



# An End-to-End Deep Learning Framework for Super-Resolution Based Inpainting

Manoj Sharma<sup>1,2(✉)</sup>, Rudrabha Mukhopadhyay<sup>3</sup>, Santanu Chaudhury<sup>1,2</sup>,  
and Brejesh Lall<sup>2</sup>

<sup>1</sup> Central Electronics Engineering Research Institute,  
Pilani 333031, Rajasthan, India

mksnith@gmail.com, schaudhury@gmail.com

<sup>2</sup> Department of Electrical Engineering, IIT Delhi, New Delhi 110016, India  
brejesh@ee.iitd.ac.in

<sup>3</sup> Heritage Institute of Technology, Kolkata 700107, India  
rudrabha@gmail.com

**Abstract.** Image inpainting is an extremely challenging and open problem for the computer vision community. Motivated by the recent advancement in deep learning algorithms for computer vision applications, we propose a new end-to-end deep learning based framework for image inpainting. Firstly, the images are down-sampled as it reduces the targeted area of inpainting therefore enabling better filling of the target region. A down-sampled image is inpainted using a trained deep convolutional auto-encoder (CAE). A coupled deep convolutional auto-encoder (CDCA) is also trained for natural image super resolution. The pre-trained weights from both of these networks serve as initial weights to an end-to-end framework during the fine tuning phase. Hence, the network is jointly optimized for both the aforementioned tasks while maintaining the local structure/information. We tested this proposed framework with various existing image inpainting datasets and it outperforms existing natural image blind inpainting algorithms. Our proposed framework also works well to get noise resilient super-resolution after fine-tuning on noise-free super-resolution dataset. It provides more visually plausible and better resultant image in comparison of other conventional and state-of-the-art noise-resilient super-resolution algorithms.

**Keywords:** Convolutional auto-encoder (CAE)  
Coupled deep convolutional auto-encoder (CDCA) · Deep learning  
Blind and non-blind image inpainting · Super-resolution (SR)  
Noise-resilient super-resolution

## 1 Introduction

Image inpainting aims to reconstruct missing regions and removal of unwanted parts of an image [1, 2]. This area of research has earned a lot of significance over the course of time. Image inpainting can be classified as blind and non-blind

based methods. Blind inpainting is a more complex problem to solve because there is no prior information provided to the network/algorithm about the exact position/location of the corrupted, missing or deteriorated regions of an image. Whereas, in non-blind inpainting the location of the regions to be filled are provided.

Recently, researchers have used deep learning algorithms for blind inpainting [3–6]. These algorithms work well for small region but filling of a big region is still an open challenge. To overcome this issue, we are proposing a new end-to-end deep learning framework for super-resolution based inpainting. Le Meur et al. [7] has shown importance of super-resolution based inpainting to fill larger regions. Aforementioned super-resolution based inpainting method was just a cascading of inpainting and super-resolution algorithm which was non-blind technique. It failed to preserve local structure/information but, our proposed framework jointly optimize both the tasks to get better results for blind inpainting.

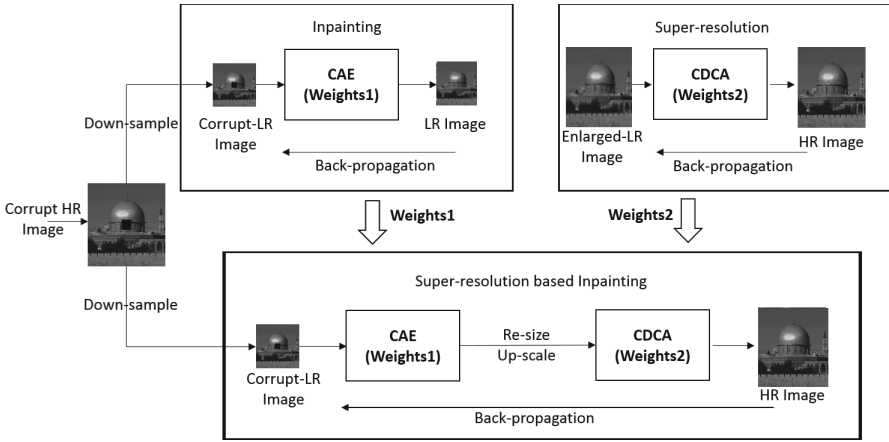
Firstly, a convolutional auto-encoder (CAE) was trained for blind inpainting with down sampled version of images and a coupled deep convolutional auto-encoder (CDCA) [8] was trained for natural image super resolution (SR) separately. This results in learned CAE kernels/filters for blind inpainting and learned CDCA kernels for natural image SR. Then both of these networks were cascaded and the resulting network was initialized as a single integral network (CAE-CDCA) with pre-trained weights of the aforementioned networks. After that, the combined CAE-CDCA was fine-tuned on a data-set having down-sampled version of images with some missing region (corrupt LR images) as input and corresponding higher resolution (HR) ground truth images as a target. Here the parameters of CAE-CDCA is optimized by minimizing final SR loss. Filters/kernels were updated at each iteration and filters were learned for inpainting and super resolution simultaneously while preserving local texture information. The simple cascaded network fails to preserve local structure and high frequency information whereas our integrated CAE-CDCA is able to preserve the local structure by filling the missing region while optimizing inpainting and SR tasks jointly. A block diagram of the proposed framework is shown in Fig. 1.

We can adopt similar framework (CAE-CDCA) with few changes to get noise-free image SR. Sharma et al. [8] has given end to end deep learning framework to get noise-resilient SR. In contrast to stacked sparse auto-encoder (SSDA) used by [8] for image denoising, deep convolutional auto-encoder (CAE) is used in our framework. Since CAE provides a better resultant de-noised image in comparison to SSDA [9]. Firstly, CAE was trained for image de-noising and CDCA was trained for super-resolution separately. This results in learned CAE weights for de-noising and learned CDCA filters for super-resolution. Then both networks were cascaded (CAE-CDCA) and this cascaded network is finetuned as an integral network with pre-trained weights. After that, combined CAE-CDCA was fine-tuned on dataset having noisy-LR images as input and corresponding HR images as a target. Here the loss gradient was back-propagated till the first layer of CAE from the last layer of CDCA. Weights/filters were updated at

each iteration and weights were learned to super-resolve and de-noising image simultaneously while preserving textural information.

The main reason behind using CDCA framework [8] of single image super-resolution (SISR) instead of other deep learning based SISR methods [9–16] is that CDCA architecture with 3 layers provides comparative results with other deep learning based SISR methods. Whereas, other architectures are very deep which require lots of computational complexity.

The rest of the paper is organized as follows. In Sect. 2 we mention related work. Section 3 covers proposed methodology for super-resolution based inpainting. Section 4 shows experimental result to show effectiveness of proposed framework and in last we conclude in Sect. 5.



**Fig. 1.** Block diagram of deep learning framework for super-resolution based inpainting.

## 2 Related Work

In the first category of inpainting, researchers targeted to segment some specific regions from a given image and fill that region with similar patterns and textures present in the background [17]. This category covers only texture synthesis task at a particular location. Second category is of diffusion based techniques [18] which fills the targeted region by diffusing the information from known neighbouring regions. These diffusion based algorithms work well for small targeted regions but fail completely or give a blurring artifact for larger targeted regions. Exemplar based inpainting [15, 17] and sparsity based inpainting [19, 20] lie in the third category. These techniques work better to fill the missing region in comparison to diffusion based techniques but, there is a need to solve complex optimization problems during usage. Le Meur et al. [21] has presented new hybrid inpainting approach which uses both diffusion and exemplar based technique but with large computational complexity. Super-resolution based inpainting method [15] was able to fill larger regions but failed to preserve the local structure.

In the above mentioned methods, it is required to know the location of the missing region. Dong et al. [22] presented a wavelet based technique for blind inpainting and Liu et al. [23] have proposed tensor completion approach for prediction of missing regions. Recently, researchers have presented deep learning frameworks [3–6] for blind inpainting tasks, where deep learning frameworks learn end-to-end mapping between input with some missing region and targeted ground truth. After learning, we have to take the inference of the learned model to inpaint test images with missing region. These algorithms works better in filling smaller regions than earlier traditional approaches but not for larger missing regions. In recent years, generative adversarial network based inpainting approaches [24, 25] have gained popularity for filling large missing regions, but these approaches fills the missing regions with arbitrary details which are quite different from the required context.

### 3 Methodology

#### 3.1 Inpainting Using CAE

Let us consider that down-sampled natural image patches with some missing region (corrupt LR image patches) represented by  $y_i$  and corresponding down-sampled ground truth HR patches be represented by  $x_i \forall i = 1, 2..n$ , where  $n$  is the total number of patches in training dataset. We normalize both the input and target patches between [0 1] as a pre-processing step. We learn the blind inpainting function  $F_1$  which converts  $y_i$  into  $x_i$ .

$$x_i = F_1(y_i, \theta_1) \quad (1)$$

Here,  $F_1$  and  $\theta_1$  are blind inpainting function and parameters respectively. These parameters are same as used in the RED10 [3] architecture. The function  $F_1$  is learned using similar convolution and de-convolution ten layer architecture. Size of patches used for training is  $l \times h$ . The function  $F_1$  is learned by minimizing the following mean square error (MSE):

$$Loss_{Inpainting} = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} \| x_i - F_1(y_i, \theta_1) \|_2^2 \quad (2)$$

#### 3.2 Super-Resolution Using CDCA

Assuming up-sampled low-resolution natural image patches are represented as  $X_i$  and their corresponding ground truth HR patches represented as  $Z_i \forall i = 1, 2..m$  where  $m$  is the total number of patches in SR training dataset. We normalize both the input and target patches between [0 1] as a pre-processing step. We learn the SR function  $F_2$  which converts  $X_i$  into  $Z_i$ .

$$Z_i = F_2(X_i, \theta_2) \quad (3)$$

Here,  $F_2$  and  $\theta_2$  are SR function and parameters respectively. The same parameters of CDCA [8] architecture is used. The function  $F_2$  is learned by using same architecture and parameters as used in [8]. Size of patches used for training is  $s.l \times s.h$ . Where  $s$  is the desired super-resolution scaling. The function  $F_2$  is learned by minimizing the following mean square error (MSE):

$$Loss_{SR} = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} \| Z_i - F_2(X_i, \theta_2) \|_2^2 \quad (4)$$

### 3.3 SR Based Inpainting Using CAE-CDCA

The proposed framework for blind inpainting is shown in Fig. 1. and it comprises of the following steps:

1. Firstly, CAE architecture was learned for coarser version image inpainting. For this, we trained a CAE on a dataset having down-sampled natural image patches with some missing region (corrupt LR image patches) as an input and corresponding down-sampled HR ground truth patches as a target. Learning at coarser level reduces the area to inpaint which helps in preserving local structure information.
2. Then, the CDCA was learned for natural image SR on a dataset with natural LR image patches as input and the corresponding SR HR patches as target.
3. After having learned the CAE filters for blind inpainting and learned the CDCA filters/kernels for image super-resolution, we cascaded both these networks and termed it as CAE-CDCA. Then CAE-CDCA was treated as one integral network with pre-trained weights as shown in Fig. 1.
4. This CAE-CDCA was fine-tuned end-to-end on a dataset which consists of natural LR image patches with some missing region as an input and corresponding HR natural image patches as target. For end-to-end fine-tuning we use  $k$  number of patches in fine-tuning dataset.

After end-to-end fine-tuning, the combined network is jointly optimized for both the task (natural image inpainting at coarser level and SR) and preserves the local structure/information at the same time.

During fine-tuning, the final loss gradients were back-propagated from final layer of the CDCA to initial layer of the CAE. We want to learn the SR based inpainting function  $F$  such that  $Z_i = F(y_i, \theta) \forall i = 1, 2..k$ . Here,  $k$  is the total number of patches in fine-tuning dataset. The kernels/filters were learned to perform image inpainting at coarser level and SR, simultaneously by minimizing the final loss.

$$Loss_{final} = \frac{1}{k} \sum_{i=1}^k \frac{1}{2} \| Z_i - F(y_i, \theta) \|_2^2 \quad (5)$$

After learning the CAE-CDCA, we can inpaint any test image by down sampling that input image and by passing through the feed forward path of CAE-CDCA (taking inference of learned CAE-CDCA).

## 4 Experimental Results

### 4.1 Datasets

To train deep CAE for blind inpainting, we generated a large training database of 0.2 million corrupted and corresponding ground truth patches (size =  $64 \times 64$ ) using imagenet [26] and few images from ETH CIL database [27]. To create blind inpainting database, we corrupt the patches by using random masks at different locations. Our CDCA framework is trained on imagenet dataset for 2x SR dataset as given in [26]. Combined CAE-CDCA framework is fine-tuned on the dataset having down-sampled corrupted patches ( $64 \times 64$ ) as an input and corresponding ground truth patches ( $128 \times 128$ ) as a target. Fine-tuning is also done on imagenet dataset. We tested our framework on BSD300 dataset [28] and on remaining images from ETH CIL dataset. ImageNet dataset was used for fine-tuning our proposed deep learning framework to get noise-free super-resolution.

### 4.2 Experiments

To train CAE for blind inpainting, we take 10 layer architecture having 5 convolutional and 5 de-convolutional layers with ReLU as the activation function as given in RED10 [3]. At each convolutional and de-convolutional layer, the kernel size is  $5 \times 5$ , feature map is 64. we set batch size: 150 and learning rate:  $10^{-4}$ .

To train CDCA for 2x SR of natural LR images, we use similar parameters as given in [8]. To learn convolutional feature maps, we take  $9 \times 9$  filter size for first layer and  $5 \times 5$  filter for the last two layers of the SR module. Feature maps used for first, second and third layers are 64, 32 and 1 respectively. We also set batch size: 150 and learning rate:  $10^{-4}$ . Fine-tuning of CAE-CDCA was done on same parameters with a learning rate of  $10^{-5}$  for the last layer and  $10^{-4}$  for all other layers.

All these deep learning frameworks were trained and tested on a HP Z640 desktop workstation with 64 GB RAM, two Intel Xeon-E-5 processors and with GTX-1080 GPU support. All the experiments are performed on gray scale images. But the same framework can be extended to work on color images.

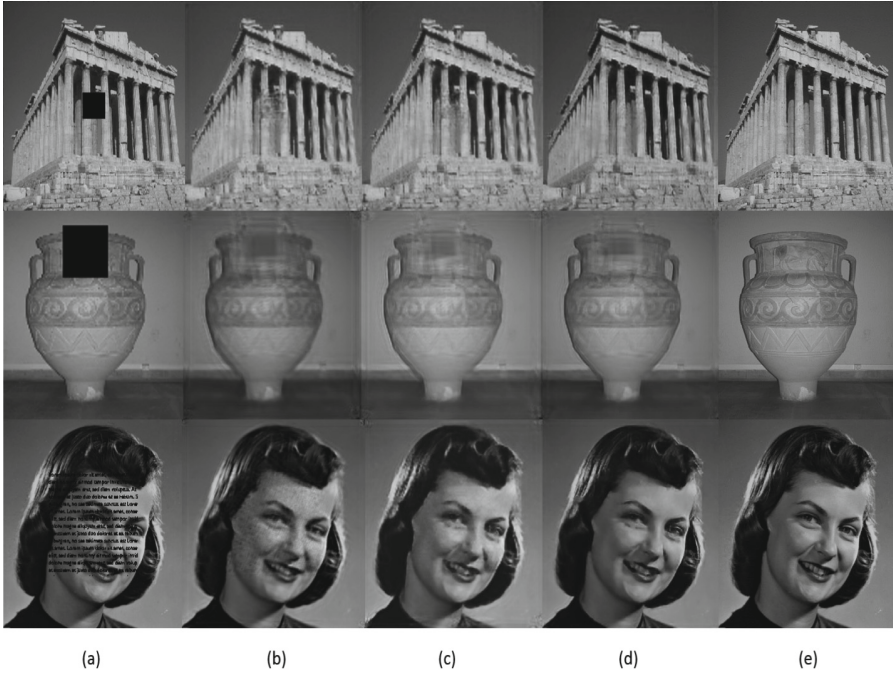
To generate noisy input, we add a different type of noises to the down-sampled ground-truth image patches using inbuilt functions in Matlab. Experiments have been conducted for noise resilient super-resolution. Proposed framework has been compared with conventional and state of art noise-free SR techniques.

### 4.3 Results

To verify effectiveness of proposed CAE-CDCA for blind inpainting, we have conducted several experiments as shown in Table 1. Proposed CAE-CDCA performs better in terms of PSNR and SSIM with comparison to state-of-the-art blind inpainting techniques. Mainly, we focus on larger missing region because state-of-the-art techniques like SSDA [5], BiCNN [4] and RED30 [3] performs

**Table 1.** Blind image inpainting results comparison on different images and dataset

Datasets	INPUT	INPUT	SSDA [5]	SSDA	BiCNN [4]	BiCNN	RED [3]	RED	Proposed	Proposed
Images	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Image1	16.63	0.7823	20.16	0.7904	20.72	0.7984	25.2316	0.8075	28.36	0.8143
Image2	13.32	0.8104	24.04	0.8375	24.87	0.8654	25.03	0.9013	26.89	0.9704
Image3	18.41	0.8252	23.67	0.8276	24.89	0.8279	28.08	0.8311	29.71	0.8358
Set5+Set14	17.19	0.7129	20.28	0.7591	25.06	0.7878	26.23	0.8110	29.87	0.8765
BSD100	19.73	0.8359	22.60	0.8471	25.01	0.8901	27.39	0.9126	30.86	0.9416



**Fig. 2.** Visual comparison of different blind inpainting algorithms for Image.1 (upper one), Image.2 (middle one) and Image.3 (lower one): (a) input, (b) BiCNN [4], (c) RED30 [3], (d) proposed, (e) ground truth.

exceptionally well for small region (missing line or region 1–10 pixel width) but for larger region, these algorithms fail to fill region with proper local texture/information. We get 3.12, 1.86, 1.64, 3.64 and 3.51 dB improvement for Image.1, Image.2, Image.3, Set5+Set14 [3] and BSD100 dataset [3] respectively from the RED30 [3]. To make fair comparison, we have also shown results of other state-of-the-art blind inpainting techniques (No comparison with non-blind inpainting techniques).

In Fig. 2, we have shown the visual comparison of different algorithms for blind inpainting. For Image.1, missing region is of size  $(32 \times 32)$  at the mid of image. Here, the texture information in resultant image of proposed method

is clearer than RED30 [3] and BiCNN [4] resultant image. In case of Image.2, missing region is of size  $(64 \times 64)$  at random location. We can visualize that, proposed method is able to restore sharp edges of pot compared to other methods. We also show result on Image.3 with text as a mask. Here our result is visually comparable with RED30 [3] but gives much better results than BiCNN [4].

We are showing results with different mask at different locations to prove generalization and robustness of proposed framework. To make comparisons more generalized for all natural images, we compare all algorithms for BSD100 dataset images with random missing region (square of  $32 \times 32$ ) at different locations. Our architecture performs inpainting at coarser version with 5 convolutional and 5 de-convolutional layers. Due to performing inpainting at down-sampled version, our framework have less computational complexity and reduced area to be inpainted (missing region). We use CDCA [8] to perform SR because it provides state-of-the-art results with less computational complexity (only 3 layer convolutional network). Thus, our combined CAE-CDCA provides better result with less complexity in comparison of RED30 [3] for blind inpainting.

We have also tested proposed framework (CAE+CDCA) for 3X and 4X SR on BSD100 dataset to fill more bigger region (missing region of size greater than  $100 \times 100$ ). Our framework get 3.61 dB and 3.87 dB improvement in PSNR from RED30 [3] for inpainting missing region of size  $(120 \times 120)$  using 3X and 4X SR, respectively. Our framework also works better than [8] for getting noise resilient SR. We get 0.87 dB, 1.19 dB and 0.98 dB improvement on BSD200 dataset for getting 2X, 3X and 4X noise (Gaussian noise with different variance) resilient super resolution, respectively.

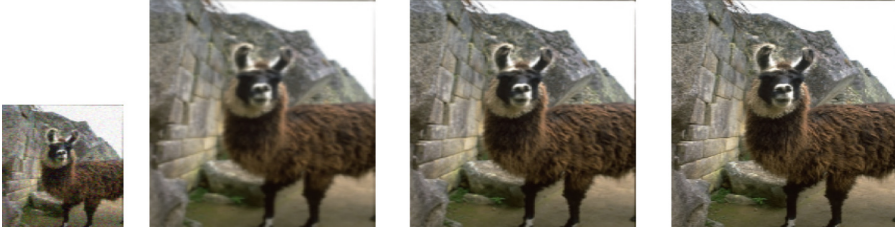
Comparison of proposed framework with conventional and state of art 3x noise resilient image SR has been shown in Table 2. The table shows that results of proposed frameworks are better than state-of-art and traditional frameworks. Here, conventional method refers to best algorithm for image de-noising followed by best algorithm for image super-resolution (i.e, CAE+CDCA). Here, the + sign shows just cascading of methods. In Table 2, CAE+CDCA is just cascading of learned CAE for de-noising and learned CDCA for super-resolution (without fine-tuning). CAE-CDCA is proposed framework with fine-tuned weights of combined network. We have also compared the result of integrated architecture with cascaded network and the result shows that proposed framework performs better than a cascaded framework. In Table 2, results are shown on Set5, Set14 and BSD200 dataset with the gaussian noise of different variance (10, 20 and 30) for 3X super-resolution. To test the robustness of our architecture, we applied other noises as well. The improvement in PSNR, when compared to conventional method, was 2.2 dB for blurring and 2.9 dB for salt and pepper noise in case of 2x noise resilient SR. The test images for experiments were obtained from, set5, set14 and BSD200 dataset [28].

In Fig. 3, we can easily visualize that high-frequency information is far better than state of art and conventional approach. This figure clearly represents that conventional techniques are not able to recover high frequency and texture information.



**Table 2.** Comparison of different methods with gaussian noise for 3x noise resilient super-resolution

Algo./Datasets	Set5	Set14	BSD200	Set5	Set14	BSD200	Set5	Set14	BSD200
	( $\sigma = 10$ )	( $\sigma = 10$ )	( $\sigma = 10$ )	( $\sigma = 20$ )	( $\sigma = 20$ )	( $\sigma = 20$ )	( $\sigma = 30$ )	( $\sigma = 30$ )	( $\sigma = 30$ )
	PSNR	PSNR	PSNR	PSNR	PSNR	PSNR	PSNR	PSNR	PSNR
SSDA+CDCA	29.11	26.47	24.65	28.59	26.28	25.06	28.17	25.96	23.08
CAE+CDCA	29.17	26.49	24.67	28.67	26.34	25.12	28.19	26.22	23.15
SSDA-CDCA	32.24	28.86	26.83	32.03	28.71	26.41	31.63	28.36	25.96
CAE-CDCA	32.53	28.99	27.06	32.49	28.99	26.94	31.72	28.48	26.33

**Fig. 3.** Noise resilient image SR (2x) comparison on lamma. (a) LR image with Gaussian noise ( $\sigma = 30$ ). (b) Conventional. (c) Proposed CAE-CDCA. (d) Original.

## 5 Conclusion

We proposed a novel end-to-end deep learning based framework for blind inpainting which is more efficient in filling larger missing regions. Our framework performs inpainting at coarser level and super-resolution simultaneously preserving the local structure/information. We did exhaustive experiments with several state-of-the-art image inpainting techniques. Our framework can be easily adopted for getting noise-free super-resolution after fine-tuning on noise-free super resolution dataset. Experimental results show the effectiveness of our framework. End-to-end optimization is the main reason behind our better results. In future, we plan to experiment with color images for blind inpainting.

## References

1. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH (2000)
2. Guillemot, C., Le Meur, O.: Image inpainting: overview and recent advances. *IEEE Signal Process. Mag.* **31**(1), 127–144 (2014)
3. Mao, X.-J., Shen, C., Yang, Y.-B.: Image Denoising Using Very Deep Fully Convolutional Encoder-Decoder Networks with Symmetric Skip Connections. *CoRR* (2016)
4. Cai, N., Su, Z., Lin, Z., Wang, H., Yang, Z., Ling, B.W.K.: Blind inpainting using the fully convolutional neural network. *Vis. Comput.* **33**(2), 249–261 (2017)

5. Xie, J., Xu, L., Chen, E.: Image denoising and inpainting with deep neural networks. In: *Advances in Neural Information Processing Systems*, pp. 341–349 (2012)
6. Fawzi, A., Samulowitz, H., Turaga, D., Frossard, P.: Image inpainting through neural networks hallucinations. In: *IEEE 12th Image, Video and Multidimensional Signal Processing Workshop, IVMSP (2016)*
7. Le Meur, O., Ebdelli, M., Guillemot, C.: Hierarchical super-resolution-based inpainting. *IEEE Trans. Image Process.* **22**(10), 3779–3790 (2013)
8. Sharma, M., Chaudhury, S., Lall, B.: Deep learning based frameworks for image super-resolution and noise-resilient super-resolution. In: *International Joint Conference on Neural Networks, IJCNN (2017)*
9. Mao, X.-J., Shen, C., Yang, Y.-B.: Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections. *CoRR (2016)*
10. Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) *ECCV 2014. LNCS*, vol. 8692, pp. 184–199. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13)
11. Kim, J., Kwon Lee, J., Mu Lee, K.: Accurate Image Super-Resolution Using Very Deep Convolutional Networks. *CoRR (2015)*
12. Wang, Z., Liu, D., Yang, J., Han, W., Huang, T.: Deep networks for image super-resolution with sparse prior. In: *IEEE International Conference on Computer Vision, ICCV (2015)*
13. Kim, J., Kwon Lee, J., Mu Lee, K.: Deeply-Recursive Convolutional Network for Image Super-Resolution. *CoRR (2015)*
14. Yang, W., Feng, J., Yang, J., Zhao, F., Liu, J., Guo, Z., Yan, S.: Deep Edge Guided Recurrent Residual Learning for Image Super-Resolution. *CoRR (2016)*
15. Ledig, C., Theis, L., Huszar, F., Caballero, J., Aitken, A.P., Tejani, A., Totz, J., Wang, Z., Shi, W.: Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial. *CoRR (2016)*
16. Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z.: Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. *CoRR (2016)*
17. Criminisi, A., Perez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Process.* **13**(9), 1200–1212 (2004)
18. Bertalmio, M., Bertozzi, A.L., Sapiro, G.: Navier-stokes, fluid dynamics, and image and video inpainting. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR (2001)*
19. Ghorai, M., Mandal, S., Chanda, B.: Patch sparsity based image inpainting using local patch statistics and steering kernel descriptor. In: *23rd International Conference on Pattern Recognition, ICPR (2016)*
20. Fadili, M.-J., Starck, J.-L., Murtagh, F.: Inpainting and zooming using sparse representations. *Comput. J.* **52**(1), 64–79 (2009)
21. Le Meur, O., Gautier, J., Guillemot, C.: Exemplar-based inpainting based on local geometry. In: *18th IEEE International Conference on Image Processing, ICIP (2011)*
22. Dong, B., Ji, H., Li, J., Shen, Z., Xu, Y.: Wavelet frame based blind image inpainting. *Appl. Comput. Harmon. Anal.* **32**(1), 268–279 (2012)
23. Liu, J., Musialski, P., Wonka, P., Ye, J.: Tensor completion for estimating missing values in visual data. *IEEE Trans. Pattern Anal. Mach. Intell. TPAMI* **35**(1), 208–220 (2013)

24. Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: feature learning by inpainting. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR (2016)
25. Yeh, R., Chen, C., Lim, T.-Y., Hasegawa-Johnson, M., Do, M.N.: Semantic Image Inpainting with Perceptual and Contextual Losses. CoRR (2016)
26. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis. IJCV* **115**(3), 211–252 (2015)
27. Image Inpainting Dataset of Computational Intelligence Lab - ETH Zürich, ETH-CIL Dataset (2012)
28. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: 8th International Conference on Computer Vision, ICCV (2001)