IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 March 2024, accepted 7 April 2024, date of publication 11 April 2024, date of current version 19 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3387413

RESEARCH ARTICLE

Multiscale Feature Attention Module Based Pyramid Network for Medical Digital Radiography Image Enhancement

WENJING XUE[1], YINGMEI WANG[1], AND ZHIEN QIN[2]
[1]School of Mathematics and Statistics, Shandong University of Technology, Zibo 255000, China
[2]Shinva Medical Instrument Company Ltd., Zibo 255086, China

Corresponding author: Yingmei Wang (yingmeiwang_sdu@163.com)

This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2022MA027, in part by the Opening Project of Guangdong Province Key Laboratory of Computational Science at Sun Yat-sen University under Grant 2021003, and in part by the Visiting and Training Fund for Teachers from Ordinary Undergraduate Universities in Shandong Province.

ABSTRACT Medical digital radiography (DR) is widely used in the clinical application. To deal with the problems of noise, edge blur, low contrast in DR images, we propose a multiscale feature attention module based pyramid enhancement network by training image blocks. The network is in the framework of a simplified U-Net, which reduces the computational load by reducing the convolution layer, and adopts Laplacian pyramid connection instead of concatenation operation to preserve the image boundary information. In addition, we embed a simple multiscale feature attention (SMFA) module between the encoder and decoder, which integrates the feature information of different scales precisely and makes the network have a stronger ability to perceive the local feature information. Our proposed algorithm is a network realization of Gauss-Laplacian pyramid decomposition with an attention module. Furthermore, we design a side feature loss function combined with mean square loss and absolute loss. We adopt batch normalization between convolution and activation operations to ensure information of all gray scale regions to be considered, which enhances the robustness of the network. We use LeakyReLu activation function and Sigmoid function in the previous layers and in the output layers respectively to preserve the negative information of multiscale details and to keep the gray scale region of the output images. Experiments with real data of different parts of human body validate the effectiveness of our algorithm, which shows that our proposed algorithm performs well on contrast enhancement, structure details preservation, and noise suppression. It has certain value of clinical application.

INDEX TERMS Medical DR image enhancement, multiscale features extraction, U-Net, pyramid network.

I. INTRODUCTION

Medical Digital Radiography (DR) images have the advantages of high resolution, wide dynamic range, fast imaging speed and small radiation dose. It has become one of the most widely used medical imaging methods in clinical practice. In order to eliminate the effect of scattered lines, most static medical DR devices use a low-density antiscatter grid, which results in grid artifacts in raw DR images. In order to reduce the influence of antiscatter grid, most manufacturers carry out a fixed height correction to remove the grid image. Fixed height correction ensures the removal of grid images at the same height for the following image post processing. However, in clinical practice, due to different doctors, different scanning positions and so on, the scanning heights are different. As a result, it is difficult to deal with grid images in the raw scanning data and at the same time to keep image details. In addition, due to noise, improper X-ray exposure and unequal thickness of human tissue [1], DR images have blurred edges and low contrast [2], [3], which will affect disease diagnosis of the doctors. To meet the needs of high-end clinical applications, DR image enhancement is of

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

research significance [4], [5] in the field of medical radiology imaging.

Traditional image enhancement algorithms have certain effects in both natural image and medical DR image enhancement (especially in natural image enhancement), such as histogram equalization (HE) [6], adaptive histogram equalization (AHE) [7], contrast-limited adaptive histgram equalization (CLAHE) [8], [9], [10], nonlinear enhancement based on wavelet decomposition [12], [13], [14], enhancement based on Gauss-Laplacian pyramid [15], [16], [17] and other algorithms. These traditional enhancement algorithms have some inherent problems. For medical DR image enhancement, histogram equalization and its improved algorithms need large computation cost with high noise sensitivity, low local contrast, and insufficient sharpening of the boundary, which reduces its clinical applicability. Adaptive unsharp mask algorithm [18] uses local variance as a parameter to enhance high-frequency components, which can enhance the edges and image details to a certain extent. However, it is also sensitive to noise and is prone to artifacts. The nonlinear enhancement algorithms based on wavelet decomposition are easy to cause serious artifacts because of difficult parameter determination. The improved algorithms based on Gauss-Laplacian pyramid need manual parameter adjustment, which is difficult to adjust and not intelligent.

Recently there are more and more deep learning [19] based methods in medical image processing and reconstruction. For example, Singh and Gupta [20] proposed an enhanced feature-level medical image fusion algorithm based on gradient decomposition, but it has the problems of noise and uneven brightness distribution. Halupka et al. [21] proposed a deep learning based optical coherence tomography (OCT) image enhancement method, but it also suffers from noise. Liang et al. [22] proposed an improved low-dose computed tomography (CT) image enhancement network based on CycleGAN. This method can improve the structural details and suppress noise and artifacts, but the GAN [23] network is difficult to train because of gradient disappearance or gradient explosion. Xia et al. [24] gave a deep residual neural network based low-dose CT image enhancement algorithm. Georgescu et al. [25] proposed a new multimodal multi-head convolutional attention module for CT and MRI image super-resolution. At present, there have been many deep learning based image enhancement methods in natural image processing, medical CT and OCT fields. However, due to the particularity and difficulty of obtaining medical DR raw data, medical DR image enhancement methods based on deep learning are still relatively few.

Aiming at the above problems, we propose a pyramid network based on a multiscale feature attention module for medical DR image enhancement, which introduces deep learning method into medical DR image enhancement. The shape of the proposed network is in the framework of a simplified U-Net with a simple multiscale feature attention module and a special skip connection, which is some kind of network realization of Gauss-Laplacian pyramid decomposition with an attention module. Experiments with real data of different body parts verify the effectiveness and generalization ability of the algorithm.

The network is small in scale with reduced network parameters to improve the training speed. The whole image can be enhanced through the training of image blocks, and the image enhancement effect is remarkable. The main contributions of our work are as follows:

- We design a small scale network in the framework of U-Net [26], [27] with a Gauss-Laplacian pyramid skip connection. We use a single-square convolution layer instead of the traditional two-square convolution layer to reduce the convolution operation. And we choose a Gauss-Laplacian pyramid connection between the encoder and the decoder to enhance the image boundary features, which makes our method a network realization of Gauss-Laplacian pyramid decomposition method.
- The proposed network contains a simplified multi-scale feature attention (SMFA) module embedded in the bottom of the encoder and decoder, which integrates much more local feature information of different scales. Hence it can improve the local feature information perception ability of the network to preserve detailed features to the maximum extent.
- We design a side penalty loss function combined with absolute loss and square loss and add a batch normalization (BN) [28] operation between the convolution and activation operations to ensure the robustness of the network and to enhance the information of every gray range of the images. Due to DR image features and BN operation, multi-scale details of DR image contains negative values. Accordingly, we adopt the LeakyReLu activation function [29], [30] at the layers other than the output layer to retain the negative information of multi-scale details, so the enhanced image can better restore more features of DR images. At the same time, we use the Sigmoid activation function [31] in the output layer to ensure that the output image and the input image have the same gray scale range.

II. RELATED WORK

A. U-NET

U-Net [26], [27] is a classical convolutional neural network architecture with a structure shape like the letter U, so it is called U-Net. It can effectively capture local and global information in images, and has a good ability of capturing deep features. Specifically, U-Net network structure consists of symmetrical encoders and decoders. The encoder part gradually reduces the size of the feature map and increases the channels through a series of double-layer convolution blocks and pooling operations. To extract abstract features of the data, each convolution block contains a convolution and a ReLu activation function. In the decoder part, the feature map is gradually restored to the size of the original input image by upsampling and double-layer convolution operation, and then the details are gradually restored.

In U-Net, the convolution blocks in the decoder usually make skip connections with the corresponding convolution blocks in the encoder to concatenate features, so the decoder can make use of the different-level feature information. Its network structure is simple and effective, and we can freely adjust the depth of the network according to the needs of the processing task.

B. BATCH NORMALIZATION AND ACTIVATION FUNCTIONS

Batch normalization (BN) [28] can ensure that the input data of each layer obey the normal distribution, which can speed up the training speed and improve the generalization ability of the network. Only γ and β parameters in the BN layer are learnable parameters. The calculation process is as follows:

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i,\ \sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i-\mu_{\mathcal{B}})^2,$$
$$\hat{x}_i \leftarrow \frac{x_i-\mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2+\epsilon}},\ y_i \leftarrow \gamma\hat{x}_i+\beta \equiv BN_{\gamma,\beta}(x_i), \quad (1)$$

where $x_i, y_i, i = 1, \cdots, m$ are the input and the final output, and $\mu_{\mathcal{B}}, \sigma_{\mathcal{B}}^2$ are their mean value and variance, respectively.

Activation function can do nonlinear processing of the linear weighted summation of each neuron in the deep neural network. At present, there are various activation functions with their own characteristics. In our experiments, the convolution of each layer except the output layer adopts the LeakyReLu activation function (as shown in Fig. 1 (a)), which responds to the negative input and retains the information of the negative input. The LeakyReLu activation function [29], [30] is formulated by

$$\sigma(x) = \begin{cases} \alpha x, & x < 0, \\ x, & x \geq 0, \end{cases} \quad (2)$$

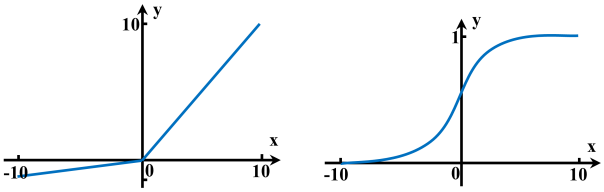
where x is the input and α is a small positive number, always set to 0.1. LeakyReLu assigns small amplitudes to all negative values instead of setting them all to 0, preventing negative information loss and effectively solving the problem of gradient disappearance.

The output layer convolution uses the Sigmoid activation function [31] (as shown in Fig. 1 (b)), a smooth curve that can map the input nonlinearly and capture more information. With its derivability, it is more convenient to use the optimization algorithm for model training and parameter updating. The Sigmoid activation function formula is as follows:

$$\sigma(x) = \frac{1}{1+e^{-x}}. \quad (3)$$

III. OUR METHOD

Inspired by the structure of U-Net and Gaussian-Laplacian pyramid, we design a pyramid network in the framework of U-Net with an attention module SMFA and name it SMFA-Pyramid network. Section III-A describes the whole SMFA-Pyramid network structure. Section III-B introduces the simple multiscale feature attention module. Section III-C gives the loss function for training the network.



(a) LeakyReLu activation function; (b) Sigmoid activation function.

FIGURE 1. Activation functions.

A. SMFA-PYRAMID NETWORK

SMFA-Pyramid network is a new image enhancement algorithm with a simple multi-scale feature attention module based on U-Net structure and Gauss-Laplacian pyramid, whose architecture is shown in Fig. 2. In our U-shape pyramid network, without changing the information extraction effect of the network, we use a single square convolution instead of the two square convolutions in U-Net, which reduces the convolution operation and improves the training speed. To ensure information of each gray range of the image to be enhanced and the robustness of the network, we introduce a BN operation between the convolution and the activation operation. According to the pixel characteristics of medical DR image, we adopt LeakyReLu activation function in all layers except the output layer to retain the multi-scale negative information of the image, which preserves the multi-scale information of medical DR image to a greater extent. And we use Sigmoid activation function in the output layer to ensure that the output image and the input image are in the same gray scale range.

For the skip connection, we use the Laplacian pyramid connection to replace the concatenation connection between channel features of the encoder and decoder in the U-Net. Let $U_i, i = 1, 2, 3$ denote the downsampling features in the encoder, and $L_i, i = 1, 2, 3$ the Laplacian pyramid features. That is to say, the downsampling features U_i of the encoder is upsampled to obtain $U_i', i = 1, 2, 3$, and then compute the difference between the features before U_i and U_i' to obtain the Laplacian pyramid features L_i, where we use a convolution layer to make their channels consistent, which is formulated by

$$\begin{cases} U_1' = \text{up}(U_1), \\ L_1 = f_3(f_2(f_1(\text{input})) - U_1'), \\ U_2' = \text{up}(U_2), \\ L_2 = f_5(f_4(U_1) - U_2'), \\ U_3' = \text{up}(U_3), \\ L_3 = f_7(f_6(U_2) - U_3'), \end{cases} \quad (4)$$

where $\text{up}(\cdot)$ is the upsampling function, $f_i(\cdot), i = 1, 2, \cdots, 7$ the Convolution+BN+LeakyReLu operations

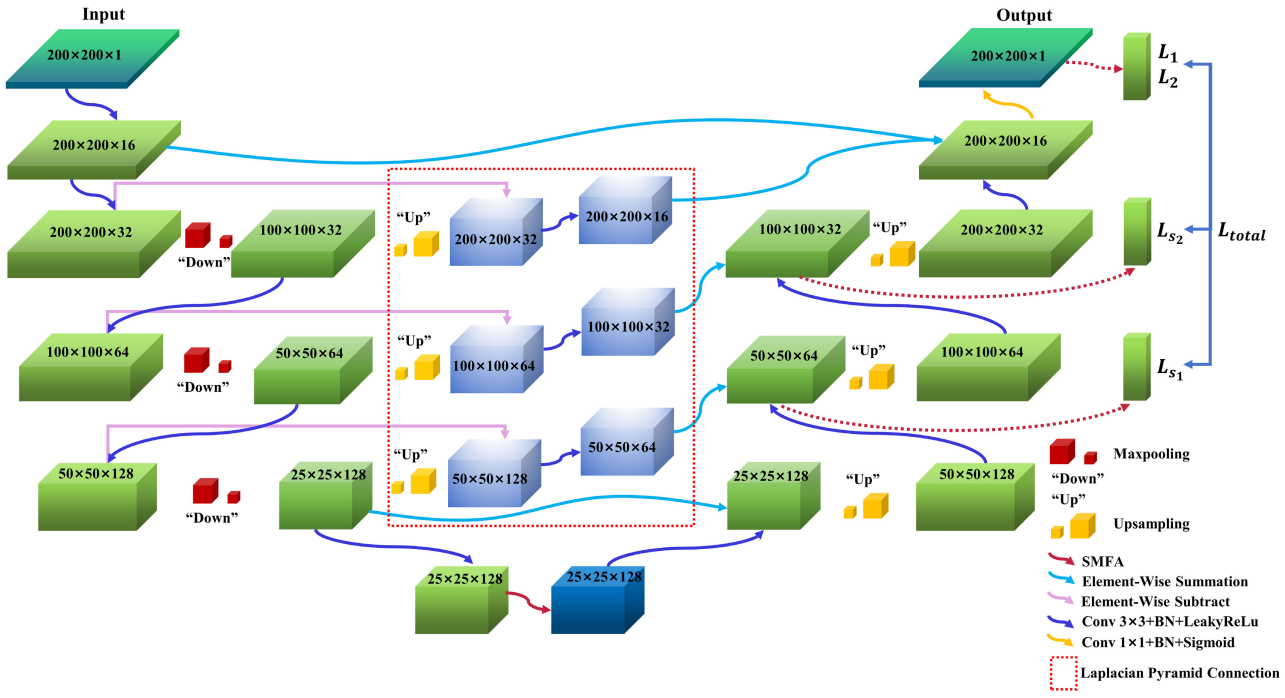


FIGURE 2. SMFA-Pyramid network structure.

with convolution kernel size of 3×3, respectively. The details are shown in Algorithm 1.

Algorithm 1 Laplacian Pyramid Connection

Input: The SMFA-Pyramid Network input, the maxpooling features U_i, $i = 1, 2, 3$;

Output: Feature map L_i, $i = 1, 2, 3$;

1: Upsampling operation is performed on U_1 to obtain U_1';
2: Perform two Convolution+BN+LeakyReLu operations f_1 and f_2 on input, and do subtraction with U_1' to get $f_2(f_1(\text{input})) - U_1'$, and then perform one Convolution+BN+LeakyReLu operation f_3 on the obtained difference to obtain L_1;
3: Upsampling operation is performed on U_2 to obtain U_2';
4: Perform one Convolution+BN+LeakyReLu operation f_4 on U_1 and subtract U_2', and then perform Convolution+BN+LeakyReLu operation f_5 on the obtained difference to obtain L_2.
5: Upsampling operation is performed on U_3 to obtain U_3';
6: Perform one Convolution+BN+LeakyReLu operation f_6 on U_2 and subtract U_3', and then perform Convolution+BN+LeakyReLu operation f_7 on the obtained difference to obtain L_3.
7: **return** L_1, L_2, and L_3.

B. SIMPLE MULTI-SCALE FEATURE ATTENTION (SMFA) MODULE

In this subsection, we introduce the proposed simple multi-scale feature attention (SMFA) module in detail. SMFA module is embedded between the encoder and decoder of our proposed SMFA-Pyramid network. It is composed of 3 branches in parallel and its basic structure is shown in Fig. 3. The right branch consists of one downsampling, one Convolution+BN+LeakyReLu operation and upsampling. The left branch mainly includes two downsamplings and two different kernel-size Convolution+BN+LeakyReLu operations. The middle branch combines features of the left and the right branches using Convolution+BN+LeakyReLu operation, element-wise product and element-wise summation, which is formulated by

$$\begin{cases} l_1(x) = \text{up}(f_9(f_8(\text{dw}(x)))), \\ l_2(x) = \text{up}(f_{11}(f_{10}(\text{dw}(\text{dw}(x))))), \\ l(x) = l_1(x) + l_2(x), \\ r(x) = \text{up}(f_{12}(\text{dw}(x))), \\ m(x) = f_{13}(x) \cdot l(x) + r(x), \end{cases} \tag{5}$$

where x is the SMFA module input, $\text{dw}(\cdot)$ the downsampling function to resize the input to half size and double its channels, $\text{up}(\cdot)$ the upsampling function to adjust the feature size and channel to the input size by bilinear interpolation, $f_i(\cdot)$, $i = 8, 9$ two Convolution+BN+LeakyReLu operations with convolution kernels of size 5×5, $f_i(\cdot)$, $i = 10, 11$ two Convolution+BN+LeakyReLu operations with convolution kernels of size 3×3, and $f_i, i = 12, 13$ are two Convolution+BN+LeakyReLu operations with convolution kernels of size 1×1. $l(x)$, $r(x)$ are the results of the left and right dashed boxes in Fig. 3 respectively. $m(x)$ is the output of the SMFA module. The details are shown in Algorithm 2.

Algorithm 2 SMFA Module

Input: Feature map x;
Output: Feature map $m(x)$;

1: Perform one maxpooling operation on x, and two Convolution+BN+LeakyReLu operations f_8 and f_9 with kernel size of 5×5, and then do upsampling to obtain l_1;
2: Perform one maxpooling operation on the feature map in Step 1 after one maxpooling operation on x, and then two Convolution+BN+LeakyReLu operations f_{10} and f_{11} with a kernel size of 3×3, and then perform upsampling to obtain l_2;
3: $l(x) = l_1 + l_2$;
4: Perform one maxpooling operation on x, one Convolution+BN+LeakyReLu operation f_{12} with a kernel size of 1×1, and one upsampling to obtain $r(x)$;
5: Perform one Convolution+BN+LeakyReLu operation f_{13} on x with kernel size of 1×1, which is multiplied with $l(x)$ by elements and added with $r(x)$ by elements to get $m(x)$.
6: **return** $m(x)$.

The function of the right branch of the SMFA module is to preserve the shallow features of DR data, and the left branch is to extract deeper features of different scales. The middle branch fuses features of different scales, and the whole module adopts LeakyReLu activation function to retain negative information of DR image, which can more accurately integrate multi-scale feature information to maximize the preservation of details and improve the performance of the SMFA module.

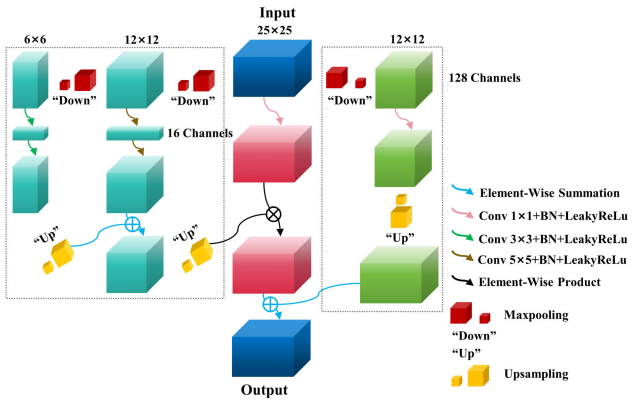


FIGURE 3. SMFA module structure.

C. LOSS FUNCTION

In order to better training the proposed network, we design a loss function L_{total} including absolute error L_1, the square error L_2 [32] and a side penalty loss L_s, which is formulated by

$$L_{\text{total}} = \alpha L_1 + \beta L_2 + \gamma L_s. \tag{6}$$

The parameters α, β, γ represent the positive weight coefficients of three loss terms respectively, which satisfies $\alpha + \beta + \gamma = 1$. In our experiments, we take $\alpha = 0.1$, $\beta = 0.4$, $\gamma = 0.5$.

1) L_1 AND L_2 LOSS FUNCTIONS

The L_1 loss function is expressed by

$$L_1 = \sum_{i=1}^{m}\sum_{j=1}^{n} |K(i,j) - I(i,j)|, \tag{7}$$

where m, n are the numbers of image rows and columns, and $K(i,j), I(i,j)$ represent the labeled high quality image and the network enhanced image respectively.

Compared with L_1 loss function, L_2 loss function is more robust and smooth, which makes the optimization algorithm more efficient in model training. The L_2 loss function is formulated by

$$L_2 = \sum_{i=1}^{m}\sum_{j=1}^{n} (K(i,j) - I(i,j))^2. \tag{8}$$

2) SIDE PENALTY LOSS FUNCTION

Considering that the differences between the side output layers of our SMFA-Pyramid decoder and the labeled images may provide some suggestion to mitigate the gradient loss problem and facilitate encoder training by back propagating the side output loss, we use two side outputs to compute the side penalty loss, which is formulated by

$$L_s = \zeta L_{s_1} + \eta L_{s_2}, \tag{9}$$

where $L_{s_i}, i = 1, 2$ are L_2 losses for two side output layers shown in Fig. 2. Parameters ζ and η represent the positive weight coefficients of the two side output losses respectively, which satisfies $\zeta + \eta = 1$. In our experiments, we choose $\zeta = 0.6$ and $\eta = 0.4$ for better enhancement performance. Gennerally, we can set ζ and η around 0.5, which has no significant impact on the experimental results.

Since the features in the side output layers are as important as those of the final output layer for the experimental results, we choose $\gamma \cdot \zeta + \gamma \cdot \eta = 0.5$ and $\alpha + \beta = 0.5$. As the L_2 loss function is more robust and smooth than L_1 loss function and L_2 loss is smaller than L_1 loss, we choose β much larger than α. According to this discipline, we set $\alpha, \beta, \gamma, \zeta$, and η empirically by extensive experiments. In our experiments, we use $L_{\text{total}} = 0.1L_1 + 0.4L_2 + 0.3L_{s_1} + 0.2L_{s_2}$.

IV. EXPERIMENTS

A. DATA PREPARATION AND EXPERIMENTAL ENVIRONMENT

In order to verify the enhancement effect of our SMFA-Pyramid network, we use the real clinical medical DR image data set provided by Shinva Medical Instrument Co., Ltd. for experiments. Several typical DR images after black and white inversion (for simplicity, we call them the raw

(a) (b) (c) (d) (e) (f)

FIGURE 4. Typical images in the training dataset: (a) (c) (e) raw images of hand and foot with different projection positions; (b) (d) (f) the corresponding labeled images of (a) (c) (e).

images) and their corresponding enhanced labeled images in the training dataset are shown in Fig. 4.

The data set contains 43 pairs of original and labeled DR images, including hand, foot, wrist, knee, ankle, elbow, patella, tibia and fibula images of different projection positions. The specific numbers of cases for different body parts are shown in Table 1. The training dataset contains 40 image pairs, and the test dataset contains 3 image pairs with size of 3060 × 3060.

TABLE 1. Our dataset statistics.

Boby parts	Number of cases	Size
Hand	4	3060 × 3060
Foot	8	3060 × 3060
Wrist	7	3060 × 3060
Knee	4	3060 × 3060
Ankle	10	3060 × 3060
Elbow	2	3060 × 3060
Patella	4	3060 × 3060
Tibia And Fibula	4	3060 × 3060

In order to improve the image enhancement effect and avoid overfitting caused by insufficient samples, we implement a series of preprocessing operations such as grid shadow elimination, data clipping, flipping, expansion on the raw data before training. Firstly, we use a Fourier domain low-frequency filter to remove grid shadow and pre-denoise the raw data. Secondly, we crop the background information unrelated to the region of interest in the raw images, which is mainly caused by DR scans without changing the opening size of the beam collimator. Finally, we cut the data processed in the first two steps into 200×200 image blocks with a step size of 100, fill the boundaries of the image blocks less than 200×200 with zeros, and flip them left and right, up and down, and diagonally. Through the above operations, the training set is extended to 78664 pairs of image block pairs, which increases the diversity and number of samples, and improves the generalization ability and robustness of the model. To evaluate the model more rationally and accurately and avoid overfitting, we add cross-validation in the training of the network. 70% (55,000 pairs of image blocks) of the training data is for model training and 30% (23,664 pairs of image blocks) of the training data is for model validation. In addition, due to the large differences in radiation dose when DR devices scan for different parts and bodies, the gray ranges of raw images are very different. In order to eliminate the impact of data range difference on the model, we normalize the range of raw and labeled images to [0, 1].

We implement all models in Windows environment, and use PyTorch (1.10.0+cu102) to run the proposed network. The hardware used in the experiment is Windows 10, with an Intel (R) Core (TM) i9-9900K CPU@3.60 GHz and an NVIDIA GeForce RTX 2080 graphics card. Adam [33] optimizer is used to optimize the model. The momentum parameters are $\beta_1 = 0.9$, $\beta_2 = 0.999$, Batch_size $= 4$, and Crop_size $= 200 \times 200$. The total epoch number is set to 300, the initial learning rate lr_{init} is set to 0.05, and the learning rate ($lr_{new} = \lambda \times lr_{init}$) is updated with the number of iterations. The factor λ [34] is determined by

$$\lambda = \delta\left(1 - \frac{epoch_i}{2 \times n}\right) + \frac{epoch_i}{2 \times n}, \tag{10}$$

where $epoch_i$ represents the ith iteration, n the number of images in the training set, and δ is the parameter to prevent the learning rate multiplication factor from being zero. In our experiments, we set $\delta = 10^{-5}$.

B. EVALUATION INDEX

We evaluate the enhanced image quality through visual effects and objective evaluation indexes including peak signal-to-noise ratio (PSNR) [35], structural similarity index (SSIM) [36], contrast-to-noise ratio (CNR) [37] and Cross Sum Modified Laplacian (XSML) [38].

PSNR is used to measure the difference between enhanced image and labeled image. It is obtained by calculating the mean square error between two images. The PSNR function is expressed by

$$\begin{aligned} \text{PSNR} &= 10 \times \log_{10}\left(\frac{\text{MAX}_K^2}{\text{MSE}}\right), \\ \text{MSE} &= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)]^2, \end{aligned} \tag{11}$$

where $K(i,j)$ and $I(i,j)$ represent the labeled high quality image and the network enhanced image respectively. MAX_K^2 represents the square of the maximum grey level of the labeled images $K(i,j)$. Generally the smaller the MSE [39] and the larger the PSNR is, and then the better the enhanced image quality is.

SSIM is used to measure the similarity between two images, formulated by

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{x,y} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{12}$$

Here μ_x and μ_y represent the mean values of images x and y, σ_x^2 and σ_y^2 represent the variance of x and y, respectively. $\sigma_{x,y}$ represents the covariance of x and y. C_1 and C_2 are two constants to avoid a zero denominator. The SSIM value is in the range of [0, 1], the closer to 1 the better the image quality.

CNR is mainly used to measure the ratio relationship between the contrast of useful information and the relative

intensity of noise. In the experiments, we use CNR to quantify the contrast and noise level of an enhanced image. The CNR function is formulated by

$$\mathrm{CNR} = \frac{|\mu_s - \mu_n|}{\sqrt{\sigma_s^2 + \sigma_n^2}},$$
$$\sigma_s^2 = E\{(|I_s|^2 - \mu_s)^2\},\ \sigma_n^2 = E\{(|I_n|^2 - \mu_n)^2\}, \quad (13)$$

where μ_s represents the average gray value of useful information in the image, μ_n the average gray value of noise in the image, σ_s^2 the variance of useful information in the image, σ_n^2 the variance of noise in the image, I_s useful information in the image, and I_n is noise information in the image. In general, the greater the value of CNR is, the higher the image contrast, the lower the noise level and the better the image quality is. Because most pixel values of useful information in the enhanced DR images and the labeled images are greater than 0.5, we choose a threshold of 0.5 to distinguish between the useful information and the noise.

To validate the effect of reducing blur, XSML is used to evaluate the sharpness of images. Generally, the larger the XSML, the better the image quality is. The Laplace operator in XSML takes into account direct neighbors in the x and y directions. In addition, diagonal neighbors are included and weighted by a factor of $\frac{1}{\sqrt{2}}$ to compensate for their greater distance to the center pixel. XSML is formulated by

$$\mathrm{XSML}(x,y) = \frac{L(x,y,z) - \bar{L}(x,y)}{\bar{L}(x,y)},$$
$$\bar{L}(x,y) = \frac{1}{N}\sum_{i=1}^{N} L(x,y,z_i),$$
$$\begin{aligned} L(x,y) &= \sum_{(i,j)\in D(x,y)} |K(i+t,j) + K(i-t,j) \\ &\quad - 2K(i,j)| + |K(i,j+t) + K(i,j-t) - 2K(i,j)| \\ &\quad + \frac{1}{\sqrt{2}}|K(i+t,j+t) + K(i-t,j-t) - 2K(i,j)| \\ &\quad + \frac{1}{\sqrt{2}}|K(i-t,j+t) + K(i+t,j-t) - 2K(i,j)|, \end{aligned} \quad (14)$$

where $D(x,y)$ is a square pixel neighborhood of the pixel (x,y). The parameter t can be adjusted according to the characteristic size of texture elements in the image. We use $t = 3$ for all computations in our experiments.

C. ANALYSIS OF EXPERIMENTAL RESULTS

In this section, we give experimental results of our proposed SMFA-Pyramid network for medical DR image enhancement. At the same time, we compare our method with AHE [7], CLAHE [8], Gauss-Laplacian pyramid [17] (GLP), LLCNN [40], U-Net [26], structure and illumination constrained GAN [41] (StillGAN) and Self-Calibrated Illumination [42] (SCI). To make the comparison more fair and justified, all comparison algorithms use the same preprocessed raw data.

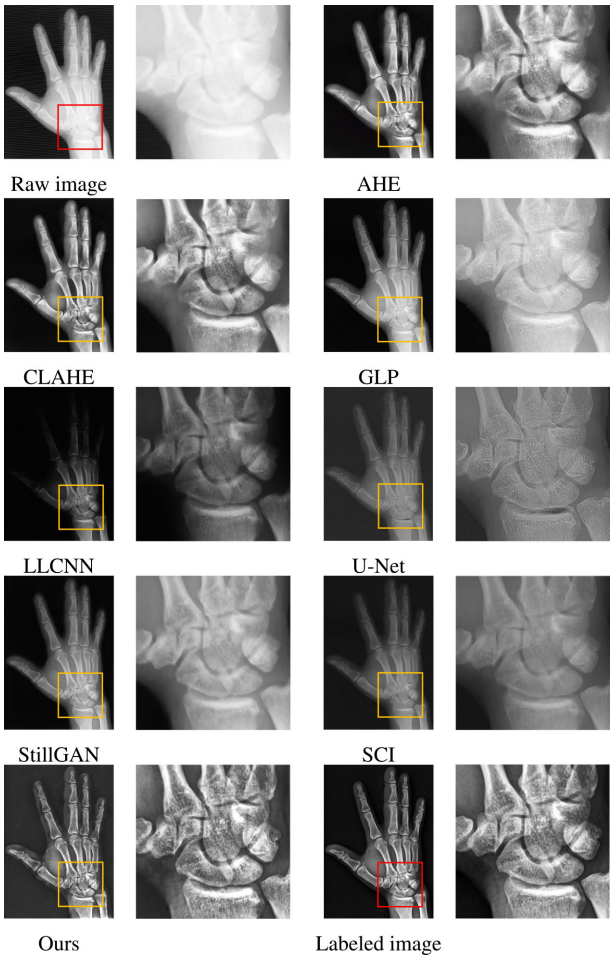


FIGURE 5. Comparison of hand experimental results.

The visual results of the comparison experiments are shown in Figs. 5 and 6. We use red boxes and orange boxes to highlight the comparison areas in the raw DR images and the enhanced images of various comparison algorithms respectively. We can see from Figs. 5 and 6 that the overall images enhanced by AHE and CLAHE algorithms are too bright, especially in the orange box area. And there are still some problems of low contrast, boundary blur and detail loss. Gauss-Laplacian pyramid algorithm uses an exponential enhancement adjustment parameter, which improves image detail enhancement to some extent, but the overall image is bright and the contrast is still not satisfactory.

LLCNN performs image enhancement based on convolutional operation without manual parameter. However, the enhanced image is dark in general, and the parts with small gray value are lost especially in finger ends and muscle parts, which may be caused by mislearning small values as the background during the training process. And the problem of blurred edges and poor contrast still exists. U-Net extracts and restores image features based on encoder and decoder without manual parameter adjustment, and has better enhancement effect on sharpening edges and maintaining details than AHE, CLAHE and LLCNN algorithms. StillGAN extracts the

image features based on adversarial network without manual adjustment, and the contrast is better than the previous four methods, but the enhanced images are too smooth and most bone trabecula information is lost. SCI is an unsupervised learning network, which is designed for enhancing low-high natural image without labeled images, but its enhanced images are too smooth and have poor contrast. Our proposed SMFA-Pyramid network can effectively combines global and local features and has the best contrast and image details among all comparison methods.

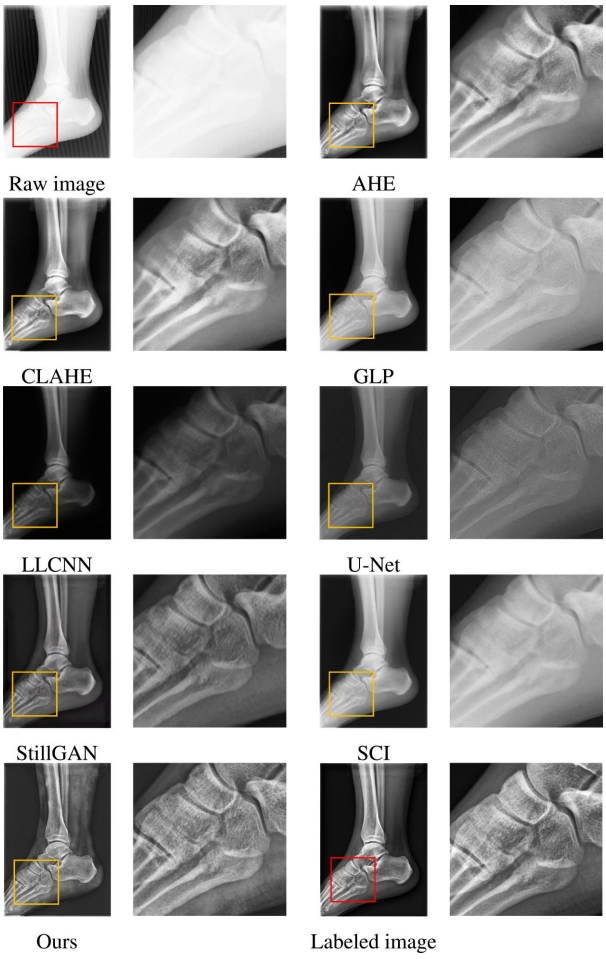
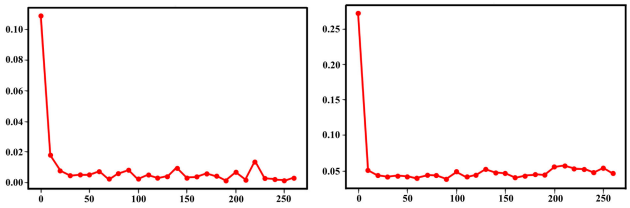


FIGURE 6. Comparison of foot experimental results.

In order to validate that there is no over-fitting problem in our proposed network, we give the cross-validation error curve during the training process, as shown in Fig. 7. We can see that the training error and the validation error can decrease quickly and get steady during the training process. That is to say, our data preprocessing and the network structure can overcome the over-fitting problem to some extent.

To further illustrate the effectiveness of the proposed algorithm, we use the labeled images as reference images to calculate the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), contrast to noise ratio (CNR) and



(a) Training error curve. (b) Verification error curve.

FIGURE 7. The cross-validation error curve.

Cross Sum Modified Laplacian (XSML) of all comparison algorithms, as shown in Tables 2 and 3.

TABLE 2. Hand PSNR, SSIM, CNR, and XSML of different algorithms.

Algorithms	PSNR/dB	SSIM	CNR	XSML
Raw image	20.69	0.740	3.413	0.236
AHE [7]	22.06	0.856	3.187	0.261
CLAHE [8]	22.23	0.858	3.245	0.277
GLP [17]	26.86	0.907	3.349	0.330
LLCNN [40]	16.48	0.586	3.910	0.170
U-Net [26]	18.15	0.616	5.319	0.237
StillGAN [41]	19.18	0.789	5.772	0.254
SCI [42]	22.89	0.777	5.395	0.247
Ours	30.75	0.979	8.486	0.436

TABLE 3. Foot PSNR, SSIM, CNR, and XSML of different algorithms.

Algorithms	PSNR/dB	SSIM	CNR	XSML
Raw image	19.78	0.712	4.171	0.479
AHE [7]	21.04	0.853	4.202	0.562
CLAHE [8]	21.10	0.866	4.216	0.599
GLP [17]	25.64	0.917	4.064	0.609
LLCNN [40]	13.26	0.525	4.423	0.274
U-Net [26]	21.15	0.849	4.584	0.547
StillGAN [41]	21.96	0.861	5.830	0.572
SCI [42]	16.55	0.579	4.837	0.387
Ours	29.98	0.928	8.017	0.639

From Tables 2 and 3, we can see that CNR of network-based methods (LLCNN, UNet, StillGAN, SCI and our proposed SMFA-Pyramid) is significantly higher compared with traditional methods (AHE, CLAHE, and GLP), which indicates that network-based methods have better performance in denoising and contrast improvement. At the same time, our method has the highest in PSNR, SSIM, CNR and XSML values. Both visual effects and objective evaluation show that our proposed SMFA-Pyramid network has the best performance in medical DR image enhancement among all comparison algorithms.

In order to verify the generalization of the proposed algorithm, we conduct tests on the raw images of human body parts not included in the training data. We find that our proposed network also performed well in enhancing DR image of cervical vertebra, lumbar vertebra and some other parts

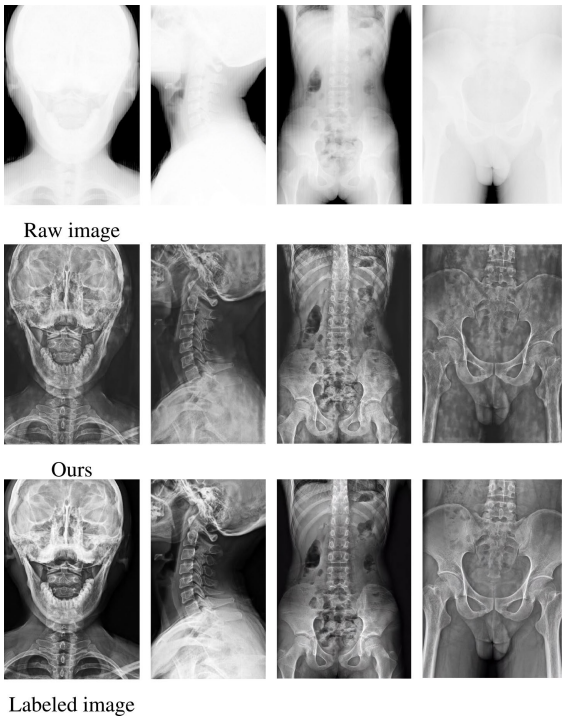


FIGURE 8. The experimental results of different parts of human body beyond the training set.

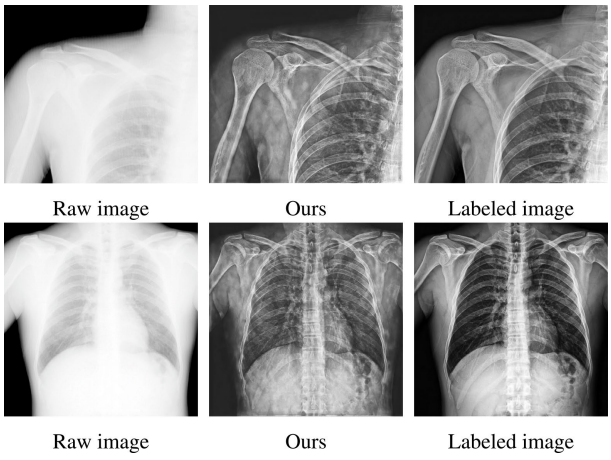


FIGURE 9. The experimental results of different parts of human body beyond the training set.

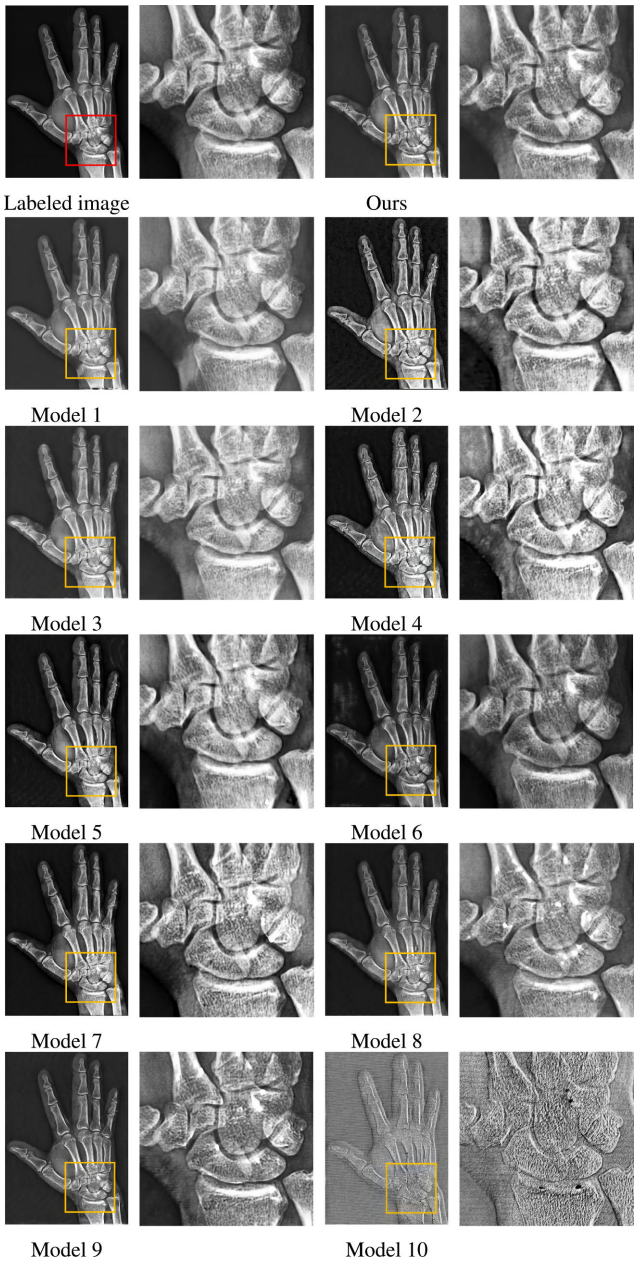


FIGURE 10. Comparison of ablation results.

difficult to enhance, as shown in Figs. 8 and 9. We can see from Fig. 8 that the proposed SMFA-Pyramid model trained with the limbs data can be partially used for the enhancement of medical DR images of other body parts mainly composed of bones and thick soft tissue. And the enhancement effect is relatively better for the parts with more bones. As shown in Fig. 9, our network still has certain shortcomings for the chest cavity, such as the poor enhancement effect on the lung textures. We believe that the performance can be improved by enlarging the dataset and optimizing the network in the near future.

V. ABLATION EXPERIMENT AND RESULT ANALYSIS

In this section, to verify advantages of the addition of BN module, the effectiveness of Pyramid module, the selection of LeakyReLu activation function, the simple multi-scale feature attention SMFA module and our designed pyramid connection and loss function, we conduct the following ablation experiments. We name our SMFA-Pyramid network without BN as Model 1, without SMFA as Model 2, and name LeakyReLu activation function in our model replaced by ReLu activation function the as Model 3. Adding the SMFA module to the first step and the second step downsampling positions in Pyramid network encoder, we name them as Model 4 and Model 5 respectively. Replacing the Laplacian

TABLE 4. Model description.

Model name	Model description	Objective
Ours	SMFA-Pyramid network	To study the impact of all modules in our model
Model 1	SMFA-Pyramid network without BN	To study the impact of BN in our model
Model 2	SMFA-Pyramid network without SMFA	To study the impact of our SMFA module
Model 3	ReLu activation function instead of LeakyReLu for SMFA-Pyramid network	To study the impact of LeakyReLu activation function in our model
Model 4	Adding the SMFA module to the first step downsampling position in Pyramid network encoder	To study the impact of SMFA module in the shallow layer of the encoder in our model
Model 5	Adding the SMFA module to the second step downsampling position in Pyramid network encoder	To study the impact of SMFA module in the deeper layer of the encoder in our model
Model 6	The encoder and decoder of SMFA-Pyramid network are skip connected by the element-wise addition	To study the impact of our pyramid skip connection
Model 7	Using only L_1 loss function in our SMFA-Pyramid network to update the parameters of the model	To study the impact of L_1 loss function in L_{total} without L_2 and L_s in our model
Model 8	Using only L_2 loss function in our SMFA-Pyramid network to update the parameters of the model	To study the impact of L_2 loss function in L_{total} without L_1 and L_s in our model
Model 9	Using L_1 and L_2 loss function in our SMFA-Pyramid network to update the parameters of the model	To study the impact of L_1 and L_2 loss function in L_{total} without L_s in our model
Model 10	Using only side penalty loss function L_s in our SMFA-Pyramid network to update the parameters of the model	To study the impact of L_s loss function in our model

TABLE 5. PSNR, SSIM, CNR, and XSML of the ablation experiments.

Algorithms	PSNR/dB	SSIM	CNR	XSML
Model 1	26.20	0.956	6.719	0.372
Model 2	19.16	0.747	5.424	0.291
Model 3	23.51	0.920	6.317	0.336
Model 4	20.55	0.823	5.541	0.309
Model 5	24.62	0.900	6.489	0.342
Model 6	23.59	0.904	6.476	0.339
Model 7	21.51	0.839	6.256	0.314
Model 8	25.28	0.859	6.611	0.338
Model 9	23.32	0.848	6.276	0.328
Model 10	8.325	0.023	2.978	0.079
Ours	30.75	0.979	8.486	0.436

pyramid connection in the SMFA-Pyramid model by the element-wise addition joint without residual, we name it as Model 6. Using only L_1 and only L_2 loss function in our SMFA-Pyramid network to update the parameters of the model, we name them as Model 7 and Model 8 respectively. Model 9 uses L_1 and L_2 loss functions and Model 10 only uses side penalty loss function L_s. We summarize all the models in the ablation experiment in Table 4 for the detailed model description and objectives. We give the visual results of the enhanced images by different models in Fig. 10 and the objective evaluation in Table 5.

From Fig. 10, we can see that Model 1 has a certain enhancement effect on the image, but there are still problems of blur and low contrast, which shows the important role of BN in our network. The enhanced image of Model 2 has obvious detail loss and blurred edges, which validates the important role of SMFA module in detail preservation. The enhanced image of Model 3 has more image information, but the contrast of the image is lower than our model and some bone trabecula details are lost, which shows that the LeakyReLu activation function plays an important role in preserving the negative information of the image. The enhanced image of Model 4 has too much noise and loses details, which indicates that SMFA module cannot extract more features in the shallow layer of the encoder. The details of the enhanced image of Model 5 have been improved than that of Model 4, but there are still some noises, poor background and white spots in the enhanced image, which leads to the loss of image information. The results of Models 4 and 5 show that the attention module SMFA should be added in the deepest layer of the encoder. The contrast and details of the image enhanced by Model 6 are obviously improved, but it still is not as good as our method, which shows that the Laplacian pyramid connection can enhance the texture and the edge information of the image.

The image enhanced by Model 7 has a lot of noises, and the image enhanced by Model 8 has the problem of blur and has white spots or black spots. Model 9 uses the loss function with L_1 and L_2, and the enhanced image is better than Models 7 and 8, but there are still some problems of details loss, the appearance of noise and many white or black spots. The image of Model 10 has lost its original visual effect, but it presents more information on the edges and textures of the image, which reflects the role of side penalty loss function in preserving details and edges. Our proposed SMFA-Pyramid network enhanced images have the best structure, details and contrast than the first ten models, which fully shows that BN, LeakyReLu, SMFA module and Laplacian pyramid connection can get better multi-scale information in medical DR image enhancement.

At the same time, it shows that our loss function can train the network better than only L_1 loss function, L_2 loss function, L_1 and L_2 loss function, and side penalty loss function, which improves the effect of image enhancement. We can see from Table 5 that our proposed SMFA-Pyramid network has the highest PSNR, SSIM, CNR and XSML values, which further demonstrates the effectiveness of our model in enhancing network performance.

VI. DISCUSSION

Aiming at medical static DR image enhancement, we propose a deep learning based model SMFA-Pyramid network. The training data of the model are all limbs of the human body. Though the generalization effect of the model is relatively good, there are still some problems in the enhancement of soft tissue details of medical DR images especially in the chest. Therefore, we are expanding the dataset and further optimizing the network structure. Furthermore, the loss function and the layer number can be optimized. The combination of deep learning based methods and traditional methods needs an intensive study.

VII. CONCLUSION

In this paper, we propose a SMFA-Pyramid deep learning model for medical static DR image enhancement. The whole image can be enhanced through the training of image blocks. Compared with U-Net network, the Pyramid structure in this model has less convolution computation and network size. And Laplacian pyramid connection can preserve image boundary feature information better. Loss function with side output information, batch normalization operation, LeakyReLu, Sigmoid activation function, and SMFA model make the model more effective and accurate for multi-scale feature expression of medical DR images for better enhancement effect. It also prove that the proposed network has good generalization in images enhancement beyond the training set. We are expanding the dataset, and then the network structure will be further optimized to solve the problem of medical DR image enhancement in thoracic, lumbar and other parts.

ACKNOWLEDGMENT

The authors sincerely thank the laboratory of Shinva Medical Instrument Company Ltd., for providing the real anonymous medical DR data for experiments.

REFERENCES

[1] W. Zhu, H. Jiang, E. Wang, Y. Hou, L. Xian, and J. Debnath, "X-ray image global enhancement algorithm in medical image classification," *Discrete Continuous Dyn. Syst.*, vol. 12, nos. 4–5, pp. 1297–1309, 2019.

[2] T. Reutershan, H. H. Effarah, A. Lagzda, and C. P. J. Barty, "Numerical evaluation of high-energy, laser-Compton X-ray sources for contrast enhancement and dose reduction in clinical imaging via gadolinium-based K-edge subtraction," *Appl. Opt.*, vol. 61, no. 6, p. C162, Feb. 2022.

[3] Z. Chen, W. Xue, T. Wang, and D. Ni, "Cardiac motion scoring based on CNN with attention mechanism," in *Proc. 3rd Int. Symp. Image Comput. Digit. Med.*, Xian, China, Aug. 2019, pp. 128–132.

[4] Q. Zhao, W. Yang, and Q. Liao, "AdaSAN: Adaptive cosine similarity self-attention network for gastrointestinal endoscopy image classification," in *Proc. IEEE 18th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2021, pp. 1855–1859.

[5] Y. Cai, J.-G. Yu, Y. Chen, C. Liu, L. Xiao, E. M. Grais, F. Zhao, L. Lan, S. Zeng, J. Zeng, M. Wu, Y. Su, Y. Li, and Y. Zheng, "Investigating the use of a two-stage attention-aware convolutional neural network for the automated diagnosis of otitis media from tympanic membrane images: A prediction model development and validation study," *BMJ Open*, vol. 11, no. 1, Jan. 2021, Art. no. e041139.

[6] J. Joseph, J. Sivaraman, R. Periyasamy, and V. R. Simi, "An objective method to identify optimum clip-limit and histogram specification of contrast limited adaptive histogram equalization for MR images," *Biocybern. Biomed. Eng.*, vol. 37, no. 3, pp. 489–497, 2017.

[7] I. S. Isa, S. N. Sulaiman, M. Mustapha, and N. K. A. Karim, "Automatic contrast enhancement of brain MR images using average intensity replacement based on adaptive histogram equalization (AIR-AHE)," *Biocybern. Biomed. Eng.*, vol. 37, no. 1, pp. 24–34, 2017.

[8] B. Subramani and M. Veluchamy, "MRI brain image enhancement using brightness preserving adaptive fuzzy histogram equalization," *Int. J. Imag. Syst. Technol.*, vol. 28, no. 3, pp. 217–222, Sep. 2018.

[9] Z. Al-Ameen, G. Sulong, A. Rehman, A. Al-Dhelaan, T. Saba, and M. Al-Rodhaan, "An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, pp. 1–12, Dec. 2015.

[10] K. Koonsanit, S. Thongvigitmanee, N. Pongnapang, and P. Thajchayapong, "Image enhancement on digital X-ray images using N-CLAHE," in *Proc. 10th Biomed. Eng. Int. Conf. (BMEiCON)*, Aug. 2017, pp. 1–4.

[11] R. Xing, M. Liu, K. Meng, and S. Mei, "Coupling technique of Haar wavelet transform and variational iteration method for a nonlinear option pricing model," *Mathematics*, vol. 9, no. 14, p. 1642, Jul. 2021.

[12] S. Mei, M. Liu, A. Kudreyko, P. Cattani, D. Baikov, and F. Villecco, "Bendlet transform based adaptive denoising method for microsection images," *Entropy*, vol. 24, no. 7, p. 869, Jun. 2022.

[13] L. Meng, M. Kexin, X. Ruyi, S. Mei, and C. Cattani, "Haar wavelet transform and variational iteration method for fractional option pricing models," *Math. Methods Appl. Sci.*, vol. 46, no. 7, pp. 8408–8417, May 2023.

[14] M. Liu, S. Mei, P. Liu, Y. Gasimov, and C. Cattani, "A new X-ray medical-image-enhancement method based on multiscale Shannon–Cosine wavelet," *Entropy*, vol. 24, no. 12, p. 1754, Nov. 2022.

[15] P. Liu and S. Mei, "Laplace multiscale image enhancement denoising algorithm," *Agricult. Netw. Inf.*, vol. 3, pp. 142–144, Jan. 2010.

[16] X. Tian, J. Wang, D. Du, S. Li, C. Han, G. Zhu, Y. Tan, S. Ma, H. Chen, and M. Lei, "Medical imaging and diagnosis of subpatellar vertebrae based on improved Laplacian image enhancement algorithm," *Comput. Methods Programs Biomed.*, vol. 187, Apr. 2020, Art. no. 105082.

[17] W. Zhu, J. Liu, M. Zhu, Q. Shao, and Y. Yan, "Research on improved algorithm of DR image enhancement based on Gauss–Laplacian pyramid," *Chin. J. Med. Instrum.*, vol. 43, no. 1, pp. 10–13, 2019.

[18] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, Mar. 2000.

[19] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges, and future directions," *IEEE Access*, vol. 7, pp. 99540–99572, 2019.

[20] S. Singh and D. Gupta, "Detail enhanced feature-level medical image fusion in decorrelating decomposition domain," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[21] K. J. Halupka, B. J. Antony, M. H. Lee, K. A. Lucy, R. S. Rai, H. Ishikawa, G. Wollstein, J. S. Schuman, and R. Garnavi, "Retinal optical coherence tomography image enhancement via deep learning," *Biomed. Opt. Exp.*, vol. 9, no. 12, p. 6205, 2018.

[22] X. Liang, L. Chen, D. Nguyen, Z. Zhou, X. Gu, M. Yang, J. Wang, and S. Jiang, "Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy," *Phys. Med. Biol.*, vol. 64, no. 12, Jun. 2019, Art. no. 125002.

[23] Y. Chen, X. H. Yang, Z. Wei, A. A. Heidari, N. Zheng, Z. Li, H. Chen, H. Hu, Q. Zhou, and Q. Guan, "Generative adversarial networks in medical image augmentation: A review," *Comput. Biol. Med.*, vol. 144, Jul. 2022, Art. no. 105382.

[24] K. Xia, Q. Zhou, Y. Jiang, B. Chen, and X. Gu, "Deep residual neural network based image enhancement algorithm for low dose CT images," *Multimedia Tools Appl.*, vol. 81, no. 25, pp. 36007–36030, Oct. 2022.

[25] M.-I. Georgescu, R. T. Ionescu, A.-I. Miron, O. Savencu, N.-C. Ristea, N. Verga, and F. S. Khan, "Multimodal multi-head convolutional attention with various kernel sizes for medical image super-resolution," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2023, pp. 2194–2204.
[26] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.
[27] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021.
[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[29] J. Xu, Z. Li, B. Du, M. Zhang, and J. Liu, "Reluplex made more practical: Leaky ReLU," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jul. 2020, pp. 1–7.
[30] A. K. Dubey and V. Jain, "Comparative study of convolution neural networks ReLU and leaky-ReLU activation functions," in *Applications of Computing, Automation and Wireless Systems in Electrical Engineering* (Lecture Notes in Electrical Engineering). Singapore: Springer, 2019, pp. 873–880.
[31] J. Wang, H. Zhu, S. H. Wang, and Y. D. Zhang, "A review of deep learning on medical image analysis," *Mob. Netw. Appl.*, vol. 26, pp. 351–380, Jun. 2021.
[32] L. Jiao and J. Zhao, "A survey on the new generation of deep learning in image processing," *IEEE Access*, vol. 7, pp. 172231–172263, 2019.
[33] F. Zou, L. Shen, Z. Jie, W. Zhang, and W. Liu, "A sufficient condition for convergences of Adam and RMSProp," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11119–11127.
[34] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U^2-Net: Going deeper with nested U-structure for salient object detection," *Pattern Recognit.*, vol. 106, Oct. 2020, Art. no. 107404.
[35] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study," *J. Comput. Commun.*, vol. 7, no. 3, pp. 8–18, 2019.
[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[37] A. Rodriguez-Molares, O. M. H. Rindal, J. D'Hooge, S.-E. Måsøy, A. Austeng, M. A. L. Bell, and H. Torp, "The generalized contrast-to-noise ratio: A formal definition for lesion detectability," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 4, pp. 745–759, Apr. 2020.
[38] A. Thelen, S. Frey, S. Hirsch, and P. Hering, "Improvements in shape-from-focus for holographic reconstructions with regard to focus operators, neighborhood-size, and height value interpolation," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 151–157, Jan. 2009.
[39] H. Marmolin, "Subjective MSE measures," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, no. 3, pp. 486–489, Jul. 1986.
[40] L. Tao, C. Zhu, G. Xiang, Y. Li, H. Jia, and X. Xie, "LLCNN: A convolutional neural network for low-light image enhancement," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
[41] Y. Ma, J. Liu, Y. Liu, H. Fu, Y. Hu, J. Cheng, H. Qi, Y. Wu, J. Zhang, and Y. Zhao, "Structure and illumination constrained GAN for medical image enhancement," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3955–3967, Dec. 2021.
[42] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, Jun. 2022, pp. 5627–5636.

WENJING XUE received the bachelor's degree from Ludong University, in 2020. She is currently pursuing the master's degree with Shandong University of Technology. Her current research interests include medical image processing and restoration.

YINGMEI WANG received the bachelor's degree in mathematics and applied mathematics from Qufu Normal University, in 2009, and the master's and Ph.D. degrees in applied mathematics from Shandong University, in 2012 and 2016, respectively. She is currently an Associate Professor with the School of Mathematics and Statistics, Shandong University of Technology. Her current research interests include medical image restoration and reconstruction.

ZHIEN QIN received the bachelor's degree from Qufu Normal University, in 2007, and the master's degree from Shandong University, in 2010. He is currently an Engineer with Shinva Medical Instrument Company Ltd. His current research interests include medical image processing software and product development.