

RESEARCH

Open Access



Anchored neighborhood deep network for single-image super-resolution

Wuzhen Shi¹, Shaohui Liu^{1*} , Feng Jiang¹, Debin Zhao¹ and Zhihong Tian²

Abstract

Real-time image and video processing is a challenging problem in smart surveillance applications. It is necessary to trade off between high frame rate and high resolution to meet the limited bandwidth requirement in many specific applications. Thus, image super-resolution become one commonly used techniques in surveillance platform. The existing image super-resolution methods have demonstrated that making full use of image prior can improve the algorithm performance. However, the previous deep-learning-based image super-resolution methods rarely take image prior into account. Therefore, how to make full use of image prior is one of the unsolved problems for deep-network-based single image super-resolution methods. In this paper, we establish the relationship between the traditional sparse-representation-based single-image super-resolution methods and the deep-learning-based ones and use transfer learning to make our proposed deep network take the image prior into account. Another unresolved problem of the deep-learning-based single-image super-resolution method is how to avoid neurons compromise to different image contents. In this paper, the image patches are anchored to the dictionary atoms to group into various categories. As a result, each neuron will work on the same types of image patches that have similar details, which makes the network more accurate to recover high-frequency details. By solving these two problems, we propose an anchored neighborhood deep network for single-image super-resolution. Experimental results show that our proposed method outperforms many state-of-the-art single-image super-resolution methods.

Keywords: Image super-resolution, Anchored neighborhood regression, Deep learning, Transfer learning, Convolutional neural network

1 Introduction

In recent years, big data [4, 7], cloud computing, and AI (artificial intelligence) are the most popular research topics. Deep learning in AI has moved from research labs to applications, especially computer vision, natural language processing, speech recognition, and many other fields [8, 21]. These applications require systems or platforms to interact with the real world by sensors [5, 6], such as cameras in computer vision. However, the bandwidth of these devices are very limited. For example, the bandwidth is about 480 Mbps for the USB 2.0 interface. If the resolution is 1920×1080 and the frame rate is 100 hz, then the bandwidth is about 5 Gbps. Moreover, the frame rate must be larger than 100 Hz in some high-speed applications. Hence, frame rate up-conversion and super-resolution are necessary in many real-time applications. Figure 1 gives

an example of super-resolution in surveillance application. In that case, the server side has high performance to process images acquired from the sensors with limited bandwidth interface. In addition, such high payload raises the scheduling problem in different communication environments [26–28].

Image super-resolution (SR) technology takes the low-resolution (LR) images as input and maps them to the corresponding high-resolution (HR) space. It has been studied for a long time but has become more prevalent with the new generation of ultra-high-definition (UHD) TVs (3840×2048). Most video content is not available in UHD resolution. Therefore, SR algorithms are needed to generate UHD content from full high definition (FHD) (1920×1080) or lower resolutions [16]. Depending on the number of the input LR images, image SR generally can be divided into single-image SR method and multiple-image SR method. In this paper, we focus on single-image SR, which aims at recovering a high-resolution image

*Correspondence: shliu@hit.edu.cn

¹Harbin Institute of Technology, Harbin 150001, China

Full list of author information is available at the end of the article

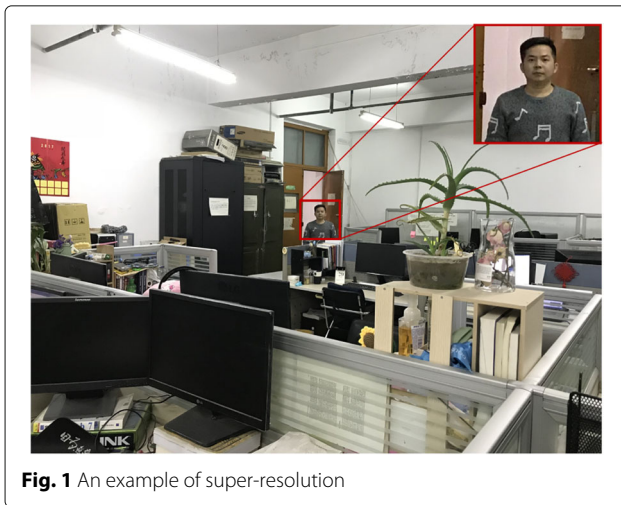


Fig. 1 An example of super-resolution

from a single low-resolution image. For convenience, we roughly subdivide the single-image SR methods into two subclasses: the non-deep-learning-based methods and the deep-learning-based ones. Most non-deep-learning-based single-image SR methods either try to find the new kinds of image prior or propose a new way to use these existing image prior, while the deep-learning-based methods always learn a simple end-to-end mapping between the LR image and the HR one.

Traditional non-deep-learning-based SR methods have demonstrated that image prior, e.g., local smoothing, nonlocal self-similarity, and sparsity, plays an important role in image SR. Neighbor embedding (NE) approaches assume that small image patches from a low-resolution image and its high-resolution counterpart form low-dimensional nonlinear manifolds with similar local geometry. Chang et al. [3] proposed a SR method based on this principle using the manifold learning method of locally linear embedding (LLE) [24]. In addition to the local linear prior, image sparsity is the most commonly used in the literature of single-image SR. Yang et al. [35] proposed the first sparse-representation-based single-image SR method that assumes the low-frequency image patches have the same sparse representation with the corresponding high-frequency image patches. On this basis, Zeyde et al. [37] proposed a more efficient dictionary learning method for both low- and high-resolution patches, which leads to significant training time savings. Other kinds of image prior, e.g., local smoothing and nonlocal self-similarity, are also well studied as the regularization term in the reconstruction-constraint-based single-image SR methods. Apart from investigating the new image prior, some traditional methods try to find out a more compact representation of the well-known image prior or a more efficient way to use these image prior for improving the image SR performance. In [30], Timofte et al. propose

an anchored neighborhood regression (ANR) for single-image SR. That is to anchor the neighborhood embedding of a low-resolution patch to the nearest atom in the dictionary and to precompute the corresponding embedding matrix. In later, they further propose an improved variant of ANR, which combines the best qualities of anchored neighborhood regression and simple functions (SF) [31]. In order to make better use of the image sparse prior, Zhang et al. [38] propose a dual dictionary to learn residual iteratively.

Recently, deep learning method has got much attention, and it is successfully applied in many low- and high-level computer vision problems. Some deep-learning-based image SR methods have also been explored. The pioneering work of deep-learning-based SR is SRCNN proposed by Dong et al. [10, 11]. They demonstrated that a convolutional neural network (CNN) can learn a mapping from low-resolution image to high-resolution one in an end-to-end manner. It does not require any engineered features that are typically necessary in traditional non-deep-learning-based methods. Soon after, they expanded this work for JPEG compressive image restoration [9]. Recently, they further proposed an improved version of SRCNN that takes the 1×1 convolution into account to reduce the network weights and result in a fast SRCNN (FSRCNN) [12]. Different from [9–12] that use the undegraded image as ground true for training, some works try to learn image residual. Kim et al. [17] proposed a very deep network to learn residual to accelerate the convergence speed. The abovementioned methods (except FSRCNN) need to upscale the low-resolution input image to the high-resolution space using a single filter, commonly bicubic interpolation, before reconstruction. To avoid adding computational complexity, Shi et al. [29] proposed to operate on the low-resolution image space and introduce an efficient sub-pixel convolution layer which learns an array of upscaling filters to upscale the final LR feature maps into the HR output. Many other new deep-learning-based single-image SR methods [15, 20, 34] and video SR methods [2, 14] have also been proposed.

Both the traditional non-deep-learning-based methods and the deep-learning-based ones have their advantages. Traditional methods always take the image prior into account and have the process of strict mathematical derivation. In comparison, deep-learning-based methods always are the end-to-end mapping, which avoids the complex optimization solving process and results in much fast running speed. How to make full use of the advantages of both of them is a very interesting problem. In this paper, one of the problems we focus on is how to take the image prior into account for the deep-learning-based single-image SR method. On the other hand, in all the previous deep-learning-based methods, each neuron works

on the whole input feature map. It has to compromise to different image contents, although it has very small receptive field. For example, to the smooth region of the image, we expect the neuron to be low-pass filter, while a high-pass filter is better to the complex texture region. However, the natural images are usually content-rich containing not only the smooth region but also the complex texture region. The neuron works on the whole image has to compromise to these totally different image contents that will affect their activity to the final output. In this paper, how to avoid the neuron compromises to different image contents is another problem we focus on.

In this paper, we propose an anchored neighborhood deep network for single-image SR. The well-known anchored neighborhood regression method show that the image high-frequency details can be computed by the input low-frequency patches multiplied by a precomputed matrix, which is well trained by a large amount of low and high patch pairs with sparse prior constraint. We design a convolution layer to mimic the matrix multiplication process and transfer the weights of each row of the well-trained matrix to one convolution filter. Different from the previous transfer learning, we transfer the weights from a matrix, which is outside any network and trained with strict image prior constraint, to the network instead of one network to the other one. Since the weights of the matrix are trained with strict image prior constraint, the convolution layer whose parameters are transferred from the well-trained matrix has took the image prior into account. Inspired by the anchored neighborhood regression single-image SR method, we first anchor the feature vectors to the nearest dictionary atom. Then, to different kinds of feature vectors, we use different convolution layers to predict their corresponding high-frequency details. Figure 2 gives an intuitional description to this process. That results in each neuron works on the same

kinds of image patches to avoid compromise to different image contents. Through solving these two problems, we successfully design an anchored neighborhood deep network for single-image SR. Experimental results show that our proposed method have comparable performance with many state-of-the-art single-image SR methods.

In the next section, we will give some background on sparse-representation-based and deep-learning-based SR methods and review the anchored neighborhood regression methods. Section 3 introduces the motivation and summarize our contribution. In Section 4, we propose our anchored neighborhood deep network in detail. Section 5 describes our experiments, where we compare the performance of our approach to other state-of-the-art methods. Section 6 concludes our work. Finally, we discuss some valuable future works in Section 7.

2 Related work

Since our proposed method is inspired by the anchored neighborhood regression and makes full use of the advantages of both sparse representation approaches and the deep-learning-based ones, we shortly review them.

2.1 Sparse representation approaches

Sparse representation try to use nonzero coefficients as few as possible to represent signal's main information. For a patch x_i , the process of finding its sparse representation vector α_i with respect to a known over-complete dictionary D is called sparse coding. As can be seen, owing to the over-completeness, the null space of D introduces additional degrees of freedom in the choice of α_i , which can be exploited to improve its compressibility. To obtain the sparse representation, sparse coding can be formulated as

$$\alpha_i = \arg \min_{\alpha_i} \|x_i - D\alpha_i\|_2^2 \quad s.t. \|\alpha_i\|_0 < L \quad (1)$$

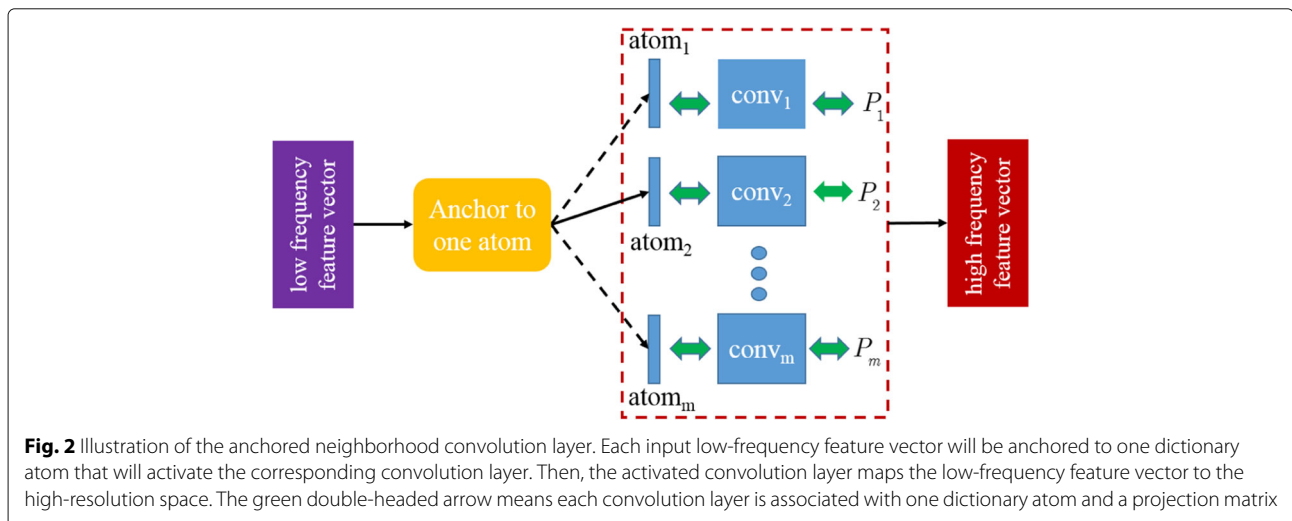


Fig. 2 Illustration of the anchored neighborhood convolution layer. Each input low-frequency feature vector will be anchored to one dictionary atom that will activate the corresponding convolution layer. Then, the activated convolution layer maps the low-frequency feature vector to the high-resolution space. The green double-headed arrow means each convolution layer is associated with one dictionary atom and a projection matrix

Though this problem is NP-hard in general, it can be approximated by a wide range of techniques [22]. In this paper, we adopt an orthogonal matching pursuit (OMP) [32] algorithm to solve this problem for its simplicity and efficiency.

The other one main problem of sparse representation is dictionary learning. Its general formulation is:

$$D, \{\alpha_i\} = \arg \min_{D, \{\alpha_i\}} \sum_k \|x_i - D\alpha_i\|_2^2 \quad (2)$$

$$s.t. \|\alpha_i\|_0 < L \quad \forall i$$

where $\{\alpha_i\}$ are the sparse representation vectors for $\{x_i\}$. There are many dictionary learning methods that have been proposed in recent year. One of the widely used dictionary learning methods is K-SVD [18], which has shown more effectivity and higher efficiency than many other state-of-the-art dictionary learning methods.

The sparse-representation-based SR method assumes the same sparse representation for low-resolution patches as their corresponding high-resolution patches. Therefore, the sparse dictionaries have to be jointly learned for low- and high-resolution image patches. Given a set of training image patch pairs X_h and X_l , the joint dictionary learning can be formulated as:

$$\arg \min_{D_l, D_h, \alpha} \frac{1}{N} \|X_h - D_h \alpha\|_2^2 + \frac{1}{M} \|X_l - D_l \alpha\|_2^2 + \lambda \left(\frac{1}{N} + \frac{1}{M} \right) \|\alpha\|_0 \quad (3)$$

where X_h and X_l , N and M are the high- and low-resolution patches and their dimensionality, respectively, and α is the coefficient vector representing the sparsity constraint.

To speed up the running time, Timofte et al. [30] proposed anchored neighborhood regression for fast single-image SR. They relaxed the L0 norm constraint to L2 norm and used part of the dictionary atoms to represent each patch. Then, the objective function will become

$$\alpha_i = \arg \min_{\alpha_i} \|x_i - D_l \alpha_i\|_2^2 + \lambda \|\alpha_i\|_2 \quad (4)$$

With the L2 norm, this turns the problem into ridge regression and gives it a closed solution. An input low-frequency patch y_i can be projected to a high-resolution space as

$$x_i = D_h \left(D_l^T D_l + \lambda I \right)^{-1} D_l^T y_i = P_i y_i \quad (5)$$

where P_i is the stored projection matrix for dictionary atom D_l^i . In summary, ANR computes offline the projection matrix P_i for each dictionary atom in the training process and anchors each patch to its most similar dictionary atom and maps it to output the high-frequency detail patch with the corresponding projection matrix P_i . In [31], Timofte et al. propose A+, an improved variant of ANR,

which combines the best qualities of ANR and SF. We refer the reader to [30, 31] for more details about ANR and A+.

2.2 Deep learning approaches

The previous deep-learning-based image SR approaches always learn an end-to-end mapping, which takes the low-resolution image as input and directly outputs the high-resolution one. The pioneer work is SRCNN [10], which is a simple three-layer network. Specifically, the first layer performs patch extraction and representation, which extracts overlapping patches from the input image and represents each patch as a high-dimensional vector. Then, the non-linear mapping layer maps each high-dimensional vector of the first layer to another high-dimensional vector, which is conceptually the representation of a high-resolution patch. At last, the reconstruction layer aggregates the patch-wise representations to generate the final output. Inspired by other successful high-level works, Kim et al. [17] proposed to increase the network depth to have a larger receptive field to predict the image details and use the residual learning method to accelerate convergence. In [34], Wang et al. designed a network to mimic the traditional sparse-representation-based SR method. However, it needs multiple layers to get the accurate sparse representation, and all the image patches use the same network structure. Many other deep-learning-based image SR methods have also been proposed.

Transfer learning in deep neural networks becomes popular since the success of deep learning in image classification [19]. The features learned from the ImageNet show good generalization ability [36] and become a powerful tool for several high-level vision problems. Many works have demonstrated that transfer of the network parameters learned from the ImageNet to their own network can improve performance. Inspired by the success of transfer learning applied in high-level vision problems, Dong et al. [9] explored several transfer settings on compression artifact reduction and demonstrated the effectiveness of transfer learning in low-level vision problems. Different from these transfer learning approaches mentioned above, we propose to transfer the parameters from a precomputed projection matrix that makes our network can make full use of the image prior and improve SR performance.

3 Motivations and contributions

Traditional non-deep-learning-based single-image SR methods try to either find new kinds of image prior or propose a new way to use these existing image prior. All these previous works demonstrated that makes full use of image priors can improve image SR performance. How to use the image prior in deep-learning-based methods is still rarely studied. So it inspires us to explore how to take image prior into account for deep-learning-based method.

Fortunately, previous works proposed by Timofte et al. [30, 31] show that the objective function with sparse prior constraint has a closed solution. Furthermore, the matrix multiplication can be easily implemented by a convolution layer. Therefore, transferring these weights of the projection matrix trained offline to a convolution layer is a very natural selection.

The neurons of these previous deep-learning-based methods work on the whole input feature map. They have to compromise to different image contents. ANR and A+ proposed by Timofte et al. [30, 31] inspire us to anchor different image patches to different dictionary atoms, and then, all the patches are naturally divided into multiple categories and each neuron will work on the similar image patches.

In this paper, contrary to previous works, we propose to transfer the weights of the matrix, which are trained offline using a large amount of patches with image prior constraint, to the weights of a convolution layer. That results in that our network has the inherent property of taking the image prior into account. Similar to ANR and A+, we anchor each feature vector to one of the dictionary atoms and then use the corresponding convolution layer to map the low-frequency input vector to predict its high-frequency detail. As a result, each neuron of our network works on the same kinds of image patches to avoid compromise to different image contents.

In short, the contributions of this work are mainly in three aspects:

- We establish a relationship between our deep-learning-based single-image SR method and the well-known sparse representation one. The transfer learning technology has been used to join the traditional approach with good ability of using image prior knowledge and the deep-learning-based approach with strong end-to-end optimization ability.
- We propose anchored neighborhood deep network for single-image SR. Compared to the previous deep-learning-based methods, the neurons in our proposed SR network pay more attention to acquire the local image information to avoid compromise to different image contents. Compared to the traditional anchored neighborhood regression methods, the traditional methods are local optimization, while our proposed network is an end-to-end global optimization.
- We give large amount of experiments to demonstrate the robustness of our proposed new single-image SR method.

4 Proposed method

Compared to previous deep-learning-based single-image SR methods, our proposed method is also an end-to-end

mapping that takes the low-resolution image as input and directly outputs the high-resolution one. The difference are mainly two aspects: we use a sparse prior constraint convolution layer to take the image sparse prior into account and use an anchored neighborhood convolution layer to avoid neurons compromise into different image contents. Therefore, we firstly introduce the sparse prior constraint convolution layer and the anchored neighborhood convolution layer that are associated with the two problems we focus on. Finally, we introduce our new network structure for single-image SR.

4.1 Sparse prior constraint layer

As shown in Eq. (4) and (5), the L2 norm sparse constraint objective function has a close solution $x_i = P_i y_i$, where projection matrix is precomputed offline by a set of low- and high-image patch pairs. If each row of the projection matrix P_i is considered as a filter, we can use a convolution layer to mimic this mapping process to predict the image detail. Here, we assume that y_i is a vector of size $n \times 1$, x_i is a vector of size $m \times 1$, and P_i is a matrix of size $m \times n$. Then, each convolution is of size $1 \times 1 \times n$, i.e., the spatial size of each convolution is 1×1 and it has n feature maps. Since the projection matrix P_i has m rows, there are m convolutions of size $1 \times 1 \times n$. It should be noted that there is no bias in each filter so that all the filters can fully mimic the matrix multiplication process.

As shown in Eq. (5), $x_i = D_h(D_l^T D_l + \lambda I)^{-1} D_l^T y_i$, where D_l and D_h are two well-trained low and high dictionaries. Since x_i is the close solution with image sparse prior constraint, we transfer the matrix weights to one convolution layer that will make our network have an inherent attribute to take image sparse prior into account and the output x_i will be a more accurate high-frequency prediction.

4.2 Anchored neighborhood layer

The ANR and A+ firstly find the neighborhoods and then calculate a separated projection matrix P_i for each dictionary atom D_i in the offline training process. As a result, given an input patch feature y_i , it just needs to anchor it to its nearest neighbor atom D_i and map it to HR space using the stored projection matrix P_i . In this paper, we use a network to mimic this process, which has an inherent attribute to make our method get better performance.

The anchored neighborhood convolution layer is outlined in Fig. 2. To each dictionary atom D_i , we calculate its projection matrix P_i using the same method as A+, which has took the image sparse prior into account. After training all projection matrices, we transfer them to different convolution layers using the method mentioned above. That is, each sub-convolution layer with respect to an atom in the anchored neighborhood layer

is a sparse prior constraint convolution layer. It should be noted that all these sub-convolution layers can be parallel implemented. To each input low-frequency feature vector, the anchored neighborhood layer will anchor it to one dictionary atom that will activate the corresponding sub-convolution layer. Then, the activated convolution layer maps the low-frequency feature vector to the high-resolution space, which executes the traditional matrix multiplication process.

Since we transfer the weights of the projection matrix P_i to the sub-convolution layer, the anchored neighborhood convolution layer has fully took the sparse image prior into account. Both ANR and A+ demonstrate the projection matrix P_i can be used to accurately predict the high-frequency details. Therefore, it is sure that our anchored neighborhood convolution layer can predict the accurate image in high frequency for the later layer to further refine. More importantly, through the anchoring process, the image patches will be divided into multiple categories, and each neuron will work on the similar feature vectors instead of the whole image that makes it avoid compromise to different image contents.

4.3 Proposed network structure

The proposed network structure is outlined in Fig. 3. It can be simply divided into four parts, i.e., feature extraction layer, anchored neighborhood convolution layer, combination layer, and deep integration subnetwork. We have used different colors to mark the corresponding part in Fig. 3.

Feature extraction. The ANR and A+ show that the features used to represent the image patches have strong influence on the performance. The most basic feature to use is the patch itself. This however does not give the feature good generalization properties. An often used similar feature is the first- and second-order derivative of the patch [3, 35]. In this paper, we use a convolution layer with

$n1$ filters of size $3s \times 3s \times 1$, where s is the magnification factor, to extract the image feature. As a result, the output feature is a $n1 \times 1$ vector. At the same time, we use the “one-hot” convolution, which means one filter extracts only one pixel in the receptive field, to extract LR patches for the later image reconstruction. The filter size of the one-hot convolution is also $3s \times 3s \times 1$.

Anchored neighborhood convolution. This layer has been introduced in detail in the Section 4.2. It is used to take image prior into account to fastly and accurately predict the image details and to make the neurons work on the local image patches to avoid compromise to different image contents. Note that the dictionary used in our experiment has 1024 atoms. Therefore, there are 1024 parallel sparse prior constraint layers in this anchored neighborhood layer.

Combination The anchored neighborhood convolution layer outputs the initial high-frequency details for each low-resolution patch. We firstly add these estimated high-frequency details to the corresponding LR patch, which is extracted by the one-hot convolution, to get the initial high-resolution feature vector. We reshape these feature vectors to get the image patches and concatenate them to output the initial high-resolution estimation. In other words, the combination layer contains a reshape and a concatenation process.

Deep integration. It has been demonstrated in the literature that the deeper the network, the better the performance. To further fuse the image local similarity details, we design a deep integration subnetwork that cascades m convolution layers, where the layers except the first and the last are of the same type: d filters of the size $f \times f \times d$, where a filter operates on $f \times f$ spatial region across d channels (feature maps). The first layer operates on the output of the combination layer, so that it has d filters of the size $f \times f \times 1$. The last layer, which outputs the final

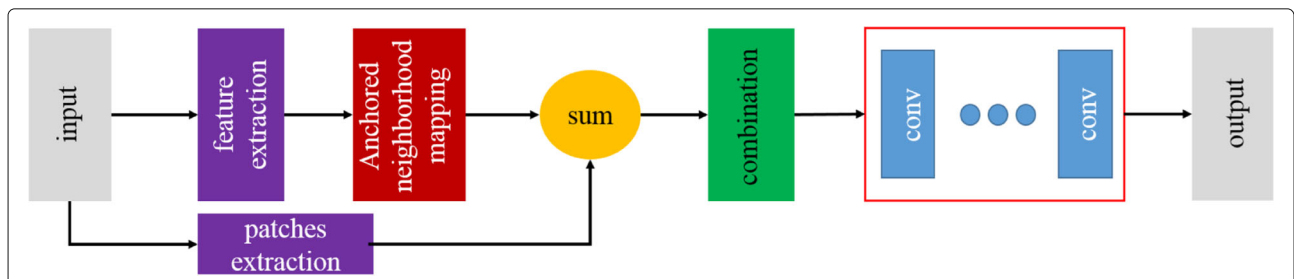


Fig. 3 Illustration of the proposed anchored neighborhood deep network structure for single image SR. It uses a convolution layer to extract low frequency feature and then uses an anchored neighborhood convolution layer to predict the high frequency details. After combining these high frequency patches, we cascade multiple convolution layers to fuse the local similarity patches and refine the image details

image estimation, consists of a single filter of size $f \times f \times d$. It can be formulated as

$$F_i(y) = \max(0, w_i * y + b_i), i \in \{1, m-1\} \quad (6)$$

$$F_m(y) = w_m * F_{m-1}(y) + b_m \quad (7)$$

where $\max(\cdot)$ represents the rectified linear unit (ReLU) operator and w_i and b_i represent the filters and biases of the i th layer respectively.

4.4 Training

We now describe the objective to minimize to find the optimal parameters of our model. Following most of deep-learning-based image restoration methods, the mean square error is adopted as the cost function of our network. Our goal is to train an end-to-end mapping f that predicts values $\hat{y} = f(x)$, where x is an input low-resolution image and \hat{y} is the estimation of the corresponding high-resolution image. Given a set of high-resolution image examples $y_i, i = 1 \dots N$, we generate the corresponding low-resolution images $x_i, i = 1 \dots N$ (in fact, we upscale them to the original size by bicubic interpolation). Then, the optimization objective is represented as

$$\min_{\theta} \frac{1}{2N} \sum_{i=1}^N \|f(x_i; \theta) - y_i\|_F^2 \quad (8)$$

where θ is the network parameter needed to be trained, $f(x_i; \theta)$ is the estimated high-resolution image with respect to low-resolution image x_i . We use the adaptive moment estimation (Adam) [18] to optimize all network parameters.

5 Experimental results and discussion

In this section, we evaluate the performance of our method on several datasets. We first describe datasets used for training and testing our method. Next, some training details are given. Finally, we show the quantitative and qualitative comparisons with five state-of-the-art methods. We name our anchored neighborhood deep network as ANNet.

5.1 Implementation details

Datasets for training and testing. It is well known that training dataset is very important for the performance of learning-based image restoration methods. A lot of training dataset can be found in the literature. For example, SRCNN [10, 11] uses a 91-image dataset and VDSR [17] uses a 291-image dataset. In this paper, we mainly follow FSRCNN [12] to use the General-100 dataset, which contains 100 bmp format images (with no compression). To further test the impact of different training datasets to the performance, we also establish our own training

Table 1 The average PSNR (dB) and SSIM comparisons with our proposed method with different filter sizes on Set5 [1]

Filter size	3	5	7	9
3	32.68/0.9106	32.70/0.9108	32.72/0.9104	32.72/0.9104
5	32.80/0.9117	32.74/0.9106	32.79/0.9110	32.78/0.9113
7	32.75/0.9105	32.78/0.9111	32.78/0.9111	32.81/0.9113
9	32.78/0.9108	32.78/0.9109	32.80/0.9114	32.82/0.9114

All models are trained on the General-100 dataset

dataset, which contains 260 bmp format images. We set the patch size as 45×45 and use data augmentation (rotation or flip) to prepare training data. Following FSRCNN and SRCNN, we use three datasets, i.e., Set5 [1] (5 images), Set14 [37] (14 images), and BSD200 [23] (200 images) for testing, which are widely used for benchmark in other works. Note that the test images are strictly separate from the training datasets.

Training strategy. For weight initialization, we use the method described in He et al. [13]. This is a theoretically sound procedure for networks utilizing rectified linear units (ReLU). For the other hyper-parameters of Adam, we set the exponential decay rates for the first and second moment estimate to 0.9 and 0.999, respectively. We train all our experiments only over 30 epochs and each epoch iterate 2000 times with a batch size of 64. The learning rate of the first 10 epochs is 0.0001, the 11 to 20 epochs is 0.00001, while that of the other 10 epochs is 0.000001. We implement our model using the MatConvNet package [33].

5.2 Investigation of different settings

To test the property of our anchored neighborhood deep network, we design a set of controlling experiments. We investigate the impact of the filter size, the network depth, and the training dataset. Since the parameters of the anchored neighborhood layer are fixed by the projection matrix trained offline, we mainly investigate different settings at the deep integration subnetwork.

Table 2 The average PSNR (dB) and SSIM comparisons with our proposed method with different depths on Set5 [1]

Image	2 layers	3 layers	4 layers
Baby	35.25/0.9240	35.26/0.9244	35.33/0.9250
Bird	35.52/0.9565	35.78/0.9584	35.97/0.9596
Butterfly	27.92/0.9166	28.39/0.9235	28.44/0.9249
Head	33.71/0.8254	33.77/0.8279	33.78/0.8286
Woman	31.51/0.9315	31.67/0.9338	31.70/0.9344
Avg.	32.78/0.9108	32.97/0.9136	33.05/0.9145

All models are trained on the General-100 dataset

Table 3 Comparison of our proposed ANNet trained with different datasets

Image	Baby	Bird	Butterfly	Head	Woman	Avg.
Small dataset	38.48/0.9656	41.07/0.9866	32.11/0.9648	35.74/0.8870	35.42/0.9702	36.56/0.9548
Big dataset	38.54/0.9655	41.29/0.9870	32.99/0.9685	35.75/0.8867	35.56/35.42	36.83/0.9557

Firstly, we investigate the impact of the filter size to the performance. In these experiments, the deep integration subnetwork just has two convolution layers. The average PSNR and SSIM values on the Set5 dataset of these experiments are shown in Table 1. The first column represents the filter size of the first convolution layer of the deep integration subnetwork, while the first row represents the filter size of the second convolution layer. Therefore, the values in the second row and the third column are the average PSNR and SSIM values of our network with the spatial size of the first and the second layers of the deep integration subnetwork being 3×3 and 5×5 , respectively. Since we use the square filter, we simplify them to one number. Table 1 shows that the larger the filter size, the better the performance. That can be attributed to it having a larger receptive field to get more useful information to predict the image details.

Next, we investigate the impact of network depth to the performance. In these experiments, the filter size of the deep integration subnetwork of the two-layer network, the three-layer network, and the four-layer network are 9-3, 9-5-3, and 9-5-5-3, respectively, where the number is the filter spatial size. The PSNR and SSIM values on the Set5 dataset of these experiments are shown in Table 2. Obviously, the four-layer network obtains the best performance not only merely on the average PSNR and SSIM but also on every single image. The three-layer network also outperforms the two-layer network. On this test dataset, the four-layer network can improve roughly on average

to 0.27 dB and 0.08 dB for the two- and three-layer networks, respectively. It reveals that the deeper the network, the better the performance, which agrees with the other researcher's finding. The good performance of the deeper network can also be attributed to the deep network having a larger receptive field to get more useful information to predict the image details.

Finally, we investigate the impact of training dataset to the performance. In our most experiments, we follow FSRCNN to use General-100 as the training dataset. To further test the impact of training dataset to the performance, we establish our own training dataset, which contains 260 bmp format images. Table 3 shows the PSNR and SSIM values on Set5 of our proposed ANNet trained with different datasets. The small dataset represents the General-100 dataset, which contains 100 images, while the big dataset represents our new established dataset that contains 260 images. On this test dataset, our network with the same setting trained with a larger dataset can improve roughly on average to 0.27 dB than that trained with a small dataset. That means a large training dataset is a good trick to improve the network performance.

5.3 Comparisons with state-of-the-art methods

We compare our ANNet with four state-of-the-art learning-based single-image SR methods, namely, the A+ [31], SRF [25], SRCNN [10, 11], and SCN [34]. A+ and SRF are the two state-of-the-art traditional non-deep-learning-based methods, while SRCNN and SCN are the

Table 4 The results of average PSNR (dB) and SSIM on the Set5 [1], Set14 [37], and BSD200 [23]

Dataset	Scale	Bicubic	A+	SRF	SRCNN	SCN	ANNet
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Set5	2	33.66/0.9299	36.55/0.9544	36.87/0.9556	36.34/0.9521	36.76/0.9545	36.73/0.9550
	3	30.39/0.9299	32.59/0.9088	32.71/0.9098	32.39/0.9033	33.04/0.9136	32.80/0.9117
	4	28.42/0.8104	30.28/0.8603	30.35/0.8600	30.09/0.8503	30.82/0.8728	30.38/0.8637
Set14	2	30.23/0.8687	32.28/0.9056	32.51/0.9074	32.18/0.9039	32.48/0.9067	32.43/0.9075
	3	27.54/0.7736	29.13/0.8188	29.23/0.8206	29.00/0.8145	29.37/0.8226	29.28/0.8217
	4	26.00/0.7019	27.32/0.7471	27.41/0.7497	27.20/0.7413	27.62/0.7571	27.40/0.7509
BSD200	2	29.70/0.8625	31.44/0.9031	31.65/0.9053	31.38/0.9287	31.63/0.9048	31.58/0.9052
	3	27.26/0.7638	28.36/0.8078	28.45/0.8095	28.28/0.8038	28.54/0.8119	28.46/0.8107
	4	25.97/0.6949	26.83/0.7359	26.89/0.7368	26.73/0.7291	27.02/0.7434	26.84/0.7365
Avg.		28.80/0.8151	30.53/0.8491	30.67/0.8505	30.39/0.8474	30.81/0.8542	30.77/0.8514

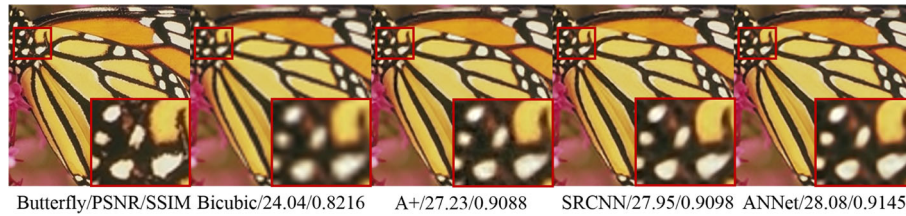


Fig. 4 Visual quality comparison of single-image SR on image butterfly from Set5 [1] with upscaling factor 3

two popular deep-learning-based single-image SR methods. In Table 4, we provide a summary of quantitative evaluation on several datasets. The results of the other four methods are the same as reported at FSRCNN [12]. The setting of our ANNet to run the experiment for comparison is the deep integration subnetwork having two layers with filter sizes 5 and 3, respectively. It is trained on the public small General-100 dataset instead on our own big dataset. Table 4 shows that our proposed ANNet outperforms A+, SRF, and SRCNN. On this setting and test dataset, our ANNet can improve roughly to 1.97, 0.38, 0.24, and 0.1 dB on average over all three test datasets, in comparison with Bicubic, SRCNN, A+, and SRF. Our ANNet are comparable with SCN because the difference of average PSNR value is just 0.04 dB. Furthermore, SCN needs multiple cascade operations to get the best performance. As discussed in the above section, we can increase our network depth or use a larger training dataset to get better performance. Some qualitative results are also given in Figs. 4, 5, and 6. Figure 4 shows the visual quality comparison of single-image SR on image butterfly from Set5 with an upscaling factor of 3. Figures 5 and 6 use the image baby and woman from Set5 as two examples with an upscaling factor 4, respectively. Obviously, our ANNet can recover more image details. All these results demonstrate our proposed ANNet is a robust single-image SR method.

6 Conclusions

In this paper, we focus on two rarely studied problems for deep-learning-based single-image super-resolution: one is how to take image prior into account for

deep-learning-based approaches, the other one is how to avoid the neuron compromising to different image contents. To the first problem, we use the transfer learning technology to transfer the weights of a projection matrix trained with strict image prior constraint to one convolution layer. To the second problem, the proposed ANNet anchors each input feature vector to one of the dictionary atoms and maps it to the high-resolution space with the corresponding convolution layer. By solving these two problems, we have proposed an anchored neighborhood deep network for single-image super-resolution. Experimental results show that our proposed method outperforms many state-of-the-art single-image super-resolution methods. Our experiments demonstrate that the deeper our network, the better the performance. Furthermore, a large training dataset is a good trick to improve the network performance, which inspires us to use a larger dataset like ImageNet to train our network for practical application.

7 Future work

One of the problems we focus on is how to take image prior into account for deep-learning-based image super-resolution. However, we have just used the image sparse prior in this paper. Apart from the sparse prior, many other kinds of image prior, e.g., local smoothing and non-local self-similarity, have been well studied in the traditional non-deep-learning-based image restoration methods. Therefore, the natural way to expand our work is to take more image prior into account and explore more effective way to use these image priors for deep-learning-based image super-resolution. On the other hand, multiple

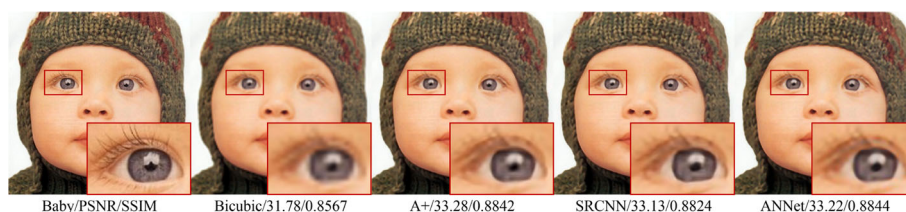


Fig. 5 Visual quality comparison of single image SR on image baby from Set5 [1] with upscaling factor 4



Baby/PSNR/SSIM Bicubic/26.50/0.8320 A+/28.66/0.8822 SRCNN/28.89/0.8837 ANNet/29.47/0.8858

Fig. 6 Visual quality comparison of single image SR on image woman from Set5 [1] with upscaling factor 4

frames input of the video super-resolution will offer more abundant image information. In the future, we will also investigate extending our proposed anchored neighborhood deep network into a spatio-temporal network to super-resolve one frame from multiple neighboring frames.

Abbreviations

AI: Artificial intelligence; Adam: Adaptive moment estimation; ANR: Anchored neighborhood regression; CNN: Convolutional neural network; FHD: Full high definition; FSRCNN: Fast SRCNN; HR: High resolution; JPEG: Joint picture expert group; K-SVD: K-singular value decomposition; LLE: Locally linear embedding; LR: Low resolution; NE: Neighbor embedding; ReLU: Rectified linear units; SF: Simple functions; SR: Super resolution; SRCNN: Super resolution CNN; USB: Universal serial bus; UHD: Ultra-high definition; VDSR: Very deep convolutional SR

Acknowledgements

We would like to acknowledge all our team members, especially Min Gao and Xinwei Gao, for their constructive suggestions on deep-learning-based image restoration and image compression. We would also like to acknowledge NVIDIA Corporation who kindly provided two sets of GPU.

Funding

This work is partially funded by the MOEMicrosoft Key Laboratory of Natural Language Processing and Speech, Harbin Institute of Technology, the Major State Basic Research Development Program of China (973 Program 2015CB351804), and the National Natural Science Foundation of China under Grant Nos. 61572155, 61672188, and 61272386.

Availability of data and materials

The training dataset of General-100 dataset [12] and the testing datasets of Set5 [1] (5 images), Set14 [37] (14 images), and BSD200 [23] (200 images) are public. Please refer to the corresponding project website for downloading these datasets. The source code of the proposed method and our self-established dataset are available from the corresponding author on reasonable request, and it will also soon be available from Github.

Authors' contributions

WS, SL, and FJ conceived and designed the study. WS performed the experiments. WS and SL wrote the paper. FJ, DZ, and ZT reviewed and edited the manuscript. All authors read and approved the manuscript.

Authors' information

Wuzhen Shi is now a PhD candidate at the Harbin Institute of Technology (HIT), Harbin, China. He received master's degree from Northwest A & F University, Yangling, Shaanxi, China, in 2014 and received bachelor's degree from Shenyang Agricultural University, Shenyang, China, in 2012. His research interest is deep-learning-based image restoration. Shaohui Liu received the BS, MS, and PhD degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 2000, 2002, and

2007, respectively. He is now an Associated Professor in the Department of Computer Science, HIT, and his research interests include data compression, pattern recognition, and image and video processing.

Feng Jiang received the BS, MS, and PhD degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 2001, 2003, and 2008, respectively. He is now an associated professor in the Department of Computer Science, HIT, and a visiting scholar in the School of Electrical Engineering, Princeton University. His research interests include computer vision, pattern recognition, and image and video processing.

Debin Zhao received the BS, MS, and PhD degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 1985, 1988, and 1998, respectively. He is now a professor in the Department of Computer Science, HIT. He has published over 200 technical articles in refereed journals and conference proceedings in the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.

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.

Author details

¹Harbin Institute of Technology, Harbin 150001, China. ²Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China.

Received: 7 January 2018 Accepted: 20 April 2018

Published online: 24 May 2018

References

1. M Bevilacqua, A Roumy, C Guillemot, ML Alberi-Morel, *Low-complexity single-image super-resolution based on nonnegative neighbor embedding*. (British Machine Vision Association, BMVA, 2012). <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=caf6296bM0b6eM48c7M8336Mbde1f8cfff1a7&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
2. J Caballero, C Ledig, A Aitken, A Acosta, J Totz, Z Wang, W Shi, in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. Real-time video super-resolution with spatio-temporal networks and motion compensation, (2017), pp. 2848–2857. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=fffb3e7Mb3c2M44e8Ma22eM42bbf72807b0&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>

3. H Chang, DY Yeung, Y Xiong, in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1. Super-resolution through neighbor embedding (Institute of Electrical and Electronics Engineers Computer Society, 2004), pp. I–I. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=32a31d81M1dbcM4166M9621Mc35ebfa5d73f&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=5&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
4. BW Chen, X He, SY Kung, Support vector analysis of large-scale data based on kernels with iteratively increasing order. *J. Supercomput.* **72**(9), 3297–3311 (2015)
5. BW Chen, M Imran, M Guizani, Cognitive sensors based on ridge phase-smoothing localization and multiregional histograms of oriented gradients. *IEEE Trans. Emerg. Top. Comput.* **99**(1), 1–1 (2016)
6. BW Chen, W Ji, Geo-conquesting based on graph analysis for crowdsourced metatrails from mobile sensing. *IEEE Commun. Mag.* **55**(1), 92–97 (2017)
7. BW Chen, L Yang, Y Gu, Privacy-preserved big data analysis based on asymmetric imputation kernels. *Futur. Gener. Comput. Syst.* **78**(2), 859–866 (2018)
8. YH Chen, GL Peng, CH Xie, W Zhang, CH Li, SH Liu, *Acclin: Bridging the gap between artificial and real bearing damages for bearing fault diagnosis*, (2018). <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=974dc1acM3428M427cMaf68M88b679d61353&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
9. C Dong, Y Deng, C Change Loy, X Tang, in *Proceedings of the IEEE International Conference on Computer Vision*. Compression artifacts reduction by a deep convolutional network (Institute of Electrical and Electronics Engineers Inc., 2015), pp. 576–584. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=f7c5d2b6M9aacM4421M96fdM8899597110c7&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
10. C Dong, CC Loy, K He, X Tang, in *European Conference on Computer Vision*. Learning a deep convolutional network for image super-resolution (Springer Verlag, 2014), pp. 184–199. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=12803a52M31fbM4d01M8dc2M7924976ea23f&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
11. C Dong, CC Loy, K He, X Tang, Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern. Anal. Mach. Intell.* **38**(2), 295–307 (2016)
12. C Dong, CC Loy, X Tang, in *European Conference on Computer Vision*. Accelerating the super-resolution convolutional neural network (Springer Verlag, 2016), pp. 391–407. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=5955183cM9d92M4761Mbcd5Ma1e5a47cc81a&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
13. K He, X Zhang, S Ren, J Sun, in *Proceedings of the IEEE international conference on computer vision*. Delving deep into rectifiers: surpassing human-level performance on imagenet classification (Institute of Electrical and Electronics Engineers Inc., 2015), pp. 1026–1034. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=e2442619Mb94fM4279Mbce2M3bc9d411de51&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
14. Y Huang, W Wang, L Wang, in *Advances in Neural Information Processing Systems*. Bidirectional recurrent convolutional networks for multi-frame super-resolution (Neural information processing systems foundation, 2015), pp. 235–243. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=72e0c168M759cM424fMa7c5M5ddad54500c7&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
15. J Johnson, A Alahi, L Fei-Fei, in *European Conference on Computer Vision*. Perceptual losses for real-time style transfer and super-resolution (Springer Verlag, 2016), pp. 694–711. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=756f3295M76c3M48f7Mb5c9M144b7fdbfa41&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
16. A Kappeler, S Yoo, Q Dai, AK Katsaggelos, Video super-resolution with convolutional neural networks. *IEEE Trans. Comput. Imaging.* **2**(2), 109–122 (2016)
17. J Kim, J Kwon Lee, K Mu Lee, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Accurate image super-resolution using very deep convolutional networks (IEEE Computer Society, 2016), pp. 1646–1654. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=5537f4a9M9a32M4c26M8acdM34670ec6fc4&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
18. D Kingma, J Ba, in *International Conference on Learning Representations (ICLR2015)*. Adam: a method for stochastic optimization, (2015). arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>
19. A Krizhevsky, I Sutskever, GE Hinton, in *Advances in neural information processing systems*. Imagenet classification with deep convolutional neural networks (Neural information processing system foundation, Canada, 2012), pp. 1097–1105. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=558ed086M3315M42beMa6bcM314dbe5adfdf&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=2&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
20. C Li, W Zhang, G Peng, S Liu, Bearing Fault Diagnosis Using Fully-Connected Winner-Take-All Autoencoder[J]. *IEEE Access*, 6103–6115 (2017). <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=596b49bcM49c2M4664M9a6cMf7ae8f1a2a24&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
21. CH Li, W Zhang, GL Peng, SH Liu, Bearing fault diagnosis using fully-connected winner-take-all autoencoder. *IEEE Access.* **6**, 6103–6115 (2017)
22. X Liu, X Wu, D Zhao, in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. Sparsity-based soft decoding of compressed images in transform domain (IEEE Computer Society, 2013), pp. 563–566. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=d9f77bf5Mb0f0M43e6M96d7M76c38c39bbfb&usageZone=resultslst&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
23. D Martin, C Fowlkes, D Tal, J Malik, in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 2. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics (Institute of Electrical and Electronics Engineers Inc., 2001), pp. 416–423. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=>

- 15a4198dM46f8M47f7Mb681M9c75fe00d020&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes
24. ST Roweis, LK Saul, Nonlinear dimensionality reduction by locally linear embedding. *Science*. **290**(5500), 2323–2326 (2000)
 25. S Schuler, C Leistner, in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Fast and accurate image upscaling with super-resolution forests (IEEE Computer Society, 2015). <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=060d721fMcba2M4e9fM95c2Mf20da6a9dd53&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 26. B Shen, N Chilamkurti, R Wang, X Zhou, S Wang, *Deadline-aware rate allocation for iot services in data center network*, (2017). <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=d61b88fbM31dfM4a19Ma0c5M2db50d3ffa3&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 27. B Shen, X Zhou, M Kim, Mixed scheduling with heterogeneous delay constraints in cyber-physical systems. *Futur. Gener. Comput. Syst.* **61**(8), 108–117 (2016)
 28. B Shen, X Zhou, R Wang, A delay-aware schedule method for distributed information fusion with elastic and inelastic traffic. *Inf. Fusion*. **36**(7), 68–79 (2017)
 29. W Shi, J Caballero, F Huszár, J Totz, AP Aitken, R Bishop, D Rueckert, Z Wang, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network (IEEE Computer Society, 2016), pp. 1874–1883. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=694d88c8Med6eM42b7Mb6d9Ma5a3a3c2b67b&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 30. R Timofte, V De Smet, L Van Gool, in *Proceedings of the IEEE International Conference on Computer Vision*. Anchored neighborhood regression for fast example-based super-resolution (Institute of Electrical and Electronics Engineers Inc., 2013), pp. 1920–1927. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=f011e4e4M1f2fM456cMac3aM36d34ba54572&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 31. R Timofte, V De Smet, L Van Gool, in *Asian Conference on Computer Vision*. A+: Adjusted anchored neighborhood regression for fast super-resolution (Springer Verlag, 2014), pp. 111–126. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=1097e68cM0485M4e62Mb4ddM4ca34a2a3887&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 32. JA Tropp, AC Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theory*. **53**(12), 4655–4666 (2007)
 33. A Vedaldi, K Lenc, in *Proceedings of the 23rd, ACM international conference on Multimedia*. Matconvnet: Convolutional neural networks for matlab (Association for Computing Machinery, Inc., 2015), pp. 689–692. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=5218517dM4c7eM4743Mb58cM6bc27cdf5a5&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 34. Z Wang, D Liu, J Yang, W Han, T Huang, in *Proceedings of the IEEE International Conference on Computer Vision*. Deep networks for image super-resolution with sparse prior (Institute of Electrical and Electronics Engineers Inc., 2015), pp. 370–378. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=17d7184aMe136M4d35M96edMb2282ecacdf7&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 35. J Yang, J Wright, TS Huang, Y Ma, Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **19**(11), 2861–2873 (2010)
 36. MD Zeiler, R Fergus, in *European conference on computer vision*. Visualizing and understanding convolutional networks (Springer Verlag, 2014), pp. 818–833. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=ec17fa35Md67dM4840M83f7M023f167dfd70&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 37. R Zeyde, M Elad, M Protter, in *International conference on curves and surfaces*. On single image scale-up using sparse-representations (Springer Verlag, Heidelberg, 2010), pp. 711–730. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=87e27abfMc848M48c3MacbfM3ecd93a93759&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>
 38. J Zhang, C Zhao, R Xiong, S Ma, D Zhao, *Image super-resolution via dual-dictionary learning and sparse representation*. (IEEE Computer Society, Washington, 2012), pp. 1688–1691. <https://www.engineeringvillage.com/search/doc/detailed.url?SEARCHID=20abcd56M257fM49e1Mbb5fM23b6731e5923&usageZone=resultslist&usageOrigin=searchresults&pageType=quickSearch&searchtype=quickSearch&CID=quickSearchDetailedFormat&DOCINDEX=1&database=1&format=quickSearchDetailedFormat&tagscope=&displayPagination=yes>

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