引文格式:王杰,柳稼航,凌信鹏,等.基于深度学习的InSAR影像局部与非局部联合相位滤波[J].武汉大学学报(信息科学版),2026,51(2):367-379.DOI:10.13203/j.whugis20240052

Citation: WANG Jie, LIU Jiahang, LING Xinpeng, et al. Deep Learning-Based Joint Local and Non-local Phase Filtering for InSAR Images[J]. Geomatics and Information Science of Wuhan University, 2026, 51(2):367-379.DOI:10.13203/j.whugis20240052

基于深度学习的InSAR影像局部与非局部联合相位滤波

王 杰[1,2] 柳稼航[1,2] 凌信鹏[1,2] 段则先[1,2]

1 南京航空航天大学航天学院,江苏 南京,210016

2 空间光电探测与感知工业和信息化部重点实验室,江苏 南京,210016

摘 要:相位滤波是合成孔径雷达干涉测量(interferometric synthetic aperture radar, InSAR)数据处理的关键技术之一,干涉图质量显著影响后续处理的精度。传统方法滤波效果差且效率低下,深度学习方法潜力大但目前精度受限。为有效提升滤波效果,结合卷积神经网络(convolutional neural network, CNN)与Transformer网络的优势,提出了一种局部与非局部联合的InSAR影像相位滤波方法。该方法一方面利用Transformer全局特征提取能力强的特点,构建了CNN与Transformer相结合的相位滤波网络(phase filtering network combining CNN and Transformer, PFCT),能够同步进行局部与非局部滤波,从而克服现有方法因忽略干涉条纹非局部特征而导致的精度限制;另一方面,设计了一种新的复合损失函数,让网络迭代优化过程中同时保持去噪与条纹结构两方面的性能。在仿真实验中,PFCT的均方误差比次优算法降低了15.5%,结构相似性指标提高了5.3%;在实测实验中,残点去除性能比次优算法提高了1.8%。结果表明,PFCT的滤波性能优于目前的主要算法,能获得更好的效果。

关键词:InSAR;相位滤波;深度学习;CNN;Transformer;损失函数

中图分类号:P237 **文献标识码**:A **收稿日期**:2024-04-25

DOI:10.13203/j.whugis20240052 **文章编号**:1671-8860(2026)02-0367-13

Deep Learning-Based Joint Local and Non-local Phase Filtering for InSAR Images

WANG Jie[1,2] *LIU Jiahang*[1,2] *LING Xinpeng*[1,2] *DUAN Zexian*[1,2]

1 College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

2 Key Laboratory of Space Photoelectric Detection and Sensing of Ministry of Industry and Information Technology, Nanjing 210016, China

Abstract: **Objectives**: Interferometric synthetic aperture radar (InSAR) acquires ground elevation information and surface deformation information by interferometrically processing multiple synthetic aperture radar (SAR) complex images with certain temporal and spatial correlations. It has been widely applied in high-precision 3D mapping, as well as in the monitoring and early warning of geological disasters such as earthquake deformation, volcanic movement, and surface subsidence. Phase filtering is one of the key technologies in InSAR data processing, and the quality of interferograms significantly affects the accuracy of subsequent processing steps such as phase unwrapping and digital elevation model generation, but filtering effectiveness is often compromised by complex noise in the interferometric phase. The traditional phase filtering methods with manual parameter adjustment have not only low filtering accuracy but also low computational efficiency. Deep learning-based phase filtering methods have demonstrated great potential in automating this process and improving performance, but current convolutional neural networks (CNN) based methods

基金项目:江苏省双创人才计划(JSSCRC2021501)。

第一作者:王杰,硕士,主要从事InSAR相位滤波与相位解缠研究。wj104777@nuaa.edu.cn

通信作者:柳稼航,博士,教授。jhliu@nuaa.edu.cn

overlook the non-local features causing limiting accuracy. Transformer-based networks, on the other hand, have strong global modeling capabilities and can extract non-local features from interference fringes, but they frequently struggle with capturing fine-grained local details and spatial structures of interference fringes, which are equally crucial for precise phase restoration. **Methods:** Deep learning is data-driven, yet there is currently no publicly available InSAR interferometric phase dataset. The proposed method generates a clean phase map using shuttle radar topography mission 30 m digital elevation model, and adds simulated noise to produce an interferometric phase map, thereby creating a dataset for training the network. Besides, to effectively improve the filtering effect, we propose a joint local and non-local phase filtering network combing CNN and Transformer (PFCT) for InSAR images. First, based on the strong global feature extraction capability of the robust Transformer network, the PFCT network structure is proposed for phase filtering. The network performs both local and nonlocal filtering at the same time, avoiding the accuracy limitation problem caused by the neglect of non-local features of interference fringes in existing CNN-based methods while also addressing the local detail deficiency typical of Transformer approaches. The architecture integrates CNN layers for capturing high-frequency details and spatial patterns with Transformer modules for modeling long-range dependencies across the fringe patterns. Then, a new complex loss function which combines point phase loss and phase gradient loss is proposed to guide the network to maintain the integrity of interference fringes while improving filtering performance. **Results:** In the simulation data experiment, the proposed method achieves a mean square error index that is 15.5% lower than the suboptimal algorithm, and the structural similarity (SSIM) index is 5.3% higher, indicating superior noise removal and structural preservation ability. These improvements are consistently observed across varying noise levels, confirming the model's robustness. The enhanced SSIM particularly underscores its effectiveness in maintaining critical fringe structures essential for accurate unwrapping. In the true data experiment, the residue removal index is 1.8% higher than the suboptimal method, demonstrating enhanced reliability in practical applications. Visual assessment of filtered interferograms also shows a marked reduction in speckle noise and better preservation of deformation fringes in areas with high phase gradients. The experiments show that the PFCT network model only slightly reduces computational efficiency, but achieves better filtering results than other methods, effectively maintaining the stripe structure while maintaining the filtering effect. **Conclusions:** The proposed method has a great filtering effect and powerful generalization ability on InSAR phase filtering across different terrains and noise conditions. By successfully hybridizing CNN and Transformer strengths, it provides a robust and efficient option for high-quality interferogram generation in modern InSAR processing workflows. Future work may focus on further optimizing the network for edge-device deployment and exploring its adaptation to multi-temporal InSAR time series analysis for dynamic monitoring applications.

Key words: InSAR; phase filtering; deep learning; CNN; Transformer; loss function

合成孔径雷达(synthetic aperture radar, SAR)遥感是一种主动式遥感,不受云雨和昼夜的影响,具有全天时、全天候工作的优点和一定的地表穿透力[1],在资源勘探、城市规划、地形测绘等领域发挥着重要作用[2]。SAR干涉测量(interferometric SAR, InSAR)是在此基础上发展而来的一种遥感测绘技术[3],通过对多幅具有一定时间以及空间相关性的SAR复数影像进行干涉处理来获取地面高程信息以及地表形变信息[4],已被广泛应用于高精度三维制图以及地震形变、火山运动、地表沉降等地质灾害的监测和预警[5-7]。然而,系统热噪声、时间去相干、空间去相干等因素对相位造成了较大干扰[8],导致相位解缠精度下降。相位滤波是抑制这些噪声的主要手段,在InSAR数据处理中具有重要作用。

为了有效地去除这些噪声,研究者们提出了多种算法,这些算法大致分为空间滤波和变换域滤波两大类,空间滤波又可以细分为局部滤波和非局部滤波两小类。局部滤波算法的基本思想是采用滑动窗口遍历干涉相位图里的所有像素,并对窗口内的像素进行平滑处理[9]。局部滤波继承了传统图像滤波的思想,并结合InSAR干涉图

局部特征的特性进行了改进,代表算法为圆周期均值滤波[10]、Lee滤波[11]等。圆周期均值滤波的噪点去除能力较好,但滤波强度不足。Lee滤波针对滤波窗口内不同的干涉条纹走向采用不同的滤波模板,并根据窗口内的噪声强度自适应改变滤波强度,在低相干区域有效滤除噪声的同时避免了高相干区域边缘条纹细节损失的问题,但受局部窗口中残差点的影响,容易产生条纹伪影等失真现象。

非局部滤波算法以文献[12]提出的非局部均值算法为核心框架,不再局限于在小窗口内对像素进行平滑处理,而是将图像分成多个图像块,计算不同图像块间的相似性并进行加权平均,以此实现滤波。文献[13]提出了非局部InSAR(nonlocal InSAR, NL-InSAR),首次运用了非局部的思想,通过各图像块的统计特性来估计各图像块的相似性,再根据计算的相似性为各图像块赋予加权系数,各图像块加权平均后即可得到最终滤波结果。文献[14]将三维块匹配(block matching 3D, BM3D)算法滤波[15]引入InSAR领域,对InSAR干涉相位进行图像块堆叠协同滤波,提出了InSAR-BM3D滤波,实现了在去噪强度和细节保留之间更优异的平衡。非局部滤波算法的滤波精度较好,但对非局部相似区域进行的搜索和处理操作需要大量的计算时间,降低了计算效率。

变换域滤波将干涉相位从空间域转换到频域或小波域[16],进行滤波处理后再转换回空间域,代表算法为Goldstein滤波,它用傅里叶变换将干涉相位转换到频域,对干涉相位的频谱进行滤波。由于正常相位在相位频谱中属于低频成分,噪声相位属于高频成分,文献[17]设计了一个低通滤波器,并设置滤波参数对整个频谱进行平滑处理。由于滤波参数固定,该算法可能出现低噪声区域滤波过度而高噪声区域滤波不足的现象。

近年来,随着深度学习在图像处理领域的快速发展[18],研究者们开始将深度学习引入到InSAR数据处理中,提出了多种基于卷积神经网络(convolutional neural network, CNN)的相位滤波网络模型,并利用模拟出来的相位数据[19-21]对网络进行训练,实现了快速、高精度相位滤波。通过改进传统图像处理的经典滤波算法DnCNN(denoising CNN)[22],文献[23]提出了IPDnCNN(interferometric phase DnCNN),它将干涉相位变换至复数域,同时对相位实部和虚部进行滤波以保护条纹边缘,但在高噪声区域滤波效果较差。多尺度特征融合和深度CNN(multi-scale feature dynamic fusion with deep CNN, MSFF-DCNN)通过提取干涉相位实部和虚部的多尺度特征,并进行动态融合以进行相位复原[24]。MONet(multiobjective neural network)针对以往网络模型中损失函数过于单一的问题,首次提出了统计损失和空间损失联合的新型损失函数[25],但该函数训练时间过长且局部细节保持能力较差,总体滤波效果不够理想。

尽管基于深度学习的相位滤波显示出了良好的潜力,目前的研究仍处于起步阶段。大多数算法采用数据驱动的有监督训练策略,但目前尚无公开公认的相位滤波数据集,这对获取高质量的训练数据提出了很大挑战。为解决这一问题,一些研究者尝试利用完全模拟的相位数据进行训练,通过随机矩阵插值放大的方法[26]对地形进行数学模拟。由于模拟地形与真实的地形特征并不完全一致,难以保证网络学习到的特征是正确的。为此,另一些研究者先使用准实测相位构建数据集,即用公开的数字高程模型(digital elevation model, DEM)和已知卫星参数推算的地形相位作为数据集中的真实值,再向其中添加高斯噪声,得到带噪声的相位图。这种高斯噪声分布与真实噪声分布相去甚远,且这些深度学习算法大多基于CNN,专注于局部特征提取,面对InSAR相位图中的复杂结构时,容易受到其有限的感受野限制,导致相位滤波效果不佳。相比之下,Transformer模型[27]以其较强的全局关系捕捉能力,在检测、分割、去噪等多项计算机视觉任务中取得了良好效果,但在相位滤波任务中容易忽视干涉影像的局部特征,导致条纹细节滤波效果较差[28]。

为探究更有效的滤波方法,本文提出了一种CNN与Transformer联合相位滤波的网络结构(phase filtering network combining CNN and Transformer, PFCT),该方法能同时进行局部滤波和非局部滤波,可有效改善相位滤波效果。为获得与真实数据保持较高一致性的训练样本,采用公开DEM和卫星参数构建数据集,通过采集真实干涉图的相干性图,并从中选取平均相干性在一定范围内的切片作为模拟干涉图的相干性图,来模拟干涉所需要的主辅单视复数(single-look complex, SLC)影像。本文设计了针对相位实部和虚部特性的特征提取模块(feature extraction block, FEB)和相位重建模块(phase recon-

struction block, PRB),建立空间-条纹结构复合损失函数来引导网络训练,使网络在提升滤波性能的同时保留原有的条纹结构,避免了单一空间损失函数带来的结构失真,并利用仿真数据与实测数据进行对比实验,结果验证了本文PFCT方法的有效性。

1 数据集构造

深度学习是由数据驱动的,但目前尚无公开的InSAR干涉相位数据集。事实上,SAR卫星数据及其干涉相位容易受到各种噪声的影响导致失相干,在没有清晰的理想相位作为参考的情况下难以直接构建数据集,因而需要先构建由理想相位和噪声相位组成的InSAR相位对数据集。

1.1 理想相位

InSAR干涉几何示意图如图1所示,S_1和S_2为两次观测时卫星天线所在的位置;P为地面观测目标;B为空间基线,即天线间的距离;h为目标P距海平面的垂直高度;H为天线S_1距海平面的垂直高度;R_1和R_2分别为天线S_1和S_2到目标P的距离;θ表示天线S_1观测方向的入射角;α表示B与水平方向的夹角。

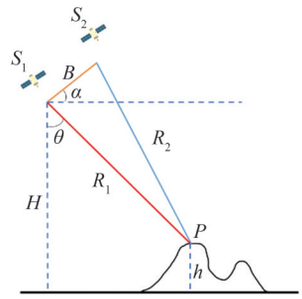


图1 InSAR干涉几何示意图

Fig. 1 InSAR Interferometric Geometry Diagram

本文选择了航天飞机雷达地形测量任务30 m分辨率的DEM来生成干净的干涉相位,传感器参数为:B=60 m,波长λ=0.056 m,θ=45°,α=45°,R_1=231 km,将地形高度h转换为干净的地形干涉相位φ_{clean}的计算公式为[29]:

$$\varphi_{\text{clean}}=\left(\frac{4\pi Bh\cos(\theta-\alpha)}{\lambda R_1\sin\theta}\right)\%(2\pi)-\pi \tag{1}$$

式中,%为求余符号。

为了使神经网络能够学习各种不同的地形特征,本文选择了具有起伏地形的DEM来制作数据集。选择图2(a)所示的四川与西藏之间的崎岖山区作为训练数据集的参考DEM,并选择图2(b)所示的埃特纳(Etna)火山周边地区作为测试数据集的参考DEM。

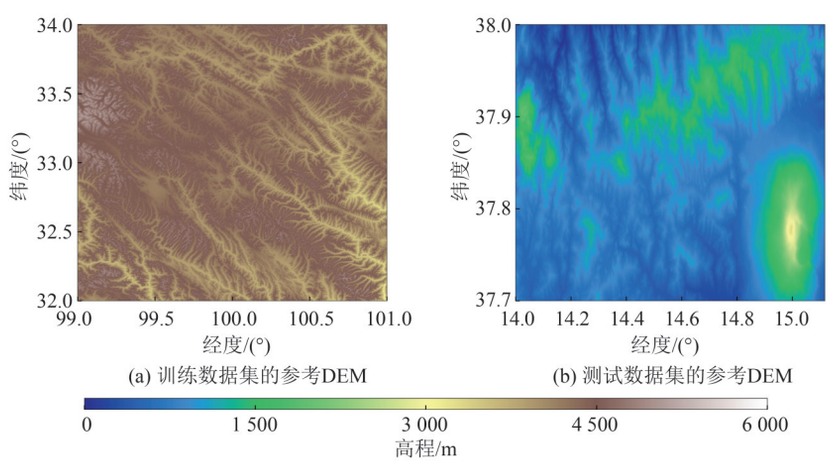


图2 数据集的参考DEM

Fig. 2 Reference DEM for Datasets

此外,为测试网络在不同条纹频率的滤波结果,通过设置45 m、60 m、90 m和150 m的空间基线B来改变相位图的条纹频率,结果如图3所示。将这些DEM数据切分成64×64的小块,并插值放大至256×256,最终得到了44 944个样本的训练数据集和2 688个样本的测试数据集。

1.2 噪声相位

借助DEM数据可以获得模拟相位的理想值,还需要模拟理想相位对应的噪声相位作为输入。为方便起见,本文直接模拟了带噪声的SLC影像对,噪声相位φ_{noisy}可以通过主SLC影像z_1与辅SLC影像z_2复共轭相乘获得,计算公式为:

$$\varphi_{\text{noisy}}=z_1\cdot z_2^* \tag{2}$$

本文采用的噪声模型是对文献[30]所提方法的一种改进,简化了其相干性图的构建方式。SLC影像对$[z_1\ z_2]^{\mathrm{T}}$可以由两个标准高斯变量

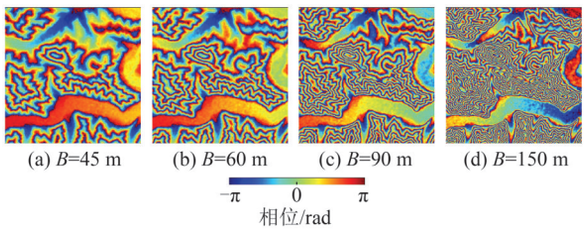


图3　不同条纹频率下的干涉相位图

Fig. 3　Interferometric Phase Maps Under Different Fringe Frequencies

u_1、u_2表示[15]，计算公式为：

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = T \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{3}$$

式中，T是$[z_1\ z_2]^{\mathrm{T}}$的协方差矩阵的Cholesky分解，计算公式为：

$$T = \begin{bmatrix} A_1 & 0 \\ A_2\rho \mathrm{e}^{-\mathrm{j}\varphi_{\mathrm{clean}}} & A_2\sqrt{1-\rho^2} \end{bmatrix} \tag{4}$$

式中，A_1和A_2分别为主辅SLC影像的振幅；ρ为主辅SLC影像的相干性。

可以看出，模拟的$[z_1\ z_2]^{\mathrm{T}}$只取决于真实相位φ_{clean}、相干性ρ和振幅A_1与A_2，其中A_1和A_2在主辅SLC影像中可以近似相等，本文用A来表示，则模拟的SLC图像对$[z_1\ z_2]^{\mathrm{T}}$的计算公式为：

$$\begin{cases} z_1 = Au_1 \\ z_2 = A\rho \mathrm{e}^{-\mathrm{j}\varphi_{\mathrm{clean}}} u_1 + A\sqrt{1-\rho^2}\, u_2 \end{cases} \tag{5}$$

相干性和振幅通常彼此具有强相关性，但与干涉相位却呈现弱相关或不相关的特性[30]。首先，从先进陆地观测卫星（advanced land observing satellite，ALOS）相控阵型L波段SAR的干涉图中随机提取4个子相干性图，并要求这4个子图的平均相干性分别分布在[0.4，0.5）、[0.5，0.6）、[0.6，0.7）、[0.7，0.8]区间内，让噪声强度分布更加均衡。对应的4个子相干性图如图4所示，平均相干性$\bar{\rho}$分别是0.44、0.54、0.62、0.76。

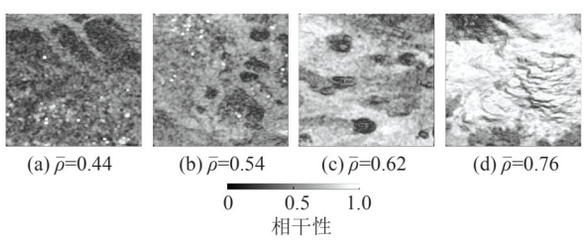


图4　相干系数图

Fig. 4　Coherence Maps

然后，将4幅相干性图乘以255来简单模拟振幅并且代入式(5)，以获得模拟的SLC图像对，噪声相位数据集的模拟流程如图5所示。最终，整个训练数据集包括44 944组理想相位和含噪相位图像对，尺寸大小都为256×256，测试数据集则包括了2 688组图像对。

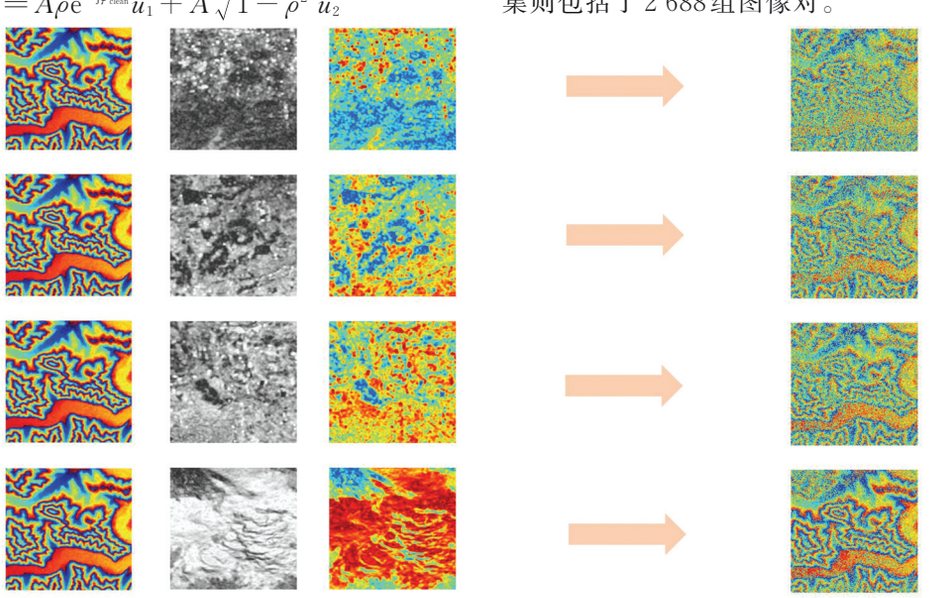


图5　噪声相位数据集的模拟流程

Fig. 5　Simulation Pipeline of Noisy Phase Datasets

2　PFCT相位滤波网络模型

U-Net[31]网络结构采用经典的编码器-解码器网络结构，将编码器部分的下采样与解码器部分的上采样跳跃连接，整体网络呈U型，密集融合了浅层与深层特征，在各图像任务中表现突出。本

文采用与U-Net相似的编码器-解码器结构,重新设计网络架构和关键模块以适应InSAR相位滤波任务,设计结果如图6所示,φ_{real}和φ_{imag}分别为噪声相位φ_{noisy}的实部和虚部,φ'_{real}和φ'_{imag}分别为滤波后相位的实部和虚部,CTB(convolution-Transformer block)为卷积-Transformer模块。

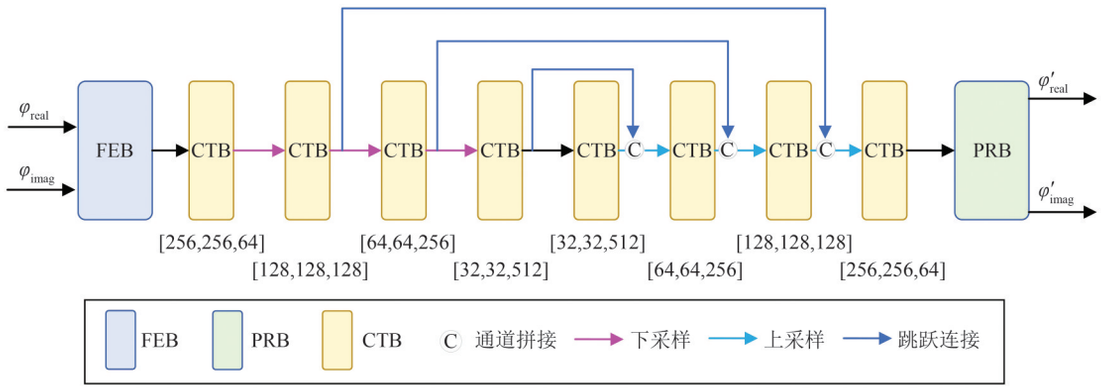


图6 PFCT网络结构

Fig. 6 PFCT Net Structure

为了在滤波时不破坏条纹边缘处的相位突变,将实数域的相位变换到复数域,以减少边缘跳变被误判为噪声的可能性,计算公式为:

$$\begin{cases}\varphi_{real}=\cos\varphi_{noisy}\\ \varphi_{imag}=\sin\varphi_{noisy}\end{cases} \tag{6}$$

将φ_{real}和φ_{imag}输入网络,最终输出φ'_{real}和φ'_{imag},滤波后的相位φ_{filted}计算公式为:

$$\varphi_{filted}=\arctan\frac{\varphi'_{imag}}{\varphi'_{real}} \tag{7}$$

整个滤波流程如图7所示。

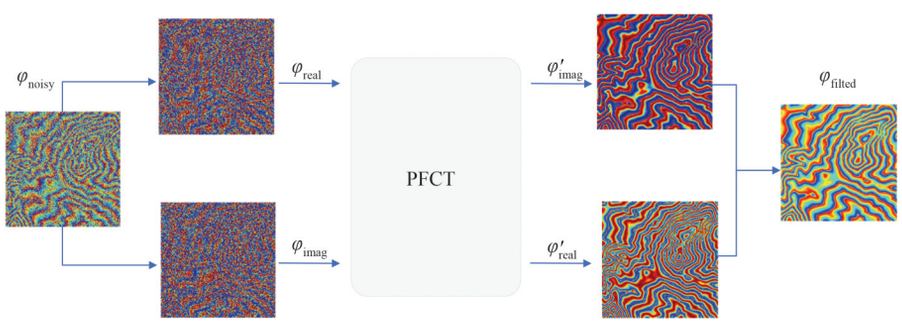


图7 滤波流程

Fig. 7 Flowchart of Filtering

2.1 FEB

相位在进入网络之前被拆分成了实部和虚部,二者之间存在着一定的关联性。FEB结构如图8所示,提取完二者浅层特征后,需要进行特征拼接来防止相位实部和相位虚部在滤波过程中解耦。相位实部和相位虚部在送入网络时,均需要进行两次特征提取,即先使用3×3卷积和双曲正切函数(tanh)进行初次特征提取,再使用3×3卷积进行第二次特征提取。为了让两个分支提取的特征保持一致,使用批归一化(BatchNorm)层对特征进行正则化,之后利用泄漏修正线性函数(LeakyReLU)进行激活,将实部和虚部激活后的特征进行拼接,送入下一个模块。

2.2 CTB

CNN通常采用固定大小的卷积核,专注于局部特征提取,但面对InSAR相位图中的复杂结构

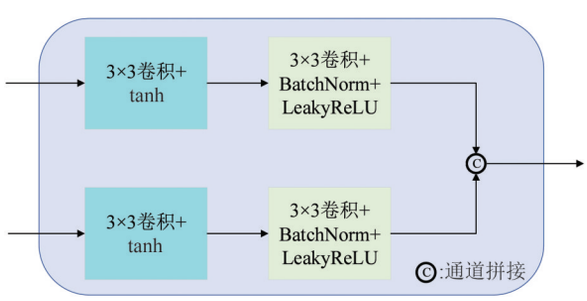


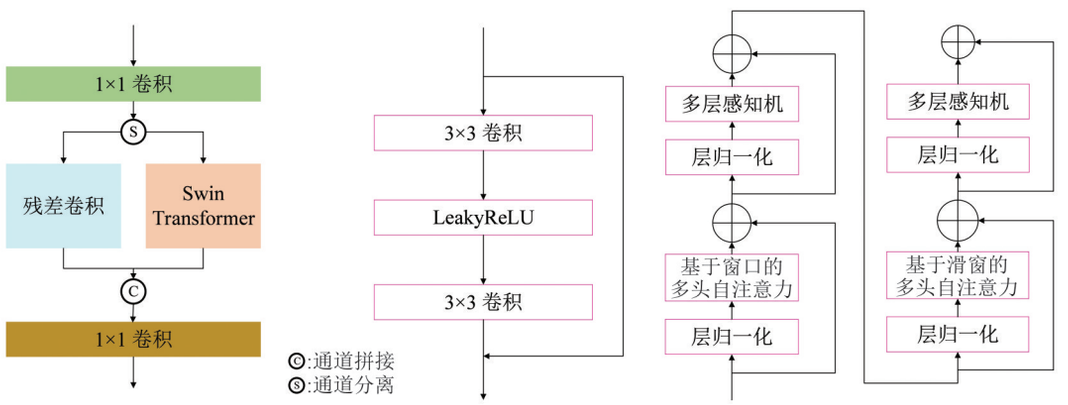
图8 FEB结构

Fig. 8 Structure of FEB

时,容易受感受野的限制,相位滤波效果不佳。相比之下,Transformer模型因其强大的自注意力机制而具备较强的全局关系建模能力,可计算任意两点的相关程度,能更好地捕捉数据间的相关性。它

通过对不同区域进行加权处理实现了非局部处理,但重建InSAR相位图的微观细节能力相对薄弱,容易忽略图像中的局部特征。Swin(shifted windows) Transformer在Transformer的基础上提出了窗口自注意力机制和滑动窗口自注意力机制,二者交替使用既减少了计算成本,又实现了信息的跨窗口交互[32]。

CTB是PFCT网络的骨干模块,整体结构如图9(a)所示,特征图先进行1×1的卷积:如果该CTB处在编码器阶段,则保持输入的特征图通道维度不变;如果处在解码器阶段,则将输入特征图的通道数减半。卷积后从通道维度对整个特征图进行平均分离操作,分离成通道数相同的两个子特征图,一个子特征图通过残差卷积模块进行局部滤波,另一个通过Swin Transformer模块进行非局部滤波,两个子模块的结构分别如图9(b)和图9(c)所示。滤波后的两个特征图在通道维度进行拼接,再通过1×1的卷积,就得到了滤波后的特征图。



(a) CTB总体结构　(b) 残差卷积子模块结构　(c) Swin Transformer子模块结构

图9　CTB结构

Fig. 9　Structure of CTB

2.3　PRB

过滤后的噪声相位实部和虚部特征图是拼接在一起的整体,需要先进行特征解耦以防止实部和虚部特征在重建相位时互相干扰,PRB整体结构如图10所示。PRB先以最后一个CTB输出的特征图作为输入,然后分为实部分支和虚部分支,每个分支都对输入的特征图进行独立处理。先通过一个3×3的卷积层保持特征图的通道数量不变,并使用tanh函数进行激活,再用1×1的卷积将通道维数降至1,这样两个分支就可以从特征图中分别重建得到滤波后的相位实部和相位虚部。

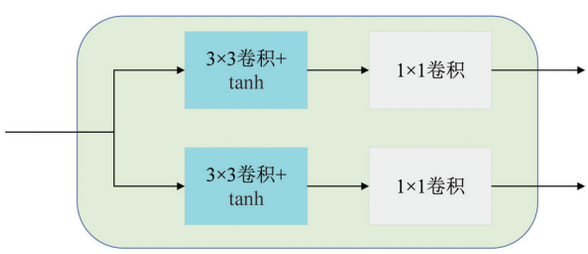


图10　PRB结构

Fig. 10　Structure of PRB

2.4　损失函数

设置一个合理的损失函数对于加速神经网络的收敛速度和优化神经网络性能至关重要。在InSAR滤波任务中,需要精确滤除每个点的相位噪声,同时保留完整的空间条纹结构。使用单一损失函数训练,往往难以达到这一目标。为此,本文设计一个复合损失函数L_{total},计算公式为:

$$L_{\text{total}}=L_{\text{point}}+L_{\text{stripe}} \tag{8}$$

式中,L_{point}为每个点的平均相位滤波误差损失;L_{stripe}为反映条纹结构的误差损失。L_{point}的计算公式为:

$$L_{\text{point}}=\frac{1}{n\times n}\sum_{j=1}^{n}\sum_{i=1}^{n}(\varphi_{\text{gt}}-\varphi_{\text{filted}})^2+\frac{1}{n\times n}\sum_{j=1}^{n}\sum_{i=1}^{n}\left|\varphi_{\text{gt_real}}-\varphi'_{\text{real}}\right|+\frac{1}{n\times n}\sum_{j=1}^{n}\sum_{i=1}^{n}\left|\varphi_{\text{gt_imag}}-\varphi'_{\text{imag}}\right| \tag{9}$$

式中,n为图像的尺寸;(i,j)为相位点的坐标;φ_{gt}为真实相位;$\varphi_{\text{gt_real}}$、$\varphi_{\text{gt_imag}}$分别为真实相位的实部和虚部。可以看出,L_{point}由φ_{gt}和φ_{filted}的L_2损失、$\varphi_{\text{gt_real}}$与φ'_{real}的L_1损失以及$\varphi_{\text{gt_imag}}$与φ'_{imag}的L_1损失组成。由于本文最终预测的相位值是由网络预测相位的实部和虚部再反正切求得,如果只采用相位间的L_2损失作为损失函数,并不能保证网络预测的实部和虚部的精确性;同理,如果采用实部和虚部的L_1损失作为损失函数,也不能保证反正切求得的相位值的准确性。而采用L_{point}作为

损失函数,不仅可以保证预测的实部和虚部的精确性,也能保证预测的相位值的准确性。

滤波方法对条纹结构的保持效果难以直接量化,但由于条纹边缘存在剧烈的相位跳变,该处的梯度值会远大于非边缘区域。因此,只要确保滤波后相位图中这些位置的梯度特征,即可认为条纹结构得到了保持。基于这个思想,本文计算了真实相位与滤波后相位的水平梯度差和垂直梯度差,并将两者的平均绝对值之和L_{stripe}作为条纹结构损失,其计算公式为:

$$L_{\text{stripe}}=\frac{1}{n\times(n-1)}\sum_{j=1}^{n-1}\sum_{i=1}^{n}\left|\nabla_v(\varphi_{\text{gt}}(i,j))-\nabla_v(\varphi_{\text{filted}}(i,j))\right|+\frac{1}{n\times(n-1)}\sum_{j=1}^{n}\sum_{i=1}^{n-1}\left|\nabla_h(\varphi_{\text{gt}}(i,j))-\nabla_h(\varphi_{\text{filted}}(i,j))\right| \tag{10}$$

式中,∇_v和∇_h为该点的垂直梯度和水平梯度,计算公式为:

$$\begin{cases}\nabla_v(\varphi(i,j))=\varphi(i,j)-\varphi(i,j+1)\\ \nabla_h(\varphi(i,j))=\varphi(i,j)-\varphi(i+1,j)\end{cases} \tag{11}$$

3 相位滤波实验与分析

3.1 实验细节

本文采用小批量训练方法,批量大小为12,使用Adam优化器,动量衰减参数$\beta_1=0.9$,$\beta_2=0.99$。训练持续了120个周期,初始学习率$\eta=0.000\,2$,并且每20个周期减半一次。本文提出的PFCT在一台CPU为Intel Core I7-13700k、GPU为GeForce RTX 3090ti的服务器上进行训练,并在CPU为Intel Core I5-9500、GPU为GeForce GTX 1660ti的主机上进行测试,两台主机开发环境均为Python 3.10.4,采用的深度学习框架均为PyTorch 1.12。

3.2 评价指标

在模拟的数据集中,准实测相位有着对应的真实相位作为参考,两者的均方误差(mean square error, MSE)和结构相似性(structural similarity, SSIM)作为客观衡量滤波效果的指标。MSE的计算公式为:

$$R_{\text{MSE}}=\frac{1}{N}\sum_{i=1}^{N}(\varphi_{\text{filted}}-\varphi_{\text{gt}})^2 \tag{12}$$

式中,N为相位图中的像素点数。R_{MSE}越趋于0,整体滤波效果越好。SSIM的计算公式为:

$$R_{\text{SSIM}}=\frac{\left(2\mu_{\varphi_{\text{gt}}}\mu_{\varphi_{\text{filted}}}+(k_1L)^2\right)\left(2\sigma_{\varphi_{\text{gt}}\varphi_{\text{filted}}}+(k_2L)^2\right)}{\left(\mu_{\varphi_{\text{gt}}}^2+\mu_{\varphi_{\text{filted}}}^2+(k_1L)^2\right)\left(\sigma_{\varphi_{\text{gt}}}^2+\sigma_{\varphi_{\text{filted}}}^2+(k_2L)^2\right)} \tag{13}$$

式中,k_1L和k_2L为用来维持稳定的常数,$k_1=0.01$,$k_2=0.03$,$L=2\pi$;$\mu_{\varphi_{\text{gt}}}$和$\mu_{\varphi_{\text{filted}}}$分别为真实相位和滤波后相位的均值;$\sigma_{\varphi_{\text{gt}}\varphi_{\text{filted}}}$为真实相位和滤波后相位的协方差;$\sigma_{\varphi_{\text{gt}}}$和$\sigma_{\varphi_{\text{filted}}}$分别为真实相位和滤波后相位的标准差。$R_{\text{SSIM}}$越接近于1,代表滤波结果与真实相位的结构越相似。

这些客观指标虽然可以衡量网络性能的优异,但在实际应用中,测量区域的实时准确DEM是不存在的,测得的干涉相位也就不存在参考,无法使用MSE和SSIM来评价网络性能,所以使用无参考指标Q[33]来评价网络对实测相位的滤波效果。Q可以通过对干涉相位梯度矩阵$\boldsymbol{G}$进行奇异值分解求得,$\boldsymbol{G}$的计算公式为:

$$\boldsymbol{G}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{U}\begin{bmatrix}s_1&0\\0&s_2\end{bmatrix}\begin{bmatrix}v_1&v_2\end{bmatrix}^{\mathrm{T}} \tag{14}$$

式中,$\boldsymbol{U}$和$\boldsymbol{V}$均为单位正交阵;$\boldsymbol{S}$为奇异值矩阵;s_1和s_2为梯度矩阵$\boldsymbol{G}$的奇异值,$s_1\geqslant s_2\geqslant 0$;与之对应的$v_1$和$v_2$分别为局部梯度场的主导方向和边缘方向[33]。$Q$的计算公式为:

$$Q=s_1\frac{s_1-s_2}{s_1+s_2} \tag{15}$$

Q越大,意味着干涉影像的视觉观感越清晰,越接近理想的无噪声状态[33]。此外,残差点的留存数(number of residues, NOR)也可以有效衡量网络的滤波效果。

3.3 实验结果与对比

本文将PFCT与一些传统相位滤波算法,如Lee滤波、Goldstein滤波、圆周期均值滤波等局部滤波和NL-InSAR、InSAR-BM3D等非局部滤波,以及近年来一些典型的基于深度学习的算法进行对比实验,比较这些算法的滤波效果。

3.3.1 仿真数据实验

PFCT与传统相位滤波算法结果对比如图11所示。由图11可知,Lee滤波、Goldstein滤波和圆周期均值滤波都有着大量的噪点未能去除,滤波效果相对较差;NL-InSAR滤波则产生了大面积的条纹结构失真;InSAR-BM3D滤波效果较其他传统算法好,但在一些局部区域出现了条纹断裂的问题;PFCT在保持了较好滤波效果的同时,条纹结构保持能力也优于其他传统算法。

不同条纹密度下不同深度学习算法的滤波

结果如图 12 和图 13 所示。可以看出，本文提出的 PFCT 针对不同条纹密度的相位干涉图都可以取得很好的滤波效果，滤波结果和真实相位最为相似。MSFF-DCNN 和 Uformer（U-shaped Transformer）针对稀疏条纹有不错的滤波效果，仅有部分细节处理不如 PFCT，但对于稠密条纹的滤波效果不佳，在局部区域出现条纹断裂或者严重失真的现象。MONet 在条纹稀疏和稠密时都会产生条纹伪影，改变了局部条纹结构，但失真程度比 MSFF-DCNN 和 Uformer 略好。总体而言，PFCT 的滤波效果优于其他算法，在不同条纹密度下都具有良好的结构和细节保持能力。

不同滤波算法在整个仿真数据集上的测试结果见表 1，PFCT 的 MSE 比次优的 MSFF-DCNN 低了 15.5%，SSIM 则高了 5.3%，展现了 PFCT 的优异性能。

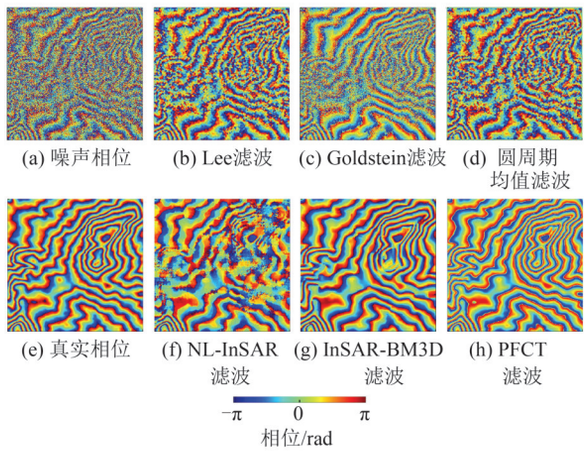


图 11　PFCT 与传统相位滤波算法结果对比

Fig. 11　Comparison of Results for Traditional Phase Filtering Algorithms and PFCT

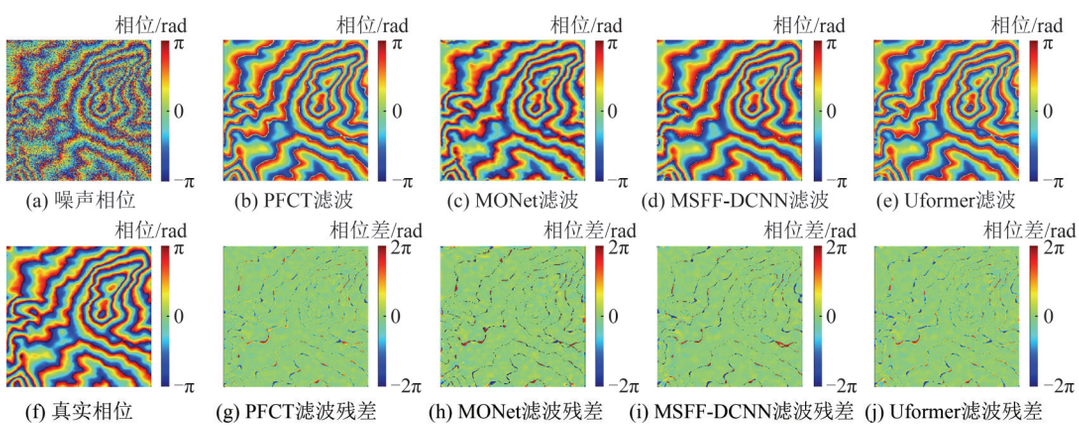


图 12　B=60 时深度学习算法滤波结果对比

Fig. 12　Comparison of Deep Learning Filtering Results When B=60

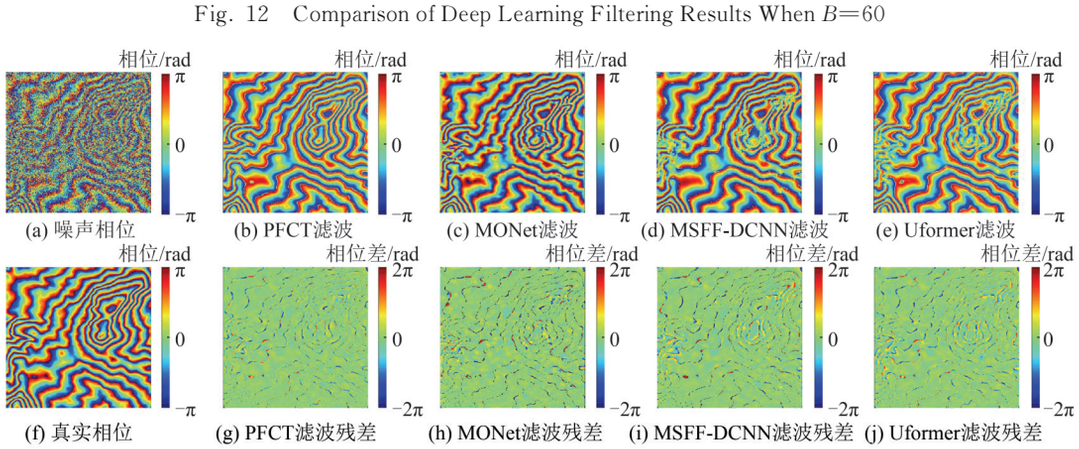


图 13　B=90 时深度学习算法滤波结果对比

Fig. 13　Comparison of Deep Learning Filtering Results When B=90

3.3.2　实测数据实验

为进一步验证 PFCT 对实际影像的滤波效果，选取由欧洲遥感卫星测出的 Etna 火山主辅 SLC 影像和由 ALOS 相控阵型 L 波段 SAR 测出的中国辽宁省局部地区及吉林省局部地区主辅 SLC 影像，截取出 768×768 大小的影像切片作为实测数据进行对比实验，这 3 组数据实测干涉相位与滤波结果分别如图 14、图 15 和图 16 所示。

由图 14 可见，传统算法中 Lee 滤波、Goldstein 滤波、圆周期均值滤波 3 类方法都有大量明显的噪点未能去除。NL-InSAR 滤波结果有大量条纹失真，滤波效果较差。InSAR-BM3D 滤波的

表1 仿真数据集的测试结果

Table 1 Test Results of Simulated Dataset

滤波方法	MSE	SSIM	NOR	耗时/s
含噪相位图	4.166	0.147	11 244	
Lee滤波[11]	1.059	0.416	13 872	0.484
Goldstein滤波[17]	2.159	0.382	16 181	0.053
圆周期均值滤波[10]	2.579	0.376	1 467	0.191
NL-InSAR[13]	2.411	0.439	3 585	17.084
InSAR-BM3D[14]	1.154	0.699	6 612	6.214
MONet[25]	1.186	0.679	975	0.018
MSFF-DCNN[24]	0.801	0.721	1 152	0.205
Uformer[28]	0.983	0.713	1 113	0.068
PFCT	0.677	0.759	855	0.091

效果是传统算法中最好的，去除了大部分噪点的同时也基本保持了原有的条纹结构，但条纹边缘不清晰，滤波结果不够平滑，深度学习算法中MONet滤波也存在类似现象。MSFF-DCNN滤波和Uformer滤波虽然残差点留存率较低，但滤波失败的区域面积同样较大，产生大面积的失真。而PFCT滤波仅在条纹结构模糊的区域滤波效果比InSAR-BM3D稍差，其余区域滤波效果相对较好，但InSAR-BM3D花费的时间是PFCT的70余倍，计算效率显著降低。

由图15可见，实测相位的条纹复杂程度远高于图14展示的Etna实测数据，滤波难度较Etna数据大。传统算法中仍然只有InSAR-BM3D滤波结果较好，其他算法的条纹结构保持效果和噪点平滑效果都差一些。深度学习算法中，还是只有PFCT能在细节保持、条纹结构保持、条纹平滑3方面取得最好的效果，而MONet、MSFF-DCNN和Uformer这3种算法滤波效果虽然不如PFCT，但比其他传统算法要好很多。

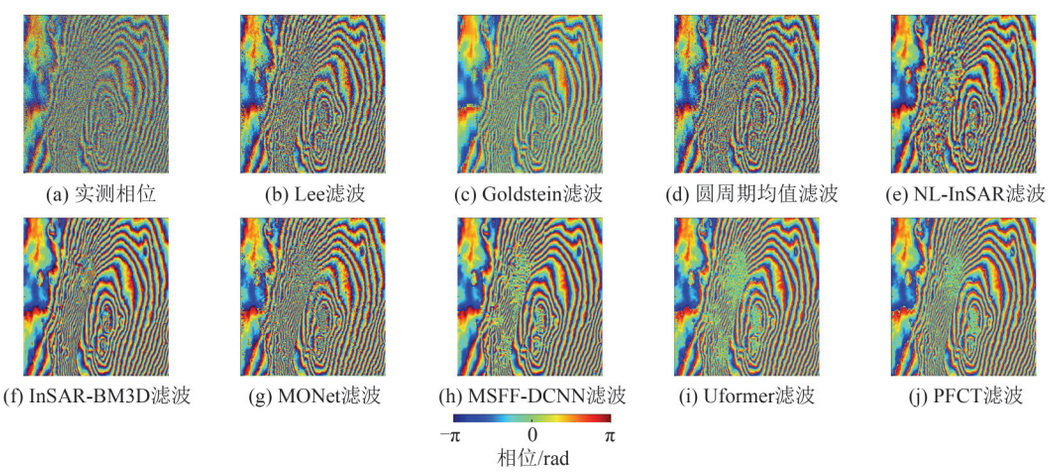


图14 Etna实测相位及滤波结果

Fig. 14 Measured Phase and Filtering Results of Etna

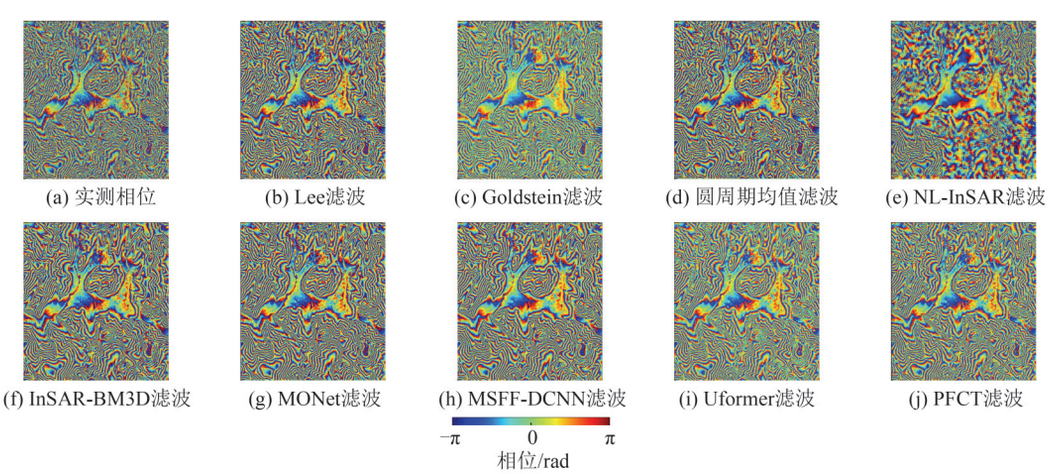


图15 辽宁省局部地区实测相位及滤波结果

Fig. 15 Measured Phase and Filtering Results in Local Areas of Liaoning Province

由图16可见，该实测相位存在大量高条纹密度区域，而且局部低相干区域增多，滤波难度比辽宁省局部地区的实测相位更大。传统算法的滤波效果相对较差，效果稍好一些的InSAR-

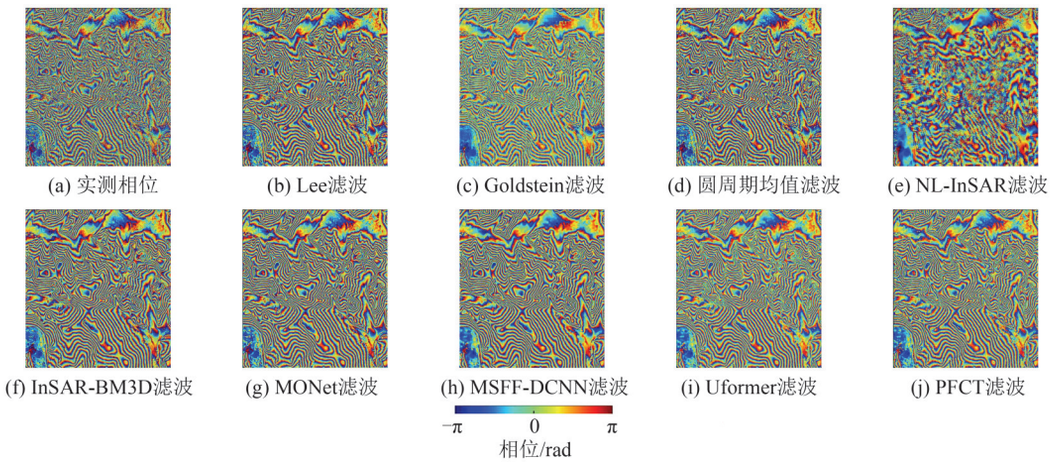


图 16　吉林省局部地区实测相位及滤波结果

Fig. 16　Measured Phase and Filtering Results in Local Areas of Jilin Province

BM3D滤波也存在局部模糊、噪点残存较多的现象。深度学习算法中，各算法都存在局部条纹断裂的现象，但PFCT断裂相对较少且其平滑性和细节保持能力均优于其他3种算法。

表2展示了3组实测数据的测试结果。对于Q指标，PFCT表现最优，表明PFCT的滤波结果在观感上更加清晰，更接近无噪声的理想状态。对于NOR指标，PFCT的残差去除效果仅比圆周期均值滤波差，但圆周期均值滤波是专门针对残差点进行处理的一种滤波算法，观感上条纹模糊，细节保持较差。总体而言，本文提出的PFCT在花费时间上比次优算法的MSFF-DCNN减少了65.4%的情况下，残点去除性能高了1.8%，展现了PFCT处理实测数据的优异性能。

表 2　实测数据的测试结果

Table 2　Test Results of Measured Data

滤波方法	NOR	Q	耗时/s
含噪相位图	301 525	28.8	
Lee滤波	133 812	32.1	4.022
Goldstein滤波	148 541	38.4	0.199
圆周期均值滤波	14 455	30.1	1.691
NL-InSAR	81 192	43.9	179.478
InSAR-BM3D	90 322	92.3	48.139
MONet	88 641	66.7	0.143
MSFF-DCNN	76 262	90.4	2.003
Uformer	77 152	89.8	0.639
PFCT	74 907	102.2	0.694

4　结　语

相位滤波是InSAR数据处理中的重要步骤，滤波质量直接影响后续相位解缠等处理的准确性。针对目前InSAR相位滤波精度受限的问题，本文首先提出一种基于真实DEM与真实相干系数图的InSAR数据集构建方法，通过真实DEM推算纯净相位，并用真实相干系数图模拟其噪声分布，减少了仿真数据与真实数据间的域间隔。然后提出一种以CNN与Transformer并行处理的CTB为骨干的PFCT，利用包含点误差损失和条纹结构误差损失的复合损失函数进行训练，通过不同条纹密度和噪声强度的仿真数据实验对模型进行性能评估。实验结果表明，PFCT的滤波效果较其他方法表现更优，在维持滤波效果的同时有效保持了条纹结构的完整性。最后，利用实测数据验证了PFCT的泛化性能。

参　考　文　献

[1] 邓云凯，禹卫东，张衡，等. 未来星载SAR技术发展趋势[J]. 雷达学报，2020，9(1)：1-33.
DENG Yunkai，YU Weidong，ZHANG Heng，et al. Forthcoming Spaceborne SAR Development[J]. *Journal of Radars*，2020，9(1)：1-33.

[2] 郭华东，张露. 雷达遥感六十年：四个阶段的发展[J]. 遥感学报，2019，23(6)：1023-1035.
GUO Huadong，ZHANG Lu. 60 Years of Radar Remote Sensing：Four-Stage Development [J]. *Journal of Remote Sensing*，2019，23(6)：1023-1035.

[3] BAMLER R，HARTL P. Synthetic Aperture Radar Interferometry[J]. *Inverse Problems*，1998，14(4)：R1-R54.

[4] 袁煜伟，李志伟，母岷政. 中国第一代全球大气再分析数据在InSAR大气改正中的应用研究：以南加州地区为例[J]. 武汉大学学报(信息科学版)，2025，50(5)：959-972.
YUAN Yuwei，LI Zhiwei，MU Minzheng. Application of China's First Generation Global Atmospheric

Reanalysis Data in InSAR Atmospheric Correction: A Case Study of Southern California[J]. *Geomatics and Information Science of Wuhan University*, 2025, 50(5): 959-972.

[5] 许文斌, 罗兴军, 朱建军, 等. InSAR火山形变监测与参数反演研究进展[J]. 武汉大学学报(信息科学版), 2023, 48(10): 1632-1642.
XU Wenbin, LUO Xingjun, ZHU Jianjun, et al. Review of Volcano Deformation Monitoring and Modeling with InSAR[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(10): 1632-1642.

[6] 赵峰, 张雷昕, 王腾, 等. 城市地表形变的双极化Sentinel-1数据极化时序InSAR技术监测[J]. 武汉大学学报(信息科学版), 2022, 47(9): 1507-1514.
ZHAO Feng, ZHANG Leixin, WANG Teng, et al. Polarimetric Persistent Scatterer Interferometry for Urban Ground Deformation Monitoring with Sentinel-1 Dual Polarimetric Data[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(9): 1507-1514.

[7] 韩炳权, 刘振江, 陈博, 等. 2022年泸定Mw 6.6地震InSAR同震形变与滑动分布[J]. 武汉大学学报(信息科学版), 2023, 48(1): 36-46.
HAN Bingquan, LIU Zhenjiang, CHEN Bo, et al. Coseismic Deformation and Slip Distribution of the 2022 Luding Mw 6.6 Earthquake Revealed by InSAR Observations[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(1): 36-46.

[8] ZEBKER H A, VILLASENOR J. Decorrelation in Interferometric Radar Echoes[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1992, 30(5): 950-959.

[9] XU G, XING M D, ZHANG L, et al. Bayesian Inverse Synthetic Aperture Radar Imaging[J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(6): 1150-1154.

[10] LANARI R, FORNARO G, RICCIO D, et al. Generation of Digital Elevation Models by Using SIR-C/X-SAR Multifrequency Two-Pass Interferometry: The Etna Case Study[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, 34(5): 1097-1114.

[11] LEE J S, PAPATHANASSIOU K P, AINSWORTH T L, et al. A New Technique for Noise Filtering of SAR Interferometric Phase Images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1998, 36(5): 1456-1465.

[12] BUADES A, COLL B, MOREL J M. A Review of Image Denoising Algorithms, with a New One[J]. *Multiscale Modeling & Simulation*, 2005, 4(2): 490-530.

[13] DELEDALLE C A, DENIS L, TUPIN F. NL-InSAR: Nonlocal Interferogram Estimation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(4): 1441-1452.

[14] SICA F, COZZOLINO D, ZHU X X, et al. InSAR-BM3D: A Nonlocal Filter for SAR Interferometric Phase Restoration[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(6): 3456-3467.

[15] DABOV K, FOI A, KATKOVNIK V, et al. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering[J]. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095.

[16] XU G, XING M D, XIA X G, et al. Sparse Regularization of Interferometric Phase and Amplitude for InSAR Image Formation Based on Bayesian Representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(4): 2123-2136.

[17] GOLDSTEIN R M, WERNER C L. Radar Interferogram Filtering for Geophysical Applications[J]. *Geophysical Research Letters*, 1998, 25(21): 4035-4038.

[18] ISOGAWA K, IDA T, SHIODERA T, et al. Deep Shrinkage Convolutional Neural Network for Adaptive Noise Reduction[J]. *IEEE Signal Processing Letters*, 2018, 25(2): 224-228.

[19] SICA F, GOBBI G, RIZZOLI P, et al. Φ-Net: Deep Residual Learning for InSAR Parameters Estimation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(5): 3917-3941.

[20] YU H W, YANG T X, ZHOU L F, et al. PDNet: A Lightweight Deep Convolutional Neural Network for InSAR Phase Denoising[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5239309.

[21] PU L M, ZHANG X L, ZHOU L M, et al. Nonlocal Feature Selection Encoder-Decoder Network for Accurate InSAR Phase Filtering[J]. *Remote Sensing*, 2022, 14(5): 1174.

[22] ZHANG K, ZUO W M, CHEN Y J, et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising[J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155.

[23] LI S, XU H P, GAO S, et al. An Interferometric Phase Noise Reduction Method Based on Modified Denoising Convolutional Neural Network[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 4947-4959.

[24] YANG W, HE Y, ZHANG L F, et al. A Deep Convolutional Neural Network with Multiscale Feature Dynamic Fusion for InSAR Phase Filtering[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 6687-6710.

[25] VITALE S, FERRAIOLI G, PASCAZIO V, et al. InSAR-MONet: Interferometric SAR Phase Denoising Using a Multiobjective Neural Network [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5239814.

[26] PU L M, ZHANG X L, ZHOU Z N, et al. A Phase Filtering Method with Scale Recurrent Networks for InSAR[J]. *Remote Sensing*, 2020, 12(20): 3453.

[27] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5998-6008.

[28] ZHENG S X, ZHANG X L, PU L M, et al. An InSAR Phase Filtering Method Based on Transformer Network [C]//IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022.

[29] FERRAIOLI G, DELEDALLE C A, DENIS L, et al. PARISAR: Patch-Based Estimation and Regularized Inversion for Multibaseline SAR Interferometry [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(3): 1626-1636.

[30] MURDACA G, RUCCI A, PRATI C. Deep Learning for InSAR Phase Filtering: An Optimized Framework for Phase Unwrapping [J]. *Remote Sensing*, 2022, 14(19): 4956.

[31] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C]//Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015.

[32] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, Online, 2021.

[33] ZHU X, MILANFAR P. Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content[J]. *IEEE Transactions on Image Processing*, 2010, 19(12): 3116-3132.