

Research Article

Neural Network-Based Ultra-High-Definition Video Live Streaming Optimization Algorithm

Yunning Feng,¹ Nan Hu^{ID,1} and Xiaosheng Yu²

¹School of Electrical and Control Engineering, Shenyang Jianzhu University, Shenyang 110168, China

²School of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Nan Hu; hunan@sjzu.edu.cn

Received 11 February 2022; Revised 20 March 2022; Accepted 29 March 2022; Published 23 April 2022

Academic Editor: Xingsi Xue

Copyright © 2022 Yunning Feng 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.

Online live streaming has been widely used in distant teaching, online live shopping, and so on. Particularly, online teaching live streaming breaks the time and space boundary of teaching and has better interactivity, which is a new distant education mode. As a new online sales model, online live shopping promotes the rapid development of Internet economy. However, the quality of live video affects the user experience. This paper studies the optimization algorithm of ultra-high-definition live streaming, focusing on superresolution technology. Convolutional neural network (CNN) is a multilayer artificial neural network designed to process two-dimensional input data. It takes advantage of CNN in image processing. This paper proposes an image superresolution algorithm based on hybrid dilated convolution and Laplacian pyramid. By mixing the dilated convolution module, the receptive field of the network can be improved more effectively to obtain more context information so that the high-frequency features of the image can be extracted more effectively. Experiment was running on Set5, Set14, Urban100, and BSD100 datasets, and the results reveal that the proposed algorithm outperforms baselines with respect to peak signal to noise ratio (PSNR), structural similarity index measurement (SSIM), and image quality.

1. Introduction

At the end of 2019, the COVID-19 pandemic spread across the world. However, with the continuous spread of the pandemic, the application of online live streaming is becoming more and more extensive, such as online live streaming for shopping and online live streaming for teaching [1, 2]. However, the quality of live video needs to be further improved. With the continuous development of multimedia technology, online live streaming, short video, and other video applications have gradually become mainstream media for people's study, life, and entertainment due to their strong social and interactive characteristics [3]. According to relevant survey [4], video media generated 60% of Internet traffic in 2016, and it will develop further. By 2020, that figure has reached 78%. However, the video system is often limited by various objective conditions, especially in the video sending end, the video collection equipment with insufficient accuracy, limited network bandwidth, and terminal with

insufficient processing capacity make it difficult for the video system to provide adequate ultra-high-definition video sources [5–7].

In order to solve the above problems, superresolution is used in the video system so that the video application with limited objective conditions can also provide high-quality video presentation [8]. For example, wearable video devices with relatively weak CPU processing capacity are often unable to support high-resolution video and can only use low-resolution video formats. But users of these devices still want high-definition video content. Therefore, using super-resolution technology to restore the video quality at the video display end can greatly improve the visual experience of users [9, 10]. In addition, video superresolution technology is also used in many fields such as medical image research, security monitoring processing, and video coding and decoding, which has very high research value. Although many video superresolution methods have been proposed, due to the characteristics of video frames and the diversity

of video scenes, their superresolution results are not completely satisfactory. Further research is needed to improve the performance of video superresolution.

Superresolution is the process of generating high-resolution images from a set of low-resolution images [11]. This process supplements the spatial pixels of the image, increases the texture details of the image, restores the high-frequency information lost in the imaging process, and makes the image exquisite in detail, natural in picture, and has a better visual effect. This process can also be seen as the reverse process from high-resolution images to multiple low-resolution images. The simplest method of video superresolution is to perform single-frame superresolution on each frame of low-resolution video directly and finally restore all the obtained high-resolution images to high-resolution video according to the sequence of video stream [12]. However, such methods do not take into account the correlation between frames in the video, and the resulting high-resolution video may have problems such as poor interframe transition and interframe flicker. When superresolution technology is used in video, not only the current low-resolution video frame to be restored but also the transition relationship between the current frame and adjacent frames should be considered [13, 14]. Based on the single-frame image superresolution technology, combined with the imaging characteristics of video sequence, video superresolution technology often uses the redundant information between adjacent frames to further improve the superresolution performance.

Convolutional neural networks (CNN) imitating biological vision mechanisms are widely used in the field of computer vision. In superresolution reconstruction, CNN takes advantage of its learning function to establish the mapping relationship between low-resolution images and high-resolution images through training [15]. Due to its unique properties, CNN not only provides good performance in feature perception but also can detect features close to human visual system observation, so it has been widely applied in the field of superresolution technology. Accordingly, the main contribution of this paper is that a hybrid dilated convolution and Laplacian pyramid-based image superresolution algorithm is proposed.

The remainder of this paper is organized as follows. Section 2 reviews related work. In Section 3, the improved image superresolution algorithm is presented. Experimental results are presented in Section 4. Section 5 concludes this paper.

2. Related Work

Video is a sequence of images that are projected onto the screen at a rate that gives it image continuity. In order to improve the image resolution and display effect, many image resolution enhancement algorithms have been proposed by relevant researchers. Traditional superresolution algorithms include image interpolation, image superresolution based on sparse representation, and image superresolution based on manifold learning. In [16], the authors proposed a new single image superresolution method, which obtained the

initial high-resolution image by feature constrained polynomial interpolation. In [17], a new random forest method for image superresolution feature enhancement was proposed, using the traditional gradient-based feature to enhance the image superresolution feature, and formulated different feature formulas in different stages of image superresolution processing. In [18], a new superresolution algorithm for vertically guided neonatal image was proposed. In [19], a single image recognition method combining comprehensive sparse coding and analytical sparse coding was proposed. In [20], the author proposed a sparse Bayesian estimation-based single image superresolution method to reconstruct the superresolution image. In [21], a manifold learning-based improved texture image superresolution algorithm was proposed.

Many superresolution methods based on neural network have been proposed. In [22], the authors developed a capsule attention and reconstruction neural network (CARNN) framework to incorporate the capsule into the image superresolution CNN. In [23], the authors constructed a full deconvolution neural network (FDNN) and used FDNN to solve the problem of single image superresolution. In order to improve the resolution of remote sensing images, in [24], a superresolution neural network called progressive residual depth neural network (PRDNN) was proposed. In [25], a new perceptual image superresolution method was proposed to gradually generate visually high-quality results by constructing a stage network. In [26], the authors used CNN to generate superresolution underwater images. In [27], the authors used the enhanced attention network to realize the superresolution of compressed images. In [28], an arbitrary scale superresolution method of medical image was proposed, which combined meta learning with generative adversarial networks. In [29], the authors used gradual strategy to train CNN and proposed an efficient superresolution model. In [30], a deformable residual convolution network for image superresolution was proposed to enhance the transformation modeling ability of CNN. The emergence of deep learning has greatly promoted the development of image superresolution reconstruction, motivating this paper.

3. Proposed Method

Inspired by the Laplacian pyramid structure, in this section, a hybrid dilated convolution and Laplacian pyramid-based image superresolution algorithm is proposed.

As shown in Figure 1, the network can be divided into $\log_2 RT$ levels, and RT is the reconstruction times. Each level can be divided into feature extraction and image reconstruction. The feature extraction of each level is composed of a hybrid dilated convolution layer and a deconvolution layer. The reconstruction consists of a convolution layer and an element-wise summation layer [31]. The feature extraction of the first level is slightly different from that of other levels. Compared with other levels, the first level has an extra convolution layer for extracting shallow features and channel transformation because of the number of channels in the input image is different from that of each level [32, 33]. Therefore, a convolution layer with input channel 1 and output channel 64 is added at the beginning of the first level to

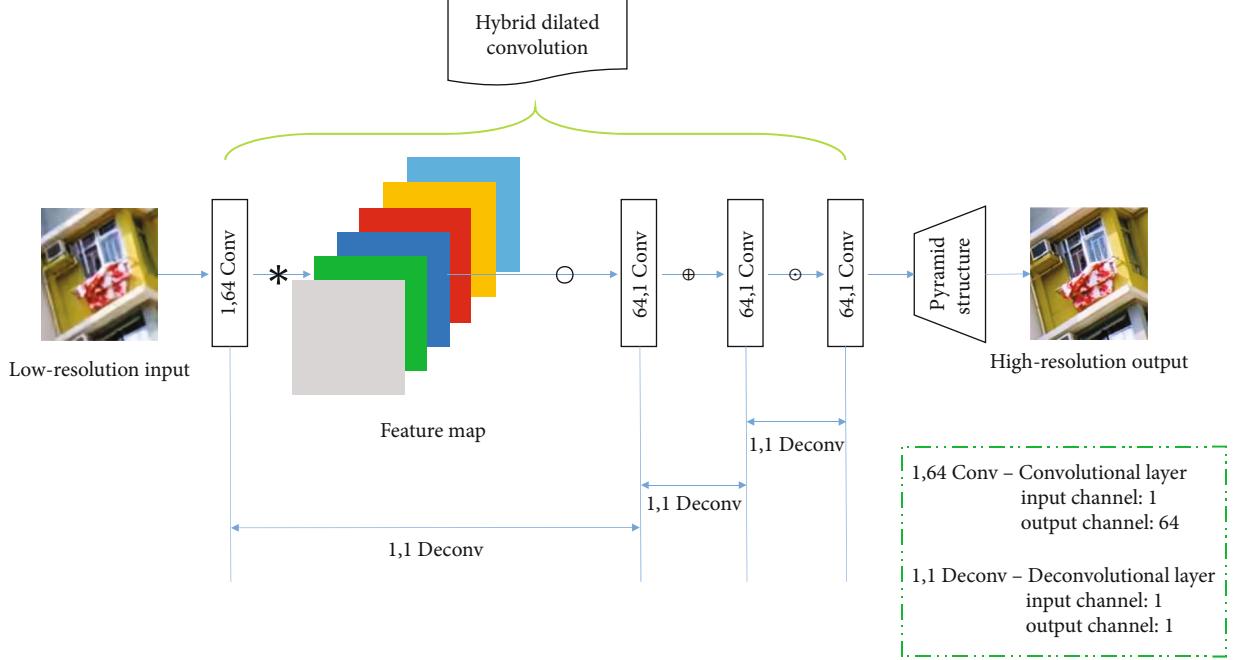


FIGURE 1: Overall network architecture of the proposed method.

extract shallow features from low-resolution images with input channel 1 and output 64 feature images. In feature extraction, high-dimensional features are extracted, and then, the extracted features are amplified by two times with a deconvolution layer, which serves as the input of feature extraction at the next level and reconstruction at this level.

The input of the image reconstruction is the output image of the reconstruction of the upper level and the output of the feature extraction of this level. The image goes through a deconvolution layer with input channel 1 and output channel 1, and the image is amplified by two times. The feature image is fused by a convolution layer with input channel 64 and output channel 1, and the original feature image with channel number 64 is transformed into a residual image with channel number 1. Finally, the amplification image and residual image are added together by element-wise summation operation to obtain the output of the reconstruction of this level and then serve as the input of the reconstruction of the next level. In this way, feature extraction and reconstruction are performed level by level. Moreover, after $\log_2 RT$ level feature extraction, high-resolution images with the required amplification of RT can be obtained.

Different from upsampling and downsampling, this section adopts the Laplacian pyramid structure step by step for image reconstruction to better balance the accuracy and efficiency of reconstruction. Let input low-resolution image from the network be I^{LR} , and C_{RT} is used to represent the convolution layer of the first layer for shallow feature extraction.

$$C_{RT}(I^{LR}) = \max(0.3 \times (\log w_c * I^{LR} + b_c), w_c * I^{LR} + b_c), \quad (1)$$

where w_c and b_c are the weight and bias of the first-level convolution kernel, $*$ is the convolution operation, the convolution kernel size of the first-level convolution is 3×3 , the input channel is 1, and the output channel is 64. Among them, $\max()$ is the activation function ReLU, and then, $C_{RT}(I^{LR})$ is passed into the network as the input of the first-level feature extraction.

Let C_{HD}^k represent the hybrid dilated convolution module of the k th level feature extraction, C_k represents its output feature map, C_D^k represents the deconvolution layer of the k th level feature extraction, and $C_{k,D}$ is its output feature map; then we have

$$C_1 = C_{HD}^1(C_{RT}(I^{LR})), \quad (2)$$

$$C_{k,D} = C_D^k(C_k), \quad k = 1, 2, \dots, \log_2 RT, \quad (3)$$

$$C_k = C_{HD}^k(C_{k-1,D}), \quad k = 1, 2, \dots, \log_2 RT. \quad (4)$$

As can be seen from equations (2) to (4), each level of Laplacian pyramid structure network feature extraction is made by hybrid dilated convolution module that is followed by deconvolution layers. The output of the hybrid dilated convolution module for deconvolution layer at the corresponding level of input and the output of the deconvolution layer to the next level is regarded as the input feature extraction at the next lower level hybrid dilated convolution of the input [34]. The hybrid dilated convolution module C_{HD}^k is composed of multiple hybrid dilated convolution blocks. The deconvolution layer can be defined as follows.

$$C_D^k(C_k) = \text{ReLU}(\log w_d \circ C_k + b_d), \quad (5)$$

where w_d and b_d are the weight and bias of the deconvolution layer and \circ is the deconvolution operation. The size of the convolution kernel of the deconvolution layer of the specific feature extraction is 4×4 , and the step size is 2 (amplified twice each time). The input channel is 64, and the output channel is 64.

At the first level in reconstruction branch, the low-resolution image I^{LR} is input to the deconvolution layer of the reconstruction, $C_{O,D}^k$ is used to represent the deconvolution layer of the reconstruction branch, and O_k represents its output; then, we have

$$O_1 = C_{O,D}^1(I^{LR}). \quad (6)$$

C_R^k represents the convolution layer of the reconstruction branch, and O_R represents its output; then, we have

$$\begin{aligned} O_R &= C_R^k(C_{k,R}), \\ C_R^k(C_{k,R}) &= w_R^k * C_{k,R} + b_R^k, \end{aligned} \quad (7)$$

where $C_{k,R}$ is the output of the feature extraction branch of this level and w_R^k and b_R^k are the weight and bias of the convolutional layer of the reconstruction branch of the k th level, respectively. While the size of the convolutional kernel of the convolutional layer is 3×3 , the input channel is 64, and the output channel is 1. The output of the reconstructed branch at this level is defined as follows.

$$O_k^R = \text{ReLU}(O_k \oplus O_R), \quad (8)$$

where \oplus represents the element-wise summation layer.

The output expression of other level is defined as follows.

$$\begin{aligned} O_k &= C_{O,D}^k(O_{k-1}), \\ C_{O,D}^k(O_{k-1}) &= w_{O,R}^k \circ O_{k-1} + b_{O,R}^k, \end{aligned} \quad (9)$$

where $w_{O,R}^k$ and $b_{O,R}^k$ are the weight and bias of the deconvolution layer of the reconstruction branch of the k th level, respectively. While the size of the convolutional kernel of the convolutional layer is 4×4 , the step size is 2, the input channel is 1, and the output channel is 1.

$$I^{SR}\left(O_{\log_2 RT}^k\right) = C_{RT}(I^{LR}) \circ C_D^k(C_k) + C_R^k(C_{k,R}) \oplus C_{O,D}^k(O_{k-1}). \quad (10)$$

Given the above, it can be concluded that the input of the reconstruction network at this level is the output of the reconstruction network at the upper level (except the first level, which is the input low-resolution image I^{LR}) and the output of the feature extraction network at this level. The whole network is extracted and reconstructed step by step. Finally, after $\log_2 RT$ times of extraction and reconstruction, high-resolution image I^{SR} of target multiple can be obtained, namely, $O_{\log_2 RT}^k$. Although the network is divided into two

parts, the whole network is trained and optimized together. Compared with the upsampling method, the proposed method can greatly reduce the time and space consumption of the algorithm.

4. Experiments

4.1. Evaluation Metrics. The evaluation index of image superresolution reconstruction algorithm can be divided into subjective evaluation index and objective evaluation index. The subjective evaluation index is mainly scored by the assessor after comprehensive consideration of all aspects of the image based on experience. This method is intuitive and can well reflect the visual quality of the reconstructed image. However, this method is affected by many factors. In order to avoid the situation of different evaluation scores for the same reconstructed quality image, we need objective evaluation indexes that can reflect subjective evaluation and not be transferred by subjective will. However, there is no objective evaluation index that can completely describe subjective evaluation, which can only reflect to a certain extent. In this paper, the two most commonly used metrics are selected for detailed introduction, respectively, peak signal to noise ratio (PSNR) and structural similarity index measurement (SSIM). Both of these two evaluation metrics use mathematical formulas to describe the similarity of two images, which are not interfered by subjective factors and are more scientific than subjective impression scoring.

PSNR evaluates the image quality by calculating the ratio of the error of the corresponding pixel point between image X to be evaluated and the standard reference image Y to the maximum pixel of the image. PSNR is intuitive, simple, and easy to understand with a small amount of calculation, but it is a pixel-by-pixel evaluation index. It is error sensitive and often differs from human visual perception because it does not take into account the visual characteristics of human eyes, such as brightness structure and other information [35]. Unlike PSNR, which only considers pixel difference between images, SSIM describes the similarity between image X to be evaluated and the standard reference image Y from three aspects: brightness, contrast, and structural information. The value range of SSIM is $(0, 1]$. The larger the value is, the better the quality of reconstructed image is. SSIM takes into account various factors to evaluate the quality of an image, including brightness, contrast, and structural information of the image, so as to measure the quality of reconstructed high-resolution images in a more comprehensive way and the evaluation results are more consistent with human vision.

4.2. Datasets. The training set used in this paper consists of 292 pictures with different resolutions, scenes, and types. However, the data amount of 292 images is not enough to support CNN learning, so data enhancement is needed for the training set. The process of enhancement is as follows. (i) The original image is image-resized using bicubic method (to save the texture, the image is only downsampling; upsampling will destroy the texture information of the image), and the scaling factor is $(1, 0.9, 0.7)$. (ii) The scaled

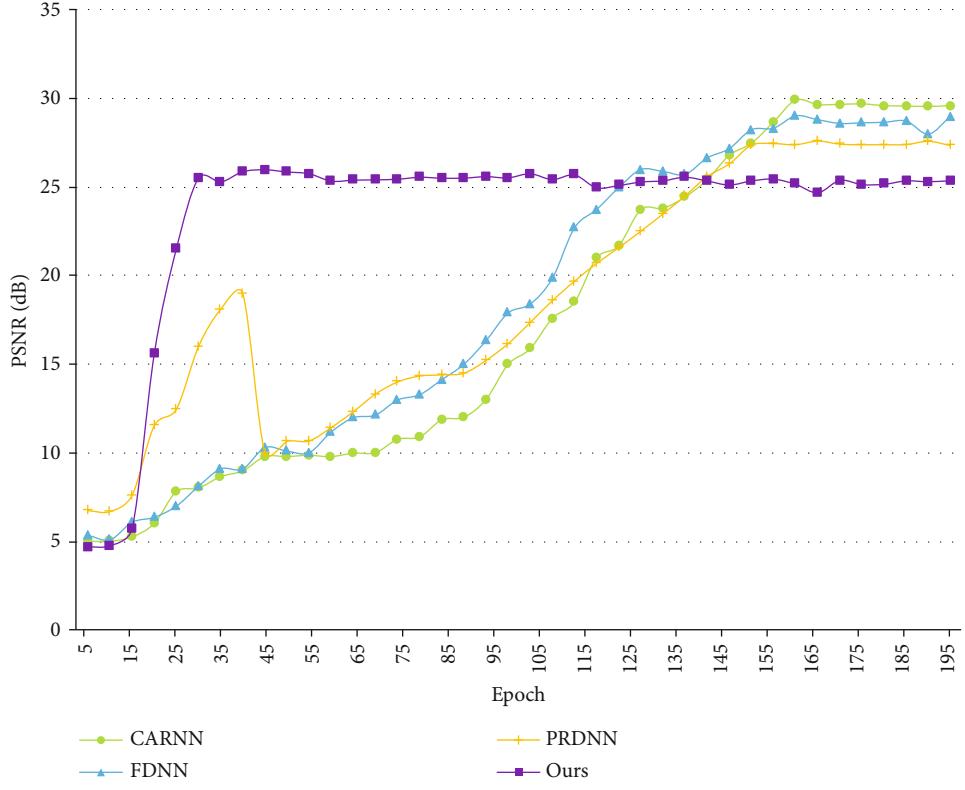


FIGURE 2: Performance comparison of PSNR.

image is conducted image-rotate with (0, 90, 180, 270) degree. (iii) Image-flip is performed on the rotated image in both horizontal and vertical directions. After image enhancement, the amount of image data (number of images) can be 48 times as much as before ($4 \times 4 \times 3$). The test sets used in this paper are Set5, Set14, Urban100, and BSD100.

4.3. Experiment Settings. The experiment was running on an Intel i7-12700KF 3.6 GHz CPU, 32 GB of RAM (3333 MHz), and NVIDIA RTX 3080 Ti, 12 GB GDDR6X GPU. $w_c = 0.35$, $b_c = 0.92$, $w_d = 0.37$, $b_d = 0.91$, $w_R^k = 0.41$, $b_R^k = 0.89$, $w_{O,R}^k = 0.42$, and $b_{O,R}^k = 0.90$, where w is set randomly and b is set according to literature [36]. The batch size is set as 64, and initial learning rate of weight is $le - 4$. Every 200 epochs, the learning rate of weights decreases by 10 times. To verify the effectiveness of superresolution, CARNN [22], FDNN [23], and PRDNN [24] were selected for performance comparison and image reconstruction.

4.4. Model Analysis. Figure 2 shows that the hybrid dilated convolution and Laplacian pyramid-based image superresolution algorithm proposed in this paper was relatively high within 150 epochs at the beginning. However, after 150 epochs, the network based on Laplacian pyramid structure had a fast convergence speed and a high PSNR value, and the learning was stable and forward. The results show that the image superresolution reconstruction based on Laplacian pyramid structure network can better learn the low-resolution image and high-resolution image mapping. It is also indicated that the step-by-step upsampling method adopted in

this paper can better extract features in different resolution spaces, alleviate the learning limitations of one-layer deconvolution when the amplification is too large, better learn the mapping from low-resolution image to high-resolution image, and obtain more high-frequency information. Moreover, a high-resolution image with sharp edges and rich texture details is obtained. As can be seen from Figure 3, the SSIM of the algorithm proposed in this paper is always at a stable level and always higher than 0.9. In contrast, the other three baselines fluctuated frequently between 0.75 and 0.92, which was unfavorable to the reconstruction process of superresolution images. This also confirms the validity of the algorithm in this paper.

To more intuitively feel the quality of image reconstruction of the model, we also show the high-resolution images reconstructed by the above compared algorithm. Four images from Set5, Set14, Urban100, and BSD100 datasets were selected, and the final experimental results are shown in Figures 4–7. Set5 dataset and Set14 dataset are low-complexity single-image superresolution datasets based on nonnegative neighborhood embedding, and the training set is used for single-image superresolution reconstruction, that is, to reconstruct high-resolution images from low-resolution images to obtain more details. Urban100 dataset has rich texture images and is generally used as a dataset for network testing, while BSD100 is a classical image dataset having 100 test images.

From Figures 4–7, it can be seen that the high-resolution image reconstructed by the algorithm proposed in this paper is more similar to the real high-resolution image, and the

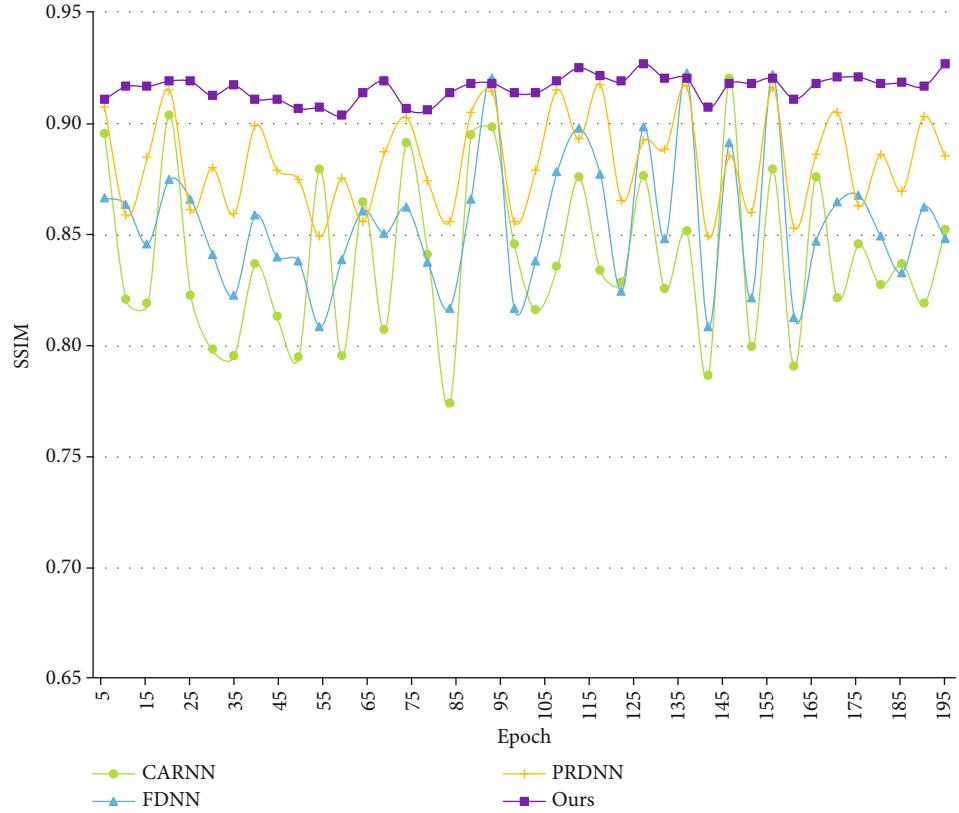


FIGURE 3: Performance comparison of SSIM.



FIGURE 4: Visual comparison of different models of Set5.

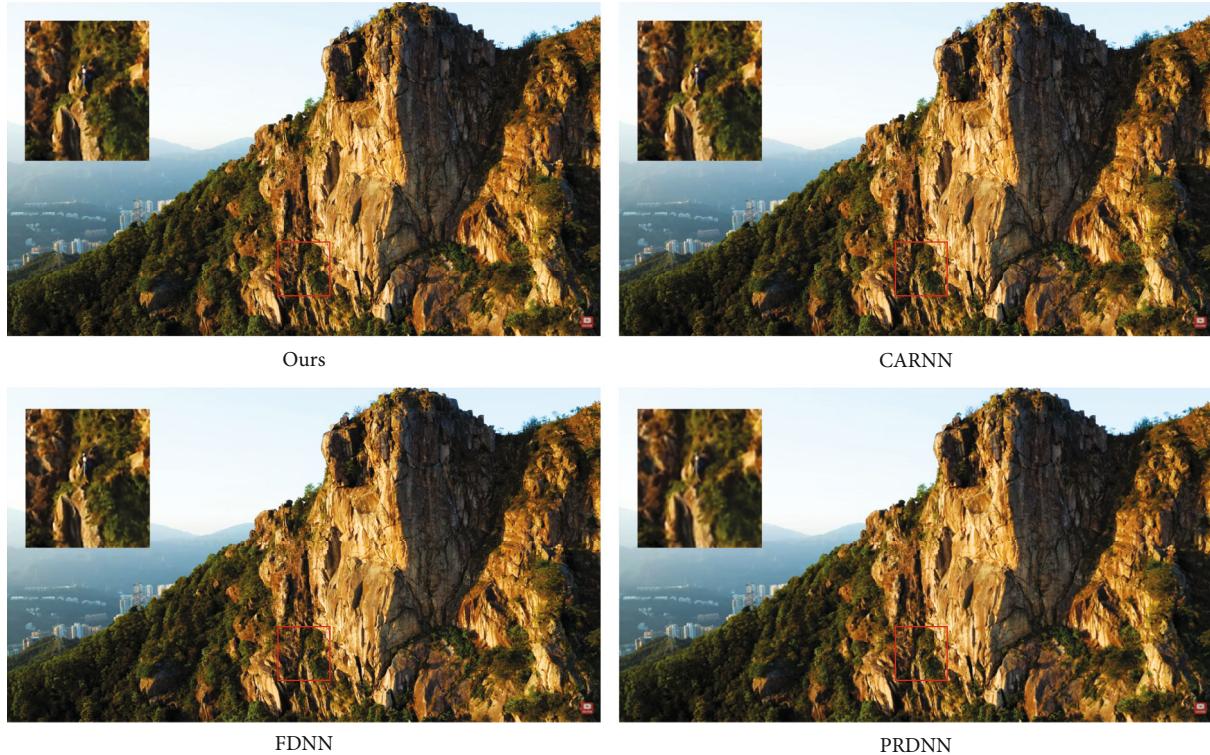


FIGURE 5: Visual comparison of different models of Set14.

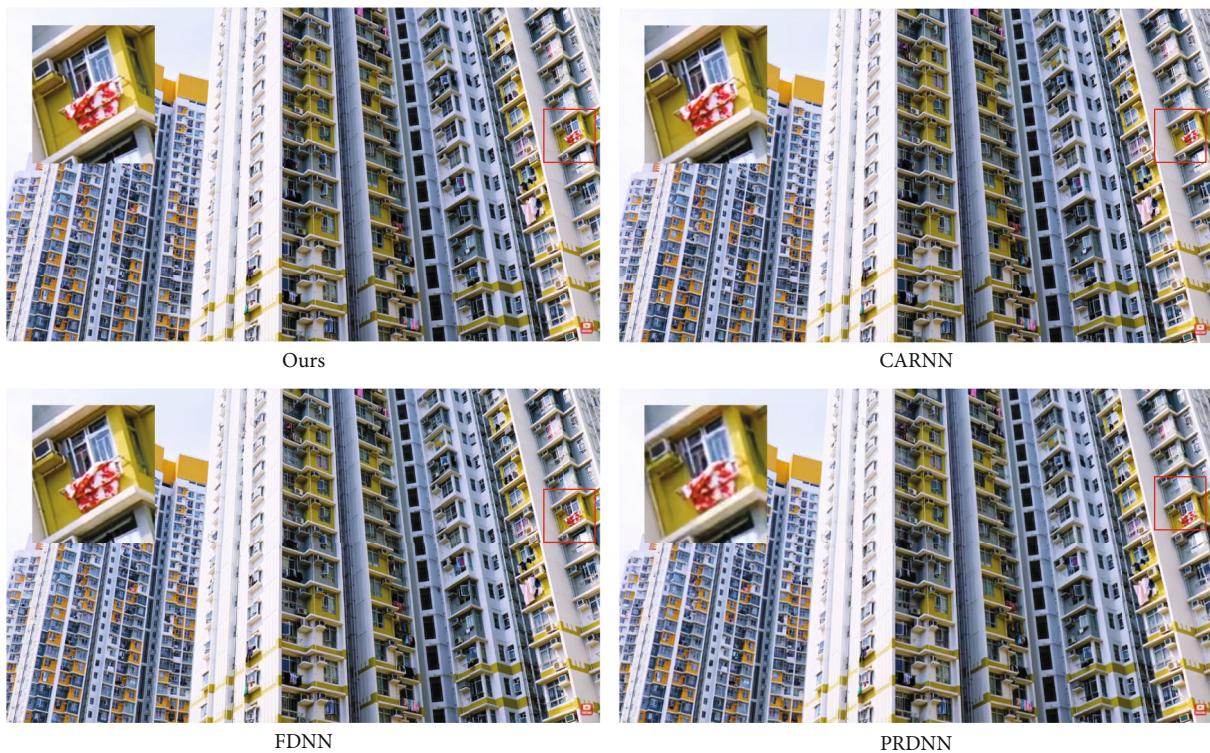


FIGURE 6: Visual comparison of different models of Urban100.

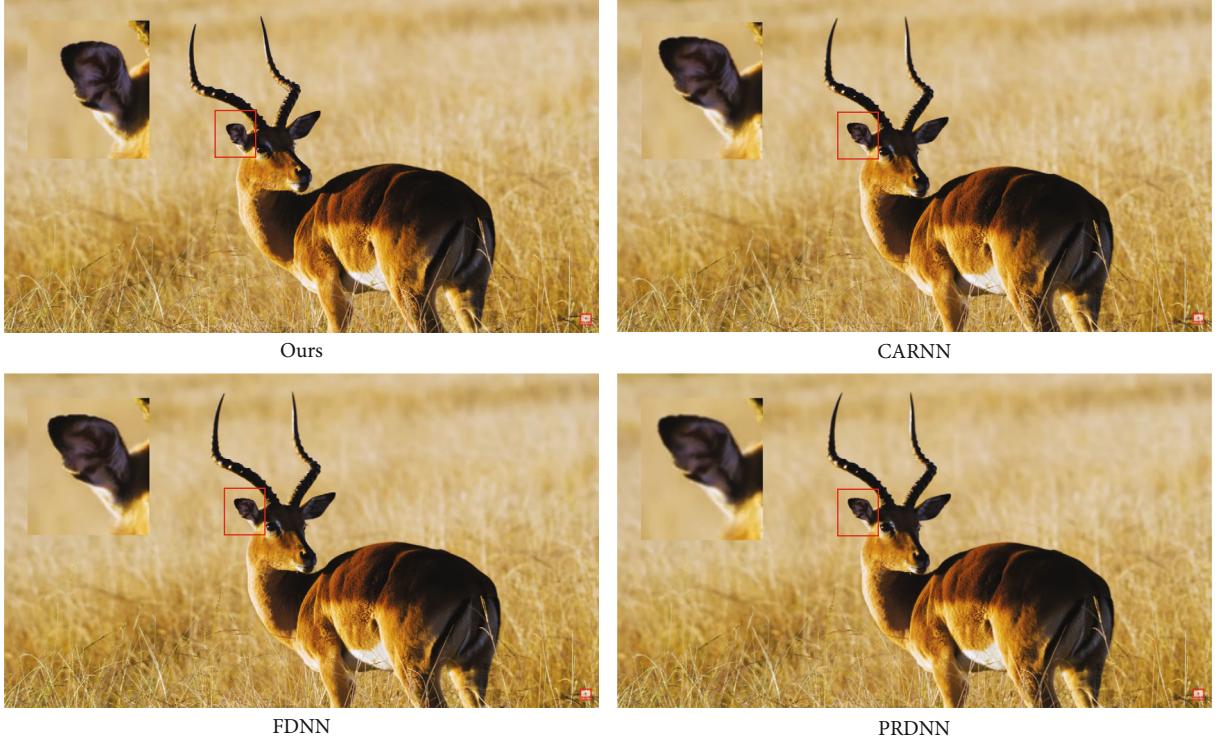


FIGURE 7: Visual comparison of different models of BSD100.

real image texture is reconstructed. Although there are flaws in performance, compared with other baselines, the performance of the algorithm proposed in this paper is the best. It can also be seen from Figure 4 that, compared with other baselines, the algorithm proposed in this paper has reconstructed the edge contour of the font better and is closer to the real high-resolution image. The experimental results show that the image superresolution algorithm based on hybrid dilated convolution and Laplacian pyramid proposed in this chapter can extract the medium- and high-frequency features of images more effectively, reduce the attenuation phenomenon when features are transmitted in the network, and reconstruct high-resolution images that are more consistent with human visual perception.

5. Conclusion

In this paper, an image superresolution algorithm based on hybrid dilated convolution and Laplacian pyramid structure is proposed. The hybrid dilated convolution module is used to extract image features, which can better expand the receptive field of the model without introducing additional parameters. In this way, the model can obtain more context information and improve the feature extraction ability of the model and will not cause grid effect. To alleviate the problems of insufficient learning ability of the upsampling layer when the reconstruction multiple is large and the large amount of calculation and long reconstruction time of the front upsampling model, the step-by-step upsampling method based on Laplacian pyramid is used to gradually

amplify the image, which balances the relationship between reconstruction time and reconstruction quality and allows the model to learn more high-frequency information. Experiments show that the proposed algorithm effectively improves the image quality.

This paper proposes a solution to the shortcomings of the existing CNN-based image superresolution reconstruction algorithm and verifies the effectiveness of the proposed method through experiments. However, there are still some problems to be solved in this field. The follow-up work will be carried out from the following aspects in the future.

- (i) For standard low-resolution and high-resolution image datasets, the existing image superresolution algorithm acquires low-resolution images by down-sampling high-resolution images from public datasets to obtain corresponding low-resolution images. However, low-resolution images acquired in this way cannot completely simulate the image degradation process, and acquired low-resolution images have similar styles. Therefore, in the future, efforts will be made to establish a complete set of low-resolution and high-resolution image data to make up for this shortcoming in the field of image superresolution reconstruction
- (ii) The existing image based on CNN superresolution algorithm is to rebuild a multiple single algorithm. When the reconstruction multiple is different, a new model has to be retrained, which is extremely inflexible. Moreover, the existing algorithms can only

- reconstruct high-resolution images with integer multiple and cannot achieve arbitrary amplification. The future will be how to realize the reconstruction of the flexible network
- (iii) Natural images are rich in prior information, while the existing CNN-based image superresolution ignores the prior information. In the future, we will focus on the exploration of prior information and make full use of the prior information of images to reconstruct high-resolution images with richer high-frequency information
- ## Data Availability
- All data used to support the findings of the study is included within this paper.
- ## Conflicts of Interest
- The authors declare no conflicts of interest in this paper.
- ## Acknowledgments
- This work was supported by Youth Program Research Projects of Liaoning Higher Education Institutions (Grant No. Lqn202014).
- ## References
- [1] T. Chen, L. Peng, J. Yang, G. Cong, and G. Li, "Evolutionary game of multi-subjects in live streaming and governance strategies based on social preference theory during the COVID-19 pandemic," *Mathematics*, vol. 9, no. 21, p. 2743, 2021.
 - [2] Y. Zhao and F. Bacao, "How does gender moderate customer intention of shopping via live-streaming apps during the COVID-19 pandemic lockdown period?," *International Journal of Environmental Research and Public Health*, vol. 18, no. 24, 2021.
 - [3] M. Zhang, F. Qin, G. Wang, and C. Luo, "The impact of live video streaming on online purchase intention," *Service Industries Journal*, vol. 40, no. 9-10, pp. 656–681, 2020.
 - [4] M. Martini, "Online distant witnessing and live-streaming activism: emerging differences in the activation of networked publics," *New Media & Society*, vol. 20, no. 11, pp. 4035–4055, 2018.
 - [5] V. M. Baskaran, Y. C. Chang, J. Loo, and K. Wong, "Design and implementation of parallel video combiner architecture for multi-user video conferencing at ultra-high definition resolution," *Multimedia Tools and Applications*, vol. 74, no. 17, pp. 6589–6622, 2015.
 - [6] H. Kim, S. Ahn, W. Kim, and S. Lee, "Visual preference assessment on ultra-high-definition images," *IEEE Transactions on Broadcasting*, vol. 62, no. 4, pp. 757–769, 2016.
 - [7] F. Zhu, Y. Ning, X. Chen, Y. Zhao, and Y. Gang, "On removing potential redundant constraints for SVOR learning," *Applied Soft Computing*, vol. 102, 2021.
 - [8] C. Zhang, Y. Niu, T. Wu, and X. Li, "Color image super-resolution and enhancement with inter-channel details at trivial cost," *Journal of Computer Science and Technology*, vol. 35, no. 4, pp. 889–899, 2020.
 - [9] S. Pang, Z. Chen, and F. Yin, "Lightweight multi-scale aggregated residual attention networks for image super-resolution," *Multimedia Tools and Applications*, vol. 81, 2022.
 - [10] P. Liu, Y. Hong, and Y. Liu, "Deep differential convolutional network for single image super-resolution," *IEEE Access*, vol. 7, pp. 37555–37564, 2019.
 - [11] X. Jiang, N. Wang, J. Xin, X. Yang, Y. Yu, and X. Gao, "Image super-resolution via multi-view information fusion networks," *Neurocomputing*, vol. 402, pp. 29–37, 2020.
 - [12] C. Liu, X. Sun, C. Chen et al., "Multi-scale residual hierarchical dense networks for single image super-resolution," *IEEE Access*, vol. 7, pp. 60572–60583, 2019.
 - [13] F. Hao, T. Zhang, L. Zhao, and Y. Tang, "Efficient residual attention network for single image super-resolution," *Applied Intelligence*, vol. 52, no. 1, pp. 652–661, 2022.
 - [14] F. Zhu, J. Gao, J. Yang, and N. Ye, "Neighborhood linear discriminant analysis," *Pattern Recognition*, vol. 123, 2022.
 - [15] P. V. Arun, K. M. Buddhiraju, A. Porwal, and J. Chanussot, "CNN-based super-resolution of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6106–6121, 2020.
 - [16] J. Liu, Y. Liu, H. Wu, J. Wang, X. Li, and C. Zhang, "Single image super-resolution using feature adaptive learning and global structure sparsity," *Signal Processing*, vol. 188, 2021.
 - [17] H. Li, K. M. Lam, and M. Wang, "Image super-resolution via feature-augmented random forest," *Signal Processing: Image Communication*, vol. 72, pp. 25–34, 2019.
 - [18] Y. Zhang, F. Shi, J. Cheng, L. Wang, P. 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.
 - [19] X. Li, G. Cao, Y. Zhang, A. Shafique, and P. Fu, "Combining synthesis sparse with analysis sparse for single image super-resolution," *Signal Processing-Image Communication*, vol. 83, p. 115805, 2020.
 - [20] Y. Yang, "Research on the single image super-resolution method based on sparse Bayesian estimation," *Cluster Computing-the Journal of Networks Software Tools and Applications*, vol. 22, no. S1, pp. 1505–1513, 2019.
 - [21] D. Mishra, B. Majhi, S. Bakshi, A. K. Sangaiah, and P. K. Sa, "Single image super resolution for texture images through neighbor embedding," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 8337–8366, 2020.
 - [22] J. T. Hsu, C. H. Kuo, and D. W. Chen, "Image super-resolution using capsule neural networks," *IEEE Access*, vol. 8, pp. 9751–9759, 2020.
 - [23] F. Cao, K. Yao, and J. Liang, "Deconvolutional neural network for image super-resolution," *Neural Networks*, vol. 132, pp. 394–404, 2020.
 - [24] J. Zhang, S. Liu, Y. Peng, and J. Li, "Satellite image super-resolution based on progressive residual deep neural network," *Journal of Applied Remote Sensing*, vol. 14, no. 3, 2020.
 - [25] Z. Hui, J. Li, X. Gao, and X. Wang, "Progressive perception-oriented network for single image super-resolution," *Information Sciences*, vol. 546, pp. 769–786, 2021.
 - [26] T. Yang, S. Jia, and H. Ma, "Research on the application of super resolution reconstruction algorithm for underwater image," *CMC-Computers Materials & Continua*, vol. 63, no. 3, pp. 1249–1258, 2020.

- [27] X. Wang, Z. Wang, X. He, C. Ren, and P. Karn, "Super-resolution of compressed images using enhanced attention network," *Journal of Electronic Imaging*, vol. 30, no. 3, 2021.
- [28] J. Zhu, C. Tan, J. Yang, G. Yang, and P. Lio, "Arbitrary scale super-resolution for medical images," *International Journal of Neural Systems*, vol. 31, no. 10, 2021.
- [29] M. Zarepoor, P. Shamsolmoali, and J. Yang, "Learning depth super-resolution by using multi-scale convolutional neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 2, pp. 1773–1783, 2019.
- [30] Y. Zhang, Y. Sun, and S. Liu, "Deformable and residual convolutional network for image super-resolution," *Applied Intelligence*, vol. 52, no. 1, pp. 295–304, 2022.
- [31] Z. Zhang, X. Wang, and C. Jung, "DCSR: dilated convolutions for single image super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1625–1635, 2019.
- [32] J. Lv, X. Wang, K. Ren, M. Huang, and K. Li, "ACO-inspired information-centric networking routing mechanism," *Computer Networks*, vol. 126, pp. 200–217, 2017.
- [33] X. Lin, J. Wu, S. Mumtaz, S. Garg, J. Li, and M. Guizani, "Blockchain-based on-demand computing resource trading in IoV-assisted smart city," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1373–1385, 2022.
- [34] L. Ma, M. Huang, S. Yang, R. Wang, and X. Wang, "An adaptive localized decision variable analysis approach to large-scale multiobjective and many-objective optimization," *IEEE Transactions on Cybernetics*, vol. 52, pp. 1–13, 2021.
- [35] H. Zheng, Z. Shi, C. Zhou, M. Haardt, and J. Chen, "Coupled coarray tensor CPD for DOA estimation with coprime L-shaped array," *IEEE Signal Processing Letters*, vol. 28, pp. 1545–1549, 2021.
- [36] W. Sun, X. Zhang, and X. He, "Lightweight image classifier using dilated and depthwise separable convolutions," *Journal of Cloud Computing-Advances Systems and Applications*, vol. 9, no. 1, 2020.