

Received August 29, 2018, accepted September 18, 2018, date of publication September 24, 2018,  
date of current version February 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2871626

# A Fast Medical Image Super Resolution Method Based on Deep Learning Network

SHENGXIANG ZHANG, GAOBO LIANG<sup>✉</sup>, SHUWAN PAN, AND LIXIN ZHENG<sup>✉</sup>

Fujian Provincial Academic Engineering Research Centre in Industrial Intellectual Techniques and Systems, College of Engineering, Huaqiao University, Quanzhou 362021, China

Corresponding author: Lixin Zheng (zlx@hqu.edu.cn)

This work was supported in part by the Science and Technology Bureau of Xiamen under Grant 3502Z20173045 and in part by the Technology Bureau of Quanzhou under Grant 2017G036.

**ABSTRACT** Low-resolution medical images can hamper medical diagnosis seriously, especially in the analysis of retina images and specifically for the detection of macula fovea. Therefore, improving the quality of medical images and speeding up their reconstruction is particularly important for expert diagnosis. To deal with this engineering problem, our paper presents a fast medical image super-resolution (FMISR) method whereby the three hidden layers to complete feature extraction is as same as the super resolution convolution neural network. It is important that a well-designed deep learning network processes images in the low resolution instead of the high-resolution space and enables the super-resolution reconstruction to be more efficient. Sub-pixel convolution layer addition and mini-network substitution in the hidden layers are critical for improving the image reconstruction speed. While the hidden layers are proposed for ensuring reconstruction quality, our FMISR framework performs significantly faster and produces a higher resolution images. As such, the technique underlying this framework presents a high potential in retinal macular examination as it provides a good platform for the segmentation of retinal images.

**INDEX TERMS** Super resolution, medical imaging, deep learning, medical diagnosis.

## I. INTRODUCTION

In medical image analysis, the typical common medical imaging systems that are utilized for expert diagnosis are nuclear Magnetic Resonance Imaging (MRI) [1], Computed Tomography (CT) [2], Positron Emission Computed Tomography (PET-CT) [3] and Ultrasound (US) [2]. However these images have low resolution, inherent noise, and lack of structural information. Due to hardware devices and existing imaging technology limitations, image super-resolution processing is favored by medical experts and researchers for its advantages of being intuitionistic, noninvasive, convenient and secure [4].

Single image super-resolution (SISR) technology principally is partitioned into three aspects: edge-based [5], image-based statistics [6]–[9], and sample-based method [10]–[16]. Among them, a sparse representation proposed by Yang *et al.* [17] has always occupied the dominant position in the field of super-resolution restoration. Yang *et al.* [18] also published a classic paper on image super-resolution, named image super-resolution via sparse representation, in 2010. This paper researched on image statistics and suggested that

image patches can be well-represented as a sparse linear combination of elements from an appropriately chosen over-complete dictionary. Inspired by this observation, we seek a sparse representation for each patch of the low-resolution input, and then utilize the coefficients of this representation to generate the high-resolution output. Theoretical results from compressed sensing suggest that under mild conditions, the sparse representation can be correctly recovered from the down-sampled signals. However, in this reconstruction algorithm, the coefficient regularization outcome is not manifest. Xiaoyan *et al.* [19] also proposed an improved algorithm that introduces the local constraint of weighting. However, when processing of some images to achieve more details, it tends to trigger excessive texture.

Propelled by the drive of big data such as ImageNet database, the image super-resolution method based on deep neural network is promising, which greatly promotes the prosperity of image reconstruction algorithm. In 2016, Zhao *et al.* [20] proposed a novel technique in order to address the problem of super-resolution (SR) for medical ultrasound (US) images. Tajbakhsh *et al.* [21] discussed

whether using full training or fine tuning in convolutional neural networks is better for medical image analysis. But one of the fatal flaws in a deep neural network is the computational cost.

Image Super-Resolution Using Deep Convolutional Networks (SRCNN) [22] is the earliest ancestor of deep learning method, which aims at recovering a high-resolution image from a single low-resolution image through Convolution Neural Network (CNN) [23]. There have been a few studies of using deep learning techniques for image restoration. The deep models in these methods are not specifically designed to be an end-to-end solution. On the contrary, the proposed SRCNN optimizes an end-to-end mapping. Because of its simple convolution neural network structure, it can be used to cope with the issue of image segmentation. Furthermore, the SRCNN is faster in terms of speed. This method can also be employed in other fields of object recognition. SRCNN firstly uses a bicubic interpolation to amplify its size. Secondly, it performs nonlinear mapping through a three-layer convolutional neural network. The resulting output is used as a reconstructed high-resolution image. The whole process can be divided into two parts: 1) Patch Extraction and Representation, as well as 2) Non-Linear Mapping and Reconstruction.

The advantage of SRCNN lies in the simplicity of its three-layer convolutional neural network, ease of convergence, low computational complexity, and ability to quickly reconstruct high-resolution image while maintaining high quality. Nevertheless, due to its relatively shallow network, the image features required for reconstruction cannot be extracted effectively; although larger cores can reduce the amount of computation, a large amount of information will be lost in each convolution, which results a poverty reconstruction image in the end. For example, the ringing effect is caused by selecting an inappropriate image model in image restoration. The direct cause of the ringing effect is the loss of information during image degradation, especially due to the loss of high frequency information.

In various experiments, the SRCNN shows extensibility and portability. The researchers investigate the impact of using different datasets on the model performance. Next, they explore different architecture designs of the network, and study the relations between super-resolution performance and factors like depth, number of filters and filter sizes. Finally, the SRCNN is extended to cope with color images and evaluate the performance on different channels.

There are many different ways for super-resolution reconstruction. Shi *et al.* [24] proposed a method called Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network (ESPCN). However, compared with the method of SRCNN, the ESPCN method is lack of contextual information after reconstruction. Meanwhile it is not enough to express the characteristics of object. In 2017, Gao *et al.* [4] utilized a deep convolutional network for medical image super-resolution that is an improved SRCNN algorithm. The reconstructed CT images

**TABLE 1. Mini-network and 5\*5 models.**

	ESPCN (5*5)	Our (3*3+3*3)
Time Complexity	$O(N^2 * (5*5*64)) = O(1600 * N^2)$	$O(N^2 * (3*3+3*3) * 64) = O(1152 * N^2)$
Parameter	$5*5+1 = 26$	$3*3+1+3*3+1 = 20$
Calculation	$25(I-4)^2$	$9(I-2)^2+9(I-4)^2$

can clearly provide an important reference for clinicians to make the correct treatment decisions. Although it is able to achieve a better quality, it costs significant time. Reducing the time to rebuild images has become a problem that needs to be solved urgently.

In this paper, we focus on shortening the time of image reconstruction and optimize the structure for speed. An efficient structure for reconstruction named fast medical image super resolution (FMISR) is proposed. It is a combined sub-pixel convolutional layer and mini-network in order to shorten the time of super-resolution. In addition, we implemented hidden layers to remain the information while training the images for improving the quality of the reconstruction. Next, we address the problem on how to obtain the quality of an image. In particular, the Peak Signal to Noise Ratio (PSNR) is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. And it is the most common and widely used objective measurement method of quality evaluation [25].

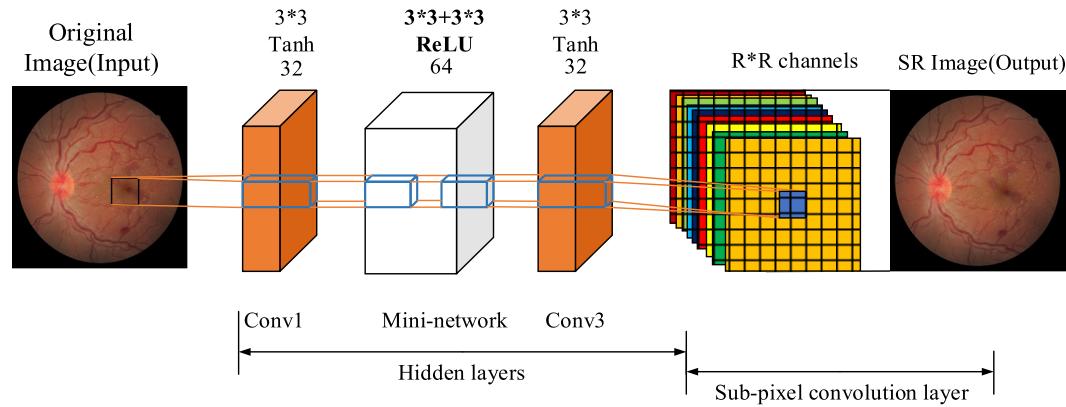
In the following, we will first present the structure of FMISR and then illustrate why it gets the speed improved and the details of every component in it. Next, we demonstrate how to conduct experiments and elaborate what the experimental results will be in the next two sections. Through discussion, we draw a conclusion for this paper and introduce the future work.

## II. FAST MEDICAL IMAGE SUPER RESOLUTION BASED ON DEEP LEARNING

As showed in Fig.1, the proposed fast medical image super resolution method is based on a well-designed deep learning network, which comprises three components, i.e., sub-pixel convolutional layer, mini-network and hidden layers. Among these components, the mini-network and the sub-pixel convolution layer are designed for improving the reconstruction speed, since the mini-network is a small convolution neural network and the sub-pixel convolution layer can be directly used as the super resolution image output layer. Then, in the following sections, these components will be introduced in detail.

### A. SUB-PIXEL CONVOLUTION LAYER

In the last layer, we applied the sub-pixel convolution layer that is proposed by Shi *et al.* [24] who implemented an efficient neural network (ESPCN) to reconstruct a low-resolution



**FIGURE 1.** Structure framework of FMISR.

image. In contrast to SRCNN [22], the reconstruction of the Shi model is directly in the low-resolution space [24]. The sub-pixel convolution layer in the model can be indirectly amplification process of the implementation of the image. It achieves a high-resolution image from low-resolution feature maps directly with one upscaling filter for each feature map ( $R \times R$  channels) as shown in Fig.1. In a convolution layer, different convolution kernel  $W$  of size  $k$  can be activated in low-resolution space. The number of activation patterns is exactly  $R \times R$  (as shown in Fig 1). In the position of activation patterns, there are  $[k/R]^2$  weights activated. These patterns are periodically activated during the convolution of the kernel across the image depending on the different sub-pixel location:  $\text{mod}(x, r)$ ,  $\text{mod}(y, r)$  where  $(x, y)$  are the output pixel coordinates in high-resolution space to rearrange the elements. The key operator of sub-pixel convolution layer is a periodic shuffling operator (PS). A periodic shuffling operator is to replace the elements of a  $H \times W \times C \times R^2$  tensor to a tensor of shape  $RH \times RW \times C$ . Here,  $C$  means color of the image. Because the periodic shuffling operator reconstructs a high-resolution image from low-resolution feature maps directly without convolution computational cost, it costs less time compared with other operators. Mathematically, this operation can be described in the following way, and the T means transfer:

$$PS(T)_{x,y,c} = T_{\left[\frac{x}{R}\right], \left[\frac{y}{R}\right], c \cdot R \cdot \text{mod}(y, R) + c \cdot \text{mod}(x, R)} \quad (1)$$

## B. MINI-NETWORK

In order to shorten the time of super-resolution, two  $3 \times 3$  convolution kernels cascade is nested in hidden layers, named the mini-network. After analyzing the SRCNN 9-5-5 model, the second layer can achieve a better feature map with the  $5 \times 5$  convolution kernel. We replace the  $5 \times 5$  convolution kernel by this mini-network in order to obtain the same consequence but on a much faster basis.

Large convolution kernel can achieve a greater receptive field, but also adds more parameters, then increases the amount of computation. Since the number of parameters is related to the convolution kernel size, a small convolution

kernel is advantageous. Note that we use the ReLU activation function to extract the non-linear feature, whereby its calculation is lower than the Tanh activation function. This is because the ReLU activation function only determine whether the input is greater than zero. The details of the image are related to the receptive field extracted by the filter. In mini-network, the two  $3 \times 3$  convolution kernels cascade ensure the same receptive field using the  $5 \times 5$  convolution kernel.

According to time complexity formula (2) and the convolution formula (3), we make a calculation comparison and parameter comparison with ESPCN shown in Table 2.

$$T = O(N^2 * K^2 * F) \quad (2)$$

$$\text{Output} = \frac{I - K - 2 * P}{S} + 1 \quad (3)$$

In time complexity formula (2),  $T$  means the time complexity;  $N$  is the size of input image;  $K$  is the convolution kernel size, where  $F$  is the number of filters. In the convolution formula (3), Input image size is  $I$ ;  $K$  is the convolution kernel size;  $P$  means padding which is the addition of an extra layer of zero between the images to make the output image of the same size as the input.  $S$  represents the step length of the filter in both horizontal and vertical directions in the original image. For convenience of discussion we assume that  $P = 0$ ,  $S = 1$ .

We computed the time complexity and the number of parameters in the mini-network and  $5 \times 5$  convolutional kernel respectively. It is obvious that mini-network has an advantage in the performance of both. We compare the results of the calculation, the value of  $I$  is greater than 10, the same  $I$  for this mini-network is smaller than the  $5 \times 5$  model in terms of multiplication and addition.

## C. HIDDEN LAYERS

Deep architecture is composed of multiple layers of parameterized non-linear modules, and the parameters of every module are subjected to learning. The more hidden layers are being added, the more features the network will be learned. Moreover, we added a new layer to exploit the inner high-frequency components compared with ESPCN.

**TABLE 2.** Parameters of hidden layers.

Details of hidden layers Parameters		
Conv1	<b>3*3*32 and Tanh</b>	
Mini-network	<b>(3*3+3*3) *64 and ReLU</b>	
Conv3	<b>3*3*32 and Tanh</b>	

In ESPCN, the convolution kernel's number is 64 and 32 respectively, the two layers applied the Tanh activation function. Our hidden layers are composed by three convolution layers. Here, when more convolution neural network layers are added, the more characteristics can be extracted from an image. The Conv1 means the first layer in hidden layers. Then, the Conv1 and Conv3 contain the same convolution kernel number (32), the same size of convolution kernel (3\*3) and the same activation function (Tanh). In the middle of the mini-network, involves two convolution kernels cascade with the size of 3\*3 convolution kernel and ReLU activation function. The parameters that we have set are shown in Table 2.

### III. EXPERIMENTS AND RESULTS

The experimental environment includes hardware devices and software configurations. The computer configuration is Intel(R) Core(TM) i7-6700CPU@3.40GHz, and the GPU is NVIDIA GeForce GTX 1050-Ti. The experimental platform is equipped with 64-bit Windows 7, Caffe, MatlabR2016a, CUDA Toolkit v8.0, and Anaconda2.

#### A. SETUP

The whole network in our setup includes the training network and testing network. In training network, the basic learning rate is  $10^{-4}$ , the weight decay is zero and momentum is 0.9 in the process of iteration. The gradient descent algorithm we choose is Mini-Batch Gradient Descent (MBGD). The pseudocode is as follows:

#### Algorithm 1 Mini-Batch Gradient Descent

```

Repeat
{
    For i = 1,11, 21,31,..., 911
    {
         $\theta_j := \theta_j - \alpha \frac{1}{10} \sum_{k=i}^{i+9} (h_\theta(x^{(k)}) - y^{(k)}) x_j^{(k)}$ 
        (for every j = 0,...,n)
    }
}

```

In this pseudocode,  $\sum_{k=i}^{i+9} (h_\theta(x^{(k)}) - y^{(k)}) x_j^{(k)}$  defines the loss function. If the number of samples in the training set is 1000, then each mini-batch is only a subset, assuming that there are 10 samples in each mini-batch. In this way, the entire training data set can be divided into 100 mini-batches. When there is a flat area in the error surface, mini-batch gradient descent can learn faster.

In every convolution layer, we implemented the Gaussian distribution to initialize weights which is showed in

**TABLE 3.** PSNR(dB) and SR-Time for different datasets in 300000 iterations.

Dataset	Scale	Bicubic	ESPCN	OUR
		PSNR	PSNR/SR-Time	PSNR/SR-Time
Brain	3	24.553	25.080/0.298s	<b>25.502/0.220s</b>
Abdomen1	3	27.815	28.829/0.261s	<b>29.891/0.248s</b>
Abdomen2	3	26.525	27.504/0.270s	<b>28.161/0.267s</b>
Knee	3	32.898	35.131/0.309s	<b>35.309/0.227s</b>
Cell	3	26.562	27.964/ <b>0.264s</b>	<b>27.980/0.274s</b>
SR-Time/s	None	None	28.901/0.280s	<b>29.368/0.240s</b>

formula (4) [26].

$$X \sim N(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

where  $\mu$  is the mean or expectation of the distribution,  $\sigma$  is the standard deviation, and  $\sigma^2$  is the variance in it.

At the same time, batch size in the training data sets is 128 and that of the testing data sets is 32. The Euclidean Loss function is used to compute the loss between predictive value and label value which is defined by [27]:

$$\text{Euclidean Loss} = \frac{1}{2N} \sum_{n=1}^N \|\hat{y}_n - y_n\|_2^2 \quad (5)$$

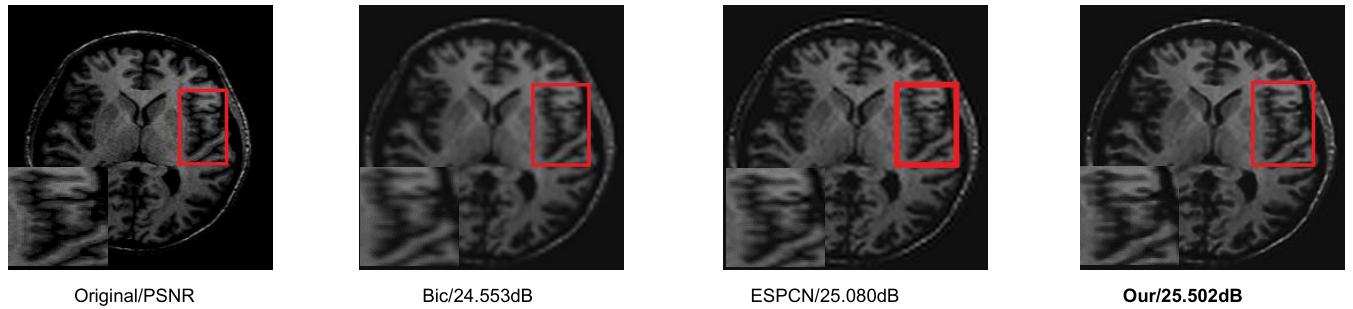
whereby  $N$  means the total number of input image,  $n$  is the number of input image,  $\hat{y}_n$  is the predictive value and  $y_n$  is the label value.

#### B. DATASET AND PROTOCOL

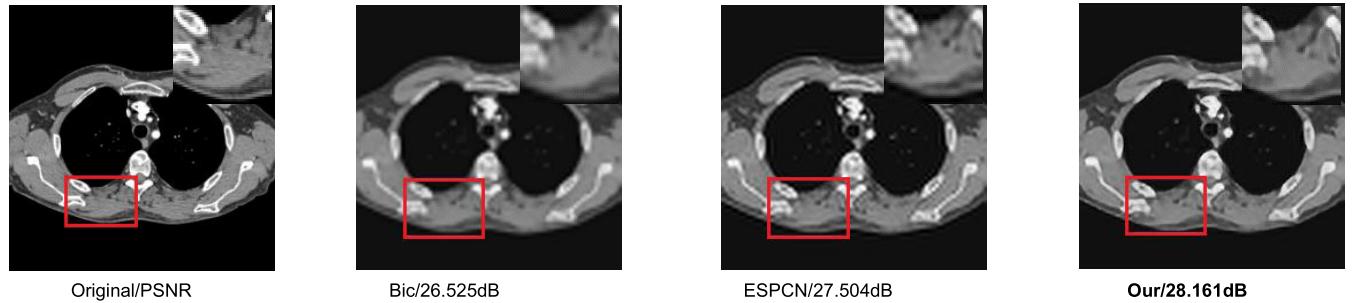
During the training phase, the publicly available benchmark datasets contain Timofte dataset [15], 91 training images and two test datasets Set5 and Set14 which provide 5 and 14 images. The Berkeley segmentation dataset [28] is BSD300 and BSD500, which provide 100 and 200 images for testing. And the super texture dataset [16] that provides 136 texture images is used.

#### C. IMAGE RECONSTRUCTION SPEED RESULT

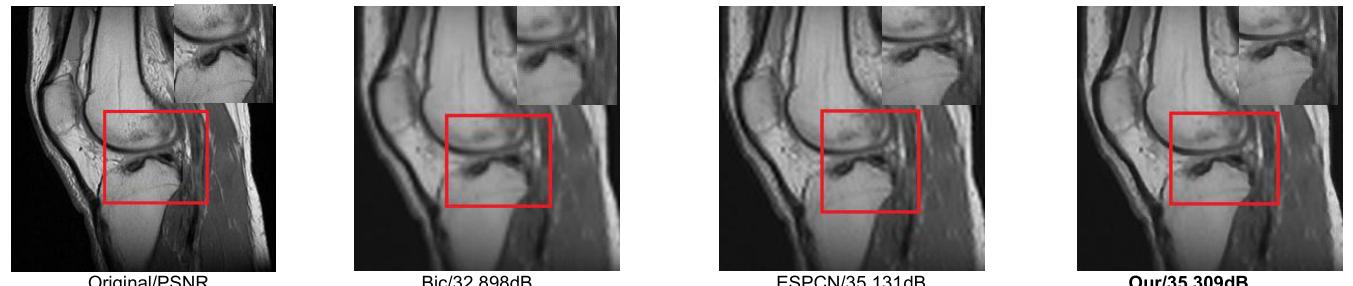
In this section, we evaluated our model's run time on publicly available dataset IDI (I Do Imaging) [30] with an upscale factor of 3. IDI is the resource location to find the free and the open source medical imaging software, whereby one can find nearly 300 software projects that are neatly categorized, ranked, and searchable. Implementation of both ESPCN and FMISR algorithms is based on the Caffe codes in the computer with GTX 1050-Ti. Therefore, the same experimental environment is guaranteed to ensure that there is only one variable. The results are presented in Table 3. Compared to ESPCN model, our model has two acceleration modules that are optimized for both the number of parameters and the design of the structure, making it a lightweight network structure.



**FIGURE 2.** The brain image from Brain dataset with an upscaling factor 3.



**FIGURE 3.** Super-resolution examples for abdomen from public dataset Abdomen1 based on an upscaling factor of 3 with their corresponding PSNR values shown under each sub-figure.



**FIGURE 4.** Knee image processing based on a public Knee dataset based on an upscaling factor 3 with their corresponding PSNR values shown under each sub-figure.

#### D. IMAGE RECONSTRUCTION QUALITY RESULT

We record the value of the PSNR [25] that is the objective criteria for measuring image distortion or noise levels. The PSNR is for grey-level (8 bits) images. In the formula (5), given input image  $f$  and the reconstructed image  $f'$ , both of size  $M * N$ , the PSNR between  $f$  and  $f'$  is defined by:

$$PSNR = 10 \log_{10} \frac{(255)^2}{MSE} \quad (6)$$

where

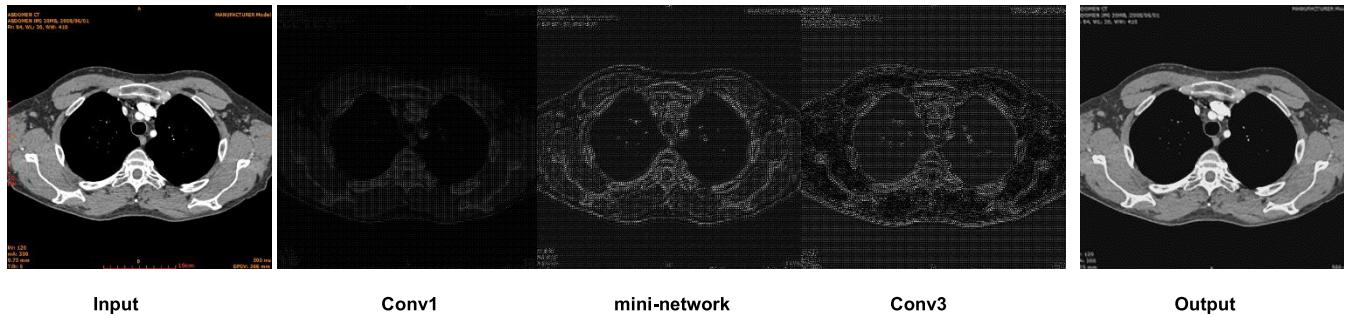
$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f - f')^2 \quad (7)$$

The larger PSNR value between two images shows a higher image quality. The common reference is 30dB, and the image deterioration is obvious below 30dB. And Fig.2, Fig.3 and Fig.4 are the visualization of super-resolution, which compared the ESPCN and our FMISR method. The process of super-resolution can also be visualized in Fig.5.

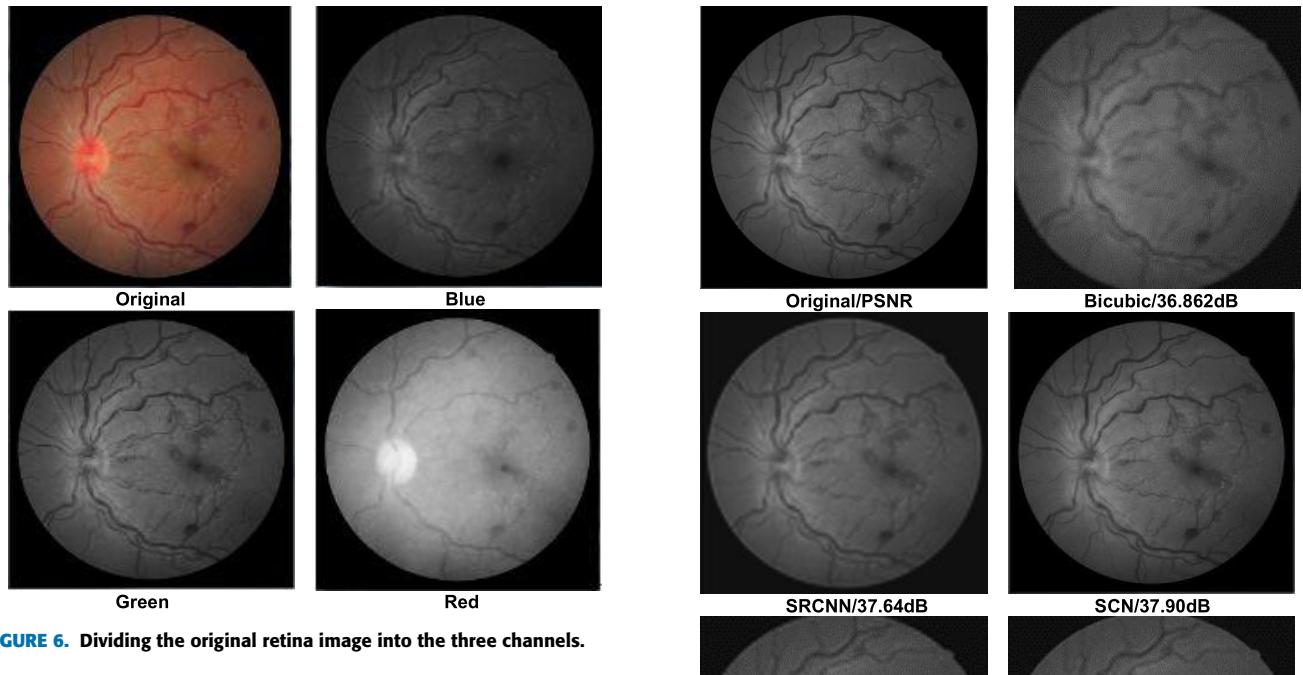
#### E. SUPER-RESOLUTION IN THE LOW RESOLUTION RETINAL IMAGE FOR THE DETECTION OF MACULA FOVEA

We studied the mechanism of diabetic retinopathy as follows. In the fundus image, the macula lies slightly below the 3.5mm disc and has a brown oval structure. It is the most sensitive part of visual acuity whereby any damage that occurs in the center of the macula may cause vision impairment or even blindness. We started to reconstruct high-resolution images from low-resolution retinal images so that the macular center will become more apparent. In this way, it is not only allowed medical experts to quickly and accurately determine the cause, but also results in having the advantage of not needing high-resolution camera.

*Step 1:* The dataset is a Digital Retinal Images for Vessel Extraction (DRIVE) [31]. Here, we truncated the pictures into a 256\*256 low-resolution picture using OpenCV and divide each picture image into three channels. The vascular profile



**FIGURE 5.** Process of the Super-Resolution technique.

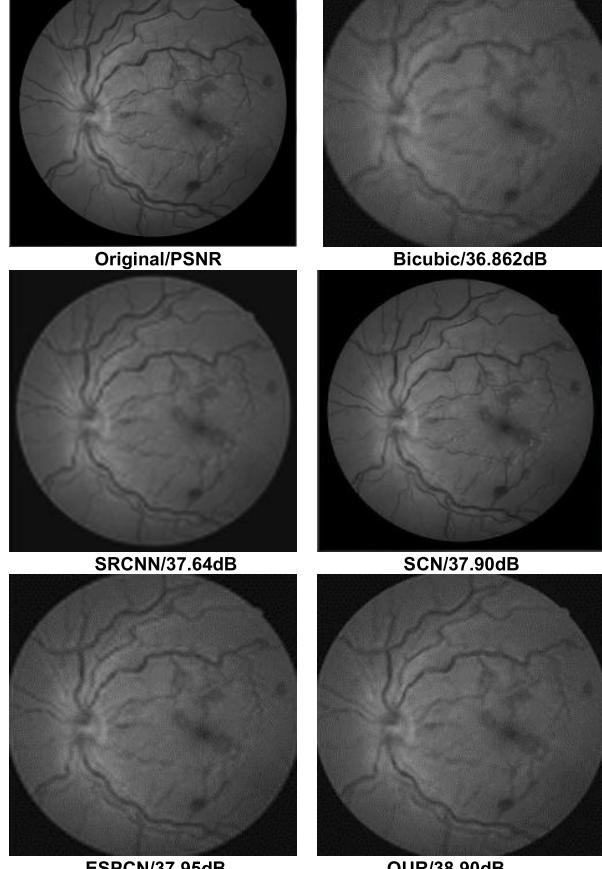


**FIGURE 6.** Dividing the original retina image into the three channels.

and edges in the green channel are much clearer and the contrast is higher; The red channel is brighter but the blood vessels are not obvious; The blue channel is less noisy compared to the red channel; however, the contrast of the blood vessel and the background is not significant enough, which is not conducive to the later segmentation work. We choose the green channel as an ideal input image. Note that the three channels are shown in Fig.6.

Step 2: Reconstruct a single retinal image using multiple reconstruction algorithms based on CNN method. In the choice of contrast method, SCN (sparse coding-based network) [33], SRCNN [22] and ESPCN [24] are selected. Because they are the algorithms by learning features to reconstruct a single image.

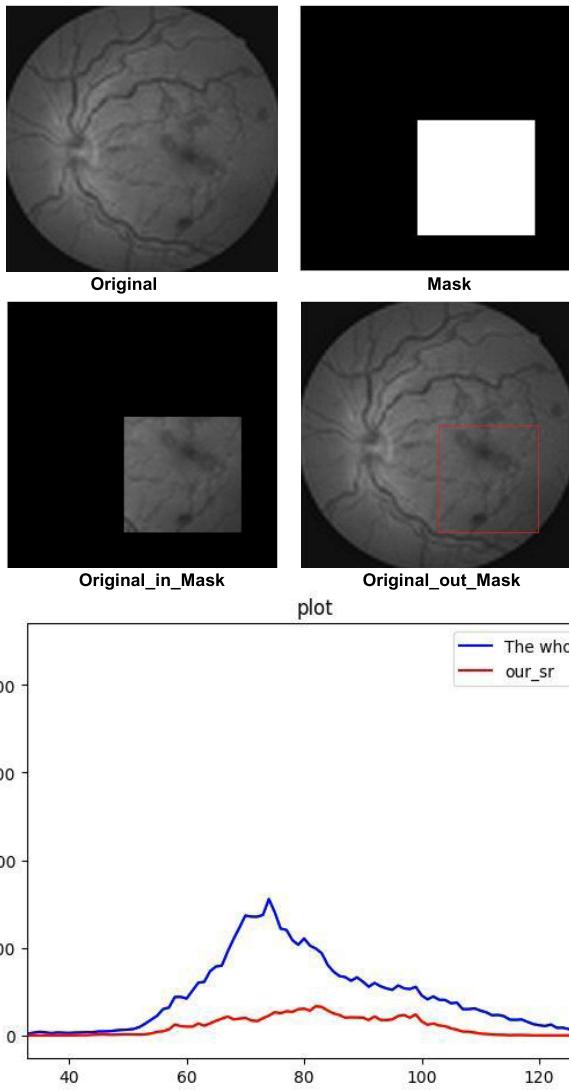
Step 3: In order to compare the advantages of our reconstruction algorithm in macular detection, we implemented masking techniques to count the high frequency information in local regions of the image. As shown in Fig.8, we utilized a mask to count the grayscale histogram of an area near the macula, and we observe the proportion and information of our algorithm in the histogram information of the entire image.



**FIGURE 7.** Different methods in reconstructing a single retinal image.

The blue curve expresses the high frequency information of the whole picture, and the red curve means the high frequency information of the area we masked.

The graph in Fig.9 shows the histograms of the gray histograms in the mask area computed using five algorithms. The red color curve is the line that pertains to our FMISR algorithm. We have found that our network structure performs better at brightness sensitive sites than other algorithms. We discussed the FMISR's PSNR value and reconstructing time using the three different CNN methods. Performance parameters are presented in Table 4.

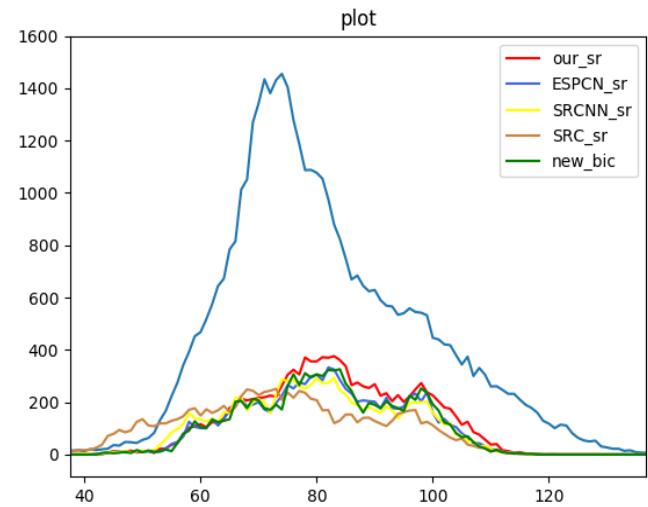
**FIGURE 8.** Grayscale histogram of an area near the macula using FMISR.

## IV. DISCUSSION

### A. IMAGE RECONSTRUCTION SPEED IMPROVED

In the traditional super-resolution reconstruction algorithms introduced above, the reconstruction time has always been an aspect neglected by researchers. Traditional algorithms have been devoted to improving algorithms to increase the quality of super-resolution reconstruction. However, with the prospering and promotion of deep learning algorithms, the SRCNN algorithm based on deep learning has broken the record of most traditional algorithms for reconstruction quality. More importantly, it has greatly improved the speed of processing. Then, researchers started to study how to achieve a faster reconstruction speed while guaranteeing the quality of reconstruction.

In the Table 4, we utilized bold fonts to indicate the figure values pertaining to the better reconstruction results. We test the time for the reconstruction of the algorithm based on five different test datasets, and the results were positive on all four of the test datasets. Our average common speed

**FIGURE 9.** Comparison curves of five types of algorithms.**TABLE 4.** Comparing performance with different super-resolution methods.

Method	Scale	Bicubic	SCN	SRCCNN	ESPCN	Our
PSNR/ dB	3	36.862	37.901	37.648	37.951	<b>38.908</b>
SR-Time/s	3	0.025	4.957	2.460	0.346	<b>0.293</b>

is increased by 0.24s, which is used to reconstruct a single image from IDI (I Do Imaging) [30] on 1050-Ti GPU. Compared with the ESPCN algorithm in the Matlab codes that are provided by [24], the time of FMISR super-resolution reconstruction of single image is reduced by 40ms.

### B. IMAGE RECONSTRUCTION QUALITY IMPROVED

The improvement of the quality of super-resolution reconstruction is an eternal topic. From the edge-based reconstruction of traditional algorithms to the reconstruction of sparse dictionaries, emphasis has been placed on quality improvement. In the algorithm published by Wang *et al.* [33] in 2010, the traditional algorithm has pushed the quality of super-resolution reconstruction to a peak. But later on, the subsequent deep learning reconstruction algorithm proposed by Dong *et al.* [22] raised a new climax. The quality of reconstruction continues to rise.

Regarding the quality of reconstructed image, we evaluate the quality of reconstructed images from two aspects. On one hand, the value of PSNR increased +0.467 dB on the dataset IDI. On the other hand, we visualized the reconstructed images in Fig 2, Fig 3 and Fig 4. The process of reconstruction can also be visualized in Fig 5. Based on the results, the subjective visual perception is superior to other methods in terms of performance.

### C. APPLY TO RETINAL IMAGE FOR THE DETECTION OF MACULA FOVEA

The super resolution reconstruction algorithm is changing rapidly and is becoming more and more widely used. In the medical field, the improved algorithm based on SRCNN proposed by Zhao *et al.* [20] has also been applied in CT medical images. We used open medical image datasets to test a variety of super-resolution algorithms. In the case of diabetic retinopathy, we found the center of the macula by super-resolution reconstruction and achieved effective results.

Applying this method to retinal image for the detection of macula fovea immensely reduce the burden of judgment on medical images by researchers and medical experts. We also compared and evaluated many reconstruction algorithms, and in our method, the results of PSNR are much higher than others in quality. The reconstruction speed is enormously faster than traditional reconstruction algorithms.

To prove the superiority of the algorithm, we implemented a set of application experiments based on the retinal images. For a better detection of macula fovea, we attempt to reconstruct an acceptable image as much as possible. We compared the three super-resolution reconstruction methods by learning, SCN [33], SRCNN [22], ESPCN [24] and a traditional interpolation-based function, Bicubic [32]. From the Fig 9, the red color line that pertains to our algorithm (FMISR) is obviously higher in terms of performance in comparison with the other lines. This shows that in the whole picture image, the high frequency information in the reconstructed image processed by the algorithm is significantly increased. Note that we can also draw conclusions from the data in Table 4. Our reconstruction model runs on a magnitude that is much faster than SCN model, SRCNN model and ESPCN model (+4.664s, +2.167s and S+0.053s improved). In terms of reconstruction quality, our model achieves a much better super-resolution performance (+1.007dB, +1.260dB and +0.957dB) in comparison to its counterparts.

### V. CONCLUSION

In this paper, we demonstrated an efficient network model specifically for medical image super-resolution method, which is based on an increased convolution layer in order to achieve better picture reconstruction. The time for image reconstruction has been significantly reduced by 50 ms compared to the ESPCN in a single low resolution retinal image reconstruction. In addition, our mini-network has a higher speed improvement in the field of super-resolution. Note that we demonstrate the positive effect of the subpixel convolution layer as well as tanh activation function. We used different activation function in the same structure, and the results show that the Tanh activation function performs incomparably better than ReLU activation.

Moreover, the optimization of the structure is essential to determine an optimal receptive field, and a good receptive field can more effectively extract the image features required. The PSNR after image reconstruction has been improved by

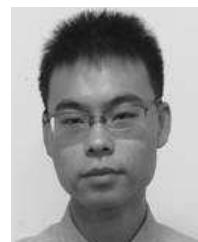
0.95dB when compared with the ESPCN based on the low resolution retinal image. In our study, we also used the masking technique to test the luminance sensitive area. According to our experimental results, our reconstruction algorithm extracts a better receptive field in the training process of the neural network where the brightness changes obviously and results in a better reconstruction effect.

With regards to the application of deconvolution layers, multiple deconvolution layers are now used for the visualization of neural network. During the reconstruction process, the selection of the picture magnification is mostly implemented using the deconvolution layers and direct linear interpolation. However, these two methods do not have a good effect on reconstruction. In future studies, the implementation of deconvolution layers in this field of image processing can be used as a research motivation. Whether there is an optimized deconvolution layer that can contain the original image information is the question of interest in the research. We present a reasonable hypothesis stating that there are some relationships between convolution layer and deconvolution layer based on information from neural network training.

### REFERENCES

- [1] R. R. Peeters *et al.*, “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. Autom. Congr. (CAC)*, Jinan, China, Oct. 2017, pp. 5310–5315, doi: [10.1109/CAC.2017.8243724](https://doi.org/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] N. Efrat, D. Glasner, A. Apartsin, B. Nadler, and A. Levin, “Accurate blur models vs. image priors in single image super-resolution,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 2832–2839.
- [7] H. He and W.-C. Siu, “Single image super-resolution using Gaussian process regression,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 449–456.
- [8] J. Yang, Z. Lin, and S. Cohen, “Fast image super-resolution based on in-place example regression,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 1059–1066.
- [9] C. Fernandez-Granda and E. J. Candès, “Super-resolution via transform-invariant group-sparse regularization,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 3336–3343.
- [10] H. Chang, D.-Y. Yeung, and Y. Xiong, “Super-resolution through neighbor embedding,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun./Jul. 2004, p. 1.
- [11] S. Wang, L. Zhang, Y. Liang, and Q. Pan, “Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 2216–2223.
- [12] K. Zhang, X. Gao, D. Tao, and X. Li, “Multi-scale dictionary for single image super-resolution,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 1114–1121.
- [13] X. Gao, K. Zhang, D. Tao, and X. Li, “Image super-resolution with sparse neighbor embedding,” *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3194–3205, Jul. 2012.

- [14] Y. Zhu, Y. Zhang, and A. L. Yuille, "Single image super-resolution using deformable patches," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2917–2924.
- [15] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. Asian Conf. Comput. Vis. (ACCV)*. Cham, Switzerland: Springer, 2014, pp. 111–126.
- [16] D. Dai, R. Timofte, and L. Van Gool, "Jointly optimized regressors for image super-resolution," in *Proc. Comput. Graph. Forum*, 2015.
- [17] J. Yang, J. Wright, T. S. 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.
- [18] 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.
- [19] Z. Xiaoyan et al., "An improved sparse representation super resolution reconstruction algorithm," *J. Chongqing Univ. Posts Telecommun. (Natural Sci. Ed.)*, 2016.
- [20] N. Zhao, Q. Wei, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Single image super-resolution of medical ultrasound images using a fast algorithm," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Apr. 2016, pp. 473–476.
- [21] N. Tajbakhsh et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016, doi: 10.1109/TMI.2016.2535302.
- [22] 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. 2015.
- [23] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [exploratory DSP]," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 145–154, Jan. 2011.
- [24] W. Shi et al., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. Comput. Vis. Pattern Recognit.*, 2016, pp. 1874–1883.
- [25] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2366–2369.
- [26] T. W. Anderson, "An introduction to multivariate statistical analysis," *Technometrics*, vol. 28, no. 2, pp. 180–181, 1958.
- [27] C. Gosling, "Encyclopedia of distances," *Reference Rev.*, vol. 24, no. 6, p. 34, 2009.
- [28] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int. Conf. Comput. Vis.*, vol. 2, Jul. 2001, pp. 416–423.
- [29] D. Dai, R. Timofte, and L. Van Gool, "Jointly optimized regressors for image super-resolution," *Eurographics*, vol. 7, p. 8, 2015.
- [30] *I Do Imaging (IDI), a Searchable Database of Free and Open Source Medical Imaging Software [DB/OL]*. Accessed: Apr. 26, 2018. [Online]. Available: <https://idointaging.com/>
- [31] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [32] C. de Boor, "Bicubic spline interpolation," *J. Math. Phys.*, vol. 41, nos. 1–4, pp. 212–218, 1962.
- [33] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang, "Deep networks for image super-resolution with sparse prior," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 370–378.



**GAOBO LIANG** was born in Xiao Gan, Hubei, China, in 1994. He is currently pursuing the M.S. degree in computer science and technology with the College of Engineering, Huaqiao University, China. From 2017 to 2018, he was a Research Assistant with the College of Engineering, Huaqiao University. His research interests include the applications of deep learning in medical image analysis and the applications of deep learning in rapid visual identification. He is currently specializing in the positioning of manufacturing components used by industrial sorting robots.



**SHUWAN PAN** was born in Yan Cheng, China, in 1982. He received the B.S. degree from Jiangsu Normal University in 2006 and the Ph.D. degree in microelectronics and solid state electronics from Xiamen University in 2011. Since 2011, he has been a Lecturer with the College of Engineering, Huaqiao University. His current research interests focus on photoelectric detection and machine vision technology.



**LIXIN ZHENG** was born in Fujian, China, in 1967. He is currently the Dean and a Professor with the College of Engineering, Huaqiao University. His research interests mainly focused on image recognition and machine vision technology. He is the Vice President of the Fujian Automation Society and the Xiamen Automation Society and the Director of the Fujian Power Supply Society. He received the bachelor's degree from the Department of Electronic Engineering, Huaqiao University, in 1987, the master's degree in testing and automation control from the Department of Mechanical Engineering, Huaqiao University, in 1990, and the Ph.D. degree in predictive control from Tianjin University in 2002. In 1997, he was in Japan for further study. He has published significantly in the fields of industrial robot vision systems. He was supported by the Fujian Excellent Talent Support Program for his research on robotics in 2007. He received the Award of the Outstanding Young and Middle-Aged Teachers of Huaqiao University in 2004 and the Outstanding Teacher of Huaqiao University Award in 2006.



**SHENXIANG ZHANG** was born in Jingzhou, Hubei, China, in 1994. He is currently pursuing the M.Eng. degree in computer science and technology with Huaqiao University. As of now, his research interests include computer vision and the applications of deep learning in medical image super-resolution reconstruction. Furthermore, he is currently specializing in the field of machine vision applied to industrial robots.