# GENERATION OF RADAR IMAGES USING GENERATIVE ADVERSARIAL NETWORKS

*A. S. Minaev, V. V. Fedorov*

FSUE «RFNC – VNIITF named after Academ. E. I. Zababakhin», Snezhinsk, Russia

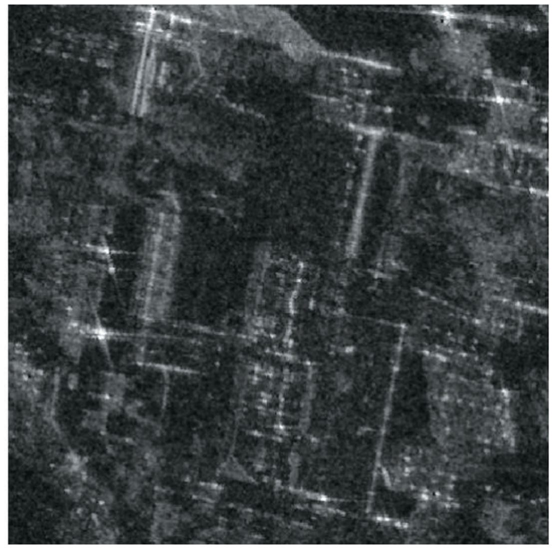
Snezhinsk Physics and Technology Institute, a branch of the Federal State Autonomous Educational Institution of Higher Education "National Research Nuclear University MEPhI," Snezhinsk, Russia

Radar imaging (SAR) is in high demand for Earth mapping and monitoring tasks due to its ability to provide all-weather, around-the-clock observation and to obtain information about the terrain and surface structure. Optical satellite images, in turn, offer high detail in object shapes and colors but are subject to weather and lighting constraints. Combining SAR and optical data allows for a more comprehensive characterization of an area; however, the volume of available radar data is often insufficient: high survey costs, limited temporal coverage, and complex physical effects complicate the collection of large datasets. In this context, the task of synthesizing plausible SAR images from optical images using generative adversarial networks (GAN) becomes highly relevant, offering the possibility to fill gaps in missing radar data and to supplement real observations.

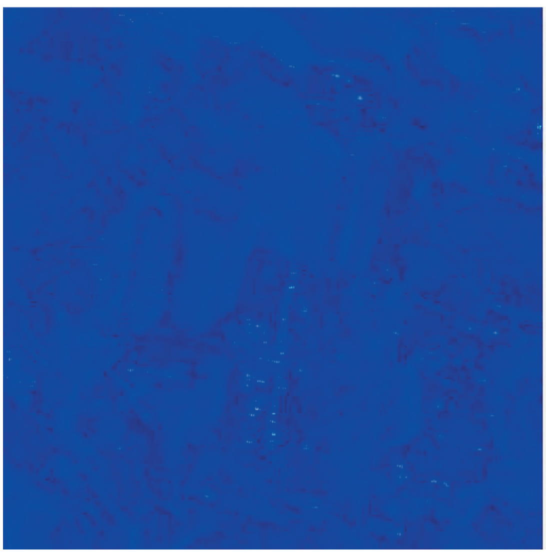
In this work, we propose a model for bidirectional translation between optical and radar domains that integrates ideas from Pix2Pix/UGATIT and attention mechanisms [1, 2]. The system includes two generators (for translating from optical to SAR and vice versa) and four discriminators: global (CAM discriminators) and local (PatchGAN-like). The inclusion of global and local discriminators makes it possible to control both the overall style of the scene and the realism of small details. To improve training stability, we use a combined loss function that includes adversarial and L1 criteria. This helps preserve the structure of objects and characteristic effects in SAR scenes (bright reflections, speckle noise) more effectively. The generators employ convolutional encoder-decoder blocks with residual layers (ResNet) as well as attention mechanisms based on CAM (Class Activation Mapping) and Swin Transformer [3]. The CAM module highlights the most significant fragments of the image by using global and maximum pooling (GAP/GMP) to form activation maps, while the Swin Transformer accounts for spatial context at different scales. ILN and AdaLIN (Instance/Layer Interpolation Normalization) provide flexibility in transitioning between channel-wise and layer-wise statistics. Particular attention is paid to compensating for shifts between optical and radar images. To address this, an optional registration module based on a convolutional network (ResUnet) can be included in the model to compute a deformation field, followed by alignment (Transformer_2D) of the generated and reference images.

Training and testing were conducted on the OpenEarthMap-SAR dataset, which contains paired optical and radar scenes (amplitude channels VV or HH) covering urban and rural landscapes in the US, Japan, and France, with a resolution of 0.15–0.5 m [4]. Of the 17,000 pairs sized 512×512, about 10% were reserved for the test set. According to the experimental results, the model demonstrates high realism of the generated SAR images: the average PSNR value reaches 19 dB, and the SSIM is 0.78. For the FID distribution metric, a value of about 42 was obtained, indicating the statistical similarity of the synthetic data to the real ones. Additional experiments using weighted metrics (Weighted SSIM, etc.) confirm stable generation quality across different landscape types. Figure 1 shows visual results for one of the test scenes, including the original optical image, the real SAR image, the corresponding CAM map, and the generated radar frame, clearly demonstrating the capabilities of the proposed approach.

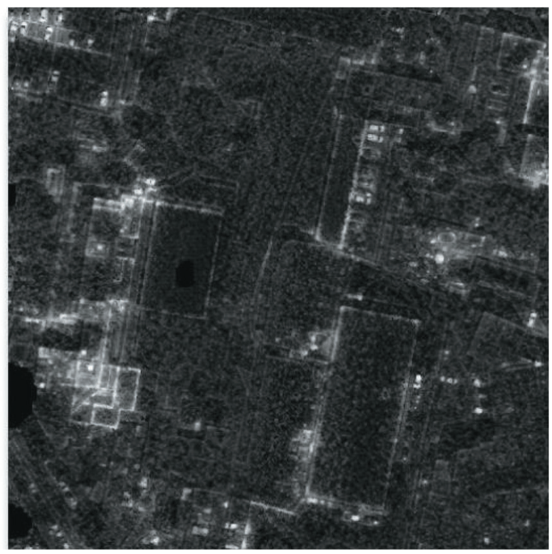
Оптическое изображение



Радиолокационное изображение



Карта сопоставления активационных классов



Сгенерированное изображение

Fig. 1. Visual results for one of the test scenes

## References


1. **Isola, P.** Image-to-image translation with conditional adversarial networks [Text] / P. Isola et al. // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – P. 1125–1134.
2. **Kim, J.** U-gat-it: unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation [Text] //arXiv preprint arXiv:1907.10830. – 2019.
3. **Liu, Z.** Swin transformer: Hierarchical vision transformer using shifted windows [Text] / Z. Liu et al. // Proceedings of the IEEE/CVF international conference on computer vision. – 2021. – P. 10012–10022.
4. **Xia, J.** OpenEarthMap-SAR: A Benchmark Synthetic Aperture Radar Dataset for Global High-Resolution Land Cover Mapping [Text] / J. Xia et al. // arXiv preprint arXiv:2501.10891. – 2025.