NTIRE 2022 Efficient SR Challenge Factsheet
Modified Residual Feature Distillation Network

Dr. Pratik Narang
BITS Pilani
Pilani, Rajasthan
pratik.narang@pilani.bits-pilani.ac.in

Usneek Singh
BITS Pilani
Pilani, Rajasthan
f20190127@pilani.bits-pilani.ac.in

Syed Sameen
BITS Pilani
Pilani, Rajasthan
p20190412@pilani.bits-pilani.ac.in

Harsh Khaitan
Kwikpic Tech. Services
Kolkata, India
harsh@kwikpic.in

1. Introduction

This factsheet template is meant to structure the description of the contributions made by each participating team in the NTIRE 2022 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 2022 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 2022 challenge report and to submit papers with their solutions to the NTIRE 2022 workshop. Detailed descriptions are much appreciated.

The factsheet, source codes/executables, trained models should be sent to **all of the NTIRE 2022 challenge organizers (Yawei Li, Kai Zhang, and Radu Timofte)** by email.

2. Email final submission guide

To: yawei.li@vision.ee.ethz.ch
kai.zhang@vision.ee.ethz.ch
radu.timofte@vision.ee.ethz.ch
cc: your_team_members
Title: NTIRE 2022 Efficient SR Challenge - TEAM_NAME

Body contents should include:

a) team name

b) team leader's name and email address

c) rest of the team members

d) team name and user names on NTIRE 2022 CodaLab competitions

e) executable/source code attached or download links.

f) factsheet attached

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.
3. Put your pretrained model under the `model_zoo` folder.
4. Modify `model_path` in `test_demo.py`. Modify the imported models.

5. `python test_demo.py`

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- Multicog
- Team leader name- Dr. Pratik Narang
- Address/ phone no/ email- Department of CSIS, BITS Pilani, +91 96543 51084 , pratik.narang[at]pilani.bits-pilani.ac.in
- Other memebers- Usneek Singh, Syed Sameen, Harsh Khaitan
- Affiliation- BITS Pilani, Pilani
- User names and entries on the NTIRE 2022 Codalab competitions (development/validation and testing phases)
  User names on Codalab:-
  - Pratik Narang- pnarang
  - Usneek Singh - Usneek
  - Syed Sameen- Syed
  - Harsh Khaitan - harshkhaitan

  Entries are by the name of Usneek and Syed.
- Best scoring entries of the team during development/validation phase
  The best entry during the validation phase had a PSNR of 28.36 https://codalab.lisn.upsaclay.fr/competitions/1865#participate-submit_results
- Link to the codes/executables of the solution(s) https://github.com/usneek/Modified-RFDN

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.

The work in this model is inspired from [2] Residual Feature Distillation Network. Three major modifications have been proposed in the network that reduces the parameters and runtime while maintaining similar accuracy for the model -

- The normal convolution replaced by multiception layers in residual block
- The l1 loss changed to Charbonnier loss for training
- The up-sample block with Pixelshuffle changed to pixel attention block inspired from [3].

4.2.1 The basic architecture of Residual Distillation Network

[2] is the winner of NTIRE-2020 super-resolution challenge. It has 3 major stages in the pipeline- first the features are extracted using a convolution layer, then the feature vector is processed through the distillation block and finally the up-sample blocks constructs the image in required scale. It is more lightweight than IMDN network which has been set as the baseline model for this competition.

4.2.2 Architectural changes

- **Multiception layers**- As described by [1], Multiception significantly reduces the number of parameters of standard convolution-based models by 32.48% on average while still preserving accuracy. In a depthwise convolution, a single convolution is applied to a single channel at a time unlike standard CNN's in which it is done for all the channels. It has fewer parameters and is computationally cheaper because of less number of operations. Multiception method involves using depthwise convolution with kernels of different sizes like 1 x 1, 3 x 3, 5 x 5, 7 x 7. The number of parameters in a multiception layer using 3 kernels is given by the equation:-

  $$p = 3c * (\sum_{j=1}^{3} K_j^2 + N) \tag{1}$$

  where Kj is the size of kernel, c and N are number of input and output channels respectively.
- **Upsampling block** The [3] talks about pixel attention network in the reconstruction stage. This network replaces the standard Pixelshuffle layer in the network

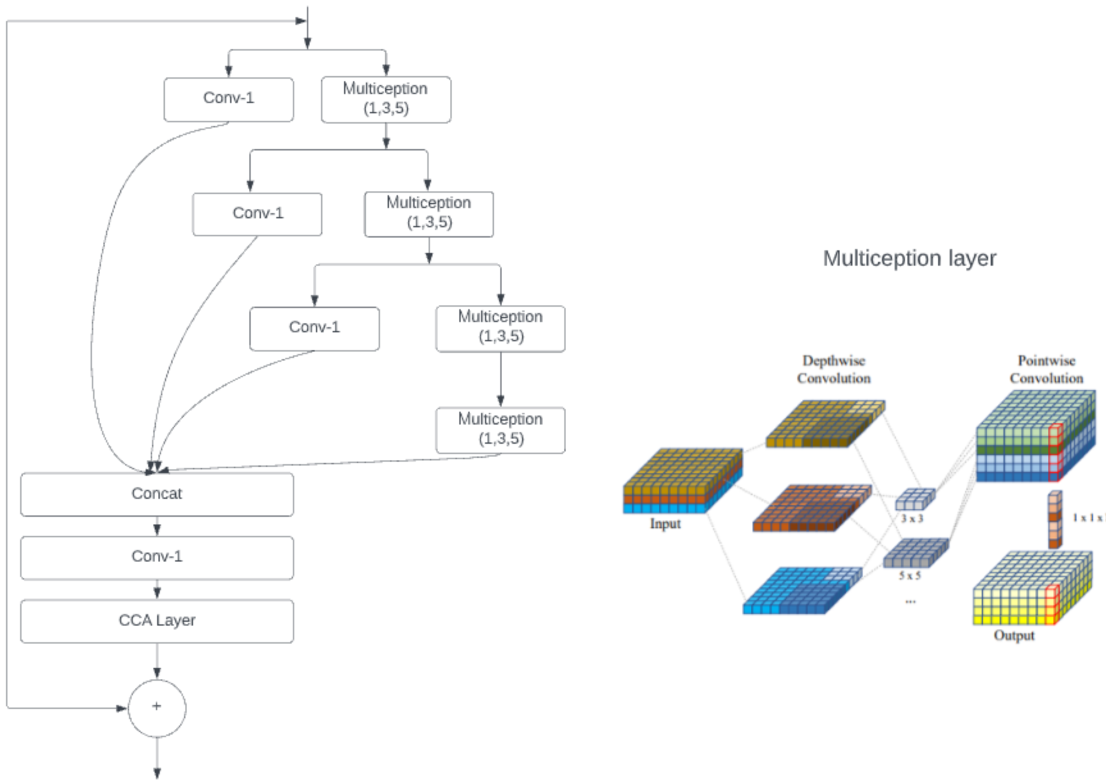

Figure 1. The residual block in modified RFDN

with nearest neighbour interpolation layer that reduces the parameters and computation cost. Moreover, previous works have shown that attention mechanism can effectively improve the performance in SR tasks. channels respectively.

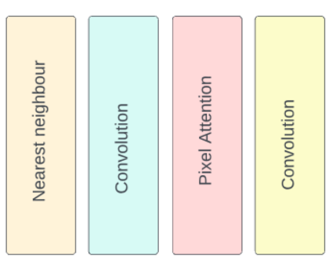

Figure 2. Structure of up-sampling block

- **Loss function while training** Smooth L1 loss is used for training. It is a combination of L1 loss and L2 loss. It perfectly avoids the flaws of L1 and l2 loss. Smooth L1 Loss is less sensitive to outliers than L2 Loss, or more robust. The magnitude of the gradient can be controlled .

4.2.3 Training Strategy

In this model, patches of size 256 x 256 are randomly cropped from the LR images as input for each training mini batch of size 64.Training data is augmented with random horizontal flips and 90 rotations.Model is trained with ADAM optimiser with learning rate set to 5 x 10-4. The learning rate is halved after every 200 iterations. The model is trained for 200 epochs. m-RFDN uses a channel number of 52 to achieve a better reconstruction quality. The networks are implemented by using PyTorch framework . DIV-2K dataset is used as training data. Flickr-2K is used as validation data.

4.2.4 Experimental Results

Using the above mentioned training details, model is trained and compared with original RFDN model on the test dataset provided in the competition. Analysis of PSNR, parameters and runtime is formulated in the table given below.

Method	PSNR	Parameters	Runtime
RFDN	28.14	311575	0.2s
m-RFDN	28.16	441568	0.2s

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2022 workshop- We plan to submit our solution to NTIRE 2022 highlighting the changes in the work of RFDN that reduces the parameters and runtime while preserving the accuracy.

- General comments and impressions of the NTIRE 2022 challenge- This challenge has motivated researchers from computer vision community to design more efficient approaches for image super-resolution. Such approaches are more practical and deploy-able on mobile devices. This challenge is big step towards bridging the gap between theory and practice.

- What do you expect from a new challenge in image restoration, enhancement and manipulation? We expect challenges to involve more practical aspects calling it as Real-World Image restoration. The real-life images are very different from the artificially created images in the dataset. The focus can also be done on interpretability of super-resolution models with large scaling factors.

- Other comments: encountered difficulties, fairness of the challenge, proposed subcategories, proposed evaluation method(s), etc- The challenge was conducted in a fair manner and proposed evaluation methods are effective.

References

[1] Guoqing Bao, Manuel B Graeber, and Xiuying Wang. Depthwise multiception convolution for reducing network parameters without sacrificing accuracy. In *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 747–752. IEEE, 2020. 2

[2] Jie Liu, Jie Tang, and Gangshan Wu. Residual feature distillation network for lightweight image super-resolution. In *European Conference on Computer Vision*, pages 41–55. Springer, 2020. 2

[3] Hengyuan Zhao, Xiangtao Kong, Jingwen He, Yu Qiao, and Chao Dong. Efficient image super-resolution using pixel attention. In *European Conference on Computer Vision*, pages 56–72. Springer, 2020. 2