

УДК 004.932.4

Andrii Didenko¹, Andrii Oliinyk², Olena Syvachuk³

¹PhD student NU “Zaporizhzhia Polytechnic”

²Doctor of Technical Sciences, professor, NU “Zaporizhzhia Polytechnic”

³senior teacher NU “Zaporizhzhia Polytechnic”

NEURAL NETWORK METHODS FOR THERMAL IMAGE SUPER RESOLUTION

Thermal data is used in many areas of life and science, as it allows to analyze the infrared spectrum, which can contain a lot of useful information about the subject under research. It has applications in healthcare, military, construction, agriculture, etc. The main disadvantage of thermography is the low resolution of the output thermal images, which significantly reduces the

ability to unlock its full potential. In order to overcome this drawback, thermal image super-resolution (TISR) methods, including transformer neural networks, are used.

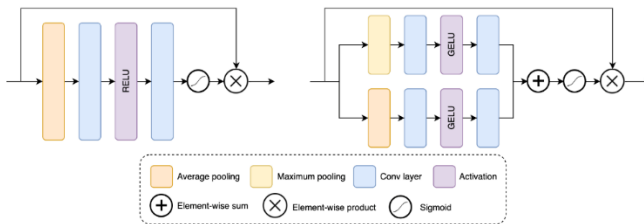
The described method was applied during the 5th TISR challenge of the PBVS-2024 workshop [1], one of the goals of which is to obtain super-resolution images at x8 scale from the reduced ones.

In recent years, visual transformer neural networks have gained popularity in ISR tasks due to their ability to capture complex dependencies and visual relationships on images. This is achieved through the attention mechanism, which analyzes spatial information across the entire image, leading to state-of-the-art results in image processing tasks. Among recent methods, SwinIR [2] and HAT [3] are worth considering due to their high performance.

SwinIR utilizes Swin Transformer [4] architecture for image restoration tasks, which uses shifted windowing schemes to manage self-attention computations within local windows, significantly reducing computational complexity while capturing both local and global contextual information effectively.

HAT architecture integrates channel attention mechanism and window-based self-attention into a transformer-based architecture specifically for the task of image restoration. This model combines the strengths of both attention types to focus on and enhance relevant features effectively.

During participation in the challenge and conducting experiments, several modifications to HAT model were introduced. Firstly, global skip connection with bicubic upsampling was added to the model architecture to enhance feature extraction and preserve information flow (HAT-bc model). Secondly, the Channel Attention mechanism was modified by adding maximum pooling layer [5] and changing ReLU to GELU activation function (HAT-ca model). Picture 1 shows the comparison between original and modified Channel Attention blocks. Table 1 represents the quantitative comparison between models on the challenge’s test set using SSIM and PSNR metrics.



Picture 1 – Original (left) and modified (right) Channel Attention block

Table 1 – Comparative results on the test set

Model	SSIM	PSNR
Baseline (bicubic)	0.7746	25.72
SwinIR	0.8202	26.90
HAT	0.8224	26.95
HAT-bc	0.8227	27.00
HAT-ca	0.8228	27.02

As a result, modifications of transformer neural network architecture, such as adding skip connection with auxiliary bicubic upsampling and enhanced attention mechanism, can improve the overall model's performance.

REFERENCES

1. PBVS 2024 Thermal Image Super-Resolution Challenge: (TISR) – Track1 [Text] [Electronic resource]. – Access mode: <https://codalab.lisn.upsaclay.fr/competitions/17013/> (Date of access 15.03.24).
2. Liang J., Cao J., Sun G., Zhang K., Van L. Gool, Timofte R., SwinIR: Image Restoration Using Swin Transformer [Text] / J. Liang J., J. Cao, G. Sun, K. Zhang, L. Gool Van, R. Timofte // IEEE/ CVF International Conference on Computer Vision Workshops (ICCVW), 2021. – P. 1833-1844.
3. Xiangy C., Xinta W., Z. Jiantao Z., Yu Q., Chao D., Activating More Pixels in Image Super-Resolution Transformer [Text] / C. Xiangy, W. Xinta, Z. Jiantao, Q. Yu, D. Chao // IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2023. – P. 22367-22377.
4. Liu Z., Lin Y., Cao Y., Hu H., Wei Y., Zhang Z., Lin S., Guo B., Swin Transformer: Hierarchical Vision Transformer using Shifted Windows: [Text] / Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin S., B. Guo // IEEE /CVF International Conference on Computer Vision (ICCV), 2021. – 204 p.
5. Woo S., Park J., J-Y. Lee J-Y., Kweon I.S., CBAM: Convolutional Block Attention Module [Text] / S. Woo, J. Park, J-Y. Lee, I.S. Kweon // European Conference on Computer Vision (ECCV), 2018. – P. 3–19.