NTIRE 2025 Image Super-Resolution ($\times 4$) Challenge Factsheet -SAKSRNet-

Sunder Ali Khowaja
Dublin City University
Dublin, Ireland.
sunderali.khowaja@dcu.ie

Ik Hyun Lee
IKLab and Tu Korea
Tu Korea, Siheung-Si, Republic of Korea
ihlee@tukorea.ac.kr

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 image super-resolution ($\times 4$).

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 decide. 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 appreciated.

The factsheet, source codes/executables, trained models should be sent to **all of the NTIRE 2025 challenge organizers (Zheng Chen, Jue Gong, Jingkai Wang, Kai Liu, Lei Sun, Zongwei Wu, Yulun Zhang, and Radu Timofte)** by email.

2. Email final submission guide

To: zhengchen.cse@gmail.com
leosun0331@gmail.com
g1017325431@gmail.com
normal.kliu@gmail.com
jingkaiwang100@gmail.com
zongwei.wu@uni-wuerzburg.de
yulun100@gmail.com
Radu.Timofte@uni-wuerzburg.de
cc: your_team_members
Title: NTIRE 2025 Image Super-Resolution ($\times 4$) 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 for comparing 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/[Team_ID]_[Model_Name]` folder
3. Put your pretrained model under the `model_zoo/[Team_ID]_[Model_Name]` folder
4. Modify `model_path` in `test.py`. Modify the imported models
5. `python test.py` (restore images, details in GitHub)
6. `python eval.py` (eval results, details in GitHub)

Please send us the command to download your code, e.g. `git clone https://github.com/sander-ali/SAKSRNet` When submitting the code, please remove the input and output images in the (any) `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: SAK_DCU
- Team leader name: Sunder Ali Khowaja
- Team leader address, phone number, and email: Dublin City University, Dublin, Ireland, +353899782455, and sunderali.khowaja@dcu.ie
- Rest of the team members: Ik Hyun Lee
- Team website URL (if any)
- Affiliation IKLab and TU Korea
- Affiliation of the team and/or team members with NTIRE 2025 sponsors (check the workshop website)
- User names and entries on the NTIRE 2025 Codalab competitions (development/validation and testing phases): avengers
- Best scoring entries of the team during the testing phase: 30.805
- Link to the codes/executables of the solution(s): To add

4.2. Method details

The network architecture of the proposed method is shown in Figure 1. The proposed image super resolution framework integrates multiple components which are designed to enhance feature extraction, preserve fine details and improve upsampling quality. The network comprises of multi-scale convolution block (MSCB) [4], gated convolution feature enhancement (GCFE) module [1], Swin transformer block (STB) [3], and Lightweight recurrent mechanism [5], which consists of gated convolution (GConv) and progressive pixel shuffle upsampling (PPSU) [6]. The components work in order to ensure high-quality image reconstruction from low-resolution inputs. We provide the detailed overview of the proposed method in the subsequent paragraphs.

The network takes low-resolution image (LR) as an input, which undergoes initial processing through hierarchical feature extraction. The Multi-scale convolution block (MSCB) is responsible for extracting multi-level spatial and contextual information by employing convolutional layers with varying strides and kernel sizes. The MSCB ensures to capture both fine and coarse details in an effective manner while improving the robustness of the extracted features. The MSCB also employs lightweight residual connection to maintain stability during training.

Once the initial feature maps are obtained, the gated convolution feature enhancement (GCFE) module processes them, accordingly. The said block replaces traditional convolutions with gated convolutions (GConv), which introduce dynamic feature selection mechanisms to improve the model's ability for suppressing noise while retaining texture and edge information. The block also considers short-term information bottlenecks to ensure efficient feature propagation and prevent over-smoothing of fine details.

The refined feature maps from the GCFE module are then processed through the swin transformer block (STB) to enhance global feature interactions. The STB leverages multi-head attention and LayerNorm to improve long-range dependency modeling. It also applies hierarchical window-based attention mechanism to refine image structures and mitigates distortions found in upsampled images.

The SAKSRNet further learns from the lightweight processing unit, which is a recurrent module included to optimize spatial feature aggregation. The unit captures spatial dependencies in a memory-efficient manner, allowing the model to refine the details across varying receptive fields in a dynamic manner.

Finally, the features that are refined are processed through progressive pixelshuffle upsamping (PPSU) module. The module progressively increases the spatial resolution of the image. In contrast to traditional upsampling methods, the PPSU employs sub-pixel convolutional layers to improve sharpness and minimize artifacts. The progressive characteristics of this module allows gradual enhancement of the reconstructed image that ensures well-preserved and fine textures in the high-resolution output.

We train the SAKSRNet using the ADAM optimizer [2] using a composite loss function comprising of L1-loss, Perceptual loss, and adversarial loss. The composite loss function deals with pixel-wise accuracy, ensuring perceptual similarity, and enhance realism [7]. The initial learning rate was set to 2×10^{-4} and leverages cosine annealing schedule, which reduces the learning rate to 10^{-6} after 500 iterations. We also employ a warm-up strategy in the first 5000 iterations to prevent sudden divergence and stabilize training. The model is trained with a batch size of 16 and a patch size of 64×64, ensuring an optimal balance between training efficiency and GPU memory usage. The network is trained on A6000 GPU cluster with 48MB memory. In order to improve the generalization, we apply various augmentation strategies, which include random rotations, vertical and horizontal flips, and Gaussian noise injection. Apart from the augmenta-

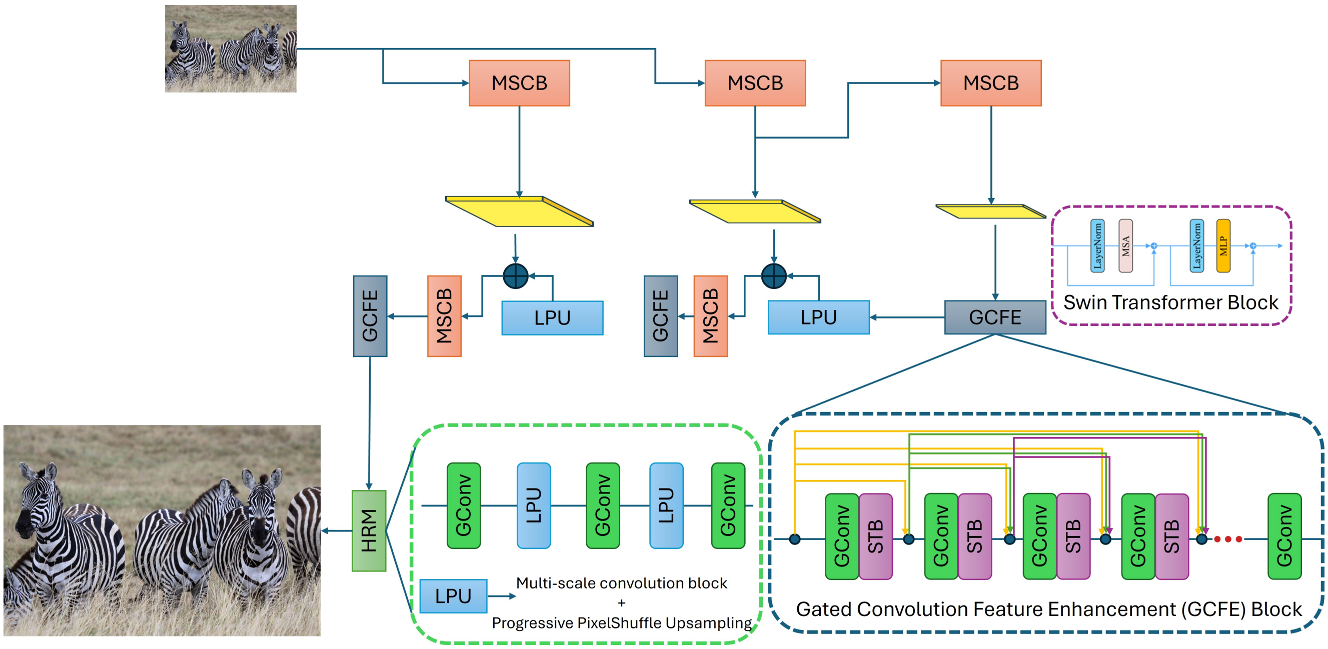

Figure 1. The network architecture for the proposed SAKSRNet

tions, we do not use any additional data to train the network.

Below we show couple of visual results in Figure 2 and Figure 3. to the show the efficacy of SAKSRNet. We have used an existing codebase of RTSR. A method that was part of NTIRE 2023 Image Super Resolution Method [7]. We intend to submit the extended version of this article to a Journal paper after making some improvements.

5. Other details

- Planned submission of a solution(s) description paper at NTIRE 2025 workshop: No, we intend to submit our work in a Journal after making some revisions to the method in order to improve the results.
- General comments and impressions of the NTIRE 2025 challenge: I have been participating in NTIRE challenges and I believe its the most organized challenge I participate in. Its always good to be a part of NTIRE 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] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua. Gated context aggregation network for image dehazing and deraining. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1375–1383, 2019. 2

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

[3] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pages 1833–1844, 2021. 2

[4] Qi Qi, Kunqian Li, Haiyong Zheng, Xiang Gao, Guojia Hou, and Kun Sun. Sguie-net: Semantic attention guided underwater image enhancement with multi-scale perception. *IEEE Transactions on Image Processing*, 31:6816–6830, 2022. 2

[5] Zhengxue Wang, Guangwei Gao, Juncheng Li, Yi Yu, and Huimin Lu. Lightweight image super-resolution with multi-scale feature interaction network. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2021. 2

[6] Dongyang Zhang, Changyu Li, Ning Xie, Guoqing Wang, and Jie Shao. Pffn: Progressive feature fusion network for lightweight image super resolution. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 3682–3690, 2021. 2

[7] Yulun Zhang, Kai Zhang, Zheng Chen, Yawei Li, Radu Timofte, Junpei Zhang, Kexin Zhang, Rui Peng, Yanbiao Ma, Licheng Jia, Huaibo Huang, Xiaoqiang Zhou, Yuang Ai, Ran He, Yajun Qiu, Qiang Zhu, Pengfei Li, Qianhui Li, Shuyuan Zhu, Dafeng Zhang, Jia Li, Fan Wang, Chunmiao Li, TaeHyung Kim, Jungkeong Kil, Eon Kim, Yeonseung Yu, Beomyeol Lee, Subin Lee, Seokjae Lim, Somi Chae, Heungjun Choi, ZhiKai Huang, YiChung Chen, YuanChun

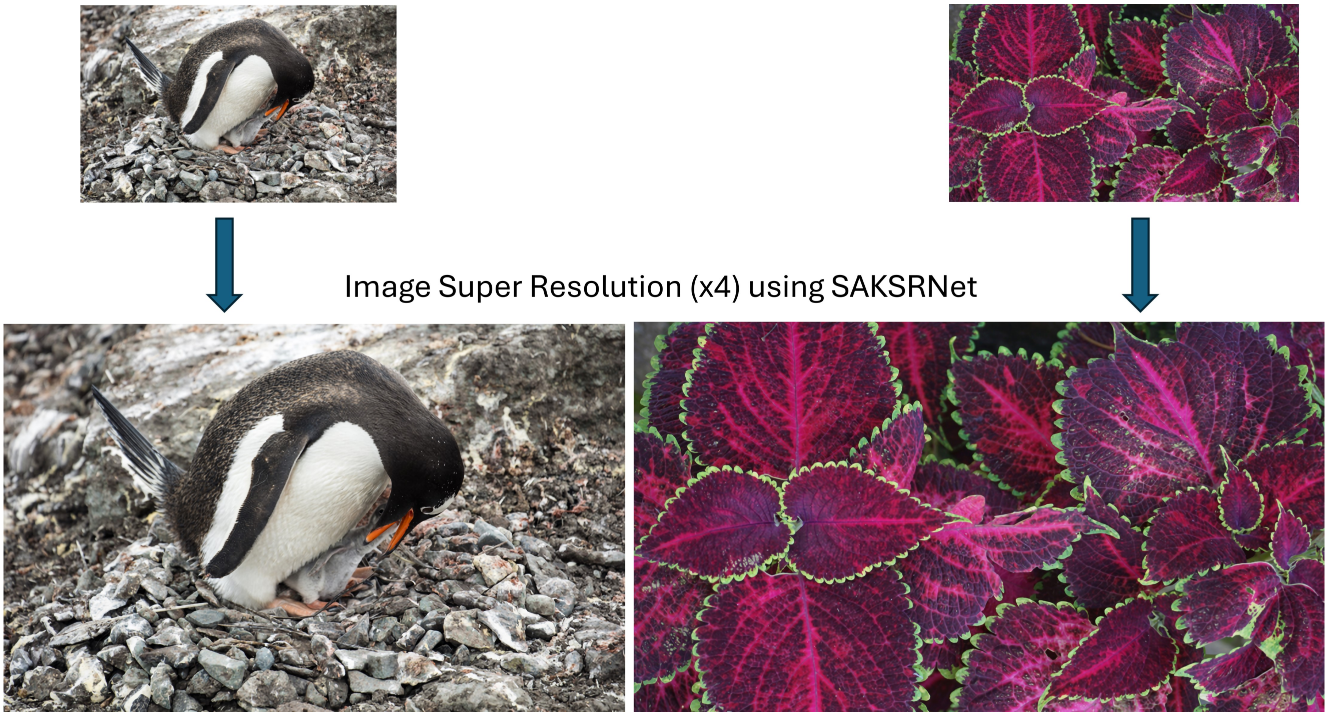

Figure 2. Enhanced x4 images from LR inputs using SAKSRNet

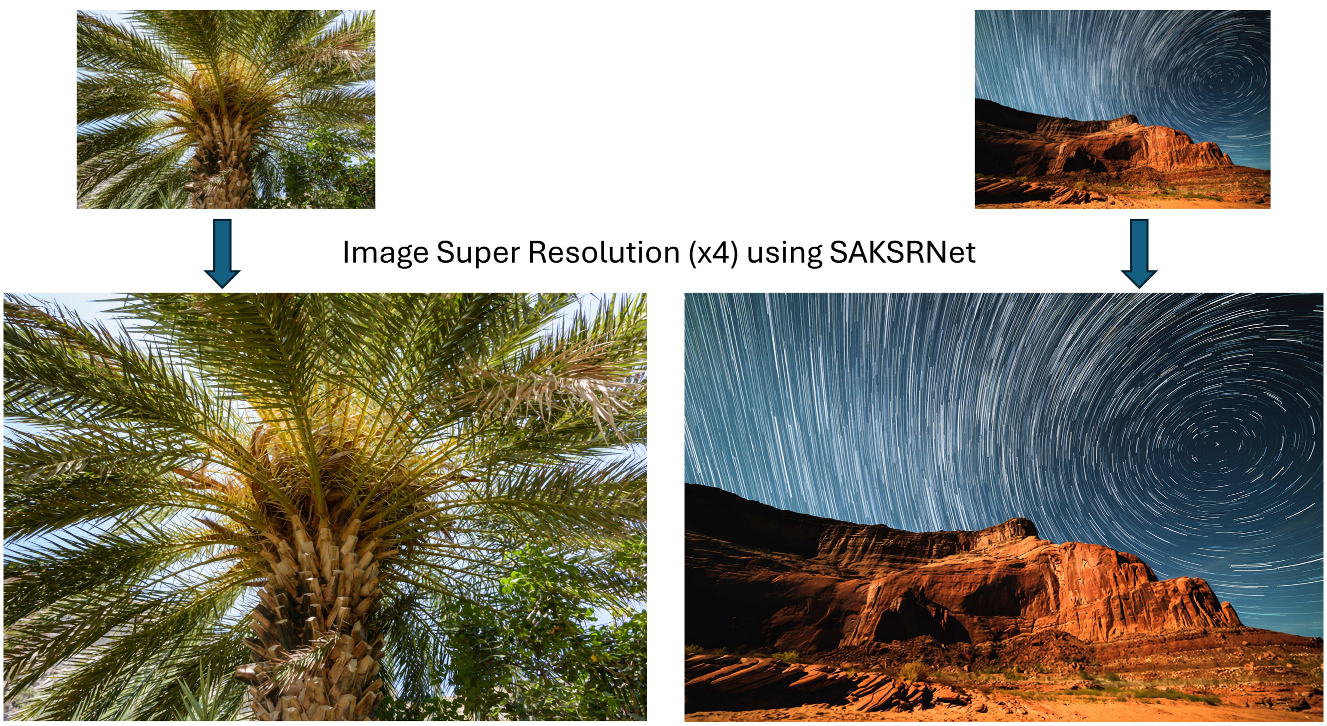

Figure 3. Enhanced x4 images from LR inputs using SAKSRNet

Chiang, HaoHsiang Yang, WeiTing Chen, HuaEn Chang, I-Hsiang Chen, ChiaHsuan Hsieh, SyYen Kuo, Ui-Jin Choi, Marcos V. Conde, Sunder Ali Khowaja, Jiseok Yoon, Ik Hyun Lee, Garas Gendy, Nabil Sabor, Jingchao Hou, Guanghui He, Zhao Zhang, Baiang Li, Huan Zheng, Suiyi Zhao, Yangcheng Gao, Yanyan Wei, Jiahuan Ren, Jiayu Wei, Yanfeng Li, Jia

Sun, Zhanyi Cheng, Zhiyuan Li, Xu Yao, Xinyi Wang, Danxu Li, Xuan Cui, Jun Cao, Cheng Li, Jianbin Zheng, Anjali Sarvaiya, Kalpesh Prajapati, Ratnadeep Patra, Pragnesh Barik, Chaitanya Rathod, Kishor Upla, Kiran Raja, Raghavendra Ramachandra, and Christoph Busch. Ntire 2023 challenge on image super-resolution (×4): Methods and results. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1865–1884, 2023. 2, 3