

基于双通道特征融合编解码网络的极化 SAR 图像分类^{*}

王蒙蒙^{1,2} 刘秀清¹ 张 衡¹ 王春乐¹ 贾小雪¹

(1. 中国科学院空天信息创新研究院 北京 100190; 2. 中国科学院大学电子电气与通信工程学院 北京 100049)

摘 要: 图像分类是极化合成孔径雷达(PolSAR)图像解译的关键。传统卷积神经网络(CNN)逐像素的分类,造成卷积的重复计算。PolSAR 图像存在丰富的信息,包括极化相干信息与极化分解信息,因此如何融合信息实现高效分类至关重要。基于极化散射特征分析,以 U_net 网络模型为基础,提出了双通道特征融合编解码网络,该网络使用注意力机制特征融合模块将极化相干信息和极化分解特征整合到语义分割框架中,辅助深度 CNN 分类器训练,实现高精度像素级的标记,同时加入空间金字塔结构有效的提取多尺度特征。该网络结构避免了逐像素切片重复计算,有效提升计算效率。利用 AIRSAR 获取的旧金山地区 PolSAR 数据和海南博鳌地区机载 PolSAR 数据进行试验研究,试验结果两个地区总体分类精度(OA)分别达到 97.11% 与 99.97%,验证了提出的分类方法的有效性与较好的应用价值。

关键词: PolSAR; 特征融合; 地物分类; 编解码; 深度学习

中图分类号: TP753 **文献标识码:** A **国家标准学科分类代码:** 510.4050

PolSAR image classification based on dual-channel feature fusion encoder-decoder network

Wang Mengmeng^{1,2} Liu Xiuqing¹ Zhang Heng¹ Wang Chunle¹ Jia Xiaoxue¹

(1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Image classification is an important application of polarized synthetic aperture radar (PolSAR) image interpretation. Traditional convolutional neural network (CNN) classifies pixel-by-pixel, resulting in repeated computation of convolution. PolSAR images have rich information, including polarization coherence feature information and polarization decomposition feature information, so it is crucial to fuse the information to achieve efficient classification. This paper proposes dual-channel feature fusion encoder-decoder network based on polarization scattering feature information analysis and the U_net network model. The network uses an attention feature fusion mechanism to integrate polarization coherence feature information and polarization decomposition feature information into a semantic segmentation framework, assists deep CNN classifier training to achieve high precision pixel-level labeling, and incorporates a spatial pyramid structure to efficiently extract multi-scale features. The network structure avoids the repetitive computation of pixel-by-pixel classification and effectively improves the computational efficiency. We use the PolSAR data acquired by AIRSAR in San Francisco area and the airborne PolSAR data in Boao area of Hainan for experimental study, and the experimental results show that the overall accuracy (OA) of the two areas is 97.11% and 99.97% respectively, which verifies the effectiveness and better application value of the classification method.

Keywords: PolSAR; feature fusion; classification; encoder-decoder; deep learning

收稿日期: 2022-09-20

^{*} 基金项目: 国家自然科学基金(61901445)、北京市自然科学基金(4192065)项目资助

0 引言

极化合成孔径雷达(PolSAR)^[1-2]系统通过发射和接收不同组合形式的极化波获取回波数据,提供不同极化方式(HH、HV、VH、VV)丰富的目标后向散射信息,为SAR图像解译提供了充足的数据基础^[3]。

近几十年来,极化SAR地物分类任务得到学者持续广泛的关注,涌现出大量优秀的地物分类算法。地物分类算法可以大致分为监督分类和无监督分类,如文献[4]使用的Wishart无监督分类方法,该方法是一个可广泛用于极化SAR影像协方差矩阵的无监督分类器,这类方法的缺陷是迭代分类后各个类别丢失了其物理散射机制类型。文献[5]提出了一个适用于全极化SAR图像非监督分类的多分量散射模型分解-Wishart算法,利用极化目标分解和统计分布理论来挖掘数据中的潜在信息,但是无监督分类效果相对于采用一定数量标签做训练的监督分类效果较差,只适用于简单的地物分类。因此监督分类也被引入PolSAR图像分类领域,Chen等^[6]提出了结合目标散射机制解释和极化特征挖掘的专家知识来辅助深度卷积神经网络(convolutional neural networks,CNN)分类器训练进行PolSAR图像分类,形成了极化特征驱动的深度卷积神经网络。Yang等^[7]基于Resnet和去噪自动编码器(denoising autoencoder,DAE),提出一种基于切片和像素的混合分类方法,结果表明该方法分类结果可以保持良好轮廓及精度,但以上方法为切片预测,切片分类结果为该切片中心像素点的类别,会导致分类效率较低。全卷积神经网络(fully convolutional networks,FCN)^[8]的提出实现了输出与输入相同分辨率的类别预测,它将全连接层改为卷积层,实现了端到端的分类。但以上方法虽然解决了像素到像素的问题,仍存在图像位置信息丢失的问题。与传统的视觉图像不同,极化SAR图像成像机理与光学图像差异较大,PolSAR图像每个像素点结构为复数矩阵,结构更为复杂,因此SAR图像分类比光学图像不同,如何利用好PolSAR的极化相干信息与极化分解信息实现高效分类亟待解决。同时由于PolSAR图像地物尺寸往往差异较大,应高效提取目标的多尺度特征。而且目前极化SAR图像标签数据集较少,因此本文针对PolSAR数据集的特点设计网络。

针对切片式重复计算的问题,基于编解码全卷积神经网络结构。同时借鉴U-net网络^[9]使用跳跃连接方式连接低层与高层的特征图来恢复原始图像结构信息,融入网络较浅的层来进行像素更加精准的定位。为了充分利用PolSAR图像的极化分解信息,引入极化分解特征辅助分类,本文提出了双通道特征融合编解码网络的极化分类算法。其编码器部分由两个网络分支组成,分别从极化相干矩阵和极化目标分解特征中提取特征,为了更高效地融入两路信息,引入注意力机制实现在空间与通道特征过滤冗余信息,以细化PolSAR图像中的重要特征并抑制不必要

的特征,以加权的方式获取两分支的特征融合,实现极化相干信息与极化分解特征的多级融合辅助。同时为了提升网络提升对大尺度地物的分辨能力,在处理高维信息时借鉴deeplabv3+网络^[10]加入经过改进的空间金字塔结构,有效地聚集具有更大感受野的多尺度特征,最终通过解码器一步步上采样实现像素级的分类结果。

1 研究方案

1.1 双通道特征融合网络

U-net网络在下采样层和上采样层之间使用跳跃连接,向网络高层提供图像的低层信息来辅助图像重构。因为经历多次卷积以及下采样操作后得到的深层特征图分辨率较低,在上采样时难以恢复。因此U-net网络共使用4次跳跃连接,而不是仅仅直接在高级语义特征上进行监督和损失反传,这样就保证了最后恢复出来的特征图融合了更多的层次的特征且定位更加精准。U-net网络主要分为编码器模块、解码器模块两部分。

1)编码器模块由5个卷积层构成,采用 3×3 卷积层提取特征,对输出数据进行批归一化处理以加快模型收敛缓解梯度弥散问题。随后采用ReLU激活函数非线性激活,然后通过大小为 2×2 ,步长为2的池化层进行下采样。

2)解码器模块,最底层的特征图先上采样,然后按通道维度与编码阶段对应大小的特征图拼接,同样经过两次 3×3 的卷积层和ReLU激活函数后,反复4次,不断逐层与编码部分传送过来的底层特征融合后上采样,最终通过 1×1 的卷积层与Softmax函数得到像素级的标记图。本文在U-net网络的基础上针对PolSAR图像特征提出特征融合编解码网络。整体网络结构如图1所示,主要改进体现在4个方面。

(1)网络输入改进为双通道,不仅包含了目标后向散射的完整极化信息,还加入了有助于分类的极化分解特征。两分支分别输入极化分解信息与极化相干信息。近年来,随着卷积神经网络的兴起,许多网络被用于极化SAR图像的分类。但是PolSAR系统是相干成像机制,PolSAR图像包含强烈的斑点噪声。斑点噪声会对自动分割算法带来影响,且缺少大量的标记样本来训练有效的深度CNN。因此如何使用较少样本达到较高的分类精度是亟待解决的问题。

引入极化分解特征,可有效提高分类性能^[11]。为了充分挖掘两种信息,充分利用极化相干信息与极化分解信息在场景理解中具有互补性的特点,使其互相补充发挥它们的优势。本文设计具有两路编码层分支的全卷积网络结构,分别处理极化相干信息与极化分解特征信息。融合网络架构的关键是特征融合方式,它用于结合两路编码器分支的特征信息。FuseNet^[12]用特征图逐元素加和方式在不同深度特征层实现逐步融合。但是这种方式存在一定的局限性,因为 T 矩阵元素与极化分解参数量纲可能不同,可能导致部分特征值被忽略。所以为了避免这种误

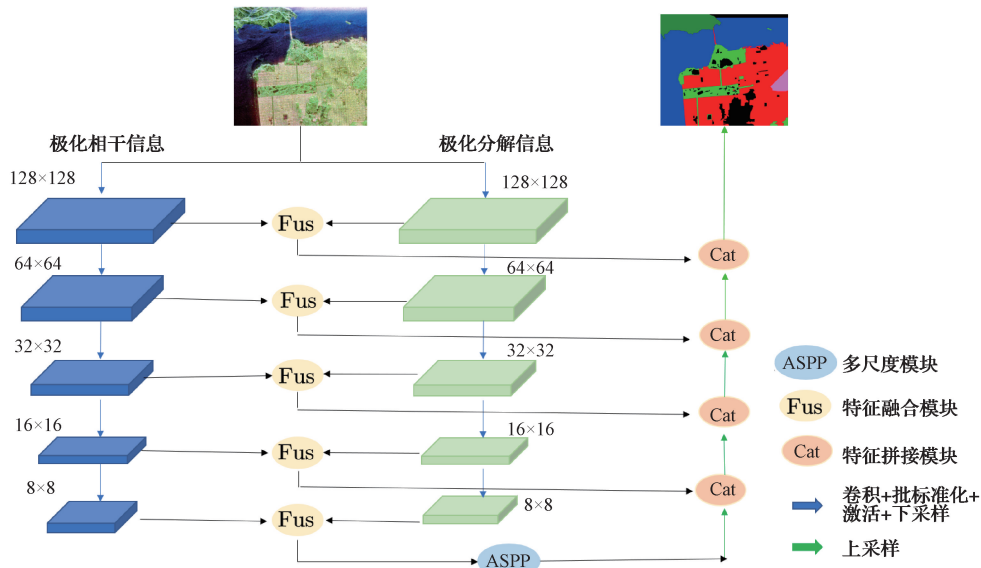


图1 整体网络结构

差,本文对特征融合方式进行改进,特征图逐元素加和改为特征图按照通道维度拼接的方式对不同深度特征层实现逐步融合。具体融合方式如图2(a)所示,两分支特征按通道维度拼接后,经过注意力机制模块,对输入每个部分赋予不同的权重,抽取出来更关键的信息后,经过两次 3×3 的卷积与 ReLu 激活函数进行非线性激活,同时输出通道数为 64,传递到解码模块。

(2)引入轻量的注意力机制(CBAM)^[13],改进融合方式提升分类性能。两条支路进行特征融合后再送入解码部分。虽然效果不错,但会使网络结构变得更加复杂。本文针对这一问题对网络进行改进,增加注意力机制,形成基于注意力机制的特征融合模块。CBAM 同时关注通道和空间维度,以精炼 SAR 图像中的重要特征并抑制不必要的特征。

注意力机制帮助网络对输入每个部分赋予不同的权重,抽取出来更加关键及重要的信息,使模型判断更加准确。CBAM 模块流程如图2(b)所示。通道注意力模块由平均池化和最大池化操作来学习互补特征,SAR 图像的特征在于像素灰度的强度,并且它们包含随机斑点噪声。本文使用最大池化从 SAR 图像中提取局部强特征,使用平均池化提取全局统计特征,降低 SAR 图像中噪声的影响,如图3所示。对于拼接后的特征即 $S \in \mathbf{R}^{C \times H \times W}$ 经过通道注意力模块得到:

$$A_s = \sigma(\Gamma(\text{AvgPool}(S) + \Gamma(\text{MaxPool}(S)))) \quad (1)$$

式中: σ 为 Sigmoid 函数; Γ 为两层神经网络。对输入特征 $S \in \mathbf{R}^{C \times H \times W}$ 基于 H 和 W 维度进行均值池化和全局最大池化,得到两个 $(C \times 1 \times 1)$ 的特征图,送入全连接层运算后相加,再经过 Sigmoid 激活操作,生成一维通道注意力图 A_c ,再将 A_c 与输入元素 S 相乘,获得通道注意力调整后

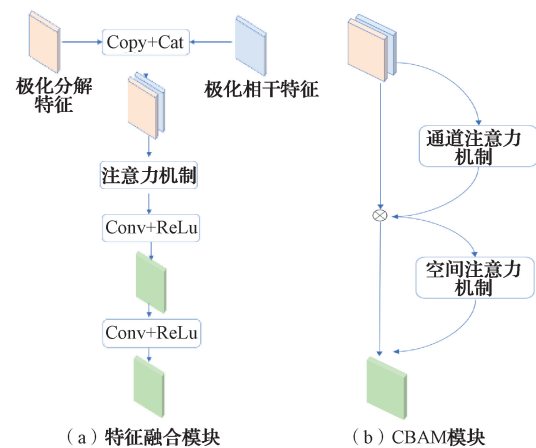


图2 特征融合模块与注意力机制模块

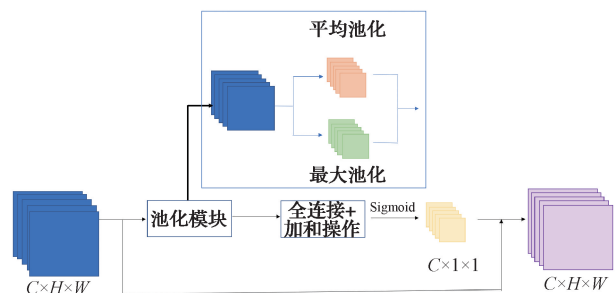


图3 通道注意力模块

的特征图 S_c :

$$S_c = A_c(S) \times S \quad (2)$$

经过通道注意力机制后得到的 S_c 相较于 S 通道维度不变,空间维度被压缩。相反空间注意力机制得到最终的输出 S_{cs} 相较于输入特征 S_c 空间维度不变,通道维度

被压缩,如图4所示,计算如式(3)。

$$A_s = \sigma(f^{7 \times 7}(\text{AvgPool}(S); \Gamma(\text{MaxPool}(S)))) \quad (3)$$

式中: σ 为 Sigmoid 函数; $f^{7 \times 7}$ 为卷积核大小为 7×7 的卷积。基于通道进行全局最大池化和均值池化,得到两个 $(1 \times H \times W)$ 的特征图,然后两个特征图进行拼接,经过卷积变为 1 通道特征,再经过一个 Sigmoid 特征图得到二维空间注意力图 A_s , 并与输入特征 S_c 相乘得到经过通道注意力机制与空间注意力机制的输出 S_{CS} :

$$S_{CS} = A_s(S_c) \times S_c \quad (4)$$

两路编码层分支能独立学习特征,CBAM 模块在通道和空间特征自适应选择的方式过滤掉冗余信息,以加权的方式获取它们的融合结果,实现极化分解信息对极化相干信息的多级辅助,不仅节约参数和计算力,还可以极大提升网络性能。

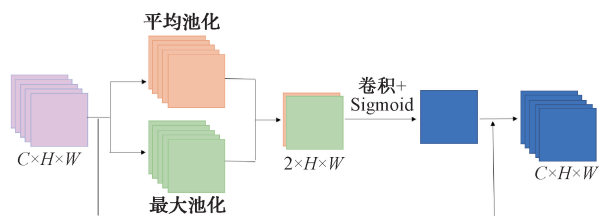


图4 空间注意力模块

(3)加入空洞卷积池化金字塔(atrous spatial pyramid pooling, ASPP)^[14]模块,增大了感受野,提升对大尺度地物的分辨能力。经过4次下采样的底层特征图通过改进的ASPP模块后作为解码器的初始输入。空洞卷积通过引入一个称为“扩张率”的参数,增加了感受野,让每个卷积核感受更大范围的特征,可以有效解决图像语义分割问题中下采样导致的图像分辨率、丢失信息的问题。

传统ASPP使用具有不同采样率的多个并行的空洞卷积层进行多尺度特征提取,即一个 1×1 的卷积分支、空洞率分别为 6、12、18 的 3 个空洞卷积层分支以及一个全局平均池化层,最后将 5 个分支提取的特征融合生成结果,增大感受野,有效获取多尺度特征和上下文信息。由于本文输入数据大小为 128×128 ,经过编码结构后输入到ASPP模块数据大小为 8×8 ,ASPP以不同的扩张率能够有效地捕捉多尺度信息,但是前一层特征图只有 8×8 ,那么当将这个 3×3 的卷积核应用于此特征图的时候,就会导致存在太多全 0 填充区域,经捕获不到全局信息没有意义,且对于 PolSAR 图像,扩张率过大会增大斑点噪声的干扰。故将空洞率减少为 2、4、6 如图5所示,可有效减少参数量同时扩大感受野,提升网络对大尺度地物的识别能力。

(4)统一过滤器深度为 64,减小网络结构,提升了分类效率。统一过滤器深度为 64 效果更令人满意,网络结构也更小,过滤器深度过深网络不易收敛,分割效率低。另外设计统一的过滤器数量(64)可以降低时间和空间复

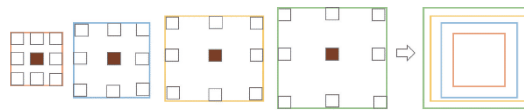


图5 改进 ASPP 模块

杂度^[15],同时降低网络结构设置对实验的硬件设备,尤其是可以降低对 GPU 的要求,是一种降低时间和空间复杂度的综合考虑。且使用统一的过滤器数量会使网络结构更加清晰,易于编码实现。

1.2 双通道极化 SAR 特征选取

为了充分挖掘极化信息,提高地物分类精度,本文设计具有两路编码层分支的全卷积网络结构。

1)通道 1:极化相干矩阵

极化散射矩阵包含了极化合成孔径雷达的极化信息。每个像素的偏振散射矩阵可以使用 2×2 极化散射矩阵表示:

$$[S] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \quad (5)$$

由式(5)可以看出,散射矩阵 $[S]$ 为一个 (2×2) 的复矩阵,通过计算矩阵中元素的相位与幅度,可以了解目标散射的多极化特性。其中,在单站情况下,满足互易定理的条件,即 $S_{hv} = S_{vh}$ 。

对于 PolSAR,根据互易条件($S_{hv} = S_{vh}$),极化相干矩阵 $T^{[16]}$ 可以表示为:

$$T_3 = \langle k k^* \rangle = \begin{bmatrix} T_{11} & T_{13} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \quad (6)$$

$k = (1/\sqrt{2})[S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv}]^T$ 是三维 Pauli 特征矢量, k^* 是 k 的共轭转置。相干矩阵 T 是一个埃尔米特半正定矩阵,其主对角线元素是实数,其他元素为复数,并且关于主对角线是共轭对称的, T 的上三角矩阵的 6 个元素可以用来完全表示 PolSAR 数据。因此,使用这 9 个元素来构造网络的输入向量^[17],可表示为:

$$V = [T_{11}, T_{22}, T_{33}, \text{real}(T_{12}), \text{imag}(T_{12}), \text{real}(T_{13}), \text{imag}(T_{13}), \text{real}(T_{23}), \text{imag}(T_{23})] \quad (7)$$

其中, $\text{real}()$ 代表取实部, $\text{imag}()$ 代表取虚部。物理散射机制的特征包含在 V 矩阵中。极化相干矩阵包含了目标后向散射的完整极化信息,为了不丢失目标的散射信息,选择极化相干矩阵作为通道 1 的输入量。

2)通道 2:极化分解特征参数

仅仅采用极化相干信息难以实现更好的信息表征,极化分解特征可以表征目标不同散射机理特征,有助于目标分类。因此选取极化分解特征参数作为通道 2 的输入。

综合多特征信息则需要考虑如何高效地实现特征信息的融合互补。为了提高分类精度,基于融合 CNN 将极化分解特征融入 PolSAR 分类网络,利用极化相干矩阵的

对角线元素 T_{11} 、 T_{22} 、 T_{33} 得到 Pauli 相干分解的 3 个散射分量^[18], 分别代表表面散射、偶次散射、体散射。基于 $H/A/\bar{\alpha}$ 分解得到极化散射参数 $\bar{\alpha}$ 、极化散射熵 H 、极化散射各向异度 A ^[19]。另外 Yamaguchi 四分量分解将极化协

方差矩阵分为表面散射、偶次散射、体散射和螺旋散射 4 种散射机制^[20]。选取了 Pauli 分解、 $H/A/\bar{\alpha}$ 以及 Yamaguchi 四分量分解组成一个 10 维的特征集, 综合上述极化相干信息与极化分解信息得到模型的特征输入(表 1)。

表 1 本文所用特征集

特征名称	物理意义	维数
极化相干矩阵 T	相干矩阵对角线元素及其实部虚部分离后的上三角元素	9
H	极化散射熵, 衡量散射过程中的极化程度, 表征分解目标的统计无序程度 $H/A/\bar{\alpha}$ 分解的参数	1
A	极化散射各向异度, 描述第二与第三散射机理的相对重要性; $H/A/\bar{\alpha}$ 分解的参数	1
$\bar{\alpha}$	极化散射参数, 表征目标的散射机理与目标的朝向极化散射角, 是对散射机理的物理解释, 取 $0^\circ, 45^\circ, 90^\circ$ 时分别代表表面散射, 体散射, 偶次散射; $H/A/\bar{\alpha}$ 分解的参数	1
Pauli 分解	分解分量表征不同地物散射类型: 单次或奇次散射机理、偶次散射机理、 45° 取向二面角散射体	3
Yamaguchi 极化分解	分解分量表征不同地物散射类型: 表面散射、偶次散射、体散射和螺旋散射	4

2 实验及分析

2.1 实验数据集

本文使用滑动窗口操作从实验图像中生成数据切片集^[21], 并从地面真实图像中生成对应的标签切片集。为了权衡性能和计算负担, 本文采用滑动窗口大小为 128, 步长为 32 的设置。此外, 学习率 lr 设置为 0.001, 批量大小设定为 8。本文所有非深度学习方法都基于 MATLAB 2020 实现, 基于深度学习的方法均在 Pytorch 框架下实现。

研究数据集采用旧金山地区与海南博鳌地区两个数据集。

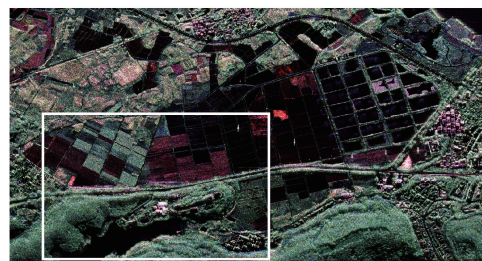
1) 旧金山地区, AIRSAR, L 波段。旧金山地区 Pauli 分解 RGB 图、地面真实地类分布如图 6 所示。原始数据的尺寸为 $900 \times 1\,024$ pixels, 空间分辨率为 10 m。该地物包括 5 类, 分别为山脉、水体、城市、有角度城市 and 植被。图 6(b) 中黑色区域为未标记类别, 每一类的样本数量如表 3 所示。由于有角度城市样本总数较少, 抽样率为 5.6%, 其他类抽样率均小于 2%。

2) 海南博鳌地区, L 波段, 机载。海南地区光学图、Pauli 分解 RGB、地面真实地类分布如图 7 所示。成像

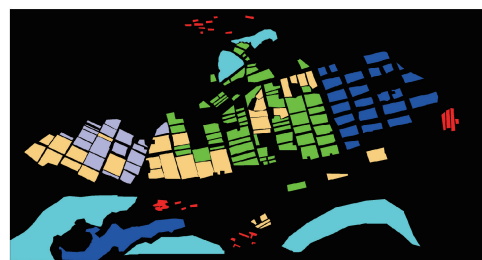
时间为 2021 年 2 月 25 日, 原始数据的尺寸为 $1\,700 \times 3\,200$ pixels, 空间分辨率为 3 m, 入射角度为 55° 。图 7(c) 中黑色背景为未标记区域, 图 7(b) 白框内各个地物类别



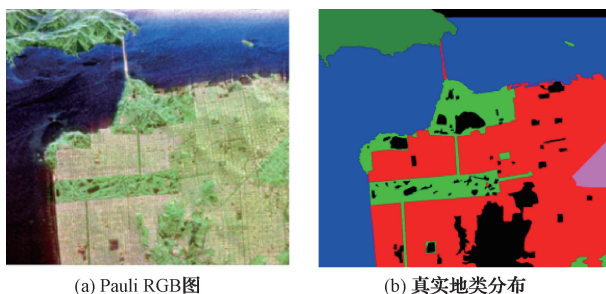
(a) 光学图



(b) Pauli RGB图



(c) 真实地类分布



(a) Pauli RGB图

(b) 真实地类分布



图 6 旧金山数据 Pauli RGB 图及真实地类分布

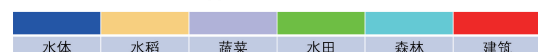


图 7 博鳌光学图、Pauli RGB 图及真实地类分布

实物图示例如图 8 所示,地物包括 6 类,分别为水体、水稻、蔬菜、水田、森林和建筑。每一类的样本数量如表 4 所示,每样抽样率为 5%。

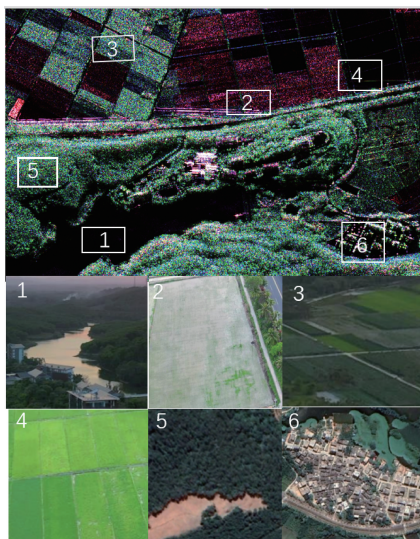


图 8 地面实物图

2.2 算法处理流程与评估指标

特征提取模块(图 9)过程如下:1)从集合中随机选择一定百分比的数据作为训练数据输入到网络,这些数据通过网络的复杂下采样部分向前传播以提取多级特征图;2)这些低分辨率特征图由上采样部分进行上采样,以生成预测的标签;3)根据损失函数计算预测的标签与对应的真实

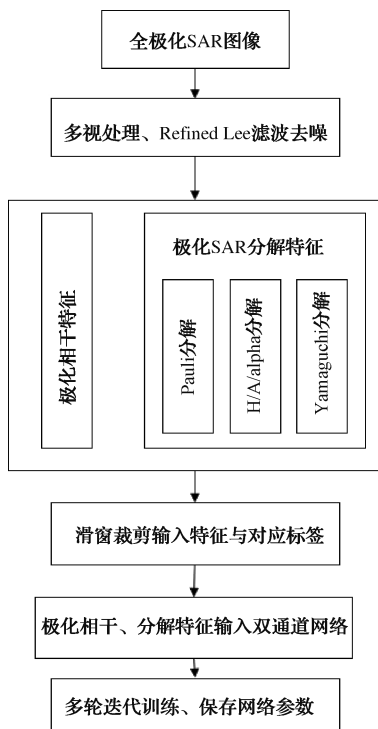


图 9 算法处理流程

标签之间的误差,迭代更新网络中的参数;4)根据某些条件,当误差值基本不变时,更新迭代将终止。在分类模块中,将整个 PolSAR 数据集输入到训练好的网络中。基于最后一个 Softmax 层的输出预测该 PolSAR 图像中每个像素的标签。值得注意的是,与预测每个图像块中心的单个标签的 CNN 模型相比,编解码网络可以实现像素级标记,并且可以减少预测期间的计算时间。

本文使用混淆矩阵法和 Kappa 分析法对极化 SAR 图像地物分类结果进行定量评价。根据混淆矩阵可以计算总体分类精度(overall accuracy, OA)、Kappa 系数等。

OA 表示整幅图像的预测正确率:

$$OA = \frac{\sum_{i=1}^c M_{ii}}{n} \quad (8)$$

Kappa 系数反映预测结果与地面真值之间的一致性:

$$Kappa = \frac{N(\sum_{i=1}^c M_{ii}) - \sum_{i=1}^c (\sum_{j=1}^c M_{ij} \sum_{j=1}^c M_{ji})}{N^2 - \sum_{i=1}^c (\sum_{j=1}^c M_{ij} \sum_{j=1}^c M_{ji})} \quad (9)$$

2.3 结果与分析

实验数据包括 AIRSAR 平台获得的旧金山地区 PolSAR 数据以及机载获取的海南博鳌地区 PolSAR 数据。在两个数据集上分别进行了 5 组实验,分别为 U_net 网络、FuseNet 网络、(M1)双通道特征融合编解码网络、(M2)双通道特征融合编解码网络+注意力机制、(M3)双通道特征融合编解码网络+注意力机制+多尺度模块即本文最终算法,证明本文提出最终算法的有效性。

1) 旧金山地区分类结果及分析

混淆矩阵如表 2 所示,可以看到,主要的混淆发生在城市 and 植被之间。但城市 and 植被分类精度分别达到 96.94% 与 94.74%,说明本文设计算法对于城市与植被的分类仍能达到较高精度。同时由于其他种类之间存在巨大的纹理特征差异,因此可以正确地区分。

表 2 本文 M3 网络旧金山地区分类结果混淆矩阵

类别	山地	水体	城市	有角度建筑	植被
山地	98.63%	1.17	0	0	0
水体	0.58	97.47%	0.95	0	1.00
城市	0	0.38	96.94%	0.17	2.51
有角度建筑	0	0	0.58	99.42%	0
植被	0	0.10	4.26	0	94.74%

不同算法分类精度如表 3 所示,分类结果如图 10 所示。M1 拼接方式双通道融合编解码网络相比于原始 U_net 网络,OA 和 Kappa 系数分别提升了 2.70%、3.46%,值得注意的是拼接的特征融合方式分类结果更好,并且类别之间的界限更加清晰。明显看出 FuseNet 加和的特征融合方法精度低于 M1 拼接的特征融合网络,推测原因如

上文所述,由于 T 矩阵元素与极化分解参数量纲可能不同,造成部分元素被忽略,分割效果相比于双通道特征融合网络较差。图 10(a)为 U_net 网络的分类结果。可以看出 A、B、C 区域植被和城市混合严重,有很多分类错误的区域。图 10(b)为 M1 双通道特征融合网络的分类结果。由于加入了有助于分类的极化分解特征,提高了分类精度,可以看出 A、B 区域正确率更高。图 10(c)为 M2 双通道特征融合网络加入注意力机制模块的分类结果,A 区域相较于 M1 网络分类结果明显正确率更高,植被分类精度明显提升,是因为注意力机制帮助模型对输入特征的每个部分赋予不同的权重,抽取更加关键及重要的信息,使模型做出更加准确的判断,提高分类精度。加入注意力机制后相比于原始 U_net 网络,OA 和 Kappa 系数分别提升了 4.30%、6.66%。本文最终算法 M3 又加入 ASPP 多尺度模块通过不同的空调率构建不同感受野的卷积核,获取多尺度物体信息,有利于边缘保持。如图 10(d)加入 ASPP 模块后,网络感受野增大,可以发现 B 区域的水体错误分类几乎消失,C 区域的城市预测误差也明显减少,并且加入多尺度模块后图像中孔洞和离散像素的数量大大减少。对于 OA、Kappa 的评价是最优的,分别达到

97.11%和 95.63%。以上证明了所提出的算法结构的有效性,该结构可以捕获更多的区别特征并有效地包含更多的空间信息。由表 3 可知,在本文最终 M3 网络精度提高的同时,所消耗的时间变化不大,更体现了本文所提出算法的高效性。

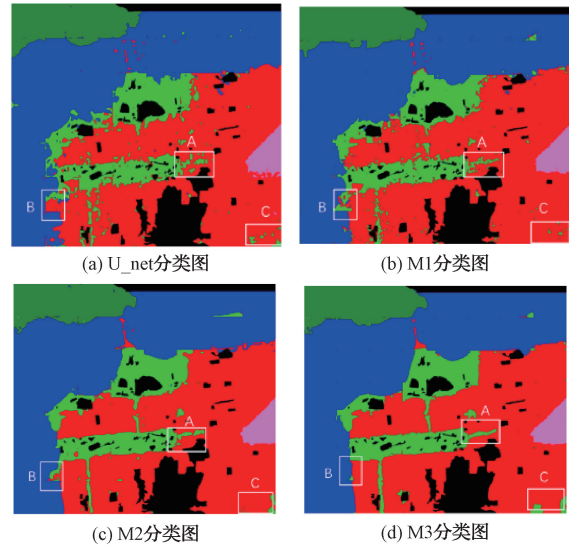


图 10 旧金山地区不同算法的分类结果

表 3 旧金山地区不同算法的分类精度

类别	样本数	采样点	精度/%				
			U_net	FuseNet	M1	M2	M3
山地	62 937	1 000	97.48	95.67	98.25	98.5	98.63
水体	338 393	1 000	94.46	93.83	94.66	96.31	97.47
城市	333 022	1 500	90.86	94.52	95.02	97.08	96.94
有角度建筑	14 552	800	98.85	99.11	98.94	99.07	99.42
植被	84712	1 000	85.18	84.52	87.86	91.73	94.74
OA/%	—	—	92.39	93.39	94.89	96.37	97.11
Kappa/%	—	—	88.61	90.04	91.68	94.51	95.63
耗时/s	—	—	2.34	1.68	1.76	1.92	1.93

2) 海南博鳌地区分类解结果

如图 11(a)所示,U-net 网络识别不同类别时表现不佳,不能很好的识别水田和水体、水田和水稻,这是由于水田与水体、水稻散射特征较为相似,只用 T 矩阵的特征分类误差较大。图 11(b)为本文提出的 M1 双通道特征融合算法,可以看到每一类精度均达到 99%以上,证明了双通道特征融合对于地物分类的有效性。从图 11(a)~(d)整体效果来看,本文最终算法 M4 双通道特征融合网络+注意力机制+ASPP 模块网络的分类图明显更接近真实地图。所有方法的评价指标如表 4 所示。所有双通道特征融合的方法都实现了超过 99%的 OA。就所有评估指标而言,所有双通道特征融合的方法对比 U_net 单通道的方

法都表现了更好的性能。虽然算法 M3 在 OA 方面相较于 M2 仅获得 0.34%的改善,但除水体之外的所有类别都显示出更高的准确率,这与图 11 的结果一致。总之,从图 11(a)~(d)和表 4 来看,对于海南博鳌 PolSAR 数据集,本文提出的算法与其他方法相比具有最佳性能,并且具有区分不同地形类别的强大能力。注意到采用双通道网络后耗时超过 U-net 网络,这是由于网络复杂度变高,考虑到精度显著提高的情况,耗时少量增加在可接受范围内。

图 12(a1)~(d1)为海南博鳌区域不同算法图像预测结果误差图,图 12(a2)~(d2)中图像从左到右对应为图 12(a1)~(d1)误差图中的 A、B、C 三个区域,其中黑色为未标记区域,绿色为预测正确区域,红色为预测错误区

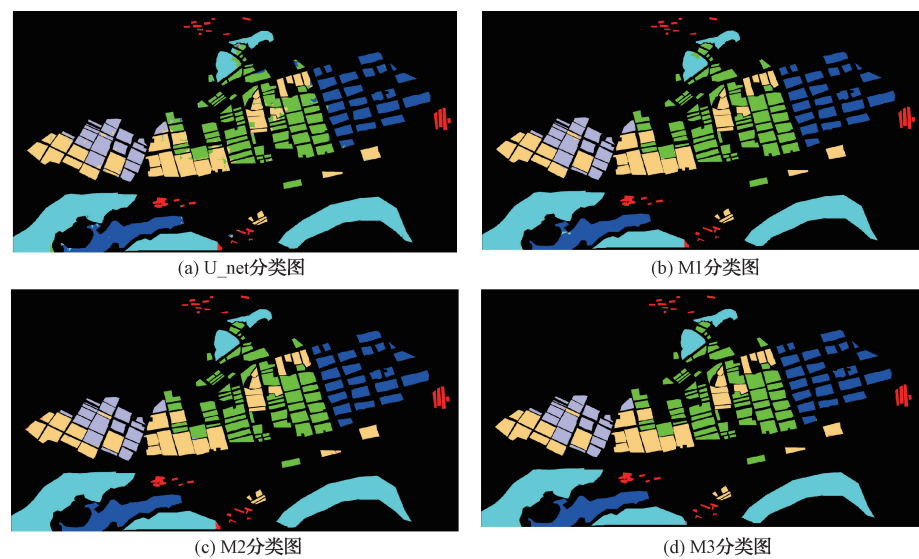


图 11 海南博鳌地区地区不同算法的分类结果

表 4 海南博鳌地区不同算法的分类精度

类别	样本数	采样点	精度/%				
			U_net	FuseNet	M1	M2	M3
水体	226 697	11 334	99.07	99.84	99.62	99.95	99.9
水稻	239 883	11 994	95.63	98.84	99.34	98.63	99.72
蔬菜	97 138	4 856	94.76	98.84	99.17	99.2	99.74
水田	252 033	12 601	94.49	97.1	99.63	99.72	99.85
森林	387 799	19 389	98.47	98.55	99.11	99.7	99.97
建筑	25 264	1 263	97.59	98.09	99.99	99.98	100
OA/%	—	—	96.88	98.91	99.38	99.53	99.87
Kappa/%	—	—	95.99	98.59	99.2	99.36	99.82
耗时/s	—	—	4.19	6.26	6.58	7.63	7.72

域。A 区域为森林和水田混淆区域,从局部误差图中可以看出,M1 改为双通道后预测错误明显减少,使用基于注意力机制的特征融合模块 M2 后,错误几乎消失,这是由于添加注意力模块后,网络更关注有助于分类的信息,提升了分类精度。B 区域为水稻和蔬菜混淆区域,使用 M1 双通道特征融合网络后,由图 12(b2)可知,分类正确的点

明显增加,但是由于水稻和蔬菜纹理较为相似,因此,图 12(c2)加入注意力机制后分类边缘误差仍然存在。C 区域为森林与建筑混淆区域,改为 M1 双通道后分类错误极化没有变化,分析原因可能是这两种类别仅仅极化相干信息就足以区分。图 12(d2)加入 ASPP 多尺度模块后 M3 网络,感受野增大,有利于边缘保持,精度也得到提升。

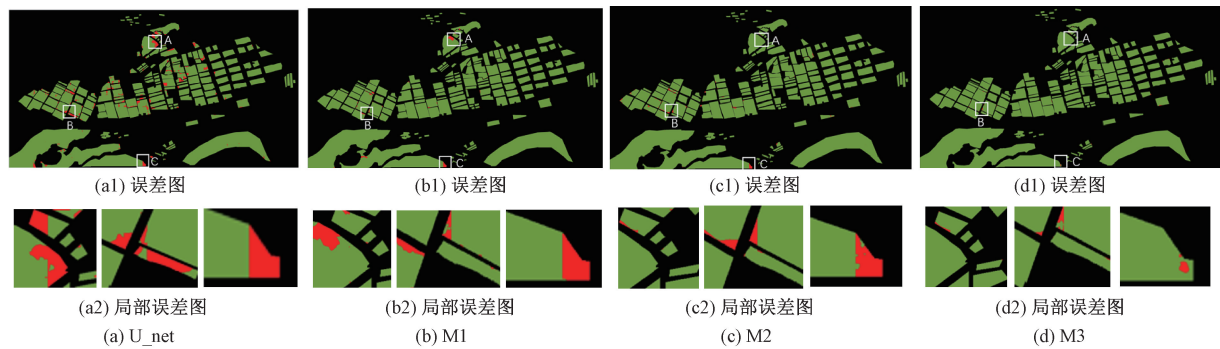


图 12 海南博鳌地区不同算法的误差图

3 结 论

本文针对极化 SAR 图像提出了双通道特征融合编解码网络极化分类算法,利用两路分支结构将极化相干信息与极化目标分解特征使用注意力机制进行特征融合,同时加入 ASPP 模块,有效的提取多尺度特征,充分地利用 PolSAR 图像信息,提高 PolSAR 图像分类性能。并在旧金山 PolSAR 数据与海南博鳌地区 PolSAR 数据上进行 5 组实验,证明算法的有效性。后续将尝试通过挖掘极化不同分解特征以减少输入特征维度,减轻网络负担提升网络训练速度,提升分类精度。

致谢:此次海南博鳌极化数据获取得到了中国科学院空天信息创新研究院航天微波遥感系统部的支持,在此表示衷心感谢!

参 考 文 献

- [1] 庞英,刘畅. 一种改进的基于超像素的 SAR 图像海陆分割方法[J]. 国外电子测量技术, 2019, 38(8): 12-18.
- [2] 李亚娟. 结合全局和局部稀疏表示的 SAR 图像目标识别方法[J]. 电子测量与仪器学报, 2020, 34(2): 165-171.
- [3] 赵泉华,王肖,李玉,等. 基于散射梯度矢量的 PolSAR 数据舰船检测方法[J]. 仪器仪表学报, 2020, 41(3): 39-48.
- [4] 钟能,杨文,杨祥立,等. 基于混合 Wishart 模型的极化 SAR 图像非监督分类[J]. 雷达学报, 2017, 6(5): 533-540.
- [5] 陈军,杜培军,谭琨. 一种改进的全极化 SAR 图像 MCSM-Wishart 非监督分类方法[J]. 国土资源遥感, 2015, 27(2): 15-21.
- [6] CHEN S W, TAO C S. PolSAR image classification using polarimetric-feature-driven deep convolutional neural network[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(4): 627-631.
- [7] YANG R, HU Z, LIU Y, et al. A novel polarimetric SAR classification method integrating pixel-based and patch-based classification[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 17(3): 431-435.
- [8] LONG, J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [9] PRIYA B H, REDDY B R. Machine learning based U-net algorithm for SAR image classification[J]. IEEE Delhi Section Conference, 2022: 1-8.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [11] XIAO D, LIU C, WANG Q, et al. PolSAR image classification based on dilated convolution and pixel-refining parallel mapping network in the complex domain [J]. Computer Vision and Pattern Recognition, 2019, arXiv:1909.10783.
- [12] BERGADO J R, PERSELLO C, STEIN A. FuserNet: End-to-end multispectral VHR image fusion and classification[J]. IEEE International Geoscience and Remote Sensing Symposium, 2018: 2091-2094.
- [13] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module[J]. Proceedings of the European Conference on Computer Vision, 2018: 3-19.
- [14] YANG Z, CHEN L, FU T, et al. Spine image segmentation based on U-Net and atrous spatial pyramid pooling[J]. Journal of Physics: Conference Series, 2022, 2209(1): 012020.
- [15] 苏健民,杨岚心,景维鹏. 基于 U-Net 的高分辨率遥感图像语义分割方法[J]. 计算机工程与应用, 2019, 55(7): 207-213.
- [16] 杨成财,余慧庄,龙郝明. 基于 SVM 和能量最小化的 PolSAR 图像分类方法[J]. 电子测量技术, 2020, 43(3): 146-152.
- [17] CAO Y, WU Y, ZHANG P, et al. Pixel-wise PolSAR image classification via a novel complex-valued deep fully convolutional network[J]. Remote Sensing, 2019, 11(22): 2653.
- [18] ZHU T, YU J, LI X, et al. SAR image classification method based on superpixel and Span-Pauli decomposition[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2015, 43(7): 77-81.
- [19] 尼格拉·吐尔逊,依力亚斯江·努尔麦麦提,王远弘,等. 基于 H/A/ α 分解全极化合成孔径雷达数据的干旱区土壤盐渍化分类[J]. 江苏农业学, 2019, 47(22): 273-279.
- [20] 李仲森. 极化雷达成像基础与应用[M]. 北京: 电子工业出版社, 2013: 133-168.
- [21] SATO M, WANG X S, CHEN S W. Uniform polarimetric matrix rotation theory and its applications[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(8): 4756-4770.

作 者 简 介

王蒙蒙, 硕士, 主要研究方向为极化 SAR 图像分类。

E-mail: wangmengmeng20@mails.ucas.ac.cn

刘秀清, 博士, 副研究员, 主要研究方向为极化 SAR 数据处理与信息提取技术。

E-mail: lucia@mail.ie.ac.cn

张衡, 博士, 副研究员, 主要研究方向为多基星载 SAR 信号处理、系统设计、多基线干涉 SAR 信号处理等。

E-mail: zhangheng@aircas.ac.cn

王春乐, 博士, 助理研究员, 主要研究方向为极化合成孔径雷达信号处理。

E-mail: clwang@mail.ie.ac.cn

贾小雪, 博士, 副研究员, 主要研究方向为合成孔径雷达信号处理技术。

E-mail: xiaoxue_snowing@163.com