일반논문 (Regular Paper)
방송공학회논문지 제24권 제4호, 2019년 7월 (JBE Vol. 24, No. 4, July 2019)
https://doi.org/10.5909/JBE.2019.24.4.633
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

고밀도 스킵 연결을 통한 재귀 잔차 구조를 이용한 단일 이미지 초해상도 기법

진 건[a)], 정 제 창[a)‡]

Single Image Super-resolution using Recursive Residual Architecture Via Dense Skip Connections

Chen Jian[a)] and Jechang Jeong[a)‡]

요 약

최근, 단일 이미지 초해상도 복원 기법(super-resolution)에서 컨볼루션 신경망 모델은 매우 성공적이다. 잔여 학습 기법은 컨볼루션 신경망 훈련의 안전성과 성능을 향상시킬 수 있다. 본 논문은 저해상도 입력 이미지에서 고해상도 목표 이미지로 비선형 매핑 학습을 위해 고밀도 스킵 연결(dense skip-connection)을 통한 재귀 잔차 구조를 이용한 단일 이미지 초해상도 복원 기법을 제안한다. 제안하는 단일 이미지 초해상도 복원 기법은 고밀도 스킵 연결 방식을 통해 재귀 잔차 학습 방법을 채택해서 깊은 신경망에서 학습이 어려운 문제를 완화하고 더 쉽게 최적화하기 위해 신경망 안에 불필요한 레이어를 제거한다. 제안하는 방법은 매우 깊은 신경망의 사라지는 변화도(vanishing gradient) 문제를 완화할 뿐만 아니고 낮은 복잡성으로 뛰어난 성능을 얻음으로써 단일 이미지 초해상도 복원 기법의 성능을 향상시킨다. 실험 결과를 통해 제안하는 알고리듬이 기존의 알고리듬 보다 결과가 더 우수함을 보인다.

Abstract

Recently, the convolution neural network (CNN) model at a single image super-resolution (SISR) have been very successful. The residual learning method can improve training stability and network performance in CNN. In this paper, we propose a SISR using recursive residual network architecture by introducing dense skip connections for learning nonlinear mapping from low-resolution input image to high-resolution target image. The proposed SISR method adopts a method of the recursive residual learning to mitigate the difficulty of the deep network training and remove unnecessary modules for easier to optimize in CNN layers because of the concise and compact recursive network via dense skip connection method. The proposed method not only alleviates the vanishing-gradient problem of a very deep network, but also get the outstanding performance with low complexity of neural network, which allows the neural network to perform training, thereby exhibiting improved performance of SISR method.

Keywords : single image super-resolution, convolution neural network, residual learning, dense skip connection

Ⅰ. Introduction

From monochrome television broadcasting, color television broadcasting to High-Definition television (HDTV) broadcasting, the development of broadcasting has undergone three times of evolution. In recent years, the television broadcasting has developed into Ultra-High-Definition television (UHDTV) with the high-resolution of 3840 2160 so that it can provide a stereoscopic effect in the visual field to satisfy the human visual system. Moreover, as the functions of cameras in smartphones developing rapidly, users can obtain high-quality pictures without digital cameras. However, due to the limitations of technological development, existing legacy video or image is mostly Full-HD video or low-resolution image. Therefore, an image processing technique that names super-resolution (SR) method [1,2,3] have been proposed that is used to reconstruct a high-resolution (HR) image from low-resolution (LR) image.

The single image super-resolution (SISR) aims at recovering an HR image from a single LR image. SISR is a class of technique that enhances the resolution of an imaging system, and the SISR algorithm has been widely used in various computer vision tasks as the core technology of general image processing that has outstanding functions and performance, such as security and surveillance imaging, medical imaging [4,5], satellite imaging, face recognition [6,7], which more image details are required on demand.

a) 한양대학교 전자통신컴퓨터공학부(Department of Electronic Computer Engineering, Hanyang University)
‡ Corresponding Author : 정제창 (Jechang Jeong)
E-mail: jjeong@hanyang.ac.kr
Tel: +82-2-2220-4370
ORCID: https://orcid.org/0000-0002-3759-3116
· Manuscript received March 29, 2019; Revised June 3, 2019; Accepted June 26, 2019.

Many SISR methods have been studied in the computer vision community. Early methods including image interpolation are the process of transferring the image from a given resolution to high resolution. There are numerous classical methods such as nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation [8]. And Lanczos resampling [9] method. Currently, learning methods are widely used to model a mapping from LR to HR patches. For example, Neighbor embedding [10,11] methods interpolate the patch subspace and external example-based method [12] that can be formulated for generic image super-resolution. The sparse coding [13] method use a learned compact dictionary based on sparse single representation. Lately, the SISR based on convolution neural network (CNN) [14,15,16] successfully learns the mapping from LR image to HR image with large datasets. The SISR based on the deep learning can obtain a high-quality HR image from LR image, and it possesses a splendid performance by comparing with the traditional SISR methods. However, there are large numbers of limitations on the existing SISR based on CNN, and it will generate a large number of parameters, also the computational calculation will be complex because the CNN model needs to take large dataset as the training data. For this reason, in this paper, we propose a new CNN model, the HR image can effectively reconstruct from LR image by using self-learning method of the feature maps through a study based on propose CNN model instead of solving it by formulas.

The organization of this thesis is as follows. The existing SR algorithm that based on the CNN methods and the deep learning model that is used in the proposed method are described in Section Ⅱ. The proposed algorithm which used the CNN model is proposed in Section Ⅲ. Section Ⅳ includes experimental results, and finally the proposed algorithm will be concluded in Section Ⅴ.

II. Related work

1. Image super-resolution using deep convolution networks (SRCNN)

In recent times, the CNN methods have presented an explosive popularity partially due to its success in image classification. They have also been successfully applied to other computer vision tasks. Among them, the CNN model which learns a mapping from LR to HR image patches have been proposed. Termed SRCNN. Fig. 1 shows the SRCNN model structure. The HR image can be decomposed into a low frequency information (LR image) and high frequency information (residual image or image details). In SRCNN, denote the interpolated image as Y. The goal is to recover from Y an image $F(Y)$ that is as similar as possible to ground truth image X, and learn a mapping F. The SRCNN method is one of the representative SR methods which uses 3-layers CNN model: patch extraction/ representation, non-linear mapping and reconstruction. Filter of spatial sizes 9×9, 1×1, 5×5 were used respectively. The first layer is expressed as an operation F_1:

$$F_1(Y) = \max(0, W_1 * Y + B_1) \tag{1}$$

The operation of the second layer is:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \tag{2}$$

The operation of last layer which produces the final high-resolution image is:

$$F(Y) = W_3 * F_2(Y) + B_3 \tag{3}$$

However, as the depth of the networks increasing, the performance can be significantly improved. But it also needs more data to prevent overfitting.

2. Deep residual learning for image recognition

Recently, the deep convolution neural networks [17] gain a series of breakthroughs for image classification [17,18]. But the deeper neural networks are more difficult to train, as the depth of the network increases, the training error and test error are not reduced. For this reason, in this thesis, proposed a residual learning network [19] structure to ease the training of networks are substantially deeper than previously methods and prevent the gradient vanishing problem. The purpose of residual learning is to improve performance and solving image recognition problems by deep neural network.

Fig. 2 shows the structure of neural network which uses the residual learning method. The convolution layer outputs

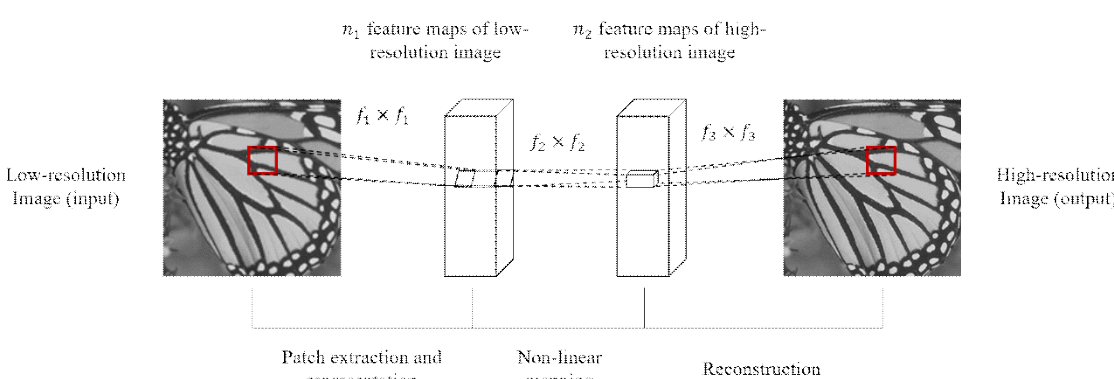

그림 1. SRCNN 모델 구조
Fig. 1. SRCNN model structure

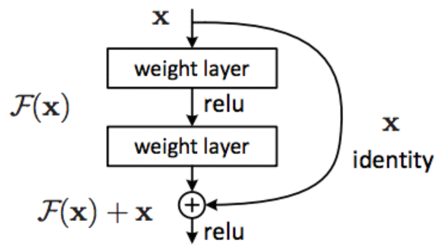

그림 2. 잔차 학습의 구조
Fig. 2. The structure of residual learning

are added to the outputs of the stacked layers, the $H(x)$ can consider as an underlying mapping to be fit by a few stacked layers, then the x can denote the inputs to the first of these layers. Let these layers approximate a residual function $F(x) = H(x) - x$. The input and output are of the same dimensions. This method added neither extra parameter nor computational complexity. The residual learning method not only in order to prevent the loss of image details, but also helps gradient flow. Formally, in this paper, the building block defined as:

$$y = F(x, \{W_i\}) + x \tag{4}$$

Here x and y are the input and output vectors of the layers considered.

3. Densely connected convolutional networks

Huang et al. [20] proposed the dense network structure [20] which composed of several dense blocks, and there are several dense layers in one dense block. The dense convolutional network (DenseNet) connects each layer to every other layer in a feed-forward fashion. The neural network structure of a dense network consists of two rules. First, the features of the current dense layer are connected to the feature of the next dense layer only on the same dense block, consequently, the *l*-th layer receives the feature-maps of all preceding layers, $x_0, \ldots, x_{l-1}$, the operation of the layer defined as:

$$x_l = H_l([x_0, x_1, \ldots x_{l-1}]) \tag{5}$$

Here, $[x_0, x_1, \ldots x_{l-1}]$ means concatenation of features before the l-th dense layer, $H_l(.)$ is the batch normalization [21], Leaky ReLU [17], and the operation of convolutional. Second, the dense layer produces features after the convolution operation, where the is referred to as the growth rate. DenseNets have several compelling advantages: prevent the vanishing-gradient problem, strengthen feature propagation, and reducing the number of parameters. An effective in very deep CNNs may potentially improve the SR reconstruction performance.

In this paper, we used the recursive residual network instead of the multi stacked layers in network, and by introducing the dense network method which concatenation of layer blocks.

Ⅲ. Proposed algorithm

1. Recursive residual block

In this section, we present the technical of our proposed CNN model. Recently, residual networks [19] exhibit excellent performance in computer vision problems. For this reason, we used the residual network in our model. Specifically, we adopt the global residual learning for our CNN model, and in our recursive residual blocks, we introduce residual learning into our recursive residual blocks by constructing the recursive block structure, and utilize every recursive residual block to concatenation. Now, we present details about our proposed recursive residual block in Fig. 3. We construct convolutional layers for each recursive residual block, the layers have the same number of 64 filter size with the kernel size as 3×3. We remove the batch normalization layers from our proposed recursive residual block because the Nah et al. [22] presented in their image deblurring task. The batch normalization layers normalize features get rid of range flexibility from networks, so it is better to remove the batch normalization layers. Therefore, Formulate proposed recursive residual block as:

$$H_{output} = H[F(H^{u-1}, W^u) + H^0] \tag{6}$$

Where $u = 1, 2, \mathrm{U}$, U is the number of residual units in a recursive residual block. H^{u-1} and $F(H^{u-1}, W^u)$ are the input and output of the u-th residual unit, F denotes the residual function. The H^0 is the result of the first convolutional layer in our recursive residual block. Different from residual learning network, we utilized the convolutional layer at the output of recursive residual block for the stability of the network, the H[ ] is the operation of the last convolutional layer.

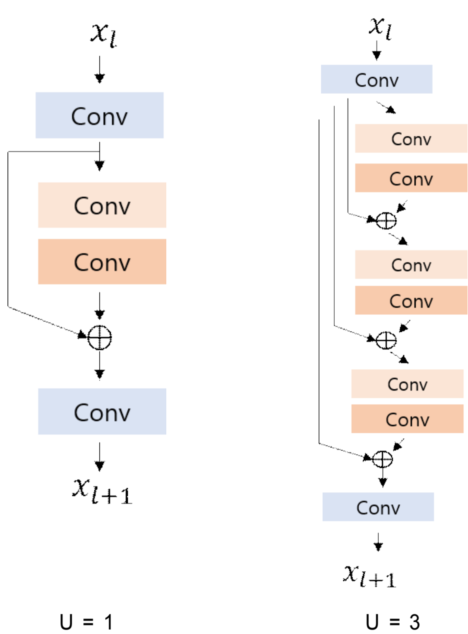

그림 3. 재귀 잔차 블록의 구조. U는 재귀 블록 내 잔차 단위 개수이다.
Fig. 3. Structures of our recursive residual block. U means number of residual units in the recursive block

2. Network structure

In this thesis, the proposed CNN model adopted the residual learning network and dense skip connection architecture. As shown in Fig 4, the proposed CNN model mainly consists three parts: the convolution layer for learning low-level features, the blocks of recursive residual architecture for learning the high-level features, and the reconstruction layer for generating the high-resolution (HR) output. The first layer adopted the convolutional layer which the feature map of 64 with 3 × 3 kernel size. After applied a convolutional layer to the input low-resolution (LR) images for learning low-level features, a set of recursive residual blocks are adopted for learning high-level features. However, as CNN become increasingly deep, the problem of vanishing gradient hampers the training of networks. Therefore, in this thesis, introduced the dense skip connection architecture. Fusion of every layer features through dense skip connection architecture can further boost the reconstruction performance. Simply stacked several recursive residual blocks, and followed by introducing dense skip connections between the recursive residual blocks. In this thesis, a total of four recursive residual blocks were used. If the large number of feature maps are directly into the last layer, it will significantly increase the computational cost and the model size, for this reason, it is necessary to reduce the number of feature maps, in the proposed CNN model, set the bottleneck layer that the feature maps is reduced to 256 with 1 × 1 kernel size in order to keep model compactness, and improve the computational efficiency. The feature maps in the HR space are used to generate HR images via last layer. Finally, the last layer which is reconstruction layer with 3 × 3 kernel size and one channel of output. Moreover, because the SR output is vastly similar to the input, in this thesis, adopted the global residual learning method into the proposed CNN model to estimate residual image from the input and output of the neural network. In this thesis, every convolutional layer is followed by PReLU activation function layer for nonlinear mapping except the last layer.

3. Training

We analyze the effective patch size and stride in the network model. Finally, by considering both the performance and network complexities, we split input images into 41× 41 patches with the strides of 41. In image classification application [17,18], performance of the network can be enhanced by training data augmentation. The most commonly used methods include cropping, flipping, and rotation [23]. In SR, rotation and flipping of input images lead to the better results. Therefore, we applied rotation and flipping on the training image patches. Finally, we used the input patches that is rotated by 90º, and flipped for data augmentation. As the network activation function, the PReLU function is used. For the training, given a training dataset $\{x^{(i)},y^{(i)}\}_{i=1}^{N}$, where N is the number of training patches and $y^{(i)}$ is the ground truth HR patch and the $x^{(i)}$ is the LR patch, the mean squared error (MSE) is used as the loss function, which is defined as:

$$L = \frac{1}{N}\sum_{i=1}^{N} \| r - f(x) \|^2 \tag{7}$$

where x denotes interpolated LR image and y means the corresponding HR image. Our goal is to learn a model f that predicts values $\hat{y} = f(x)$. We define the residual image $r = y - x$. We train all of the CNN models during 30 epochs. Training is carried out by using Adam optimizer, and we set the learning rate to $1e-4$, and then decreased by a factor of 10 every 10 epochs.

Ⅳ. Experiment results

We used the windows 10, 64-bit operating system, intel core i5-3570 in MATLAB R2018a and PyCharm Community edition 2018 with GPU 1060 for this experiment. We subjectively compared the conventional interpolate patch subspace methods, classical SR methods based on CNN and our proposed method on several datasets. For the network training, we use 291 image set [24] as training data. In addition, data augmentation that is rotation and flip used. And Set5 [11] image set as validation dataset. Finally, the Set5, and Set14 [25] are used as test dataset to evaluate the performance of the proposed network. In this paper, the SISR method utilizes the proposed CNN model that to reconstruct the HR image from the LR image. In this thesis, we compare our method with four SR methods: A+ [12], SRCNN [14], FSRCNN [15], and VDSR [16]. In our proposed method, the number of residual units

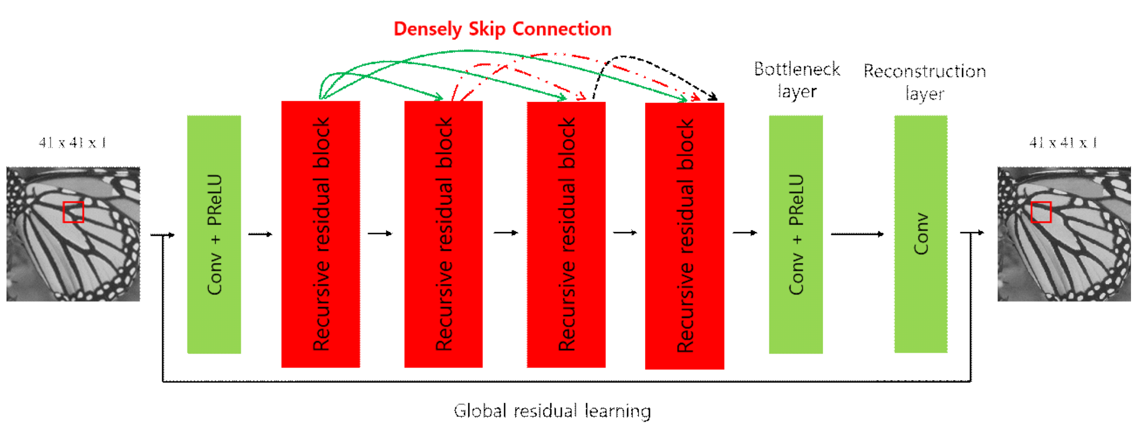

그림 4. 본 논문에서 제안된 CNN 모델의 구조
Fig. 4. The architecture of our proposed CNN model

in the recursive residual block is 1, when the number of residual units in recursive residual block set as 3, we denote our method as our+. The performance of the result image should be evaluated qualitatively and quantitatively. The primary method used when qualitatively evaluating images is typically to evaluate directly with the human eye. Subjective image quality evaluates is one of the effective qualitative evaluation methods because the human eye is very sensitive to image defects. However, the result images evaluate directly with human eye may not be objective. Therefore, the quality of result images also should be evaluated qualitatively. The Fig. 5, and Fig. 6 shows the results of subjective image quality comparison when used conventional methods and proposed method. Comparing the result in Fig. 5 and Fig. 6, it is clear that the SR using the proposed method have more effective effect than which using the conventional method. Moreover, we adopt the peak signal-to-noise ratio (PSNR) and Structural Similarity (SSIM) to obtain the quantitative of the test datasets images and compare the result with classical method results.

In this thesis, the CNN model uses recursive residual architecture by introducing the dense skip connections. The information used for reconstruction that the receptive filed sets as the 41 41 for our network. The training uses batches of size 16. Momentum and weight decay parameters are set to 0.9 and 0.0001. Moreover, we used the one residual unit and three residual units in our recursive residual block of our network, and compared the results. We summarize the results of the super-resolution based on interpolation method, learning paradigms method, classical deep learning method, and our proposed method. in the Table Ⅰ. According to the experimental results, the results of our work have good performance in objective evaluation (PNSR and SSIM). Our CNN model is effective in easing the difficulty of training network because the model uses the residual learning between HR and LR image patches,

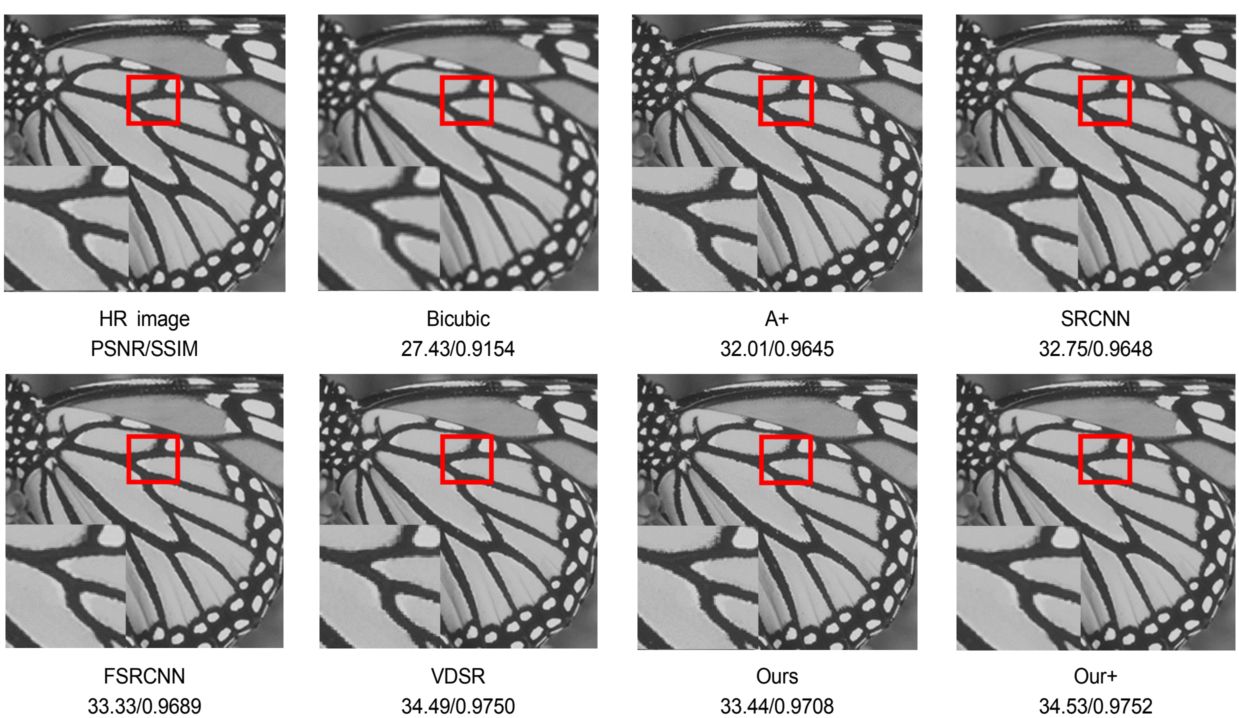

그림 5. Set5의 “butterfly” 영상의 2배 확대 실험 결과
Fig. 5. The result image of “butterfly” from Set5 with an upscaling factor=2

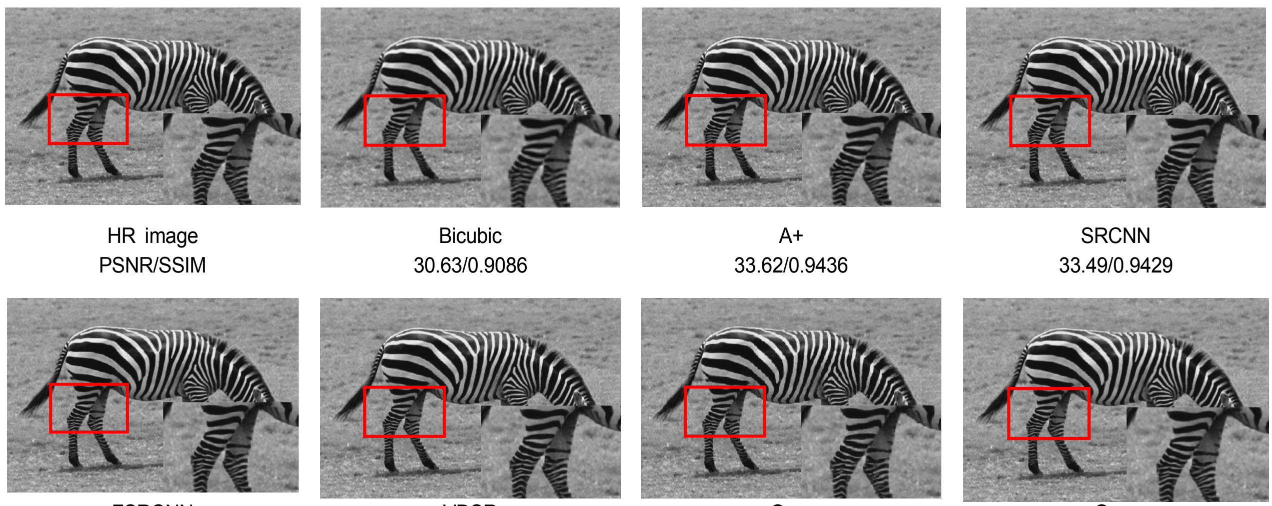

그림 6. Set14의 “zebra” 영상의 2배 확대 실험 결과
Fig. 6. The result image of “zebra” from Set14 with an upscaling factor=2

표 1. Set5, Set14의 2배, 3배, 4배 확대 결과 평균 PSNR
Table 1. Average PSNR for scale 2, 3, and 4 on datasets Set5, and Set14

Dataset	scale	Bicubic PSNR/SSIM	A+[12] PSNR/SSIM	SRCNN[14] PSNR/SSIM	FSRCNN[15] PSNR/SSIM	VDSR[16] PSNR/SSIM	Ours PSNR/SSIM	Ours+ PSNR/SSIM
Set5[11]	2	33.66/0.9299	36.54/0.9544	36.66/0.9542	37.00/0.9558	37.53/0.9587	37.23/0.9575	37.55/0.9590
	3	30.390.8682	32.58/0.9088	32.75/0.9090	33.16/0.9140	33.66/0.9213	33.45/0.9156	33.72/0.9222
	4	28.42/0.8104	30.28/0.8603	30.48/0.8628	30.71/0.8657	31.35/0.8838	30.89/0.8664	31.32/0.8842
Set14[25]	2	30.240.8688	32.28/0.9056	32.42/0.9063	32.63/0.9088	33.03/0.9124	32.85/0.9098	33.02/0.9108
	3	27.55/0.7742	29.13/0.8188	29.28/0.8209	29.43/0.8242	29.77/0.8314	29.68/0.8264	29.80/0.8321
	4	26.00/0.6675	27.32/0.7491	27.49/0.7503	27.59/0.7535	28.01/0.7674	27.64/0.7541	28.04/0.7672

comparing to the conventional methods, the proposed method also can reduce the difficulty of the training model and the residual learning method can carriers rich image details. The experimental results show that the performance of the proposed method is evaluated by objective comparison and subjective comparison with existing methods. Although the objective quality comparison results still have some inferior performance in some cases, the proposed method shows that the result of image quality evaluation is higher than that of the existing methods in most images. In the subjective quality comparison, noticeable improvement can visually be found in the result images.

Ⅴ. Conclusion

In this paper, in order to show the better performance of the super-resolution based on CNN, we propose a new network structure for SISR based on CNN method. Comparing to the conventional SISR methods, it utilizes the recursive residual network architecture via dense skip connections into the CNN model for significantly boots of the model performance. Through experiment, the results prove that using our CNN model can make result images obtain more details. Moreover, it presents the significant impacts on the results of SISR with good performance results than previous methods.

참 고 문 헌 (Reference)

[1] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International Journal of Computer Vision,* Vol.40, No.11, pp. 25-47. 2000.

[2] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," *Proceeding of International Conference on Computer Vision*, 2009.

[3] M. Irani, and S. Peleg, "Improving resolution by image registration," *Graphical models and Image Processing*, pp.231-239, 1991.

[4] S. Peled, and Y. L. Cun, "Super-resolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* , Vol.45, No.1, pp. 29–35, 2001.

[5] W. Shi, J. Caballero, C. Ledig, X. Zhuang, W. Bai, K. Bhatia, A. Marvao, T. Dawes, D. ORegan, and D. Rueckert, "Cardiac image super-resolution with global correspondence using multi-atlas patchmatch," *Proceeding of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp.9–16. 2013.

[6] W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Transaction on Image Processing*, Vol.21, No.1, pp.327–340, July 2011.

[7] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE Transaction on Image Processing*, Vol.12, No.5, pp.597–606, 2003.

[8] D. Y. Han, "Comparison of commonly used image interpolation methods," *Proceeding of International Conference on Computer Science and Electronics Engineering*, Mar 2013.

[9] C. E. Duchon, "Lanczos filtering in one and two dimensions," Journal of Applied Meteorology, pp.1016-1022, 1979.

[10] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," *Proceeding of the IEEE conference on Computer Vision and Pattern Recognition (CVPR),* 2004.

[11] C. G. Marco Bevilacqua, Aline Roumy, and M. L. A. Morel, "Low complecity single-image super-resolution based on nonnegative neighbor embedding," *Proceeding of IEEE British Machine Vision Conference,* 2012.

[12] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," *Proceeding of Asian Conference on Computer Vision (ACCV)*, Nov 2014.

[13] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transaction on Image Processing*, Vol.19, No.11, Nov 2010.

[14] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolution networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol.38, No.2, Feb 2016.

[15] C. Dong, C. C. Loy, K. He, and X. Tang, "Accelerating the super-resolution convolutional neural network." *Proceeding of European Conference on Computer Vision (ECCV),* 2016.

[16] J. Kim, J. Kwon, and K. Mu. Lee, "Accurate image super-resolution using very deep convolution networks," *Proceeding of the IEEE conference on Computer Vision and Pattern Recognition (CVPR),* 2016.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet calssification with deep convolutiuon neural network," *Neural Information Processing Systems (NIPS)*, 2012.

[18] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. L. Cun, "Overfeat: Intergrated recognition, localization and detection using convolution networks," *Proceeding of IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, Dec 2013.

[19] K. He, X. Y. Zhang, S. Q. Ren, and J. S, "Deep residual learning for image recognition," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Dec 2015.

[20] G. H, Z. L, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *Proceeding of IEEE conference on Computer Vision and Pattern Recognition (CVPR),* 2017.

[21] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *Procceding of International Conference on Machine Learning(ICML)*, pp 1097-1105, 2012.

[22] Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," *Proceeding of IEEE conference on Computer Vision and Pattern Recognition (CVPR),* Dec. 2016.

[23] R. Timofte, R. Rothe, and L. Cool, "Seven ways to improve example-based single image super resolution," *Proceeding of IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1865–1873, Jun 2016.

[24] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceeding of IEEE conference on Computer Vision and Pattern Recognition (CVPR),* 2001.

[25] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," *Proceeding of International Conference on Curves and Surfaces.* pp.711-730, 2012.

저 자 소 개

진 건

- 2013년 9월 : 대전대학교 전자통신공학과 학사
- 2017년 9월 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 석사과정
- ORCID : https://orcid.org/0000-0002-3831-222
- 주관심분야 : 영상처리, 딥러닝

정 제 창

- 1980년 2월 : 서울대학교 전자공학과 학사
- 1982년 2월 : KAIST 전기전자공학과 석사
- 1990년 : 미국 미시간대학 전기공학과 공학박사
- 1980년 ~ 1986년 : KBS 기술연구소 선임연구원 (디지털 및 뉴미디어 연구)
- 1990년 ~ 1991년 : 미국 미시간대학 전자컴퓨터공학부 연구교수 (영상 및 신호처리 연구)
- 1991년 ~ 1995년 : 삼성전자 HDTV 연구개발 담당 수석연구원
- 1995년 ~ 현재 : 한양대학교 전자컴퓨터통신공학과 교수 (영상통신 및 신호처리 연구실)
- ORCID : https://orcid.org/0000-0002-3759-3116
- 주관심분야 : 영상처리, 영상압축