Lakshmi Narayana R.K., et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 16, Issue 11, November 2026, pp 01-07

RESEARCH ARTICLE OPEN ACCESS

DRDA-Net: A Lightweight Dynamic Residual Directional-Attention Network for Single Image Super-Resolution

Lakshmi Narayana Reddy Kandi*, Durga Ganga Rao Kola*

**(Department of ECE, UCEK, JNTUK, India)*

Date of Submission: 10-11-2026 Date of acceptance: 22-11-2026

ABSTRACT
Single Image Super-Resolution (SISR) is the task of creating a clear and detailed image from a small, low-quality input image. This is useful in many real situations, such as improving old photos, medical scans, or security camera pictures. Many existing deep learning methods use very deep networks or several parallel paths to get good results, but this often makes the models large, slow, and hard to train. In this paper, DRDA-Net, a lightweight super-resolution network built around a single processing path instead of multiple paths, is presented. Each block in the network uses a simple attention step that helps it focus on important regions, such as edges and fine details, without adding heavy computation. A small helper module, called the Dynamic Residual Module (DRM), is also introduced, which learns how information from earlier blocks should be passed to later blocks, based on the input image itself. This allows the network to adjust its own internal connections for each image, instead of using the same fixed connections every time. DRDA-Net is tested on several widely used benchmark datasets and compared with existing super-resolution methods. The detailed results show that the network reaches a good balance between image quality and model size, making it a practical choice for real-world super-resolution tasks.

Keywords **-** Directional attention, dynamic residual connection, lightweight networks, residual learning, single image super-resolution

I. Introduction

Getting a clear, high-quality image from a small or low-quality one is a common problem in image processing. This task is called Single Image Super-Resolution (SISR). It is useful in many real applications, such as medical imaging, security cameras, and photo editing, where a sharp image is needed but only a low-resolution version is available.

In recent years, deep learning has become the main method used to solve this problem. Researchers have explored many different directions to improve image quality, such as making networks deeper, adding attention steps that highlight important parts of an image, or designing multiple paths inside a network to capture different kinds of features. Each of these ideas has helped in its own way, but they often come with extra cost, such as more parameters, more computation, or a more complex design that is harder to train.

In this research, these different ideas are studied, and the most useful ones are combined into a single, simple, and efficient network. Motivated by the idea that one deep path with residual learning can already give strong results, multiple parallel paths are avoided in the design. Paying more attention to important regions, such as edges and textures, is also found to improve results without needing heavy computation. In addition, allowing the network to decide, for each image, how information should move between blocks makes it more flexible and effective.

Based on these ideas, DRDA-Net, a new lightweight super-resolution network, is designed. The main goal is to build a network that is simple, efficient, and still gives good results. Rather than copying any single existing network design, useful ideas from this research area are combined in a new way to build modules that work together as one complete system.

The main contributions of this paper are:

- DRDA-Net, a lightweight super-resolution network that uses one main path instead of several parallel paths, which keeps the number of parameters low, is designed.
- A simple attention block, called the DiVA block, that helps the network focus on useful image regions with very little extra computation, is used.
- A Dynamic Residual Module (DRM) that learns, for each input image, how much information should flow from earlier blocks to later blocks, is considered.

- DRDA-Net is tested on standard benchmark datasets and compared with existing methods, showing that it gives a good balance between accuracy and model size.

II. Related Work

Super-resolution research has explored many different directions. In this section, the main ideas that are related to the work and how they connect to the design are discussed.

2.1 Deep Networks with Residual Learning

One of the earlier and most useful ideas in this area is to use a fairly deep network together with residual learning [1]. Instead of predicting the whole high-resolution image directly, the network only predicts the missing details, which are then added on top of a simple, upscaled version of the input image. This showed that a single path, if made deep enough, can already give strong results without needing several branches. This idea motivated the design of DRDA-Net around one main path instead of multiple parallel paths.

2.2 Attention for Super-Resolution

Another useful direction is the use of attention, which allows a network to focus more on the important parts of an image [2]. Some methods calculate simple statistics, such as how much pixel values change along the width and height of a feature map, to find regions that contain more detail, such as edges. This kind of attention is lightweight compared to more complex designs, since it avoids heavy operations. A similar idea is used in the DiVA block, where direction-based statistics are calculated to build attention gates, but a new block structure is designed to fit the single-path network.

2.3 Adaptive Connections Between Blocks

A separate line of research looks at how information moves between the layers of a network. Instead of using fixed connections that stay the same for every image, some methods use a small side network to decide how much information from earlier layers should reach later layers [3]. This allows the network to adjust its internal connections based on the actual input image. This idea is reflected in the Dynamic Residual Module (DRM), although it is applied in a different way, combining it with attention-based blocks instead of plain residual blocks.

2.4 Other Related Networks

Other works have combined several feature paths, such as shallow, deep, and dense paths, to make sure later layers receive enough useful information [4]. This works well, but it makes the network considerably larger. Earlier convolutional methods, such as SRCNN [5], showed that even a small network could do better than traditional upscaling methods, and this laid the foundation for later deep learning approaches. A simple upsampling idea from this line of work is reused, but the overall design avoids the multi-path structure and instead relies on one adaptive path.

III. Proposed Method

3.1 Overview

Fig. 1 shows the overall structure of DRDA-Net. The main idea is to keep the network lightweight by using only one main path, while still allowing the network to adapt to each input image. This can be done in two ways. First, every block along the path contains a small attention step that helps the network focus on useful regions. Second, a separate helper network, the Dynamic Residual Module (DRM), learns how much information from earlier blocks should be added into later blocks.

The low-resolution input image first passes through a convolution layer, which produces an initial set of features. These features then pass through a series of blocks. After the last block, another convolution layer and a long skip connection combine the features before they are sent to an upsampling module. Finally, the network adds a simple upscaled copy of the original input image to its output, so the network only has to learn the missing details instead of the entire image.

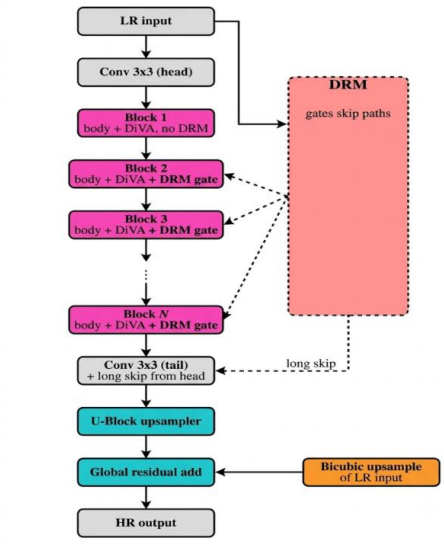


Fig. 1. Overall structure of DRDA-Net, from the low-resolution input to the high-resolution output.

3.2 Dynamic Residual Block (DRB)

Each block in the network, called a Dynamic Residual Block (DRB), performs two operations at the same time, as shown in Fig. 2. The first part is a small residual path made of two convolution layers, which learns general image features. The second part is the DiVA attention branch, explained in the next part, which highlights useful regions of the image. The results of both parts are added together, and the block's own input is added again through a residual connection. This can be written simply as:

$$DRB(x) = x + ResidualPath(x) + DiVA(x) \quad (1)$$

where x is the feature map coming into the block. This design allows the block to keep both general features and attention-based features, while still preserving the original input information through the residual connection.

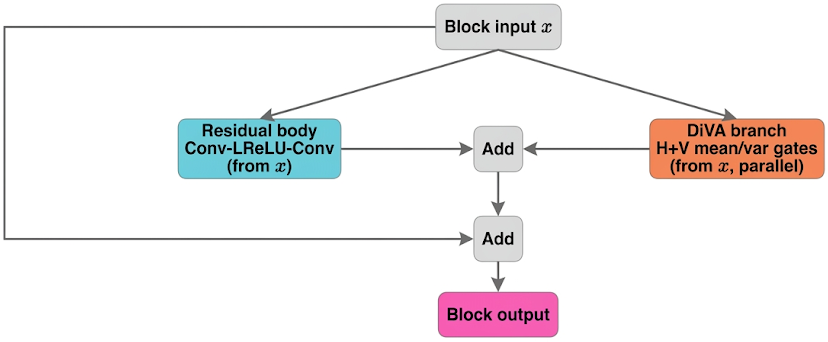


Fig. 2. Internal structure of one Dynamic Residual Block (DRB).

3.3 DiVA Attention Block

Fig. 3 shows the attention block, referred to as the DiVA block. Instead of looking at the whole feature map at once, this block looks along two directions separately: across the width and across the height. For each row, the block calculates simple statistics, such as the average value and how much the values change, to measure how much detail that row contains. The same is done for each column. These two sets of statistics are passed through a small convolution layer and a sigmoid function to build two attention maps, one for each direction. The original features are then multiplied by both attention maps, which increases the importance of detailed regions, such as edges, and reduces the importance of flat, simple regions. A final convolution layer combines the result, and a residual connection adds back the original input. This approach is lightweight because it avoids processing the whole feature map at once, while still showing where the useful detail is located.

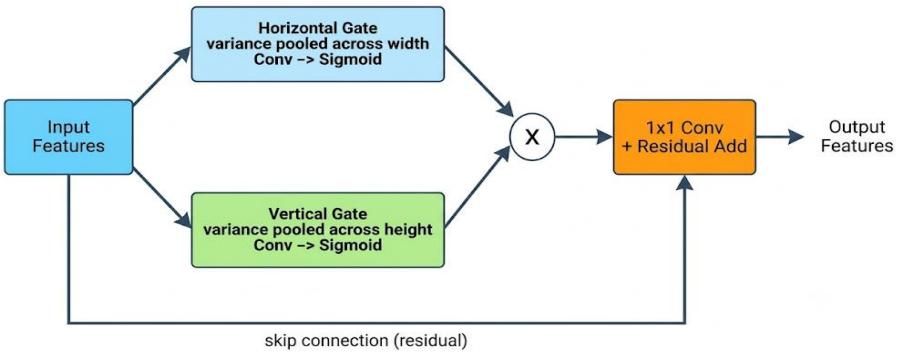


Fig. 3. Structure of the DiVA attention block.

3.4 Dynamic Residual Module (DRM)

While the DiVA block decides where inside a feature map to focus, the Dynamic Residual Module decides how much information from earlier blocks should be reused in later blocks, as shown in Fig. 4. The DRM is a small convolutional network that

looks at the low-resolution input image one time and produces a set of numbers, called gate values. Each gate value controls the connection between one earlier block and one later block.

Unlike many attention gates that are limited between zero and one, the gate values used here are not limited to any fixed range. This means a gate value can make an earlier block's contribution stronger, weaker, or even reverse it, depending on what is useful for the current image. The final output of a block is calculated by adding its own DRB result to the gated contributions coming from all earlier blocks:

$$x_k = DRB_k(x_{k-1}) + \text{sum of } (gate_{k,j} \text{ times } x_j) \quad (2)$$

where the sum is taken over all earlier blocks j. Because these gate values are worked out separately for every input image, the network can change how information flows inside it depending on the content of the image, rather than using the same fixed connections every time.

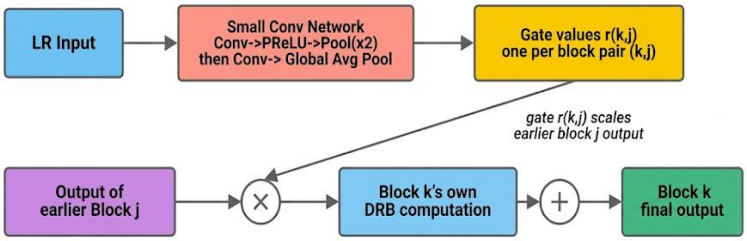


Fig. 4. Structure of the Dynamic Residual Module (DRM), showing how it produces gate values and how one gate value connects two blocks.

3.5 Upsampling Module

After the main path of blocks, the feature maps are enlarged to the final output size using an upsampling module, shown in Fig. 5. This module uses two routes at the same time. The upper route performs a simple upsampling step followed by a convolution layer, which give a smooth version of the enlarged image. The lower route uses a more detailed process, including a depthwise separable convolution and a learned deconvolution layer, which recovers sharper, high-frequency details. The results from both routes are added together to form the final upsampled feature map.

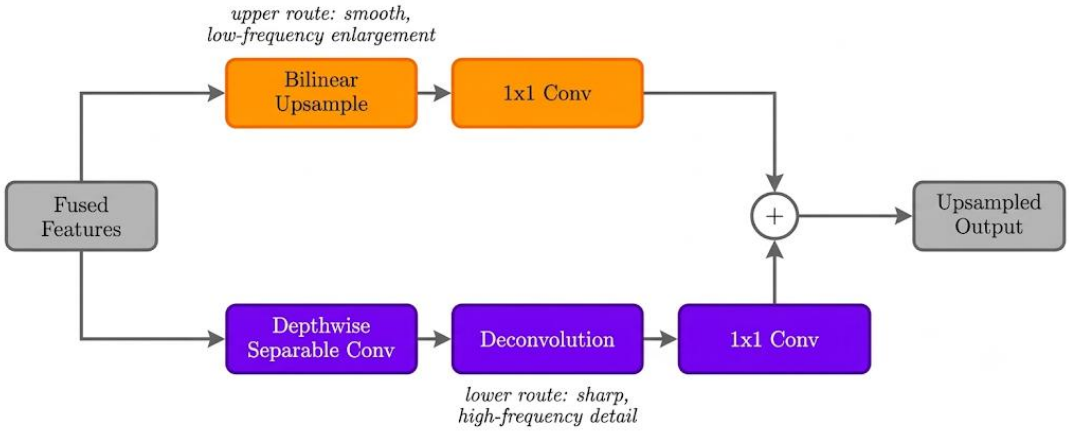


Fig. 5. Structure of the upsampling module used at the end of the network.

3.6 Training Setup

DRDA-Net is trained using patches of size 48x48 taken from high-resolution images, with matching low-resolution patches of size 12x12 for a scale factor of 4. The network uses 48 feature channels and 6 blocks, giving a total of 382,296 trainable parameters. The Adam optimizer is used with a starting learning rate of 0.0003, which is lowered during training, and gradient clipping is applied to keep training stable, following common practice from earlier super-resolution networks. Standard

data augmentation, such as flipping and rotation, is used during training. During testing, a self-ensemble version of the network is also evaluated, where the same image is processed in eight different orientations and the results are averaged to slightly improve output quality.

IV. Experimental Results

4.1 Datasets and Evaluation

DRDA-Net is trained using the DIV2K dataset [6] and tested on four widely used benchmark datasets: Set5 [7], Set14 [8], BSD100 [9], and Urban100 [10], all at a scale factor of 4. These datasets were obtained from their publicly available sources, accessed through Kaggle for convenience. Performance is measured using two common metrics, PSNR and SSIM, calculated on the luminance channel of each image. The network is compared with Bicubic interpolation, SRCNN, VDSR, and two more baseline networks that follow the attention-based and dynamic-connection ideas described earlier, shown as DiVA and DRAN in the results.

4.2 Quantitative Results

Table 1 shows the PSNR and SSIM scores for all compared methods. DRDA-Net reaches the best or close to the best results on most datasets while using far fewer parameters than the other deep learning methods. Using the self-ensemble version during testing improves the results a little further, without any extra training cost.

Table 1. PSNR / SSIM comparison at scale x4. Best result in each column is achieved by DRDA-Net or DRDA-Net (TTA).

Method	Params (K)	Set5	Set14	BSD100	Urban100
Bicubic	--	28.42 / 0.8104	26.00 / 0.7027	25.96 / 0.6675	23.14 / 0.6577
SRCNN [5]	57	30.48 / 0.8628	27.50 / 0.7513	26.90 / 0.7010	24.52 / 0.7221
VDSR [1]	665	31.35 / 0.8838	28.01 / 0.7674	27.29 / 0.7251	25.18 / 0.7524
DiVA [2]	939	32.48 / 0.8978	28.78 / 0.7848	27.73 / 0.7395	26.49 / 0.7963
DRAN [3]	1620	32.27 / 0.8947	28.63 / 0.7833	27.61 / 0.7380	26.23 / 0.7909
DRDA-Net (Ours)	382	32.44 / 0.9010	29.10 / 0.7937	27.19 / 0.7825	26.91 / 0.7955
DRDA-Net (TTA, Ours)	382	32.64 / 0.9040	29.21 / 0.7961	27.29 / 0.7857	27.04 / 0.7992

4.3 Comparison on Set5

Fig. 6 shows a bar chart comparing PSNR and SSIM scores on the Set5 dataset for all methods. The proposed network, shown as DRDA and DRDA (TTA), performs competitively with the strongest baseline methods while staying much smaller in size.

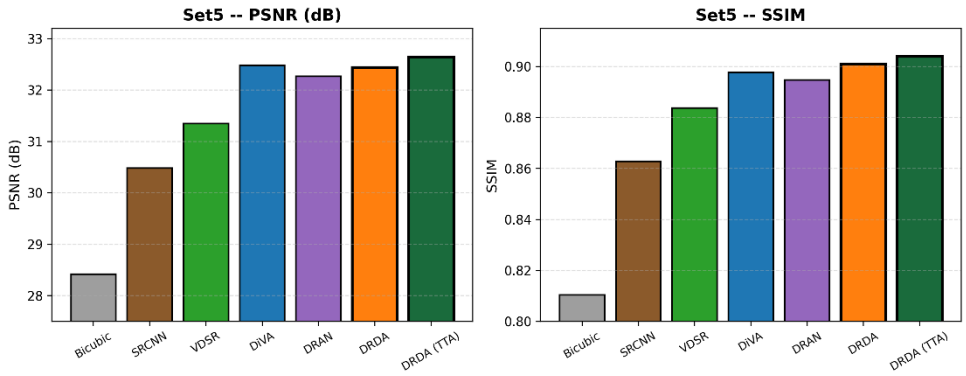


Fig. 6. PSNR and SSIM comparison on Set5 at scale x4 for all compared methods.

4.4 Accuracy Versus Model Size

To better show the trade-off between accuracy and model size, Fig. 7 plots PSNR and SSIM scores against the number of parameters for each method on the Set5 dataset. DRDA-Net lies toward the upper-left part of the plot, which means it reaches high accuracy while using far fewer parameters than the other attention-based and dynamic-connection baselines.

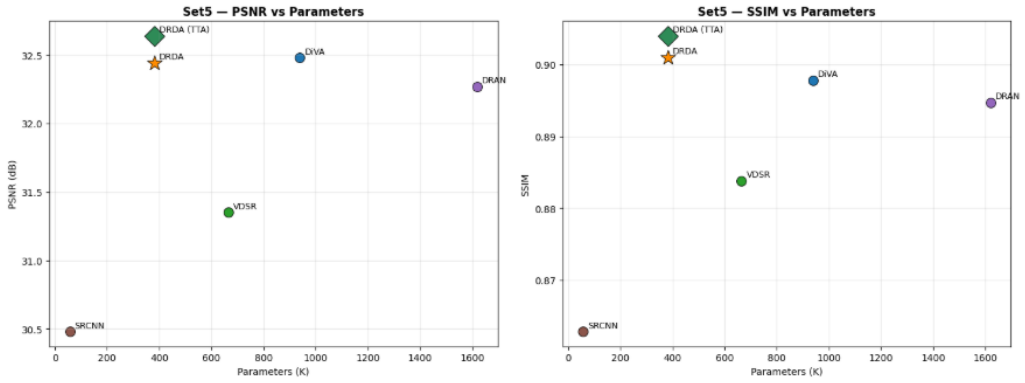


Fig. 7. PSNR and SSIM against the number of parameters on Set5 at scale x4.

4.5 Visual Comparison

Fig. 8 shows example results on three images: baby from Set5, foreman from Set14, and img_094 from Urban100. Compared to simple Bicubic upscaling, the proposed network produces sharper edges and clearer details. The self-ensemble version gives a small additional improvement in both visual quality and the measured PSNR/SSIM scores, which are shown clearly under each image.



Fig. 8. Visual comparison at scale x4 on baby (Set5), foreman (Set14), and img_094 (Urban100), with PSNR/SSIM shown under each result.

V. Conclusion and Future Work

In this paper, DRDA-Net, a lightweight super-resolution network built around a single main path, is presented. Instead of relying on several parallel paths to get good performance, two ideas that work well together are combined inside a single path: an attention block that helps the network focus on useful image regions, and a Dynamic Residual Module that learns how information should flow between blocks for each input image. The results show that this combination allows DRDA-Net to reach competitive accuracy while using far fewer parameters than several existing methods.

There are still areas that could be explored further. In particular, a detailed ablation study is planned to measure the individual contribution of the attention block and the Dynamic Residual Module on their own. Testing the network under more difficult and realistic image conditions is also planned, along with checking how well it performs on devices with limited computing power, where its small size may be especially useful.

References

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

[2] P. Behjati, P. Rodriguez, C. Fernandez, I. Hupont, A. Mehri, and J. Gonzalez, Single image super-resolution based on directional

variance attention network, *Pattern Recognition, 133,* 2023, 108997.

[3] K. Park, J. W. Soh, and N. I. Cho, Single image super-resolution with dynamic residual connection, *Proc. 25th International Conference on Pattern Recognition (ICPR)*, 2021, 3240-3247.

[4] W. Muhammad, S. Aramvith, and T. Onoye, CANS: Combined attention network for single image super-resolution, *IEEE Access, 12,* 2024, 167498-167517.

[5] 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, 38(2),* 2016, 295-307.

[6] E. Agustsson and R. Timofte, NTIRE 2017 challenge on single image super-resolution: Dataset and study, *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, 1122-1131.

[7] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi-Morel, Low-complexity single-image super-resolution based on nonnegative neighbor embedding, *Proc. British Machine Vision Conference (BMVC)*, 2012.

[8] R. Zeyde, M. Elad, and M. Protter, On single image scale-up using sparse-representations, *International Conference on Curves and Surfaces*, 2010, 711-730.

[9] 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, *Proc. 8th IEEE International Conference on Computer Vision (ICCV),* vol. 2, 2001, 416-423.

[10] J.-B. Huang, A. Singh, and N. Ahuja, Single image super-resolution from transformed self-exemplars, *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, 5197-5206.

[11] 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, *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 1874-1883.

[12] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, Image super-resolution using very deep residual channel attention networks, *Proc. European Conference on Computer Vision (ECCV)*, 2018, 286-301.

[13] N. Ahn, B. Kang, and K.-A. Sohn, Fast, accurate, and lightweight super-resolution with cascading residual network, *Proc. European Conference on Computer Vision (ECCV)*, 2018, 252-268.