IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 25, 2019, accepted February 7, 2020, date of publication February 17, 2020, date of current version March 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2974278

Wavelet-Based Enhanced Medical Image Super Resolution

FARAH DEEBA[1], (Member, IEEE), SHE KUN[1], FAYAZ ALI DHAREJO[2,3], (Member, IEEE), AND YUANCHUN ZHOU[2,3]

[1]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China
[3]University of Chinese Academy of Sciences, Beijing 100190, China

Corresponding author: Yuanchun Zhou (zyc@cnic.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61836013, and in part by the National Key Research Development and Plan of China under Grant 2016YFB501901.

ABSTRACT Low-resolution medical images can seriously interfere with the medical diagnosis, and poor image quality can lead to loss of detailed information. Therefore, improving the quality of medical images and accelerating the reconstruction is of particular importance for diagnosis. To solve this problem, we propose a wavelet-based mini-grid network medical image super-resolution (WMSR) method, which is similar to the three-layer hidden-layer-based super-resolution convolutional neural network (SRCNN) method. Due to the amplification characteristics of wavelets, a stationary wavelet transform (SWT) is used instead of a discrete wavelet transform (DWT). Also, due to the nature of redundant (scale-by-scale) wavelets, it is possible to retain additional information about the image and restore high-resolution images in detail. For a large amount of training data, wavelet sub-band images, including approximation and frequency subbands are combined into a predefined full-scale factor. The mapping between the wavelet sub-band image and its approximate image is then determined. In order to ensure the reproducibility of the image, a method of adding a sub-pixel layer is proposed to realize the hidden layer, and replacing the small mini-grid-network on the hidden layer is of considerable significance to speed up the image recovery speed. Experimental results on the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) show that the model has better performance.

INDEX TERMS Medical images, super-resolution (SR), deep learning, wavelet learning.

I. INTRODUCTION

The medical imaging system provides detailed information about the anatomy of human organs and the functions of human organs. Typical conventional medical imaging systems for expert diagnoses, such as MRI, CT, PET-CT, and Ultrasound [1]–[3]. However, these images are usually in low quality and lack of internal information. Due to hardware and current imaging technology limitations, medical professionals and researchers prefer image super-resolution processing technology for medical diagnosis [4]. The Single Image Super-Resolution (SISR) problem is considered very complex in theory because the number of unknown variables in the High-Resolution (HR) image is better than in the Low-Resolution (LR) image. To solve this problem, scientists have introduced several techniques in the field of Super-resolution, which is mainly divided into three aspects based on edges [5], interpolation-based methods [6], and sample-based methods [7]–[9].

The sparse-based implementation method first proposed by Yang [10], and it often takes the first place in the field of high-quality rehabilitation. Later, Yang *et al.* [11] introduced an improved popular technique for image super-quality through sparse representation. In that article, the authors believe that the image block can be well represented by appropriate dictionary selection. Inspired by this observation, we look for a scarce view for each low-resolution input patch and then use the coefficients of this view to producing high-resolution output. Nowadays, researchers are more interested in the neural network, deep learning-based approach for solving the SISR problem, due to the enormous capacity (payload) of the neural network model, and

The associate editor coordinating the review of this manuscript and approving it for publication was Yong Yang.

holistic learning. These neural networks help to acquire the functionality applied in previous methods and also improve many deep learning algorithms. These improved Deep neural network DNN methods are cost-effective and significantly reduced with sufficient quality.

Our proposed method is inspired by the category of the wavelet domain SISR algorithms [12]–[15]. Many of these algorithms provide convenient performance results. However, their computational cost is too high. In [12], the author introduces a wavelet dictionary learning algorithm that learns a compact dictionary for single image super-resolution (SISR). Later, a related method, the DWT dictionary learning method [13] was introduced, which was inspired by DWT technology. Deeba [14] used wavelet properties in conjunction with coupled dictionary learning methods. Most of these algorithms produce excellent results. However, their calculation costs are high. As deep learning algorithms grow, acceptable quality increases significantly, and computational costs reduce.

For a deep convolution network, the most reliable method is SRCNN [8], which aims to extract high-resolution image from the low-resolution image by CNN. SRCNN [8] in the wavelet domain used to improve excellent visual effects [15]. In SRCNN [8], the authors used three-layer network architectures to learn the complex nonlinear mapping between HR and LR image patches. After that, a deep network architecture was proposed by [16], authors use residual images for training instead of using HR and LR images, and adjustable gradient clipping to increase the convergence of their algorithm. Besides, the same author [8] proposed an accelerated version of the SRCNN [8] algorithm called the Fast Super-Resolution Convolutional Neural Network (FSRCNN) [17] algorithm, which can obtain better results without interpolation between LR and HR images. This will reduce the mapping in the feature learning steps. Shi [18] proposed the first real-time image and video super-resolution using the sub-pixel convolutional neural network named as ESPCN [18]. However, compared to the SRCNN method, the ESPCN method lacks context information after reconfiguration. Super-resolution with multiple degradation algorithm SRMD by deep network model was proposed by [19], with the use of degradation maps obtained using the dimensional reduction analysis of necessary components (PCA) and then stretching. In this way, they learned one network model for many scales.

Inspired by the three-layer network SRCNN [8], Gao *et al.* [4] use a deep convolution network to achieve super-resolution of a single image on a medical data set. The treatment diagnosis can be further improved by reconstructing these data sets. Although better quality can be obtained, it usually costs more to restore the HR images. Shortening the image recovery time has become an issue and should be resolved instantaneously. The authors in [15] introduced the DWT wavelet domain-based deep learning method and achieved good results, but the author did not make full use of the potential of deep learning and wavelet performance. Authors [16] introduced the fast medical image super-resolution by speeding the network of [18].

Our proposed wavelet-based approach integrates the benefits of end-to-end network learning with the potential of model capacity [16] along with wavelet properties such as redundancy, directionality, sparsity [20], [21] and etc. To apply the wavelet analysis, we choose the SWT wavelet instead of the DWT due to its upsampling property to preserve more context information. Our proposed wavelet-based mini-grid network for the medical image super-resolution method focuses on the faster reconstruction as well as enhancing the visual quality more significant for better diagnosis purposes. Therefore, in order to improve the visual quality of the algorithm, we adopted SWT and combined the network with sub-pixel convolution and mini-grid network to reduce the super-resolution time. Specifically, we implemented three hidden layers to maintain information while training the image.

For model training, we designed a wavelet domain deep neural network architecture that trains the network between approximate wavelet subbands and their corresponding subband images. It can be clearly seen from the experiment that by adding wavelets in the end-to-end network, the visual performance of the image can be improved at a reasonable cost, thereby realizing the task of the SISR. By using a public dataset, the proposed method is well compared to the latest algorithms. The quality of the visual measurements is analyzed by the peak signal-to-noise ratio (PSNR) and structural similarity index measurements (SSIM).

Below we first explain how to propose and improve the speed of the network in the wavelet domain. Then, we outline how we conducted the experiment and described the experimental results in the next section. At the end of this article, we summarized the proposed approach and stated the future work.

II. PROPOSED METHOD

In this paper, a mini-grid wavelet-based model is proposed, due to unique characteristics of wavelets, including wavelet sub-bands are sparse, and they exploit multi-scale modeling. We chose the SWT wavelet because of its upscaling features, so the size of the wavelet subband remains the same while retaining the details of the wavelet subbands. Figure1 shows DWT and SWT decomposition, and the technique used in these [22]–[25] regarded the LR image as a wavelet approximation image of the corresponding HR image.

$$A_s(p,q) = \sum_{I=j}^{P} \sum_{j=1}^{Q} h_{I=1}^{1} h_{j=1}^{2} A_s(I,j) \quad (1)$$

$$H_s(p,q) = \sum_{I=j}^{P} \sum_{j=1}^{Q} h_{I=1}^{1} h_{j=1}^{2} A_s(I,j) \quad (2)$$

$$V_s(p,q) = \sum_{I=j}^{P} \sum_{j=1}^{Q} g_{I=1}^{1} h_{j=1}^{2} A_s(I,j) \quad (3)$$

$$D_s(p,q) = \sum_{I=j}^{P} \sum_{j=1}^{Q} g_{I=1}^{1} g_{j=1}^{2} A_s(I,j) \quad (4)$$

Here wavelet analysis filters are represented as $h_p^1, h_q^2, g_p^1, g_q^2$ where $A_{s-1}(p,q)$, $H_{q-1}(p,q)$, $V_{q-1}(p,q)$, and $D_{s-1}(p,q)$, are wavelet sub bands approximation, horizontal,

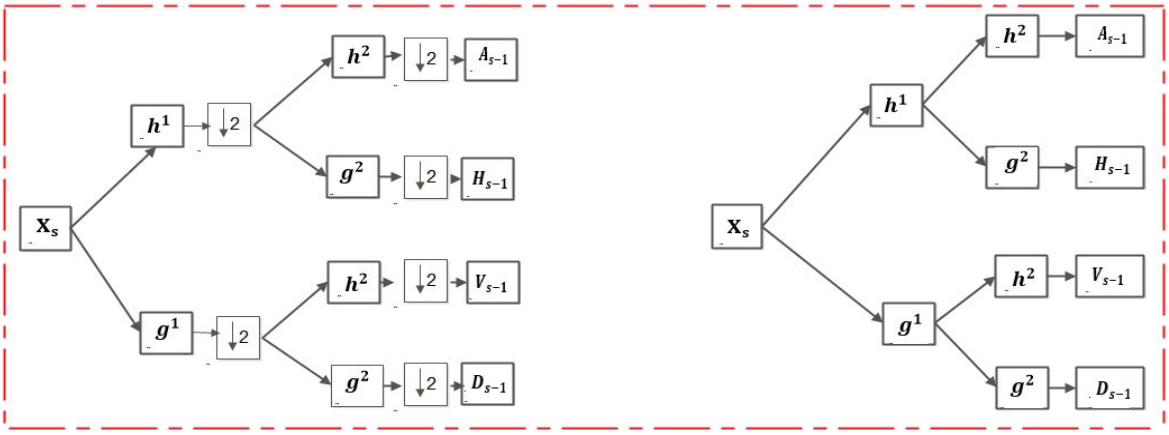

FIGURE 1. Wavelet decomposition (a) DWT decomposition (b) SWT decomposition.

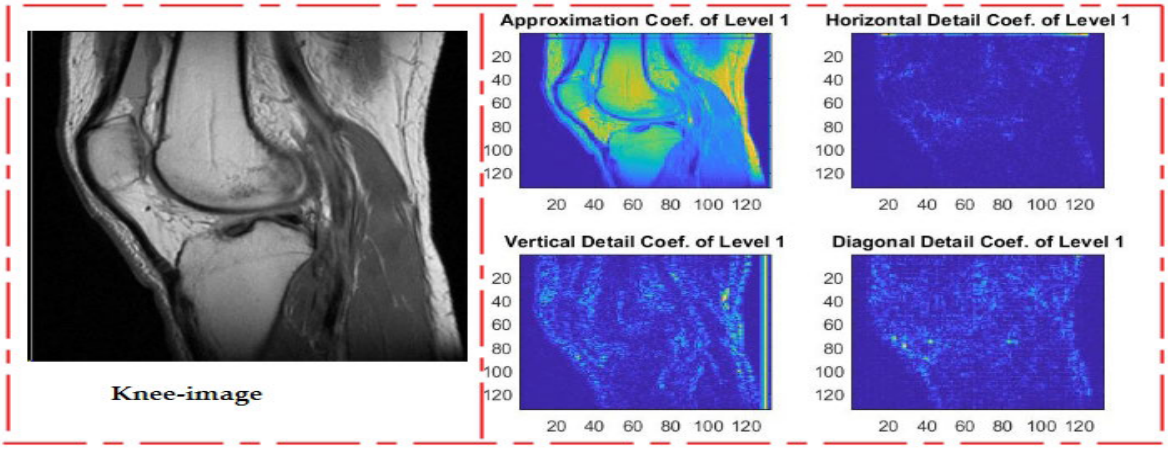

FIGURE 2. Visual results of wavelet decomposition, original, approximation, horizontal, vertical and diagonal sub-bands.

vertical, and diagonal, respectively. The visual decomposition is represented in Figure 2. The wavelet synthesis equation is written as

$$\begin{aligned} A_{s+2}(p,q) = & \sum_{I=jj}^{p}\sum_{j=1}^{Q} h_{I-p}^{\sim 1} h_{j-q}^{\sim 2} \tilde{A}_s(I,j) A_{s+2}(p,q) \\ & + \sum_{I=jj}^{P}\sum_{j=1}^{Q} h_{I-p}^{\sim I} g_{j-q}^{\sim 2} H_s^{\sim}(I,j) \\ & \times \sum_{I=jj}^{P}\sum_{j=1}^{Q} g_{I-p}^{\sim I} g_{I-q}^{\sim 1} h_{j-n}^{\sim 2} V_s^{\sim}(I,j) \\ & + \sum_{I=1}^{P}\sum_{j=1}^{Q} g_{I-p}^{\sim I} h_{j-q}^{\sim 2} D_s^{\sim}(I,j) \end{aligned} \tag{5}$$

It can be seen from Figure 2 that the wavelet subband is very sparse, which reveals the obvious directionality of the image. Further application of dimensions will lead to the lack of fine features of this directionality. A first-order inverse transform is performed to obtain an HR image. Figure2 depicts the strong dependency among wavelet coefficients toward wavelet sub-bands.

In image processing, reducing the size is a well-known problem, and many methods have been introduced for this [26]–[28]. In [26], the authors proposed a method that combines the dimensionality reduction cycle using principal component analysis (PCA) and multidimensional scaling capabilities. This approach can detect non-linear degrees of option underlying complicated common observations such as an individual signature or images faces in varying viewing positions. The authors of [27] described the KPCA method is the best among three (principal component analysis (PCA), independent component analysis (ICA), and kernel principal component analysis(KPCA)) for the vector support machine (SVM) feature extraction. Authors in [28], proposed an optimization algorithm with geometry represents higher dimensional data, allows for a reduction in computational measurement over previous PCA and MDS models. Our proposed method studies a single network model at several scales, where wavelet domain decomposition is used before training the network, and then the image of the wavelet subband is used as the input to the training.

The proposed method is changed prior to the neural network and wavelet-based processes, based on the following perspectives.

- SWT wavelet decomposition is utilized in our proposed method to evaluate the wavelet coefficients.
- We propose a deep network architecture similar to the fast medical image super-resolution method based on

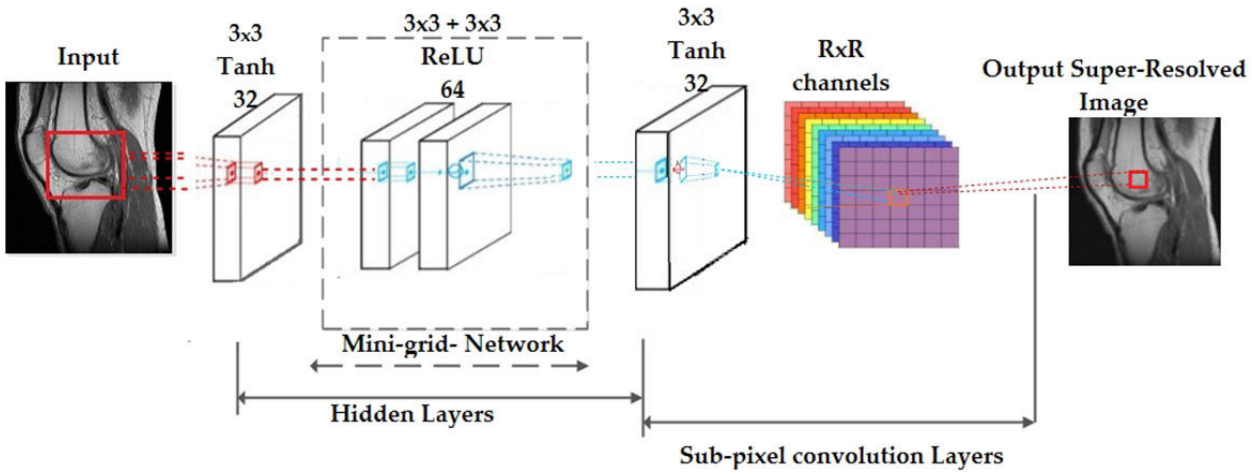

FIGURE 3. Structure framework of the proposed model.

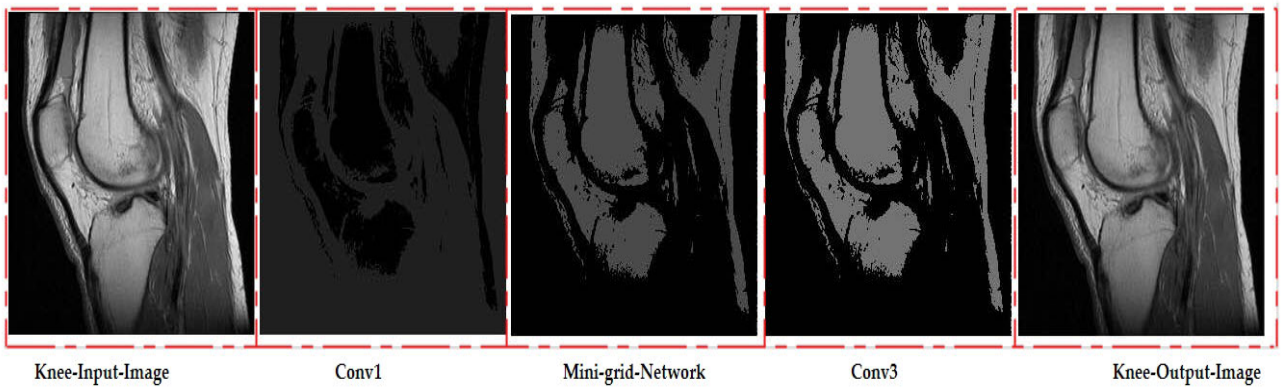

FIGURE 4. Process of the proposed Super-Resolution method.

deep learning networks [29]. Still, we train the network on wavelet domain images instead of residuals. However, the author [15] used DWT in conjunction with a three-layer neural network inspired by SRCNN [8].

- We designed a deep neural network based on the wavelet domain, proposed a super-resolution accelerated wavelet domain, deep neural network model, to determine sparse output, improveaccuracy, speed of reconstruction, and training efficiency.

Figure 3 shows the structure framework of the proposed model which depicts a proposed method for a fast medical reconstruction method based on a three-layer deep learning network called a ''mini-grid-network,'' hidden layer, and sub-pixel layer. Because the ''mini-grid-network'' is a kind of small convolution neural network, sub-pixel convolution can be used as the output layer of the super-resolution image directly. The output sparsity can be determined by wavelet, thus improving the accuracy of image reconstruction, while figure 4 represents our three-layered network-based output.

A. SUB-PIXEL CONVOLUTION LAYER

The sub-pixel convolutional layer is used as the last layer in the proposed model. Shi [18] introduced the sub-pixel convolutional layer to reconstruct a low-resolution image. As shown in Figure 3, by using an upscaling filter for each feature map (R*R channel), the sub-pixel convolution layer can obtain a high-resolution image directly from the low-resolution feature map. Several kernels W with size k can be activated in a low-resolution subspace.

In the convolutional layer, W is the kernel of size K and can be activated in low-resolution space. R*R is the number of active patterns, as shown in Figure 3. Activation weights are being activated for the active pattern position $[k/R]^2$. These patterns are regularly activated during the convolution of the image according to the location of the different sub-pixels mod (x, y), mod (y, r). Here (x, y) represents the high-resolution output pixel coordinates used to rearrange the elements.

B. MINI-GRID- NETWORK

Two convolution kernels of size 3*3 are nested in hidden layers to reduce the time named mini-grid-network. After analyzing model SRCNN (9-5-5), the feature map is achieved better with the second layer having the configuration of 5*5 convolution kernel. In the proposed method, a 5*5 convolution kernel is replaced by this mini grid-network to achieve the same results much faster as in [29]. To achieve the greater

TABLE 1. MINI-GRID –Network.

Mini-Network	ESPCN[25](5×5)	(3x3+3x3)
Time Complexity	$O(N^2(5 \times 5 \times 64))$	$O(N^2(3 \times 3 + 3 \times 3))$
Parameter	5×5+1= 26	3×3+2+3×3+2= 22
Calculations	$25(N-4)^2$	$9(N-2)+9(N-4)^2$

TABLE 2. Parameters of hidden layers.

HIDDEN LAYER PARAMETERS		
Conv1	3×3×32	Tanh activation function
Mini-grid-Network	(3×3+3×3)×64	ReLUactivation function
Conv3	3 ×3×32	Tanhactivation function

susceptibility, large convolution kernels will be used with increased numbers of parameters, but it also increases the number of calculations. As convolution kernel size is directly proportional to the number of parameters, so considering this small size of the kernel is favorable. In the proposed method, we utilized the ReLU function instead of the Tanh activation function due to its lower calculation property. ReLU only determines that the input is greater than zero. The same receptive field is achieved by 3*3 convolution kernels by mini-network.

$$T = O(N^2 * K^2 * F) \tag{6}$$

$$\text{output} = \frac{N - K - 2 * P}{S} + 1 \tag{7}$$

Here T represents time, the size of the input image is represented by N, kernel size is K and number of the filter is represented by F, padding is represented by P and step length in horizontal as well as in vertical direction is represented by S. we used padding $P = 0$ and $S = 1$.

Table1 shows the calculation of time complexity and parameters in the mini-grid-network.

C. HIDDEN LAYERS

The deep network consists of multiple layers, each parameter of each layer is utilized for feature learning purpose, as the layers increase the feature learning rate is also increased. Additionally, mini-grid-network is added for exploiting the speed of the network due to its good quick performance on the network [29]. The proposed model comprised of three layers, conv1, and con3 comprised of 32 kernel size with 3*3 conv size with Tanh activation function, while conv2 is named as mini-grid-network.

This mini-grid-network consists of two conv layers with a 3*3 conv kernel and ReLU activation function. Table2 listed all parameters.

III. EXPERIMENTAL SETUP FOR TRAINING AND TESTING

As our model is wavelet-based, so we used the one-level wavelet decomposition before training and utilized approximation image and its corresponding sub-band images for training the model. Wavelets have redundancy property across each scale, provide approximation wavelet sub-band as input at a certain scale, reconstruction of approximation image can be perfectlyreconstructed due to its wavelet redundancy property.

LR image is decomposed into one level wavelet decomposition represented by x, and sub-bands (horizontal, vertical, and diagonal) sub-bands represented by Y. Here in our proposed method, we learn the relationship between LR approximation image and wavelet sub-band (horizontal, vertical, and diagonal) images. One problem with SRCNN [8] network is that the details about the input image should be saved after getting the output image. Those learned features are utilized and the input image is removed. If the network is so deep with several layers, that will be an end to end learning in this case and increase the overload and also requires memory. Because of this, the vanishing gradient problem [30] occurs, and it needs a solution. In the proposed network, we solve this problem by learning the wavelet coefficients.

In our proposed network model parameters are adjusted with the mentioned values, the learning rate is 10^{-4} is set, momentum is 0.9, weight decay is 0. Mini- Batch Gradient Descent algorithm is chosen in our process.

Mini-Batch Gradient Descent

Repeat:
{
For i = 1, 11, 21,31,…911
{

$$\boldsymbol{\theta_j} := \boldsymbol{\theta_j} - \alpha \frac{1}{10} \sum_{k=i}^{i+9} (\mathbf{h_\theta}\left(\mathbf{x^{(k)}}\right) - \mathbf{y^{(k)}}) \mathbf{x_j^{(k)}})$$

(for every j = 0,…..,n)
}
}

From above, $\sum_{k=i}^{i+9} (h_\theta\left(x^{(k)}\right) - y^{(k)}) x_j^{(k)})$ represents the loss function. In the proposed method Gaussian distribution is implemented for each convolution layer to initialize the weights, described in the equation below.

$$X \sim N\left(x \mid \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e - \frac{(x-\mu)^2}{2\sigma^2} \tag{8}$$

σ represents the standard deviation, σ^2 is variance in it, and μ is the mean of the distribution. For the training dataset, batch size sets to the value of 128, while for testing, we use 32 batch size. To calculate the loss between label value and predictive values Euclidean loss [31] function is used and described as below.

$$\text{Euclidean Loss} = \frac{1}{2N} \sum_{n=1}^{N} \left\| Y'_n - Y_n \right\|_2^2 \tag{9}$$

Here N and n represents the total input image and number of the input image, Y'_n is predictive value, and Y_n is label value. Wavelets have redundancy property across each

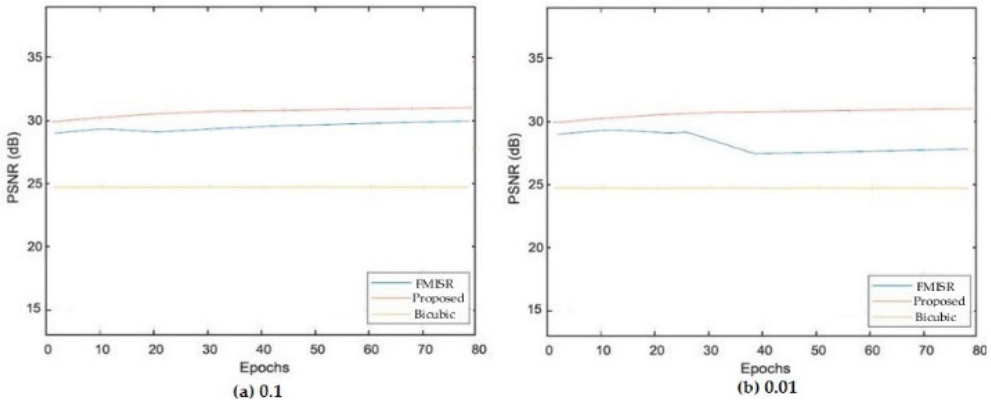

FIGURE 5. Performance curve comparison with Bicubi, FMISR [29], and proposed methods at0.1 and 0.01 learning rates.

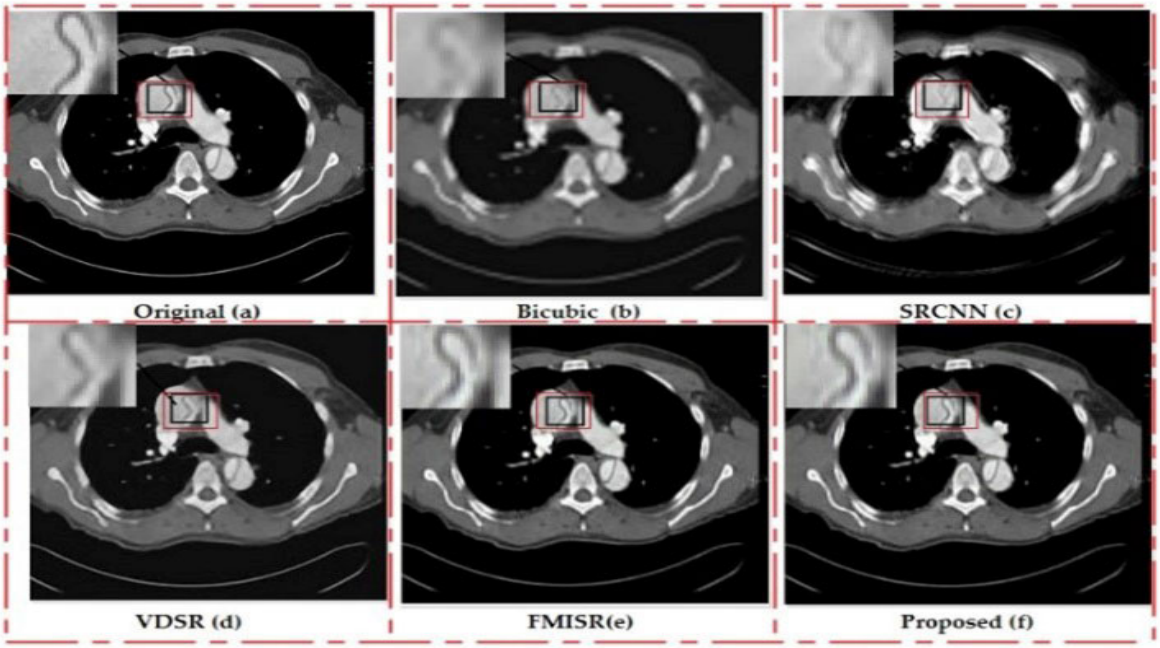

FIGURE 6. Abdomen image based on a scaling factor of 2, in which (a) original, (b) Bicubic, (c) SRCNN [8], (d) VDSR [16], (e) FMISR [29], and (f) proposed approach sub- figures respectively.

scale, provide approximation wavelet sub-band as input at a specific scale, reconstruction of approximation image can be perfectly reconstructed due to its wavelet redundancy property. For SISR, we learn the corresponding mapping between approximation image and its corresponding coefficients by employing the redundancy property of wavelet. As can be seen from Table 3, the algorithm provides good results by applying a deep neural network architecture in the wavelet domain. We use a depth of 20 weights listed in Table.1. Performance Table for proposed and FMISR [29] algorithm for Knee image(PSNR), is shown in Table3, and performance curves for different learning rates (0.1, 0.01) at scale parameter of 2 is represented in Figure 5. Hence, our approach provides better performance at mentioned learning rates.

For training and testing purposes, our experiment environment setup consists of a windows based machine with intel(R) Core(TM) i5-7300HQ CPU @ 3.40GHz, NVIDIA GeForce GTX 1080-Ti. Additionally, Matlab 2017 with CUDA Toolkit and Anaconda is utilized for the setup.

Computational time concerning different comparable algorithms is calculated and shown in table 4.

As can be seen from the Table 4, our method is faster and better than the bicubic, SRCNN [8], and VDSR [16] methods. Still, due to the same small network and hidden layer, the calculation time of our approach is almost similar to that of FMISR[29] but still can provide a better calculation of level 2.

IV. DATASET FOR TRAINING AND TESTING

In our proposed method, a publicly available data Shenzhen Hospital X-ray Set [33] containing 662 X-ray images were used for training purpose. All images were resized to 512 × 512 size, and we tested the models on three different

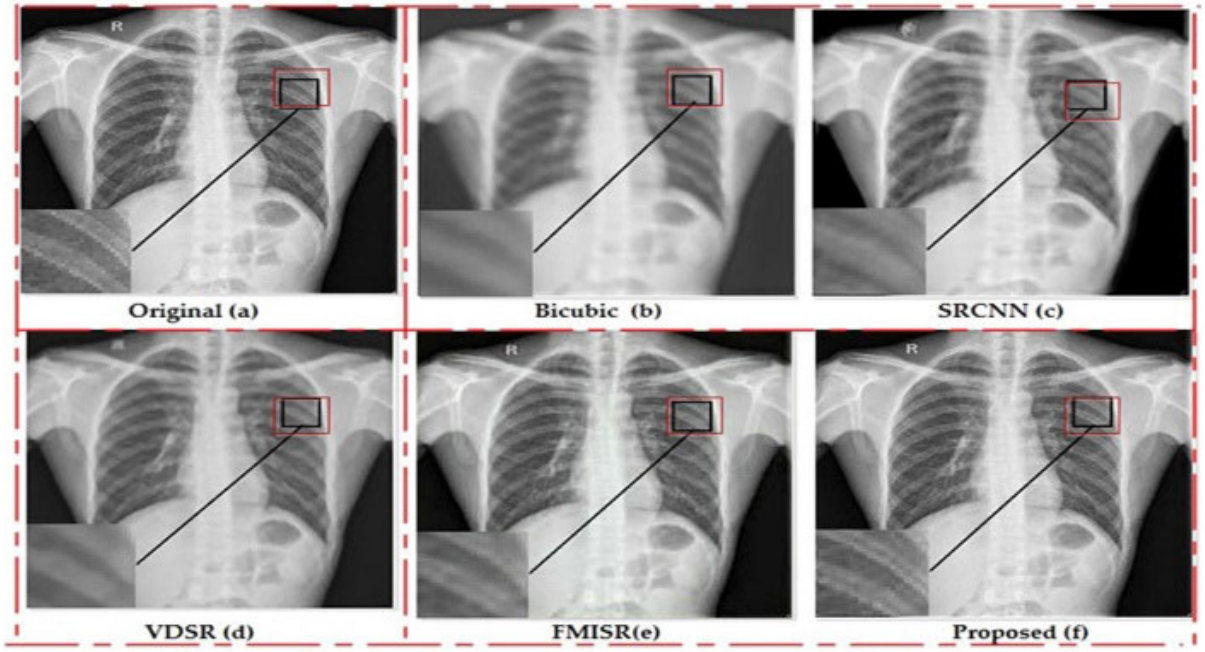

FIGURE 7. Chest image based on a scaling factor of 2, in which (a) original, (b) Bicubic, (c) SRCNN [8], (d) VDSR [16], (e) FMISR [29], and (f) proposed approach sub-figures respectively.

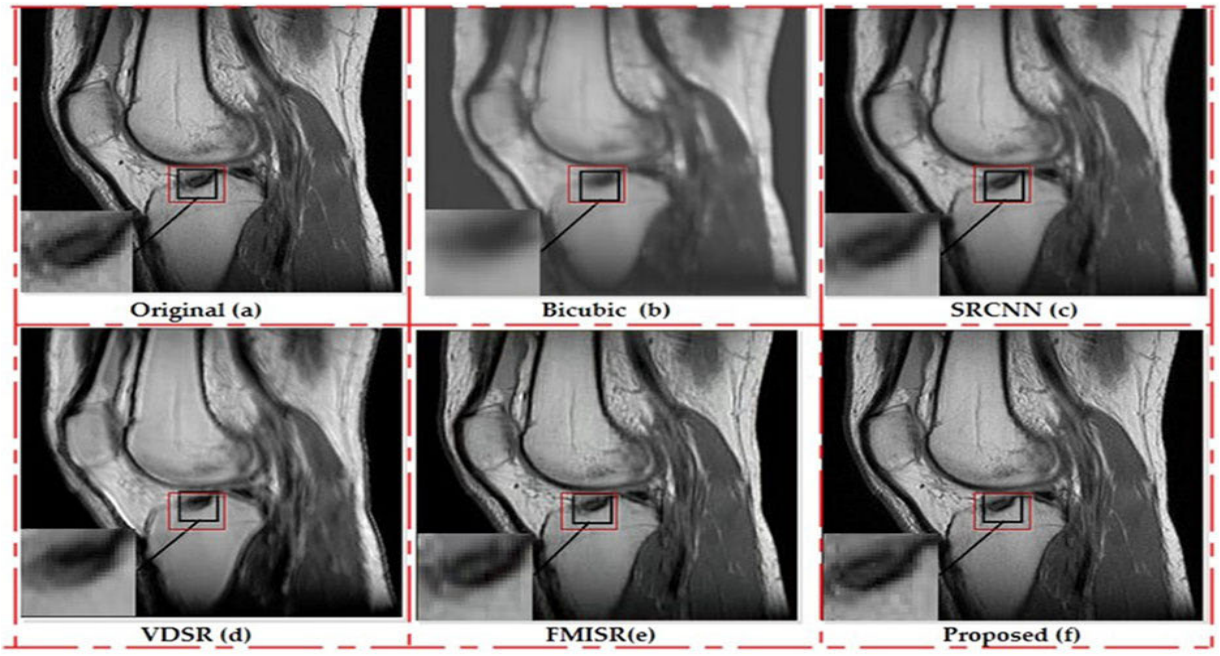

FIGURE 8. Knee image based on a scaling factor of 2 in which (a) original, (b) Bicubic, (c) SRCNN [8], (d) VDSR [16], (e) FMISR [29], and (f) proposed approach sub-figures respectively.

publicly available datasets, including Montgomery County X-ray, Teeth, Abdomen1, and knee images. Quantitative analysis was performed according to PSNR [14] and SSIM [14]. The mathematical definition is as follows.

A. PEAK SIGNAL-TO-NOISE RATIO (PSNR)

The Peak signal-to-noise ratio (PSNR) is used for the quantitative performance measure. Given a true image (original HR image) F and its estimated$\acute{F}$, with M×N pixels size, The PSNR is described as

$$\mathrm{PSNR}(\mathrm{F},\acute{\mathrm{F}}) = 10\log_{10}\frac{255^2}{\mathrm{MSE}(\mathrm{F},\acute{\mathrm{F}})} \tag{10}$$

where MSE (F,$\acute{F}$) represents the mean–square error between two images (F,$\acute{F}$). Given (F,$\acute{F}$), MSE defined as

$$\mathrm{MSE}\left(\mathrm{F},\acute{\mathrm{F}}\right) = \frac{1}{\mathrm{MN}}\sum_{\mathrm{i}=1}^{\mathrm{M}}\sum_{\mathrm{j}=1}^{\mathrm{N}}\left(\mathrm{F}_{\mathrm{ij}} - \acute{\mathrm{F}}_{\mathrm{ij}}\right)^2 \tag{11}$$

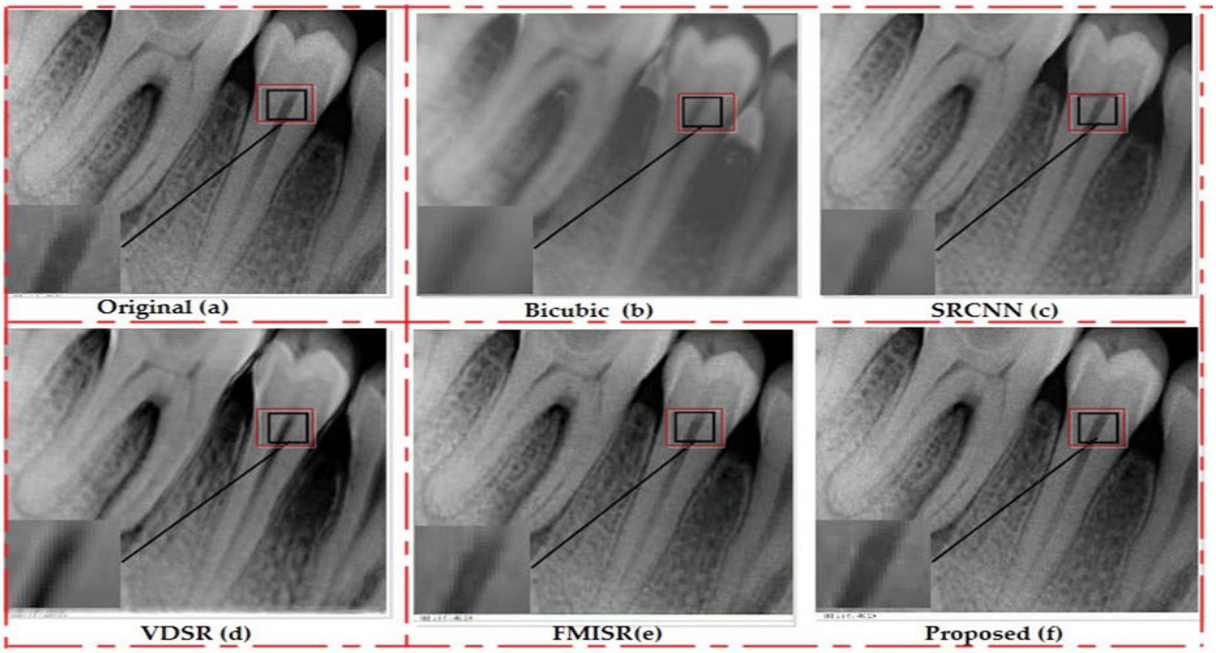

FIGURE 9. Teeth image based on a scaling factor of 2, in which (a) original, (b) Bicubic, (c) SRCNN [8], (d) VDSR [16], (e) FMISR [29], and (f) proposed approach sub-figures respectively.

TABLE 3. Table 3 Shows the Performance (peak signal to noise ratio or PSNR) for the proposed and FMISR [29] network ("Knee" mage, scale 2).

(a) 0.1 rate of learning			
Epoch	**FMISR**	**Proposed**	**Variance**
20	**28.15**	**30.11**	**1.96**
40	29.21	30.72	1.51
60	**29.27**	**31.04**	**1.77**
80	29.49	31.35	1.86
(b) 0.01 rate of learning			
Epoch	**FMISR**	**Proposed**	**Variance**
20	27.19	29.03	1.84
40	**27.87**	**28.92**	**1.05**
60	27.81	28.12	0.31
80	**27.95**	**30.07**	**2.12**

TABLE 4. Computational time for different method.

Time Computation					
Method	Bicubic	SRCNN	VDSR	FMISR	Proposed
Scale	2	2	2	2	2
SR-Time/s	0.017	2.371	2.101	0.257	**0.254**

An error is calculated between the real HR image and the reconstructed HR image. The higher the PSNR value, the better the reconstruction image.

B. STRUCTURAL SIMILARITY INDEX MEASUREMENT (SSIM)

For high-resolution quality evaluation of reconstructions, SSIM (structural similarity index) is extensively used. Wang and Zhou [32] and the mathematical representation of the SSIM index defined as:

$$\text{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{12}$$

where μ_x, μ_y are the average of x and y respectively, d2 σ_x^2, σ_y^2 are the variance of x and y respectively, σ_{xy} is the covariance of x and y, C1, C2 are the constants. Our model is compared with four different methods included the bicubic technique, SRCNN [8], VDSR[16], and FMISR[29], which is represented in Table 5 respectively. Trained models of these compared algorithms are provided by authors. The proposed algorithm gives better results than compared algorithms.

Figures 6-9 represent the comparative visual resultswith the scale 2 parameters. The proposed wavelet domain-based mini-grid-network for medical image super-resolution provides sharper edges and textures based visual results at scale 2, represented by Figures 6-9, respectively.

TABLE 5. Quantitative comparison values in terms of PSNR (Left) and SSIM (right).

Dataset	Scale	Bicubic	SRCNN[9]	VDSR[16]	FMISR[29]	Proposed
Brain	2	26.31/0.821	27.50/0.845	28.33/0.860	28.64/0.868	**29.08/0.876**
	3	23.90/0.814	24.92/0.828	25.16/0.849	26.52/0.861	**26.81/0.869**
	4	21.63/0.805	22.40/0.811	23.88/0.834	23.960.854	**24.29/0.865**
Chest	2	26.30/0.825	26.65/0.841	27.54/0.871	28.32/0.892	**30.01/0.908**
	3	23.92/0.818	24.75/0.827	26.29/0.867	27.06/0.872	**28.430/0.894**
	4	21.51/0.803	22.14/0.816	23.29/0.854	24.82/0.849	**26.28/0.882**
Knee	2	27.24/0.828	27.58/0.849	28.19/0.878	29.49/0.8901	**31.65/0.925**
	3	23.98/0.821	24.96/0.836	26.34/0.871	27.08/0.893	**28.21/0.919**
	4	21.76/0.814	22.54/0.825	23.91/0.862	24.89/0.869	**26.32/0.898**
Teeth	2	24.53/0.796	24.84/0.811	25.86/0.826	26.01/0.839	**27.32/0.856**
	3	22.40/0.783	22.68//0.797	23.73/0.815	24.17/0.822	**25.10/0.843**
	4	20.07/0.768	20.18/0.784	21.81/0.799	22.20/0.817	**23.67/0.829**

V. CONCLUSION

In this paper, we propose an effective wavelet-based deep neural network model to achieve super-resolution of a single image. In the experiments, we used medical datasets of four types of images (abdomen, X-rays, knees, and teeth). Compared with other deep neural network methods, the proposed network expands the convolutional layer to obtain a more realistic image reconstruction, thereby significantly reducing the computing time based on the mini-grid-network. To shorten the time of image reconstruction, we optimize the speed structure by combining sub-pixel convolution layers and ''mini-grid-network.'' Besides, we implemented a hidden layer to preserve information when training images to improve the quality of reconstruction. By using wavelets, many useful features of neural networks are used in SISR tasks such as large model capacity, end-to-end learning, and high performance in the wavelet domain. The SWT used instead of DWT due to its upscaling property experimental analysis is carried out to validate the efficiency of the proposed model. In the future, this work can be extended by applying other wavelet transforms, for example, Multiresolution discrete wavelet transform, and dual-tree complex wavelet transform.

REFERENCES

[1] R. R. Peeters, P. Kornprobst, M. Nikolova, S. Sunaert, T. Vieville, G. Malandain, R. Deriche, O. Faugeras, M. Ng, and P. Van Hecke, ''The use of super-resolution techniques to reduce slice thickness in functional MRI,'' *Int. J. Imag. Syst. Technol.*, vol. 14, no. 3, pp. 131–138, 2004.

[2] M. Elad and A. Feuer, ''Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images,'' *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, Dec. 1997.

[3] J. A. Kennedy, O. Israel, A. Frenkel, R. Bar-Shalom, and H. Azhari, ''Super-resolution in PET imaging,'' *IEEE Trans. Med. Imag.*, vol. 25, no. 2, pp. 137–147, Feb. 2006.

[4] Y. Gao, H. Li, J. Dong, and G. Feng, ''A deep convolutional network for medical image super-resolution,'' in *Proc. Chin. Automat. Congr. (CAC)*, Oct. 2017, pp. 5310–5315, doi: 10.1109/CAC.2017.8243724.

[5] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, ''Gradient profile prior and its applications in image super-resolution and enhancement,'' *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529–1542, Jun. 2011.

[6] B. S. Morse and D. Schwartzwald, ''Image magnification using level-set reconstruction,'' in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Dec. 2001, pp. 8–14.

[7] R. Timofte, V. De Smet, and L. Van Gool, ''A+: Adjusted anchored neighborhood regression for fast super-resolution,'' in *Proc. Asian Conf. Comput. Vis.* Singapore: Springer, 2014, pp. 111–126.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, ''Image super-resolution using deep convolutional networks,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[9] J.-B. Huang, A. Singh, and N. Ahuja, ''Single image super-resolution from transformed self-exemplars,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5197–5206.

[10] J. Yang, J. Wright, T. Huang, and Y. Ma, ''Image super-resolution as sparse representation of raw image patches,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[11] J. Yang, J. Wright, T. S. Huang, and Y. Ma, ''Image super-resolution via sparse representation,'' *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[12] M. Nazzal and H. Ozkaramanli, ''Wavelet domain dictionary learning-based single image superresolution,'' *Signal, Image Video Process.*, vol. 9, no. 7, pp. 1491–1501, Jan. 2014.

[13] S. Ayas and M. Ekinci, ''Single image super resolution based on sparse representation using discrete wavelet transform,'' *Multimedia Tools Appl.*, vol. 77, no. 13, pp. 16685–16698, Oct. 2017.

[14] F. Deeba, S. Kun, W. Wang, J. Ahmed, and B. Qadir, ''Wavelet integrated residual dictionary training for single image super-resolution,'' *Multimedia Tools Appl.*, vol. 78, no. 19, pp. 27683–27701, Jun. 2019.

[15] N. Kumar, R. Verma, and A. Sethi, ''Convolutional neural networks for wavelet domain super resolution,'' *Pattern Recognit. Lett.*, vol. 90, pp. 65–71, Apr. 2017.

[16] J. Kim, J. K. Lee, and K. M. Lee, ''Accurate image super-resolution using very deep convolutional networks,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1646–1654.

[17] Y. Wang, L. Xie, S. Qiao, Y. Zhang, W. Zhang, and A. L. Yuille, ''Multi-scale spatially-asymmetric recalibration for image classification,'' in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 509–525.

[18] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, ''Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1874–1883.

[19] K. Zhang, W. Zuo, and L. Zhang, ''Learning a single convolutional super-resolution network for multiple degradations,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3262–3271.

[20] E. Matsuyama, D.-Y. Tsai, Y. Lee, M. Tsurumaki, N. Takahashi, H. Watanabe, and H.-M. Chen, ''A modified undecimated discrete wavelet transform based approach to mammographic image denoising,'' *J. Digit. Imag.*, vol. 26, no. 4, pp. 748–758, Dec. 2012.

[21] Y. Chen and Z. Cao, "Change detection of multispectral remote-sensing images using stationary wavelet transforms and integrated active contours," *Int. J. Remote Sens.*, vol. 34, no. 24, pp. 8817–8837, Oct. 2013.
[22] S. S. Kim, I. K. Eom, and Y. S. Kim, "Image interpolation based on statistical relationship between wavelet subbands," in *Proc. IEEE Multimedia Expo Int. Conf.*, Jul. 2007, pp. 1723–1726.
[23] K. Kinebuchi, D. D. Muresan, and T. W. Parks, "Image interpolation using wavelet based hidden Markov trees," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, May 2001, pp. 1957–1960.
[24] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, "Wavelet algorithms for high-resolution image reconstruction," *SIAM J. Sci. Comput.*, vol. 24, no. 4, pp. 1408–1432, 2003.
[25] J. Tian, L. Ma, and W. Yu, "Ant colony optimization for wavelet-based image interpolation using a three-component exponential mixture model," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12514–12520, Sep. 2011.
[26] J. B. Tenenbaum, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000.
[27] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee, and Q. M. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, nos. 1–2, pp. 321–336, Sep. 2003.
[28] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, 2003.
[29] S. Zhang, G. Liang, S. Pan, and L. Zheng, "A fast medical image super resolution method based on deep learning network," *IEEE Access*, vol. 7, pp. 12319–12327, 2019.
[30] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
[31] C. Gosling, "Encyclopedia of distances," *Reference Rev.*, vol. 24 no. 6, pp. 34–34, 2010, doi: 10.1108/09504121011067175.
[32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[33] S. Eger, S. Candemir, S. Antani, Y. X. Wáng, P. X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quant. Imag. Med. Surg.*, vol. 4, no. 6, pp. 475–477, 2014, doi: 10.3978/j.issn.2223-4292.2014.11.20.

FARAH DEEBA (Member, IEEE) received the B.E. and M.E. degrees from the Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan, in 2010 and 2014, respectively.

She has three years of experience in teaching (graduate and postgraduate level) at Hamdard University, Karachi, Pakistan. She is currently pursuing the Ph.D. degree with the School of Information and Software Engineering, University of Electronic Science and Technology, China. She also worked collaboratively as a Research Scientist with the Computer Network Information Center, Chinese Academy of Sciences, Beijing, China. Her research interests include super-resolution, image processing, deep learning, remote sensing, and machine learning.

SHE KUN received the B.Sc. degree from the University of Electronic Science and Technology, China, the M.S. degree from the Southwest Institute of Communications, and the Ph.D. degree in computer science from the University of Electronic Science and Technology, China. He is currently a Professor with the University of Electronic Science and Technology of China, Chengdu, China. His research interests include intelligent computing, cloud computing and big data, network security, and network engineering.

FAYAZ ALI DHAREJO (Member, IEEE) received the B.E. degree in electronic engineering from the Quaid-e-Awam University of Engineering, Science and Technology (QUEST), Pakistan, in 2016, and the M.S. degree from the School of Information and Software Engineering, University of Electronic Science and Technology of China (UESTC), in 2018.

Since September 2018, he has been with the Computer Network Information Center, Chinese Academy of Sciences. His research interests are image processing, deep learning, remote sensing, and machine learning.

YUANCHUN ZHOU received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, in 2006. He is currently a Professor with the Computer Network Information Center, Chinese Academy of Sciences. He has published more than 90 articles. His main research interests include data mining, data intelligence, massive data, and data intensive computing.