

一种基于密集深度分离卷积的SAR图像水域分割算法

张金松* 邢孟道 孙光才

(西安电子科技大学雷达信号处理国家重点实验室 西安 710071)

(西安电子科技大学信息感知技术协同创新中心 西安 710071)

摘 要: SAR图像的水域分割在舰船目标检测、灾害监测等军事和民用领域具有重要意义。针对传统水域分割算法鲁棒性差、难以准确进行分割等问题,该文首先建立了基于高分三号的SAR图像水域分割数据集,并基于深度学习技术提出了基于密集深度分离卷积的分割网络架构,该网络以SAR图像作为输入,通过密集分离卷积和扩张卷积提取图像高维特征,并构造基于双线性插值的上采样解码模块用于输出分割结果。在水域分割数据集上的实验结果表明,与传统方法相比,该方法不仅在分割准确度上有大幅提高,在算法的鲁棒性和分割速度上也具有部分优势,具备较好的工程实用价值。

关键词: 合成孔径雷达; 水域分割; 深度学习; 密集分离卷积; 特征提取

中图分类号: TN958

文献标识码: A

文章编号: 2095-283X(2019)03-0400-13

DOI: 10.12000/JR19008

引用格式: 张金松, 邢孟道, 孙光才. 一种基于密集深度分离卷积的SAR图像水域分割算法[J]. 雷达学报, 2019, 8(3): 400-412. doi: 10.12000/JR19008.

Reference format: ZHANG Jinsong, XING Mengdao, and SUN Guangcai. A water segmentation algorithm for SAR image based on dense depthwise separable convolution[J]. *Journal of Radars*, 2019, 8(3): 400-412. doi: 10.12000/JR19008.

A Water Segmentation Algorithm for SAR Image Based on Dense Depthwise Separable Convolution

ZHANG Jinsong* XING Mengdao SUN Guangcai

(National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

(Collaborative Innovation Center of Information Sensing and Understanding,
Xidian University, Xi'an 710071, China)

Abstract: Water segmentation of real SAR images is of great significance in military and civilian applications such as ship target detection and disaster monitoring. To solve the issues of poor robustness and inaccurate segmentation of traditional water segmentation algorithms, this paper first establishes a SAR water segmentation dataset based on the GF3 satellite and then presents a segmentation network architecture based on depthwise separable convolution. The network takes real SAR images as inputs, extracts high-dimensional features through depthwise separable and dilated convolutions, constructs an up-sampling and decoding module based on bilinear interpolation, and then outputs the corresponding segmentation results. The segmentation results of a water segmentation dataset show that the proposed segmentation method remarkably improves the segmentation accuracy, the segmentation robustness and running speed than traditional method. Therefore, the findings demonstrate the excellent practical engineering value of the proposed algorithm.

Key words: Synthetic Aperture Radar (SAR); Water segmentation; Deep learning; Dense separable convolution; Feature extraction

收稿日期: 2019-01-14; 改回日期: 2019-04-08; 网络出版: 2019-06-19

*通信作者: 张金松 jinsongxd@163.com

*Corresponding Author: ZHANG Jinsong, jinsongxd@163.com

基金项目: 国家重点研发计划(2017YFC1405600), 国家自然科学基金创新群体基金(61621005)

Foundation Items: The State Key Research Development Program (2017YFC1405600), The Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61621005)

责任编辑: 关键 Corresponding Editor: GUAN Jian

1 引言

合成孔径雷达(Synthetic Aperture Radar, SAR)具有全天时、全天候、高分辨和不受云雾影响等优点,被广泛地应用于农业、水洋、军事等领域^[1]。实现SAR图像的准确水域分割具有重要意义,以舰船检测为例,经验信息假设舰船目标位于海面、河流、湖泊等典型水域上,由于目标本身散射后向系数较大,回波信号信噪比高,在纯水域上检测舰船目标相对容易^[2]。而对于水陆交接区域而言,如图1(a)所示,由于陆地散射系数大于水域散射系数,导致原始图像中存在复杂的陆地背景干扰,直接利用舰船检测方法对此类图像进行检测,在实现水域舰船检测的同时(图1(a)绿色圈所示),也会导致地面区域出现多个舰船目标(图1(a)红色圈所示),即造成检测漏警率高、检测准确率低现象,从而严重影响舰船检测的准确度。因此,有必要针对SAR图像水陆交接区域研究相应分割算法,并将分割算法作为舰船目标的检测预处理步骤,如图1(b)所示,在分割后的水域场景(绿色区域)内检测舰船目标,以期提高舰船目标检测正确率^[3-5]。

现有水域分割算法可以概述为以下几类:阈值分割方法、聚类方法和边缘检测方法以及基于马尔科夫随机场的分割方法。(1)阈值分割方法利用1个或多个阈值对SAR图像按图像灰度进行划分,灰度值在同一阈值范围内的为同一类型地物^[6-8]; (2)聚类分割方法采用相似度统计方法对SAR图像中的具有相似特性的像素点进行分类^[9]; (3)边缘检测方法利用水域和陆地区域后向散射系数的差异,提取水域和陆地交界区域的灰度值变化实现分割^[10]; (4)马尔科夫随机场方法基于贝叶斯理论,通过对每个像素点邻域建立马尔科夫随机模型,并计算邻域内的统计信息估计当前像素点的概率分布^[11]。上述传统方法具有各自的优势,例如阈值分割和边缘检测具有计算量小、速度快特点,聚类方法和马尔科夫随

机场方法在简单场景SAR图像上可获得非常高的分割准确度。但上述方法容易受噪声影响,且算法参数调整过程需要过多人工参与,适用于简单场景,在复杂真实的SAR图像上,无法兼容水域分割的准确性、鲁棒性和实时性等要求^[3]。

在深度学习用于图像分割方面,文献[12]率先开发了经过端到端训练的全卷积网络(Fully Convolutional Network, FCN),用于图像分割,FCN采用任意大小的图像作为输入,可生成与原图像相同大小的分割图像。文献[13]将文献[12]的FCN扩展到生物显微镜图像分割,创建了U-net网络,该网络由两部分组成:卷积部分用于计算特征,扩张部分用于对图像进行空间定位,下采样部分具有类似FCN的结构,提取具有 3×3 卷积的特征,上采样或扩张部分使用上卷积减少特征图的数量,同时增加尺寸和宽度,最后采用 1×1 卷积处理特征图以生成分割结果,从而完成对输入图像的逐像素分类。此后,U-net架构在FPN, PSPNet, Deeplabv3+等分割模型^[14-17]中得到了广泛的扩展。

本文基于深度学习技术,首先收集了原始高分三号星载SAR滑动聚束模式数据,对数据做预处理操作,建立水域分割数据集。然后研究了深度分离卷积和扩张卷积的算法原理,基于残差连接网络架构和密集连接网络架构构建了基于深度分离卷积的SAR图像特征提取网络,在此基础上,结合特征提取网络和扩张卷积以及双线性插值上采样方法,构建了基于编码-解码架构的SAR图像水域分割网络架构。测试集数据以及其他模式下的SAR图像的分割结果,证明了所提方法的有效性以及工程的可实现性。

2 数据集

深度学习服从统计学规律,其假设测试数据和训练数据服从同一概率分布,即当数据量足够大时,便可以用神经网络拟合的权重参数来准确对测



图1 水域分割对舰船检测意义

Fig. 1 Significance of water segmentation for ship detection

试数据预测,数据量越大,网络泛化性能也越为优异。为了实现准确的水域分割,本文选取高分三号卫星数据作为数据源。高分三号是我国首颗自主研制的C频段多极化SAR卫星,具有高分辨率、大成像幅宽、高辐射精度、多成像模式和长时工作等特点^[18,19]。该卫星共有12种成像模式,其中5种分辨率较高的成像模式对应的成像参数如表1所示。

表 1 高分三号成像模式

Tab. 1 The imaging modes of GF3 satellite

工作模式	分辨率(m)	极化方式	成像幅宽(km)
滑块聚束(SL)	1	单极化	10
超精细条带(UFS)	3	单极化	30
精细条带1(FSI)	5	双极化	50
精细条带2(FSII)	10	双极化	100
标准条带1(QPSI)	8	全极化	30

本文收集了10幅高分三号聚束模式的SAR图像,用来训练和测试水域分割算法的性能。由于2维采样率不一致导致SAR图像出现形变,采用多视处理对原始SAR图像进行几何校正。同时原始SAR图像能量值起伏明显,采用对数函数对原始图像进行量化,并根据量化结果手动调整图像灰度值至0~255像素。标注工具采用LabelMe开源工具包进行标注,标注过程中由手工记录陆地和水域的交接边缘,陆地区域标记为类别“0”,水域标记为类别“1”。另外由于各区域面积差异较大,宽幅陆地区域内出现的面积较小、对舰船检测影响可以忽略的水域统一标记为陆地,对应的标记结果为与原始大图尺寸完全一致的标签图像。由于原始图像尺寸较大,直接采用图像进行分割会造成计算量较大、运算速度慢,因此对原始图像和对应标签进行4倍线性降采样处理,在损失较少图像信息的前提下,保证图像检测准确度。此外,由于卷积神经网络对输入图像的尺寸限制,以513像素为步长将原始图像裁剪成尺寸为513×513像素的小图像,尺寸不足的地方以0像素填充,填充后的区域默认为水域。另外深度卷积神经网络具有旋转敏感性,无法捕获输入图像在多个角度的特征,以30°为间隔对裁剪后的图像进行多次旋转。在旋转的同时,做了左右翻转以及上下翻转操作对数据进行扩充。除上述步骤外,不再采用其它方法对SAR图像进行调整和扩充。在裁剪和扩充过程中,图像的数量和尺寸变化情况图像如表2所示。为了对分割算法进行训练、验证和测试,数据集中17102张图片作为训练数据集,3019张图片作为验证数据集,1059张图片作为测试集。

表 2 数据集图像组成

Tab. 2 The composition of dataset

图像类型	数量	图像尺寸(像素)
原始图像	10	$\approx 10,000 \times 10,000$
裁剪图像	480	513 × 513
扩充图像	21180	513 × 513

3 特征提取网络

3.1 深度可分离卷积

传统深度神经网络中的卷积层采用的卷积方式如图2(a)所示,其利用每个卷积核同时对特征图的2维平面特征和通道特征进行卷积,在尺寸维度和通道维度存在参数耦合,会造成卷积核参数较多,计算复杂。

本文所采用的卷积网络以深度分离卷积网络为基本架构^[20],结构如图2(b)所示,其将常规卷积分解为两个独立的卷积过程:深度卷积和逐点卷积。其中,深度卷积采用维度为 $m \times n \times 1 \times C_1$ 的卷积核对输入特征图进行滤波,经卷积获取的中间特征图的通道个数与输入特征图的维度一致;而分离卷积则只考虑通道特征,不考虑图像的两维尺寸特征,直接采用 $1 \times 1 \times C_1 \times C_2$ 的卷积层对中间特征图进行滤波, C_1 表示每个卷积核的通道个数,与中间特征图的个数一致, C_2 表示卷积核的个数。深度分离卷积最终可以得到与常规卷积方式类似的输出特征图,同时解耦了特征图在2维平面特征和通道特征的相关性。

3.2 扩张卷积

在卷积神经网络中,接收域代表着卷积输出特征图中的每个像素点对卷积输入特征图中像素点的抽象能力。如图3(a)所示,利用尺寸为 3×3 的卷积核对输入特征图进行特征提取,输出特征图的接收域同样为 3×3 。高的接收域代表着更优异的上下文信息提取能力,常规卷积网络一般采用卷积层串联池化层的做法,以期提高卷积结果的接收域,但对于水域分割来说,池化层的降分辨率操作会造成对原始输入图像的信息丢失,因此本文在卷积过程中采用扩张卷积网络提取特征^[21],从而兼顾卷积网络的高分辨率和高特征提取能力。扩张卷积过程如图3(b)所示。

可见,图3(b)所示的扩张卷积在不增加卷积核参数数量的情况下,充分融合了特征图的周围像素信息,将输出特征图的接收域由 3×3 像素扩展到 5×5 像素,更有利于水陆分离任务中像素点分类的准确性。

3.3 特征提取网络

本节采用3.2节所述的深度可分离卷积和扩张卷积构建SAR图像特征提取网络,网络结构如图4所示。

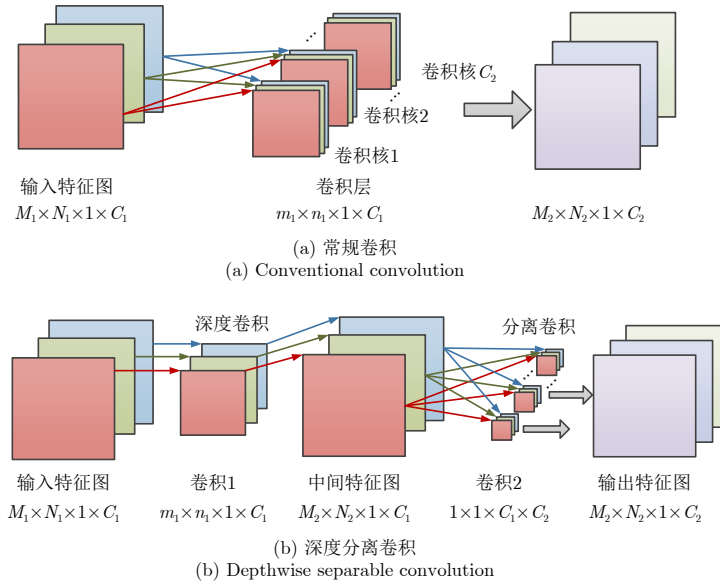


图2 常规卷积和深度分离卷积结构对比图

Fig. 2 Comparison of conventional convolution and depthwise separable convolution

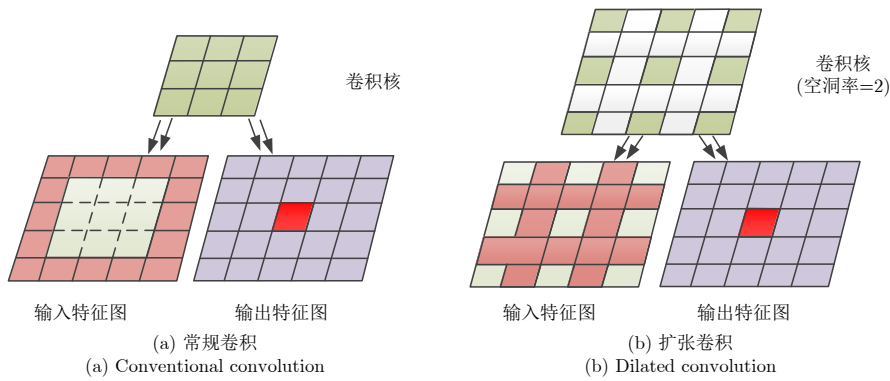


图3 常规卷积和扩张卷积结构对比图

Fig. 3 Comparison of conventional convolution and dilated convolution

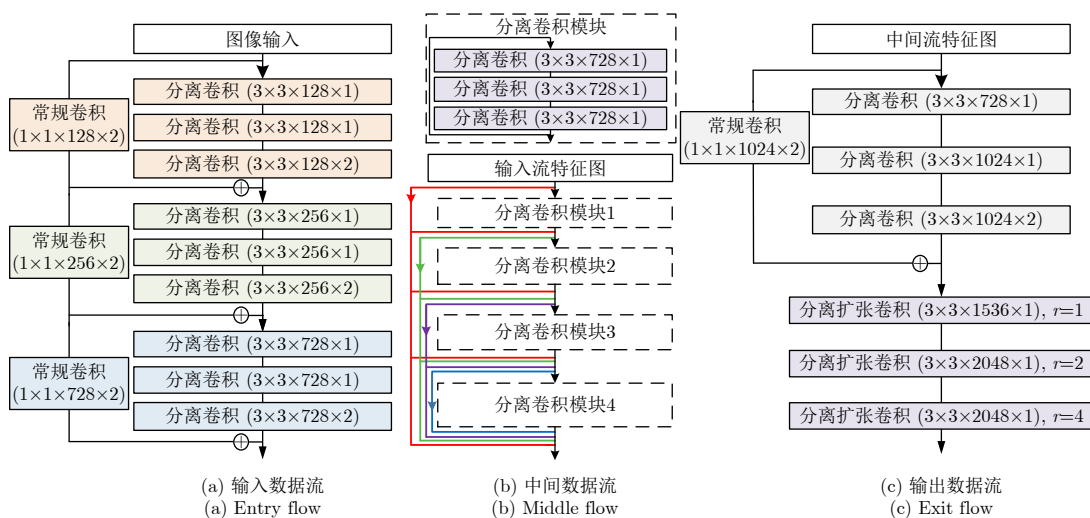


图4 特征提取网络结构示意图

Fig. 4 The structure of feature extraction network

该特征提取网络由3部分组成:输入数据流、中间数据流和输出数据流^[22]。

3.3.1 输入数据流

输入流网络结构如图4(a)所示,其采用3个卷积模块,每个卷积模块中包含不同的卷积层,以第1个卷积模块为例,其以原始SAR图像 x_{input} 作为输入,采用残差网络连接^[23]:主干支路采用3个深度分离卷积层,卷积层的结构为4维向量 $[h, w, n, s]$,其中 $[h, w]$ 表示卷积核的尺寸, n 表示采用卷积核的个数, s 表示卷积核的步长,卷积层的计算方式与图3(b)保持一致,卷积结果表示为 $f(x)$,该支路用来提取 x_{input} 的高维非线性特征;跨层连接支路则采用常规卷积对 x_{input} 做尺寸降维、通道升维,保证卷积后的特征图 $g(x)$ 维度与主干支路输出特征图 $f(x)$ 维度一致。原始SAR图像经过输入数据流后,输出特征图尺寸降低为输入图像的1/8,特征图的个数则增加到728。

3.3.2 中间数据流

中间数据流网络结构如图4(b)所示,其由4个相同的分离卷积模块组成,每个分离卷积模块由基于残差连接的3个分离卷积层组成,卷积层的步长 $s=1$ 。为了提高各分离卷积模块之间的特征映射关系,本文提出了基于密集连接的分离卷积网络,其网络结构如图4(b)所示。设 x_{l-1} 为上层分离卷积模块的输出, h_l 为表示对分离卷积模块做非线性变换函数,则对于传统逐层连接方式而言,第 l 层的输出特征图 \mathbf{X}_l 则表示为

$$\mathbf{X}_l = h_l(\mathbf{X}_{l-1}) \quad (1)$$

3.3.1节所述的残差连接方式则表示为

$$\mathbf{X}_l = h_l(\mathbf{X}_{l-1}) + \mathbf{X}_{l-1} \quad (2)$$

密集连接网络与残差网络区别在于,残差网络将上层网络输入特征图 x_{l-1} 和对其进行非线性变换 $h_l(\mathbf{X}_{l-1})$ 的求和作为下层网络输入,而密集连接网络则将所有前层网络模块的输出特征图做通道并联,并将并联后的特征图作为下层网络输入^[24],其计算方式可以表述为

$$\mathbf{X}_l = h_l([\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_{l-1}]) \quad (3)$$

3.3.3 输出数据流

输出数据流结构如图4(c)所示,其包含2个深度分离卷积模块,第1个卷积模块与上节所述残差模块结构类似,第2个卷积模块联合深度可分离卷积和扩张卷积,扩张卷积的扩张率依次为(1, 2, 4)。另外,第2个分离卷积模块的各层间不采用残差连接,而是直接采用级联方式。

本节所述的特征提取网络中除残差连接的跨层连接支路外,其他所有的深度分离卷积层均采用非

线性修正单元函数作为激活函数^[25]。另外,由于网络经过每层深度分离卷积后,网络的权值参数和输出特征图的概率分布会发生变化,导致网络训练时在反向传播过程中出现梯度消失现象,无法训练得到局部最优值。针对此问题,该网络结构采用了块归一化层对每层输出特征图的分布进行了重新调整^[26]。

4 水域分割网络

原始SAR图像经过特征提取网络的多次卷积之后,虽然特征图通道个数增多,但特征图的分辨率信息会下降。为了实现准确的水域分割,必须通过卷积网络逐层恢复图像分辨率信息,以得到与原始输入SAR图像尺寸完全一致的分割结果。本文基于该思想,提出了基于编码-解码结构的水域分割网络,通过双线性插值进行上采样并添加多个卷积层获得与原始SAR图像逐像素对应的分割结果,网络结构如图5所示。

该网络由编码架构和解码架构2部分组成,其中编码架构的意义在于,通过深度分离卷积和扩张卷积等形式,提取输入图像在多个分辨率条件下的特征表示,而解码架构则利用提取的多维特征信息,构建深度神经网络,完成对输入图像中逐像素的类别判断^[16]。

4.1 编码架构

编码架构由特征提取网络和空洞金字塔池化模块构成。其中,特征提取部分采用上节所述的特征提取网络,输入的SAR图像在经过特征提取模块后,特征图尺寸减小为原始输入图像的1/16,而特征图的个数则增加到2048。

扩张空间金字塔模块的主要意义在于增加所提取的高维特征图的上下文信息概括能力。如图5所示,该模块采用多个不同的扩张率(6, 12, 18)对特征图进行扩张卷积,使特征图中的每个像素点充分融合周围像素点的特征,即获取更多的上下文信息,多个扩张率的选择能提取多尺度的特征信息。该模块在采用扩张卷积对特征图进行多尺度特征变化的同时,也采用了 1×1 的卷积层和全局平均池化以保留特征图的自身原始信息。为了综合各个卷积层提取的特征图,该模块按通道维度对各特征图进行合并,由于合并后的特征图维度较大,采用 1×1 卷积对该特征图进行降维,并获得编码架构最终的输出特征图。

4.2 解码架构

解码架构如图5所示,解码架构将特征提取网络中输入数据流的第2个分离卷积模块提取的特征图作为低维特征输入,将输出数据流提取的特征图和编码结构的末端输出特征图作为高维特征输入。

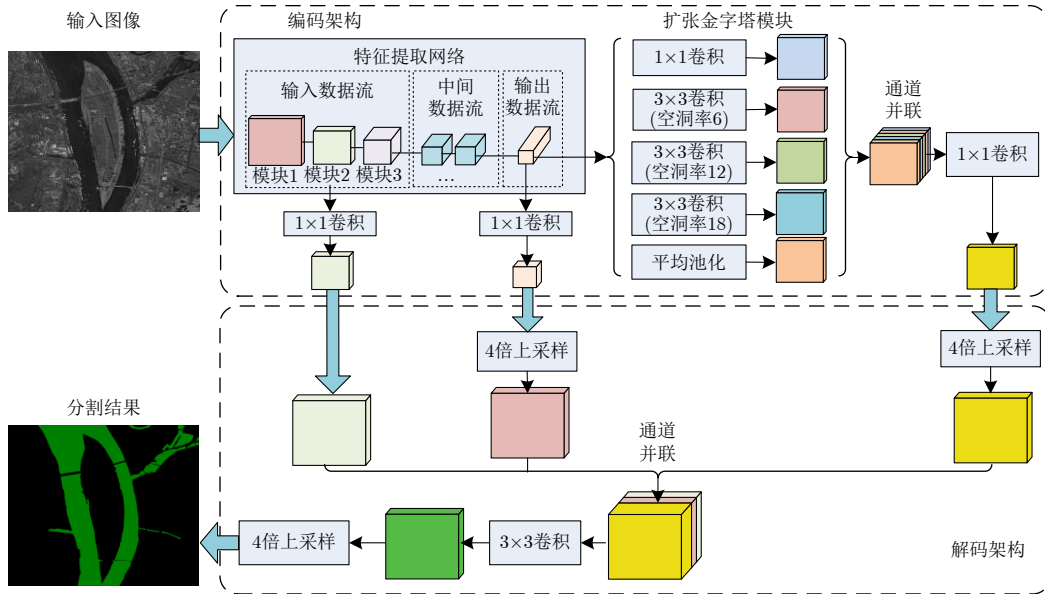


图5 基于编码-解码结构的SAR图像水域分割网络示意图

Fig. 5 The structure of encoder-decoder network for water segmentation

网络的具体结构为：首先对低维特征图进行 1×1 卷积以降低特征图的通道个数，获取的特征图假设为 $f_1(\mathbf{X})$ ，然后采用双线性差值对高维特征图进行4倍上采样，得到特征图 $f_2(\mathbf{X})$ ， $f_1(\mathbf{X})$ 与 $f_2(\mathbf{X})$ 维度一致，均为原始输入图像尺寸的1/4。然后按通道维度对3组特征图进行并联，并采用 3×3 卷积对合并后的特征图进行特征提取和通道维度降维，获得特征图 $g(\mathbf{X})$ ，然后再次利用双线性差值对该特征图进行4倍上采样，即可获得最终的分割结果。

4.3 网络损失函数

为了将网络的输出映射到每个类别的概率值，输出层采用softmax作为激活函数^[27]

$$\hat{p}_{(m,n)}^k = \frac{e^{X_{(m,n)}^k}}{\sum_{k=1}^K e^{X_{(m,n)}^k}} \quad (4)$$

式中， $X_{(m,n)}^k$ 表示输入图像在 (m,n) 位置的对应第 k 个类别的网络输出值， K 表示样本类别数量， $\hat{p}_{(m,n)}^k$ 表示经过softmax变换后对应第 k 个类别的概率输出值。网络采用交叉熵损失函数作为优化的目标函数^[28]，设每次训练过程采用的训练图像个数为 L ，输入图像尺寸为 M, N ，则目标函数为

$$E = \sum_{m,n=1}^{M,N} \sum_{l=1}^L \sum_{k=1}^K p_{(m,n)}^{l,k} \lg \hat{p}_{(m,n)}^{l,k} \quad (5)$$

式中， $p_{(m,n)}^{l,k}$ 表示网络在图像在 (m,n) 位置对类别 l 的真实标签， $\hat{p}_{(m,n)}^{l,k}$ 表示对应的概率值。由于在水域分割任务中，仅存在“水域”和“陆地”2类目标，设 $p_{(m,n)}^l$ 表示网络在 (m,n) 位置对“水域”类

别的真实标签， $\hat{p}_{(m,n)}^l$ 为对应的预测概率值，则目标函数可以表示为

$$E = \sum_{m,n=1}^{M,N} \sum_{l=1}^L p_{(m,n)}^l \lg (1 - \hat{p}_{(m,n)}^l) + (1 - p_{(m,n)}^l) \lg (1 - \hat{p}_{(m,n)}^l) \quad (6)$$

5 实验结果及分析

下面给出实验结果对比及相关分析。

5.1 网络训练结果

5.1.1 训练参数设置

学习速率决定了反向传播过程中梯度下降的速度，本文采用多项式算法调整学习速率，设第 λ_i 代学习速率为

$$\varepsilon_i = \varepsilon_0 \left(1 - \frac{\lambda_i}{\lambda_{\text{global}}}\right)^\beta \quad (7)$$

式中， $\varepsilon_0 = 7e^{-3}$ 代表初始学习速率， $\lambda_{\text{global}} = 5e^5$ 代表网络全局训练步长， $\beta = 0.9$ 用来调整学习速率的降低速度。在网络训练过程中，采用带动量的梯度下降算法，已知第 i 代权值为 θ_i ，则第 $i+1$ 代权值更新量表示为

$$\Delta\theta_{i+1} = \alpha\Delta\theta_i - \varepsilon_i \frac{\partial E}{\partial \theta_i} - \eta\varepsilon_i\theta_i \quad (8)$$

式中， $\alpha = 0.9$ 表示动量系数，用来表征第 i 代权值更新量 $\Delta\theta_i$ 对本次更新量的影响程度， $\partial E / \partial \theta_i$ 表示网络损失函数 E 对网络权值 θ_i 的影响程度， $\eta = 4e^{-5}$ 表示权值正则化系数，用来对网络权值正则化。第 $i+1$ 代更新后的权值表示为

$$\theta_{i+1} = \theta_i + \Delta\theta_{i+1} \quad (9)$$

另外, 由于显存空间限制, 每次训练输入的图像数量设为4。网络采用基于Imagenet数据集^[20]训练好的模型作为分割网络中特征提取网络的训练初始值, 其他参数则采用随机初始化。

5.1.2 训练结果

网络训练结果如图6所示, 其中图6(a)表示训练过程中学习速率随训练步长变化结果, 与式(7)所示变化规律保持一致。图6(b)则表示训练过程中损失函数变化情况, 红线表示具体变化, 绿线表示损失值包络。由图中可见, 随着训练步数增加, 损失函数逐渐收敛到最优值。

在网络训练完成后, 本文在测试集中任意选取了1幅包含陆地和水域场景在内的SAR图像, 如图7(a)所示, 图7(b)为分割网络对输入图7(a)的预测分割结果, 图中蓝色区域表示水域, 黑色表示陆地区域, 可见图像中水域和陆地边缘均被有效分割, 且保持了完整的边缘轮廓。为了更直观理解网络的工作原理, 输入图7(a)在网络传播过程中的高维特征图, 如图7(c)到图7(g)所示。

5.2 数据集和网络结构对分割性能影响

在第2节数据集部分, 考虑到较小的原始图像数据量, 为提高水域分割网络的准确度, 本文采用旋转和翻转等操作对原始SAR图像进行扩充, 并利用扩充后数据集训练、验证和测试网络结构。为了验证数据扩充对网络的水域分割性能的影响, 本文详细比较了利用所提分割网络对不同的数据所产生的分割结果, 结果如表3所示。

由表3可见, 采用未扩充的数据训练分割网络, 像素准确度仅为0.9569, 分割交并比为0.9497; 而分别采用旋转和翻转操作均能提高分割准确度。分

割性能最为优异的是本文采用的旋转加翻转的数据扩充方式, 使像素准确度和分割交并比分别提高了0.03和0.035, 达到了较高的分割准确度。因此数据扩充对分割性能具有着重要影响。

在3.3节特征提取网络部分, 为了增强特征提取网络对SAR图像的特征提取能力和误差的反向传播能力, 本文分别将残差连接和密集连接应用到特征提取网络的不同数据流部分。残差连接和密集连接对分割网络的性能影响如表4所示。

表4中的结果均为采用扩充后的数据集训练的网络模型的分割结果。其中直连方式代表特征提取网络内部不采用任何跨层连接方式。残差连接代表如图4所示的网络连接, 但中间数据流不再采用密集连接; 密集连接方式则与其相反。而残差加密集连接即本文所提的连接方式与图4保持完全一致。由表4可见, 分别采用残差连接和密集连接都能有效提高分割的像素准确度和交并比。而采用所提的残差加密集的连接方式, 使像素准确度和分割交并比分别提高了0.057和0.055, 性能明显提升。

5.3 分割准确度对比

为了证明本文算法的分割有效性, 本文采用了多种分割方法和所提算法进行对比。首先在测试数据集中选取了2幅包含水域和陆地在内的SAR图像, 如图8(a)所示。输入图像对应的真实标签如图8(b)所示, 标签中的黑色区域表示分割后的水域场景, 白色表示分割后的陆地区域。针对传统分割方法, 选取了FCM方法^[9]、MRF方法^[11]、基于形态学操作改进的OTSU方法^[4]和Levelset^[30]方法, 针对深度学习方法, 选取了Unet方法^[13]和Deeplabv3+^[16]方法, 各方法的分割结果如图8(c)到图8(i)所示。

由分割结果中可见, 传统方法能有效提取图像

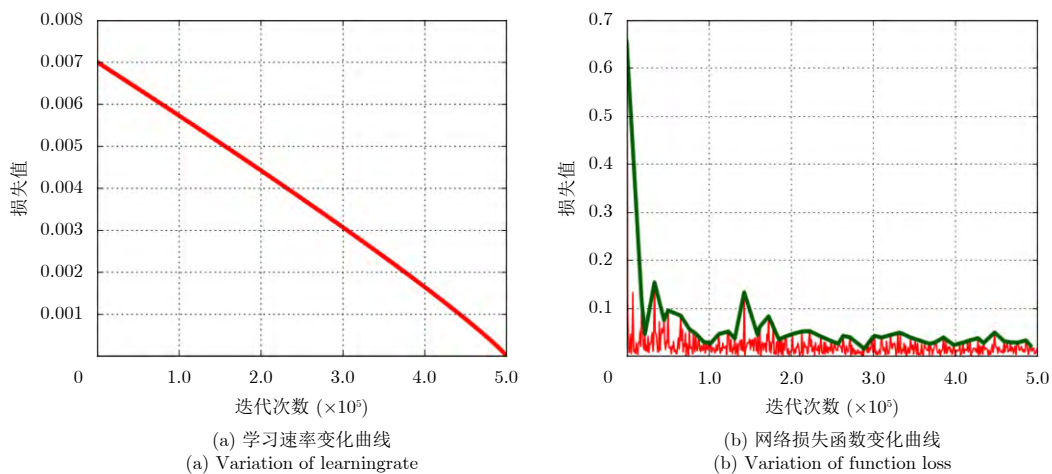


图6 分割网络训练结果示意图

Fig. 6 Training results of segmentation network

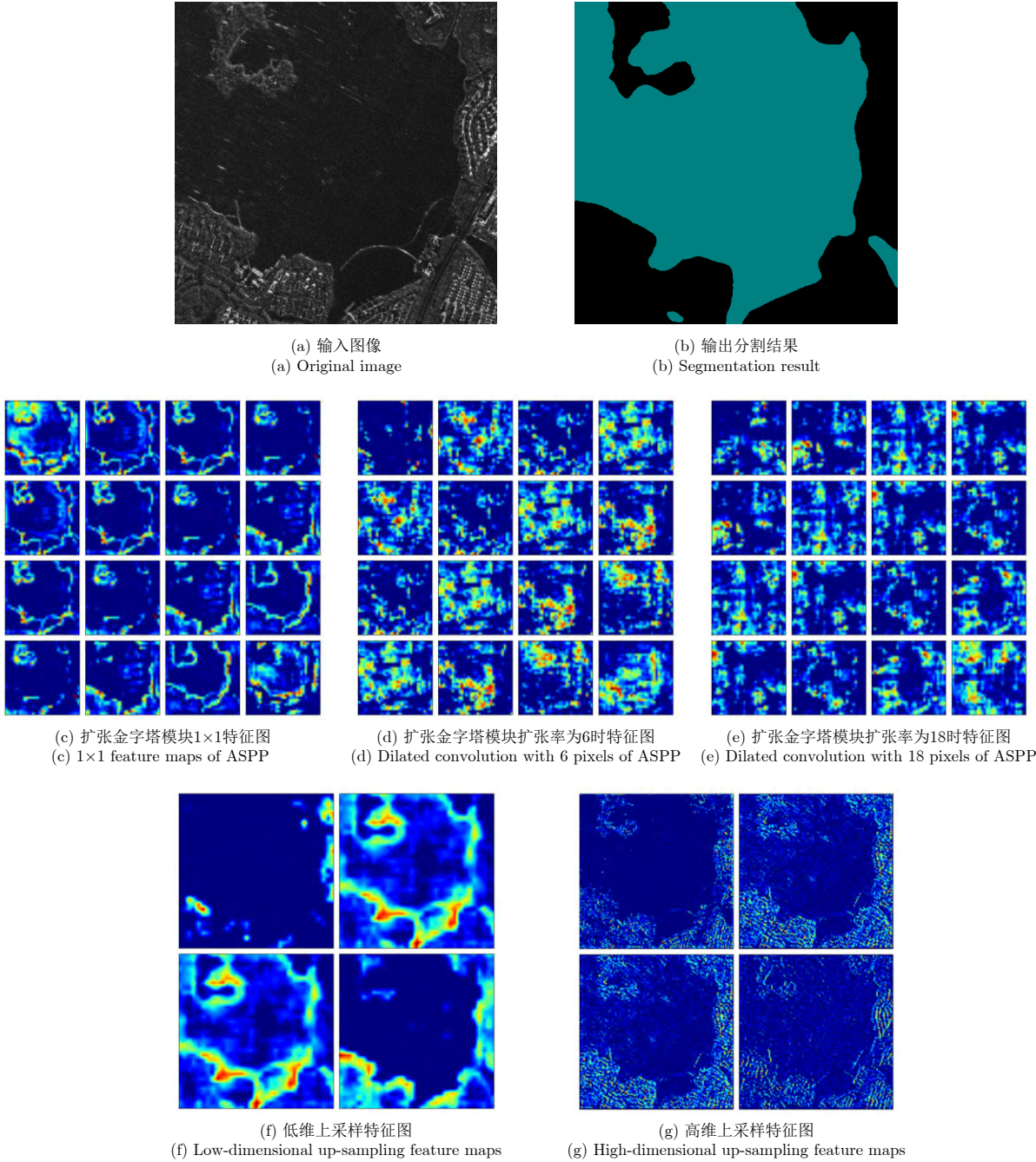


图 7 网络分割结果示意图

Fig. 7 Segmentation results of segmentation network

表 3 数据扩充对分割性能的影响

Tab. 3 Segmentation effects of data augmentation

扩充方法	像素准确度	交并比
未扩充	0.9569	0.9497
旋转	0.9806	0.9758
翻转	0.9620	0.9603
旋转+翻转	0.9887	0.9844

表 4 网络结构对分割性能的影响

Tab. 4 Segmentation effects of network structure

连接方式	像素准确度	交并比
直连	0.9312	0.9289
仅残差	0.9703	0.9681
仅密集	0.9679	0.9638
残差+密集	0.9887	0.9844

中的轮廓变化特征，但对陆地区域存在过分割现象，且水域和陆地区域分割结果不明确，分割结果

中错分类的像素点个数较多。而深度学习方法则普遍能准确地提取水域和陆地区域，相比于传统分割

方法,性能大幅提升,但也存在一些问题,例如Unet方法^[13]对水域和陆地交界轮廓不明确,丢失了完整的边界信息,而Deeplabv3+方法^[16]虽然分割效果较好,但仍然存在边界信息缺失现象。相比于Unet和Deeplabv3+,本文所提方法将密集分离卷积网络应用于SAR图像的特征提取,并且采用扩张金字塔池化模块和多个特征上采样模块逐层恢复输入图像的分辨率信息和语义信息,图8(i)所示的分割结果证明了所提方法在实现准确水域分割的基础上,也极大程度地保留了水域和陆地边界信息,且分割结果基本与图8(b)所示的分割标签保持一致,说明了所提方法具有较高的水域分割准确度。

为了定量分析各算法的分割效果,对比了像素准确度(Pixel Accuracy, PA)、分割交并比(Intersection over Union, IoU)以及分割速度等多个指标,结果如表5所示。

由表5可见,在像素精度和交并比上,深度学

习方法均取得了比传统方法较好的结果,且本文所提方法在2个评估指标上均领先于其他分割算法,与图8所示分割结果保持一致。此外,本文所提方法在运行速度上也具有一定优势,对于像素为 513×513 的裁剪后的输入小图像,处理速度为0.14 s/张,对于像素约为 $10,000 \times 10,000$ 的原始宽幅SAR图像,处理速度为3.5 s/张,可见所提方法相比于其他深度学习方法在以增加部分计算量为代价的前提下,兼顾了分割准确度和分割速度,具备一定的工程实用价值。

5.4 分割鲁棒性分析

由于不同模式下雷达工作参数有所差异,水域和陆地场景的后向散射系数会出现变化,进而导致SAR图像特征出现差异。为了证明所提算法的鲁棒性,采用上述利用滑动聚束(SL)模式的SAR图像训练好的网络模型,对多模式以及多极化方式下的星载SAR图像分割效果进行测试,结果如图9和图10所示。

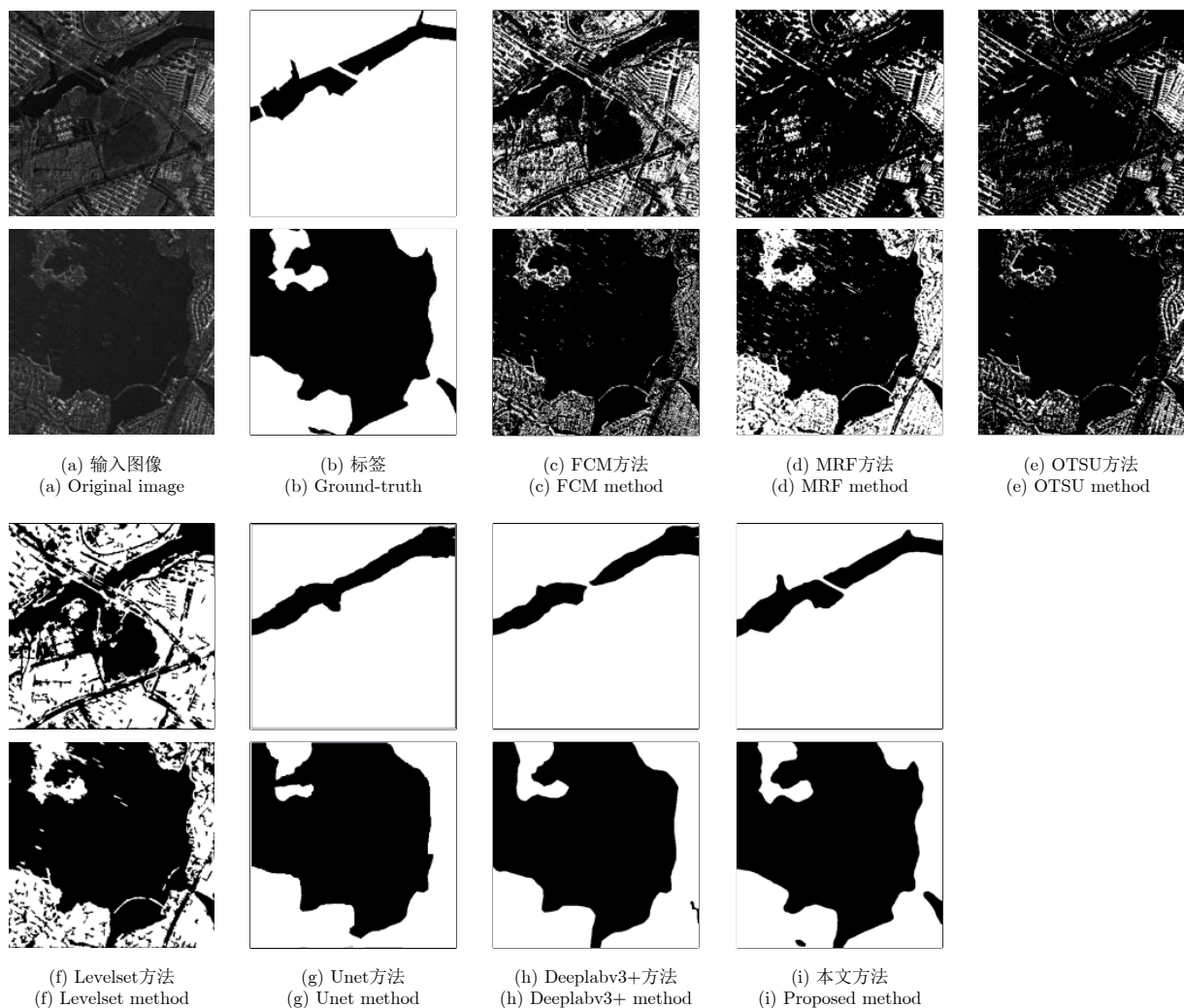


图8 各方法分割结果对比图

Fig. 8 Segmentation results of different methods

表 5 各水域分割算法性能对比
Tab. 5 Segmentation performance of different methods

方法类别	具体方法	像素准确度	交并比	小图速度(s)	大图速度(s)
传统方法	FCM	0.6710	0.4644	8.24	206.0
	MRF	0.5961	0.5430	2.29	57.25
	OTSU	0.6303	0.6108	0.06	1.50
	Levelset	0.7134	0.6868	3.41	85.25
	Unet	0.9533	0.9496	0.07	1.75
深度学习	DeepLabv3+	0.9672	0.9566	0.10	2.50
	所提方法	0.9887	0.9844	0.14	3.50
	理想值	1.0000	1.0000	—	—

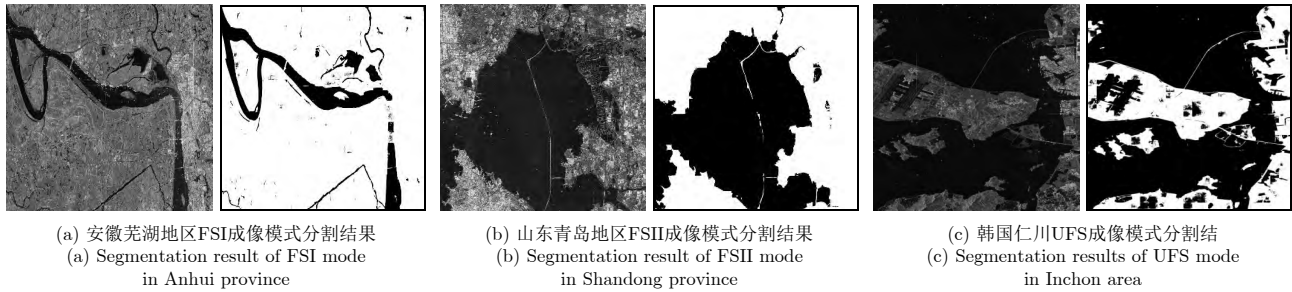


图 9 不同工作模式分割结果对比图
Fig. 9 Segmentation results of different imaging modes

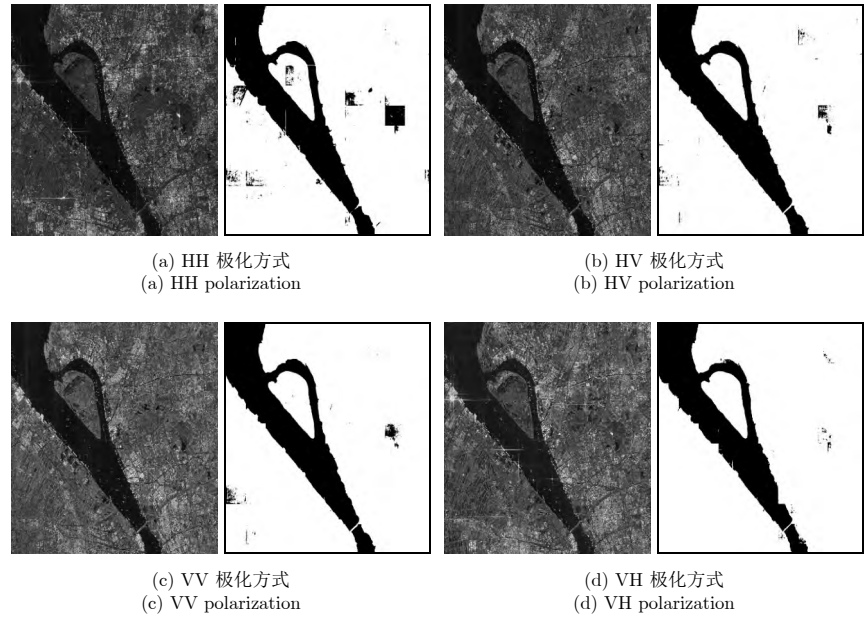


图 10 不同极化方式分割结果对比图
Fig. 10 Segmentation results of different imaging polarizations

由于原始图像尺寸较大，首先将原始图像无重叠裁剪到 513×513 像素，然后利用所提的分割算法对裁剪后的图像进行分割，并将分割结果按原始裁剪位置重新合并成大图。由图9可见，所提分割

模型在不同地区的多种工作模式下，取得了较好的分割效果，除了图9(c)中将机场区域错划分为水域外，图中的水域和陆地区域均被有效分割。另外，同一地区不同极化模式对应的原始SAR图像及对应

的分割结果如图10所示, 由于不同极化方式下部分陆地区域图像灰度值变化明显等, 所提方法将此类陆地区域错误划分成了水域, 导致分割性能下降。除了这部分区域外, 所提方法在不同极化方式下对其他区域均实现了有效的水域分割。

本文所提分割方法在各工作模式和极化方式下的交并比分割结果如表6所示, 可见分割结果与图9和图10分割结果保持一致, 充分说明所提方法具有较好的鲁棒性, 即利用滑动聚束模式、HH极化方式的SAR图像进行训练所获取的分割模型, 能有效实现多种成像模式和多种极化方式下的水域分割。

表 6 本文方法对多模式多极化下SAR图像的IoU分割结果
Tab. 6 IoU under multi-mode and multi-polarization
by the proposed method

工作模式/极化方式	HH	HV	VH	VV
SL (1 m)	0.9844	—	—	—
UFS (3 m)	0.9240	—	—	—
FSI (5 m)	0.9365	0.9542	—	—
FSII (10 m)	0.9549	0.9454	—	—
QPSI (8 m)	0.9605	0.9684	0.9686	0.9717

6 结论

本文针对实测SAR图像中的水域分割难题, 提出一种基于深度分离卷积和编码-解码架构的分割网络模型。首先收集了高分三号SAR图像, 并根据人工标注建立水域分割数据集, 研究了深度可分离卷积和扩张卷积两种新的卷积结构, 在此基础上, 提出基于密集分离卷积的特征提取网络, 并利用双线性插值进行上采样和利用卷积层重构上采样特征图, 最终得到与输入SAR图像逐像素对应的分割结果。实验结果证明本文所提方法在SAR图像水域的准确度、鲁棒性和实时性上, 均取得了较为理想的分割效果, 具备较好的工程实用价值。

参 考 文 献

- [1] 吴一戎. 多维度合成孔径雷达成像概念[J]. 雷达学报, 2013, 2(2): 135–142. doi: 10.3724/SP.J.1300.2013.13047.
WU Yirong. Concept of multidimensional space joint-observation SAR[J]. *Journal of Radars*, 2013, 2(2): 135–142. doi: 10.3724/SP.J.1300.2013.13047.
- [2] 艾加秋, 齐向阳, 禹卫东. 改进的SAR图像双参数CFAR舰船检测算法[J]. 电子与信息学报, 2009, 31(12): 2881–2885. doi: 10.3724/SP.J.1146.2008.01707.
AI Jiaqiu, QI Xiangyang, and YU Weidong. Improved two parameter CFAR ship detection algorithm in SAR images[J]. *Journal of Electronics & Information Technology*, 2009, 31(12): 2881–2885. doi: 10.3724/SP.J.1146.2008.01707.
- [3] 牛世林, 郭拯危, 李宁, 等. 星载SAR水域分割研究进展与趋势分析[J]. 聊城大学学报: 自然科学版, 2018, 31(2): 72–86.
NIU Shilin, GUO Zhengwei, LI Ning, *et al.* Research progress and trend analysis of water extraction by spaceborne SAR[J]. *Journal of Liaocheng University: Natural Science Edition*, 2018, 31(2): 72–86.
- [4] 安成锦, 牛照东, 李志军, 等. 典型Otsu算法阈值比较及其SAR图像水域分割性能分析[J]. 电子与信息学报, 2010, 32(9): 2215–2219. doi: 10.3724/SP.J.1146.2009.01426.
AN Chengjin, NIU Zhaodong, LI Zhijun, *et al.* Otsu threshold comparison and SAR water segmentation result analysis[J]. *Journal of Electronics & Information Technology*, 2010, 32(9): 2215–2219. doi: 10.3724/SP.J.1146.2009.01426.
- [5] 李智, 曲长文, 周强, 等. 基于SLIC超像素分割的SAR图像海陆分割算法[J]. 雷达科学与技术, 2017, 15(4): 354–358. doi: 10.3969/j.issn.1672-2337.2017.04.003.
LI Zhi, QU Changwen, ZHOU Qiang, *et al.* A sea-land segmentation algorithm of SAR image based on the SLIC superpixel division[J]. *Radar Science and Technology*, 2017, 15(4): 354–358. doi: 10.3969/j.issn.1672-2337.2017.04.003.
- [6] AMITRANO D, CIERVO F, DI MARTINO G, *et al.* Modeling watershed response in semiarid regions with high-resolution synthetic aperture radars[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(7): 2732–2745. doi: 10.1109/jstars.2014.2313230.
- [7] OTSU N. A threshold selection method from gray-level histograms[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62–66. doi: 10.1109/TSMC.1979.4310076.
- [8] SUI H G and XU C. Automatic extraction of water in high-resolution SAR images based on multi-scale level set method and Otsu algorithm[C]. *Proceedings of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, Australia, 2012: 453–457. doi: 10.5194/isprsarchives-XXXIX-B7-453-2012.
- [9] LIU Chun, YANG Jian, YIN Junjun, *et al.* Coastline detection in SAR images using a hierarchical level set segmentation[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(11): 4908–4920. doi: 10.1109/jstars.2016.2613279.
- [10] 侯彪, 胡育辉, 焦李成. SAR图像水域的改进Shearlet边缘检测[J]. 中国图象图形学报, 2010, 15(10): 1549–1554. doi: 10.11834/jig.20101019.
HOU Biao, HU Yuhui, and JIAO Licheng. Improved shearlet edge detection for waters of SAR images[J]. *Journal of Image and Graphics*, 2010, 15(10): 1549–1554. doi: 10.11834/jig.20101019.

- [11] LIU Zhongling, LI Fei, LI Ning, *et al.* A novel region-merging approach for coastline extraction from sentinel-1A IW mode SAR imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(3): 324–328. doi: 10.1109/LGRS.2015.2510745.
- [12] SHELHAMER E, LONG J, and DARRELL T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640–651. doi: 10.1109/TPAMI.2016.2572683.
- [13] RONNEBERGER O, FISCHER P, and BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention, Munich, Germany, 2015: 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 936–944. doi: 10.1109/CVPR.2017.106.
- [15] ZHAO Hengshuang, SHI Jianping, QI Xiaojuan, *et al.* Pyramid scene parsing network[C]. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 6230–6239. doi: 10.1109/CVPR.2017.660.
- [16] 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. doi: 10.1109/TPAMI.2017.2699184.
- [17] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* Semantic image segmentation with deep convolutional nets and fully connected CRFs[EB/OL]. arXiv: 1412.7062, 2014.
- [18] 张庆君. 高分三号卫星总体设计与关键技术[J]. 测绘学报, 2017, 46(3): 269–277. doi: 10.11947/j.AGCS.2017.20170049.
ZHANG Qingjun. System design and key technologies of the GF-3 satellite[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(3): 269–277. doi: 10.11947/j.AGCS.2017.20170049.
- [19] 丁赤飏, 刘佳音, 雷斌, 等. 高分三号SAR卫星系统级几何定位精度初探[J]. 雷达学报, 2017, 6(1): 11–16. doi: 10.12000/JR17024.
DING Chibiao, LIU Jiayin, LEI Bin, *et al.* Preliminary exploration of systematic geolocation accuracy of GF-3 SAR satellite system[J]. *Journal of Radars*, 2017, 6(1): 11–16. doi: 10.12000/JR17024.
- [20] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 1800–1807. doi: 10.1109/CVPR.2017.195.
- [21] YU F and KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. arXiv preprint arXiv: 1511.07122, 2015.
- [22] SZEGEDY C, VANHOUCKE V, IOFFE S, *et al.* Rethinking the inception architecture for computer vision[C]. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 2818–2826. doi: 10.1109/CVPR.2016.308.
- [23] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778. doi: 10.1109/CVPR.2016.90.
- [24] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, *et al.* Densely connected convolutional networks[C]. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2261–2269. doi: 10.1109/CVPR.2017.243.
- [25] NAIR V and HINTON G E. Rectified linear units improve restricted boltzmann machines[C]. Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 2010: 807–814.
- [26] IOFFE S and SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[EB/OL]. arXiv preprint arXiv: 1502.03167, 2015.
- [27] MIKOLOV T, SUTSKEVER I, CHEN Kai, *et al.* Distributed representations of words and phrases and their compositionality[C]. Proceedings of the 26th International Conference on Neural Information Processing Systems, Lake Tahoe, USA, 2013: 3111–3119.
- [28] HANSEN L K and SALAMON P. Neural network ensembles[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993–1001. doi: 10.1109/34.58871.
- [29] DENG Jia, DONG Wei, SOCHER R, *et al.* ImageNet: A large-scale hierarchical image database[C]. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, 2009: 248–255. doi: 10.1109/CVPR.2009.5206848.
- [30] VESE L A and CHAN T F. A multiphase level set framework for image segmentation using the Mumford and shah model[J]. *International Journal of Computer Vision*, 2002, 50(3): 271–293. doi: 10.1023/a:1020874308076..

作 者 简 介



张金松(1995-), 男, 山东德州人, 西安电子科技大学信号与信息处理专业博士研究生, 研究方向为SAR图像解译, 深度学习及SAR成像。

E-mail: jinsongxd@163.com



孙光才(1984-), 男, 湖北孝感汉川人。西安电子科技大学副教授, 博士生导师, 主要研究方向为多通道波束指向SAR 成像和 SAR 动目标成像。

E-mail: rsandsgc@126.com



邢孟道(1975-), 男, 浙江嵊州人。西安电子科技大学教授, 博士生导师, 主要研究方向为雷达成像、目标识别和天波超视距雷达信号处理。

E-mail: xmd@xidian.edu.cn