



Research Article

A Novel Face Super-Resolution Method Based on Parallel Imaging and OpenVINO

Zhijie Huang ¹, Wenbo Zheng,^{2,3} Lan Yan,^{3,4} and Chao Gou ¹

¹School of Intelligent Systems Engineering, Sun Yat-Sen University, Guangzhou 510275, China

²School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China

³The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

⁴School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence should be addressed to Chao Gou; gouchao@mail.sysu.edu.cn

Received 11 December 2020; Revised 26 January 2021; Accepted 29 January 2021; Published 15 February 2021

Academic Editor: Yu Liu

Copyright © 2021 Zhijie Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Face image super-resolution refers to recovering a high-resolution face image from a low-resolution one. In recent years, due to the breakthrough progress of deep representation learning for super-resolution, the study of face super-resolution has become one of the hot topics in the field of super-resolution. However, the performance of these deep learning-based approaches highly relies on the scale of training samples and is limited in efficiency in real-time applications. To address these issues, in this work, we introduce a novel method based on the parallel imaging theory and OpenVINO. In particular, inspired by the methodology of learning-by-synthesis in parallel imaging, we propose to learn from the combination of virtual and real face images. In addition, we introduce a center loss function borrowed from the deep model to enhance the robustness of our model and propose to apply OpenVINO to speed up the inference. To the best of our knowledge, it is the first time to tackle the problem of face super-resolution based on parallel imaging methodology and OpenVINO. Extensive experimental results and comparisons on the publicly available LFW, WebCaricature, and FERET datasets demonstrate the effectiveness and efficiency of the proposed method.

1. Introduction

Image super-resolution (SR) is a classical issue in the engineering field of image processing technology and computer vision [1]. The face super-resolution problem is one of the important branches of the super-resolution problem. With the advent of the age of intelligence, face super-resolution has been gradually applied in many applications such as face recognition, intelligent surveillance, and identity recognition.

Face super-resolution is a mathematical inverse problem, which is an “ill-posed” problem of producing a high-resolution face image with critical information or a good visual effect (close to the real face image) from a low-resolution face image [2]. At present, there are two categories of reconstruction methods: interpolation-based method and learning-based method [3].

The interpolation-based method is the first method proposed to solve the face super-resolution problem. It essentially regards the face super-resolution problem as an image improperly posed problem and solves the problem based on the prior information of low-resolution images. The classical methods include nearest neighbor interpolation method, bilinear interpolation method, and bicubic interpolation method [4]. Nearest neighbor interpolation method is simple and computed fast because the interpolation value is obtained by calculating the gray value of the nearest neighbor [5]. In bilinear interpolation method, it gets the interpolated pixel values by interpolating the value of its surrounding pixels bilinearly to realize the linear interpolation in the horizontal and vertical directions [6]. Different from the nearest neighbor interpolation method and the bilinear interpolation method, 16 neighborhood pixels of the point to be interpolated are used for cubic interpolation in

bicubic interpolation method [7]. Although the above-mentioned traditional methods have good real-time computational simplicity, the reconstructed images are unsatisfactory due to aliasing, blocking, and blurring artifacts in many cases. In recent years, some advanced interpolation-based methods have been proposed. For example, Jiang et al. [8] presented a missing strength intensity interpolation method based on smooth regression with local structure prior (LSP), named SRLSP for short. Wei [9] proposed an image super-resolution reconstruction method using the high-order derivative interpolation associated with fractional filter functions. However, since the interpolation methods rely on the prior information of low-resolution images, the reconstruction results will drop rapidly while inputting a small image.

Another category of method is the machine learning-based method, which essentially uses machine learning to learn the mapping between low-resolution face images and high-resolution images. The method based on statistical dictionary is one of the learning-based methods, which builds a dictionary and learns the connection between low-resolution images and high-resolution images. For example, Yang et al. [10] proposed a reconstruction method based on a sparse representation model, which is used to learn low-resolution and high-resolution image blocks by training low-resolution dictionary image blocks and high-resolution dictionary image blocks to get the sparse similarity of the reconstructed image finally. However, the method based on statistical learning needs to specify the number of dictionary elements and the noise variance model in advance, so it is difficult to obtain the optimal parameters of natural images in real life. As for other types of learning-based methods, such as position-patch-based method, Ma et al. [11] proposed a position-patch-based super-resolution method, which used the same position image patches and the same weights of each training image to hallucinate and reconstruct the high-resolution face image. Huang and Wu [12] introduced an approach to reconstruct high-resolution face image directly from a single low-resolution face image by using multiple local linear transformations and Procrustes analysis without using large redundant low-resolution and high-resolution patch databases. Jiang et al. [13] proposed a simple and effective scheme, local constraint representation (LCR). It is a method that not only incorporates locality constraints to maintain locality but also exploits the sparse property of the redundant data representation. In [14], a method based on single image is proposed to estimate the high-resolution embedding and enhance the local compatibility and smoothness constraints of each block in the target high-resolution image by using the training image pair and overlapping. It is noteworthy that there are categories of single image-based super-resolution and multiple images based super-resolution, such as TLcR-RL [15], a context-patch-based face hallucination approach that can fully exploit contextual information of image patches to obtain a high-resolution face image. There are essential differences between the two methods. In this paper, we only discuss the single image-based super-resolution.

Deep learning methods show great potentiality in terms of face super-resolution. Dong et al. [16] validated that CNN can achieve an end-to-end learning effectively from low-resolution images to high-resolution images and proposed the single-image reconstruction convolutional neural networks (SRCNN), and in 2016, Dong et al. [17] proposed a method (FSRCNN) based on compact hourglass-shaped CNN structure to reconstruct HR image by introducing a deconvolution layer. It is an improvement on the previous SRCNN [16] to achieve faster and better image super-resolution. Liang et al. [18] proposed a framework consisting of a deep convolutional neural network with image gradient priors, named SRCNN-Pr, which exploits various image priors during the training phase of a deep CNN. Mao et al. [19] introduced an approach based on deep convolutional neural network (DSRCNN), which combines the adversarial automatic with two depixelate layers to reconstruct the pixelated image and map it to a higher resolution. Lim et al. [20] proposed a multiscale deep super-resolution method by removing unnecessary modules in the traditional residual networks, which reconstructs high-resolution images with different upscaling factors in a single model. SRCNN is a pioneering work of deep learning in super-resolution reconstruction. All kinds of improvement methods based on SRCNN achieve better performance directly or indirectly by increasing the network structure. This type of deep learning method has been successfully applied to deal with the problem of face super-resolution reconstruction and achieves good results [3]. However, there are some limitations in the following aspects:

- (1) The reconstruction of the network needs to rely on a large number of face images, but it is time-consuming to get a large number of face images from real-world scenes.
- (2) The mean-square loss function has a forced prediction and label to match exactly, i.e., "either 0 or 1," which results in overconfidence of the model.
- (3) It is highly dependent on GPU devices and the inference speed is too slow on normal CPU devices.

In particular, we take the following strategies to address these three issues, respectively:

- (1) According to the parallel image theory [21, 22], we propose to learn from a large number of available caricature faces. In this paper, we use the Web-Caricature [23] database as an example to design model and conduct experiments.
- (2) We introduce a center loss function for constructing networks to classify the unmarked categories in the training set and realize the identification of features.
- (3) We apply the latest OpenVINO toolkit [24–26] provided by Intel to optimize the model on the 6th-generation CPU with a relatively cheap price. It shows that our proposed method could be used on a general mobile device for real-time applications.

In summary, we propose a deep face super-resolution algorithm based on parallel image theory and OpenVINO,

termed as **SRCNN-CL**, which is a sample improvement on the light SRCNN and yields better experimental results than the original networks in both qualitative and quantitative results. We use OpenVINO to speed up the inference of the model and measure this acceleration with quantitative assessments. In particular, we build a three-layer convolutional neural network to learn the mapping from low-resolution face images to high-resolution images. During this process, we use the center loss function to overcome the inability to discriminate features, and then we train the model based on parallel image theory. Finally, we use the OpenVINO toolkit for acceleration.

All in all, there are three contributions in this paper:

- (1) We propose a face super-resolution method based on parallel imaging theory with high robustness and efficiency.
- (2) We propose an OpenVINO-based face super-resolution reconstruction method, which accelerates the processing speed of the model and meets the real-time applications.
- (3) We improve the conventional network structure and propose a center loss function, which improves the ability of model to distinguish features. Compared with existing methods on the benchmark databases, better experimental results are obtained in terms of both qualitative and quantitative results.

2. OpenVINO Tools

In May 2018, Intel launched a toolkit called OpenVINO [27]. The toolkit can quickly deploy applications and solutions in simulating human vision. Especially on the convolutional neural network- (CNN-) based method, we can use this toolkit to span the Intel hardware of the computer vision (CV) workload and get a maximum performance. OpenVINO is mainly composed of model optimizer and inference engine. The model optimizer mainly imports models trained in standard frameworks such as Caffe and TensorFlow and transforms and optimizes them into a format that can be used by Intel tools (especially inference engines) [28]. The inference engine is an engine that runs deep learning models. It contains a set of libraries to integrate inference with applications easily.

Specifically, the model optimizer is a cross-platform command-line tool. It could facilitate the conversion between training and deployment environments, analyze the static models, and adjust deep learning models to achieve optimal execution on endpoint target devices. After using the model optimizer to create the intermediate representation, the inference engine is used to infer the input data. The inference engine is a C++ library with a set of C++ classes for inferring input data (images) and obtaining results [29]. The C++ library provides an API to read intermediate representations, set input and output formats, and execute models on the device. Figure 1 illustrates a typical work-flow for deploying a trained deep learning model.

The following is the summary of steps to optimize and deploy the trained models:

- (i) Configure the model optimizer for the framework.
- (ii) According to the trained network topology, weights, and bias values, transform the training model to generate an optimized intermediate representation (IR) of the model.
- (iii) Use the inference engine in the target environment to test the model in an intermediate representation format by verifying the application or the sample application.
- (iv) Integrate the inference engine in the application to deploy the model in the target environment.

3. Proposed Method

Parallel images [21] are a branch of parallel vision [22], which can provide large-scale and diverse image data (including real images and virtual images) for parallel vision research. At first, the virtual image is used to expand and supplement the real image to obtain the parallel image “big data” that combines virtual and real. Then, various visual models are learned and evaluated through computational experiments, and finally, the visual models are optimized online by parallel execution to realize the intelligent perception and understanding of complex environments [22, 30].

3.1. Parallel Imaging for Face Super-Resolution. Inspired by the parallel image theory [22], aiming at the problem of face super-resolution, we expand and supplement the corresponding real-world face images by using caricature face images on the Internet to obtain parallel face images that combine virtual and real information. Then, we learn and evaluate the computational experiment of the reconstruction model and optimize the visual model online to realize the intelligent perception and understanding of complex scenes with the help of parallel execution finally. Figure 2 shows the parallel vision frame structure of face super-resolution.

We use artificial web crawlers to download the corresponding face caricature images from the Internet, use the face super-resolution model to perform computational experiment and parallel execution of parallel vision, and learn to extract “little knowledge” applied to face super-resolution. In the process of computational experiments and parallel execution, we introduce the OpenVINO framework to accelerate and optimize the face super-resolution model. Figure 3 shows the technical flow of parallel imaging.

3.2. Network Architecture. In order to get the high-resolution images which meet the requirements better, we propose to apply the bicubic interpolation method to obtain low-resolution images before training network and to perform corresponding preprocessing in the testing phase.

In our network, the first layer of convolution operation is used for feature extraction and it could be expressed as

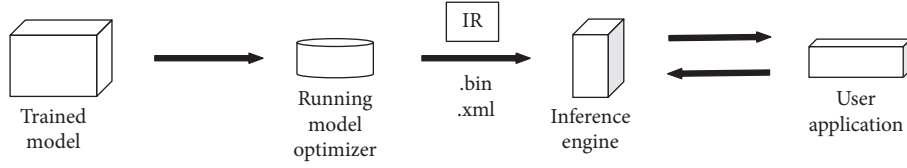


FIGURE 1: Basic framework of OpenVINO.

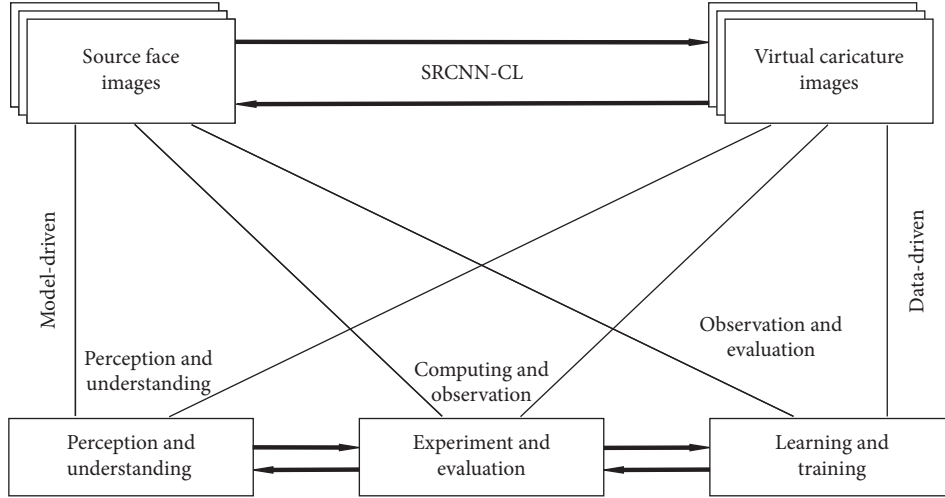


FIGURE 2: Basic framework of parallel vision.

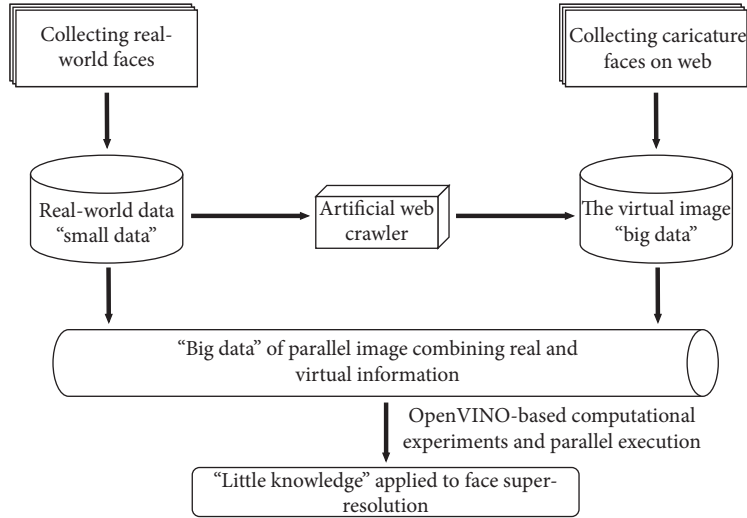


FIGURE 3: Technological flowchart of parallel imaging.

$$C_1(X) = \max(0, W_1 \otimes X + B_1) + \frac{\min(0, W_1 \otimes X + B_1)}{a}, \quad (1)$$

where B_1 is the bias, W_1 is the filter with its size, $W_1 = A \times D_1 \times m_1$, A is the number of channels of the input image, D_1 is the space size of the filter, m_1 is the number of filters, \otimes is the convolution operation, and the parameter a is uniformly distributed.

In the first layer, we extract features from each image block through a convolution operation. We map the features

extracted from the first layer to the second layer. The operation of the second layer could be expressed as

$$C_2(X) = \max(0, W_2 \otimes C_1 + B_2) + \frac{\min(0, W_2 \otimes C_1 + B_2)}{a}, \quad (2)$$

where B_2 is the bias, W_2 is the filter with its size, $W_2 = D_1 \times D_2 \times m_2$, D_2 is the space size of the filter, and m_2 is the number of filters.

In the third layer of our network, we get the high-resolution images finally. The operation of the third layer could be expressed as

$$C_3 = W_3 \otimes C_2(X) + B_3, \quad (3)$$

where B_3 is the bias, W_3 is the filter with its size, $W_3 = D_2 \times D_3 \times m_2 \times c$, D_3 is the space size of the filter, and c is the number of channels of the output image.

In order to learn the end-to-end mapping function $C = \{C_1, C_2, C_3\}$, the traditional method uses the mean square (MSE) method to estimate the network parameters $\Psi = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ as

$$\text{MSE}(\Psi) = \frac{1}{n} \sum_{i=1}^n \|C(Y, \Psi) - X_i\|^2, \quad (4)$$

where n is the number of training samples.

According to Wen's theory [31], the network is used as a feature extractor to extract features for recognition. In this case, the training set contains the categories of test samples and the performance of the model degrades less. For the face super-resolution, because of the impossibility to collect all the face images, it is natural that some faces are not in the training set. Thus, the features extracted by the network are not only separable but also highly discriminative. Further, it is obviously inappropriate to use the MSE loss function. Therefore, we introduce the center loss function to construct the network:

$$\text{Loss} = - \sum_{i=1}^m \log \frac{e^{W_i \otimes Y_i + B_i}}{\sum_{j=1}^n e^{W_j \otimes X_j + B}} + \frac{\lambda}{2} \sum_{i=1}^n \|C(Y, \Psi) - X_i\|^2, \quad (5)$$

where m is the number of categories.

In general, we have built a three-layer convolutional network as shown in Figure 4.

4. Experimental Results

4.1. Data Preparation and Parameter Setting. Our work uses labeled faces in the wild (LFW) database [32] and Web-Caricature [23] database collected through the network as training set and testing set. We select 60 images of LFW database and 40 images of WebCaricature database randomly and take a total of 100 images of them as the training set. We randomly select 20 images in the face recognition technology (FERET) [33] database as the test set. An example of the training sample set is shown in Figure 5.

In our experiment, we choose the simplest three-layer convolution neural network (9-1-5). We set m_1 to 64, m_2 to 32, and μ to 0.9. The weight of each layer in the network is initialized to a Gaussian distribution with a mean value of 0 and a standard deviation of 0.0001. The deviation of each layer is initialized to 0, and the learning rate is always fixed to 10^{-4} during network training. The convergence diagram of network training is shown in Figure 6. It shows that the model has converged. In order to solve the problem of different magnification factors, we respectively downsample the data in the training set with magnification factors of 2, 3, and 4, and finally, we get a training set containing 400 face images.

4.2. Quantitative Evaluation Protocol. In order to evaluate the performance of the super-resolution algorithm and the test speed of the model, we use peak signal-to-noise ratio (PSNR) and Image Structure Similarity Index (SSIM) to evaluate the performance of the super-resolution algorithm. We evaluate the acceleration effect of OpenVINO by parallel speedup ratio.

Peak signal-to-noise ratio represents the ratio of the variance of the image signal to noise. The higher the value, the better the reconstruction effect of the image. For the original image I and reconstructed image K with $M \times N$ size, the specific calculation formula is as follows:

$$\text{MSE} = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m \|K(i, j) - I(i, j)\|^2, \quad (6)$$

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}^2}{\text{MSE}},$$

where MAX indicates the maximum value of the image color and the value of the 8-bit image is 255.

Image structure similarity is an index to measure the similarity of two images. The value range is $[0, 1]$. The closer the value of SSIM is to 1, the more similar the structure of the two images. The specific calculation formula is as follows:

$$\text{SSIM}(I, K) = \frac{(2\mu_I \mu_K + \mathfrak{C}_1)(2\sigma_{IK} + \mathfrak{C}_2)}{(\mu_I^2 + \mu_K^2 + \mathfrak{C}_1)(\sigma_I^2 + \sigma_K^2 + \mathfrak{C}_2)}, \quad (7)$$

where μ_I is the average of I , μ_K is the average of K , σ_I^2 is the variance of I , σ_K^2 is the variance of K , σ_{IK} is the covariance of I , and K , \mathfrak{C}_1 , and \mathfrak{C}_2 are the constants.

For the execution of the model, we hope that tasks are evenly distributed among the cores of the central processing unit (CPU) without introducing additional workload for each core. If we can successfully achieve the goal, when running programs on the P core system, each core runs a process or thread, the running speed of the parallel program is P times the speed of the serial program. Multiprocess and multithreads always introduce some costs. For example, sharing memory programs usually have critical sections and need to use some mutual exclusion mechanisms, such as mutex. It is a cost for parallel programs but serial programs to call mutex because it will force parallel programs to execute critical section code serially. Distributed memory programs usually need to transfer data across the network, which is slower than accessing data in local memory. In contrast, serial programs do not have these additional overheads. So, it is very difficult to find a parallel program with an ideal speedup ratio. In addition, with the number of processes increasing, the overhead increases. More threads means more threads need to access the critical section, and more processes means more data need to be transferred across the network. Therefore, we simply define the speedup of a parallel program as

$$\rho = \frac{T_{\text{ParallelExecution}}}{T_{\text{SerialExecution}}}. \quad (8)$$

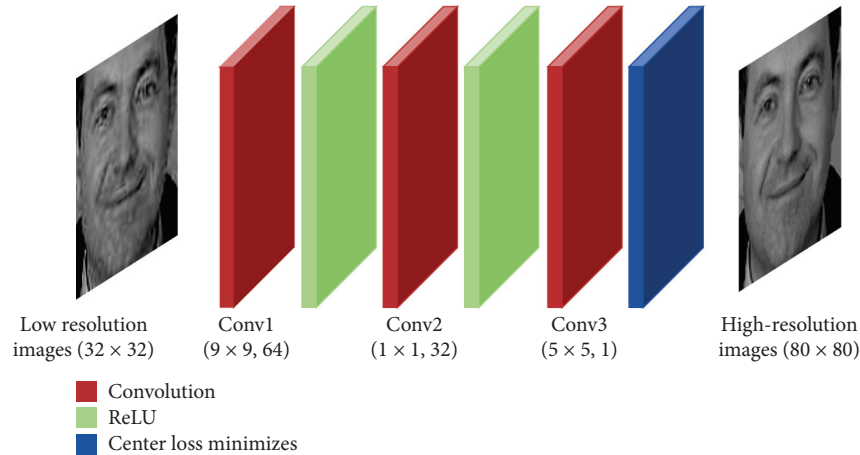


FIGURE 4: The structure of CNN for super-resolution.



FIGURE 5: A case of an experimental training sample set.

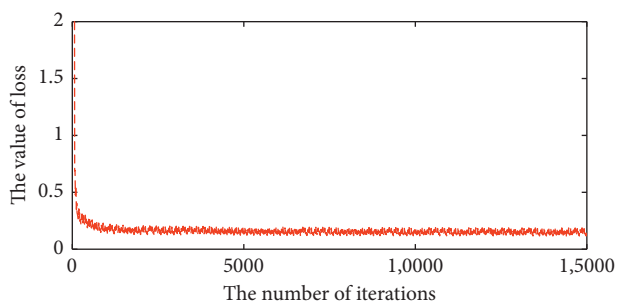


FIGURE 6: The flowchart of the model of the loss function.

4.3. Analysis of Experimental Results. After reconstructing low-resolution face images obtained by interpolation, it shows a better visual result in terms of detailed reconstruction with our method. Since the human eyes, nose, and mouth are the regions with large signal energy in the face image [42], they are also the most important representations of the human face. Therefore, in this paper,

we have chosen to illustrate the face reconstruction image with a magnification factor of 4, as shown in Figure 7.

We compared our method with fifteen face image super-resolution methods which are LcR [13], SSR [41], LLT [12], Wang's [40], NE [14], SR [10], WSR [34], CDA [39], bicubic [35], MSR [11], Yang's [10], Glanser's [36], DPSR [37], SRCNN [16], and SRAHF [38]. The result is shown in Table 1.

According to Table 1, our method is optimal and sub-optimal in terms of SSIM and PSNR compared with many algorithms. Especially, compared with the original SRCNN, the PSNR and SSIM are improved by approximately 1.3 dB and 0.02, respectively. Therefore, it can be proved that the algorithm in this paper is highly robust and the model has improved robustness after introducing the parallel image theory.

We compared the test speed of SRCNN with our method before and after using OpenVINO. The result is shown in Table 2. It can be obtained from Table 2 that after using OpenVINO, the parallel speedup ratios of SRCNN



FIGURE 7: Comparison of face image super-resolution methods' results on FERET by different methods. (a) Original HR images. (b) LR input images. (c) LcR [13]. (d) SSR [41]. (e) LLT [12]. (f) Wang's [40]. (g) NE [14]. (h) SR [10]. (i) WSR [34]. (j) CDA [39]. (k) Bicubic [35]. (l) MSR [11]. (m) Yang's [10]. (n) Glanser's [36]. (o) DPSR [37]. (p) SRCNN [16]. (q) SRAHF [38]. (r) Ours.

TABLE 1: Comparison of face super-resolution methods' results on FERET database.

Methods	PSNR (dB)	SSIM
LLT [12]	26.3577	0.8864
WSR [34]	31.9449	0.8954
Bicubic [35]	24.7071	0.8622
MSR [11]	31.9216	0.8987
Yang's [10]	29.0257	0.8579
Glanser's [36]	24.7513	0.8246
DPSR [37]	29.5572	0.8743
SRCNN [16]	30.3913	0.8832
SRAHF [38]	30.1527	0.8792
LcR [13]	32.3186	0.9032
CDA [39]	32.2249	0.9011
Wang's [40]	31.3323	0.8893
NE [14]	31.1459	0.8794
SR [10]	31.2309	0.8798
SSR [41]	32.5439	0.9074
Ours	31.6786	0.9085

TABLE 2: Parallel speedup ratio comparison.

Methods	Test time (frame/second)	Parallel speedup ratio (%)
SRCNN	5	—
Ours	5	—
SRCNN + OpenVINO	35	700
Ours + OpenVINO	38	760

and our method reach 700% and 760%, respectively. It can be proved that OpenVINO accelerates the test speed of our model.

5. Conclusion

This paper proposes an improved three-layer end-to-end convolutional neural network, which is used to reconstruct low-resolution face images to obtain corresponding high-resolution images. This paper introduces the parallel image theory to improve the robustness of our model, and it accelerates the test speed of our model by using OpenVINO. In addition, this paper uses a method of cotraining with multiple magnification factors to reduce the number of required network models. The final experimental results show that the method in this paper can cope with the super-resolution reconstruction of face images with different magnification factors and achieves better than traditional reconstruction methods.

On the other hand, through our experiment, the parallel image theory can effectively improve the robustness within the framework of existing model. In the parallel execution stage of the parallel image, OpenVINO can be used to accelerate the execution speed of our model, which makes the parallel vision model faster and better for real-time application. Therefore, our next research direction is to apply OpenVINO and parallel image theory to other computer vision models.

Data Availability

The training data used to support the findings of this study were collected from the LFW, WebCaricature, and FERET public databases.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Zhijie Huang and Wenbo Zheng contributed equally to this study.

Acknowledgments

This work was supported in part by the National Key R&D Program of China (2020YFB1600400), in part by the National Natural Science Foundation of China (61806198), and in part by the Key Research and Development Program of Guangzhou (202007050002).

References

- [1] K. Nguyen, C. Fookes, S. Sridharan, M. Tistarelli, and M. Nixon, "Super-resolution for biometrics: a comprehensive survey," *Pattern Recognition*, vol. 78, pp. 23–42, 2018.
- [2] Y. Zhang, F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "Longitudinally guided super-resolution of neonatal brain magnetic resonance images," *IEEE Transactions on Cybernetics*, vol. 49, no. 2, pp. 662–674, 2019.
- [3] K. Li, S. Yang, R. Dong, X. Wang, and J. Huang, "Survey of single image super-resolution reconstruction," *IET Image Processing*, vol. 14, no. 11, pp. 2273–2290, 2020.
- [4] D. Huang and H. Liu, "A short survey of image super resolution algorithms," *Journal of Computer Science Technology Updates*, vol. 2, no. 2, pp. 19–29, 2015.
- [5] R. R. Schultz and R. L. Stevenson, "A bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.
- [6] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 6, pp. 508–517, 1978.
- [7] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 10, no. 10, pp. 1521–1527, 2001.
- [8] J. Jiang, C. Chen, J. Ma, Z. Wang, and R. Hu, "Srlsp: a face image super-resolution algorithm using smooth regression with local structure prior," *IEEE Transactions on Multimedia*, vol. 19, no. 1, pp. 27–40, 2016.
- [9] D. Wei, "Image super-resolution reconstruction using the high-order derivative interpolation associated with fractional filter functions," *IET Signal Processing*, vol. 10, no. 9, pp. 1052–1061, 2016.
- [10] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [11] X. Ma, J. Zhang, and C. Qi, "Hallucinating face by position-patch," *Pattern Recognition*, vol. 43, no. 6, pp. 2224–2236, 2010.
- [12] H. Huang and N. Wu, "Fast facial image super-resolution via local linear transformations for resource-limited applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 10, pp. 1363–1377, 2011.
- [13] J. Jiang, R. Hu, Z. Wang, and Z. Han, "Noise robust face hallucination via locality-constrained representation," *IEEE Transactions on Multimedia*, vol. 16, no. 5, pp. 1268–1281, 2014.
- [14] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004 CVPR*, vol. 1, p. 1, Washington, DC, USA, June 2004.
- [15] J. Jiang, Y. Yu, S. Tang, J. Ma, A. Aizawa, and K. Aizawa, "Context-patch face hallucination based on thresholding locality-constrained representation and reproducing learning," *IEEE Transactions on Cybernetics*, vol. 50, no. 1, pp. 324–337, 2020.
- [16] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016a.
- [17] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proceedings of the European Conference on Computer Vision*, pp. 391–407, Amsterdam, Netherlands, October 2016b.
- [18] Y. Liang, J. Wang, S. Zhou, Y. Gong, and N. Zheng, "Incorporating image priors with deep convolutional neural networks for image super-resolution," *Neurocomputing*, vol. 194, pp. 340–347, 2016.
- [19] H. Mao, Y. Wu, J. Li, and Y. Fu, "Super resolution of the partial pixelated images with deep convolutional neural network," in *Proceedings of the 24th ACM International Conference on Multimedia*, pp. 322–326, Amsterdam, Netherlands, October 2016.
- [20] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 136–144, Honolulu, HI, USA, July 2017.
- [21] W. Kunfeng, L. Yue, W. Yutong, and W. Fei-Yue, "Parallel imaging: a unified theoretical framework for image generation," in *Proceedings of the 2017 Chinese Automation Congress (CAC)*, pp. 7687–7692, Jinan, China, October 2017.
- [22] K. Wang, C. Gou, N. Zheng, J. M. Rehg, and F.-Y. Wang, "Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives," *Artificial Intelligence Review*, vol. 48, no. 3, pp. 299–329, 2017.
- [23] J. Huo, W. Li, Y. Shi, Y. Gao, and H. Yin, "Webcaricature: a benchmark for caricature recognition," in *Proceedings of the British Machine Vision Conference*, Newcastle upon Tyne, England, September 2018.
- [24] N. A. Andriyanov, "Analysis of the acceleration of neural networks inference on intel processors based on openvino toolkit," in *Proceedings of the 2020 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, pp. 1–5, Svetlogorsk, Russia, July 2020.
- [25] Z. Jin and H. Finkel, "Analyzing deep learning model inferences for image classification using openvino," in *Proceedings of the 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pp. 908–911, New Orleans, LA, USA, May 2020.
- [26] V. Kustikova, E. Vasiliev, A. Khvatov et al., "Intel distribution of openvino toolkit: a case study of semantic segmentation," in *Proceedings of the International Conference on Analysis of Images, Social Networks and Texts*, pp. 11–23, Kazan, Russia, July 2019.
- [27] J. Yrjänäinen, X. Ni, B. Adhikari, and H. Huttunen, "Privacy-aware edge computing system for people tracking," in

- Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP)*, pp. 2096–2100, Abu Dhabi, UAE, October 2020.
- [28] R. D. Castro-Zunti, J. F. Yépez Rodríguez, and S.-B. Ko, “License plate segmentation and recognition system using deep learning and openvino,” *IET Intelligent Transport Systems*, vol. 14, no. 2, pp. 119–126, 2020.
- [29] A. Demidovskij, Y. Gorbachev, M. Fedorov et al., “Openvino deep learning workbench: comprehensive analysis and tuning of neural networks inference,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, Seoul, Korea (South), October 2019.
- [30] C. Gou, H. Zhang, K. Wang, F.-Y. Wang, and Q. Ji, “Cascade learning from adversarial synthetic images for accurate pupil detection,” *Pattern Recognition*, vol. 88, pp. 584–594, 2019.
- [31] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, “A discriminative feature learning approach for deep face recognition,” in *Proceedings of the European Conference on Computer Vision*, pp. 499–515, Springer, Amsterdam, Netherlands, October 2016.
- [32] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua, “Labeled Faces in the Wild: A Survey,” *Advances in Face Detection and Facial Image Analysis*, pp. 189–248, Springer International Publishing, Cham, Switzerland, 2016.
- [33] K. Baek, B. A. Draper, J. R. Beveridge, and K. She, “Pca vs. ica: a comparison on the feret data set,” *Journal of Colloid and Interface Science JClS*, pp. 824–827, 2002.
- [34] Z. Wang, R. Hu, S. Wang, and J. Jiang, “Face hallucination via weighted adaptive sparse regularization,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 5, pp. 802–813, 2014.
- [35] F. Yeganli, M. Nazzal, and H. Ozkaramanli, “Super-resolution using multiple structured dictionaries based on the gradient operator and bicubic interpolation,” in *Proceedings of the 2016 24th Signal Processing and Communication Application Conference (SIU)*, pp. 941–944, Zonguldak, Turkey, May 2016.
- [36] D. Glasner, S. Bagon, and M. Irani, “Super-resolution from a single image,” in *Proceedings of the 2009 IEEE 12th International Conference on Computer Vision*, pp. 349–356, Kyoto, Japan, September 2009.
- [37] Y. Zhu, Y. Zhang, and A. L. Yuille, “Single image super-resolution using deformable patches,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2917–2924, Columbus, Ohio, July 2014.
- [38] L.-J. Deng, W. Guo, and T.-Z. Huang, “Single image super-resolution by approximated heaviside functions,” *Information Sciences*, vol. 348, pp. 107–123, 2016.
- [39] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, “Coupled deep autoencoder for single image super-resolution,” *IEEE Transactions on Cybernetics*, vol. 47, no. 1, pp. 27–37, 2017.
- [40] X. Wang and X. Tang, “Hallucinating face by eigen-transformation,” *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 35, no. 3, pp. 425–434, 2005.
- [41] J. Jiang, J. Ma, C. Chen, X. Jiang, and Z. Wang, “Noise robust face image super-resolution through smooth sparse representation,” *IEEE Transactions on Cybernetics*, vol. 47, no. 11, pp. 3991–4002, 2017.
- [42] M. J. Arcaro, P. F. Schade, J. L. Vincent, C. R. Ponce, and M. S. Livingstone, “Seeing faces is necessary for face-domain formation,” *Nature Neuroscience*, vol. 20, no. 10, pp. 1404–1412, 2017.