

A Two-Stage Method Using SegFormer and Convolutional Networks for Noise-Resistant InSAR Phase Unwrapping

Xiaohan Yang¹(220233419@seu.edu.cn), Xin Tian^{1*}(tianxin@seu.edu.cn), Mi Jiang^{2*}(jiangmi@csu.edu.cn)

^{1, 1*} School of Transportation, Southeast University, Nanjing, 211189

^{2*} School of Geosciences and Information Physics, Central South University, Changsha, Hunan, 410083

Abstract

This study presents an InSAR phase unwrapping method that integrates SegFormer with convolutional networks. To tackle phase discontinuities caused by noise, decorrelation, and abrupt topographic changes, a two-stage unwrapping strategy is employed. The first stage uses a filtering module based on the simplified fully convolutional network, OKNet, to preprocess noisy interferograms, effectively reducing noise while maintaining fringe continuity. In the second stage, an enhanced SegFormer network, named U-SegFormer, performs the phase unwrapping. Experimental results show that this method significantly outperforms traditional filtering and unwrapping approaches in terms of RMSE on simulated data with high noise and phase discontinuities, offering new insights into deep learning-based InSAR phase unwrapping methodologies.

Introduction & Objective

InSAR technology extracts surface information by analyzing phase differences captured at different times and locations. The images are formed by the conjugate multiplication of two complex SAR images, with phases constrained to $[-\pi, \pi]$, necessitating phase unwrapping for continuity. This process is NP-hard due to noise and topographical influences. Deep learning has shown exceptional performance in optical image processing and effectively addresses unwrapping challenges in low-coherence and discontinuous regions.

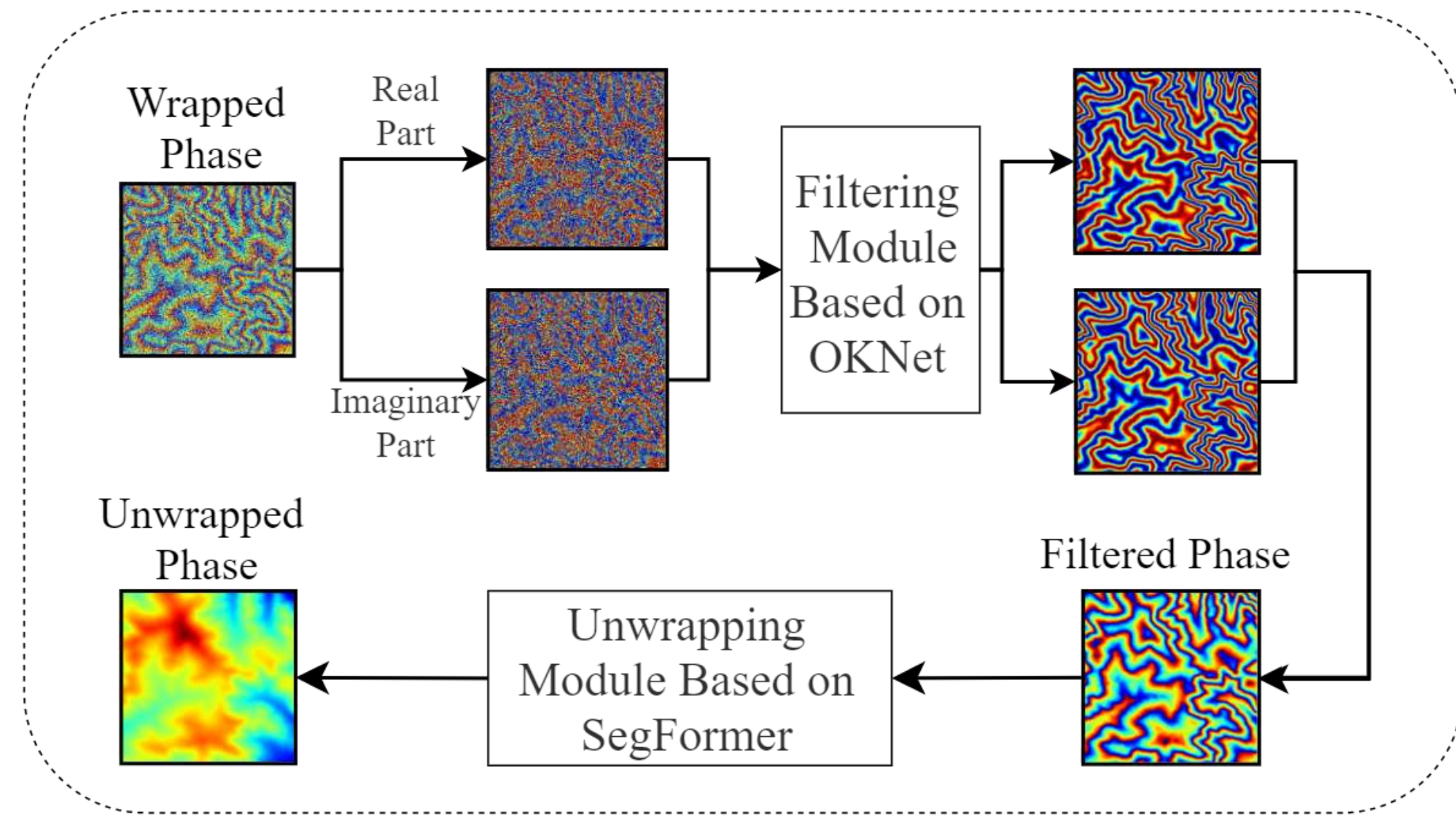


Figure 1: Overall Workflow of the Method

As shown in the figure1, We propose a two-stage deep learning method: first, a simplified fully convolutional network, OKNet[1], for Phase Filtering; second, an enhanced SegFormer network, U-SegFormer[2], for phase unwrapping.

Methods

The filtering module we used is based on the fully convolutional network OKNet, originally designed for image dehazing. This module comprises three branches: a local branch for capturing fine-grained details, a large-scale branch for extracting broad contextual features, and a global branch for modeling global dependencies through Dual-Domain Channel Attention (DCAM) and Frequency-based Spatial Attention (FSAM).

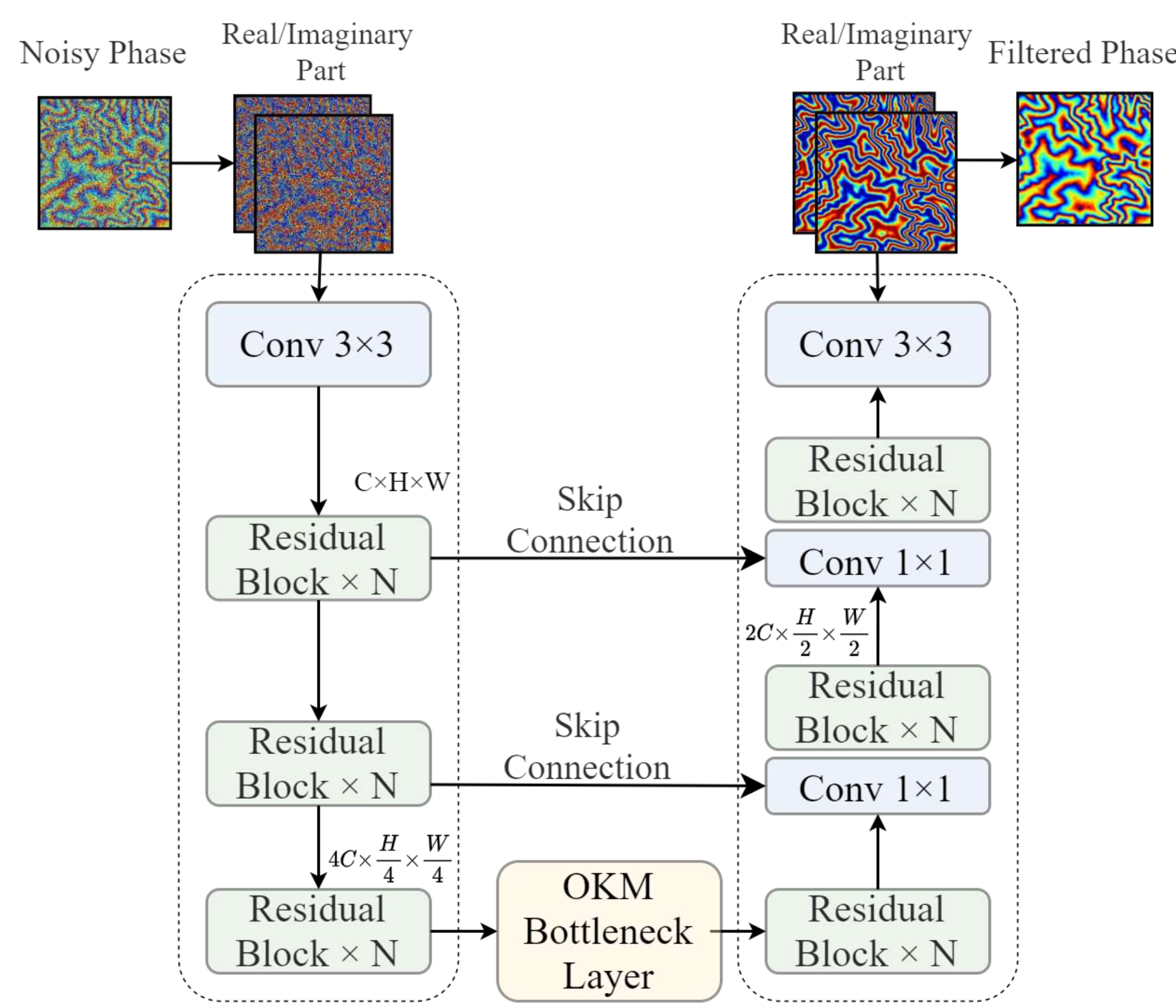


Figure 2. OKNet

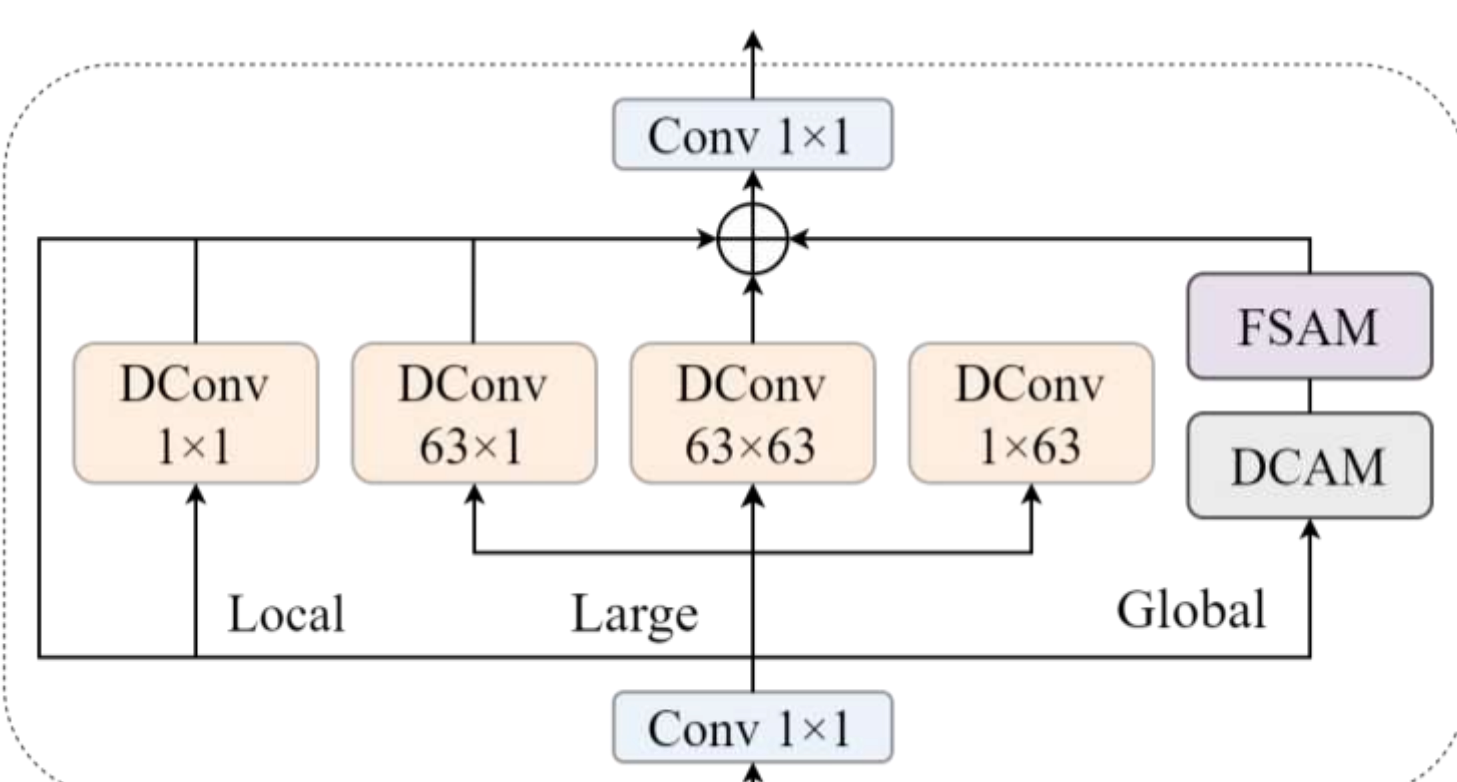


Figure 3. OKM Bottleneck

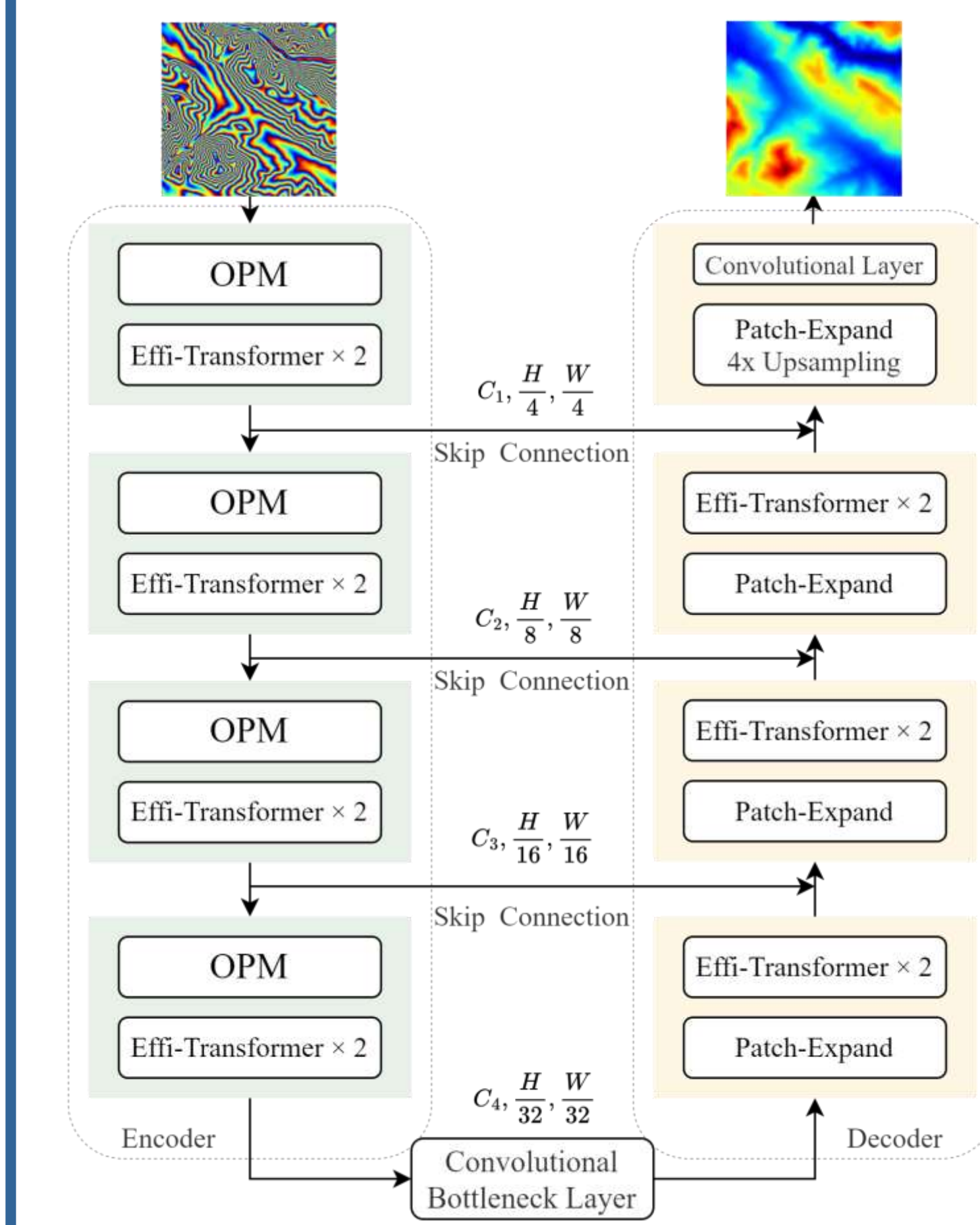


Figure 4. Structure of U-Segformer

Methods

The interferometric phase patch processed by the filtering module will subsequently undergo phase unwrapping using an improved SegFormer-based network.

SegFormer is a lightweight transformer model designed for processing natural images. However, its lightweight multi-layer perceptron decoder is too simplistic to capture the detailed texture features of InSAR phase patches, failing to meet the precision requirements for phase unwrapping tasks.

Inspired by U-Net, we add skip connection structures to the decoder module; introduced the Effi-Transformer module for feature extraction; and implemented a Patch Expanding Module (PEM) for upsampling, modifying SegFormer into a U-shaped architecture to achieve better results.

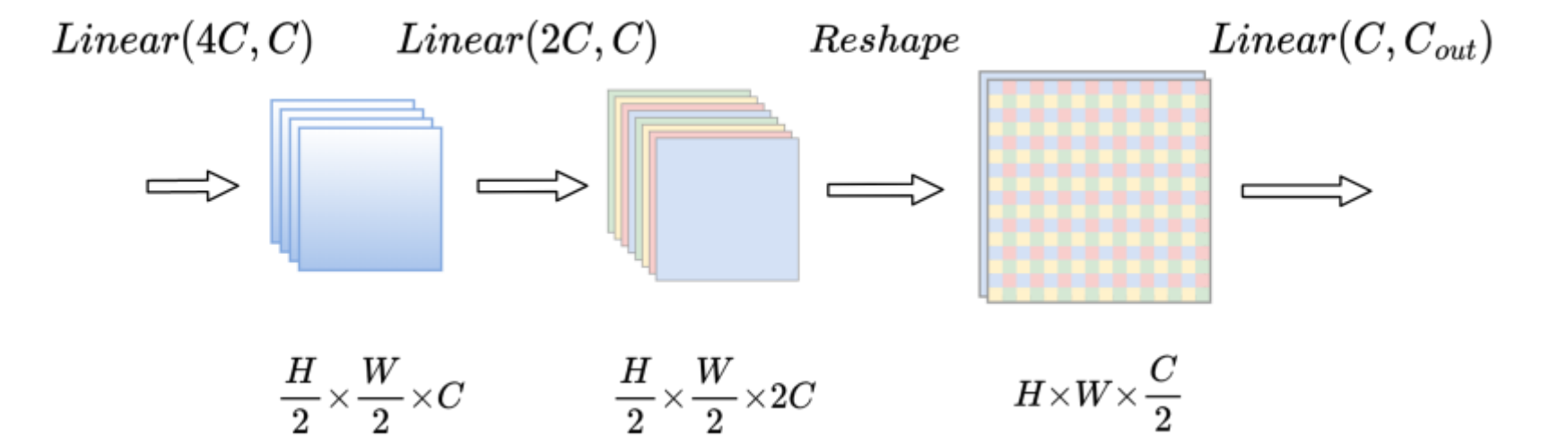


Figure 5. Patch Expanding Module

Results & Discussion

Apply U-SegFormer unwrapping method on simulated data, comparing it with Minimum Cost Flow(MCF), Least Squares(LS), and Phase Unwrapping via Max Flows(PUMA).

Method	RMSE	SSIM
LS	7.6240	0.8838
PUMA	3.8067	0.9314
MCF	3.1506	0.9308
U-SegFormer	2.5982	0.8851

As shown in the table and Figure 6, U-SegFormer exhibits good unwrapping accuracy and RMSE performance for high noise and discontinuity datasets, though it slightly lags behind MCF and PUMA in structural similarity.

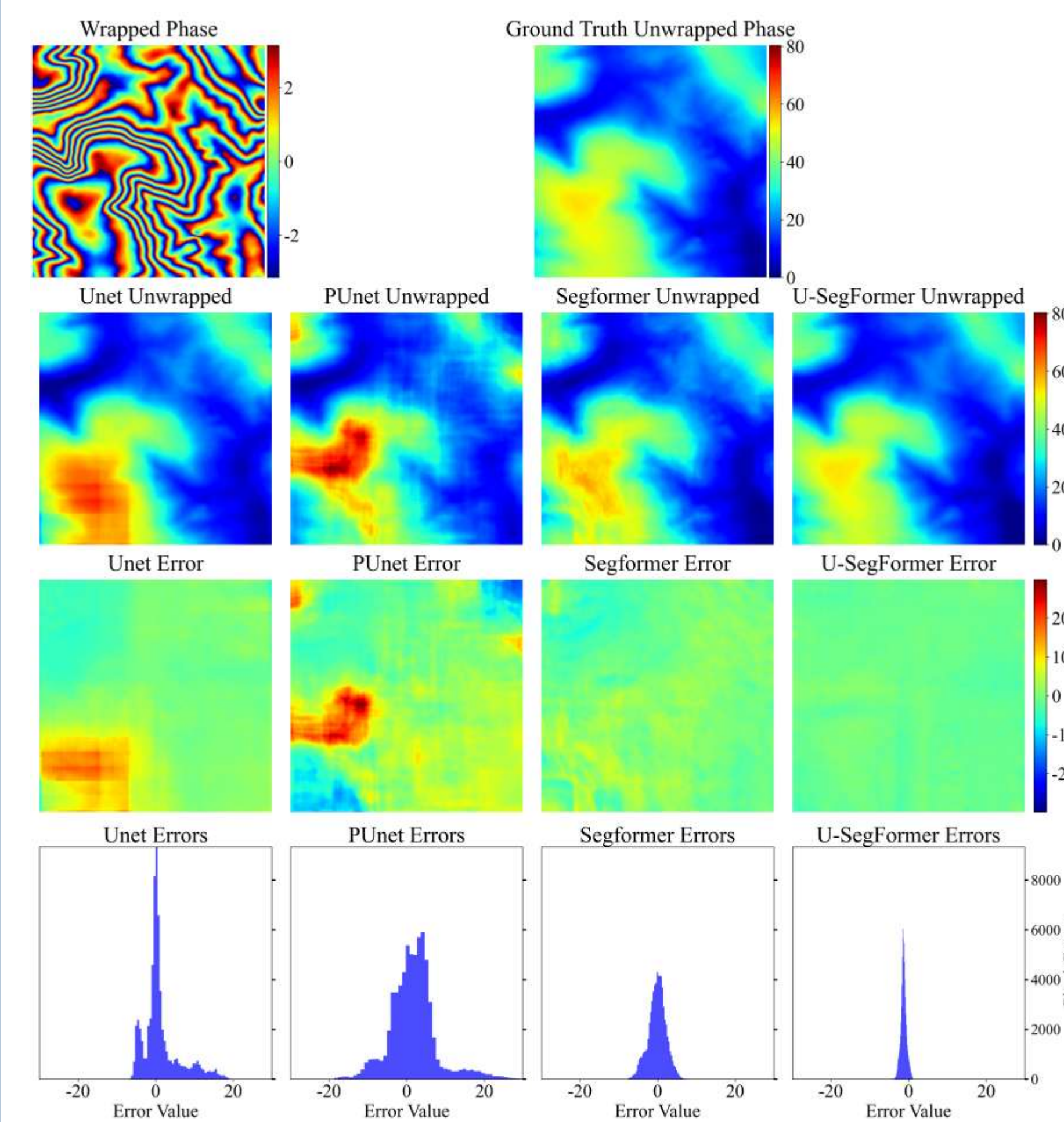


Figure 7: Comparison of U-SegFormer with deep learning Methods

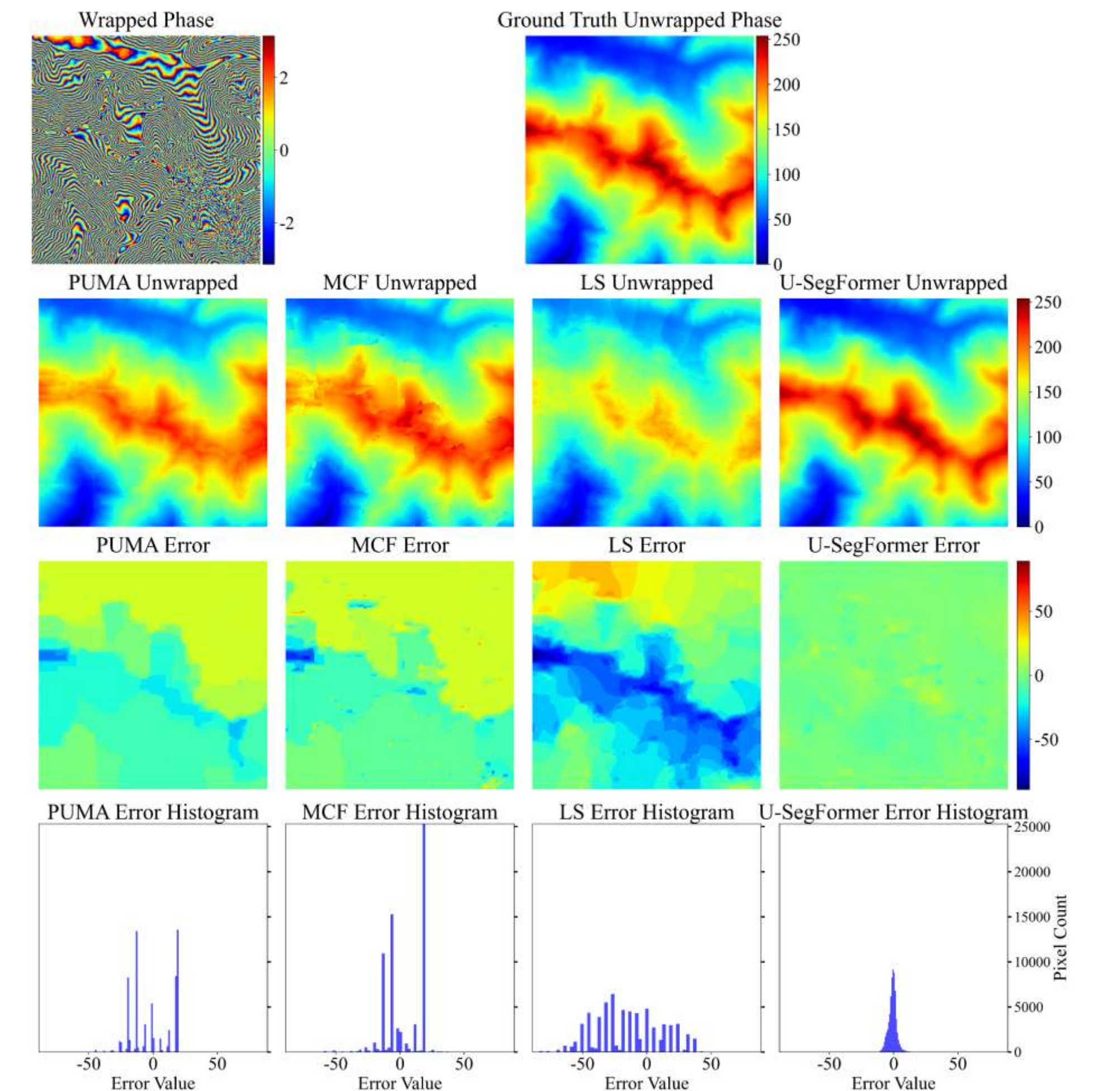


Figure 6: Comparison of U-SegFormer with Traditional Methods

In another set of simulated data, we compared the phase unwrapping results of U-SegFormer with those obtained from the classical image segmentation network U-Net, the one-step phase unwrapping network PUnet[3], and the unaltered original SegFormer. The results are presented in Figure 7 and the table below.

Method	RMSE	SSIM
PUnet	14.3741	0.5276
Unet	9.4343	0.7790
SegFormer	5.1609	0.7001
U-SegFormer	4.4575	0.7905

The results indicate that, compared to several other deep learning methods, U-SegFormer achieved the best performance on the simulated test dataset.

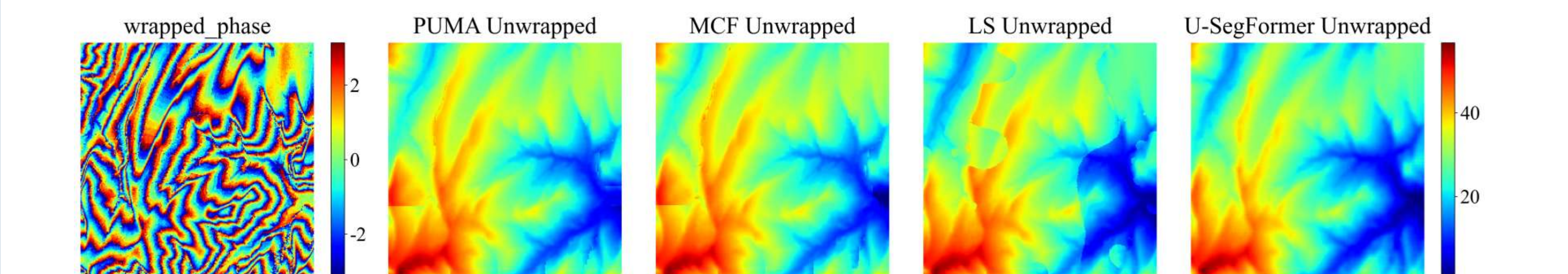


Figure 8: Comparison of Unwrapped Results for Real Data

Figure 8 shows a TanDEM-X single-pass data from Article[4], illustrating that for the wrapped phase in real scenarios, U-SegFormer can produce a more continuous unwrapped phase compared to traditional methods.

Conclusion

Testing results on simulated and real datasets show that this SegFormer-based one-step unwrapping network achieves relatively continuous unwrapped results, but its structural similarity is slightly lower than that of traditional methods, indicating a potential compromise in fringe consistency. Future experiments will combine the network with traditional methods to enhance accuracy by predicting pixel discontinuities.

References

- [1] Y. Cui, W. Ren, and A. Knoll, "Omni-Kernel Network for Image Restoration," in Proceedings of the AAAI Conference on Artificial Intelligence, Mar. 2024, pp. 1426–1434.
- [2] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers." arXiv, Oct. 28, 2021.
- [3] Z. Wu, T. Wang, Y. Wang, R. Wang, and D. Ge, "Deep Learning for the Detection and Phase Unwrapping of Mining-Induced Deformation in Large-Scale Interferograms," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–18, 2022.
- [4] L. Zhou and H. Yu, "InSAR-DLP: A benchmark dataset for deep learning-based synthetic aperture radar interferometry phase unwrapping [Software and Data Sets]," IEEE Geoscience and Remote Sensing Magazine, vol. 12, no. 2, pp. 118–124, Jun. 2024.

Acknowledgement : This work was conducted as part of the Dragon 6 cooperation programme between the European Space Agency (ESA) and the Ministry of Science and Technology (MOST) of China. The authors acknowledge the support and coordination provided by the Dragon Programme Office