NTIRE 2023 Efficient SR Challenge Factsheet
Large Kernel Distillation Network

Chengxing Xie[1], Xiaoming Zhang[1], Linze Li[1], Haiteng Meng[1], Tianlin Zhang[2], Tianrui Li[1], Xiaole Zhao[1]
[1]Southwest Jiaotong University, China
[2]National Space Science Center, Chinese Academy of Science, China
zxc0074869@gmail.com, zxmswjtu@163.com

1. Introduction

This factsheet template is meant to structure the description of the contributions made by each participating team in the NTIRE 2023 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 2023 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 2023 challenge report and to submit papers with their solutions to the NTIRE 2023 workshop. Detailed descriptions are much appreciated.

The factsheet, source codes/executables, trained models should be sent to **all of the NTIRE 2023 challenge organizers (Yawei Li, Yulun Zhang, and Radu Timofte)** by email.

2. Email final submission guide

To: yawei.li@vision.ee.ethz.ch
yulun100@gmail.com
timofte.radu@gmail.com
cc: your_team_members
Title: NTIRE 2023 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 2023 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: Set5 Baby
- Team leader name: Xiaoming Zhang
- Team leader address, phone number, and email: Southwest Jiaotong University, China; +86 18782162105; zxmswjtu@163.com
- Rest of the team members: Chengxing Xie, Linze Li, Haiteng Meng, Tianlin Zhang,Tianrui Li, Xiaole Zhao
- Team website URL (if any)
- Affiliation
- Affiliation of the team and/or team members with NTIRE 2023 sponsors (check the workshop website)
- User names and entries on the NTIRE 2023 Codalab competitions (development/validation and testing phases): xiaomingdd233, stella_von
- Best scoring entries of the team during development/validation phase: 29.01 dB on DIV2K Validation
- Link to the codes/executables of the solution(s)

4.2. Method details

You should describe your proposed solution in detail. This part is equivalent to the methodology part of a conference paper submission. The description should cover the following details.

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

  As shown in Figure 1, we propose large kernel distillation network (LKDN), with a network topology similar to BSRN [3], which mainly includes four parts: shallow feature extraction, multiple stacked feature distillation blocks, multi-layer feature fusion, and image reconstruction block.

  In the first stage, the input image is replicated N times and passed through a blueprint convolution (BSConv) as shown in Figure 2b to extract shallow features.

  The deep features are then progressively extracted through stacked large kernel distillation blocks (LKDB), as shown in Figure 2a. To boost the representation capacity of the backbone and reduce the time consumption taken by skip-connections, as Figure 2c demonstrates, re-parameterized blueprint convolution and re-parameterizable skip connections are introduced to construct the RBSB (Re-parameterized Blueprint Shallow Block). Additionally, Large Kernel Attention (LKA) is introduced to increase the network's receptive field and improve its performance as shown in Figure 2d. Then, the information interaction across channels is achieved through a 1×1 convolution. Finally, a Pixel-Norm [4] layer is added to stabilize the training of the LKDB.

  A 1×1 convolution is used to fuse the multi-layer features extracted by the previous LKDB in the channel dimension.

  Finally, the SR image is generated by the reconstruction module, which includes re-parameterized 3×3 convolution and non-parametric sub-pixel operation, and its detail is in Figure 2e.

  The LDKN consists of 5 LKDBs, with 42 intermediate channels.

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

- Training strategy

  The training process consists of two stages: the initial training stage and fine-tuning stage.

  In the initial training stage: we first randomly crop HR patches with a size of 256×256 from the HR image and set the mini-batch size to 128. We augment the training data with random horizontal flips and 90 rotations. The original LKDN is trained by minimizing the common-used L1 loss function through the Adam optimizer [2] with $\beta 1 = 0.9$, $\beta 2 = 0.99$ and $\epsilon = 10^{-8}$, the learning rate is set to a constant 1×10^{-3}, and a total of 9.5×10^5 iterations are trained.

  In the fine-tuning stage: The patch size of HR images and the batch size are set to 480×480 and 64 respectively. Then LKDN is fine-tuned by minimizing the L2 loss function, and the learning rate is set to 2×10^5, and a total of 5×10^4 iterations are trained.

- Experimental results Table 1 demonstratex the qualitative results on benchmark datasets of LKDN on SR($\times 4$).

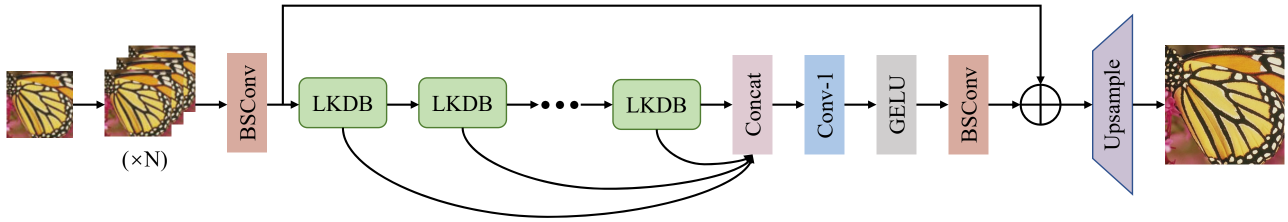

Figure 1. The architecture of large kernel distillation network (LKDN).

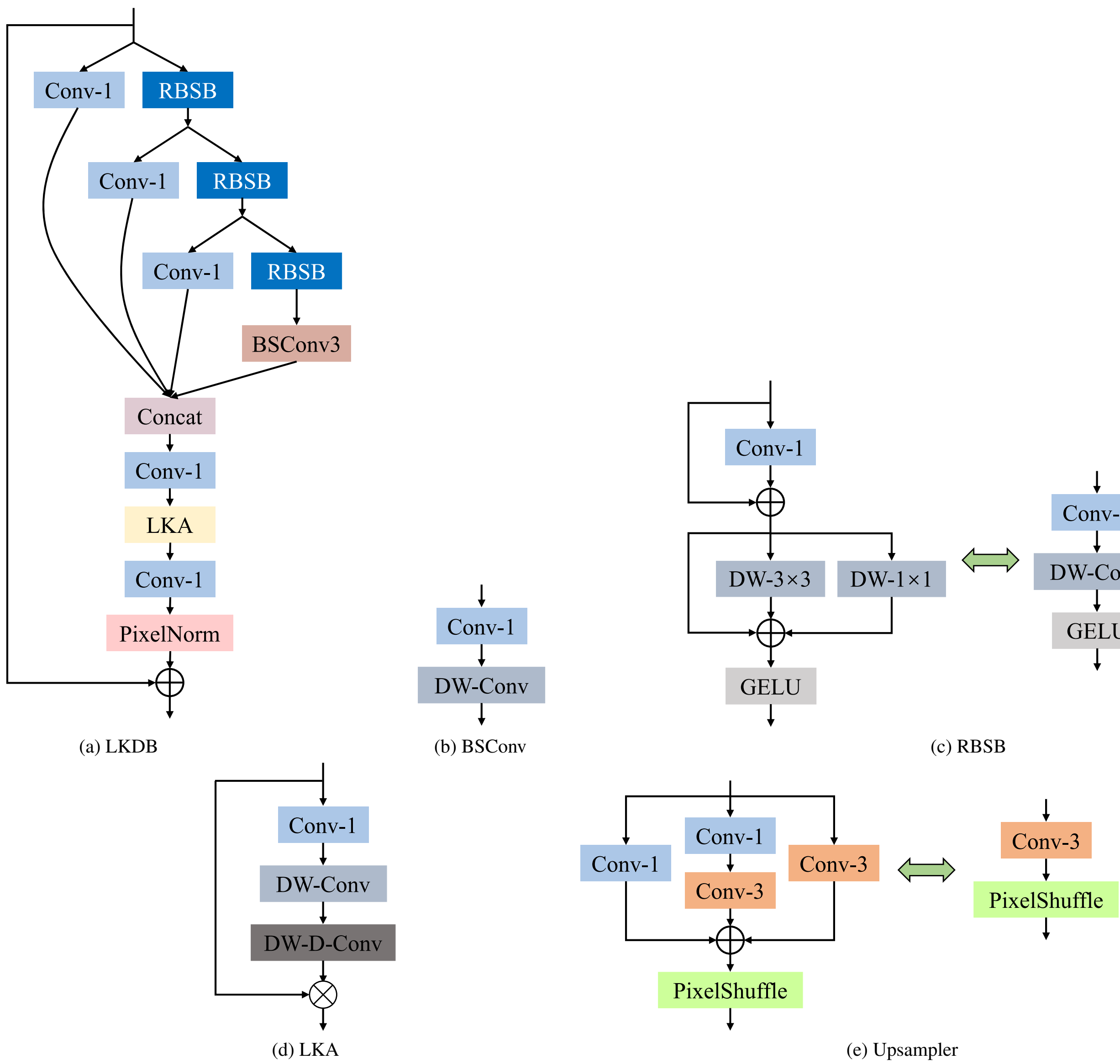

Figure 2. (a) LKDN: Large Kernel Distillation Block; (b) BSConv: Bluprint Convolution; (c) RBSB: Re-parameterized Bluprint Shallow Block; (d) LKA: Large Kernel Attention; (e) Re-parameterized Upsampler.

- References

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

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

Table 1. Quantitative results on benchmark datasets, PSNR(dB) / SSIM.

Method	Params	Multi-Adds	Set5	Set14	B100	Urban100	Manga109
LKDN	129.177K	8.29G	32.10 / 0.8938	28.62 / 0.7820	27.59/ 0.7371	26.07 / 0.7845	30.50 / 0.9077

Table 2. Other metrics of LKDN.

Model	val PSNR	val Time	Params	FLOPs	Acts	Mem	Conv
LKDN	29.01dB	78.49ms	0.129M	8.29G	202.70M	652.41M	86

Table 2 shows the metrics of LKDN.

- Which pre-trained or external methods / models have been used (for any stage, if any)

- Which additional data has been used in addition to the provided NTIRE training and validation data (at any stage, if any)

  The training data consists of 800 images from DIV2K [1] and the first 2000 images from LSDIR.

- Training description

- Testing description

- Quantitative and qualitative advantages of the proposed solution

- 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

- 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.

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2023 workshop.

- General comments and impressions of the NTIRE 2023 challenge.

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

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

References

[1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *CVPRW*, pages 126–135, 2017. 4

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

[3] 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 *cvprw*, pages 833–843, 2022. 2

[4] Lin Zhou, Haoming Cai, Jinjin Gu, Zheyuan Li, Yingqi Liu, Xiangyu Chen, Yu Qiao, and Chao Dong. Efficient image super-resolution using vast-receptive-field attention. In *ECCVW*, pages 256–272, 2023. 2