

基于多层深度特征融合的极化合成孔径 雷达图像语义分割

胡涛*, 李卫华**, 秦先祥***

空军工程大学信息与导航学院, 陕西 西安 710077

摘要 针对传统特征表征能力较弱的问题,提出了一种基于多层深度特征融合的极化合成孔径雷达图像语义分割方法;利用经过预训练的 VGG-Net-16 模型提取表征能力更强的多层图像特征,再将各层深度特征分别用于训练对应的条件随机场模型,最后将多个条件随机场模型的输出结果进行融合,实现了最终的图像语义分割。结果表明:与基于传统经典特征的方法相比,所提方法取得了最高的总体分类精度,说明所提方法采用的融合特征具有比传统特征更强的表征能力。

关键词 图像处理; 多层深度特征融合; 语义分割; 条件随机场; 卷积神经网络

中图分类号 TP391

文献标识码 A

doi: 10.3788/CJL201946.0210001

Semantic Segmentation of Polarimetric Synthetic Aperture Radar Images Based on Multi-Layer Deep Feature Fusion

Hu Tao*, Li Weihua**, Qin Xianxiang***

Institute of Information and Navigation, Air Force Engineering University, Xi'an, Shaanxi 710077, China

Abstract Aiming at the problem that the traditional feature representation ability is weak, we propose a polarization synthetic aperture radar image semantic segmentation method based on the multi-layer deep feature fusion. The pre-trained VGG-Net-16 model is used to extract multi-layer image features with strong representation ability, and then deep features of each layer are used to train the corresponding conditional random field model. The output results of multiple conditional random field models are finally merged to realize the final semantic segmentation of the images. The results show that compared with the methods based on classical features, the proposed method achieves the highest overall accuracy, indicating that the fusion features used by the proposed method have stronger representation ability than traditional features.

Key words image processing; multi-layer deep feature fusion; semantic segmentation; conditional random field; convolutional neural network

OCIS codes 100.4996; 100.5010

1 引 言

极化合成孔径雷达(PolSAR)图像语义分割的目的在于描述 PolSAR 图像不同区域的语义信息,由于 PolSAR 图像中语义类别具有分布复杂和多样可变等特点,因此有效实现 PolSAR 图像语义分割已成为极具挑战性的课题^[1]。

传统的 PolSAR 图像语义分割方法主要由两部

分组成:特征提取和分类器设计。传统方法通常基于 PolSAR 图像的散射特性来提取特征^[2-3],然后将提取到的特征输入到合适的分类器中。这些方法提取的特征一般是针对具体问题进行设计的手工特征,对先验知识的依赖程度较高。

目前,深度学习技术已取得了巨大进展,尤其是一些处理图像分类的方法,如卷积神经网络(CNN)等。AlexNet^[4]、VGG-Net^[5]和 ResNet^[6]等具有高

收稿日期: 2018-09-20; 修回日期: 2018-10-10; 录用日期: 2018-10-22

基金项目: 国家自然科学基金(41601436, 61403414, 61703423)、陕西省自然科学基金基础研究计划(2018JM4029, 2016JQ6070)

* E-mail: hu1862965@163.com; ** E-mail: lwh_kgd@163.com; *** E-mail: qinxianxiang@163.com

性能的 CNN 模型将多维图像数据直接作为输入,有效避免了传统方法繁琐的特征提取过程,被广泛应用于图像处理和计算机视觉领域。近年来,很多研究者开始将 CNN 用于 PolSAR 图像语义分割,得到了较好的性能^[7-8],但与基于传统人工特征的方法相比,这些方法中的 CNN 模型通常有很多参数,导致训练时间远超传统方法的训练时间。因此,在兼顾效率的前提下,提高传统 PolSAR 图像语义分割方法的分割精度显得尤为重要。

为了提高传统 PolSAR 图像语义分割方法的精度,本文在“特征提取+分类器设计”这一研究思路的基础上,提出了一种基于多层深度特征融合的 PolSAR 图像语义分割方法。首先采用经过预训练的 VGG-Net-16 模型提取图像不同层次的特征,避免了训练时间过长的问题,然后利用不同层次的深度特征分别训练差异性的条件随机场(CRF)模型,最后根据融合规则将不同层次深度特征对应的输出结果进行融合,得到最终的语义分割结果。

2 相关理论

2.1 分层深度特征

VGG-Net-16 模型是一种用于实现图像分类任务的 CNN 模型^[5],其中的“16”为该模型需要学习参数的层数。VGG-Net-16 模型主要由 5 个卷积层(Conv1~Conv13 共 13 层)和 3 个全连接层组成。从 Conv1 到 Conv5,每组卷积层分别包含 2、2、3、3、3 层卷积,每个卷积层都使用尺寸为 3×3 的卷积核。在 ImageNet 数据集上训练后,VGG-Net-16 模型中每个卷积层都可以作为一个特征提取器,提取目标不同卷积层的特征表达。

2.2 CRF 模型

CRF 模型是 Lafferty 等^[9]提出的一种概率图模型,可以表示为无向图模型 $G = \{V, E\}$,其中 V 为图中节点的集合, E 为节点间无向边的集合。

假设图像观测向量为 \mathbf{x} ,类别标签为 \mathbf{y} ,标记场的后验概率 $P(\mathbf{y}|\mathbf{x})$ ^[10]为

$$P(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{Z(\mathbf{x}, \boldsymbol{\theta})} \prod_{c \in C} \psi_c(y_c|\mathbf{x}, \boldsymbol{\theta}), \quad (1)$$

式中: $\boldsymbol{\theta}$ 为模型参数; $Z(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ 为配分函数; y_c 为势团 C 对应的标签; ψ_c 为定义在势团 C 上的势函数。

当只考虑单位置和双位置势函数时,(1)式可进一步改写为

$$P(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{Z(\mathbf{x}, \boldsymbol{\theta})} \exp \left[\sum_{i \in K} \psi_i(\mathbf{x}_i, \mathbf{y}, \boldsymbol{\theta}_\Lambda) + \sum_{i \in K} \sum_{j \in N_i} \psi_{ij}(y_i, y_j, \mathbf{x}, \boldsymbol{\theta}_\Gamma) \right], \quad (2)$$

式中: K 为所有节点的集合; N_i 为节点 i 的邻域; ψ_i 和 ψ_{ij} 分别为 CRF 模型的单位位置势函数和双位置势函数;模型参数 $\boldsymbol{\theta} = \{\boldsymbol{\theta}_\Lambda, \boldsymbol{\theta}_\Gamma\}$,其中 $\boldsymbol{\theta}_\Lambda$ 和 $\boldsymbol{\theta}_\Gamma$ 分别为单位置势函数和双位置势函数的权重向量; \mathbf{x}_i 为节点 i 处的特征向量。

3 基于多层深度特征融合的 PolSAR 图像语义分割方法

针对传统图像语义分割方法受限于人工特征表征能力较弱的问题,本课题组提出了一种基于多层深度特征融合的 PolSAR 图像语义分割方法,采用 VGG-Net-16 模型提取图像不同层级的深度特征,针对不同的特征分别训练对应的 CRF 模型,然后融合多个 CRF 模型的结果作为最终的输出。

3.1 PolSAR 图像预处理

对于 PolSAR 数据,每个像素点用 \mathbf{T} 矩阵的 9 维向量表示为

$$\mathbf{V} = (T_{11}, T_{22}, T_{33}, \text{real}(T_{12}), \text{imag}(T_{12}), \text{real}(T_{13}), \text{imag}(T_{13}), \text{real}(T_{23}), \text{imag}(T_{23})), \quad (3)$$

式中: T_{mn} 为对应位置的协方差矩阵元素,其中 $m, n = 1, 2, 3$; $\text{real}(\cdot)$ 和 $\text{imag}(\cdot)$ 分别为求实部和求虚部操作。为了避免相干斑噪声的影响,实验数据经过 Lee 滤波处理^[11]。本研究将在普通光学图像上预训练的 VGG-Net-16 模型作为特征提取器,其要求的输入为 3 个通道的归一化图像。因此,对滤波后的 PolSAR 图像采用文献[4]中的方法实现归一化,然后进行主成分分析(PCA)变换^[12]实现数据降维,将前 3 个主成分作为输入图像进行特征提取。

3.2 多层深度特征的提取

本研究为图像的每个像素点选择以其为中心且尺寸为 $W \times H \times 3$ 的邻域图像块参与训练,其中 W 和 H 分别为图像块的宽度和高度,3 为通道数。VGG-Net-16 模型要求输入尺寸为 $224 \text{ pixel} \times 224 \text{ pixel} \times 3 \text{ pixel}$ 的图像,因此对所有图像块采用双线性插值上采样到 $224 \text{ pixel} \times 224 \text{ pixel} \times 3 \text{ pixel}$ 。对于无法提取完整图像块的图像边缘像素,采用零填充策略进行补充^[7]。将所有插值后的图像块输入到 VGG-Net-16 模型中,得到图像块不同层级的特征表达,选择 Conv3-3、Conv4-3 和 Conv5-3 层的特征图,对

各层特征图分别求均值,并将均值作为图像块中心像素点的特征。VGG-Net-16 模型提取的 Conv3-3、Conv-3 和 Conv5-3 层特征分别被用来训练对应的 CRF 模型,得到对应的分类结果。

3.3 CRF 模型的建立

CRF 模型针对不同应用场景的需要定义不同的势函数。目前较为常见的势函数有多项式势函数、支持向量机(SVM)和多元对数回归(MLR)^[13]等。MLR 不要求观测图像满足正态分布,因此选择 MLR 作为单位位置势函数,即

$$\psi_i(y_i, \mathbf{x}, \boldsymbol{\theta}_A) = \sum_{l=1}^L \delta(y_i=l) \lg P(y_i=l | \mathbf{x}, \boldsymbol{\theta}_A), \quad (4)$$

$$P(y_i=l | \mathbf{x}, \boldsymbol{\theta}_A) = \begin{cases} \frac{\exp[(\boldsymbol{\theta}_A)_l^T \mathbf{x}_i]}{1 + \sum_{k=1}^{L-1} \exp[(\boldsymbol{\theta}_A)_k^T \mathbf{x}_i]}, & \text{if } l < L \\ \frac{1}{1 + \sum_{k=1}^{L-1} \exp[(\boldsymbol{\theta}_A)_k^T \mathbf{x}_i]}, & \text{if } l = L \end{cases}, \quad (5)$$

式中: δ 为 Kronecker 函数; L 为图像包含的类别数量; $l \in \{1, 2, \dots, L\}$; $(\boldsymbol{\theta}_A)_k$ 为第 k 类参数,所有 $(\boldsymbol{\theta}_A)_k$ 首尾连接得到的参数 $\boldsymbol{\theta}_A$ 在参数估计中确定。

双位置势函数 ψ_{ij} 不仅应考虑邻域中相邻位置 i 和 j 的特征对这两个点的语义标签 y_i 和 y_j 的作用,还应考虑对应语义标签之间的相互关系。本研究中的双位置势函数选择广义 Ising/Potts 模型^[14],即

$$\psi_{ij}(y_i, y_j, \mathbf{x}, \boldsymbol{\theta}_1) = \sum_{k, l \in \{1, 2, \dots, L\}} (\boldsymbol{\theta}_1)_{kl}^T \mathbf{u}_{ij}(\mathbf{x}) \delta(y_i=k) \delta(y_j=l), \quad (6)$$

式中: $\mathbf{u}_{ij}(\mathbf{x})$ 为联合特征向量,体现特征向量间的差异对标号的影响,向量 $\mathbf{u}_{ij}(\mathbf{x})$ 的权重 $\boldsymbol{\theta}_1$ 在参数估计中确定; $\boldsymbol{\theta}_1$ 为所有 $(\boldsymbol{\theta}_1)_{kl}$ ($k, l=1, 2, \dots, L$) 首尾连接得到的权重向量。

确定 CRF 模型后,需要在训练阶段估计模型参

数 $\boldsymbol{\theta}$ 。CRF 模型参数的估计方法有很多,如梯度上升和树重置信传播算法(TRW)等^[15-16],本研究采用限制优化方法来估计模型参数^[17]。参数确定后,可以得到每个像素点属于某个语义类别的概率,搜索使(2)式中的后验概率最大的最优解 \mathbf{y}^* ,即 $\mathbf{y}^* = \arg \max_{\mathbf{y}} \lg P(\mathbf{y} | \mathbf{x}, \boldsymbol{\theta})$,从而确定每个像素的类别标签,得到 VGG-Net-16 模型的不同层特征对应的语义分割精度。

3.4 融合规则

VGG-Net-16 模型提取的 Conv3-3、Conv4-3 和 Conv5-3 层特征分别用来训练同一个 CRF 模型,得到 3 个有差异的 CRF 子模型,利用(7)式对子分类器的结果进行融合:

$$P(\mathbf{y} | \mathbf{x}) = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3, \quad (7)$$

式中: P_1 、 P_2 和 P_3 分别为根据 Conv3-3、Conv4-3 和 Conv5-3 层特征得到的 CRF 模型后验概率; λ_1 、 λ_2 和 λ_3 分别为对应的自适应权值系数。根据得到的 3 个子 CRF 模型的总体分类精度(OA)^[7],归一化得到对应的自适应权值系数。在得到融合后的后验概率 $P(\mathbf{y} | \mathbf{x})$ 后,搜索其最大概率对应的语义类别,即为最终的语义分割结果。

4 实验设计与结果分析

为了验证本课题组所提方法的有效性和可靠性,选取 2 组实验数据集进行实验,原始数据采用如图 1 所示的由星载合成孔径雷达(ESAR)系统拍摄的德国 Oberpfaffenhofen 地区的 L 波段数据,以及由美国喷气推进实验室(JPL)机载合成孔径雷达(AIRSAR)系统获取的荷兰 Flevoland 地区 L 波段数据。这 2 组实验数据被广泛用于评估 PolSAR 图像语义分割算法的性能,其中 Oberpfaffenhofen 数据的图像大小为 1300 pixel×1200 pixel, Flevoland 数据的图像大小为 750 pixel×1024 pixel。

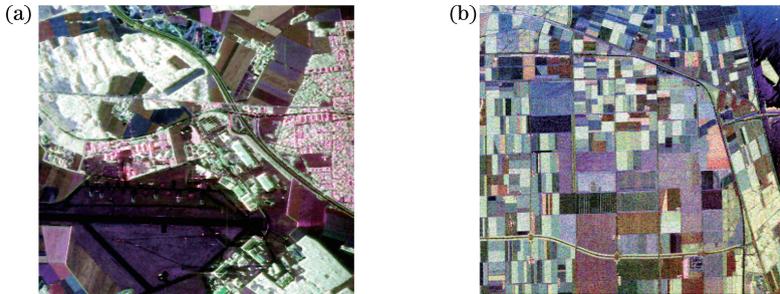


图 1 原始数据。(a) Oberpfaffenhofen 数据;(b) Flevoland 数据
Fig. 1 Original data. (a) Oberpfaffenhofen data; (b) Flevoland data

在实验中选择以下方法与所提方法进行对比：基于 Cloude 分解和 Freeman 分解所得特征的 CRF 分类方法(方法 1)；基于 Freeman 分解和协方差矩阵对角线元素所得特征的 CRF 分类方法(方法 2)；将上述 2 种方法中的特征串联融合所得特征的 CRF 分类方法(方法 3)；基于 CNN 的方法^[7](方法 4)。

表 1 所示为实验中经典方法所用到的特征。所

表 1 对比方法中用到的特征

Table 1 Features used in comparison methods

Cloude decomposition	Freeman decomposition	Covariance matrix diagonal element
$H, \alpha, A, \lambda_1, \lambda_2, \lambda_3$	P_s, P_d, P_v	C_{11}, C_{22}, C_{33}

4.1 基于 Flevoland 数据的实验结果

第 1 组实验数据为 Flevoland 数据。图 2(a)所示为 Flevoland 数据对应的地物分布参考图^[19]，包括 11 类地物，分别为豆类、森林、油菜籽、裸土、马铃薯、甜菜、小麦、豌豆、苜蓿、草地和水域，空白区域为未标记类别。

为了便于比较，文中涉及的实验都选择带有标记数据所占比例为 10% 的数据作为训练数据，所有

提方法的深度特征提取在深度学习平台 MatConvNet^[18]上完成。3 种经典方法与所提方法都采用同样的 CRF 模型，参数估计过程中的最大迭代次数设置为 1100。文献[7]中的迭代次数设置为 1100。4 组实验选取相同的训练样本和测试样本进行对比。分类性能综合评估指标为 OA、训练时间与测试时间。所有实验均在配置为 Intel Core i7 2.80 GHz 的处理器和内存为 8 GB 的计算机上完成。

带有标记的数据作为测试数据。基于大量的对比实验，将所提方法中图像块的宽度和高度设置为 $W = H = 12$ ，通过使用 VGG-Net-16 模型提取的 3 层特征训练得到的各子分类器得到 3 个 OA，分别为 80.88%、85.47% 和 87.02%；然后对其进行归一化处理，得到自适应权重系数的取值分别为 $\lambda_1 = 0.319$ ， $\lambda_2 = 0.337$ 和 $\lambda_3 = 0.344$ 。实验结果如图 2 所示。

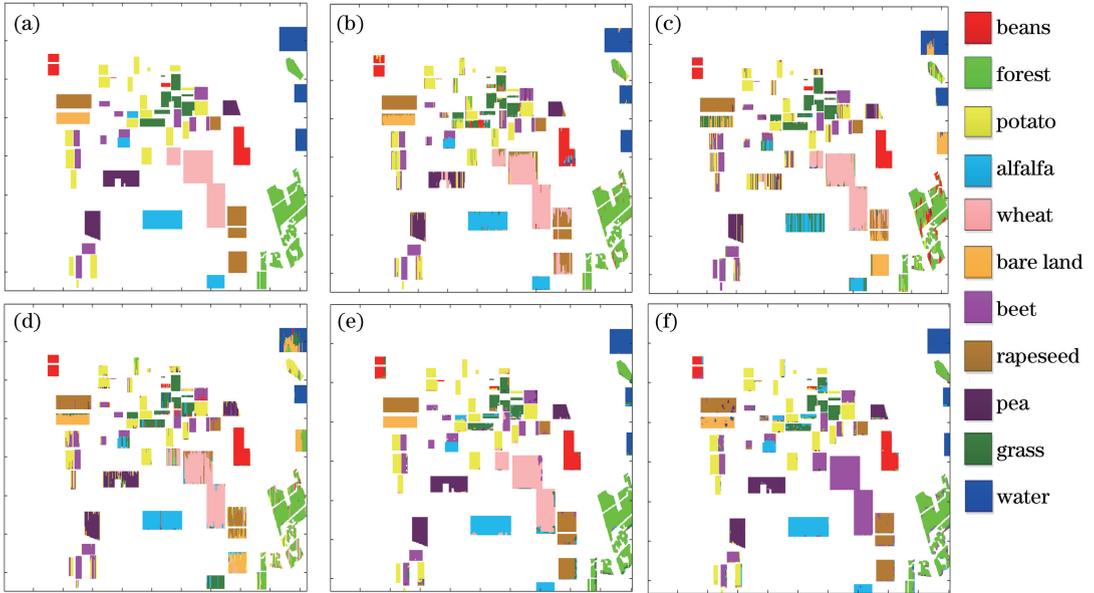


图 2 Flevoland 数据分类结果对比。(a)地物分布参考图^[19]；(b)方法 1；(c)方法 2；(d)方法 3；(e)方法 4；(f)所提方法

Fig. 2 Comparison of Flevoland data classification results. (a) Ground truth^[19]；(b) method 1；

(c) method 2；(d) method 3；(e) method 4；(f) proposed method in this paper

表 2 所示为定量评价结果。结合图 2 和表 2 可以看出：所提方法的整体效果优于 3 种经典方法，OA 达到了 92.22%；所提方法所有类别的分类精度都大于 80% (大部分大于 90%)，并且苜蓿、小麦和油菜籽等的分类精度均比 3 种经典方法的更高，这说明所提方法采用的多层深度融合特征具有比传统

特征更强的表征能力。方法 4 是目前性能较好的一种基于 CNN 的 PolSAR 图像语义分割方法。与方法 4 相比，所提方法的 OA 稍低，原因可能是所提方法使用的特征提取模型是预训练模型，对 OA 存在一定程度的影响，但所提方法需要训练的参数少于方法 4 的训练参数，训练时间和测试时间都远比方

表 2 Flevoland 数据下不同方法的性能对比

Table 2 Performance comparison of different methods under Flevoland data

Class	Classification accuracy for method 1 /%	Classification accuracy for method 2 /%	Classification accuracy for method 3 /%	Classification accuracy for method 4 /%	Classification accuracy for the proposed method in this paper /%
Beans	97.12	83.33	96.70	89.84	87.84
Forest	75.91	94.02	73.33	92.22	89.22
Potato	68.77	84.08	82.14	95.55	87.13
Alfalfa	60.99	89.22	71.92	95.15	96.51
Wheat	93.40	88.15	0.864	92.95	98.94
Bare land	51.42	87.13	90.38	99.89	91.26
Beet	91.34	90.34	89.55	90.56	84.84
Rapeseed	57.23	78.29	62.70	93.29	91.40
Pea	58.91	82.14	82.00	98.79	95.59
Grass	96.28	77.43	83.89	83.20	94.26
Water	72.14	97.04	52.63	94.90	99.41
OA /%	75.15	87.04	77.87	93.38	92.22
Training time /s	579	565	632	2851	1266
Test time /s	3.20	3.00	3.40	13.20	8.50

法 4 的训练时间和测试时间短,说明所提方法具有更高的实时性。此外,所提方法的特征维度较高,因此训练时间和测试时间比 3 种经典方法的训练时间和测试时间更长。

4.2 基于 Oberpfaffenhofen 数据的实验结果

第 2 组实验数据为德国 Oberpfaffenhofen 地区的 ESAR 图像,图 3(a)所示为对应的地物分布参考图^[20]。Oberpfaffenhofen 图像包含 3 类语义类别,分别为建筑区域、林地和开放区域,空白区域为未标记类别。为了便于比较,文中涉及的实验都选择带有标记数据所占比例为 10% 的数据作为训练数据,所有带有标记的数据作为测试数据。在提出的方法中,将图像块的宽度和高度设置为 $W=H=12$ 。通过使用 VGG-Net-16 模型提取的 3 层特征训练得到的各子分类器得到 3 个 OA,分别

为 60.85%、69.69% 和 85.23%,归一化后得到自适应权值系数的取值分别为 $\lambda_1=0.282$, $\lambda_2=0.323$ 和 $\lambda_3=0.395$ 。基于 Oberpfaffenhofen 数据的实验结果如图 3 所示。由图 3 可知,所提方法对 Oberpfaffenhofen 数据的所有语义类别都实现了较好的分割结果,比方法 1、方法 2、方法 3 的结果更接近于地物分布参考图,但稍差于方法 4 的分割结果。此外,还计算了各方法中每类语义类别的分类精度,并用 OA、训练时间和测试时间进行综合性能评估,结果如表 3 所示。通过对比图 3 和表 3 可以发现,所提方法的分类精度为 89.10%,远高于 3 种经典方法的分类精度,这进一步表明融合的深度特征相较于传统特征具有更强的表征能力。此外,虽然所提方法的 OA 低于方法 4 的 OA,但其训练时间和测试时间相对更短。

表 3 Oberpfaffenhofen 数据下不同方法的性能对比

Table 3 Performance comparison of different methods under Oberpfaffenhofen data

Class	Classification accuracy for method 1 /%	Classification accuracy for method 2 /%	Classification accuracy for method 3 /%	Classification accuracy for method 4 /%	Classification accuracy for the proposed method in this paper /%
Building area	69.65	64.52	71.25	85.60	92.22
Woodland	89.54	89.60	70.00	93.13	83.94
Open area	62.23	84.32	87.41	95.42	94.86
OA /%	69.15	80.41	80.00	92.61	89.10
Training time /s	684	659	779	3398	1476
Test time /s	3.40	3.30	3.80	16.90	9.70

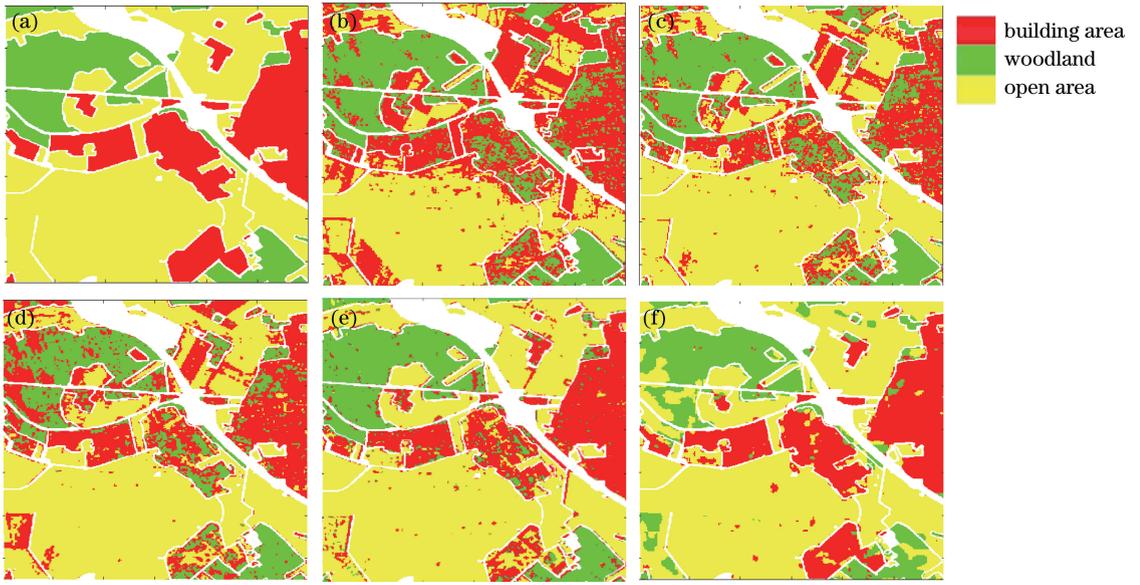


图 3 Oberpfaffenhofen 数据分类结果对比。(a)地物分布参考图^[20];(b)方法 1;(c)方法 2;(d)方法 3;(e)方法 4;(f)所提方法
 Fig. 3 Comparison of Oberpfaffenhofen data classification results. (a) Ground truth^[20]; (b) method 1;
 (c) method 2; (d) method 3; (e) method 4; (f) proposed method in this paper

4.3 VGG-Net-16 模型不同特征组合对语义分割精度的影响

为了进一步分析 VGG-Net-16 模型中不同卷积层的深度特征对所提方法的影响,在 Flevoland 实验数据下,通过组合不同卷积层的特征进行实验,得到不同卷积层特征组合所对应的语义分割结果。在实验中,采用与第 3 节中相同的方式,以实验图像的像素点为中心的图像块代表该像素点,经插值后输入到 VGG-Net-16 模型中,提取深度特征,然后将取均值后的深度特征输入到 CRF 中,得到对应特征层的结果,根据(7)式对不同卷积层特征组合中包含的特征层对应的结果进行融合,得到不同卷积层特征组合的结果。实验中,图像块宽度和高度均设置为 $W=H=12$ 。表 4 所示为利用 VGG-Net-16 模型的

Conv3-3、Conv4-3 和 Conv5-3 层特征在不同组合情况下得到的语义分割结果。由表 4 可知,在 Flevoland 数据集下,随着特征层数增加,单层深度特征对应的精度不断提高,原因是深的卷积层可以提取图像更加抽象的特征信息。此外,在两组深度特征进行组合的情况下,组合中包含的特征层数越多,相应的分割精度越高,任意两组特征进行组合得到的分割精度都比组合内单层特征得到的分割精度高,说明将不同的特征进行融合可以有效地提高分割精度。本研究提出的方法将 Conv3-3、Conv4-3 和 Conv5-3 层特征进行融合,取得了比任意单层特征或两两组合特征更高的分割精度,这说明所提方法将不同的特征组合后可以有效地融合不同层特征包含的从具体到抽象的细节信息。

表 4 不同特征组合时的分类精度

Table 4 Classification with different combination of features

Conv3-3	Conv4-3	Conv5-3	Conv3-3+ Conv4-3	Conv3-3+ Conv5-3	Conv4-3+ Conv5-3	Proposed method in this paper	%
80.88	85.47	87.02	87.66	89.97	90.11	92.22	

5 结 论

结合 PolSAR 的实际应用需求,本课题组提出了一种基于多层深度特征融合的 PolSAR 图像语义分割方法,该方法利用 VGG-Net-16 模型提取图像不同层次的深度特征,然后采用 CRF 模型对不同层特征进行分类,最后将不同层特征对应的结果进行

融合,实现了最终的语义分割。实验结果表明,在利用 VGG-Net-16 模型提取特征进行图像语义分割时,将不同层深度特征进行融合,得到了比单层深度特征和传统特征更好的性能,说明所提方法可以有效融合多层深度特征不同层级的信息,比传统方法特征具有更强的特征表征能力。此外,与基于 CNN 方法的结果进行对比后发现,所提方法具有更高的

效率,进一步验证了所提方法的有效性。

参 考 文 献

- [1] Yang W, Zhang X, Chen L J, *et al.* Semantic segmentation of polarimetric SAR imagery using conditional random fields [C] // IEEE International Geoscience and Remote Sensing Symposium, July 25-30, 2010, Honolulu, Hawaii, USA. New Jersey: IEEE, 2010: 1593-1596.
- [2] 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.
- [3] Wan J H, Zang J X, Liu S W. Fusion and classification of SAR and optical image with consideration of polarization characteristics [J]. Acta Optica Sinica, 2017, 37(6): 0628001.
万剑华, 臧金霞, 刘善伟. 顾及极化特征的 SAR 与光学影像融合与分类 [J]. 光学学报, 2017, 37(6): 0628001.
- [4] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [5] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. Computer Science, 2014, 32(74): 2965-2971.
- [6] He K M, Zhang X Y, Ren S Q, *et al.* Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [J]. Proceedings of the IEEE, 2015, 28: 1026-1034.
- [7] Zhou Y, Wang H P, Xu F, *et al.* Polarimetric SAR image classification using deep convolutional neural networks [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12): 1935-1939.
- [8] Hou B, Kou H D, Jiao L C. Classification of polarimetric SAR images using multilayer autoencoders and superpixels [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(7): 3072-3081.
- [9] Lafferty J, McCallum A, Pereira F. Conditional random fields: probabilistic models for segmenting and labeling sequence data [C] // Proceedings of the 18th International Conference on Machine Learning, Morgan Kaufmann, San Francisco, CA. San Francisco: Morgan Kaufmann, 2001: 282-289.
- [10] LI S Z. Markov random field modeling in computer vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(10): 939-954.
- [11] Han P, Han B B. Lee filter of PolSAR image based on typical scattering difference index [J]. Systems Engineering and Electronics, 2018, 40(2): 287-294.
韩萍, 韩宾宾. 基于典型散射差异指数的 PolSAR 图像 Lee 滤波 [J]. 系统工程与电子技术, 2018, 40(2): 287-294.
- [12] Mika S, Schölkopf B, Smola A, *et al.* Kernel PCA and de-noising in feature spaces [C] // Proceedings of Conference on Advances in Neural Information Processing Systems. Cambridge: The MIT Press, 1999: 536-542.
- [13] Lee C H, Schmidt M, Murtha A, *et al.* Segmenting brain tumors with conditional random fields and support vector machines [C] // Proceedings of the 1st International Conference on Computer Vision for Biomedical Image Applications. Heidelberg: Springer, 2005: 469-478.
- [14] Kumar S, Hebert M. Discriminative fields for modeling spatial dependencies in natural images [C] // Proceedings of Conference on Advances in Neural Information Processing Systems. Cambridge: The MIT Press, 2004: 1531-1538.
- [15] Domke J. Learning graphical model parameters with approximate marginal inference [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(10): 2454-2467.
- [16] Parikh D, Batra D. CRFs for image classification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(5): 460-472.
- [17] Krishnapuram B, Carin L, Figueiredo M, *et al.* Sparse multinomial logistic regression: fast algorithms and generalization bounds [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(6): 957-968.
- [18] Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for MATLAB [C] // Proceedings of the 23rd ACM International Conference on Multimedia, October 26-30, 2015, Brisbane, Australia. New York: ACM, 2015: 689-692.
- [19] Yu P, Qin A K, Claudi D A. Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty [J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 50(4): 1302-1317.
- [20] Liu B, Hu H, Wang H Y, *et al.* Superpixel-based classification with an adaptive number of classes for polarimetric SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(2): 907-924.