NTIRE 2025 Efficient SR Challenge Factsheet
Frequency-Guided Multi-level Dispersion Network for Efficient Image Super-Resolution

1. Factsheet Information

1.1. Team details

- Team name:
  XUPTBoys
- Team leader name:
  Jing Wei
- Team leader address, phone number, and email
  Address:
  Xi'an University of Posts and Telecommunications, Xi'an, China
  Phone Number: +86 19530292952
  Email: freedomwj@126.com
- Rest of the team members
  Members:
  Jing Wei[1](freedomwj@126.com),Mengyang Wang[1]
  Ruilong Guo[1],Qian Wang[1,2]
- Affiliation
  [1]Xi'an University of Posts and Telecommunications, Xi'an, China
  [2]National Engineering Laboratory for Cyber Event Warning and Control Technologies

- Affiliation of the team and/or team members with NTIRE 2025 sponsors (check the workshop website)
  None
- User names and entries on the NTIRE 2025 Codalab competitions (development/validation and testing phases)
  User Name: jianai; **Entries:** 8
  User Name: longlong; **Entries:** 2
- Best scoring entries of the team during development/validation phase:
  Testing: 27.03dB
  Validation: 26.91dB
- Link to the codes/executables of the solution(s) https://github.com/jianai-skr/FMDN

1.2. Method details

The XUPTBoys team proposed the Frequency-Guided Multilevel Dispersion Network (FMDN), as shown in Fig. 1.FMDN adopts a similar basic framework to [2, 6, 7, 10].

Based on the above analysis, they propose the new Frequency-Guided Multi-level Dispersion Block(FMDB) and the new Frequency-Guided Multi-level Dispersion Block Basic(FMDB-B) as the base block of AGDN.As shown in Fig. 2 they use Hierarchical Variance-guided Spatial Attention(HVSA),Reallocated Contrast-Aware Channel Attention (RCCA) as alternatives to Enhanced Spatial Attention (ESA) [8] and Contrast-Aware Channel Attention (CCA) [3], Frequency-Guided Residual block (FRB), Asymmetric FeedForward Network (AFFN), Multilevel Residual Convolution (MRConv) and Multilevel Residual Convolution Basic(MRConv-B).The difference between FMDB and FMDB-B is that the former uses MRConv, while the latter uses MRConv-B.

In HVSA,the effects of multilevel branching and local variance on performance are examined. Small-window multilevel branches fail to capture sufficient information, while local variance within a single branch can create significant weight disparities. To address these issues, [10] was enhanced to introduce the D5 and D7 branches, which effectively utilize local variance to capture information-rich regions while balancing performance and complexity. In RCCA,this approach improves the traditional channel attention mechanism by not only reallocating weights across channels but also better managing shared information among them. Introduces complementary branches with 1×1 convolutions and GELU activation functions, which help redistribute complementary information, improving the uniqueness of each channel.In FRB,it enhances feature representation using convolutional layers and GELU activation. It normalizes input, extracts features with depthwise convolutions of different kernel sizes, and combines them through residual connections to preserve spatial information for effective image processing.In AFFN,it applies layer normalization and a 1x1 convolution to expand feature

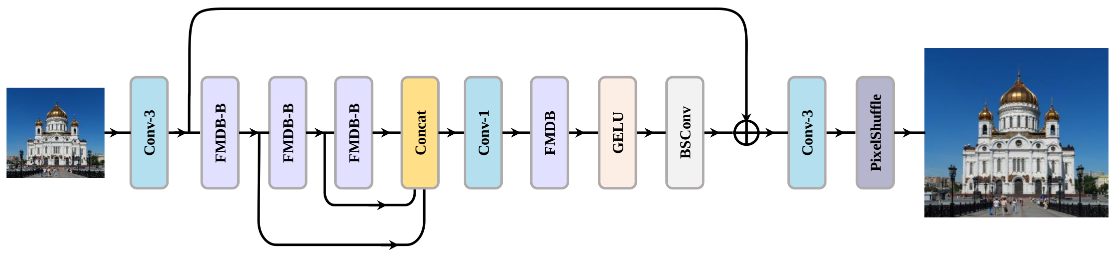

Figure 1. The whole framework of Frequency-Guided Multi-level Dispersion Network (FMDN)

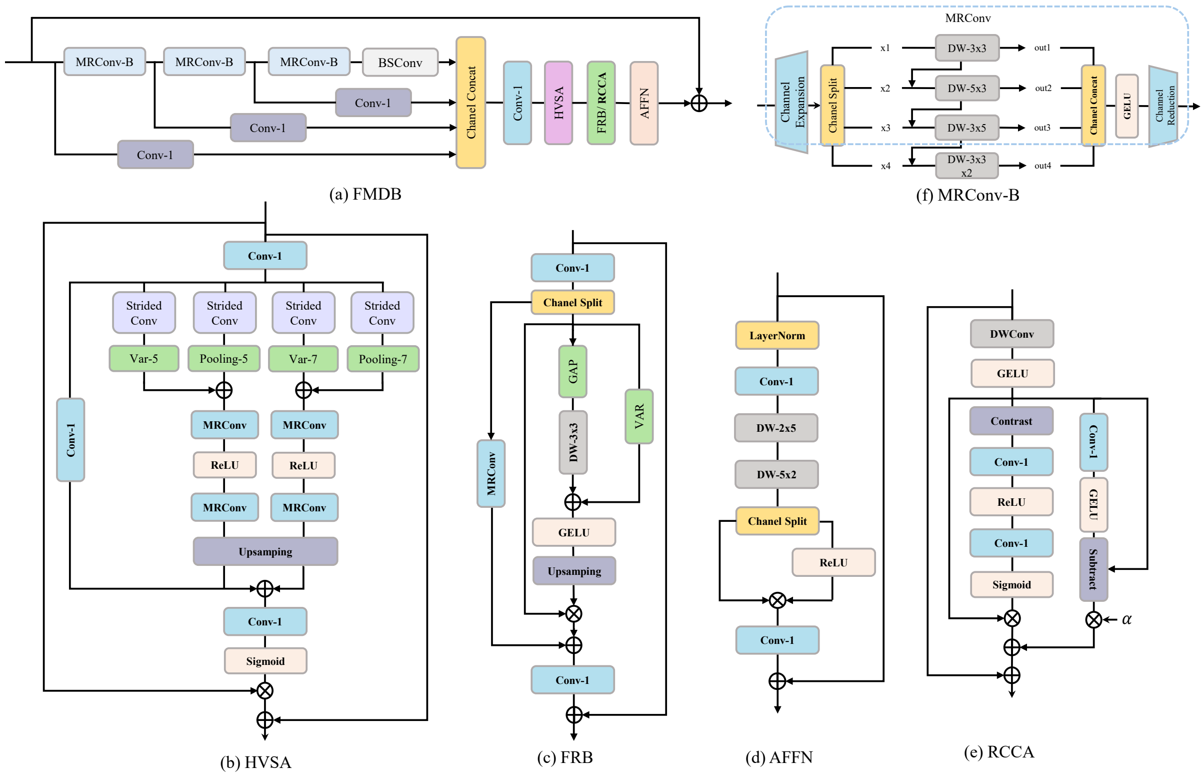

Figure 2. The details of each component. (a) FMDB: Frequency-Guided Multi-level Dispersion Block; (b) HVSA: Hierarchical Variance-guided Spatial Attention; (c) FRB: Frequency-Guided Residual block; (d) AFFN: Asymmetric FeedForward Network; (e)RCCA: Re-allocated Contrast-aware Channel Attention; (f) MRConv-B/MRConv: Multilevel Residual Convolution Basic and Multilevel Residual Convolution

dimensions. It then uses two depthwise convolutions with different kernel sizes, combines the results with GELU activation, and projects the output back to the original dimension with a residual connection.In MRConv and MRConv-B,MRConv and MRConv-B use convolution kernels of different sizes for parallel convolution, and finally activate the output using GELU and combine it with residual connections, effectively preserving spatial information

1.3. Training strategy

The proposed FMDN has 3 FMDB-Basic blocks and 1 FMDB block, in which the number of feature channels is set to 24. The details of training steps are as follows:

1. Pretraining on the DIV2K [1] and and Flickr2K [11]. HR patches of size 256 × 256 are randomly cropped from HR images, and the mini-batch size is set to 64. The model is trained by minimizing the L1 loss function [9] with the Adam optimizer [4]. The initial learning rate is set to 2×10^{-3} and halved at $\{100k, 500k, 800k, 900k, 950k\}$-iteration. The total number of iterations is 1000k.

2. Finetuning on 800 images of DIV2K and the first 10k images of LSDIR [5]. HR patch size and mini-batch

Table 1. The performance of FMDN.

Method	#Params [M]	#FLOPs [G]	Val PSNR [dB]	Test PSNR [dB]
Baseline [12]	0.276	16.70	26.93	27.01
FMDN	**0.071**	**3.39**	**26.91**	**27.03**

size are set to 384 × 384 and 64, respectively. The model is fine-tuned by minimizing L2 loss function [9]. The initial learning rate is set to 5×10^{-4} and halved at $\{500k\}$-iteration. The total number of iterations is 1000k.

1.4. Experimental results

In this section, we show the performance of FMDN in some aspects. As shown in Tab. 1, FMDN obtained a performance of 26.91 dB in the validation set and 27.03 dB in the test set with few parameters.

References

[1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 126–135, 2017. 2

[2] Zheng Hui, Xinbo Gao, Yunchu Yang, and Xiumei Wang. Lightweight image super-resolution with information multi-distillation network. In *Proceedings of the 27th acm international conference on multimedia*, pages 2024–2032, 2019. 1

[3] Zheng Hui, Xinbo Gao, Yunchu Yang, and Xiumei Wang. Lightweight image super-resolution with information multi-distillation network. In *Proceedings of the 27th acm international conference on multimedia*, pages 2024–2032, 2019. 1

[4] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 2

[5] Yawei Li, Kai Zhang, Jingyun Liang, Jiezhang Cao, Ce Liu, Rui Gong, Yulun Zhang, Hao Tang, Yun Liu, Denis Demandolx, et al. Lsdir: A large scale dataset for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1775–1787, 2023. 2

[6] Zheyuan Li, Yingqi Liu, Xiangyu Chen, Haoming Cai, Jinjin Gu, Yu Qiao, and Chao Dong. Blueprint separable residual network for efficient image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 833–843, 2022. 1

[7] Jie Liu, Jie Tang, and Gangshan Wu. Residual feature distillation network for lightweight image super-resolution. In *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 41–55. Springer, 2020. 1

[8] Jie Liu, Wenjie Zhang, Yuting Tang, Jie Tang, and Gangshan Wu. Residual feature aggregation network for image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2359–2368, 2020. 1

[9] Zhaoyang Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Proceedings of the ieee/cvf international conference on computer vision. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10012–10022, 2021. 2, 3

[10] Yanyu Mao, Nihao Zhang, Qian Wang, Bendu Bai, et al. Multi-level dispersion residual network for efficient image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2023. 1

[11] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, and Lei Zhang. Ntire 2017 challenge on single image super-resolution: Methods and results. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 114–125, 2017. 2

[12] Yan Wang. Edge-enhanced feature distillation network for efficient super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 777–785, 2022. 3