblurring methods generally rely on empirical statistics of the natural image [1–3] and additional information, such as using the latent prior [4, 5] to constrain this non-convex optimization problem. Furthermore, with the help of predicting the latent edge, these operations usually apply the strong boundaries on blur kernel estimation [6–8]. These implicit or explicit intermediate image properties are computationally expensive and increase the complexity of the estimation process. Recently, the deep neural

network has been applied to image restoration. The CNN-based methods [9–16] are developed to solve the deblurring problem to restore the intermediate properties or the blur kernels. In addition, the framework which utilizes the end-to-end model for direct latent image prediction has also been proposed.

Face deblurring problem has attracted considerable attention due to its wide range of application. While due to the characteristic of faces with less details or explicit edge (i.e., smooth skin with less facial features), some clear image prior-based methods on account of empirical statistics of natural images may not be applied to some specific problems (i.e., face or text deblurring).

To summarize, in this work, we first propose an end-to-end convolutional neural network model to learn effective features from the blurred face images, and then estimate a latent one. To constrain the network, we introduce to utilize a transfer learning framework to learn the multiple features. In addition, we adopt well-established deep networks to obtain extremely expressive features and achieve high quality results. Specifically, we also utilize the generative adversarial network (GAN) to optimize image realistic.

2 Related works

For image deblurring, there are many algorithms proposed to solve this problem. In this section, we summarize the existing methods and put this work in proper context.

*Correspondence: ciom_xtf1@bit.edu.cn

¹School of Optics and Photonics, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing, Beijing, China

²Key Laboratory of Photoelectronic Imaging Technology and System, 5 South Zhongguancun Street, Haidian District, Beijing, Beijing, China

2.1 Image prior and edge prediction

Image deblurring problem is often formulated as an ill-posed problem, which is solved via constraining by assuming of latent prior. A relevant method is based on statistical prior (i.e., heavy-tail distribution). Fergus et al. [1] adopt a mixture of Gaussian to model the statistical prior of the image. Furthermore, Shan et al. [2] and Levin et al. [7] propose to describe the gradient distribution as Hyper Laplacian prior and a piecewise function for image deblurring respectively. To solve the problem, here the maximum a posteriori (MAP) method utilizes the sparse statistical priors to constrain the model.

Except for statistical latent prior, numerous different priors also have been proposed for describing the latent properties. Krishnan et al. [17] introduce a sparsity prior of the clear images. Xu et al. [18] propose to describe the gradient prior via a L_0 constraint. In [19], Michaeli and Irani utilize the patch recurrence to model the image prior. These methods often estimate the blur kernel and clear image alternately via the MAP-based method. These methods often via MAP to estimate the blur kernel and clear image alternately. In addition to the image priors, some methods estimate the blur kernel and clear images via explicitly obtaining salient edge [6, 18, 20]. Summarizing the above discussion, these methods depend on generic priors, statistic of natural images. While the utilized coarse-to-fine model is computationally expensive, and it may not perform well that images do not contain such substantially information in some subjects.

2.2 Convolutional neural networks

Recently, convolutional neural networks have been widely used in image processing. Compared with the aforementioned methods, CNN-based image deblurring methods could be summarized as follows. First, learn effective priors for image deconvolution. Schuler et al. [9] adopt a multi-layer perception (MLP) to process the images with defocus blur. Xu et al. [10] propose a singular value decomposition (SVD)-based network to achieve deconvolution with outlier, but it needs to fine-tune the sub-network for each case. Zhang et al. [11] use the convolutional neural network to learn the effective prior of images and deblur the image in a half-quadratic optimization. Second, in contrast to the non-blind deblurring problem, estimating the blur kernel via a CNN-based model also be proposed in some methods. Yan et al. [21] learn a classifier to distinguish the type of the blur kernels and then estimate the parameter of the kernels for each type in two sub-network. Sun et al. [12] propose to use classification network to describe the linearly non-uniform blur kernel and combine Markov random field (MRF) to optimize these patch-wised blur kernels in an approximatively traversed way. Third, in terms of an end-to-end model, CNN-based methods achieve speed gained which is time-

consuming in some existing algorithms. Chakrabarti et al. [13] and Schuler et al. [14] propose to estimate the blur kernel and achieve deblur images in a two-stage framework in space and frequency domain respectively. Furthermore, in view of text and face deblurring, such end-to-end models [15] also are introduced to solve these specific tasks. Due to less texture, such state-of-the-art methods do not perform well on face deblurring.

2.3 Multi-task learning

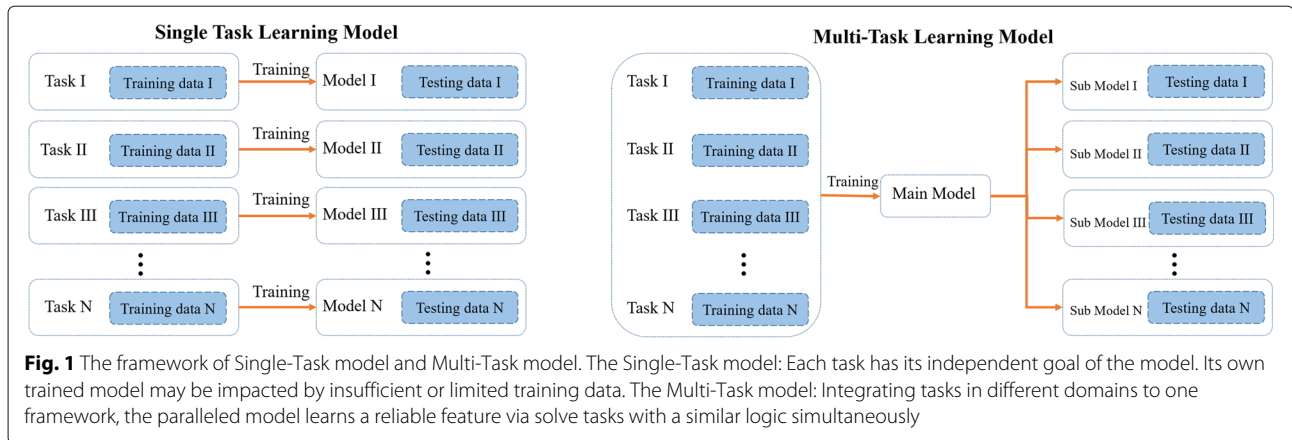
Transfer learning bridging tasks in such domains to target domain is utilized in machine learning and computer vision. This algorithm where a model trained on the source domain or data is purposed to refine the target model. It is exploited to assist the generalization in source task to improve a significant performance in target task. The transfer learning is classified to inductive transfer learning, transductive transfer learning, and unsupervised transfer learning based on the kind of source and target tasks (domains). Multi-task learning is an inductive transfer learning method to solve multiple tasks at the same time. It can result in improving the learning efficiency and prediction accuracy of multiple tasks in the model. Multi-task learning has been widely used in many examples in computer vision (i.e., semantic segmentation [22], classification [23], detection [24, 25], and depth regression). Inspired by these works, we exploit multi-task framework for multiple features learning. We demonstrate that the proposed multi-task learning method would present better constraint compared to the single-task learning.

3 Methods

3.1 Multi-task learning

To describe the following multi-task learning model, We first summarize the single-task learning method. Most machine learning task is a single task. As shown in Fig. 1, each task has its own training model, with training on the specific data, the trained model which is appropriate to a task is independent of others. In a multi-task model, it can optimize more than one task in paralleled learning. For an inductive transfer learning model, the inductive bias is introduced to optimize the model. In the case of multiple tasks model, not only the target task but also the source tasks will impact upon the inductive bias. With the help of domain-specific information that related tasks achieved from the training set, the main (target) tasks utilize this inductive bias to improve the generalization performance of this main task.

Single-task learning which constrained by a L_1 or a L_2 norm for the image content will converge to the homologous solution of the image. In our CNN-based deblurring method, we propose a multi-task framework to propagate the image and structure into the network simultaneously. In contrast to the one task model, we learn a network via



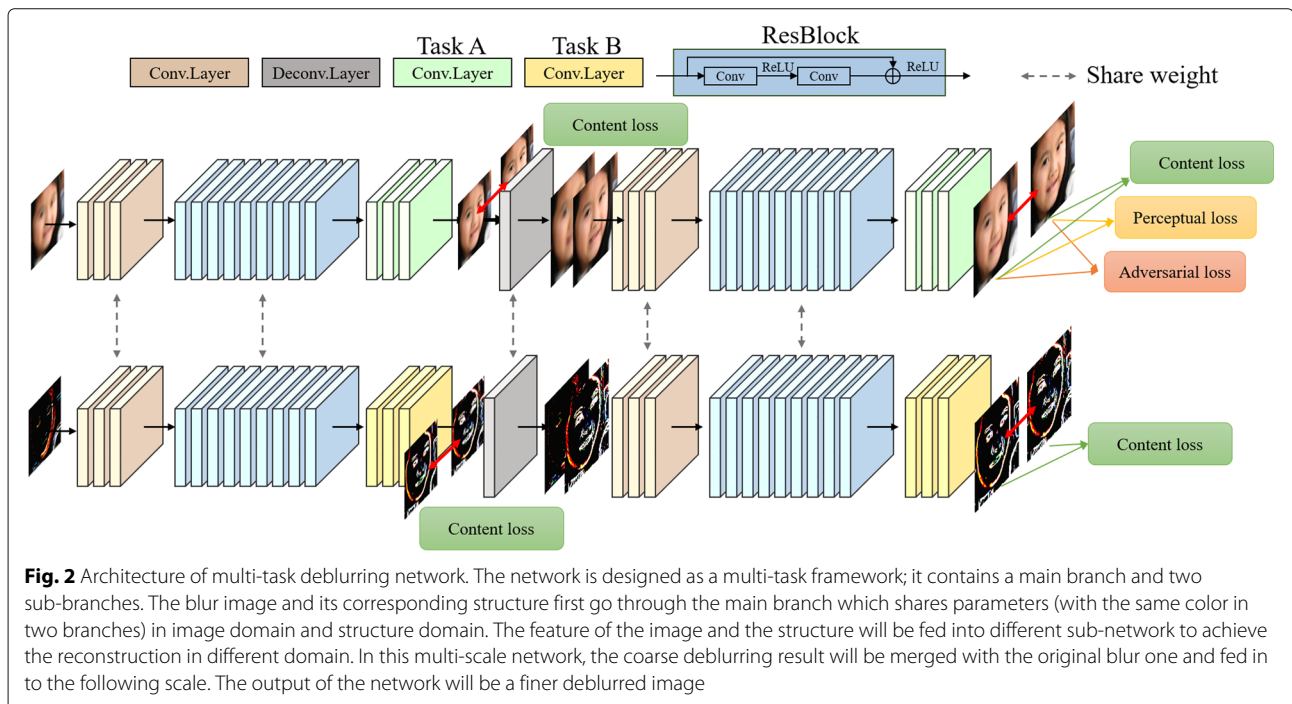
sharing weights between multiple tasks, it facilitates to enforce the sparsity across tasks through norm regularization.

As shown in Fig. 2, the main network (orange and blue layers) shares the weights in image domain and its structure domain. For a single-task model, it focuses on learning features only for one specific task. To achieve the goal of multiple feature learning, the multi-task model learns the feature representation via constraining each task simultaneously. In this way, the latent feature will be estimated with rich structure prior.

3.2 Multi-scale face deblurring network

The aim of face deblurring is to restore clear images with more explicit structure and facial details. Most state-of-the-art methods are employed to estimate the blur

kernel and latent images alternately. Here, a coarse-to-fine framework is applied to solve the iterative problem. As the multi-scale model has been embedded in such framework and it succeeds in gaining implicit or explicit prior from each scale. To address this problem, we employ the multi-scale model to the convolutional neural network framework. For the multi-scale model, the degraded image will be downsampled to a coarse size, and the introduced blur will meanwhile be rescaled to alleviative. It shows in Fig. 2 that we first deblur the image on a coarse scale; it learns to present the features on a lower level. For the upper scale, when combining the coarse reconstruct result, the upper level will accept the lower feature and formulate the coarse-to-fine architecture.



3.3 Synthesis loss function

3.3.1 Content loss

To learn an end-to-end deblurring network, we need to estimate the deblurred images and corresponding structure. Basically, we use a pixel-wise loss to facilitate the convergence of the multi-scale deblurring network. The utilized L_1 criterion describes the difference between latent image I and restored image $\mathcal{G}_1(B)$ for each scale. Especially, in this multi-task deblurring network, the error between structure ∇ of clear images and deblurred images is also considered. The loss is defined as

$$\begin{cases} \mathcal{L}_c = \frac{1}{W \times H} \sum_{n=1}^2 \sum_{x=1}^W \sum_{y=1}^H \|\mathcal{G}_1(B_{x,y}) - I_{x,y}\|_1, \\ \mathcal{L}_{\nabla c} = \frac{1}{W \times H} \sum_{n=1}^2 \sum_{x=1}^W \sum_{y=1}^H \|\mathcal{G}_2(\nabla B_{x,y}) - \nabla I_{x,y}\|_1 \end{cases} \quad (2)$$

Where n indicates the scale of this coarse-to-fine deblurring network. W and H indicate the spatial dimension of the image. Although this loss will assist a global constraint to the content consistency. For the face images deblurring, the goal of this specific problem is to restore the faces with more facial features (e.g., mouths, eyes). Only use L_1 loss in image domain will lead to an overly smooth result. But it is a basic constraint for image reconstruction. The proposed multi-task model with multiple content loss will introduce more structures and details via constraint in structure domain.

3.3.2 Perceptual loss

To restore more details of faces, we introduce to use the perceptual loss on our network. The perceptual loss has been utilized in style transfer and super resolution problems [26–28]. Gatys et al. [29] give an analysis of the texture synthesis based on the feature spaces of convolutional neural networks. The perceptual loss utilizes the obtained high-dimensional features from a high-performing convolutional neural network can assist to restore the image with more natural textures. In the method, we are aiming at achieving more facial features; here, we use the pre-trained VGG19 network [30] for this specific problem. The perceptual loss define as

$$\mathcal{L}_p = \sum_l \|\phi_l(\mathcal{G}_1(B)) - \phi_l(I)\|_2^2. \quad (3)$$

Where the ϕ_l denotes the activation at the l th layer of the pre-trained feature extracting network. In the paper, we choose the Conv1-2, Conv2-2, Conv3-2, Conv4-2, and Conv5-2 layers to acquire the features and compute the perceptual loss.

3.3.3 Adversarial loss

The adversarial loss [31] has been adopted in super-resolution [27], image deblurring [15], and related problems [32]. Ledig et al. have demonstrated that the generative

adversarial networks (GAN) can improve the generative model with a more realistic result. To train the generative adversarial model, it is formulated to solve a min-max problem as

$$\min_{\mathcal{G}_1} \max_{\mathcal{D}} [\log \mathcal{D}(I)] + [\log(1 - \mathcal{D}(\mathcal{G}_1(B)))], \quad (4)$$

$$\mathcal{L}_{\text{adv}} = -\log \mathcal{D}(\mathcal{G}_1(B)).$$

Where the \mathcal{D} denotes a discriminator. We utilize the discriminator of DC-GAN as \mathcal{D} net. We train the discriminator network, it is utilized to distinguish the reconstructed image and the real image. On the other hand, to fool the discriminator, the deblurring network is applied to generate a deblurred image which is more realistic. The adversarial loss \mathcal{L}_{adv} represents that deblurred image allows maximum flexibility for fooling the discriminator.

3.3.4 Overall loss

The overall synthesis loss function is formulated as

$$\mathcal{L} = \mathcal{L}_c + \mathcal{L}_{\nabla c} + \mathcal{L}_p + \mathcal{L}_{\text{adv}} \quad (5)$$

In this work, we utilize the \mathcal{L}_c , $\mathcal{L}_{\nabla c}$, and \mathcal{L}_p in each level of the multi-scale deblurring network and add the \mathcal{L}_{adv} on the final level which the deblurred image with the original size.

4 Experimental

4.1 Implementation detail

We use a multi-scale network for the deblurring problem as shown in Fig. 2. This network has two scales which share weight between each level. The basic model for each scale begins with three convolutional layers; we add six ResBlocks and three convolutional layers after the beginning unit. Especially, except the first convolutional layer sets the kernel size to 11×11 to increase the receptive field, all the convolutional layers have the kernel size of 5×5 with 64 channels. The input of each scale contains six channels, that is to say, we combine the blurred image and deblurred result of lower scale as the input of the upper scale. Here, note that we copy the blurred image (i.e., six channels) as the input of the first scale. In this multi-scale model, we utilize a deconvolutional layer to achieve upsampling. For this multi-task framework, except the last three convolutional layers which utilized to perform reconstruction, the basic deblurring network shares the weight between two tasks (i.e., structure deblurring sub-network and image deblurring sub-network).

We implement the multi-task deblurring network on the Pytorch platform and train the network on NVIDIA Titan X GPU. We set the batch size of 16 and learning rate to $1e^{-4}$. To guarantee the convergence of the multi-task framework, we firstly train the multi-task deblurring network with a content loss for about 5 days. we then add the perceptual loss and adversarial loss individually for joint

training. Specifically, first, we train this multi-task deblurring network using the loss (2) for 100,000 iterations. Second, we embed the perceptual loss (3) for 50,000 iterations. Finally, we add the adversarial loss (4) and jointly train this network for 50,000 iterations.

4.2 Datasets

We use images with the size of $128 \times 128 \times 3$ in all our training and testing experiments. For the multi-scale framework, the image will be downsampled to $64 \times 64 \times 3$ for the first level of the network. Here, we collect 6464 face images from the Helen dataset [33] (2000 images), CMU PIE dataset [34] (2164 images), and CelebA [35] dataset (2300 images) for training. To generate images for training and testing this deblurring network, we synthesize 20000 motion blur kernels based on random 3D camera trajectories generative model [36]. We set the blur kernel size from 13×13 to 27×27 , and convolute them with

images to generate blur images. That is to say, we utilize 130 million face images to train this network. For testing this deblurring network, we choose other 200 images from CelebA and Helen datasets respectively, synthesize other 80 blur kernels, and generate 16,000 blur images to test this model.

4.3 Ablation study

4.3.1 Multi-scale learning

In our work, we use a multi-scale network to formulate this coarse-to-fine deblurring model. To validate the performance of the proposed model, we evaluate the effect of the multi-scale framework. Here, we first train a baseline model which only with a single-scale network for face deblurring. We then build another multi-scale model with two level and compare these two networks. These models are optimized only by content loss. Figure 3 shows the deblurring results for the proposed model with different

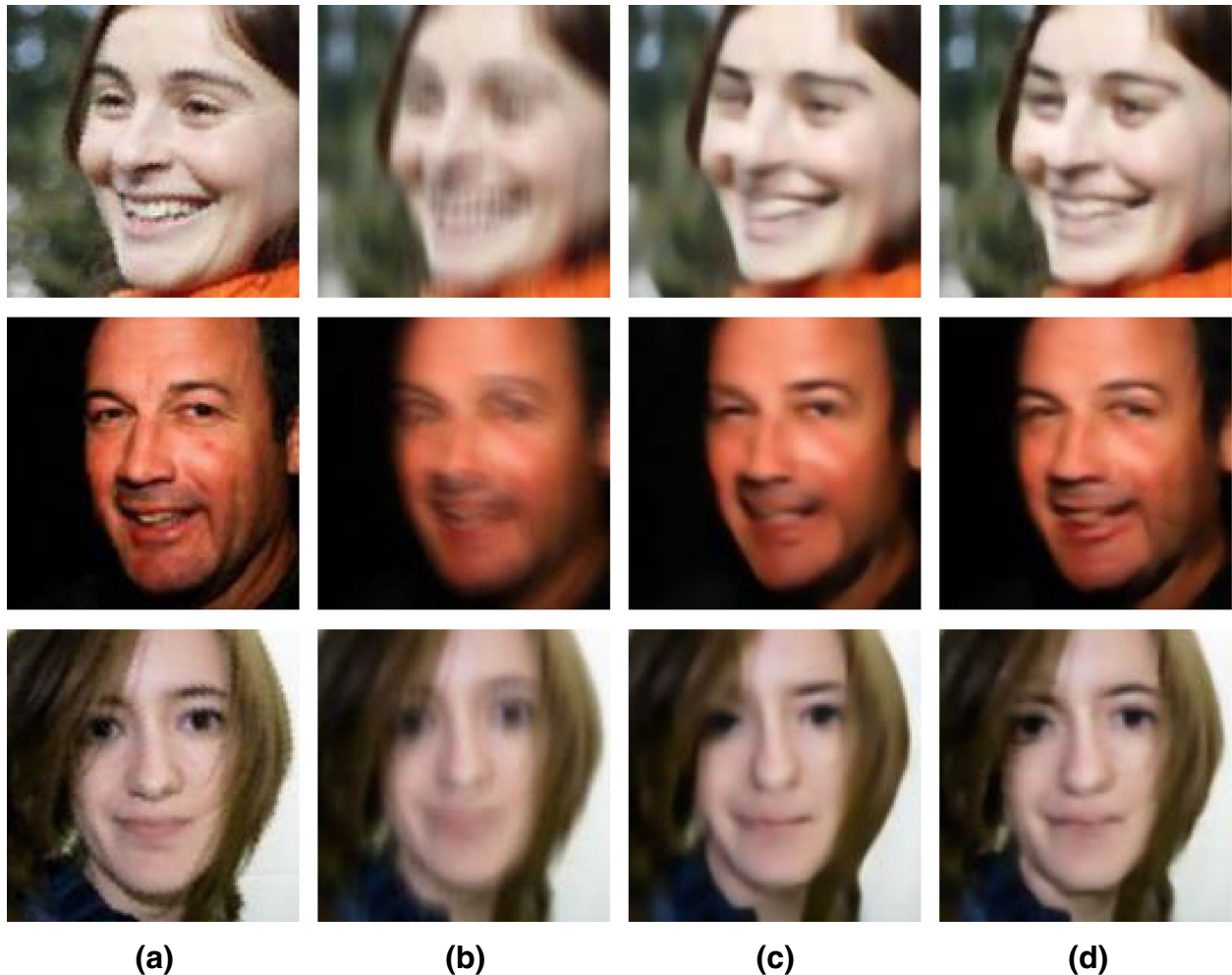


Fig. 3 Ablative study for multi-scale framework. **a** Ground truth images **b** Blurred images **c** Deblurred results via a single-scale network. **d** Deblurred results via a multi-scale network

Table 1 Quantitative evaluation for multi-task framework

Model	Loss	Helen		CelebA	
		PSNR	SSIM	PSNR	SSIM
Multi-scale (1 scale)	Content loss	22.70	0.824	22.19	0.844
Multi-scale (2 scales)	Content loss	23.30	0.847	22.57	0.856
Multi-task (2 scales)	Content loss	23.69	0.852	23.04	0.859

scales. As shown in Fig. 3c, the deblurring result which obtained via a single-scale model produces fewer detail results. The proposed multi-scale architecture is designed as a coarse-to-fine model. Especially, the first level is utilized as a coarse deblurring model; it can achieve a slight reconstruction for the degraded images. In addition, with the help of recovered implicit or explicit feature acquired from the lower level, we feed the original blur images and the upsampled deblurring result into the second level of the model and optimized it with the same

constraint. Therefore, the model is sufficient to achieve a great convergence. Figure 3d shows that the coarse-to-fine model optimized by the content loss and coarse deblurred image lead to an accurate convergence with facial features. Table 1 also demonstrates that the multi-scale framework performs well than the single-task model.

4.3.2 Transfer learning

To further demonstrate the potential of the proposed multi-task learning model, we conduct an analysis on ablation study. Here, we set the multi-scale model constrained by the content loss as the baseline model and train another multi-task model which guides original images and corresponding latent structure simultaneously. Figure 4c illustrates an inevitably produced smooth and ambiguous deblurring result considering the obtained implicit or explicit feature from the blurred images. The goal of image deblurring is to restore a sharp image with more details (facial features, accurate structures, etc.). If



Fig. 4 Ablative study for transfer learning model. **a** Ground truth images. **b** Blurred images. **c** Deblurred results via a single-task model. **d** Deblurred result via a multi-task model

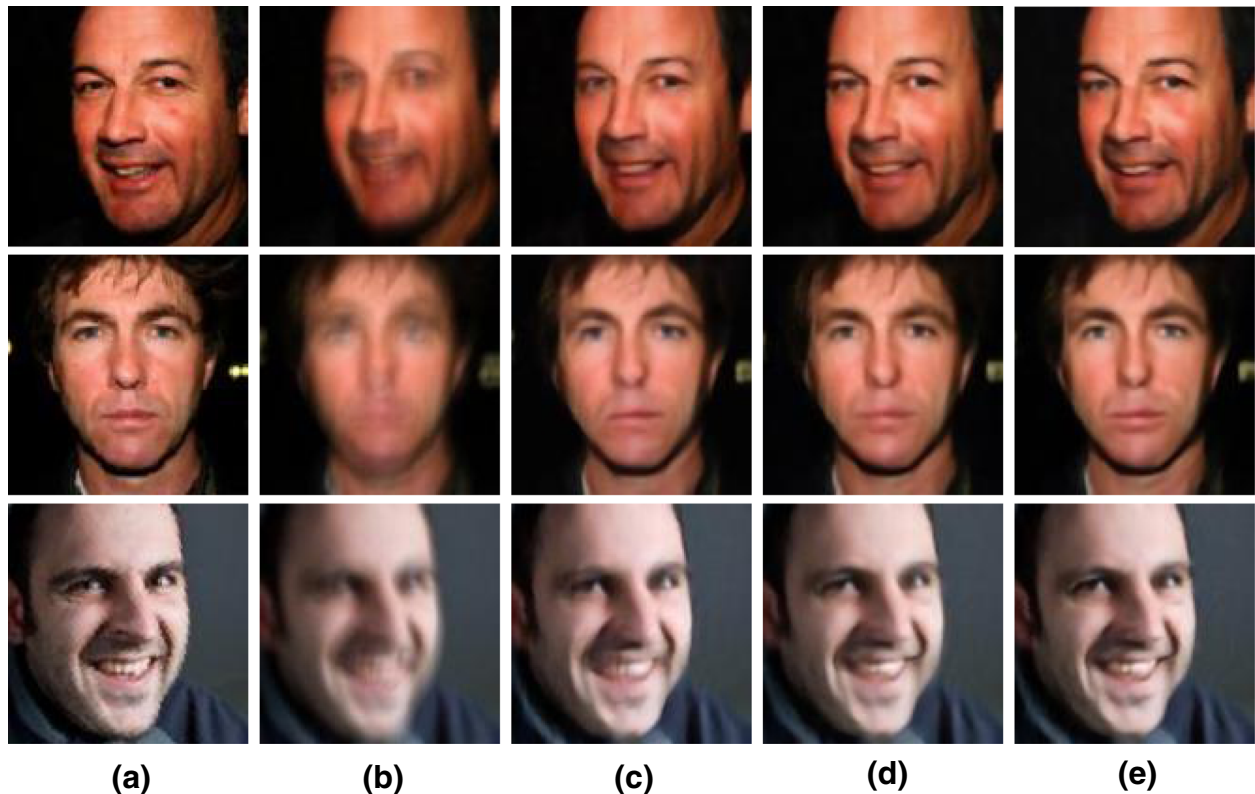


Fig. 5 Effects of additional losses. **a** Ground truth images. **b** Blurred images. **c** Deblurred results (content loss). **d** Deblurred results (content loss + perceptual loss). **e** Deblurred results (content loss + perceptual loss + adversarial loss)

more accurate structures could be obtained, it will constrain the network to a better convergence. We train the network with two goals which include images and structures. The main network learns the capacity of recovering the content and structures simultaneously. The following deblurring sub-network will utilize these features to reconstruct an accurate deblurring image with more details. As shown in Fig. 4d and Table 1, the proposed method performs well for face deblurring. For example, Fig. 4 shows that deblurring face images via the multi-task model can reconstruct more accurate shape of the facial feature; furthermore, more texture can also be restored.

4.3.3 Additional synthesis analysis

In addition to the proposed multi-scale and multi-task model, by constraining the network with perceptual loss and adversarial loss, it will achieve a more realistic result. The visual examples are provided in Fig. 5, and a quantitative evaluation is also shown in Table 2. If the network is just optimized via content loss (L_1), it will achieve a solution with high PSNR or SSIM. However, it will fall to recover more texture and high frequency content from the degraded images. Here, we utilize the pre-trained VGG19 [30] network which can express the sufficient and efficient features of images to constrain the network in feature

domain. To extract the semantic features from such specific layers of the VGG19 [30] network, the perceptual loss can assist to preserve more details and texture from the blur images. In addition to the perceptual loss, we also improve to learn the reconstruction by employing the feedback of adversarial loss. As shown in Fig. 5 and Table 2, the perceptual loss and adversarial loss encourage the deblurring process to a better optimized solution. Figure 5e also demonstrates that it will increase more details (e.g., the tooth and eyes).

5 Result and discussion

We have investigated the effect of the proposed model for image reconstruction via an ablative study. We also compare the performance of the deblurring network with state-of-the-art methods. Here, we give a quantitative and

Table 2 Quantitative evaluation for constraints

Approach	Helen		CelebA	
	PSNR	SSIM	PSNR	SSIM
Content loss	23.69	0.852	23.04	0.859
+ Perceptual loss	24.21	0.857	23.48	0.864
+ Adversarial loss	24.25	0.864	23.56	0.871

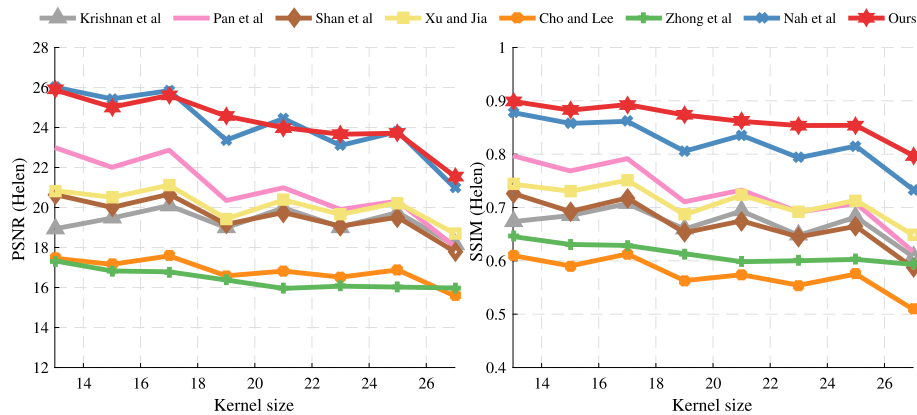


Fig. 6 Quantitative evaluation on Helen dataset. Here, we choose 100 clear images from Helen dataset and convolute these images with 80 blur kernels to generate 8000 blurred images for testing

qualitative evaluation in this section. For this specific reconstruction problem, we also demonstrate our face deblurring method on identity recognition ability.

5.1 Comparisons with state-of-the-arts

We first evaluate the image quality on PSNR and SSIM. We provide seven deblurring algorithms to compare with the proposed method. As shown in Figs. 6 and 7, we compare with the state-of-the-art methods on different size of blur kernels. We evaluate the result on Helen and CelebA dataset, the proposed algorithm performs favorably against the state-of-the-art methods. Here, we also provide a qualitative comparison which is present in Fig. 8. As the MAP-based methods [2, 6, 17, 18, 37, 38] synthesize the deblurred results based on the latent prior. That is to say, an unsatisfactory implicit or explicit prior will directly introduce a failure deblurring result. For example, as shown in Fig. 8d, the face deblurring method [37] depends on the structure to constrain the

iterative solution. The inaccurate prior will lead to an ambiguous result. Furthermore, due to the deconvolution process, the deblurring result will lead to heavy ringing artifacts. We also compare our algorithm with other CNNs-based method (e.g., Fig 8i). Our method performs well on the latent structure and details, as the robust features which are obtained via the multi-task model.

5.2 Face recognition

To further demonstrate the potential of the proposed face deblurring method, we conduct an analysis of the facial feature. We first exploit the identity distance [39] to evaluate the consistency between deblurred face images and their corresponding ground truth images. It is defined as the Euclidean distance of identity feature acquired via a deep convolution model. Except for PSNR and SSIM, it can also evaluate the consistency (i.e., similarity) of the deblurred result. The curve in Fig. 9 depicts the identity

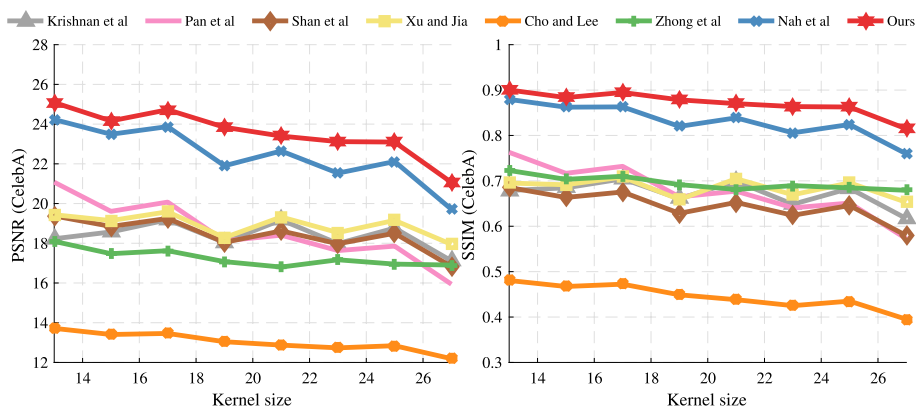


Fig. 7 Quantitative evaluation on CelebA dataset. Here, we choose 100 clear images from CelebA dataset and convolute these images with 80 blur kernels to generate 8000 blurred images for testing

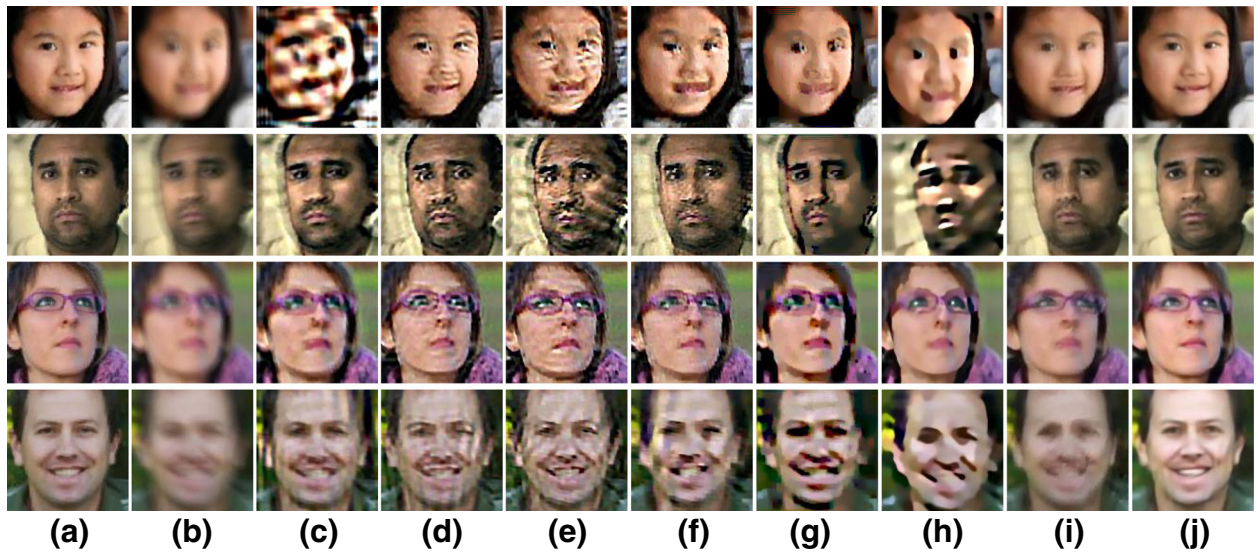


Fig. 8 Comparison with state of the art methods. **a** Ground truth images. **b** Blurred images. **c** Krishnan et al. [17], **d** Pan et al. [37], **e** Shan et al. [2], **f** Xu et al. [18], **g** Cho and Lee [6], **h** Zhong et al. [38], **i** Nah et al. [15], **j** Ours

distance of the state-of-the-art algorithms and the proposed method. The lower distance (error) demonstrates that our method can better match the original images with more accurate facial features. In addition, we also test our method for face recognition. For probe data, we choose 100 face images (8000 deblurred results) with different identities from the CelebA dataset. For each identity, we randomly choose other 9 images (i.e., 900 face images) to generate a gallery data. As the deblurred images do not perform well or reconstruct with such artifacts, it will impact on the recognition. For example, Fig. 8e, h shows that deblurring the image via the state-of-the-art methods

will lead to such ringing or extremely smooth artifacts; the limited facial feature will be faded. During the recognition, the face detection process and similarity evaluation based on identity distance and the ungraded features will result in the poor performance. As shown in Fig. 8j and Table 3, our deblurring result with less artifacts and preserving more details is effective for face detection and recognition.

5.3 Execution time

We also give an evaluation on execution time; it shows in Table 4 that the running time of CNNs-based methods

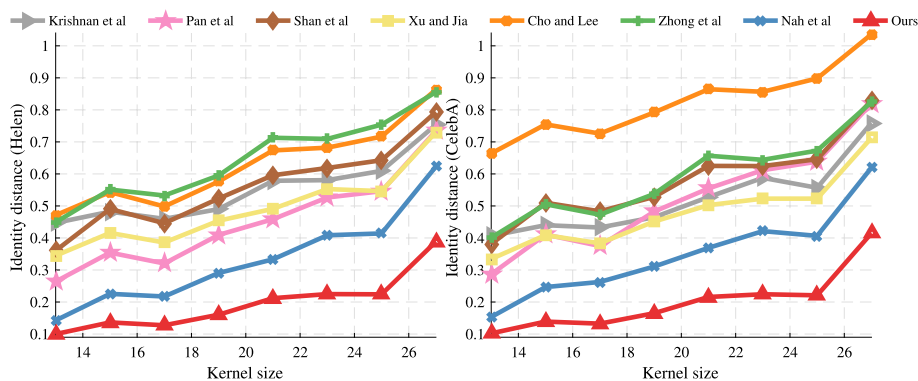


Fig. 9 Evaluation on face identity distance. There are 100 identities chosen from CelebA dataset and we randomly choose other 9 images for each identity. That is to say, the probe data and gallery data are composed of 8000 deblurred face images and 900 images respectively. The curve demonstrates that the proposed method performs favorably against state-of-the-art methods

Table 3 Quantitative evaluation for face detection and recognition on the CelebA dataset

Method	Detection (%)	Top 1 (%)	Top 3 (%)	Top 5 (%)
Clear images	100	71	84	89
Blurred images	77.4	29.1	43.4	51.3
Krishnan et al. [17]	80.0	33.8	48.9	56.6
Pan et al. [37]	78.9	42.0	55.7	62.2
Shan et al. [2]	76.0	32.4	46.9	54.0
Xu et al. [18]	82.5	41.1	55.4	62.1
Cho and Lee [6]	52.2	17.2	27.3	32.5
Zhong et al. [38]	69.5	27.6	41.6	48.5
Nah et al. [15]	86.0	40.1	55.3	62.4
Ours	92	55	69	75

could be significantly improved. For MAP-based methods, the execution time is limited by the alternately iterative solution. In addition, as the proposed framework learns the robust feature via a multi-task network, it makes a trade-off between the task and the redundant parameters. It reports that the proposed framework offers a significant advantage over the state-of-the-art methods in terms of speed.

6 Conclusions

In this work, we propose a deep convolutional neural network to solve the face deblurring problems. The proposed method learns the implicit and explicit features via a multi-task model. It extends the network to learn the robust features in the image domain and structure domain simultaneously via sharing the weight. Due to the exploitation of multi-task framework and multi-scale network, the learned network can achieve fast convergence. It performs favorably against the state-of-the-art methods. In addition, the extensive evaluation also shows that the proposed method is effective for deblurred face detection and recognition.

Table 4 Comparison of execution time. We report the average execution time on 10 images with the size of 128×128

Method	Implementation	CPU / GPU	S
Krishnan et al. [17]	MATLAB	CPU	2.52
Pan et al. [37]	MATLAB	CPU	8.11
Shan et al. [2]	C++	CPU	16.32
Xu et al. [18]	C++	CPU	0.31
Cho and Lee [6]	C++	CPU	0.41
Zhong et al. [38]	MATLAB	CPU	8.07
Nah et al. [15]	MATLAB	GPU	0.09
Ours	PYTHON	GPU	0.02

Abbreviations

CNN: Convolutional neural network; GAN: Generative adversarial networks; MAP: Maximum a posteriori; MLP: Multi-layer perception; SVD: Singular value decomposition; MRF: Markov random field; VGG: Visual geometry group; PSNR: Peak signal to noise ratio; SSIM: Structural similarity index measure

Acknowledgments

The authors would like to thank Tingfa Xu for the support.

Funding

This work was supported by the Major Science Instrument Program of the National Natural Science Foundation of China under Grant 61527802.

Availability of data and materials

All data are fully available without restriction.

Authors' contributions

ZS and TX conceived of the multi-task deblurring method. SJ was responsible for the programming. JG and JZ verified the analytical methods. ZS wrote the manuscript, and all authors revised the final manuscript. In addition, TX is the corresponding author. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 19 September 2018 Accepted: 11 January 2019

Published online: 29 January 2019

References

- R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, W. T. Freeman, Removing camera shake from a single photograph. *ACM TOG (Proc. SIGGRAPH)*, 787–794 (2006)
- Q. Shan, J. Jia, A. Agarwala, High-quality motion deblurring from a single image. *ACM TOG (Proc. SIGGRAPH)*. **27**(3), 73–17310 (2008)
- A. Levin, Y. Weiss, F. Durand, W. T. Freeman, in *Efficient marginal likelihood optimization in blind deconvolution*. Conference on Computer Vision and Pattern Recognition (IEEE, 2011)
- L. Sun, S. Cho, J. Wang, J. Hays, in *Edge-based blur kernel estimation using patch priors*. International Conference on Computational Photography (IEEE, 2013)
- J. Pan, D. Sun, H. Pfister, M. Yang, in *Blind image deblurring using dark channel prior*. Conference on Computer Vision and Pattern Recognition (IEEE, 2016)
- S. Cho, S. Lee, Fast motion deblurring. *ACM TOG (Proc. SIGGRAPH Asia)*. **28**(5), 145–11458 (2009)
- A. Levin, Y. Weiss, F. Durand, W. T. Freeman, in *Understanding and evaluating blind deconvolution algorithms*. Conference on Computer Vision and Pattern Recognition (IEEE, 2009)
- L. Xu, J. S. J. Ren, Q. Yan, R. Liao, J. Jia, in *Deep edge-aware filters*. International Conference on Machine Learning, (2015), pp. 1669–1678
- C. J. Schuler, H. C. Burger, S. Harmeling, B. Schölkopf, in *A machine learning approach for non-blind image deconvolution*. Conference on Computer Vision and Pattern Recognition (IEEE, 2013)
- L. Xu, J. S. J. Ren, C. Liu, J. Jia, in *Deep convolutional neural network for image deconvolution*. Neural Information Processing Systems, (2014)
- J. Zhang, J. Pan, W.-S. Lai, R. W. H. Lau, M.-H. Yang, in *Learning fully convolutional networks for iterative non-blind deconvolution*. Conference on Computer Vision and Pattern Recognition (IEEE, 2017)
- J. Sun, W. Cao, Z. Xu, J. Ponce, in *Learning a convolutional neural network for non-uniform motion blur removal*. Conference on Computer Vision and Pattern Recognition (IEEE, 2015)
- A. Chakrabarti, in *A neural approach to blind motion deblurring*. European Conference on Computer Vision, (2016)
- C. J. Schuler, M. Hirsch, S. Harmeling, B. Schölkopf, Learning to deblur. *TPAMI*. **38**(7), 1439–1451 (2016)

15. S. Nah, T. Hyun Kim, K. Mu Lee, in *Deep multi-scale convolutional neural network for dynamic scene deblurring*. Conference on Computer Vision and Pattern Recognition (IEEE, 2017)
16. Z. Shen, W. Lai, T. Xu, J. Kautz, M. Yang, *Deep semantic face deblurring*, (2018)
17. D. Krishnan, T. Tay, R. Fergus, in *Blind deconvolution using a normalized sparsity measure*. Conference on Computer Vision and Pattern Recognition (IEEE, 2011)
18. L. Xu, S. Zheng, J. Jia, in *Unnatural L0 sparse representation for natural image deblurring*. Conference on Computer Vision and Pattern Recognition (IEEE, 2013)
19. T. Michaeli, M. Irani, in *Blind deblurring using internal patch recurrence*. European Conference on Computer Vision, (2014)
20. J. Pan, Z. Hu, Z. Su, M. Yang, in *Deblurring text images via l0-regularized intensity and gradient prior*. Conference on Computer Vision and Pattern Recognition (IEEE, 2014)
21. R. Yan, L. Shao, Blind image blur estimation via deep learning. *TIP*. **25**(4), 1910–1921 (2016)
22. I. Misra, A. Shrivastava, A. Gupta, M. Hebert, in *Cross-stitch networks for multi-task learning*. Conference on Computer Vision and Pattern Recognition (IEEE, 2016)
23. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun, Overfeat: Integrated recognition, localization and detection using convolutional networks. *CoRR* (2013). <https://doi.org/abs/1312.6229>
24. Z. Zhang, P. Luo, C. C. Loy, X. Tang, in *Facial landmark detection by deep multi-task learning*. European Conference on Computer Vision, (2014)
25. L. Trottier, P. Giguère, B. Chaib-draa, Multi-task learning by deep collaboration and application in facial landmark detection. *CoRR* (2017). <https://doi.org/abs/1711.00111>
26. J. Johnson, A. Alahi, L. Fei-Fei, in *Perceptual losses for real-time style transfer and super-resolution*. European Conference on Computer Vision, (2016)
27. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, in *Photo-realistic single image super-resolution using a generative adversarial network*. Conference on Computer Vision and Pattern Recognition (IEEE, 2017)
28. L. Sun, J. Hays, Super-resolution using constrained deep texture synthesis. *CoRR* (2017). <https://doi.org/abs/1701.07604>
29. L. A. Gatys, A. S. Ecker, M. Bethge, in *Texture synthesis using convolutional neural networks*. Neural Information Processing Systems, (2015)
30. K. Simonyan, A. Zisserman, in *Very deep convolutional networks for large-scale image recognition*. ICLR, (2015)
31. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, in *Generative adversarial nets*. Neural Information Processing Systems, (2014)
32. R. Huang, S. Zhang, T. Li, R. He, in *Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis*. ICCV, (2017)
33. V. Le, J. Brandt, Z. Lin, L. Bourdev, T. S. Huang, in *Interactive facial feature localization*. European Conference on Computer Vision, (2012)
34. T. Sim, S. Baker, M. Bsat, in *International Conference on Automatic Face and Gesture Recognition*. The cmu pose, illumination, and expression (pie) database (IEEE, 2002)
35. Z. Liu, P. Luo, X. Wang, X. Tang, in *Deep learning face attributes in the wild*. International Conference on Computer Vision (IEEE, 2015)
36. G. Boracchi, A. Foi, Modeling the performance of image restoration from motion blur. *TIP*. **21**(8), 3502–3517 (2012)
37. J. Pan, Z. Hu, Z. Su, M. Yang, in *Deblurring face images with exemplars*. European Conference on Computer Vision, (2014)
38. L. Zhong, S. Cho, D. N. Metaxas, S. Paris, J. Wang, in *Handling noise in single image deblurring using directional filters*. Conference on Computer Vision and Pattern Recognition (IEEE, 2013)
39. F. Schroff, D. Kalenichenko, J. Philbin, in *FaceNet: A unified embedding for face recognition and clustering*. Conference on Computer Vision and Pattern Recognition (IEEE, 2015)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)