Research on Super-resolution Reconstruction Optimization Method Based on Deep Learning

Mei Zhang [a)], Huan Liu [b)]

School of information, Guizhou University of Finance and Economics, Guiyang 550025, China

[a)] 409348604@qq.com, [b)] 942833674@qq.com

Abstract. In order to reconstruct low-resolution images into high-resolution images and compensate for the loss of information between high-resolution images and low-resolution images, this paper presents an effective super-resolution reconstruction optimization method, which analyzes the advantages and disadvantages of existing super-resolution constructing methods, and throws out problems that need to be resolved, and designs reasonable solutions.

Key words: image, super-resolution, reconstruction.

INTRODUCTION

Image Super-Resolution (SR) is an effective method to improve image quality in order to meet the demand of high-resolution images [1]. Its core is to construct lower resolution (LR) images to high resolution (HR) images. In recent years, super-resolution reconstruction technology based on machine learning has developed rapidly, breaking the deficiencies of super-resolution reconstruction results sensitive to amplification factors, especially the development of sparse coding theory has brought many inspirations to the SR method [2]. The application of deep learning model in the field of super-resolution reconstruction has brought further improvement to the reconstruction effect. However, the existing deep learning super-resolution reconstruction algorithm has a high complexity, and the parameters to be estimated are large in scale and usually initialized to random values. Usually, a large amount of iterative training is required to learn meaningful parameters. The amount of natural image data for SR training should not be too small, otherwise it will affect the generalization ability of the model. Therefore, the training process based on deep learning SR algorithm is inefficient, and it is difficult to achieve effective training on ordinary computers. In some applications that require super-resolution reconstruction of images in special areas, the model needs to be adjusted according to the visual characteristics of the image reconstruction and the actual needs. The expensive training cost limits the implementation of such improvement measures.

TRADITIONAL RESEARCH METHODS AND PROBLEMS

According to the different sources of training data, the SR methods based on machine learning can be divided into external sample-based methods and internal sample-based methods, and deep learning-based SR methods.

SR Methods Based on Internal Sample

The theoretical basis of the super-resolution reconstruction method based on the internal sample is that the texture of the image sample tends to reappear in the original image or the scaled version of the image, by searching in the self-image and other samples generated by it, usually You can find materials that are more consistent with the characteristics of the sample itself. Yang [2] uses the internal similarity of image samples to perform super-

resolution reconstruction based on self-images; The literature [3] uses cross-proportionally generated images to construct more training data; The literature [4] proposes the DNC model of the depth learning method, which uses a similar block search and weighted combination of internal sample block sequences for different amplification factors. However, this kind of reconstruction method is limited to the case where no repeatable samples can be found in the image itself or in the scaled image, and the generalization ability is limited.

SR Methods Based on External Sample

The super-resolution reconstruction method based on external samples usually uses a series of general sample elements to predict the high-resolution image loss information (high frequency components). The literature [5] adopts low-rank matrix recovery and joint learning methods to map the low-rank components of the original high-resolution images and low-resolution images block features into the unified space, and then complete reconstruction based on neighborhood embedding. The literature [6] is based on compressible sensing theory and jointly trains dual dictionaries for high-resolution images and low-resolution images. The literature[7], the high-resolution images and low-resolution images are represented by two feature spaces, which extends SR methods based on sparsely encoded to a more universal situation; The literature[8] allows the mapping relationship between high-resolution and low-resolution sparse representation coefficients; The literature[9] combines sparse dictionary learning with neighborhood embedding (NE) strategies and proposes a fixed Neighborhood regression method; The literature[10] allows deformation of image blocks, and sets sparse dictionary elements to contain only singular primitives, making learned sparse dictionaries more expressive. The sparse coding based super-resolution reconstruction method can usually obtain better visual effects, but the processing efficiency is lower in practical applications.

SR Methods Based on Deep Learning

With the continuous development of deep learning in the field of computer vision, super-resolution reconstruction methods based on deep learning have gained more attention in recent years. In literature [11], the depth-restricted Boltzmann machine (RBM) model is applied to the super-resolution reconstruction of images, and the sparse expression coefficients shared by high-resolution and low-resolution images are learned by the RBM model. The literature [4] uses joint local dynamic coding. The machine (CLA) model constrains the reconstruction within the layer, cascades all stacked CLA models globally, uses the inverse conduction algorithm to suppress the global range of errors, and trains the deep network; In literature [12], an end-to-end constructional neural network model was trained on natural image data sets to perform super-resolution reconstruction based on external examples, called the SRCNN model. This model achieved better results using only a 3-layer network structure. Reconstruct the effect, but the cost of training is relatively high. This dissertation starts from the premise of ensuring the reconstruction quality and improves the training efficiency of the model. A secondary super-resolution reconstruction method based on CNN and ELM is proposed.

SECONDARY RECONSTRUCTION METHOD BASED ON CNN AND ELM

The Brief Introduction of Extreme Learning Machine

ELM (Extreme Learning Machine) is a single hidden-layer feed-forward neural network characterized by the initialization of parameters between the input layer and the hidden layer to take random values. ELM has fast learning speed, simple application method, high training accuracy, and can handle two categories, multiple categories, and regression. In this paper, ELM is used for super-resolution reconstruction of images. Compared with other regression models, the ELM model has low complexity and can process more training data in an acceptable time, and the generalization ability and reconstruction effect are even greater. Outstanding. The structure of the ELM model is shown in Figure 1.
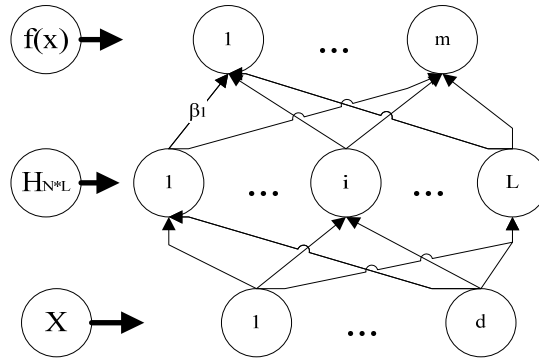
FIG. 1. Structure of the ELM model

ELM includes three levels: input layer x, hidden layer $H_{N \times L}$, and output layer $f(x)$. The d-input nodes correspond to the d-dimensional feature vector $(x \in R^d, x = (x_1, ..., x_d)^T)$ of the input layer. The feature vector x is mapped to a vector $(G(a_1, b_1, x), ..., G(a_L, b_L, x))^T$ in the hidden layer. $G(a_i, b_i, x)$ Is the output of the $i-th$ additive hidden node, which formula is as follows:

$$G(a_i, b_i, x) = g(a_i \cdot x + b_i), a_i \in R^d, b_i \in R \tag{1}$$

Where g(x) represents the activation function. The feature vector of the L dimension of the hidden layer can obtain an m-dimensional vector after linear transformation. The m-output nodes correspond to the m classes of the output layer. The formula is as follows:

$$f(x) = \sum_{i=1}^{L} \beta G(a_i, b_i, x), \beta_i \in R^m \tag{2}$$

The ELM training set is $\{(x_j, t_j) | x_j \in R^d, t_j \in R^m, j = 1, ..., N\}$, where x_j the feature vector is, t_j is the label of x_j. In the training phase, each instance x_j is passed to the ELM as a vector of input layers. t_j as the expected output of the output layer, there is an input instance x_j corresponding to each tag vector t_j. In the regression problem, t is a real number, and the value of t_j directly represents the model's response to input x_j The parameter values $(a_i, b_i, i = 1, ..., L)$ of the hidden layer nodes are initialized at random, and the variables $(\beta_i, ..., \beta_L)$ can be calculated by the following formula:

$$H\beta = T \tag{3}$$

Witch $H = \begin{bmatrix} G(a_1, b_1, x_1) & ... & G(a_L, b_L, x_1) \\ ... & ... & ... \\ G(a_1, b_1, x_N) & ... & G(a_L, b_L, x_N) \end{bmatrix} = [h(x_1), \quad ... \quad , h(x_N)]_{L \times N}^{T}$

$$h(x) = \begin{bmatrix} G(a_1, b_1, x) \\ ... \\ G(a_L, b_L, x) \end{bmatrix}_{L \times 1} \tag{4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ ... \\ \beta_L^T \end{bmatrix}_{L \times m} , \quad T = \begin{bmatrix} t_1^T \\ ... \\ t_L^T \end{bmatrix}_{N \times m}$$

The least squares solution of β^* can be solved using MP generalization and has a minimum norm:

$$\beta^* = HT \tag{5}$$

The final output equation of ELM can be obtained:

$$f(x) = \beta^{*T} h(x) \tag{6}$$

The ELM training and testing process in this article is shown in Figure 2.

In the deep learning network parameter training, a certain range of Gaussian random parameters is usually used as the initialization of the model, and the random initial value may cause the model to fall into a local optimal solution or fail to converge during training. Many deep learning model trainings include pre-training and tuning. The pre-trained model is equivalent to a supervised learning model to a certain extent. It can effectively avoid over-fitting, the model expression has a good distribution, and the model parameters have certain sparseness.

Based on the above analysis, this paper uses the following optimization for the super-resolution reconstruction model:

1) Using the ISODATA algorithm based on image structure features to cluster training datasets;

2) For different types of image training sets, the combination of pre-training and tuning is used for model training;

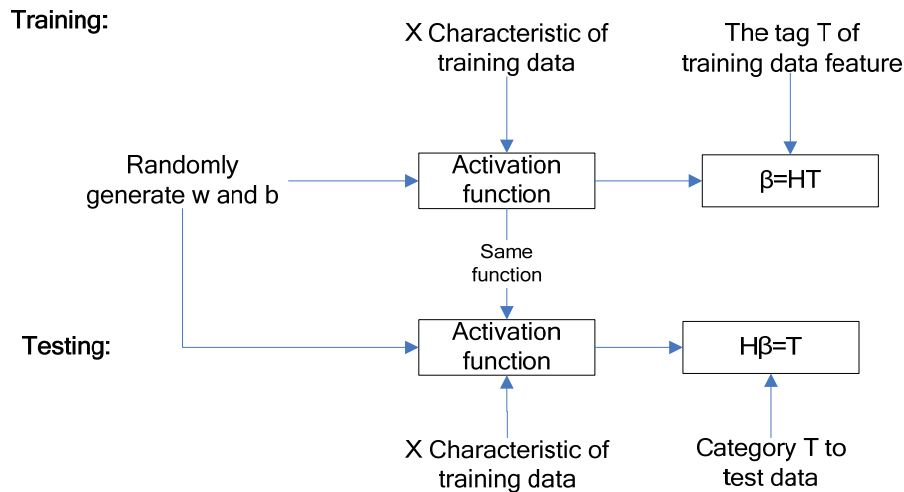3) Reduce the parameter scale of the CNN network to accelerate learning efficiency.

FIG. 2. ELM training and testing process

QUADRATIC RECONSTRUCTION ALGORITHM FLOW BASED ON CNN AND ELM

The super-resolution reconstruction method proposed in this paper combines the advantages of the CNN model for learning the features of the image structure and the ELM model for learning the pixel-level features of the image. The image reconstructed from the CNN is then incorporated into the trained ELM model. Secondary reconstruction, obtaining the final SR image, the overall framework of the method in this paper as shows.

Figure 3 describes in detail the overall process of algorithm training and testing. It is worth noting that during the training process, the two algorithm models CNN and ELM are independent of each other. Specifically, when the CNN training is performed, high- and low-resolution image block sample pairs having a gain coefficient are used, and after the ISODATA classification, a training process of pre-training and tuning is performed. In the training of ELM, the high and low-resolution training data samples are the entire image, and the amplification factor is less than K, so as to enhance the effect of image reconstruction.
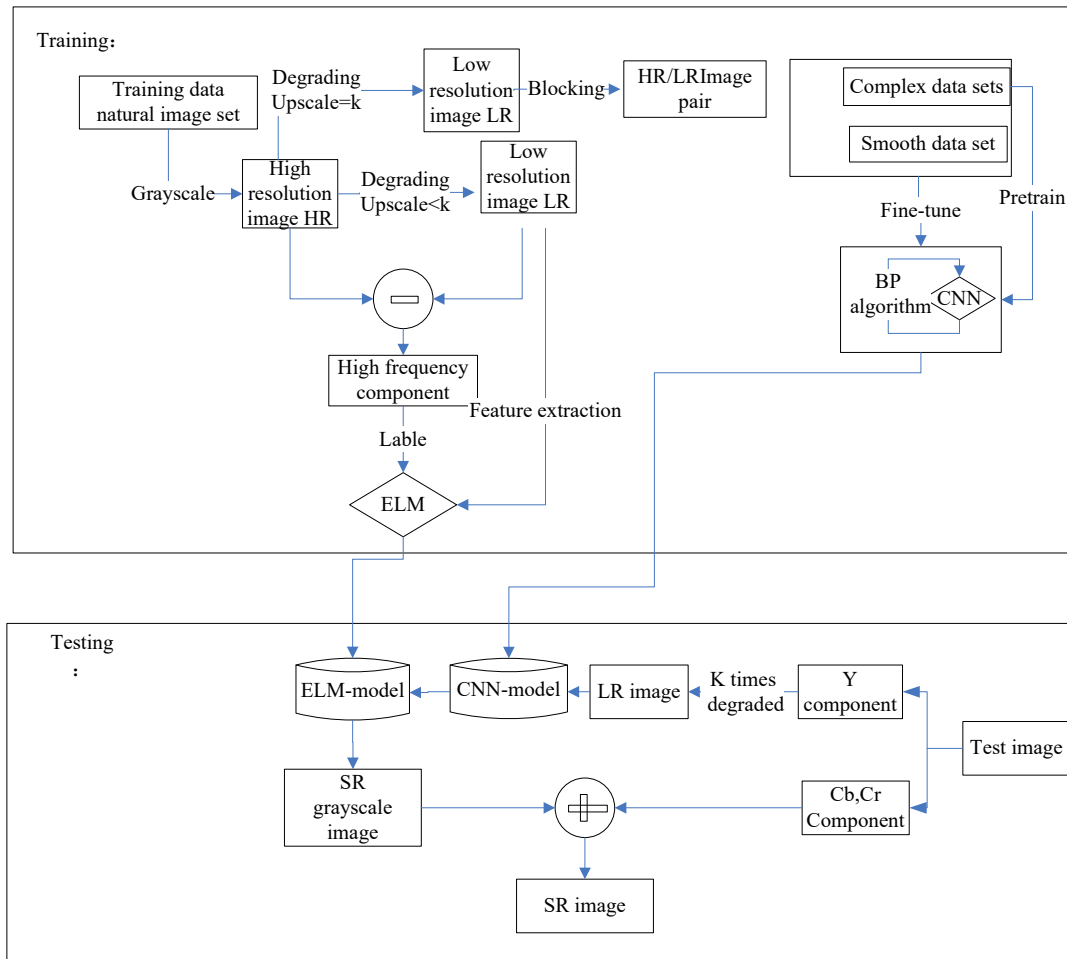

FIG. 3. Overall flow of the CNN+ELM algorithm

SUMMARY

This dissertation focuses on the problem of low training efficiency of deep learning super-resolution model, proposes the method of secondary super-resolution reconstruction combining CNN and ELM, and designs the overall algorithm flow based on secondary reconstruction of CNN and ELM. On the basis of the CNN super-resolution optimization algorithm, this paper is less than the global amplification factor, the low-resolution image pairs feature-level feature extraction at the pixel level, use the ELM model with higher training efficiency to learn the missing high-frequency components of the low-resolution image, and weight-superimpose the component and the CNN-reconstructed results. Implement secondary reconstruction of the input image.

ACKNOWLEDGEMENTS

REFERENCES

1. T.S. Huang, R. Tsati, "Multi-frame image restoration and registration", Advances in Computer Vision and Image Processing, 1984, 1:317-339.
2. Yang C, Huang J, Yang M, etal. Exploiting self-similarities for single frame super-resolution//Proceedings of the10th Asian Conference on Computer Vision. Queenstown, New Zealand, 2010: 497-510.
3. Freedman G, Fattal R. Image and video upscaling from local self-examples. ACM Transactions on Graphics (TOG), 2011, 30(2), Article12, 1-11.
4. Cui Z, Chang H, Shan S, etal. Deep network cascade for image super-resolution//Proceedings of the IEEE International Conference on European Conference on Computer Vision (ECCV). Zurich, Switzerland, 2014:49-64.
5. Chen Xiao-xuan, Qi Chun, Single-image super-resolution via low-rank matrix recovery and joint learning, Chinese Journal of Computers, 2014, 37(6):1372-1379.
6. Yang J, Wright J, Shuang T. Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 2010, 19(11):2861-2873.
7. Yang J, Wang Z, Lin Z. Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing, 2012, 21(8):3467-3478.
8. He L, Qi H, Zaretzki R, etal. Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution//Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Portland, USA, 2013:345-352.
9. R. Timofte, V. De and L. V. Gool, Anchored neighborhood regression for fast example-based super-resolution//Proceedings of the IEEE International Conference on International Conference on Computer Vision (ICCV), Sydney, Australia,2013:1920-1927.
10. Zhu Y, Zhang Y, Yuille A, etal. Single image super-resolution using deformable patches//Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, USA,2014:2917-2924.
11. Gao J, Guo Y, Yin M, et al. Restricted Boltzmann machine approach to couple dictionary training for image super-resolution//Proceedings of the IEEE International Conference on Image Processing (ICIP), Melbourne, Australia,2013:499-503.
12. Dong C, Changeloy C, He K. Learning a deep constructional network for image super-resolution//Proceedings of the IEEE International Conference on European Conference on Computer Vision (ECCV). Zurich, Switzerland,2014:184-199.