IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2021, accepted January 12, 2021, date of publication January 19, 2021, date of current version January 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052946

Lightweight Feedback Convolution Neural Network for Remote Sensing Images Super-Resolution

JIN WANG[1,2], (Senior Member, IEEE), YIMING WU[1], LIU WANG[1], LEI WANG[3], OSAMA ALFARRAJ[4], AND AMR TOLBA[4,5]

[1]Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410000, China
[2]School of Information Science and Engineering, Fujian University of Technology, Fuzhou 350118, China
[3]School of Civil Engineering, Changsha University of Science and Technology, Changsha 410000, China
[4]Department of Computer Science, Community College, King Saud University, Riyadh 11437, Saudi Arabia
[5]Department of Mathematics and Computer Science, Faculty of Science, Menoufia University, Shebin El-Kom 32511, Egypt

Corresponding author: Lei Wang (leiwang@csust.edu.cn)

This work was supported in part by the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia through the Research Group under Grant RG-1438-070, in part by the National Natural Science Foundation of China under Grant 61772454 and Grant 62072056, and in part by the National Key Research and Development Program of China under Grant 2019YFC1511000.

ABSTRACT There are lots of image data in the field of remote sensing, most of which have low-resolution due to the limited image sensor. The super-resolution method can effectively restore the low-resolution image to the high-resolution image. However, the existing super-resolution method has both heavy computing burden and number of parameters. For saving costs, we propose the feedback ghost residual dense network (FGRDN), which considers the feedback mechanism as the framework to attain lower features through high-level refining. Further, for feature extraction, we replace the convolution of the residual dense blocks (RDBs) with ghost modules (GMs), which can remove the redundant channels and avoid the increase of parameters along with the network depth. Finally, the spatial and channel attention module (SCM) is employed in the end of the RDB to learn more useful information from features. Compared to other SOTA lightweight algorithms, our proposed algorithm can reach convergences more rapidly with fewer parameters, and the performance of the network can be markedly enhanced on the image texture and object contour reconstruction with better peak signal-to-noise ratio (PSNR) and structural similarity (SSIM).

INDEX TERMS Remote sensing, super-resolution, feedback mechanism, ghost module, attention mechanism.

I. INTRODUCTION

High-resolution remote sensing images play an important role in civil engineering, urban planning, agricultural monitoring, and military reconnaissance. However, the limitations of image sensors have resulted in low resolution of pictures taken by remote sensing satellites and wireless sensors. So, how to restore low-resolution remote sensing images to high-resolution images through algorithms has become a current research hotspot. Super-resolution [1]–[4] not only restores sharpness from lower-resolution images, but also saves the cost of replacing or upgrading camera sensors because the super-resolution method is to process the photos taken. Therefore, super-resolution technology has great research value in remote sensing image processing [5].

In the field of computer vision, single-image super-resolution (SISR) is an ill-posed, low-level problem [6]–[9] that is often integrated with other industrial fields [10]–[15]. In resent years, the field of SISR has been widely used deep learning algorithms [16]–[19]. The first deep learning-based SISR algorithm was the super-resolution using convolutional neural networks (SRCNN), proposed by Dong *et al.* [20]. This method uses a three-layer convolution structure and achieves a better reconstruction effect than the traditional method. As network depth increased, the number of parameters also increased; this, however, takes up a lot of computing resources, and leads to over-fitting or under-fitting problems. The residual network (Res-Net) was then proposed by

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues.

He *et al.* [21] in the field of image recognition, using a short-cut connection between network blocks, which greatly increased the network depth. This not only solves the over-fitting or under-fitting problem in deep convolutional networks, thereby making it easier to optimize the network, but also enables the network to learn deeper features. The very deep super-resolution (VDSR) method was proposed by Kim *et al.* [22], This method first introduced the global residual into the SR field and increased the number of network layers to 20 by using the structural-adjustable gradient clipping technique. Subsequently, Lim *et al.* [23] proposed the enhanced deep super-resolution (EDSR) network, which deepened the network by stacking multiple residual units, introduced global residual learning, and finally built a network model with a special reconstruction magnification scale. At the same time, based on EDSR, Lim *et al.* constructed a multi-scale enhanced deep super-resolution (MDSR) network by adding processing modules with different reconstruction magnification scales at the front and rear ends of the network.

At the same time, the connection between the convolutions has become a research focus. The SISR using dense-connection network (SR-Dense Net) was proposed by Tong *et al.* [24]. They added dense skip connections between each convolutional layer, enabling the feature information extracted by the different levels of convolutional layers to be fully utilized. Moreover, Zhang *et al.* [25], fusing the residual connects and dense connect, proposed the residual dense network (RDN). The residual connection was introduced based on the dense unit, while the advantages of residual connection and dense connection were combined to create residual dense blocks (RDBs). RDBs based on hierarchical feature fusion integrate the feature information both locally and globally in the network; this means that the feature information of each layer in the deep network can be more fully integrated and utilized, which is helpful for realizing a better reconstruction effect of the network. However, as the network depth increases, the calculation and parameters also increase, leading to the inefficiency of the image super-resolution algorithm.

Some deep learning-based super-resolution methods have reduced the number of network parameters in order to adapt to sensor devices while still ensuring the image recovery effect. Hui *et al.* [26] proposed the information distillation network (IDN) for super-resolution by means of information distillation, cleverly utilizing the concept of information distillation to divide the feature extraction module into an information enhancement unit and a compression unit that respectively extract and remove redundant information from features. Hui *et al.* [27] subsequently proposed the information multi-distillation network (IMDN). In the basic block, feature extraction convolution is sliced channel-wise. The original convolution task is decomposed into subsequent convolution tasks with fewer channels in each layer, and the channel attention mechanism is then added to reduce the number of parameters of the original convolution and achieve better performance than IDN. Li *et al.* [28] introduced the concept of the feedback mechanism [29] into the super-resolution context, accordingly proposing the super-resolution feedback network (SRFBN). The main framework of this network adopts RNN to pass the high-level features of the last iteration into the low-level features of the current iteration. RNN is used to reduce the number of network parameters while ensuring the extraction of high-level features. Li *et al.* [30] then proposed the gated multiple feedback network (GMFN), which passed the high-level features of the different levels of this iteration into the low-level features of the next iteration. While maintaining the reconstruction effect of model depth, the network fuses the high and low-level features.

Our proposed method is based on the feedback mechanism used in GMFN to pass the high-level features of this output into the low-level feature extraction in the next iteration. The network causes the low-level features refined by the high-level features and builds the feedback framework. The network also introduces the Ghost module [31] (GM) from the field of image classification. This module abandons the redundant channels in feature mapping, then learns the information of the remaining channels in order to supplement the discarded redundant channels. It also cuts down the parameters while maintaining the high-lever information. Accordingly, we improved the RDB by replacing the convolution operation with GM.

The research method of attention mechanism comes from the study of human eye system. In science of perception, because of vision information bottlenecks, people usually concentrate on a part of vision information and neglect others. Accordingly, at the end of the RDB module, we added the spatial and channel attention module (SCM) to increase the network's ability to extract useful feature information. The attention mechanism parameters added by GM reduce the overall number of parameters to even less than the number in the GMFN network model. The main contributions in this paper can be summarized as follows:

1) We propose the FGRDN. The features extracted in the high layer features can refine the features in the next iteration shallow layer. While the deeper features extracted layer can be extracted without increasing the network depth.
2) The ghost residual dense blocks (GRDBs) are proposed, which replace the feature extraction convolutional layer in the RDB in the deep feature extraction. This can achieve the feature extraction function of ordinary convolution, while at the same time also removing the redundant channel in the feature channel and reducing the number of parameters in the network.
3) The SCM is joined at the end of each GRDB. This can make the network get more useful information in LR images. It can also enhance the feature mapping expression ability, effectively recover more texture, and contour details, and obtain a better high-resolution image reconstruction effect.

The rest of this paper is organized as follows. In Section II, we introduce the relevant research related to this paper. In Section III, we explain the structure of the entire network and the composition of each module. In Section IV, we introduce the experiments, along with a comparison with existing lightweight algorithms. The finally Section summarize our proposed method.

II. RELATED WORK

A. FEEDBACK MACHANISM

The concept of the feedback mechanism [29] is based on the method of image classification. In particular, feedback-based forecasting has two requirements: 1. network iteration (like RNN); 2. in each iteration, the posterior information is reintroduced into the network, while the loss is instantiated by using the recursive network model and associating with each iteration. Moreover, the feedback mechanism has the following advantages: 1. predictions can be made early, while the loss and outputs can be calculated separately in each iteration; 2. the result of feature prediction extracted by the network conforms to the hierarchical method of extracting from rough features to fine features, and the high-level features purify the low-level by the feedback mechanism. The SRFBN [28] is first time use the feedback mechanism in the field of SISR and concatenates the last features of the previous iteration into the coarse feature extraction of this iteration. The GMFN [30] is optimized from SRFBN, the backbone of the network is converted to the stack of RDB for deep feature extraction, and the gated feedback module (GFM) is used to link several deep RDB operations and pass them to the coarse feature extraction in the next iteration. By means of feedback mechanism, more information is transferred from the deep to the shallow features, so that the shallow features can be refined. However, this feedback mechanism has some shortcomings: for example, the loss does not converge when the number of iterations is too high, while the backbone network lacks the ability to extract high-level feature information. We changed the iteration times of the feedback network to two and used the attention mechanism to increase the extraction of high-level feature information from the backbone network, thereby achieving a good reconstruction effect.

B. ATTENTION MECHANISM

The basic idea behind the attention mechanism involves letting the network focus on the important information in the feature and ignore other irrelevant information. There is a large body of research on the attention mechanism in the field of super-resolution. The attention mechanism can be bifurcated into channel and spatial attention according to its different characteristics. The channel attention mechanism retains channel features that contain more high-level information through the learned channel weights. The network of residual channel attention (RCAN) was proposed by Zhang *et al.* [32], which added the squeeze-excitation (SE) module [33] to the end of the local residual block by employing the resistant-in-residue (RIR) network structure and recalibrated the feature channels using the correlation between channels. The spatial attention mechanism transforms the spatial information in the original image into another space by learning the high-level feature information between pixels and retaining the key information. The super-resolution residual attention module (SRRAM) [34] fuses channel attention and spatial attention in one module and uses depth-wise convolution to conduct the feature extraction of spatial attention. At the same time, the SE module is used to adjust the attention of the input feature channel. Finally, the spatial and channel attention modules are added; these modules are activated via sigmoid function, after which element multiplication is done with input features. SRRAM achieves a good super-resolution image reconstruction effect by combining spatial and channel attention; however, the network adopts feedforward propagation caused by the simple residual module stack, which reduces the connection between low-level and high-level features. By adding an attention mechanism into feedback communication, we use the high-level features to affect the low-level features, while at the same time extracting more effective high-level information through the attention mechanism.

C. GHOST MODULE

The concept of the GM was first proposed by Han *et al.* [31] as an image classification method. The GM uses the characteristics of deep learning to discard the feature map in the redundant channels, under the premise of ensuring that network parameters are reduced. In more detail, let us assume that the number of input channels is T. The GM decomposes the convolution operation into two parts. The first is the primary convolution, which uses ordinary convolution to extract the features and outputs the number of feature channels as S ($0 < S < T$). The second is a cheap operation, which carries out a series of simple linear operations on the output features of the first part; here, the number of output feature channels is T-S. Finally, the two outputs are concatenated to ensure that the number of output feature channels is equal to the number of input features. Accordingly, we replace the GM in the RDB feature extraction module convolution operation, thereby reducing a large number of network parameters while ensuring the recovery quality of the super-resolution image.

III. PROPOSED METHOD

This section will provide a detailed introducing to the proposed FGRDN from 6 parts: Part A introduces the whole operation process of our FGRDN network. Part B provides an introduction to the gated feedback module (GFM). Part C discusses the GRDBs. Part D describes the GM redundancy removal, while Part E presents the spatial and channel attention module (SCM) Finally, Part F presents the design of the loss function.

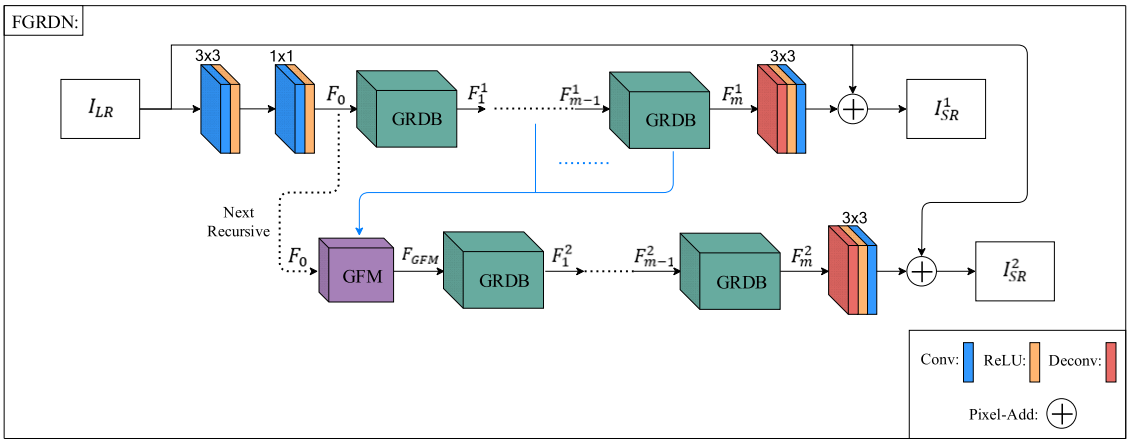

FIGURE 1. The FGRDN network framework contains two iterations. Here, F_1^1 and F_1^2 represent the output of the first GRDB in two iterations, respectively.

A. FEEDBACK GHOST RESIDUAL DENSE NETWORK

The network structure design is inspired by the GMFN proposed by Li *et al.* [30]. Our network is shown in the Figure 1 with two iterations. In the first iteration, the RGB gamut in the low-resolution image I_{LR} is segmented into three-layer feature channels via mean-shift. At the same time, a bilinear interpolation algorithm is used to carry out an up-sampling operation for the mean-shift three-layer feature channel in order to prepare for the subsequent image reconstruction. Next, the features extracted from the low-level layer are input into GRDBs for deep layer feature extraction. The feature extracted from deep layer and the I_{LR} up-sampled by bilinear interpolation are input into the image reconstruction module. The output of the matching SISR I_{SR}^1 is only used for the calculation of the loss function.

The second iteration takes the output of the last four GRDBs in the first iteration, then inputs the GFM. It combines the high-level features of the last iteration with the low-level features and passes into the high-level feature extraction module through the same reconstruction module. The output of this iteration is the super-resolution image I_{SR}^2; this output is not only used to calculate the loss function, but also acts as the super-resolution image output of the whole network.

As for the formula expression of the overall framework, we will explain it in three parts: low-level feature extraction, high-level feature extraction and image reconstruction.

1) LOW-LEVEL FEATURE EXTRACTION

Low-level feature extraction mainly works to extract shallow features from the input low-resolution image. It contains two convolution operation layers; the convolution kernel sizes are 3×3 and 1×1 respectively, and the number of channels is 64 layers. The above operation procedure is shown as follow:

$$F_0 = f_{shallow}\left(I_{LR}\right) \tag{1}$$

where I_{LR} represents the input low-resolution image, $f_{shallow}(\cdot)$ represents the shallow feature extraction function, and F_0 represents the low-level feature extraction.

2) HIGH-LEVEL FEATURE EXTRACTION

The high-level feature extraction module contains m GRDBs. The main purpose is to extract the edge and texture details in feature mapping through these GRDBs. The operation of the two iterations are represented as follows:

$$F_m^1 = F_{m-1}^1(F_{m-2}^1(\cdots F_2^1\left(F_1^1(F_0)\right)\cdots)), \quad (t=1) \tag{2}$$

$$F_m^2 = F_{m-1}^2(F_{m-2}^2(\cdots F_2^2\left(F_1^2(F_{GFM}(F_0))\right)\cdots)), \quad (t=2) \tag{3}$$

where F_{GFM} represents the gated feedback module, while $F_1^1 \sim F_m^1$ represents the output from the first GRDB to the last GRDB in the first iteration. $F_1^2 \sim F_m^2$ respectively denote the first GRDB to the last GRDB in the second iteration; moreover, m is the number of GRDBs ($m = 8$), while t is the number of iterations ($t = 2$).

3) IMAGE RECONSTRUCTION

We use a transpose and a convolution with 3×3 kernel to realize the part of image reconstruction. So, the extracted features and low-resolution images are then added to each other using the results of interpolation sampling. The output is the reconstructed super-resolution image, which can be represented as as follow:

$$I_{SR}^t = f_{RB}\left(F_m^t\right) + f_{UP}\left(I_{LR}\right) \tag{4}$$

where f_{UP} represents the interpolation up-sampling operation, F_m^t represents the output of the high-level feature extraction. F_{RB} is the image reconstruction function, and I_{SR}^t represents the super-resolution image output of the t^{th} iteration.

B. GATING FEEDBACK MODULE

The gating feedback module (GFM) originates from GMFN [30]. As shown in Figure 3, the GFM consists of two

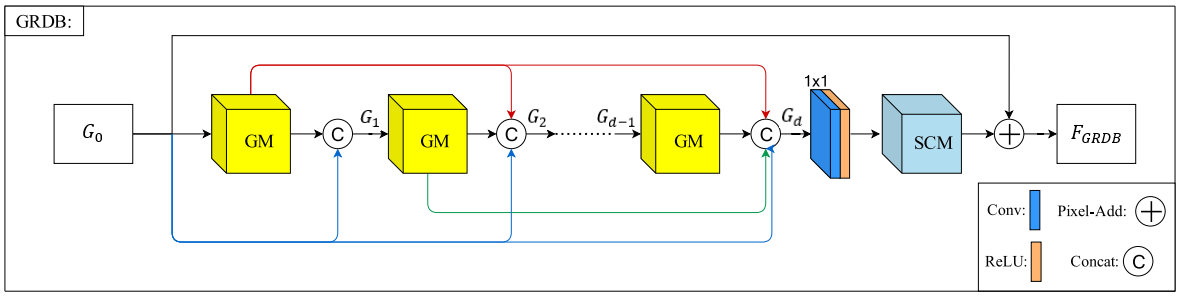

FIGURE 2. The ghost residual dense block illustrated.

concatenation operations and two convolution kernels of size 1×1. The purpose of this module is to adaptively select and enhance the high-level information from the multiple high-level features of this iteration, as well as to refine the low-level information by using convolution operation to select meaningful high-level information. In the next step, the refined low-level information is passed into the next iteration through concatenate operations. To be more specific, the output of the last four GRDBs of the previous iteration is imported into the join operation; subsequently, the feature map after the join is convolved with 1×1. The number of feature channels is compressed from the 64×4 layer after the join to the 64-layer via 1×1 convolution. The convoluted feature map is then introduced into this iteration, after which the shallow feature output of this iteration is concatenated. Next, the connected feature map is convolved with 1×1. Through this convolution operation, the number of feature layers is compressed from 64×2 layers to 64 layers. Finally, the depth feature extraction part of this iteration can be expressed as follow:

$$F_{GFM} = f_{refine}\left(F^1_{m-b}, \cdots, F^1_m\right), \quad (0 < b < m \,\&\, b \in Z) \tag{5}$$

where $F^1_{m-b} \sim F^1_m$ represents the output from the m-b GRDBs in the previous iteration to the last GRDB ($b = 4$), while $f_{refine}(\cdot)$ is the feature refining function.

C. GHOST RESIDUAL DENSE BLOCK

The key components of the residual dense block are the residual connection of the module and the dense connection between feature extraction convolutions. The residual connection can minimize the gradient explosion and gradient disappearance phenomena, increasing the network depth, while the dense connection can enhance the information correlation between the feature extraction convolution. As shown in Figure 2, based on residual dense block (RDB), the GM replaces the original 3×3 feature to extract the convolution, while the SCM is added at the end of the network. The subsequent dth GMs are densely connected, and the first d-1 feature maps are input via concatenate operations. The output is a 64-layer feature map. There are a total of d GMs in a GRDB (d = 8), and the output of the last GM outputs

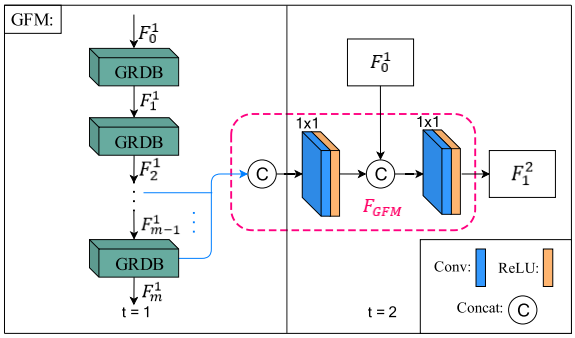

FIGURE 3. The gating feedback module illustrated.

passes the SCM. After multiplying the output feature of the attention module by the residual factor 0.2, and then adding the GRDB input, the added result is the output of GRDB. The above operation procedure can be showed as follow:

$$F_{GRDB} = \alpha(F_{SCM}(W_{1\times1}(G_d))) + G_0 \tag{6}$$

where G_0 is the input of the GRDB module, while $W_{1\times1}$ represents the convolution with a convolution kernel of 1×1, including the ReLU activation function. Finally, F_{SCM} represents the spatial channel attention module. Moreover, α represents the weighting factor of the dense residual module ($\alpha = 0.2$), while G_d represents the output of the last GM module ($d = 8$), which can be represented as follow:

$$F_{GRDB} = \alpha(F_{SCM}(W_{1\times1}(G_d))) + G_0 \tag{7}$$

where $G_1 \sim G_{d-1}$ represents the output from the first GM to the d-1^{th} GM module, F_{GM} is the operation of the GM module, and concat represents the cascading operation.

D. GHOST MODULE

The GM is composed of a 1×1 and a 3×3 convolution process as shown in Figure 4. The module is designed to remove the redundant feature mapping channels and then learn the redundant channels discarded from the retained feature map via convolution operation. Finally, the learned feature map is connected with the original feature map to determine the number of channels to the input. GM can not only make the convolution operation abandon similar feature mapping

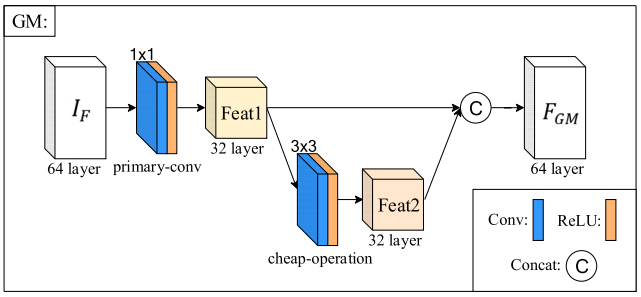

FIGURE 4. The Ghost Module illustrated.

channels, but also focuses more on useful feature mapping channels. Meanwhile, it can reduce the number of parameters and computational overhead of the network, making it suitable for lightweight networks. It can be represented in as follows:

$$F_{GM} = concat\left(Feat1, Feat2\right) \tag{8}$$

$$Feat1 = W_{1\times1}\left(I_F\right) \tag{9}$$

$$Feat2 = W_{3\times3}\left(Feat1\right) \tag{10}$$

where I_F represents the input feature map of the GM module, while $Feat1$ represents primary convolution, which comprises the convolution with convolution kernel size 1×1 and ReLU activation function. $Feat2$ represents a cheap operation consisting of a 3×3 convolution kernel and ReLU activation function. Finally, F_{GM} represents the output of the GM.

E. SPATIAL AND CHANNAL ATTENTION MODULE

As shown in Figure 5, we refer to the attention mechanism in SRRAM [34] to integrate the spatial attention and channel attention into one module. Channel attention adopts the squeeze-extraction (SE) method. First, the input features are averaged via global average pooling, and the number of channels is compressed via 1×1 convolution. The ReLU function is then activated to preserve the channel containing the information. Finally, the number of channels is reduced by means of 1×1 convolution. Spatial attention uses the depth-wise convolution to filter the information pixel by pixel. Add the output of spatial and channel attention to the pixel-multiple, and the result is adjusted via sigmoid function once according to the weight of each channel and each pixel,

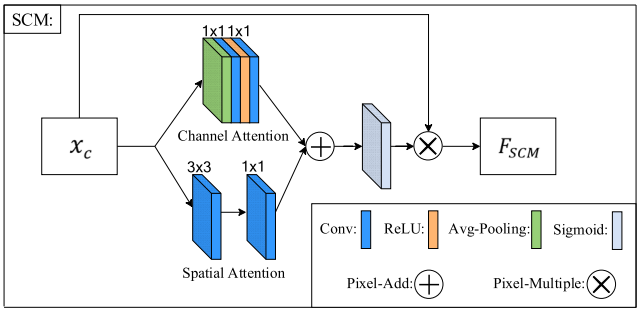

FIGURE 5. The spatial and channel attention module illustrated.

with a range of 0-1. The next step is to multiply the output of the Sigmoid function with the input of the attention module to preserve the important channels and spatial pixels. The above process can be obtained as follows:

$$F_{SCM} = x_c \cdot \sigma\left(F_{SA} + F_{CA}\right) \tag{11}$$

$$F_{SA} = W_P\left(W_G\left(x_c\right)\right) \tag{12}$$

$$F_{CA} = W_U\left(\delta\left(W_D\left(F_{GAP}\left(x_c\right)\right)\right)\right) \tag{13}$$

where F_{SCM} represents the output of the SCM, x_c represents the input of the SCM, and σ represents the sigmoid function. Moreover, F_{SA} represents the output of the spatial attention operation, W_P represents the pixel-to-pixel convolution operation with a 1×1 convolution, while W_G represents the group convolution with a convolution kernel size of 3×3 and a group of 64. Furthermore, F_{CA} represents the output of the channel attention operation, W_U represents the compressed convolution operation that reduces 64 channels to 4, and the convolution kernel is 1×1, W_D represents the expansion of the number of channels back to 64, the convolution kernel is 1×1, represents the ReLU activation function. Finally, F_{GAP} represents the global average pooling operation, which can be represented as follow:

$$F_{GAP} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c\left(i, j\right) \tag{14}$$

where H and W represent the height and width of the input feature map respectively, while i and j represent the horizontal and vertical coordinates of pixel points.

F. LOSS FUNCTION

We choose L1 loss to optimize our network. Referring to the loss function calculation of the feedback mechanism, we use the super-resolution image output of two iterations and the original high-resolution image I_{HR} to calculate the L1 loss function, then take the average value, which can be expressed as follow:

$$\mathcal{L}\left(\Theta\right) = \frac{1}{T} \sum_{t=1}^{T} \left|\left|I_{HR} - I_{SR}^{t}\right|\right|_1 \tag{15}$$

where Θ represents the network parameters, T denotes the number of iterations. I_{HR} and I_{SR}^{t} represent the original high-resolution images and super-resolution reconstruction images respectively.

IV. EXPERIMENTAL RESULTS AND DISCUSSINGS

A. PARAMETER SETTINGS

We train all networks with a batch size of 16. To fully exploit the contextual information from the LR images, we feed the RGB image patches with different patch sizes based on the upscaling factor. The network parameters are initialized using the He *et al.* [35] method. Adam [36] is employed to optimize the network parameters with an initial learning rate of 0.0002; the learning rate is then multiplied by 0.5 every 200 epochs. We used the Ubuntu 16.08 operating system. Memory size is 128GB. The CPU uses Intel Xeon E5-2670. The GPU uses Nvidia Titan X. The model was trained under the GPU version of Pytorch1.6.0.

TABLE 1. The PSNR/SSIM mean values of various models at ×2 up-scale factor in the standard test set.

Module	Up-scale factor	Set5 [38] PSNR/SSIM	Set14 [39] PSNR/SSIM	B100 [40] PSNR/SSIM	Urban100 [41] PSNR/SSIM	Manga109 [42] PSNR/SSIM
VDSR [22]	2	37.53/0.9587	33.03/0.9142	31.90/0.8960	30.76/0.9140	37.22/0.9750
DRCN [43]		37.63/0.9588	33.04/0.9118	31.85/0.8942	30.75/0.9133	37.55/0.9732
LapSRN [45]		37.52/0.9591	32.99/0.9124	31.80/0.8952	30.41/0.9103	37.27/0.9740
DRRN [44]		37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188	37.88/0.9749
MemNet [46]		37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195	37.72/0.9740
IDN [26]		37.83/0.9600	33.30/0.9148	32.08/0.8958	31.27/0.9196	38.01/0.9749
SRMDNF [47]		37.79/0.9601	33.32/0.9159	32.05/0.8958	31.33/0.9204	38.07/0.9761
CARN [48]		37.76/0.9590	33.52/0.9166	32.09/0.8978	31.92/0.9256	38.36/0.9765
SRRAM [34]		37.82/0.9592	33.48/0.9171	32.12/0.8983	32.05/0.9264	-/-
IMDN [27]		38.00/0.9605	33.63/0.9177	32.19/0.8996	32.17/0.9283	38.88/0.9774
FGRDN (ours)		37.97/0.9605	33.57/0.9179	32.12/0.8991	32.24/0.9293	38.58/0.9767

TABLE 2. The PSNR/SSIM mean values of various models at ×3 up-scale factor in the standard test set.

Module	Up-scale factor	Set5 [38] PSNR/SSIM	Set14 [39] PSNR/SSIM	B100 [40] PSNR/SSIM	Urban100 [41] PSNR/SSIM	Manga109 [42] PSNR/SSIM
VDSR [22]	3	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279	32.01/0.9340
DRCN [43]		33.82/0.9226	29.76/0.8311	28.80/0.7963	27.15/0.8276	32.24/0.9343
LapSRN [45]		33.81/0.9220	29.79/0.8325	28.82/0.7980	27.07/0.8275	32.21/0.9350
DRRN [44]		34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378	32.71/0.9379
MemNet [46]		34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
IDN [26]		34.11/0.9253	29.99/0.8354	28.95/0.8013	27.42/0.8359	32.71/0.9381
SRMDNF [47]		34.12/0.9254	30.04/0.8382	28.97/0.8025	27.57/0.8398	33.00/0.9403
CARN [48]		34.29/0.9255	30.29/0.8407	29.06/0.8034	28.06/0.8493	33.50/0.9440
SRRAM [34]		34.30/0.9256	30.32/0.8417	29.07/0.8039	28.12/0.8507	-/-
IMDN [27]		34.36/0.9270	30.32/0.8417	29.09/0.8046	28.17/0.8519	33.61/0.9445
FGRDN (ours)		34.41/0.9274	30.37/0.8432	29.08/0.8053	28.23/0.8542	33.62/0.9449

B. DATASETS

As shown in Table 4, the training set used in the experiment is DIV2K [37], which is derived from the high-quality image dataset used for the image repair task in the NITRE Challenge in 2017. It comprises 800 training sets, 100 verification sets and 100 test set images, each of these images have 2K resolution. The image types in DIV2K include figures, artifacts, environments (cities, villages), flora and fauna, and natural scenery. We use only 800 images from the training set for training. Each image was then amplified in three different ways: rotating the picture randomly by 90°, 180° and 270°; flipping the picture horizontally or vertically; or reducing the original image scale by 0.6∼0.9.

The test set uses Set5 [38], Set14 [39], BSD100 [40], Urban100 [41], and Manga109 [42], all widely used super-resolution benchmark test sets, to evaluate the model performance. Urban100 contains 100 images of challenging urban scenes, with densely detailed high-frequency fe atures. Manga109 contains 109 cartoon pictures, with high-frequency and low-frequency information and text information, that test the network of text and image processing ability.

C. PERFORMANCE COMPARISON

The above test sets are used for testing, and the following common performance evaluation indicators are used: Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM), which can be expressed as follows:

$$PSNR = 10 \cdot log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{16}$$

$$SSIM\ (x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{17}$$

The PSNR is widely used to objectively evaluate the distortion of reconstructed images. PSNR calculates the similarity based on the mean square error of two images, and the unit is decibel (dB). The larger the value of PSNR, the less distortion and the better effect of super-resolution image. The objects in natural scenes often have strong structural features, which are independent of illumination information. The SSIM estimates the similarity between the original image and the distorted image by combining the structural information, brightness information and contrast information of the image. The larger the SSIM value, the better the effect of super-resolution image.

All RGB images were converted to the YCbCr color space, and we selected the Y channel (luminance) for training and testing of the super-resolution model. We will use the network with different magnification training to test the corresponding magnification and compare the test results with the advanced super-resolution model developed in recent years. Table 1, 2, 3 present the quantitative comparison results

TABLE 3. The PSNR/SSIM mean values of various models at ×4 up-scale factor in the standard test set.

Module	Up-scale factor	Set5 [38] PSNR/SSIM	Set14 [39] PSNR/SSIM	B100 [40] PSNR/SSIM	Urban100 [41] PSNR/SSIM	Manga109 [42] PSNR/SSIM
VDSR [22]	4	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524	28.83/0.8870
DRCN [43]		31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510	28.93/0.8854
LapSRN [45]		31.54/0.8852	28.09/0.7700	27.32/0.7275	25.21/0.7562	29.09/0.8900
DRRN [44]		31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638	29.45/0.8946
MemNet [46]		31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630	29.42/0.8942
IDN [26]		31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632	29.41/0.8942
SRMDNF [47]		31.96/0.8925	28.35/0.7787	27.49/0.7337	25.68/0.7731	30.09/0.9024
CARN [48]		32.13/0.8937	28.60/0.7806	27.58/0.7349	26.07/0.7837	30.47/0.9084
SRRAM [34]		32.13/0.8932	28.54/0.7800	27.56/0.7350	26.05/0.7834	-/-
IMDN [27]		32.21/0.8948	28.58/0.7811	27.56/0.7353	26.04/0.7838	30.45/0.9075
FGRDN (ours)		32.25/0.8953	28.60/0.7824	27.58/0.7364	26.13/0.7877	30.57/0.9097

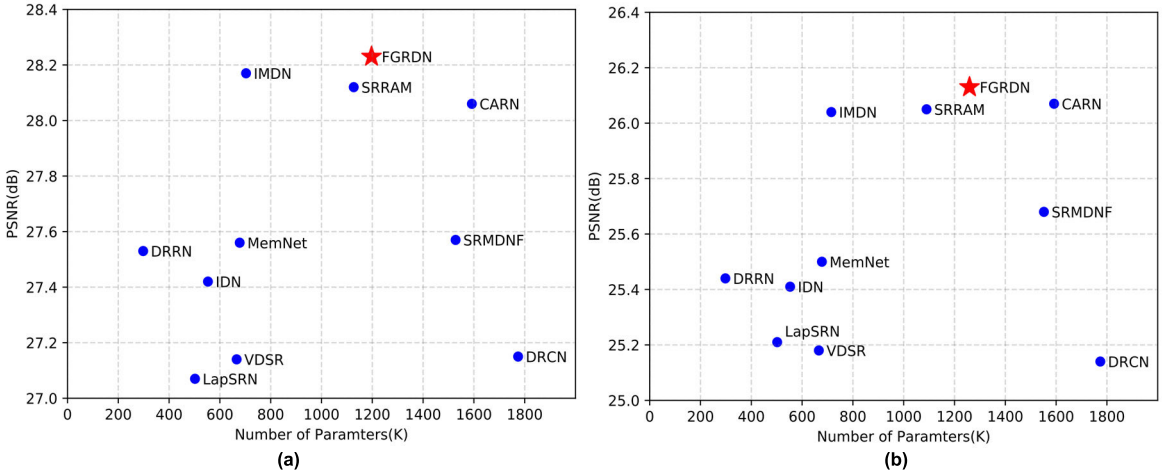

FIGURE 6. Performance and parameter comparison reconstructed for images in Urban100: (a) ×3 up-scale factor; (b) ×4 up-scale factor. FGRDN achieves the best PSNR with a relatively lower number of model parameters compared with the deep learning-based methods VDSR [22], DRCN [43], DRRN [44], LapSRN [45], IDN [26], MemNet [46], SRMDNF [47], IMDN [27], SRRAM [34], and CARN [48].

TABLE 4. The information of datasets.

Usage	Datasets	Number	Source
Training	DIV2K	800	ETH Zurich
Test	Set5	5	INRIA
Test	Set14	14	Israel Institute of Technology
Test	BSD100	100	University of California, Berkeley
Test	Urban100	100	University of Illinois
Test	Manga109	109	University of Tokyo

PSNR and SSIM using different super-resolution models at different up-scale factor of ×2, ×3 and ×4. Red and blue colors indicate the best and second best performance, respectively.

As can be seen from Table 1, compared to other models at an up-scale factor of ×2, our network does not always obtain the best effect. Using the GM can reduce the quantity and remove the channels of the redundancy, while each channel in the lower magnification retains more high-frequency detail, which leads to some important features of the channel having been lost in removing redundant steps, to reach the best results under ×2 magnification.

Moreover, as can be seen from the data in Table 2 and Table 3 below, most of the PSNRs and SSIMs in our network in the above test set either exceeded those of other super-resolution models or achieved second place. As we all know, as magnification increases, the difficulty of image super-resolution reconstruction will also increase. Our network is superior to most others under the conditions of ×3 and ×4 magnification. In the test set Manga109, the PSNR value of this network is 2.99dB higher than that of the classical model SRCNN and 0.13dB higher than that of the latest model IMDN under ×4 magnification condition. This shows that our network can transmit the high-frequency details that are difficult to learn in high magnification to this iteration through the feedback mechanism, thereby increasing the learning depth of the high-frequency information and achieving good reconstruction effects under high magnification.

D. NUMBER OF PARAMETER COMPARISON

To build a lightweight SR model, it is necessary to have fewer network parameters while also ensuring the image reconstruction effect. From Figure 6, we can observe that compared with other, most advanced methods, such as IDN [26], CARN [48], SRRAM [34] and IMDN [27], the proposed

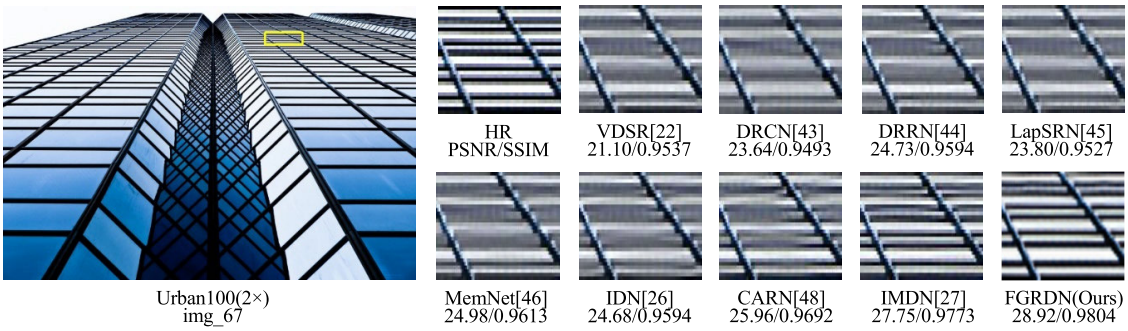

FIGURE 7. Visual effect comparison of img_67 image super-resolution reconstruction results under ×2 up-scale in Urban100 [41].

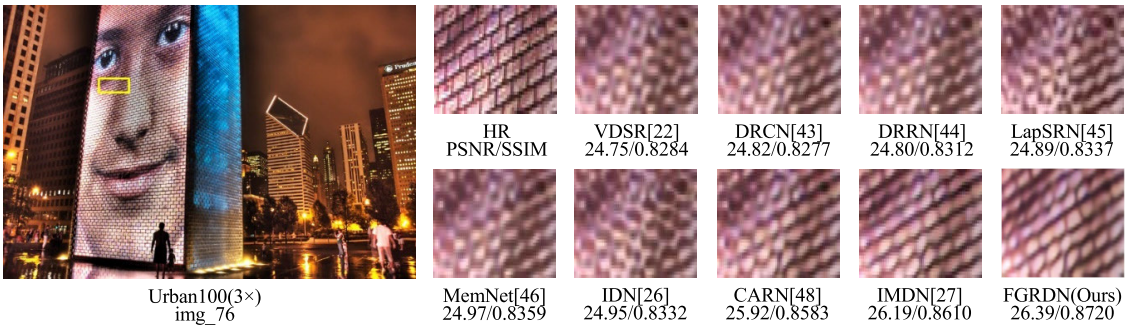

FIGURE 8. Visual effect comparison of img_76 image super-resolution reconstruction results under ×3 up-scale in Urban100 [41].

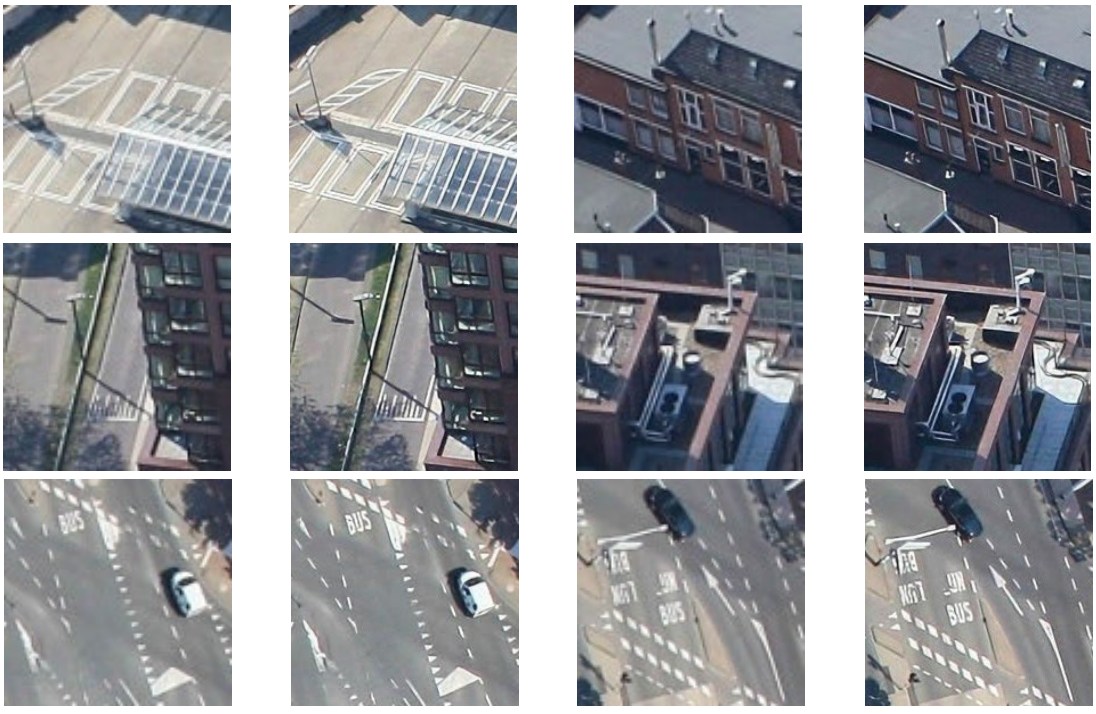

FIGURE 9. Visual effect of ground truth compared with super-resolution reconstruction results under ×2 up-scale in ITCVD [49].

FGRDN has fewer parameters and better image reconstruction performance, achieving a tradeoff between performance and model size.

E. VISUAL EFFECT COMPARISON

As can be seen from Figure 7, the img_67 image with the ×2 up-scale factor is a building with a glass frame structure,

which contains a lot of high-frequency information both laterally and diagonally. Most of the reconstructed models, such as VDSR [22], DRCN [43], DRRN [44], LapSRN [45], MemNet [46], IDN [26], and CARN [48], fail to restore the clear horizontal black lines. Moreover, although the model IMDN [27] can clearly recover the horizontal black lines, some of these lines have a jagged appearance, while the details are not sufficiently recovered. By contrast, our model was able to reconstruct more detail with a clearer horizontal black line.

As can be seen from Figure 8, moreover, the img_76 image of the $\times 3$ up-scale factor is a display combination building with face display. The vertical and diagonal high-frequency details are denser, which makes the extraction of high-frequency details more challenging for the models.

Accordingly, most models fail to clearly restore the slanted display border lines. As shown in model IDN, the low-frequency details are identified as high-frequency details, while the light-colored area is mistakenly restored to the display frame, which seriously affects the end visual result. By contrast, our model was able to restore more high-frequency detail and yield a clear image with a black border.

F. VISUAL EFFECT OF REMOTE SENSING IMAGES

We use the ITCVD [49] data set as our ground truth test data set. The ITCVD contains 135 images taken by an aircraft flying at an altitude of about 330m above Enschede in the Netherlands. We use the FGRDN model with $\times 2$ up-scale factor to perform super-resolution operation on six scenes of three images. Figure 9 shows the comparison between the ground truth image and the super-resolution image. The restoration effect of FGRDN on traffic lines and buildings is satisfactory. This can prove that our FGRDN has a great positive effect on the restoration of remote sensing images.

V. CONCLUSION

In this paper, we propose the feedback ghost residual dense network (FGRDN) of single-image super-resolution. The feedback mechanism is employed by the framework to enhance the correlation between the high-level and the low-level features. At the same time, we add an SCM to each GRDB, making the network more focused on high-level feature extraction. Finally, the ghost module replaces the convolutional feature extraction layer in the residual dense block; this module removes redundant feature channels and reduces the number of parameters while ensuring the reconstruction effect. After comparing some of the SOTA lightweight SISR algorithms in terms of the selected evaluation criteria and visual effects, our proposed approach obtains a better reconstruction effect. Future work will further reduce the number of parameters while continuing to ensure the effectiveness of the reconstruction; moreover, it will also try to train networks with a higher up-scale factor.

REFERENCES

[1] J. Qin, X. Sun, Y. Yan, L. Jin, and X. Peng, "Multi-resolution space-attended residual dense network for single image super-resolution," *IEEE Access*, vol. 8, pp. 40499–40511, 2020, doi: 10.1109/ACCESS.2020.2976478.

[2] P. Liu, Y. Hong, and Y. Liu, "Deep differential convolutional network for single image super-resolution," *IEEE Access*, vol. 7, pp. 37555–37564, 2019, doi: 10.1109/ACCESS.2019.2903528.

[3] K.-W. Hung, Z. Zhang, and J. Jiang, "Real-time image super-resolution using recursive depthwise separable convolution network," *IEEE Access*, vol. 7, pp. 99804–99816, 2019, doi: 10.1109/ACCESS.2019.2929223.

[4] H. Wang, Q. Hu, C. Wu, J. Chi, and X. Yu, "Non-locally up-down convolutional attention network for remote sensing image super-resolution," *IEEE Access*, vol. 8, pp. 166304–166319, 2020, doi: 10.1109/ACCESS.2020.3022882.

[5] S. Lei, Z. Shi, X. Wu, B. Pan, X. Xu, and H. Hao, "Simultaneous super-resolution and segmentation for remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. IGARSS*, Jul. 2019, pp. 3121–3124, doi: 10.1109/IGARSS.2019.8900402.

[6] Y. Chen, L. Liu, J. Tao, R. Xia, Q. Zhang, K. Yang, J. Xiong, and X. Chen, "The improved image inpainting algorithm via encoder and similarity constraint," *Vis. Comput.*, Jul. 2020, doi: 10.1007/s00371-020-01932-3.

[7] Y. Chen, J. Tao, L. Liu, J. Xiong, R. Xia, J. Xie, Q. Zhang, and K. Yang, "Research of improving semantic image segmentation based on a feature fusion model," *J. Ambient Intell. Humanized Comput.*, May 2020, doi: 10.1007/s12652-020-02066-z.

[8] J. Zhang, Z. Xie, J. Sun, X. Zou, and J. Wang, "A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection," *IEEE Access*, vol. 8, pp. 29742–29754, 2020.

[9] Y. Chen, V. Phonevilay, J. Tao, X. Chen, R. Xia, Q. Zhang, K. Yang, J. Xiong, and J. Xie, "The face image super-resolution algorithm based on combined representation learning," *Multimedia Tools Appl.*, Nov. 2020, doi: 10.1007/s11042-020-09969-1.

[10] J. Wang, W. Wu, Z. Liao, R. S. Sherratt, G.-J. Kim, O. Alfarraj, A. Alzubi, and A. Tolba, "A probability preferred priori offloading mechanism in mobile edge computing," *IEEE Access*, vol. 8, pp. 39758–39767, 2020.

[11] W. Lu, Y. Zhang, S. Wang, H. Huang, Q. Liu, and S. Luo, "Concept representation by learning explicit and implicit concept couplings," *IEEE Intell. Syst.*, early access, Sep. 2, 2020, doi: 10.1109/MIS.2020.3021188.

[12] J. Wang, Y. Tang, S. He, C. Zhao, P. K. Sharma, O. Alfarraj, and A. Tolba, "Logevent2vec: Logevent-to-vector based anomaly detection for large-scale logs in Internet of Things," *Sensors*, vol. 20, no. 9, p. 2451, 2020.

[13] J. Zhang, S. Zhong, T. Wang, H. Chao, and J. Wang, "Blockchain-based systems and applications: A survey," *J. Internet Technol.*, vol. 21, no. 1, pp. 1–14, 2020, doi: 10.3966/160792642020012101001.

[14] J. Wang, Y. Yang, T. Sherratt, R. S. Wang, and J. Zhang, "Big data service architecture: A survey," *J. Internet Technol.*, vol. 20, pp. 393–405, Mar. 2020.

[15] J. Wang, Y. Zou, L. Peng, L. Wang, O. Alfarraj, and A. Tolba, "Research on crack opening prediction of concrete dam based on recurrent neural network," *J. Internet Technol.*, vol. 21, pp. 1161–1169, Jul. 2020.

[16] L. Pan, J. Qin, H. Chen, X. Xiang, C. Li, and R. Chen, "Image augmentation-based food recognition with convolutional neural networks," *Comput., Mater. Continua*, vol. 59, no. 1, pp. 297–313, 2019.

[17] M. Okhovvat and M. R. Kangavari, "A mathematical task dispatching model in wireless sensor actor networks," *Comput. Syst. Sci. Eng.*, vol. 34, no. 1, pp. 5–12, 2019.

[18] C. W. Hung, W. L. Mao, and H. Y. Huang, "Modified PSO algorithm on recurrent fuzzy neural network for system identification," *Intell. Automat. Soft Comput.*, vol. 25, no. 2, pp. 329–341, 2019.

[19] W. Lu, X. Zhang, H. Lu, and F. Li, "Deep hierarchical encoding model for sentence semantic matching," *J. Vis. Commun. Image Represent.*, vol. 71, Aug. 2020, Art. no. 102794, doi: 10.1016/j.jvcir.2020.102794.

[20] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[22] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1646–1654.
[23] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1132–1140.
[24] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4809–4817.
[25] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jan. 21, 2020, doi: 10.1109/TPAMI.2020.2968521.
[26] Z. Hui, X. Wang, and X. Gao, "Fast and accurate single image super-resolution via information distillation network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 723–731.
[27] Z. Hui, X. Gao, Y. Yang, and X. Wang, "Lightweight image super-resolution with information multi-distillation network," 2019, *arXiv:1909.11856*. [Online]. Available: http://arxiv.org/abs/1909.11856
[28] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, and W. Wu, "Feedback network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3862–3871.
[29] A. R. Zamir, W. Wu, L. Sun, W. B. Shen, and B. E. Shi, "Feedback networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1808–1817.
[30] Q. Li, Z. Li, G. Jeon, X. Yang, L. Lu, and K. Liu, "Gated multiple feedback network for image super-resolution," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Cardiff, U.K., Sep. 2019, p. 188.
[31] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586.
[32] 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. Eur. Conf. Comput. Vis. (ECCV)*, Oct. 2018, pp. 294–310.
[33] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020.
[34] J.-H. Kim, J.-H. Choi, M. Cheon, and J.-S. Lee, "MAMNet: Multi-path adaptive modulation network for image super-resolution," 2018, *arXiv:1811.12043*. [Online]. Available: http://arxiv.org/abs/1811.12043
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
[36] D. P. Kingma and B. Jimmy, "Adam: A method for stochastic optimization," presented at Int. Conf. Learn. Represent., San Diego, CA, USA, May 2015.
[37] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1122–1131.
[38] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L.-A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. Brit. Mach. Vis. Conf.*, Sep. 2012, pp. 1–10.
[39] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. 7th Int. Conf. Curves Surf.*, Jun. 2010, pp. 711–730.
[40] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis. ICCV*, Jul. 2001, pp. 416–425.
[41] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5197–5206.
[42] Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamasaki, and K. Aizawa, "Sketch-based manga retrieval using manga109 dataset," *Multimedia Tools Appl.*, vol. 76, no. 20, pp. 21811–21838, Oct. 2017.
[43] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1637–1645.
[44] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2790–2798.
[45] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5835–5843.
[46] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4539–4547.
[47] K. Zhang, W. Zuo, and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3262–3271.
[48] N. Ahn, B. Kang, and K.-A. Sohn, "Fast, accurate, and lightweight super-resolution with cascading residual network," in *Proc. ECCV*, Sep. 2018, pp. 252–268.
[49] M. Yang, W. Liao, X. Li, and B. Rosenhahn, "Vehicle detection in aerial images," in *Proc. Int. Conf. Image Process. (ICIP)*, Athens, Greece, Oct. 2018, pp. 3079–3083.

JIN WANG (Senior Member, IEEE) received the B.S. and M.S. degrees from the Nanjing University of Posts and Telecommunications, China, in 2002 and 2005, respectively, and the Ph.D. degree from Kyung Hee University, South Korea, in 2010. He is currently a Professor with the Changsha University of Science and Technology. He has published more than 300 international journal and conference papers. His research interests include wireless sensor networks and network performance analysis and optimization. He is a member of ACM.

YIMING WU received the B.S. degree from Jilin Agricultural University, China, in 2018. He is currently pursuing the master's degree with the 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.

LIU WANG received the B.S. degree from the Jiangxi University of Science and Technology, China, in 2019. She is currently pursuing the master's degree with the Changsha University of Science and Technology, Changsha, China. Her research interests include computer vision, image inpainting, and deep learning. She is good at Python.

LEI WANG received the Ph.D. degree in bridge construction from the Changsha University of Science and Technology, China, in 2008. He is currently a Professor with the School of Civil Engineering, Changsha University of Science and Technology, China. He worked as a Visiting Scholar with Arizona State University, from 2013 to 2014, supported by the China Scholarship Council. His research interest includes reliability and safety evaluation of bridge structures.

OSAMA ALFARRAJ received the master's and Ph.D. degrees in information and communication technology (ICT) from Griffith University, in 2008 and 2013, respectively. His doctoral dissertation investigates the factors influencing the development of e-Government in Saudi Arabia and it is a qualitative investigation of the developers' perspectives. He is currently an Associate Professor of ICT with King Saud University, Riyadh, Saudi Arabia. His research interests include electronic commerce, M-government, the Internet of Things, cloud computing, and big data analytics

AMR TOLBA received the M.Sc. and Ph.D. degrees from the Department of Mathematics and Computer Science, Faculty of Science, Menoufia University, Egypt, in 2002 and 2006, respectively. He is currently an Associate Professor with the Faculty of Science, Menoufia University. He is on leave from Menoufia University to the Department of Computer Science, Community College, King Saud University (KSU), Saudi Arabia. He has authored or coauthored more than 50 scientific papers in top ranked (ISI) international journals and conference proceedings. His main research interests include socially aware networks, vehicular ad hoc networks, the Internet of Things, intelligent systems, and cloud computing.