

基于深度学习的合成孔径雷达图像去噪算法

付相为¹, 单慧琳^{1,2*}, 吕宗奎¹, 王兴涛²¹南京信息工程大学电子与信息工程学院, 江苏 南京 210044;²无锡学院电子信息工程学院, 江苏 无锡 214105

摘要 针对合成孔径雷达在生成图像时产生的固有斑点噪声问题,提出了一种基于多尺度注意力级联卷积神经网络的去噪算法,用多尺度卷积网络与注意力机制来实现图像的特征提取,将级联网络结构作为主网络中的特征增强部分,并加入批量归一化来防止模型中出现过拟合。实验结果表明,所提算法相较于其他传统图像去噪算法,峰值信噪比、结构相似性分别平均提高了 0.75 dB~14.45 dB 和 0.01~0.16,在图像熵上也优于其他算法,并且能较好地恢复图像中的细节信息。

关键词 图像处理; 合成孔径雷达; 卷积神经网络; 图像去噪; 多尺度注意力

中图分类号 TP751.1 文献标志码 A

DOI: 10.3788/AOS221437

1 引言

合成孔径雷达(SAR)是一种捕捉微波的传感器^[1],其原理就是通过波形的反射来进行图像建立,从而解决了传统光学的遥感雷达在采集图像时受到天气、空气杂质等环境影响的问题。SAR应用最多的就是变换检测(CD)^[2],变换检测就是针对某一目标进行动态获取图像信息,包括图像预处理、生成差异图和分析计算差异图三个步骤,应用于自然灾害估计、资源管理分配和土地地貌特征测量等领域中。然而,在变换检测过程中,SAR图像中固有的斑点噪声会降低其变换检测的性能。因此,图像去噪方法已成为变换检测中预处理的基本方法。如何从有噪声的SAR图像中恢复出一幅干净的图像是当前迫切需要解决的问题。

目前SAR图像中去噪主要分为两个方向,一个是图像滤波处理^[3],另一个是非相干多视处理^[4]。图像滤波处理主要通过滤波器进行噪声滤除,针对SAR图像中的固有斑点噪声问题研究人员进行了深入研究。Donoho等^[5-6]提出了小波阈值收缩法,将噪声与信号进行小波变换,在小波域中由于小波变换本身就对信号能量有着较强的数据集中能力,故在频域上可以有效地区分信号与噪声,再通过设定的阈值进行滤波处理,以达到去噪的效果,后来这一理论也成为了传统去噪算法中的基础部分。非局部均值滤波^[7]也是空间滤波去噪中的主流方法,其算法区别于阈值滤波,是一种空间全局算法,即将整幅图像中相似的灰度信息转换成

当前的像素灰度,其相似度用欧氏距离来表示,通过全幅图像的灰度加权平均来进行有效去噪,代表性算法有SAR-BM3D^[8]。MuraliMohanBabu等^[9]对文献[8]进行了改进,使用双边采样,并在对BM3D中图像块进行采样后进行小波离散变换,这在某种程度上引入了小波阈值的思想,故去噪图像中能保留更丰富的信息。之后,一种稀疏表示的模型被应用于SAR图像去噪领域中,其稀疏表达对斑点噪声的去除效果表现较好^[10],但因算法过于庞大,占用计算资源较多,故未得到广泛应用。另外,Rudin等^[11]提出了一种全变分的模型进行去噪,在产生阶梯效果的同时,在斑点噪声方面也取得了不错的去噪效果。文献[12-13]对全变分去噪方法进行了改良优化,在尽可能减少阶梯效果的前提下,保证了算法模型的自适应性,并且很好地保留了图像边缘信息。在低秩聚类算法方面,Gu等^[14]提出了WNNM。由于自然界中图像的低秩性与图像噪声的高秩性有着本质的差别,故WNNM通过低秩矩阵分解和核范数最小化的方法来解决图像低秩矩阵的优化问题,从而利用图像的非局部相似性来进行图像去噪,与BM3D算法相比,该方法的去噪效果有着质的提升。虽然这些算法在去除SAR斑点噪声方面都取得了不错的效果,但是在实际应用中计算速率较低,且占用较多的计算资源。另外一种主流的SAR图像去噪方法就是非相干多视去噪,其原理是将图像分解为多个多普勒带宽,每一个多普勒带宽分别使用不同的合成孔径,最后将分解的图像进行叠加。该方法有效地

收稿日期: 2022-07-07; 修回日期: 2022-08-04; 录用日期: 2022-08-25; 网络首发日期: 2022-09-04

基金项目: 国家自然科学基金(62071240,62106111)

通信作者: *shanhuilin@nuist.edu.cn

去除了斑点噪声,但分解图像会造成分辨率下降,分辨率的下降将会对变换检测中的其他任务造成很大的影响。

近年来,随着深度学习的快速发展,其在图像去噪方面有着非常好的表现。针对普通自然图像的去噪,研究人员进行了许多研究。Zhang 等^[15]提出了经典的去噪网络 DnCNN,该网络是端对端去噪机制,利用残差网络学习与批量归一化(BN)的方法来加速网络的收敛,在特征提取方面有着较好的鲁棒性,同时也取得了不错的去噪效果。随后,Yue 等^[16]采用对抗生成网络(GAN)框架对模型进行训练。基于 GAN 模型训练的不稳定性成为对抗网络中的最大限制,故需要更长的时间来收敛。上述算法在普通图像去噪中产生了不错的效果。由于 SAR 图像噪声本身是乘性噪声,故研究人员针对 SAR 图像去噪问题提出了 SAR-CNN^[17]。然而,由于噪声的特殊性,算法对其训练样本有着一定的要求,SAR-CNN 只选取了非相干多视的图像,故在噪声样本训练中有着先天的劣势,去噪效果一般。Thakur 等^[18]提出了 AGSDNet,该网络通过融合注意力机制与梯度的方式来进行去噪,虽然取得了较好的去噪结果,但是模型本身具有较高的冗余性,需要大量的数据集作为支撑,网络收敛速率一般。

综上所述,传统的 SAR 图像去噪算法一般都运用了全局去噪思想,其原理都是利用图像中的全局相似信息来进行处理判断。在图像分辨率较高的情况下,算法需要进行平滑等一系列预处理,再通过每个图像块的邻域处理完成像素的区分,故需要占用大量的计算资源,并且在实际应用中具有一定的时空局限性,进而不能高效地完成去噪任务。这些算法虽然提升了去噪效果,但是也增加了时间成本。在深度学习方面,上述算法虽然表现尚佳,但是在网络收敛速度、模型的冗余度和精度方面仍有提升空间。针对这些问题,本文提出了一种基于多尺度注意力的级联卷积神经网络(MALNet)的去噪算法,该网络采用了多尺度非对称卷积核和注意力的思想,多尺度非对称卷积核相较于单一的卷积核有着很好的图像感受野,即它能够从不同的尺度来进行图像信息的收集,以便提取到更为精详的图像细节特征。随后,网络对不同尺度的卷积核进行维度拼接,引入注意力机制对拼接后的特征图进行特征的注意力划分,从而使得整个模型对图像的主要特征有良好的增强能力。网络中间采用密集级联层(DCB)来对特征进行进一步强化,最后通过差减计算来实现图像的恢复与重建。

2 MALNet 网络结构

2.1 非对称卷积

目前大多数卷积神经网络(CNN)中使用的是规则形卷积结构,如 3×3 、 4×4 、 5×5 和 $n \times n$ 。规则形的卷积核结构在特征提取时虽然表现出不错的效

果,但是在一些图像边缘信息和图像模糊信息的提取方面仍有不足。本文使用的是非对称卷积结构,该结构由 Tian 等^[19]于 2022 年提出,即非对称卷积网络(ACNet)。该网络对恢复有噪声的低分辨率图像非常有效,非对称卷积相较于现有的正方形卷积,可以进行模型压缩和加速。已有的研究表明,该结构可以将标准的 $d \times d$ 卷积分解为 $1 \times d$ 卷积和 $d \times 1$ 卷积,以减少参数量。该设计背后的理论比较简单,即若二维卷积核的秩为 1,则运算可等价地转换为一系列一维卷积。然而,由于深度网络中卷积核学习到的特征值是遵循图像中主要特征信息分布的,其内部秩比实际中的高,故直接将特征图像应用于卷积核会导致显著的信息损失。Jin 等^[20]运用结构约束使二维卷积分离,在获得不错精度的条件下,运算时间约缩短为原来的 1/2。Denton 等^[21]基于奇异值分解(SVD)分解找到了一种无限逼近低秩矩阵,然后对上层网络进行精细化以提升恢复性能。Jaderberg 等^[22]通过最小化重构误差,成功地将水平核和垂直核应用于网络学习中。另一方面,非对称卷积(conv)也被广泛地应用于网络结构设计中,如在 Inception-v3 中, 7×7 卷积被 1×7 卷积和 7×1 卷积代替。本文中使用的 $\text{conv}(3 \times 3)$ 、 $\text{conv}(3 \times 1)$ 、 $\text{conv}(1 \times 3)$ 和 $\text{conv}(5 \times 5)$ 这 4 个卷积核来进行图像的初步特征提取。最后,将它们进行融合拼接(concat),其网络可以表达为

$$F(x) = \text{concat} \left\{ \text{ReLU}[\text{conv}(5 \times 5)] + \text{ReLU}[\text{conv}(3 \times 1)] + \text{ReLU}[\text{conv}(1 \times 3)] + \text{ReLU}[\text{conv}(3 \times 3)] \right\}. \quad (1)$$

每一个卷积核后都需有 ReLU 激活函数,将图像数值滤波至 $[0, \infty]$ 的范围内,以防止图像在反向传播时产生梯度爆炸和消失。非对称卷积核结构示意图如图 1 所示。

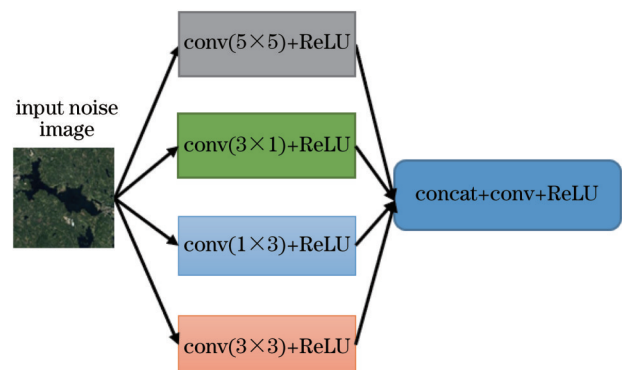


图 1 非对称卷积核结构示意图

Fig. 1 Schematic diagram of asymmetric convolution kernel

2.2 注意力机制

注意力机制本质上就是网络自主学习出一组权重系数,以“动态加权”的方式来强调所感兴趣的区域,并抑制不相关背景区域的机制。目前主流使用的注意力

机制为通道注意力、空间注意力和自注意力。通道注意力指先构建出当前模型不同通道的权重系数,再通过学习来捕获最佳的模型通道,并自主削减其余通道的权重系统,从而实现强化重要特征和抑制非重要特征,典型结构为SEnet^[23]。空间注意力主要是提升关键区域的特征表达,本质上是讲原始图片中的空间信息通过空间转换模块变换到另一个空间中并且保留关键

信息,为每个位置生成权重掩膜(mask)并加权输出,从而实现感兴趣的特定目标区域的增强,并弱化不相关的背景区域。本文引入的是当前比较主流的注意力机制,即CBAM^[24],结合了空间与通道双重注意力机制,故效果更佳,结构如图2所示。其中,MaxPool为最大池化,AvgPool为平均池化,sig为Sigmoid激活函数, \oplus 为特征融合相加, \otimes 为特征数值相乘。

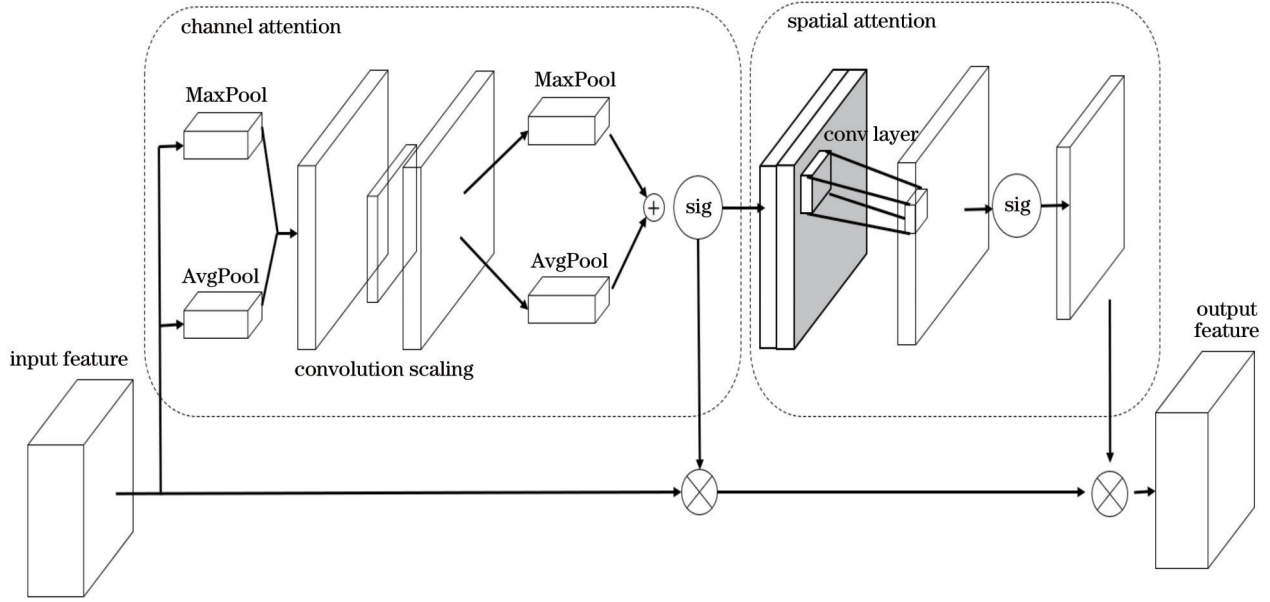


图2 CBAM结构图

Fig. 2 Architecture of CBAM

2.3 密集级联层

级联网络是一个自动训练和自动添加隐含单元的网络,其网络具有学习速率高、网络神经元与深度自定义、反向传播有效化等优点。受到DensenNet^[25]启发,本文设计了这种DCB结构,如图3所示,其中 $X_0 \sim X_4$ 分

别表示网络中残差连接的节点位置。网络在前向传播时,将每一模块的输出作为下一模块和下一个模块的输入,这种密集连接使得参数在各个卷积层之间相互影响、相互迭代,再通过最后的广义平均池化获取到图像的全局感受野,使得网络能够更好地学习到图像信息。

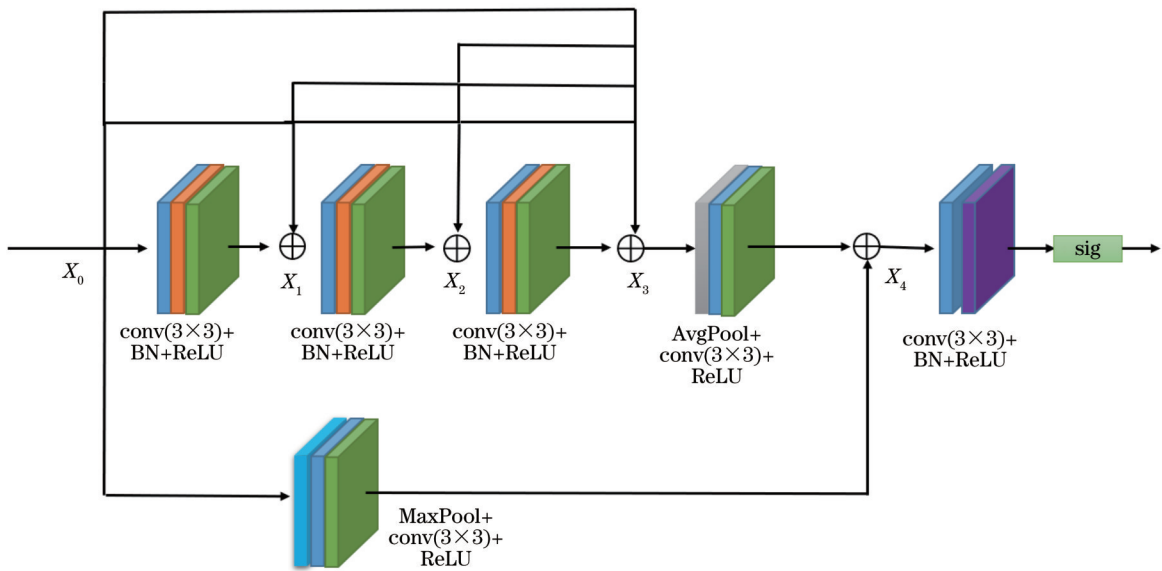


图3 DCB结构图

Fig. 3 Architecture of DCB

3 去噪机制

3.1 去噪方法

本节介绍了设计的去噪网络 MALNet, 如图 4 所示。MALNet 主要由多尺度非规则卷积核、注意力模块、级联块 DCB 组成。所提网络采用端对端的设计结构, 即输入为噪声图, 输出为清洁图。去噪机制为先利

用非对称卷积核和注意力模块获取带噪声图像的基础特征, 再经过级联块 (DCB) 自主学习图像的残差信息, 最后反馈到最后一层的输出特征中, 去噪后重建图像。MALNet 网络参数如表 1。其中, $(1, 0)$ 表示以行的方式来填充, $(0, 1)$ 表示以列的方式来填充, X_1-X_2 表示 X_1 和 X_2 中间位置; $[F(0)-F(1)]_1$ 表示 $F(0)$ 与 $F(1)$ 间的网络层数为 1。

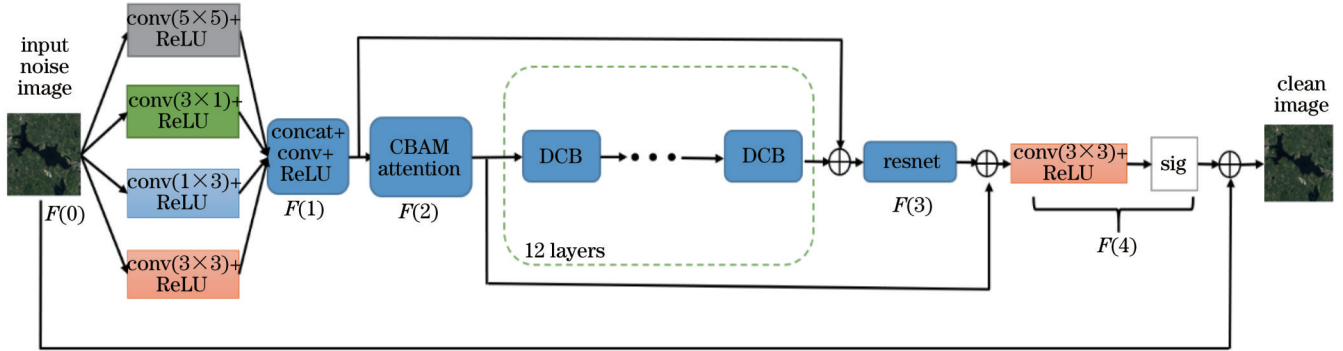


图 4 MALNet 结构

Fig. 4 Architecture of MALNet

表 1 MALNet 结构参数

Table 1 Structural parameters of MALNet

Local	Kernel size	Stride	Padding	Channel	Pooling
$[F(0)-F(1)]_1$	5×5	1	2	64	
$[F(0)-F(1)]_2$	3×1	1	$(1, 0)$	64	
$[F(0)-F(1)]_3$	1×3	1	$(0, 1)$	64	
$[F(0)-F(1)]_4$	3×3	1	1	64	
$F(1)$	3×3	1	1	32	
$F(2)$				32	
X_0	3×3	1	1	32	
X_1-X_2	3×3	1	1	32	
X_2-X_3	3×3	1	1	32	
X_3-X_4	3×3	1	1	32	AvgPool(2×2)
X_0-X_4	3×3	1	1	32	MaxPool(3×3)
$F(3)$				64	
$F(4)$	3×3	1	1	1	

3.2 损失函数

由于 SAR 图像的特殊性, 其噪声图像大多是乘性噪声, 故对所提模型训练的数据集需要进行加噪预处理。根据方程 $y(a, b) = n(\sigma, \delta) \times x(a, b)$ 对图像进行加噪, 其中 $y(a, b)$ 表示输入噪声图像, $x(a, b)$ 表示未加噪的图像, $n(\sigma, \delta)$ 表示方差为 δ 、均值为 σ 的噪声。输入噪声图像经过 MALNet 后输出残差图像 (Rim), 表达为 $\text{MALNet}[y(a, b)] = \text{Rim}(a, b) = y(a, b) - n(\sigma, \delta)$ 。随后, 使用给定的均方误差 (MSE) 来训练去噪网络, 其损失函数的表达式为

$$L(w, b) = \frac{1}{N} \left\{ \text{MALNet}[y(a, b)] - n(\sigma, \delta) \right\}^2, \quad (2)$$

式中: w 和 b 是网络中学习到的权重和偏置。

3.3 评价指标

进行定性和定量实验以评估和演示所提 MALNet 模型在去噪方面的性能。将 WNNM、SAR-BM3D 和 SAR-CNN 三种算法与所提方法进行了比较。通过肉眼观察去噪后图像的清晰和完整状况。为了公平比较, 使用三种算法文献中作者提供的默认设置去对比, 并使用峰值信噪比 (PSNR)、结构相似指数 (SSIM) 和图像熵作为客观评价指标来衡量其他去噪方法和所提模型的去噪效果。假设原始干净的图像 X 的大小为

$N \times M$, Y 为去噪图像。PSNR 的计算公式为

$$R_{\text{PSN}} = 10 \lg \left\{ \frac{NMf_{\max}^2}{\sum_{i=1}^N \sum_{j=1}^M [X(i,j) - Y(i,j)]^2} \right\}, \quad (3)$$

式中: f_{\max} 为输入图像的最大强度, 对于一些 8 位灰度图像, 有 256 个可能的灰度值, 故 $f_{\max} = 256$ 。PSNR 可衡量模型的去噪效果。此外, 还可以使用其他指标来评价原图与去噪图像之间的差别。SSIM 是一种质量评估指标, 可用来衡量两幅图像之间的相似性。假设 x 和 y 是计算 SSIM 的两个非负图像信号, SSIM 的计算公式为

$$M_{\text{ssi}}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (4)$$

表 2 实验平台参数

Table 2 Parameters of experimental platform

Hardware	Description
CPU	12th Gen intel(R) Core(TM) i9-12900KF @ 3.19 GHz
GPU	NVIDIA GeForce RTX 3090
Memory	32 GB
System	Windows 10
Video memory	24 GB
Language framework	Anconda + Python3.8.13 + Pytorch1.11.0

4.2 数据集

所提 MALNet 网络需要噪声图像和对应的无噪声图像作为训练集, 由于 SAR 这样的图像数据集难以获取, 选用的数据集是西北工业大学的 NWPU-RESISC45 Dataset^[26], 故需要手工合成图像数据集。合成数据集的过程为: 1) 选取 NWPU-RESISC45 Dataset^[26] 中 1000 张像素为 256 pixel \times 256 pixel 的图像作为预选数据集, 该数据集包括了飞机场、海岸线、雨林、梯田、桥梁、船只、公路、陆地、雪山和山地每类图像各 100 张, 将其作为训练集; 2) 选取每类图像各 20 张, 共 200 张作为验证集; 3) 训练时将均值为 0、方差为 σ 的高斯噪声以乘性的方式加到 1200 张干净图像中, 完成噪声图像的数据集制作。对比实验包括 4 组不同方差 ($\sigma = 20, \sigma = 30, \sigma = 40, \sigma = 50$) 噪声水平的图像, 验证数据集使用与训练数据集相同的加噪方法。

4.3 实验过程

实验中使用了 1000 张 SAR 清洁图像, 训练初将每张像素为 256 pixel \times 256 pixel 的图像以步长为 10 pixel 滑动平移的方式切割成为 466 张像素为 40 pixel \times 40 pixel 的窗格图像, 并使用翻转、平移、缩放、旋转、镜像来作为数据增强的方式, 最后将图像作为输入来训练 MALNet 网络。网络训练的 epoch 设置为 150, BN 中设置动量为 0.95, batchsize 设置为 64, 使

式中: μ_x 和 μ_y 分别为图像 x 和图像 y 的平均强度; σ_x 和 σ_y 分别为图像 x 和图像 y 的标准差; σ_{xy} 是图像 x 和图像 y 的协方差; c_1 和 c_2 是常数值。局部参数 ($\mu_x, \mu_y, \sigma_x, \sigma_y, \sigma_{xy}$) 在一个 8×8 的方形窗口内, 并以像素平移的方式计算。PSNR 与 SSIM 是评价图像优劣的基本指标, 图像熵可作为衡量图像中信息能量的计算方法, 其计算公式为

$$H = \sum_{p=0}^{255} \frac{f(p, q)}{S^2} \lg \frac{f(p, q)}{S^2}, \quad (5)$$

式中: p 为像素灰度值; q 为邻域灰度值; $f(p, q)$ 为灰度空间出现的频次; S 为图像尺寸。

4 实验结果分析

4.1 实验平台

实验平台搭建的相关参数如表 2 所示。

用 Adam^[27] 优化器来优化网络权重系数, 引入动态调整学习率 (lr), 设置其初值为 0.01, 后续依次分别在 epoch 为 15、55、115 时乘以 0.3 来调整学习率的大小, 其他情况均为默认设置, 整个模型经过 150 轮训练, 数据迭代 1092188 次后收敛, 其损失曲线如图 5 所示。测试时网络能适应任意大小的噪声图像, 所提 MALNet 使用 Python 来对图像进行高斯噪声乘性处理, 并使用 Pytorch 框架来实现模型算法。

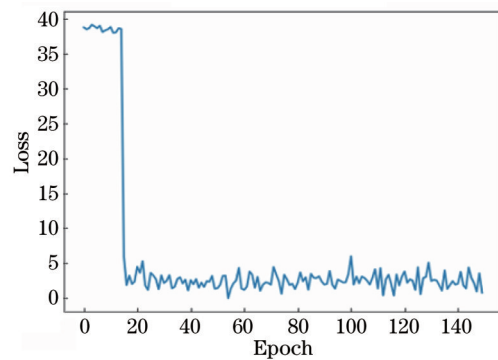


图 5 训练模型的损失曲线

Fig. 5 Loss curve of training model

4.4 密集级联层深度实验

为了实现 MALNet 最好的去噪效果, 针对不同深度的 DCB 进行实验, 即消融实验。对单一背景类型图

像下不同 DCB 层数的网络与多背景类型图像下不同 DCB 层数的网络分别进行了训练和测试,前者为纯训练模式,后者为混合训练模式。两种模式各训练 150 轮,取噪声水平为 $\sigma = 40$,其余设置均不变,纯训练模式下选用了飞机图像进行训练,混合模式下选用 10 种类型的图像进行训练,最后在纯训练模式下和混合训练模式下分别对飞机图像进行同一张图片

的 10 次验证并计算 10 次的平均 PSNR 值,如图 6 所示。可以发现:在纯训练模式和混合训练模式下的模型中,DCB 层数为 12 时的 PSNR 值分别为 29.3 dB 和 28.5 dB,均达到最高。虽然纯训练模式下的最高值比混合训练模式高出 0.8 dB,但是两者均在 12 层 DCB 时取得最高值,故 12 层仍是所提网络中最佳的 DCB 层数。

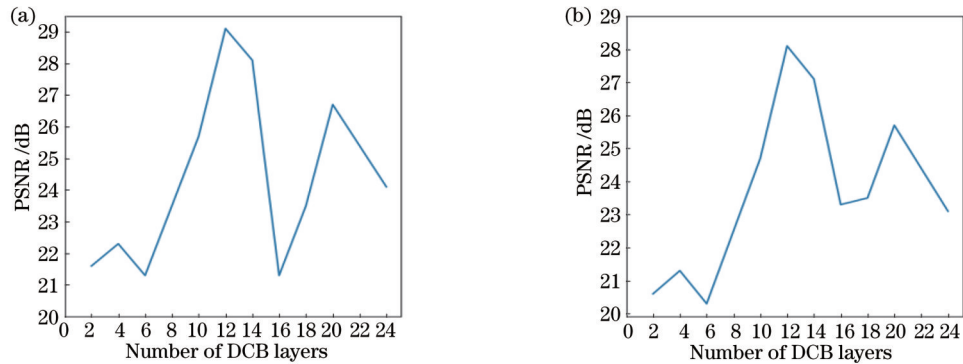


图 6 不同模式下不同 DCB 层的 PSNR 值。(a)单一背景图像;(b)多背景图像

Fig. 6 PSNR values of different DCB layers under different modes. (a) Single background image; (b) multiple background image

4.5 不同背景图像下实验分析

在 4.4 节的条件下,由图 6 可知:当整个网络处于纯训练模式(单一背景类型图像)下时,PSNR 值能达到 29.3 dB;当整个网络处于混合训练模式(多背景类型图像)下时,PSNR 值为 28.5 dB。可以发现,当网络只考虑单一类型图像时确实能有效提高网络的去噪性能,但在实际应用中 SAR 噪声图像复杂多样,故纯训练模式下无法保证整个算法能有较好的泛化能力和鲁棒性。因此,多类型图像训练在实际应用中较为适用,并且其去噪性能依然较好。

4.6 实验对比分析

对比了三个去噪算法,选取了飞机、山体和海岸作为验证图像,飞机图像去噪效果图、海岸图像去噪效果图和山体图像去噪效果图如图 7~9 所示,展示了当前噪声等级为 $\sigma = 30$ 时不同算法去噪结果的视觉效果对比。图 7~9 中 6 幅图依次为无噪声图像、噪声图像、传统方法 WNNM 的去噪效果图、SAR-BM3D 的去噪效果图、SAR-CNN 的去噪效果图和所提 MALNet 的去噪效果图。WNNM 的去噪效果图中有很多没有去除干净的瑕疵,并且纹理丢失相当严重,视觉效果模糊。

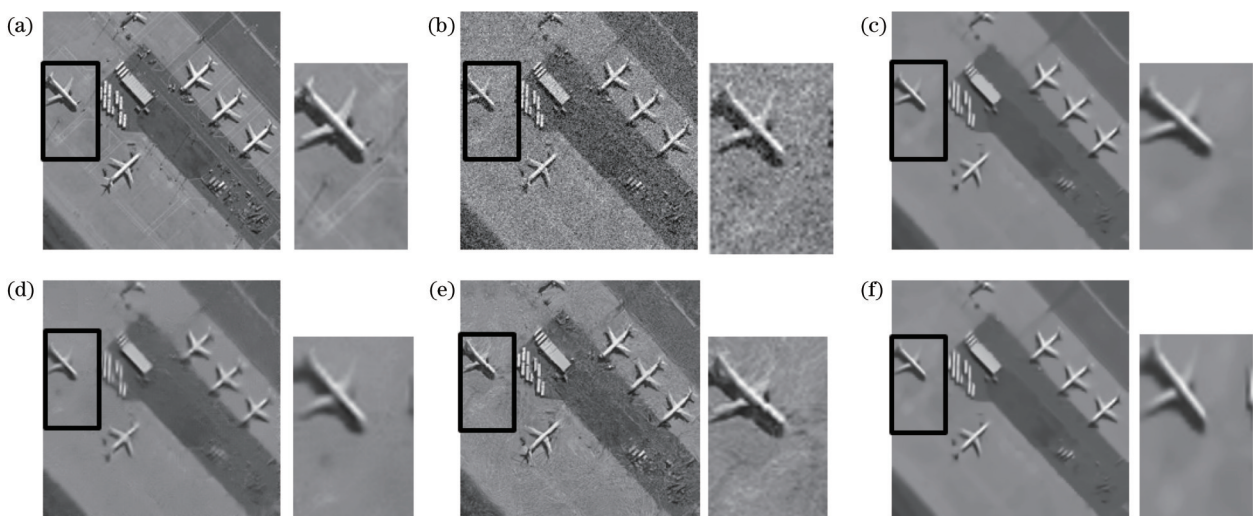


图 7 飞机场图像去噪效果对比。(a)原图;(b)带噪声图像;(c) WNNM 去噪图像;(d) SAR-BM3D 去噪图像;(e) SAR-CNN 去噪图像;(f) MALNet 去噪图像

Fig. 7 Denoising effect comparison of airport image. (a) Original image; (b) noisy image; (c) denoised image obtained by WNNM; (d) denoised image obtained by SAR-BM3D; (e) denoised image obtained by SAR-CNN; (f) denoised image obtained by MALNet

SAR-BM3D的去噪效果图中保留了一些细节,但是飞机机身很模糊,机尾部分的大部分边缘信息被抹去,海岸图像中海浪边缘信息丢失得较严重。SAR-CNN的去噪效果图中飞机跑道边缘信息没有被很好地保留,并且旁边的摆放物体比较模糊,虽然飞机机翼恢复出来了,但是最下方的飞机整体仍然与参考图像相差甚远,所恢复出来的小物件是模糊的。在另外两种场景

的对比图中,SAR-CNN的去噪效果图中噪点较多,一些纹理线条褶皱得较严重,画面撕裂感很强,故其去噪效果一般。所提MALNet在恢复图像细节和结构方面展现出较好的效果,海岸中的海浪边缘被有效保留,山体的棱角较为清晰,噪点去除效果较好,获得了比其他方法更好的视觉效果。

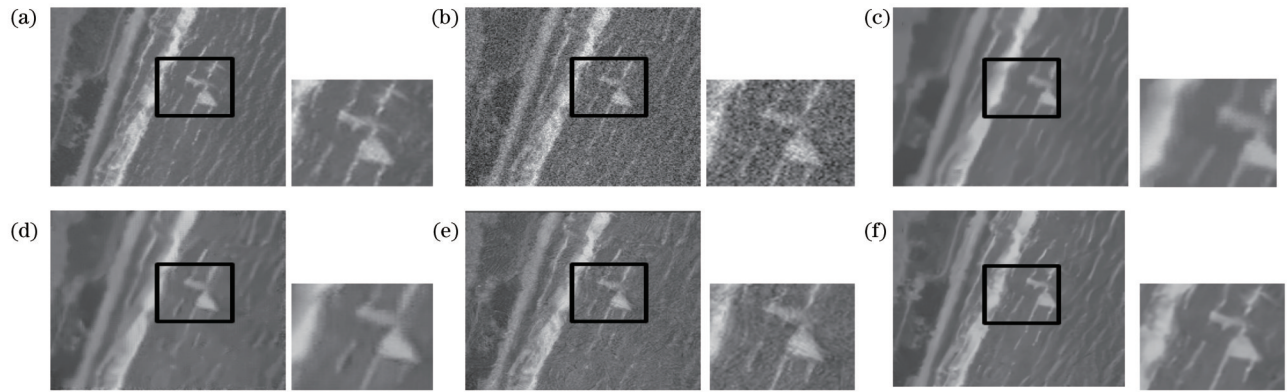


图 8 海岸图像去噪效果对比。(a)原图;(b)带噪声图像;(c)WNNM去噪图像;(d)SAR-BM3D去噪图像;(e)SAR-CNN去噪图像;(f)MALNet去噪图像

Fig. 8 Denoising effect comparison of coast image. (a) Original image; (b) noisy image; (c) denoised image obtained by WNNM; (d) denoised image obtained by SAR-BM3D; (e) denoised image obtained by SAR-CNN; (f) denoised image obtained by MALNet

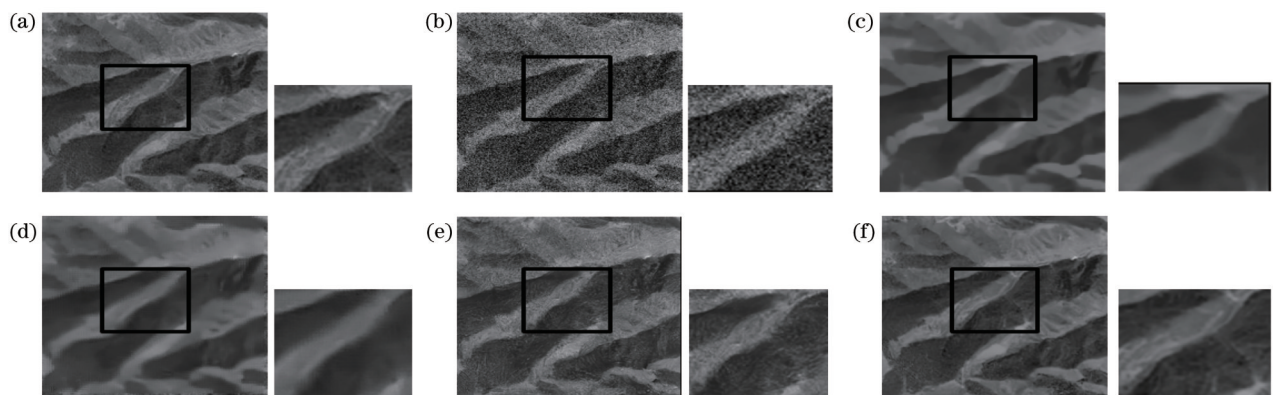


图 9 山体图像去噪效果对比。(a)原图;(b)带噪声图像;(c)WNNM去噪图像;(d)SAR-BM3D去噪图像;(e)SAR-CNN去噪图像;(f)MALNet去噪图像

Fig. 9 Denoising effect comparison of mountain image. (a) Original image; (b) noisy image; (c) denoised image obtained by WNNM; (d) denoised image obtained by SAR-BM3D; (e) denoised image obtained by SAR-CNN; (f) denoised image obtained by MALNet

实验的主要目的是测试所提MALNet在数据集上的实验性能。表3和表4为不同方法在不同图像上的PSNR和SSIM结果,表3和表4中的行分别为噪声参数为 $\sigma=20$ 、 $\sigma=30$ 、 $\sigma=40$ 、 $\sigma=45$ 、 $\sigma=50$ 时不同算法所产生的PSNR和SSIM数值,列分别为同一方法在不同噪声水平下的PSNR和SSIM数值,所包含的方法有SAR-BM3D模型、SAR-CNN模型、WNNM模型和所提MALNet。从表3可以看出:所提MALNet的平均PSNR值比SAR-BM3D高出约9.25 dB,比SAR-CNN高出约0.75 dB,比WNNM高出约14.45 dB;除

了噪声水平在 $\sigma=50$ 时,MALNet的平均PSNR值比SAR-CNN少了0.01dB,其余情况下所提模型在每个噪声水平上得到的PSNR值均比其他三种方法高,尤其是当噪声参数为20时,所提方法的PSNR比SAR-CNN算法最大高出2.56 dB。从表4可以看出,MALNet的SSIM大部分都是所对比方法中的最高值,只有在噪声参数为50时,略低于SAR-CNN,并且所得到的平均SSIM也是最高的。因此,综合PSNR和SSIM两个客观评价指标可知,所提网络的去噪性能是优于所对比方法的。

在表 5 中,每行表示同一张图像在同一噪声水平下不同算法的图像熵值。人工合成的乘性噪声具有特殊性,噪声越强,图像熵越大,衡量去噪效果的好坏以低图像熵为基本目标。计算得出三类图像在 5 种不同噪声水平下各类算法去噪图像的平均信息熵, WNNM 为 7.113492 bit/pixel, SAR-BM3D 为 6.842258 bit/pixel, SAR-CNN 为 7.499375 bit/pixel, MALNet 为 6.691700 bit/pixel。所提算法相较于 WNNM、SAR-BM3D、SAR-CNN 分别优出约 0.42179 bit/pixel、0.15056 bit/pixel 和 0.80768 bit/pixel。只有在山体图像上,当噪声参数为 50 时,所提算法略逊于 SAR-BM3D。

表 3 各类型 SAR 图像在不同噪声水平下各算法的去噪水平 (PSNR)

Table 3 Denoising level (PSNR) of each algorithm for each type of SAR image under different noise levels unit: dB

Image	σ	WNNM	SAR-BM3D	SAR-CNN	MALNet
Airport	20	28.87	29.28	31.85	33.42
	30	26.62	27.13	30.14	31.11
	40	26.35	26.52	29.86	29.97
	45	25.28	25.93	27.77	29.31
	50	25.17	24.01	27.67	27.82
Mountain	20	28.15	29.13	31.10	33.76
	30	26.53	27.65	30.03	31.17
	40	25.72	26.04	28.53	29.59
	45	25.14	25.99	27.17	27.85
	50	24.05	25.92	27.31	26.62
Coast	20	28.22	30.21	31.01	33.83
	30	26.63	28.52	29.43	31.58
	40	26.55	27.73	28.59	29.98
	45	25.39	26.24	26.87	28.93
	50	24.37	25.55	26.12	26.11

5 结 论

针对合成孔径雷达图像中的噪声,提出了一种新的去噪模型 MALNet。该模型采用端到端架构,不需要单独的子网或手动干预。所提解决方案包括三个模块,即多尺度非对称卷积模块、基于注意机制的特征提取模块和基于密集级联网络的特征增强模块。模型还加入了批量归一化和全局平均池化以提升自适应性,无需大量的数据集即可完成收敛,图像数据在训练 150 轮后可完成收敛,训练效率方面表现突出,可移植性较好。另外,去噪结果表明,相比于其他去噪算法, MALNet 不仅能够产生比较好的去噪效果,还具有很好的鲁棒性,网络在降噪与保留细节上实现了良好的平衡性。所提算法在不同噪声水平图像上的去噪效果进一步表明, MALNet 可以提供感知上令人满意的去

表 4 各类型 SAR 图像在不同噪声水平下各算法的去噪水平 (SSIM)

Table 4 Denoising level (SSIM) of each algorithm for each type of SAR image under different noise levels

Image	σ	WNNM	SAR-BM3D	SAR-CNN	MALNet
Airport	20	0.7929	0.8324	0.8923	0.9257
	30	0.7538	0.7826	0.8621	0.8969
	40	0.7063	0.7505	0.8669	0.8671
	45	0.6818	0.6867	0.7926	0.8309
	50	0.6821	0.6027	0.7818	0.7892
Mountain	20	0.7867	0.8354	0.9012	0.9354
	30	0.7128	0.7891	0.9010	0.9194
	40	0.6957	0.7497	0.8725	0.8963
	45	0.6708	0.7134	0.7971	0.8361
	50	0.6684	0.6808	0.8211	0.7655
Coast	20	0.7911	0.8561	0.9036	0.9253
	30	0.7187	0.7899	0.8947	0.8987
	40	0.7064	0.7587	0.8809	0.8827
	45	0.6817	0.7308	0.7437	0.8817
	50	0.6512	0.6915	0.7018	0.7939

表 5 各类型 SAR 图像在不同的噪声水平下各算法的图像熵

Table 5 Image entropy of each algorithm for each type of SAR image under different noise levels

unit: bit·pixel⁻¹

Image	σ	WNNM	SAR-BM3D	SAR-CNN	MALNet
Airport	20	6.5312	6.3112	6.8529	6.1879
	30	6.8254	6.7709	7.2964	6.3542
	40	7.1596	6.9606	7.5976	6.8793
	45	7.3177	7.2318	7.2151	6.9074
	50	7.8325	7.6321	7.4647	7.5258
Mountain	20	6.4219	6.3864	7.0158	6.0872
	30	6.8297	6.5913	7.2479	6.3173
	40	7.3541	7.1291	7.8931	6.7966
	45	7.5627	7.3945	7.9315	6.9037
	50	7.9861	7.4837	8.0147	7.5813
Coastal	20	6.5207	6.2186	6.9615	6.2157
	30	6.7896	6.4517	7.3541	6.7959
	40	7.1584	6.9526	7.7987	6.8394
	45	7.6983	7.0118	7.8424	7.0288
	50	7.9527	7.3189	7.8951	7.8198

噪结果,并且在 PSNR、SSIM 和图像熵指标上都优于其他同类算法。MALNet 的灵活性、高效性和有效性为 SAR 图像去噪提供了一种新的解决方案。

参 考 文 献

[1] 保铮,孙长印,邢孟道.机动目标的逆合成孔径雷达成像原理与算法[J].电子学报,2000,28(6):24-28,39.

- Bao Z, Sun C Y, Xing M D. Principle and algorithm of inverse synthetic aperture radar imaging for maneuvering targets[J]. *Acta Electronica Sinica*, 2000, 28(6): 24-28, 39.
- [2] Najafi A, Hasanlou M, Akbari V. Land cover changes detection in polarimetric SAR data using algebra, similarity and distance based methods[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, XLII-4/W4: 195-200.
- [3] Nair R R, David E, Rajagopal S. A robust anisotropic diffusion filter with low arithmetic complexity for images[J]. *EURASIP Journal on Image and Video Processing*, 2019, 2019(1): 48.
- [4] 陈鑫祥, 蒲冰鑫, 俞建, 等. 基于多视图图像点云的建筑物屋顶分割算法研究[J]. *地理空间信息*, 2022, 20(5): 33-37.
- Chen X X, Pu B X, Yu J, et al. Research on building roof segmentation algorithm from multiple view stereo point clouds [J]. *Geospatial Information*, 2022, 20(5): 33-37.
- [5] Donoho D L. De-noising by soft-thresholding[J]. *IEEE Transactions on Information Theory*, 1995, 41(3): 613-627.
- [6] Donoho D L, Johnstone I M. Adapting to unknown smoothness via wavelet shrinkage[J]. *Journal of the American Statistical Association*, 1995, 90(432): 1200-1224.
- [7] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 20-25, 2005, San Diego, CA, USA. New York: IEEE Press, 2005: 60-65.
- [8] Parrilli S, Poderico M, Angelino C V, et al. A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(2): 606-616.
- [9] MuraliMohanBabu Y, Subramanyam M V, GiriPrasad M N. A modified BM3D algorithm for SAR image despeckling[J]. *Procedia Computer Science*, 2015, 70: 69-75.
- [10] Foucher S. SAR image filtering via learned dictionaries and sparse representations[C]//IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium, July 7-11, 2008, Boston, MA, USA. New York: IEEE Press, 2008: 229-232.
- [11] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D*, 1992, 60: 259-268.
- [12] 丁红波, 王珍珠, 刘东. 激光雷达信号去噪方法的对比研究[J]. *光学学报*, 2021, 41(24): 2401001.
- Ding H B, Wang Z Z, Liu D. Comparison of de-noising methods of LiDAR signal[J]. *Acta Optica Sinica*, 2021, 41(24): 2401001.
- [13] 刘吉, 黄晓慧, 武锦辉, 等. 基于正余弦分解的自适应全变分散斑去噪方法[J]. *中国激光*, 2020, 47(10): 1004004.
- Liu J, Huang X H, Wu J H, et al. Adaptive total variation speckle denoising method based on sine-cosine decomposition[J]. *Chinese Journal of Lasers*, 2020, 47(10): 1004004.
- [14] Gu S H, Zhang L, Zuo W M, et al. Weighted nuclear norm minimization with application to image denoising[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 2862-2869.
- [15] 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.
- [16] Yue Z S, Zhao Q, Zhang L, et al. Dual adversarial network: toward real-world noise removal and noise generation[M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12355: 41-58.
- [17] Chierchia G, Cozzolino D, Poggi G, et al. SAR image despeckling through convolutional neural networks[C]//2017 IEEE International Geoscience and Remote Sensing Symposium, July 23-28, 2017, Fort Worth, TX, USA. New York: IEEE Press, 2017: 5438-5441.
- [18] Thakur R K, Maji S K. AGSDNet: attention and gradient-based SAR denoising network[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4506805.
- [19] Tian C W, Xu Y, Zuo W M, et al. Asymmetric CNN for image superresolution[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 52(6): 3718-3730.
- [20] Jin J, Dundar A, Culurciello E. Flattened convolutional neural networks for feedforward acceleration[EB/OL]. (2014-12-17) [2022-02-04]. <https://arxiv.org/abs/1412.5474>.
- [21] Denton E, Zaremba W, Bruna J, et al. Exploiting linear structure within convolutional networks for efficient evaluation [C]//Advances in Neural Information Processing Systems 27, December 8-13, 2014, Montreal, Quebec, Canada. Cambridge: The MIT Press, 2014: 1269-1277.
- [22] Jaderberg M, Vedaldi A, Zisserman A. Speeding up convolutional neural networks with low rank expansions [EB/OL]. (2014-05-15) [2021-02-04]. <https://arxiv.org/abs/1405.3866>.
- [23] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [24] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[EB/OL]. (2018-07-17) [2022-02-04]. <https://arxiv.org/abs/1807.06521>.
- [25] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [26] Cheng G, Han J W, Lu X Q. Remote sensing image scene classification: benchmark and state of the art[J]. *Proceedings of the IEEE*, 2017, 105(10): 1865-1883.
- [27] Kingma D P, Ba J. Adam: a method for stochastic optimization [EB/OL]. (2014-12-22) [2022-02-04]. <https://arxiv.org/abs/1412.6980>.

Synthetic Aperture Radar Image Denoising Algorithm Based on Deep Learning

Fu Xiangwei¹, Shan Huilin^{1,2*}, Lü Zongkui¹, Wang Xingtao²

¹*School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, Jiangsu, China;*

²*School of Electronic & Information Engineering, Wuxi University, Wuxi 214105, Jiangsu, China*

Abstract

Objective Synthetic aperture radar (SAR) is a kind of sensor to capture microwaves. Its principle is to establish images through the reflection of waveforms, so as to solve the problem that traditional optical remote sensing radars are affected by weather, air impurities, and other environmental factors when collecting images. The most widely used SAR is change detection (CD). CD refers to the dynamic acquisition of image information from a certain target, which includes three steps: image preprocessing, generation of difference maps, and analysis and calculation of difference maps. It is applied to the estimation of natural disasters, management and allocation of resources, and measurement of land topographic characteristics. However, in the process of CD, the inherent speckle noise in SAR images will reduce the performance of CD. Therefore, the image denoising method has become a basic method of preprocessing in CD. How to restore a clean image from a noisy SAR image is an urgent problem to be solved.

Methods Traditional denoising algorithms of SAR images generally use the global denoising idea whose principle is to use the global similar information in images to perform processing and judgment. In the case of the high resolution of images, these algorithms need a series of preprocessing such as smoothing and then complete pixel distinction through the neighborhood processing of each image block. The algorithms usually occupy huge computing resources and have certain spatial and temporal limitations in practical applications. In addition, they cannot efficiently complete the denoising task. In terms of deep learning, some algorithms perform well, but there is still room for improvement in network convergence speed, model redundancy, and accuracy. To solve these problems, this paper proposes a denoising algorithm based on a multi-scale attention cascade convolutional neural network (MALNet). The network mainly uses the idea of multi-scale irregular convolution kernel and attention. Compared with a single convolution kernel, a multi-scale irregular convolution kernel has an excellent image receptive field. In other words, it can collect image information from different scales to extract more detailed image features. Subsequently, the convolution kernels of different scales are concat layers in the network, and an attention mechanism is introduced into a concat feature map to divide the attention of the features so that the whole model has a positive enhancement ability for the main features of the image. In the middle of the network, the dense cascade layer is used to further strengthen the features. Finally, the image restoration and reconstruction are realized by network subtraction.

Results and Discussions In this paper, qualitative and quantitative experiments are carried out to evaluate and demonstrate the performance of the proposed MALNet model in denoising. The WNNM, SAR-BM3D, and SAR-CNN algorithms are compared with our proposed method. The clear state and complete signs of the denoised images are visually observed. In order to make a fair comparison, we use the default settings of the three algorithms provided by the authors in the literatures. Peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and image entropy are used as objective evaluation indexes. The PSNR, SSIM, and image entropy are calculated as error metrics.

Three denoising algorithms are compared, and airport, mountain, and coast are selected as verification images. The denoising effects of airport images (Fig. 7), coast images (Fig. 8), and mountain images (Fig. 9) are analyzed. It shows the visual effect comparison of denoising results of different algorithms. In Figs. 7–9, 6 figures are successively noise-free image, noise image, denoised image obtained by WNNM, denoised image obtained by SAR-BM3D, denoised image obtained by SAR-CNN and denoised image obtained by MALNet. It is obvious that the WNNM denoised image has many defects that are not removed completely, and the texture loss is quite serious. SAR-BM3D denoised image retains some details, but the aircraft fuselage is very vague, and the tail part has gotten most of the edge information erased. Although the aircraft wing in the SAR-CNN denoised image is recovered, the whole aircraft at the bottom is still far from the reference image, and the recovered small objects are blurred.

It can be seen from Table 3 that the average PSNR value of the proposed MALNet is about 9.25 dB higher than that of SAR-BM3D, about 0.75 dB higher than that of SAR-CNN, and about 14.45 dB higher than that of WNNM. Moreover, in terms of the noise level, MALNet is 0.01 dB less than that SAR-CNN. The PSNR value of the proposed

MALNet model at each noise level is higher than that of other algorithms. Especially, when the noise parameter is 20, the proposed method is 2.56 dB higher than that of the SAR-CNN algorithm. In terms of structural similarity (Table 4), it can be seen that the SSIM of MALNet is mostly the highest among all methods. Only when the noise parameter is 50, it is slightly lower than that of SAR-CNN, but the average SSIM is still the highest. The average information entropy of the denoised images by the four algorithms is 7.113492 bit/pixel for WNNM, 6.842258 bit/pixel for SAR-BM3D, 7.499375 bit/pixel for SAR-CNN, and 6.6917 bit/pixel for MALNet. The proposed algorithm outperforms WNNM, SAR-BM3D, and SAR-CNN by 0.42179 bit/pixel, 0.15056 bit/pixel, and 0.80768 bit/pixel, respectively. Therefore, in terms of the three objective evaluation indexes of PSNR, SSIM, and image entropy, the proposed network in this paper has better denoising performance than other comparison methods.

Conclusions In this paper, a new denoising model MALNet is proposed for solving the noise in SAR images. This model uses an end-to-end architecture and does not require separate subnets or manual intervention. The solution includes three modules, i. e., multi-scale irregular convolution module, feature extraction module based on attention mechanism, and feature enhancement module based on dense cascade network. The model also adds batch normalization and global average pooling to improve its adaptability. It can complete convergence without massive data sets. The image data can complete convergence after 150 rounds of training. The training efficiency is outstanding, and the portability is positive. The experimental results show that compared with those of other traditional image denoising algorithms, the PSNR and SSIM of the proposed algorithm are improved by 0.75 dB–14.45 dB and 0.01–0.16, respectively. The proposed algorithm is superior to other algorithms in image entropy and can better recover the details of images.

Key words image processing; synthetic aperture radar; convolutional neural network; image denoising; multi-scale attention