# NTIRE 2025 Image Denoising ($\sigma = 50$) Challenge Factsheet -Pixel Purifiers-

Deepak Kumar Tyagi
SRI-B
deepak.tyagi@samsung.com

Aman Kukretti
SRI-B
a.kukretti@samsung.com

Gajender Sharma
SRI-B
gajender.s@samsung.com

Sriharsha Koundinya
SRI-B
harsha.raj@samsung.com

Asim Manna
SRI-B
asim.m@partner.samsung.com

March 2025

## 1 Team details

- **Team name:** [Pixel Purifiers]
- **Team leader name:** [Deepak tyagi]
- **Team leader address, phone number, and email:** [2870, PHOENIX BUILDING, BAGMANE CONSTELLATION BUSINESS PARK, Ferns City, Doddanekkundi, Bengalur'u, Karnataka 560037, +919167516374, deepak.tyagi@samsung.com]
- **Rest of the team members:** [Aman Kukretti, Gajender Sharma, Sriharsha Koundinya, Asim Manna]
- **User names and entries on the NTIRE 2025 Codalab competitions:** [Username-deepak.tyagi]
- **Best scoring entries during development/validation phase:**
**Development Phase**: 30.69dB
**Testing Phase**: 29.84dB
- **Link to the codes/executables of the solution(s):** [GITHUB LINK-CLICK HERE]

# 2 Methodology

## 2.1 Architecture

Our proposed denoising method is based on Restormer [1] which is an efficient Transformer and its architecture uses multi-Dconv head transposed attention block (MDTA) for channel attention and gated-Dconv feed-forward network (GDFN) for Feed forward network. MDTA block applies Self Attention across channels rather than the spatial dimension to compute cross-covariance across channels to generate an attention map encoding the global context implicitly. Additionally, depth-wise convolutions are used to emphasize on the local context before computing feature covariance to produce the global attention map. GDFN block introduces a novel gating mechanism and depth-wise convolutions to encode information from spatially neighboring pixel positions, useful for learning local image structure for effective restoration.

## 2.2 Experiments

We conducted extensive experiments to evaluate the effectiveness of our approach:

- **Datasets:** The model is trained on both DIV2K and LSDIR [2] datasets with L1 loss as criterion.
- **Experimental Setup:** For training strategy, we use the flip and rotation for augmentation and used batch size of 256x256. The optimizer and the scheduler of the learning rate used are AdamW and CosineAnnealingRestartCyclicLR with initial value 1e-4. Model inference are done on 256x256 patch size. 8 Nvidia V100 (32GB) GPU were used for training the models.
- **Evaluation Metrics:** Metrics such as PSNR and SSIM to assess model performance.
- **Running Complexity:**
  Number of parameters (M) - 26.112
  Flops (G) - 140.99
- **Learning Rate Scheduling:** CosineAnnealingRestartCyclicLR with initial value 1e-4 is taken as Learning rate.
- **Regularization:** Weight decay were applied to prevent overfitting.
- **Data Augmentation:** Augmenting the training data patches helped improve the generalization capabilities of the models and model did not overfit to the data. Augmentation used: Left flipping , Right flipping, rotation in [90,180,270] degrees,etc.

Ensemble techniques played a crucial role in our experiments by effectively boosting performance through the aggregation of diverse model predictions, leading to more robust and accurate results.

### 2.3 Hard Dataset Mining

To improve PSNR further, we use hard datset mining:

Hard dataset mining [3] technique is used on trained model where training patches with loss greater than certain threshold are used for retraining the model. The learning rate used for this transfer learning is 100x smaller than actual learning rate such that model remains generalised.

### 2.4 Ensemble Techniques

To further boost accuracy and robustness, we employ ensemble techniques:

- **Self Ensemble :** Test time Augmentation ensemble [4] can improve model accuracy significantly and we used multiple flips and rotations for image before model inference. The outputs are averaged to create final output image.

## References


[1] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," 2022.

[2] Y. Li, K. Zhang, J. Liang, J. Cao, C. Liu, R. Gong, Y. Zhang, H. Tang, Y. Liu, D. Demandolx, R. Ranjan, R. Timofte, and L. Van Gool, "Lsdir: A large scale dataset for image restoration," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1775–1787, 2023.

[3] Y. Becker, R. Z. Nossek, and T. Peleg, "Sdat: Sub-dataset alternation training for improved image demosaicing," 2024.

[4] R. Timofte, R. Rothe, and L. V. Gool, "Seven ways to improve example-based single image super resolution," 2015.