



Edge Assisted Asymmetric Convolution Network for MR Image Super-Resolution

Wanliang Wang^(✉), Fangsen Xing, Jiacheng Chen, and Hangyao Tu

College of Computer Science and Technology, Zhejiang University of Technology,
Hangzhou 310023, China
zjutwwl@zjut.edu.cn

Abstract. High-resolution magnetic resonance (MR) imaging is beneficial for accurate disease diagnosis and subsequent analysis. Currently, the single image super-resolution (SR) technique is an effective and less costly alternative technique to improve the spatial resolution of MR images. Structural information in MR images is crucial during clinical diagnosis, but it is often ignored by existing deep learning MR image SR technique. Consequently, we propose edge assisted feature extraction block (EAFEB), which can efficiently extract the content and edge features from low-resolution (LR) images, allowing the network to focus on both content and geometric structure. To fully utilize the features extracted by EAFEB, an asymmetric convolutional group (ACG) is proposed, which can balance structural feature preservation and content feature extraction. Moreover, we design a novel contextual spatial attention (CSA) method to facilitate the network focus on critical information. Experiment results in various MR image sequences, including T1, T2, and PD, show that our Edge Assisted Asymmetric Convolution Network (EAACN) has superior results relative to recent leading SR models.

Keywords: Magnetic Resonance Imaging · Super-Resolution · Edge Assisted · Asymmetric Convolution · Contextual Spatial Attention

1 Introduction

Image super-resolution (SR) [1] is a process of extrapolating a high-resolution (HR) image from one or more low-resolution (LR) images. It has ability to recover image details and is now widely used in medical images [2], remote sensing [3], surveillance, and security [4]. HR Magnetic resonance (MR) images with clear details are difficult to obtain directly from medical instruments due to limitations in instrumentation, scanning time, body motion, and interference from noise during imaging. Specifically, HR MR images are essential for disease diagnosis and facilitate intelligent analysis, such as detection, registration, and segmentation [5]. Therefore, SR techniques are critical to MR images.

Deep learning-based SR methods have gradually been applied to MR image SR in recent years, which are more flexible and have better evaluation metrics as well as visual performance than traditional MR SR methods. The existing deep learning MR image SR has been studied primarily in the depth and width

of the network. In particular, the number of pixels, image types, and shapes in MR images is much lower than in natural images, and deep networks are likely to be ineffective, over-fitted, and unable to obtain a clear edge structure.

According to the mechanism of the human visual system (HVS), the human eye is most sensitive to edge information compared to other components in medical images. Besides, MR images have large background regions with little information and target regions containing complex tissue textures. The edge-based approach can divide background and target regions, reconstruct the structural features of target regions well. As a result, We specifically design an edge feature extraction branch in edge assisted feature extraction block (EAFEB) to obtain edge features. To fully utilize the features extracted by EAFEB, we propose an asymmetric convolutional group (ACG), which can balance structural feature preservation and content feature extraction. Meanwhile, to alleviate the burden of the model and allow it to focus on processing critical information, we incorporate contextual spatial attention (CSA) in nonlinear mapping processing.

In summary, our contributions are four-fold:

(1) Edge assisted feature extraction block (EAFEB) is proposed to efficiently extract the content and edge features from low-resolution images, allowing the network to focus on both content and geometric structure with minimal computational effort. (2) A novel Asymmetric Convolutional Group (ACG) is designed to preserve structural features while further extracting content features. (3) We propose Contextual Spatial Attention (CSA), which encompasses a broader perceptual domain and allows the network to focus on critical information. (4) Qualitative and quantitative experimental results in MR image datasets demonstrate the superiority of our model over other advanced models.

2 Related Work

2.1 Edge Assisted Image Super-Resolution

Edge information has been used in many previous SR tasks. Yang et al. [6] introduce a deep edge-guided recurrent residual network to progressively recover high-frequency details. A soft edge-assisted network proposed by Fang et al. [7], integrate image edge prior knowledge into the model. Ma et al. [8] designed a structure-preserving branch to alleviate structural distortions in the trunk branch. Pablo et al. [9] propose a set of one-layer edge SR architectures that have superior SR quality at the same speed as bicubic interpolation. Edge information is crucial information in images, as well as for MR images. Therefore, we design an edge feature extraction branch and also implement preservation of feature structure in Nonlinear mapping, details will be described in Sect. 3.

2.2 MR Image Super-Resolution

MR is a safe, radiation-free medical imaging technique that highlights multiple details of tissue through multiple sequences of imaging. Nevertheless, MR images suffer from low spatial resolution and artifacts. Recently, advanced methods based on deep learning [2, 10] are used for MR image SR. For example,

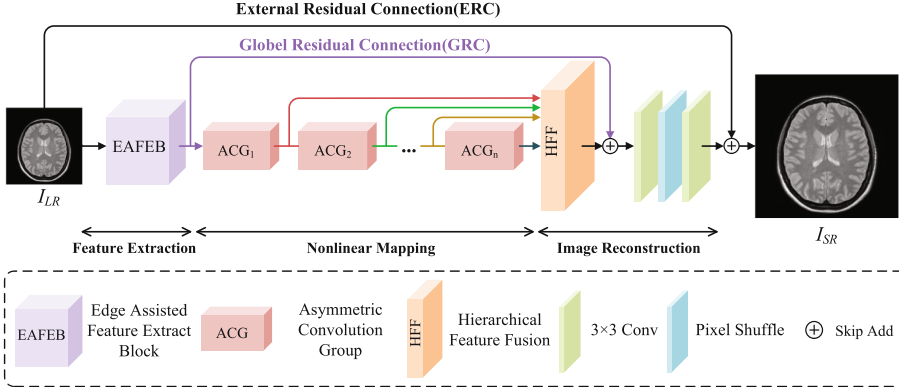


Fig. 1. Framework of the Edge Assisted Asymmetric Convolution Network (EAACN).

Feng et al. [11] proposed T^2Net , which uses joint learning of MR image reconstruction and SR tasks with a transformer to learn correlations between tasks, thus improving the performance of SR. Du et al. [12] proposed an anisotropic MR image reconstruction method, which based on residual learning with long and short skip connections, can effectively restore high-frequency details of MR images. For efficient processing of hierarchical features discriminatively on different channels, Zhao et al. [13] further propose a channel splitting network (CSN).

3 Methods

3.1 Network Architecture

The goal of MR SR is to make the output SR image as close as possible to the real HR image. As shown in Fig. 1, our EAACN network mainly contains three parts: Shallow Feature Extraction, Nonlinear Mapping, and Image Reconstruction.

Shallow Feature Extraction. In the shallow feature extraction section, shallow features $F_{shallow}$ will be extracted by EAFEB from the input LR image:

$$F_{shallow} = EAFEB(I_{LR}) \tag{1}$$

where I_{LR} is original LR input. EAFEB’s multi-channel output includes both content and edge information features, which is beneficial to guarantee the retention of content features and structural features of I_{LR} in the following network.

Nonlinear Mapping. The nonlinear mapping section contains several ACGs, each of which is directly connected to ensure a smooth flow of features. $F_{shallow}$ is ACG’s initial input. To enable the network to fully exploit global features of both shallow and deep features, we use hierarchical feature fusion (HFF) to fuse output features of all ACGs and extract the effective features of each stage:

$$F_{HFF} = HFF[A_{1,o}, A_{2,o}, \dots, A_{n,o}] \tag{2}$$

where $A_{n,o}$ denotes output of n-th ACG, $[\dots]$ implies concatenation. HFF is hierarchical feature fusion, which consists of 1×1 convolutional layer and CSA block. Convolutional is used to modify the channel dimension, while CSA is used to refine spatial dimension. In addition, global residual connection (GRC) is adopted to alleviate network training difficulty, output $F_{mapping}$ can be formulated as:

$$\mathbf{F}_{\text{mapping}} = \mathbf{F}_{\text{HFF}} + \mathbf{F}_{\text{shallow}} \quad (3)$$

Image Reconstruction. In the final section of the entire network, $F_{mapping}$ is upscaled via the Sub-Pixel Convolutional [14] layer, and the output I_{SR} can be represented as follows with the addition of External Residual Connection (ERC):

$$\mathbf{I}_{\text{SR}} = S_{\uparrow}(\mathbf{F}_{\text{mapping}}) + \hat{\mathbf{I}}_{\text{LR}} \quad (4)$$

here S_{\uparrow} is composed of a sub-pixel convolutional layer followed by a 3×3 convolutional layer, \hat{I}_{LR} is interpolated results of the original input LR image. In our model, the bicubic approach is used to implement the interpolation.

Training Objective. Given an MR dataset $\{I_{LR}^i, I_{HR}^i\}_i^N$, where I_{HR}^i represents the ground truth of I_{LR}^i and N denotes the total number of training sets. L_1 loss is utilized to train the model in order to optimize the EAACN network:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|H_{EAACN}(I_{LR}^i, \theta) - I_{HR}^i\|_1 \quad (5)$$

where θ indicates the parameter setting in EAACN.

3.2 Edge Assisted Feature Extraction Block

Figure 2 depicts the detailed structure of the EAFEB block. EAFEB block takes a single-channel grayscale image as input and outputs an a-channel feature map, its two branches handle the extraction of content features and extraction of edge information, respectively. The output channel size of each branch is set to $a/2$, and the final output is obtained by concatenating the results of the two branches.

The content feature extraction branch achieves channel dimension increase and content information extraction, its output F_C can be expressed as:

$$\mathbf{F}_C = E_C(\mathbf{I}_{\text{LR}}) \quad (6)$$

where E_C denotes the content feature extraction branch, which consists of 3×3 convolution, ReLU and 1×1 convolution. Inspired by WDSR [15], to extract more useful features and enable a smoother flow of features, wide convolutional is used. The first 3×3 convolution of the content feature extraction module increases the number of feature channels to $2a$, instead of the output channel $a/2$ of this branch. Following the ReLU function, 1×1 convolution fuses the multidimensional features, and the final output channel is $a/2$.

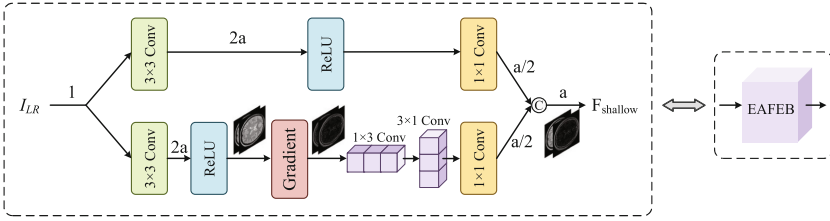


Fig. 2. Framework of the proposed Edge Assisted Feature Extraction Block (EAFEB).

The edge information extraction branch achieves channel dimension increase and edge feature extraction. F_R is obtained by 3×3 convolution and ReLU:

$$\mathbf{F}_R = \text{ReLU}(\text{Conv}_{3 \times 3}(\mathbf{I}_{LR})) \quad (7)$$

where $\text{Conv}_{n \times m}$ indicates the convolution with kernel size $n \times m$ and ReLU is the activation function. Then, from the F_R , we extract the edge information map $\text{Edge}(F_R)$ by calculating the difference between adjacent pixels:

$$\begin{aligned} F_{R_x}(X) &= F_R(x+1, y) - F_R(x-1, y) \\ F_{R_y}(X) &= F_R(x, y+1) - F_R(x, y-1) \\ \nabla F_R(X) &= (F_{R_x}(x), F_{R_y}(x)) \\ \text{Edge}(F_R) &= \|\nabla F_R\|_2 \end{aligned} \quad (8)$$

$\text{Edge}(\cdot)$ is implemented by two convolutions with fixed parameters, which compute gradient information in the horizontal and vertical directions of coordinate point $X = (x, y)$, respectively.

The edge feature $\text{Edge}(F_R)$ has values close to zero in most regions and only takes values at the edge areas. To limit the impact of noise on edge feature map even further, 1×3 and 3×1 convolutions were refined horizontally and vertically, respectively, to obtain the enhanced edge feature map F_S :

$$\mathbf{F}_S = \text{Conv}_{3 \times 1}(\text{Conv}_{1 \times 3}(\text{Edge}(\mathbf{F}_R))) \quad (9)$$

The multidimensional edge feature maps are then fused by 1×1 convolution, and its output channel is $a/2$. Finally, content features and edge features are aggregated in channel dimension, output F_{shallow} of EAFEB is thus given by:

$$\mathbf{F}_{\text{shallow}} = [\mathbf{F}_C, \text{Conv}_{1 \times 1}(\mathbf{F}_S)] \quad (10)$$

Without a doubt, the most significant task in network design for MR images is the maintenance of structural information. Incorporating edge feature into F_{shallow} can enrich F_{shallow} with structural information, making subsequent networks easier to maintain structural features. Furthermore, EAFEB is placed at the very beginning of the network rather than in each group or block in the nonlinear mapping part as it reduces the number of parameters, allowing edge feature extraction module to be performed only once.

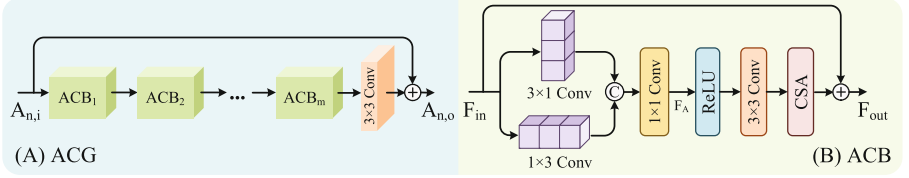


Fig. 3. Architecture of the proposed Asymmetric Convolution Group (ACG) and Asymmetric Convolution Block (ACB).

3.3 Asymmetric Convolution Group

In this section, we will discuss the ACG module, which has excellent perception for both edge and content features. Figure 3-A depicts the overall structure of ACG, which consists of multiple directly connected Asymmetric Convolution Blocks (ACB) and local residual connections. ACB is a lightweight feature extraction block that can improve the content and structural features.

In ACB block, as shown in Fig. 3-B, input features are first extracted by asymmetric convolution in both horizontal and vertical directions, afterwards concatenation and 1×1 convolution are adopted to combine and refine the extracted feature results:

$$\mathbf{F}_A = \text{Conv}_{1 \times 1}[\text{Conv}_{1 \times 3}(\mathbf{F}_{in}), \text{Conv}_{3 \times 1}(\mathbf{F}_{in})] \quad (11)$$

where F_{in} is ACB block's input and F_A is ACB's stage output. We design this method to extract the feature F_A rather than directly using 3×3 convolution for two reasons: 1. There are fewer parameters in this method. 2. It is more effective at extracting edge structure features. Later, ReLU is introduced to improve the ACB block feature stream's nonlinear ability, and 3×3 convolution is then used to improve the perception of content features.

Furthermore, based on the spatial information of the input feature map, we propose a contextual spatial attention block that can dynamically adjust the network's attention to important spatial information. Next, to improve the stability of the network training, residual connection is added, and the resulting ACB output F_{out} is shown below:

$$\mathbf{F}_{out} = \text{CSA}(\text{Conv}_{3 \times 3}(\text{ReLU}(\mathbf{F}_A))) + \mathbf{F}_{in} \quad (12)$$

3.4 Contextual Spatial Attention

In order to improve the effectiveness of ACB, we incorporate an attention mechanism that allows the network to focus more on crucial information computation. The perceptual field of the attention module should be expanded in the design of the CSA module, and the number of parameters should also be considered.

To control the number of parameters, the CSA network's computational effort is focused on low dimensions and small scales. At the beginning of the CSA, an 1×1 convolutional layer is adopted to compress the number of channels.

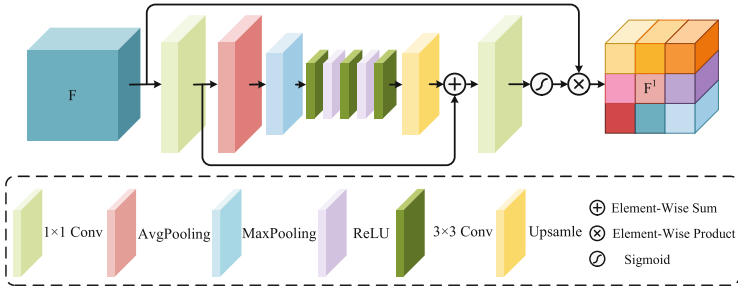


Fig. 4. Detailed implementation of Contextual Spatial Attention (CSA).

Besides, to increase the receptive fields and reduce the computational overhead of the subsequent convolution steps, an avg-pooling layer and a max-pooling layer (kernel size 5, stride size 2) are used (Fig. 4).

Furthermore, a convolutional group is designed for adding learnable parameters to achieve feature full utilization of CSA, it consists of three convolutional layers. To widen the perceptual field, we set the middle convolutional layer’s dilation rate to 2. Then, the output features are upscaled to their original size. We also add a residual connection to obtain a more stable output. After that, we use an 1×1 convolutional layer to recover the number of channels and a sigmoid layer to obtain the attention mask. Finally, attention mask is element-wise producted with input features to focus significant spatial regions of the inputs.

4 Experiments

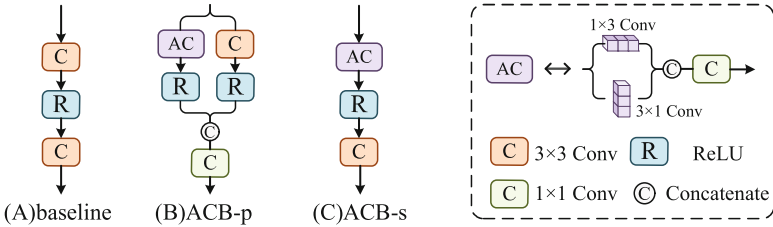
4.1 Datasets and Implementation Details

Three MR image sequences from the public IXI dataset, proton density-weighted imaging (PD), T1-weighted imaging (T1), and T2-weighted imaging (T2), are used for experiment. 576 3D volumes for each MR image type are utilized, and each volume is clipped to the size of $240 \times 240 \times 96$ (*height* \times *width* \times *depth*), the same as [13]. We divide them randomly into 500 training sets, 70 testing sets, and 6 validation sets, respectively. Each 3D volume can be regarded as 96 240×240 2D images in the vertical direction, 10 images chosen as training samples using interval sampling. Thus, in every MR image sequences, we gained $50 \times 100 = 5000$ training images, $70 \times 10 = 700$ testing images, and $6 \times 10 = 60$ validation images in every MRI sequence.

The configuration setting of EAACN is shown in Fig. 1, the number of ACG and ACB is set to 10 and 8, respectively. Data augmentation is performed on the 5000 training images, which are randomly rotated by three angles (90° , 180° , 270°) and horizontally flipped. In each training batch, we feed 12 randomly cropped 60×60 patches into our EAACN. L_1 loss function and ADAM optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$ are used for training. Initial learning rate is set to 10^{-4} and halved at every 6×10^4 iteration. We implement our network using the PyTorch framework with a 3090 GPU.

Table 1. Ablation investigation of EAFEB in EAACN. We report best PSNR/SSMI values on PD sequences. The maximal values of each row are **highlighted**.

Method	AvgParameters/M	SR×2	SR×3	SR×4
FEB	7.03	38.212/0.9794	33.000/0.9477	30.274/0.9113
EAFEB	7.18	38.239/0.9795	33.028/0.9480	30.304/0.9119

**Fig. 5.** Architecture of various structures in ACB, A is the baseline, B and C are the parallel and serial models, respectively.

4.2 Model Analysis

We study the effects of several components in EAACN, including Edge Assisted Feature Extraction Block (EAFEB), Asymmetric Convolution Block (ACB), and Contextual Spatial Attention (CSA).

Effect of EAFEB. To demonstrate the effectiveness of the EAFEB, we redesign a feature extraction module called FEB, which differs from EAFEB in that it only has a content extraction branch. The LR features extracted by FEB have the same output feature dimension as EAFEB. We conduct ablation experiments in the PD×2, PD×3, and PD×4, respectively. As shown in Table 1, when compared to the standard FEB block, the EAFEB block improves SR performance by an average of 0.028 dB in PSNR values and a slight increase in the number of parameters. This demonstrates EAFEB’s effectiveness in fusing content and edge features.

Effect of ACB. Our proposed ACB can use various basic structures, and we design three different combination convolutional structures for future specific exploration of the module’s capabilities (see Fig. 5). To simplify the representation, only necessary connections have been drawn. Figure 5-A depicts the model’s general infrastructure. We replace only one convolutional block with AC to improve edge structure information retention while still maintaining a sense of content information. It is worth noting that the AC has fewer parameters than the convolutional block. As shown in Fig. 5-B and Fig. 5-C, parallel and serial modes are used separately. Larger multiples of the SR task would be more dependent on the structure retention, so SR×4 is chosen for the ablation test.

The quantitative comparison results are reported in Table 2, replacing convolutional blocks with AC improves performance, demonstrating hybrid edge and

Table 2. Ablation investigation of ACB in EAACN. We report best PSNR/SSMI values. The maximal values of each row are **highlighted**.

Method	Parameters/M	PD	T1	T2
baseline	7.54	30.231/0.9113	28.268/0.8482	29.634/0.9042
ACB-p	7.22	30.262/0.9115	28.290/0.8488	29.658/0.9041
ACB-s	7.22	30.304/0.9119	28.301/0.8487	29.710/0.9045

Table 3. Ablation investigation of Spatial Attention in EAACN. We report best PSNR/SSMI values in SR \times 2. The maximal values of each row are **highlighted**.

Method	Parameters/M	PD	T1	T2
w/o SA	6.32	38.145/0.9793	34.882/0.9619	37.681/0.9761
w/o SA+CBAM	6.37	38.157/0.9793	34.936/0.9621	37.712/0.9761
w/o SA+CSA	7.07	38.239/0.9795	34.977/0.9624	37.772/0.9763

content feature extraction’s superiority. The serial ACB-s perform best, which is capable of balancing edge and content perception, with an average 0.06 dB increase over the three datasets when compared to the base structure.

Effect of CSA. Attention mechanisms play an important role in networks, and to demonstrate the effectiveness of CSA in EAACN, we compared it to the previous spatial attention mechanisms CBAM [16], the results of which are shown in Table 3. Removing or replacing CSA results in a considerable performance decrease (PSNR of 0.093 dB and 0.061 dB, respectively). Figure 6 depicts the experimental PSNR performance graph of attention mechanism used in EAACN. As can be observed, CSA has a more steady PSNR curve than CBAM due to its larger perceptual field. Besides, CSA can significantly increase network performance compared to CBAM.

4.3 Comparison with Other Methods

To further demonstrate the effectiveness of our proposed network, we compare EAACN with several SOAT SR networks evaluated by objective evaluation metrics, containing VDSR [17], SRResNet [18], EDSR [19], SeaNet [7], CSN [13] and W^2 AMSN [10]. These models are trained and tested using their default parameter settings with MR image sequences datasets. As shown in Table 4, EAACN+ has the best performance by using self-ensemble, EAACN achieves better performance on each sequence than other methods.

Performance, parameters, and flops volume analysis of these networks are shown in Fig. 7. Compare to the proximity performance performer, EDSR, CSN, and W^2 AMSN, our network has a smaller number of parameters and flops. As for compare to another edge prior network, SeaNet, our EAACN shows a significant performance improvement with a similar number of parameters.

Figure 8 shows the visual results of these methods after recovering images at different sequences and different resolutions. As can be seen from the arrow pointing at the enlarged part of the figure, the structure and content of our network recovery are the closest to the real image. For example, the bottom row

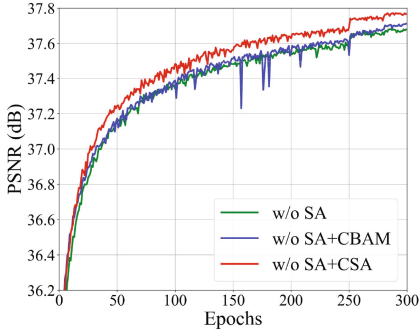


Fig. 6. The performance comparison with attention block. T2 sequence (SR \times 2) is measured.

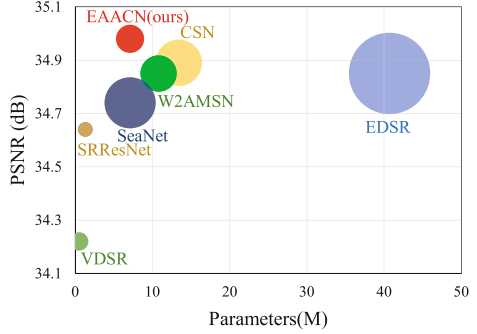


Fig. 7. Performance, parameters, and flops. The number of flops is reflected by the size of the bubbles. T1 sequence (SR \times 2) is measured.

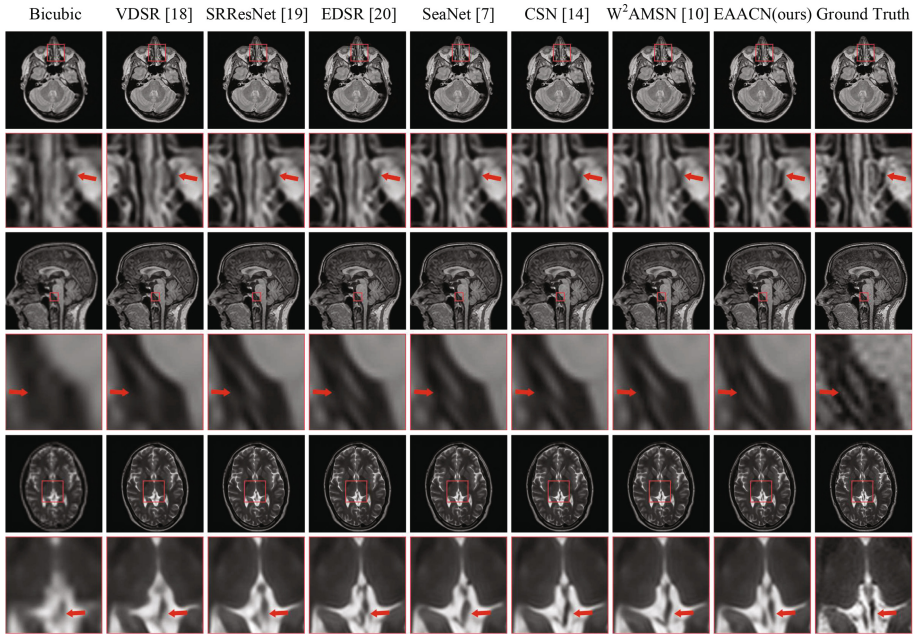


Fig. 8. Visual comparison for PD (top), T1 (middle) and T2 (bottom) with SR \times 2 (top), SR \times 3 (middle) and SR \times 4 (bottom), respectively. The places indicated by red arrows are complex positions. (Color figure online)

in Fig. 8 shows the T2 sequence with SR \times 4. The red arrow points to the location of the black pathway, which is recovered completely by our method, while others cannot accurately restore it.

Table 4. Quantitative results with PD, T1, and T2 sequence. Maximal and second ones are **highlighted** and underlined.

Method	Scale	PD	T1	T2
		PPSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	×2	31.36/0.9394	30.55/0.9108	30.95/0.9304
VDSR [17]	×2	36.59/0.9737	34.22/0.9567	36.41/0.9716
SRResNet [18]	×2	37.71/0.9778	34.64/0.9602	37.32/0.9748
EDSR [19]	×2	38.15/0.9792	34.85/0.9617	37.72/0.9761
SeaNet [7]	×2	37.93/0.9784	34.74/0.9609	37.53/0.9755
W ² AMSN [10]	×2	38.08/0.9790	34.85/0.9617	37.62/0.9759
CSN [13]	×2	38.13/0.9792	34.89/0.9619	37.70/0.9760
EAACN (ours)	×2	<u>38.24/0.9795</u>	<u>34.98/0.9624</u>	<u>37.77/0.9763</u>
EAACN+(ours)	×2	38.42/0.9800	35.08/0.9630	37.95/0.9768
Bicubic	×3	27.59/0.8679	27.00/0.8033	27.37/0.8527
VDSR [17]	×3	31.06/0.9272	29.49/0.8816	31.10/0.9271
SRResNet [18]	×3	32.38/0.9416	30.11/0.8960	32.07/0.9382
EDSR [19]	×3	32.84/0.9465	30.37/0.9017	32.47/0.9424
SeaNet [7]	×3	32.57/0.9438	30.24/0.8993	32.29/0.9410
W ² AMSN [10]	×3	32.79/0.9460	30.33/0.9015	32.42/0.9419
CSN [13]	×3	32.81/0.9463	30.36/0.9025	32.46/0.9421
EAACN (ours)	×3	33.03/0.9480	<u>30.50/0.9049</u>	<u>32.63/0.9435</u>
EAACN+(ours)	×3	33.29/0.9499	30.68/0.9072	32.90/0.9454
Bicubic	×4	25.46/0.7943	25.17/0.7071	25.43/0.7764
VDSR [17]	×4	28.52/0.8783	27.27/0.8101	28.33/0.8764
SRResNet [18]	×4	29.78/0.9015	27.94/0.8356	29.29/0.8966
EDSR [19]	×4	30.21/0.9107	28.21/0.8464	29.60/0.9031
SeaNet [7]	×4	30.03/0.9073	28.12/0.8430	29.60/0.9024
W ² AMSN [10]	×4	30.09/0.9087	28.13/0.8437	29.52/0.9017
CSN [13]	×4	30.17/0.9094	28.20/0.8453	29.58/0.9026
EAACN (ours)	×4	<u>30.30/0.9119</u>	<u>28.30/0.8487</u>	<u>29.71/0.9045</u>
EAACN+(ours)	×4	30.65/0.9157	28.54/0.8536	30.05/0.9086

5 Conclusion

One of the main issues with deep learning-based super-resolution of MR images is the difficulty in recovering clearer structural features, which is critical information in diagnostic process. In this work, we present EAACN for super-resolution of 2D MR images. To maintain the sharp edges and geometric structure of MR images, we innovatively add an edge assisted feature extraction block. An asymmetric convolutional group is adopted in order to allow the network to keep geometric structure while extracting content information. In addition, contextual spatial attention is proposed with a great perceptual field and effective results. Extensive experiments prove that EAACN is superior to state-of-the-art models in both quality and quantity. We believe that our method has the potential to be applied in other types of medical images as well, such as CT and PET.

Acknowledgment. This work is partly supported by the National Natural Science Foundation of China (No. 61873240) and the Open Project Program of the State Key Lab of CAD&CG (Grant No. A2210), Zhejiang University. Thanks to Zheng Wang from Zhejiang University City College for his partial contribution to this project.

References

1. Chen, J., Wang, W., Xing, F., et al.: Residual adaptive dense weight attention network for single image super-resolution. In: IJCNN 2022, pp. 1–10 (2022)
2. Zhang, Y., Li, K., Li, K., et al.: MR image super-resolution with squeeze and excitation reasoning attention network. In: CVPR 2021, pp. 13425–13434 (2021)
3. Ni, N., Wu, H., Zhang, L.: Hierarchical feature aggregation and self-learning network for remote sensing image continuous-scale super-resolution. *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5 (2022)
4. Farooq, M., Dailey, M.N., Mahmood, A., Moonrinta, J., Ekpanyapong, M.: Human face super-resolution on poor quality surveillance video footage. *Neural Comput. Appl.* **33**(20), 13505–13523 (2021). <https://doi.org/10.1007/s00521-021-05973-0>
5. Xia, J., Li, X., Chen, G., et al.: A new hybrid brain MR image segmentation algorithm with super-resolution, spatial constraint-based clustering and fine tuning. *IEEE Access* **8**, 135897–135911 (2020)
6. Yang, W., Feng, J., Yang, J., et al.: Deep edge guided recurrent residual learning for image super-resolution. *IEEE Trans. Image Process.* **26**(12), 5895–5907 (2017)
7. Fang, F., Li, J., Zeng, T.: Soft-edge assisted network for single image super-resolution. *IEEE Trans. Image Process.* **29**, 4656–4668 (2020)
8. Ma, C., Rao, Y., Cheng, Y., et al.: Structure-preserving super resolution with gradient guidance. In: CVPR 2020, pp. 7766–7775 (2020)
9. Michelini, P.N., Lu, Y., Jiang, X.: Super-resolution for the masses. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1078–1087 (2022)
10. Wang, H., Hu, X., Zhao, X., et al.: Wide weighted attention multi-scale network for accurate MR image super-resolution. *IEEE Trans. Circ. Syst. Video Technol.* **32**(3), 962–975 (2022)
11. Feng, C.-M., Yan, Y., Fu, H., Chen, L., Xu, Y.: Task transformer network for joint MRI reconstruction and super-resolution. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12906, pp. 307–317. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87231-1_30
12. Du, J., et al.: Super-resolution reconstruction of single anisotropic 3D MR images using residual convolutional neural network. *Neurocomputing* **392**, 209–220 (2020)
13. Zhao, X., Zhang, Y., Zhang, T., et al.: Channel splitting network for single MR image super-resolution. *IEEE Trans. Image Process.* **28**(11), 5649–5662 (2019)
14. Shi, W., et al.: Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: CVPR 2016, pp. 1874–1883 (2016)
15. Yu, J., Fan, Y., Yang, J., et al.: Wide activation for efficient and accurate image super-resolution. arXiv preprint [arXiv:1808.08718](https://arxiv.org/abs/1808.08718) (2018)
16. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1
17. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: CVPR 2016, pp. 1646–1654 (2016)

18. Ledig, C., Theis, L., Huszár, F., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR 2017, pp. 105–114 (2017)
19. Lim, B., Son, S., Kim, H., et al.: Enhanced deep residual networks for single image super-resolution. In: CVPR Workshops 2017, pp. 1132–1140 (2017)