IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 36, No. 6, December, 2022, pp. 939-952
Journal homepage: http://iieta.org/journals/ria

Study of Deep Learning-based models for Single Image Super-Resolution

Omar Soufi*, Fatima-Zahra Belouadha

Mohammadia School of Engineers, Mohammed V University in Rabat, AMIPS research team, Rabat 10090, Morocco

Corresponding Author Email: omar_soufi2@um5.ac.ma

https://doi.org/10.18280/ria.360616

Received: 5 November 2022
Accepted: 18 December 2022

Keywords:
CNN, deep learning, GAN, images, neural networks, SISR, super-resolution, systematic review

ABSTRACT

The super-resolution of images has seen remarkable progress, especially with the use of deep learning models. This technique allows having a better-quality image from one or more low-resolution versions. Super-resolution, therefore, aims at enriching a low-resolution image with additional pixel density and high-frequency detail. This paper presents a comprehensive empirical study based on a systematic review of deep learning-based models for single image super-resolution (SISR), exploring the set of techniques offered by deep learning technology and used for SISR. In this paper, we present a global and complete state of the art on deep learning model based on reference metrics (mainly Peak Signal to Noise Ratio -PSNR- and Structural SIMilarity -SSIM-) in the field of computer visualization and image reconstruction. This study was done on several deep learning designs with 90 different models tested on 7 reference datasets in the computer vision domain. Thus, our goal is to present a benchmark to demonstrate the performance and limitations of these models as well as to guide future research in the field of super image resolution to develop efficient algorithms. Moreover, our study covers different neural network architectures (Generative Adversarial Networks -GAN-, Convolutional Neural Networks -CNN-, and Recurrent Neural Networks -RNN-...), using different techniques and technologies.

1. INTRODUCTION

The rapid development of technology makes the image central to the evolution of computing so computer vision is the branch of artificial intelligence (AI) that deals with how computers can gain high-level understanding from images, it seeks to understand and automate the tasks that the human's visual system can perform [1]. Thus, computer vision is important for event management, object tracking (autonomous vehicle for example) [2], image reconstruction, and others.

Super-Resolution (SR) is increasingly in demand due to the need for higher-quality images for sensitive domains such as spatial remote sensing [3], medical imaging [4], environmental monitoring [5], and other domains requiring high-resolution images. Moreover, this technique improves other deep learning processes for vision using images such as classification [6] and segmentation [7].

To have images of better resolution the intuitive idea is to replace the image sensor, but this solution is expensive and sometimes impossible as for example the case of satellite images where the acquisition sensor is airborne. Therefore, the solution is to use SR methods.

This article is an extension of our article [8] benchmarking some models for SR, which was limited to three types of architectures, and only according to the scaling factor.

This paper focuses on SR through the presentation of an empirical study of the most recent algorithms that have demonstrated better performances than the standard method, which is the bicubic interpolation through experimental results and a synthetic summary of metrics and criteria. Indeed, relevant comparisons are made from different angles.

The principle of super-resolution consists in generating a high resolution (HR) or very high resolution (VHR) image from a low resolution (LR) image by passing through a method, thus finding details not existing on the original image.

The development of neural network architecture and associated techniques has allowed us to have today a multitude of SR algorithms that have shown better performance. These algorithms differ according to the type of dataset used for training, the architecture used, the loss function, the depth of the network, the performance metrics, and the scaling factor. Faced with the array of these algorithms, the user finds it impossible to choose an efficient algorithm for a given image in each context. Thus, this study aims to address the following issues:

- Guide the user in choosing an efficient algorithm for SR based on these constraints.
- Why one algorithm is better suited for SR with better performance.
- Present a reliable and credible comparison of deep learning models for SR.

This study aims to be a reference in design and model selection for SR. This study was done through an understanding of the SR problem, combined with an understanding of image descriptors and a thorough analysis of learning models for SR. The objective is to guide the selection of existing models for super-resolution or to advance research in the field of SR in order to identify features that can improve the techniques and architectures in the implementation and design of deep learning models to do image super-resolution in an efficient way.

Our approach aims at guiding the correct choice of model

for SR, which uses the "No Free Lunch" theorem [9] that states that no Machine Learning model is efficient for any problem.

The main contributions of our research work are summarized in the following:

- Proposal of a framework for the study of SISR.
- Presentation of a taxonomy of existing models, main datasets, performance metrics, and methods.
- Proposal of a complete study of the main existing SISR learning-based methods according to their architectures.
- Global comparison of the evaluated models for each scaling factor.
- Assistance to users in choosing a suitable model.
- Presentation of some challenges and researcher's orientations for future work.

This document is structured as follows. The preliminary and related works of SISR are briefly introduced in Section II. In Section III, the Methodology that is, evaluation and empirical study for SISR learning based are described. Section IV presented the result of the study by the scale factor. The discussions and orientations are presented in Section V. Finally, Section VI concludes this work and analyzes future research directions of SISR.

2. PRELIMINARY AND RELATED WORKS

A digital image is a 3D matrix composed of pixels (which means PICture Element). To visualize an image, we assign to each band a colored filter by additive synthesis of the 3 primary colors Red, Green, and Blue so we speak of RGB image.

An image is characterized by its definition which represents the total number of pixels (Height X Width), its resolution which represents the number of points per inch, its depth which represents the number of bands, the coding of the colors which represents the number of bits by pixel (the coding RGB 24 bits is the most used), its format which represents the way in which the image is stored according to the algorithm of compression used.

SISR is an exciting research topic [10, 11] given its practical importance for image texture enhancement. This technique consists in applying a degradation function on an HR image in order to obtain an LR image. Thus, these algorithms can be classified into specific models (graphic illustration [12]), face [13], scenes [14], etc.), and other generic models for the processing of complex images.

$$LR= D(HR,f) \quad (1)$$

D is the degradation function and f is the scaling factor which is the parameter of this function. So, the SR consists in learning the function of Eq. (1) inverse in order to reconstruct the HR image from an LR version by applying the function.

$$HR=SR(LR,f) \quad (2)$$

SR involves both images and videos [15, 16] such that image SR is mainly divided into SISR (single image SR) and MISR (multiple image SR) [17]. SISR provides a higher-quality image based on a single input image, while MISR provides a high-resolution image from a set of merged images of the same image [18].

SISR is used more than MISR because of the performance, simplicity of processing, and support from researchers as well as the flexibility of use [19].

To date, traditional SISR algorithms are mainly divided into three categories: interpolation-based methods that use pixel adjustment based on spatial structure (Neighborhood) [20], reconstruction-based methods that sample scenes from an image sequence [21], and learning-based methods [22] as shown in Figure 1 The classification of reconstruction-based and interpolation-based algorithms has been done in other studies [23, 24], these methods have average performance given several limitations deep learning has solved [11]. We note that among the classical methods bicubic interpolation is the most used, which makes it still used as standard input for several deep learning models.

The field of super-resolution of images does not cease developing because of its practical applications in many fields. This paper focuses on deep learning methods as these algorithms are gaining in performance and attention from researchers today due to their better reconstruction capabilities [19] and most SISR methods are based on deep learning [25]; without forgetting the reference method which is bicubic interpolation [26] and which is part of interpolation based.

Deep learning models for super-resolution have been classified according to several; so, for our study, we adopt the classification by type of learning model. Thus, Figure 2 shows an example of super-resolution by different models

Figure 1. Classification of image super-resolution method

Figure 2. Super Resolution of an image by several models [27]

A historical problem for SISR is that it is an ill-posed inverse problem (i.e., one must know the degradation function (1) in order to apply the inverse function (2) to have a high-resolution image), this assumption is not always achievable, it is also noted that the quality of recovery of the target image is influenced by the learning of features within the training images taken as samples [28].

The development of neural networks and associated technologies allows us today to have powerful algorithms for deep learning allowing SISR in an efficient way with better performances thanks to the technological development offering several architectures like CNN [29], GAN [30, 31], and others (Table 1); as well as many adapted loss functions [32].

Several works have performed benchmarks for SR, but they address this problem from different angles, which remains useless for users to guide their choice of model for SR. For example, Chen et al. focused on real-world single image super-resolution (RSISR) to address the problem of SR degradation on synthetic data [19]. Other works are not recent and address methods that are outdated today as the case for [23, 33]. Zhang

et al. [34] worked on Ultra High Definition (UHD), introducing two datasets UHDSR4K and UHDSR8K which remain limited to this type of image. Liu et al. [35] have worked on recent networks but this work focuses on the optimization capacity of the architectures by addressing only 12 networks.

Table 1. Classification of SISR models by architecture

CNN	CNN	Residual Networks	RNN	GAN	Attention Networks	Random Forest
CAR [36]	MFSRCNN [54]	BTSRN [70]	BSRN [83]	4PP-EUSR [93]	ABPN [102]	FAFR [100]
CSNLN [37]		CARN [71]	DBPN-RES-MR64-3[84]	Edge-informed SR [94]	DRLN [72]	JMPF+ [93]
CSRCNN [38]	MDSR [55]	CARN-M [71]	D-DBPN [84]	HiFaceGAN [13]	HAN+ [103]	
CMSC [39]	MWCNN [56]	DRLN+ [72]	DRCN [72]	Nearest neighbors [77]	SelNet [104]	
DnCNN [40]	N3Net [57]	EDSR [73]	DRRN [85]	ESRGAN [82]	SRRAM [105]	
FALSR-A [41]	DnCNN-3[40]	IKC [74]	HBPN [86]	ProSR [95]	PASSRnet [106]	
FSRCNN [42]	Perceptual Loss [58]	REDNet [75]	DSRN [87]	RFN [96]	SAN [107]	
CNF [43]	ESPCN [59]	LCSCNet [76]	LapSRN [77]	SFT-GAN [97]	STSR [108]	
CSCN [44]	RC-Net [60]	LapSRN [77]	MemNet [88]	SPSR [98]	SwinIR [27]	
Deep CNN Denoiser [45]	ScSR [61]	PMRN+ [78]	GMFN [89]	S-RFN [99]		
IMDN [46]	SRCNN [29]	RCAN [79]	SCN [90]	SRGAN [30]		
ENet-E [47]	SRMDNF [62]	RDN [80]	SRFBN [91]	Super-FAN [100]		
IA [48]	SRNTT-l2[63]	RL-CSC [81]	NLRN [92]	BSRGAN [101]		
IDN [49]	ZSSR [64]	SRResNet [30]	SESR [92]			
RED30[50]	WaveletCNN [65]	ESRGAN [82]				
SPBP-L+ [51]	PFF [66]					
LFFN-S [52]	CRAN [67]					
VDSR [53]	AdderNets [68]					
	SRWarp [69]					

Table 2. Classification of SISR models by scale factor

X2	X2	X3	X3	X4	X4	X4	X8/X16
BTSRN	SRCNN	Bicubic	LFFN-S	4PP-EUSR	FSRCNN	REDNet	Bicubic
CARN	SRMDNF	BTSRN	MDSR	ABPN	GMFN	RFN	SRCNN
CMSC	SRRAM	CARN	MemNet	Bicubic	HAN+	RL-CSC	FSRCNN
CNF	VDSR	CMSC	MWCNN	BSRN	HBPN	SAN	MFSRCNN
CSCN	ZSSR	CNF	PMRN+	BTSRN	HiFaceGAN	SCN	SCN
D-DBPN	CAR	Deep CNN Denoiser	RCAN	BTSRN	IA	SelNet	VDSR
DnCNN	DRLN+	DnCNN	RDN	CAR	IDN	SESR	LapSRN
DRCN	HAN+	DRCN	RED30	CARN	IKC	SFT-GAN	MemNet
DRLN	CSNLN	DRLN	REDNet	CMSC	IMDN	SPSR	MSLapSRN
DRRN	PMRN+	DRLN+	SCN	CNF	JMPF+	SRC.AX	EDSR
EDSR	HBPN	DRRN	SelNet	CSNLN	LapSRN	SRCNN	D-DBPN
FSRCNN	SRFBN	EDSR	SRCNN	DBPN-MR64	LFFN-S	SRFBN	RCAN
IA	SPBP-L+	FSRCNN	SRFBN	D-DBPN	LopSRN	S-RFN	DRLN
IDN	IMDN	HAN+	SRMDNF	DnCNN	Manifold Simplification	SRGAN	DBPN-MR64-
LapSRN	MWCNN	IDN	SRRAM	DnCNN-3	MDSR	SRGAN + Residual	DRLN+
MDSR	FALSR-A	IKC	STSR	DRCN	MemNet	SRMDNF	HAN+
MemNet	CARN	IMDN	VDSR	DRLN	MWCNN	SRNTT-l2	ABPN
RCAN	RED30	LapSRN	ZSSR	DRLN+	nearest neighbors	SRRAM	HBPN
RDN	LFFN-S	LCSCNet		DRRN	NLRN	SRResNet	CSRCNN
REDNet	DRCN			DSRN	PASSRnet	Super-FAN	DeepRED
SCN	DnCNN-3			Edge-informed SR	Perceptual Loss	SwinIR	ABPN
ScSR	N3Net			EDSR	ProSR	VDSR	
SelfExSR	VDSR			ENet-E	RCAN	Wavelet CNN	
SelNet	CSRCNN			ESPCN	RC-Net	ZSSR	
CRAN	RC-Net			ESRGAN	RDN	PFF	
AdderNets	IKC			FAFR	RED30		
SRWarp	Deep CNN Denoiser						

3. METHODOLOGY (EVALUATION AND EMPIRICAL STUDY)

Our methodology in this work is based on a systemic review [109] of deep learning models for SISR through a thorough scientific review in order to influence decisions in the selection and design of neural networks for SR. This approach respects the principles of knowledge synthesis characterizing systemic review as stated in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [110], in order to provide the essentials of learning-based SISR, also exploiting the results of several studies on state-of-the-art super-resolution neural networks with varied datasets.

This study provides quantitative and qualitative comparisons between the set of models studied in terms of visual and perceptual quality, this comparison is complemented by other criteria such as design details and model parameterization mainly for the first 3 networks for each scaling factor.

In this study, we proceed to in-depth systematic analysis in order to study the technical and architectural constraints of the selected models in terms of deep learning techniques combined with the specific domain of SISR.

Our empirical study in this paper is based on 3 pillars, namely: the choice of deep learning models for SISR, the choice of datasets to train and test these models, and finally the choice of performance measurement metrics. This research synthesis consists in evaluating SISR algorithms from the same framework (metrics, hypotheses, and datasets), in order to have a fair and unbiased analysis; in addition to using diverse data.

3.1 Choice of models

Through this study we made an inventory of more than 200 neural networks for SR, the study led to the choice of 90 most recent networks with a date later than 2016 with 61% of the networks having a date later than 2019. We found that these models can be divided into 6 categories representative of all the deep learning models used today for SR that we classify in Table 1, based on network architecture for SR. These same methods are also distinguished by the network parameterization, depth, and learning optimization conditions. We group the selected networks by the scale factor allowed by this model (Table 2) and then apply our analysis approach to generate the considered performance measures.

3.2 Datasets

In addition to the intrinsic features of the design of the neural network architecture for SRs, the dataset used for its training is determined since the network learns based on the models of the latter; thus, the choice of datasets for this study was made following a thorough analysis. Indeed, these models are trained and tested differently on 43 opensource and famous data science datasets using images (Set5, Set14, BSD100, Urban100, SunHay80, Urban300, Urban500, VID4, 91-Image, IV2K, Waterloo, MCL-V, GOPRO, CelebA, FFHQ 256*256, Sintel, FlyingChairs, DND, RENOIR, NC, FFHQ 1024*1024, SIDD(M), RSR, Vimeo-90k, BSD200, Manga109, VggFace2, PIRM, Celeb-HQ, CUFED5, MiddleBurry, KITTI 2012, KITTI 2015, DIV8K, USR-248,MiddleButy, Sun80, CUFED5, FFHQ 512*512, DIV8K, DIV2K, FFHG, WebFace). These datasets differ primarily in two factors, namely genericity and complexity. Genericity characterizes the use of the dataset, i.e., its content, which can be specific to a domain of use (such as datasets for faces [13], scenes [14], graphic illustrations [11]) or generic and can be used without specifying the domain of use. The complexity determines whether an image is primitive or complex, a primitive image is an image with textures (group of pixels) representing a simple basic pattern (e.g. simple edges and corners) [111], and a complex one containing more details and complicated textures with high-frequency details depending on the objects in the image which is also reflected in the image definition (the total number of pixels). The compression format is also important, so it is in our interest to use a lossless compression format in order to have enough information for the learning process, as well as to use only 06 image datasets designed to evaluate the performance of SR algorithms after a complete study of the constitution and characteristics of each dataset; to evaluate the performance of the recently proposed SISR-based deep learning models.

Table 3. Characteristics of the images used

Dataset	Size	Avg. Resol.	Avg. Pixels	Format	Encoding
Set5 [112]	5	313×336	113,491	PNG	RVB 24
Set14 [113]	14	512×512	230,203	PNG	RVB 24
BSD100 [114]	100	432×370	154,401	PNG	RVB 24
Urban100 [115]	100	984×797	774,314	PNG	RVB 24
Manga109 [116]	109	826×1169	966,011	PNG	RVB 24
DIV2K [117]	1000	1972×1437	2,793,250	PNG	RVB 24

This table (Table 3) represents the characteristics of the datasets used in terms of the number of images of the dataset at the origin (the number of images determines the number of examples seen by the network at the time of training so it determines the capacity of the network for the SR) without considering the images that are generated from these datasets to have LR images or to the data-augmentation which allows improving the performance of the models without additional calculation [118]; the choice of these datasets considered a gradual factor in order to prove this effect on the SR. The number of pixels represents the overall size of the pixels in the dataset; thus, we tried to choose graduated definitions to see the effect of the definition on the reconstruction. We chose the same PNG format [119] which guarantees better lossless compression. We also note that one dataset contains natural scenes (Set5, Set14, and BSDS100) while another contains urban scenes with frequency details at all levels (Urban100). Thus, the choice of these datasets for our empirical study was made on the basis of the key characteristics of the images that compose them.

3.3 Performance measurement metrics

To measure the performance of deep learning models for SR we use 2 types of metrics which are quantitative and qualitative metrics. Qualitative metrics can use the perception

of human subjects, so we only use quantitative metrics in this paper due to their demonstrability and reliability.

In this study, we use PSNR and SSIM metrics as evaluation criteria for the selected algorithms, these metrics are considered as a reference in the field of visualization model evaluation.

The PSNR (Peak Signal to Noise Ratio) [120] measures the reconstruction quality of a digital image. It is used to evaluate the reconstruction quality of the super-resolved image.

The PSNR is inversely proportional to the logarithm of the mean square error (MSE) (3) [121] between the low-resolution image and the super-resolved image.

$$\mathrm{MSE}(x,y) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(x_{ij} - y_{ij}\right)^2 \tag{3}$$

$$\mathrm{PSNR}(x,y) = 10\cdot\log_{10}\left(\frac{L^2}{MSE(x,y)}\right) \tag{4}$$

MN (3) is the image definition (size in pixels) and L (4) is the maximum possible value for a pixel (for 8-bit RGB images, it is 255).

The PSNR is used only to evaluate the correspondence between the super-resolved image and the original image, the latter does not take into account the visual quality of the reconstruction of the image, so we can have a model that generates images with missing high-frequency details, which forces us to combine this metric with the SSIM.

SSIM (structural SIMilarity) [120] measures the similarity between two digital images. This metric originally used to measure the visual quality of compressed images has been extended to measure the visual quality of super-resolved images compared to the low-resolution image. SSIM measures the structural similarity between the two images since the human eye is sensitive to structural changes within the image.

SSIM is computed on a set of windows of an image. The metric between two windows x and y of size NxN is as follows:

$$\mathrm{SSIM}(x,y) = A(x,y)\cdot B(x,y)\cdot C(x,y) \tag{5}$$

$$A(x,y) = \frac{2\mu_x\mu_y + E_1}{\mu_x^2 + \mu_y^2 + E_1} \tag{6}$$

$$B(x,y) = \frac{2\sigma_x\sigma_y + E_2}{\sigma_x^2 + \sigma_y^2 + E_2} \tag{7}$$

$$C(x,y) = \frac{\sigma_{xy} + E_3}{\sigma_x\sigma_y + E_3} \tag{8}$$

With x the ground truth image; y the super-resolved image; μ_x the mean of x; μ_y the mean of y; σ_x^2 the variance of x; σ_y^2 the variance of y; σ_{xy} the covariance of x and y; $E_1 = (K_1L)^2$; $E_2 = (K_2L)^2$; $E_3 = \frac{E_2}{2}$; L the dynamic of the pixel values, which is 255 for the 8-bit coding; $K_1 = 0{,}01$; $K_2 = 0{,}03$.

E_1, E_2 and E_3 are intended to stabilize the ratio when the denominator is very low.

We apply the formula of Eq. (5) to the luminance on windows of size 8x8 in order to evaluate the visual quality of the whole image, moving pixel by pixel. However, in order to reduce the computational complexity, we use only a subset of these windows (reduction by a factor of two in both dimensions). This reduces the complexity of the calculation.

The use of the PSNR and SSIM metrics allows us to measure the performance of the RH algorithms, but they do not reflect the overall quality of image reconstruction, which led the authors [122] to carry out a comparative study. But the considered framework and the choice that was made for the selection of the models force us to consider these two metrics for the fair and balanced evaluation, however, we can consider other metrics that are less used for the SR as the perceptual score that has a relationship with the metric indices [24].

4. RESULTS OF THE STUDY

We present the experimental results of this study by scale factor used for SR, presenting each time the results obtained in terms of PSNR and SSIM, but presenting only the 3 best networks by metric in terms of visual quality (PSNR) for each dataset. (PS in the tables refers to PSNR while SS refers to SSIM).

4.1 Scale factor of 2

The CAR [123] model showed better performance on small and medium datasets of its architecture that uses adaptive and guided content reordering to preserve essential information using an end-to-end learning pipeline (including the degradation function which gives strength to this model), for the DIV2K dataset this last one is not the best because it used it for testing and not for training at the time of learning. We also notice that the latter builds the edges well and produces sharp images. For the DRLN+ [72] model It improves visual quality by using Laplacian attention to learn complex inter and intra-level features through structure, it is a way to focus on important information, and that is why it was noticed for Manga9 which is a complex dataset with high-frequency details. It is better for low-resolution images. But it is not good for images with very fine details. This same principle is taken up by HAN+ [103] but this time for the correlations between the convolution layers which gives very fine high-frequency structures, it is also better for natural images. HBPN [86] has shown performances for large datasets by exploring spatial correlations between layers which allows the reconstruction of fine textures. Not to forget CSNLN [37] which is good for the faithful reconstruction of natural, complex, and high-frequency repeating features (Table 4).

4.2 Scale factor of 3

DRLN+ and HAN+ which keeps these performances also for the scaling factor of 3 for the reasons we have mentioned before. We also have SRFBN [91] which stands out on DIV2K and is a feedback network between information representation levels for faithful SR with several degradation mechanisms. This network is suitable for images with several degradation mechanisms (Table 5).

Table 4. Top 3 networks of scale factor 2

Set5		Set14		BSD100		Urban100		D1V2K		Mangal09	
Method	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**
CAR	38,94/0,97	CAR	35,61/0,93	CAR	34,78/0,90	CAR	35,24/0,91	HAN+	39,02/0,97	DRLN+	39,75/0,99
DRLN+	38,34/0,96	DRLN+	34,43/0,92	DRLN+	33,83/0,90	DRLN+	33,54/0,94	DRLN+	38,65/0,97	HAN+	39,62/0,98
HAN+	38,33/0,93	CSRCNN	34,34/0,92	HAN+	32,47/0,90	HAN+	33,53/0,94	HBPN	38,55/0,97	CSNLN	39,37/0,99

Table 5. Top 3 networks of scale factor 3

Set5		Set14		BSD100		Urban100		D1V2K		Manga109	
Method	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**
DRLN+	34,86/0,93	DRLN+	30,80/0,85	HAN+	29,41/0,81	DRLN+	29,36/0,87	HAN+	35,04/0,93	DRLN+	34,94/0,95
HAN+	34,85/0,99	HAN+	30,79/0,85	DRLN+	29,40/0,85	HAN+	29,21/0,87	SRFBN	34,89/0,90	HAN+	34,87/0,95
SRFBN	34,75/0,98	DRLN	30,73/0,92	DRLN	29,36/0,88	DRLN	29,21/0,90	DRLN+	34,43/0,96	DRLN	34,71/0,98

Table 6. Top 3 networks of scale factor 4

Set5		Set14		BSD100		Urban100		D1V2K		Manga109	
Method	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**	**Method**	**PS/SS**
SwinIR	32,93/0,90	SwinIR	29,15/0,79	CAR	29,15/0,87	CAR	29,28/0,87	ABPN	32,87/0,90	ABPN	31,79/0,92
CAR	32,82/0,91	CAR	29,09/0,79	DRLN+	27,87/0,86	HBPN	27,30/0,85	CSNLN	32,21/0,88	DRLN+	31,78/0,92
HAN+	32,75/0,90	SAN	29,05/0,93	SAN	27,86/0,86	SAN	27,23/0,84	SRGAN	32,17/0,88	DBPN-RES-MR64-3	31,7/0,90

4.3 Scale factor of 4

For this scale factor we notice better performance for the SwinIR [27] model for small datasets, the latter uses transformers and is used both for SR, de-noising, and JPEG compression artifacts reduction, it is a fast network but uses a lot of training data, it can be used for real data but suffers from several limitations.

In addition, we note the superiority of the CAR and DRLN+ models, but also the SAN [107] model which also adopts the correlation between layers, by the second order channel attention principle (SOCA). This staged network tries to produce realistic images (Table 6).

4.4 Scale factor of 8

Table 7. Top 3 networks of scale factor 8

Set5		Set14		BSD100		Urban100		D1V2K		Manga109	
Method	**PSSS**	**Method**	**PSSS**	**Method**	**PSSS**	**Method**	**PSSS**	**Method**	**PSSS**	**Method**	**PSSS**
MFSRCNN	29,03 0,82	DBPN-RES-MR64-3	25,41 0,65	DRLN+	25,06 0,60	DRLN+	23,24 0,65	DRLN+	28,23 0,72	DBPN-RES-MR64-3	25,71 0,81
DBPN-RES-MR64-3	27,51 0,79	DRLN+	25,4 0,65	DBPN-RES-MR64-3	25,05 0,60	DBPN-RES-MR64-3	23,2 0,65	HAN+	28,18 0,72	DRLN+	25,55 0,80
HAN+	27,47 0,79	HAN+	25,39 0,65	HAN+	25,04 0,60	HAN+	23,2 0,65	DBPN-RES-MR64-3	28,18 0,71	HAN+	25,54 0,80

In addition to HAN+ and DRLN+, we note the best performances of the DBPN-Res-MR64-3 [84] model for an SR X8, this network is a combination of dense connection, residual learning, and recurrent network, it is based on the exploitation of iterative up and down sampling layers using the dependency between the LR and HR image within a deep back-projection network allowing remarkable results for this scale factor. It is a particular network for large-scale factors (X8) but requires a significant amount of time for learning (Table 7).

4.5 Scale factor of 16

The study framework fixed at the beginning forces us to select only one network for this scaling factor which is ABPN [60]. This network requires a larger learning field, so we added the DIV8K dataset [124] which allowed an acceptable visual result. This back-projection learning model exploits the cross-

correlation by an attention mechanism, but the consideration of the scale factor of 16 allowed to loss the realism within the reconstructed images (Table 8).

Table 8. Results of the scaling factor 16

SET5		SET14		BSD 100		URBAN 100		MANGA109		DIV2K		DIV8K	
PS	**SS**	**PS**	**SS**	**PS**	**SS**	**PS**	**SS**	**PS**	**SS**	**PS**	**SS**	**PS**	**SS**
23,42	0,61	22,17	0,53	22,72	0,51	20,39	0,52	21,25	0,67	24,38	0,64	26,71	0,65

5. DISCUSSION AND ORIENTATIONS

5.1 Discussion

Historically and since the evolution of neural networks, CNNs are the most used for SR, we note that several techniques have been proposed for SISR such as the use of very deep networks and recursive networks in addition to connection hopping, residual networks, multi-scale processing, texture synthesis, overhead networks, hierarchical features, dense connections, attention networks, and attention mechanisms. Moreover, GANs allow for gaining perceptual and sometimes visual quality through loss and opposition functions.

The analysis carried out in this article has allowed us to observe a considerable evolution in the performance and accuracy of SR models given the technological possibilities offered by deep learning. This development generates a considerable processing capacity given the complexity of the proposed algorithms but this constraint is still supported today because of the availability of storage and datasets diversified in complexity and size. However, a balance has to be made between the SR task to be performed and the size and structure of the training data, since the performance depends on this balance. Thus, choosing the best algorithm for the task and the training data is often a difficult task and remains unclear. This constraint leads to a historical problem for the SR task which is the non-generalization of the learning networks on real-world datasets since the learning was done on simulated images.

Indeed, the analysis performed allowed us to deduce that the content adaptive image downscaling technique allowed to have an optimal performance by producing better quality HR images from LR images with potentially detailed information [123], as well as for the correlations between layers [103], without forgetting the learning of inter and intra level features which helps to improve the accuracy [72].

Thus, preserving the essential features of the image is the key to any SISR algorithm, and this process requires mastering the pipeline between the LR and HR image in both directions.

We also notice that a large part of the selected models uses a simple and uniform degradation in the form of sub-sampling to generate LR images from the corresponding HR images which result in a simulated dataset. This way of doing things is only suitable for real datasets as their degradations are more complex, which produces less efficient SISR algorithms for practical applications [125].

He also notes that if we have images with less high-frequency detail, we can have a higher PSNR.

To this end, we note that the exploitation of correlations between intermediate layers is less used in most of the models studied. This technique has given remarkable performances in terms of reconstruction quality by recovering fine textures which improves the perceptual quality contrary to networks that focus on the implementation of larger or deeper architectures.

The study carried out shows that we must go through a thorough analysis according to the methodology proposed in this article according to the objective to be achieved in order to make a better choice of the adapted network, which can be completed by other metrics and other aspects. We note that there are models that are best suited for the case of primitive images with less detail and are not suited to the case of complex images, so the choice must be conditioned by the type of image. Therefore, we can deduce that the SR is determined according to the richness in high-frequency details, which is confirmed by other studies [23]. Thus, images with less contrasted pixels lead to higher PSNR values and vice versa.

In our experience, the network input is a determining factor for having higher PSNR and SSIM metrics. For models that use a traditional input, they cannot have a better reconstruction because of the statistical properties that prevent less frequent patterns from being reconstructed on the super-resolved image.

In addition, the model evaluation metric must be well chosen according to the training domain for SR, as we can validate the model through the Spearman rank correlation coefficient [126] which proved this performance for image-adapted metrics. According to [23] we can use 4 other metrics for SISR, namely the weighted peak signal-to-noise ratio (WPSNR), the multi-scale structure similarity index (MSSSIM) [127], the noise quality measure (NQM) [128] and the information fidelity criterion (IFC) [48]. But the metric should be adapted according to the dataset and the SR task to be performed as there are metrics that focus on the edges than on the center and vice versa adapted, other metrics that are adapted to human perception or for natural scenes, as well as other metrics, work well on high-frequency images while others work on low-frequency images. Indeed, traditional reduction methods are designed for better human visual perception, which makes them unsuitable for performance measurement metrics that use signal distortion. In this particular case, we cannot correctly recover images via SR because we can have beautiful images but at the expense of realism.

The WPSNR metric significantly improves performance since it uses perceptual contrast sensitivity in terms of spatial frequency [129] to compute weights, unlike PSNR which considers weights to be identical throughout the learning process.

This study determines the scaling factor of 4 as acceptable performance for the SISR methods evaluated in this paper, since exceeding this limit results in significant complexity and performance degradation.

The degradation method is quite important because it determines how the reconstruction will be done especially the noise and blur that sometimes block the SR process.

Thus, most of the methods aim to have a high PSNR but produce blurred images, GANs try to solve this problem by

generating images with synthetic textures but can generate false textures. Today the attention mechanism in neural networks offers a better track for SR especially if combined with correlation learning between feature maps. Moreover, transformers are also a promising alternative.

In order to guide the user's choice we propose to use CAR [123] for small and medium size datasets and if we don't know the degradation process within the image, on the other hand, DRLN [72] is better for small images and with high-frequency details, moreover, HAN [103] fits well for natural images with high-frequency content, however, ABPN [60] converges well for a large dataset, thus, CSNLN [37] fits for natural, complex and with repeating features, finally, SRFBN [91] is better for images with multi-degradation. So, the scaling factor is important, that's why we have presented the results by scaling factor if the user knows before the desired scaling factor he goes directly to the selected networks in this category, otherwise, he can try a factor of 4 if the result is satisfactory otherwise, he tries a factor of 2.

5.2 Orientations

Despite the development made in the field of SISR, there are still some limitations in the existing models, in this section, we present practical directions to consider for the implementation of deep learning algorithms for image SR, in order to improve the existing algorithms or to guide future researchers in this field to solve open SR problems. Indeed, when applying deep learning algorithms to solve a very specific problem such as SISR, both common deep learning techniques and model learning techniques are also very important to consider, as well as specific knowledge of image and SR characteristics in order to have optimal and efficient architectures for SR.

Indeed, to solve the problem of designing larger or deeper architectures (see Discussion) and to remain faithful to complex image features, the proposed model must be powerful in terms of learning correlations between image patterns to reduce the computational burden (by exploiting digital image properties), incorporating residual group structures. Furthermore, to solve the problem of image realism, it is necessary to work with a real-world dataset with authentic degradations (e.g., working with multi-resolution cameras without generating LR images from HR or changing the image acquisition sensor lens). Experiments have shown a remarkable increase in performance on this type of dataset as well as a great ability to extend to other datasets. Therefore, it is necessary to expand the dataset and investigate new strategies for learning the SISR model [19].

Learning methods for SR must take into account not only the degradation of image resolution, but also the blurring and noise in the images. In addition, post-processing methods can be used to enhance the SR process to improve the contrast in SISR methods, these methods should generate a low computational load so as not to burden the learning process [34, 130].

Thus, in this study, a comparison was made based on objective and subjective criteria (SSIM, PSNR, and architecture). However, other criteria can be considered such as the number of parameters, the memory footprint, and the learning time of the model which are directly related to the number of adjustable parameters. We can also consider the execution time of the algorithm. It is necessary to work with images without compression or with lossless compression, in order to preserve the intrinsic characteristics of the images. Ideally, one should work with a completely unsupervised pipeline, so as not to assume how the images are scaled.

6. CONCLUSION AND PERSPECTIVES

In this paper, we conducted a comprehensive and in-depth study and analysis of the performance of deep learning models important for SISR. The study carried out considers the same framework of analysis to standardize the study and have a balanced and fair comparison, this survey has focused on 90 algorithms that are grouped by scale factors and tested on 7 datasets. The results obtained allowed us to classify the scale factor models, analyzing in detail the first 3 networks selected to take advantage of the techniques used. A detailed and systematic evaluation has been carried out through quantitative experiments to identify the strengths and limitations of these models considering reference metrics. This study tried to be balanced and fair by respecting the principles of a systematic review. As such it is noted that a considerable increase in accuracy and performance has been seen recently for SISR models but there are still other opportunities to be exploited and other limitations to be overcome. Thus, improvements to existing algorithms and other key factors for the design of new SISR architectures have been proposed.

At the end of this study, we can propose some perspectives for neural networks for SISR, among others:

- The combination between the understanding of fundamental deep learning techniques (design, training, optimization, ...) as well as the specific problems of SISR (image descriptor, degradation method, ...).
- Considering the correlations between the layers of the neural network allows for better learning and efficient optimization, and the correlation between the data (image channels) allows for a faithful mapping between the LR and HR images.
- The degradation process (subsampling) must preserve the essential characteristics of the original image (high and low-frequency details) for a better reconstruction.
- The SISR algorithm must be able to control the image SR process in both directions, i.e., both the learning of the degradation method and the reconstruction method.
- The design and implementation of optimal architectures in terms of complexity and computation time as well as in several parameters in order to facilitate the use of these architectures in the real world and on a larger scale. Thus, a more complex model facilitates overlearning within the network.
- Consideration of domain-adaptive evaluation criteria as well as designing objective and loss functions (pixel loss, content/perceptual loss, conflicting loss, texture loss, total variation loss) consistent with the domain specification, or combining several functions, such that the loss function determines the nature of the SR algorithm.
- The adoption of the best network design strategies by having simple networks.

REFERENCES

[1] https://fr.wikipedia.org/wiki/Vision_par_ordinateur, accessed on Sept. 17, 2022.

[2] Ahmad, I., Pothuganti, K. (2020). Design &

implementation of real time autonomous car by using image processing & IoT. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, pp. 107-113. https://doi.org/10.1109/ICSSIT48917.2020.9214125

[3] Aburaed, N., Panthakkan, A., Al-Saad, M., El Rai, M.C., Al Mansoori, S., Al-Ahmad, H., Marshall, S. (2020). Super-resolution of satellite imagery using a wavelet multiscale-based deep convolutional neural network model. In Image and Signal Processing for Remote Sensing XXVI, 11533: 305-311. https://doi.org/10.1117/12.2573991

[4] Yamashita, K., Markov, K. (2020). Medical image enhancement using super resolution methods. In International Conference on Computational Science, Amsterdam, The Netherlands, pp. 496-508. https://doi.org/10.1007/978-3-030-50426-7_37

[5] Farooq, M.A., Khan, A.A., Ahmad, A., Raza, R.H. (2020). Effectiveness of state-of-the-art super resolution algorithms in surveillance environment. In Conference on Multimedia, Interaction, Design and Innovation, pp. 79-88. https://doi.org/10.1007/978-3-030-74728-2_8

[6] Iftenea, M., Liub, Q., Wangc, Y. (2017). Very high resolution images classification by fusing deep convolutional neural networks. The 5th International Conference on Advanced Computer Science Applications and Technologies (ACSAT 2017), pp. 172-176.

[7] Wang, L., Li, D., Zhu, Y., Tian, L., Shan, Y. (2020). Dual super-resolution learning for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 3774-3783. https://doi.org/10.1109/CVPR42600.2020.00383

[8] Soufi, O., Aarab, Z., Belouadha, F.Z. (2022). Benchmark of deep learning models for single image super-resolution (SISR). In 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, pp. 1-8. https://doi.org/10.1109/IRASET52964.2022.9738274

[9] Gómez, D., Rojas, A. (2016). An empirical overview of the no free lunch theorem and its effect on real-world machine learning classification. Neural Computation, 28(1): 216-228. https://doi.org/10.1162/NECO_a_00793

[10] Blau, Y., Mechrez, R., Timofte, R., Michaeli, T., Zelnik-Manor, L. (2018). The 2018 PIRM challenge on perceptual image super-resolution. In European Conference on Computer Vision, Munich, Germany, pp. 334-355. https://doi.org/10.1007/978-3-030-11021-5_21

[11] Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.H., Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia, 21(12): 3106-3121. https://doi.org/10.1109/TMM.2019.2919431

[12] Kopf, J., Lischinski, D. (2011). Depixelizing pixel art. In ACM Transactions on Graphics, 30(4): 1-8. https://doi.org/10.1145/1964921.1964994

[13] Choi, J.S., Kim, M. (2017). A deep convolutional neural network with selection units for super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, pp. 154-160. https://doi.org/10.1109/CVPRW.2017.153

[14] Sun, L., Hays, J. (2012). Super-resolution from internet-scale scene matching. In 2012 IEEE International conference on computational photography (ICCP), Seattle, WA, USA, pp. 1-12. https://doi.org/10.1109/ICCPhot.2012.6215221

[15] Cheng, X., Chen, Z. (2020). Video frame interpolation via deformable separable convolution. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(7): 10607-10614. https://doi.org/10.1609/aaai.v34i07.6634

[16] Ranganarayana, K., Rao, G.V. (2022). Modified ant colony optimization for human recognition in videos of low resolution. Revue d'Intelligence Artificielle, 36(5): 731-736. https://doi.org/10.18280/ria.360510

[17] Laghrib, A., Hadri, A., Hakim, A., Raghay, S. (2019). A new multiframe super-resolution based on nonlinear registration and a spatially weighted regularization. Information Sciences, 493: 34-56. https://doi.org/10.1016/j.ins.2019.04.029

[18] Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P. (2004). Fast and robust multiframe super resolution. IEEE transactions on image processing, 13(10): 1327-1344. https://doi.org/10.1109/TIP.2004.834669

[19] Chen, H., He, X., Qing, L., Wu, Y., Ren, C., Sheriff, R. E., Zhu, C. (2022). Real-world single image super-resolution: A brief review. Information Fusion, 79: 124-145. https://doi.org/10.1016/j.inffus.2021.09.005

[20] Zhou, F., Yang, W., Liao, Q. (2012). Interpolation-based image super-resolution using multisurface fitting. IEEE Transactions on Image Processing, 21(7): 3312-3318. https://doi.org/10.1109/TIP.2012.2189576

[21] Hardie, R.C., Barnard, K.J., Armstrong, E.E. (1997). Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. IEEE transactions on Image Processing, 6(12): 1621-1633. https://doi.org/10.1109/83.650116

[22] Dong, C., Loy, C.C., He, K., Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2): 295-307. https://doi.org/10.1109/TPAMI.2015.2439281

[23] Yang, C.Y., Ma, C., Yang, M.H. (2014). Single-image super-resolution: A benchmark. In European Conference on Computer Vision, Zurich, Switzerland, pp. 372-386. https://doi.org/10.1007/978-3-319-10593-2_25

[24] Yang, J., Wright, J., Huang, T.S., Ma, Y. (2010). Image super-resolution via sparse representation. IEEE transactions on image processing, 19(11): 2861-2873. https://doi.org/10.1109/TIP.2010.2050625

[25] Gavade, A., Sane, P. (2014). Super resolution image reconstruction by using bicubic interpolation. In National Conference on Advanced Technologies in Electrical and Electronic Systems, 10: 1.

[26] Zhang, K., Liang, J., Van Gool, L., Timofte, R. (2021). Designing a practical degradation model for deep blind image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, pp. 4791-4800. https://doi.org/10.1109/ICCV48922.2021.00475

[27] Schulter, S., Leistner, C., Bischof, H. (2015). Fast and accurate image upscaling with super-resolution forests. In Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, MA, USA, pp. 3791-3799. https://doi.org/10.1109/CVPR.2015.7299003

[28] Johnson, J., Alahi, A., Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In European conference on computer vision, 694-711. https://doi.org/10.1007/978-3-319-46475-6_43
[29] Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, pp. 136-144. https://doi.org/10.1109/CVPRW.2017.151
[30] Sun, W., Chen, Z. (2020). Learned image downscaling for upscaling using content adaptive resampler. IEEE Transactions on Image Processing, 29: 4027-4040. https://doi.org/10.1109/TIP.2020.2970248
[31] Kuyoro, A., Nzenwata, U.J., Awodele, O., Idowu, S. (2022). GAN-based encoding model for reversible image steganography. Revue d'Intelligence Artificielle, 36(4): 561-567. https://doi.org/10.18280/ria.360407
[32] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition, Amsterdam, The Netherlands, pp. 4681-4690. https://doi.org/10.1109/CVPR.2017.19
[33] Zhang, K., Li, D., Luo, W., Ren, W., Stenger, B., Liu, W., Yang, M.H. (2021). Benchmarking ultra-high-definition image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, pp. 14769-14778. https://doi.org/10.1109/ICCV48922.2021.01450
[34] Wang, Z., Chen, J., Hoi, S.C. (2020). Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence, 43(10): 3365-3387.
[35] Liu, H., Ruan, Z., Zhao, P., Dong, C., Shang, F., Liu, Y., Timofte, R. (2022). Video super-resolution based on deep learning: A comprehensive survey. Artificial Intelligence Review, 55: 5981-6035. https://doi.org/10.1007/s10462-022-10147-y
[36] Zhang, K., Liang, J., Van Gool, L., Timofte, R. (2021). Designing a practical degradation model for deep blind image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 4791-4800. https://doi.org/10.48550/arXiv.2103.14006
[37] Choi, J.H., Kim, J.H., Cheon, M., Lee, J.S. (2018). Lightweight and efficient image super-resolution with block state-based recursive network. arXiv Preprint arXiv:1811.12546. https://doi.org/10.48550/arXiv.1811.12546
[38] Zhang, K., Zuo, W., Gu, S., Zhang, L. (2017). Learning deep CNN denoiser prior for image restoration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 3929-3938. https://doi.org/10.1109/CVPR.2017.300
[39] Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.H., Liao, Q. (2019). LCSCNet: Linear compressing-based skip-connecting network for image super-resolution. IEEE Transactions on Image Processing, 29: 1450-1464. https://doi.org/10.1109/TIP.2019.2940679
[40] Timofte, R., Rothe, R., Van Gool, L. (2016). Seven ways to improve example-based single image super resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1865-1873. https://doi.org/10.1109/CVPR.2016.206
[41] Kim, J., Lee, J.K., Lee, K.M. (2016). Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1646-1654. https://doi.org/10.48550/arXiv.1511.04587
[42] Han, W., Chang, S., Liu, D., Yu, M., Witbrock, M., Huang, T.S. (2018). Image super-resolution via dual-state recurrent networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 1654-1663. https://doi.org/10.1109/CVPR.2018.00178
[43] Wang, Z., Liu, D., Yang, J., Han, W., Huang, T. (2015). Deep networks for image super-resolution with sparse prior. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, pp. 370-378. https://doi.org/10.1109/ICCV.2015.50
[44] Li, Z., Zhao, C., Zhang, H., Zhang, Z., Wu, X. (2021). Application of multi-scale fusion super-resolution algorithm in UAV detection. Optics Journal, 42(3): 462-473. https://doi.org/10.5768/JAO202142.0302003
[45] Li, H., Lam, K.M., Li, D. (2018). Joint maximum purity forest with application to image super-resolution. Journal of Electronic Imaging, 27(4): 043005. https://doi.org/10.1117/1.JEI.27.4.043005
[46] Kim, J.H., Choi, J.H., Cheon, M., Lee, J.S. (2020). MAMNet: Multi-path adaptive modulation network for image super-resolution. Neurocomputing, 402: 38-49. https://doi.org/10.1016/j.neucom.2020.03.069
[47] Plötz, T., Roth, S. (2018). Neural nearest neighbors networks. Advances in Neural Information Processing Systems, 1095-1106. https://doi.org/10.48550/arXiv.1810.12575
[48] Zhang, J., Wang, Z., Zheng, Y., Zhang, G. (2021). Cascaded convolutional neural network for image super-resolution. In International Conference on Artificial Intelligence and Security, Dublin, Ireland, pp. 361-373. https://doi.org/10.1007/978-3-030-78615-1_32
[49] Sajjadi, M.S., Scholkopf, B., Hirsch, M. (2017). Enhancenet: Single image super-resolution through automated texture synthesis. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, pp. 4491-4500. https://doi.org/10.1109/ICCV.2017.481
[50] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L. (2017). Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Transactions on Image Processing, 26(7): 3142-3155. https://doi.org/10.1109/TIP.2017.2662206
[51] Ren, H., El-Khamy, M., Lee, J. (2017). Image super resolution based on fusing multiple convolution neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, pp. 54-61. https://doi.org/10.1109/CVPR.2016.182
[52] Gu, J., Lu, H., Zuo, W., Dong, C. (2019). Blind super-resolution with iterative kernel correction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 1604-1613. https://doi.org/10.1109/CVPR.2019.00170

[53] Hu, Y., Gao, X., Li, J., Huang, Y., Wang, H. (2018). Single image super-resolution via cascaded multi-scale cross network. arXiv preprint arXiv:1802.08808. https://doi.org/10.48550/arXiv.1802.08808
[54] Zhang, Z., Wang, Z., Lin, Z., Qi, H. (2019). Image super-resolution by neural texture transfer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 7982-7991. https://doi.org/10.1109/CVPR.2019.00817
[55] Wang, Y., Perazzi, F., McWilliams, B., Sorkine-Hornung, A., Sorkine-Hornung, O., Schroers, C. (2018). A fully progressive approach to single-image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, pp. 864-873. https://doi.org/10.1109/CVPRW.2018.00131
[56] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 4681-4690. https://doi.org/10.1109/CVPR.2017.19
[57] Shocher, A., Cohen, N., Irani, M. (2018). "zero-shot" super-resolution using deep internal learning. In Proceedings of the IEEE conference on computer vision and pattern recognition, Lake City, UT, USA, pp. 3118-3126. https://doi.org/10.48550/arXiv.1712.06087
[58] Liu, Z.S., Wang, L.W., Li, C.T., Siu, W.C., Chan, Y.L. (2019). Image super-resolution via attention based back projection networks. In 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea (South), pp. 3517-3525. https://doi.org/10.1109/ICCVW.2019.00436
[59] Kong, S., Fowlkes, C. (2018). Image reconstruction with predictive filter flow. arXiv preprint arXiv:1811.11482. https://doi.org/10.48550/arXiv.1811.11482
[60] Ma, C., Rao, Y., Cheng, Y., Chen, C., Lu, J., Zhou, J. (2020). Structure-preserving super resolution with gradient guidance. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 7769-7778. https://doi.org/10.1109/CVPR42600.2020.00779
[61] Huang, H., He, R., Sun, Z., Tan, T. (2017). Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, pp. 1689-1697. https://doi.org/10.1109/ICCV.2017.187
[62] Cheng, X., Li, X., Yang, J., Tai, Y. (2018). SESR: Single image super resolution with recursive squeeze and excitation networks. In 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, pp. 147-152. https://doi.org/10.1109/ICPR.2018.8546130
[63] Zhang, Y., Wei, D., Qin, C., Wang, H., Pfister, H., Fu, Y. (2021). Context reasoning attention network for image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, pp. 4278-4287. https://doi.org/10.1109/ICCV48922.2021.00424
[64] Choi, J.H., Kim, J.H., Cheon, M., Lee, J.S. (2020). Deep learning-based image super-resolution considering quantitative and perceptual quality. Neurocomputing, 398: 347-359. https://doi.org/10.1016/j.neucom.2019.06.103
[65] Song, D., Wang, Y., Chen, H., Xu, C., Xu, C., Tao, D. (2021). Addersr: Towards energy efficient image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, pp. 15648-15657. https://doi.org/10.1109/CVPR46437.2021.01539
[66] Hui, Z., Li, J., Gao, X., Wang, X. (2021). Progressive perception-oriented network for single image super-resolution. Information Sciences, 546: 769-786. https://doi.org/10.1016/j.ins.2020.08.114
[67] Son, S., Lee, K.M. (2021). SRWarp: Generalized image super-resolution under arbitrary transformation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, pp. 7782-7791. https://doi.org/10.48550/arXiv.2104.10325
[68] Fan, Y., Shi, H., Yu, J., Liu, D., Han, W., Yu, H., Huang, T.S. (2017). Balanced two-stage residual networks for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, pp. 161-168. https://doi.org/10.1109/CVPRW.2017.154
[69] Mao, X., Shen, C., Yang, Y.B. (2016). Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Advances in Neural Information Processing Systems, 29. https://doi.org/10.48550/arXiv.1603.09056
[70] Tai, Y., Yang, J., Liu, X. (2017). Image super-resolution via deep recursive residual network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 3147-3155. https://doi.org/10.1109/CVPR.2017.298
[71] Tai, Y., Yang, J., Liu, X., Xu, C. (2017). Memnet: A persistent memory network for image restoration. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, pp. 4539-4547. https://doi.org/10.1109/ICCV.2017.486
[72] Haris, M., Shakhnarovich, G., Ukita, N. (2018). Deep back-projection networks for super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 1664-1673. https://doi.org/10.1109/CVPR.2018.00179
[73] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y. (2018). Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 2472-2481. https://doi.org/10.1109/CVPR.2018.00262
[74] Plötz, T., Roth, S. (2018). Neural nearest neighbors networks. Advances in Neural Information Processing Systems, 1095-1106. https://doi.org/10.48550/arXiv.1810.12575
[75] Wang, Z., Liu, D., Yang, J., Han, W., Huang, T. (2015). Deep networks for image super-resolution with sparse prior. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, pp. 370-378. https://doi.org/10.1109/ICCV.2015.50
[76] Fu, C.H., Chen, H., Zhang, Y.L. (2014). Chan, Single image super resolution based on sparse representation and adaptive dictionary selection. In: 2014 19th International Conference on Digital Signal Processing,

Hong Kong, China, pp. 449-453. https://doi.org/10.1109/ICDSP.2014.6900704
[77] Dong, C., Loy, C.C., He, K., Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In European conference on Computer Vision, Zurich, Switzerland, pp. 184-199. https://doi.org/10.1007/978-3-319-10593-2_13
[78] Hui, Z., Gao, X., Yang, Y., Wang, X. (2019). Lightweight image super-resolution with information multi-distillation network. In Proceedings of the 27th acm International Conference on Multimedia, New York, NY, United States, pp. 2024-2032. https://doi.org/10.1145/3343031.3351084
[79] Dai, T., Cai, J., Zhang, Y., Xia, S.T., Zhang, L. (2019). Second-order attention network for single image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 11065-11074. https://doi.org/10.1109/CVPR.2019.01132
[80] Bulat, A., Tzimiropoulos, G. (2018). Super-fan: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with gans. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 109-117. https://doi.org/10.48550/arXiv.1712.02765
[81] Yang, W., Wang, W., Zhang, X., Sun, S., Liao, Q. (2019). Lightweight feature fusion network for single image super-resolution. IEEE Signal Processing Letters, 26(4): 538-542. https://doi.org/10.1109/LSP.2018.2890770
[82] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R. (2021). Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, pp. 33-1844. https://doi.org/10.1109/ICCVW54120.2021.00210
[83] Zhang, K., Zuo, W., Zhang, L. (2018). Learning a single convolutional super-resolution network for multiple degradations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 3262-3271. https://doi.org/10.1109/CVPR.2018.00344
[84] Niu, B., Wen, W., Ren, W., Zhang, X., Yang, L., Wang, S. (2020). Single image super-resolution via a holistic attention network. In European Conference on Computer Vision, Glasgow, UK, pp. 191-207. https://doi.org/10.1007/978-3-030-58610-2_47
[85] Yang, L., Wang, S., Ma, S., Gao, W., Liu, C., Wang, P., Ren, P. (2020). Hifacegan: Face renovation via collaborative suppression and replenishment. In Proceedings of the 28th ACM International Conference on Multimedia, Seattle, WA, USA, pp. 1551-1560. https://doi.org/10.1145/3394171.3413965
[86] Liu, Y., Zhang, X., Wang, S., Ma, S., Gao, W. (2020). Progressive multi-scale residual network for single image super-resolution. arXiv Preprint arXiv:2007.09552. https://doi.org/10.48550/arXiv:2007.09552
[87] Ahn, N., Kang, B., Sohn, K.A. (2018). Fast, accurate, and lightweight super-resolution with cascading residual network. In Proceedings of the European Conference on Computer Vision (ECCV), 252-268. https://doi.org/10.48550/arXiv.1803.08664
[88] Liu, P., Zhang, H., Zhang, K., Lin, L., Zuo, W. (2018). Multi-level wavelet-CNN for image restoration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, pp. 773-782. https://doi.org/10.1109/CVPRW.2018.00121
[89] Mei, Y., Fan, Y., Zhou, Y., Huang, L., Huang, T.S., Shi, H. (2020). Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 5690-5699. https://doi.org/10.1109/CVPR42600.2020.00573
[90] Wang, L., Wang, Y., Liang, Z., Lin, Z., Yang, J., An, W., Guo, Y. (2019). Learning parallax attention for stereo image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 12250-12259. https://doi.org/10.1109/CVPR.2019.01253
[91] Dong, C., Loy, C.C., Tang, X. (2016). Accelerating the super-resolution convolutional neural network. In European Conference on Computer Vision, Amsterdam, pp. 391-407. https://doi.org/10.1007/978-3-319-46475-6_25
[92] Zhang, M., Liu, Z., Yu, L. (2018). Image Super-Resolution via RL-CSC: When Residual Learning Meets Convolutional Sparse Coding. arXiv Preprint arXiv:1812.11950. https://doi.org/10.48550/arXiv.1812.11950
[93] Liu, P., Zhou, X., Yang, J., Fang, R. (2019). Image Restoration Using Deep Regulated Convolutional Networks. arXiv preprint arXiv:1910.08853. https://doi.org/10.48550/arXiv.1910.08853
[94] Johnson, J., Alahi, A., Li, F. (2016). Perceptual losses for real-time style transfer and super-resolution. In European Conference on Computer Vision, Amsterdam, The Netherlands, pp. 694-711. https://doi.org/10.1007/978-3-319-46475-6_43
[95] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y. (2018). Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 2472-2481. https://doi.org/10.1109/CVPR.2018.00262
[96] Li, Q., Li, Z., Lu, L., Jeon, G., Liu, K., Yang, X. (2019). Gated multiple feedback network for image super-resolution. arXiv Preprint arXiv:1907.04253. https://doi.org/10.48550/arXiv.2007.09552
[97] Glasner, D., Bagon, S., Irani, M. (2009). Super-resolution from a single image. In 2009 IEEE 12th international conference on computer vision, Kyoto, Japan, pp. 349-356. https://doi.org/10.1109/ICCV.2009.5459271
[98] Wang, X., Yu, K., Dong, C., Loy, C.C. (2018). Recovering realistic texture in image super-resolution by deep spatial feature transform. In Proceedings of the IEEE conference on computer vision and pattern recognition, Lake City, UT, USA, pp. 606-615. https://doi.org/10.1109/CVPR.2018.00070
[99] Li, H., Lam, K.M., Wang, M. (2019). Image super-resolution via feature-augmented random forest. Signal Processing: Image Communication, 72: 25-34. https://doi.org/10.1016/j.image.2018.12.001
[100] Caballero, J., Ledig, C., Aitken, A., Acosta, A., Totz, J., Wang, Z., Shi, W. (2017). Real-time video super-resolution with spatio-temporal networks and motion

compensation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 4778-4787. https://doi.org/10.1109/CVPR.2017.304

[101] Liu, Z.S., Wang, L. W., Li, C.T., Siu, W.C. (2019). Hierarchical back projection network for image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 0-0. https://doi.org/10.48550/arXiv.1906.06874

[102] Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA pp. 624-632. https://doi.org/10.1109/CVPR.2017.618

[103] Li, Z., Yang, J., Liu, Z., Yang, X., Jeon, G., Wu, W. (2019). Feedback network for image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 3867-3876. https://doi.org/10.1109/CVPR.2019.00399

[104] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778. https://doi.org/10.1109/CVPR.2016.90

[105] Liu, D., Wen, B., Fan, Y., Loy, C.C., Huang, T.S. (2018). Non-local recurrent network for image restoration. Advances in Neural Information Processing Systems. https://doi.org/10.48550/arXiv.1806.02919

[106] Liang, M., Du, J., Li, X., Xu, L., Liu, H., Li, Y. (2013). Spatio-temporal super-resolution reconstruction based on robust optical flow and Zernike moment for dynamic image sequences. In 2013 IEEE International Symposium on Industrial Electronics, Taipei, Taiwan, pp. 1-6. https://doi.org/10.1109/ISIE.2013.6563661

[107] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 4681-4690. https://doi.org/10.1109/CVPR.2017.19

[108] Yu, Y., Si, X., Hu, C., Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation, 31(7): 1235-1270. https://doi.org/10.1162/neco_a_01199

[109] Anwar, S., Barnes, N. (2020). Densely residual laplacian super-resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence. 44(3): 1-1. https://doi.org/10.1109/TPAMI.2020.3021088

[110] Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and pattern Recognition Workshops, Honolulu, HI, USA, pp. 136-144. https://doi.org/10.1109/CVPRW.2017.151

[111] Zeyde, R., Elad, M., Protter, M. (2012). On single image scale-up using sparse-representations. In International Conference On Curves And Surfaces, Avignon, France, pp. 711-730. https://doi.org/10.1007/978-3-642-27413-8_47

[112] Martin, D., Fowlkes, C., Tal, D., Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, Vancouver, BC, Canada, pp. 416-423. https://doi.org/10.1109/ICCV.2001.937655

[113] Huang, J.B., Singh, A., Ahuja, N. (2015). Single image super-resolution from transformed self-exemplars. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, pp. 5197-5206. https://doi.org/10.1109/CVPR.2015.7299156

[114] Aizawa, K., Fujimoto, A., Otsubo, A., Ogawa, T., Matsui, Y., Tsubota, K., Ikuta, H. (2020). Building a manga dataset "manga109" with annotations for multimedia applications. IEEE MultiMedia, 27(2): 8-18. https://doi.org/10.1109/MMUL.2020.2987895

[115] Agustsson, E., Timofte, R. (2017). NTIRE 2017 challenge on single image super-resolution: dataset and study. In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, pp. 1122-1131. https://doi.org/10.1109/CVPRW.2017.150

[116] Yoo, J., Ahn, N., Sohn, K.A. (2020). Rethinking data augmentation for image super-resolution: A comprehensive analysis and a new strategy. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 8375-8384. https://doi.org/10.1109/CVPR42600.2020.00840

[117] Walsh, N., Muellner, L. (1999). DocBook: the definitive guide. O'Reilly Media, Inc..

[118] Hore, A., Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. In 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, pp. 2366-2369. https://doi.org/10.1109/ICPR.2010.579

[119] Blau, Y., Mechrez, R., Timofte, R., Michaeli, T., Zelnik-Manor, L. (2018). The 2018 PIRM challenge on perceptual image super-resolution. In European Conference on Computer Vision, Munich, Germany, pp. 334-355. https://doi.org/10.1007/978-3-030-11021-5_21

[120] Rethlefsen, M.L., Page, M.J. (2022). PRISMA 2020 and PRISMA-S: common questions on tracking records and the flow diagram. Journal of the Medical Library Association: JMLA, 110(2): 253-253. https://doi.org/10.5195/jmla.2022.1449

[121] Yan, B., Bare, B., Ma, C., Li, K., Tan, W. (2019). Deep objective quality assessment driven single image super-resolution. IEEE Transactions on Multimedia, 21(11): 2957-2971. https://doi.org/10.1109/TMM.2019.2914883

[122] Yang, J., Lin, Z., Cohen, S. (2013). Fast image super-resolution based on in-place example regression. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, pp. 1059-1066. https://doi.org/10.1109/CVPR.2013.141

[123] Zhang, K., Liang, J., Van Gool, L., Timofte, R. (2021). Designing a practical degradation model for deep blind image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 4791-4800. https://doi.org/10.48550/arXiv.2103.14006

[124] Lamour, J., Timofte, R. (2019). Div8k: Diverse 8k resolution image dataset. In 2019 IEEE/CVF International Conference on Computer Vision Workshop

(ICCVW), Seoul, Korea (South), pp. 3512-3516. https://doi.org/10.1109/ICCVW.2019.00435

[125] Cai, J., Zeng, H., Yong, H., Cao, Z., Zhang, L. (2019). Toward real-world single image super-resolution: A new benchmark and a new model. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), pp. 3086-3095. https://doi.org/10.1109/ICCV.2019.00318

[126] Moore, D.S., McCabe, G.P. (1989). Introduction to the Practice of Statistics. WH Freeman/Times Books/Henry Holt & Co. New York: W.H. Freeman, Macmillan Learning, 2017.

[127] Wang, Z., Simoncelli, E.P., Bovik, A.C. (2003). Multiscale structural similarity for image quality assessment. In The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2: 1398-1402. https://doi.org/10.1109/ACSSC.2003.1292216

[128] Damera-Venkata, N., Kite, T.D., Geisler, W.S., Evans, B.L., Bovik, A.C. (2000). Image quality assessment based on a degradation model. IEEE Transactions on Image Processing, 9(4): 636-650. https://doi.org/10.1109/83.841940

[129] Robson, J.G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. Josa, 56(8): 1141-1142. https://doi.org/10.1364/JOSA.56.001141

[130] Hui, Z., Li, J., Gao, X., Wang, X. (2021). Progressive perception-oriented network for single image super-resolution. Information Sciences, 546: 769-786. https://doi.org/10.1016/j.ins.2020.08.114