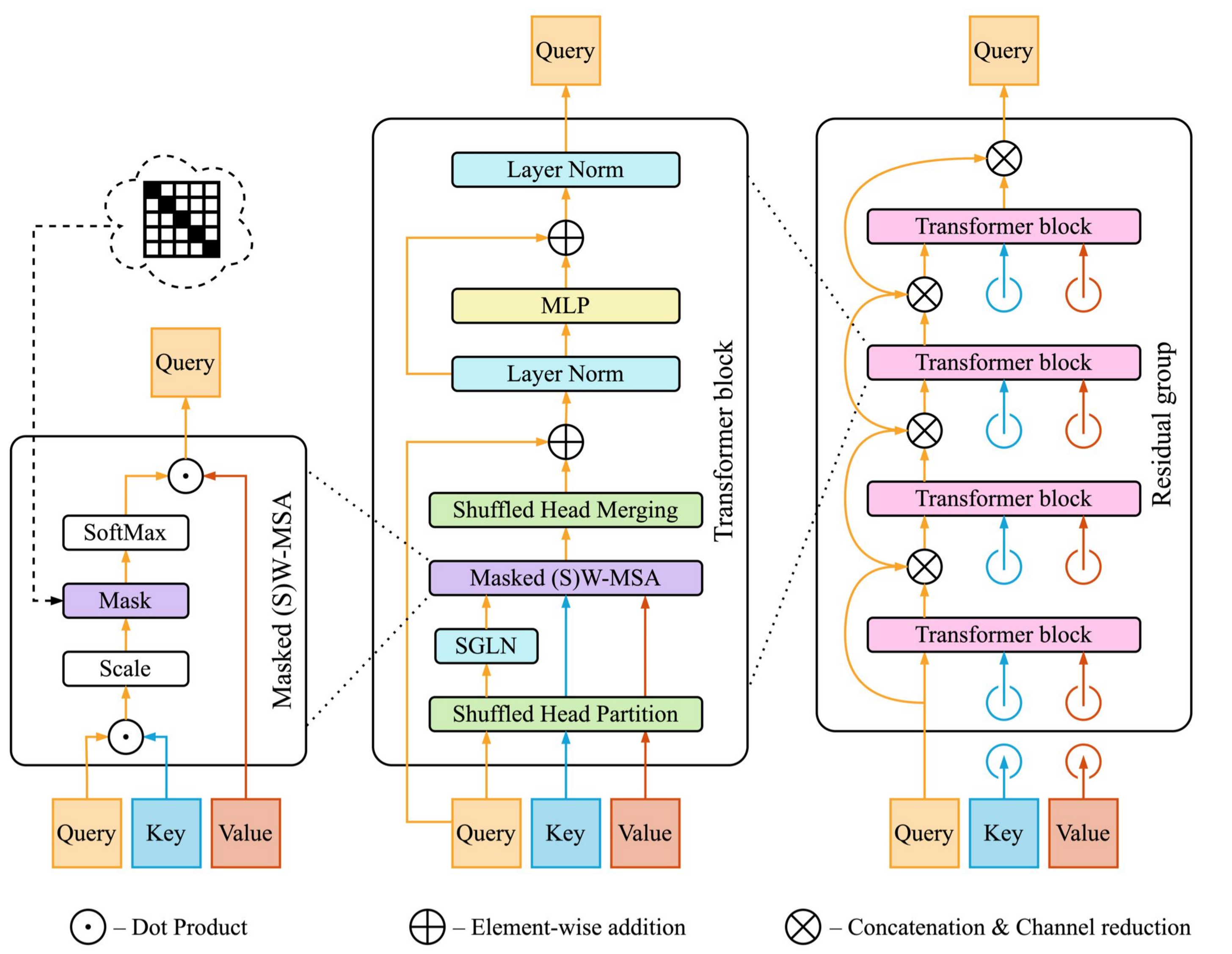
UNIVERSITY OF TARTU
Institute of Computer Science

PerkinElmer
For the Better

# SwinIA: Self-Supervised Blind-Spot Image Denoising without Convolutions


**Pavel Chizhov, Mikhail Papkov**SUP

Institute of Computer Science, University of Tartu
{pavel.chizhov,mikhail.papkov}@ut.ee
SUP – supervisor

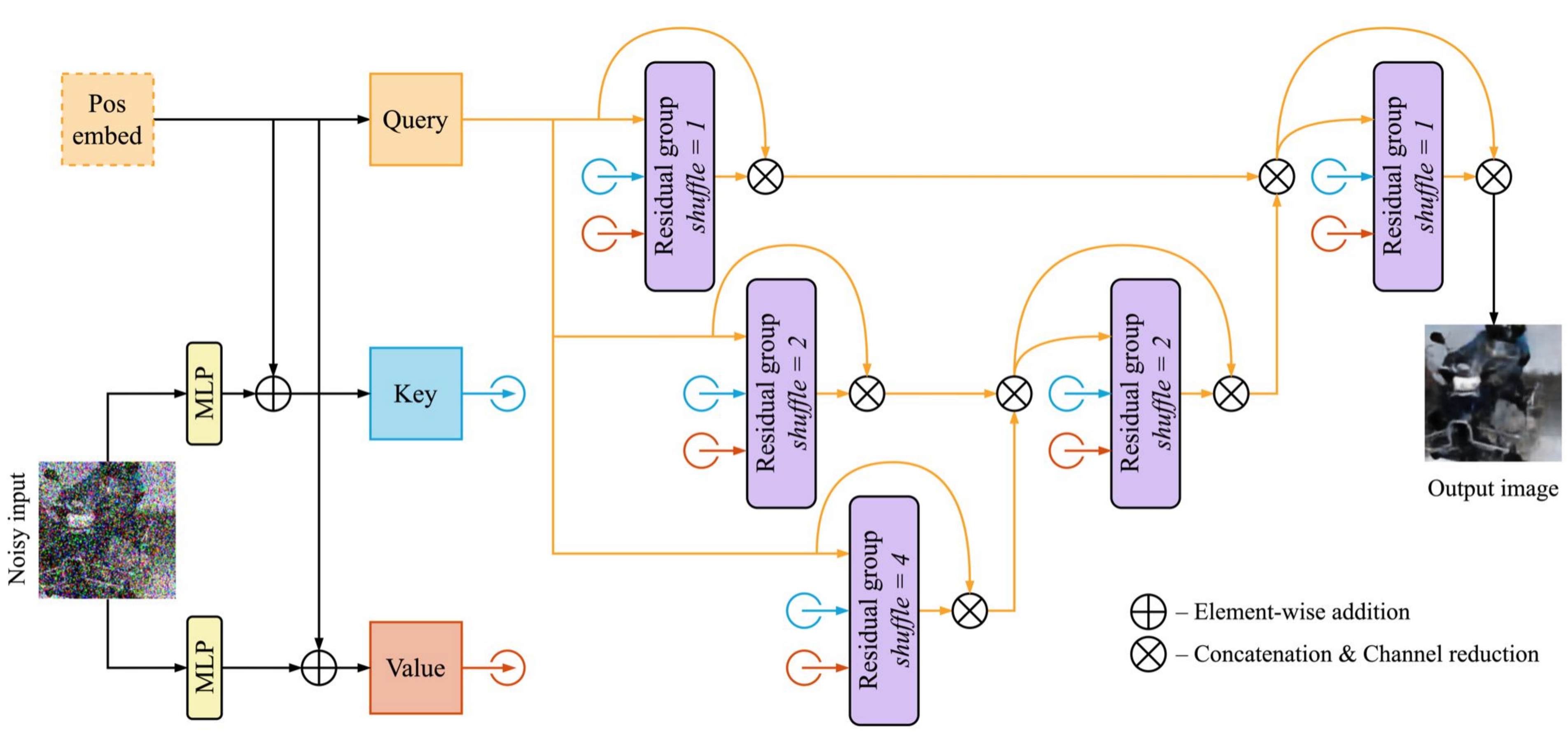




Figure 1. SwinIA model architecture. **Left:** a residual group with the transformer block details. **Right:** the complete U-shaped architecture with residual groups [1].

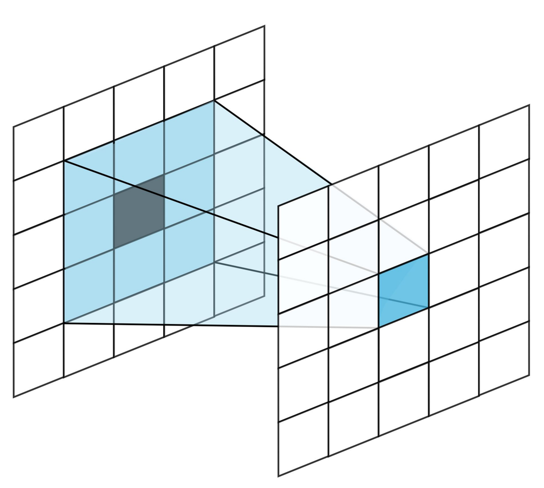


Figure 2. Blind-spot denoising

**Goal:** restore the underlying signal for a noisy image *without* clean images, *i.e.*, ground truth.

**Common tool:** blind spot network (BSN) – reconstruct pixel from its neighborhood (Figure 2).

**Challenge:** hard to avoid learning of the identity function.

## Why not just cut out the center from CNN kernels?

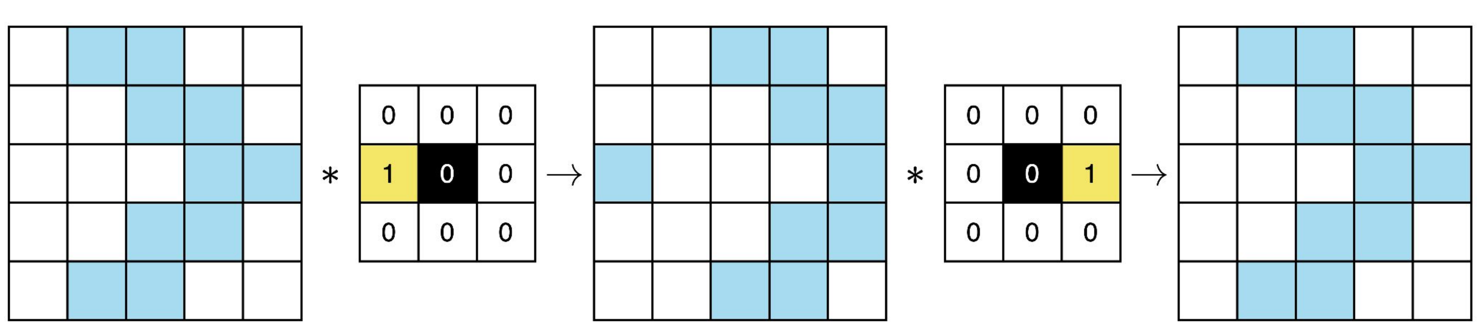


Figure 3. A counterexample to a convolutional BSN with nullified kernel center. The network learns a simple identity function via permutation in two consecutive filters.

## Popular methods

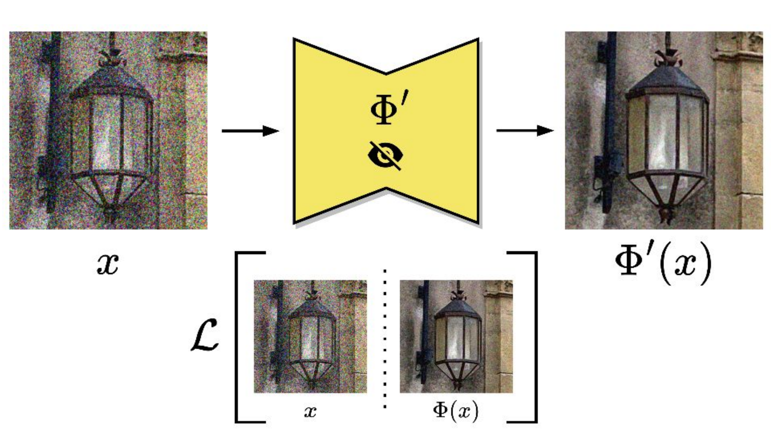



Figure 4. "True" blind-spot methods.

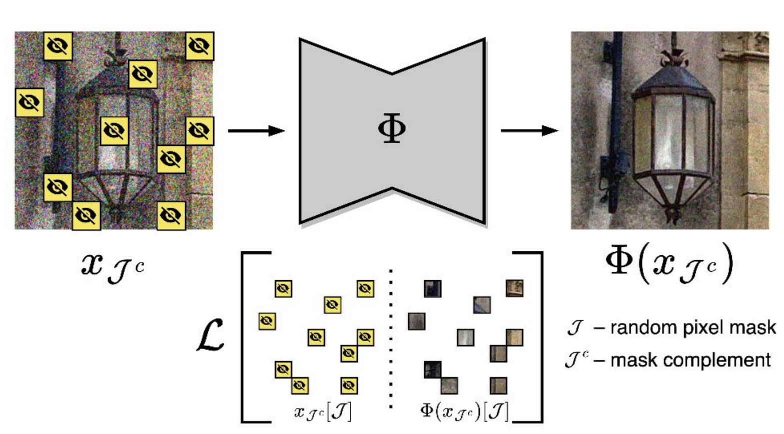



Figure 5. Mask-based methods.

| Method | Idea | Base | #passes |
|---|---|---|---|
| Laine19 | Image rotations | CNN | 4 |
| Honzatko | Dilated convolutions | CNN | 1 |
| DBSN | Knowledge distillation* | CNN | 2 |
| **SwinIA (ours)** | **Masked attention** | **T** | **1** |

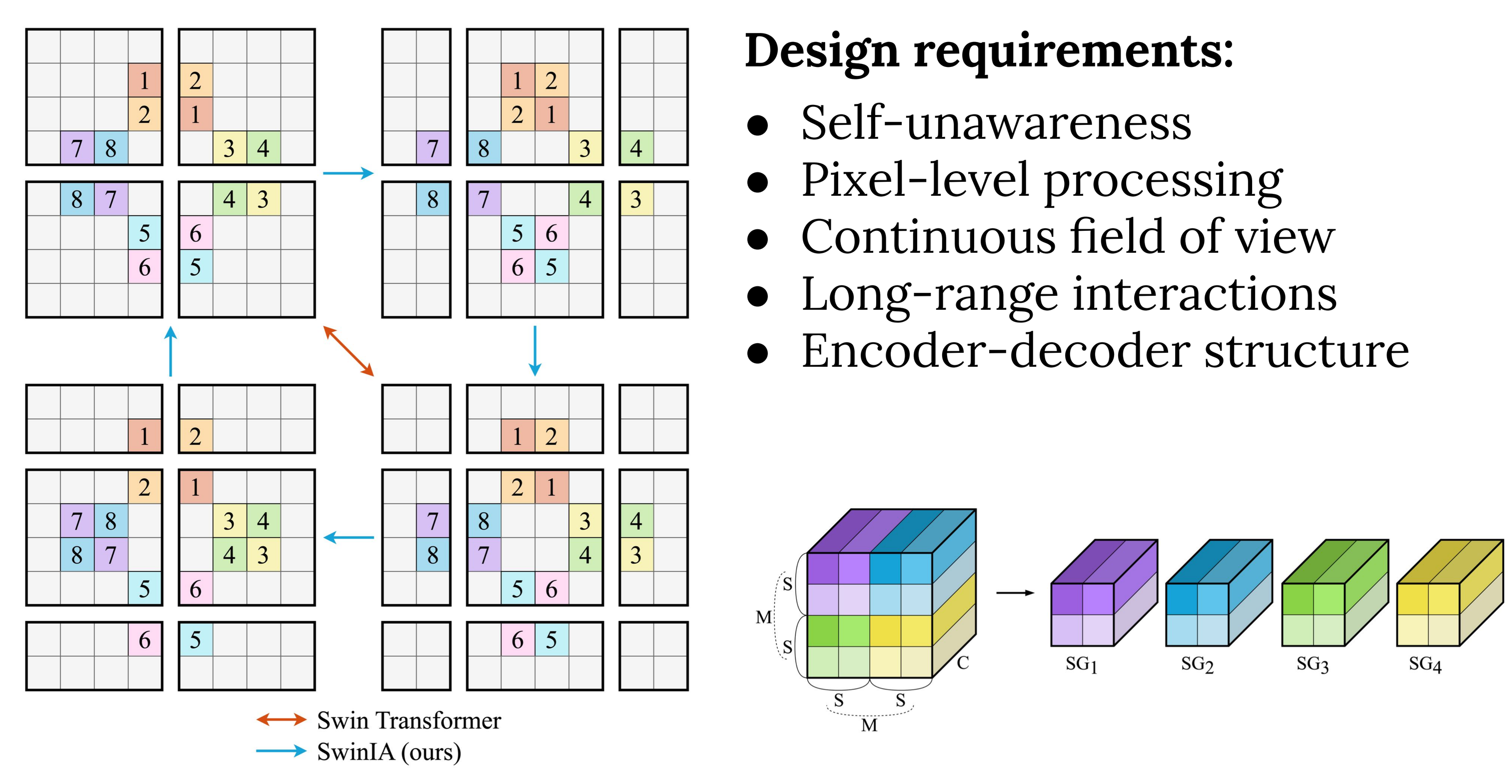
Table 1. "True" blind-spot methods.

| Method | Base | #passes |
|---|---|---|
| Noise2Self/ Noise2Void | CNN | 1 |
| Noise2Same | CNN | 2 |
| Blind2Unblind | CNN | 5 |
| CADT | T + CNN | 2 |

Table 2. Mask-based methods.

* – unpaired learning with clean images

## Design




**Design requirements:**

- Self-unawareness
- Pixel-level processing
- Continuous field of view
- Long-range interactions
- Encoder-decoder structure

Figure 6. SwinIA design decisions. **Left:** Window shifting approaches. Pairs of adjacent pixels that never participate in the same self-attention in Swin are enumerated 1 – 8. **Right:** Shuffle group partition example for shuffle S = 2 [1].

## Results

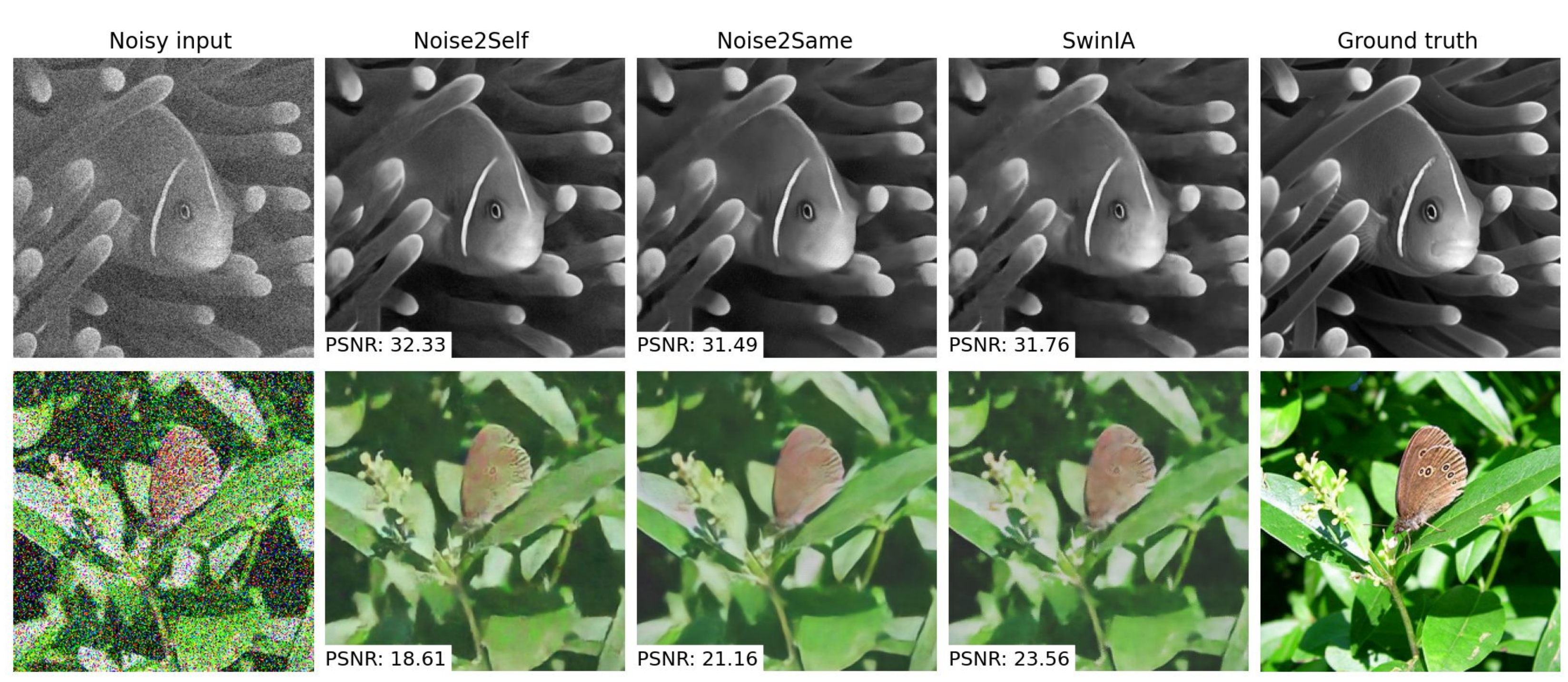



Figure 7. Results visualization example [1].

Evaluation on a total of 11 datasets: sRGB/Grayscale images with synthetic/natural noise of single/mixed modalities.

State of the art compared to other true BSNs on grayscale images and state of the art overall on sRGB images with mixed synthetic noise.

## Key takeaway

The first self-supervised image denoising model to combine:

- True blind-spot property
- Noise model unawareness
- Fully-transformer architecture
- Autoencoder training
- Single forward pass


## Acknowledgements

We thank PerkinElmer Inc. for support and High Performance Computing Center of the Institute of Computer Science at the University of Tartu for the provided computing power.


## References


[1] Mikhail Papkov and Pavel Chizhov. SwinIA: Self-Supervised Blind-Spot Image Denoising with Zero Convolutions. 2023. arXiv: 2305.05651 [cs.CV]