NTIRE 2025 Efficient SR Challenge Factsheet
-title of the contribution-

Shengyun Zhong
Northeastern University, USA
shengyunzhong2002@gmail.com

Mingyang Wu
Texas A&M University, USA
mingyang@tamu.edu

Renjie Li
Texas A&M University, USA
renjie@tamu.edu

Yushen Zuo
The Hong Kong Polytechnic University, Hong Kong
zuoyushen12@gmail.com

Zhengzhong Tu
Texas A&M University, USA
tzz@tamu.edu

1. Introduction

This factsheet template is meant to structure the description of the contributions made by each participating team in the NTIRE 2025 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 2025 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 2025 challenge report and to submit papers with their solutions to the NTIRE 2025 workshop. Detailed descriptions are much appreciated.

The factsheet, source codes/executables, trained models should be sent to **all of the NTIRE 2025 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
cshguo@gmail.com
zongwei.wu@uni-wuerzburg.de
timofte.radu@gmail.com
CC:
your_team_members

Title: NTIRE 2025 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 the 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 2025 CodaLab competitions

e) Code, pre-trained 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 training data) should also be explained in detail.

4.1. Team details

- Team name: TACO_SR
- Team leader: Shengyun Zhong
- Address: 935N 72nd St, 98103, Seattle, WA, USA
  Phone: +1 (510)916-9891
  email: shengyunzhong2002@gmail.com
- Rest of members: Mingyang Wu, Renjie Li, Yushen Zuo, Zhengzhong Tu
- Team URL: https://taco-group.github.io/
- Affiliation:
  Shengyun Zhong: Northeastern University, USA
  Mingyang Wu: Texas A&M University, USA
  Renjie Li: Texas A&M University, USA
  Yushen Zuo: The Hong Kong Polytechnic University, Hong Kong
  Zhengzhong Tu: Texas A&M University, USA
- Affiliation of the team and/or team members with NTIRE 2025 sponsors: None
- User names: ShelvinZhong
  Entries(development/validation phases): 6
  Entries(testing phases): 2
- Link to the codes/executables of the solution(s): https://github.com/ShelvinZhong/TenInOneSR_EVAL

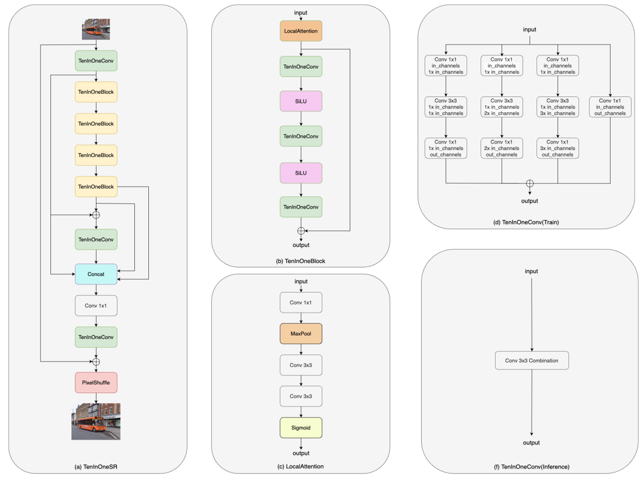

Figure 1. The architecture of our proposed TenInOneSR.

4.2. Method details

Method. The overall architecture of their network is showed in the Figure(a), inspired by SPAN [2] and PFDNLite [1]. Motivated by the design of the Conv3XC module in SPAN, they introduce two additional parallel branches with varying channel expansion ratios, resulting in a novel convolution module termed TenInOneConv, which fuses multiple convolution kernels into a single equivalent kernel to improve inference efficiency. Furthermore, to enhance the model's capability in capturing local texture and detail features, the LocalAttention module inspired by PFDNLite is integrated, allowing the network to better focus on informative regions within feature maps.

TenInOneSR employs four TenInOneBlock modules. Each of these blocks (detailed in Figure (b)) begins with a LocalAttention module, which enhancing the network's ability to capture fine details. Subsequently, each block applies three cascaded TenInOneConv layers, interleaved with the SiLU activation function, to perform hierarchical feature refinement. The block concludes with a residual connection, allowing better gradient flow.

Notably, the behavior of the TenInOneConv differs between the training and inference phases. During training (Figure (d)), TenInOneConv operates in a multi-branch configuration. It introduces three parallel convolutional branches with different channel expansion ratios (gains set as 1, 2, and 3), along with an additional skip connection. This multi-scale feature extraction enables the network to better aggregate complementary spatial features.

In the inference stage (Figure (f)), for computational efficiency and faster runtime, these multiple convolution kernels are fused into a single equivalent convolution kernel. Specifically, the parallel branches and skip connection weights are mathematically combined to form one unified

3×3 convolutional kernel, significantly accelerating inference without compromising performance.

Training Details. The proposed architecture is trained on two NVIDIA RTX Titan GPUs with a total of 48 GB memory.

In the first training stage, the DIV2K dataset is augmented by a factor of 85$\times$ and registered into the LSDIR format, resulting in a large-scale training set containing 152,991 high-resolution RGB images. During this stage, training is conducted with 64 randomly cropped 256$\times$256 patches per batch, using common augmentations such as random flipping and rotation. The model is optimized using the Adam optimizer with L1 loss for a total of 100,000 iterations. The learning rate is initialized at 5×10^{-4} and decayed by half every 20,000 iterations.

In the second stage, we keep the training strategy and hyperparameters unchanged, except for increasing the input patch size to 384$\times$384 and reducing the batch size to 32 to fit GPU memory. Then another 100,000 training iterations are conducted to further improve the model's performance on higher-resolution textures.

In the second stage, we keep the training strategy and hyperparameters unchanged, except for increasing the input patch size to 384×384 and reducing the batch size to 32 to fit GPU memory.Then another 100,000 training iterations are conducted to further improve the model's performance on higher-resolution textures.

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2025 workshop.
- General comments and impressions of the NTIRE 2025 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] Bin Ren, Yawei Li, Nancy Mehta, Radu Timofte, and et al. The ninth ntire 2024 efficient super-resolution challenge report. In *CVPRW*, 2024. 2

[2] Cheng Wan, Hongyuan Yu, Zhiqi Li, Yihang Chen, Yajun Zou, Yuqing Liu, Xuanwu Yin, and Kunlong Zuo. Swift parameter-free attention network for efficient super-resolution. In *CVPRW*, 2024. NTIRE 2024 ESR Challenge. 2