NTIRE 2024 Efficient SR Challenge Factsheet
-Efficient Enhanced Residual Network-

Jingyi Zhang
Hefei University of Technology
Hefei, China
jingyizhang0806@163.com

Baiang Li
Hefei University of Technology
Hefei, China
ztmotalee@gmail.com

Qi Zhu
University of Science and Technology of China
Hefei, China
zqcrafts@mail.ustc.edu.cn

Xiaogang Xu
Zhejiang Lab, Zhejiang University
Hangzhou, China
xiaogangxu00@gmail.com

Dan Guo
Hefei University of Technology
Hefei, China
guodan@hfut.edu.cn

Chunle Guo*
Nankai University
Tianjin, China
guochunle@nankai.edu.cn

1. Introduction

This factsheet template is meant to structure the description of the contributions made by each participating team in the NTIRE 2024 challenge on efficient image super-resolution.

Ideally, all the aspects enumerated below should be addressed. The provided information, the codes/executables and the achieved performance on the testing data are used to decide the awardees of the NTIRE 2024 challenge.

Reproducibility is a must and needs to be checked for the final test results in order to qualify for the NTIRE awards.

The main winners will be decided based on overall performance and a number of awards will go to novel, interesting solutions and to solutions that stand up as the best in a particular subcategory the judging committee will decided. Please check the competition webpage and forums for more details.

The winners, the awardees and the top ranking teams will be invited to co-author the NTIRE 2024 challenge report and to submit papers with their solutions to the NTIRE 2024 workshop. Detailed descriptions are much appreciated.

The factsheet, source codes/executables, trained models should be sent to **all of the NTIRE 2024 challenge organizers (Yawei Li, Bin Ren, Nancy Mehta, and Radu Timofte)** by email.

2. Email final submission guide

To: yawei.li@vision.ee.ethz.ch
bin.ren@unitn.com
nancy.mehta@uni-wuerzburg.de
timofte.radu@gmail.com
cc: your_team_members

Title: NTIRE 2024 Efficient SR Challenge - TEAM_NAME - TEAM_ID

To get your TEAM_ID, please register at Google Sheet. Please fill in your Team Name, Contact Person, and Contact Email in the first empty row from the top of sheet. Body contents should include:

a) team name

b) team leader's name and email address

c) rest of the team members

d) user names on NTIRE 2024 CodaLab competitions

e) Code, pretrained model, and factsheet download command, e.g. `git clone ...`, `wget ...`

f) Result download command, e.g. `wget ...`

- Please provide different urls in e) and f)

Factsheet must be a compiled pdf file together with a zip with .tex factsheet source files. Please provide a detailed explanation.

3. Code Submission

The code and trained models should be organized according to the GitHub repository. This code repository provides the basis to compare the various methods in the challenge. **Code scripts based on other repositories will not be accepted.** Specifically, you should follow the steps below.

1. Git clone the repository.
2. Put your model script under the `models` folder. Name your model script as `[Your_Team_ID]_[Your_Model_Name].py`.
3. Put your pretrained model under the `model_zoo` folder. Name your model checkpoint as `[Your_Team_ID]_[Your_Model_Name].[pth or pt or ckpt]`
4. Modify `model_path` in `test_demo.py`. Modify the imported models.
5. `python test_demo.py`

Please send us the command to download your code, e.g. `git clone [Your repository link]` When submitting the code, please remove the LR and SR images in `data` folder to save the bandwidth.

4. Factsheet Information

The factsheet should contain the following information. Most importantly, you should describe your method in detail. The training strategy (optimization method, learning rate schedule, and other parameters such as batch size, and patch size) and training data (information about the additional trainning data) should also be explained in detail.

4.1. Team details

- Team name

  Lasagna

- Team leader name

  Jingyi Zhang

- Team leader address, phone number, and email

  9th Floor, Block A, Science And Education Building, Hefei University of Technology (Emerald Lake Campus), 485 Danxia Road, Hefei 230601, China,

  (+86)13738264339,

  jingyizhang0806@163.com

- Rest of the team members

  Baiang Li, Hefei University of Technology

  Qi Zhu, University of Science and Technology of China

  Xiaogang Xu, Zhejiang Lab, Zhejiang University

  Dan Guo, Hefei University of Technology

  Chunle Guo*, Nankai University

- Team website URL (if any)

- Affiliation

  Hefei University of Technology, University of Science and Technology of China, Zhejiang Lab, Zhejiang University, Nankai University

- Affiliation of the team and/or team members with NTIRE 2024 sponsors (check the workshop website)

- User names and entries on the NTIRE 2024 Codalab competitions (development/validation and testing phases)

  jingyiii, 2(test)

- Best scoring entries of the team during development/validation phase

  best scoring entry:2, score:27.03

- Link to the codes/executables of the solution(s) Team25

4.2. Method details

- General method description (How is the network designed.)

  Although the EDSR [1] network achieves high performance in the field of single-image super-resolution, its network structure is relatively bulky. Therefore, we have adopted a strategy of reducing the number of blocks to lighten the model. However, merely decreasing the number of blocks are not able to meet the performance standards set by the competition. Therefore, we have made certain adjustments to its structure by incorporating an ESA [2] after each block, and to reduce the parameters, we have eliminated several convolutional layers. Additionally, we appended a convolutional layer at the end of the ERB block for the purpose of fine-tuning. This convolutional layer was not included at the onset of the training process but was incorporated during the final fine-tuning phase to participate in the computation. Simultaneously, in order to increase inference speed and reduce GPU memory usage, we have removed the residual connections in the

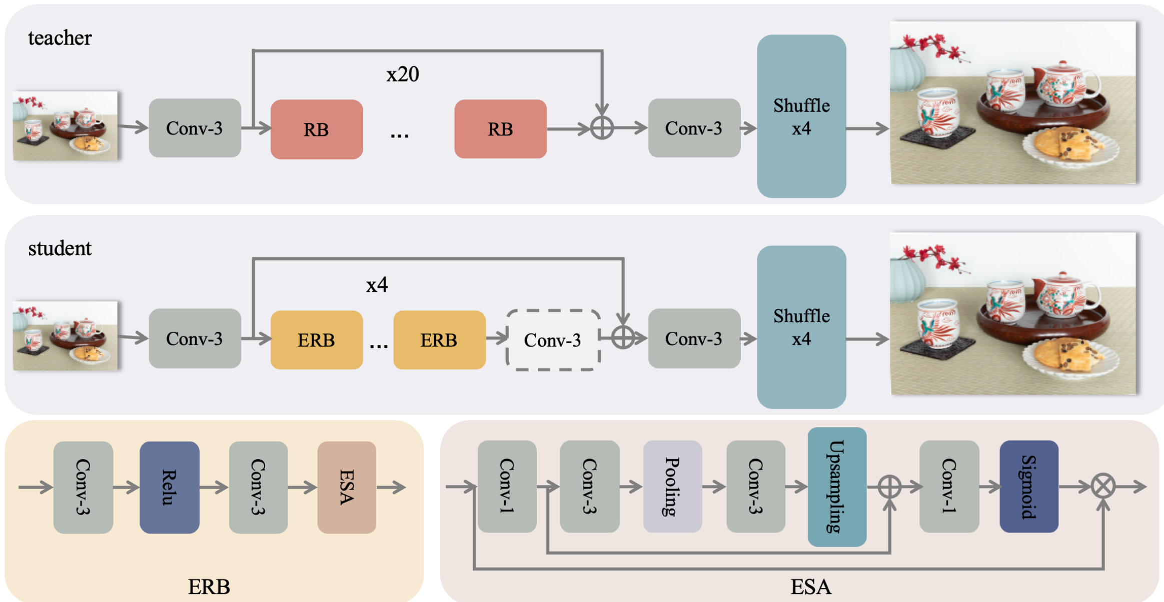

Figure 1. The structure of the proposed EERN.

RB modules of EDSR. Experiments show that removing residual connections has minimal impact on the module's performance. Moreover, considering that excessive changes in the number of feature channels are detrimental to performance in lightweight networks, we have opted to replace two successive shuffle2 operations with a single shuffle4 operation to avoid the final convolutional reduction of channels from 256 to 3. Lastly, we trained the network using a knowledge distillation approach. Due to the mismatch in the number of feature channels between the teacher model and the student model, we did not utilize the loss of feature maps. Instead, the loss was calculated based on the outputs of both models.

- Representative image / diagram / pipeline of the method(s)

  see Fig 1.

- Training strategy

  The number of ERB modules and the number of it's feature channels were set to 4 and 84, respectively. We trained a total of 1700 epochs to bring the model to convergence. The process was divided into three stages. In the first stage, we initiated the learning rate at 1e-4, employing a cosine annealing strategy to decrease the learning rate to 5e-7 by the 400th epoch. During this cycle, we utilized L1 loss, with the dataset limited to DIV2K. In the second stage, the starting learning rate was set to 1e-5, again using a cosine annealing method to reduce the learning rate to 5e-7 by the 300th epoch. Moreover, in this cycle, PSNR loss was employed for fine-tuning, and the dataset was expanded to include both DIV2K and folders named 0001000 to 0010000 from the LSDIR dataset. In the final stage, the convolutional layer was introduced and initialized to zero. This stage comprised a total of 1000 epochs, with all other settings being identical to those of the second stage.

- Experimental results

  valid PSNR: 26.902440 db

  test PSNR: 26.993437 db

- References

Additionally, you can refer to the following items to detail your description.

- Total method complexity (number of parameters, FLOPs, GPU memory consumption, runtime)

  tested on NVIDIA GeForce RTX 3090 GPU*1

  parameters(M): 0.6574

  FLOPs(G): 41.2126

  GPU memory consumption(M): 549.762

  runtime(milliseconds): 17.728811

- Which pre-trained or external methods/models have been used (for any stage, if any)
  None.

- Which additional data has been used in addition to the provided NTIRE training and validation data (at any stage, if any)
  None.

- Training description
  Our training strategy has been discussed in previous section.

- Testing description
  The result can be directly obtained using our pretrained model.

- Quantitative and qualitative advantages of the proposed solution
  Our solution is efficient and quickly enough while maintaining performance.

- Results of the comparison to other approaches (if any)

- Results on other benchmarks (if any)

- Novelty degree of the solution and if it has been previously published.
  Our method propose a distillation strategy which can effectively improve the student model's performance through teacher loss.

- It is OK if the proposed solution is based on other works (papers, reports, Internet sources (links), etc). It is ethically wrong and a misconduct if you are not properly giving credits and hide this information.

  Our code drew inspiration from SRN [3]. After finetuning and modifying the model structure, we employed new parameter settings and training strategies.

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2024 workshop.
  We will include more experimental results and propose a more efficient solution for the workshop paper.

- General comments and impressions of the NTIRE 2024 challenge.
  Well done.

- What do you expect from a new challenge in image restoration, enhancement and manipulation?
  I am expecting of single HDR image reconstruction task.

- Other comments: encountered difficulties, fairness of the challenge, proposed subcategories, proposed evaluation method(s), etc.
  None.

References

[1] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017.

[2] 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.

[3] 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 1970–1980, 2023.