

基于精调的膨胀编组-交叉CNN的PolSAR地物分类

肖东凌 刘畅*

(中国科学院电子学研究所 北京 100190)

(中国科学院大学 北京 100049)

摘要: 在极化合成孔径雷达(PolSAR)地物分类研究中, 基于实数CNN的分类算法无法充分利用PolSAR图像丰富的通道相位信息, 并且在逐像素预测中存在大量冗余计算, 导致分类效率较低。针对以上问题, 该文采用一种复数域下的像素映射深度模型, 实现低采样率下精确且高效的PolSAR地物分类。为充分使用PolSAR数据的通道相位信息, 该文基于一种编组-交叉卷积网络(GC-CNN)将分类模型推广到复数域, 并利用网络提取的复数特征及其对应的相位和幅度来实现更精确的分类; 为加快分类速度, 该文还采用了一种精调的膨胀编组-交叉卷积网络(FDGC-CNN)来实现像素到像素的直接映射, 并进一步提升了分类精度。在基于AIRSAR平台的16类地物数据和E-SAR平台的4类地物数据的实验中, 该文采用的FDGC-CNN模型相较于SVM分类器和实数CNN模型, 能够更准确和更高效地实现多类别地物分类, 全局分类精度分别为96.94%和90.07%、总耗时4.22 s和4.02 s。

关键词: 极化合成孔径雷达(PolSAR); 地物分类; 膨胀卷积; 编组-交叉卷积

中图分类号: TN958

文献标识码: A

文章编号: 2095-283X(2019)04-0479-11

DOI: 10.12000/JR19039

引用格式: 肖东凌, 刘畅. 基于精调的膨胀编组-交叉CNN的PolSAR地物分类[J]. 雷达学报, 2019, 8(4): 479-489. doi: 10.12000/JR19039.

Reference format: XIAO Dongling and LIU Chang. PolSAR terrain classification based on fine-tuned dilated group-cross convolution neural network[J]. *Journal of Radars*, 2019, 8(4): 479-489. doi: 10.12000/JR19039.

PolSAR Terrain Classification Based on Fine-tuned Dilated Group-cross Convolution Neural Network

XIAO Dongling LIU Chang*

(*Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China*)

(*University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: In the study of the terrain classification based on the Polarimetric Synthetic Aperture Radar (PolSAR), the algorithms based on general CNN do not fully utilize the phase information in different channels, and the pixel-by-pixel classification strategy with extensive redundant computation is inefficient. To mitigate these problems, a deep pixel-to-pixel mapping model in the complex domain is used for achieving a fast and accurate PolSAR terrain classification at a low sampling rate. To completely utilize the phase information, this study uses Group-Cross CNN to extend the original model to the complex domain allowing complex-number input signals and significantly improving the classification accuracy. In addition, to speed up the algorithm, a Fine-tuned Dilated Group-Cross CNN (FDGC-CNN) was adopted to directly achieve pixel-to-pixel mapping as well as improve accuracy. We verified the adopted model on two PolSAR images comprising 16 classes terrains from the AIRSAR and 4 classes terrains from the E-SAR. According to our model, the overall classification accuracies were 96.94% and 90.07% respectively while the running time was 4.22 s and 4.02 s respectively. Therefore, FDGC-CNN achieved better accuracy with higher efficiency compared to SVM and traditional CNN.

Key words: PolSAR; Terrain classification; Dilated convolution; Group-cross convolution

收稿日期: 2019-03-04; 改回日期: 2019-03-19; 网络出版: 2019-05-05

*通信作者: 刘畅 cliu@mail.ie.ac.cn *Corresponding Author: LIU Chang, cliu@mail.ie.ac.cn

基金项目: 国家自然科学基金(61471340), 国家重点研发计划(2017YFB0503001)

Foundation Items: The National Natural Science Foundation of China (61471340), The State Key Research Development Program (2017YFB0503001)

责任编辑: 代大海 Corresponding Editor: DAI Dahai

1 引言

极化合成孔径雷达(Polarimetric Synthetic Aperture Radar, PolSAR)凭借全天时、全天候和高分辨率成像的能力以及丰富的地物目标散射信息和通道相位信息,被广泛运用于地物分类、目标检测和相干变化检测等领域。在地物分类研究中,表面平滑的地物纹理特征区分度不高,而PolSAR丰富的通道相位信息则能有效地区分纹理差异不大的地物,因此PolSAR数据在该领域受到了极大的关注。

传统的PolSAR图像处理算法按照特征提取、训练分类器两个基本步骤独立展开^[1-4],特征和分类器之间的交互较弱,模型连贯性较差。相较于传统的图像算法,卷积神经网络(Convolution Neural Network, CNN)能根据具体的任务来自适应地捕获特征,并联合特征网络训练分类器,具有更好的灵活性和连贯性,因此能够得到更加准确的分类结果。目前,CNN已被广泛运用到PolSAR图像处理领域,例如Zhang等人^[5]用复元素CNN来实现PolSAR地物分类;Xu等人^[6]利用CNN实现PolSAR图像相干变化的检测;Bi等人^[7]用基于图的CNN以及Chen等人^[8,9]采用极化特征驱动的CNN算法实现精确的PolSAR地物分类。

在基于PolSAR数据的地物分类研究中,深度分类网络的训练数据一般是后向散射矩阵 \mathbf{S} 以及由 \mathbf{S} 变换得到的协方差矩阵 \mathbf{C} 和相干矩阵 \mathbf{T} ^[10-13],而非实数CNN模型输入的3通道RGB图像。 \mathbf{C} 矩阵和 \mathbf{T} 矩阵的非对角元素都是复数,且 \mathbf{S} 、 \mathbf{C} 和 \mathbf{T} 矩阵都包含能有效区分不同地物的通道相位信息。其中PolSAR的后向散射矩阵 \mathbf{S} 如式(1)所示。

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} = \begin{bmatrix} |S_{HH}| e^{j\phi_{HH}} & |S_{HV}| e^{j\phi_{HV}} \\ |S_{VH}| e^{j\phi_{VH}} & |S_{VV}| e^{j\phi_{VV}} \end{bmatrix} \quad (1)$$

其中, S_{HH} 和 S_{VV} 项包含同极化通道回波功率,而 S_{HV} 和 S_{VH} 项则包含交叉极化通道回波功率。协方差矩阵 \mathbf{C} 矩阵的表示见式(2)。式中 $\langle \cdot \rangle$ 为时间或空间集合平均。

$$\mathbf{C} = \begin{bmatrix} \langle |S_{HH}|^2 \rangle & \sqrt{2} \langle S_{HH} S_{HV}^* \rangle & \langle S_{HH} S_{VV}^* \rangle \\ \sqrt{2} \langle S_{HV} S_{HH}^* \rangle & 2 \langle |S_{HV}|^2 \rangle & \sqrt{2} \langle S_{HV} S_{VV}^* \rangle \\ \langle S_{VV} S_{HH}^* \rangle & \sqrt{2} \langle S_{VV} S_{HV}^* \rangle & \langle |S_{VV}|^2 \rangle \end{bmatrix} \quad (2)$$

为了充分利用PolSAR丰富的通道相位信息,Zhang等人^[5]将CNN网络底层实现中的变量修改为复数;然而该算法在实现上较为繁琐,并且模型的迁移推广性不强,如推广到膨胀卷积神经网络

(Dilated CNN)^[14]、全卷积神经网络(FCN)^[15]等复杂的CNN模型。因此本文基于一种编组-交叉卷积神经网络(Group-Cross CNN, GC-CNN)来模拟复数乘法运算并支持复数信号的输入和传递。该结构实质上是两组并行的交叉且交互的实数CNN网络,因此更容易实现和迁移推广到复杂的网络中。

在算法效率方面,逐像素点(pixel-by-pixel)分类的算法^[5-9]在提取相邻像素特征时存在大量的重复运算,因此分类较大尺寸图像的时间开销很大。而像素到像素(pixel-to-pixel)映射的算法^[14,15],如FCN和Dilated CNN则能有效剔除冗余计算,实现端到端的高效预测。其中基于FCN的算法^[15-17]由于在解码网络上采样的过程中丢失了大量的边缘细节信息,导致分类结果边缘粗糙;而基于Dilated CNN的算法^[14]通过带孔洞的卷积核实现尺寸不变的输出特征图,这样既能实现快速的端到端分类又能保持目标的细节信息,因此本文采用基于Dilated CNN的膨胀编组-交叉卷积网络(Dilated Group-Cross CNN, DGC-CNN)来提升分类效率。

在DGC-CNN的训练中,需要和输入图像尺寸相同的标注图,这对于PolSAR图像来说是困难的,因为带标注的PolSAR图像较少,且大部分带标注的PolSAR数据都存在未标注的区域。因此,本文用总采样率小于1.5%的标注像素预训练GC-CNN,并将其参数直接迁移到匹配的DGC-CNN中来实现效率提升。考虑到参数的直接迁移会导致DGC-CNN分类精度的损耗,本文沿用训练GC-CNN时的训练像素对迁移后的DGC-CNN进行精调,来捕获更多空间上下文语义特征,恢复并进一步提升GC-CNN模型的分类精度,文中称该算法为精调的膨胀编组-交叉卷积神经网络(Fine-tuned Dilated Group-Cross CNN, FDGC-CNN)。

综上,本文的结构安排如下:第2节介绍卷积和编组-交叉卷积单元;第3节详细介绍编组-交叉卷积神经网络(GC-CNN)和精调的膨胀编组-交叉卷积网络(FDGC-CNN);第4节为实验部分,基于AIRSAR平台的16类规则地物数据和E-SAR平台4类不规则地物数据,验证了本文采用的网络和算法的准确性和高效性。

2 编组-交叉卷积单元

2.1 CNN原理

CNN启发于视觉成像原理的研究,它通过使用不同的卷积核来模拟可视皮层细胞对边缘和方向的感知,随后连接非线性激活函数得到图像一系列复杂的特征表达^[18]。CNN模型包含卷积层、池化层和激活层,其中卷积层由多组卷积核组成;第 i 组卷积核的输出 I_i' 的数学表达式如下

$$I'_j = \sum_{i=1}^c \left[\sum_{(p+i, q+j) \in W} I_{p+m, q+n}^i K_{m, n}^{j i} \right] + b_j, \quad (3)$$

$$K^j \in \mathbb{R}^{c \times s \times s}$$

其中, c 和 s 分别为输入图像通道数和卷积核窗口尺寸; $I_{m, n}^i$ 表示输入图像的第 i 通道、坐标为 (m, n) 的值; $K_{m, n}^{j i}$ 表示第 j 组卷积核 K_j 的第 i 通道、坐标为 (m, n) 的值; b_j 为第 j 组卷积核的偏置项; W 为卷积在图像上滑动的区域, (p, q) 为滑动窗口的左上角坐标; 之后将 I'_j 输入激活层进行非线性表示, 增加特征的非线性可分性, 常见的激活函数有Sigmoid, tanh和ReLU^[19]以及其它相关变种如CReLU^[20], PReLU^[21]等; 最后, 池化层对激活后的输出进行降采样, 增加了每个像素的感受野, 使空间形状灰度特征汇聚到更抽象的特征空间, 并且减少了下一层卷积的计算量。

在多层CNN模型中, 浅层卷积捕获局部的边缘、角点特征, 深层卷积捕获更抽象且强语义的高维特征。常见的多层CNN特征模型如VGG^[22], GoogLeNet^[23], ResNet^[24], DenseNet^[25]等已用于目标检测、分类和分割等诸多研究领域。

2.2 编组-交叉卷积原理

PolSAR通道之间的相位信息反映散射地物的回波时延, 携带着地物的空间结构信息, 对提升地物分类精度起着重要的作用。为了不丢失PolSAR的通道相位信息, 本文采用编组-交叉卷积神经网络(GC-CNN)来实现复数信号的输入、传输和输出。该网络将输入信号的实部通道和虚部通道编组, 并分别输入两组并行且交叉交互的实数CNN网络, 将实数模型推广到复数域, GC-CNN结构如图1所示。

复数域下的输入图像 I 由实部和虚部组成, 因此本文将编组为两组相同通道数的实数图像组, 分别表示实部和虚部。设复数输入图像 I 的表达式如下

$$I = I_{\text{real}} + jI_{\text{imag}}, I \in \mathbb{C}^{m \times n \times c}, \quad (4)$$

$$I_{\text{real}} \in \mathbb{R}^{m \times n \times c}, I_{\text{imag}} \in \mathbb{R}^{m \times n \times c}$$

其中, c 为 I 的通道数, I_{real} 和 I_{imag} 分别表示实部和虚部的图像组。同样, 编组-交叉卷积核也被划分为两组实数信号, 分别表示实部和虚部; 窗口尺寸为 $s \times s$ 的编组-交叉卷积层如式(5)所示。

$$K = K_{\text{real}} + jK_{\text{imag}}, K \in \mathbb{C}^{c \times s \times s \times o}, \quad (5)$$

$$K_{\text{real}} \in \mathbb{R}^{c \times s \times s \times o},$$

$$K_{\text{imag}} \in \mathbb{R}^{c \times s \times s \times o}$$

其中, c, o 分别是卷积层的输入和输出维度, $K_{\text{real}}, K_{\text{imag}}$ 分别是实部卷积核和虚部卷积核。

实数卷积运算实质上是多组3维卷积核和图像的内积, 本文记该运算为Conv(\cdot); 同时编组-交叉卷积GC-Conv(\cdot)是2组实数卷积的交叉结合。第 o 组编组-交叉的输出 I'_o 可表示为

$$I'_o = \text{GC-Conv}(I, K_o)$$

$$= \text{GC-Conv}(I_{\text{real}} + jI_{\text{imag}}, K_{o, \text{real}} + jK_{o, \text{imag}})$$

$$= \text{Conv}(I_{\text{real}}, K_{o, \text{real}}) - \text{Conv}(I_{\text{imag}}, K_{o, \text{imag}})$$

$$+ j[\text{Conv}(I_{\text{imag}}, K_{o, \text{real}}) + \text{Conv}(I_{\text{real}}, K_{o, \text{imag}})] \quad (6)$$

从式(6)可以看出编组-交叉卷积核的输出可以拆分成实部输入 I_{real} 和实部卷积核 $K_{o, \text{real}}$ 的卷积结果、虚部输入 I_{imag} 和虚部卷积核 $K_{o, \text{imag}}$ 的卷积结果、实部输入 I_{real} 和虚部卷积核 $K_{o, \text{imag}}$ 的卷积结果、虚部输入 I_{imag} 和实部卷积核 $K_{o, \text{real}}$ 的卷积结果, 共4个通道; 之后对4通道实数特征图进行线性合并, 即组合成编组-交叉的输出 $I' = I'_{\text{real}} + jI'_{\text{imag}}$ 。

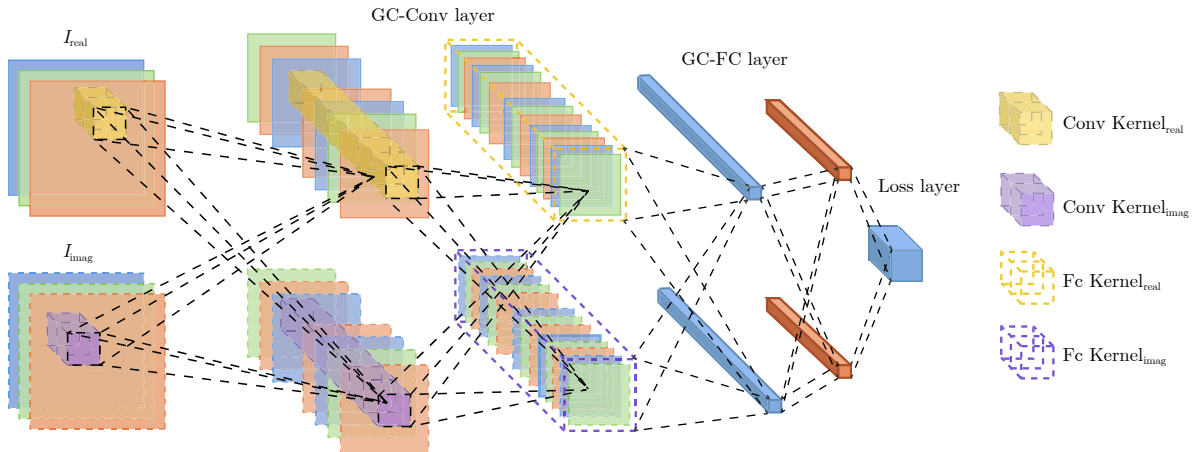


图1 编组-交叉卷积神经网络结构示意图

Fig. 1 Structure of group-cross convolution neural network

值得注意的是, ReLU激活函数也应推广到复数域, 记作 ReLU' (见图2)。具体形式如下所示。

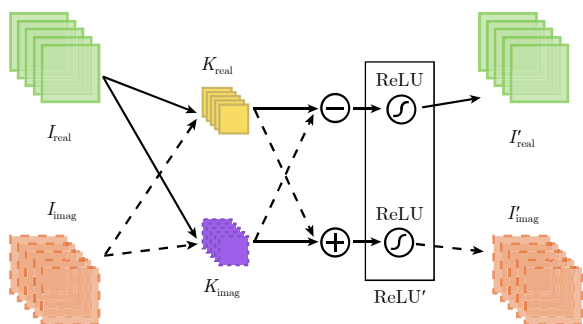


图 2 编组-交叉卷积层结构示意图

Fig. 2 Structure of group-cross convolution layer

$$\begin{aligned} \mathbf{I}' &= \text{ReLU}'(\mathbf{I}) = \text{ReLU}(I_{\text{real}}) + j\text{ReLU}(I_{\text{imag}}) \\ &= \max\{0, I_{\text{real}}\} + j\max\{0, I_{\text{imag}}\} \end{aligned} \quad (7)$$

3 精调的膨胀编组-交叉卷积神经网络

本节将详细地介绍编组-交叉卷积神经网络(GC-CNN)和精调的膨胀编组-交叉卷积神经网络(DGC-CNN)。

3.1 编组-交叉卷积神经网络

算法首先以低采样率($<2\%$)在PolSAR图像中随机采样, 并取 10×10 的像素窗作为中心像素的特征区域, 输入3层的GC-CNN。GC-CNN的参数集中在全连接层, 在训练样本不足的情况下, 模型容易过拟合, 因此, 采用dropout和 L_2 正则来泛化模型。为了使该网络参数能够很好地迁移到下一节介绍的DGC-CNN中, 采用无边缘填充(Padding)的卷积方式进行特征提取, 从而使得在每层 3×3 卷积

层后, 输入图像尺寸都将减少2个像素, 如图3(a1)所示。因此, 10×10 的特征像素窗通过2层 3×3 编组-交叉卷积层后, 可以得到尺寸为 1×1 的复数特征图。最后将复数特征图输入 1×1 编组-交叉卷积层以提取高维复数特征向量, 用于中心像素点的分类。GC-CNN的具体结构如图3(b)所示。为将复数特征图转换为实数的分类结果, 模型采用图4所示的结构将复数特征图及其相位和幅度转换为实数输出。

该结构将复数输出特征图的实部和虚部及其对应的幅度和相位在通道上拼接, 并输入 4×1 的全连接层来获得各类别的分类概率。最后计算分类损失, 更新GC-CNN参数。

在模型的训练过程中, 对于样本数较少的地物, 网络学习较为困难。针对此问题, 算法采用RetinaNet^[26]中的焦点损失函数(focal loss)来减少易分样本的损失, 进而使网络的训练更关注样本数较少的类别以及训练误差较大的样本点。记复数输出特征图为 $\mathbf{I}' = \lambda e^{j\varphi}$, focal loss可由式(8)和式(9)计算

$$p(\mathbf{I}') = \text{Softmax}(w_r \Re(\mathbf{I}') + w_i \Im(\mathbf{I}') + w_m \lambda + w_p \varphi) \quad (8)$$

$$L_{\text{focal}}(\mathbf{I}', \mathbf{I}) = -\alpha (1 - p_{\mathbf{I}}(\mathbf{I}'))^\gamma \lg(p_{\mathbf{I}}(\mathbf{I}')) \quad (9)$$

其中, \mathbf{I} 表示当前样本的真实标注, $p_{\mathbf{I}}(\mathbf{I}')$ 为输出类别 \mathbf{I} 的预测概率。

3.2 精调的膨胀编组-交叉卷积神经网络

GC-CNN在提取相邻像素特征时存在大量的冗余计算, 且分类效率较低。为了提升GC-CNN的分类速度, 本文将GC-CNN的参数迁移到相同参数规模的膨胀编组-交叉卷积网络(DGC-CNN)中, 如

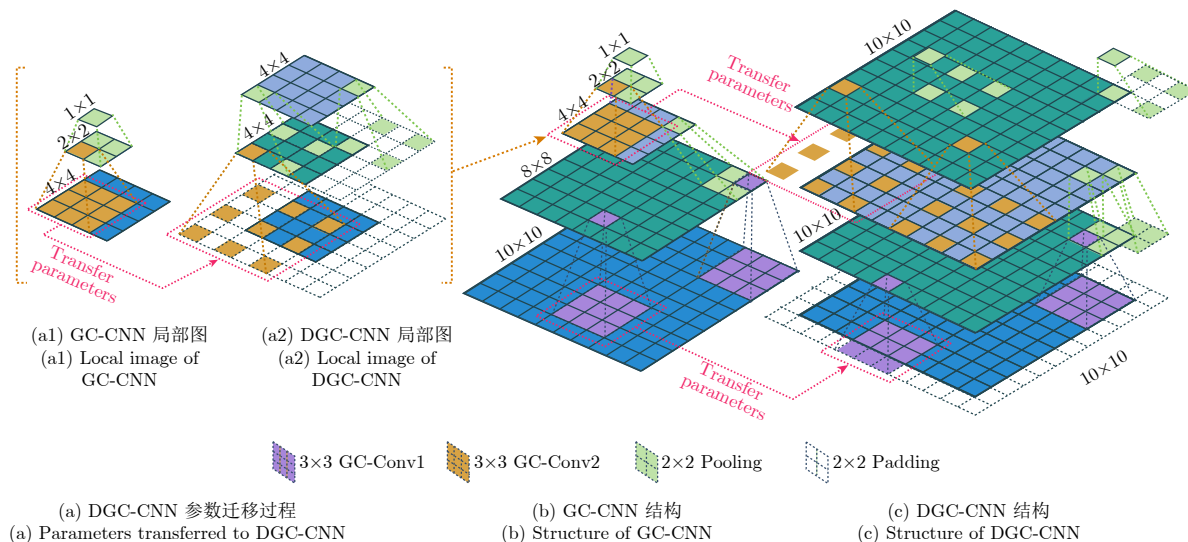


图 3 膨胀编组-交叉卷积神经网络结构示意图

Fig. 3 Structure of Dilate Group-Cross CNN (DGC-CNN)

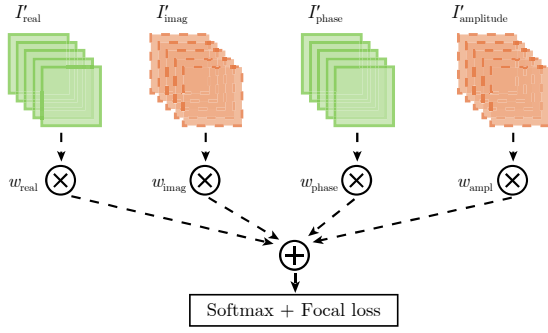


图4 GC-CNN模型损失计算结构

Fig. 4 Loss calculation of GC-CNN

图3(c)所示, 实现输入输出像素的直映射, 从而大幅度减少冗余计算, 提升分类效率。

在GC-CNN中, 为了捕获上下文语义特征, 需要将输入图像进行池化降采样, 使特征图中每个像素点都具有比降采样前大4倍的上下文感受野和更强的空间语义表达能力, 但与此同时图像局部的边缘细节信息丢失严重。膨胀卷积可以解决这个问题, 它通过对普通卷积窗进行空洞填充(hole padding), 使卷积输出的特征图能在尺寸不变的情况下和普通卷积池化后的特征图具备相同的感受野, 即稠密的特征图(dense map), 如图3(a2)所示。因此本文将GC-CNN的第2层卷积参数迁移到空洞维度为1的DGC-CNN中。为了和GC-CNN中的池化层对应, DGC-CNN中的池化核步长和边缘填充尺寸设置为1个像素, 且第2层池化核的空洞维度为1, 从而实现如图3(c)所示的端到端的映射分类。图3(b)、图3(c)中红色虚线为GC-CNN的第1,2层卷积核参数迁移至DGC-CNN对应位置的示意图。

由于DGC-CNN模型直接由GC-CNN参数迁移激活, 分类精度上较GC-CNN会有所降低。因此, 本文使用精调的膨胀交叉卷积神经网络(FDGC-CNN)来捕获更多上下文语义特征, 进一步提升分类精度, 其中精调算法流程图如图5所示。FDGC-CNN和全卷积神经网络FCN相同, 在训练中都需要和输入图像尺寸一致的先验标注图, 这对于先验标注较为稀缺的PolSAR数据来说是不可行的。为解决这一问题, 本文沿用预训练GC-CNN时采样的像素点来实现FDGC-CNN的训练。在FDGC-CNN的训练过程中, 只计算训练图像块中采样像素点的损失, 并用于微调网络参数。具体的精调训练步骤为: 用 128×128 的滑窗在输入复数图像上以步长为64进行滑动采样, 作为DGC-CNN的输入数据; 之后在训练过程中生成损失权重矩阵 \mathbf{W}_{loss} , 用于剔除图像块中非采样像素点的损失, 其中 \mathbf{W}_{loss} 矩阵的采样像素点权重为1, 其余像素点权重为0; 最后将 \mathbf{W}_{loss}

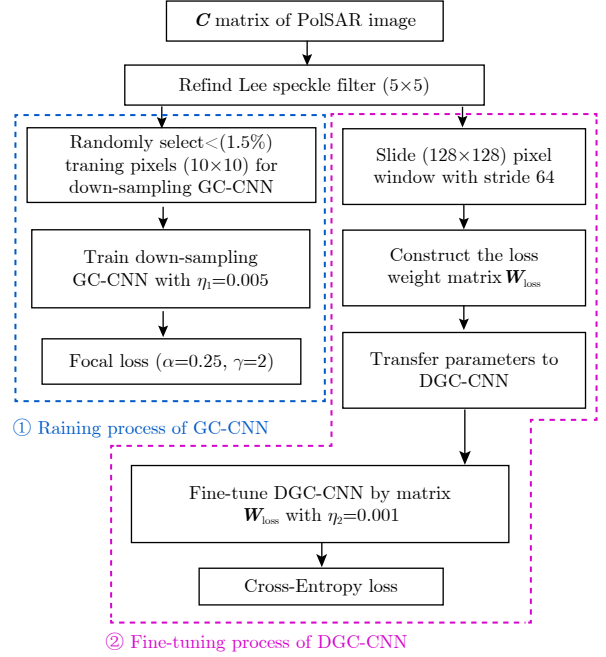


图5 DGC-CNN精调流程图

Fig. 5 Flowchart of DGC-CNN fine-tuning process

和训练预测结果加权, 得到采样像素点的平均分类损失, 并精调网络参数。

4 实验结果和分析

为了验证FDGC-CNN模型的分类型能力, 本节选择参数规模相当的CNN, GC-CNN, Dilated CNN, DGC-CNN与FDGC-CNN进行对比试验; 实验的输入是PolSAR数据的协方差矩阵 \mathbf{C} , 根据互易原理, 单波段 \mathbf{C} 矩阵可以转化为9维的实数特征向量 $\mathbf{Input}_{\text{real}}$, 如式(10)所示。其中 C_{11}, C_{22}, C_{33} 分别为协方差矩阵 \mathbf{C} 的对角元素, C_{12}, C_{13}, C_{23} 分别为 \mathbf{C} 的非对角复数元素。

$$\mathbf{Input}_{\text{real}} = [C_{11}, \Re(C_{12}), \Im(C_{12}), C_{22}, \Re(C_{13}), \Im(C_{13}), C_{33}, \Re(C_{23}), \Im(C_{23})] \quad (10)$$

推广到复数域下的模型, 网络输入为6维复数特征向量 $\mathbf{Input}_{\text{real}} = [C'_{11}, C'_{22}, C'_{33}, C_{12}, C_{13}, C_{23}]$, 为了构造GC-CNN网络实部和虚部两组实数输入图像, 实验将 \mathbf{C} 的对角元素进行极小值填充并得到 $C'_{11}, C'_{22}, C'_{33}$, 具体形式为 $C_{11}+j(1e-8), C_{22}+j(1e-8), C_{33}+j(1e-8)$ 。

为了让实数和复数域模型的网络参数规模相近, 本文设置GC-CNN的第1层卷积通道数为12, 参数自由度(Degree of Freedom, DoF)为8788; 设置CNN的第1层卷积通道数为17, 参数自由度为8891。GC-CNN和CNN的参数自由度如图6所示。CNN和GC-CNN的参数DoF差异很小, 利于实验说明被推广到复数域下的网络的性能优势。

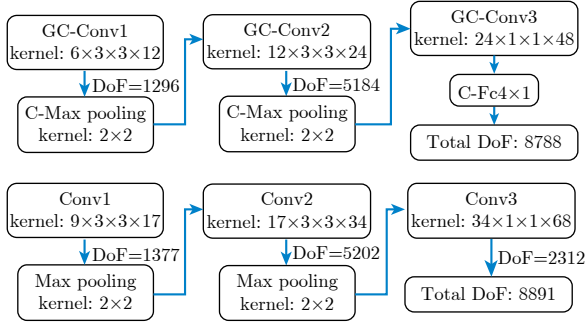


图 6 CNN和GC-CNN参数自由度结构图

Fig. 6 The DoF of CNN and GC-CNN

实验采用像素分类准确率(Overall Accuracy, OA)、度量分类区域和真实地物重合程度的频权交并比(Frequency Weighted Intersection over Union, FWIoU)以及一致性检验指标Kappa系数3个评价指标对试验中各个网络算法进行评估,记 N 为分类混淆矩阵, N_{ij} 为地物 i 被分为地物 j 的像素点数量。OA如式(11)所示

$$OA = \frac{\sum_{i=0}^{cls} N_{ii}}{\sum_{i=0}^{cls} \sum_{j=0}^{cls} N_{ij}} \quad (11)$$

FWIoU如式(12)所示

$$FWIoU = \sum_{i=0}^{cls} \left(\frac{\sum_{j=0}^{cls} N_{ij}}{\sum_{i=0}^{cls} \sum_{j=0}^{cls} N_{ij}} \times \frac{N_{ii}}{\sum_{i=0}^{cls} N_{ij} + \sum_{j=0}^{cls} N_{ji} - N_{ii}} \right) \quad (12)$$

Kappa系数的计算如式(13)所示

$$Kappa = \frac{\sum_{i=1}^{cls} \left[N_{total} N_{ii} - \sum_{j=1}^{cls} N_{ij} \sum_{j=1}^{cls} N_{ji} \right]}{N_{total}^2 - \sum_{i=1}^{cls} \left(\sum_{j=1}^{cls} N_{ij} \sum_{j=1}^{cls} N_{ji} \right)} \quad (13)$$

4.1 基于AIRSAR平台的Flevoland-Netherlands地区PolSAR数据实验

实验采用美国NASA/JPL实验室基于AIRSAR平台的Flevoland-Netherlands地区L, P, C全波段PolSAR数据。该数据广泛地用于PolSAR地物分类

研究,总共包含了14类农作物和2类人造目标。其中,数据L波段部分的Pauli伪彩图如图7(a)所示,图像尺寸为1279×1024;16类真实地物^[27]和对应的标签图例见图7(b)和图7(c),图7(b)中的黑色区域不参与模型的训练和评估。

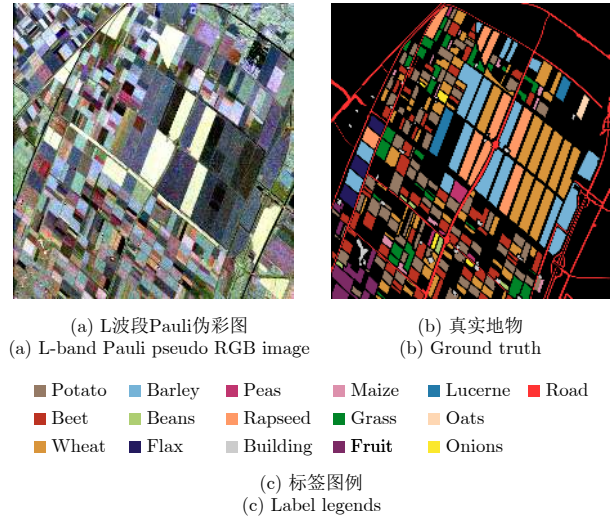


图 7 Flevoland-Netherlands地区L, P, C波段PolSAR数据

Fig. 7 L, P, C-band PolSAR image data over the Flevoland-Netherlands region

16类地物包括potato, beans, wheat, fruit等14类庄稼以及道路和少量建筑。在预训练CNN和GC-CNN时,由于不同地物类别的样本数差异很大,小样本地物类别的学习更加困难。因此,实验为每类地物采样数目相近的样本用于训练,每类地物采样样本数详见表1第2行,各地物的采样率如表1第1行括号中数值所示,除Peas外,每类地物的采样率都低于6%,训练样本总采样率为1.5%。在Zhang等人^[5]的实验中,分类精度随采样率升高而提升,并推荐在实际运用中将总采样率设置为5%~10%。Hou^[28]和Guo^[29]分别采用5%和10%的采样率实现精度相当的地物分类。本文设置总采样率为1.5%,旨在更低的总采样率下验证所提出模型的学习能力和泛化能力。

参考Chen等人^[8]的实验过程,本文对表1中每组实验随机进行了10次采样和训练,每次训练的超参数相同,表1的评估结果取10次实验的平均值。由于输入数据为L, P, C全波段极化SAR图像,因此CNN网络的输入为27维实数特征矢量,GC-CNN网络的输入为18维复数特征矢量。CNN和GC-CNN各迭代60轮后的收敛曲线如图8所示。

可见,由于编组-交叉卷积能够充分利用PolSAR数据丰富的相位信息,GC-CNN能更快和更稳定地收敛,并且能得到更好的分类效果;从表1中的

表 1 Flevoland-Netherlands地区PolSAR图像分类结果表
Tab. 1 Classification results of Flevoland-Netherlands region

Class (%)	1 (2.0)	2 (4.9)	3 (0.9)	4 (0.6)	5 (1.6)	6 (1.5)	7 (1.0)	8 (7.2)	9 (5.2)	10 (5.0)
Train	600	400	600	600	600	700	700	300	240	100
Total	29630	8031	66148	90681	38015	46583	68427	4155	4611	1982
SVM	92.47	97.89	97.52	98.77	95.28	92.46	96.34	96.57	89.42	95.67
CNN	97.81	99.35	98.99	98.93	99.96	97.05	99.21	96.21	83.65	88.27
Dilated CNN	97.66	99.48	98.79	98.79	99.91	96.78	98.68	94.12	79.03	84.79
GC-CNN	96.72	99.60	99.03	99.00	99.96	97.51	99.32	96.59	93.20	95.62
DGC-CNN	96.71	99.49	98.80	98.99	99.95	97.43	99.17	96.58	91.93	95.05
FDGC-CNN	96.63	99.29	98.90	98.59	99.95	98.65	99.59	97.11	91.32	97.01

Class (%)	11 (4.4)	12(4.5)	13(5.4)	14(4.9)	15(4.4)	16(1.1)	OA	Kappa	FWIoU	Time (s)
Train	600	100	100	400	200	400	-	-	-	-
Total	13485	2206	1838	8203	4536	42861	-	-	-	-
SVM	98.34	93.02	98.89	98.56	28.54	57.68	93.27	92.44	87.04	669.00
CNN	97.81	96.48	99.93	99.56	86.69	69.36	95.72	95.31	92.77	220.31
Dilated CNN	97.81	91.10	99.83	99.67	78.99	63.02	95.16	94.34	91.56	3.94
GC-CNN	97.83	98.05	99.98	99.57	87.04	72.79	96.49	95.92	93.64	401.56
DGC-CNN	97.86	97.37	99.82	99.45	82.36	70.13	96.14	95.50	93.04	4.32
FDGC-CNN	96.49	98.85	99.98	99.73	89.61	77.99	96.94	96.44	94.29	4.32

注: Grass(1); Flax(2); Potato(3); Wheat(4); Rapseed(5); Beet(6); Barly(7); Peas(8); Maize(9); Bean(10); Fruit(11); Onion(12); Oats(13); Lucerne(14); Building(15); Road(16)

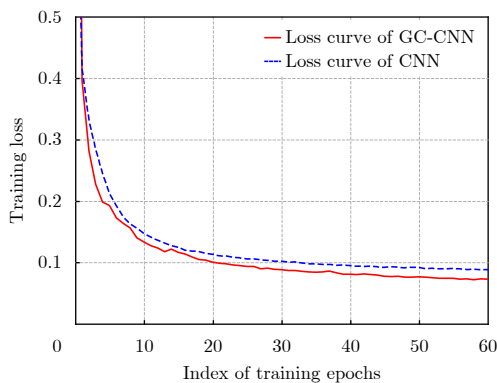


图 8 GC-CNN和CNN训练loss收敛图

Fig. 8 Training loss convergence curves of GC-CNN and CNN

CNN和GC-CNN两行可以看出，GC-CNN在OA, Kappa和FWIoU 3项指标上都高过CNN大约0.6%，并且在Maize, Peas, Beans和Onion等样本数较少的类别上具有更好的分类结果。然而在参数规模相当的情况下，GC-CNN的卷积乘法运算量是CNN的2倍，因此，在预测过程中GC-CNN的分类速度较CNN变慢了约2倍。

CNN和GC-CNN训练结束后，将具参数分别迁移至Dilated CNN和DGC-CNN中，从表1中CNN和Dilated CNN两行可以看出，Dilated CNN的分类速度较CNN加速了大约55倍，但由于

直接迁移CNN的参数，分类精度有0.56%的损耗。DGC-CNN和GC-CNN相比，以损失0.3%分类精度为代价，大幅度提升了分类速度，且分类精度较Dilated CNN有1.1%的提升。

在精调过程中，由于像素到像素(pixel-to-pixel)映射方式的训练使输入图像块内的采样像素被固定在一个批次(batch)中计算损失并更新网络参数，引入了先验的空间信息。因此，FDGC-CNN能够捕获更多的上下文语意特征，进一步提升分类精度。表1最后一行显示，FDGC-CNN相较于其他算法能够大幅度提升道路的分类精度，并且对于一些小样本地物如Onion, Peas, Beans和Beet等能够实现更加准确的分类。在AIRSAR平台的Flevoland-Netherlands地区实验中，FDGC-CNN实现了最精确和相对高效的地物分类。

图9列出了各网络的分类结果图。从图9中可见，每个网络都能实现边缘保持良好的连贯分类。图中的白框指出，编组-交叉卷积能在小类别上实现更加准确的分类，从而提升整体分类精度。另外，FDGC-CNN能够弥补DGC-CNN的精度损耗，并且在Road, Beans, Beet上分类结果更加连贯。综上，本文采用的FDGC-CNN模型在Flevoland-Netherlands极化数据上兼顾了准确率和分类效率，得到了最优的地物分类结果。

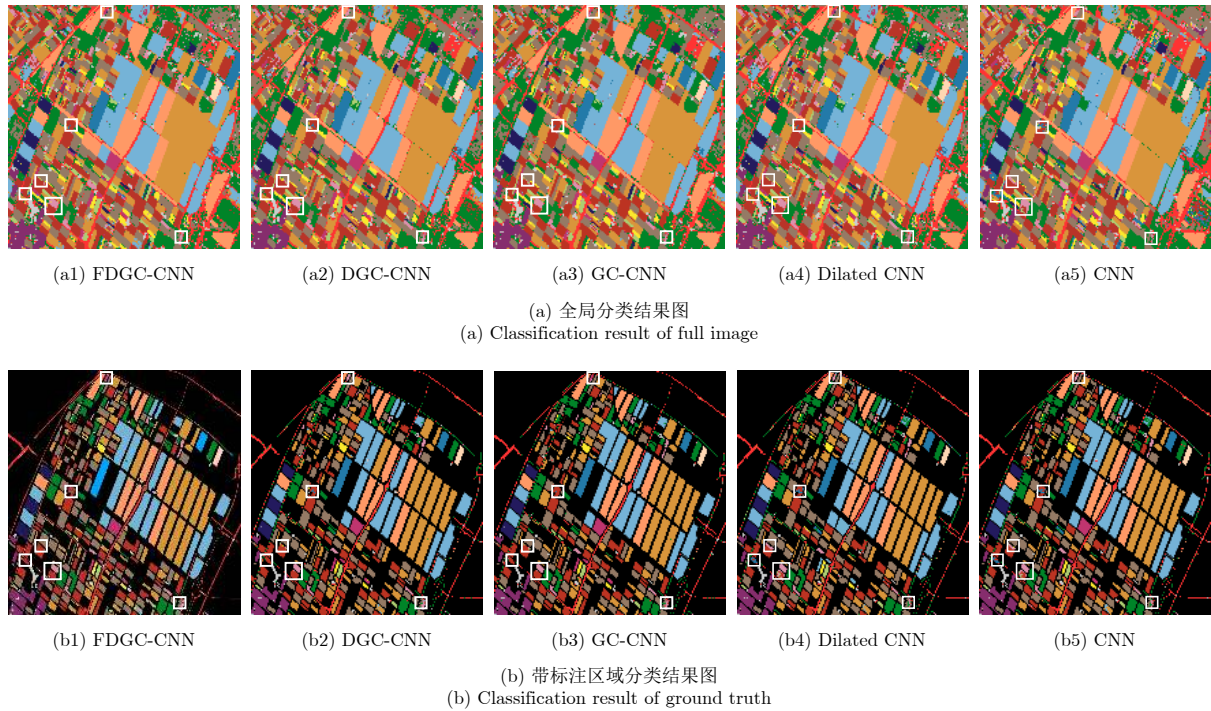


图 9 Flevoland-Netherlands地区分类结果图

Fig. 9 Classification images of the Flevoland-Netherlands region

4.2 基于E-SAR平台的Oberpfaffenhofen地区L波段PolSAR数据实验

为验证FDGC-CNN在表面不规则地物上的分类性能, 本文采用德国E-SAR平台的L波段PolSAR数据进行实验。该数据包含表面粗糙的城区(Urban)和灌木(WoodLand)以及平滑的耕地(FarmLand)和道路(Road), 总共4类地物。数据尺寸为1300×1200, 其Pauli伪彩图如10(a)所示, 真实地物和标注图例分别如图10(b)、图10(c)所示, 其中真实地物根据Google Earth光学图像通过人工标注得到。4类地物的采样样本数和实验总样本数如表2第1、第2列所示, 单类采样率详见表2第1列括号中的数值, 实验总采样率为0.94%。

表2第4, 第6列指出, GC-CNN在表面不规则

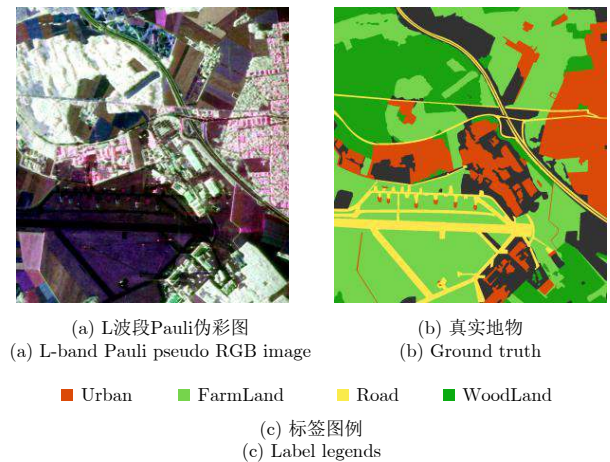


图 10 Oberpfaffenhofen地区L波段PolSAR数据

Fig. 10 L-band PolSAR image data over the Oberpfaffenhofen region

表 2 Oberpfaffenhofen地区PolSAR图像分类结果表

Tab. 2 Classification results of oberpfaffenhofen region

Class (%)	Train	Total	CNN	Dilated CNN	GC-CNN	DGC-CNN	FDGC-CNN
WoodLand (1.0)	3000	290915	88.14	88.33	90.40	89.67	90.55
FarmLand (0.5)	3000	622739	94.02	94.23	93.54	94.27	95.30
Urban (1.1)	3000	274684	82.78	79.93	85.80	82.36	86.03
Roads (1.8)	2500	137400	78.89	76.45	78.82	76.08	73.65
OA	-	-	88.83	88.13	89.72	88.90	90.07
Kappa	-	-	81.40	79.26	82.80	81.20	83.36
FWIoU	-	-	80.70	80.44	81.72	80.42	82.11
Time (s)	-	-	262.47	2.60	513.43	3.21	4.02

的WoodLand和Urban地物上的分类精度相较CNN模型有较大提升，这主要因为PolSAR通道相位表征的是回波时延，GC-CNN通过充分学习通道相位信息来提取不规则地物的空间几何特征，因此能进一步提升不规则地物的分类精度。Dilated CNN和DGC-CNN在参数迁移后都损失了大约0.7%的总分类精度，但分类耗时缩短了大约100倍。精调后的FDGC-CNN模型在OA、Kappa和FWIoU

指标上都优于DGC-CNN和GC-CNN，兼顾了分类准确率和效率。

图11为Oberpfaffenhofen地区实验分类结果图。GC-CNN在白框区域较CNN能够实现更连贯的分类，空洞区域较少。FDGC-CNN相较于GC-CNN在FarmLand上的离散区域更少，说明精调过程能够结合训练图像块中像素的空间信息来约束平滑地物中的离散区域，进而提升分类精度。

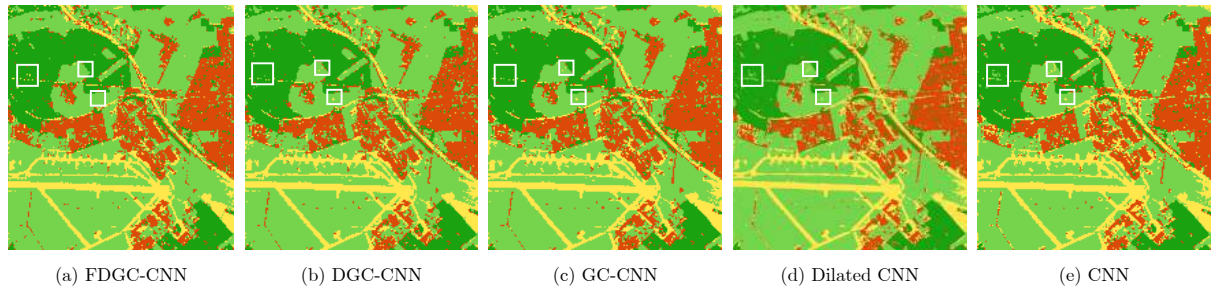


图 11 Oberpfaffenhofen地区分类结果图
Fig. 11 Classification images of the Oberpfaffenhofen

4.3 实验分析

从2组实验的分类精度评估指标曲线(如图12所示)可以看出，基于编组-交叉卷积的模型能够利用相位信息有效地提升网络的分类精度。将CNN和GC-CNN的参数迁移到Dilated CNN和DGC-CNN后，分类精度轻微下降，但分类速度大幅度提升。精调的FDGC-CNN能学习更多的上下文语义特征，从而修正一些样本数较少或具有较强结构特征地物的错分像素，并能根据像素的空间关联信息剔除规则地物区域中的离散点和空洞点，最终在GC-CNN和DGC-CNN的总体分类精度上实现了进一步的提升。综上，两组实验验证了FDGC-CNN模型能够在少量先验标注下实现高精度高效率的极化SAR地物分类。

5 结论

本文通过FDGC-CNN算法实现了在低采样率下精确且高效的PolSAR地物分类。该算法首先用低采样率采样得到的像素点预训练GC-CNN，随后将网络参数迁移到DGC-CNN中，在分类精度降低的代价下实现像素到像素(pixel-to-pixel)直接映射的快速分类。最后通过精调参数迁移后的DGC-CNN以捕获更多上下文强语义特征，进而修正错分样本，提升分类精度。基于AIRSAR平台以及E-SAR平台的PolSAR数据的实验表明，在低采样率下，FDGC-CNN相较于SVM分类器和CNN模型取得了精度和效率上的大幅度提升，兼顾了分类精度和效率，具有很强的实用性和部署性。

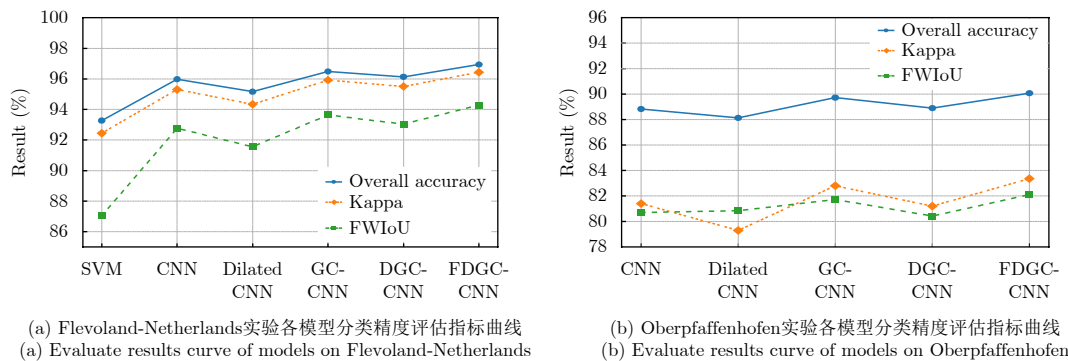


图 12 Flevoland-Netherlands和Oberpfaffenhofen地区实验各模型评估结果曲线
Fig. 12 Evaluate results curve of models on Flevoland-Netherlands and Oberpfaffenhofen

参 考 文 献

- [1] YANG Wen, ZHONG Neng, YANG Xiangli, *et al.* Riemannian sparse coding for classification of PolSAR images[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 2016: 5698–5701. doi: 10.1109/IGARSS.2016.7730488.
- [2] 张祥, 邓喀中, 范洪冬, 等. 基于目标分解的极化SAR图像 SVM监督分类[J]. 计算机应用研究, 2013, 30(1): 295–298. doi: 10.3969/j.issn.1001-3695.2013.01.076.
ZHANG Xiang, DENG Kazhong, FAN Hongdong, *et al.* PolSAR SVM supervised classification method combining with polarimetric target decomposition[J]. *Application Research of Computers*, 2013, 30(1): 295–298. doi: 10.3969/j.issn.1001-3695.2013.01.076.
- [3] ZHANG Hongsheng, ZHANG Yuanzhi, and LIN Hui. Urban land cover mapping using random forest combined with optical and SAR data[C]. 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 2012: 6809–6812. doi: 10.1109/IGARSS.2012.6352600.
- [4] ZHANG Guangyun and JIA Xiuping. Simplified conditional random fields with class boundary constraint for spectral-spatial based remote sensing image classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2012, 9(5): 856–860. doi: 10.1109/LGRS.2012.2186279.
- [5] ZHANG Zhimian, WANG Haipeng, XU Feng, *et al.* Complex-valued convolutional neural network and its application in polarimetric SAR image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(12): 7177–7188. doi: 10.1109/TGRS.2017.2743222.
- [6] 徐真, 王宇, 李宁, 等. 一种基于CNN的SAR图像变化检测方法[J]. 雷达学报, 2017, 6(5): 483–491. doi: 10.12000/JR17075.
XU Zhen, WANG R, LI Ning, *et al.* A novel approach to change detection in SAR images with CNN classification[J]. *Journal of Radars*, 2017, 6(5): 483–491. doi: 10.12000/JR17075.
- [7] BI Haixia, SUN Jian, and XU Zongben. A graph-based semisupervised deep learning model for PolSAR image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(4): 2116–2132. doi: 10.1109/TGRS.2018.2871504.
- [8] CHEN Siwei and TAO Chensong. PolSAR image classification using polarimetric-feature-driven deep convolutional neural network[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(4): 627–631. doi: 10.1109/LGRS.2018.2799877.
- [9] CHEN Siwei, TAO Chensong, WANG Xuesong, *et al.* PolSAR target classification using polarimetric-feature-driven deep convolutional neural network[C]. 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 2018. doi: 10.1109/IGARSS.2018.8518529.
- [10] KONG J A, SWARTZ A A, YUEH H A, *et al.* Identification of terrain cover using the optimum polarimetric classifier[J]. *Journal of Electromagnetic Waves and Applications*, 1988, 2(2): 171–194.
- [11] LEE J S, GRUNES M R, and KWOK R. Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution[J]. *International Journal of Remote Sensing*, 1994, 15(11): 2299–2311. doi: 10.1080/01431169408954244.
- [12] KOUSKOULAS Y, ULABY F T, and PIERCE L E. The Bayesian Hierarchical Classifier (BHC) and its application to short vegetation using multifrequency polarimetric SAR[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(2): 469–477. doi: 10.1109/TGRS.2003.821066.
- [13] 邹焕新, 罗天成, 张月, 等. 基于组合条件随机场的极化SAR图像监督地物分类[J]. 雷达学报, 2017, 6(5): 541–553. doi: 10.12000/JR16109.
ZOU Huanxin, LUO Tiancheng, ZHANG Yue, *et al.* Combined conditional random fields model for supervised PolSAR images classification[J]. *Journal of Radars*, 2017, 6(5): 541–553. doi: 10.12000/JR16109.
- [14] YU F and KOLTUN V. Multi-scale context aggregation by dilated convolutions[J]. *arXiv preprint arXiv: 1511.07122*, 2015.
- [15] LONG J, SHELHAMER E, and DARRELL T. Fully convolutional networks for semantic segmentation[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015: 3431–3440. doi: 10.1109/CVPR.2015.7298965.
- [16] BADRINARAYANAN V, KENDALL A, and CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *arXiv preprint ar-Xiv: 1511.00561*, 2015.
- [17] RONNEBERGER O, FISCHER P, and BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. Proceedings of 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015: 234–241.
- [18] BENGIO Y, LECUN Y, NOHL C, *et al.* LeRec: A NN/HMM hybrid for on-line handwriting recognition[J]. *Neural Computation*, 1995, 7(6): 1289–1303. doi: 10.1162/neco.1995.7.6.1289.
- [19] NAIR V and HINTON G E. Rectified linear units improve restricted boltzmann machines[C]. Proceedings of the 27th International Conference on Machine Learning (ICML-10),

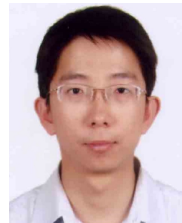
- Haifa, Israel, 2010: 807–814.
- [20] SHANG W, SOHN K, ALMEIDA D, *et al.* Understanding and improving convolutional neural networks via concatenated rectified linear units[C]. International Conference on Machine Learning, 2016: 2217–2225.
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 1026–1034. doi: 10.1109/ICCV.2015.123.
- [22] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. *arXiv preprint arXiv: 1409.1556*, 2014.
- [23] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Identity mappings in deep residual networks[C]. Proceedings of the 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 630–645.
- [24] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 770–778. doi: 10.1109/CVPR.2016.90.
- [25] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, *et al.* Densely connected convolutional networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 2261–2269. doi: 10.1109/CVPR.2017.243
- [26] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2999–3007. doi: 10.1109/ICCV.2017.324.
- [27] HOEKMAN D H and VISSERS M A M. A new polarimetric classification approach evaluated for agricultural crops[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(12): 2881–2889. doi: 10.1109/TGRS.2003.817795.
- [28] HOU Biao, KOU Hongda, and JIAO Licheng. 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. doi: 10.1109/JSTARS.2016.2553104.
- [29] GUO Yanhe, WANG Shuang, GAO Chenqiong, *et al.* Wishart RBM based DBN for polarimetric synthetic radar data classification[C]. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 2015: 1841–1844. doi: 10.1109/IGARSS.2015.7326150.

作者简介



肖东凌(1995–), 男, 四川眉山人, 硕士生, 研究方向为SAR图像目标检测识别、PolSAR图像地物分类。

E-mail: xdluestc@outlook.com



刘畅(1978–), 男, 山东烟台人, 研究员, 博士生导师。2006年在中国科学院电子学研究所获得博士学位, 现担任中国科学院电子学研究所研究员、博士生导师。主要研究方向为SAR系统及其相关SAR成像处理技术等。

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