Single image super-resolution: a comprehensive review and recent insight

Article *in* Frontiers of Computer Science (electronic) · February 2023

DOI: 10.1007/s11704-023-2588-9

2 authors, including:

Hanadi Al-Mekhlafi
Tianjin University

Front. Comput. Sci., 2024, 18(1): 181702
https://doi.org/10.1007/s11704-023-2588-9

REVIEW ARTICLE

Single image super-resolution: a comprehensive review and recent insight

Hanadi AL-MEKHLAFI, Shiguang LIU (✉)

College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

Abstract Super-resolution (SR) is a long-standing problem in image processing and computer vision and has attracted great attention from researchers over the decades. The main concept of SR is to reconstruct images from low-resolution (LR) to high-resolution (HR).It is an ongoing process in image technology, through up-sampling, de-blurring, and de-noising. Convolution neural network (CNN) has been widely used to enhance the resolution of images in recent years. Several alternative methods use deep learning to improve the progress of image super-resolution based on CNN. Here, we review the recent findings of single image super-resolution using deep learning with an emphasis on distillation knowledge used to enhance image super-resolution., it is also to highlight the potential applications of image super-resolution in security monitoring, medical diagnosis, microscopy image processing, satellite remote sensing, communication transmission, the digital multimedia industry and video enhancement. Finally, we present the challenges and assess future trends in super-resolution based on deep learning.

Keywords super-resolution, deep learning, single-image, interpolation-based, learning-based, reconstruction-based

1 Introduction

Super-resolution (SR) is a technique that recovers high-resolution images from observed low-resolution images. It is one of the most important classes used in image processing systems that aim to improve the resolution of given images and videos. Super-resolution plays an important role in different applications, such as medical satellite imaging, imaging systems, astronomical imaging, and surveillance. In recent years, there are remarkable innovations in the field of image super-resolution with deep learning (DL). For instance, example-based image super-resolution methods have been used to increase the resolution of an image through the concept of prior knowledge under learning models. These methods either exploit internal similarities of the same image [1−3], or learn mapping functions from external low- and high-resolution exemplar pairs [4−12]. At this stage, the practical application of high-resolution images is limited to various imaging equipment and complex external environments. It is difficult for generated images to qualify for the requirements of practical applications. With the increasing popularity of neural networks, the image is an important carrier of two-dimensional information that flows in the neural network. The image transmission process is becoming increasingly difficult and expensive as the number of generated images rises, resulting in image information loss and required high storage capacity. Therefore, if the network bandwidth and storage space are limited, the images will require compression, transmission, and storing to reduce overall delays and storage costs. These procedures inevitably cause significant degradation in image quality. To enhance SR methods and reduce memory, researchers have proposed end-to-end trainable layers by substituting predetermined upsampling embedded at the model's end to conduct low dimensional space for the majority of computation efficiency. In this survey, we provide comprehensive insights on the recent progress of SR technology with distillation knowledge SR methods. Recently, some reviews have previously been published on deep learning-based image super-resolution. Yang et al. [13] and Zhou et al. [14] reviewed single image super-resolution (SISR) based on deep learning methods. Related work by Ha et al. [15] deeply discussed the state-of-the-art single-image super-resolution (SISR) models and categorized them using the CNN and GAN methods. Zhang et al. [16] studied the scope of image super-resolution methods using CNN for space applications by specifically reviewing the SRCNN [17], VDSR [18], FSRCNN [19], and DRCN [20] methods. Li et al. [21] also studied state-of-the-art image SR methods, focusing on real-time methods based on GANs and CNNs. Wang et al. [22] also extensively reviewed state-of-the-art image SR methods based on deep learning, classifying existing methods into supervised SR, unsupervised SR, and domain-specific SR by introducing performance evaluation metrics and benchmark datasets. While Syed et al. [23] surveyed single-image super-resolution based on deep learning and introduced initial classical methods used for image super-resolution. Their survey categorized image SR methods into classical methods, supervised and unsupervised methods, and domain-specific super-resolution methods,

Received September 16, 2022; accepted February 1, 2023

E-mail: lsg@tju.edu.cn

thereby providing an introduction of performance SR methods for specific applications and benchmark. Their review focused on an overview of both classical and deep learning-based methods. Liu et al. [24] presented findings on state-of-the-art blind image SR, and classified existing models based on data used to solve the SR approach and degradation modeling. Zhu et al. [25] studied SR methods based on attention mechanisms. These reviews did not cover the domain of super-resolution as a whole, so this survey bridge the research gap by giving a brief summary of super-resolution using deep learning-based methods with knowledge distillation. Hence, the purpose of this paper is to provide a collection of recent advances in super-resolution approaches based on deep learning. Our contributions can be summarized as follows:

- This paper presents a brief tutorial of image super-resolution techniques which may help to know clearly the super-resolution tasks;
- It also compares between the different super-resolution stages and shows the steps to build the structure for the super-resolution method;
- This work highlights the most recent advances in different image SR methods and reveals the extensions for deep learning based on image super-resolution;
- This survey deeply focuses on distillation knowledge SR which is research gaps in other reviews;
- Finally, this work discusses the difficulties and challenge of image super-resolution and presents potential solutions for future perspectives.

The organization of this paper is as follows: Section 2 presents the problem definition of SISR, and Section 3 describes comprehensive improvement and extensions deep learning-based super resolutions. Section 4 covers SR's application fields, and Section 5 highlights the quality evaluation of an image. Finally, in Section 6, we summarize the main challenges and future directions of image SR. The taxonomy for SISR has been summarized in Fig. 1.

2 Problem formulation

Basically, the super-resolution technique tries to rebuild an HR image from an LR observation. In recent decades, this technique has been a hot topic in the field of image processing. High-resolution images are essential for many applications that need zooming technology, such as medical imaging, surveillance, and satellite imaging. Deep learning is a part of machine learning dealing with artificial neural networks. Thus, SR attempts to correct the problem of minute details within traditional algorithms, such as an inability to eliminate defects and artifacts in upscaling approaches.

Super-resolution of an image is a complex problem, which is the root cause of the difficulty of super-resolution tasks, making accuracy and perceived quality urgent issues to be addressed and solved. Accuracy is the closest connection between a constructed super-resolution image and a target one based on the premise that perceived quality is for enhancing the visual quality of the generated image. In addition, as the computing capabilities of mobile devices continue to increase, it is becoming increasingly important to deploy super-resolution neural networks on mobile terminals [26,27].

SISR aims to generate an image of high-resolution from an image of low-resolution, as shown in Fig. 2(a). We will denote I_{lr} as the image of low-resolution (LR) and I_{hr} as the image of high-resolution (HR); I_{lr} is modeled as the output of the following degradation procedure (1):

$$I_{lr} = DFBI_{hr} + N, \tag{1}$$

where I_{hr} is the reconstructing HR image, D represents downsampling, F represents geometric deformation, B represents blur and N the noise process as shown in Fig. 2(b).

3 Image super-resolution techniques

Super-resolution of an image not only plays a crucial role in human life and work but can also promote the continuous development of the computer vision field. It always piqued the attention of scholars in the graphics and computer vision field due to its huge application prospects and profound research

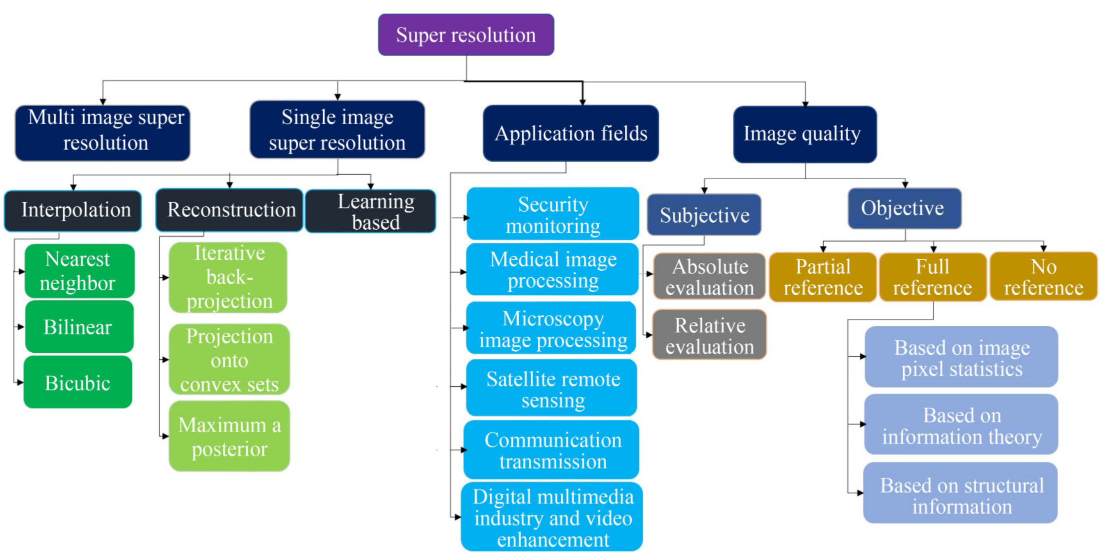

Fig. 1 The taxonomy for single image super-resolution

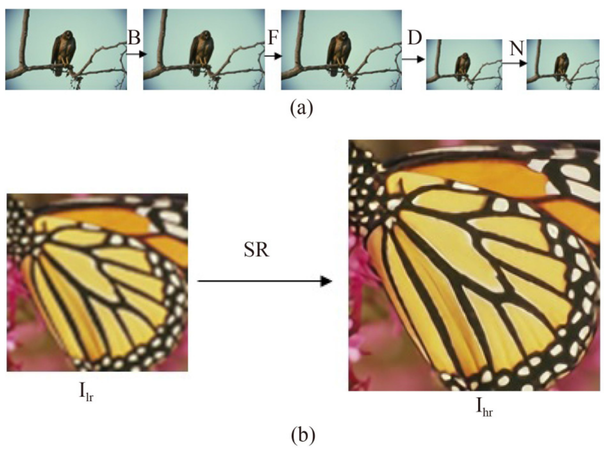

Fig. 2 Illustration of image super-resolution. (a) Degradation model; (b) low and high super-resolution

significance. Image super-resolution has been continuously developed, starting from the initial interpolation-based method to the reconstruction-based algorithm, as illustrated in Fig. 1.

Recently, a big wave of deep learning has brought new life to image super-resolution, leading to long-term developments in deep learning on super-resolution tasks. The image interpolation algorithm plays a role in adjusting the image dimension and image restoration in reconstructing a degraded image. A reconstruction algorithm is essential for image super-resolution [28]. The algorithms described therein assume that if a target image has high-resolution, then a degraded image with low-resolution will also include certain geometric displacements related to the target imaging. They typically utilize the variations between degraded LR image and targeted image with HR, and are thus based on reconstruction.

3.1 Super-resolution of single image based on interpolation method

Interpolation is the simplest and fastest method to provide image super-resolution on the sampled visual data acquired from the sensor [29]. Within this process, image interpolation can estimate new pixels through pixel interpolation, which is the simplest way to enhance image resolution [30].The method based on interpolation processes the image using a low-pass interpolation check. Interpolation algorithms commonly consist of bicubic, bilinear and nearest-neighbor interpolation.

3.1.1 Nearest neighbor interpolation

Nearest neighbor interpolation means that the value of an interpolated pixel can be represented by the closest neighboring pixel [31]. By comparing the values of neighboring pixels, this algorithm chooses the one with the closest approximation to the value produced $x \geqslant 0$:

$$p(x) = \begin{cases} p(\lfloor x \rfloor), & \text{if } x - \lfloor x \rfloor < 0.5, \\ p(\lfloor x \rfloor + 1), & \text{otherwise}, \end{cases} \tag{2}$$

where $\lfloor . \rfloor$ is known as a floor operator, and it represents the largest integer that is less than or equal to the argument in the given equation [32].

As a result of nearest neighbor lacks subpixel accuracy and produces sharp discontinuities, especially when dealing with arbitrary transformations of scale and rotation. Compared to some interpolation algorithms, this algorithm has the advantage of being less complex and very simple, and drawbacks such as blurring, as well it retains the original noise details in the reconstructed image, which can be advantageous in some image processing tasks.

3.1.2 Bilinear interpolation

The basic concept of bilinear interpolation is to introduce unknown locations using horizontal and vertical values. It estimates the value of the interpolated pixel by bilinearly interpolating the values of the pixels around it. Bilinear interpolation is similar to the previous method, except that the nearest neighbor method uses a value to determine the pixel value which will be interpolated [30].

For any (x, y) coordinate in a 2D image f, the following bilinear interpolation equations are [33,34]:

$$f_{y1} = f11 + \tfrac{f21 - f11}{x2 - x_1}(x - x_1),$$
$$f_{y2} = f12 + \tfrac{f22 - f12}{x2 - x_1}(x - x_1),$$
$$f_{x,y} = fy1 + \tfrac{fy2 - fy1}{y2 - y_1}(y - y_1),$$

where

$$\begin{aligned} x_1 &= \lfloor x \rfloor, & f11 &= f(x_1 + y_1), \\ x_2 &= \lfloor x \rfloor + 1, & f12 &= f(x_1 + y_2), \\ y_1 &= \lfloor y \rfloor, & f21 &= f(x_2 + y_1), \\ y_2 &= \lfloor y \rfloor + 1, & f22 &= f(x_2 + y_2). \end{aligned}$$

The main disadvantages of bilinear interpolation are poor image detail preservation and the generation of significant aliasing artifacts for images that have been rotated. This method is utilized in some tasks as an intermediate transformation in practice, mostly because it does not produce undershoot artifacts.

3.1.3 Bicubic interpolation

The bicubic interpolation method is more complex compared with the other methods because it involves more pixel values, which are 16 pixels around the unknown pixel that perform cubic interpolation [35]. Bicubic interpolation takes the sum of the convolution weights of these 16 pixels to calculate the overall pixel value. Although the bicubic interpolation requires more calculations and takes longer to perform, the restoration effect is far superior to the last two methods, so it is commonly used for different applications.

When speed is not a problem, bicubic interpolation is usually utilized over nearest neighbor or bilinear interpolation into image resembling as the method of choice. In comparison to bilinear interpolation, which only considers pixels (2×2), bicubic interpolation accounts for 16 pixels (4×4). With bicubic interpolation, images that look like the original are smoother and have less interpolation distortion. In many image processing tasks, bicubic interpolation is the commonly used method. Bicub interpolation can be expressed as a general equation, which is:

Assume that the values of the function f as well as its derivatives f_x, f_y, and f_{xy} are known at all four corners $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ of the unit square, respectively. Then, the interpolated for this method as:

$$p(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij}x^i y^j. \tag{3}$$

The interpolation problem involves calculating 16 coefficients a_{ij}. By plugging in values for the function into $p(x,y)$, Four equations can be obtained as follows [36]:

$f(0,0) = p(0,0) = a_{00}$,

$f(1,0) = p(1,0) = a_{00} + a_{10} + a_{20} + a_{30}$,

$f(0,1) = p(0,1) = a_{00} + a_{01} + a_{02} + a_{03}$,

$f(1,1) = p(1,1) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij}$.

In conclusion, the super-resolution method based on interpolation is fast and straightforward, with a small amount of calculation and real-time performance required. However, the interpolation method cannot be applied to complex scenes because it has better reconstruction ability for flat areas, but struggles to recover high-frequency information for scenes with more complex textures leading to blurry in the generated high-resolution images due to smooth edge production, as summarizing in Table 1.

3.2 Super-resolution of single image method based on reconstruction

Super-resolution of images based on the reconstruction method has become of considerable interest during the last decade. Its primary idea is to generate a high-resolution image from low-resolution degraded images. The reconstruction-based image super-resolution method uses motion estimation and prior knowledge to reconstruct a target high-resolution image [30]. Commonly used reconstruction-based super-resolution (SR) approaches include iterative back-projection (IBP), projection onto convex sets (POCS), and maximum a posteriori (MAP).

3.2.1 Iterative back-projection (IBP)

IBP was first used for image SR tasks by Irani and Peleg [37]. The estimation of HR image results can be mapped onto an LR degraded image to acquire an LR simulation image. The IBP method uses an iterative technique to continuously correct the simulation errors in the reconstructed image. IBP uses this method to consistently reduce reconstruction errors until convergence is achieved [30].

When estimating the HR image, IBP uses gradient-based optimization with a constant step to fine-tune the initial HR picture selection. In IBP, the goal of the cost function is to achieve a minimal difference between the simulated and observed values. The estimated HR solution can be obtained by an iterative formulation described by the equation.

Table 1 Comparison between several interpolation methods

Interpolation	Description
Nearest neighbor	It computes the new value pixel to be interpolated through is the closest neighbor pixel.
Bilinear	It computes the new value pixel to be interpolated is the value of its surrounding pixels bilinearly.
Bicubic	It computes 4×4 pixels around the unknown pixel to do cubic interpolation.

$$I_{HR}^{(i)} = I_{HR}^{(n-1)} + \frac{1}{p}\sum_{k=1}^{p}((((LR_k S\, i_k^{(n-1)}) \uparrow d) * H_{BP})W_k^T). \tag{4}$$

$I_{HR}^{(i)}$, $I_{HR}^{(n)}$ are the HR images estimated in $(n-1)^{\text{th}}$ and n^{th} iteration, respectively. H_{BP} is the back-projection kernel (BPK) which is isotropic in nature. $\uparrow d$, $\downarrow d$ are the up-sampling and down-sampling operation, respectively, by a factor d, and W_k^T is the backward warping transformation matrix. The equation of $S\, i_k^{(n-1)}$ as follows:

$$S\, i_k^{(n-1)} = ((((I_{HR}^{(n-1)}) \downarrow d) * B_k)W_k). \tag{5}$$

This algorithm's major advantages can be represented by its ability to use a variety of degradations and to process sequence data with cheap computational costs. In contrast, its limitations are to cause obvious abnormalities, such as ringing, in the digitally reconstructed image's borders. In addition, the high-frequency components of the digital image are highly impacted in high blurring and noisy conditions, which causes the performance to degrade [38].

3.2.2 Projection onto convex sets (POCS)

POCS was first used for image super-resolution tasks by Stark et al. [39]. The POCS method constrains a feasible solution space for the target image by utilizing the convex set to describe each constraint. The intersection of these constraints explains high-resolution images [30].

This efficient and straightforward approach is a collection theory for reconstructing images. The feasible region of the reconstructed image comprises intersection-consistent projective convex sets, and convex constraint sets gaven a versatile spatial observation model and a strong capability to embed prior knowledge.

This algorithm is iterative, with each step projecting a point from the solution space onto the convex set at the nearest available point [40]. Convergence to the intersection of the convex constraint sets is achieved after a finite number of iterations.

3.2.3 Maximum a posteriori (MAP)

MAP was also applied to image super-resolution tasks [41,42]. In order to solve super-resolution images, Markov Random Fields are utilized by MAP as prior knowledge. MAP uses low-resolution images as a condition and the maximum posterior probability to solve super-resolution images.

It is also commonly employed for super-resolution (SR) field-based reconstruction [30]. Thus, reconstruction-based methods often have good stability and adaptability. Still, they also have some drawbacks: the optimization of this method is more complicated, the solution time is longer, and the calculation complexity is higher. The reconstruction-based method requires multiple images from the same scene, and there is a small offset between these images. These limitations are often difficult to simulate in practical applications.

Given the LR observations LR_k and the a priori probability distribution of f, this work aims to determine the MAP estimate [43]. The ICM (iterated conditional modes) technique provides a computationally viable alternative for calculating the MAP estimate [44].

$$p(f_i \setminus LR, f_{\eta_i})\ p(LR_i \setminus f_i).p(f_i \setminus f_{\eta_i}). \quad (6)$$

For MAP methods, it's simple to use prior information in the HR image reconstruction phase. Additionally, a global optimization can be attained if we assume Gaussian noise and a convex prior model.

3.3 Learning based single image super-resolution

The learning-based image SR method is determined by the training data set for detecting the relationship in both low-resolution and high-resolution images. In the test phase, the function of learned mapping could potentially obtain the high-resolution target image based on the low-resolution input image. These methods have received widespread attention due to their high image reconstruction performance and subsequent results.

3.3.1 Traditional methods

Recently, learning-based super-resolution of an image has become a hot research topic which was first popularized by Freeman et al. [6]. The local linear embedding was introduced in related works to generate SR images using a linear combination of neighbors. The ANR [8] method first learned HR and LR dictionaries from high-resolution image blocks. A+ [45] is an improved version of ANR that uses the L2 norm instead of the L1 norm to speed up the learning and optimization process by enhancing the reconstruction quality. These researchers also introduced a tree-based SR method and which utilizes an extreme learning machine. Table 2 summarizes some of these traditional SR methods.

3.3.2 Deep learning-based methods

Deep learning has been used to learn the predicted image SR values involving techniques that have achieved significant improvements. Currently, super-resolution based deep learning has achieved better results than traditional methods. Moreover, convolutional neural network (CNN) has become essential for visual tasks, such as object detection, image generation, and image annotation. From there, researchers then designed a constant convolution layer of the feature map size to finally obtain an I_{hr}.

Lightweight network

For SR-based deep learning, previous works such as that of Dong et al. [17] initially proposed three layers of SRCNN which analyzed the endwise relationship between LR and HR images. Kim et al. [18] expanded the network's depth by growing a deeper network to approximate 20 layers in VDSR [18] over SRCNN [17], and used the deep features to learn the image's abstract features.

For the depth image SR (DSR) problem, Wen et al. [46] proposed deep color-guided cascade coarse-to-fine CNN. They used an edge-preserving filter kernel by deep CNN to estimate the ideal filters. Then, smaller filtering kernels are used to improve results and get a better HR depth image.

To reduce the computational cost, FSRCNN [19] directly inputs low-resolution images into the neural network and using up-sampling convolution to obtain image size features of high-resolution, and then generating images of high-resolution through convolutional layers. This method performs convolution in small features, which greatly reduces the calculation amount. Similarly, the input images with low-resolution have a larger receptive field. This method has gradually become a common structure in image super-resolution networks.

In order to reduce network parameters, DRRN [47] uses a recursive neural network and MemoryNet structure to learn the dependencies between different levels. The above methods interpolate a low resolution to a target-resolution image size, which then produces a constant convolution layer of a feature map size to finally achieve a high-resolution image. Kim et al. [20] proposed that DRCN uses the convolutional network with Deeply-Recursive. DRRN [47] used deep recursive ResNet in another work. It builds a recursive block structure in which several residual units are stacked to bring recursive learning into the residual branch.When comparing DRRN and DRCN, DRRN gets to share the weights among these residual units while DRCN gets to share the weights among the convolutional layers. As well, DRCN supervises every recursion to help backpropagation and prevent very deep models from exploding gradients.

For hierarchical features mechanism, Hui et al. [48] presented the IMDN method for extracting and aggregating hierarchical features based on channel attention. To reuse different hierarchical features, Ahn et al. [49] first proposed a cascading mechanism on a residual network (CARN). CBPN [50] uses cascading up and down-sampling layers to reconstruct features from both the LR and HR spaces simultaneously. This method known universally as multi-hierarchical feature fusion (MSRN) was first proposed by Li et al. [51]. Liu et al. [52] introduced the feature distillation connection (FDC), essentially similar to channel splitting. They created a residual feature distillation network (RFDN) after rethinking the information multi-distillation network (IMDN). Li et al. [53] summarized that an image super-resolution was based on multi-scale channel attention (MCSN). Their method used a multi-scale feature fusion block (MSFFB) to extract multi-scale features from filters. Additionally, the channel shuffle attention mechanism (CSAM) was introduced, which improves feature selection by promoting the flow of information across feature channels. In conclusion, a global feature fusion connection (GFFC) was proposed to enhance feature utilization. Table 3 summarizes these lightweight SR methods.

Attention mechanism

The attention mechanism is one of the networks that widely

Table 2 Example for traditional methods

Method	Concept	Details	Dataset(scale): PSNR/SSIM	Time
ANR	Anchored Neighborhood Regression	Sparse dictionary and regressor anchored	Set14(×3): 28.65/0.8093	0.69
A+	Adjusted Anchored Neighborhood Regression	Decision tree and extreme learning machine	Set14(×3): 29.13/0.8188	0.69

Table 3 Different lightweight methods

Method	Concept	Keywords	Parameters	Mult-Adds	Dataset(scale):PSNR/SSIM
SRCNN	Super-Resolution Convolutional Neural Network	Deep convolutional neural network	57K	52.7G	Set14(×4): 27.50/0.7513
VDSR	Very Deep Super-Resolution	Residual learning	665K	612.6G	Set14(×4): 28.01/0.7674
DRRN	Deep Recursive Residual Network	Recursive and residual connections	297K	6,796.9G	Set14(×4): 28.21/0.7720
DRCN	Deeply-Recursive Convolutional Network	Recursive-supervision and skip-connections	1,774K	17,974.3G	Set14(×4): 28.02/0.7670
FSRCNN	Fast Super-Resolution Convolutional Neural Network	Re-design the SRCNN, deconvolution layer	12K	4.6G	Set14(×4): 27.61/0.7550
IMDN	Information Multi-Distillation Network	Channel attention	715K	58.53G	Set14(×4): 28.58/0.7811
RFDN	Residual Feature Distillation Network	Feature distillation connection	550K	33.13G	Set14(×4): 27.50/0.7513
CARN	Cascading mechanism on A Residual Network	Cascading mechanism	1,592K	90.9G	Set14(×4): 28.60/0.7806
CBPN	Compact Back-Projection Network	Cascading up- and down-sampling layers	1,197K	97.9G	Set14(×4): 28.63/0.7813
MSRN	Multi-Scale residual Network for image Super-Resolution	Multi-hierarchical Feature fusion	4219K	349.9G	Set14(×4): 28.60/0.7751
MCSN	Multi-Scale Channel attention Network	Multi-scale feature fusion, channel shuffle attention mechanism and global feature fusion connection	1581K	728.4G	Set14(×4): 28.73/0.7847

used to improve the performance of SR. Zhang et al. initially introduced the channel-attention-based convolution neural network method to solve the single image super-resolution problems, namely very deep residual channel attention networks (RCAN) [54]. Similarly, Hu et al. [55] proposed the Squeeze-and-Excitation Networks (SENet).

Single image super-resolution with a recursive squeeze and excitation networks (SESR) were later introduced by Cheng et al. [56] to deal with channels, which achieved good results. Since attention is not bound to channels, Roy et al. [57] suggested that the Concurrent spatial and channel SE in fully CNN have obtained positive results in image segmentation. As such, Dai et al. [58] introduced a second-order Attention Neural network (SAN) for image super-resolution. Choi et al. [59] presented a selecting unit (SU) with the super-resolution network SelNet.

Laplacian pyramid attention (LA) is proposed by Anwar et al. [60] for super-resolution (DRLN) in which RNAN [61] creates an image restoration network with local and global feature attention. To extract the feature map's global features, RNAN employs a non-local structure by using multiple stacked convolutional layers to learn local features as an attention mechanism, thereby assigning each pixel on a feature map to a specific weight. Hu et al. [62] introduced the Spatial Feature Modulation and Channel-wise Network for SISR (CSFM), combining the spatial and channel attention mechanisms to benefit from both. The Residual Attention Module (RAM) for SISR was proposed by Kim et al. [63], which combines spatial and channel attention, but with some adjustments to make it more suitable for the super-resolution task.

Image super-resolution based on residual feature aggregation (RFANet) was proposed by Liu et al. [64] to improve spatial attention. Zheng et al. [65] introduced an upsampling attention network (UAN) for feature extraction and reconstruction. In their work, residual attention groups (RAGs) used non-local skip connections and various residual feature attention blocks (RFABs). Table 4 summarizes these SR methods based on attention mechanism.

Residual and dense learning

Residual learning has been excessively considered to combat the issue of declining feature reuse. EDSR [66] achieved great success by further expanding the network's width and removing the useless batch normalization layer, subsequently becoming the NTIRE 2017 super-resolution task champion by proving MDSR which is a multimagnification factor version of EDSR.

As shown in Fig. 3 many of the newly proposed super-

Table 4 Different attention methods

Method	Concept	Keywords	Parameters	Mult-Adds	Dataset(scale): PSNR/SSIM
RCAN	Residual channel attention network	Channel-attention-based convolution neural network	16M	–	Set14(×4): 28.87/0.7889
SENet	Squeeze-and-Excitation Network	Stacking a collection of Squeeze-and-Excitation blocks	28.1M	3.87G	ImageNet:–/–
SESR	Squeeze and Excitation Super-Resolution	Recursive squeeze and excitation networks	624K	–	Set14(×4): 28.32/0.784
ScSE	Spatial and channel Squeeze and Excitation	Concurrent spatial and channel SE blocks	3.3×10^4	–	–
SAN	Second-order Attention Neural network	Second-order channel attention, non-locally enhanced residual group	15.7M	–	Set14(×4): 28.92/0.7888
SelNet	Deep Network with Selection units	selection unit(SU)	1,417K	83.1G	Set14(×4): 28.49/0.7783
DRLN	Densely Residual Laplacian Network	Laplacian pyramid attention	3.4×10^4	–	Set14(×4): 28.94/0.7900
RNAN	Residual Non-Local Attention Networks	Non local attention	7.5M	–	Set14(×4): 28.83/0.7878
CSFM	Channel-wise and Spatial Feature Modulation	Spatial and channel attention mechanisms	(12-13)M	–	Set14(×4): 28.87/0.7886
RAM	Residual Attention Module	Residual spatial and channel attention	1389K	–	Set14(×4): 33.57/–
RFANet	Residual Feature Aggregation Network	Feature aggregation	11M	–	Set14(×4): 28.88/0.7894
UAN	Upsampling attention network	Non-local and local skip connection in residual attention groups (RAGs)	15.7M	–	Set14(×4): 28.92/0.7789

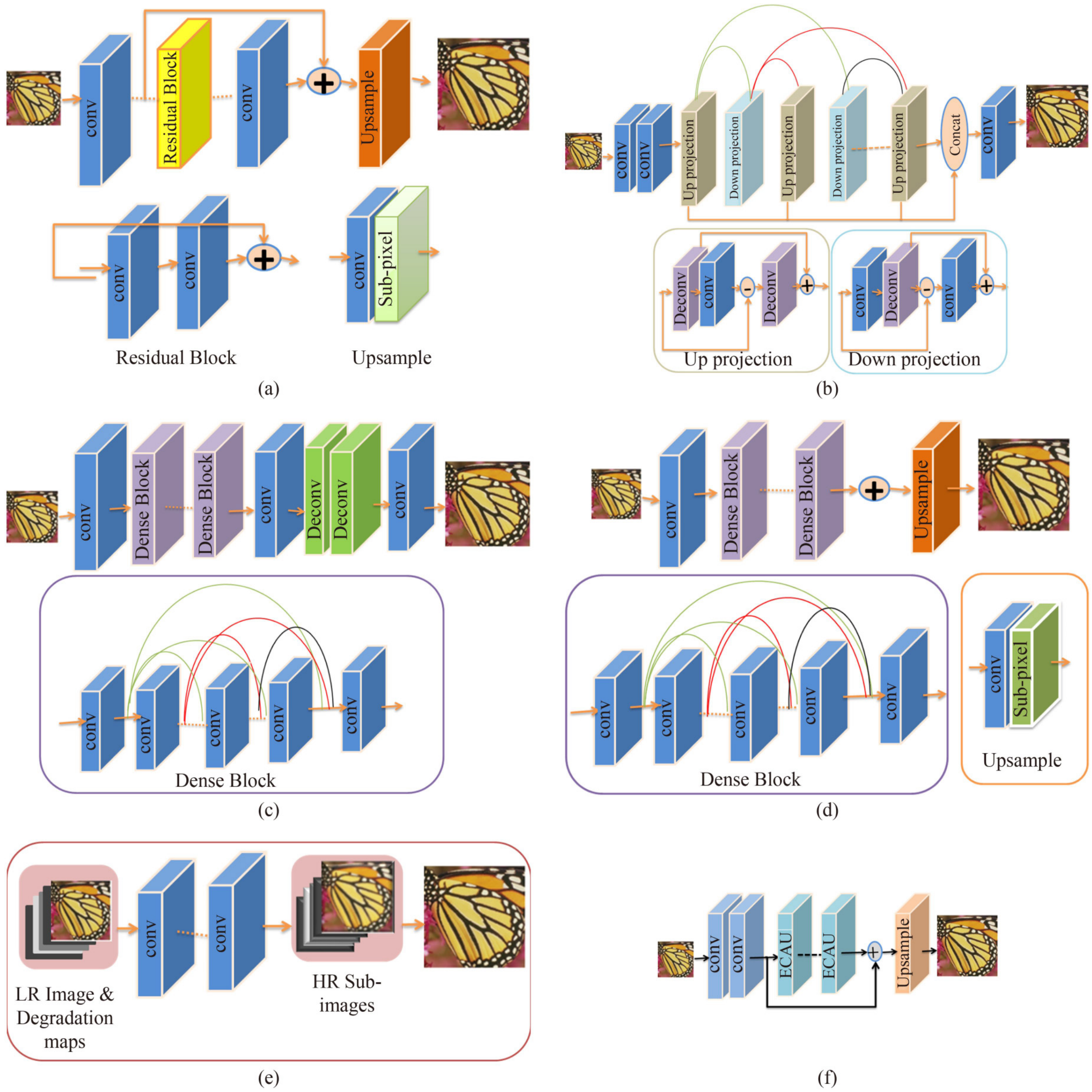

Fig. 3 Residual and dense learning methods. (a) EDSR, (b) D-DBPN, (c) SRDenseNet, (d) RDN, (e) SRMD, and (f) KARN

resolution methods are based on the EDSR and SRMDNF [67] by learning the mapping relationship between low-resolution images and high-resolution images in different downsampling modes. Zhang et al. [68] used the dense connection method RDN to extract hierarchical features and generate high-quality super-resolution images. Tong et al. [69] utilized image SR via dense skip connections and achieved a good result. Wang et al. [70] discovered the shared aims of the residual networks and dense networks, thereby proposing a Mixed Link network (MixNet) that develops its performance and decreases the number of parameters.

Recently, different methods of up-sampling layers have been continuously proposed, including bicubic upsampling, deconvolution, and sub-pixel convolution. In Table 5, ESPCNN [71] presented sub-pixel convolution in the super-resolution method. The feature map channels are exchanged for the feature map's space size in sub-pixel convolution. Each pixel on the channel is arranged according to a fixed method to obtain a larger spatially sized feature map. D-DBPN [72] uses an error feedback mechanism to gradually correct errors in high resolution reconstruction. Furthermore, it utilizes iterative upsampling and downsampling to learn how to map low-resolution images to high-resolution images and vice versa. In answer to this, Dun et al. [73] proposed a Kernel-Attended Residual Network (KARN), asserting that KARN has the best performance for feature fusion and feature representation. In addition, they observed that a multi-channel fusion block (MCFB) restores a large amount of textual feature information as well as a kernel attended block (KAB) for improving their network. As such, a space-feature re-calibration block (SFRB) was introduced that merged the calibration into spatial features.

Generative adversarial networks

Ledig et al. first used generative adversarial networks to image SR tasks (SRGAN [74]). Within the parameters of SRGAN, it is believed that a certain difference exists between the image

Table 5 Different residual and dense learning methods

Method	Concept	Details	Parameters	Mult-Adds	Dataset(scale):PSNR/SSIM
EDSR/ MDSR	Enhanced Deep Super-Resolution / Multi-scale Deep Super-Resolution	Removing unnecessary modules in conventional residual networks/ Using different upscaling factors	EDSR: 43M / MDSR: 8M	EDSR: 2890.0G/ MDSR: 407.5G	Set14(×4): 28.80/0.7876
SRMDNF	Super-Resolution Multiple Degradation Noise-Free	Multiple degradation, convolutional neural network	1555K	89.3G	Set14(×4): 28.35/0.7787
RDN	Residual Dense Network	Residual dense block	23M	–	Set14(×4): 28.81/.7871
SR-DenseNet	Super-Resolution Dense Network	Dense skip connections	–	–	Set14(×4): 28.50/0.7782
MixNet	Mixed Link Network	Residual network, dense network and Dual Path Network	48.5M	–	CIFAR-10: 4.19
ESPCNN	Efficient Sub-Pixel Convolutional Neural Network	Sub-pixel convolution	58K	–	Set14(×4): 27.73/–
KARN	Kernel-Attended Residual Network	Multi-channel fusion, kernel-attended and space-feature re-calibration	15M	–	Set14(×4): 28.73/0.785
D-DBPN	Dense-Deep Back-Projection Networks	Dense connections with back-projection units	10M	5715.4G	Set14(×4): 28.82/0.786

distribution generated by the MSE-based super resolution network and the actual image distribution. Thus, generative adversarial networks have been used to reduce this difference. By optimizing adversarial losses, the generative network produces images with high-resolution, which are more realistic. Despite having images generated by SRGAN with low PSNR values, they contain rich high-frequency details.

An unsupervised method is used by ZSSR [75] to study mapping from low to high-resolution in real test images. ESRGAN [76] solves the shortcomings of artifacts in SRGAN generated super-resolution images by removing the batch processing layer and using the techniques of small weight initialization, RaGAN [77], and residual scale [78], as shown in Fig. 4. Zhang et al. presented their findings on RankSRGAN [79], proposing that the generator could be optimized in perceptual metrics direction. A GAN composition of pseudo-paired SR and unpaired kernel correction networks was then proposed by Maeda [80]. Table 6 summarizes the details of these SR methods based on the GAN technique.

Deep learning-based blind super-resolution methods

Researchers have recently focused on blind super-resolution approaches [81–88], which aim to rebuild low-resolution images that are degraded in unknown and complex ways. Explicit degradation representations are used in the first category of blind SR approaches, which typically have two components: conditional restoration and degradation prediction. The two components listed above can be done

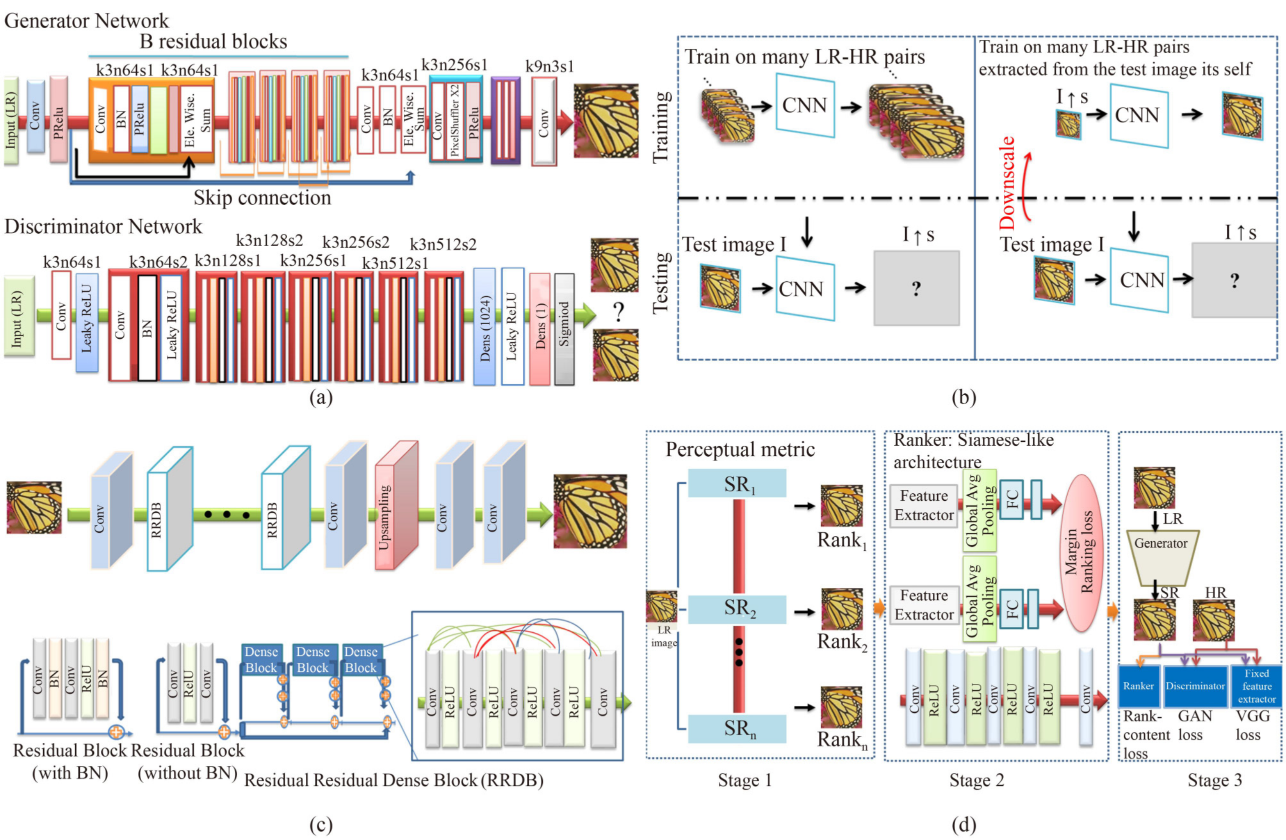

Fig. 4 Generative adversarial networks (GANs). (a) SRGAN [74], (b) ZSSR [75], (c) ESRGAN [76], and (d) RankSRGAN [79]

Table 6 Different GAN methods

Method	Concept	Keywords	Parameters	Mult-Adds	Dataset(scale): PSNR/SSIM
SRGAN	Super-Resolution Generative Adversarial Network	Generative adversarial network	1.5M	127.8G	Set14(×4): 26.02/0.7397
ZSSR	Zero-Shot Super-Resolution	Unsupervised learning	–	–	Set14(×4): 28.01/0.7651
ESRGAN	Enhanced SRGAN	Generative adversarial networks	–	–	Set14:28.88/0.7896
RaGAN	Relativistic average GAN	Generative adversarial networks	–	–	CIFAR-10: –/–
Rank-SRGAN	Super-Resolution Generative Adversarial Networks with Ranker	Ranker and Perceptual metrics	19,194 (VGG16)	–	Set14(×4): 26.57/–
Unpaired GAN	Unpaired SR using a Generative Adversarial Network	pseudo-paired SR network and unpaired kernel correction network	–	–	DIV2K(×4): 21.32/0.5541

separately [81] or together (iteratively) [84,85,87], using predefined degradation representations (such as degradation types and levels). Furthermore, artifacts will inevitably result from inaccurate degradation estimations.

Another classification for conditional restoration and degradation prediction is obtaining training pairs as close to real-world data as possible before training a unified network.

In Table 7, Bell-Kligler et al. [81] proposed KernelGAN method for training an internal generative adversarial network (GAN) on a single image. Liang et al. [86] introduced KernelGAN-FKP based on KernelGAN and Double-DIP [84,89] to integrate a flow-based kernel before entering the framework. Cornillère et al. [90] presented SRSVD as a method for evaluating the non-blind SR model output and optimizing kernel latent variables by minimizing the discriminator error. Kernels must be optimized patch by patch for spatially variant SR, which can be potentially inefficient and ineffective.

As illustrated in Fig. 5, Gu et al. [82] introduced IKC for kernel estimation using paired training data. They estimated the PCA feature of the kernel using a CNN network and then corrected it iteratively through alternating optimization.

Knowledge distillation

Knowledge distillation (KD) has garnered significant interest as a type of model compression and acceleration in which a smaller model (student) is effectively learned from a large model (teacher). Hinton et al. [91] were the first to propose this structure, which they termed "knowledge distillation". It is effectively learned from a large model (teacher) [92], as illustrated in Fig. 6. The community's interest in it is rapidly growing. Knowledge distillation aims to improve the student model's performance and is commonly used to compress networks. However, the semantic information is not used in training SISR models, making the ground-truth images identical to the original high-resolution images. Numerous lightweight super-resolution models have been made and applied to mobile devices such as cameras.

To distill from teacher super-resolution networks as in Fig. 7, Gao et al. [93] computed various statistical maps of feature

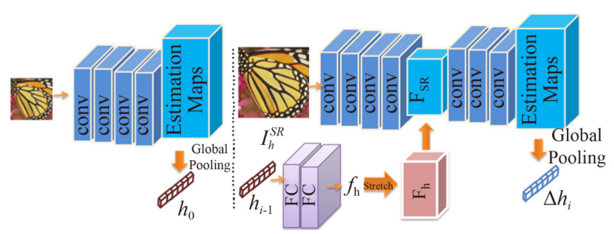

Fig. 5 Example for Blind super resolution method: IKC (Gu et al. [82])

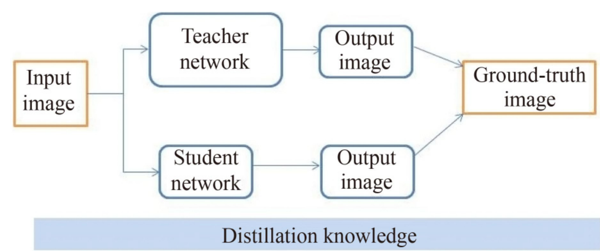

Fig. 6 Overview of distillation knowledge SR structure

Table 7 Different Blind super resolution methods

Method	Concept	Keywords	Details	Parameters	Dataset(scale): PSNR/SSIM
KernelGAN	Kernel generative adversarial network	Internal-GAN	● It estimated kernel based on the image patch recurrence property. ● GAN used to generate images. ● From the generator is derived the blur kernel. ● A generator is used to re-downscale the LR image, and a discriminator is used to ensure cross-scale patch similarity.	151K	KernelGAN+USRNet for DIV2K(×4): 23.69/0.6539
FKP	Flow-based kernel prior	Kernel prior	● It improved KernelGAN [81] and Double-DIP [84,89]. ● it used USRNet [88] to generate the output of SR based on the kernel estimation.	143K	KernelGAN-FKP + USRNet for DIV2K(×4): 25.46/0.7229
SRSVD	Blind Image Super-Resolution with Spatially Variant Degradations	Kernel discriminator	● Degradation-aware generator network used to generate HR image. ● A kernel discriminator used to identify the errors. ● kernel parameters reconstructed using a kernel discriminator to analyze artifacts that built by a Degradation-aware generator network.	<2M	BSD100 : 29.92/0.846
IKC	Iterative Kernel Correction	Blur kernel estimation	● It is based on the predict-and-correct principle. ● It used spatial feature transform layers for multiple blur kernels.	–	Set14(×4): 28.26/0.7688

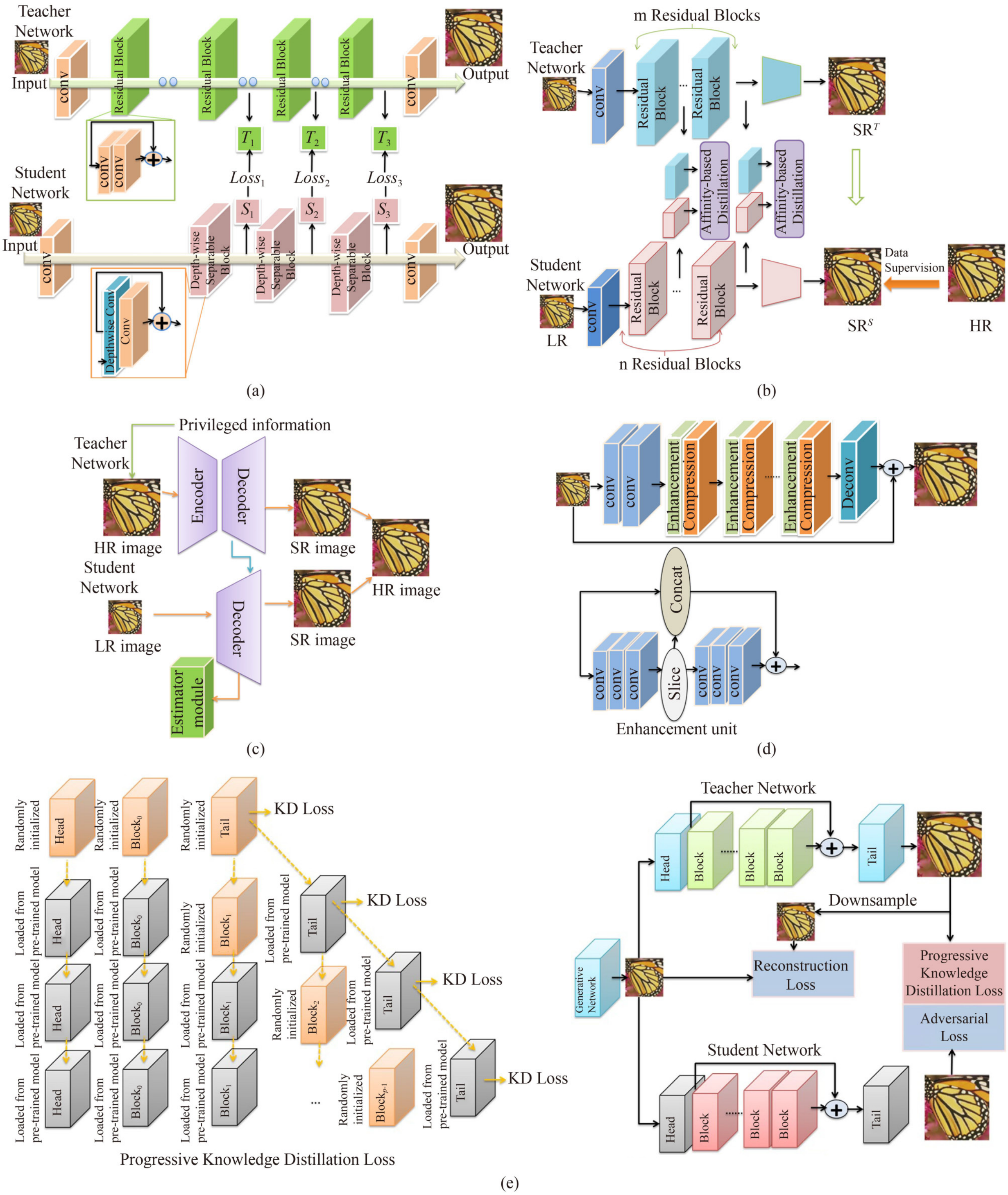

Fig. 7 Knowledge distillation methods. (a) TNSR/SNSR [93], (b) FAKD [94], (c) PISR [95], (d) IDN [97], and (e) DFKD [98]

maps. He et al. [94] introduced a feature affinity-based knowledge distillation (FAKD) for the SR method, which uses feature map correlation to improve distillation performance. Lee et al. [95] used ground truth high resolution images as privileged information and feature distillation to improve the performance of a convenient super resolution network. Zhang et al. [96] proposed the concept of the learnable pixel-wise importance map by utilizing teacher network prediction to initialize the importance map and improve the performance of lightweight super-resolution networks.

Furthermore, Hui et al. [97] and Jiang et al. [99] improved the performance of the lightweight SR network and presented new structures to perform distillation between different parts of the approaches. Zhang et al. [98] introduced a new data-free

Table 8 Different knowledge distillation methods

Method	Concept	Details	Parameters	Mult-Adds	Dataset(scale): PSNR/SSIM
TNSR/ SNSR	Teacher Network for SR / Student Network for SR	● To propagate knowledge, the features are extracted from the networks as Satistical maps.	TNSR=805.83K and SNSR=81.67K	TNSR=2.593G and SNSR=0.271G	TNSR-Set14(×4): 28.43/- and SNSR-Set14(×4): 27.43
FAKD	Feature Affinity-based Knowledge Distillation	● It used feature map correlation. ● It introduced distilling the feature-affinity matrix of the strategy of teacher-student network. ● It improved the performance of RCAN [54] and SAN [58] using the teacher network (TN) and student network (SN).	Tn-RCAN=15.59M and SN-RCAN= 5.17M	Tn-RCAN=36.80G and SN-RCAN= 12.93G	TN-RCAN-Set14(×4): 28.851/0.7885 and SN-RCAN-Set14(×4): 28.750/ 0.7859
PISR	Privileged Information for SR	● Ground-truth high-resolution (HR) images as privileged information. ● It improved the performance of FSRCNN.	13K	4.6G	Set14(×4): 27.77/0.7615
AIL	Adaptive importance learning	● Learnable pixel-wise importance map. ● It improved the performance of VDSR [18].	(VDSR-f32+AIL)= 166K	(VDSR-f32+AIL)= 642M	(VDSR-f32+AIL)-Set14(×4): 0.14/0.0041
IDN	Information Distillation Network	● Stacked information distillation. ● The key of IDN is the information distillation block. ● The local long and short path characteristics were extracted by combining the enhancement and compression units.	553K	89.0G	Set14(×4): 28.25/0.7730
DDRN	Deep distillation recursive network	● Multi-scale purification unit, ultra-dense residual blocks	297K	6,796.9G	Set14(×4): 28.21/0.7720
DFKD	Data-free knowledge distillation	● Iterative and progressive training strategy. ● It developed EDSR [66] and VDSR [18] models.	–	–	EDSR-Set14(×4): 28.33/0.7758 and VDSR-Set14(×4): 27.73/0.7617

knowledge distillation framework for SR. In addition, a generator network was used to approximate the original training data from the given teacher network. Unlike classification networks, whose outputs are probability distributions, the inputs and outputs of SR models are images with similar patterns. Table 8 summarizes these SR methods based on knowledge distillation.

4 The application fields of SR

The research on image super-resolution has essential significance and good application prospects. As such, image super-resolution can be found in wide applications in several areas:

4.1 SR in security monitoring

Safety monitoring is a powerful tool to ensure people's safety and crackdown crimes. At present, there are numerous safety monitoring devices in the cities with poor imaging quality, creating a growing demand for high-resolution imaging in the security field [100]. Because high-resolution images have more detailed information, this can potentially assist with the fight against crime. For example, it can provide vital image recognition information related to the physical descriptions of criminal suspects, vehicle license plate details, and identifying possible environmental hazards to name a few. However, updating and upgrading numerous safety monitoring equipment is often too expensive.

Image super-resolution can potentially solve these problems at a lower cost at the algorithm level. In addition, the security monitoring field also includes various visual tasks such as target detection, pedestrian detection, anomaly detection and vehicle recognition. Image super-resolution can be used at the preprocessing stage to provide more detailed information and improve the accuracy of these important tasks [101,102].

4.2 SR in medical images

In the diagnosis of diseases, super-resolution has become an effective tool for enhancing the resolution of medical images by making the shape and size of pixels clearer, particularly in CT [103], MRI [104], and B-ultrasound [105]. High-resolution medical images can provide more detailed and exact information of the target organ in the body such as the size and shape.

The image super-resolution method can provide greater clarity for medical images by improving the accuracy and reliability of a diagnosis. In addition, the existing medical imaging devices have different imaging effects, which is not conducive to learning reliable classifications or other artificial intelligence algorithms on these images [106]. By using an image super-resolution algorithm, these images can be transformed into images with the same size and more high-frequency information to improve the learnability and effect of artificial intelligence algorithms for medical images [107,108].

4.3 SR in microscopy image processing

Recently, SR has also played a vital role in microscopic image processing, with many works in this field exploring the biological structures and cell organelles of human beings and other organisms. Furthermore, by using these techniques and super-resolution strategies, we can now study the biological activity of medicine and its kinetic properties [109]. This has enormous potential for revolutionary application in this field by studying the molecular mechanism in life animal models using techniques like Lineage Tracing Methods [110] and Intravital Imaging [111]. Fluorescence microscopy is one of the most important and beneficial techniques in cell biological experiments, which uses super-resolution methods [112].

As such, the 2014 Nobel Prize in Chemistry was awarded to the Super-resolved fluorescence microscopy field [113].

Fluorescence microscope has been commonly used as an essential tool in the examination of the signaling in molecular pathways, living cells and tissues. Therefore, the fluorescence range might be increased with super-resolutions [114]. Concluding that SR is critically needed for diagnostic settings and it may play a pivotal role to study the pathophysiological conditions through live imaging super-resolution

4.4 SR in Satellite remote sensing field

As aerospace technology advances, satellite remote sensing plays an increasingly important role in the field of aerospace [115]. Because of the long-distance and the imaging equipment, remote sensing images' resolution is relatively low, which has caused adverse effects on land forest cover detection, wetland species detection and urban development planning. High-resolution remote sensing images can potentially contain more details of urban layout, improve the accuracy of maps and navigation, and make the calculation of area coverage more accurate [116].

In addition, it is difficult to maintain the imaging equipment on aircrafts, so further research of super-resolution images could have significant value in the remote sensing image field.

4.5 SR in the communication transmission field

The Internet has developed rapidly over several decades and has become an essential component of human life. However, there are challenges with transmitting high-quality and large-sized high-resolution images requiring large amounts of transmission due to bandwidth limitations, the consumption of network resources and increases in network delays, which creates a poor user experience [117].

The image super-resolution method can enlarge and restore low-resolution images that occupy small bandwidth during transmission by enhancing the image and video resolution. The SR method can reduce the cost of communication transmission [118], improve the quality of images and videos, and positively change user perception. In addition, the storage of image data overhead has doubled with increasingly rapid data transmissions. If you only store the low-resolution version of the image, you can use the SR approach for improving the image resolution when using the image, which can save and reduce the overall storage space of the image.

4.6 Digital multimedia industry and video Enhancement

With the continuous emergence of display devices such as ultra-high-definition televisions and smartphones, the demand for high-definition and ultra-high-definition images has also increased. However, the acquisition of high-definition or ultra-high-definition images and videos is usually difficult and expensive.

The use of image or video super-resolution methods can generate pictures and videos that match the resolution of the display device through display effects on Ultra HD devices. The image super-resolution method can use a cloud server or local algorithms to increase the image's size and quality obtained by these mobile camera-enabled devices, thereby providing users with more of their favorite images containing more texture details [71,119–121].

4.7 The field of computer vision

Image SR can be thought of as a task in image processing. The application of image-based artificial intelligence in daily life is becoming more widespread, bringing great convenience to work and life in the areas of object detection, image classification, pedestrian and vehicle detection, and target tracking to name a few.

Many other tasks in computer vision can benefit from image SR at a preprocessing stage [122]. Image super-resolution can make the generated image contain more detailed information, such as detailed edge information of objects, the characteristics and orientation of textures in the image, and interaction information between the subject and the environment in the image. This rich information can enhance the performance of tasks like classification, anomaly detection, object detection and image generation [123].

5 Image quality evaluation standards

Image quality evaluation is one of the basic image processing techniques and has a wide range of applications. There are two types of quality evaluations: subjective and objective quality evaluations [124,125]. A subject of subjective quality evaluation is when a person is scored based on the viewing experience, which is most accurate according to human perception. However, this is disadvantageous in that the evaluation is troublesome to prepare and often inefficient in application. Therefore, objective quality evaluation, which is based on image statistical information, is more commonly used in answer to this problem.

5.1 Subjective evaluation of image quality

A subjective evaluation of image quality involves only qualitative assessments made by individuals as observers. Untrained or well-trained observers are the most common options, and this strategy gives statistical outcomes higher significance. A sufficient number of observers are involved in the image subjective assessment to confirm that the assessment has statistical value [126]. There are two main types of subjective evaluation methods: relative and absolute evaluations.

5.1.1 Absolute evaluation

Absolute evaluation is based on the observer's knowledge and understanding of the absolute image quality based on overall evaluation performance. When calculating the direct quality evaluation value, the absolute evaluation can be used for an original image via the Double Stimulus Continuous Scale (DSCQS) method. This method plays an image that will be evaluated using a certain rule for a specific time, such as leaving a certain amount of time for the observer to score after playing; from there, they will average all the given scores as the final sequence. From a global perspective, there are provisions for the evaluation scale where the image quality can be divided into levels and expressed numerically. It is also known as the "full scale" of an image evaluation [127,128].

5.1.2 Relative evaluation

In the relative evaluation, there is no original image used as a reference. It can be achieved via comparing and evaluating a batch of images with each other to be evaluated by detecting the sequence of each image, and giving the corresponding evaluation value. In this regard, the relative evaluation uses the Single Stimulus Continuous Quality Evaluation (SSCQE) method [128,129]. This involves evaluating a batch of images in a specific sequence. While viewing the images, the observer assigns the corresponding evaluation scores to the images to be evaluated. This is similar to subjective absolute evaluation, in that subjective relative evaluation also has a scoring system, which is referred to as the “group goodness scale” of image evaluation.

5.2 Objective image quality evaluation

The objective image quality evaluation uses a calculation method to determine a numerical value which quantitatively analyzes the image quality according to its content, structure, and naturalness. It can be divided into three types [130]: full-reference (FR), reduced-reference (RR) and no-reference (NR).

5.2.1 Full reference

The concept of full reference image quality evaluation can be defined as a compression of the images to be evaluated using reference images and analyzing the image distortion to achieve a high-quality evaluation [131]. There are three key aspects to the overall objective evaluation regarding the quality of a reference image: pixel statistics, information theory, and structural information.

Based on image pixel statistics

Pixels of images based on a statistical analysis, known as the Mean Square Error and Peak-Signal to Noise Ratio, are related to the evaluation of quality method. They measure the statistical quality of the evaluated image by calculating the differences between the gray values in the corresponding pixels of the evaluated image in relation to the reference image. This is known as a ***mean square error***, which is a reference-type evaluation index. For evaluation images such as F' and $F \in R_{M\times N}$, M and N are the positions of the pixels used for the constructed and reference images. The definition of MSE is as follows:

$$\mathrm{MSE} = \frac{1}{M\times N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}|(F'(i,j)-F(i,j))^2|. \qquad (7)$$

MSE is the square of the average of the differences in pixel positions i, j for the constructed and reference images such as $F'(i,j)$ and $F(i,j)$. This is known as a ***peak signal-to-noise ratio***, which calculates the global information contained in the image quality reconstruction. PSNR is a way of making calculations based on MSE. Therefore, like MSE, it calculates the difference between pixels. The calculation of PSNR is as follows:

$$\mathrm{PSNR} = 10\times\log\left(\frac{L\times L}{MSE}\right), \qquad (8)$$

where L is the largest pixel value, usually ranging in pixel values from 0-255. The unit of PSNR is a decibel (dB).

Both PSNR and MSE measure the image quality by calculating the size of the pixel error between the image to be evaluated and the reference image. Typically, a higher PSNR value will result in greater image quality and less distortion in the assessed and reference images. These two methods are relatively simple, easy to implement, and widely used in image denoising and the like. However, this type of algorithm is based on the global statistics of image pixel values and does not consider the local visual factors of the human eye, so it is impossible to grasp the local quality of the image

Based on information theory

This forms the basis of information entropy in information theory, where mutual information is widely used to evaluate image quality. In recent years, Sheikh and Bovik have proposed two algorithms: Information Fidelity Criterion (IFC) [132] and Visual Information Fidelity (VIF) [133]. The mutual information is computed to reference and evaluate images in order to determine its overall image quality. These two methods have a certain theoretical support, which expands the relationship between the image and the human eye in terms of information fidelity, but these methods do not respond to the structural information of the image.

Based on structural information

The concept of structural information is based on years of research work related to image processing, image compression, and image visual quality evaluation [134]. Researchers believe that the primary role of human vision is to extract background structure information, and that the human vision system is capable of doing so in a highly adaptable manner. Therefore, the measurement of the structural distortion of the image could be the closest approach regarding the perceived quality of the image. Based on this, an objective criterion for image quality was proposed for the first time that conforms to the characteristics of the human visual system and is subsequently referred to as given-Structure Similarity (SSIM) [130]. ***Structural similarity*** (SSIM) measures the structural similarity of images. Unlike MSE and PSNR, SSIM uses the mean and variance to measure the similarity of the test image and reference image structure under global statistical characteristics. The calculation method of SSIM is as follows:

$$\mathrm{SSIM} = \left(\frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}\right). \qquad (9)$$

Among them, μ_x and μ_y represent the mean of x,y, respectively. σ_x and σ_y denote the variance of the image x, y, respectively. $C_1$and C_2 are two smaller constants, preventing the denominator from being zero. SSIM constructs the structural similarity between the reference image and the evaluated image according to the correlating image pixels. The larger the SSIM value, the better the image quality. The index algorithm is simple to implement, and the evaluation parameters are relatively reliable. In this field of study, numerous researchers have made improvements in combination with the human visual system. These calculation methods are currently widely used in all aspects of image processing [129,130].

5.2.2 Partial reference

Partial reference is also known as semi-reference. It uses part of the characteristic information in the ideal image as a reference to compare and analyze the image to be evaluated by obtaining the image quality evaluation result. The referenced information is a feature derived from an image which must first extract a portion of the feature information from both the picture to be evaluated and the ideal image, then compared to the extracted portion of information which determines the quality of the image [135,136]. Some reference methods can be divided into methods based on original image features, digital watermarking methods, and statistical models in the Wavelet Domain. Because part of the reference quality evaluation depends on some features of the image, compared with the overall image, the amount of data has dropped significantly. At present, applications are more concentrated in image transmission systems.

5.2.3 No reference

The no-reference method is also known as the first evaluation method. Because general ideal images are difficult to obtain, this quality evaluation method which deviates from the dependence on ideal reference images is widely used. No-reference methods are generally based on statistical image characteristics [137,138]. The ***perception index*** is an evaluation method without reference images. Researchers found [139] that simply increasing the values of structural similarity and peak signal-to-noise ratio does not always improve an image’s perceived quality. With deep learning-based super-resolution tasks, training with MSE as a loss usually yields the values of structural similarity and peak signal-to-noise ratio that are higher. Still, the edges of the generated images often appear too smooth, resulting in a significant loss of high-frequency information. PI is a method of calculating the perceived quality of an image by using the statistical characteristics of the image itself. PI combines two measurement methods without reference images: NIQE [140] and MA [141]. PI [142] was used as one of the evaluation indicators for the 2018 PIRM Challenge on Perceptual Image Super-resolution. Perception Index and human subjective evaluation correlate to ensure the content of the generated image matches that of the original image. Similarly, PI often needs to be combined with MSE to evaluate image reconstruction quality.

6 Challenges and future directions

The super-resolution of a single image can be based on an input image with a low-resolution, and the reconstructed image contains real details. High-resolution images save textures and keep image content accurate to enhance the graph at the algorithm-level quality of images and research image data storage and transmission, improve social security, and other fields of computer vision. SR is of great significance with a wide range of applications. The difficulties and challenges of super-resolution are to increase the resolution of a single image. Several obstacles can be represented in the practical super-resolution methods, which may be related to the task of fine-tuning the required parameters, blind super-resolution and evaluation metrics, and the like.

Task of fine-tuning the required parameters The performance of super-resolution algorithms can be influenced by parameters. It has been observed as a computationally efficient way to determine parameter values for a non-robust model of Super-Resolution not. For example, GAN-based SISR models require a wide range of training parameters for enhancing SR performance. These characteristics are not applicable in several fields. Lightweight as network design has been widely used to reduce computational redundancy, and CARN [49], CBPN [50], MSRN [51], and MCSN [53] are examples of Lightweight Networks.

Blind super-resolution (BSR) The purpose of BSR is to rebuild low-resolution images that have been degraded in unknown and complex ways. The unknown parameters of the imaging system may be estimated from the measured data using a blur estimation algorithm. Two stages have been recently used to solve this problem: A- Estimating blur kernel from observed LR image and a B- Restoring SR image based on an estimated kernel. Both depend on two independently trained models. KernelGAN [81], KernelGAN-FKP [83], SRSVD [89], and IKC [85] represent the recent methods of Blind super-resolution.

Evaluation metrics Evaluation metrics represent one of the key components for image quality measurements. PSNR, SSIM, and PI have been used as quality metrics. PSNR can be used for the evaluation of non-realistic smooth surfaces, while SSIM represents the evaluation of structures, brightness, and contrast to imitate human perception. PI needs to be simultaneously combined with MSE to evaluate image reconstruction quality. Together, these metrics are not fully elicited by the SR quality. Further metrics are required to assess SISR quality.

The computation efficiency The computational complexity of optimization procedures in SR approaches has attracted a lot of interest in learning a mapping from LR image to HR image. A deep network typically increases SR performance, but it also has a high computational cost and limits the applications available on mobile devices. Several efforts have been undertaken to reduce the model size by distilling information and reusing features effectively to address this issue. Many lightweight SR networks effectively minimize the model size and redundant computation. Redundant computation is still present and prevents them from achieving higher levels of computational efficiency. In the future, researchers are anticipated to put more effort on reducing the size of neural networks to improve the SISR method.

There are also other challenges related to flexibility. However, super-resolution researchers are still working to develop robust methods in terms of automatic parameter selection,and memory usage. Future work may cover the methods and techniques that are applied in video super-resolution.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (Grant No. 62072328).

References

1. Freedman G, Fattal R. Image and video upscaling from local self-examples. ACM Transactions on Graphics, 2011, 30(2): 12
2. Glasner D, Bagon S, Irani M. Super-resolution from a single image. In: Proceedings of the 12th IEEE International Conference on Computer Vision. 2009, 349−356
3. Yang J, Lin Z, Cohen S. Fast image super-resolution based on in-place example regression. In: Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2013, 1059−1066
4. Bevilacqua M, Roumy A, Guillemot C, Morel M L A. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In: Proceedings of British Machine Vision Conference (BMVC). 2012, 135.1−135.10
5. Chang H, Yeung D-Y, Xiong Y. Super-resolution through neighbor embedding. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.; vol. 1. 2004, I–I
6. Freeman W T, Pasztor E C, Carmichael O T. Learning low-level vision. International Journal of Computer Vision, 2000, 40(1): 25–47
7. Jia K, Wang X, Tang X. Image transformation based on learning dictionaries across image spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(2): 367–380
8. Timofte R, De Smet V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution. In: Proceedings of 2013 IEEE International Conference on Computer Vision. 2013, 1920−1927
9. Yang J, Wang Z, Lin Z, Cohen S, Huang T. Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing, 2012, 21(8): 3467–3478
10. Yang J, Wright J, Huang T, Ma Y. Image super-resolution as sparse representation of raw image patches. In: Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. 2008, 1−8
11. Yang J, Wright J, Huang T S, Ma Y. Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 2010, 19(11): 2861–2873
12. Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. In: Proceedings of the 7th International Conference on Curves and Surfaces. 2010, 711−730
13. Yang W, Zhang X, Tian Y, Wang W, Xue J H, Liao Q. Deep learning for single image super-resolution: a brief review. IEEE Transactions on Multimedia, 2019, 21(12): 3106–3121
14. Zhou L, Feng S. A review of deep learning for single image super-resolution. In: Proceedings of International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). 2019, 139–142
15. Ha V K, Ren J, Xu X, Zhao S, Xie G, Vargas V M. Deep learning based single image super-resolution: A survey. In: Proceedings of the 9th International Conference on Brain Inspired Cognitive Systems. 2018, 106−119
16. Zhang H, Wang P, Zhang C, Jiang Z. A comparable study of CNN-based single image super-resolution for space-based imaging sensors. Sensors, 2019, 19(14): 3234
17. Dong C, Loy C C, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295–307
18. Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016, 1646−1654
19. Dong C, Loy C C, Tang X. Accelerating the super-resolution convolutional neural network. In: Proceedings of the 14th European Conference on Computer Vision. 2016, 391–407
20. Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016, 1637−1645
21. Li X, Wu Y, Zhang W, Wang R, Hou F. Deep learning methods in real-time image super-resolution: a survey. Journal of Real-Time Image Processing, 2020, 17(6): 1885–1909
22. Wang Z, Chen J, Hoi S C H. Deep learning for image super-resolution: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3365–3387
23. Bashir S M A, Wang Y, Khan M, Niu Y. A comprehensive review of deep learning-based single image super-resolution. PeerJ Computer Science, 2021, 7: e621
24. Liu A, Liu Y, Gu J, Qiao Y, Dong C. Blind image super-resolution: A survey and beyond. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022: 1–9
25. Zhu H, Xie C, Fei Y, Tao H. Attention mechanisms in CNN-based single image super-resolution: a brief review and a new perspective. Electronics, 2021, 10(10): 1187
26. Tian J, Ma K K. A survey on super-resolution imaging. Signal, Image and Video Processing, 2011, 5(3): 329–342
27. Shah A J, Gupta S B. Image super resolution-a survey. In: Proceedings of the 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking. 2012, 1−6
28. Ghesu F C, Köhler T, Haase S, Hornegger J. Guided image super-resolution: a new technique for photogeometric super-resolution in hybrid 3-D range imaging. In: Jiang X, Hornegger J, Koch R, eds. Pattern Recognition. Cham: Springer, 2014, 227−238
29. Moitra S. Single-image super-resolution techniques: a review. International Journal for Science and Advance Research in Technology, 2017, 3(4): 271–283
30. Huang D, Liu H. A short survey of image super resolution algorithms. Journal of Computer Science Technology Updates, 2015, 2(2): 19–29
31. Schultz R R, Stevenson R L. A Bayesian approach to image expansion for improved definition. IEEE Transactions on Image Processing, 1994, 3(3): 233–242
32. Wang Y, Wan W, Wang R, Zhou X. An improved interpolation algorithm using nearest neighbor from VTK. In: Proceedings of 2010 International Conference on Audio, Language and Image Processing. 2010, 1062−1065
33. Titus J, Geroge S. A comparison study on different interpolation methods based on satellite images. International Journal of Engineering Research & Technology, 2013, 2(6): 82–85
34. Parsania P, Virparia D. A review: Image interpolation techniques for image scaling. International Journal of Innovative Research in Computer and Communication Engineering, 2014, 2(12): 7409–7414
35. Li X, Orchard M T. New edge-directed interpolation. IEEE Transactions on Image Processing, 2001, 10(10): 1521–1527
36. Gavade A B, Sane P. Super resolution image reconstruction by using bicubic interpolation. In: Proceedings of National Conference on Advanced Technologies in Electrical and Electronic Systems. 2014, 1
37. Irani M, Peleg S. Improving resolution by image registration. CVGIP: Graphical Models and Image Processing, 1991, 53(3): 231–239
38. Nayak R, Patra D. Enhanced iterative back-projection based super-resolution reconstruction of digital images. Arabian Journal for Science and Engineering, 2018, 43(12): 7521–7547
39. Stark H, Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections. Journal of the Optical Society of America A, 1989, 6(11): 1715–1726
40. Fan C, Wu C, Li G, Ma J. Projections onto convex sets super-resolution reconstruction based on point spread function estimation of low-resolution remote sensing images. Sensors, 2017, 17(2): 362
41. Schultz R R, Stevenson R L. Improved definition video frame

enhancement. In: Proceedings of 1995 International Conference on Acoustics, Speech, and Signal Processing (ICASSP). 1995, 2169−2172
42. Schultz R R, Stevenson R L. Extraction of high-resolution frames from video sequences. IEEE Transactions on Image Processing, 1996, 5(6): 996–1011
43. Besag J. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society: Series B: (Methodological), 1986, 48(3): 259–279
44. Homem M R P, Martins A L D, Mascarenhas N D A. Super-resolution image reconstruction using the discontinuity adaptive ICM. 1994. Poster, https://www.dca.fee.unicamp.br/~ting/misc/anais-poster-2007/posters/33629.pdf
45. Timofte R, De Smet V, Van Gool L. A+: Adjusted anchored neighborhood regression for fast super-resolution. In: Proceedings of the 12th Asian Conference on Computer Vision. 2014, 111−126
46. Wen Y, Sheng B, Li P, Lin W, Feng D D. Deep color guided coarse-to-fine convolutional network cascade for depth image super-resolution. IEEE Transactions on Image Processing, 2019, 28(2): 994–1006
47. Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, 2790−2798
48. Hui Z, Gao X, Yang Y, Wang X. Lightweight image super-resolution with information multi-distillation network. In: Proceedings of the 27th ACM International Conference on Multimedia. 2019, 2024−2032
49. Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proceedings of the 15th European Conference on Computer Vision. 2018, 256−272
50. Zhu F, Zhao Q. Efficient single image super-resolution via hybrid residual feature learning with compact back-projection network. In: Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop. 2019, 2453−2460
51. Li J, Fang F, Mei K, Zhang G. Multi-scale residual network for image super-resolution. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). 2018, 517−532
52. Liu J, Tang J, Wu G. Residual feature distillation network for lightweight image super-resolution. In: Proceedings of European Conference on Computer Vision. 2020, 41−55
53. Li W, Li J, Li J, Huang Z, Zhou D. A lightweight multi-scale channel attention network for image super-resolution. Neurocomputing, 2021, 456: 327–337
54. Zhang Y, Li K, Li K, Wang L, Zhong B, Fu Y. Image super-resolution using very deep residual channel attention networks. In: Proceedings of 15th European Conference on Computer Vision (ECCV). 2018, 294−310
55. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 7132−7141
56. Cheng X, Li X, Yang J, Tai Y. SESR: Single image super resolution with recursive squeeze and excitation networks. In: Proceedings of the 24th International Conference on Pattern Recognition (ICPR). 2018, 147−152
57. Roy A G, Navab N, Wachinger C. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks. In: Proceedings of 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. 2018, 421−429
58. Dai T, Cai J, Zhang Y, Xia S T, Zhang L. Second-order attention network for single image super-resolution. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019, 11057−11066
59. Choi J S, Kim M. A deep convolutional neural network with selection units for super-resolution. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017, 1150−1156
60. Anwar S, Barnes N. Densely residual Laplacian super-resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(3): 1192–1204
61. Zhang Y, Li K, Li K, Zhong B, Fu Y. Residual non-local attention networks for image restoration. 2019, arXiv preprint arXiv: 1903.10082
62. Hu Y, Li J, Huang Y, Gao X. Channel-wise and spatial feature modulation network for single image super-resolution. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(11): 3911–3927
63. Kim J H, Choi J H, Cheon M, Lee J S. Ram: Residual attention module for single image super-resolution. 2018, arXiv preprint arXiv: 1811.12043v1
64. Liu J, Zhang W, Tang Y, Tang J, Wu G. Residual feature aggregation network for image super-resolution. In: Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020, 2356−2365
65. Zheng Z, Jiao Y, Fang G. Upsampling attention network for single image super-resolution. In: Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2021, 399−406
66. Lim B, Son S, Kim H, Nah S, Lee K M. Enhanced deep residual networks for single image super-resolution. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017, 1132−1140
67. Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 3262−3271
68. Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y. Residual dense network for image super-resolution. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 2472−2481
69. Tong T, Li G, Liu X, Gao Q. Image super-resolution using dense skip connections. In: Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). 2017, 4809–4817
70. Wang W, Li X, Yang J, Lu T. Mixed link networks. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. 2018, 2819−2825
71. Shi W, Caballero J, Huszár 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. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, 1874−1883
72. Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 1664−1673
73. Dun Y, Da Z, Yang S, Xue Y, Qian X. Kernel-attended residual network for single image super-resolution. Knowledge-Based Systems, 2021, 213: 106663
74. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, 105−114
75. Shocher A, Cohen N, Irani M. Zero-shot super-resolution using deep internal learning. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018, 3118−3126
76. Wang X, Yu K, Wu S, Gu J, Liu Y, Dong C, Qiao Y, Loy CC. ESRGAN: Enhanced super-resolution generative adversarial networks.

In: Proceedings of European Conference on Computer Vision. 2018, 63−79
77. Jolicoeur-Martineau A. The relativistic discriminator: a key element missing from standard GAN. In: Proceedings of the 7th International Conference on Learning Representations. 2019
78. Szegedy C, Ioffe S, Vanhoucke V, Alemi A A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence. 2017, 4278−4284
79. Zhang W, Liu Y, Dong C, Qiao Y. RankSRGAN: Generative adversarial networks with ranker for image super-resolution. In: Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. 2019, 3096−3105
80. Maeda S. Unpaired image super-resolution using pseudo-supervision. In: Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020, 288−297
81. Bell-Kligler S, Shocher A, Irani M. Blind super-resolution kernel estimation using an internal-GAN. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. 2019, 26
82. Gu J, Lu H, Zuo W, Dong C. Blind super-resolution with iterative kernel correction. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019, 1604−1613
83. Michaeli T, Irani M. Nonparametric blind super-resolution. In: Proceedings of IEEE International Conference on Computer Vision. 2013, 945−952
84. Ren D, Zhang K, Wang Q, Hu Q, Zuo W. Neural blind deconvolution using deep priors. In: Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020, 3338−3347
85. Zhang K, Gool L V, Timofte R. Deep unfolding network for image super-resolution. In: Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020, 3214−3223
86. Liang J, Zhang K, Gu S, Gool L V, Timofte R. Flow-based kernel prior with application to blind super-resolution. In: Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021, 10596−10605
87. Luo Z, Huang Y, Li S, Wang L, Tan T. Unfolding the alternating optimization for blind super resolution. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. 2020
88. Wang L, Wang Y, Dong X, Xu Q, Yang J, An W, Guo Y. Unsupervised degradation representation learning for blind super-resolution. In: Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021, 10576−10585
89. Gandelsman Y, Shocher A, Irani M. “Double-DIP”: Unsupervised image decomposition via coupled deep-image-Priors. In: Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019, 11018−11027
90. Cornillère V, Djelouah A, Wang Y. Sorkine-Hornung O, Schroers C. Blind image super-resolution with spatially variant degradations. ACM Transactions on Graphics, 2019, 38(6): 166
91. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, 770–778
92. Gou J, Yu B, Maybank S J, Tao D. Knowledge distillation: a survey. International Journal of Computer Vision, 2021, 129(6): 1789–1819
93. Gao Q, Zhao Y, Li G, Tong T. Image super-resolution using knowledge distillation. In: Proceedings of the 14th Asian Conference on Computer Vision (ACCV). 2018, 527−541
94. He Z, Dai T, Lu J, Jiang Y, Xia S T. Fakd: feature-affinity based knowledge distillation for efficient image super-resolution. In: Proceedings of 2020 IEEE International Conference on Image Processing (ICIP). 2020, 518−522
95. Lee W, Lee J, Kim D, Ham B. Learning with privileged information for efficient image super-resolution. In: Proceedings of the 16th European Conference on Computer Vision. 2020, 465−482
96. Zhang L, Wang P, Shen C, Liu L, Wei W, Zhang Y, Van Den Hengel A. Adaptive importance learning for improving lightweight image super-resolution network. International Journal of Computer Vision, 2020, 128(2): 479–499
97. Hui Z, Wang X, Gao X. Fast and accurate single image super-resolution via information distillation network. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 723−731
98. Zhang Y, Chen H, Chen X, Deng Y, Xu C, Wang Y. Data-free knowledge distillation for image super-resolution. In: Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021, 7848–7857
99. Jiang K, Wang Z, Yi P, Jiang J, Xiao J, Yao Y. Deep distillation recursive network for remote sensing imagery super-resolution. Remote Sensing, 2018, 10(11): 1700
100. Lu T, Yang W, Wan Y. Super-resolution for surveillance video via adaptive block-matching registration. In: Proceedings of 2015 International Conference on Artificial Intelligence and Industrial Engineering. 2015, 1−3
101. Shao J, Chao F, Luo M, Lin J. A super-resolution reconstruction algorithm for surveillance video. Journal of Forensic Science and Medicine, 2017, 3(1): 26
102. Ghazali N N A N, Zamani N A, Abdullah S N H S, Jameson J. Super resolution combination methods for CCTV forensic interpretation. In: Proceedings of the 12th International Conference on Intelligent Systems Design and Applications (ISDA). 2012, 853−858
103. Yu H, Liu D, Shi H, Yu H, Wang Z, Wang X, Cross B, Bramler M, Huang T S. Computed tomography super-resolution using convolutional neural networks. In: Proceedings of 2017 IEEE International Conference on Image Processing (ICIP). 2017, 3944−3948
104. Greenspan H, Oz G, Kiryati N, Peled S. MRI inter-slice reconstruction using super-resolution. Magnetic Resonance Imaging, 2002, 20(5): 437–446
105. Yu J, Lavery L, Kim K. Super-resolution ultrasound imaging method for microvasculature *in vivo* with a high temporal accuracy. Scientific Reports, 2018, 8(1): 13918
106. Isaac J S, Kulkarni R. Super resolution techniques for medical image processing. In: Proceedings of 2015 International Conference on Technologies for Sustainable Development (ICTSD). 2015, 1−6
107. Park S C, Park M K, Kang M G. Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine, 2003, 20(3): 21–36
108. Kouame D, Ploquin M. Super-resolution in medical imaging: An illustrative approach through ultrasound. In: Proceedings of 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 2009, 249−252
109. He S, Tian Y, Feng S, Wu Y, Shen X, Chen K, He Y, Sun Q, Li X, Xu J, Wen Z L, Qu J Y. In vivo single-cell lineage tracing in zebrafish using high-resolution infrared laser-mediated gene induction microscopy. eLife, 2020, 9: e52024
110. Jones M G, Khodaverdian A, Quinn J J, Chan M M, Hussmann J A, Wang R, Xu C, Weissman J S, Yosef N. Inference of single-cell phylogenies from lineage tracing data using Cassiopeia. Genome Biology, 2020, 21: 92
111. Andresen V, Pollok K, Rinnenthal J L, Oehme L, Günther R, Spiecker H, Radbruch H, Gerhard J, Sporbert A, Cseresnyes Z, Hauser A E,

Niesner R. High-resolution intravital microscopy. PLoS One, 2012, 7(12): e50915
112. Small A, Stahlheber S. Fluorophore localization algorithms for super-resolution microscopy. Nature Methods, 2014, 11(3): 267–279
113. Orrit M. Celebrating optical nanoscopy. Nature Photonics, 2014, 8(12): 887
114. Singh A, Sidhu J S. Super resolution applications in modern digital image processing. International Journal of Computer Applications, 2016, 150(2): 6–8
115. UI Hoque M R, Burks R, Kwan C, Li J. Deep learning for remote sensing image super-resolution. In: Proceedings of the 10th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). 2019, 286–292
116. Zhu H, Tang X, Xie J, Song W, Mo F, Gao X. Spatio-temporal super-resolution reconstruction of remote-sensing images based on adaptive multi-scale detail enhancement. Sensors, 2018, 18(2): 498
117. Guo J, Gong X, Wang W, Que X, Liu J. SASRT: semantic-aware super-resolution transmission for adaptive video streaming over wireless multimedia sensor networks. Sensors, 2019, 19(14): 3121
118. Mithra K, Vishvaksenan K S. Security and resolution enhanced transmission of medical image through IDMA aided coded STTD system. In: Proceedings of 2017 International Conference on Communication and Signal Processing (ICCSP). 2017, 2061−2065
119. Hayat K. Multimedia super-resolution via deep learning: a survey. Digital Signal Processing, 2018, 81: 198–217
120. Bishop C M, Blake A, Marthi B. Super-resolution enhancement of video. In: Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics. 2003, 25–32
121. Del Gallego N P, Ilao J. Multiple-image super-resolution on mobile devices: an image warping approach. EURASIP Journal on Image and Video Processing, 2017, 2017(1): 8
122. Hassaballah M, Hosny K M. Recent Advances in Computer Vision: Theories and Applications. Cham: Springer, 2018
123. Liu D, Soran B, Petrie G, Shapiro L. A review of computer vision segmentation algorithms. 2012Lecture notes 53, 2012.
124. Liu F, Wang J, Zhu S, Gleicher M, Gong Y. Visual-quality optimizing super resolution. Computer Graphics Forum, 2009, 28(1): 127–140
125. Hirst D, Rilliard A, Aubergé V. Comparison of subjective evaluation and an objective evaluation metric for prosody in text-to-speech synthesis. In: Proceedings of the 3rd ESCA/COCOSDA Workshop on Speech Synthesis. 1998, 1−4
126. A. R. Reibman and T. Schaper, “Subjective performance evaluation for super-resolution image enhancement,” in Second Int. Wkshp on Video Proc. and Qual. Metrics (VPQM’06), 2006 https://ece.uwaterloo.ca/~z70wang/publications/ICIP12b.pdf
127. Opozda S, Sochan A. The survey of subjective and objective method for quality assessment of 2D and 3D images. Theoretical and Applied Informatics, 2014, 26(1−2): 39−67
128. Th. Alpert and J.-P. Evain, “Subjective quality evaluation - the SSCQE and DSCQE methodologies,” EBU technical review, Spring, See tech. ebu. ch/docs/techreview/trev_271-evain. pdf website, 1997, 12−20
129. Int. Telecommun. Union, Methodology for the Subjective Assessment of the Quality of Television Pictures ITU-R Recommendation BT.500-9, Tech. Rep., See itu. int/rec/R-REC-BT. 500-11-200206-S/en website, 2002
130. Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600–612
131. Yan B, Bare B, Ma C, Li K, Tan W. Deep objective quality assessment driven single image super-resolution. IEEE Transactions on Multimedia, 2019, 21(11): 2957–2971
132. Sheikh H R, Bovik A C, De Veciana G. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Transactions on Image Processing, 2005, 14(12): 2117–2128
133. Sheikh H R, Bovik A C. A visual information fidelity approach to video quality assessment. In: The first international workshop on video processing and quality metrics for consumer electronics; vol. 7. 2005, 2117–2128
134. Yeganeh H, Rostami M, Wang Z. Objective quality assessment for image super-resolution: a natural scene statistics approach. In: Proceedings of the 19th IEEE International Conference on Image Processing. 2012, 1481−1484
135. Wang X, Jiang G, Yu M. Reduced reference image quality assessment based on Contourlet domain and natural image statistics. In: Proceedings of the 5th International Conference on Image and Graphics. 2009, 45−50
136. Al Madeed N, Awan Z, Al Madeed S. Image quality assessment-a survey of recent approaches. In: Proceedings of the 8th International Conference on Computer Science, Engineering and Applications. 2018, 143−156
137. Keshk H M, Abdel-Aziem M M, Ali A S, Assal M A. Performance evaluation of quality measurement for super-resolution satellite images. In: Proceedings of 2014 Science and Information Conference. 2014, 364−371
138. Zhu X, Cheng Y, Peng J, Wang R, Le M, Liu X. Super-resolved image perceptual quality improvement via multifeature discriminators. Journal of Electronic Imaging, 2020, 29(1): 013017
139. Blau Y, Michaeli T. The perception-distortion tradeoff. In: Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, 6228−6237
140. Mittal A, Soundararajan R, Bovik A C. Making a "completely blind" image quality analyzer. IEEE Signal Processing Letters, 2012, 20(3): 209–212
141. Ma C, Yang C Y, Yang X, Yang M H. Learning a no-reference quality metric for single-image super-resolution. Computer Vision and Image Understanding, 2017, 158: 1–16
142. Blau Y, Mechrez R, Timofte R, Michaeli T, Zelnik-Manor L. The 2018 PIRM challenge on perceptual image super-resolution. In: Proceedings of European Conference on Computer Vision. 2018, 334−355

Hanadi AL-Mekhlafi received the Master degree from School of Software Engineering, South China University of Technology, China. She is currently studying for PhD degree in the College of Intelligence and Computing, Tianjin University, China. Her research interests include image/video processing, computer graphics, deep learning.

Shiguang Liu (Senior Member, IEEE) received the PhD degree from the State Key Laboratory of CAD and CG, Zhejiang University, China. He is currently a Professor with the College of Intelligence and Computing, Tianjin University, China. His research interests include image/video processing, computer graphics, visualization, and virtual reality.