Lightweight Single Image Super-Resolution Convolution Neural Network in Portable Device

**Jin Wang[1], Yiming Wu[1], Shiming He[1], Pradip Kumar Sharma[2], Xiaofeng Yu[3*],
Osama Alfarraj[4], and Amr Tolba[4,5]**

[1] School of Computer and Communication Engineering, Changsha University of Science and Technology
Changsha, Hunan 410114 China
[e-mail: jinwang@csust.edu.cn, wuyiming_cs@stu.csust.edu.cn, smhe_cs@csust.edu.cn]
[2] Department of Computing Science, University of Aberdeen
Aberdeen, 999020 UK
[e-mail: pradip.sharma@abdn.ac.uk]
[3] School of Business, Nanjing University
Nanjing, Jiangsu 210093 China
[e-mail: xiaofengyu@nju.edu.cn]
[4] Computer Science Department, Community College, King Saud University, Riyadh, 11437, Saudi Arabia
[e-mail: oalfarraj@ksu.edu.sa, atolba@ksu.edu.sa]
[5] Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Shebin-El-kom32511, Egypt
[*] Corresponding author: Xiaofeng Yu

Abstract

Super-resolution can improve the clarity of low-resolution (LR) images, which can increase the accuracy of high-level compute vision tasks. Portable devices have low computing power and storage performance. Large-scale neural network super-resolution methods are not suitable for portable devices. In order to save the computational cost and the number of parameters, Lightweight image processing method can improve the processing speed of portable devices. Therefore, we propose the Enhanced Information Multiple Distillation Network (EIMDN) to adapt lower delay and cost. The EIMDN takes feedback mechanism as the framework and obtains low level features through high level features. Further, we replace the feature extraction convolution operation in Information Multiple Distillation Block (IMDB), with Ghost module, and propose the Enhanced Information Multiple Distillation Block (EIMDB) to reduce the amount of calculation and the number of parameters. Finally, coordinate attention (CA) is used at the end of IMDB and EIMDB to enhance the important

information extraction from Spaces and channels. Experimental results show that our proposed can achieve convergence faster with fewer parameters and computation, compared with other lightweight super-resolution methods. Under the condition of higher peak signal-to-noise ratio (PSNR) and higher structural similarity (SSIM), the performance of network reconstruction image texture and target contour is significantly improved.

---

1. Introduction

Image is an important carrier to record people's lives and historical scenes. There are a lot of image data in urban governance, such as traffic flow prediction, remote sensing, and evidence collection. With the proliferation of high-resolution display devices, low-resolution images can't get good visual effect on high-resolution devices. Single image super-resolution can save the cost of upgrading imaging equipment to achieve high resolution image reconstruction performance Therefore, single image super-resolution has high research value.

Single image super-resolution (SISR) is a kind of low-level and ill-posed image processing problem [1-5]. Since real low-resolution (LR) images have innumerable high-resolution (HR) images corresponding to them, traditional interpolation methods cannot extract high-frequency information from low-resolution images, resulting in the failure to obtain pleasant visual effect. The Super-Resolution Convolution Neural Network (SRCNN) proposed by Dong et al. [6] was the first method to apply three-layer convolutional neural network, and achieved better performance than the traditional super-resolution method.

However, the network structure is usually deepened because of the better super-resolution performance pursuing. Even if the reconstruction block use the deconvolution and the sub-pixel convolution in the end of network, the computational cost and the number of parameters will inevitably increase. For example, the convolution depth of the Residual Super-resolution Network (EDSR) proposed by Kim et al. [7] reaches 150 layers, and the Residuals Attention Mechanism Network (RCAN) proposed by Zhang et al. [8] reaches 415 layers. However, the convolution depth of the lightweight cascading residuals super-resolution network (CARN) [9], is only 34 layers. The lightweight super-resolution algorithm greatly reduces the computational cost and the number of parameters on the premise of maintaining a reasonable reconstruction performance. Therefore, it has a high availability in real-time super resolution equipment and terminal equipment with poor computing power.

Hui et al. [10] proposed lightweight image SR and information multiple distillation network (IMDN). In the IMDB, channel split operation is carried out for the feature extraction convolution, and the input feature channels are divided into feature channels with a fixed number of smaller channels. It not only ensures a good reconstruction performance but also

reduces the amount of calculation and the number of parameters.

To further improve IMDN, we consider the three ways: Ghost module [11], Feedback mechanism [12], and Coordinate attention (CA) mechanism [13]. We propose the Enhanced Information Multiple Distillation Network (EIMDN) based on the IMDN with the three methods. First, we propose the Enhanced Information Multiple Distillation Block (EIMDB) which use the Ghost module to replace the 3×3 convolution from IMDB feature extraction. And we propose the middle feature extraction and a deep feature extraction module, which consists of EIMDB and IMDB, respectively. Then, we replace the contrast aware channel attention (CCA) mechanism in IMDB with CA mechanism, which can achieve the performance of channel attention and retain the long-term spatial dependence. Finally, we use the feedback mechanism to combine the deep feature extraction with the shallow feature extraction and fed into the middle feature extraction module to improve the context connection between features. Compared with the existing lightweight super-resolution methods, our method achieves a better super-resolution reconstruction performance while maintaining a low number of parameters and computation. The main contributions in this paper can be summarized as follows:

- We propose EIMDB and use Ghost module to replace the convolution operation of the feature extraction part in IMDB, which further reduces the number of parameters and the amount of calculation.
- We propose middle feature extraction and high feature extraction, respectively using EIMDB and IMDB, and using feedback mechanism to improve the context connection between features. And we introduced the CA mechanism to replace the CCA mechanism in IMDB, which improved the reconstruction performance.
- We propose the EIMDN. Compared with the existing lightweight super-resolution methods, we reduce the number of parameters and the computational cost while ensuring the reconstruction performance of the network.

The rest of this paper is organized as follows. The related work is present in Section 2. The Section 3 is explained the architecture of our network and submodule composition. The Section 4 present the design of experiments, along with a comparison with existing lightweight methods. The final Section summarizes our network and further work.

2. Related Work

2.1 Single Image Super-Resolution Based on Deep Learning

The (SRCNN) was proposed by Dong et al. [6] for the first time, introduced three-layer convolutional neural network into the field of image super-resolution, and achieved better results than traditional methods. The Deep Recursive Convolution Network (DRCN) proposed by Kim et al. [14] introduced the recursive neural network to super-resolution and made the feature extraction part repeat learning through recursive structure without in-creasing the number of network parameters. However, early deep learning super-resolution network use the up-scaling image as the network input, which leads to large feature map scale and increases the number of parameters and computation of the network. Then, the Faster Super-resolution Convolutional Neural Network (FSRCNN) proposed by Dong et al. [15] and the Sub-pixel Convolutional Network (ESPCN) proposed by Shi et al. [16] introduced the methods of deconvolution and sub-pixel convolution, respectively to extract the features of low-resolution images directly through the network. The two methods reduce the computational cost due to the large feature map.

With the deepening of the network, the overfitting of convolutional neural network limited the training of image reconstruction. The EDSR proposed by Lim et al. [7] used global residual connection and local residual connection, which further deepens the network and achieved a better reconstruction performance while reducing the training overfitting. The RCAN model proposed by Zhang et al. [8] added the Squeeze-Excitation channel attention mechanism to the residual blocks, which improved the image reconstruction performance further. Recently, Image Processing Transformer (IPT) proposed by Chen et al. [17] used Transformer [34] to extract deep features and achieve multiple low-level computer vision tasks such as super-resolution image rain-removal, defogging and denoising. The IPT used ImageNet [35] dataset and 32 V100 GPUs for training, which made the image reconstruction reaching one of the best performances.

Although the large-scale SISR network can achieve great image reconstruction performance, the large-scale network not only need a long time of training and a huge amount of computing resources, but also cannot be applied to portable terminal devices and low computing power of the Internet of Things devices.

2.2 Lightweight Single Image Super-Resolution

The lightweight super-resolution method was proposed to apply to portable terminal devices. The CARN proposed by Ahn et al. [12] used global and local cascade connection to improve the contextual learning between features under the same network depth structure, and proposed a CARN-M network with about 60% fewer parameters than CARN. The Information distillation network (IDN) proposed by Hui et al. [18] introduces the concept of information distillation, dividing the functional extraction module into information enhancement unit and compression unit, extracting important information and deleting redundant information, respectively. The IMDB also proposed by Hui et al. [13] splits the feature channels successively, and uses information distillation on the different feature channels, respectively. The CCA mechanism is added at the end of the module, which plays the role of extracting important channel information. Therefore, we use feedback mechanism, Ghost module, and CA attention mechanism to improve the IMDN.

3. Our Proposed Method -- EIMDN

3.1 Network Architecture

As shown in **Fig. 1**, the network architecture is divided into five parts: shallow feature extraction, middle feature extraction, deep feature extraction, image reconstruction and feedback mechanism. We will introduce our network structure in the following five parts. We define $I_{LR} \in \mathbb{R}^{h \times w \times 3}$ the input low-resolution image, $I_{SR} \in \mathbb{R}^{H \times W \times 3}$ the corresponding output super-resolution image, and $I_{HR} \in \mathbb{R}^{H \times W \times 3}$ the ground truth (original high-resolution image) of the input image I_{LR}. The W and H represent the width and high of the I_{HR}. THE w and h represent the width and high of the I_{LR}. Our goal is to reconstruct the high-resolution image I_{SR} from the input I_{LR} by our proposed end to end EIMDN network.
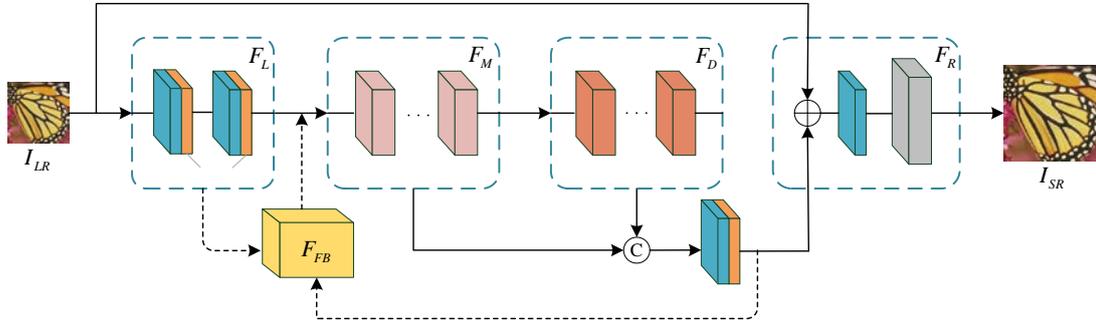
Fig. 1. The EIMDN architecture, where I_{LR} and I_{SR} represent low-resolution images of the input network and super-resolution images of the output network respectively, F_L represents the shallow feature extraction module, F_M represents the middle feature extraction module, F_D represents the deep feature extraction module, F_R represents the image reconstruction module, and F_{FB} represents the feedback mechanism.

3.1.1 Shallow feature extraction

Shallow feature extraction increases the number of channels use convolution from the red, green, and blue (RGB) channels of low-resolution images, so that more channels can be imputed in the middle and deep feature extraction. Specifically, we use a 3×3 convolution and a 1×1 convolution, and the number of output channels are 256 and 64, respectively, which can be represented as follow:

$$F_L = conv_{1\times1}\big(conv_{3\times3}(I_{LR})\big), \tag{1}$$

where F_L represents the shallow feature extraction module output, $conv_{1\times1}$ and $conv_{3\times3}$ represent the 1×1 and 3×3 kernel size of convolution operation, respectively, and I_{LR} represents the low-resolution images of the input.

3.1.2 Middle feature extraction

We propose the concept of middle feature extraction, aiming to obtain the middle feature information in the feature map through the EIMDB module with fewer parameters and less the computational cost. We used N EIMDB modules to extract middle features. The detail of EIMDB is introduced in Section 3.2, and middle feature extraction can be represented as follow:

$$F_M = F_{EIMDB}^N \Big(\cdots F_{EIMDB}^i \Big(\cdots \Big(F_{EIMDB}^1 (F_L) \Big) \cdots \Big) \cdots \Big), \tag{2}$$

where F_M represents the middle feature extraction module output, and F_{EIMDB}^i represents the i_{th} EIMDB output in the middle feature extraction ($0 < i \leq N$).

3.1.3 Deep feature extraction

Although the parameters and the calculation cost of the EIMDB are greatly reduced because the EIMDB is not good at extracting deep information. Therefore, we use the M IMDB modules as the deep feature extraction module to extract the deeper information. Combine with middle feature extraction, we not only decrease the parameters and the calculational cost, but also ensure the extraction effect of feature information. The deep feature extraction can be represented as follow:

$$F_D = F_{IMDB}^M \Big(\cdots F_{IMDB}^j \Big(\cdots \Big(F_{IMDB}^1 (F_M) \Big) \cdots \Big) \cdots \Big), \tag{3}$$

where F_D represents the deep feature extraction module output, and F_{IMDB}^j represents the j_{th} IMDB in the deep feature extraction output $(0 < j \leq M)$.

3.1.4 Image reconstruction

The deep feature extraction and middle feature extraction output feature map are firstly integrated through concatenate and 1×1 convolution, and then the pixel-add operation is carried out with the output of shallow feature extraction, followed by a 3×3 convolution and a sub-pixel convolution, with the output channels of 64 and 3, respectively. The reconstruction part can be represented as follow:

$$I_{SR} = F_R = f_{sub}(conv_{1 \times 1}(concat(F_M, F_H), I_{LR}), \tag{4}$$

where I_{SR} represents super-resolution image output, which is image reconstruction module output F_R, $f_{sub}(\cdot)$ represents the subpixel convolution operation, and $concat(\cdot)$ represents the concatenate operation.

3.1.5 Feedback mechanism

We introduce the feedback mechanism in the Super-resolution network based on feedback mechanism (SRFBN) [19]. The feedback mechanism combines the deep feature extraction with the shallow feature extraction in order to improve the contextual relevance of the features without increasing the number of parameters. In the ablation study in 4.4, we discuss the influence of feedback mechanism on the recovery effect. The feedback mechanism can be represented as follow:

$$F_{FB} = conv_{1 \times 1}\left(concat\left(conv_{1 \times 1}\left(concat(F_M, F_H)\right), F_L\right)\right), \tag{5}$$

where F_{FB} represents the feedback mechanism module output, and theoutput channels by the two 1×1 convolution operation is 64.

3.2 Information Multiple Distillation Block (IMDB)

As shown in **Fig. 2**, the IMDB splits the feature map into four 3×3 convolutions for feature extraction. After each convolution, some channels are split into inputs to the next level of convolution, while the remaining channels are retained. Specifically, the number of input channel layers of the four convolutions are 64, 48, 48, 48, and the number of output channel layers are 64, 64, 64, 16, respectively. After the first three convolution outputs, the channel layers are divided into 48 and 16 by split operation, which are respectively passed into the next convolution operation and reservation. The reservation can be regarded as the feature of refinement. Then, the 16 layers of channels retained by the four convolutions are reconstructed into 64 layers by concatenate operation. Finally, the CA mechanism is introduced to extract important features. IMDB can be shown as follows:

$$F_{refine1}^i, F_{coarse1}^i = f_{split}\left(conv_{3 \times 3}\left(F_{IMDB_{in}}^i\right)\right), \tag{6}$$

$$F_{refine2}^i, F_{coarse2}^i = f_{split}\left(conv_{3 \times 3}\left(F_{coarse_1}^i\right)\right), \tag{7}$$

$$F_{refine3}^i, F_{coarse3}^i = f_{split}\left(conv_{3 \times 3}\left(F_{coarse_2}^i\right)\right), \tag{8}$$

$$F_{refine4}^i = conv_{3 \times 3}\left(F_{coarse3}^i\right), \tag{9}$$

$$F_{IMDB}^i = conv_{1 \times 1}\left(F_{CA}\left(concat\left(F_{refine_1}^i, F_{refine_2}^i, F_{refine_3}^i, F_{refine_4}^i\right)\right)\right), \tag{10}$$

where $F_{IMDB\_in}^i$ represents the input of the i_{th} IMDB, F_{refine}^i and F_{coarse}^i respectively represent the feature map to be cut and feature map for further feature extraction, and $f_{split}(\cdot)$

represents the feature channel segmentation operation, F_{CA} represents the CA mechanism, which is illustrated in section 3.3.2, and F_{IMDB}^i represents the output of the i_{th} IMDB module.
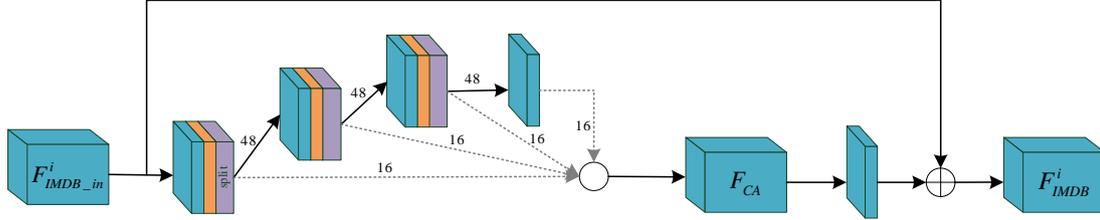

Fig. 2. The IMDB architecture.

3.3 Enhanced Information Multiple Distillation Block (EIMDB)

The IMDB four convolution of feature extraction with Ghost module operation are replaced, as shown in **Fig. 3**, which decreased the parameters and the amount of calculation. At the same time, CA mechanism is used to replace CCA mechanism in the IMDB, which not only retains the channel attention, but also increases the spatial attention in both high and weight directions in images. The EIMDB can be shown as follows:

$$F_{refine\_1}^i, F_{coarse\_1}^i = f_{split}(F_{GM}(F_{EIMDB\_in}^i)), \qquad (11)$$

$$F_{refine\_2}^i, F_{coarse\_2}^i = f_{split}(F_{GM}(F_{coarse\_1}^i)), \qquad (12)$$

$$F_{refine\_3}^i, F_{coarse\_3}^i = f_{split}(F_{GM}(F_{coarse\_2}^i)), \qquad (13)$$

$$F_{refine\_4}^i = F_{GM}(F_{coarse\_3}^i), \qquad (14)$$

$$F_{EIMDB}^i = conv_{1\times1}(F_{CA}(concat(F_{refine\_1}^i, F_{refine\_2}^i, F_{refine\_3}^i, F_{refine\_4}^i))), \qquad (15)$$

where $F_{EIMDB\_in}^i$ represents the input of the i_{th} EIMDB, F_{refine}^i and F_{coarse}^i respectively represent the feature map to be cut and feature map for further feature extraction, and $f_{split}(\cdot)$ represents the feature channel segmentation operation. F_{CA} represents the CA attention mechanism, and F_{EIMDB}^i represents the output of the i_{th} EIMDB module.
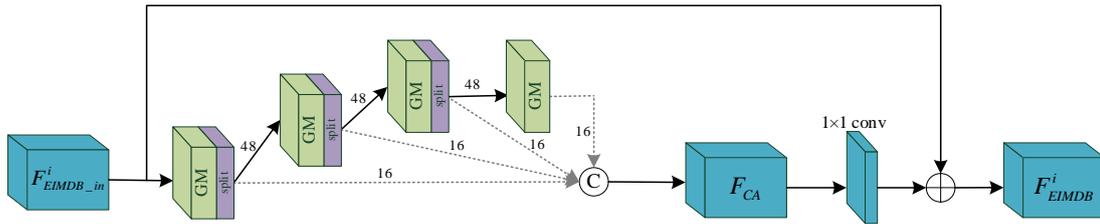

Fig. 3. The EIMDB architecture.

3.3.1 Ghost module

Some lightweight image recognition networks (such as MoblieNet [37] and MoblieNeXt [38]) usually use depth-wise convolution instead of traditional convolution because of group convolution. The group convolution reduces the computational cost and the number of parameters by reducing the correlation between feature channels. Therefore, Ghost module using group convolution aims to remove the redundancy of feature channels. We use Ghost module to replace the feature extraction convolution part in IMDB.

As shown in **Fig. 4**, Ghost module is composed of two parts: primary convolution and cheap operation. Assume that the number of input feature channels is D and the number of output feature channels of Ghost module is K. Primary convolution removes redundant feature channels by convolution kernel size of 1×1. The number of feature channels after 1×1 convolution is d $(0 < d < D)$. The primary convolution is shown as follow:

$$Feat_1 = conv_{1 \times 1}(I_F), \tag{16}$$

where $Feat_1$ represents the feature map after primary convolution, and I_F represents the input of Ghost module.

Cheap operation learns useful information from retained feature map by group convolution, and the number of output channels is $K - d$. Then concatenate operation combines the feature map of the two parts as the output of Ghost module, which is shown as follows:

$$Feat_2 = g\_conv_{3 \times 3}(Feat_1), \tag{17}$$
$$F_{GM} = concat(Feat_1, Feat_2), \tag{18}$$

where $Feat_2$ represents the feature map after cheap operation, $g\_conv_{3 \times 3}$ represents the kernel size of 3×3 group convolution, and F_{GM} represents the output of Ghost module.
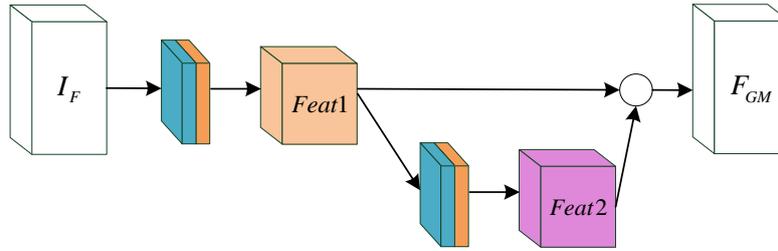

Fig. 4. The Ghost module architecture.

3.3.2 Coordinate attention

We use the CA mechanism to replace the original CCA in IMDB. As shown in **Fig. 5**, the CA mechanism firstly combine the vertical and horizontal input feature channels into two-dimension aware feature map by coordinate global average pooling. Then, the aware feature map is used to extract the vertical and horizontal dependencies, respectively. By the extraction of channel attention. The aware feature maps carry the coordinate attention information in the input features by horizontal and vertical pixel-multiplication respectively.

The coordinate global average pooling of CA mechanism is represented as follow:

$$F^X_{pool}(w) = \frac{1}{W}\sum_{0 \leq i < W} x_c(h, i), \tag{19}$$
$$F^Y_{pool}(h) = \frac{1}{H}\sum_{0 \leq i < H} x_c(j, w), \tag{20}$$

where $F^X_{pool}(w)$ and $F^Y_{pool}(h)$ represent one-dimensional pooling with width as the direction and length as the direction respectively, W and H are the width and length of input feature mapping respectively, and $x_c(\cdot)$ represents the value of pixel points at a fixed position.

The CA mechanism can not only extract the important channel information, but also obtain the spatial location information, with fewer increasing of the computational cost and the number of parameters. The necessity of the CA mechanism is discussed in Section 4.4.1.
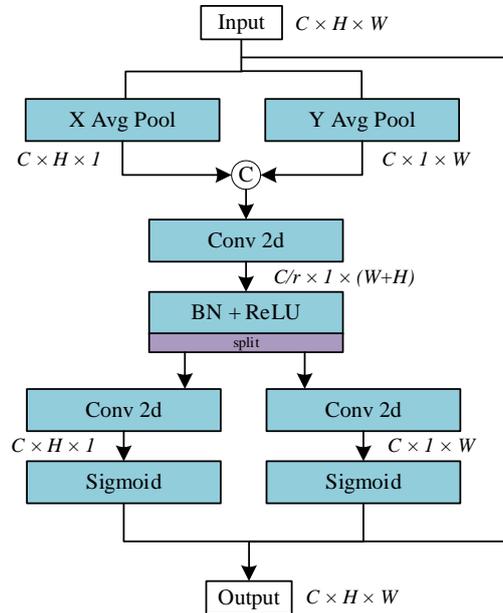
Fig. 5. Architecture of CA mechanism, where C is the number of channels, W and H are width and height respectively, and the reduction ratio $r = 8$.

3.4 Loss Function

The L_1 loss function is chosen to optimize the network. Using the loss function of the feedback mechanism, we calculate the L_1 loss function using the super-resolution image output of the two iterations I_{SR}^t. The I_{HR} is original high-resolution image, and average it as:

$$\mathcal{L}(\Theta) = \frac{1}{T}\sum_{t=1}^{T} ||I_{HR} - I_{SR}^t||_1, \tag{21}$$

where the Θ is network parameter, $T = 2$ is the overall number of iterations, t is the network iteration, I_{SR}^t and I_{HR} represent the super-resolution image and, original high-resolution image respectively.

4. Experimental Results and Analysis

4.1 Datasets

Table 1 shows the information of datasets. The training set is DIV2K [20], including 800 images with 2K resolution of people, handmade products, buildings (cities, villages), natural scenery, plants and animals. The training set is amplified in three ways: rotating the image randomly by 90°, 180° and 270°; horizontally or vertically flipping the image; or down-sampling the HR image by 0.6~0.9. Five widely used super-resolution benchmarks, Set5 [21], Set14 [22], BSD100[23], Urban100 [24], and Manga109 [25], are used to evaluate the model performance. The Urban100 contains 100 challenging images of city scenes with dense high-frequency feature details. Manga109 is 109 comic cover pictures, with high frequency and low frequency information as well as text information, to test the model's comprehensive processing ability of text and picture.

Table 1. The information of datasets

Usage	Datasets	Number	Source
Training	DIV2K [20]	800	ETH Zurich
Test	Set5 [21]	5	INRIA
Test	Set14 [22]	14	Israel Institute of Technology
Test	BSD100 [23]	100	University of California, Berkeley
Test	Urban100 [24]	100	University of Illinois
Test	Manga109 [25]	109	University of Tokyo

4.2 Implementation Details

The experiment uses the method of He et al. [36] to initialize the network parameters. The initial learning rate is 10^{-4}, which is multiplied by 0.5 every 200 epochs, and the total iteration is 1000 times. Adam [26] algorithm ($\beta1=0.9$, $\beta2=0.999$) is used to optimize the network parameters. We set the image patch-size as 48×48 and the batch-size as 16. The experimental environment is Pytorch1.6.0 of GPU version, the GPU is used for training with NVIDIA RTX 2070 Super, and the operating system is Ubuntu16.08. The CPU uses Inter i5-10400F. The Memory size is 32GB.

The evaluation criteria commonly use peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). We calculate the PSNR and SSIM on the luminance channel (i.e., Y channel of the YCbCr channels converted from the RGB channels):

$$PSNR = 10 \cdot log_{10} \left(\frac{MAX_I^2}{MSE} \right),$$ (19)

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)},$$ (20)

where the MSE is the mean square error, and the MAX_I is the maximum value that represents the color of an image point, the x and y are I_{HR} and I_{SR}, respectively, the μ and σ are mean value and standard deviation, respectively, and the c is constant term.

From the experiment, we propose two models, EIMDN-L and EIMDN-S as our network. The number of EIMDB (N) and IMDB (M) used by EIMDN-L is N=6 and M=6, respectively; the number of EIMDB and IMDB used by EIMDN-S is N=3 and M=3, respectively. The EIMDN-L+ represent the EIMDN-L used the DIV2K and Flickr2K for training set. Our EIMDN-L+ is the best model of performance evaluation.

4.3 Performance Evaluation

Table 2 shows the EIMDN-S and EIMDN-L compared with the state-of-the-art (SOTA) lightweight SR methods. It is well known that with the increase of up-scaling factors, the difficulty of image super-resolution reconstruction is increase. Therefore, a lot of lightweight SR methods cannot get better performance in height up-scaling factors. The EIMDN-L is better than most models at up-scaling factor ×3 and ×4. In the up-scaling factor ×4, the PSNR value of the EIMDN-L in the test set Manga109 is higher than the VDSR and the CARN 2.35dB and 0.11dB, respectively. The EIMDN-L can deepen the learning depth of high-frequency information by feedback mechanism, which is difficult to extract in high magnification. It can achieve good reconstruction performance in high up-scaling factors. Through CA mechanism, more high frequency information in channel and spatial can be filtered and retained. The performance of our EIMDN-L in Urban100 test set is significantly higher than other data sets, because Urban100 test set contains urban buildings with more high-frequency details, and the EIMDN-L is good at reconstruct the building details. Therefore, the Urban 100 test set have better performance by our EIMDN-L

Table 2. The SOTA super-resolution methods PSNR / SSIM mean values compared with our methods at ×2, ×3 and ×4 up-scaling factor in the benchmark test set. Red and blue color indicate the best and second-best performance, respectively (The EIMDN-L+ use **blur** type)

Method	Scale	Set5 PSNR/SSIM	Set14 PSNR/SSIM	B100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
VDSR [27]	×2	37.53/0.9587	33.03/0.9142	31.90/0.8960	30.76/0.9140	37.22/0.9750
DRCN [7]		37.63/0.9588	33.04/0.9118	31.85/0.8942	30.75/0.9133	37.55/0.9732
LapSRN [28]		37.52/0.9591	32.99/0.9124	31.80/0.8952	30.41/0.9103	37.27/0.9740
DRRN [29]		37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188	37.88/0.9749
MemNet [30]		37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195	37.72/0.9740
IDN [18]		37.83/0.9600	33.30/0.9148	32.08/0.8958	31.27/0.9196	38.01/0.9749
SRMDNF [31]		37.79/0.9601	33.32/0.9159	32.05/0.8958	31.33/0.9204	38.07/0.9761
CARN [12]		37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256	38.36/0.9765
CARN-M [12]		37.53/0.9583	33.26/0.9141	31.92/0.8960	30.83/0.9233	36.34/0.9696
SRRAM [32]		37.82/0.9592	33.48/0.9171	32.12/0.8983	32.05/0.9264	-/-
ESRN-F [33]		37.93/0.9602	33.56/0.9171	32.16/0.8996	31.99/0.9276	-/-
IMDN [13]		38.00/0.9605	33.63/0.9177	32.19/0.8996	32.17/0.9283	38.88/0.9774
(Ours) EIMDN-S		37.87/0.9600	33.42/0.9162	32.05/0.8981	31.72/0.9242	38.36/0.9763
(Ours) EIMDN-L		37.99/0.9605	33.68/0.9191	32.14/0.8993	32.19/0.9289	38.56/0.9768
(Ours) EIMDN-L+		38.11/0.9608	33.76/0.9193	32.22/0.9002	32.46/0.9316	39.13/0.9779
VDSR [27]	×3	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279	32.01/0.9340
DRCN [7]		33.82/0.9226	29.76/0.8311	28.80/0.7963	27.15/0.8276	32.24/0.9343
LapSRN [28]		33.81/0.9220	29.79/0.8325	28.82/0.7980	27.07/0.8275	32.21/0.9350
DRRN [29]		34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378	32.71/0.9379
MemNet [30]		34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
IDN [18]		34.11/0.9253	29.99/0.8354	28.95/0.8013	27.42/0.8359	32.71/0.9381
SRMDNF [31]		34.12/0.9254	30.04/0.8382	28.97/0.8025	27.57/0.8398	33.00/0.9403
CARN [12]		34.29/0.9255	30.29/0.8407	29.06/0.8034	28.06/0.8493	33.50/0.9440
CARN-M [12]		33.99/0.9236	30.08/0.8367	28.91/0.8000	26.86/0.8263	32.07/0.9404
SRRAM [32]		34.30/0.9256	30.32/0.8417	29.07/0.8039	28.12/0.8507	-/-
ESRN-F [33]		34.32/0.9268	30.35/0.8410	29.09/0.8046	28.11/0.8512	-/-
IMDN [13]		34.36/0.9270	30.32/0.8417	29.09/0.8046	28.17/0.8519	33.61/0.9445
(Ours) EIMDN-S		34.32/0.9265	30.27/0.8408	29.03/0.8038	27.94/0.8478	33.43/0.9430
(Ours) EIMDN-L		34.47/0.9274	30.40/0.8434	29.09/0.8050	28.22/0.8535	33.69/0.9449
(Ours) EIMDN-L+		34.57/0.9285	30.48/0.8447	29.19/0.8074	28.51/0.8600	34.08/0.9471
VDSR [27]	×4	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524	28.83/0.8870
DRCN [7]		31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510	28.93/0.8854
LapSRN [28]		31.54/0.8852	28.09/0.7700	27.32/0.7275	25.21/0.7562	29.09/0.8900
DRRN [29]		31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638	29.45/0.8946
MemNet [30]		31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630	29.42/0.8942
IDN [18]		31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632	29.41/0.8942
SRMDNF [31]		31.96/0.8925	28.35/0.7787	27.49/0.7337	25.68/0.7731	30.09/0.9024
CARN [12]		32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837	30.47/0.9084
CARN-M [12]		31.92/0.8903	28.42/0.7762	27.44/0.7304	25.63/0.7688	29.09/0.8946
SRRAM [32]		32.13/0.8932	28.54/0.7800	27.56/0.7350	26.05/0.7834	-/-
ESRN-F [33]		32.15/0.8940	28.59/0.7804	27.59/0.7354	26.11/0.7851	-/-
IMDN [13]		32.21/0.8948	28.58/0.7811	27.56/0.7353	26.04/0.7838	30.45/0.9075
(Ours) EIMDN-S		32.07/0.8932	28.48/0.7791	27.49/0.7339	25.85/0.7787	30.23/0.9044
(Ours) EIMDN-L		32.19/0.8953	28.64/0.7826	27.57/0.7367	26.13/0.7876	30.58/0.9087
(Ours) EIMDN-L+		32.41/0.8973	28.74/0.7851	27.64/0.7385	26.32/0.7925	30.90/0.9119

In the up-scaling factor ×2, our method does not always achieve the best effect compared with other models. This indicate that although the Ghost module adopted by our EIMDB can reduce the number of parameters and remove the redundant channels of feature, each channel

retains more high-frequency details at a lower up-scaling factor. This cause some important feature channels to be lost in the removal of redundant steps and does not achieve the best results at up-scaling factor ×2.

The EIMDN-S achieve weaker reconstruction performance than the EIMDN-L because of lower parameters and computational cost. But compared with the SOTA, the EIMDN-S still maintains a good reconstruction performance.

4.4 Ablation Experiment

We discuss two ablation experiments: The influence of feedback mechanism and CA mechanism; The influence of different number of IMDB and EIMDB.

4.4.1 The influence of feedback mechanism and CA mechanism

We compare the performance of four network models: Baseline (IMDN-S with neither feedback nor CA mechanism as our benchmark for comparison), Without Feedback (EIMDN-S without feedback mechanism), Without CA (EIMDN-S using CCA mechanism rather than CA mechanism), EIMDN-S in the case of up-scaling factor ×3 of Set5 test set.

As shown in **Table 3**, compared with the Baseline network, the PSNR and SSIM of the Without CA network and the Without Feedback network increase 0.09 dB, 0.05 dB, and 0.0005, 0.0003, respectively. The proposed EIMDN-S has the best performance, which is 0.1dB and 0.0008 higher than the PSNR and SSIM of the benchmark network. The experiment proves that the feedback mechanism and CA mechanism can improve the performance of image reconstruction.

Table 3. Influence of adding feedback mechanism and CA mechanism on network reconstruction performance

Module	CA	Feedback	Set5	
			PSNR	**SSIM**
Baseline			34.22	0.9257
Without Feedback	√		34.27	0.9260
Without CA		√	34.31	0.9262
EIMDN-S	√	√	34.32	0.9265

4.4.2 The influence of different number of IMDB and EIMDB

We discuss the influence of the number of EIMDB and IMDB in the network. As shown in **Table 4**, we set 6 network the number of EIMDB and IMDB as N=6 M=0, the N=3 M=3, the N=4 M=4, the N=5 M=5, the N=6 M=6 and the IMDN in 5 test sets and up-scaling factor ×3.

Table 4. The number of EIMDB (N) and IMDB (M) in network PSNR and SSIM performance comparison (scale: ×3)

| Module	Param. (K)	Set5		Set14		Urban100		Manga109	
		PSNR	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
N=6	224	34.03	0.9242	30.02	0.8353	27.45	0.8357	32.80	0.9380
N=3 M=3	502	34.32	0.9265	30.27	0.8408	27.94	0.8478	33.43	0.9430
N=4 M=4	627	34.39	0.9268	30.32	0.8413	28.08	0.8503	33.51	0.9436
IMDN	703	34.36	0.9270	30.32	0.8417	28.17	0.8519	33.61	0.9445
N=5 M=5	753	34.41	0.9272	30.38	0.8428	28.16	0.8523	33.58	0.9442
N=6 M=6	879	34.47	0.9274	30.40	0.8434	28.22	0.8535	33.69	0.9449
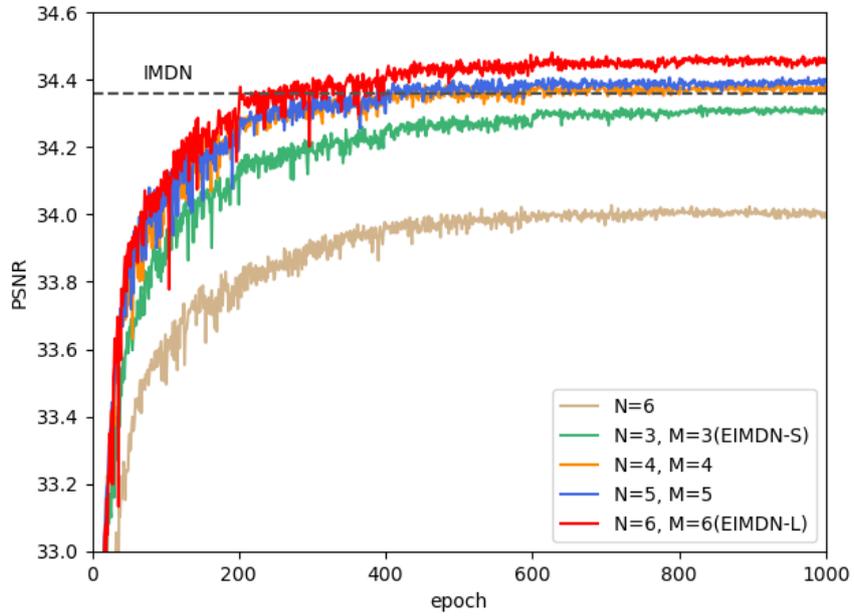
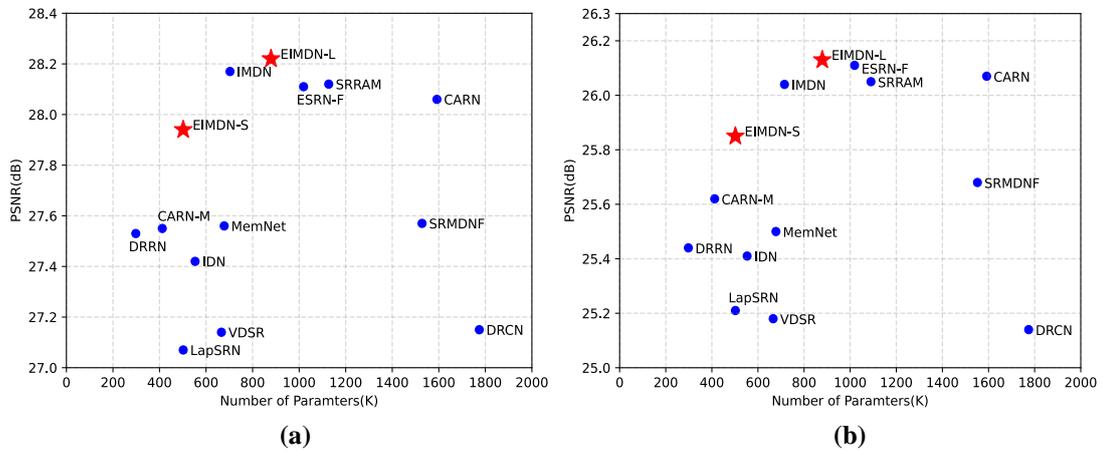**Fig. 6.** The influence of the number of EIMDB and IMDB in the network on the Set5 (scale: ×3).

Fig. 7. The EIMDN-L and EIMDN-S (red star) parameters and performance comparison with SOTA method in Urban100 dataset and up-scaling factor **(a)** ×3 , **(b)** ×4.
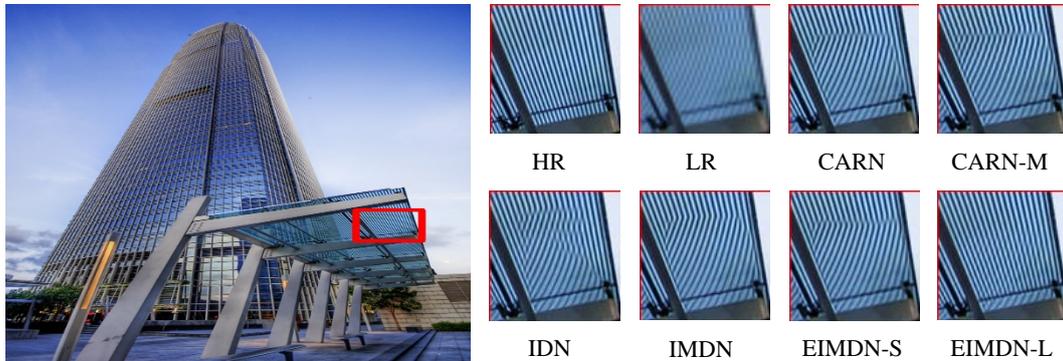
Fig. 8. The EIMDN-S and EIMDN-L visual results compare with SOTA image super-resolution reconstruction results with up-scaling ×2 in "Img_046" of Urban100.

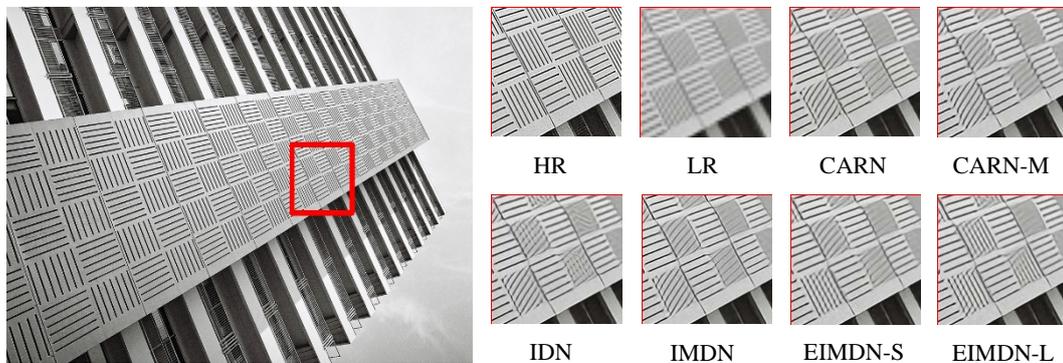**Fig. 9.** The EIMDN-S and EIMDN-L visual results compare with SOTA image super-resolution reconstruction results with up-scaling ×3 in "Img_092" of Urban100.

Fig. 10. The EIMDN-S and EIMDN-L visual results compare with SOTA image super-resolution reconstruction results with up-scaling ×4 in "Barbara" of Set14.

With the increase of the number of EIMDB and IMDB, the number of parameters will increase slightly, and the performance of image reconstruction will also be improved. In the case of N=5 and M=5, the performance surpasses that of IMDN on PSNR and SSIM.

The N=6 module proves the necessity of setting middle feature extraction module and deep feature extraction module. As shown in **Fig. 6**, although the EIMDB can the number of huge parameters is reduced, the performance of image reconstruction is significantly reduced. Therefore, we propose a middle feature extraction module and a deep feature extraction

module to extract features at different levels, which can decrease parameters and calculation amount to pursue the performance of super-resolution image reconstruction.

4.5 The Parameters and The Computational Cost Comparison

4.5.1 The parameters comparison

For lightweight SR, it is important that decrease the network parameters to guarantee the reconstruction performance. As shown in **Fig. 7**, compared with the SOTA method in Urban100 dataset, our EIMDN-L achieved a good PSNR effect with a slight increase in the number of parameters. EIMDN-S also achieves better results in the case of lower number of parameters, and achieves the tradeoff between image super-resolution reconstruction and model size.

4.5.2 The computational cost comparison

We use the multi-add operation proposed by CARN [12] to evaluate the computational complexity of the model. Specifically, Multi-adds is the number of composites multiply-accumulate operations for a single image. We assume the HR image size to be 720p (1280 × 720) to calculate multi-adds. As shown in **Table 5**, compared with the SOTA methods, our EIMDN-S maintains a lower the computational cost. However, our EIMDN-L has slightly increased the computational cost while achieving better reconstruction results.

Table 5. Comparison of computational cost amount of the model. (Unit: G)

Scale	VDSR	DRRN	IDN	CARN	CARN-M	IMDN	EIMDN-S	EIMDN-L
×2	612.6	6796.9	124.6	222.8	91.2	158.8	139.2	242.7
×3	612.6	6796.9	56.3	118.8	46.1	71.5	62.9	108.9
×4	612.6	6796.9	32.3	90.9	32.5	40.9	36.1	62.0

4.6 Visual Effect Comparison

The IDN, CARN-M, CARN and IMDN methods are selected to compare the visual effects of reconstructed images from Set14 and Urban100 datasets under the conditions of up-scaling factor ×2, ×3 and ×4. As shown in **Fig. 8**, our EIMDN-L does a good job of restoring the correct texture of glass grating from the "Img_046" image in Urban100. The EIMDN-S is like other lightweight methods for reconstruction. As shown in **Fig. 9** and **Fig. 10**, from the "Img_062" and the "Barbara", we can also observe that our reconstruction effect is more favorable and more details can be recovered. Our EIMDN-S has similar performance compared with the SOTA methods, but the EIMDN-L method is significantly superior to the contrast method.

5. Conclusion and Future Work

We propose a lightweight SISR network, EIMDN with Ghost module and CA attention mechanism based of IMDN, which reduces the number of parameters and improves the extraction of key features. We also used a feedback mechanism to enhance the fusion between low and high features and further enhance the reconstruction effect. Compared with the existing SOTA method, our method achieves good results on ×3 and ×4 up-scaling factors, which are 0.05dB and 0.09dB higher than IMDN on Urban100 dataset, respectively. Future

work will use Transformer to enhance our network performance or apply our network in different dataset (like medical or remote sensing dataset). Also, we will apply our method to IoT for edge computing or portable devices.

Acknowledgement

References

[1]   M. Long, F. Peng, and Y. Zhu, "Identifying natural images and computer-generated graphics based on binary similarity measures of PRNU," *Multimedia Tools and Applications*, vol. 78, no. 1, pp. 489-506, May. 2019. Article (CrossRef Link)

[2]   Y. Gui, and G. Zeng, "Joint learning of visual and spatial features for edit propagation from a single image," *Visual Computer*, vol. 36, no. 3, pp. 469-482, Jun. 2020. Article (CrossRef Link)

[3]   D. Zhang, T. Yin, G. Yang, M. Xia, L. Li, and X. Sun, "Detecting image seam carving with low scaling ratio using multi-scale spatial and spectral entropies," *Journal of Visual Communication and Image Representation*, vol. 48, no. 1, pp. 281-291, Feb. 2017. Article (CrossRef Link)

[4]   X. Zhang, F. Peng, and M. Long, "Robust coverless image steganography based on DCT and LDA topic classification," *IEEE Transactions on Multimedia*, vol. 20, no. 12, pp. 3223-3238, Oct. 2018. Article (CrossRef Link)

[5]   J. Wang, Y. Wu, L. W, L. W, O. Alfarraj, and A. Tolba, "Lightweight feedback convolution neural network for remote sensing images super-resolution," *IEEE Access*, vol. 9, no. 1, pp. 15992-16003, Feb. 2021. Article (CrossRef Link)

[6]   C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, Nov. 2016. Article (CrossRef Link)

[7]   B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1132-1140, 2017. Article (CrossRef Link)

[8]   Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. of European Conference on Computer Vision*, pp. 294-310, 2018. Article (CrossRef Link)

[9]   N. Ahn, B. Kang, and K. Sohn, "Fast, accurate, and lightweight super-resolution with cascading residual network," in *Proc. of European Conference on Computer Vision*, pp. 252–268, 2018. Article (CrossRef Link)

[10] Z. Hui, X. Gao, Y. Yang, and X. Wang, "Lightweight image super-resolution with information multi-distillation network," in *Proc. of ACM the 27th International Conference on Multimedia*, pp. 2024-2032, 2019. Article (CrossRef Link)

[11] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: more features from cheap operations," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1577-1586, 2020. Article (CrossRef Link)

[12] A. R. Zamir, W. Wu, L. Sun, W. B. Shen, and B. E. Shi, "Feedback networks," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1808-1817, 2017. Article (CrossRef Link)

[13] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," *arXiv preprint arXiv:2103.02907*, 2021. Article (CrossRef Link)

[14] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1637-1645, 2016. Article (CrossRef Link)

[15] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. of European Conference on Computer Vision*, pp. 391-407, 2016. Article (CrossRef Link)

[16] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1874-1883, 2016. Article (CrossRef Link)

[17] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, and W. Gao, "Pre-trained image processing transformer," *arXiv preprint arXiv:2012.00364*, 2021. Article (CrossRef Link)

[18] Z. Hui, X. Wang, and X. Gao, "Fast and Accurate single image super-resolution via information distillation network," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 723-731, 2018. Article (CrossRef Link)

[19] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, and W. Wu, "Feedback network for image super-resolution," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3862-3871, 2019. Article (CrossRef Link)

[20] E. Agustsson, and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: dataset and study," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1122-1131, 2017. Article (CrossRef Link)

[21] B. Marco, R. Aline, G. Christine, and L. Marie, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. of British Machine Vision Conference*, pp. 135.1-135.10, 2012. Article (CrossRef Link)

[22] Z. Roman, E. Michael, and P. Matan, "On single image scale-up using sparse-representations," in *Proc. of Springer Curves and Surfaces*, pp. 711-730, 2010. Article (CrossRef Link)

[23] D. Martin, C. Fowlkes, D. Tal, J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. of IEEE International Conference on Computer Vision*, pp. 416-425, 2001. Article (CrossRef Link)

[24] J. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp.5197-5206, 2015. Article (CrossRef Link)

[25] Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamasaki, and K. Aizawa, "Sketch-based manga retrieval using manga109 dataset," *Multimedia Tools and Applications*, vol. 76, no. 20, pp. 21811-21838, Nov. 2017. Article (CrossRef Link)

[26] D. P. Kingma, and B. Jimmy, "Adam: a method for stochastic optimization," in *Proc. of IEEE International Conference on Learning Representations*, 2015. Article (CrossRef Link)

[27] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1646-1654, 2016. Article (CrossRef Link)

[28] W. S. Lai, J. B. Huang, N. Ahuja, and M. H. Yang, "Deep laplacian pyramid networks for fast and accurate super resolution," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5835-5843, 2017. Article (CrossRef Link)

[29] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual Network," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2790-2798, 2017. Article (CrossRef Link)

[30] Y. Tai, J. Yang, X. Liu, and C. Xu, "Memnet: A persistent memory network for image restoration," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4539–4547, 2017. Article (CrossRef Link)

[31] K. Zhang, W. Zuo, and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3262–3271, 2018. Article (CrossRef Link)

[32] J. Kim, J. Choi, M. Cheon, and J. Lee, "MAMNet: Multi-path Adaptive Modulation Network for Image Super-Resolution," *arXiv preprint arXiv:1811.12043*, 2018. Article (CrossRef Link)

[33] D. Song, C. Xu, X. Jia, Y. Chen, C. Xu, and Y. Wang, "Efficient residual dense block search for image super-resolution," in *Proc. of Educational Advances in Artificial Intelligence*, vol. 34, no. 7, pp. 12007-12014, 2020. Article (CrossRef Link)

[34] A. Vaswani, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. of Neural Information Processing Systems*, pp. 5998-6008, 2017. Article (CrossRef Link)

[35] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and F. Li, "ImageNet: a large-Scale hierarchical image database," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248-255, 2009. Article (CrossRef Link)

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *Proc. of IEEE International Conference on Computer Vision*, pp. 1026-1034, 2015. Article (CrossRef Link)

[37] A. G. Howard, M. Zhu, B. Chen, D. Kalenicheko, W. Wang, T. Weyand, M. Andreeto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprint arXiv:1704.04861*, 2017. Article (CrossRef Link)

[38] D. Zhou, Y. Chen, J. Feng, and S. Yan, "Rethinking Bottleneck Structure for Efficient Mobile Network Design," in *Proc. of European Conference on Computer Vision*, pp. 680-697, 2020. Article (CrossRef Link)

Jin Wang received the B.S. and M.S. degree from Nanjing University of Posts and Telecommunications, China in 2002, 2005 respectively. He received Ph.D. degree from Kyung Hee University Korea in 2010. Now, he is a professor at Changsha University of Science and technology. He has published more than 300 international journal and conference papers. His research interests mainly include wireless sensor network, network performance analysis and optimization. He is a senior member of IEEE and a member of ACM.

Yiming Wu received the B.S. degree in Jilin Agricultural University, China, in 2018. He is a postgraduate student in Changsha University of Science and Technology, Changsha, China. His research interests include computer vision, super resolution, and deep learning. He is good at Python and C.

Shiming He received the Ph.D. degree in computer application from Hunan University, China, in 2013. She is an associate professor in School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, China. Her current research interests include privacy preserving, wireless network and mobile computing.

Pradip Kumar Sharma received the Ph.D. degree in computer science engineering from Seoul National University of Science and Technology, in 2019. He is an assistant professor in Cybersecurity in the Department of Computing Science at the University of Aberdeen, Aberdeen, UK, He is listed in the world's Top 2% Scientists for citation impact during the calendar year 2019 by Stanford University. Also, he received a top 1% reviewer in computer science by Publons Peer Review Awards 2018 and 2019, Clarivate Analytics.

Xiaofeng Yu received his Ph.D. degree from Nanjing University, China, in 2007. He is currently an associate professor in the School of Business, Nanjing University. His current research interests include strategy, technologies and applications of Electronic Business.

Osama Alfarraj received the master's and Ph.D. degrees in information and communication technology from Griffith University, in 2008 and 2013, respectively. He is currently an Associate Professor of computer sciences with King Saudi University, Riyadh, Saudi Arabia. His current research interests include eSystems (eGov, eHealth, and ecommerce), cloud computing, and big data. He has served as a Consultant and a member of the Saudi National Team for Measuring E-Government, Saudi Arabia, for two years.

Amr Tolba received the M.Sc. and Ph.D. degrees from Mathematics and Computer Science Department, faculty of science, Menoufia University, Egypt, in 2002 and 2006, respectively. He is currently an Associate Professor at the Faculty of Science, Menoufia University, Egypt. He is currently on leave from Menoufia University to the Computer Science Department, Community College, King Saud University (KSU), Saudi Arabia. Dr Tolba serves as a technical program committee (TPC) member in several conferences. He has authored/coauthored over 90 scientific papers in top ranked (ISI) international journals. His main research interests include socially aware networks, vehicular ad-hoc networks, Internet of Things, intelligent systems, Big Data, recommender systems, and cloud computing.