International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V71I8P201
Volume 71 Issue 8, 1-12, August 2023

Review Article

Deep Learning in Medical Image Super-Resolution: A Survey

Vaishali Patel[1], Anand Mankodia[2]

[1,2]*U V Patel College of Engineering, Ganpat University, Gujarat, India.*

[1]*Corresponding Author : vbp02@ganpatuniversity.ac.in*

Received: 24 June 2023 Revised: 23 July 2023 Accepted: 06 August 2023 Published: 15 August 2023

Abstract - *Various deep learning (DL) algorithms, specific to convolutional neural networks (CNN), have grown exponentially and have become a methodology of choice for medical image processing. Many parameters degrade the medical images, like hardware limitations, the physical condition of the patient, the relative motion of objects, the limited scanning time of the image acquisition system, and many more. Obtaining medical images with the desired resolution to provide better assistance for the detection and proper diagnosis of the disease is still changeling task while limiting the level of radiation. In this aspect, a super-resolution is not a bad idea to obtain a high-resolution image using the information available in a degraded version of the image. This paper reviews the convolutional neural networks specific to medical imaging. They differ in terms of pre and post-upsampling, the density of CNN layers, and loss functions incorporated in the architecture. This paper provides a summary of DL-based super-resolution algorithms in terms of the quantitative evaluation, loss function and activation function. It also classifies available reference datasets used for medical image super-resolution. It was found that the DL-based medical image super-resolution achieves excellent quantitative and qualitative outcomes. However, this review also uncovers that deep learning techniques are complex in structure, computationally expensive and require large amounts of training datasets. A description of various medical image super-resolution deep convolutional neural networks, current challenges faced in this field, and directions for future work are provided.*

Keywords - *Convolution Neural Network, Deep learning, Medical imaging, Super-resolution.*

1. Introduction

Image Super Resolution (SR) is the subtopic of the Image Processing method, which is used to obtain high-resolution (HR) images from one or more low-resolution (LR) observational images. High-resolution images not only provide pleasing pictures to the viewer but also provide valuable discernible details for subsequent analysis.

Image super-resolution technology is receiving much attraction in various applications such as medical imaging, surveillance, satellite imaging, and identification of biometric information, where high-resolution images are desirable for human interpretation and further analysis. However, the focus of this study is on medical imaging, where an image with the desired resolution is required by a healthcare person for better visualization, detection and diagnosis of diseases.

When a camera captures a continuous scene, it introduces three major artifacts: aliasing, blur, and additive noise. The alias is due to the low sampling rate. This results in a lack of high-frequency image content. High frequencies have edges and texture information, resulting in ultimate artifacts at image edges. Blurring is caused by the relative movement among the image and the digital camera. Atmospheric noises, such as rainy surroundings or dusty air, introduced extra noise in the picture [1].

A schematic diagram of linking LR images to HR images is shown in Figure 1 [24]. Equation 1 can relate the HR image to the LR image.

$$Y = DBMX + n \tag{1}$$

Where,
Y stands for LR images;
D, downsampling matrix for downsampling L1N1×L2N2 HR image into N1×N2 LR image;
B, blur matrix- image point spread function;
M, warping matrix describing motion that occurs during the acquisition of LR image;
X stands for desired HR image;
n stands for additional noise.

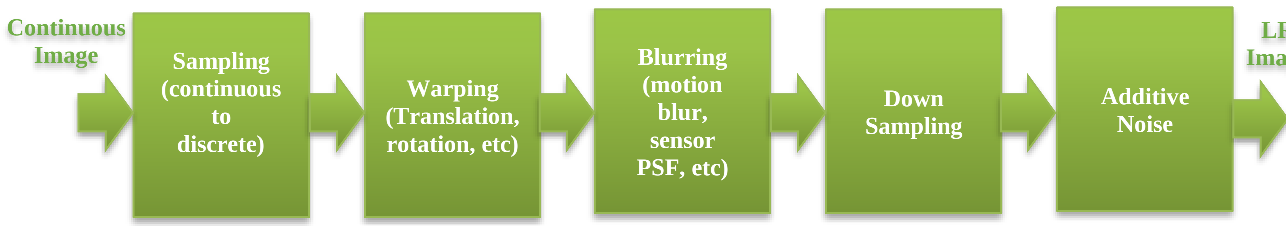


Fig. 1 A schematic diagram linking LR image to HR image

The purpose of super-resolution is to oppose the effect of the downsampling, blurring, and warping that associates the LR image with the favored HR image. Super-resolution methods can be largely categorized into types primarily based on entering input images. Multi-image super-resolution takes more than one LR image captured by distinctive cameras or at a special time of the same scene. The single-image super-resolution method uses the handiest LR image to reconstruct the preferred HR image.

2. Evolvement of Super-Resolution

The problem of super-resolution has been tried to address for a long and as a result of this, there exist several various algorithms in the literature. Interpolation is one of the arithmetically simple and traditional approaches to enlarging the image at the desired scale. Interpolation-based methods are popular for real-time application development due to less computational time and simplicity of implementation [2]. Interpolations such as bilinear, nearest neighbor, bi-cubic, and spline are the simplest. They use weighted sums from neighborhood pixels to estimate the weight of unknown pixels. This technique does not use any prior knowledge and is also non-adaptive, so it encounters a problem of aliasing, blurring, glitches, and ringing effects, especially at the edges. As the research progress, many edge-directed interpolation methods have been proposed in the literature that generates fine edge details [3, 4, 6, 58]. Edge-directed interpolation algorithms segment the image into structure area and texture area. Both structure and texture regions are interpolated separately. A hybrid approach of interpolation with wavelet transform, discrete wavelet transform, iterative interpolation, and convolutional neural network was proposed, combining different approaches to generate super-resolved images [7, 8, 9, 10, 11]. An aggregate of different approaches makes the high-resolution reconstructed image as bright as the original image, with the high-frequency details sharper than conventional algorithms.

Example-based learning is another approach to obtaining a super-resolved image. Pioneer's work on an example-based approach is proposed in [59], which creates a training set containing low-resolution patches and corresponding high-resolution patches. The closest neighbor of the input patch is determined in the low-resolution space, with its corresponding high-resolution patch used for reconstruction. However, the proposed method calls for excessive computational load for looking at the patch. It gives correct outcomes only when the data's resolution or noise degradations shape those of the images to which it is implemented. Numerous SR methods based on this category were proposed in research papers [13, 14, 15, 16, 17, 18]. The performance of this type of algorithm depends on the number of images, characteristics of images present, and size of image patches present in dataset dictionaries. It is also time costly. Due to high time consumption, these techniques are rarely used in real-time applications.

The iterative back projection (IBP) approach is one of the simpler tactics to get SR reconstructed images. In this technique, the HR image is anticipated by back-projecting the error between the simulated and captured LR images [61]. Repeat this manner numerous times to evaluate the HR image by minimizing the cost function and back projecting error at each step. The main advantages of this method are fast convergence and low complexity. This requires less iteration. The major drawback is that back projection is isotropic; hence it projects the error isotropically in the smooth area as well as at the edges. Thus it creates a ringing and chessboard effect. However, to deal with these problems, improved IBP were proposed in the literature. Research papers in [20, 21, 22] used different filtrations techniques like the Bi-lateral filter, Infinite Symmetrical Exponential Filter, and Canny and Gabor Edge detection filters with IBP to treat the smooth area and edges. Noise robust IBP was proposed in [23] and used Principal component analysis (PCA) transform as a cost function to estimate the clean reconstruction and minimize noise's effect. These hybrid approaches of IBP tried to address the problem of the ringing effect and chessboard effect considerably.

Today and in the last decades, Deep learning network has shown popularity in Image processing and so in Image Super Resolution too. To solve the super-resolution problem, various DL-based models have been proposed in the past, some of which have become stepping stones for future research in SR technology. In 2014, the first super-resolution image algorithm based on a convolutional neural network (CNN) was developed. SRCNN (Super-Resolution Convolutional Neural Network) developed by Dong et al.

[31] pioneered the development of CNN-based algorithms. Therefore, many researchers have continuously improved this CNN-based algorithm through various methods to this day. All options presented depend on the type of loss function used, the upsampling module deployed, and the network design strategy employed. A wide range of deep learning algorithms have been used to remedy SR problems, ranging from the early Convolutional Neural Networks. (CNN) approaches to recently promising SR approaches based on Generative Adversarial Networks.

3. Deep Learning Architecture

Deep learning has made great leaps forward in image processing tasks such as image segmentation, classification, detection and the resulting super-resolution of images. DL approaches have powerful representation capabilities and can use multilayer artificial neural networks (ANN) to explore high-level data features. Figure 2 represents a shallow three - layers ANN network. The primary layer is called the input layer; the last layer is referred to as the output layer, and in among is referred to as the hidden layer. There is only one hidden layer here.

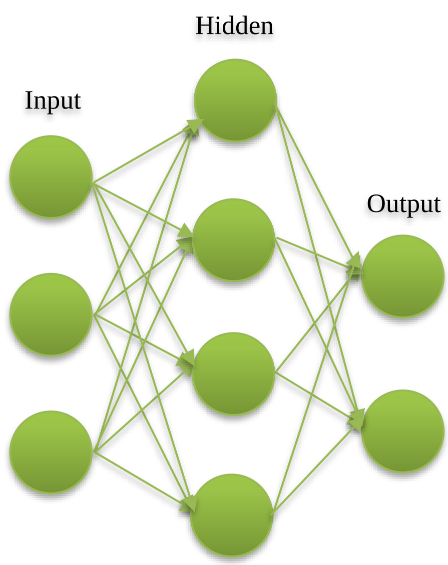


Fig. 2 Illustration of Shallow ANN

Hidden layers can be more than one to make the network deep for better feature extraction. For example, a 5- layers DL network has 3 hidden layers. Each neuron in the network consists of a linear transformation followed by a point activation function [25]. Many DL algorithms have been proposed in the literature, including Convolutional neural network (CNN), recursive neural network (RNN), residual network (ResNet), dense networks, auto-encoders, generative adversarial networks (GAN) and many more [26,27,28]. A recursive network repetitively uses the same module with a shared weight matrix to predict output from the given input.

This method incorporates repetitive fields without requiring additional parameters to estimate high-level features. Image Super-resolution is a task to transform the input image into a target image, where the target image is highly correlated with the input image. Using this fact, the residual network mainly learns the residuals between LR and HR images via skip connection and high-frequency content is learned through the main path of the network. The dense network uses the previous layer all feature maps generated in the previous layer using dense blocks as inputs of its preceding layers.

Auto-encoder and GAN are unsupervised neural networks for image processing. A deep auto-encoder consists of two symmetrical deep-belief networks that act as encoders and the decoders representing data. Auto-encoder converts high-dimensional information into low-dimensional information. As such, the auto-encoder is particularly useful for noise removal, characteristics and feature extraction, compression, and comparable tasks. A GAN-based neural network has two networks named generator and discriminator. The generative network estimates the new pixels from LR image pixels, while the discriminative network distinguishes between estimated super-resolved image pixels and original image pixels.

CNN is one of the most effective DL architectures, which learns features itself widely used for image classification, retrieval, reconstruction and detection tasks as the accuracy of results generated is very high. Several variations are introduced in the literature; however, a simple architecture is designed using a set of hidden layers for feature extraction. The hidden layers are grouped and named as per the task but mainly as convolution and pooling modules. The convolution module convolved the image with the mask to extract features. The pooling module reduces the number of parameters and selects only needed features to produce output which results in a reduction in memory requirement, and processing becomes faster.

3.1. Deep Learning based Super-Resolution Algorithm

The SRCNN model is the first to apply a DL-based algorithm to address the inverse problem [31]. SRCNN is one of the simplest algorithms shown in Figure 3, combining three operational functions: feature extraction and their representation, nonlinear mapping, and reconstruction. As a part of preprocessing in this model, LR images are first resized to the favored length using bicubic interpolation. The first patch extraction layer extracts patches from the LR image, representing every patch as a multidimensional vector. The bicubic interpolated image is convolved with a 9x9 kernel, expanding the three-channel image into 64 feature maps at the first layer.

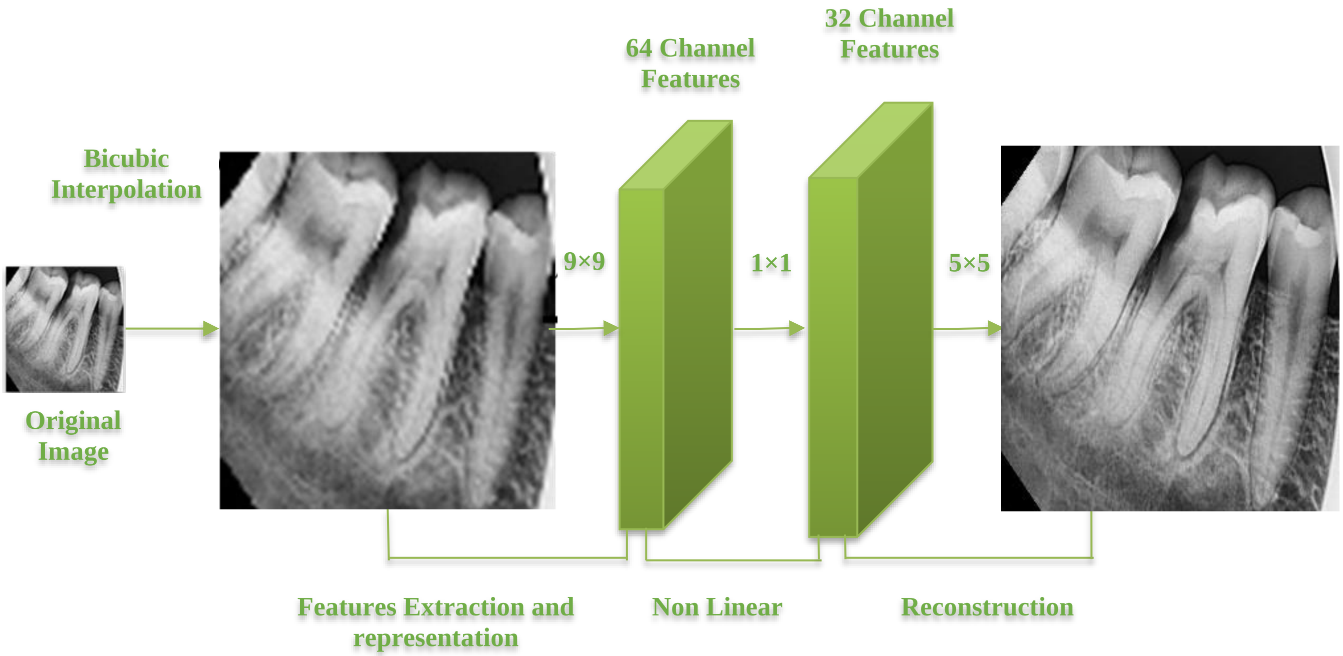


Fig. 3 Basic block diagram for DL based Image Super-resolution

Nonlinear mapping maps each multidimensional vector onto another multidimensional vector nonlinearly. A second layer applies a 1x1 kernel and generates 32 feature maps. Each multidimensional vector is conceptually an illustration of the HR patch to be used for reconstruction. At last, all the reconstructed HR patches are aggregated to shape the very last HR image. The last layer applies a 5x5 kernel to generate the output image. A ReLU activation function accompanies both the first and second convolutional layers.

As shown in Figure 3, SRCNN algorithms for image super-resolution generate the SR image the same size as the ground reality image. Therefore, preprocessing or upscaling is performed on the image before feeding it into the DL network. Depending upon the upsampling operation used in the model, Deep learning Super-resolution methods can be roughly divided into four categories; Pre-upsampling, Post-upsampling, Iterative Up-and-Down Sampling and Progressive upsampling Super-resolution.

4. Deep Learning Super-resolution for Medical Imaging

Super-resolution algorithm aims to achieve high-quality images by increasing image resolution so that greater details are clearer for clinical usage and accurate dis-ease diagnosis in medical imaging modalities along with computed tomography, magnetic resonance imaging and X-Rays (radiographic) imaging. Several factors affect the spatial resolution of an image depending on the modality. Super-resolution algorithms can increase image resolution without changing the imaging protocols. This segment explores the usage of deep learning networks exceptional for super-resolution (SR) and their applications in medical imaging.

4.1. Image Super-Resolution Techniques for Medical Imaging

"Ming Zhao et al." presented a super-resolution algorithm for visually generating better cardiovascular ultrasound images using a Generative Adversarial Network (LSRGAN) with a Laplacian Pyramid framework. Traditional GAN uses a sigmoid cross-entropy loss function in the discriminator. This method uses the least squares loss function for the discriminator. The method generates Cardiac MRI images with high-frequency detail. However, the network structure is very complex, and the reconstruction speed is slow. [29].

"Heng Liu et al." have proposed a GAN network in which cycle loss is calculated. Additionally, it includes pixel loss, perceptual feature loss, and adversarial loss with cycle consistency loss to ensure that the generated image is together with all details. It can maintain perceptual consistency with corresponding LR images [30]. Ultrasound images: CCA-US and US-CASE were used, which gave better results compared to SRCNN [31] and SRGAN [32].

"Ahlad Kumar et al." proposal for image super-resolution architecture based on deep learning in the domain of Tchebichef transform. It has been observed that the task of using the Tchebichef transform domain for SR exploits both the high and low-frequency representations of the images, simplifying the super-resolution task. This method improves the quality of X-ray and CT images of COVID-19, provides better image quality and is useful for clinical diagnosis [33].

"Jin Zhu et al." presents a technique at any magnification scale for super-resolution of medical images. This method combines meta-learning with generative adversarial networks for rendering medical images at any required scale to increase resolution. Brain images from different datasets were evaluated. This technique achieves similar constancy performance and the best perception quality at the cutting edge of technology [34].

"Kuan Zhang et al." presents a perceptually tuned GAN network. This framework optimizes Super-Resolution on 3D MRI datasets using perceptually adapted generative adversarial networks (GANs) and generates thinner high-resolution slices [35].

Faezehsadat Shahidi proposed a super-resolution algorithm for breast cancer histopathology images. This network is equipped with enhanced wasserstein gradient penalty and perceptual loss with GAN instead of residual block, improved wide residual blocks and plain self-attention layer before up-sampling is used in the generator network. In comparison with bicubic, SRGAN and A-SRGAN, the model achieved finer detail and promising results in terms of quantitative metrics. Two different methods are applied, weight and batch normalizations. The model is best suited for weight normalization [36].

"Defu Qiu et al." proposed a deep learning network with a couple of residual networks for the super-resolution reconstruction of medical images. Multiple residual blocks are connected via multiple jumps to build multiple enhanced residual modules to ensure the maximum gathering of deep features from adjacent convolution layers. A deep residual network connected by multi-level jumpers is trained by stochastic gradient descent. Additionally, an adjustable learning rate strategy is used to train super-resolution reconstruction. Common datasets such as Set 5, Set 14, and Urban100, Head, Brain, and Knee were evaluated [37].

"Hongbo Zhu et al." proposed deep generative adversarial network (GAN) architecture based on a deep grammar model called Functional-Realistic GAN (FRGAN). The logical features are collected using a region proposal network. The LUNA16 dataset of the lung image database consortium (LIDC) and the self-collected dataset were used for evaluation. The architecture is too complex, leading to higher computational and storage costs [38].

"J. Andrew et al." introduced an advanced autoencoder-based deep learning technique for brain magnetic resonance images. Besides the usual autoencoder, the proposed autoencoder model additionally includes convolution layers and feature activation functions for feature extraction using skip connection and a deconvolution layer for up-sampling [39].

"Zhaoyang Songa et al. achieved a progressive super–resolution of COVID-CT images using a back-projection network. The progressive back projection network has two levels: a residual up-projection and a down-projection stage. Each level is composed of back-projection, deep feature extraction and up-scaling modules. The role of Up-projection and Down-projection residual modules is to reconstruct error. The residual attention module extracts deep features to build a super-resolution model [40].

"Farah Deeba et al." proposed enhanced medical image super-resolution based on wavelet. It is similar to the SRCNN three-layer structures, but wavelet coefficients are used as the CNN input instead of image patches. X-ray, dental, abdominal and knee images are used as datasets [41].

"Yan Lv et al." proposed a deep network to resolve retinal images super using Dense and ReZero residual networks (DRRN) [42]. First of all, densely coupled modules are used to reap extra feature statistics. After that, local and global residual learning is used to learn features to improve the resolution of retinal images. Retinal image datasets like STARE, DRIVE, and CHASE_DB1 are used to evaluate the algorithm. This technique gives good outcomes compared to SRCNN [31] and FSRCNN [43].

"S. Zhang et al." proposed a deep learning network-based speedy medical image Super - revolution method [44]. The hidden layer consists of three convolution layers. Conv1 is the first layer of the hidden layers, and Conv3 is the third layer of the hidden layers. Both contain the same number of convolution kernels of the same size and have the same Tanh activation function—the middle of the mini-lattice, involving two cascaded convolutions with the ReLU activation function involved. Images of the brain, abdomen, knee and cell were used. This method uses exclusive activation functions in the same structure, and the results show that the Tanh activation function characteristic is higher than ReLU activation.

"Lina Zhang et al." demonstrated a GAN-based SR network for medical images [56]. With the benefits of GAN, deconvolution techniques are employed to recover HR representations from the LR images rather than the straightforward bilinear interpolation methods.

4.2. Medical Imaging Source Dataset for Image Super Resolution

In this study, it was found that there are different types of image datasets available in the literature. They all differ in terms of the types of images a dataset contains, image resolution, image quality, and image size. These datasets are broadly classified as common datasets and specific field datasets. Common datasets generally contain images like babies, birds, people, animals, buildings, plants, flowers and many more natural images. Some publically available benchmark datasets are reviewed in [51, 45]. Specific field dataset contains images specific to particular research filed, such as remote sensing datasets, thermal image datasets, face image datasets, underwater images data and medical image datasets.

Deep learning researchers working to solve the inverse problem with respect to super-resolution specific to medical images are having great challenges with enough datasets to train deep learning systems. However, some medical imaging datasets are mentioned in Table 1 that researchers in the field have utilized during their experiments. Table 1 shows the datasets identified during this study.

4.3. Evaluation Metrics, Loss Functions, Activation Functions and Comparative Study

4.3.1. Evaluation Metrics

Image quality can be assessed by subjective (qualitative) or objective (quantitative) methods.

Methods that use human observers to estimate image quality are called subjective image quality estimation methods [46]. The Subjective image quality is related to the characteristics of human visual perception.

Since this method is performed on human subjects, it can accurately and more reliably estimate the visual quality of images, but it is time-consuming, expensive and practically it is difficult to incorporate into image processing algorithms for optimization [62].

The main purpose of the objective evaluation is to obtain a quantified value that appropriate measurements will generate. The objective evaluation method is to use computational techniques to reflect the visual perception of the human eye and give numerical evaluation results.

Objective quality assessment metrics measure the similarity between perfect quality image (original, ground truth) and reconstructed image (Super-resolution image).

The maximum value of peak signal-to-noise ratio (PSNR) is the most popular statistic for measuring the quality of the super-resolved image obtained through super-resolution algorithms [48]

Table 1. Dataset specification identified during this study

Publication	Dataset Specification
Zaho et al. [29]	Cardiac magnetic resonance images: 2560 images of size 1024×1024
Heng Liu et al. [30]	B-mode ultrasound images of the common carotid artery (CCA-US): 84 images and US-CASE:125 ultrasound images of liver, heart and mediastinum
Ahlad Kumar et al. [33]	Chest X-ray and CT images of COVID-19: 646 images of variable size. Dataset is continuously updating.
Jin Zhu et al. [34]	Early-stage Alzheimer's Disease (OASIS-brains): 416 images of size 176×208×176 Brain tumour segmentation dataset (BraTS): 210 images with glioblastoma (GBM/HGG) and 75 images with lower grade glioma (LGG) of size 240×240×155 Automated Cardiac Diagnosis Challenge (ACDC): 150 images of various sizes various, from 174×208 to 184×288 COVID-CT dataset: 632 images of size 512×512
Kuan Zhang et al. [35]	3D MRI dataset of size 256×256×1170
Faezehsadat Shahidi [36]	Breast Cancer Histopathological Image Classification (BreakHis): 7909 images of various magnification factors like 40×, 100×, 200×, and 400× Cancer Metastases in Lymph Nodes (Camelyon): 400 complete slide images of size 218000×95000
DefuQiu et al. [37]	Head, Brain, and Knee datasets
J. Andrew et al. [39]	Brain tumor segmentation challenge (BraTS2017):285 images of size 256×256 ×3 MRBrain18: 31 images of size 240×240
Zhaoyang Song et al. [40]	COVID-CT: 349 images of different sizes
Farah Deeba et al. [41]	Shen-zhen Hospital Chest X-ray: 662 images of various size Montgomery County X-ray: 138 of size either 4020×4892 or 4892×4020
Shengxiang Zhang [44]	Brain, Abdomen, Knee, and Cell images from I Do Imaging (IDI)
Yan Lv et al. [42]	STARE: 400 Fundus images of size 605×700 DRIVE: 40 Fundus images of 565×584 CHASE_DB1: 28 Fundus images of size 999×960 HRF: 15 Fundus images of size 3504×2336

X is the original image (Reference Image/Ground Truth Image), Y is the super-resolution image, and the size is M × N. Where M stands for the height of the image, and N stands for the image's width. Mathematically, the PSNR is given by the following equation (2).

$$\mathrm{PSNR} = 10 \log_{10}\left(\frac{\mathrm{L}^2}{\mathrm{MSE}}\right) \tag{2}$$

Among them, L is the dynamic variety of image pixel intensities; for a grayscale image with a depth of intensity of 8 bits/pixel, $L=2^8-1=255$. Higher PSNR values generally indicated better image quality.

The human visual perception system is very capable of recognizing structural information from images. Therefore, it is possible to distinguish between sample and reference images using a matrix miming this behaviour. The structural similarity index (SSIM) measures the structural similarity between images based on three independent image characteristics: brightness, contrast, and structure [49, 63]. SSIM is expressed by equation (3). SSIM is 1 when two images are identical, so values close to 1 would generally indicate high reconstruction quality.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

$$C1 = (K1L)^2$$
$$C2 = (K_2L)^2$$

L is the dynamic range of pixel values (8-bit grayscale image is 255)
K1, K2 <<1 are approximately about 0.01 and 0.03.
μx and μy represent, respectively, the mean intensities of the original and reconstructed images.
σx and σy are the variances of original and reconstructed images, respectively.

4.3.2. Loss Functions

The loss function is an integral part of the DL network used to minimize reconstruction error to obtain an optimized solution. Simply, it measures how algorithms react while training them on different datasets. Broadly, loss functions are categorized into two types: (a) Classification loss functions: used in DL networks addressing the problem where the answer comes in the form of the probability of the input data falling in certain pre-defined categories and (b) Regression loss functions: used to optimize reconstruction error in DL network which is predicting continuous value. Basically, Image Super-resolution is a regression problem that involves predicting a ground truth image continuously. Let us get a brief idea of loss functions.

L2 loss function/ Mean Squared Error (MSE)

The MSE is calculated by averaging the squared differences between the super-resolved image and the source image given by equation (4).

$$\mathrm{MSE} = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} \left(r(x,y) - f(x,y)\right)^2}{mn} \tag{4}$$

Here, f and r are the super-resolved and source images, respectively. L2 loss is good for the model to achieve a high signal-to-noise ratio and is often used for SR due to its simple mathematical calculation. MSE is a quadratic function which always guarantees that gradient descent will converge to a global minimum. However, the error is quadratic, so large differences are heavily penalized compared to small differences, which deteriorates the network's performance [51, 52].

L1 loss function/ Mean Absolute Error (MAE)

MAE is calculated by averaging the absolute differences between predicated and original values, as given by equation (5).

$$\mathrm{MAE} = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} |r(x,y) - f(x,y)|}{mn} \tag{5}$$

This shows that the L1 loss is equivalent to degenerated likelihood function, which removes randomness from the learning process. The L1 loss function is more robust and does not affect outliers [55]. The L2 loss function, on the other hand, tries to fit the model to these outliers and is, therefore, very sensitive to outliers values in the dataset.

Mean Bias Error (MBE)

MBE is calculated by averaging the actual differences between predicated values and original values, not absolute values, s as given by equation (6)

$$MBE = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} (r(x,y) - f(x,y))}{mn} \tag{6}$$

Here, positive and negative errors cancel each other, which is why it is one of the lesser-used loss functions in the DL network.

Perceptual Loss Function

The PSNR of feedforward convolutional neural networks is typically increased by training them with a per-pixel loss between the source and output images. However, the loss per pixel does not capture the difference in perceptual between the output image and the source image. This type of loss function is used in [29, 30, 33, 39, 40, 42] for different medical imaging modalities. The difference between high-level visual feature representations recovered from pre-trained CNN is recently used to calculate perceptual loss functions [32, 54]. HR images are generated by means of minimizing a loss function. This approach has been carried out in [34, 35, 36]. These methods produce high-quality images, however gradual to build.

4.3.3. Activation Functions

The Activation function is one of the important blocks of NN. The vary name of activation function suggests that they determine whether a neuron should be activated or not. Since images are highly nonlinear functions, different types of activation functions are used in NN designed for image processing to add nonlinearity.

Table 2 lists the different activation functions with their specifications. For DL-based super-resolution, rectified linear unit (ReLu) and its variants gain popularity due to their simple implementation and remarkable performance.

There is no empirical rule for choosing an activation function. The choice of the activation function is based on the characteristics of the problem that are going to be addressed for easy and quicker convergence of the network. The ReLU activation function is widely used in SR.

This section discusses evaluation metrics, activation functions, and loss functions for DL-based Super-resolution algorithms. A comparative study of various medical image super-resolution algorithms is given in Table 3.

5. Discussion

With the advancement of DL technology in the field of super-resolution, researchers have applied various existing DL architectures to the application of medical image super-resolution. Later, researchers developed new architectures to improve the quality of different medical imaging modalities.

5.1. Super-Resolution Medical Imaging Current Challenges

Despite using preprocessing and different augmentation techniques to generate datasets, the lack of high-quality reference image datasets is one of the recent challenges for SR deep learning methods in medical imaging.

Obtaining ground-truth HR medical images is challenging due to the various modalities' specific limitations.

As discussed in the previous section, finding an appropriate loss function which results in quick conversion as well as improves perceptual quality is also a big challenge for medical SR.

The performance of a DL network depends not only on the depth of the network, padding techniques, dropout nodes, and many layers but also on how well they are combined. Haphazardly increment in the depth results in enormous calculations, excessive training time, and memory consumption.

One main challenge observed is the integration of medical image super-resolution algorithms into the real-time clinical environment. Efforts can be made in the future to address the existing shortcomings for improved estimation of HR image is described in the next part.

5.2. Future Trends

Limitations of existing algorithms can be mitigated by adding twists or designing new algorithms to improve existing techniques. Some directions for future work that can be incorporated into medical image super-resolution algorithms are given in the following lines.

Table 2. Activation function used in DL-based algorithm of Super-resolution

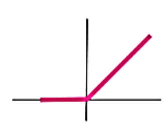
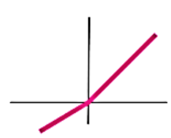
Activation Function	Formula	Range	Shape	Remark
Rectified Linear Unit (ReLu)	$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$	0,+∞		It helps with better prediction and less overfitting because not all of the neurons are activated at the same time. Unlike tanh or sigmoid, it does not experience a vanishing gradient. Computationally effective, requiring less training time and achieving convergence more quickly.
Leaky ReLu	$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases}$	-∞,+∞		This is an upgraded version of ReLu that tackles the issue of dying neurons by incorporating a modest slope with enabling negative values. It is simple and effective for computing.
Parametric leaky ReLU (PReLU)	$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases}$	-∞,+∞		PReLU) is a variation of Leaky ReLU, where α is not a hyperparameter but becomes a parameter that can be modified by backpropagation like any other parameter.

Table 3. Comparative analysis of existing algorithms for Medical Image Super-resolution

Publication	SR Method	Loss function	Scaling Factor	PSNR	SSIM
Zaho et al. 2020 [29]	LPGAN: Laplacian Pyramid based on Generative Adversarial Networks	l2	2	36	0.94
			4	31	0.84
			8	27	0.7
Heng Liu et al. 2021[30]	CycleGAN: self-supervised learning strategy to get LR-HR pairs	l1	4	32.491/ 35.222	0.876/ 0.919
Ahlad Kumar et al. [33]	Tchebichef transform domain based deep learning architectures	l2	2	43.43	0.9806
			3	40.79	0.9681
			4	39.83	0.9351
Jin Zhu et al.[34]	MIASSR: Meta-learning using GANs	Pixel-wise L1 loss, adversarial loss, and perception loss	Arbitrary scales (1, 4)	36.46	0.95
Kuan Zhang et al. [35]	SOUP-GAN	Perceptual loss	2	26	0.96
			3	23	0.92
			4	21	0.87
Faezehsadat Shahidi [36]	WA-SRGAN: wide residual blocks and self-attention layer	Pixel-wise L1 loss, adversarial loss, and perception loss	2	28.74	0.97
			4	26.19	0.79
			8	21.55	0.49
DefuQiu et al. [37]	MIRN: improved residual network	Euclidean Loss	2	36.220	×
J. Andrew et al. [39]	Autoencoder-based SR approach	l2	2	29.023	0.884
Zhaoyang Song et al. [40]	CNN: Progressive back-projection network	l1	4	28.31	0.783
Farah Deeba et al. [41]	Wavelet-Based: medical image super-resolution	Euclidean Loss	2	27.32	0.856
			3	25.1	0.843
			4	23.67	0.829
Shengxiang Zhang [44]	FMISR: a fast method for super-resolution medical images based on deep learning networks	Euclidean Loss	3	29.368/ 0.240	×
Yan Lv et al. [42]	DRRN: dense residual networks	l2	3	46.91	×
			4	43.74	

Finding the most appropriate loss function for super-resolution medical images is of great importance for improving the model's overall performance.

In future work, researchers may try to remodel and optimize the loss function to improve the network structure to speed up the network reconstruction.

Most of the current algorithms use spatial-domain information to obtain a super-resolved image. Future work can be extended by exploring different transform domains and proposing a hybrid approach for image super-resolution.

Deep learning networks require multiple paired datasets for training. However, many are not publicly available, which restricts the research efforts. Researchers should share datasets publically for healthy research work to extend the application of medical image super-resolution in a clinical setting.

6. Conclusion

The central thought is to provide an extensive overview of the field of medical image super-resolution. This article discusses the details of various image super–resolution techniques from beginning to end. DL-based medical image super-resolution is a fast-growing field with numerous applications in different medical imaging modalities. There is always a trade-off between performance and network complexity for medial image super-resolution. A comparative analysis between different SR techniques has been demonstrated. Several types of medical imaging datasets used by many researchers were reviewed. Despite the exponential progress of DL in medical image processing in the last few years, several shortcomings of existing techniques are ponied out and bring forth future research directions. In summary, to design an algorithm for medical image super-resolution with less complexity, a run time that works well in a real-time clinical environment is the need of the hour.

References

[1] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution Image Reconstruction: A Technical Overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21-36, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[2] Watchara Ruangsang, and Supavadee Aramvith, "Efficient Super-resolution Algorithm Using Overlapping Bicubic Interpolation," *IEEE 6th Global Conference on Consumer Electronics,* pp. 1-2, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Xin Li, and M.T. Orchard, "New Edge-Directed Interpolation," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1521–1527, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Hassan Aftab, Atif Bin Mansoor, and Muhammad Asim, "A New Single Image Interpolation Technique for Super Resolution," *IEEE International MultitopicConference*, pp. 592-596, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[5] V. Banupriya, and S. Anusuya, "Improving Classification of Retinal Fundus Image using Flow Dynamics Optimized Deep Learning Methods," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 9, no. 12, pp. 39-48, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] R. Priyadharsini, and T. Sree Sharmila, "An Efficient Edge Preserving Interpolation Method for Underwater Acoustic Image Resolution Enhancement," *Archives of Acoustics*, vol. 47, no. 2, pp. 267-274, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Zhi-Song Liu, Wan-Chi Siu, and Jun-Jie Huang, "Image Super-resolution via Hybrid NEDI and Wavelet-based Scheme," *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 1131-1136, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[8] Shalini Dubey, Shalini Sahu, and Pankaj Bazal, "Single Image Super Resolution using Interpolation and Discrete Wavelet Transform," *International Journal of Trend in Scientific Research and Development*, vol. 2, no. 6, pp. 241-249, 2018. [CrossRef] [Publisher Link]

[9] Na Sun, and Huina Li, "Super-resolution Reconstruction of Images Based on Interpolation and Full Convolutional Neural Network and Application in Medical Fields," *IEEE Access*, vol. 7, pp. 186470-186479, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] Yunfeng Zhang et al., "A Single-Image Super-Resolution Method Based on Progressive-Iterative Approximation," *IEEE Transactions on Multimedia*, vol. 22, no. 6, pp. 1407-1422, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Cui Zhou, and Jinghong Zhou, "Single-Frame Remote Sensing Image Super-Resolution Reconstruction Algorithm Based on Two-Dimensional Wavelet," *IEEE 3rd International Conference on Image, Vision and Computing*, pp. 360-363, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Priscilla Whitin, and V. Jayasankar, "A Novel Deep Learning-Based System for Real-Time Temperature Monitoring of Bone Hyperthermia," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 1, pp. 187-196, 2023. [CrossRef] [Publisher Link]

[13] Yi Tang, Jun-Hua Chen, and Zuo Jiang, "Piece-Wise Kernel Regression for Example-Based Super-Resolution," *International Conference on Machine Learning and Cybernetics,* pp. 143-148, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Faezeh Yeganli, and Kuldeep Singh, "Finger-Print Image Super-Resolution via Gradient Operator based Clustered Coupled Sparse Dictionaries," *IEEE International Symposium on Innovations in Intelligent Systems and Applications,* pp. 1-4, 2019. [Google Scholar] [Publisher Link]

[15] Yi Tang, and Ling Shao, "Pairwise Operator Learning for Patch-Based Single-Image Super-Resolution," *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 994–1003, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shuiping Gou et al., "Image Super-Resolution Based on the Pairwise Dictionary Selected Learning and Improved Bilateral Regularization," *IET Image Processing*, vol. 10, no. 2, pp. 101–112, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yapeng Tian et al., "Anchored Neighborhood Regression Based Single Image Superresolution from Self-Examples," *IEEE International Conference on Image Processing*, pp. 2827-2831, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[18] Kaibing Zhang et al., "Single Image Super-Resolution with Multiscale Similarity Learning," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 24, no. 10, pp. 1648-1659, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[19] M. N. Rajesh, B. S. Chandrasekar, "Deep Learning-Based Semantic Segmentation Models for Prostate Gland Segmentation," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 2, pp. 157-171, 2023. [CrossRef] [Publisher Link]

[20] Shengyang Dai et al., "Bilateral Back-Projection for Single Image Super Resolution," *IEEE Conference on Multimedia and Expo*, pp. 1039-1042, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[21] Vaishali Patel et al., "Hybrid Approach for Single Image Super Resolution using ISEF and IBP," *International Conference on Communication Systems and Network Technologies*, pp. 495-499, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[22] Rujul R Makwana, and Nita D Mehta, "Single Image Super-Resolution via Iterative Back Projection Based Canny Edge Detection and a Gabor Filter Prior," *International Journal of Soft Computing & Engineering*, vol. 3, no.1, pp. 379-384, 2013. [Google Scholar] [Publisher Link]

[23] Jun-Sang Yoo, and Jong-Ok Kim, "Noise-Robust Iterative Back-Projection," *IEEE Transactions on Image Processing*, vol. 29, pp. 1219-1232, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Damber Thapa et al., "A Performance Comparison among Different Super-Resolution Techniques," *Computers & Electrical Engineering*, vol. 54, pp. 313-329, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[25] NasimulNoman, "A Shallow Introduction to Deep Neural Networks," *Deep Neural Evolution*, pp. 35 – 63, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[26] Yoong Khang Ooi, and Haidi Ibrahim, "*Deep Learning Algorithms for Single Image Super-Resolution: A Systematic Review*," *Electronics*, vol. 10, no. 7, p. 867, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] Ying Liu et al., "Single Image Super Resolution Techniques Based on Deep Learning: Status, Applications and Future Directions," *Journal of Image and Graphics*, vol. 9, no. 3, 2021.[Google Scholar] [Publisher Link]

[28] Hong Lin et al., "Generative Adversarial Image Super-Resolution Network for Multiple Degradations," *IET Image Processing*, vol. 14, no. 17, pp. 4520-4527, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[29] Ming Zhao et al., "Super-Resolution of Cardiac Magnetic Resonance Images using Laplacian Pyramid Based on Generative Adversarial Networks," *Computerized Medical Imaging and Graphics*, vol. 80, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[30] Heng Liu et al., "Perception Consistency Ultrasound Image Super-Resolution via Self-Supervised CycleGAN," *Neural Computing and Applications*, vol. 35, pp. 12331-12341, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[31] Chao Dong et al., "Image Super-Resolution using Deep Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[32] Christian Ledig et al., "Photo-Realistic Single Image Super-Resolution using a Generative Adversarial Network," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 105-114, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[33] Ahlad Kumar, Harsh Vardhan Singh, and Vijeta Khare, "Tchebichef Transform Domain-Based Deep Learning Architecture for Image Super-Resolution," *IEEE Transactions on Neural Networks and Learning Systems*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] Jin Zhu et al., "MIASSR: An Approach for Medical Image Arbitrary Scale Super-Resolution," *arXiv preprint Image and Video Processing*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[35] Kuan Zhang et al., "SOUP-GAN: Super-Resolution MRI Using Generative Adversarial Networks," *Tomography*, vol. 8, no. 2, pp. 905-919, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[36] Faezehsadat Shahidi, "Breast Cancer Histopathology Image Super-Resolution Using Wide-Attention GAN with Improved Wasserstein Gradient Penalty and Perceptual Loss," *IEEE Access*, vol. 9, pp. 32795-32809, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[37] Defu Qiu et al., "Multiple Improved Residual Networks for Medical Image Super-Resolution," *Future Generation Computer Systems*, vol. 116, pp. 200-208, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[38] Hongbo Zhu et al., "Functional-Realistic CT Image Super-Resolution for Early-Stage Pulmonary Nodule Detection," *Future Generation Computer Systems*, vol. 115, pp. 475-485, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] J. Andrew et al., "Super-Resolution Reconstruction of Brain Magnetic Resonance Images via Lightweight Autoencoder," *Informatics in Medicine Unlocked*, vol. 26, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[40] Zhaoyang Songa et al., "Progressive Back-Projection Network for COVID-CT Super-Resolution," *Computer Methods and Programs in Biomedicine*, vol. 208, 2021.[CrossRef] [Google Scholar] [Publisher Link]

[41] Farah Deeba et al., "Wavelet-Based Enhanced Medical Image Super Resolution," *IEEE Access*, vol. 8, pp. 37035-37044, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[42] Yan Lv et al., "Fusing Dense and ReZero Residual Networks for Super-Resolution of Retinal Images," *Pattern Recognition Letters*, vol. 149, pp. 120-129, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[43] Chao Dong, Chen Change Loy, and Xiaoou Tang, "Accelerating the Superre Solution Convolutional Neural Network," *European Conference on Computer Vision*, pp. 391–407, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[44] Shengxiang Zhang et al., "A Fast Medical Image Super Resolution Method Based on Deep Learning Network," *IEEE Access*, vol. 7, pp. 12319-12327, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[45] Garima Pandey, and Umesh Ghanekar, "A Conspectus of Deep Learning Techniques for Single-Image Super-Resolution," *Pattern Recognition and Image Analysis*, vol. 32, no. 1, pp. 11-32, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[46] Minakshi Gogoi, and Mala Ahmed, "Image Quality Parameter Detection: A Study," *International Journal of Computer Sciences and Engineering*, vol. 4, no. 7, pp. 110-116, 2016. [Publisher Link]

[47] Varsha Nemade, Sunil Pathak, and Ashutosh Kumar Dubey, "Hybrid Deep Convolutional Neural Network Approach for Detecting Breast Cancer in Mammography Images," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 5, pp. 102-119, 2023. [CrossRef] [Publisher Link]

[48] Kai Li et al., "Survey of Single Image Super-Resolution Reconstruction," *IET Image Processing*, vol. 14, no. 11, pp. 2273-2290, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[49] Zhou Wang et al., "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transaction on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[50] G. R. Meghana, Suresh Kumar Rudrahithlu, and K. C. Shilpa, "Detection of Brain Cancer using Machine Learning Techniques a Review," *SSRG International Journal of Computer Science and Engineering*, vol. 9, no. 9, pp. 12-18, 2022. [CrossRef] [Publisher Link]

[51] Zhihao Wang, Jian Chen, and Steven C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3365-3387, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[52] Zhaowen Wang et al., "Deep Networks for Image Super-Resolution with Sparse Prior," *IEEE International Conference on Computer Vision*, pp. 370-378, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[53] B. Suryakanth, and S A. Hari Prasad, "3D CNN-Residual Neural Network Based Multimodal Medical Image Classification," *International Journal of Engineering Trends and Technology*, vol. 70, no. 10, pp. 371-380, 2022. [CrossRef] [Publisher Link]

[54] Justin Johnson, Alexandre Alahi, and Li Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," *European Conference on Computer Vision*, pp. 694–711, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[55] Bee Lim et al., "Enhanced Deep Residual Networks for Single Image Super-Resolution," *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, vol. 1, pp. 1132-1140, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[56] Lina Zhang, Haidong Dai, and Yu Sang, "Med-SRNet: GAN-Based Medical Image Super-Resolution via High-Resolution Representation Learning," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.[CrossRef] [Google Scholar] [Publisher Link]

[57] R. Tamilaruvi et al., "Brain Tumor Detection in MRI Images using Convolutional Neural Network Technique," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 9, no. 12, pp. 198-208, 2022. [CrossRef] [Publisher Link]

[58] Ruchika Dhawan, and Umesh Ghanekar, "Single-Image Super-Resolution Using Rational Fractal Interpolation and Adaptive Wiener Filtering," *Proceedings of First International Conference on Computational Electronics for Wireless Communications*, pp. 477-486, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[59] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-Based Super-Resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[60] M. Sakthivadivu, and P. Suresh Babu, "Analytical and Empirical Survival Study on Natural Image Compression and Classification using Machine Learning Techniques," *International Journal of Computer Trends and Technology*, vol. 70, no. 8, pp. 21-29, 2022. [CrossRef] [Publisher Link]

[61] Michal Irani, and Shmuel Peleg, "Motion Analysis for Image Enhancement: Resolution, Occlusion and Transparency," *Journal of Visual Communication and Image Representation*, vol. 4, no. 4, pp. 324-335, 1993. [CrossRef] [Google Scholar] [Publisher Link]

[62] Shahrukh Athar, and Zhou Wang, "A Comprehensive Performance Evaluation of Image Quality Assessment Algorithms," *IEEE Access*, vol. 7, pp. 140030-140070, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[63] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440–345, 2006. [CrossRef] [Google Scholar] [Publisher Link]