

6 Super-Resolution of 3D Magnetic Resonance Images of the Brain

Enrique Domínguez, Domingo López-Rodríguez, Ezequiel López-Rubio, Rosa Maza-Quiroga, Miguel A. Molina-Cabello and Karl Thurnhofer-Hemsi

Department of Computer Science, Universidad de Málaga
Biomedical Research Institute of Málaga (IBIMA)

CONTENTS

6.1	Introduction.....	157
6.2	Super-Resolution: Definition.....	158
6.3	Previous Works.....	159
6.3.1	Traditional Methods.....	159
6.3.2	Deep Neural Networks.....	160
6.4	Improved Deep Learning Methods.....	161
6.5	Examples and Comparisons.....	164
6.5.1	Methods.....	164
6.5.2	Datasets.....	165
6.5.3	Measures.....	166
6.5.4	Results.....	167
6.6	Conclusions.....	170
	References.....	174

6.1 INTRODUCTION

Magnetic Resonance (MR) imaging is commonly used in medical procedures and diagnoses since it is a non-invasive technique that provides excellent soft-tissue contrast images. It does not require the use of ionizing radiation, together with full three-dimensional capabilities. Moreover, MR imaging allows performing functional, diffusion, and perfusion imaging.

Magnetic resonance images (MRI) with high-resolution (HR) can provide rich anatomical details, crucial for reliable computer-aided radiological diagnosis and image post-processing. Notwithstanding continuous improvements in the acquisition technology, it is common to detect artifacts in the obtained image. The most

common of these are blurring and the appearance of noise, limiting the quality of the produced images.

Besides the practical and operational limits to the acquisition time, the specific MR technology and the tissues' magnetic properties cause the image resolution to fall in the order of millimeters.

Hardware limitations, high signal-to-noise ratios (SNR), and patient motion (even the occasioned by heartbeat or by the patient's breathing) also limit image resolution. Note that MRI data is usually obtained with different voxel sizes in the typical clinical setting depending on the acquired modality. In particular, the in-plane resolution is usually higher than the out-plane one (i.e., in the slice direction), producing non-isotropic voxel sizes (i.e., rectangular voxels).

There are some recent studies on critical problems due to the use of low-resolution (LR) images. In Mulder et al., (2019), the authors investigated voxel geometry's influence by imaging simulated elliptical structures with voxels varying in shape and size. For each reconstructed structure, the authors calculated and analyzed the differences in volume and similarity between the labeled volume and the ellipsoid's predefined dimensions. As a result, the authors find that larger voxels typical of coarser-resolution images, and increasing anisotropy, end in more significant deviations of both volume and shape measures, clearly demonstrating the anatomical inaccuracies introduced in LR images.

Small but clinically important lesions in the brain may be challenging to visualize or characterize correctly when inspecting low-resolution MRIs.

Besides, the increasing use of functional imaging to examine intratumoral heterogeneity has led to a clinical need for improved spatial resolution for these inherently LR sequences.

Therefore, in some cases, the acquired images need an upsampling to decrease the voxel size and obtain higher resolution images, which will be post-processed or analyzed.

6.2 SUPER-RESOLUTION: DEFINITION

The formation of the LR image from an HR image follows a degradation model $I_{LR} = D(I_{HR}; \omega)$, where the degradation operator D acts on the HR image and whose parameters are represented by ω .

This degradation function D is usually defined as a sequential composition of blurs, downsampling, and the addition of noise. A complete model considers all these choices. The result is the following model (Zhang et al., 2018):

$$I_{LR} = (b * I_{HR}) \downarrow_s + \eta_\sigma \quad (6.1)$$

where $b * I$ represents the convolution between a blur kernel b and a latent HR image I , \downarrow_s is a subsequent downsampling operation with scale factor s , and η usually is additive white Gaussian noise with standard deviation (noise level) σ .

Image super-resolution aims at reconstructing the corresponding HR images from the LR images. We present a formal definition of the problem below.

Let us consider a three-dimensional (low-resolution) image I_{LR} where, without loss of generality, we will assume that voxel coordinates are linearly spaced in a cubic grid with $(0, 0, 0)$ as one vertex and the opposite vertex is given by (D_x, D_y, D_z) . Voxel spacing is characterized by a vector (h_x, h_y, h_z) , where h_s indicates the distance along the s -axis between the centers of two adjacent voxels. This means that voxel coordinates can be expressed in the form $(n_x h_x, n_y h_y, n_z h_z)$ for some $n_x, n_y, n_z \in N$.

The problem of super-resolution consists of determining an image I_{HR} with spacing $(\alpha h_x, \alpha h_y, \alpha h_z)$ for some $\alpha \in (0, 1)$ and such that $I_{HR} = I_{LR}$ in all the voxels where their corresponding grids coincide. The lower the value of α , the finer the resolution of I_{HR} is. The inverse of α , $\frac{1}{\alpha}$, is usually referred to as zoom factor and generally takes integer values. Using the “times” notation, commonly used zoom factors are $2x, 3x, 4x\dots$

Note that this problem is ill-posed (López-Rubio, 2016) since, as it is defined, the solution is not unique. Multiple high-resolution images I_{HR} can be degraded to the same I_{LR} .

6.3 PREVIOUS WORKS

This section aims to give a comprehensive and brief overview of previous works in image super-resolution (SR). The process of generating high-resolution (HR) images from low-resolution (LR) images can be performed using a single image or multiple images. Most of the existing literature surveys are mainly focused on single image super-resolution, which has been extensively studied and reviewed in the following subsections. Additionally, in the next section, we provide an overview of recent advances in SR for higher-dimensional images (such as 3D scans).

SR methods can be divided into two main categories: traditional and deep learning methods. The former methods have been studied for decades, but now they are outperformed by the deep learning-based approaches, which have shown promising results in other fields in artificial intelligence, such as object classification and detection, natural language processing, audio signal processing, etcetera.

6.3.1 TRADITIONAL METHODS

The classical way of obtaining an SR image is by polynomial interpolation, which represents an arbitrary continuous function underlying the discrete samples that make up the LR image. In general, traditional SR methods are broadly classified as techniques based on the frequency domain, interpolation, or regularization, where the last two approaches are based on the spatial domain.

The methods based on the frequency domain are an intuitive way to enhance the details of the images by extrapolating the high-frequency information of LR images. These methods consist of transforming the input LR image(s) to the frequency domain, estimating the HR image, and then transforming back the HR image to the spatial domain. In general, we can divide these methods into two groups according to the transformation employed: algorithms based on Fourier or

wavelet transformation. Detailed explanations of these techniques and a list of different provided approaches in the literature are described in Nasrollahi & Moeslund, (2014).

The approaches based on the interpolation construct an HR image by projecting all the information available from LR images. Usually, this reconstruction process consists of the following steps: image registration, multi-channel restoration, image fusion, and image interpolation. These techniques and the numerous algorithms available in the literature for image interpolation and SR are addressed in (Abd El-Samie et al., 2012), where several chapters are devoted to each stage, including simulation experiments along with the MATLAB code.

Regularization-based methods consist of incorporating the prior knowledge of the unknown HR image by using a regularization strategy. This information can be extracted from the LR images, which is contained in the probability distribution of the unknown signal (HR image). From the Bayesian point of view, the HR image can be estimated by applying Bayesian inference to exploit the information provided by the LR images and the prior knowledge of the HR image. More details about the different regularization strategies provided in the literature and the related references are described in (Tian & Ma, 2011).

In the literature, a great variety of classical SR methods have been proposed, where most of them are surveying at several works (Abd El-Samie et al., 2012; Nasrollahi and Moeslund, 2014; van Ouwwerkerk, 2006; Shah and Gupta, 2012; Tian & Ma, 2011). These surveys also provide different taxonomies covering all the types of traditional SR techniques and comparative discussions of the different methods.

6.3.2 DEEP NEURAL NETWORKS

SR models based on deep learning (DL) have been actively explored in recent years due to the rapid development of DL techniques, which have outperformed the state-of-the-art algorithms on various SR benchmarks. A great variety of DL methods have been proposed in the literature, which can be classified according to the most distinctive features in their model designs (e.g., network architecture, loss function, learning principles and strategies, etcetera.). Additionally, several taxonomies have been recently proposed (Anwar et al., 2020; Wang et al., 2020), covering the recent advances of SR techniques based on DL systematically and comprehensively.

Most of the existing SR works based on DL are focused on supervised learning, i.e., these models are trained with both LR images and the corresponding HR images. A taxonomy for these models grouped into nine categories is proposed in (Anwar et al., 2020), where a comparison between these models in terms of network complexity, memory footprint, model input and output, learning details, type of loss functions, and other architectural differences are also presented.

The Super-Resolution Convolutional Neural Network (SRCNN) was the pioneering work (Dong et al., 2014) using DL techniques that inspired several later works. Basically, it consists of convolutional layers where each layer is stacked together linearly with Rectified Linear Unit (ReLU). The functionality of these layers is different, and it varies from the feature extraction of the first layer to the

aggregation of the features maps to the final HR image of the last layer. This model is trained by minimizing the difference between the reconstructed HR image and the ground truth HR image using Mean Squared Error (MSE). This is the earliest and simplest network design, which can be categorized as linear networks due to the linear network architecture. In this direction, several works can be found in the literature: a fast version (FSRCNN) which improves the speed and quality achieving a real-time rate (24 fps) of computation (Dong et al., 2016), and another fast SR approach using an efficient sub-pixel convolution capable of processing 1080p videos in real-time (Shi et al., 2016).

In addition to the above models, other further improvements have been proposed, which essentially differ in some components such as model frameworks, up-sampling methods, network design, or learning strategies. These works are addressed in (Anwar et al., 2020), which can be divided into several categories, including linear networks, residual networks, recursive networks, progressive reconstruction designs, densely connected networks, multi-branch designs, attention-based networks, multiple-degradation handling networks, and generative adversarial networks (GAN).

In general, most of the state-of-the-art SR models can be attributed to a combination of multiple strategies like the channel attention mechanism, sub-pixel up-sampling, residual learning, etcetera. These models are mostly focused on supervised learning, i.e., learning with LR-HR image pairs. However, in some real-world scenarios, it is not easy to collect images with different resolutions. Thus, datasets are constructed by performing some predefined degradations on HR images (and obtaining the paired LR image). The main drawback of these datasets is that the trained SR models actually learn a reverse process of the predefined degradation instead of the real-world LR-HR mapping. To avoid this behavior, unsupervised SR is needed, where only unpaired LR-HR images are provided for training. A summarized table including some representative models with their key strategies is provided in Wang et al. (2020), where unsupervised SR models are also discussed.

Apart from the above general SR works, there exist some other popular domain-specific applications where SR can serve to advance in those fields. Medical imaging is one of these fields which is rapidly evolving in increased resolution devices, demonstrating the potential of SR research into practical medical applications. Besides the existing SR works in medical imaging (Greenspan, 2009), recent promising advances to apply SR techniques in medical imaging applications are reviewed in the next section.

6.4 IMPROVED DEEP LEARNING METHODS

The enhancement of the results of deep learning super-resolution methods for 3D MRIs can be accomplished in two distinct ways. We consider these alternatives next.

First, it is possible to modify the architecture of the deep network in order to improve its super-resolution performance. There are several possibilities to do this, including the insertion or deletion of neural layers, the modification of the

number of channels in the neural layers, a change in the stride parameter of the convolutional layers, etcetera. A novel strategy to obtain a better performing architecture consists of changing the loss function, as proposed in Thurnhofer-Hemsi et al. (2019) and Thurnhofer-Hemsi et al. (2020a). The standard loss function for super-resolution deep learning networks is the squared Euclidean norm of the difference between the high-resolution version of the original image (the ground truth) and its reconstruction that the network produces as output (the estimation). While the squared Euclidean norm has demonstrated its suitability for many purposes, some improvements can be obtained by employing other alternatives.

A possible alternative is the family of L_p norms. The members of this family are characterized by a parameter p that specifies to which power the absolute value of the error is raised. The value $p = 2$ corresponds to the usual squared Euclidean norm. On the one hand, higher values ($p > 2$) give more importance to extreme values of the absolute error. This is not convenient since there can be a small number of voxels in the MRI with gross errors or irrelevant features, which could have an outsized influence on the learning of the deep network parameters. On the other hand, smaller values ($p < 2$) give less importance to outliers than the standard squared Euclidean norm. Therefore, they are more promising to yield appropriate loss functions for super-resolution since they provide robustness with respect to outliers. Consequently, in Thurnhofer-Hemsi et al. (2019), the usage of L_p norms with $p < 2$ for 3D MRI super-resolution is proposed. It must be remembered that the mathematical definition of norm includes three conditions that must be fulfilled: non-negativity, linearity with respect to a multiplying factor, and the triangle inequality. The last condition is only held for $p \geq 1$. Given these considerations, L_p norms with values of p in the range $1 \leq p \leq 2$ are studied.

The derivation of a learning rule for an L_p norm is based on the definition of a loss function as the sum of the L_p norms of the absolute errors for all samples. Then the partial derivatives of the loss function are calculated with respect to the synaptic weights. Finally, the stochastic gradient descent method is employed in order to obtain a procedure that adaptively modifies the synaptic weights to minimize the loss function.

The usage of an L_p norm with $p < 2$ as the loss function can increase the resilience of the deep network to outliers. Nevertheless, there is the risk that the learning procedure does not pay enough attention to significant errors. In order to reconcile both requirements, it is proposed in Thurnhofer-Hemsi et al. (2020a) that a multiobjective optimization is carried out to learn the synaptic weights of the deep network. This is accomplished by setting two goals to be optimized. The first one is the minimization of the standard squared Euclidean norm, while the second one is the minimization of a suitable L_p norm with $p < 2$. In order to implement this method in practice, the scalarization strategy is advocated to obtain a single loss function out of the two norms. Two variants of scalarization can be applied to this purpose. The first one is weighted sum scalarization (Gass & Saaty, 1955). This is a straightforward approach where the multiobjective loss function is given as the weighted average of the squared Euclidean norm and the L_p norm. Here the weighting factors for the norms must be adjusted from the data. The second

scalarization variant is weighted Chebyshev scalarization (Gong, 2011). This is a more elaborate approach that requires the estimation of an ideal point, which is an ordered pair formed by the minimum of the first norm and the minimum of the second norm, both computed over the full domain of both norms. The multi-objective loss function is then defined as the weighted Chebyshev distance to that point, where again we have two weighting factors that must be adjusted. For the purposes of combining the squared Euclidean norm and an L_p norm, the (0,0) point can be employed as the ideal point since neither of them can be negative. After the multiobjective loss function has been selected, the methodology proceeds as before. In other words, the partial derivatives of the loss function with respect to the synaptic weights are calculated, followed by its minimization by the stochastic gradient descent method.

The second family of methods to enhance the performance of deep convolutional neural network-based super-resolution of 3D MRIs is based on the combination of the output images obtained for shifted versions of the original input image. The rationale behind this strategy is that the behavior of the convolutional network is slightly different for those versions of the input, so that merging the associated outputs may remove the defects which are present in a minority of such outputs. In other words, a consensus image is built from the individual output images. Since the output images are shifted by the same displacement vector as its associated shifted input, it is necessary to undo the shift on each output image prior to the construction of the consensus image.

Two approximations can be distinguished for this shifting and consensus strategy. The first one is presented in Thurnhofer-Hemsi et al. (2018), and it is called random shifting. Here the displacement vector is modeled as a random variable that is uniformly distributed over a cube of possible integer displacements. Then the mathematical expectation of the unshifted output image is estimated over the distribution of the displacement vector, which is to be taken as the consensus final image. This estimation is carried out by averaging the unshifted images for a sample of randomly drawn displacements. As the number of samples tends to infinity, by the law of large numbers, it is known that such estimation converges to the true value of the mathematical expectation. Therefore, the more samples that are considered, the more accurate the consensus image is expected to be. This is because the mathematical expectation of the unshifted output image has no dependency on a specific displacement vector, while the unshifted images depend on the displacement vectors that have been used to produce them. Under this approach, two parameters must be optimized, namely the size of the cube of possible integer displacements and the number of samples.

The second approximation to shifting is given in Thurnhofer-Hemsi et al., (2020b). In this case, an underlying function is assumed to exist that takes the displacement vector as an argument and outputs the high-resolution 3D MRI. This underlying function is defined on the set of possible displacement vectors. In order to approximate the function, a regular lattice of displacement vectors is considered. This means that the procedure has two tunable parameters, namely the pixel stride that defines the distance between two consecutive points in the

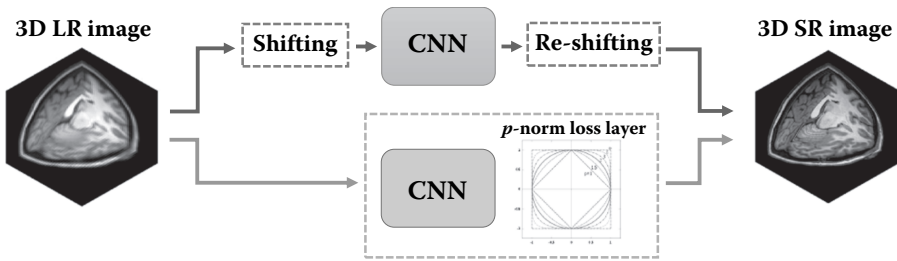


FIGURE 6.1 General overview of the latest super-resolution proposals.

lattice and the side of the cube of points in the lattice that will be considered. After these parameters are determined, the final consensus image is built by averaging the output images corresponding to the displacements that are associated with the points of the lattice within the chosen cube. This way, a zeroth-order approximation of the underlying function is implemented.

As shown above, the L_p norm and the multiobjective optimization approaches involve the proposal of new deep learning schemes, as shown in Figure 6.1. On the contrary, the random shifting and the regular shifting approaches are metamethods that can be applied to any 3D super-resolution deep learning architecture, which implies that they can enhance the results of new architectures as they are proposed.

6.5 EXAMPLES AND COMPARISONS

This Subsection describes several experimental issues such as the methods and datasets that can be employed in the experiments, the measures used to yield the performance of a selected method, or how to assess a fair analysis about the comparison between the competitor methods. Additionally, experiments have been carried out in order to evaluate the performance of each method.

6.5.1 METHODS

As previously mentioned, many algorithms that face the field of super-resolution of 3D Magnetic Resonance images can be found in the literature. Here some of them are presented:

- Spline. Bicubic spline interpolation as implemented in MATLAB (Mathworks Inc.).
- NLMU (non-local means upsampling) (Manjón et al., 2010). A sub-sampling coherence constraint combined with a data-adaptive patch-based reconstruction is used to recover some high-frequency information. It has been written in MATLAB.
- LRTV (low-rank total variation) (Shi et al., 2015). This method employs low-rank regularization and total variation techniques to integrate both local and global information for image reconstruction. It has been developed in MATLAB.

- SRCNN3D (Pham et al., 2017). A three-dimensional convolutional neural deep network has been trained with patches of HR brain images. It predicts a mapping from the LR space to the missing high-frequency components. This method is implemented in Python by using the Caffe package (Jia et al., 2014). A pre-trained network is available on its website. This model has been trained with ten images (in particular, images 33–42) from the Kirby dataset (Landman et al., 2011) over 470,000 iterations. Other parameters of that training process were a learning rate of 0.0001, a batch size of 256, and a momentum of 0.9, while the chosen model optimization was the stochastic gradient descent.
- SRCNN3D+RegSS (Thurnhofer-Hemsi et al., 2020b). A convolutional neural network is combined with a regularly spaced shifting mechanism over the input image. It has been implemented in Python.
- SRCNN3D+RndS (Thurnhofer-Hemsi et al., 2018). This method is a previous version of the SRCNN3D+RegSS approach based on the use of a random shifting technique.
- SRCNN3D-i50K-WCS, SRCNN3D-i50K-WSS (Thurnhofer-Hemsi et al., 2020a). They use a combination of L_p-norms in the loss layer where two multiobjective optimization techniques are employed to combine two cost functions: Weighted sum scalarization (WSS) and Weighted Chebyshev scalarization (WCS). The chosen convolutional neural network is the SRCNN3D method. In this case, that was trained over 50,000 iterations.
- SRCNN3D-i50K-p1.9, SRCNN3D-i50K-p2 (Thurnhofer-Hemsi et al., 2019). They are previous versions of SRCNN3D-i50K-WCS and SRCNN3D-i50K-WSS methods. They are based on the use of a p-norm loss layer to improve the learning process. Two versions with $p = 1.9$ and $p = 2.0$ are used in the experiments. Again, the chosen network is SRCNN3D, and it was trained over 50,000 iterations.

6.5.2 DATASETS

The experiments should be performed on well-known datasets, public if possible. Some datasets which can be found in the literature are:

- Kirby 21 (Landman et al., 2011). In particular, two T1-weighted MRI images of this dataset (10 and 11) were used in the experiments. These data were acquired on a 3-T MR scanner with a $1.0 \times 1.0 \times 1.2 \text{ mm}^3$ voxel resolution over a field-of-view (FOV) of $240 \times 204 \times 256 \text{ mm}$ acquired in the sagittal plane.
- OASIS (Marcus et al., 2007). Specifically, experiments use two T1 images of this dataset (images 1 and 2 of the cross-sectional data). These data were acquired on a 1.5-T Vision scanner with a $1.0 \times 1.0 \times 1.25 \text{ mm}^3$ voxel resolution over a FOV of $256 \times 256 \text{ mm}$.
- IBSR (Worth, 2010). In the subsequent experiments, an image from this dataset is used. The features of this image are a size of $256 \times 256 \times 128$, with $1.5 \times 1.0 \times 1.0 \text{ mm}^3$ voxel resolution.

- CIMES. A T1-weighted image was acquired at the Medical Research Center of the University of Malaga (CIMES) using a 3-T MR scanner with a $0.93 \times 0.93 \times 1.0 \text{ mm}^3$ voxel resolution over a FOV of $256 \times 256 \text{ mm}$.

Images used in the experimental test step must be different from those used to train the methods to establish a fair comparison.

All these datasets are formed by HR images. Due to the aim of described methods is to generate an HR image from an LR image, the HR images from the datasets are downsampled to obtain an LR version of them. The procedure to achieve these LR datasets is as follows. First, in order to avoid fractional values, HR images are cropped regarding the zoom factor to be applied. Then, a three-dimensional Gaussian filter with a standard deviation equal to 1 is applied. Finally, the LR image is generated by carrying out a downsampling method. This method can be an interpolation function such as the nearest neighbor approach, bicubic or bilinear interpolations, among others (Molina-Cabello et al., 2020). Some of these mentioned methods can be used through the `imresize3` function of MATLAB. In particular, a cubic interpolation has been used in the experiments shown subsequently.

6.5.3 MEASURES

In order to compare the goodness of several competitor methods between them, several quality measures may be used to evaluate the performance of each method.

From a quantitative point of view, some quantitative measures may be considered. In particular, in the experiments carried out in this work, three well-known measures have been selected:

- Peak Signal-to-Noise Ratio (PSNR). It is measured in decibels (dB). Higher is better.

$$PSNR = 10 \left(\frac{peak^2}{\|Y - \hat{Y}\|^2} \right) \quad (6.2)$$

- Structural Similarity Index (SSIM) (Wang et al., 2004). It measures the structural similarities between images. Higher is better. It is defined as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6.3)$$

- Bhattacharyya coefficient (BC) (Bhattacharyya, 1946). It is focused on the closeness of the two discrete pixel probability distributions P and \hat{P} corresponding to the ground truth (GT) and restored images with values in the range $[0,255]$:

$$BC = \sum_{j=0}^{255} P(j)\hat{P}(j) \quad (6.4)$$

Other measures may be evaluated, such as CPU time.

Regarding a visual or qualitative point of view, besides the restored image produced by each competitor method, the residual image is also employed in the comparison. This residual image r is computed as the difference in absolute value between the original HR image h and the restored one s :

$$r = |h - s| \quad (6.5)$$

Due to that difference might be close to zero, a darker residual image implies better performance. In the shown experiments, color maps were adjusted to obtain better visualization and discrimination between competitors.

6.5.4 RESULTS

Next, some experimental comparisons between the previously mentioned state-of-art methods are summarized.

First of all, we measure the performance from a quantitative point of view. Thus, employing the MPRAGE, IBSR, and CIMES images, we computed the mean and standard deviation of the PSNR, SSIM, and BC, and the results are presented in Figures 6.2 and 6.3. Within the deep networks, we should differentiate two groups of methods: SRCNN3D and SRCNN3D-i50K. The first has been trained for a larger number of iterations, 470,000, while the second for only 50,000. That's because the authors used a pre-trained model, which required a long time to be completed. Therefore, the results of the SRCNN3D-i50K based methods may not be good enough as the others, so a specific analysis is going to be done.

The first thing the reader can observe is the considerable difference between the deep neural network-based methods (SRCNN3D) and traditional methods (Spline, LRTV, NLMU). In Figure 6.2, where the results for zoom factor 2 are depicted, the general accuracy of the latter is clearly behind the newest ones in the three metrics. The shifting techniques proposed by Thurnhofer-Hemsi et al. (2018, 2020b) reported the best results for PSNR and SSIM and the second best with the BC. Nevertheless, the WCS and WSS-based methods (Thurnhofer-Hemsi et al., 2020a) also reached high PSNR values with a significantly low number of iterations, which means that the training procedure can be improved a lot. Although the SSIM and BC are not the best, they are also good. The best traditional method seems to be NLMU, although LRTV also yielded similar results.

In Figure 6.3, performances are reported for zoom factor 3. Here, the performance of the multiobjective optimization methods is remarkable, especially in terms of PSNR reaching 30dB, and BC, very close to 1. The SSIM of the traditional methods has improved, which means that they reconstruct adequately more complex problems (larger-scale factors), although they are still below in the rest of the measures.

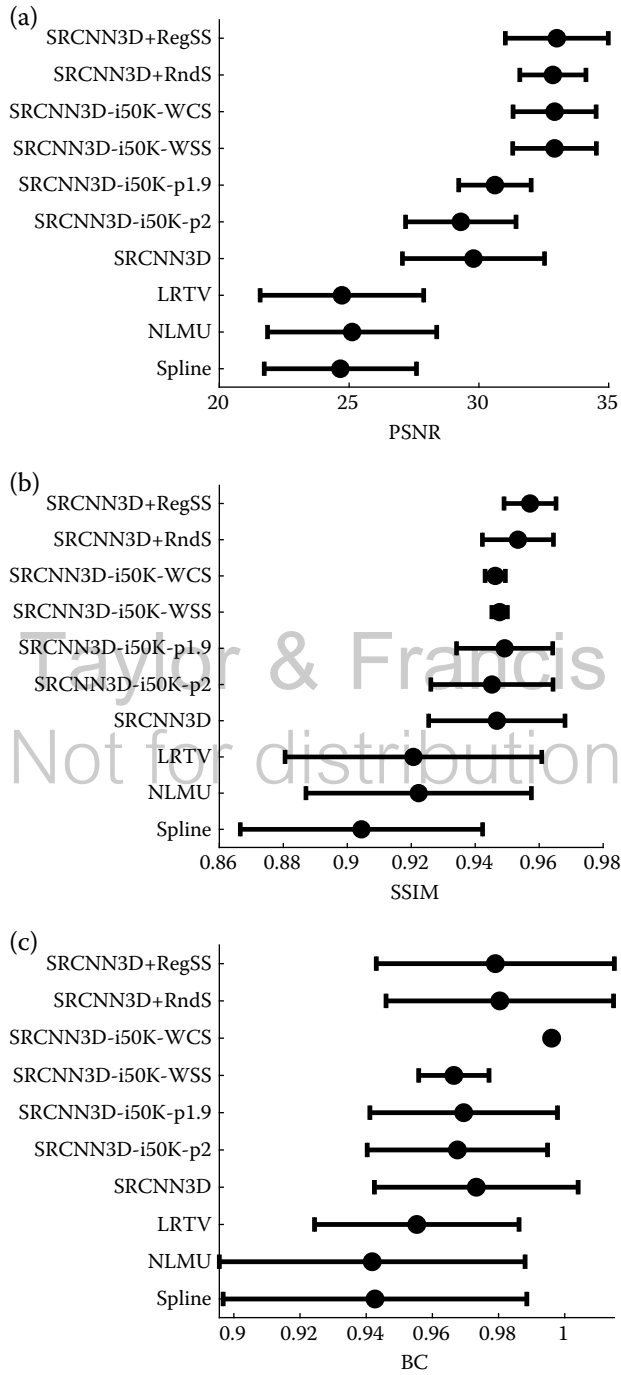


FIGURE 6.2 Comparison of the PSNR, SSIM, BC (higher is better) for the ten methods. Mean and standard deviation of the results for all the test images using zoom factor 2.

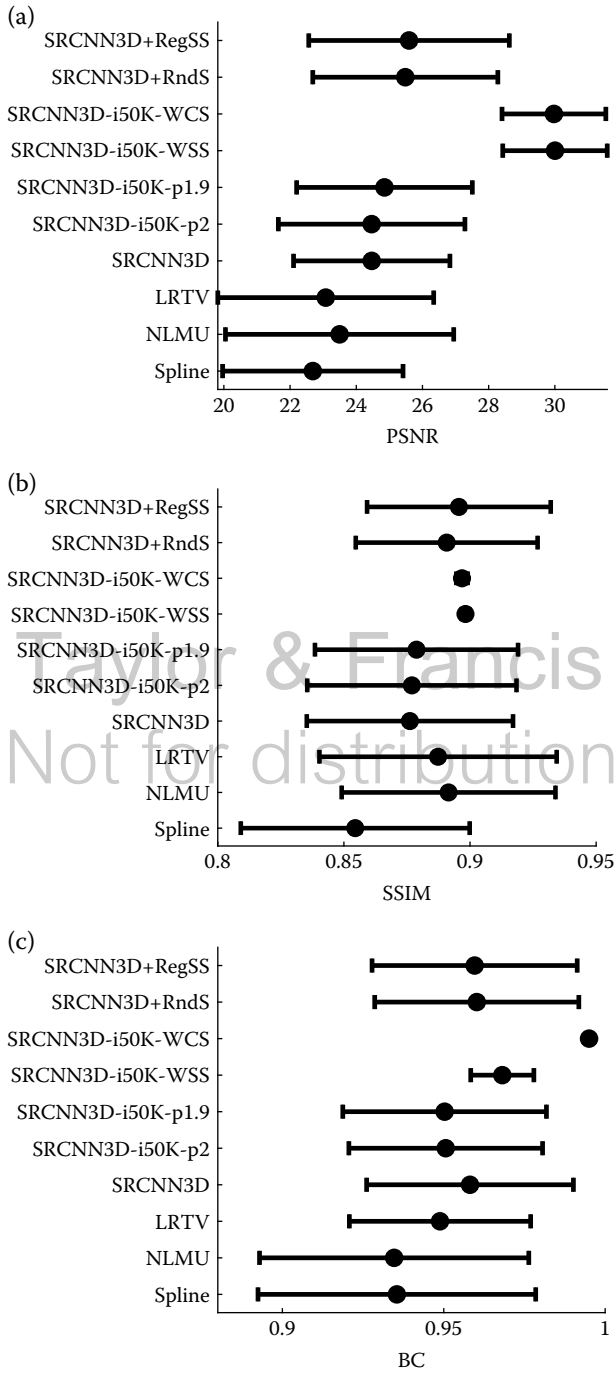


FIGURE 6.3 Comparison of the PSNR, SSIM, BC (higher is better) for the ten methods. Mean and standard deviation of the results for all the test images using zoom factor 3.

If we carry out a detailed inspection of the types of methods, the comparisons of the L_p norm conclude that $p = 1.9$ yielded better results and the traditional Euclidean norm for all cases and reconstruction scales. On the other hand, the regular shifting technique (SRCNN3D+RegSS) is slightly better than the random version, which might be stretchily related to the brain structures. The combination of the shifting method with the L_p norm may produce better results.

From a qualitative point of view, the outcomes of the compared methods are presented in Figures 6.4, 6.5, and 6.6. The outputs of the SRCNN3D-i50K-WSS and SRCNN3D-i50K-WCS methods are not available since both methods were trained with these images but not tested. First, Figure 6.4 depicts a sectional 3D view of the KKI2009-11-MPRAGE T1-weighted image. Restored and residual images are presented, as well as the original HR image and the downsampled one. Spline and LRTV methods generate blurred super-resoluted images, which can be clearly observed in Figures 6.4 (c) and (d). In order to differentiate better the rest of the outcomes, the residual images are more appropriate. With them, the NLMU method can also be discarded, as well as the L_2 norm-based version (SRCNN3D and SRCNN3D-i50K-p2). The best restorations are obtained by the image shifting methods.

Second, Figures 6.5 and 6.6 depict the restored and residual images, respectively, for a coronal section of the CIMES image using a zoom factor equal to 3. The main difficulty of this case is that the number of voxels to be reconstructed is higher, so the differences between the methods are very tiny. In this example, the regular shifting does not provide the best outcome, which is yielded by the L_p norm-based methods. Concretely, $p = 1.9$ seems to be the most accurate. Nevertheless, the random shifting also worked well, but not the traditional techniques.

6.6 CONCLUSIONS

Obtaining HR by LR using single images has been widely considered. On the other hand, approaches in SR for multidimensional images show more promising results. At present, traditional methods of SR are outperformed by the deep learning-based method. SR based on DL shows multiple encouraging taxonomies in last year. DL's pioneering work utilized convolutional networks (SRCNN), and the recent future shows improvements with works that vary in some components like GAN's works.

Many 2D models were developed in the last years, but 3D models are not plenty studied. This chapter reports four recent 3D methodologies that upgrade the current state of the art. Globally, there exist two ways to enhance a model: modify the model architecture and vary the input images.

Two robust three-dimensional super-resolution methods for MRIs were presented based on the use of a p -norm loss layer ($1 \leq p \leq 2$), at first, and combinations of L_p norm cost functions ($p < 2$) using the weighted sum *and* the weighted Chebyshev scalarizations, in the second, alternately of the conventional Euclidean formulation ($p = 2$). Qualitatively, restored images look more refined and with less structural degradation. Future research with deeper neural networks might increase the efficiency with lower values of p . Besides, adding more p -norms might improve quality images. Nevertheless, complex optimization problems must be solved, which consists of the arduous task of choosing p -norm.

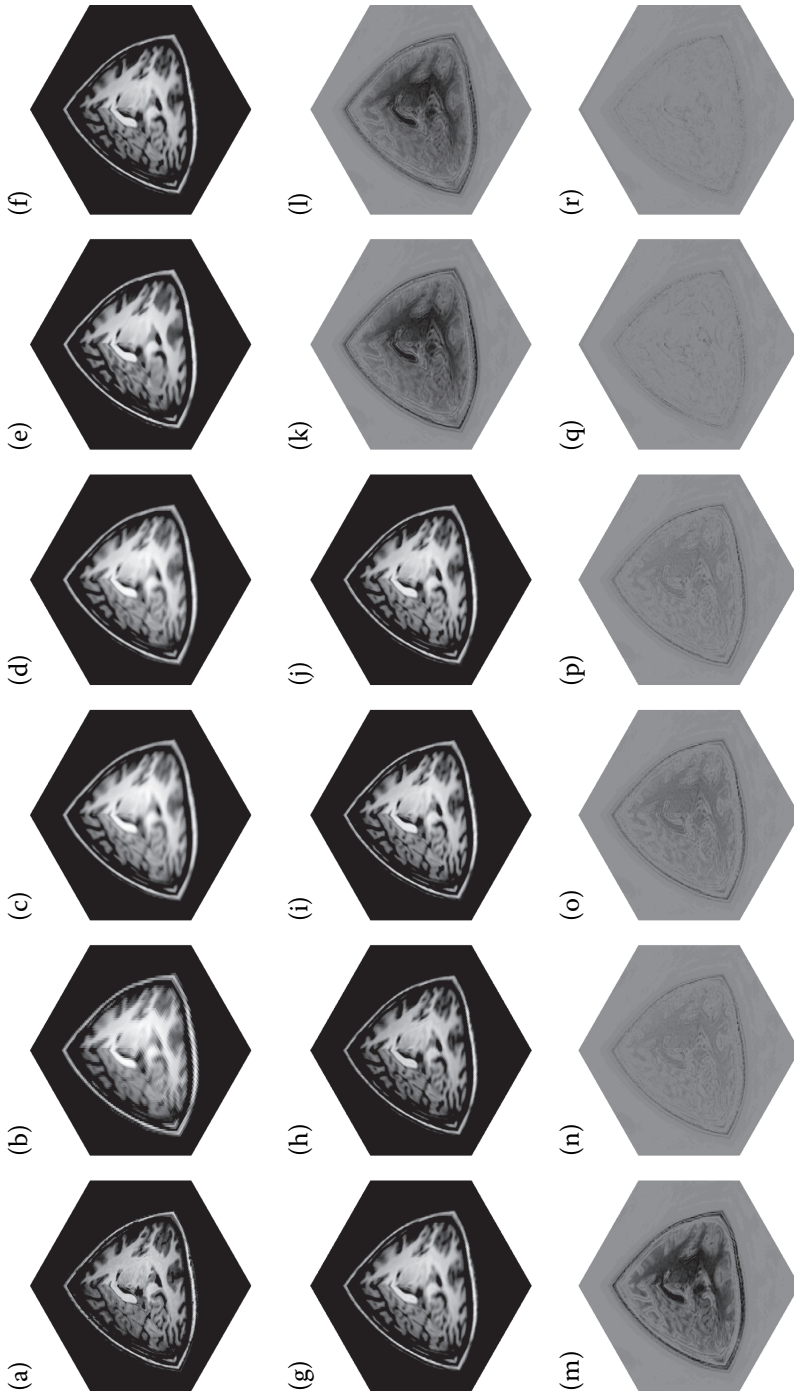


FIGURE 6.4 Qualitative results for KKI2009-11-MPRAGE image for eight methods, applied with zoom factor 2. Three-dimensional images are shown, where the XY plane corresponds to a slice of the axial view, XZ to a slice of the sagittal view, and YZ to a slice of the coronal view.

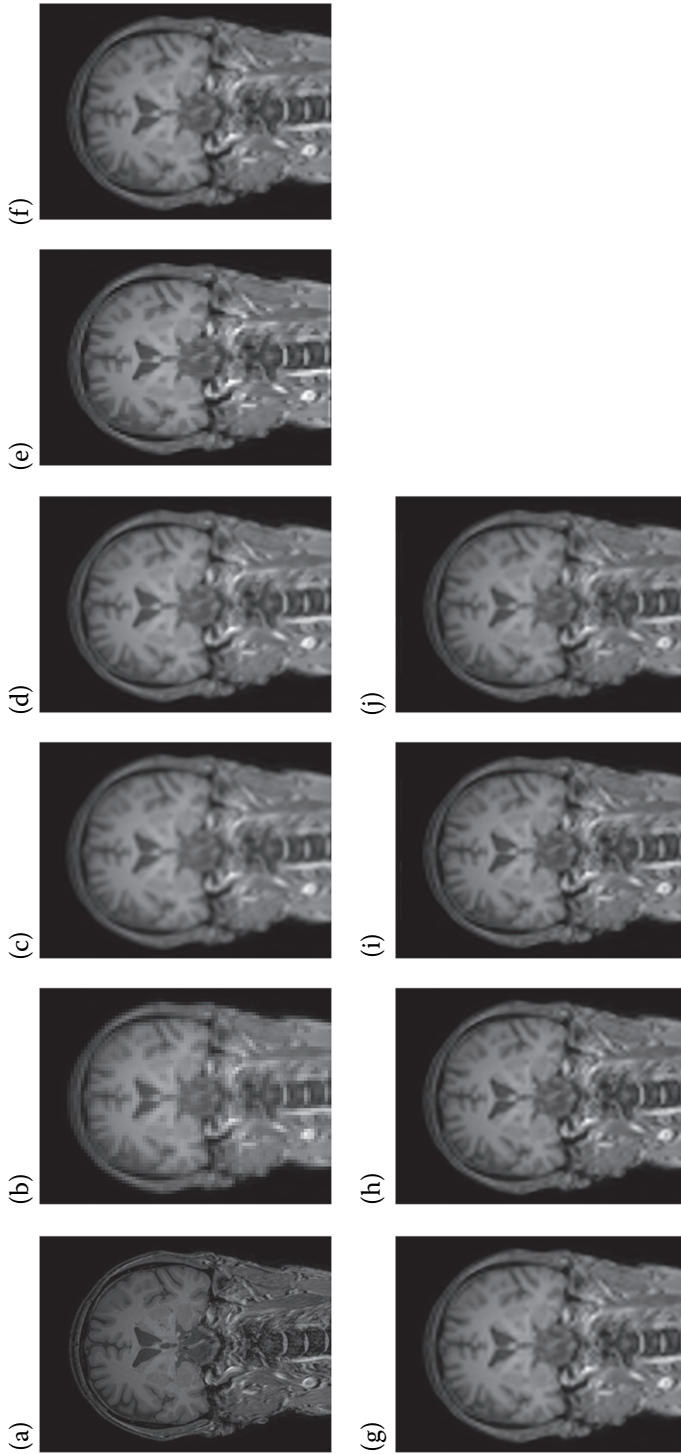


FIGURE 6.5 Restored T1-weighted images from CAMES for eight methods, applied with zoom factor 3. A coronal view is shown. The first row displays the original HR image and the input LR image.

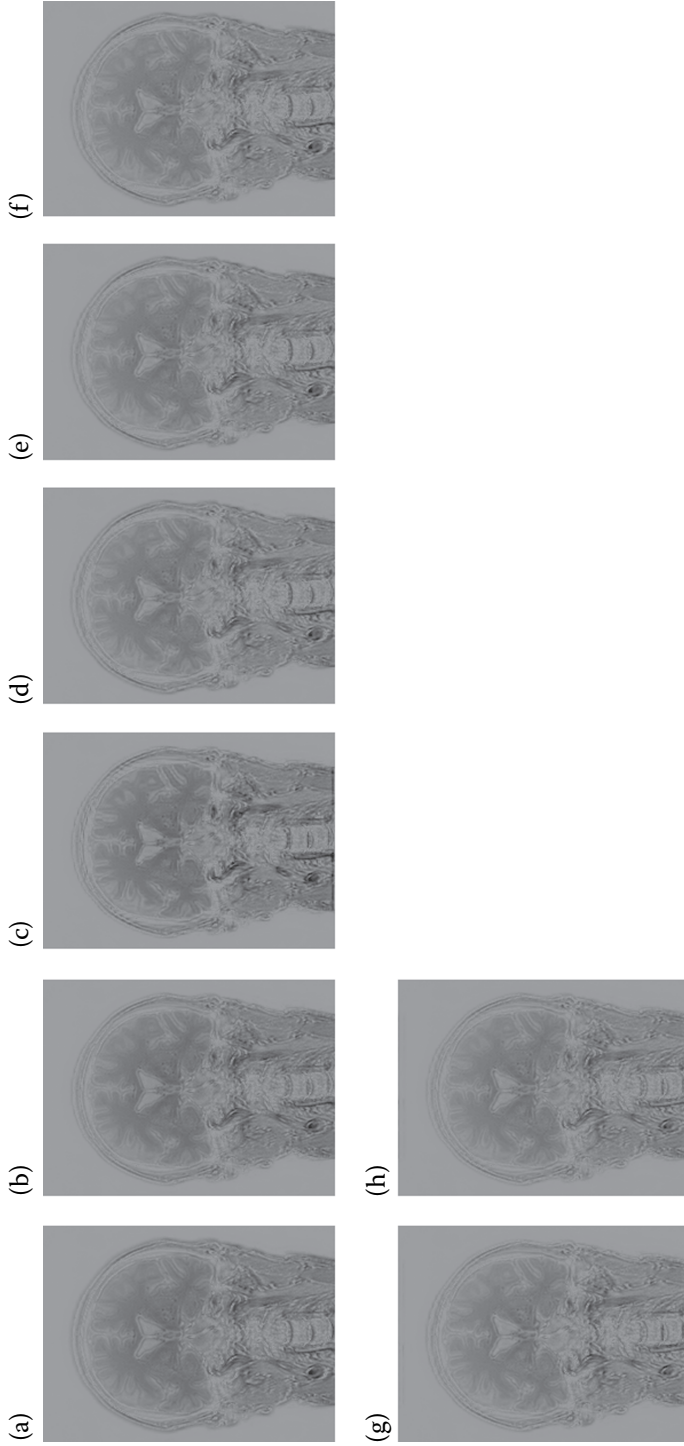


FIGURE 6.6 Residual T1-weighted images from CIMES for eight methods, applied with zoom factor 3. The coronal view is shown.

On the other hand, two MRI super-resolution methods that merge two methodologies have been presented: restoring a low-resolution image by convolutional neural network processing and quality increasing by random, in the first, and regular, in the second, shifting to the input images. The over-smoothing was weathered by using different zoom factors. Blooming results were reached that overcome other state-of-art methods. Future lines of study incorporate the development of additional tuned filtering methods performed on the three-dimensional shift space.

L_p norm and multiobjective optimization generate new DL architectures, and the random/regular shifting approach improves DL results. Machine learning could draw on the applicability of this method. Other neural networks could include these methodologies to improve the quality of the outputs. The line most directly is applying other neural networks to improve the image's quality in widespread tasks like noise removal or segmentation.

The problems of acquiring quality MRI images are widely known, like health budget for a specific machine, with high-quality long MRI sessions fewer patients are intervened, patients' claustrophobic fear of spending a long time in the resonance machine, and some more.

The finding of a methodology that allows improving the quality of an image will reduce healthcare costs and interventions duration and, therefore, increase the number of patients review daily. It supposes an immediate impact both for healthcare, including costs and physicians, and the patient.

REFERENCES

- Abd El-Samie, F.E., Hadhoud, M., & El-Khamy, S. (2012). *Image Super-Resolution and Applications*. CRC Press.
- Anwar, S., Khan, S., & Barnes, N. (2020). A Deep Journey into Super-resolution. *ACM Computing Surveys (CSUR)*, 53, 1–34.
- Bhattacharyya, A. (1946). On a Measure of Divergence Between Two Multinomial Populations. *Sankhyā: The Indian Journal of Statistics*, 7(4), 401–406.
- Dong, C., Loy, C.C., He, K., & Tang, X. (2014). Learning a Deep Convolutional Network for Image Super-Resolution. *ECCV*, 8692, 184–199.
- Dong, C., Loy, C.C., & Tang, X. (2016). Accelerating the Super-Resolution Convolutional Neural Network. *ECCV*, 391–407. https://doi.org/10.1007/978-3-319-46475-6_25.
- Gass, S., & Saaty, T. (1955). The Computational Algorithm for the Parametric Objective Function. *Naval Research Logistics Quarterly*, 2, 39–45.
- Gong, X.H. (2011). Chebyshev Scalarization of Solutions to the Vector Equilibrium Problems. *Journal of Global Optimization*, 49, 607–622.
- Greenspan, H. (2009). Super-Resolution in Medical Imaging. *The Computer Journal*, 52, 43–63.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R.B., Guadarrama, S., & Darrell, T. (2014). Caffe: Convolutional Architecture for Fast Feature Embedding. *Proceedings of the 22nd ACM international conference on Multimedia*, 675–678.
- Landman, B., Huang, A., Gifford, A., Vikram, D.S., Lim, I., Farrell, J., Bogovic, J., Hua, J., Chen, M., Jarso, S., Smith, S., Joel, S., Mori, S., Pekar, J., Barker, P., Prince, J., & Zijl, P. (2011). Multi-parametric Neuroimaging Reproducibility: A 3-T Resource Study. *NeuroImage*, 54, 2854–2866.

- López-Rubio, E. (2016). Superresolution from a Single Noisy Image by the Median Filter Transform. *SIAM Journal on Imaging Sciences*, 9, 82–115.
- Manjón, J., Coupé, P., Buades, A., Fonov, V., Collins, D., & Robles, M. (2010). Non-local MRI Upsampling. *Medical Image Analysis*, 14(6), 784–792.
- Marcus, D., Wang, T.H., Parker, J., Csernansky, J., Morris, J., & Buckner, R. (2007). Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience*, 19, 1498–1507.
- Molina-Cabello, M.A., García-González, J., Luque-Baena, R.M., & López-Rubio, E. (2020). The Effect of Downsampling–Upsampling Strategy on Foreground Detection Algorithms. *Artificial Intelligence Review*, 53, 4935–4965.
- Mulder, M., Keuken, M., Bazin, P., Alkemade, A., & Forstmann, B. (2019). Size and Shape Matter: The Impact of Voxel Geometry on the Identification of Small Nuclei. *PLOS ONE*, 14(4).
- Nasrollahi, K., & Moeslund, T. (2014). Super-Resolution: A Comprehensive Survey. *Machine Vision and Applications*, 25, 1423–1468.
- Ouwerkerk, J. (2006). Image Super-resolution Survey. *Image and Vision Computing*, 24, 1039–1052.
- Pham, C., Ducournau, A., Fablet, R., & Rousseau, F. (2017). Brain MRI Super-resolution Using Deep 3D Convolutional Networks. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 197–200.
- Shah, A.J., & Gupta, S. (2012). Image Super Resolution-A Survey. *2012 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking*, 1–6.
- Shi, F., Cheng, J., Wang, L., Yap, P., & Shen, D. (2015). LRTV: MR Image Super-Resolution with Low-Rank and Total Variation Regularizations. *IEEE Transactions on Medical Imaging*, 34, 2459–2466.
- Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A., Bishop, R., Rueckert, D., & Wang, Z. (2016). Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1874–1883.
- Thurnhofer-Hemsi, K., López-Rubio, E., Roé-Vellvé, N., Domínguez, E., & Molina-Cabello, M.A. (2018). Super-Resolution of 3D Magnetic Resonance Images by Random Shifting and Convolutional Neural Networks. *2018 International Joint Conference on Neural Networks (IJCNN)*, 4008–4015.
- Thurnhofer-Hemsi, K., López-Rubio, E., Roé-Vellvé, N., & Molina-Cabello, M.A. (2019). Deep Learning Networks with p-norm Loss Layers for Spatial Resolution Enhancement of 3D Medical Images. *IWINAC*, 287–296. https://doi.org/10.1007/978-3-030-19651-6_28.
- Thurnhofer-Hemsi, K., López-Rubio, E., Roé-Vellvé, N., & Molina-Cabello, M.A. (2020a). Multiobjective Optimization of Deep Neural Networks with Combinations of Lp-norm Cost Functions for 3D Medical Image Super-resolution. *Integrated Computer-Aided Engineering*, 27, 233–251.
- Thurnhofer-Hemsi, K., López-Rubio, E., Domínguez, E., Luque-Baena, R.M., & Roé-Vellvé, N. (2020b). Deep Learning-based Super-resolution of 3d Magnetic Resonance Images by Regularly Spaced Shifting. *Neurocomputing*, 398, 314–327.
- Tian, J., & Ma, K. (2011). A Survey on Super-resolution Imaging. *Signal, Image and Video Processing*, 5, 329–342.
- Wang, Z., Bovik, A., Sheikh, H.R., & Simoncelli, E.P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13, 600–612.

- Wang, Z., Chen, J., & Hoi, S. (2020). Deep Learning for Image Super-resolution: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10). <https://doi.org/10.1109/TPAMI.2020.2982166>.
- Worth, A.J. (2010). MGH CMA Internet Brain Segmentation Repository (IBSR). <http://www.cma.mgh.harvard.edu/ibsr/>.
- Zhang, K., Zuo, W., & Zhang, L. (2018). Learning a Single Convolutional Super-Resolution Network for Multiple Degradations. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3262–3271.

Taylor & Francis
Not for distribution