ORIGINAL PAPER

A New General Maximum Intensity Projection Technology via the Hybrid of U-Net and Radial Basis Function Neural Network

Zhen Chao[1,2,3] · Wenting Xu[3]

Abstract

Maximum intensity projection (MIP) technology is a computer visualization method that projects three-dimensional spatial data on a visualization plane. According to the specific purposes, the specific lab thickness and direction can be selected. This technology can better show organs, such as blood vessels, arteries, veins, and bronchi and so forth, from different directions, which could bring more intuitive and comprehensive results for doctors in the diagnosis of related diseases. However, in this traditional projection technology, the details of the small projected target are not clearly visualized when the projected target is not much different from the surrounding environment, which could lead to missed diagnosis or misdiagnosis. Therefore, it is urgent to develop a new technology that can better and clearly display the angiogram. However, to the best of our knowledge, research in this area is scarce. To fill this gap in the literature, in the present study, we propose a new method based on the hybrid of convolutional neural network (CNN) and radial basis function neural network (RBFNN) to synthesize the projection image. We first adopted the U-net to obtain feature or enhanced images to be projected; subsequently, the RBF neural network performed further synthesis processing for these data; finally, the projection images were obtained. For experimental data, in order to increase the robustness of the proposed algorithm, the following three different types of datasets were adopted: the vascular projection of the brain, the bronchial projection of the lung parenchyma, and the vascular projection of the liver. In addition, radiologist evaluation and five classic metrics of image definition were implemented for effective analysis. Finally, compared to the traditional MIP technology and other structures, the use of a large number of different types of data and superior experimental results proved the versatility and robustness of the proposed method.

Keywords Maximum intensity projection · Convolutional neural network · Radial basis function neural network · Projection synthesis

Introduction

In the 3D visualization of medical imaging, the maximum intensity projection (MIP) method can provide more intuitive images, the display effect of which is similar to that afforded by X-ray imaging. The algorithm principle is simple, and the amount of calculation is small. More importantly, MIP can be displayed in real time, so it is widely used in the field of medical imaging, such as displaying the three-dimensional structure of blood vessels, details of bronchi, and so forth [1–3]. In addition, the MIP technology plays an important role in the optimal diagnosis of diseases or postprocessing of medical imaging. For instance, Zheng et al. [4] applied MIP images to convolutional neural networks (CNNs) to improve the effectiveness of automatic lung nodule detection. The annotations of 2D MIP images were adopted to enhance the performances of neural networks to better

Zhen Chao and Wenting Xu are regarded as co-first author.

✉ Zhen Chao
chaozhen@sdfmu.edu.cn

1 College of Artificial Intelligence and Big Data for Medical Sciences, Shandong First Medical University & Shandong Academy of Medical Sciences, Huaiyin District, 6699 Qingdao Road, Jinan 250117, Shandong, China

2 Present Address: Research Lab for Medical Imaging and Digital Surgery, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

3 Department of Radiation Convergence Engineering, College of Health Science, Yonsei University, 1 Yonseidae-gil, Wonju, Gangwon 26493, South Korea

segment the linear structures of 3D magnetic resonance angiography (MRA) images [5]. Furthermore, Harvey et al. [6] explored the effects of MIP images of cerebral CT angiography (CTA) based on photon counting CT and proved that the quality of CTA-MIP images is better than that of traditional CT images. Recently, the world has undergone tremendous changes due to the coronavirus disease 2019 (COVID-19). In order to solve this worldwide difficulty as soon as possible, the first important thing for scientists is to better understand the nature of this virus from different perspectives. The MIP technology plays a key role in better understanding and diagnosis of COVID-19 [7–10].

The MIP technology treats the voxel of the three-dimensional volume data as a small light source. Then, according to the theory of image space rendering, the light source emits rays along a certain direction. When the rays pass through the data field and encounter the maximum density value, the maximum value is projected in each pixel on the corresponding screen to form the final projection [11–13]. Therefore, based on this projection logic, MIP has several limitations, such as the overlap of blood vessels, bones, and internal organs or inconspicuous visualization of small blood vessels and trachea due to unrelated impurities with similar intensity. Accordingly, several improved algorithms have been proposed [14–17]; however, these distant algorithms are mainly aimed at reducing the amount of calculation or increasing the depth of projection. Therefore, after more than 10 years or even 20 years, these algorithms have not been widely used. In contrast, nowadays, the traditional MIP algorithm is still being used in 3D visualization of medical imaging. Based on these considerations, we assumed that the biggest function of MIP technology is to display the stenosis, dilation, and morphological direction of blood vessels or trachea. By now, the developed post processing-based high-performance software can display high-quality MIP images well [18]. If the three-dimensionality of the image needs to be displayed, the volume rendering (VR) method can replace the projection method. Therefore, in further research, it would be necessary to understand how to use a certain algorithm to generate the better-quality MIP images, which can more clearly show the tiny details. This "high quality-higher quality" research logic or direction can better promote clinically high-precision diagnosis. However, to the best of our knowledge, when deep learning has been becoming the mainstream in the recent years, very few studies have addressed the aforementioned issue.

Furthermore, artificial intelligence technology has been widely used in medical image processing, and various network structures have been proposed to perform image synthesis, image fusion, and so forth [19–22]. The logic of image synthesis or image fusion is to transform a medical image or several medical images into the desired medical image through neural network technology. In simple terms, the MIP technology aims to synthesize a projection image from volumetric data or continuous 2D data images.

Inspired by this, the purpose of the present study is to explore the potential of using neural networks to synthesize or generate projection images and to compare the results with those afforded by the traditional MIP technology. Accordingly, we propose a new network structure based on the hybrid of U-net and radial basis function neural network (U-RBFNN), which is a combination of deep neural network and shallow neural network. To this end, we first adopted a U-net to extract features of input images and generate enhanced or improved output images. Then, these output images were used as parallel inputs of radial basis function (RBF) neural network to synthesize the final projection image. The characteristic of the proposed U-RBFNN is to fully combine the learning capabilities of different types of neural networks. First, the feature prior points are extracted by the convolutional neural network, and then point-to-point fusion is performed through the shallow neural network whose running is based on the data points; finally, the learning abilities are superimposed. Additionally, our aim is to get higher-quality images than traditional MIP images. Accordingly, for training, we adopted a transfer learning style. We first performed a certain blur processing on the initial images and then used traditional MIP images as the gold standard for supervised training. For a better comparison, in addition to comparing the traditional MIP images, we also used two neural network networks based on image fusion for better observation. For experimental datasets, in order to increase the robustness and generality of the proposed algorithm, we used three open datasets to verify the proposed algorithm by using different human body parts and different slab thicknesses. Finally, in addition to applying radiologist subjective observation, to objectively compare the generated resultant image, five metrics were implemented to evaluate the performance of different methods. The results demonstrated that the performance of the proposed algorithm was significantly better than that of the traditional MIP technology and the other two neural network-based structures. Overall, the contributions of this study can be summarized as follows:

1. The present study is the first to introduce the neural network technology to synthesize the maximum intensity projection (MIP) images to achieve superior image performance.
2. This study is the first to combine the convolution-based U-net neural network and the radial basis function neural network (RBFNN). The results proved that the CNN can effectively activate the intelligence of the shallow neural network and achieve good effects.
3. The results of applying a large number of open-databases demonstrated robustness, generality, and applicability of the proposed algorithm.

Methods and Materials

The entire flow chart of the proposed hybrid of U-net and radial basis function neural network (U-RBFNN) is shown in Fig. 1. The data basis for the generation of the maximum intensity projection (MIP) images required a specific 3D volumetric data or continuous 2D data. Accordingly, we firstly substituted the 2D images used for MIP into the U-net network, and the output images with increased detail could be obtained after feature extraction and up-sampling. Subsequently, we substituted these output images into RBFNN in parallel and, finally, the MIP image was obtained by effectively running RBFNN.

Radial Basis Function Neural Network

Radial basis function neural network (RBFNN) is a single hidden layer, a feedforward neural network based on function approximation proposed in the late 1980s. The structure of classic RBFNN includes the following three layers: input layer, hidden layer, and output layer (see Fig. 2). With the maturity of technology, RBFNN has received considerable attention from researchers in various fields due to its simple structure, strong nonlinear approximation ability, and good generalization ability. It is widely used in many research
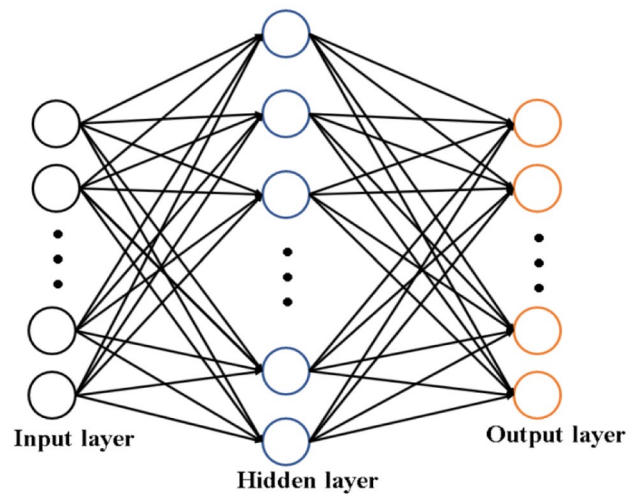
Fig. 2 The structure of the traditional radial basis function neural network (RBFNN), which includes three layers: input layer, hidden layer, and output layer

fields, including pattern classification, function approximation, and data mining [23–25]. In RBFNN, the Gaussian function is the most commonly used radial basis function to effectively activate the logical relationship between the input layer and the hidden layer [24]. The expression of the Gaussian function G is as follows (see Eq. 1):
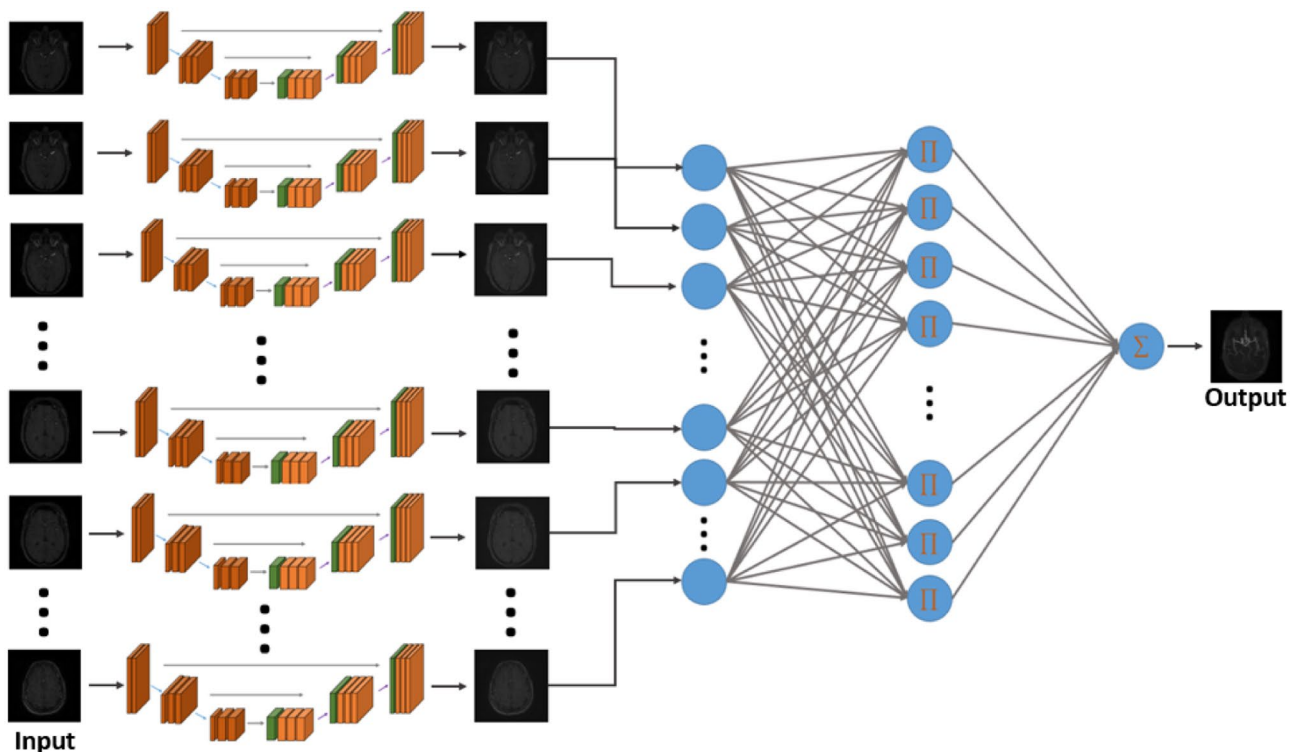
Fig. 1 The detailed flow chart of the proposed method

$$G(x_i, c_{pi}, \sigma_{pi}) = exp(-\frac{1}{2\sigma_{pi}^2} \| x_i - c_{pi} \|^2) \qquad (1)$$

where x is input variable, i is ith neuron of input layer; p is the pth neuron of hidden layer, which also corresponds to the pth Gaussian function. c_{pi} and σ_{pi} are the center and variance of the pth Gaussian function ith input neuron, and $\| \|$ is Euclidean norm.

After activating the correlation of each hidden layer neuron and each input layer neuron, the corresponding relationship between a certain hidden layer and the entire input layer was obtained using Eq. (2). Finally, the logical relationship between hidden layer neurons and output layer neurons was the linear weighting style (see Eq. 3).

$$R(x, \mu_p, \sigma_p) = \prod_{i=1}^{k} G(x_i, c_{pi}, \sigma_{pi}) \qquad (2)$$

$$Y_o = \sum_{p=1}^{l} \omega_{po} R(x, c_p, \sigma_p) \qquad (3)$$

where $R(x, \mu_p, \sigma_p)$ is the value of the pth hidden neuron. k is the number of input neurons. Y_o is the value of the oth output neuron. l is the number of hidden neurons. ω_{po} is the connection weight of the pth neuron of the hidden layer and the oth neuron of the output layer.

Accordingly, each input neuron of RBFNN was actually a variable value. In a previous study, we succeeded in substituting medical images of different modalities into RBFNN for medical image fusion [26]. From the point of view of variable points, we used the pixels at the same position of different modal images through neural network technology to synthesize the pixels at the corresponding

positions of the fused image. Therefore, following the same logic, in this study, our aim in the present study was to substitute the pixels at the same position of the continuous 2D data into the RBFNN in parallel. The output layer of the neural network is a neuron which represents a pixel point at the corresponding position of the MIP image. In turn, the entire projected image could be obtained. In addition, according to our previous results [25, 26], in medical image processing, one of the key aspects necessary to improve the intelligence of RBFNN is to effectively select or calculate the feature points of the pending images which are the neurons that make up the input layer of the neural network. Based on this, in the present study, we applied U-net neural network as a priori processing for the feature extraction of the input layer of RBFNN.

U-Net Neural Network

From the image level, to better highlight the detailed features of the image, many previous studies have applied U-net and got good effects [27, 28]. The structure of the U-net in this study, including an encoder path and a decoder path, is shown in Fig. 3. The encoder part adopted typical convolutional neural network structure, with 4 down-sampling and 5 layers of convolutions. Each group of convolution consisted of two convolution operations. The size of the convolution kernel was 3×3, followed by corresponding batch normalization operation and rectified linear unit (ReLU) activation function. In the 2nd layer to the 5th layer, there were corresponding down-sampling max pooling operation with a convolution kernel of 2×2 and a stride of 2. The purpose of down-sampling was that the size of the feature map continued
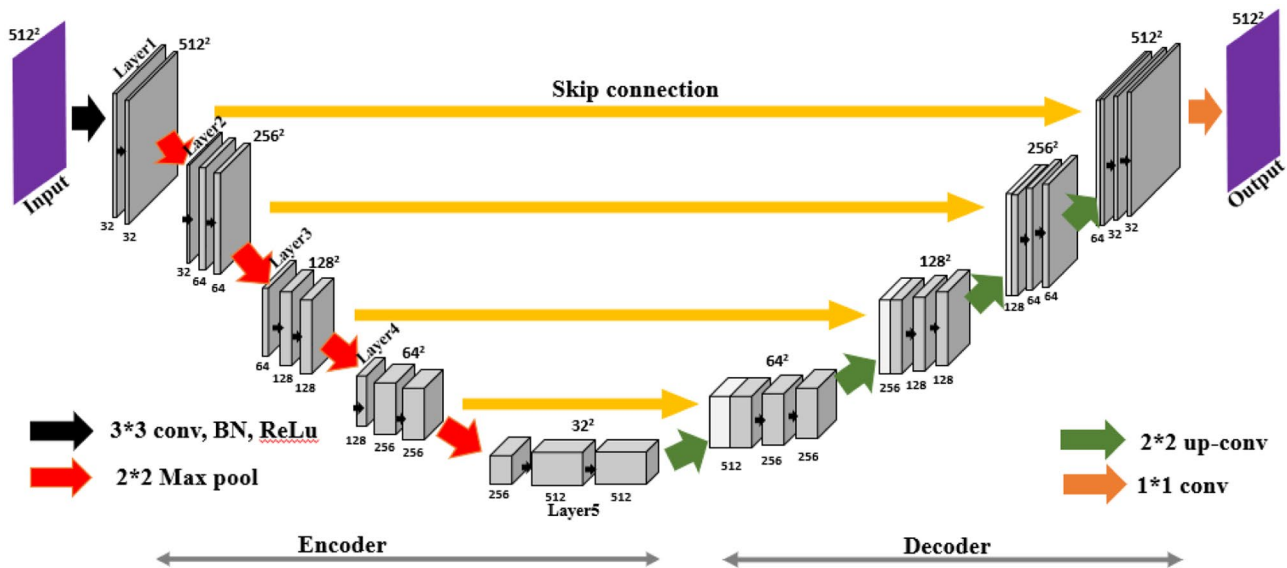

Fig. 3 The structure of U-net for activating the intelligence of RBFNN

shrinking, and the number of channels went on increasing. The decoder part was expansion processing; the up-sampling operation (the size of the convolution kernel is 2×2) was performed first; then, it was superimposed with the feature map of the corresponding layer on the down-sampling path on the left to perform convolution processing. In addition, other convolution operations were consistent with the corresponding up-sampling part on the left side, until the final image was output through a 1×1 convolution. Of note, in order to ensure the consistency of image size, we implemented zero padding.

Experimental Data

To increase the robustness and versatility of the proposed method, we adopted three open datasets. To explore morphological changes of the weak and small bronchi of the lung parenchyma, we used The Lung Image Database Consortium and Image Database Resource Initiative (L-IDRI) dataset [29]. Subsequently, Cancer Imaging Archive-pancreas dataset (CIA-P) was selected to observe the projection of the shape details of the blood vessels in the liver and pancreas [30]. Finally, Information Extraction form Images-Magnetic resonance angiography (IXI-MRA) dataset was adopted to check the morphological changes of blood vessels in the brain [31]. To increase the diversity of experimental data effects based on different slab thicknesses, the number of 2D data required for each projection based on L-IDRI, CIA-P, and IXI-MRA was selected as 10–25, 20–35, and 70–100, respectively, depending on whether it was the

training dataset or the test dataset. In addition, for L-IDRI data, in order to more effectively perform projection processing without being disturbed by irrelevant factors, we implemented the extraction of lung parenchyma. First, we transformed the image into the binary image according to the threshold and the "Find" function [32]. This was done to roughly distinguish the internal and external information of the lung parenchyma. The threshold range was determined according to the mean of the maximum pixel value and the minimum pixel value. Subsequently, the maximum connected component [33] was calculated to connect as many small components as possible. Furthermore, for the remaining gaps in the lung parenchyma, we used erosion processing [34] to fill the entire lung parenchyma. Next, in order to prevent unnecessary details from interfering with the final extraction of lung parenchyma, we removed the interfering details of original image through hole filling [35] and the Find function. Finally, based on the corroded binary image and the processed original image, we successfully extracted the lung parenchymal by the Find function. The whole processing is depicted in Fig. 4.

Training Processing

In medical image processing based on deep learning, the determination or characterization of the ground truth data has always been a challenging task [36, 37]. In the present study, we aimed to obtain a better MIP image than the traditional ones through the proposed neural network. Therefore, we had no gold standard to rely on. Accordingly, targeted
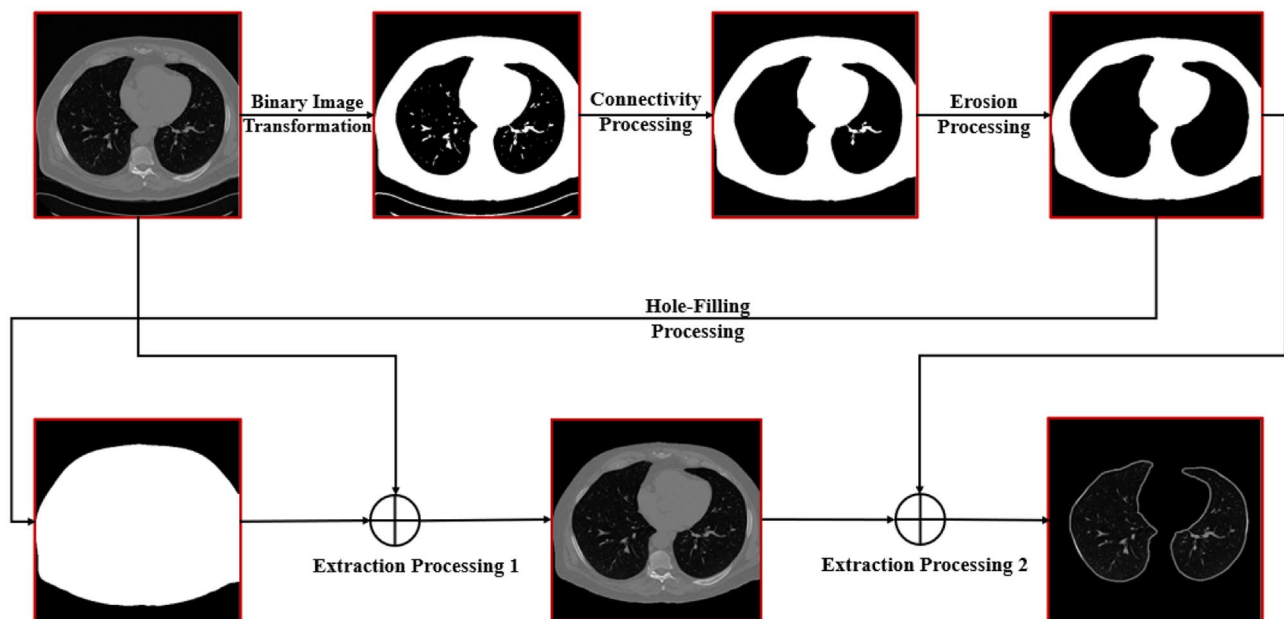

Fig. 4 The whole process of lung parenchyma extraction

transfer learning was carried out. To train U-net, according to the characteristics of the key organ structure, gray value, and contour stripe trend, and so forth, we carefully selected 2000, 2000, and 1000 data based on L-IDRI dataset, CIA-P dataset, and IXI-MRA dataset, respectively. Furthermore, we augmented the dataset by implementing blurring and adding noise locally or globally. Accordingly, the processed data were used as the original input data for training U-net, while the unprocessed data were used as the ground truth data for training U-net. The Adam optimization algorithm [38] was performed to train the U-net. We chose the learning rate as 10^{-3} to 10^{-6}. Specifically, the original learning rate was 10^{-3}, and it decreased exponentially with the increase of epochs until 10^{-6}. The batch size was selected as 5. In order to prevent overfitting, the early stop based on 8 epochs was implemented. Additionally, the loss function based on the mean square error (MSE) was selected [27]. To train RBFNN, we adopted the classic gradient descent method (GDM) [39]. We set the learning rate of shallow network and iteration number range to 0.01 and 200–300, respectively. The loss function of RBFNN was also selected based on classic MSE [25]. More importantly, the original 2D image used to synthesize the MIP image was used as the input of the initial neural network. MIP images synthesized by traditional method were used as ground truth of RBFNN.

The numbers of ground truth data used for training based on the L-IDRI dataset, CIA-P dataset, and IXI-MRA dataset were selected as 200, 200, and 250, respectively.

Evaluation Metrics

As mentioned above, to the best of our knowledge, none of the previous studies has applied neural network algorithm to MIP technology. Therefore, in order to better highlight the proposed algorithm, in addition to comparing with traditional MIP (T-MIP), we also compared the proposed method with two image fusion–based deep convolutional neural networks: IFCNN-1 [40] and IFCNN-2 [20]. In subjective observation, the radiologists can directly evaluate the shape of a key part under the projection, the stripe texture, and so forth to judge the quality of the projection. In addition to subjective evaluation, the application of objective indicators is also essential and necessary. It is well known that the most common and classic metrics for evaluating image quality are image sharpness, image contrast, and so on. Since there are no reference images, in the present study, we adopted the following five classic no-reference image quality metrics: Histogram Entropy (HISE) [41], image contrast (CONT) [42], Brenner's (BREN) [43], Tenengrad (TENG) [41], and Tenengrad variance (TENV) [44]. For all metrics, the larger was the value, the higher was the image quality.
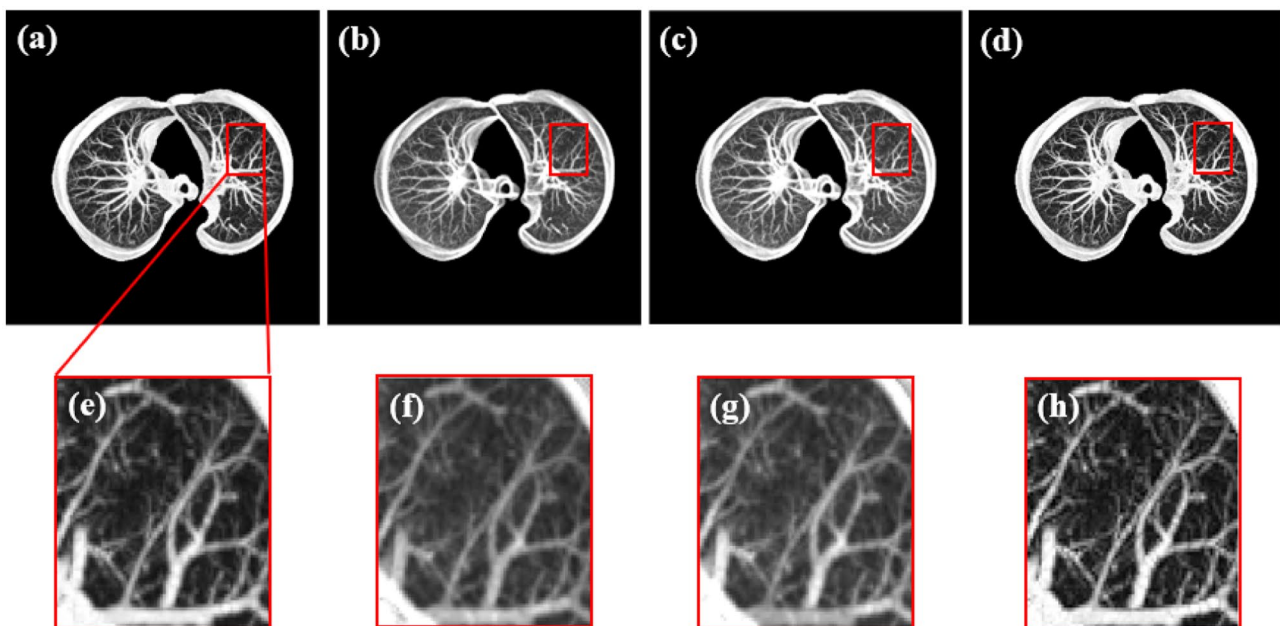

Fig. 5 Projection performance of L-IDRI Data-1 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** the magnified images of the region of interest (ROI) in respective (**a–d**)
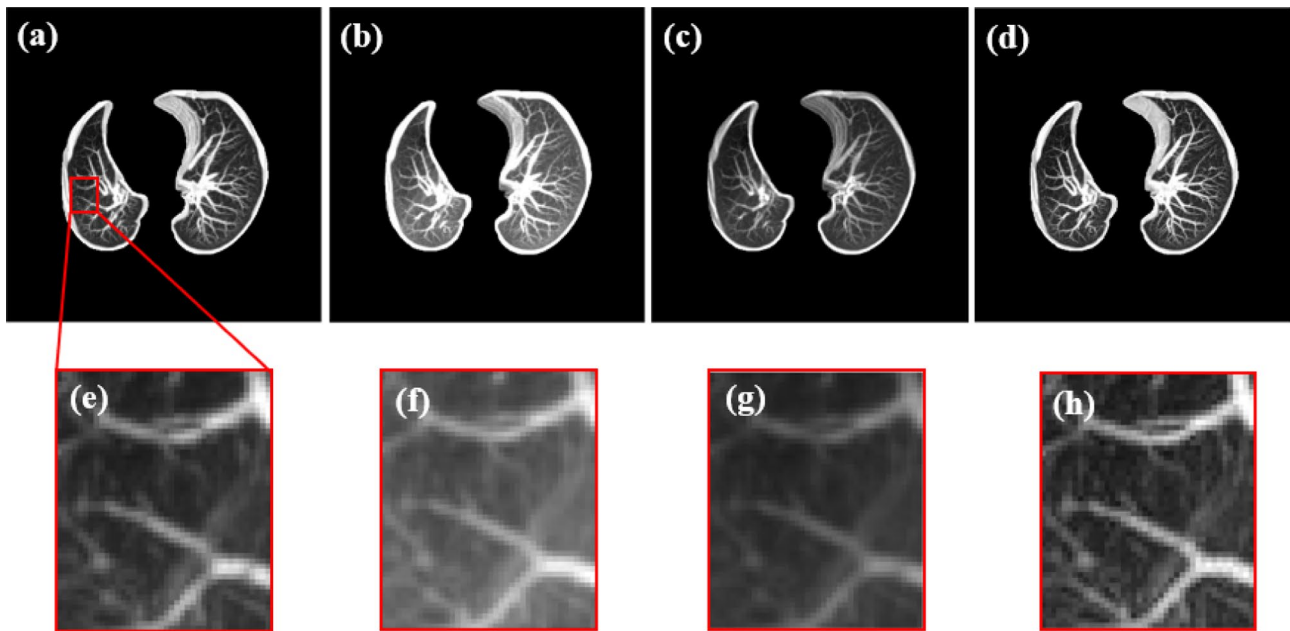
Fig. 6 Projection performance of L-IDRI Data-2 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** the magnified images of the region of interest (ROI) in respective (**a-d**)

Results

Results on Visual Observation

For the L-IDRI dataset, CIA-P dataset, and IXI-MRA dataset, we tested 1000, 1000, and 250 data, respectively.

Specifically, 3 groups of L-IDRI data (labeled L-IDRI Data-1, L-IDRI Data-2, and L-IDRI Data-3), 2 grounds of CIA-P data (labeled CIA-P Data-1, CIA-P Data-2), and 2 grounds of IXI-MRA data (labeled IXI-MRA Data-1 and IXI-MRA Data-2) are shown in Figs. 5, 6, 7, 8, 9, 10 and 11. In Fig. 5, annotation letters (a–d) denote
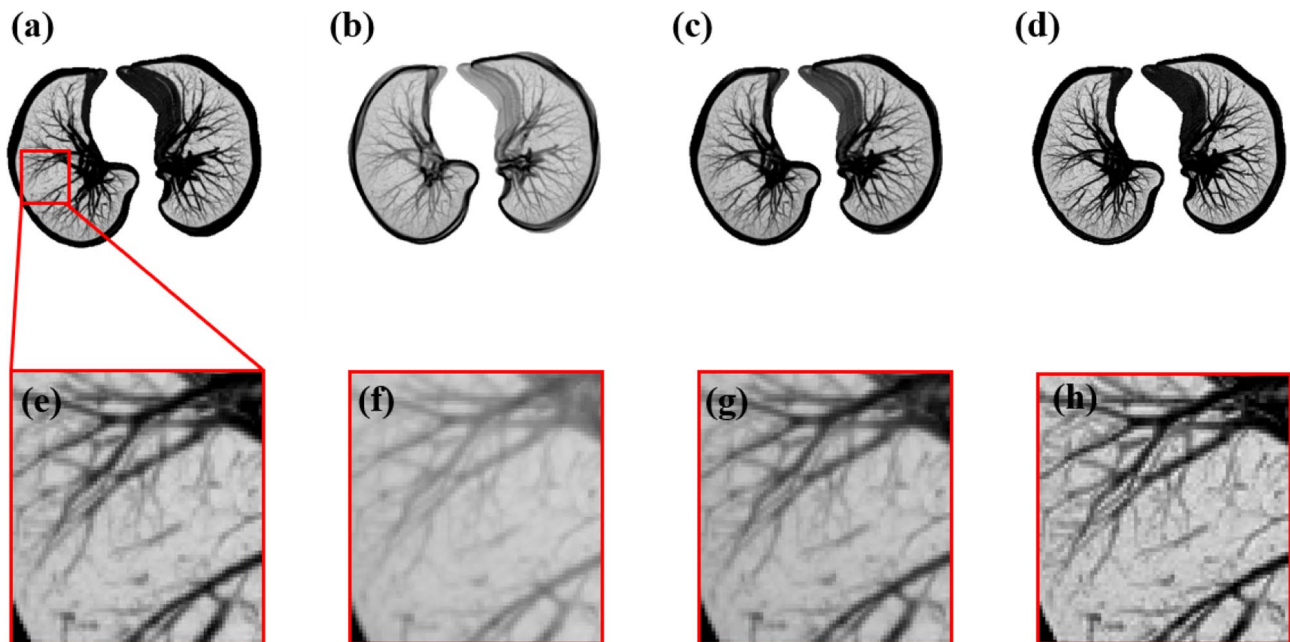

Fig. 7 Projection performance of L-IDRI Data-3 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** the magnified images of the region of interest (ROI) in respective (**a–d**)

Fig. 8 Projection performance of CIA-P Data-1 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** The magnified images of the region of interest (ROI) in respective (**a–d**)
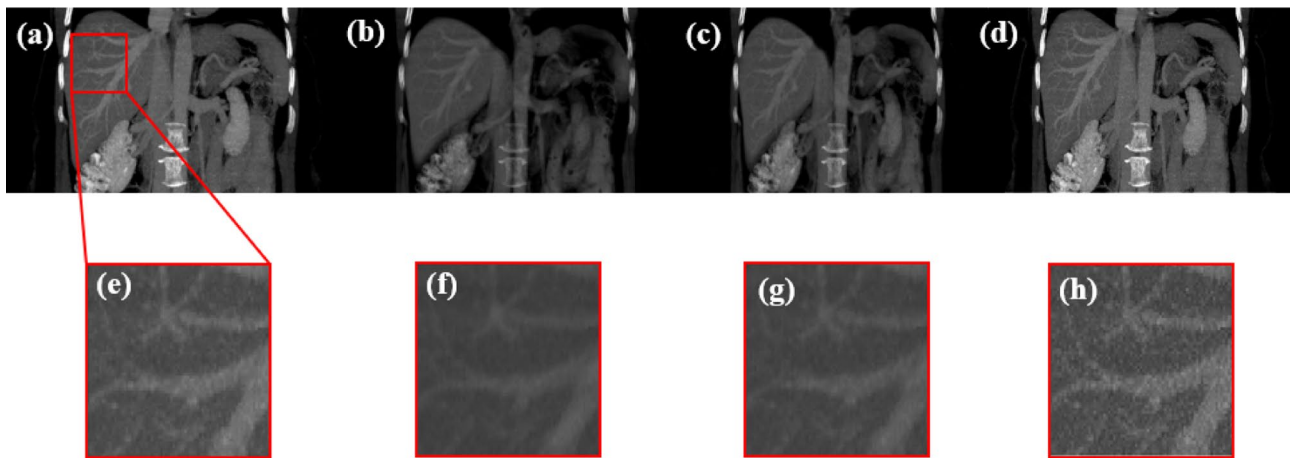
the projection images by implementing traditional MIP (T-MIP) method, two image fusion convolutional neural network methods (IFCNN-1 and IFCNN-2), and the proposed method, respectively. e–h of the second row denotes the magnified images of region of interest (ROI) in respective (a–d) for making it easier to observe the performances of several methods. Figures 6, 7, 8, 9, 10 and 11 follow the layout format of Fig. 5. In terms of visually judging image quality, the most important criterion is image definition. In addition, we also adopted fidelity to observe whether the projection images were distorted; texture continuity to observe whether some trachea, blood vessels, and other structures in the image were intact; and detail imaging to observe whether more details, especially, tiny structural details, were displayed. In order to increase authority and reduce subjective differences, we randomly selected 10 pieces of data from each dataset, and all the data were

judged and compared through four experienced radiologists. Additionally, we adopted the 5-point evaluation method; specifically, for the four standards, radiologists would score from 0 to 5, with 0 points being the lowest and 5 points being the highest. For the different types of data, the scoring situations are shown in Tables 1, 2 and 3.

Subsequently, we analyzed different datasets in further detail. In the three datasets of the L-IDRI dataset in Figs. 5, 6 and 7, we clearly observed that the performances of the two methods based on IFCNN were the worst, the images were severely distorted, and the bronchus, especially the tiny bronchus, was blurred. Compared to the traditional MIP method and the proposed method, the proposed method showed a superior performance in definition and small tracheal detail imaging. As can be seen in Fig. 7 that shows image inversion processing, the proposed method was clearly superior as compared to the
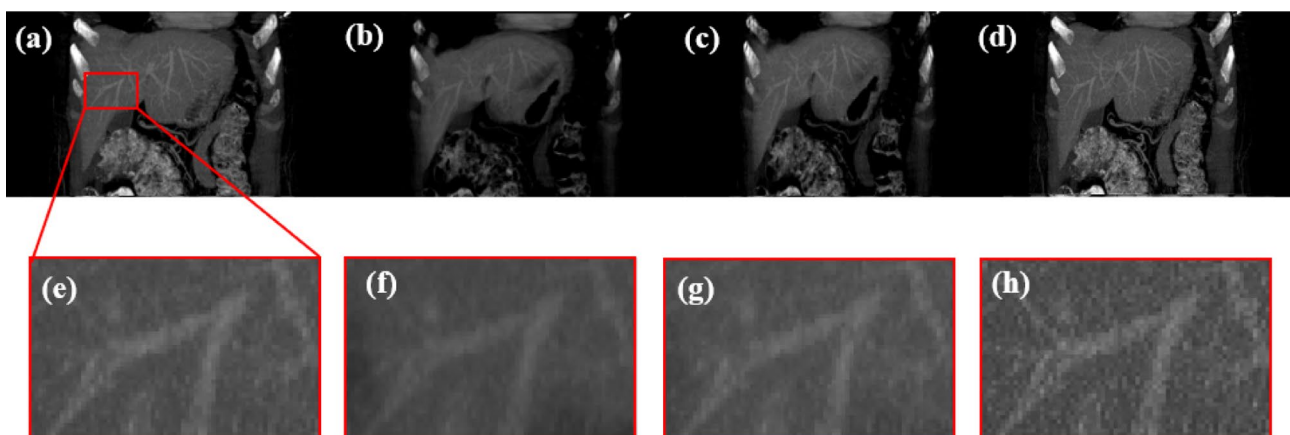

Fig. 9 Projection performance of CIA-P Data-2 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** The magnified images of the region of interest (ROI) in respective (**a–d**)

Fig. 10 Projection performance of IXI-MRA Data-1 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** The magnified images of the region of interest (ROI) in respective (**a–d**)
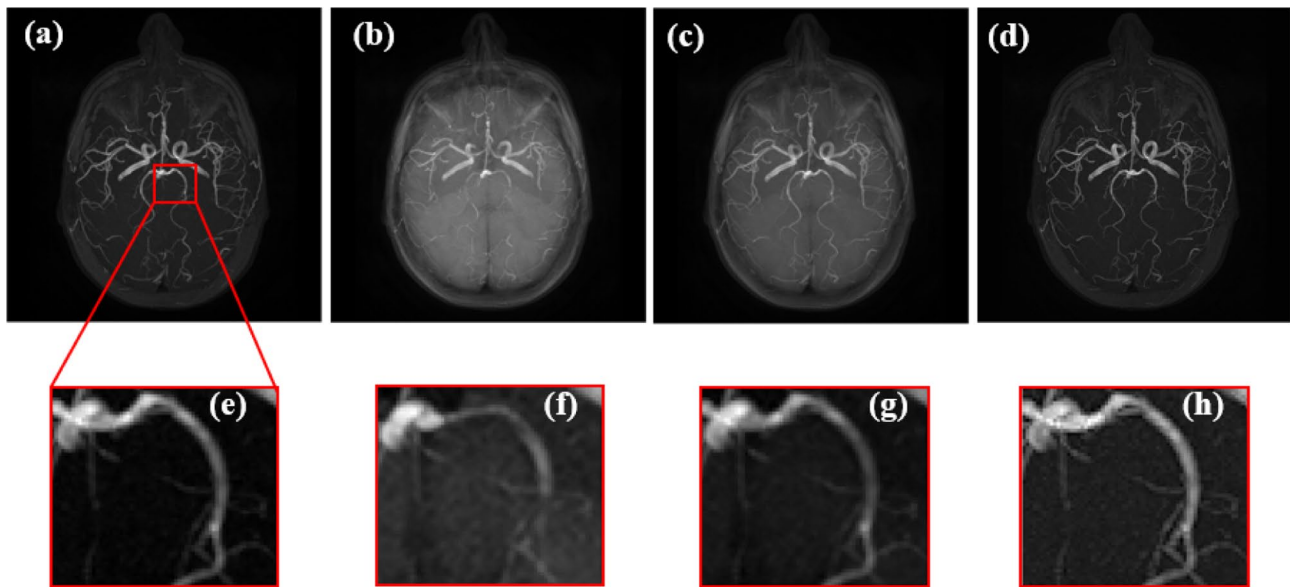
other three methods. In addition, according to the scores of four radiologists, as compared with the other three methods, the proposed method got the highest score (see Table 1). For the CIA-P dataset, our aim was to observe the projection of blood vessels in the liver. Since the gray values of blood vessels and other tissues around the liver had relatively small differences, it was very important to

more clearly visualize blood vessels to observe related lesions. As can be seen in Figs. 8 and 9, while some small blood vessels were clearly displayed in the image produced by the proposed method, this did not occur in the other three methods. The comprehensive performance of our method was also optimal based on the 5-point evaluation standard (see Table 2). Finally, we selected thick
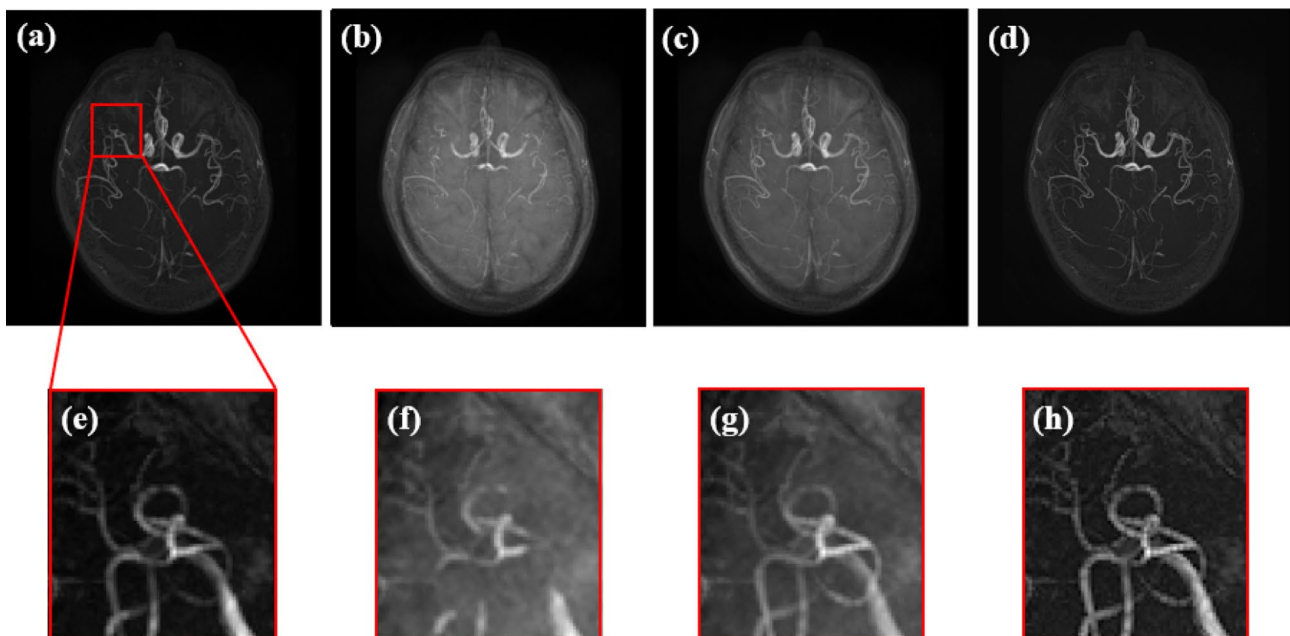

Fig. 11 Projection performance of IXI-MRA Data-2 based on four methods. **a–d** The projection images by implementing T-MIP method, IFCNN-1, IFCNN-2 method, and the proposed method, respectively. **e–h** The magnified images of the region of interest (ROI) in respective (**a–d**)

Table 1 The 5-point evaluation table based on the L-IDRI data, where four radiologists scored four metrics. DE, FI, TC, DI, and GT stand for definition, fidelity, texture continuity, detail imaging and grand total, respectively. Bold represents the best performances

	Radiologist 1					Radiologist 2				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	4	5	3	3	15	5	4	5	4	18
IFCNN-1	2	3	2	2	9	2	2	4	3	11
IFCNN-2	2	3	2	2	9	3	3	3	3	12
Proposed	5	5	5	4	**19**	5	4	5	5	**19**
	Radiologist 3					Radiologist 4				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	5	4	5	4	18	4	5	5	3	17
IFCNN-1	3	2	4	3	12	3	3	3	2	11
IFCNN-2	3	4	4	3	14	4	3	3	3	13
Proposed	5	5	5	4	**19**	5	5	5	4	**19**

slab thickness to project the blood vessels of the brain. As shown in Figs. 10 and 11, for both blood vessel continuity and microvascular imaging, our proposed method yielded more details without distortion of the image. The results in Table 3 further confirm the superior performance of the proposed method.

Results on No-Reference Image Quality Metrics

In addition to subjective observation, another important criterion is the evaluation of objective metrics. As mentioned above, for the three datasets, we tested 1000 L-IDRI dataset, 1000 CIA-P data, and 250 IXI-MRA data, respectively. Since there is no real standard image, we analyzed the definition of the images produced by the four methods. Based on this, in order to objectively evaluate a total of 2250 data, we adopted the following five definition-based metrics: histogram entropy (HISE), image contrast (CONT), Brenner's (BREN), Tenengrad (TENG), and Tenengrad variance (TENV). Tables 4, 5 and 6 show the means and standard deviations of respective test datasets based on five metrics.

Taken together, the proposed methods almost achieved the highest metrics value performance. Except for the CIA-P

dataset and IXI-MRA dataset, the two methods based on IFCNN achieved better HISE performance. Combined with the image performance, we conclude that the most likely cause was that the image distortion–produced artifacts, which interfered with the judgment of HISE. In short, as compared to the traditional MIP technology, the proposed method produced the higher quality projection images based on the three datasets. In addition, all experimental results were statistically significant ($p < 0.01$).

Discussion

Analysis of the Performances of Two IFCNN Structures for Projection Synthesis

As described in "Results", regardless of subjective observation or objective analysis, the two convolutional neural network methods based on image fusion (IFCNN-1 and IFCNN-2) did not achieve good results for synthesizing MIP images. Simply put, the overall framework of IFCNN-1 is based on the classic fully connected convolutional neural network. IFCNN-2 uses the full convolutional network structure and feature map fusion processing.

Table 2 The 5-point evaluation table based on the CIA-P data, where four radiologists scored four metrics. DE, FI, TC, DI, and GT stand for definition, fidelity, texture continuity, detail imaging, and grand total, respectively. Bold represents the best performance

	Radiologist 1					Radiologist 2				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	4	4	3	4	15	5	5	3	4	17
IFCNN-1	2	1	3	1	7	3	2	2	2	9
IFCNN-2	2	3	3	3	11	3	3	3	2	11
Proposed	4	4	4	5	**17**	5	4	5	4	**18**
	Radiologist 3					Radiologist 4				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	4	4	5	4	17	4	5	4	5	**18**
IFCNN-1	2	4	3	3	12	3	3	2	2	10
IFCNN-2	3	5	4	3	15	4	4	3	2	13
Proposed	4	4	5	5	**18**	5	5	4	4	**18**

Table 3 The 5-point evaluation table based on the IXI-MRA data, where four radiologists scored four metrics. DE, FI, TC, DI, and GT stand for definition, fidelity, texture continuity, detail imaging, and grand total, respectively. Bold represents the best performance

	Radiologist 1					Radiologist 2				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	5	4	3	2	14	4	4	2	3	13
IFCNN-1	3	1	2	1	7	1	0	1	1	3
IFCNN-2	2	3	3	2	10	1	1	3	3	8
Proposed	5	5	5	5	**20**	5	4	5	4	**18**
	Radiologist 3					Radiologist 4				
	DE	FI	TC	DI	GT	DE	FI	TC	DI	GT
T-MIP	5	4	3	3	15	4	5	3	3	15
IFCNN-1	3	1	1	2	7	3	1	2	2	8
IFCNN-2	3	1	2	3	9	4	2	3	2	11
Proposed	4	4	5	4	**17**	5	5	4	5	**19**

These two methods achieved good results based on image fusion technology with fewer input images. Conversely, IFCNN-1 and IFCNN-2 performed poorly based on the MIP technology that requires many parallel input images. This finding can be attributed to the fact that, first of all, the principle of MIP technology is to project the details or elements with the maximum density, the pertinence of feature extraction which is more accurate than that of image fusion. The layer-by-layer convolution processing based on the window or patch-level may cause blurred recognition. In addition, due to the large amount of input images, many parallel network structures are needed, so training neural networks is a considerable challenge. Under existing conditions, neural networks cannot effectively learn.

Analysis of the Performances of the Proposed Network Structures for Projection Synthesis

In a previous study, we successfully applied the radial basis function neural network based on point-level to medical image fusion. Based on these results, in the present study, we adopted a shallow, simple neural network to apply the MIP technology under the situation when the performances of the parallel and complex network structures were not good for synthesizing projection images. As is widely known, the traditional shallow neural network (SNN) is not as popular as the convolutional neural network in two-dimensional or even multi-dimensional image processing. Due to the simple network structure, the SNN cannot effectively learn the complex image structure. However, in the present study, we found that, after targeted feature point extraction, radial basis function neural network (RBFNN) was then implemented to analyze and process specific feature points, and combining effective learning, the intelligence of RBFNN can also be effectively activated. More importantly, as there is no need to train feature points and RBFNN is based on pixel (point)-level, the training burden is much smaller. In the field of medical imaging, labeled data are actually not abundant [45]. Therefore, it is necessary to investigate the training of small sample data. Based on this, we adopted U-net network to provide feature points for RBFNN. Furthermore, in order to show more intuitively why we chose the network structure based on the hybrid of U-net and RBFNN, we conducted another set of comparative experiments based on the traditional MIP method (T-MIP), the hybrid structure of U-net and traditional MIP (U-MIP), meaning that the U-net was first implemented to obtain feature images or enhanced images, and then the projection image was obtained based on the traditional MIP method of projecting these enhanced images, RBFNN-only structure, and the hybrid structure of U-net and RBFNN structure (proposed method). For all structures, we also tested 1000 L-IDRI data. Figure 12 shows the comparison effect of one of the L-IDRI dataset based on the four methods, where a–d denote the T-MIP method, the RBFNN-only method, the U-MIP method, and the proposed method, respectively. As can be seen in Fig. 12, the proposed hybrid structure based on U-net and RBFNN is the best in terms of intuitive effects, such as detail imaging

Table 4 For the four methods, the mean and standard deviation (SD) performances of the five metrics based on the L-IDRI dataset. Bold represents the best performance

	HISE	CONT ($\times 10^2$)	BREN ($\times 10^4$)	TENG ($\times 10^5$)	TENV ($\times 10^{12}$)	p-value
	Mean $\pm$ standard deviation (SD)					$p < 0.01$
T-MIP	0.86 ± 0.1	2.31 ± 0.3	2.07 ± 0.2	5.41 ± 0.4	7.74 ± 0.7	
IFCNN-1	0.71 ± 0.2	2.12 ± 0.4	1.85 ± 0.2	2.87 ± 0.5	5.79 ± 0.8	
IFCNN-2	0.73 ± 0.1	2.45 ± 0.4	2.25 ± 0.4	3.18 ± 0.3	6.95 ± 0.7	
Proposed	$\mathbf{0.87} \pm 0.1$	$\mathbf{3.32} \pm 0.3$	$\mathbf{3.51} \pm 0.3$	$\mathbf{7.66} \pm 0.4$	$\mathbf{16.7} \pm 0.6$	

Table 5 For the four methods, the mean and standard deviation (SD) performances of the five metrics based on the CIA-P dataset. Bold represents the best performance

	HISE	CONT ($\times 10^2$)	BREN ($\times 10^3$)	TENG ($\times 10^5$)	TENV ($\times 10^{12}$)	p-value
	Mean ± standard deviation (SD)					$p < 0.01$
T-MIP	0.99 ± 0.2	1.99 ± 0.3	9.54 ± 0.2	2.87 ± 0.5	8.76 ± 0.6	
IFCNN-1	1.01 ± 0.3	1.54 ± 0.4	5.71 ± 0.3	2.12 ± 0.4	2.54 ± 0.7	
IFCNN-2	**1.02** ± 0.3	1.67 ± 0.3	6.39 ± 0.3	2.30 ± 0.6	2.78 ± 0.6	
Proposed	1.00 ± 0.3	**2.97** ± 0.3	**14.4** ± 0.3	**4.27** ± 0.5	**20.1** ± 0.6	

Table 6 For the four methods, the mean and standard deviation (SD) performances of the five metrics based on the IXI-MRA dataset. Bold represents the best performance

	HISE	CONT ($\times 10^1$)	BREN ($\times 10^2$)	TENG ($\times 10^3$)	TENV ($\times 10^9$)	p-value
	Mean ± standard deviation (SD)					$p < 0.01$
T-MIP	0.69 ± 0.2	3.91 ± 0.2	2.92 ± 0.5	7.39 ± 0.6	1.82 ± 0.7	
IFCNN-1	0.72 ± 0.3	2.32 ± 0.3	3.17 ± 0.3	1.59 ± 0.6	1.66 ± 0.4	
IFCNN-2	**1.03** ± 0.4	2.28 ± 0.2	6.01 ± 0.4	8.32 ± 0.5	4.35 ± 0.6	
Proposed	0.88 ± 0.2	**6.77** ± 0.2	**6.35** ± 0.5	**14.1** ± 0.6	**5.73** ± 0.5	

and image definition. Not surprisingly, based on the objective judgment of five image definition metrics, the proposed hybrid structure of U-net and RBFNN yielded the best performance for 1000 sets of test data (see Fig. 13).

Analysis of the Performances of Different Numbers of Hidden Layer Neurons Based on RBFNN

Furthermore, due to some factors, such as the different types of data, the choice of different slab thickness when inputting data, and the different amount of adopted data, and so forth, the choice of the number of hidden layer neurons in the RBFNN is not fixed. The optimal hidden layer neuron mainly depends on the training data, input variables, and other factors; however, the amount of calculation and generalization effect should also be considered [46, 47]. For L-IDRI dataset, we selected 50 as the number of hidden neurons. For the selection of the number of hidden layers, we conducted a set of comparative experiments based on the different number of hidden layer neurons: 25, 30, 35, 40, 45, 50, 60, 70, and 100, respectively. The actual output

data from the respective trained neural network were compared to the ideal output data (gold standard) through mean square error (MSE), peak signal to noise ratio (PSNR), and structural similarity (SSIM) [48]. The larger was the value of PSNR and SSIM, the better (i.e., closer to the ideal output images) was the image quality, while MSE was the opposite. For a more intuitive observation, we selected the respective mean values for the metric values of all test data. The comparison based on different numbers of neurons is shown in Fig. 14. From the overall performances, 50 obtained the best performance. Finally, combined with image factors, we chose 50 as the optimal hidden layer neuron.

Significance of Proposed Method and Further Study

In the present study, we aimed to establish whether quality of the projection images obtained through the proposed neural network would be better than the traditional projection map, which would help doctors to more accurately and timely determine small lesions. In the field of medical imaging, today's medical equipment and post-processing method based on software can produce high-quality images; however, in some cases, they still
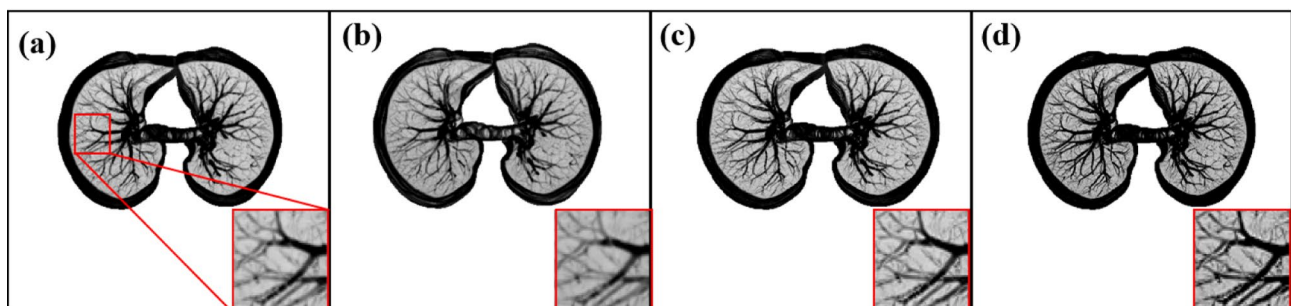
Fig. 12 The image performances based on L-IDRI data. **a–d** The MIP based on traditional method (T-MIP), the radial basis function neural network-only method (RBFNN), the MIP based on U-net neural network (U-MIP), and the proposed method, respectively
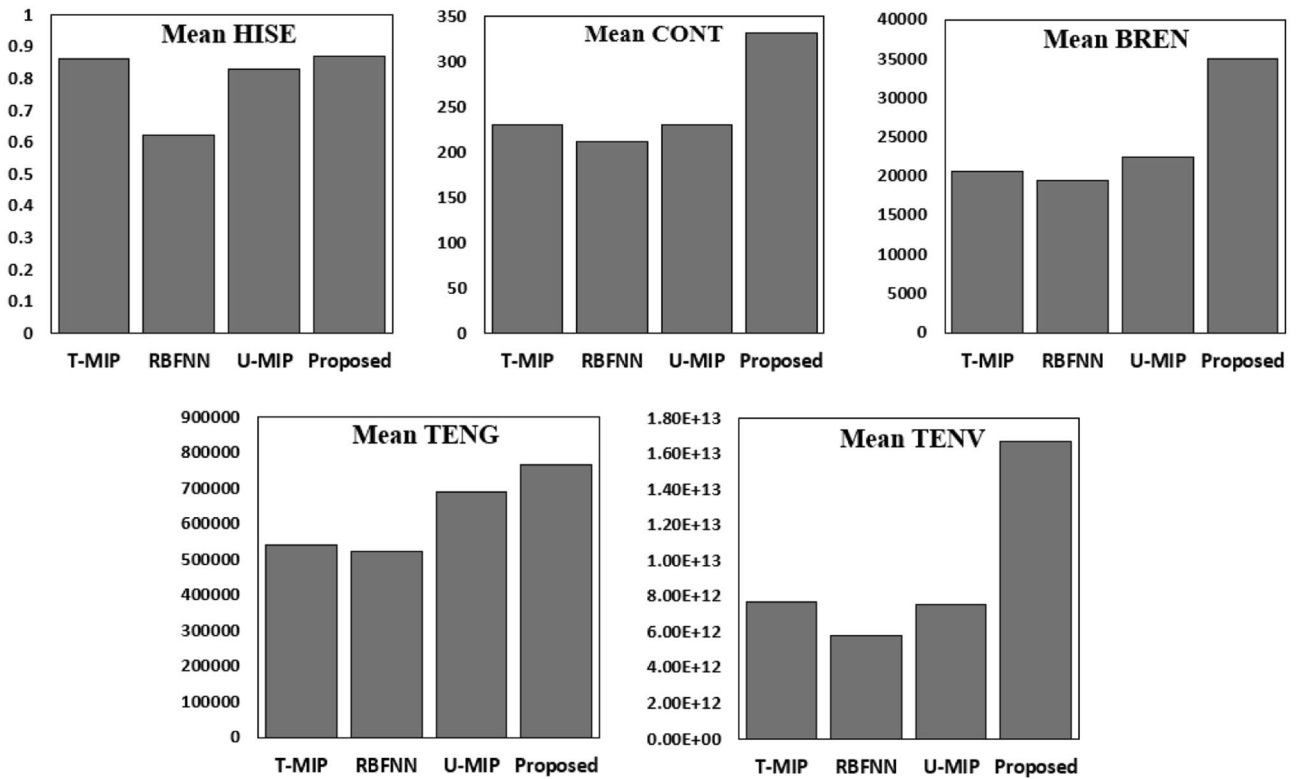
Fig. 13 For L-IDRI dataset, the objective comparison between the proposed three structures and the traditional MIP based on five no-reference quality evaluation metrics. The values of all metrics are the mean values of the test data
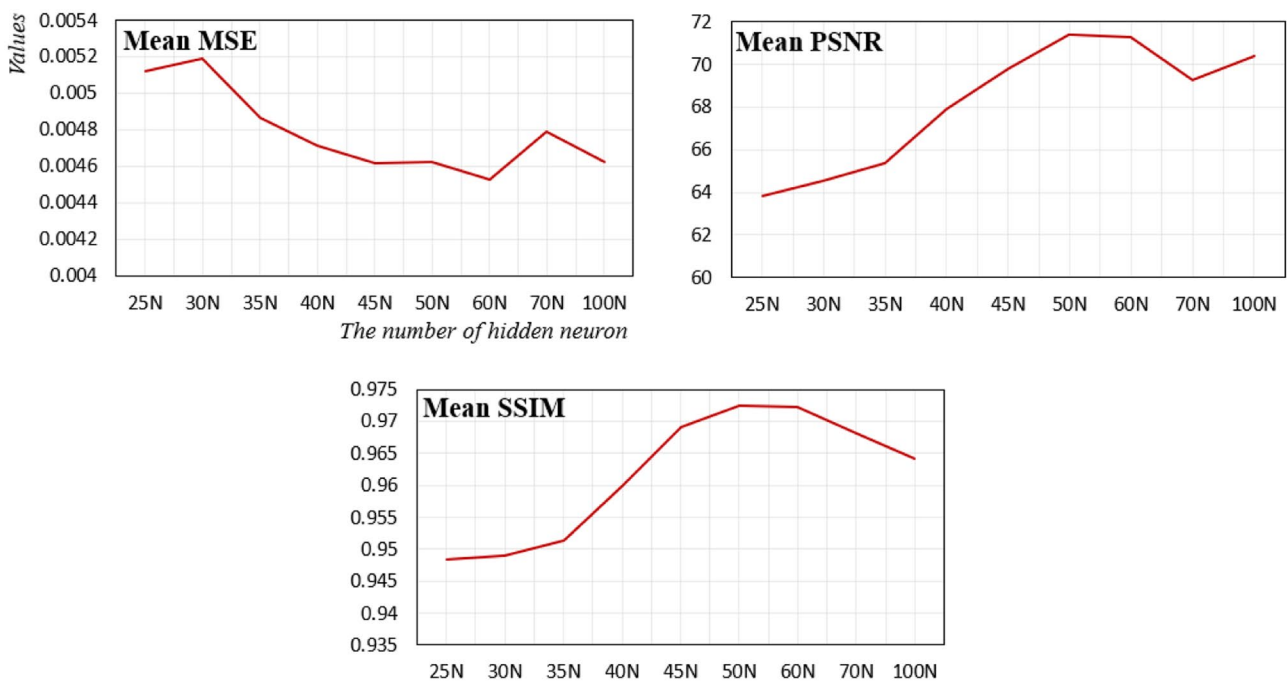

Fig. 14 For L-IDRI dataset, based on mean square error (MSE), peak signal to noise ratio (PSNR), and structural similarity (SSIM), the performances of different numbers of hidden layer neurons

cannot meet the requirements for accurate diagnosis. Therefore, we believe that the use of artificial intelligence technology to further improve high-quality images is a future development direction. Accordingly, we made a preliminary attempt in the field of synthetic projection maps. To the best of our knowledge, our study is the first to apply neural network technology to synthesize projection images. In further research, we will improve the network structure and training methods, apply more clinical data, and strive to obtain higher quality MIP images.

Conclusion

In the present study, we aimed to obtain higher-quality maximum intensity projection (MIP) images to help radiologists precisely diagnose diseases. To this end, we proposed a hybrid structure based on U-net network and radial basis function neural network (RBFNN) to synthesize MIP images. Compared to the traditional MIP method and other network structures, through the doctor's judgment and objective metric analysis, the qualities of the images obtained by the proposed method were found to be optimal. In addition, the application of a large amount of data also demonstrated robustness and generality of the proposed method.

Declarations

Ethics Approval Because of the use of open dataset, this manuscript does not deal with ethical issue.

Conflict of Interest The authors declare no competing interests.

References

1. Lacout, Alexis, et al: Pancreatic involvement in hereditary hemorrhagic telangiectasia: assessment with multidetector helical CT. Radiology 254(2):479–484,2010
2. Wang, Mao Qiang, et al: Benign prostatic hyperplasia: cone-beam CT in conjunction with DSA for identifying prostatic arterial anatomy. Radiology 282(1):271–280,2017
3. Huber, Adrian, et al: Performance of ultralow-dose CT with iterative reconstruction in lung cancer screening: limiting radiation exposure to the equivalent of conventional chest X-ray imaging. Eur Radiol 26(10):3643–3652,2016
4. Zheng, Sunyi, et al: Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection. IEEE Trans Med Imaging 39(3):797–805,2019
5. Koziński, Mateusz, et al: Tracing in 2D to reduce the annotation effort for 3D deep delineation of linear structures. Med Image Anal 60:101590,2010
6. Harvey, Evan Cary, et al: Impacts of photon counting CT to maximum intensity projection (MIP) images of cerebral CT angiography: theoretical and experimental studies. Phys Med Biol 64(18):185015,2019
7. Zou, Sijuan, Xiaohua Zhu: FDG PET/CT of COVID-19. Radiology 200770,2020
8. Albano, Domenico, et al: Incidental findings suggestive of COVID-19 in asymptomatic patients undergoing nuclear medicine procedures in a high-prevalence region. J Nucl Med 61(5):632–636,2020
9. Ciccarese, Federica, et al: Diagnostic accuracy of North America Expert Consensus Statement on reporting ct findings in patients with suspected COVID-19 infection: an Italian single center experience. Radiology: Cardiothoracic Imaging 2(4):e200312,2020
10. Dietz, Matthieu, et al: COVID-19 pneumonia: relationship between inflammation assessed by whole-body FDG PET/CT and short-term clinical outcome. Eur J Nucl Med Mol Imaging 1–9,2020
11. Tamm, Eric P, et al: Update on 3D and multiplanar MDCT in the assessment of biliary and pancreatic pathology. Abdom Imaging 34(1):64–74,2009
12. Prevedel, Robert, et al: Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. Nat Methods 11(7):727–730,2014
13. Li, Wang-jia, et al: Effect of slab thickness on the detection of pulmonary nodules by use of CT maximum and minimum intensity projection. Am J Roentgenol 213(3):562–567,2019
14. Sakas, Georgios, Marcus Grimm, Alexandros Savopoulos: Optimized maximum intensity projection (MIP). EUROGRAPHICS Workshop on Rendering Techniques. Springer, Vienna, 1995
15. Schreiner, Steven, and Robert L. Galloway: A fast maximum-intensity projection algorithm for generating magnetic resonance angiograms. IEEE Trans Med Imaging 12(1):50-57,1993
16. Fang, Laifa, et al: Fast maximum intensity projection algorithm using shear warp factorization and reduced resampling. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 47(4):696–700,2002
17. Díaz Iriberri, José, Pere Pau Vázquez Alcocer: Depth-enhanced maximum intensity projection. 8th IEEE/EG International Symposium on Volume Graphics 2010
18. Zhang, Daming, et al: Quick evaluation of lower leg ischemia in patients with peripheral arterial disease by time maximum intensity projection CT angiography: A pilot study 2020

19. Qu, Liangqiong, et al: Synthesized 7T MRI from 3T MRI via deep learning in spatial and wavelet domains. Med Image Anal 62:101663,2020

20. Zhang, Yu, et al: IFCNN: A general image fusion framework based on convolutional neural network. Information Fusion 54:99–118,2020

21. Zhou, Tao, et al: Hi-net: hybrid-fusion network for multi-modal MR image synthesis. IEEE Trans Med Imaging 2020

22. Singh, Sneha, R. S. Anand: Multimodal medical image fusion using hybrid layer decomposition with CNN-based feature mapping and structural clustering. IEEE Trans Instrum Meas 69(6):3855–3865,2019

23. Zhang, Dequan, et al: Hybrid learning algorithm of radial basis function networks for reliability analysis. IEEE Trans Reliab 2020

24. Zhao, Zhitao, et al: Prediction of interfacial interactions related with membrane fouling in a membrane bioreactor based on radial basis function artificial neural network (ANN). Bioresour Technol 282:262–268,2019

25. Chao, Zhen, Hee-Joung Kim: Removal of computed tomography ring artifacts via radial basis function artificial neural networks. Phys Med Biol 64(23):235015,2019

26. Chao, Zhen, Dohyeon Kim, Hee-Joung Kim: Multi-modality image fusion based on enhanced fuzzy radial basis function neural networks. Phys Med 48:11-20,2018

27. Park, Junyoung, et al: Computed tomography super-resolution using deep convolutional neural network. Phys Med Biol 63(14):145011,2018

28. Gómez, Pablo, et al: Low-light image enhancement of high-speed endoscopic videos using a convolutional neural network. Med Biol Eng Comput 57(7):1451–1463,2019

29. Armato III, Samuel G, et al: The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38(2):915-931,2011

30. Clark, Kenneth, et al: The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. J Digit Imaging 26(6):1045–1057,2013

31. Jiang, Dongsheng, et al: Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network. Jpn J Radiol 36(9):566–574,2018

32. Sigmon, Kermit, Timothy A. Davis: Matlab primer. CRC Press, 2004

33. He, Lifeng, et al: The connected-component labeling problem: A review of state-of-the-art algorithms. Pattern Recogn 70:25–43,2017

34. Tambe, Sagar B, et al: Image processing (IP) through erosion and dilation methods. 2013

35. Liu, Soulan, et al: A computationally efficient denoising and hole-filling method for depth image enhancement. Real-Time Image and Video Processing. International Society for Optics and Photonics Vol 9897,2016

36. Sahiner, Berkman, et al: Deep learning in medical imaging and radiation therapy. Med Phys 46(1):e1-e36,2019

37. Gsaxner, Christina, et al: PET-train: Automatic ground truth generation from PET acquisitions for urinary bladder segmentation in CT images using deep learning. 2018 11th Biomedical Engineering International Conference (BMEiCON). IEEE, 2018

38. Kingma, Diederik P, Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014

39. Karayiannis, Nicolaos B: Reformulated radial basis neural networks trained by gradient descent. IEEE Transactions on Neural Networks 10(3):657–671,1999

40. Liu, Yu, et al: Multi-focus image fusion with a deep convolutional neural network. Information Fusion 36 :191–207,2017

41. Krotkov, Eric, J-P. Martin: Range from focus. *Proceedings.* 1986 IEEE International Conference on Robotics and Automation. Vol 3 IEEE, 1986

42. Nanda, Harsh, Ross Cutler: Practical calibrations for a real-time digital omnidirectional camera. CVPR Technical Sketch 20(2),2001

43. Brenner, John F, et al: An automated microscope for cytologic research a preliminary evaluation. J Histochem Cytochem 24(1):100–111,1976

44. Pech-Pacheco, José Luis, et al: Diatom autofocusing in brightfield microscopy: a comparative study. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000 Vol 3 IEEE, 2000

45. Ilse, Maximilian, et al: Diva: Domain invariant variational autoencoders. Medical Imaging with Deep Learning. PMLR, 2020

46. Vujicic, Tijana, et al: Comparative analysis of methods for determining number of hidden neurons in artificial neural network. Central European Conference on Information and Intelligent Systems. Faculty of Organization and Informatics Varazdin, 2016

47. Aljarah, Ibrahim, et al: Training radial basis function networks using biogeography-based optimizer. Neural Comput Applic 29(7):529–553,2018

48. Hore, Alain, Djemel Ziou: Image quality metrics: PSNR vs. SSIM." 2010 20th International Conference on Pattern Recognition. IEEE, 2010