Research on Image Sharpness Enhancement Technology based on Depth Learning

Wenbao Lan*, Chang Che

School of Civil Engineering, Harbin University, Harbin, 150086, China

Abstract—**Image technology is widely used in security, traffic, monitoring, and other social activities. However, these images carrying detailed information will have feature distortion due to various external physical factors in social ingestion, transmission, and storage, resulting in poor image quality and clarity. Resolution determines the definition of an image. Super-resolution reconstruction is the process of transforming a low-resolution picture into a high-resolution image. To enhance the image clarity, this experiment introduces the advantages and disadvantages of the Super Resolution Convolutional Network (SRCNN) and Fast Super Resolution Convolutional Neural Network (FSRCNN) model and then constructs an image super resolution method based on DSRCNN. The algorithm consists of two sub-network blocks, an enhancement block and a purification block. The model first uses two Convolutional Neural Networks (CNN) to obtain complementary low-frequency information that improves the model's learning ability; next, it employs an enhancement block to fuse the image features of two paths via residual operation and sub-pixel convolution to prevent the loss of low-resolution image information; finally, it employs a feature purification block to refine high-frequency information that more accurately represents the predicted high-quality image. It is found that the PSNR and SSIM of the DSRCNN model can reach 33.43dB and 0.9157dB, respectively.**

Keywords—*Super resolution convolutional network; fast super resolution convolutional neural network; dual super-resolution convolutional network ; deep learning; image definition; super resolution; image enhancement*

I. INTRODUCTION

Pictures captured in social life are often affected by the light, shelter, poor shooting equipment, and other factors. Such shootings have poor visibility, low contrast, high noise, and other shortcomings [1]. These shortcomings make the image definition fail to reach the ideal state, which affects the regular operation of some links to social activities. Picture sharpness is directly related to the image resolution. High-resolution images have more detail and are of superior quality [2]. Processing a series of low-resolution images to obtain high-resolution images is called super-resolution reconstruction. Traditional image super-resolution reconstruction is mainly based on interpolation, machine learning, and multiple images. Researchers value the super-resolution reconstruction technology based on deep learning algorithms because of the ongoing changes in computing technology [3]. The deep learning algorithm mainly learns the high-level semantic features of data through the transformation of multi-level nonlinear structure. It can achieve reasonable and effective prediction and analysis of data [4]. Classical depth learning algorithms include SRCNN, FSRCNN, VDSR, etc. These super-resolution reconstruction technologies can improve image clarity, but there are still some limitations in operation efficiency and visual effects [5]. Therefore, this experiment suggests an image definition improvement technique based on deep learning algorithms to improve the picture quality with inadequate definition, aiming to add reference value to image applications in social activities such as security, traffic, monitoring, etc.

II. RELATED WORK

At present, the application of depth learning is very broad, and it has been dabbled in various research fields. Wang S et al. used depth learning algorithm to recognize tumor images, and introduced the whole process of depth learning algorithm in pathological image segmentation in detail. The research shows that the deep learning algorithm has obvious advantages in recognition accuracy, overall computational efficiency and generalization performance. In addition, integrating in-depth learning into the field of pathological image analysis is a far-reaching experiment in medicine [6]. Huang and other researchers found that traditional machine learning methods had limited prediction ability for landslide prone environmental factors, so they proposed a FC-SAE algorithm based on deep learning. To predict categorization, the system employs both high-level and low-level information about environmental elements. The simulation results show that the prediction rate and accuracy rate of FC-SAE can reach 85.4% and 85.2%, higher than other algorithms of the same type, so it has a good application prospect [7]. Sharma and other scholars explored deep neural networks based on machine learning. For the goal of conducting research on the use of machine learning and deep learning in different sectors, they examined and improved the current machine learning and deep learning algorithms. The final survey results show that the accuracy of deep learning algorithms has been significantly improved [8]. Zhang and other researchers conducted research on deep learning algorithms in the field of geological engineering. This paper mainly studies the application of FNN, RNN, CNN and GAN depth learning algorithms in geotechnical engineering, compiles a detailed summary containing literature, cases and depth learning algorithms, and prospects the application of depth learning algorithms in geotechnical engineering [9]. Aslan and others found that malware is constantly developing and seriously endangering the network security environment. Based on this malignant situation, they propose to use deep learning algorithm to detect malicious software. The suggested deep learning algorithm is a novel hybrid architecture that combines two comprehensive pre-training network models, which makes it different from conventional approaches. The

experimental findings demonstrate that the model can classify malware with an accuracy of 97.78% [10]. Bamisile and other scholars used machine learning algorithm and depth learning algorithm to predict solar radiation. SVR, RF, and their learning models as well as ANN, CNN, and RNN deep learning models are each utilized in the experiment to forecast solar radiation, and the outcomes of these predictions are compared. The final results show that the deep learning model has better accuracy and takes less time than the machine learning model in solar radiation prediction [11].

For the research of image enhancement technology, most researchers have made contributions to it. An underwater image enhancement approach based on MLLE was suggested by Zhang and other researchers to address the issues of poor color and low visibility of underwater photos. The method first adjusts the color and detail of the image by the principle of minimum color loss, then uses the integral graph to calculate the mean and variance of the image block, and finally introduces a color balance strategy to balance the color difference of the image. The outcomes of the experiment demonstrate that this technique is very applicable to picture enhancement [12]. Ike C and other researchers proposed a single image super-resolution method based on wavelet. The experiment combines wavelet transform with local regularization anchored domain regression model and applies it to image restoration. Experimental results show that the proposed method is effective and superior in denoising, deblurring and super-resolution reconstruction tasks [13]. Zhang and other scholars proposed a method that can enhance the image contrast while improving the embedded data, namely the improved RDH-CE model. The model offers an adaptive pixel modification technique and extracts several characteristics to aid K-means clustering. According to the final testing findings, this technique has a greater accuracy level for complicated assessment [14]. Liu and other researchers used MSRCR algorithm and guided filtering method to de fog the image. To provide a better visual impact, this approach may successfully address the issues of poor picture lighting, color distortion, and edge loss while preserving the color saturation of the image [15]. A band restricted biphasic technique was suggested by Su et al. to enhance the quality of the reconstructed picture. By using this technique, the precise spatial frequency components impacted

by the spatial offset of DPHs may be removed, preventing the picture quality from being compromised. According to the outcomes of the simulation experiment, the strategy enhances the definition of the rebuilt picture by 36.84% [16]. To solve the serious distortion problem of underwater images, Liu et al. proposed an image enhancement method based on object guided dual confrontation contrast learning, which uses contrast prompts in the training phase. The experiment demonstrates that the suggested strategy enhances the image's visual quality and increases the detector's precision [17].

Deep learning algorithm is favored by researchers in various fields for its convenient and powerful performance, and its application is also more extensive. In recent years, the research on image enhancement technology has never stopped, and many scholars have used different algorithms to explore it. In conclusion, deep learning algorithm and image enhancement technology have made considerable progress. It is an innovative attempt to apply the depth learning algorithm to the image definition enhancement technology, so this experiment starts to study this topic.

III. CONSTRUCTION OF IMAGE DEFINITION ENHANCEMENT MODEL BASED ON DEPTH LEARNING ALGORITHM

A. Research on Image Super-resolution Model based on Depth Learning

The resolution of an image determines its definition, and super resolution is the resolution of an image that has been enhanced using a variety of efficient techniques. Super Resolution refers to the technique of processing a number of low-resolution photographs to produce high resolution images (SP), which mainly includes the reconstruction of a single image and the reconstruction of multiple images [18-19]. Super Resolution Convolutional Network (SRCNN) is a milestone for depth learning algorithm to achieve super-resolution image reconstruction [20]. SRCNN uses a three-layer convolutional neural network to achieve image reconstruction, that is, to preprocess the low-resolution image through cubic interpolation. Fig. 1 illustrates the three components that make up the SRCNN model, namely block precipitation and representation, nonlinear mapping and reconstruction. The three modules correspond to three convolution operations in the network.

n_1 feature mapping of low resolution images

n_2 feature mapping of high resolution images

$f_1 \times f_1$

1×1

$f_3 \times f_3$

Low resolution image (input)

High resolution images (output)

Block precipitation and indication

Nonlinear mapping

Reconstruction
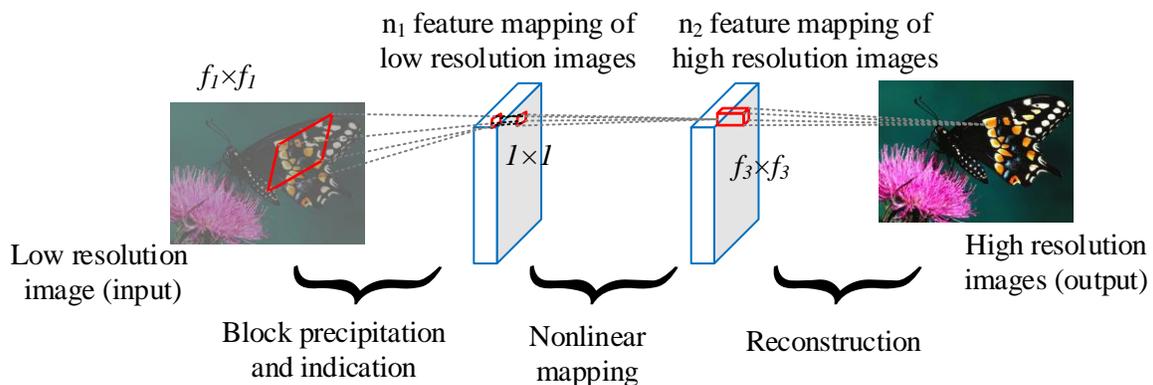
Fig. 1. Network structure of SRCNN.

B represents low resolution image; A represent real high-resolution images; the mapping function of the network is $f(\cdot)$ .Consequently, the equation for block precipitation and module expression in SRCNN is as follows:

$$f_1(B) = \max(0, W_1 * B + \delta_1) \tag{1}$$

In Formula (1), W_1 represents the weight of the convolution core of the block precipitation module; δ_1 indicates the offset of the convolution core of the block precipitation module. $\max$ represents the corresponding ReLU activation function; $f_1(B)$ represents block precipitation. The nonlinear mapping of the network is called "convolution+activation", and its functional expression is:

$$f_2(B) = \max(0, W_2 * f_1(B) + \delta_2) \tag{2}$$

In Formula (2), W_2 represents the weight of the nonlinear mapping module's convolution kernel, B_2 represents the offset of the convolution kernel of the nonlinear mapping module; $f_2(B)$ represents a nonlinear mapping. The network reconstruction process is the third convolution operation. Unlike the two modules mentioned above, the activation function is no longer used in this process. Its function expression is:

$$f(B) = W_3 * f_2(B) + \delta_3 \tag{3}$$

In Formula (3), W_3 represents the weight of the convolution core of the reconstruction module, δ_3 represents the offset of the convolution core of the reconstruction module, and $f(B)$ represents the reconstruction. The advantage of SRCNN model is that it contains fewer parameters than other deep convolution neural networks. As a result, by using fewer learning parameters, it may significantly minimize the number of calculations required by the model, thus improving the overall model operation efficiency. The loss function of SRCNN is the mean variance between the super resolution image $f(B)$ and the original real high resolution image A . The specific calculation formula is as follows:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \left\| f(B_i; \theta) - A_i \right\|^2 \tag{4}$$

In Formula (4), n represents the number of samples in the training set, $f(B_i, \theta)$ represents the minimum reconstructed image, and A_i represents the high-resolution image. The loss of mean variance represents the loss mean of a batch, which is an evaluation index of commonly used image high-resolution algorithms. Although the classic SRCNN network has been able to repair low resolution images with different coefficients, the SRCNN network still has limitations that cannot meet the real-time requirements. Thus, FSRCNN network came into being. FSRCNN network consists of five parts: feature extraction, compression, mapping, expansion and transpose

convolution [21-22]. The FSRCNN network is used to extract and compress the image features, and then the activation function pair is used for nonlinear mapping. The activation function of FSRCNN may activate the gradient of the negative half axis in addition to the SRCNN's activation function. The precise formula may be written as follows:

$$g(x_i) = \max(x_i, 0) + a_i * \min(0, x_i) \tag{5}$$

In Formula (5), x_i represents the input signal of the activation function in the i -th channel; a_i represents the coefficient of the negative part. The activation function has more stable performance. The calculation complexity of FSRCNN model can be expressed as:

$$O\left\{\left(25d + sd + 9ms^2 + ds + 81d\right)S_{LR}\right\} = O\left\{\left(9ms^2 + 2sd + 106d\right)S_{LR}\right\} \tag{6}$$

In Formula (6), d represents the characteristic layer before compression; s represents the compressed feature layer; And both meet $s < d$. m indicates the number of layers of the volume. It can be seen from Formula (6) that the computational complexity of FSRCNN model is in linear proportion to LR image, and its computational complexity is also less than SRCNN. The loss function calculation formula of FSRCNN can be expressed as:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left\| f(B_s^i; \theta) - A^i \right\|_2^2 \tag{7}$$

In Formula (7), A^i represents the i -th high-definition image in the data set; B_s^i represents the i -th resolution image in the dataset. $f(B_s^i; \theta)$ represents the output of the model. Compared with SRCNN network, FSRCNN network is mainly improved in three aspects, namely, the change of feature dimension, the replacement of convolution core with more mapping layers and the sharing of mapping layers. Although SRCNN and FSRCNN both improve the quality of super resolution images, they are generally slow and not suitable for real application scenarios. The research on the application of depth learning algorithm to the improvement of super-resolution image quality needs to be further deepened.

B. Improvement of Image Definition Enhancement Model based on Fusion of Depth Learning Algorithm

In this experiment, a dual super resolution CNN (DSRCNN) is designed based on SRCNN and FSRCNN models. The network consists of two sub network modules, enhancement blocks and feature purification blocks. Two sub network blocks enhance the super resolution performance of the network by extracting complementary low-frequency features; the enhancement block gathers several high-frequency characteristics via residual operation and sub-pixel convolution; the feature purification block uses multiple stacked convolutions to refine high-frequency features [23-24]. The process of extracting complementary low-frequency features from two sub network blocks of DSRCNN model is as follows:

$$O_{TSEB} = f_{TSEB}\left(I_{LR}\right) = \left(Cat\left(O_{(TSEB_1)}, O_{(TSEB_2)}\right)\right) \quad (8)$$

In Formula (8), O_{TSEB_1} and O_{TSEB_2} represent the output of two subnetworks, and *Cat* Crepresents the connection operation. In the second stage, the two subnetworks select the combination of reusing residual operation and convolution to enhance the effect of local level information. The formula can be expressed as:

$$O_{TSEB_k} = R\left(O_{L1} + O_{L2} + O_{L5} + O_{L7} + O_{Li} + \cdots\right) \quad (9)$$

In Formula (9), $O_{(TSEB_k)}$ represents the output of k sub network, O_{Li} represents the output of the i-th convolution layer, and R represents ReLU function. The purpose of the model's enhancement block is to combine the original picture data and gain deeper data. This module first obtains the original image and deeper information, and obtains the characteristics of the first layer in the sub network:

$$O_{(EB_s)} = Subp\left(R(O_{L1})\right)$$
$$O_{(EB_d)} = Subp\left(O_{TSEB}\right) \quad (10)$$

In Formula (10), $O_{(EB_s)}$ stands for the shallow layer's high-frequency properties, and $O_{(EB_d)}$ for the deep layer's high-frequency features. *Subp* represents sub-pixel convolution technology. Then, the second level features are obtained from the features in the residual operation fusion formula (10), and its functional expression is:

$$O_{EB} = f_{EB}\left(O_{TSEB}\right) = R\left(O_{(EB_s)} + O_{(EB_d)}\right) \quad (11)$$

In Formula (11), $O_{(EB)}$ represents the output of the enhancement block, serves as the purification block of the feature, $+$ represents a residual operation, and f_{EB} represents the function of the enhancement block. The feature purification block equation can be expressed as:

$$I_{SR} = f_{FLB}\left(O_{EB}\right) = C\left(R\left(C\left(R\left(C\left(R\left(C\left(R\left(C(O_{EB})\right)\right)\right)\right)\right)\right)\right)\right) \quad (12)$$

Formula (12) constructs a five layer feature purification block, which contains four convolution operations and activation functions, and can be used to build high-quality predicted images. In Formula (12), C represents the convolution operation, and I_{SR} represents the predicted SR image. The DSRCNN's execution procedure may be stated as follows:

$$I_{SR} = f\left(f_{EB}\left(f_{TSEB}\left(I_{LR}\right)\right)\right) = f_{DSRCNN}\left(I_{LR}\right) \quad (13)$$

In Formula (13), f_{TSEB} represents the function of two subnets, f_{EB} represents the function of enhancement block, f_{FLB} represents the function of feature purification block, and

I_{LR} represents the given LR image. DSRCNN mainly obtains complementary low-frequency information through dual CNN to improve the learning ability of the model. Then, to prevent the loss of low resolution picture information, the enhancement block merges the image characteristics of two separate image routes using residual operation and sub-pixel convolution. In order to properly reflect the anticipated high-quality picture, high-frequency information is refined using the feature purification block [25].The above process is described in detail in Fig. 2.
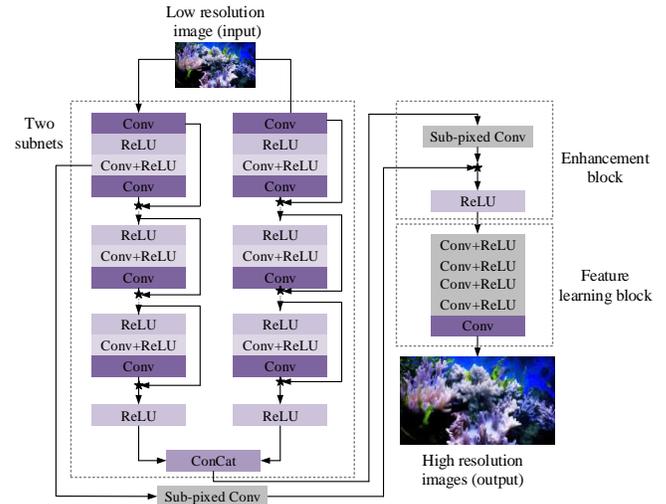
Fig. 2.   Network architecture of DSRCNN.

The difference between the actual high-quality picture and the anticipated SR image serves as the mean square error, which is how the loss function of DSRCNN is still stated. The specific function expression is:

$$l(p) = \frac{1}{2N} \sum_{(k=1)}^{N} \left\| f_{DSRCNN}\left(I_{LR}^{k} - I_{HR}^{k}\right) \right\|^2 \quad (14)$$

In Formula (14), I_{LR}^{k} represents a given low resolution image, I_{HR}^{k} represents a given high resolution image, N represents the total number of training samples, and p represents the parameter set for training a DSRCNN. The value of $l(p)$ is calculated from the training pair $\left\{I_{LR}^{k}, I_{HR}^{k}\right\}$. PSNR and SSIM can be selected as quantitative and objective evaluation indicators for model evaluation. Suppose the length and width of a noise image are M and N respectively, K represents the noise image, I represents the clean image, and their mean variance formula is:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[I(i,j) - K(i,j)\right]^2 \quad (15)$$

In Formula (15), i represents the feature in the i-th column of the picture, j represents the feature in the j-th column of the picture, and the formula of PSNR can be expressed as:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (16)$$

PSNR is measured in dBs. The less the value is, the less the picture will be warped. The PSNR index, which bases its judgment of picture quality on the associated pixel error, ignores the visual properties of the human eye. Because the human eye has a high visual sensitivity to space, brightness, color, etc., the evaluation results of PSNR are often inconsistent with the subjective feelings of the human eye. Structural similarity requires difference processing for target images X and Y. μ_X and μ_Y respectively stand for the mean values of images X and Y, σ_X and σ_Y respectively stand for the variances of images X and Y, and σ_{XY} stands for the covariance of images X and Y. The formula is as follows:

$$\mu_X = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} X(i,j)$$

$$\sigma_X = \left(\frac{1}{M \times N - 1} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(X(i,j) - \mu_X \right)^2 \right)^{\frac{1}{2}}$$

$$\sigma_{XY} = \frac{1}{M \times N - 1} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(X(i,j) - \mu_X \right) \left(Y(i,j) - \mu_Y \right) \quad (17)$$

SSIM assesses how similar pictures are in terms of structure, contrast, and brightness. The contrast formula is as follows:

$$l(X,Y) = \frac{2\mu_X \mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}$$

$$c(X,Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}$$

$$s(X,Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \quad (18)$$

In Formula (18), c_1, c_2 and c_3 are three constants, and there is no coefficient of any variable between them, but they cannot be removed. The formula of SSIM can be expressed as:

$$SSIM(x,y) = \frac{\left(2\mu_x \mu_y + c_1 \right) \left(2\sigma_{xy} + c_2 \right)}{\left(\mu_x^2 + \mu_y^2 + c_1 \right) \left(\sigma_x^2 + \sigma_y^2 + c_2 \right)} \quad (19)$$

SSIM has a value range of 0 to 1. The picture distortion decreases with increasing value, and image quality and clarity increase. In the calculation process, it is also necessary to cut the global pixels of the image, and the structural similarity obtained is the global structural similarity.

IV. PERFORMANCE VERIFICATION OF IMAGE DEFINITION ENHANCEMENT MODEL BASED ON DSRCNN

To more accurately verify the image super-resolution reconstruction performance of the model, four data sets, Set5, Set14, BSD100 and Urban100, were selected for the experiment, and the contrast experiments were carried out under the conditions that the image magnification coefficient

was 2 times, 3 times and 4 times. Set5 consists of five images: children, birds, butterflies, flowers and women. 14 photos make up the data set known as "Set14," which is often used to evaluate the effectiveness of image super-resolution models. The BSD100 dataset contains 100 daily life events, which are often used for image denoising and super-resolution testing. Urban100 is an image data set with the theme of street view, in which the images show complex color characteristics, and their objects have clear shapes and corners. The conditions under which Set5 and Set14's photos were taken are the same. BSD100 and Urban100 both provide 100 photos in natural color. The experimental training set and test set each include 100 natural photos, and the image data in the data set DIV2K is chosen. Table I shows the settings of other experimental parameters.

TABLE I.   EXPERIMENTAL PARAMETER SETTING TABLE

Experimental parameters	Set value	Experimental parameters	Set value
Batch size	64.000	Training step	600000
Parameters Beta1	0.900	Epsilon	100000000
Parameters Beta2	0.999	Learning rate of model training 1	0.0004
Initial learning rate	0.001	Learning rate of model training 2	0.00005

Fig. 3 shows the visual test results of SRCNN model, FSRCNN model and DSRCNN model on visual graphics. To create these visual graphics, select a specific area of the anticipated super-resolution image as the observation area. The data in the figure shows that the PSNR and SSIM values of SRCNN model are 27.002 and 0.928; PSNR and SSIM values of FSRCNN model are 30.783 and 0.8666; The PSNR and SSIM values of the DSRCNN model are 33.783 and 0.9688. It can be observed from the above data that in the three models, since DSRCNN's PSNR and SSIM values are the greatest, its images exhibit the least amount of distortion. This demonstrates that the DSRCNN model performs superior to the other two models.

(a) SRCNN
(27.002/0.928)

Image
(PSNR/SSIM)

(b) FSRCNN
(30.783/0.8666)

(c) DSRCNN
(33.175/0.9688)
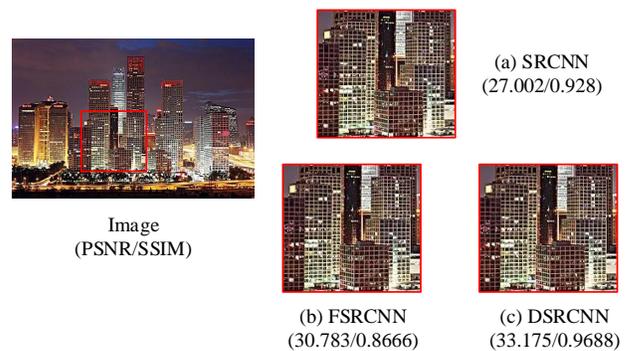
Fig. 3.   Visual diagram of three models.

Table II shows the comparison of PSNR and SSIM values of different modules of DSRCNN in Urban100 dataset with twice the image magnification factor. The results in the table demonstrate that local residual learning is efficient for the development of model performance as the PSNR and SSIM values of the DSRCNN network are better than those of the DSRCNN model with residual learning method. It is observed

that the DSRCNN model is superior to the DSRCNN with only one subnet, and the two subnets in the table model contribute to the performance improvement of the model. The PSNR value of DSRCNN with enhancement block is 0.454dB higher than that of DSRCNN without enhancement block, which verifies the improvement of network performance by enhancement block. The PSNR and SSIM values of the DSRCNN model with feature purification block are 0.16dB and 0.011dB higher than those of the DSRCNN model without feature purification block, respectively, which proves that the feature purification block can improve the network performance. As a result, the three modules in the DSRCNN model proposed in the experiment all have different contributions to their own performance improvement, and this model can improve the image quality.

To better compare the performance of the DSRCNN model proposed in this paper, in addition to the comparison with the SRCNN model, the experiment also selects six other popular depth learning models, a total of 10 models, to conduct experiments to compare and analyze their performance. The popular models selected are Bicubic model, CNF model, MenNet model, CARN-M model, WaveRestNet model and LESRCNN-S model. Fig. 4 displays the variations in PSNR and SSIM values for the chosen model on the Set5 dataset at 2, 3, and 4 times the original picture size. When the image is magnified twice, the DSRCNN model can have the next highest PSNR value, 37.74dB. When the image magnification is 3 times, the DSRCNN model has the highest PSNR value, 34.17dB. When the image magnification is 4 times, the DSRCNN model achieves the highest SSIM value of 08907. On the whole. The DSRCNN model shows relatively stable performance on the Set5 dataset.

Fig. 5 displays the changes in PSNR and SSIM values for several models using the Set14 dataset with picture magnifications of 2, 3, and 4. When the image is magnified twice, the PSNR value of the DSRCNN model is 33.43, which is the highest among all models. At the same time, its SSIM

value is 0.9157, which is the highest among all models. When the magnification of the image is 3 times, the PSNR of the DSRCNN model is still the highest among all models, 30.24, and its SSIM value is 0.8402, which also keeps the highest record. When the image magnification is 4 times, the PSNR value of the DSRCNN model is 28.46, and the SSIM value is 0.7796. Although it has decreased, it still maintains the highest value of all models. On the Set14 dataset, it is clear that the DSRCNN model performs well.

Fig. 6 displays the variations in PSNR and SSIM values for several models on the BSD100 dataset at 2, 3, and 4 times the original picture size.When the image magnification is 2 times, the PSNR value of the DSRCNN model is 32.05 and the SSIM value is 0.8978. When the image magnification is 3 times, the PSNR value of the DSRCNN model is 29.01 and the SSIM value is 0.8029, ranking first among all models. When the image magnification is 4 times, the PSNR value of the DSRCNN model is 27.50, and the SSIM value is 0.7341, which is also in the first place. This shows that on the BSD100 dataset, the DSRCNN model still has better performance than other models.

TABLE II.   COMPARISON OF PSNR VALUE AND SSIM VALUE OF DIFFERENT MODULES IN DSRCNN ON URBAN100 DATASET WITH TWICE IMAGE MAGNIFICATION FACTOR

Methods	Urban100
	PSNR/SSIM
DSRCNN with one sub-network ,without RLO	31.675/0.9236
DSRCNN with one sub-network and EB_S, without RLO,	31.221/0.9181
DSRCNN with one sub-network EB_S andRO,without RLO	31.061/0.9170
DSRCNN with EB_S , without RLO	31.648/0.9237
DSRCNN without RLO	31.700/0.9237
DSRCNN	31.834/0.9253
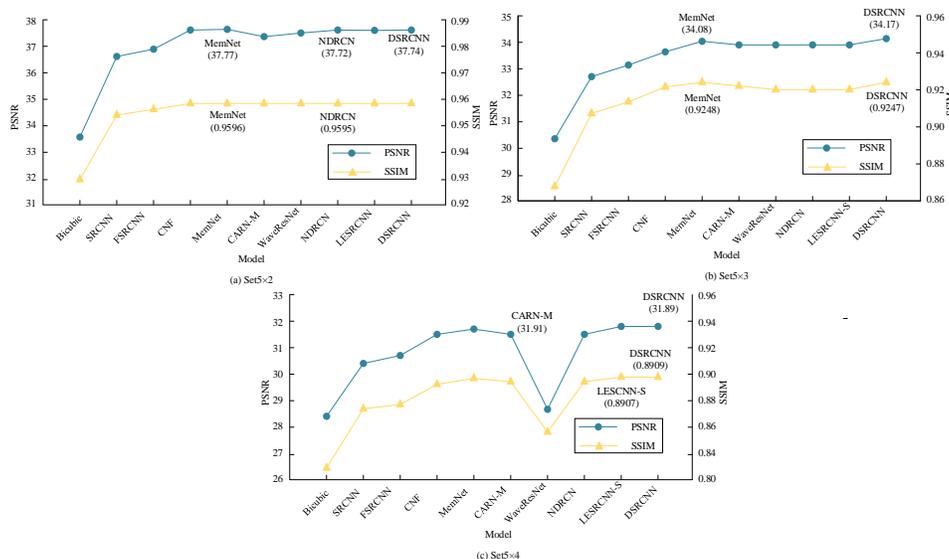
Fig. 4.   Changes of PSNR and SSIM values of different models on Set5.
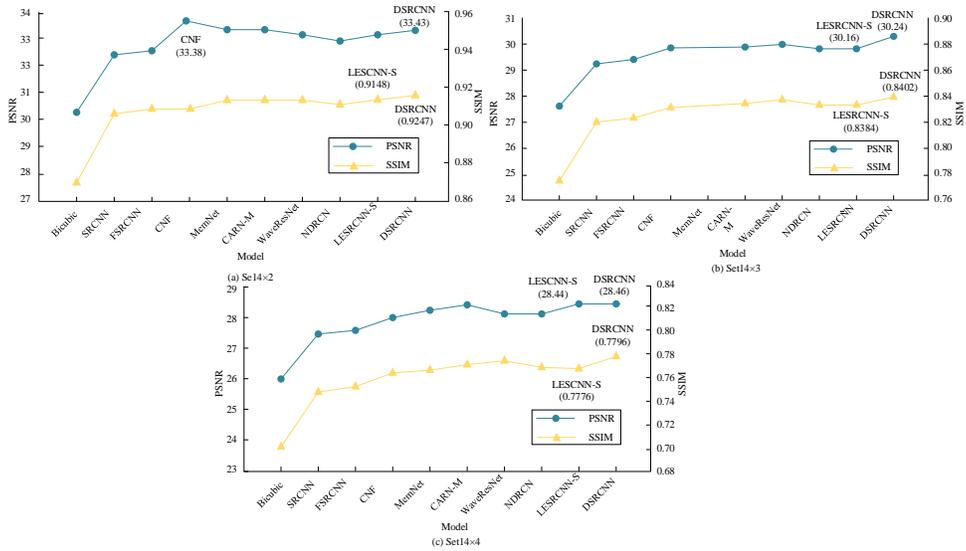
Fig. 5.   Changes of PSNR and SSIM values of different models on Set14.
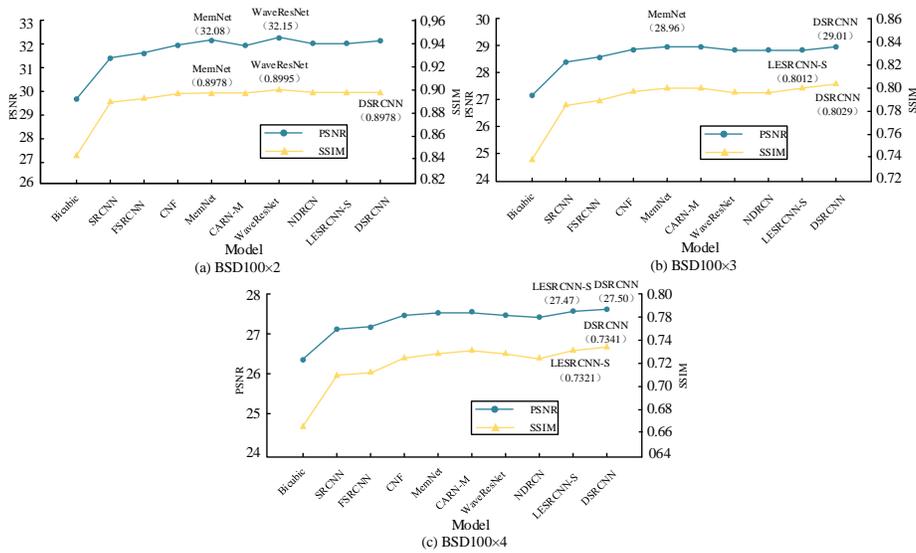
Fig. 6.   Changes of PSNR and SSIM values of different models on BSD100.

Fig. 7 displays the variations in PSNR and SSIM values between models using the Urban100 dataset with picture magnifications of 2, 3, and 4. When the image magnification is 2 times, the PSNR value of the DSRCNN model is 31.83, and the SSIM value is 0.9252, which is the highest among all models. When the image magnification is 3 times, the PSNR value of the DSRCNN model is 27.76, and the SSIM value is 0.8483, ranking the highest among all models. When the image magnification is 4 times, the PSNR value of the DSRCNN model is 25.94, and the SSIM value is 0.7815, which is always the first of all models. This shows that the DSRCNN model also has good performance on the Urban100 dataset.

Fig. 8 shows that on the dataset Urban100, when the magnification factor is 4 times the visual effect of different models, the data in the figure shows that the PSNR and SSIM values are highest for the DSRCNN model. To test the effectiveness of the DSRCNN model based on the deep learning algorithm, four distinct datasets are chosen. The

results showed that the PSNR and SSIM values of DSRCNN model fluctuated in different data sets for different models, but among the 10 chosen models, image processing performance as a whole was best. This shows that the DSRCNN model can handle the image processing, making the final image less distorted, improving its separation rate, and thus enhancing the image clarity.

Table III shows the network complexity analysis results of DSRCNN and the four models selected in the experiment. It can be seen from the table that the number of parameters of DSRCNN model is much smaller than that of the other models on the Urban100 dataset with 2, 3 and 4 times of image magnification. In addition, its number of operations (Flpos) is also less than other models. The comparison results further reflect the superior data processing ability and performance of DSRCNN. At the same time, it also verifies the innovation of DSRCNN for image data processing methods.
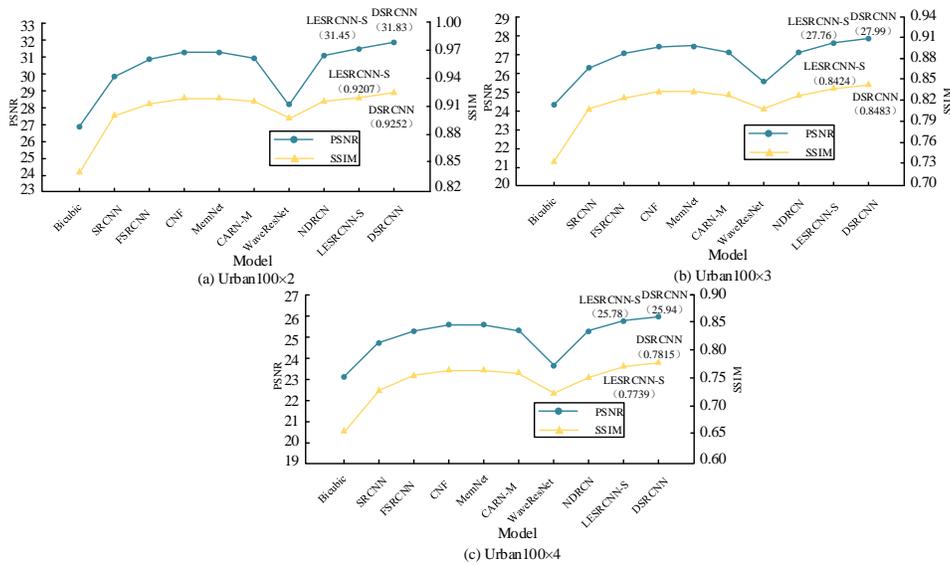
Fig. 7. Changes of PSNR and SSIM values of different models on Urban100.
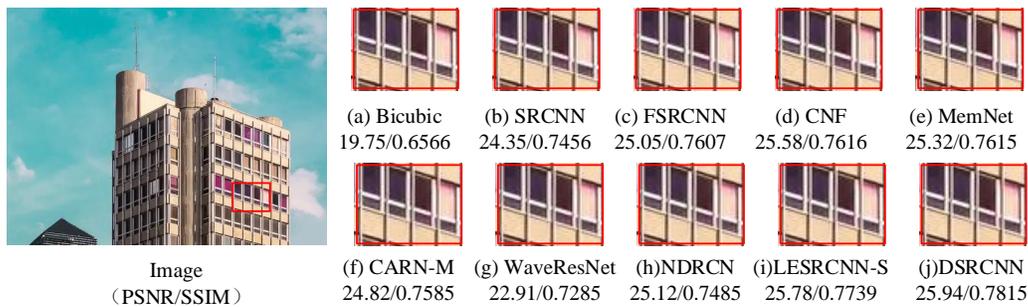
Fig. 8. Visual effects of different models.

TABLE III. COMPARISON OF MODEL PARAMETERS

Methods	Scale×2		Scale×4		Scale×6	
	Parameter quantity(K)	Flops(G)	Parameter quantity(K)	Flops(G)	Parameter quantity(K)	Flops(G)
WaveRe Net	664	10.90	675	11.23	789	12.35
LESRCNN-s	557	9.12	603	10.12	667	11.24
SRCNN	879	29.06	978	30.14	1023	31.23
FSRCNN	789	11.09	856	12.03	964	12.99
DSRCNN	550	5.56	563	6.23	574	7.02

V. CONCLUSION

Image technology is widely used in security, traffic, monitoring, and other social activities because it can carry detailed information. To improve the distortion in the process of image acquisition, transmission and storage, the DSRCNN model for image super-resolution reconstruction is established from the perspective of image super-resolution reconstruction and depth learning technology. This experiment aims to improve the image definition. In social ingestion, transmission, and storage, image features will be distorted due to various external factors, resulting in poor image quality and clarity. In this experiment, a DSRCNN model is built for image super-resolution reconstruction from the viewpoint of image super-resolution reconstruction and depth learning techniques to improve the definition of the picture. At first, the performance of the DSRCNN model, SRCNN model, and FSRCNN is compared. The PSNR and SSIM values of the SRCNN model are 27.002 and 0.928. The FSRCNN model is 30.783 and 0.8666, the DSRCNN model is 33.783 and 0.9688, and so on, which verifies that the DSRCNN model has superior performance. To further explore the DSRCNN model, four different data sets are selected to verify the model performance. Under various picture magnifications, the PSNR and SSIM values of DSRCNN models in various data sets likewise change. In the data set Set14, when the magnification is 2, the DSRCNN model has the highest PSNR and SSIM values, 33.43 and 0.9157, respectively. Therefore, the DSRCNN model can effectively process the image, reduce image distortion, and improve its resolution, enhancing image clarity.

This experiment basically achieved the purpose of the experiment. In the future, we can consider putting forward effective solutions to the visual shape differentiation in the DSRCNN model.

REFERENCES

[1] Y. Chen, W. Chen, S. Chandra Pal, A. Saha, I. Chowdhuri, B. Adeil, S. Janizadeh, A. Dineva, X. Wang, A. Mosavi, "Evaluation efficiency of hybrid deep learning algorithms with neural network decision tree and boosting methods for predicting groundwater potential," Geocarto International, vol. 37, no. 19, pp. 5564-5584, 2022.

[2] T. S. Kumar, T. Senthil, "Construction of hybrid deep learning model for predicting children behavior based on their emotional reaction," Journal of Information Technology, vol. 3, no. 1, pp. 29-43, 2021.

[3] D. A. Neu, J. Lahann J, P. Fettke, "A systematic literature review on state-of-the-art deep learning methods for process prediction," Artificial Intelligence Review, vol. 55, no. 2, pp. 801-827, 2022.

[4] M. T. Rasheed, D. Shi, "LSR: Lightening super-resolution deep network for low-light image enhancement," Neurocomputing, vol. 505, pp. 263-275, 2022.

[5] X. Y. Kong, L. Liu, Y. S. Qian, "Low-light image enhancement via poisson noise aware retinex model," IEEE Signal Processing Letters, vol. 28, pp. 1540-1544, 2021.

[6] S. Wang, D. M. Yang, R. Rong, X. Zhan, G. Xiao, "Pathology image analysis using segmentation deep learning algorithms," The American journal of pathology, vol. 189, no. 9, pp. 1686-1698, 2019.

[7] F. Huang, J. Zhang, C. Zhou, Y. Wang, J. Huang, L. Zhu, "A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction," Landslides, vol. 17, no. 1, pp. 217-229, 2020.

[8] N. Sharma, R. Sharma, N. Jindal, "Machine learning and deep learning applications-a vision," Global Transitions Proceedings, vol. 2, no. 1, pp. 24-28, 2021.

[9] W. Zhang, H. Li, Y. Li, H. Liu, Y. Chen, X. Ding, "Application of deep learning algorithms in geotechnical engineering: a short critical review," Artificial Intelligence Review, vol. 54, no. 8, pp. 5633-5673, 2021.

[10] O. Aslan, A. A. Yilmaz, "A new malware classification framework based on deep learning algorithms," Ieee Access, vol. 9, pp. 87936-87951, 2021.

[11] O. Bamisile, A. Oluwasanmi, C. Ejiyi, N. Yimen, S. Obiora, Q. Huang, "Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions," International Journal of Energy Research, vol. 46, no. 8, pp. 10052-10073, 2022.

[12] W. Zhang, P. Zhuang, H. Sun, G. Li, S. Kwong, C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," IEEE Transactions on Image Processing, vol. 31, pp. 3997-4010, 2022.

[13] C Ike, N Muhammad. Separable property-based super-resolution of lousy image data. Pattern Analysis and Applications, 2020, 23: 1407-1420.

[14] T. Zhang, T. Hou, S. Weng, F. Zou, H. Zhang, C. Chang, "Adaptive reversible data hiding with contrast enhancement based on multi-histogram modification," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 8, pp. 5041-5054, 2022.

[15] K. Liu, X. Li, "De-hazing and enhancement method for underwater and low-light images," Multimedia Tools and Applications, vol. 80, no. 13, pp. 19421-19439, 2021.

[16] X. Sui, Z. He, G. Jin, D. Chu, L. Cao, "Band-limited double-phase method for enhancing image sharpness in complex modulated computer-generated holograms," Optics Express vol. 29, no. 2, pp. 2597-2612, 2021.

[17] R. Liu, Z. Jiang, S. Yang, X. Fan, "Twin adversarial contrastive learning for underwater image enhancement and beyond," IEEE Transactions on Image Processing, vol. 31, pp. 4922-4936, 2022.

[18] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, B. Xu, "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)," European radiology, vol. 31, no. 8, pp. 6096-6104, 2021.

[19] N. Sharma, R. Sharma, N. Jindal, "Machine learning and deep learning applications-a vision," Global Transitions Proceedings, vol. 2, no. 1, pp. 24-28, 2021.

[20] Y. R. Shrestha, V. Krishna, G. von Krogh, "Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges," Journal of Business Research, vol. 123, pp. 588-603, 2021.

[21] A. P. Pandian, "Performance evaluation and comparison using deep learning techniques in sentiment analysis," Journal of Soft Computing Paradigm (JSCP), vol. 3, no. 2, pp. 123-134, 2021.

[22] Y. Y. Zhao, Q. J. Huang, "Image enhancement of robot welding seam based on wavelet transform and contrast guidance," Int. J. Innov. Comput. Inf. Control, vol. 18, pp. 149-159, 2022.

[23] Z. Zhao, B. Xiong, L. Wang, Q. Ou, L. Yu, F. Kuang, "Retinexdip: A unified deep framework for low-light image enhancement," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1076-1088, 2021.

[24] F. Rizani, "Image quality improvement using image processing method image brightness contrast and image sharpening," Multica Science and Technology (mst) Journal, vol. 1, no. 1, pp. 6-12, 2021.

[25] K. Zhang, K. Fang, W. Chen, F. Sun, Y. Song, L. Han, "Controlling the spreading of nanoliter-scale droplets on the fibers of fabrics for enhancing image quality and ink utilization," ACS Applied Materials & Interfaces, vol. 13, no. 50, pp. 60581-60589, 2021.