Efficient Module Based Single Image Super Resolution for Multiple Problems

Dongwon Park[†] Kwanyoung Kim[†] Se Young Chun
School of ECE, Ulsan National Institute of Science and Technology, 44919, Ulsan, South Korea
dong1@unist.ac.kr, cubeyoung@unist.ac.kr, sychun@unist.ac.kr

Abstract

Example based single image super resolution (SR) is a fundamental task in computer vision. It is challenging, but recently, there have been significant performance improvements using deep learning approaches. In this article, we propose efficient module based single image SR networks (EMBSR) and tackle multiple SR problems in NTIRE 2018 SR challenge by recycling trained networks. Our proposed EMBSR allowed us to reduce training time with effectively deeper networks, to use modular ensemble for improved performance, and to separate subproblems for better performance. We also proposed EDSR-PP, an improved version of previous ESDR by incorporating pyramid pooling so that global as well as local context information can be utilized. Lastly, we proposed a novel denoising / deblurring residual convolutional network (DnResNet) using residual block and batch normalization. Our proposed EMBSR with DnResNet and EDSR-PP demonstrated that multiple SR problems can be tackled efficiently and effectively by winning the 2nd place for Track 2 (×4 SR with mild adverse condition) and the 3rd place for Track 3 (×4 SR with difficult adverse condition). Our proposed method with EDSR-PP also achieved the ninth place for Track 1 (×8 SR) with the fastest run time among top nine teams.

1. Introduction

The goal of image super resolution (SR) problem is to design an algorithm to map from low resolution (LR) image(s) to a high resolution (HR) image. Conventional SR was to yield a HR image from a multiple of LR images (*e.g.*, video) considering a number of LR image degradation operators such as blurring and noise. This type of SR has been well studied [16] and fundamental performance limit for it has been analyzed [17]. In medical imaging, generating a high signal-to-noise ratio (SNR) image from a multiple of low SNR images has also been well studied with similar model based approaches as conventional SR problems [3].

† Equal contribution

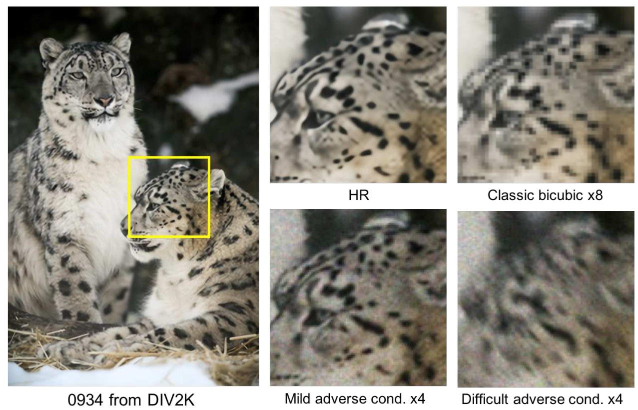

Figure 1: An example of given images for NTIRE 2018 challenge on super-resolution [18]. The goal of challenge was to design algorithms to map from low resolution images (Classic bicubic ×8, Mild adverse condition ×4 or Difficult adverse condition ×4) to a high resolution image (HR).

In contrast, a SR problem using a single LR image is challenging since high frequency information in a HR image is lost or degraded due to aliasing during sampling process. Because there was no effective way to extrapolate high frequency information, single image SR problem was usually considered as an interpolation problem [16]. An example based SR method was proposed based on Bayesian belief propagation [6] and a patch based SR method was proposed by combining a conventional multiple image based SR and an example based SR [7].

Deep neural network has applied to many image processing and computer vision problems and has shown significantly improved performance over conventional methods [12]. There have been several works on single image SR problems and several deep neural networks were proposed such as SRCNN [5], VDSR [10], SRResNet [13], and EDSR [15]. EDSR achieved state-of-the-art performance for ×4 SR problem in terms of peak SNR (PSNR) and structural similarity index (SSIM) and won the NTIRE 2017 challenge [1] for SR problems. NTIRE 2017 SR consisted of two Tracks for known (bicubic) and unknown blurs and for each Track, there were three different downsampling rates ($\times 2$, $\times 3$, $\times 4$). EDSR outperformed other previous networks including SRResNet for all public dataset including DIV2K, NTIRE 2017's new dataset [15].

NTIRE 2018 SR challenge [18] is more challenging than its previous challenge by having 4 Tracks: Track 1 with $\times 8$ SR problem and with known blur (bicubic) and Tracks 2, 3, 4 with $\times 4$ SR problems and with mild to severe noise and/or unknown blur. Figure 1 shows examples of given images for the ground truth and for Tracks 1, 2, 3 that our team participated in. Mild noise was observed in given $\times 4$ downsampled images for Track 2 and similar level of noise was observed in given $\times 4$ downsampled images, but with relatively severe unknown blur for Track 3.

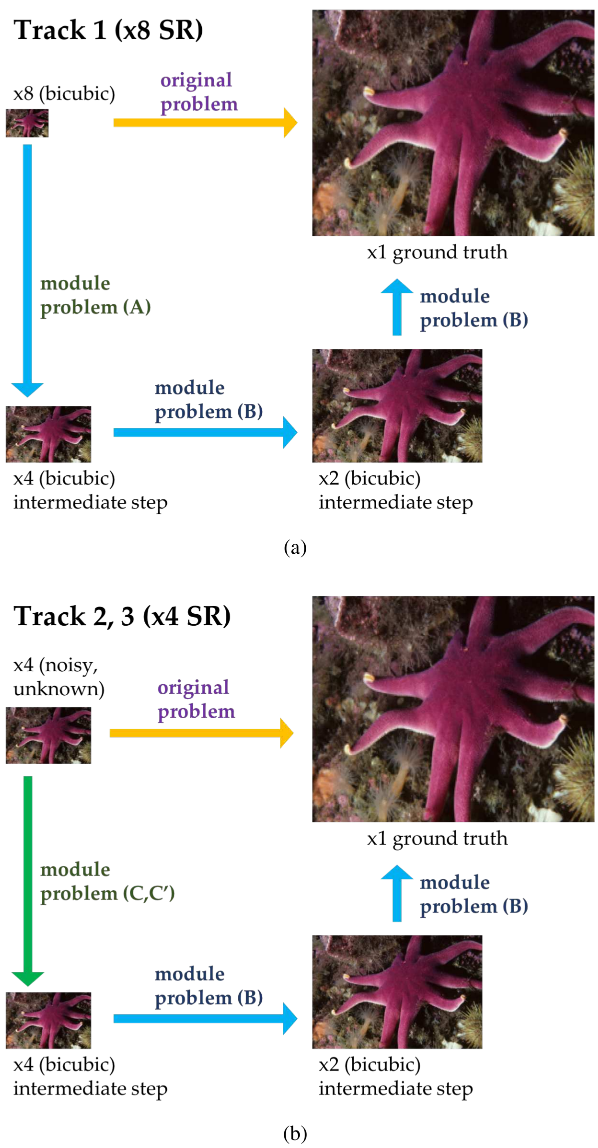

Figure 2: (a) Module based approach for Track 1 SR problem. (b) Module based approach for Tracks 2, 3 SR problems. The solution for module problem (B) can be efficiently recycled among different SR problems in all Tracks.

In this article, we propose an efficient module based approach for tackling multiple SR problems in Tracks 1, 2, 3 of NTIRE 2018 SR challenge [18]. We decomposed the original problems in Tracks 1, 2, 3 into subproblems as shown in Figures 2a (Track 1) and 2b (Tracks 2, 3), identified state-of-the-art methods for subproblems as baselines, and efficiently recycled trained deep neural networks for subproblems among all problems in different Tracks. Utilizing intermediate goals for $\times 8$ SR is not new [11] and solving multiple problems together for efficiency is not a new concept [20]. This approach could also be sub-optimal in terms of the overall cost function optimization. However, our proposed method is different from previous works in 1) module based training scheme to save training time for entire networks for Tracks 1, 2, 3 by recycling and to use effectively deeper convolutional networks with more feature map channels in the midst of limited computation and memory resource, in 2) ensemble output of each module for each subproblem to improve the performance further without increasing network complexity, and in 3) separating the problem of SR (increasing the resolution) from the problem of denoising and deblurring (Tracks 2, 3).

We also proposed new deep neural networks to improve the performance for subproblems. For SR problems in module problems (A) and (B) shown in Figures 2a and 2b, EDSR [15] was chosen as our baseline network. In this article, we proposed EDSR-PP by adding pyramid pooling layers [22] to EDSR for further performance improvement with DIV2K dataset. For denoising and deblurring problems in module problem (C, C′) as illustrated in Figure 2b, we adopt DnCNN [21], one of the state-of-the-art methods for denoising and deblurring problem, as our baseline network. We proposed a novel denoising and deblurring network called DnResNet based on residual block structure [8] and showed significant performance improvement over the baseline DnCNN.

Our models were trained using DIV2K training dataset [1] and were evaluated with DVI2K validation and test dataset. In NTIRE 2018 SR challenge, our proposed methods won the 2nd place (out of 18 teams) for Track 2 and the 3rd place (out of 18 teams) for Track 3 with EMBSR using EDSR-PP and DnResNet (team name: BMIPL_UNIST). This demonstrated that our proposed module based approach can efficiently and effectively solve multiple problems. Our EMBSR method with EDSR-PP also achieved the ninth place (out of 24 teams) for Track 1 with the fastest run time among top nine teams. Here is the summary of this article's contributions:

- Modular approach for efficient training with effectively deeper network, improved performance with modular ensemble, and novel problem decomposition.

- EDSR-PP: improved EDSR with pyramid pooling.
- DnResNet: novel architecture for denoising / deblurring based on residual block.

2. Related Works

Deep learning based super resolution. Dong *et al.* used convolutional neural network (CNN) for SR problem (SRCNN) and achieved significant improvement in performance over other conventional non-deep leaning based methods [5]. An LR image is upscaled using bicubic interpolation and then CNN was applied to restore HR details. Soon after, Kim *et al.* proposed a deep neural network using residual learning (VDSR) and showed improved PSNR performance over SRCNN [10]. In this method, CNN was trained not to yield a HR image, but a residual image for the difference between an interpolated LR image and the ground truth HR image. VDSR also used a deeper CNN network than SRCNN. Lai *et al.* proposed a Laplacian pyramid super resolution network (LapSRN) that combines multiple models and uses progressive reconstruction from $\times 8$ to $\times 4$ to $\times 2$ to HR ($\times 1$) [11]. Legit *et al.* proposed SRResNet using residual blocks [8] to significantly increase the size of the receptive field and to include local context information so that state-of-the-art performance for $\times 4$ SR problem can be obtained in terms of PSNR and SSIM [13]. SRGAN was also proposed with the same network structure as SRResNet, but with different training based on a discriminator network. SRGAN yielded visually pleasing outputs while PSNR of SRGAN was lower than that of SRResNet since SRResNet yielded an average of many possible outputs while SRGAN yielded one of many possible outputs.

Recently, Lim *et al.* won the NTIRE 2017 challenge [1] for SR problems using so-called EDSR (Enhanced Deep Super-Resolution network) that enhanced SRResNet by eliminating batch normalization and by stacking deeper layers (residual blocks from 16 to 32, filter channels from 64 to 256) [15]. EDSR also used L1 loss instead of L2 loss for better PSNR. NTIRE 2017 SR consisted of two Tracks for known (bicubic) and unknown blurs and for each Track, there were three different downsampling rates ($\times 2$, $\times 3$, $\times 4$). EDSR won the 1st place for NTIRE 2017 by outperforming SRResNet for all public dataset including DIV2K, NTIRE 2017's new dataset [15].

Deep learning based denoising and deblurring. Patch based denoising methods yielded superior denoising results compared to conventional denoising techniques [4], but they are usually slow in computation and have so called rare patch issue so that these are less effective for unique patterns in an image. Recently, there have been several attempts to outperform patch based denoisers such as BM3D using deep learning based approaches. Jain and Seung demonstrated that denoising is possible using CNN [9]. Burger *et al.* proposed a multi layer perceptron based denoiser and showed that it is challenging, but possible to obtain good denoising performance over conventional state-of-the-art methods such as BM3D [2]. Xie *et al.* proposed a deep network for denoising and inpaing [19]. Recently, Lefkimmiatis investigated a combined method of conventional non-local patch based denoiser and deep learning based denoiser [14]. Zhang *et al.* proposed a so-called DnCNN with multiple CNN blocks (similar to VDSR) to yield a residual (Gaussian noise) and to yield superior performance to other denoisers including BM3D [21]. In particular, DnCNN has greatly improved the performance of denoising and deblurring tasks with a simple deep convolution layer and residual learning.

3. Method

3.1. Modular Approach

We decomposed the original problems in NTIRE 2018 SR Tracks 1, 2, 3 [18] into subproblems as illustrated in Figures 2a (Track 1) and 2b (Tracks 2, 3) and efficiently recycled trained deep neural networks for a number of subproblems. Figure 3 illustrates our detailed network architectures for all problems in Tracks 1, 2, 3, called efficient module based super resolution (EMBSR) network. This modular approach allows us to train networks module-by-module and to efficiently recycle trained modules for mul-

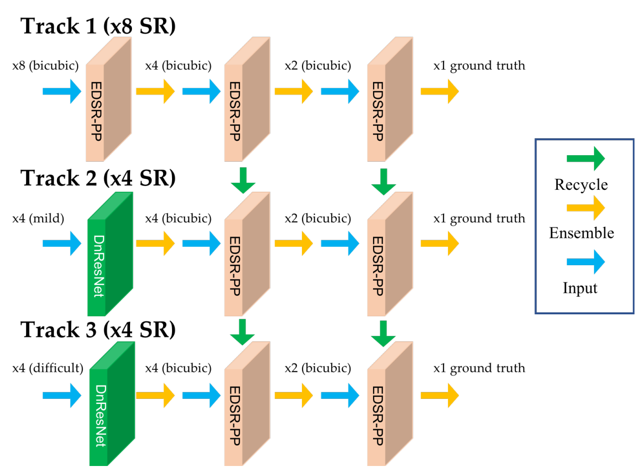

Figure 3: Modular approach for multiple SR problems. Among 9 modules, 5 modules required long training while 4 modules can be recycled with short fine tuning.

tiple SR problems (see Figure 3 to see that among 9 modules, only 5 modules require long training, while 4 modules can recycle already trained networks with relatively short fine tuning). This modular architecture also yielded effectively deeper networks with more feature map channels when limited computation and memory resource are available. Each module can generate ensemble output for each subproblem to increase the PSNR performance without increasing the complexity of networks. Lastly, modular approach allowed us to separate SR subproblems from the problem of denoising and deblurring for Tracks 2, 3. Due to this separation, significant performance improvement was achieved by utilizing optimal deep networks for different problems ($e.g.$, EDSR for SR problem and DnCNN for denoising/deblurring problem) and by aligning an input image and an intermediate target image ($\times 4$ bicubic downsampled image) for training denoiser/deblur networks.

Our EMBSR network for Track 1 ($\times 8$ bicubic) consists of three EDSR-PP networks as illustrated in the top of Figure 3. For training each module network, we downsampled ground truth images using bicubic downsampling to generate target images for each module ($\times 2$ bicubic downsampled images, $\times 4$ bicubic downsampled images). Then, all EDSR-PP modules were trained with given input $\times 8$ bicubic downsampled images and generated $\times 4$ bicubic downsampled images, input $\times 4$ bicubic downsampled images and generated $\times 2$ bicubic downsampled images, and $\times 2$ bicubic downsampled images and ground truth images. A solution for Track 1 ($\times 8$ single image SR) was created by concatenating three trained modules. Note that ensemble output is possible by having 8 variants of an input image (4 rotations $\times$ 2 left-right flips) for each neural network module. This procedure substantially improved performance. Further fine tuning is also possible. Each module is trained with perfect bicubic downsampled input images, but the ensemble output of each module contains errors from them. In EBMSR for Track 1, the second EDSR-PP module can be re-trained using ensemble output images of the first EDSR-PP module and then the third EDSR-PP module can be re-trained using ensemble output images of the re-trained second EDSR-PP module, sequentially. In our simulations, training each EDSR-PP module took about 3 days for 300 epochs and re-training each module took about 1 day for 100 epochs.

Our EMBSR network for Tracks 2, 3 is similar to the EMBSR network for Track 1, but with replacing the first EDSR-PP module with DnResNet module, as illustrated in the middle and bottom of Figure 3, respectively. The second and third "trained" EDSR-PP modules for Track 1 can be recycled in Tracks 2, 3 as shown in Figure 3 (green arrows). The first DnResNet module for tackling Track 2 can be trained using given input training data and target $\times 4$ bicubic downsampled images. Image registration between input

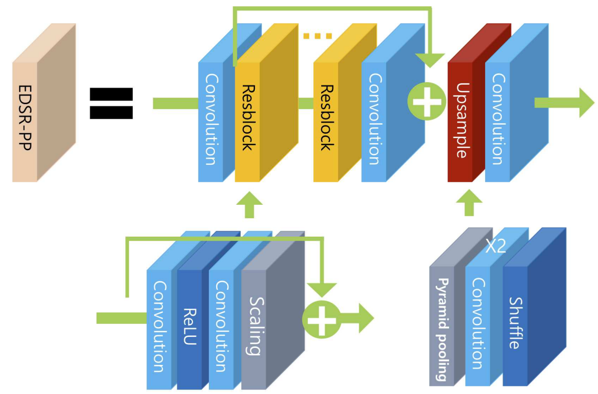

Figure 4: An illustration of our proposed EDSR-PP. Upsampling lay of the original EDSR [15] was replaced with pyramid pooling structure.

and target images using a translation motion was critical to significantly improve the performance of DnResNet as well as baseline DnCNN. For Track 3, similar approach can be applied. Then, solutions for Tracks 2, 3 can be obtained by concatenating trained DnResNet and two other trained EDSR-PP networks. Further improvement was achieved by sequentially re-training the second EDSR-PP module using ensemble output images of the first DnResNet module, and then fine tuning the third EDSR-PP module using ensemble output images of the re-trained second ESDR-PP module for both Tracks 2 and 3.

3.2. SR Module: EDSR-PP (Pyramid Pooling)

We propose a new SR network, EDSR-PP, based on a state-of-the-art SR network, EDSR [15]. EDSR-PP incorporates pyramidal pooling [22] into the upsampling layer of the original EDSR as illustrated in Figure 4. The number of residual blocks in EDSR-PP was 32 and the same network architecture was used for all SR modules in Tracks 1, 2, and 3. Typically, the receptive field size of deep learning based image processing corresponds to how much context information is included. The deeper the CNN network is, the larger the receptive field size is. However, in CNN based deep networks for image processing, this receptive field size may not be large enough to receive global context information. Pyramid pooling [22] is a recent method to resolve this issue so that both local and global context information can be utilized for image segmentation problems. We incorporated it into EDSR for SR problem. In contrast to the up-sampling layer of EDSR, pyramid pooling firstly executes average pooling and performs convolution for each of the four pyramid scales. Then, these are concatenated in the existing feature map. This process allows both local and global context information to be utilized. Four pyramid scales were used in our EDSR-PP with 1×1, 2×2,

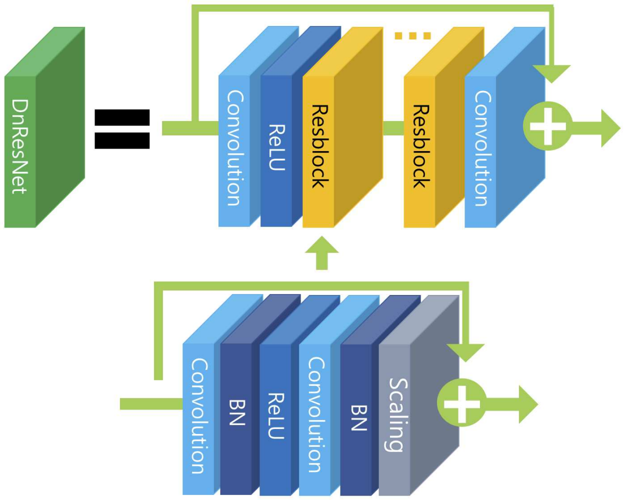

Figure 5: An illustration of our proposed DnResNet. Unlike DnCNN that uses CNN layers [21], residual blocks (Resblock) were used as a basic building block.

3×3, and 4×4 and our proposed EDSR-PP yielded better performance than EDSR.

3.3. Denoising / Deblurring Module: DnResNet

We also propose a novel denoising / deblurring network, DnResNet, based on one of the state-of-the-art methods, DnCNN [21] for denoising / deblurring problem. DnCNN uses residual learning (skip connection between input and output) and multiple convolution blocks with convolution - batch normalization - ReLU layers. Our DnResNet simply replaces all convolution blocks with our residual blocks as shown in Figure 5. Using residual blocks further increased receptive fields efficiently without concatenating more deep convolution layers. DnCNN used 64 feature map channels while our DnResNet used 128 feature map channels.

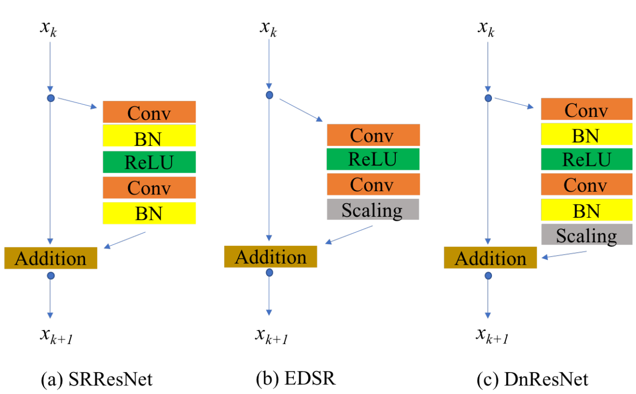

Figure 6: Comparison of residual blocks for SRResNet [13], EDSR [15], and our DnResNet.

For residual blocks, EDSR removed batch normalization layers from and added 0.1 scaling to the residual block of SRResNet as shown in Figure 6 for improved performance and numerical stability of training in SR problem. However, we found that it is advantageous to keep batch normalization layers for denoising and deblurring problems. So, we modified the residual block of EDSR by adding two batch normalization layers again. Note that our residual block is equivalent to the original residual block of SRResNet except for 0.1 residual scaling. Note also that our proposed DnResNet utilized similar residual blocks as SRResNet, but overall network architectures are quite different. Our proposed DnResNet with residual blocks outperformed DnCNN with convolution blocks for denoising and deblurring problems.

4. Experiment

4.1. Dataset

The DIV2K dataset from the NTIRE 2018 challenge was used in all simulations of this article. DIV2K is a high quality (2K resolution) image data set from the NTIRE 2017 challenge [1] . For the same ground truth HR images, $\times 8$ bicubic downsampled images were provided for Track 1, $\times 4$ downsampled images with unknown blur kernels and mild noise were provided for Track 2, and $\times 4$ downsampled images with unknown, difficult blur kernels and noise were provided for Track 3. For each track, 800 training images, 100 validation images, and 100 test images were given. In this article, we only use 10 images (801 to 810).

4.2. Training and Alignment

Training procedures are described in Section 3.1. Mini batch size was 16 and patch size was 48×48. For individual module training, 300 epochs were run with learning rates of 10^{-4} for 1 to 100 epochs and 10^{-5} for 101 to 300 epochs. It took about 3 days to run 300 epochs for each module network. Re-training learning rate was set to 10^{-5} for 100 epochs.

We found that given input images of Tracks 2 and 3 and $\times 4$ bicubic downsampled ground truth images are not well aligned. In principle, these misalignment should be taken care of by deep neural networks during training. However, aligning input and target images as much as possible helped to achieve improved performance. Given input images of Tracks 2 / 3 and $\times 4$ bicubic downsampled ground truth images were aligned using image intensity based image registration tool in MATLAB with translation motion only. Bicubic interpolation was used for sub-pixel accuracy.

4.3. DIV2K Validation Set Results

Table 1 shows performance results for DIV2K validation set, comparing various SR methods such as bicubic in-

Table 1: PSNR (dB) results of different methods for DIV2K validation data set: SRCNN [5], VDSR [10], EDSR [15], and our proposed EMBSR.

	Bicubic	SRCNN	VDSR	EDSR	**EMBSR**
$\times 2$	31.01	33.05	33.66	35.12	**35.87**
$\times 4$	26.66	27.70	28.17	29.38	**29.89**
$\times 8$	24.51	-	-	26.00	**26.22**

Table 2: Performance comparison between architectures on the DIV2K validation set (PSNR in dB).

DnCNN [21]	DnResNet	DnCNN [21] (aligned)	DnResNet (aligned)
21.005	25.359	29.439	**30.281**

terpolation, SRCNN [5], VDSR [10], EDSR [15] and our proposed EMBSR. Our EDSR-PP based EMBSR method yielded improved PSNR results for SR problems with different scales ($\times 2$, $\times 4$, and $\times 8$) over other methods including state-of-the-art EDSR method. Note that EMBSR is equivalent to EDSR-PP for $\times 2$. Thus, EDSR-PP outperformed EDSR by 0.75 dB for $\times 2$ and this result demonstrated that our proposed SR module, EDSR-PP, yielded state-of-the-art SR performance.

Table 2 showed that our proposed DnResNet outperformed current state-of-the-art denoising / deblurring method, DnCNN [21], with both misaligned and aligned data set. It seems that aligning given input and target images was critical to achieve high performance in denoising and deblurring. Trained EDSR-PP modules and DnResNet modules can be used to tackle multiple SR problems in the multiple tracks of NTIRE 2018 SR challenge.

4.4. Results of NTIRE 2018 SR Challenge

We have submitted enhanced images of DIV2K test data set to NTIRE 2018 SR challenge, Tracks 1, 2, and 3 [18]. Table 3 shows PSNR, SSIM and run time results for the top nine teams including our team (BMIPL_UNIST) using our proposed EMBSR method. Our team won the ninth place out of 24 teams with PSNR 25.331, SSIM 0.7026, and run time 2.52 sec. Note that PSNR difference between the 1st place and ours was 0.124 dB and SSIM difference was 0.0062, but we achieved these results with the fastest run time among all top nine teams. Figure 7 shows qualitative results for bicubic interpolation, EDSR, and our EMBSR. Both EDSR and EMBSR yielded similarly good results, but EMBSR yielded higher PSNR than EDSR with slightly sharper images for some examples (see 0820 $\times$ 8 from DIV2K results).

Table 3: Results of NTIRE 2018 SR challenge, Track 1, $\times 8$ bicubic downsampling (PSNR in dB).

Method	**PSNR**	**SSIM**	**Run Time**
1st method	25.455	0.7088	50
2nd method	25.433	0.7067	20
3rd method	25.428	0.7055	6.75
4th method	25.415	0.7068	11.65
5th method	25.360	0.7031	7.31
6th method	25.356	0.7037	6.99
7th method	25.347	0.7023	5.03
8th method	25.338	0.7037	14.52
Ours	**25.331**	**0.7026**	**2.52**

Table 4: Results of NTIRE 2018 SR challenge, Track 2, $\times 4$ unknown downsampling with mild blur and noise (PSNR in dB).

Method	**PSNR**	**SSIM**
1st method	24.238	0.6186
Ours	**24.106**	**0.6124**
3rd method	24.028	0.6108

Table 5: Results of NTIRE 2018 SR challenge, Track 3, $\times 4$ unknown downsampling with difficult blur and noise (PSNR in dB).

Method	**PSNR**	**SSIM**
1st method	22.887	0.5580
2nd method	22.690	0.5458
Ours	**22.569**	**0.5420**

Our proposed EMBSR methods achieved excellent performance in Tracks 2 and 3. Table 4 shows PSNR and SSIM results for the top three teams including our team (BMIPL_UNIST) for Track 2, unknown $\times 4$ downsampling with image degradation due to mild blur and noise. Our team won the 2nd place out of 18 teams with PSNR 24.106 and SSIM 0.6124 in Track 2. Figure 8 shows qualitative results for bicubic interpolation, EDSR, and our EMBSR. Our EMBSR yielded significantly better image quality than EDSR quantitatively (Table 4) and qualitatively (Figure 8).

Table 5 shows PSNR and SSIM results for the top three teams including our team (BMIPL_UNIST) for Track 3, unknown $\times 4$ downsampling with image degradation due to difficult blur and noise. Our team won the 3rd place out of 18 teams with PSNR 22.569 and SSIM 0.5420 in Track 3. Figure 9 shows qualitative results for bicubic interpolation, EDSR, and our EMBSR. EDSR does not seem to deal with multiple problems (SR, denoising, deblurring) well while our EMBSR efficiently tacked SR problem with multiple sources of image degradations. It seems that modular approach allows to use appropriate networks for different problems for improved performance.

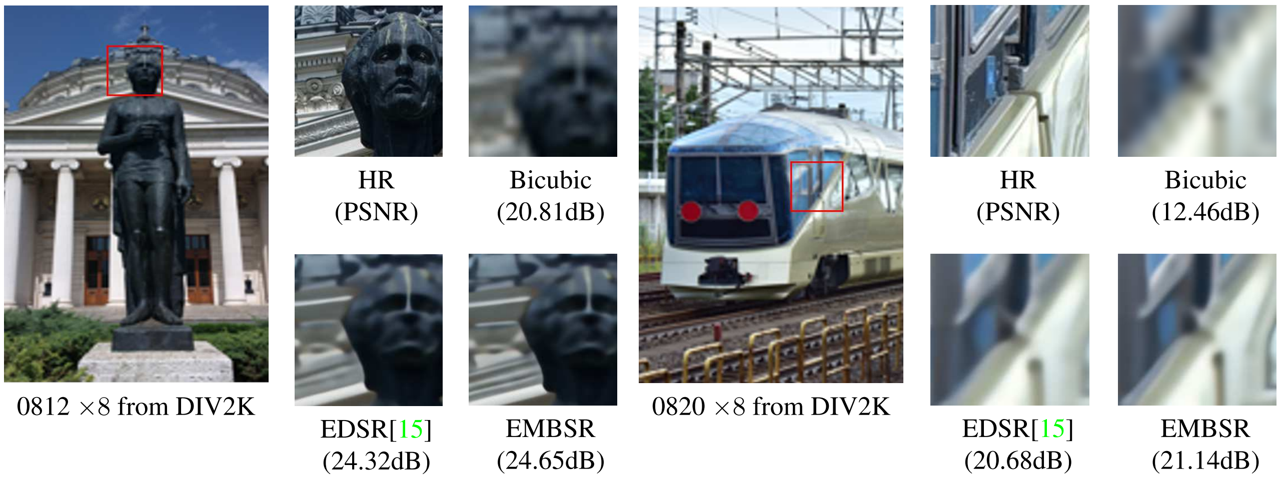

Figure 7: Results of Track 1 in NTIRE 2018 SR challenge (bicubic downsampling ×8). Our EMBSR yielded better PSNR and slightly sharper images than EDSR.

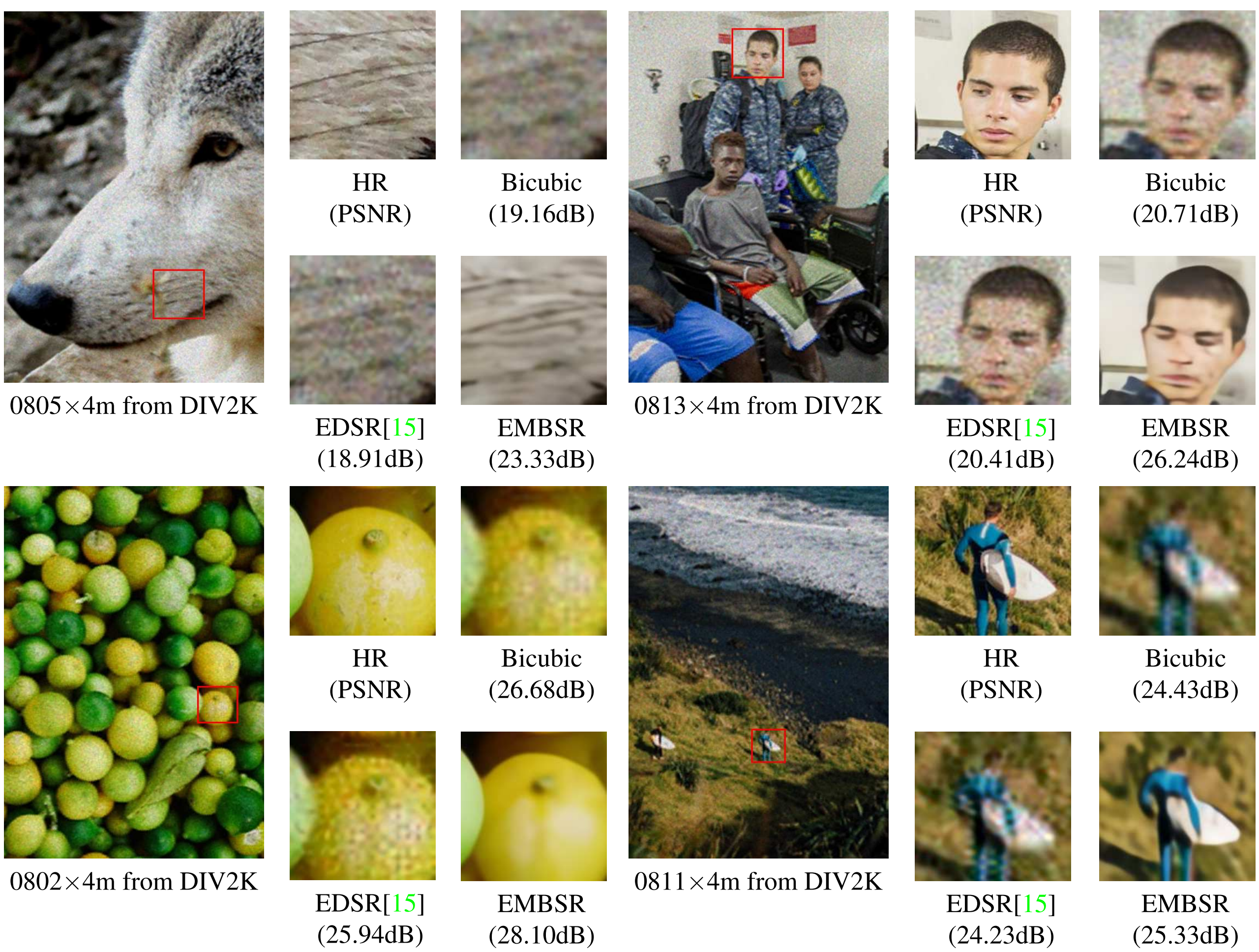

Figure 8: Results of Track 2 in NTIRE 2018 SR challenge (unknown downsampling ×4 with mild blur and noise). Our EMBSR yielded superior PSNR and image quality to EDSR and efficiently tacked SR problem with mild image degradation.

Figure 9: Results of Track 3 in NTIRE 2018 SR challenge (unknown downsampling $\times 4$ with mild blur and noise). Our EMBSR yielded superior PSNR and image quality to EDSR. EDSR does not seem to deal with multiple problems (SR, denoising, deblurring) well while our EMBSR efficiently tacked SR problem with multiple sources of image degradation.

5. Conclusion

We proposed EMBSR using modular approaches with EDSR-PP for SR and DnResNet for denoising / deblurring. Modular approach allowed us to train our networks efficiently for multiple SR problems by recycling trained networks, to use modular ensemble for improved performance, and to deal with multiple sources of image degradation efficiently. We also proposed EDSR-PP, an improved version of previous ESDR by incorporating pyramid pooling so that global as well as local context information can be utilized. Lastly, we proposed a novel denoising / deblurring residual convolutional network (DnResNet) using our residual blocks based on DnCNN. The effectiveness of our proposed methods for multiple SR problems with mixed image degradation sources was demonstrated with NTIRE 2018 SR challenge by winning the 2nd place of Track 2, the 3rd place of Track 3, and the ninth place of Track 1 with the fastest run time.

Acknowledgment

This work was partly supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(NRF-2017R1D1A1B05035810), by the Technology Innovation Program or Industrial Strategic Technology Development Program (10077533, Development of robotic manipulation algorithm for grasping/assembling with the machine learning using visual and tactile sensing information) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea), and by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. R0190-15-2054, Development of Personal Identification Technology based on Biomedical Signals to Avoid Identity Theft).

References

[1] E. Agustsson and R. Timofte. NTIRE 2017 Challenge on Single Image Super-Resolution - Dataset and Study. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, pages 1122–1131, 2017. 1, 2, 3, 5

[2] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: Can plain neural networks compete with BM3D? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2392–2399, 2012. 3

[3] S. Y. Chun and J. A. Fessler. Noise properties of motion-compensated tomographic image reconstruction methods. *IEEE transactions on medical imaging*, 32(2):141–152, Feb. 2013. 1

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, Aug. 2007. 3

[5] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a Deep Convolutional Network for Image Super-Resolution. In *European Conference on Computer Vision (ECCV)*, pages 184–199, 2014. 1, 3, 6

[6] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, Oct. 2000. 1

[7] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *IEEE International Conference on Computer Vision (ICCV)*, pages 349–356, 2009. 1

[8] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. 2, 3

[9] V. Jain and S. Seung. Natural image denoising with convolutional networks. In *Advances in Neural Information Processing Systems 21 (NIPS)*, pages 769–776, 2009. 3

[10] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1646–1654, 2016. 1, 3, 6

[11] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 624–632, 2017. 2, 3

[12] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. 1

[13] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 105–114, 2017. 1, 3, 5

[14] S. Lefkimmiatis. Non-local Color Image Denoising with Convolutional Neural Networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5882–5891, 2017. 3

[15] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced Deep Residual Networks for Single Image Super-Resolution. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1132–1140. IEEE, 2017. 1, 2, 3, 4, 5, 6, 7, 8

[16] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, May 2003. 1

[17] D. Robinson and P. Milanfar. Statistical performance analysis of super-resolution. *IEEE Transactions on Image Processing*, 15(6):1413–1428, May 2006. 1

[18] R. Timofte, S. Gu, J. Wu, L. Van Gool, L. Zhang, M.-H. Yang, et al. Ntire 2018 challenge on single image super-resolution: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018. 1, 2, 3, 6

[19] J. Xie, L. Xu, and E. Chen. Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems 25 (NIPS)*, pages 341–349, 2012. 3

[20] K. Zhang, D. Tao, X. Gao, X. Li, and Z. Xiong. Learning Multiple Linear Mappings for Efficient Single Image Super-Resolution. *IEEE Transactions on Image Processing*, 24(3):846–861, Jan. 2015. 2

[21] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning Deep CNN Denoiser Prior for Image Restoration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3929–3938, 2017. 2, 3, 5, 6

[22] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. Pyramid scene parsing network. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2881–2890, 2017. 2, 4