Enhancing Image Quality via Style Transfer for Single Image Super-Resolution

Xin Deng, *Student Member, IEEE*

***Abstract*—Recently, by feat of the Generative Adversarial Network (GAN), single image super-resolution (SISR) has achieved great breakthroughs in enhancing the perceptual image quality. However, since the network is trained by minimizing the perceptual loss, the GAN based SISR method (SRGAN) [1] results in images with very low objective quality, i.e., peak signal-to-noise ratio (PSNR). In this letter, we aim to solve this problem in an image style transfer way, to generate an image with similar perceptual quality as SRGAN, but with much higher objective quality. Moreover, we propose a threshold-based method to automatically alter the objective and perceptual quality of the reconstructed image through adjusting only one parameter. Experimental results show that our method can achieve more than 1.6 dB PSNR improvement over SRGAN with similar Mean Opinion Score value. Also, with the same objective quality, our method can provide significantly better perceptual results than other state-of-the-art SISR methods.**

***Index Terms*—Image quality, single image super-resolution (SISR), style transfer.**

I. Introduction

SINGLE image super-resolution (SISR) is an inverse problem, which aims to infer a high-resolution (HR) image from a single low-resolution (LR) image. SISR has attracted extensive attention from both academic and industrial communities and has wide applications in fields like medical imaging, infrared imaging, and face recognition [2]–[5].

Enhancing the restored image quality is the final goal of all SISR algorithms. Most SISR algorithms [6]–[18] aim to improve the reconstructed objective quality through minimizing the mean squared error (MSE) between the restored image and the ground truth. However, minimizing MSE is a pixel-wise operation, which can lead to overly smooth results. To make the reconstructed image rich in high-frequency details, Ledig *et al.* [1] proposed a sophisticated perceptual loss to replace the traditional MSE loss, and used generative adversarial network (GAN) to perform super-resolution (SRGAN). However, since the perceptual loss is defined using the high-level features of a well-trained VGG neural network [19], minimizing the perceptual loss cannot guarantee the objective quality that is calculated on the pixel space.

Manuscript received November 26, 2017; revised January 18, 2018; accepted February 5, 2018. Date of publication February 13, 2018; date of current version March 21, 2018. This work was supported by the Imperial CSC scholarship. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Vasileios Mezaris.

The author is with the Electrical and Electronic Engineering department, Imperial College London, London SW7 2AZ, U.K. (e-mail: x.deng16@imperial.ac.uk).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2018.2805809

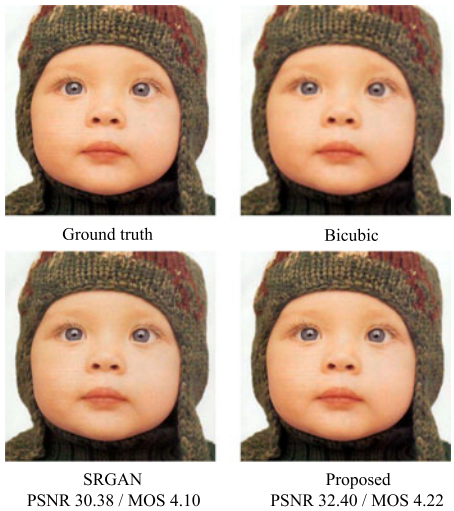


Fig. 1. Examples of 4 × super-resolved images using SRGAN [1] and the proposed approach.

In this letter, we propose a method to relieve the objective quality drawbacks of SRGAN method, to super-resolve an HR image with high objective and perceptual quality simultaneously. Here, the problem is twofold, 1) how to keep the perceptual quality superiority of SRGAN method, and 2) how to improve the objective quality of SRGAN method. For the first task, as we know, the perceptual quality is mostly influenced by the richness of high-frequency details. Thus, for guaranteeing the perceptual quality advantage of SRGAN, we should keep the high-frequency details as much as possible. Secondly, for improving the objective quality, since we can easily obtain super resolved images with high objective quality [8], [16], [17], the key insight is whether we can take advantage of these images to improve the objective quality of SRGAN. Inspired by the recent work of artistic style transfer [20], [21], we treat the high-frequency details of SRGAN method as the style information and the image with high objective quality as the content information, and use image style transfer algorithm to combine them to generate an image with rich high-frequency details and high objective quality simultaneously. Fig. 1 shows an example of the super-resolved images of SRGAN [1] and our method, and we can see that they are perceptually the same, i.e., with similar mean opinion score (MOS) value, but ours provides more than 2 dB higher value in objective quality, i.e., PSNR.

The remainder of this letter is organized as follows. Section II reviews the related work on SISR. Section III introduces the proposed approach. The experimental results are shown in Section IV to verify the effectiveness of our approach, and the conclusions are drawn in Section V to finish this letter.

II. Related Work

In this section, we review the recent works of SISR. Most SISR methods are based on supervised learning, which learn the relationships between LR and HR features through optimizing a defined loss function. According to the loss type, we can classify the SISR works into two categories: MSE loss based and perceptual loss based methods.

MSE loss based methods: are the conventional and prevalent methods to do SISR, with the aim to enhance the objective quality of reconstructed images. They can be further classified into two categories: the traditional machine learning based and the state-of-the-art neural network based methods. The representative works based on traditional machine learning include sparse coding [6], [22], [23], K-SVD [24], A+ [7], Self-Ex [11], random forests [25], etc. Specially, [6] and [24] used dictionary learning to learn over-completed dictionaries with the assumption that the LR and HR counterpart share the same sparse coefficients. The calculation of sparse coefficients is very time-consuming. To handle this problem, A+ [7] proposed to prestore the projections from LR to HR without calculating the sparse coefficients, leading to a vast computational speed-up. Random forests [25] modeled the SISR as a regression problem and directly learned the mapping from LR to HR feature, thus, the computational time can also be saved.

Different from traditional machine learning, the neural network based methods rely on training a deep neural network to get the mappings from LR to HR patches. The representative works are SRCNN [8], VDSR [16], DRCN [26], MSCN [27], ESPCN [17], and SRResNet [1]. SRCNN [8] is the first influential work to show the effectiveness of convolutional neural networks (CNN) in SISR. But its CNN architecture is very shallow, i.e., only three layers, which restricts its performance. VDSR [16] increased the network depth to 20 layers, and only the residual information was learned, leading to a faster convergence speed and higher accuracy. DRCN [26] adopted a deeply-recursive CNN to do SISR. By reusing the weight parameters, it can exploit a large image context.

Perceptual loss based methods: aim to generate a photo-realistic HR image by minimizing the perceptual loss. The representative works are Bruna *et al.* [28], Johnson *et al.* [29], and SRGAN [1]. The key insight is that the perpixel loss, i.e., MSE, is too rigid to measure the perceptual similarity between two images. Thus, they use the perceptual loss instead, which measures the high-level feature representation differences extracted from a pretrained deep neural network, e.g., VGG16. The loss function determines that these methods can generate HR images with high perceptual quality, but low objective quality. In this letter, we propose a method to compensate this drawback, aiming to improve the objective quality of SRGAN method without affecting its perceptual quality too much. This is achieved by using a tailored image style transfer algorithm with the assistance of a restored image with high objective quality.

III. Proposed Method

Inspired by image style transfer, we propose a novel method to enhance the image quality in SISR. As shown in Fig. 2, our framework consists of three steps as follows.

1) Upscaling the LR image by a perceptual loss based SISR method (i.e., SRGAN in this letter) to obtain HR style image with high perceptual quality, and upscaling the LR image by a MSE loss based SISR method to obtain HR content image with high objective quality.

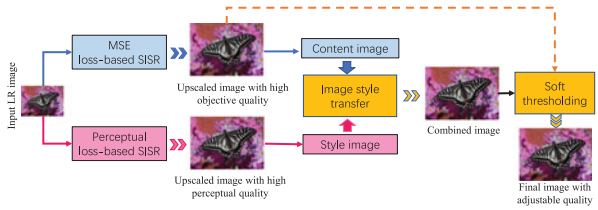


Fig. 2. Framework of the proposed image quality enhancement method.

2) Generate a combined image with the high-frequency information of SRGAN and the content information of MSE loss based method, via a tailored style transfer algorithm.
3) Adjust the objective and perceptual quality of the reconstructed image automatically through a proposed soft-thresholding based algorithm.

Content and style images: Given an LR image, we can use a MSE-based SISR method to reconstruct an HR image I_o and use a perceptual-based SISR method to restore another HR image I_p. Since we wish to generate an image which should contain the global characteristics of I_o and the high-frequency details of I_p, we use I_o as the content image and I_p as the style image in the style transfer process. Here, we regard the high-frequency details in I_p as the "style information".

Image style transfer: Gatys *et al.* [20] proposed an image style transfer algorithm using CNN to synthesize an artistic image with a natural image. Here, we just tailor it to better address our problem, with two main adjustments. First, instead of initializing the synthesized image with random white noise, we initialize it with image I_o which is generated by a MSE-based SISR method. This way, we can guarantee the uniqueness of the final synthesized image and accelerate the convergence speed. Second, we match the content information in a lower layer "conv 2–2", instead of "conv 4–2" in [20]. This is because our task is to improve the objective quality, which means we need to maintain the fine structures and match the pixel information with the content image well. Except these two changes, the whole CNN structure of [20] is kept. For specific structure information, please refer to [20].

To generate the synthesized image I_x, we need to minimize the joint loss of style and content images following [20]:

$$I_x = \min_{I_x} \alpha L_{\text{content}}(I_x, I_o) + \beta L_{\text{style}}(I_x, I_p) \tag{1}$$

where α and β are weighting factors for style and content loss, respectively. Typically, the ratio between α and β is set to 1×10^{-3}. Suppose $F^l \in \mathbb{R}^{N_l \times M_l}$ are the feature representations in the lth layer, with N_l as the number of feature maps and M_l as the size of feature map. F^l_{ij} is the activation of the ith filter at position j in layer l. With F^l and G^l as the feature representations of I_x and I_o at layer l, the total content loss is given by the content loss at the lth layer,

$$L_{\text{content}}(I_x, I_o) = \frac{1}{2} \sum_i \sum_j (F^l_{ij} - G^l_{ij})^2. \tag{2}$$

The Gram matrix $A^l \in \mathbb{R}^{N_l \times M_l}$ is used to represent the style information at the lth layer. Each element $A^l_{i,j}$ is defined as the inner product between the feature maps i and j which is calculated as follows:

$$A^l_{i,j} = \sum_k F^l_{i,k} F^l_{j,k}. \tag{3}$$

TABLE I
COMPARISON RESULTS OF DIFFERENT METHODS ON PSNR, SSIM [33] AND MOS

Set5	Bicubic	A+	SRCNN	ESPCN	SRResNet	SRGAN	Ours
PSNR	28.42	30.28	30.25	30.76	32.06	29.40	**31.30**
SSIM	0.8102	0.8603	0.8624	0.8679	0.8927	0.8347	**0.8771**
MOS	1.74	2.60	2.67	3.06	3.55	3.85	**3.95**
Set14	Bicubic	A+	SRCNN	ESPCN	SRResNet	SRGAN	Ours
PSNR	26.09	27.32	27.44	27.67	28.59	26.11	**27.72**
SSIM	0.7042	0.7491	0.7511	0.7585	0.7811	0.6951	**0.7547**
MOS	1.23	1.86	2.01	2.37	2.85	3.75	**3.67**
BSD100	Bicubic	A+	SRCNN	ESPCN	SRResNet	SRGAN	Ours
PSNR	25.96	26.77	26.90	27.04	27.61	25.18	**26.65**
SSIM	0.6678	0.7085	0.7102	0.7174	0.7370	0.6409	**0.7032**
MOS	1.35	1.78	1.95	2.22	2.56	3.46	**3.44**

With A^l and B^l as the Gram matrices for I_x and I_p, the style loss at the lth layer can be calculated by,

$$L^l_{\text{style}}(I_x, I_p) = \frac{1}{4N_l^2 M_l^2} \sum_i \sum_j (A^l_{i,j} - B^l_{i,j})^2 \tag{4}$$

and the total style loss is the weighted sum through all L layers,

$$L_{\text{style}}(I_x, I_p) = \sum_{l=1}^{L} w_l L^l_{\text{style}}. \tag{5}$$

Here, w_l is the weighting factor of style loss at the lth layer. Typically, $w_l = 1/L$. To find the optimal image I_x in (1), we adopt L-BFGS [30], like [20], to solve the gradient-based optimization problem in (1).

Soft-thresholding: The above transfer algorithm can generate an image I_x with high objective and perceptual quality simultaneously. However, the problem is that once the image is generated, its objective and perceptual quality are both fixed. To make it flexible, we hereby propose a quality adjustment method based on soft-thresholding algorithm. The adjusted image I_s can be obtained through the following formulation:

$$I_s = I_o + S_\lambda(I_x - I_o). \tag{6}$$

In (6), $S_\lambda(\cdot)$ is a pixel-wise soft-thresholding operation which is defined as follows with λ as the threshold:

$$S_\lambda(y) = \text{sign}(y) \cdot (|y| - \lambda)_+ \tag{7}$$

where $(|y| - \lambda)_+$ is defined as

$$(|y| - \lambda)_+ = \begin{cases} 0, & \text{if} \quad |y| < \lambda \\ |y| - \lambda, & \text{if} \quad |y| \geq \lambda. \end{cases} \tag{8}$$

Generally, when increasing the value of threshold λ, image I_s tends to have higher objective quality and lower perceptual quality. This is because larger λ value can remove more high-frequency details and, thus, decreases the perceptual quality. But since these high-frequency details can affect the objective quality, removing them helps improve the objective quality. In an extreme case, when $\lambda = \max(I_x - I_o)$, we have $I_s = I_o$. To keep a good balance between the objective and subjective quality, we should select a λ value between $\max(I_x - I_o)$ and $\min(I_x - I_o)$. However, for different images, the λ value can be quite different, because the pixel values vary greatly across different images. To avoid that, we develop a three-step technique to select λ as follows:

Step 1: Remove all the zero elements in the matrix $I_x - I_o$, and put the remaining K elements in a vector $v \in R^{K \times 1}$.

Step 2: Sort all the elements in v in an ascending order, with the sorted vector denoted as S_v.

Step 3: The threshold $\lambda = S_v(\text{round}(K * \gamma))$, where $\gamma \in (0, 1]$, and the round $(\cdot)$ operation is to round the element to its nearest integer.

After that, the selection of λ turns into the selection of γ ranging from 0 to 1, which is resilient to the image content. The same as λ, a larger value of γ tends to result in an image with higher objective quality and lower perceptual quality.

IV. EXPERIMENTAL RESULTS

A. Experimental Settings

The testing datasets are Set 5 [31], Set14 [24], and BSD100 [32]. All experiments are performed at 4 $\times$ upscaling factor, like [1]. The number of layer L when calculating the style loss in (5) is 5, including layer Relu1-1, Relu2-1, Relu3-1, Relu4-1, and Relu5-1. The iteration time is 1000. When compared with others, the threshold γ is set to be 0.55. In the style transfer process, we use the image super-resolution result of SRGAN as style image, and result of SRResNet as content image. The comparison methods include A+ [7], SRCNN [8], ESPCN [17], SRResNet [1], and SRGAN [1]. All their results are generated through running the source code provided online.

For the MOS evaluation, we invited 15 subjects to rate the super-resolved images with seven versions, including bicubic, A+, SRCNN, ESPCN, SRResNet, SRGAN, and the ground truth images. The images are played in a random order on a 23″ HP LCD monitor with resolution 1920 $\times$ 1080. After the viewing, the subjects are required to rate the image quality with five rating levels, i.e., 1(extremely bad), 2(bad), 3(normal), 4(good), and 5(extremely good). For each dataset, the ratings are performed separately, and the final result is averaged among all the testing images in each dataset.

B. Superiority Over SRGAN

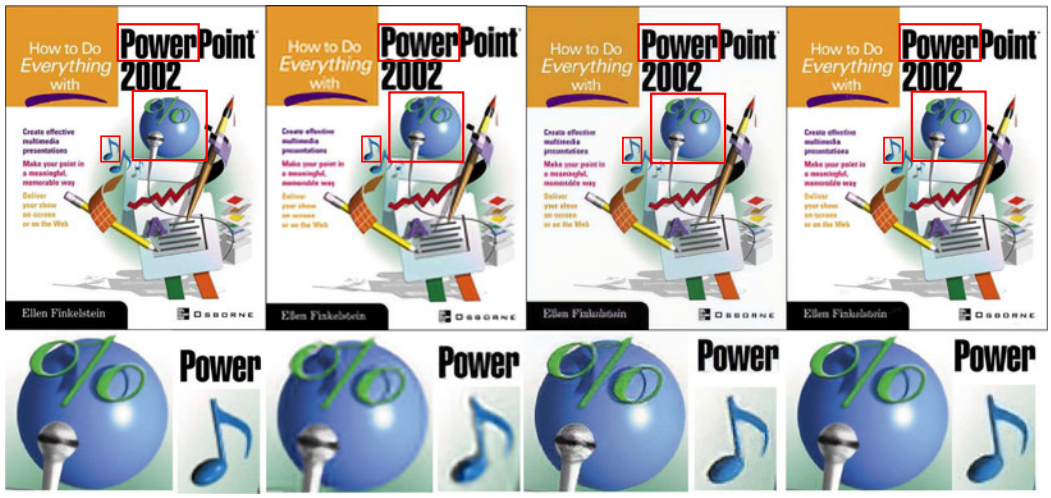
As shown in Table I, our method improves drastically over SRGAN in the objective quality. Specially, our PSNR value is on average 1.90 dB, 1.61 dB, and 1.47 dB higher than SRGAN for Set 5, Set 14 and BSD100, respectively. For the perceptual quality, we achieve nearly the same MOS value as SRGAN. In Set 5, we even perform better than SRGAN. Figs. 3 and 4 compare the visual quality of images generated by SRGAN and our method. We can see that our method provides visually pleasant images while SRGAN has some noisy edges.

C. Superiority Over Other Methods

For Set 5 and Set 14, our method outperforms SRCNN [8] and ESPCN [17] methods in both objective and perceptual quality. As in Table I, for objective quality, we are 1.05 dB and 0.54 dB higher than SRCNN and ESPCN for Set 5. For the perceptual quality, as shown in Fig. 3, the other two methods have overly-smooth results while our method provides sharper and finer edges. For the SRResNet method which we use to generate the content image, the superiority of our method mainly lies in the perceptual quality. From Table I, we can see that our MOS value is much higher than that of SRResNet.

(a) Ground truth (b) SRCNN (22.53 dB) (c) EPSCN (22.85 dB) (d) SRGAN (21.82 dB) (e) Ours (22.85dB)

Fig. 3. Image *Comic* reconstructed using different approaches for 4 × upscaling.



(a) Ground Truth (b) SRCNN (23.83dB) (c) SRGAN (24.34dB) (d) Ours (26.46dB)

Fig. 4. Image *PPT3* reconstructed using different approaches for 4 × upscaling.

Fig. 5. Reconstructed *Lenna* image with different γ. The first row shows the reconstructed image, and the second row shows the details added to the corresponding image in the first row.

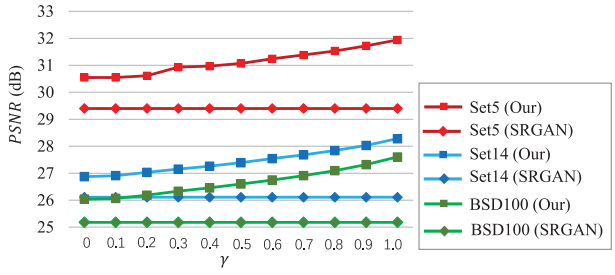


Fig. 6. Effects of parameter γ on the PSNR performance.

D. Effects of Thresholding

By adjusting the value of γ, we can easily change the objective or perceptual quality following the tendency that larger value of γ can increase the objective quality and decrease the perceptual quality. This tendency is clearly shown in Figs. 5 and 6. From Fig. 6, we can see that the PSNR value increases monotonously with the increment of γ. Moreover, regardless of γ value, our method consistently outperforms SRGAN in PSNR. Fig. 5 shows the reason why decreasing γ can improve the perceptual quality. As can be seen, when γ becomes smaller, more high-frequency details are added to the image. Thereby, the perceptual quality can be improved.

V. Conclusion

In this letter, we propose a novel image quality enhancement method to improve the objective quality of GAN based super-resolution method. The key insight is to regard the SRGAN generated image as style image, and another resolved image with high objective quality as content image. Then, we use the image style transfer algorithm to fuse them to generate an image with high objective and perceptual quality simultaneously. To make the quality adjustable, we also propose a soft-threshold based method, which is able to adjust the objective and perceptual quality through only one parameter. Experimental results verify the effectiveness of our method quantitatively and qualitatively.

REFERENCES

[1] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4681–4690.

[2] K. Nasrollahi and T. B. Moeslund, "Super-resolution: A comprehensive survey," *Mach. Vis. Appl.*, vol. 25, no. 6, pp. 1423–1468, 2014.

[3] Y. Zhang *et al.*, "Reconstruction of super-resolution lung 4D-CT using patch-based sparse representation," in *Proc. 2012 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 925–931.

[4] W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 327–340, Jan. 2012.

[5] K. Choi, C. Kim, M.-H. Kang, and J. B. Ra, "Resolution improvement of infrared images using visible image information," *IEEE Signal Process. Lett.*, vol. 18, no. 10, pp. 611–614, Oct. 2011.

[6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[7] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. Asian Conf. Comput. Vis.*, 2014, pp. 111–126.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 184–199.

[9] T. Peleg and M. Elad, "A statistical prediction model based on sparse representations for single image super-resolution," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2569–2582, Jun. 2014.

[10] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. A. Morel, "Single-image super-resolution via linear mapping of interpolated self-examples," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5334–5347, Dec. 2014.

[11] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5197–5206.

[12] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang, "Deep networks for image super-resolution with sparse prior," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 370–378.

[13] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng, and L. Zhang, "Convolutional sparse coding for image super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1823–1831.

[14] Y. Romano, J. Isidoro, and P. Milanfar, "Raisr: Rapid and accurate image super resolution," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 110–125, Mar. 2017.

[15] D. Liu, Z. Wang, B. Wen, J. Yang, W. Han, and T. S. Huang, "Robust single image super-resolution via deep networks with sparse prior," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3194–3207, Jul. 2016.

[16] J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1646–1654.

[17] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1874–1883.

[18] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, "Coupled deep autoencoder for single image super-resolution," *IEEE Trans. Cybernet.*, vol. 47, no. 1, pp. 27–37, Jan. 2017.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.

[20] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2414–2423.

[21] F. Luan, S. Paris, E. Shechtman, and K. Bala, "Deep photo style transfer," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 6997–7005.

[22] S. Yang, M. Wang, Y. Chen, and Y. Sun, "Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4016–4028, Sep. 2012.

[23] S. Yang, Z. Liu, M. Wang, F. Sun, and L. Jiao, "Multitask dictionary learning and sparse representation based single-image super-resolution reconstruction," *Neurocomputing*, vol. 74, no. 17, pp. 3193–3203, 2011.

[24] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Int. Conf. Curves Surfaces.*, 2010, pp. 711–730.

[25] J.-J. Huang, W.-C. Siu, and T.-R. Liu, "Fast image interpolation via random forests," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3232–3245, Oct. 2015.

[26] J. Kim, J. Kwon Lee, and K. Mu Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1637–1645.

[27] D. Liu, Z. Wang, N. Nasrabadi, and T. Huang, "Learning a mixture of deep networks for single image super-resolution," in *Proc. Asian Conf. Comput. Vis.*, 2016, pp. 145–156.

[28] J. Bruna, P. Sprechmann, and Y. LeCun, "Super-resolution with deep convolutional sufficient statistics," arXiv:1511.05666, 2015.

[29] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.

[30] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550–560, 1997.

[31] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. Alberi-Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 1–10.

[32] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. IEEE Int. Conf. Comput. Vis.,* 2001, vol. 2, pp. 416–423.

[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.