NTIRE 2025 Efficient SR Challenge Factsheet
-EECNet: Edge Enhanced Convolutional Network for Efficient Super-Resolution-

Rui Chen
Shenzhen International Graduate School, Tsinghua University
ShenZhen, China
chenr24@mails.tsinghua.edu.cn

1. Team details

- Team name:
  ChanSR
- Team leader name:
  Rui Chen
- Team leader address and email:
  Shenzhen International Graduate School, Tsinghua University, China;
  chenr24@mails.tsinghua.edu.cn
- Affiliation:
  Shenzhen International Graduate School, Tsinghua University, China
- Affiliation of the team and/or team members with NTIRE 2025 sponsors (check the workshop website):
  No affiliation
- User names and entries on the NTIRE 2025 Codalab competitions (development/validation and testing phases):
  a_bai
- Best scoring entries of the team during the development/validation phase:
  26.92(valid) / 27.03(test)
- Link to the codes/executables of the solution(s):
  https://github.com/chenr2020/EECNet

2. Method details

2.1. General method description

We propose the Edge Enhanced Convolutional Network (EECNet) for the efficient super-resolution task. The network architecture is inspired by the design of SRN [9], while fully exploring the capacity of reparameterizable convolution. The whole architecture is shown in Fig.1, which mainly consists of eight Edge Enhanced Blocks (EBBs) and a pixel shuffle module. Reparameterizable convolutions are utilized in the EEB, aiming to improve the super-resolution capability without introducing any additional parameter overhead during the inference stage. Meanwhile, the network is optimized by the pixel-wise loss such as L1 loss or L2 loss, along with the FFT loss [1].

2.2. Edge Enhanced Block

Inspired by EFDN [8] and FMEN [2], we propose an Edge Enhancement Convolution (EEC) to maximize the representation capacity of single convolutional layers. Recognizing the critical role of edge information in super-resolution and the difficulty for models to autonomously learn sharp edge filters, we introduce a predefined High-Pass Filter (HPF) branch to explicitly capture edge details. Following ECBSR [10], we incorporate a learnable scaling factor for the filter parameters.

High-Pass Filter Branch: Our method extracts high-frequency features through a simple yet effective process: First, we apply a Gaussian blur kernel with $\sigma = 1$ to the input feature maps to obtain smoothed versions. The Gaussian blur kernel is formulated as:

$$\mathbf{K}_{blur} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{1}$$

The high-Pass Filter are then derived by subtracting these blurred features from the original input through a skip connection, as Fig2. Since the skip connection operation can be mathematically equivalent to a $3\times\ 3$ Convolution kernel, we merge these two branches to construct our pre-

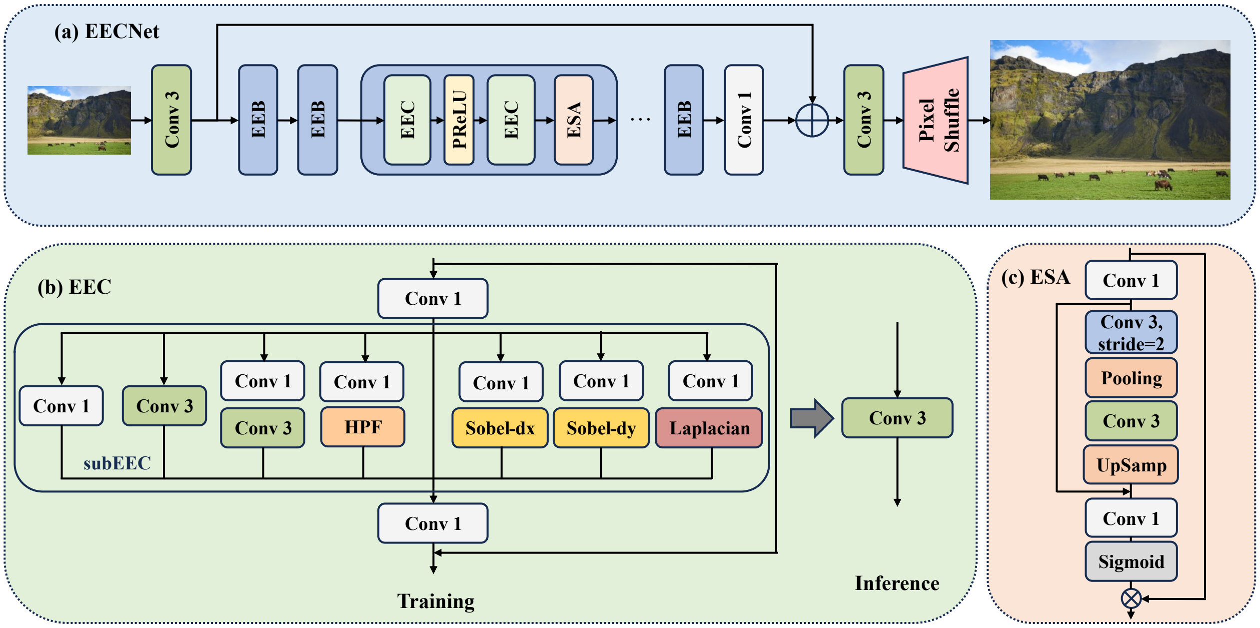

Figure 1. Network architecture of our EECNet.

defined High-Pass Filter, formulated as:

$$\mathbf{K}_{hpf} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \mathbf{K}_{blur} = \frac{1}{16}\begin{bmatrix} -1 & -2 & -1 \\ -2 & 12 & -2 \\ -1 & -2 & -1 \end{bmatrix} \tag{2}$$

Specifically, the input feature X is first processed by a $C \times C \times 1 \times 1$ convolution, then the high-frequency features are extracted using our high-pass filter. Each channel of the intermediate feature is first processed by the high-pass filter, then scaled by a channel-wise scaling factor like ECBSR [10]. The process is formulated as:

$$\mathbf{F}_{hpf} = (\mathbf{S}_{hpf} \cdot \mathbf{K}_{hpf}) \otimes (\mathbf{K}_x * \mathrm{X} + \mathbf{B}_x) + \mathbf{B}_{hpf} \tag{3}$$

where $\mathbf{K}_x$ and $\mathbf{B}_x$ are the weight and bias of $1{\times}1$ convolution. $\mathbf{S}_{hpf}$ and $\mathbf{B}_{hpf}$ are the scaling parameter and bias

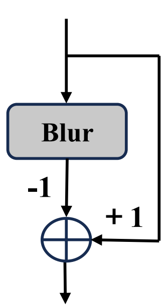

Figure 2. High-Pass Filter

with shape $C \times 1 \times 1 \times 1$. $\otimes$ and $*$ represent depth-wise convolution (DWConv) and normal convolution. $\cdot$ indicates channel-wise broadcasting multiplication. $(\mathbf{S}_{hpf} \cdot \mathbf{D}_x)$ have shape $C \times 1 \times 3 \times 3$.

Edge Enhanced Convolution: We integrate the proposed high-pass filter branch into the EDBB module of EFDN [8], creating the subEEC module. As subEEC can be mathematically equivalent to a standard 3×3 convolution, we replace the original 3×3 convolution in RRRB module of FMEN [2] with our subEEC to obtain the final EEC architecture. Notably, to ensure valid re-parameterization, we initialize the bias term of the first convolution layer as zero to compensate for the zero-padding operation in subEEC.

ESA Enhancement: To better capture global spatial information, we adopt the simplified Efficient Spatial Attention mechanism from SRN [9], whose structure is shown in Fig1(c). Compared with the original ESA, this implementation removes the 1×1 convolution layer and reduces computational complexity by employing only a single 3×3 convolution in the convolutional group.

2.3. Training strategy

The proposed EECNet contains eight EEBs, in which we set the number of feature maps to 32. Also, the channel number of the ESA is set to 16 similar to [4]. Throughout the entire training process, we use the Adam optimizer [3], where $\beta1$ = 0.9 and $\beta2$ = 0.999. The model is trained for 1000k iterations in each stage. Input patches are randomly cropped and augmented. Data augmentation strategies included horizontal and vertical flips, and random rotations of 90, 180, and 270 degrees. Model training was performed using Pytorch 1.12.0 [6] on RTX 3090. Specifically, the training strategy consists of several steps as follows.

1. In the starting stage, we train the model from scratch on the 800 images of DIV2K [7] and the first 10k images of LSDIR [5] datasets. The model is trained for total 10^6 iterations by minimizing L1 loss and FFT loss [1]. The HR patch size is set to 256×256, while the mini-batch size is set to 64. We set the initial learning rate to 1×10^{-3} and the minimum one to 1×10^{-5}, which is updated by the Cosine Annealing scheme.

2. In the second stage, we increase the HR patch size to 384, while the mini-batch size is set to 32. The model is fine-tuned by minimizing the L1 loss and FFT loss. We set the initial learning rate to 5×10^{-4} and the minimum one to 1×10^{-6}, which is updated by the Cosine Annealing scheme.

3. In the last stage, the model is fine-tuned with 480×480 HR patches, however, the loss function is changed to minimize the combination of L2 loss and FFT loss. Other settings are the same as Stage 2.

2.4. Experimental results

	PSNR(valid)	PSNR(test)	Params(M)	FLOPs(G)
EFDN	26.93	27.01	0.276	16.70
EECNet	26.92	27.03	0.209	11.59

Table 1. Comparison of EFDN and EECNet

We compare the differences between the EECNet and the EFDN on the DIV2K_LSDIR_valid&test datasets. As can be seen from the table, our model EECNet has a great advantage in the number of Params and FLOPs. Durning the testing phase, our proposed model scored 27.03.

3. Other details

- Planned submission of a solution(s) description paper at NTIRE 2025 workshop.

  We may submit a solution description paper at NTIRE 2025 workshop.

- General comments and impressions of the NTIRE 2025 challenge.

  The NTIRE 2025 challenge provides a great platform for advancing image restoration research. It encourages innovation and offers a benchmark for new techniques in the field.

- What do you expect from a new challenge in image restoration, enhancement, and manipulation?

  Perhaps we can consider introducing a competition focused on mobile device super-resolution. This would allow us to explore how well models perform in the context of resource-limited devices.

- Other comments: encountered difficulties, fairness of the challenge, proposed subcategories, proposed evaluation method(s), etc.

  So far everything seems fine, I don't have any suggestions.

References

[1] Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4641–4650, 2021. 1, 3

[2] Zongcai Du, Ding Liu, Jie Liu, Jie Tang, Gangshan Wu, and Lean Fu. Fast and memory-efficient network towards efficient image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 853–862, 2022. 1, 2

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

[4] Fangyuan Kong, Mingxi Li, Songwei Liu, Ding Liu, Jingwen He, Yang Bai, Fangmin Chen, and Lean Fu. Residual local feature network for efficient super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 766–776, 2022. 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 Workshops*, 2023. 3

[6] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 3

[7] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, Lei Zhang, et al. NTIRE 2017 challenge on single image super-resolution: Methods and results. In *CVPR Workshops*, 2017. 3

[8] 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. 1, 2

[9] Yucong Wang and Minjie Cai. A single residual network with esa modules and distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1971–1981, 2023. 1, 2

[10] Xindong Zhang, Hui Zeng, and Lei Zhang. Edge-oriented convolution block for real-time super resolution on mobile devices. In *Proceedings of the 29th ACM international conference on multimedia*, pages 4034–4043, 2021. 1, 2