



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860,CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目： 基于深度学习的 InSAR 影像局部与非局部联合相位滤波
作者： 王杰, 柳稼航, 凌信鹏, 段则先
DOI: 10.13203/j.whugis20240052
收稿日期: 2024-04-08
网络首发日期: 2024-05-15
引用格式: 王杰, 柳稼航, 凌信鹏, 段则先. 基于深度学习的 InSAR 影像局部与非局部联合相位滤波[J/OL]. 武汉大学学报(信息科学版),
<https://doi.org/10.13203/j.whugis20240052>



网络首发: 在编辑部工作流程中,稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定,且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件,可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定;学术研究成果具有创新性、科学性和先进性,符合编辑部对刊文的录用要求,不存在学术不端行为及其他侵权行为;稿件内容应基本符合国家有关书刊编辑、出版的技术标准,正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性,录用定稿一经发布,不得修改论文题目、作者、机构名称和学术内容,只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约,在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版,以单篇或整期出版形式,在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z),所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20240052

引用格式：

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), 2024, DOI: 10.13203/j.whugis20240052 (WANG Jie, LIU Jiahang, LING Xinpeng, et al. Deep Learning-Based Joint Local and Non-local InSAR Image Phase Filtering Method[J]. Geomatics and Information Science of Wuhan University, 2024, 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¹

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：相位滤波是干涉合成孔径雷达 (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, 本文结合卷积神经网络 (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。首先, 基于 Transformer 全局特征提取能力强的特点, 提出了面向相位滤波的 PFCT 网络结构。该网络可以同时进行局部滤波与非局部滤波, 避免了现有方法因忽视干涉条纹非局部特征而导致的精度受限问题。其次, 提出了一种新的复合损失函数用于同时训练网络的去噪能力和条纹结构保持能力。在仿真实验中, 本文方法的均方误差 (mean square error, MSE) 比次优算法降低了 15.5%, 结构相似性 (structural similarity, SSIM) 指标提高了 5.3%。在实测实验中, 本文方法的残点去除性能比次优方法提高了 1.8%。结果表明, 本文方法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

Deep Learning-Based Joint Local and Non-local InSAR Image Phase Filtering Method

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 Space Photoelectric Detection and Sensing of Industry and Information Technology, Nanjing 210016, China

Abstract: Objectives: Phase filtering is one of the key technologies in interferometric synthetic aperture radar (InSAR) data processing, and the quality of interferograms significantly affects the accuracy of subsequent processing steps. The traditional phase filtering methods with manual parameter adjustment not only have low filtering accuracy but also low computational efficiency. Deep

¹ 收稿日期: 2024-04-08

基金项目: 江苏省双创人才计划项目(JSSCRC2021501)。

第一作者: 王杰, 硕士生, 研究方向为 InSAR 相位滤波与相位解缠。wj104777@nuaa.edu.cn

通信作者: 柳稼航, 博士, 教授。jhliu@nuaa.edu.cn

learning-based phase filtering methods have great potential, but current convolutional neural networks (CNN) based methods overlook the non-local features causing limiting accuracy. Transformer-based networks have strong global modeling capabilities and can extract non-local features from interference fringes, but they struggle with handling local features of interference fringes. **Methods:** To effectively improve the filtering effect, this paper proposes a joint local and non-local phase filtering method for InSAR images combined with the advantages of convolutional neural network (CNN) and Transformer network. First, based on the strong global feature extraction capability of the robust Transformer network, a 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 nonlocal features of interference fringes in the existing methods. Then, a new complex loss function is proposed to guide the network to maintain the integrity of interference fringes while improving filtering performance. **Results:** In the simulation data experiment, the mean square error index was 15.5% lower than the suboptimal algorithm, and the structural similarity index was 5.3% higher. In the true data experiment, the residue removal index was 1.8% higher than the suboptimal method. 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.

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, 不受云雨和昼夜影响, 具有全天时、全天候工作的能力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]。干涉合成孔径雷达 (interferometric synthetic aperture radar, InSAR) 是在此基础上发展出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、预警^[5-7]。然而, 由于系统热噪声、时间去相干、空间去相干等因素对相位造成较大干扰^[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, 在干涉合成孔径雷达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在高相干区域边缘条纹细节损失问题。但它受局部窗口中的残差点影响, 容易产生条纹伪影等失真现象。

非局部滤波算法以 Buades 等^[12]提出的 Non-local Mean 算法为核心框架, 不再局限于在小窗口对像素进行平滑, 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。代表算法为 NL-InSAR 和 InSAR-BM3D。Deledalle 等^[13]提出了 NL-InSAR 算法, 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。Sica 等^[14]将 BM3D 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。非局部滤波算法能够取得较好的滤波精度, 但非局部相似区域的搜索和处理操作需要大量的计算时间, 牺牲了计算效率。

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, Goldstein 等^[17]设计一个低通滤波器, 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强噪声区域滤波不足的现象。

近些年来, 随着深度学习技术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^[22], Li 等^[23]提出了 IPDnCNN。它将干涉相位变换至复数域, 同时对相位实部和相位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。MSFF_DCNN 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 针对以往的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, 目前的研究还处于起步阶段。大多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, 一些研究者们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^[26]的方法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。这种高斯噪声分布也与真实噪声分布相去甚远。同时, 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, 相位滤波效果不佳。相较之下, Transformer^[27]模型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, 条纹细节滤波效果较差^[28]。

为探究更有效的滤波方法, 本文提出了一种基于 CNN 与 Transformer 的 InSAR 影像相位滤波方法 PFCT。该方法能同时进行局部滤波和非局部滤波, 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, 以此模拟干涉所需要的主辅单视复数 (single-look complex, SLC) 影像。设计了针对相位实部和虚部特性的特征提取模块和相位重建模块, 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。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 相位滤波方法的有效性。

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

图 1 为 InSAR 干涉几何示意图， 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，即天线间距离， h 为目标距海平面垂直高度， H 为天线 S_1 距海平面垂直高度， R_1 和 R_2 代表了天线 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 表示基线与水平方向的夹角。

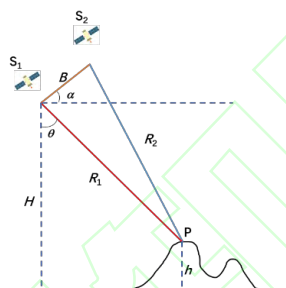


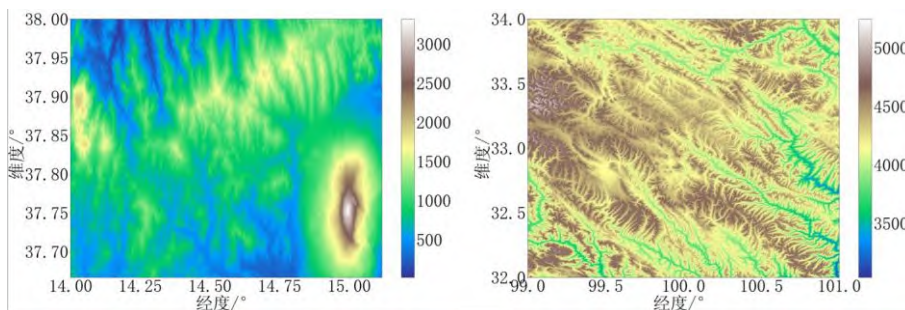
图 1 InSAR 干涉几何示意图
Fig.1 InSAR Interferometric Geometry Diagram

本文选择了航天飞机雷达地形测量任务(shuttle radar topography mission, SRTM)30m 分辨率的 DEM 来生成干净的干涉相位，传感器参数为空间基线 B 为 60m，波长 λ 为 0.056m，天线 S_1 观测方向的入射角 θ 为 45° ，基线与水平方向的夹角 α 为 45° ，影像中心到天线 S_1 传感器的距离 R_1 为 231km。用公式 (1)^[29] 可以将地形高度 h 转换为干净的地形干涉相位 φ_{clean} ：

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

其中 % 符号为求余符号。

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 Etna 火山周边地区作为测试数据集的参考 DEM。



(a)训练集参考 DEM

(b)测试集参考 DEM

图 2 数据集的参考 DEM

Fig.2 Reference DEM for Datasets

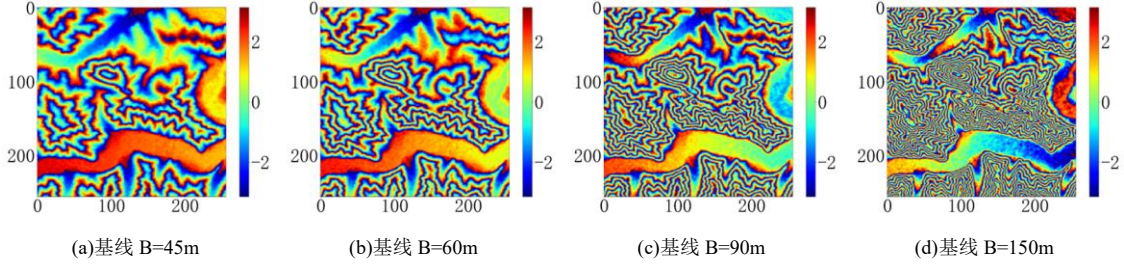


图 3 不同条纹频率下的干涉相位图

Fig.3 Interferometric Phase Maps with Different Fringe Frequencies

此外，为测试网络在不同的条纹频率的滤波结果，通过设置 45 米、60 米、90 米和 150 米的空间基线来改变相位图的条纹频率，如图 3 所示。将这些 DEM 数据切分成 64×64 的小块，并插值放大到 256×256 ，最终得到了 44944 个样本的训练数据集和 2688 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，还需要模拟它们对应的噪声相位作为输入。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：

$$\varphi_{\text{noisy}} = z_1 \cdot z_2^* \quad (2)$$

本文采用的噪声模型是对文献[30]所提方法的一种改进，简化其相干性图的构建方式。SLC 影像对 $(z_1, z_2)^T$ 可以由两个标准高斯变量 $(u_1, u_2)^T$ 表示^[15]：

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \mathbf{T} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (3)$$

其中， \mathbf{T} 是 $(z_1, z_2)^T$ 的协方差矩阵的 Cholesky 分解，可以由公式(4)表示。

$$\mathbf{T} = \begin{pmatrix} A_1 & 0 \\ A_2 \rho e^{-j\varphi_{\text{clean}}} & A_2 \sqrt{1 - \rho^2} \end{pmatrix} \quad (4)$$

可以看出模拟的 SLC 图像对 $(z_1, z_2)^T$ 只取决于上一节计算得到的真实相位 φ_{clean} ，相干性 ρ 和振幅 A_1, 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。所以，我们可以通过公式(5)得到模拟的 SLC 图像对 $(z_1, z_2)^T$ ：

$$\begin{cases} z_1 = Au_1 \\ z_2 = A\rho e^{-j\varphi_{\text{clean}}} u_1 + A\sqrt{1 - \rho^2} u_2 \end{cases} \quad (5)$$

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]。从 ALOS-PALSAR 的干涉图中随机提取四个子相干性图，并要求这四个子图的平均相干性分布在(0.4, 0.5)，(0.5, 0.6)，(0.6, 0.7)，(0.7, 0.8)区间内，让噪声强度分布更均衡。图 4 所示为对应的子相干性图，平均相干性分别是 0.44，0.54，0.62，0.76。

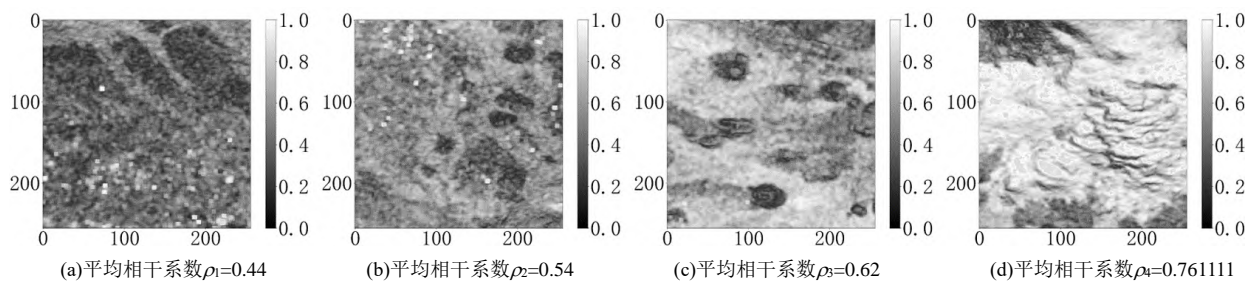


图4 用来模拟数据集的相干系数图
Fig.4 Coherence Maps Used to Simulate Datasets

然后，将这几幅相干性图乘上 255 来简单模拟振幅并且代入公式(5)以获得模拟的 SLC 图像对。图 5 所示是噪声相位数据集的模拟流程，最终，整个训练数据集包括 44944 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 2688 组图像对。

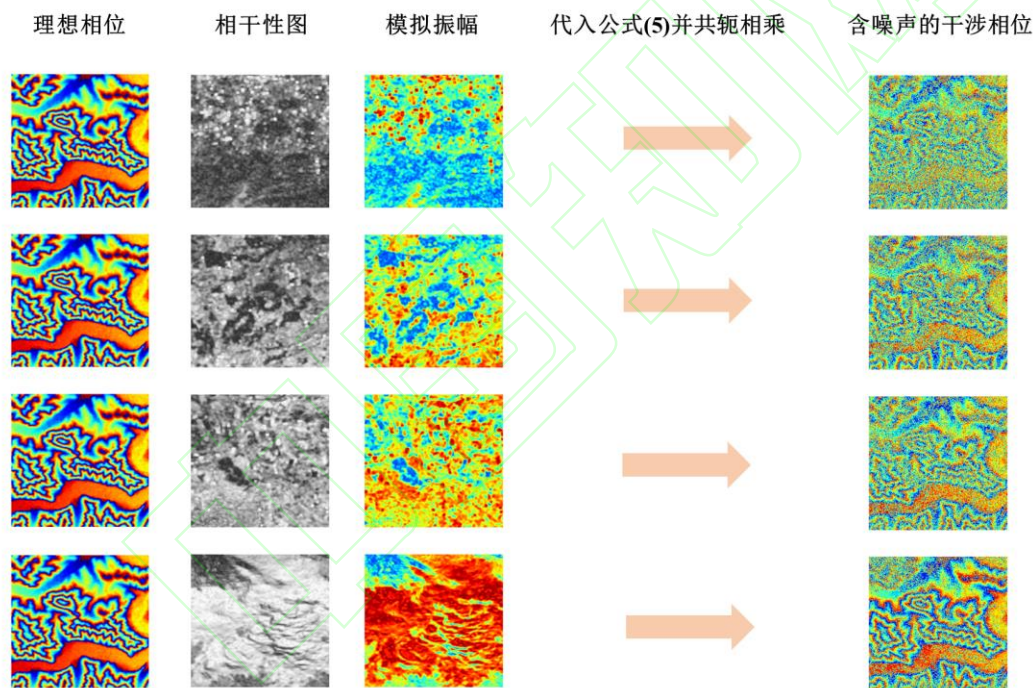


图5 噪声相位数据集的模拟流程
Fig.5 The Process of Simulating Noisy Phase Datasets

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，该网络将编码器部分的下采样与解码器部分的上采样跳跃连接，整体网络呈现 U 型，密集融合了浅层特征与深层特征，在各图像任务中表现突出。本文采用和 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。

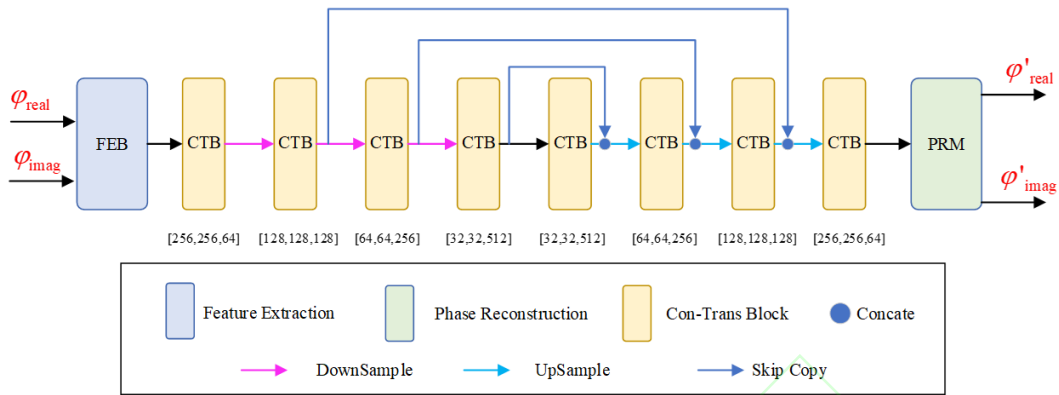


图 6 PFCT 网络结构
Fig.6 PFCT Net Structure

为了在滤波时不破坏条纹边缘处的相位突变，采用公式(6)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：

$$\begin{cases} \varphi_{\text{real}} = \cos(\varphi_{\text{noisy}}) \\ \varphi_{\text{imag}} = \sin(\varphi_{\text{noisy}}) \end{cases} \quad (6)$$

将噪声相位 φ_{noisy} 的实部 φ_{real} 和虚部 φ_{imag} 输入网络，最终输出滤波后相位的实部 φ'_{real} 和虚部 φ'_{imag} ，通过公式(7)就可以得到滤波后的相位 $\varphi_{\text{filtered}}$ ：

$$\varphi_{\text{filtered}} = \arctan \frac{\varphi'_{\text{imag}}}{\varphi'_{\text{real}}} \quad (7)$$

整个滤波流程如图 7 所示：

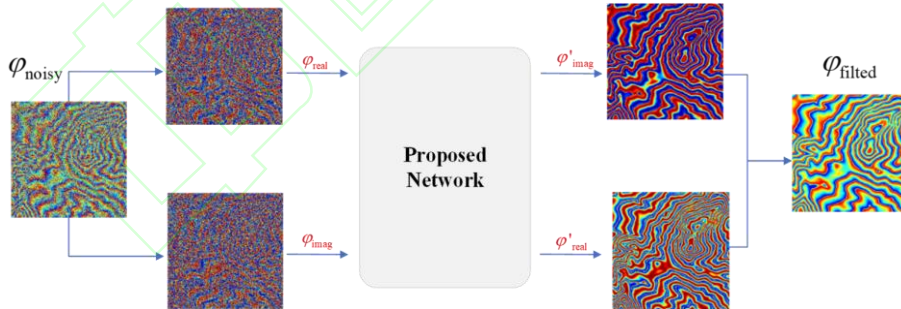


图 7 滤波流程
Fig.7 Filtering Process

2.1 特征提取模块

相位在进入网络之前被拆分成了实部和虚部，但实部和虚部存在着一定关联性。图 8 所示的特征提取模块提取完两者浅层特征后，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 Tanh 函数进行初次特征提取，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，再使用 BatchNorm 层对特征进行正则化，之后利用 LeakyReLU 函数进行激活，将实部和虚部激活后的特征进

行拼接，送入下一个模块。

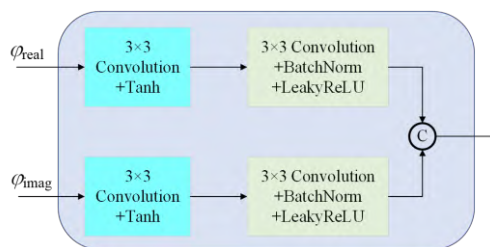


图 8 特征提取模块
Fig.8 Feature Extraction Modul

2.2 CTB 模块

CNN 通常采用固定大小的卷积核，专注于局部特征提取，但对于 InSAR 相位图中复杂的结构时，容易受到感受野的限制，相位滤波效果不佳。相比之下，Transformer 模型以其强大的自注意力机制而具备较强的全局关系捕捉能力，可以计算任意两点间的相关程度，能更好地捕获不同区域之间的关系。通过对不同区域进行加权处理，实现了非局部处理。但在处理 InSAR 相位图中的微观细节方面相对薄弱，容易忽略图像中的局部特征。Swin 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，交替使用这两种自注意力机制，既减少了计算成本又实现了信息的跨窗口交互。

CTB 模块是 PFCT 网络的骨干模块，整体模块结构如图 9(a)所示。特征图先进行 1×1 的卷积，如果该 CTB 模块处在编码器阶段，则保持输入的特征图通道维度不变，如果其处在解码器阶段，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，两个子模块的结构如图 9 (b)和图 9 (c)所示。滤波后的两份特征图在通道维进行拼接，最后再通过 1×1 的卷积，就得到了滤波后的特征图。

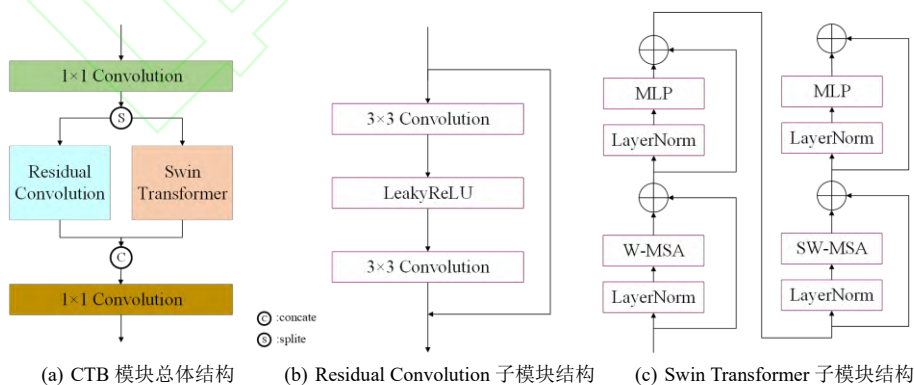


图 9 CTB 模块结构
Fig.9 The structure of CTB Module

2.3 相位重建模块

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，整体模块结构如图 10 所示。该模块以最后一个 CTB 模块输出的特征

图作为输入，然后分为实部分支和虚部分支，每个分支都独立对输入的特征图进行处理。首先通过一个 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出滤波后的相位实部和相位虚部。

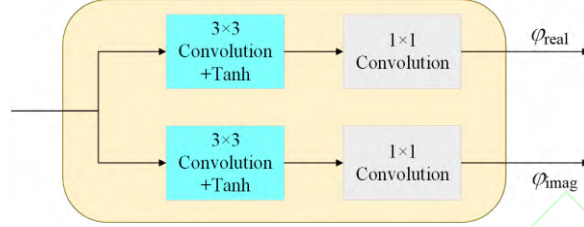


图 10 相位重建模块
Fig.10 Phase Reconstruction Module

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} ，如公式(8)所示。

$$L_{\text{total}} = L_{\text{point}} + L_{\text{stripe}} \quad (8)$$

L_{total} 由两个子损失函数组成，分别是每个点的平均相位滤波误差损失 L_{point} 和反映条纹结构误差损失的 L_{stripe} 。 L_{point} 可以由公式(9)计算出来：

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

其中 (i,j) 为相位点的坐标， n 为图像的尺寸。可以看出， L_{point} 由真实相位 φ_{gt} 和滤波后相位 $\varphi_{\text{filtered}}$ 的 L2 损失、真实相位和滤波后相位的实部 $\varphi_{\text{gt_real}}$ 和 φ'_{real} 与虚部 $\varphi_{\text{gt_imag}}$ 和 φ'_{imag} 的 L1 损失三项组成。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 L2 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 L1 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。

由于滤波方法对条纹结构的保持效果难以直接量化，但在条纹边缘处巨大的相位跳变导致该点的水平梯度或垂直梯度远大于非边缘点，因而只需要将这些位置的点在滤波后相位图中保持其梯度，就可以视作条纹结构得到了保持。基于这个思想，本文计算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作为条纹结构损失，即 L_{stripe} 。

L_{stripe} 可以由公式(10)计算出来：

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

其中 ∇_v 和 ∇_h 是求该点的垂直梯度和水平梯度，计算过程见公式(11):

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

3 实验与分析

3.1 实验细节

本研究提出的网络使用包含 12 个样本的小批量进行训练，使用 Adam 优化器，其参数为 $\beta_1 = 0.9$ 和 $\beta_2 = 0.99$ 。训练持续了 120 个周期，初始学习率设定为 $\eta = 0.0002$ ，并且每 20 个周期减半一次。本研究提出的网络在一台 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也均为 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，两者的均方差（mean square error, MSE）和结构相似性（structural similarity, SSIM），可以作为客观衡量滤波效果的指标。其中，MSE 和 SSIM 可以由式(12)和式(13)求得：

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

其中，N 代表相位图中的像素点数，MSE 值越趋于 0，整体滤波效果越好。

$$\text{SSIM} = \frac{\left(2\mu_{\varphi_{\text{gt}}} \mu_{\varphi_{\text{filtered}}} + (k_1 L)^2\right) \left(2\sigma_{\varphi_{\text{gt}} \varphi_{\text{filtered}}} + (k_2 L)^2\right)}{\left(\mu_{\varphi_{\text{gt}}}^2 + \mu_{\varphi_{\text{filtered}}}^2 + (k_1 L)^2\right) \left(\sigma_{\varphi_{\text{gt}}}^2 + \sigma_{\varphi_{\text{filtered}}}^2 + (k_2 L)^2\right)} \quad (13)$$

其中， $k_1 L$ 和 $k_2 L$ 是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 $\mu_{\varphi_{\text{filtered}}}$ 代表真实相位和滤波后相位的均值， $\sigma_{\varphi_{\text{gt}}}$ 和 $\sigma_{\varphi_{\text{filtered}}}$ 代表真实相位和滤波后相位的标准差， $\sigma_{\varphi_{\text{gt}} \varphi_{\text{filtered}}}$ 代表真实相位和滤波后相位的协方差。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优异，但在实际应用中，实际测量区域不存在实时更新保持准确的参考 DEM，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 $Q^{[32]}$ 来评价网络对实测相位的滤波效果。无参考指标 Q 可以通过对干涉相位梯度矩阵 \mathbf{G} 进行奇异值分解计算得到：

$$\mathbf{G} = \mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{U} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} [\mathbf{v}_1 \quad \mathbf{v}_2]^T \quad (14)$$

其中, s_1 和 s_2 便是梯度矩阵 \mathbf{G} 的奇异值, $s_1 \geq s_2 \geq 0$, 与之对应的 v_1, v_2 分别代表了局部梯度场的主导方向和边缘方向^[32]。由公式(15)计算得到 Q 的值, Q 越高, 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^[32]:

$$Q = s_1 \frac{s_1 - s_2}{s_1 + s_2} \quad (15)$$

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

本文将提出的方法与一些经典的传统相位滤波算法, 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, 以及一些近年来典型的基于深度学习的算法进行对比试验, 比较这些算法的滤波效果。

3.3.1 仿真数据实验

图 11 展示了不同传统相位滤波方法的结果。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 滤波效果比其它传统算法好一些, 但在一些局部区域出现了条纹断裂的问题。

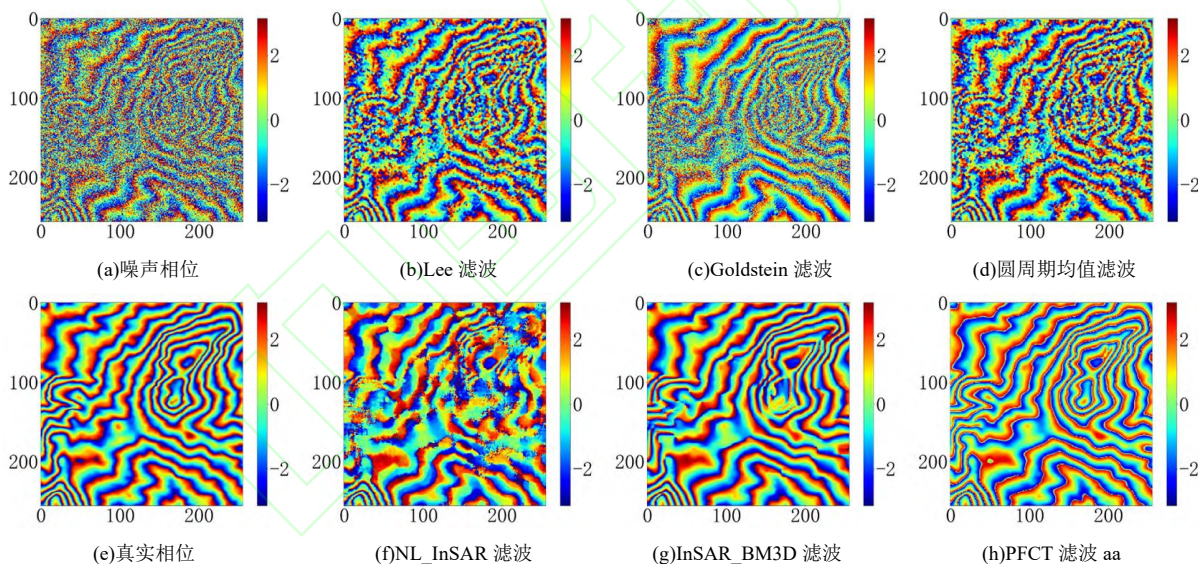
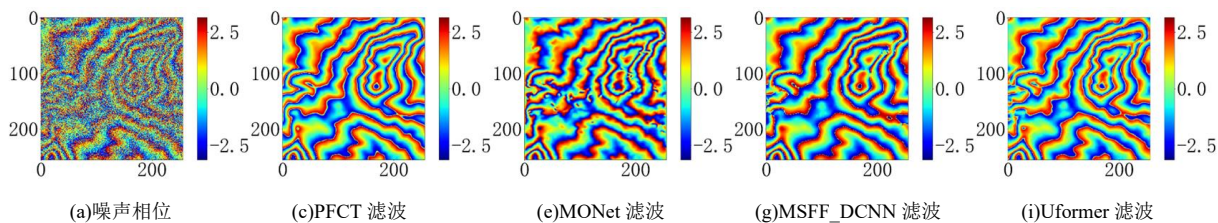


图 11 与传统方法滤波结果的对比

Fig.11 Comparison of Filtering Results with Traditional Methods



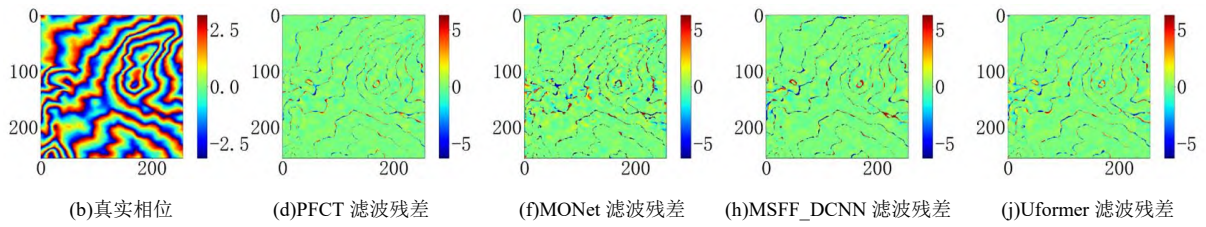


图 12 参数 B=60 时深度学习方法滤波结果

Fig.12 Deep Learning Filtering Results under Parameter B=60

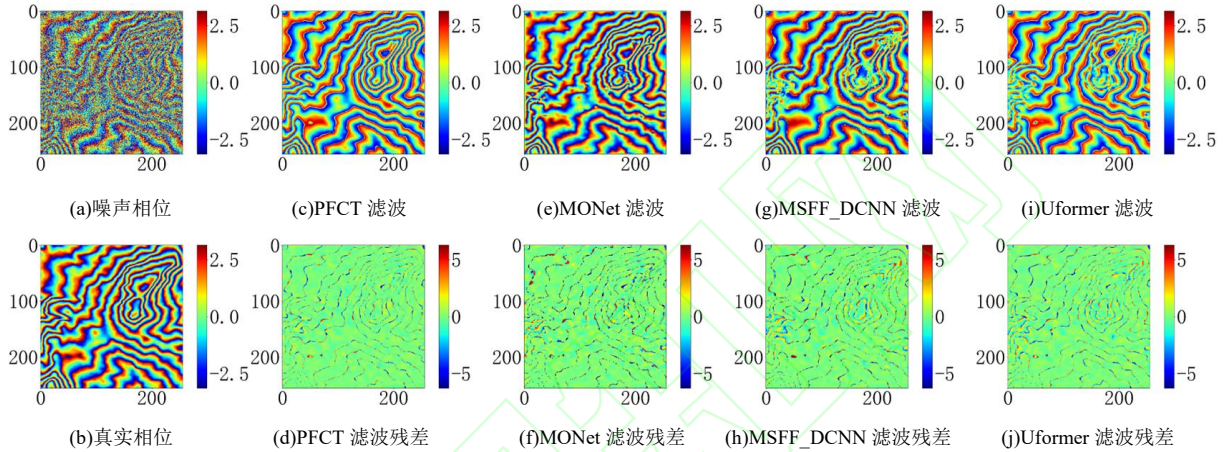


图 13 参数 B=90 时深度学习方法滤波结果

Fig.13 Deep Learning Filtering Results under Parameter B=90

图 12 和图 13 展示了不同条纹密度下不同深度学习算法的滤波结果。可以看出，本文提出的 PFCT 网络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 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，但是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。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 的滤波效果优于其它算法，在不同条纹密度下都具有良好的结构和细节保持能力。在整个仿真数据集上的测试结果见表 1，PFCT 算法的 MSE 指标比次优算法的 MSFF_DCNN 算法低了 15.5%，SSIM 指标则高了 5.3%，展现了 PFCT 算法的优异性能。

表 1 仿真数据的测试结果

Tab.1 Testing Results of Simulated Datasets

滤波方法	MSE	SSIM	NOR	Time/s
含噪相位图	4.166	0.147	11244	\
Lee 滤波 ^[11]	1.059	0.416	13872	0.484
Goldstein 滤波 ^[17]	2.159	0.382	16181	0.053
圆周期均值滤波 ^[10]	2.579	0.376	1467	0.191
NL_InSAR ^[13]	2.411	0.439	3585	17.084
InSAR_BM3D ^[14]	1.154	0.699	6612	6.214
MONet ^[25]	1.186	0.679	975	0.018

MSFF_DCNN ^[24]	0.801	0.721	1152	0.205
Uformer ^[28]	0.983	0.713	1113	0.068
PFCT	0.677	0.759	855	0.091

3.3.2 实测数据实验

为进一步验证本文方法对实际影像的滤波效果，选取由 ERS 测出的 Etna 火山主辅 SLC 影像和由 ALOS PALSAR-1 测出的辽宁局部地区以及吉林局部地区主辅 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，图 14、图 15 和图 16 展示了分别展示了这三组数据实测干涉相位与滤波结果。

观察图 14 的 Etna 实测相位滤波结果，传统算法中 Lee 滤波、Goldstein 滤波、圆周期均值滤波三类方法都有大量明显的噪点未能去除，NL_SAR 滤波结果有大量条纹失真，滤波效果较差。InSAR_BM3D 滤波的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原有条纹的结构，但条纹边缘不清晰，滤波结果不够平滑。深度学习方法中，MONet 滤波同样存在类似现象，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 算法稍差一些，其余区域区域滤波效果相对较好。但 InSAR_BM3D 算法花费的时间是 PFCT 的七十多倍，计算效率显著低于 PFCT。

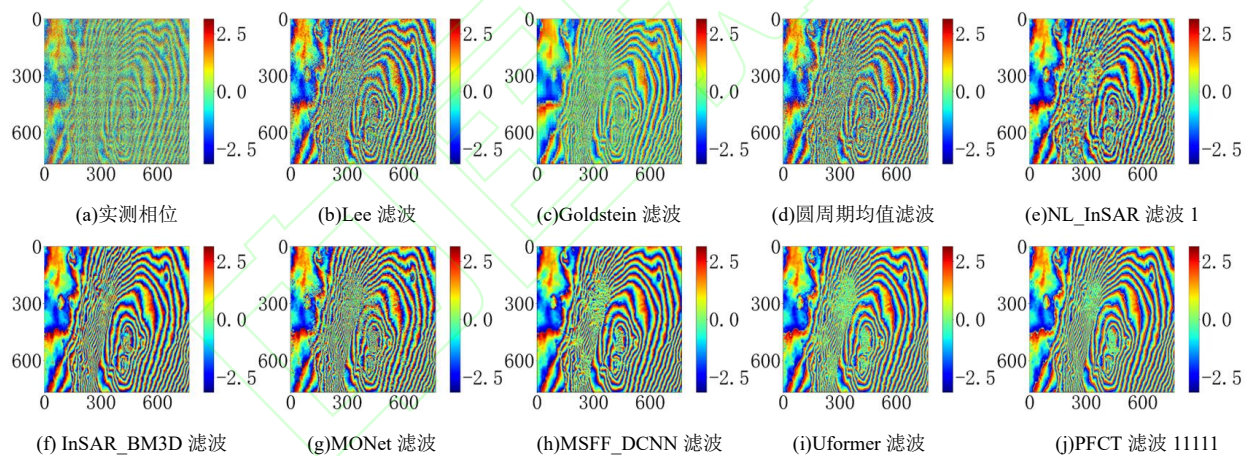
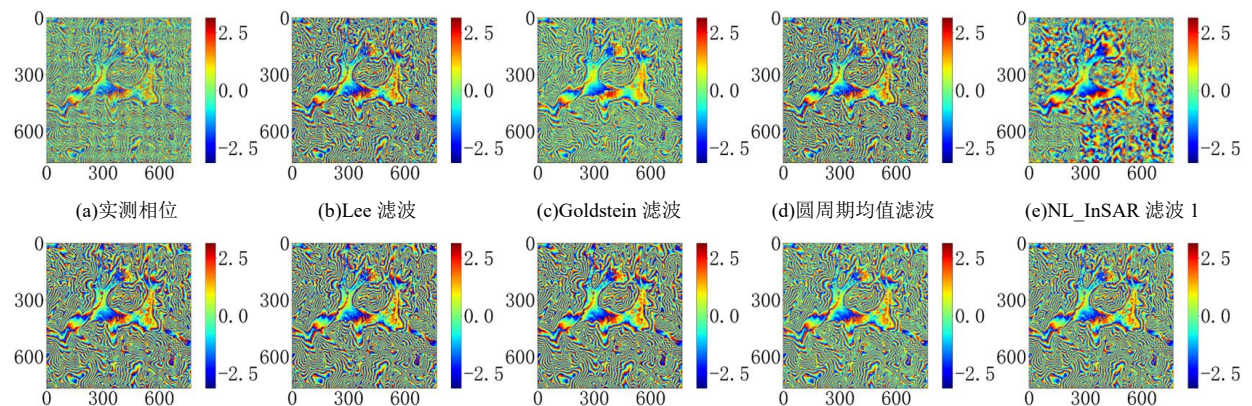


图 14 Etna 实测相位滤波结果

Fig.14 The Filtering Results of Etna Measured Phase



(f) InSAR_BM3D 滤波 (g) MONet 滤波 (h) MSFF_DCNN 滤波 (i) Uformer 滤波 (j) PFCT 滤波 11111

图 15 辽宁地区实测相位滤波结果

Fig.15 The Filtering Results of Measured Phase in Liaoning Region

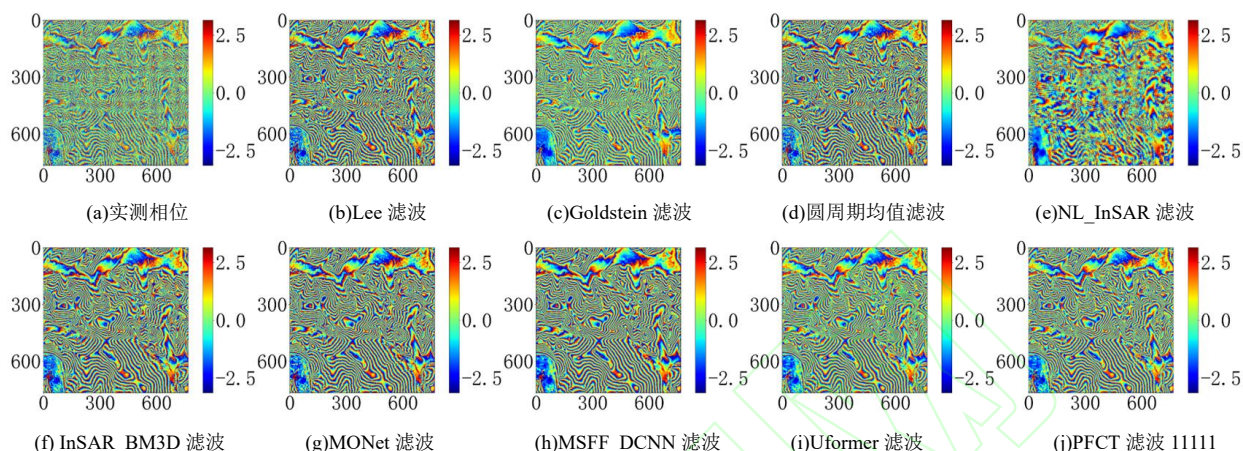


图 16 吉林地区实测相位滤波结果

Fig.16 The Filtering Results of Measured Phase in Jiling Region

图 15 展示了辽宁局部地区实测相位以及滤波结果，可以明显看出实测相位的条纹复杂程度远高于图 14 展示的 Etna 实测数据，滤波难度比 Etna 数据更大一些。传统算法中仍然只有 InSAR_BM3D 算法滤波结果较好，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 算法能在细节保持、条纹结构保持、条纹平滑三方面取得最好的效果，而 MONet、MSFF_DCNN 和 Uformer。这三种算法滤波效果虽然不如 PFCT，但比其他传统方法要好很多。

图 16 展示了吉林局部地区实测相位以及滤波结果，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，滤波难度比辽宁实测相位更高一些。传统算法的滤波效果相对较差，稍好一些的 InSAR_BM3D 算法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 算法断裂相对较少，其平滑性和细节保持能力优于其他三种方法。

表 2 展示了三组实测数据的测试结果。PFCT 算法在无参考指标 Q 取得了最优，表明 PFCT 滤波算法的滤波结果观感上更加清晰，更接近无噪声的理想状态。在 NOR 指标上，PFCT 算法仅比圆周期均值滤波差，其他算法的残差去除效果都不如 PFCT，而且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，本研究提出的 PFCT 算法在花费时间比次优的 MSFF_DCNN 算法减少了 65%的情况下，残点去除性能高了 1.8%，展现了 PFCT 算法在处理实测数据时的优异性能。

表 2 实测数据的测试结果

Tab.2 Testing results of real data

滤波方法	NOR	Q	Time/s
含噪相位图	301525	28.8	\
Lee 滤波	133812	32.1	4.022
Goldstein 滤波	148541	38.4	0.199
圆周期均值滤波	14455	30.1	1.691

NL_InSAR	81192	43.9	179.478
InSAR_BM3D	90322	92.3	48.139
MONet	88641	66.7	0.143
MSFF_DCNN	76262	90.4	2.003
Uformer	77152	89.8	0.639
PFCT	74907	102.2	0.694

4 结论

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 模块为骨干的 PFCT 相位滤波网络模型, 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 网络模型较其它方法更优的滤波效果, 在维持滤波效果的同时有效保持了条纹结构的完整性。最后再用实测数据验证了 PFCT 网络模型的泛化性能。

参考文献

- [1] Deng Yunkai, Yu Weidong, Zhang Heng, et al. Forthcoming Spaceborne SAR Development[J]. *Journal of Radars*, 2020, 9(1): 1-33.(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.)
- [2] Guo Huadong, Zhang Lu. 60 Years of RadarRemote Sensing: Four-stage Development[J]. *Journal of Remote Sensing*, 2019, 23(6): 1023-1035.(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.)
- [3] Bamler R, Hartl P. Synthetic Aperture Radar Interferometry[J]. *Inverse Problems*, 1998, 14(4): R1-R54.
- [4] Zhu Jianjun, Li Zhiwei, Hu Jun. Research Progress and Methods of InSAR for Deformation Monitoring[J]. *Acta Geodaetica et Cartographica Sinica*, 2017,46(10): 1717-1733.(朱建军, 李志伟, 胡俊. InSAR 变形监测方法与研究进展[J]. 测绘学报, 2017, 46(10): 1717-1733.)
- [5] 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.(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.)
- [6] 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.(赵峰, 张雷昕, 王腾, 等. 城市地表面变化的双极化 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.)
- [7] 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.(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.)
- [8] Zebker H, Villasenor J. Decorrelation in Interferometric Radar Echoes[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1992, 30: 950-959.
- [9] Xu G, Xing M, 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/XSAR 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, 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*, 2014, 53(4): 2123-2136.
- [17] Goldstein R M, Werner C L. Radar Interferogram Filtering for Geophysical Applications[J]. *Geophysical Research Letters*, 1998, 25: 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, Yang T, Zhou L, et al. PDNet: A lightweight deep convolutional neural network for InSAR phase denoising[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-9.
- [21] Pu L, Zhang X, Zhou L, 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, Chen Y, 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, 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, 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: 1-14.
- [26] Pu L, Zhang X, Zhou Z, 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, Zhang X, Pu L, et al. An InSAR Phase Filtering Method Based on Transformer Network[C]// 2022 IEEE International Geoscience and Remote Sensing Symposium(IGARSS), 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*, 2017, 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]// The 18th Medical Image Computing and Computer-Assisted Intervention(MICCAI), Munich, Germany, 2015.
- [32] 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.

网络首发:

标题: 基于深度学习的 InSAR 影像局部与非局部联合相位滤波

作者: 王杰, 柳稼航, 凌信鹏, 段则先

收稿日期: 2024-04-08

DOI:10.13203/j.whugis20240052

引用格式:

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), 2024, DOI: 10.13203/j.whugis20240052 (WANG Jie, LIU Jiahang, LING Xinpeng, et al. Deep Learning-Based Joint Local and Non-local InSAR Image Phase Filtering Method[J]. Geomatics and Information Science of Wuhan University, 2024, DOI: 10.13203/j.whugis20240052)

网络首发文章内容和格式与正式出版会有细微差别，请以正式出版文件为准！

您感兴趣的其他相关论文：

基于“三查”体系的丹巴县滑坡隐患早期识别与监测

付豪, 李为乐, 陆会燕, 许强, 董秀军, 郭晨, 谢毅, 王栋, 刘刚, 马志刚

武汉大学学报(信息科学版), 2024, 49(5): 734-746.

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20220418>

兰州新区平山造城竣工后地表形变 InSAR 时序监测与分析

龚翔, 许强, 蒲川豪, 陈婉琳, 修德皓, 纪续

武汉大学学报(信息科学版), 2024, 49(2): 236-245.

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20210553>

一种利用卷积神经网络的干涉图去噪方法

陶立清, 黄国满, 杨书成, 王童童, 盛辉军, 范海涛

武汉大学学报(信息科学版), 2023, 48(4): 559-567.

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20200589>