



## TARGET IMAGE PROCESSING BASED ON SUPER-RESOLUTION RECONSTRUCTION AND DEEP MACHINE LEARNING ALGORITHM

YANG LIN<sup>\*</sup>, PING ZHANG<sup>†</sup>, HE ZHANG<sup>‡</sup> AND GUOPING SONG<sup>§</sup>

**Abstract.** In dictionary-based single-frame image reconstruction algorithms, dictionaries rely on the design of artificial shallow features and are limited in their ability to represent image features. Therefore, this paper proposes a high-accuracy reconstruction method based on deep learning feature dictionary. This algorithm first uses a deep network to learn high-resolution and low-resolution training example images with deep features; Then co-train the feature dictionary under the super dense framework of the sparse dictionary; Finally, a single low-resolution image can be input and a super-resolution reconstruction can be performed using a dictionary. From the theoretical analysis, the introduction of deep network to extract the deep-level features of the image and its use in dictionary training is more beneficial to complement the high-frequency information in the low-resolution image. Experiments show that the proposed method achieves the best results in terms of both the peak signal-to-noise ratio and the gradient energy function of the reconstructed images. This shows that compared with traditional interpolation methods and some deep learning methods, the proposed method can recover image details to a high degree while preserving the original image damage information. This proves that the subjective visual and objective evaluation indicators of the algorithm presented in this article are higher than those of the comparative algorithm.

**Key words:** Deep learning, Dictionary learning, Super resolution, Deep level feature extraction, Single frame image

**1. Introduction.** In the era of the Internet, information transmission is essential. The transmission methods include sound, video, images, text, etc. The importance of images in information transmission is self-evident. When computers process information, the source of images is generally composed of photographic equipment based on spatial sampling and quantized values [8]. The amount and credibility of information displayed in an image are closely related to its resolution. Image size refers to the number of pixels in an image, while resolution refers to the pixel density per unit. At the same time, the semantic information expression is more complete. The above information is a key element in a series of operations such as image analysis, processing, and classification. In the field of computer vision, image analysis and processing occupy an important position. Image classification, facial recognition, autonomous driving, etc all belong to the category of image processing, with the difference being that single frame and multi frame images have different processing speed requirements. But when doing image processing, a clear image is very important, and the image super-resolution here can be used as image preprocessing and other work to reconstruct clarity.

Image super-resolution reconstruction is the process of processing low resolution images through certain algorithms to obtain high-resolution or ideal resolution images. High resolution images, due to their higher pixel density, can display more detailed information and present finer image quality. Using more powerful photography equipment can help obtain higher resolution images, but as a result, costs cannot be effectively controlled. So the super-resolution algorithm plays a huge role in saving costs and obtaining high-resolution images [10].

At the same time, using high-resolution images for network transmission can cause problems such as slow transmission speed, network congestion, and even transmission failure. In this way, both the smoothness of task completion and user satisfaction cannot be solved. So we can use smaller low resolution images during transmission and restore them using super-resolution algorithms on the client side, which can greatly alleviate network pressure and compression costs. In addition, many industries use super-resolution algorithms for end-to-end image processing, and industries that require image information such as satellite and drone photography,

---

<sup>\*</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China (Corresponding Author)

<sup>†</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

<sup>‡</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

<sup>§</sup>Jilin Distance Education Technology Innovation Center, Changchun, Jilin, 130022, China

driving information collection, and medical examinations have high popularity of super-resolution algorithms [19]. For example, it is necessary to magnify confidential photos taken by spy satellites, distinguish distant object categories from front and rear car cameras, capture clear license plate numbers from intersection cameras, and capture details of distant scenery by photographers, all of which require image super-resolution technology. In addition, it can also help us deal with a series of problems such as noise interference and motion blur when using images collected through low-end equipment. From this, it can be seen that image super-resolution reconstruction has a wide range of applications and high value. While it brings help to various industries, it also has improvement and research value [1].

**2. Reference.** Super-resolution imaging technology, which aims to create a high-resolution image from a single low-resolution image or multiple low-resolution images, has become a useful tool for improving the accuracy of nuclear resonance imaging. Traditional image solvers often use translation methods such as bilinear translation, bicubic translation, and nearest neighbor translation. However, because the adjacent pixel values in the image are not constant, the image returned by the definition process is blurred [12]. Ultra-high-resolution multi-images can extract information from low-resolution multi-images to create high-resolution images, but the reconstruction process requires complex registration work, high time and space, and is not easy, which is not seen in practical applications.

Deep learning is one of the many learning machines that have made significant progress in ultra-high-resolution image processing, playing an important role in tasks such as object tracking, satellite imagery, and medical imaging. Deep learning involves training multiple image models, building a learning neural network, extracting correlations between high-resolution images and low-resolution images, and using this correlation graph to obtain the frequency information needed for updated images [14].

The main purpose of super-resolution images is to recover high-quality or high-resolution images from bad samples of low-quality or low-resolution images. In this research project, Muhammad, W proposed a new concept to reduce the number of redundant networks and improve operational speed, inspired by ResNet and Xception networks [13]. Fukami K et al. proposed a new method for data generation inspired by super-resolution and intermediate techniques using machine learning techniques to control recovery through spatio-temporal data flow. For the current machine learning based on updated data, we use neural network-based descending cross-connections/multiple models to integrate water mixture properties into their network structure [2]. The design of remote sensing surveys requires the use of special satellite images. However, these images may not be available at any time. Therefore, the modification method requires remote viewing of the image at that time. In this paper, Isa, SM proposed a machine learning model to transform Landsat-8 images into Sentinel-2 images. The inspiration for this model comes from the development of super-problem models based on deep learning. Studies have shown that the proposed model can predict Sentinel-2 images that are quantitatively and qualitatively similar to the original images [6].

In addition, multiple research groups have used different deep learning neural networks to perform three-dimensional virtual reconstruction of fluorescence microscope images to accurately locate the axial position of single molecules, or to better explore various life processes of cells through color separation processing of polychromatic images.

### 3. Methods.

#### 3.1. Related theories.

**3.1.1. Principle and limitations of ScSR algorithm.** At present, the Sparsecoding superresolution (ScSR) algorithm is the most classic among dictionary building and learning based superresolution algorithms. This algorithm uses a natural image library as the training dataset to train dictionary pairs  $D_h$  (High Resolution Dictionary) and  $D_l$  (Low Resolution Dictionary) that match high-resolution (HR) and low-resolution (LR) images; Then, for each input low resolution block  $y$ , find its sparse representation coefficient  $\tilde{a}$  in  $D_l$ , and the corresponding high resolution feature blocks based on  $D_h$ . will be combined through these coefficients [9]; Finally, output high-resolution feature block  $x$ . The reconstruction principle of this algorithm can be expressed as

$$\tilde{a} = \min_a \|\tilde{D}a - \tilde{y}\|_2^2 + \lambda \|a\|_1 \quad (3.1)$$

In the formula:

$$\tilde{D} = \begin{bmatrix} FD_1 \\ \beta PD_h \end{bmatrix}, \tilde{y} = \begin{bmatrix} Fy \\ \beta\omega \end{bmatrix},$$

$D$  represents an overcomplete dictionary,  $F$  represents a feature extraction operator. Using coefficient  $\tilde{a}$  and dictionary  $D_h$ , the HR image block is obtained, which can be represented as

$$x = D_h \hat{a} \quad (3.2)$$

By combining the HR image blocks of various buildings, the HR image can be calculated. However, the accuracy of the predicted HR image is related to the accuracy of the dictionary atoms involved in the superresolution, and the key is whether the features of the extracted image describe the image pattern. In the ScSR algorithm, high-sensitivity features sensitive to the human eye are selected for image feature extraction and representation, and high-sensitivity features are extracted in the first step and the second step. The gradient operator is

$$\begin{cases} f_1 = [-1, 0, 1] \\ f_2 = f_1' \\ f_3 = [-1, 0, -2, 0, 1] \\ f_4 = f_3' \end{cases} \quad (3.3)$$

In the formula,  $f_1$  and  $f_3$  extract the horizontal gradient features  $f_2$  and  $f_4$  of the image, and extract the vertical gradient features of the image. The four gradient features are ordered as a vector based on the block features of the LR image. During the reconstruction of the over-resolved problem, the ScSR algorithm selects large gradient feature values as candidate dictionary atoms [4].

However, for some texture images with complex data sets, the change in texture gradient value is smaller than the edge pattern gradient value. If the atoms of the candidate dictionary are selected according to the above properties, the resulting dictionary will be focused on the edge structure and thus the structure of the building image.new and rare super-resolution images [5]. In addition, when there is a difference in image type between the super-resolution reconstruction image and the training set image, the candidate dictionary atoms are selected according to structure or the edge gradient value during super-resolution reconstruction is inappropriate phase, thus affecting the image quality of the super-resolution reconstruction. Therefore, the ScSR supersolving algorithm, which constructs a dictionary according to pseudo-rules, is based on a shallow dictionary, and the dictionary has some global constraints. Deep training of image models is a good way to improve the quality of reconstructed images to create a deep dictionary and improve image feature representation.

**3.1.2. Deep learning feature dictionary.** Features are an important factor in machine learning algorithms and have a decisive impact on the final model. If the training sample images have good feature expression ability, usually the computational model can obtain satisfactory results. In recent years, with the increasing complexity of target images, people have begun to pay attention to how to construct more efficient feature dictionaries to achieve more accurate descriptions of image edge information, rich textures, and geometric structures [7]. Deep networks can mine deep knowledge of data, extract deeper features of training sample images using deep networks, break through the limitations of artificial rule features, and improve the expression ability and adaptability of dictionaries.

Therefore, this paper combines the advantages of PCANet deep network and ScSR distributed dictionary algorithms to propose a high-resolution image restoration algorithm based on PCANet dictionaries. The image blocking algorithm is shown in Figure 3.1. In the training phase, it is estimated that the features of low-resolution and high-resolution images have similar expressions in their book translation. Then, a PCANet deep network is used to obtain the deep level features of the training sample images, and a pair of overcomplete feature word dictionaries  $D_h$  and  $D_l$  are obtained through dictionary joint training. Among them,  $D_h$  represents a high-resolution image feature dictionary, and  $D_l$  represents a low-resolution image feature dictionary [20].

In the super-resolution reconstruction stage, the PCANet method is also used to extract deep level features from low resolution images, solve for the sparse representation of each low resolution feature block in the

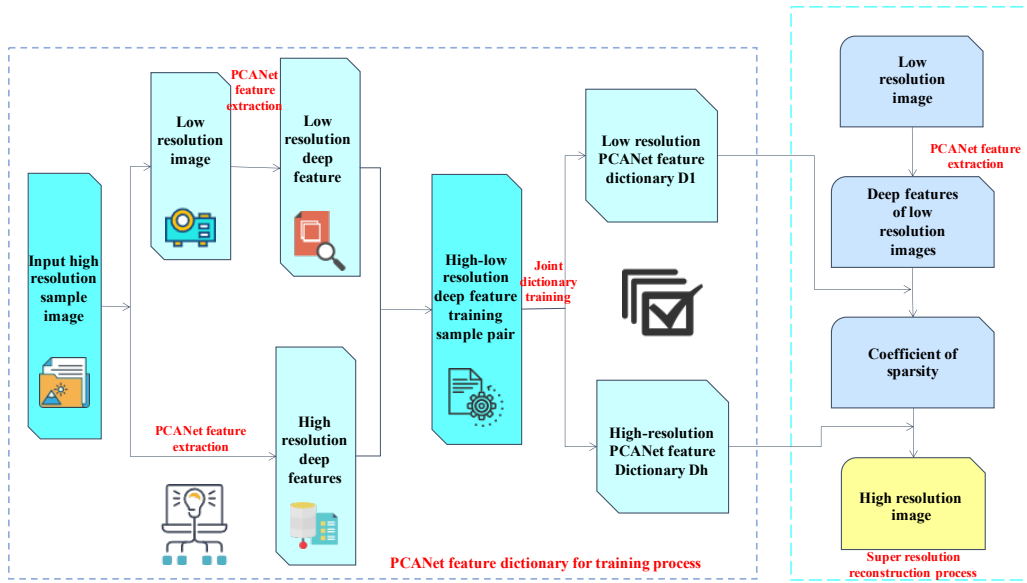


Fig. 3.1: Overview of single image superresolution algorithm based on PCANet feature dictionary

dictionary  $D_l$  of the image to be super-resolved, and directly apply the sparse representation coefficient of the low resolution image feature block on  $D_l$  to  $D_h$ , thereby obtaining the corresponding high-resolution image feature blocks, and ultimately achieving super-resolution reconstruction of the low resolution image. This algorithm uses PCANet deep network mining to train sample image features, which can obtain richer feature information than non deep networks. The deep level feature dictionary constructed on this basis also improves the expression ability of the feature dictionary, achieving effective improvement in the quality of reconstructed images [17].

### 3.2. Algorithm in this article.

**3.2.1. Image preprocessing.** Firstly, blur down sample  $K$  high-resolution images in the training data, and then enlarge them to the same size as the high-resolution image to obtain a low resolution image corresponding to the high-resolution image, forming a training sample pair:  $T = \{X_h, X_l\}$ , where  $X_h = \{x_h^i\}_{i=1}^K$  is the high-resolution feature and  $X_l = \{x_l^i\}_{i=1}^K$  is the low resolution feature. Then, calculate the blocking matrix for each training sample. Taking high-resolution training samples as an example, select a  $k_1 \times k_2$  sliding window (usually a square window with a side length of 3, 5 or 7 pixels), and each  $E * E$  sized image will be transformed into  $m \times n \quad k_1 \times k_2$  sized image blocks after extracting local features through the sliding window; Then average these  $m \times n$  image blocks and complete the feature extraction operation for a single image. Perform the above operation on all  $N$  high-resolution images to obtain a new data matrix  $X$  with  $N \times m \times n$  columns. Each column of the matrix represents an image block with a total of  $k_1 \times k_2$  elements [11].

The  $i$ -th high-resolution training sample obtained can be expressed as equation (3.4), the block matrix of the overall high-resolution sample can be expressed as equation (3.5), and the block matrix calculation of the low resolution sample is the same as that of the high-resolution sample, expressed as equation (3.6).

$$\bar{X}_{hi} = [\bar{x}_{hi,1}, \bar{x}_{hi,2}, \dots, \bar{x}_{hi,mn}] \tag{3.4}$$

$$X_h = [\bar{X}_{h1}, \bar{X}_{h2}, \dots, \bar{X}_{hK}] \in R^{k_1 k_2 \times Kmn} \tag{3.5}$$

$$X_l = [\bar{X}_{l1}, \bar{X}_{l2}, \dots, \bar{X}_{lK}] \in R^{k_1 k_2 \times Kmn} \tag{3.6}$$

**3.2.2. Feature extraction.** Based on the above training sample matrix, PCANet feature extraction is performed, and this feature is used as the training sample feature in the ScSR model for PCANet feature

dictionary training. The PCANet deep learning network completes image feature extraction in three layers. Taking high-resolution training images as an example, a detailed feature extraction process is provided below.

(1) *1st layer.* The first layer feature extraction process is the process of constructing a PCA filter and performing convolution, as shown in the first rectangular box (Firststage) in Figure 3.1. Assuming that the number of filters required in the  $i$ -th layer is  $L_1$ , a series of standard orthogonal matrices are searched for to minimize the reconstruction error through equation (3.7). By extracting the eigenvectors corresponding to the first  $L_1$  maximum eigenvalues of the covariance matrix  $X_h$ , a feature mapping matrix is formed, which is the PCA filter, as shown in equation (3.8).

$$\min_{V \in R^{k_1 k_2 \times L_1}} \|X_h - VV^T X_h\|_F^2 \quad S.T. V^T V = I_{L_1} \quad (3.7)$$

$$W_{hL}^1 = \text{mat}_{k_1 k_2} (q_{hL} X_h X_h^T) \in R^{k_1 k_2} \quad L = 1, 2, \dots, L_1 \quad (3.8)$$

Rearrange each column of the  $L_1$  feature vectors to obtain 1 patch, which results in an  $L_1$  of  $k_1 \times k_2$  window. Then, for each image, perform a convolution using these  $L_1$  windows [18]. The main information of the training sample image can be retained through the first layer PCA filter, represented as

$$I_{hi}^L = I_{hi}^L \times W_{hL}^1 \quad i = 1, 2, \dots, K \quad (3.9)$$

(2) *2nd layer.* The PCA mapping process of the second layer is shown in the second rectangular box (Secondstage) in Figure 3.1. Similar to the blocking operation of the sample image in the first layer, the output result of the PCA mapping in the first layer is used as the input of the second layer. Similarly, the matrix is subjected to block sampling, cascading, and zero averaging operations in the second layer, and the results are represented as follows

$$Y_h^L = [\bar{Y}_{h1}^L, \bar{Y}_{h2}^L, \dots, \bar{Y}_{hK}^L] \in R^{k_1 k_2 \times K_{\bar{m}\bar{n}}} \quad (3.10)$$

$$Y_h = [Y_h^1, Y_h^2, \dots, Y_h^{L_1}] \in R^{k_1 k_2 \times L_1_{\bar{m}\bar{n}}} \quad (3.11)$$

Similarly, the PCA filter consists of eigenvectors corresponding to the covariance matrix, and the filter is

$$W_{hL'}^2 = \text{mat}_{k_1 k_2} (q_{hL'} (Y_h Y_h^T)) \in R^{k_1 k_2} \quad L' = 1, 2, \dots, L_2 \quad (3.12)$$

Due to the presence of  $L_1$  filter kernels in the first layer, the second layer performs the same step of feature extraction for each feature output from the previous layer, resulting in  $L_1$  feature outputs. Finally, for each sample image, PCANet will output  $L_1 \times L_2$  feature matrices, as shown in equation (3.13). In terms of structure, the feature extraction process of the two-layer PCA is very similar, and PCANet can also be expanded into a deep network structure containing more layers as needed [3].

$$O_{hi}^{L'} = \{I_{hi}^L \times W_{hL'}^2\}_{L'=1}^{L_2} \quad (3.13)$$

(3) *Output layer.* The output layer mainly targets each output matrix in the second layer, which is binarized to only contain integers 1 and 0. Then, the matrix is binarized and hashed, as shown in the formula:

$$T_{hi}^L = \sum_{L'=1}^{L_2} 2^{L'-1} H(O_{hi}^{L'}) = \sum_{L'=1}^{L_2} 2^{L'-1} H(I_{hi}^L \times W_{hL'}^2) \quad L = 1, 2, \dots, L_1 \quad (3.14)$$

In the formula,  $T_{hi}^L$  is the result of hash encoding of high-resolution training image features,  $2^{L'-1}$  is the transformation coefficient that causes each pixel value in the image to become a numerical value between 0 and 255. The function  $H(\cdot)$  is similar to a unit step function, which quantifies and increases the difference between each feature. Then, the above results are histogram encoded to complete PCANet feature extraction of a high-resolution sample image, and the extracted results can be represented as

$$F_{hi} = [\text{Bhist}(T_{hi}^1), \dots, \text{Bhist}(T_{hi}^{L_1})]^T \in R^{(2^{L_1})_{L_1 B}} \quad (3.15)$$

Similarly, it can be inferred that after the same steps as high-resolution training images, the PCANet feature  $F_{li}$  of low-resolution training images can be represented as

$$F_{li} = \left[ \text{Bhist} (T_{li}^1), \dots, \text{Bhist} (T_{li}^{L_1}) \right]^T \in R^{(2^{L_1})L_1 B} \quad (3.16)$$

In the formula,  $F_{hi}$  and  $F_{li}$  represent the feature extraction results of the sample image; Bhist represents the histogram encoding process;  $B$  represents the number of segmented image blocks in the sample image [16].

**3.2.3. Dictionary training.** This article adopts a joint sparse encoding approach to train dictionaries under the ScSR framework. The goal of the algorithm is to obtain a pair of dictionary pairs  $D_h$  and  $D_l$  that can describe the complex features of the sample, so that the PSAllet features  $F_{hi}$  and  $F_{li}$  generated by  $K$  on the image have the same sparse representation on  $D_h$  and  $D_l$ , and  $F_{hi}$  and  $F_{li}$  have the same representation coefficients, i.e

$$\{D_h, a\} = \underset{D_h, a}{\operatorname{argmin}} \|F_{hi} - D_h \cdot a\|_2^2 + \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.17)$$

$$\{D_l, a\} = \underset{D_l, a}{\operatorname{argmin}} \|F_{li} - D_l \cdot a\|_2^2 + \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.18)$$

In order to ensure that high-resolution image features and low-resolution image features have the same sparse expression form regarding their respective dictionaries, a joint training strategy is adopted for equations (3.17) and (3.18), namely

$$\{D_l, D_h, a\} \underset{D_l, D_h, a}{\operatorname{argmin}} \|F_{hi} - D_h \cdot a\|_2^2 + \frac{1}{M} \|F_{li} - D_l \cdot a\|_2^2 + \left( \frac{1}{N} + \frac{1}{M} \right) \sum_{i=1}^K \lambda_i \|a_i\|_1 \quad (3.19)$$

In the equation,  $N$  and  $M$  represent the dimensionality of rearranging the element values of high and low resolution feature image blocks into column vectors, respectively.  $1/N$  and  $1/M$  are used to balance the costs of  $D_h$  and  $D_l$  in equations (3.17) and (3.18). For convenience in solving, equation (3.19) is written as

$$\{D_c, a\} = \underset{D_c, a}{\operatorname{argmin}} \|F_c - D_c \cdot a\|_2^2 + \sum_{i=1}^K \lambda'_i \|a_i\|_1 \quad (3.20)$$

$$D_c = \begin{bmatrix} \frac{1}{\sqrt{N}} D_h \\ \frac{1}{\sqrt{M}} D_l \end{bmatrix}, X_c = \begin{bmatrix} \frac{1}{\sqrt{N}} F_h \\ \frac{1}{\sqrt{M}} F_l \end{bmatrix}, \lambda'_i = \left( \frac{1}{N} + \frac{1}{M} \right) \lambda_i \quad (3.21)$$

Equation (3.21) is solved using an iterative method. Firstly, by giving a dictionary  $D_c$ , solve for each pair of training sample data  $F_{ci}$ ; The sparse representation coefficient  $a_i$  on  $D_c$  is obtained, resulting in a sparse representation matrix  $\alpha = \{a_i\}_{i=1}^K$ . Finally, the dictionary  $D_c$  is updated based on  $\alpha$ .

**3.2.4. Image reconstruction.** After obtaining the high resolution and low resolution image feature dictionaries for  $D_h$  and  $D_l$ , it is necessary to solve the classical optimization problem for each low resolution test image

$$\hat{a} = \min_a \|\tilde{D}a - \tilde{y}\|_2^2 + \lambda \|a\|_1 \quad (3.22)$$

In the formula:

$$\tilde{D} = \begin{bmatrix} FD_1 \\ \beta PD_h \end{bmatrix}, \tilde{y} = \begin{bmatrix} Fy \\ \beta \omega \end{bmatrix},$$

$F$  is the PCANet feature,  $P$  is the overlap area between the current high-resolution image block and the reconstructed high-resolution image block. This strategy can reduce blocky effects,  $\omega$  is the number of pixels

in the overlap area, and  $\beta$  measures the matching between the input low-resolution feature block and the high-resolution overlap area. Solve equation (3.22) to obtain the coefficient expression coefficient  $\alpha$  for each low resolution image, and apply it to  $D_h$  to obtain the initial high-resolution feature information image  $X_0 = D_h\alpha$ . Solve the following optimization problem for  $X_0$  to ensure that the final high-resolution image meets the reconstruction constraints, namely

$$X^* = \min_X \|HX - Y\|_2^2 + \lambda \|X - X_0\|_2^2 \quad (3.23)$$

In the equation,  $H$  is the image degradation operator ( $H$  is related to the imaging process, and in this experiment, this operator only represents downsampling of the test image). By using the gradient descent method to solve equation (3.23), a super-resolution result image can be obtained.

**3.2.5. Algorithm steps.** The algorithm in this article first extracts PCAdet features from low resolution images, and trains dictionaries based on these deep level features to achieve super-resolution reconstruction. The algorithm is shown below.

---

**Algorithm 1** Single frame image super-resolution reconstruction based on PCAdet feature dictionary

---

**Input:** Low resolution image  $Y$ , high and low resolution dictionary pairs  $D_h, D_l$

**Output:** High-resolution image  $X^*$

- 1: Perform  $2x$  unsampling interpolation on low resolution images with  $Y' = S(Y)$ , where  $S$  is the unsampling operator
  - 2: Extracting PCANet Features  $F_1$  from Low Resolution Images
  - 3: Establish the relationship between dictionary  $D_h, D_l$ , and  $F_1$  according to equation (3.22)
  - 4: **for**  $i = 1$  to  $T$  **do**
  - 5:     Calculate  $\hat{a} = \min_a \|\tilde{D}1a - \tilde{y}\|_2^2 + \lambda \|a\|_1$  to obtain the sparsity coefficient  $\hat{a}$  of the low resolution PCANet feature with respect to the low resolution dictionary  $D_l$
  - 6:     Based on the sparsity coefficient  $\hat{a}$  and high-resolution dictionary  $D_h$ , estimate the feature information of high-resolution images as  $X_0 = D_h\alpha$
  - 7:     Synthesize the obtained high-resolution image feature information with the low-resolution image to obtain the initial super-resolution result image
  - 8: **end for**
  - 9: Using the gradient descent method, calculate  $X^* = \min_X \|HX - Y\|_2^2 + \lambda \|X - X_0\|_2^2$ . Obtain the super-resolution result image  $X^*$  that is closest to the initial high-resolution image
- 

## 4. Experimental Results and Analysis.

**4.1. Experimental dataset.** At present, there are few publicly available breast cancer MRI image data sets, so this experimental group has built a breast cancer MRI image data set. This dataset is provided by Liaoning Cancer Hospital and includes a total of 260 cases, including 17 male cases and 243 female cases, covering the age range of 32-78 years. In the experiment, CE3 images of patients were used, with a total of 48 slices per case. All black and poor quality slice images were removed from the slices, resulting in a self built dataset of 1310 slices. The obtained image is downsampled using the bicubic downsampling method at a rate of 4 times, resulting in a low resolution image.

**4.2. Training details and parameters.** Before training, the images were standardized to accelerate the convergence speed of the network. During training, a low resolution image with a size of  $128 \times 128 \times 3$  was input, and the number of training sessions per batch was set to 4. The optimizer selected Adam and set the learning rate to  $1 \times 10^{-4}$ . The loss function was selected as Mean Absolute Error (MAE). The training process was conducted 1200 iterations, lasting approximately 34 hours. All experiments were conducted using the Keras framework on NVIDIA RTX2080ti [15].

Table 4.1: Image subjective evaluation form

Grade	Absolute measurement scale	Detail	Score
1	Excellent	The best in the group	10
2	Good	Better than the average in the group	8
3	Average	Group average	6
4	Fair	Worse than the average in the group	4
5	Poor	Worse in the group	2

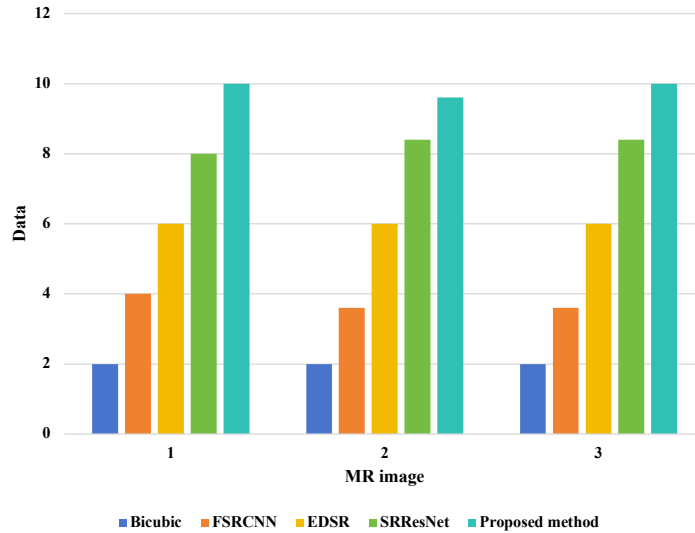


Fig. 4.1: Comparison of subjective evaluation values of super-resolution reconstruction methods

**4.3. Experimental results and discussion.** After the training is completed, high-resolution images are downsampled 4 times to obtain low resolution images as test images. The low resolution images are then input into FSRCNN, EDSR, SRResNet, and the proposed network for testing. In addition to deep learning methods, the images are also subjected to bicubic interpolation. In order to qualitatively analyze the quality of image super-resolution, image quality was evaluated from both subjective and objective perspectives. Subjectively, the reconstructed image quality under different methods is evaluated based on the subjective evaluation table of the image. This evaluation method divides the image into 5 levels based on its relative and absolute quality, and the relative and absolute measurement scales of each level are shown in Table 4.1. According to this evaluation indicator, the volunteers rated the reconstructed images, and the results are shown in Figure 4.1. From Figure 4.1, it can be seen that compared to the images recovered by other methods, the reconstructed images by the proposed method are more in line with human subjective visual perception, and the restoration effect of image details and textures is better than other methods.

Objectively, PSNR and gradient energy values were selected as the main evaluation indicators. Figures 4.2 and 4.3, respectively represent the PSNR and gradient energy values of each image under different reconstruction methods. From Figures 4.2 and 4.3, it can be seen that the proposed model achieves the best results in peak-to-noise ratio and gradient power function of the reconstructed image. It shows that compared with traditional annotation methods and some deep learning methods, the proposed model can recover the maximum amount of image details while preserving the context information.

Figure 4.4a-c shows the comparison of the super-resolution reconstruction results PSNR, SSIM, and computational cost (Time) for each experiment. From the PSNR numerical results, the interpolation method has



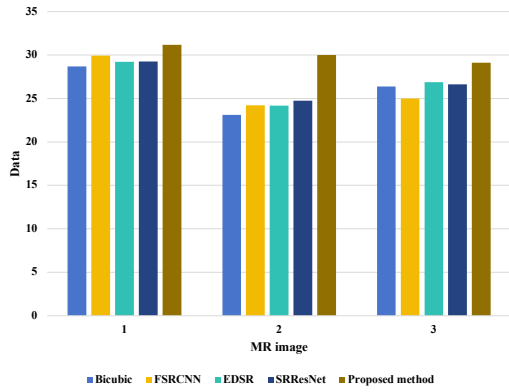


Fig. 4.2: Comparison of PSNR values of various super-resolution reconstruction

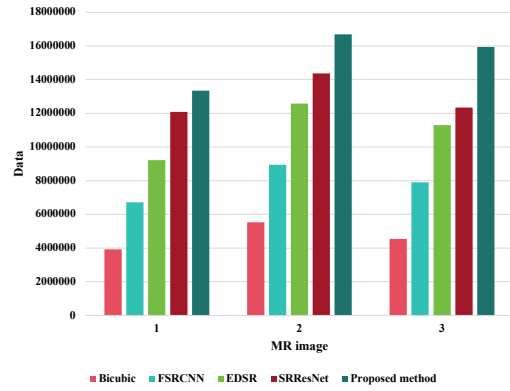
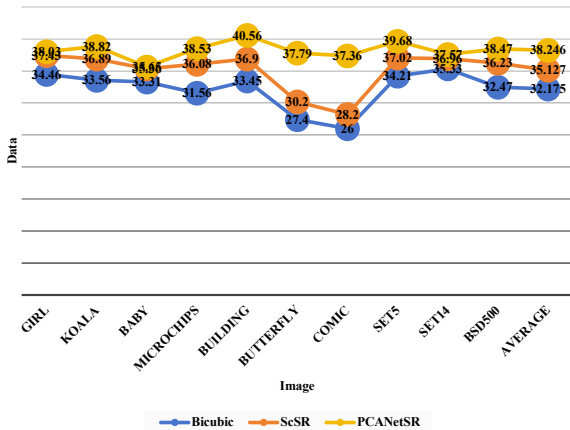
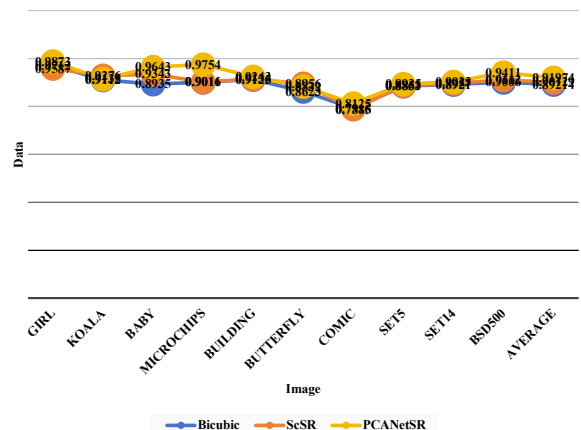


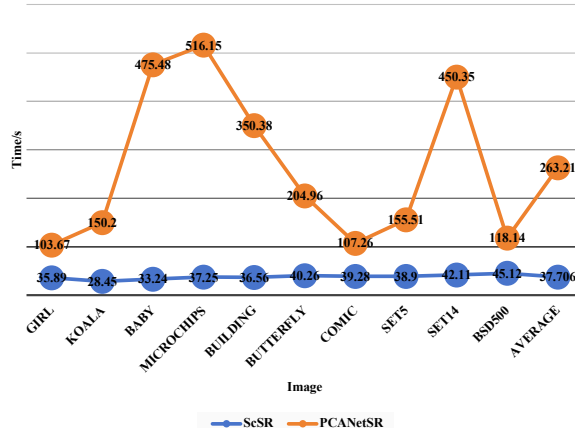
Fig. 4.3: Comparison of energy gradient values of various super-resolution reconstruction



(a) PSNR



(b) SSIM



(c) Time/s

Fig. 4.4: Objective evaluation index of the reconstructed images by different methods

a mean PSNR of 32.17, the ScSR method has a mean PSNR of 35.12, and the algorithm in this paper has a mean PSNR of 38.24, which is the highest in experimental testing. From the time numerical results of ScSR and our algorithm, it can be seen that the mean time of ScSR method is 37.70, while the mean time of our algorithm is 263.21. Compared to this method, our computational cost is slightly higher. Through analysis, it can be seen that due to the stronger ability of the algorithm in extracting deep level features of images compared to general artificial rule features to capture image details, the scale of features involved in dictionary learning and reconstruction is larger, so the algorithm has a slightly longer running time than ScSR. However, from the above objective evaluation indicators PSNR, SSIM, and subjective visual effects, it can be seen that the obtained super-resolution image quality has significantly improved. Therefore, the algorithm in this paper still has certain advantages in overall efficiency.

**5. Conclusion.** In this paper, we present a new algorithm for super-resolution reconstruction of single-frame images based on deep learning feature dictionary. In this project, we propose a novel method based on depth feature, which uses high-resolution dictionary and low-resolution dictionary to improve the performance of texture and complex structure. Experimental results show that the method is effective in both subjective and objective assessment. Objectively, the SNR and the gradient energy function are the best. From a subjective point of view, the proposed approach improves the quality of the image. Both the edge profile and the internal organs can be seen clearly, which is closer to the original high resolution image. It can be seen from the experiment that this method can make up for the disadvantages of long scan time and poor imaging quality in MRI equipment. Further work is needed to improve the training rate while maintaining super-resolution.

**Acknowledgement.** Scientific and Technologic Development Program supported by JILIN province (YDZJ202202CXJD038).

#### REFERENCES

- [1] F. AN AND J. WANG, *Image super-resolution reconstruction algorithm based on significant network connection-collaborative migration structure*, Digital Signal Processing, 127 (2022), p. 103566.
- [2] K. FUKAMI, K. FUKAGATA, AND K. TAIRA, *Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows*, Journal of Fluid Mechanics, 909 (2021), p. A9.
- [3] ———, *Super-resolution analysis via machine learning: a survey for fluid flows*, Theoretical and Computational Fluid Dynamics, (2023), pp. 1–24.
- [4] H. GAO AND K. OGAWARA, *Face reconstruction algorithm based on lightweight convolutional neural networks and channel-wise attention*, IEEEJ Transactions on Image Electronics and Visual Computing, 10 (2022), pp. 90–97.
- [5] S. GROSCHE, K. FISCHER, F. BRAND, J. SEILER, AND A. KAUP, *Enhanced image reconstruction from quarter sampling measurements using an adapted very deep super resolution network*, in Proceedings of the IEEE International Conference on Image Processing, Abu Dhabi, United Arab, 2020, IEEE, pp. 256–260.
- [6] S. M. ISA, SUHARJITO, G. P. KUSUMA, AND T. W. CENGGORO, *Supervised conversion from landsat-8 images to sentinel-2 images with deep learning*, European Journal of Remote Sensing, 54 (2021), pp. 182–208.
- [7] Y. V. KISTENEV, V. SKIBA, V. V. PRISCHEPA, D. A. VRAZHNOV, AND A. V. BORISOV, *Super-resolution reconstruction of noisy gas-mixture absorption spectra using deep learning*, Journal of Quantitative Spectroscopy and Radiative Transfer, 289 (2022), p. 108278.
- [8] I. KOKTZOGLU, R. HUANG, W. J. ANKENBRANDT, M. T. WALKER, AND R. R. EDELMAN, *Super-resolution head and neck mra using deep machine learning*, Magnetic Resonance in Medicine, 86 (2021), pp. 335–345.
- [9] C. KONG, J. CHANG, Z. WANG, Y. LI, AND W. BAO, *Data-driven super-resolution reconstruction of supersonic flow field by convolutional neural networks*, AIP Advances, 11 (2021).
- [10] Q. KUANG, *Single image super resolution reconstruction algorithm based on deep learning*, in Proceedings of the International Conference on Artificial Intelligence, Computer Networks and Communications, vol. 1852, Yunnan, China, 2021, IOP Publishing, p. 032039.
- [11] R. LIN, L. WANG, AND T. XIA, *Research on image super-resolution technology based on sparse constraint segnet network*, Journal of Physics: Conference Series, 1952 (2021), p. 022005.
- [12] B. LIU AND J. CHEN, *A super resolution algorithm based on attention mechanism and srgan network*, IEEE Access, 9 (2021), pp. 139138–139145.
- [13] W. MUHAMMAD, S. ARAMVITH, AND T. ONOYE, *Multi-scale xception based depthwise separable convolution for single image super-resolution*, Plos One, 16 (2021), p. e0249278.
- [14] K. NING, Z. SU, Z. ZHANG, AND G.-J. KIM, *An image reconstruction algorithm based on frequency domain for deep subcooling of melt drops*, Journal of Internet Technology, 22 (2021), pp. 1273–1285.
- [15] V. STUMPO, J. M. KERNBACH, C. H. VAN NIFTRIK, M. SEBÖK, J. FIERSTRA, L. REGLI, C. SERRA, AND V. E. STAARTJES, *Machine learning algorithms in neuroimaging: An overview*, Machine Learning in Clinical Neuroscience: Foundations and Applications, (2022), pp. 125–138.

- [16] S. F. SWEERE, I. VALTCHANOV, M. LIEU, A. VOJTEKOVA, E. VERDUGO, M. SANTOS-LLEO, F. PACAUD, A. BRIASSOULI, AND D. CÁMPORA PÉREZ, *Deep learning-based super-resolution and de-noising for xmm-newton images*, Monthly Notices of the Royal Astronomical Society, 517 (2022), pp. 4054–4069.
- [17] T. S. VENKATESH, R. SRIVASTAVA, P. BHATT, P. TYAGI, AND R. K. SINGH, *A comparative study of various deep learning techniques for spatio-temporal super-resolution reconstruction of forced isotropic turbulent flows*, in Proceedings of the ASME International Mechanical Engineering Congress and Exposition, vol. 10, Online, 2021, American Society of Mechanical Engineers, p. V010T10A061.
- [18] S. WANG, X. YAO, AND G. WANG, *Leakage characteristics of a sealing structure with damage holes based on super-resolution infrared thermography technology*, Measurement Science and Technology, 32 (2020), p. 025203.
- [19] B. XIE AND F. NIU, *Super-resolution reconstruction algorithm for aerial image data management based on deep learning*, Distributed and Parallel Databases, 40 (2022), pp. 699–716.
- [20] X. YANG, C. WU, D. ZHOU, AND T. LI, *Fast image super-resolution based on limit gradient embedding cascaded forest*, Circuits, Systems, and Signal Processing, (2022), pp. 1–20.

*Edited by:* B. Nagaraj M.E.

*Special issue on:* Deep Learning-Based Advanced Research Trends in Scalable Computing

*Received:* Oct 4, 2023

*Accepted:* Nov 21, 2023