NTIRE 2023 Image Denoising ($\sigma = 50$) Challenge Factsheet
Dual-View U-Net Based Transformer for High-Resolution Image Restoration

Huabin Yang, Zhongjian Zhang, Yanru Zhang
University of Electronic Science and Technology of China
Chengdu, China
huabinyang12@gmail.com
zhongjianzhang6972@gmail.com
yanruzhang@uestc.edu.cn

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 image denoising with noise level $\sigma = 50$.

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 Image Denoising 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 noisy and denoise 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
  yiriyou
- Team leader name
  Huabin Yang
- Team leader address, phone number, and email
  Chengdu, China;
  15902888694;
  huabinyang12@gmail.com
- Rest of the team members
  Zhongjian Zhang and Yanru Zhang
- Team website URL (if any)
  None
- Affiliation
  University of Electronic Science and Technology of China
- Affiliation of the team and/or team members with NTIRE 2023 sponsors (check the workshop website)
  None
- User names and entries on the NTIRE 2023 Codalab competitions (development/validation and testing phases)
  User name:yanghuabin;
  Entries for development/validation phase: 13
  Entries for testing phase: 2
- Best scoring entries of the team during development/validation phase
  The best scoring entry is 13 (PSNR: 28.92 SSIM: 0.81) during developing phase
- Link to the codes/executables of the solution(s)
  GitHub Respository

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.)
- Representative image / diagram / pipeline of the method(s)
- Training strategy
- Experimental results
- 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)
- 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)
- 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.

We propose a dual-view U-Net based Transformer for high-resolution image restoration. As shown in Figure 1, our model consists of two sub-modules. Each of them models information from channel and pixel dimensions at different scales, respectively. Each up-sampling or down-sampling operation is followed by a **L**ocally-**e**nhanced **Win**dow (LeWin) Transformer block [1]. Benefiting from the hierarchical architecture of U-Net and the window

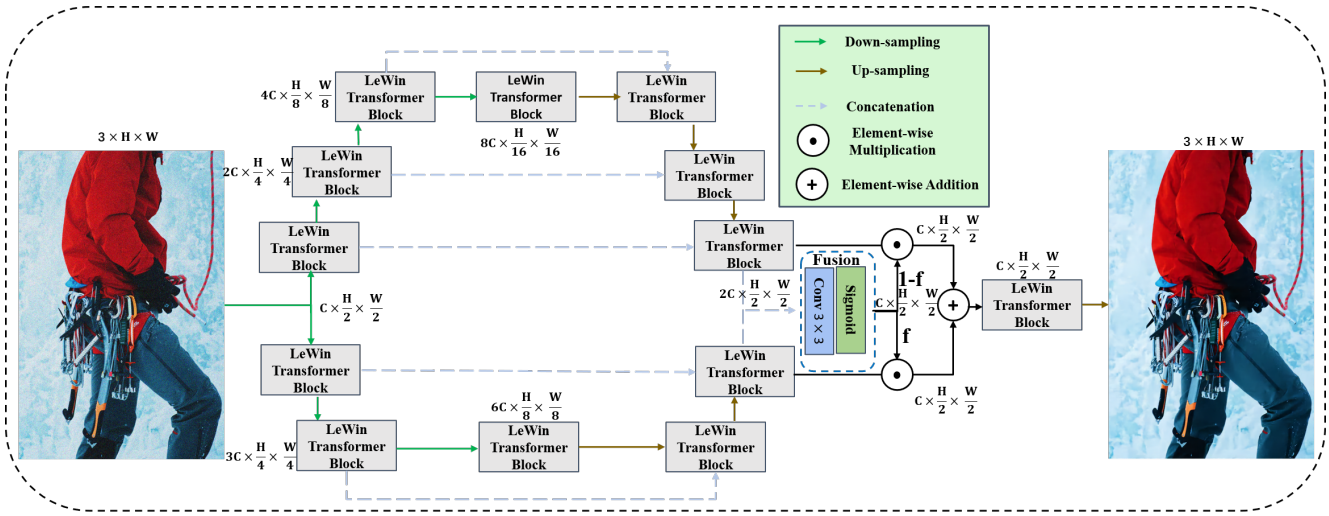

Figure 1. The architecture of our proposed framework.

mechanism, the LeWin Transformer block is capable of capturing long-range dependencies at low-resolution feature maps while reducing the computational cost. Specifically, we use 4×4 convolution block with a stride of 2 and a padding of 1 for down-sampling, and 2×2 transposed convolution with stride of 2 for up-sampling. The top sub-module has three sub-layers, we triple the channels at each sub-layer along the down-sampling path and do the opposite along the up-sampling path. However, the bottom sub-module has only two sub-layers, we expand the channels hierarchically by a factor of 2 along the down-sampling path and reduce the channel capacity by half along the up-sampling path. After that, the output feature maps of the two sub-modules are concatenated and fed into the fusion block, which consists of a convolution and a sigmoid activation. Finally, an optimal combination between two feature maps can be attained and used to recover the clean image,

$$f = sigmoid(W\,[x, y] + b), \tag{1}$$

$$z = f \odot x + (1 - f) \odot y \tag{2}$$

where $[\,,\,]$ denotes tensor concatenation, $\odot$ denotes element-wise multiplication, W, b are trainable parameters, f denotes the learned weights, x and y denote the input feature maps and z denotes the weighted sum of feature maps.

In our implementation, the objective function is the L1-Loss, the batch size is 1, the window size is 8, the optimizer is Adam with $\beta_1 = 0.9$ and $\beta_2 = 0.99$, the learning rate is 2e-3 with a decay factor of 0.5 for every 20 epochs, the total training epochs is 200. Our proposed framework achieved 27.94 on PSNR and 0.82 on SSIM in the testing phase.

In our experiments, no pre-trained models or data other than the training data and validation data provided by NTIRE organizers were used. All the experiments were conducted on a sever with an Inter Xeon Gold 6142M CPU and an NVIDIA GeForce RTX 3090 GPU.

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2023 workshop.
  None
- General comments and impressions of the NTIRE 2023 challenge.
  None
- What do you expect from a new challenge in image restoration, enhancement and manipulation?
  None
- Other comments: encountered difficulties, fairness of the challenge, proposed subcategories, proposed evaluation method(s), etc.
  None

References

[1] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general u-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17683–17693, 2022.