

# 基于深度学习和层次语义模型的极化 SAR 分类

石俊飞<sup>1,2,3</sup> 刘芳<sup>1,3</sup> 林耀海<sup>3,4</sup> 刘璐<sup>2,3</sup>

**摘要** 针对复杂场景的极化合成孔径雷达 (Synthetic aperture radar, SAR) 图像, 堆叠自编码器模型能够自动学习高层特性, 有效表示城区、森林等复杂地物的结构, 然而, 却难以保持图像的边界和细节. 为了克服该缺点, 本文结合深度自编码器和极化层次语义模型 (Polarimetric hierarchical semantic model, PHSM), 提出了新的无监督的极化 SAR 图像分类算法. 该方法根据极化层次语义模型, 将复杂的极化 SAR 图像划分为聚集、匀质和结构三大区域. 对聚集区域, 采用堆叠自编码器模型进行高层特征表示, 并构造字典得到稀疏特征进行分类; 对匀质区域, 采用层次模型进行分类; 对于结构区域, 进行线目标保留和边界定位. 实验结果表明, 该算法通过不同的分类策略优势互补, 能够得到区域一致性好且边界保持的分类结果.

**关键词** 叠自编码器, 极化层次语义模型, 极化 SAR 分类, 区域划分, 层次分割

**引用格式** 石俊飞, 刘芳, 林耀海, 刘璐. 基于深度学习和层次语义模型的极化 SAR 分类. 自动化学报, 2017, 43(2): 215–226

**DOI** 10.16383/j.aas.2017.c150660

## Polarimetric SAR Image Classification Based on Deep Learning and Hierarchical Semantic Model

SHI Jun-Fei<sup>1,2,3</sup> LIU Fang<sup>1,3</sup> LIN Yao-Hai<sup>3,4</sup> LIU Lu<sup>2,3</sup>

**Abstract** Stacked auto-encoder model can effectively represent the complex terrain structures, such as the urban and the forest, by automatically learning high-level features. However, it has difficulty in preserving details and edges. In order to overcome this shortcoming, a new unsupervised polarimetric synthetic aperture radar (PolSAR) classification method is proposed by combining the deep learning and the polarimetric hierarchical semantic model (PHSM). According to the PHSM, a PolSAR image is partitioned into aggregated, homogeneous and structural regions. For aggregated regions, a stacked auto-encoder model is applied to learn high-level features, and further the sparse representation and classification is constructed by learning a dictionary with high-level features. For homogeneous regions, hierarchical segmentation and classification is applied. In addition, edges are located and line objects are preserved for structural regions. Experimental results demonstrate that the proposed method can obtain good performance in both region homogeneity and edge preservation.

**Key words** Stacked auto-encoder, polarimetric hierarchical semantic model (PHSM), polarimetric synthetic aperture radar (SAR) image classification, region partition, hierarchical segmentation

**Citation** Shi Jun-Fei, Liu Fang, Lin Yao-Hai, Liu Lu. Polarimetric SAR image classification based on deep learning and hierarchical semantic model. *Acta Automatica Sinica*, 2017, 43(2): 215–226

收稿日期 2015-10-21 录用日期 2016-04-18  
Manuscript received October 21, 2015; accepted April 18, 2016  
国家重点基础研究发展计划 (973 计划) (2013CB329402), 国家自然科学基金 (61573267, 61571342, 61572383), 国家自然科学基金青年科学基金项目 (31300473), 教育部“长江学者和创新团队发展计划”(IRT1170), 高等学校学科创新引智计划 (B07048), 福建省自然科学基金 (2014J01073) 资助  
Supported by National Basic Research Program of China (973 Program) (2013CB329402), National Science Foundation of China (61573267, 61571342, 61572383), Youth Fund of National Natural Science Foundation of China (31300473), the Program for Cheung Kong Scholars and Innovative Research Team in University (IRT1170), the Fund for Foreign Scholars in University Research and Teaching Programs (B07048), Natural Science Foundation of Fujian Province (2014J01073)

本文责任编辑 柯登峰  
Recommended by Associate Editor KE Deng-Feng  
1. 西安电子科技大学计算机学院 西安 710071 2. 西安理工大学计算机科学与工程学院 西安 710048 3. 西安电子科技大学智能感知与图像理解教育部重点实验室 西安 710071 4. 福建农林大学计算机与信息学院 福州 350002  
1. College of Computer Science and Technology, Xidian Uni-

极化合成孔径雷达 (Synthetic aperture radar, SAR) 图像是电磁波在水平和垂直极化方式下进行的地物成像, 因此含有更多的极化信息. 近期, 随着雷达技术的发展, 极化 SAR 图像的处理已经成为研究的热点. 极化 SAR 地物分类是图像处理的关键步骤, 是人们进行图像理解和解译的前提. 传统的极化 SAR 图像分类方法主要通过目标分解和统计分布来实现. 极化数据的目标分解方法有很多, 如 Cloude 分解<sup>[1]</sup>、Freeman 分解<sup>[2]</sup>. 统计分布模型主要有 Wishart 分布<sup>[3]</sup>、K 分布<sup>[4]</sup>、G0 分布<sup>[5]</sup> 及 KummerU 分布<sup>[6]</sup>. 结合目标分解和分布模型提出  
versity, Xi'an 710071 2. School of Computer Science and Technology, Xi'an University of Technology, Xi'an 710048 3. Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, Xidian University, Xi'an 710071 4. School of Computer Science and Technology, Fujian Agriculture and Forest University, Fuzhou 350002

了经典  $H/\alpha$ -Wishart 分类方法<sup>[7]</sup>, 该方法根据 Cloude 分解进行初始分类, 并用 Wishart 分类器进行迭代调整, 能够对图像进行精确的分类. 然而, 由于没有考虑图像的空间关系, 这些方法容易受噪声影响, 得到椒盐噪声式的分类结果.

近来, 一些加入空间信息的图像处理方法<sup>[6, 8-11]</sup> 用来对极化 SAR 图像进行分类, 如基于 Mean shift (MS) 和 Markov random field (MRF)<sup>[8]</sup> 的方法, 该方法在 Mean shift 分割的基础上加入 MRF 空间邻域信息, 能够得到区域一致性好的结果. 另外, 基于层次分割的方法<sup>[6]</sup> 通过定义距离测度进行区域合并, 得到较一致的区域. 同时, Ersahin 等<sup>[10]</sup> 提出了两阶段谱聚类方法, 该算法利用轮廓特征进行初始划分, 再依据 Wishart 测度进行进一步的分类. 这些算法能够有效地抑制斑点噪声, 提高了分类的区域一致性. 然而, 极化 SAR 图像地物繁多、场景复杂、尺度不一. 由于没有考虑语义信息, 这些方法很难将聚集地物分为语义一致的区域. 聚集地物是指由同类目标聚集在一起形成的地物, 如森林、城区等. 这种地物的特点为目标和地面的散射回波形成强烈的亮暗变化, 且这种变化重复出现. 由于聚集地物内部强烈的亮暗变化和地物散射特性的较大差异, 各种底层特征都难以将其合并为语义一致的区域.

针对这种复杂地物, Liu 等<sup>[12]</sup> 提出了极化 SAR 的层次语义模型, 该模型能够将极化 SAR 图像划分为聚集、结构和匀质三种区域. 这样, 根据地物的特性, 不同的分类方法可以自适应地对不同区域进行分类. 对于匀质区域, 由于结构比较单一, 传统的分割和分类方法<sup>[6, 9-11]</sup> 能够很好地分类. 对于结构区域, 主要是边界定位和线目标保持. 对于聚集区域, 一幅极化 SAR 图像, 可能存在多种聚集地物类型, 如何区分不同聚集地物, 并对聚集区域赋予类标, 是本文研究的重点. 对于聚集区域, 同一区域内应该含有相同的地物结构, 而不同区域之间的地物结构可能不同. 因此, 问题的关键是如何表示各种聚集区域的地物结构, 并进行分类.

深度学习<sup>[13-14]</sup> 能够学习图像的结构, 得到高层的特征, 对复杂地物能够很好地表示, 因此, 在自然图像处理中得到广泛的应用. 深度模型有很多, 自编码模型<sup>[15]</sup>、卷积神经网络<sup>[16]</sup>、限制玻尔兹曼机<sup>[17]</sup>、反卷积网络<sup>[18]</sup> 等. 然而, 对于极化 SAR 图像分类, 深度学习方法的应用还很少. 另外, 由于极化 SAR 图像缺乏训练样本, 本文选用深度自编码器作为无监督的特征学习方法. 自编码器通过自身的重构进行权值学习, 得到更加抽象的特征. 然而, 深度学习由于不断地概括抽象, 难以保持边界细节. 本文采用层次语义模型, 只对聚集区域进行深度

特征学习, 而对结构和匀质区域进行精细分割, 避免了深度学习的缺点.

本文首次将深度学习和层次语义模型结合, 应用在极化 SAR 图像的分类上, 提出了一种新的无监督的深度学习方法. 该算法不仅克服了深度学习的缺点, 同时根据不同区域的特点进行分类, 得到区域一致性好且边界精准的分类结果. 该方法有三个创新点: 首先, 深度学习能够学到图像高层结构特征, 然而, 难以精确定位边界. 为了克服该缺点, 本文结合深度学习和层次语义模型, 将极化 SAR 图像分为聚集、结构和匀质三种区域类型. 其次, 对聚集区域, 由于地物结构复杂, 本文采用深度自编码对地物结构进行特征学习, 并构建字典得到区域的稀疏特征表示, 再用谱聚类方法<sup>[10]</sup> 进行分类. 最后, 对匀质区域, 本文采用层次分割方法进行合并, 得到一致的区域和精准的边界. 对结构区域, 进行边界定位和线目标保持. 三幅真实的极化 SAR 图像用来进行实验, 实验结果表明该方法不仅能够得到一致的区域, 同时能够保持边界.

本文的内容安排如下: 第 1 节主要介绍深度自编码模型; 第 2 节介绍极化层次语义模型; 第 3 节详细给出了本文提出的方法; 实验结果和分析在第 4 节; 最后一节总结了本文提出的算法.

## 1 深度自编码模型

人类的视觉具有层次认知功能, 能够有效捕捉不同地物的复杂结构. 当人眼看一幅图像时, 信号传入大脑, 负责视觉的 V1 区域通过对图像进行边角检测的稀疏表示之后传入 V2 区域, V2 区域通过对稀疏特征进行概念抽象, 得到更高层轮廓和结构特征<sup>[19]</sup>. 近年来, 深度学习能够部分模拟大脑的 V1 和 V2 区的层次认知功能<sup>[20-21]</sup>, 因此得到广泛的应用. 深度学习源于人工神经网络, 通过多层神经网络进行学习, 得到图像的结构特征.

堆叠自动编码器<sup>[22]</sup> 是由多个自动编码器堆叠形成的深度网络, 可以进行无监督的特征学习. 自编码器通过编码和解码操作能够自适应地学习网络权值, 主要用于特征的学习和降维. 由于缺乏标记样本, 无监督的分类方法更适用于极化 SAR 图像, 因此本文选择堆叠自编码进行特征学习.

图 1 所示为单层自动编码器的网络结构, 该网络包含输入层 - 隐层 - 输出层. 自编码网络通过要求输出和输入相等来训练调节网络权值, 通过自学习的方法来进行无监督的特征学习. 层 2 即为层 1 的一种特征表示, 层 3 为重构数据, 无监督的学习过程是通过最小化重构误差得到网络权重. 当多个自编码堆叠在一起, 就形成了堆叠自编码器. 堆叠自编码器的学习是每层单独训练的, 每层代表图像的一

种表示, 且每层学习的特征作为下层的输入, 高层是低层的更抽象表示, 通过维数不断减少得到输入数据的最主要成分, 因此, 它是一个降维的过程, 学习的高层特征可以进一步对图像进行分类和识别.

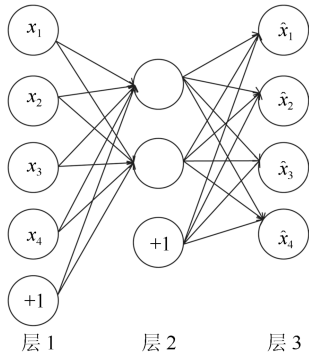


图 1 单层自动编码器的网络结构

Fig.1 Network structure of single-level auto-encoder

## 2 极化层次语义模型

极化层次语义模型<sup>[12]</sup> 是图像在语义层面上的稀疏表示. 基于 Marr 的视觉计算理论<sup>[23]</sup>, Liu 等<sup>[12]</sup> 提出了极化层次语义模型, 该模型是在初始素描模型<sup>[24]</sup> 的基础上发展得到的, 包含两层语义: 第一层是一幅素描图, 它刻画了一幅极化 SAR 图像的结构信息, 将图像变化的部分用素描线勾勒出来, 素描图是由有方向和长度特性的素描线段构成. 其获得过

程如下: 首先, 由极化边缘检测算子得到极化 SAR 图像的变化部分; 然后, 使用线段匹配追踪算法得到素描线段, 同时, 去掉噪声引起的伪线段, 得到极化素描图.

第二层为区域图, 它是在素描图的基础上进一步提取得到的. 区域图将一幅极化 SAR 图像划分为聚集区域、匀质区域和结构区域. 具体过程如下: 聚集地物形成的线段比较聚集, 而边界和线目标的线段比较稀疏, 因此, 根据线段的拓扑结构和语义含义, 将线段划分为聚集线段和孤立线段. 其中, 聚集线段表示聚集地物的结构变化. 孤立线段表示线目标和地物边界. 然后, 对不同线段类型提取其所在的地物区域, 进而将图像划分为聚集、结构和匀质区域.

聚集区域是具有聚集地物结构的区域, 例如城区、森林等. 这些区域由地物目标聚集在一起而形成, 而传统的极化 SAR 分类方法很难将其分为语义上一致的区域. 匀质区域一般是匀质地物对应的区域, 例如农田、水域、裸地等. 而结构区域一般对应于边界或者线目标所在区域, 这些区域有强烈的明暗变化, 会形成素描线. 极化层次语义模型对进一步的图像分割、分类和识别有重要的指导作用.

图 2 为层次语义模型的示例. 图 2 (a) 和 (d) 为 San Francisco 部分地区和 Ottawa 地区的全极化图像, 以 Pauli 基  $|HH - VV|$ 、 $|HH + VV|$  和  $|HV|$

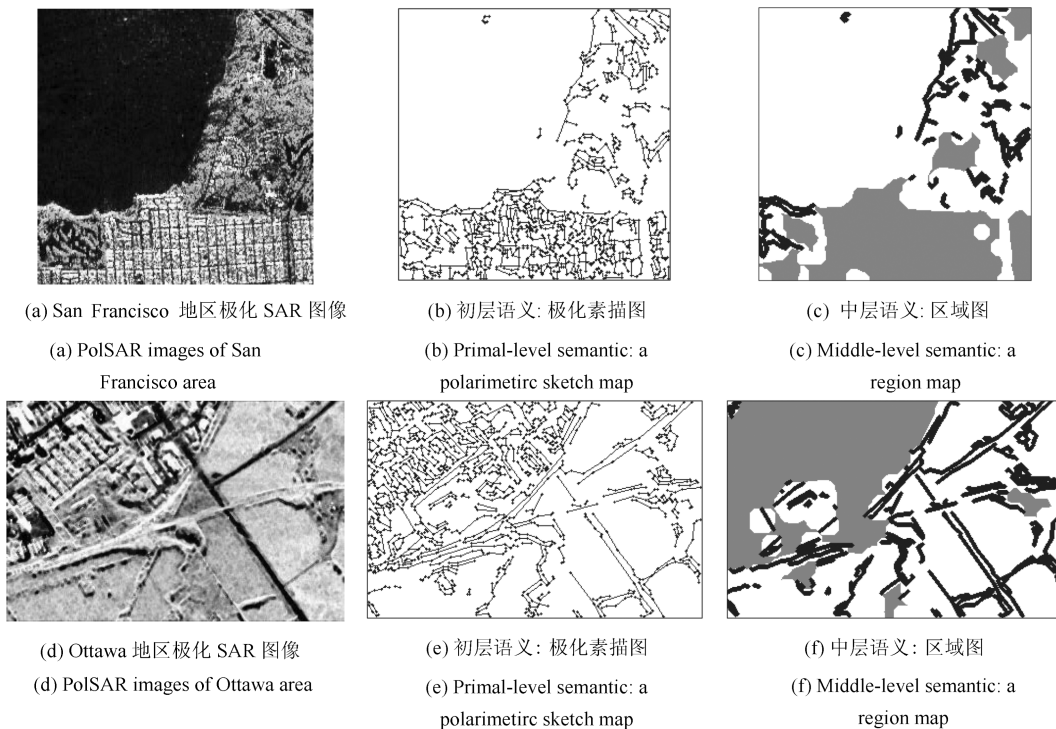


图 2 层次语义模型示例图

Fig.2 Example of hierarchical semantic model

为 RGB 颜色通道显示而成. 图 2(b) 和 (e) 为对应的初层语义: 极化素描图. 可以看出素描线段所在位置为图像变化的部分, 能够有效地刻画图像的结构信息, 是图像的稀疏表示. 图 2(c) 和 (f) 为对应的区域图, 区域图将图像划分为三个区域, 其中, 灰色为聚集区域, 白色为匀质区域, 黑色为结构区域. 区域图是极化 SAR 图像在语义上的划分, 是更稀疏的图像表示.

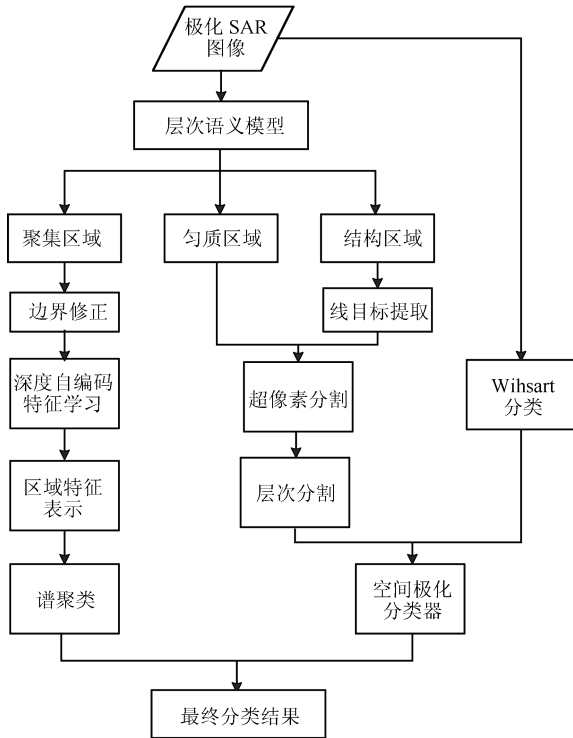


图 3 本文算法示意图

Fig. 3 Algorithm framework of the proposed method

### 3 本文算法

基于极化层次语义模型, 本文提出一种新的基于深度学习的极化 SAR 分类方法. 如图 3 所示, 基于区域图, 极化 SAR 图像被分为聚集、匀质和结构三类区域, 对不同区域采用不同的分类方式: 1) 对聚集区域, 本文采用深度自编码进行特征学习, 得到区域的特征表示, 并用谱聚类算法得到区域类别; 2) 对于匀质区域, 本文采用基于 Wishart 最大似然的层次分割方法, 并进一步和极化分类结果进行融合分类; 3) 对于结构区域, 本文首先进行线目标和边界定位, 对边界两边的超像素和相邻匀质区域进行合并.

#### 3.1 聚集区域的深度自编码模型

根据区域图我们发现, 图像的边界或线目标主要存在于结构区域, 而对于聚集区域, 主要的难点是如何判别不同区域的类别是否一致. 相同的地物类

型应该具有相似的结构, 而不同的地物类型结构差异较大, 如城区和森林的结构具有很大差异, 而两个不同位置的城区则结构相似. 因此, 对聚集地物的结构表示能够较好地刻画聚集区域. 深度学习是一种有效的特征学习方法, 能够很好地学习地物的结构, 因此, 本文对每个区域进行深度学习, 得到的高层特征可以表示该区域的特征. 不同地物类型的结构差异大, 得到的特征具有可分性.

#### 3.1.1 区域边界修正

聚集区域能够将复杂地物的亮暗变化分为一致区域, 然而, 由于聚集区域是通过聚集线段进行形态学操作得到, 因此, 其边界并不精准, 不代表地物真实的边界. 相反, 图像的过分割方法会将聚集地物分为很多过分割的区域, 然而由于精细分割, 却能够得到精准的边界. 因此, 融合聚集区域和过分割结果能够进一步修正聚集区域边界.

初始分割: 极化功率图 (SPAN) 为三个极化通道的功率之和, 用来进行初始分割. 初始分割将 SPAN 图划分为一些小区域, 这些区域称为超像素. 每个区域的类别被认为是一致的. 常用的超像素提取方法有很多, 如分水岭、均值漂移、水平集等. 本文选择均值漂移算法<sup>[25]</sup>, 因为该方法能够得到精细的边界, 同时减少了超像素个数.

边界修正: 为了进一步修正聚集区域的边界, 本文将聚集区域和均值漂移分割结果进行融合, 使用最大投票策略得到聚集区域的精准边界. 首先, 将聚集区域投影到过分割图上, 对聚集区域内部的超像素直接合并, 对聚集区域边界的超像素, 如果聚集区域覆盖超像素过半数目, 则将该超像素合并到聚集区域中, 扩充边界; 如果覆盖数目较少, 则将超像素从聚集区域中去掉, 缩回边界. 通过边界融合, 得到较为精确的聚集区域边界.

#### 3.1.2 区域特征学习

采样: 每个聚集区域内部属于同一类, 因此, 对所有像素点进行学习是耗时和无用的, 选一些代表点进行学习既可以刻画聚集地物的结构, 又可以减少计算量. 另外, 聚集区域的结构具有稀疏性, 以城区为例, 建筑物和地面会形成亮暗变化的散射特性, 并且这种变化重复出现, 构成城区. 因此, 本文通过隔点采样方式对每个聚集区域进行采样, 得到样本进行学习. 这些样本的类别即代表区域的类别. 因为聚集区域大小不一, 像素个数不同, 本文采用最小区域像素个数  $\omega$  作为采样数. 太小的区域难以包含聚集结构, 这里, 首先去掉小区域, 保留较大区域进行采样. 通过采样方式使得原本上百万的输入样本减少至上万甚至上千, 大大减少了计算复杂度.

特征提取: 堆叠自编码模型可以学习图像结构,



输入特征应为每个像素点所在的图像块. 本文对每个像素点选取一个  $N \times N$  的图像块, 把图像块作为输入来进行深度学习, 最后把学到的该图像块的特征作为中心像素特征. 窗口大小应该能够包含聚集区域的变化结构, 如城区或森林. 对于中低分辨的极化 SAR 图像, 聚集区域的结构基元较小, 因此, 本文采用  $13 \times 13$  的窗口. 对于极化 SAR 数据, 每个像素点用  $T$  矩阵的 9 维特征向量来表示如下:

$$V = \{T_{11}, T_{22}, T_{33}, \text{real}(T_{12}), \text{imag}(T_{12}), \text{real}(T_{13}), \text{imag}(T_{13}), \text{real}(T_{23}), \text{imag}(T_{23})\} \quad (1)$$

其中,  $\text{real}(\cdot)$  和  $\text{imag}(\cdot)$  分别为求实部和虚部操作子. 则图像块的特征为图像块内所有像素点特征的列向量. 即为  $N \times N \times 9$  维的输入特征. 这样, 极化散射特征能够得到充分利用.

区域稀疏表示: 训练好网络后, 最后一层的高层特征表示该像素的特征. 对一个聚集区域, 每个像素的特征不尽相同, 然而, 每个聚集区域作为同一类, 应该使用一个统一的特征进行表示, 因此, 本文采用词袋模型<sup>[26]</sup>的思想, 构造视觉字典, 每个聚集区域向视觉字典进行投影, 能够得到字典上的稀疏表示. 具体过程如下: 对所有像素点进行  $K$ -means 聚类, 得到  $M$  个聚类中心, 这些聚类中心即为视觉字, 为了得到完备的视觉字典,  $M$  要远远大于区域数. 每个区域的像素特征向字典上投影, 得到的直方图统计特征作为区域的稀疏特征表示. 根据学到的字典, 同类区域向字典投影相似性越大, 而不同类型区域的投影差异越大.

### 3.1.3 区域谱聚类

谱聚类算法是一种图划分方法, 包括结点和边上的权重, 权值为结点之间的相似性, 该算法能够将一幅图划分为几个子图, 使得子图内部尽可能相似, 子图之间距离尽可能远. 本文将每个聚集区域看作一个结点, 区域间的相似性作为结点间的权值. 谱聚类的关键为定义相似性矩阵. 由于区域为稀疏表示,

特征维数较高, 采用巴氏 (Bhattacharyya) 系数<sup>[27]</sup>用来度量区域  $P$  和区域  $Q$  之间的相似性, 定义为

$$\rho(P, Q) = \sum_{u=1}^N \sqrt{f_P^u \cdot f_Q^u} \quad (2)$$

其中,  $f_P$  和  $f_Q$  分别为区域  $P$  和  $Q$  的归一化的特征.  $u$  为  $f_P$  的第  $u$  个元素. 巴氏系数的几何含义为向量  $f_P$  和  $f_Q$  夹角的余弦值. 根据相似性矩阵, 得到 Laplacian 矩阵, 进行特征值分解, 将得到的特征向量进行  $K$ -means 聚类. 如图 4 所示, 图 4(a) 为 Ottawa 地区的极化 SAR 图像. 图 4(b) 为区域图, 灰色部分表示聚集区域, 左上角为城区, 右下角有一些小片的树丛. 通过对聚集区域边界修正和深层自编码聚类, 得到图 4(c) 的聚类结果. 可以看出, 这些聚集区域被分为两类, 大片区域为城区, 小块区域为树丛. 位置不相邻的小片树丛也能够分为一类.

### 3.2 结构区域边界定位

结构区域的线段有两种含义, 边界和线目标, 对于线目标, 我们希望能够保留而不被合并. 对于边界, 需要精确定位.

线目标的特点为垂直于线段方向的灰度值具有两次突变, 根据这个特点, 将满足条件的线段划分为线目标. 同时, 对于大于 3 个像素宽的线目标, 将检测出平行的两条线段, 因此, 距离相近且平行的线段也认为是线目标.

对于代表边界的结构区域, 本文通过极化边缘检测方法<sup>[12]</sup>得到精细的边界, 将边界两侧的区域进行超像素分割, 将这些超像素与相邻的匀质区域进行合并.

### 3.3 匀质区域的层次分割和分类

#### 3.3.1 层次分割

深度学习通过学习复杂图像的空间结构, 能够得到区域一致性良好的分类效果. 然而, 图像边界却很难保持. 对于匀质区域, 区域内部结构比较单一,

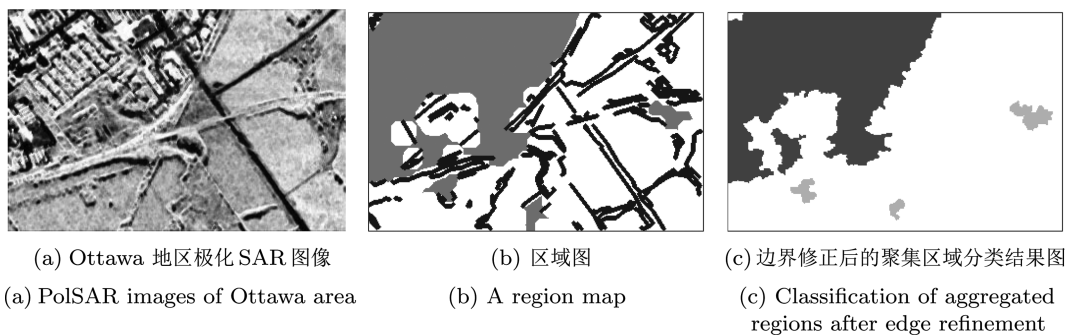


图 4 聚集区域分类示例图

Fig. 4 Example of classification of aggregated regions

底层的特征能够将其较好地表示, 分割的重点是不同区域之间的边界, 因此, 本文对匀质区域采用层次分割方法<sup>[4]</sup>进行分割. 首先, 采用均值漂移方法对匀质区域进行过分割得到超像素, 然后, 用层次分割对超像素进行迭代合并. 因为极化 SAR 协方差矩阵  $C$  满足 Wishart 分布, 本文采用最大似然的合并策略, 对于超像素  $S_i$  和  $S_j$ , 该测度定义如下<sup>[4]</sup>:

$$SC_{i,j} = MLL(S_i) + MLL(S_j) - MLL(S_i \cup S_j) = L(m_i + m_j) \ln |C_{S_i \cup S_j}| - Lm_i \ln |C_{S_i}| - Lm_j \ln |C_{S_j}| \quad (3)$$

其中,  $L$  是视数,  $m_i$  是超像素  $S_i$  内的像素个数,  $C_{S_i}$  是  $S_i$  的平均协方差矩阵,  $C_{S_i \cup S_j}$  为  $s_i \cup s_j$  的平均协方差矩阵, 其中,  $\cup$  为并操作.  $MLL(S_i)$  为超像素  $S_i$  的最大似然,  $\ln$  为求自然对数操作,  $|\cdot|$  为求行列式操作. 每次迭代时, 层次分割算法将测度值最小的两个超像素块进行合并. 此外, 匀质区域合并阈值为达到区域个数  $U$ .

### 3.3.2 空间极化分类器

在相干斑一致性假设下, 极化数据满足 Wishart 分布, Wishart 分类器<sup>[7]</sup>通过考虑极化统计特性进行分类, 得到广泛的应用. 本文对极化数据进行  $K$ -means 初始分类, 并使用 Wishart 进行迭代优化, 得到精细的分类结果. 由于没有考虑空间信息, 该算法容易受噪声影响, 分类结果的区域一致性较差.

匀质区域通过层次分割得到区域一致性好的分割结果, 然而, 不相邻区域的类别赋予是个难题. 空间极化分类器<sup>[28]</sup>能够融合分割和分类结果, 得到区域一致性好且边界精准的分类结果. 层次分割得到一致的区域, 而 Wishart 分类得到精细的类别, 因此, 空间极化分类器根据最大投票策略, 对匀质区域中的每个超像素, 赋予对应的分类图中的数目最多的类别. 这样得到的分类结果区域一致性好, 且不相邻区域的类别也能准确标记. 空间极化分类过程如图 5 所示, 给定分割图和基于像素的分类图, 分割图有 4 个超像素区域, 基于像素点的分类结果有三类, 将分割图映射到分类图上, 对每个分割区域中的类别使用最大投票策略, 最大数目的类别赋给这个超像素区域, 得到区域一致的分类结果.

## 4 实验结果和分析

### 4.1 实验数据和设置

实验数据: 为了验证本文算法的有效性, 对几幅真实的极化 SAR 数据进行实验, 这些数据来自不同波段不同卫星. 第一幅是 NASA/JPL AIRSA 卫星拍摄的 L 波段 San Francisco 地区的 4 视全极

化 SAR 数据, 大小为  $900 \times 700$ ; 第二幅为 Convair Ottawa 地区的单视全极化 SAR 数据, 大小为  $222 \times 3429$ ; 第三幅为 RadarSAT-2 C 波段西安地区极化 SAR 图像, 大小为  $512 \times 512$ , 分辨率为 8 m.

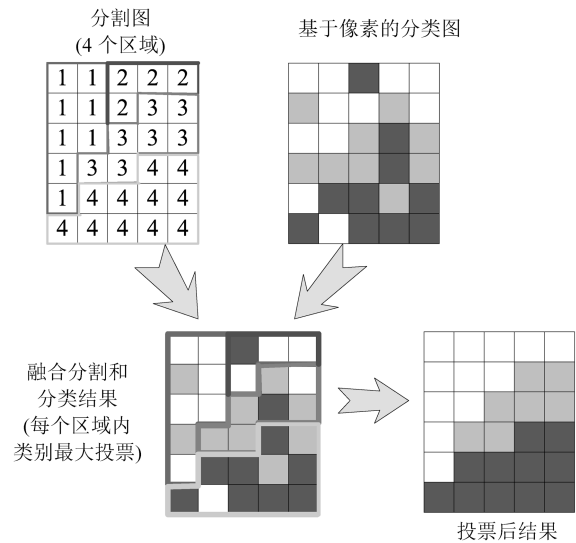


图 5 空间极化分类过程示意图

Fig. 5 Procedure of spatial-polarimetric classification

对比算法: 为了验证本文算法的优势, 三个相关的算法用来进行对比: 1) Wishart 分类算法<sup>[7]</sup>, 该算法在初始聚类后, 根据 Wishart 测度进行迭代优化, 得到最终分类结果; 2) Wishart MRF (Markov random field) 方法<sup>[29]</sup>, 该算法在 Wishart 分类中加入 MRF 邻域先验项, 引入空间信息, 获得空间一致的分类结果; 3) 堆叠自编码器 (Stacked auto-encoder, SAE) 分类方法<sup>[22]</sup>. 通过堆叠自编码器无监督地进行特征学习, 对特征进行稀疏表示后进行  $K$ -means 聚类. 该算法用来验证结合深度学习和层次语义模型的有效性.

网络结构设计: 网络的层数与数据样本个数和数据复杂度有关, 深度学习处理非常复杂的数据时, 层数会较多, 达到 10 层以上. 对于普通数据量, 4~8 层网络就能对原始数据进行很好的拟合. 对于复杂极化 SAR 图像, 根据图像的稀疏特性, 采样方式使得输入样本大大减少, 因此, 网络层数设置为 5.

每层的学习是对原始数据的不断抽象, 维数也不断减少, 图像块选取为  $13 \times 13$ , 那么, 对于  $13 \times 13 \times 9$  的输入维数, 如图 6 所示, 本文将 5 层的网络节点分别设置为 1089, 729, 441, 225 和 81, 分别代表不同窗口下的抽象特征.

另外, 字典个数设定为  $M = 10 \times \gamma$ ,  $\gamma$  为聚集区域个数. 为了避免欠分割, 匀质区域合并阈值  $U = S/2$ . 其中,  $S$  为均值漂移得到的超像素个数. 所有实验在硬件配置为 Intel Core i3 3.20 GHz 处

理器和 4 GB 内存的计算机上运行。

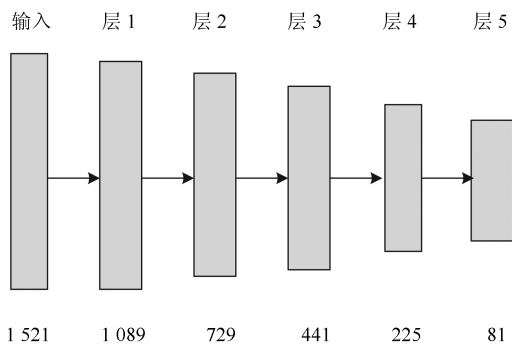


图 6 网络结构设计示意图

Fig. 6 Example of the network structure

#### 4.2 合成图像实验结果和分析

图 7(a) 为合成的极化 SAR 图像, 该图像由城区、海洋和森林三类地物构成。图 7(b) 为对应的标准类标图, 该图像是由 San Francisco 地区极化 SAR 图像的三类地物合成得到, San Francisco 地区图像如图 8(a) 所示。

图 7(c)~(f) 分别为 Wishart、Wishart MRF、SAE 和本文算法的分类结果图。从图 7 中可以看出, 本文算法能够得到更好的区域一致性和边界保持结果。Wishart 分类能够得到精细的结果, 然而城

区和森林两类被混淆。Wishart MRF 方法中森林出现大部分错分。在图 7(e) 中, 由于 SAE 学到的是高层的空间结构特征, 对边界部分难以精确定位, 因此, 圆形边界出现错分, 森林边界也难以保持。本文提出的算法在城区和森林都能够得到较为一致的分类结果, 圆形边界也能够较好的保持, 由于 Wishart 分类很难精确地分类森林边界, 文中算法对森林边界也难以完全保持, 但与其他算法相比, 本文算法能够得到更好的分类结果。

表 1 展示了本文算法和三个对比算法在合成图像的分类结果统计。其中计算了三类地物的分类精度、平均精度和 Kappa 系数。从表 1 中可以看出, 文中算法在分类精度达到 96.05%, 和其他三类算法相比, 分别提高了 12.42%、12.64% 和 0.87%。另

表 1 不同算法的分类结果统计 (%)

Table 1 Classification accuracies for different algorithms (%)

	Wishart	Wishart MRF	SAE	本文算法
城区	61.33	92.47	96.11	<b>92.94</b>
海洋	96.05	<b>99.23</b>	93.64	97.09
森林	93.52	58.52	95.79	<b>98.13</b>
平均精度	83.63	83.41	95.18	<b>96.05</b>
Kappa 系数	78.48	76.75	84.64	<b>94.30</b>

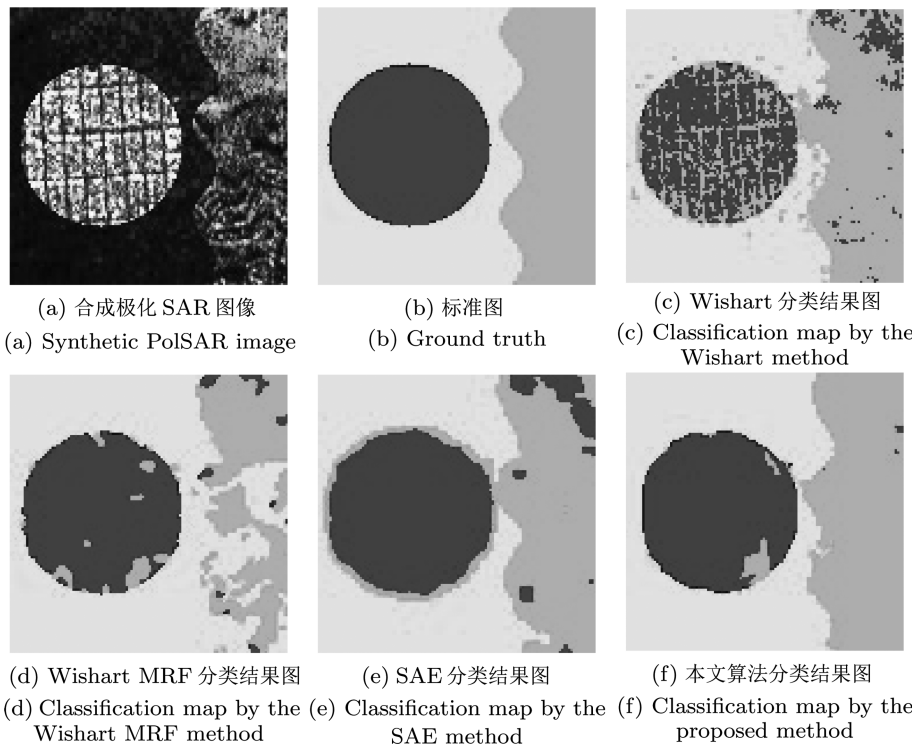


图 7 合成极化 SAR 图像分类结果图

Fig. 7 Classification maps of synthetic PolSAR image

外, Kappa 系数也高达 94.30%, 比其他算法分别高出 15.82%, 17.55% 和 9.66%。

文中算法的混淆矩阵如表 2 所示, 从表 2 中可以看出, 主要的错分来自于城区和森林的混淆. 另外, 表 3 给出了各类算法的运行时间. 可以看出, 文中算法和 SAE 方法使用的时间较长, 因为深度网络的学习比较耗时, 但文中算法比 SAE 时间短且能够得到更优的分类结果.

表 2 文中算法的混淆矩阵 (%)

Table 2 Confusion matrix for the proposed method (%)

	城区	海洋	森林
城区	92.94	2.70	4.36
海洋	0.40	97.09	2.51
森林	0.32	1.55	98.13

表 3 不同算法的运行时间 (s)

Table 3 Running time for different algorithms (s)

	Wishart	Wishart MRF	SAE	本文算法
时间	9.79	68.41	109.55	98.82

### 4.3 San Francisco 数据实验结果和分析

图 8 (a) 显示了 San Francisco 地区的极化 SAR 图像伪彩图. 从图 8 中可以看出, 这幅图像较为复杂, 含有多种地物类型, 左上角为山脉, 山脉周围为海洋, 海洋上的桥梁为金门大桥, 桥下面为山地, 其中有个高尔夫球场. 右下部分为城区, 城区中间有草地, 一些道路和人工目标也出现在城区中. 对这些复杂的场景进行精确的分类是具有挑战性的. 因为缺乏标记图, 为了便于理解, 图 8 (b) 给出了对应的 Google Earth 光学图像.

三个对比算法和本文提出的方法的实验结果显示在图 8 (c)~(f). 其中, 图 8 (c)~(e) 分别为 Wishart、Wishart MRF 和 SAE 算法分类结果, 图 8 (f) 为本文算法的结果. 从图 8 (f) 中可以看出, 与图 8 (c) 和 (d) 相比, 本文算法能够得到更加一致的区域, 能够将城区划分为语义上一致的区域, 同时, 不相邻的几块城区也能够正确标记. Wishart 分类结果虽然能够得到精细的分类, 但由于噪声的影响, 分类结果的区域一致性较差, 城区被分为多类的混合. 然而, 为了进一步的图像理解, 城区应该被分为

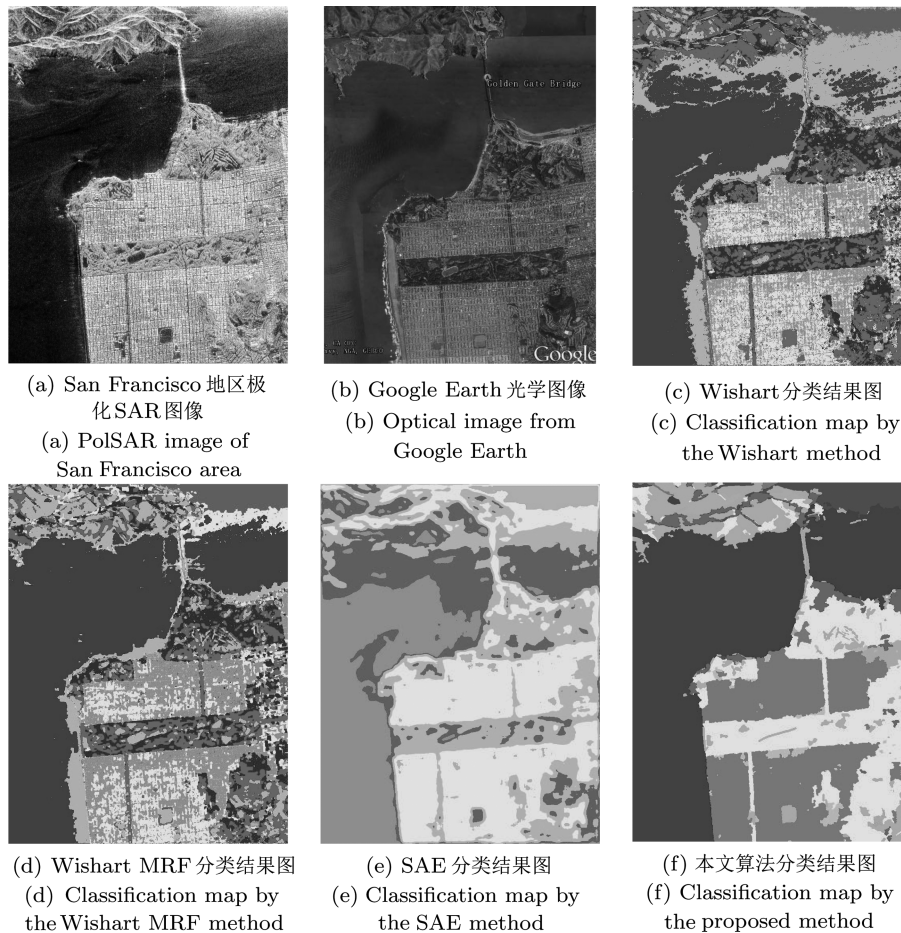


图 8 San Francisco 地区极化 SAR 图像分类结果图

Fig. 8 Classification maps of the PolSAR image on San Francisco area

语义上一致的区域. 图 8(d) 加入了 MRF 后, 空间一致性得到改善, 但城区、山地等仍然有很多过分割的小块. 图 8(e) 通过深度学习自动学习高层特征, 能够得到区域一致性较好的结果, 但由于高层特征不能精确定位边界, 使得边界模糊泛化. 与图 8(e) 相比, 本文得到更加准确的边界. 因此, 本文算法通过结合深度学习和层次语义模型, 使优势共存, 避免了单个算法的缺点, 得到区域一致性好且边界保持的分类结果.

#### 4.4 Ottawa 地区实验结果和分析

Ottawa 地区的单视极化 SAR 数据在方位方向进行 10 视处理, 得到  $222 \times 342$  的 10 视极化 SAR 数据, 处理后的 Ottawa 地区伪彩图如图 9(a) 所示, 该图像左上角为城区, 城区下方有一条铁路, 右侧为裸地, 其中有一些道路和小树丛.

图 9(b)~(e) 分别为 Wishart、Wishart MRF、SAE 和本文算法的分类结果图. 从图 9(b) 和 (c) 中可以看出, 城区被分为很多混杂的类或部分丢失. 从图 9(d) 中可以看出, 深度学习区域一致性好, 但丢失了一些线目标, 同时泛化了道路和边界. 图 9(e) 能够将城区划分为一致的类别, 同时, 将聚集区域中的城区和树丛分为不同的类. 另外, 道路和线目标也能够得到较好的保持.

#### 4.5 西安地区实验结果和分析

西安地区的极化 SAR 图像如图 10(a) 所示, 图像以渭河为中心, 左上角为城区, 右下角有小片城

区、村庄和大片裸地. 右上角有横跨于渭河之上的桥梁. 平行于桥梁的为一条铁路. 右侧有一条细的河流穿过. 图 10(b) 为对应的光学图像, 来源于 Google Earth. 光学图像和极化 SAR 图像不是同一时期获得, 因此会有少量差异, 但大体相同. 由于多种地物类型的存在, 对该图像的分类具有一定的难度.

图 10(c)~(f) 分别为三个对比算法和本文算法的分类结果图. 图 10(c) 为 Wishart 分类结果图, 图 10(d) 为 Wishart MRF 分类结果图, 图 10(c) 和 (d) 能够得到较为精细的分类结果, 然而, 裸地部分有所丢失, 且城区部分被分为多类, 而人眼视觉能够将这些类别整合为城区, 因此, 高层特征和语义信息的加入有助于对图像地物的理解和识别. 图 10(e) 通过自动学习高层特征, 对特征进行聚类得到的分类结果, 该结果区域一致性好, 城区部分较为一致, 然而, 河流右侧的边界丢失, 且右下角的细河流也基本消失. 图 10(f) 为本文提出的算法结果, 从图中可以看出, 本文算法对城区部分能够得到语义一致的分类结果, 同时对河流边界也能够较好地保持.

#### 4.6 参数分析

在堆叠自编码算法中, 图像块的大小是特征学习的重要参数. 图像块大小选取的原则是既能反映该类地物的结构, 又要避免包含多种结构. 另外, 图像块大小选取还与图像分辨率和结构复杂程度有关, 低分辨图像和结构简单的地物, 较小的图像块就能

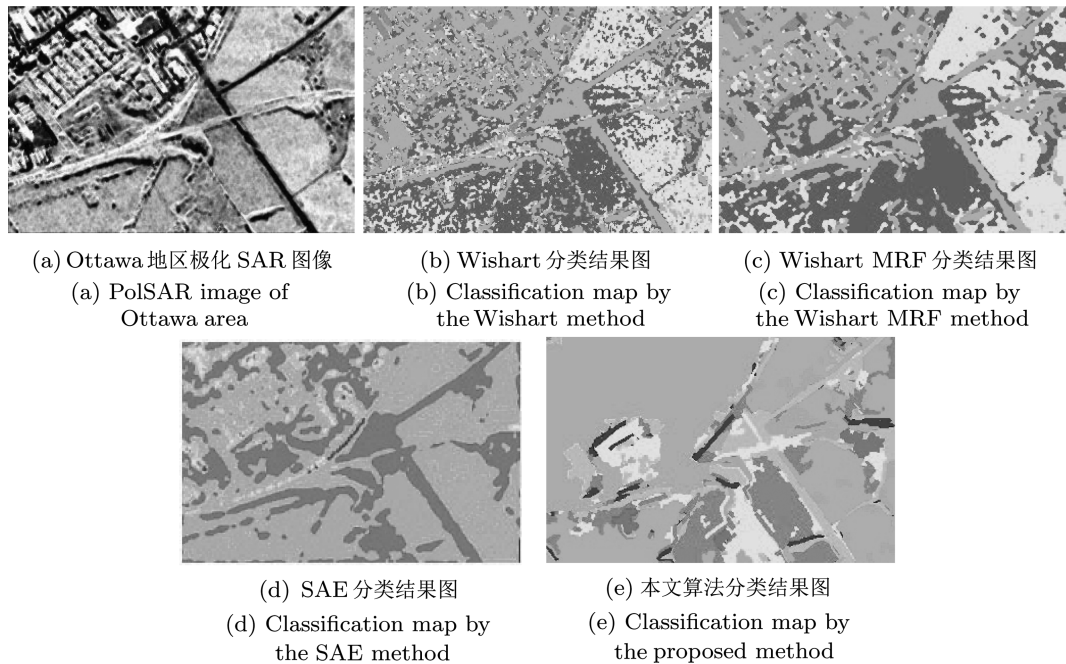


图 9 Ottawa 地区极化 SAR 图像分类结果图

Fig. 9 Classification maps of the PolSAR image on Ottawa area

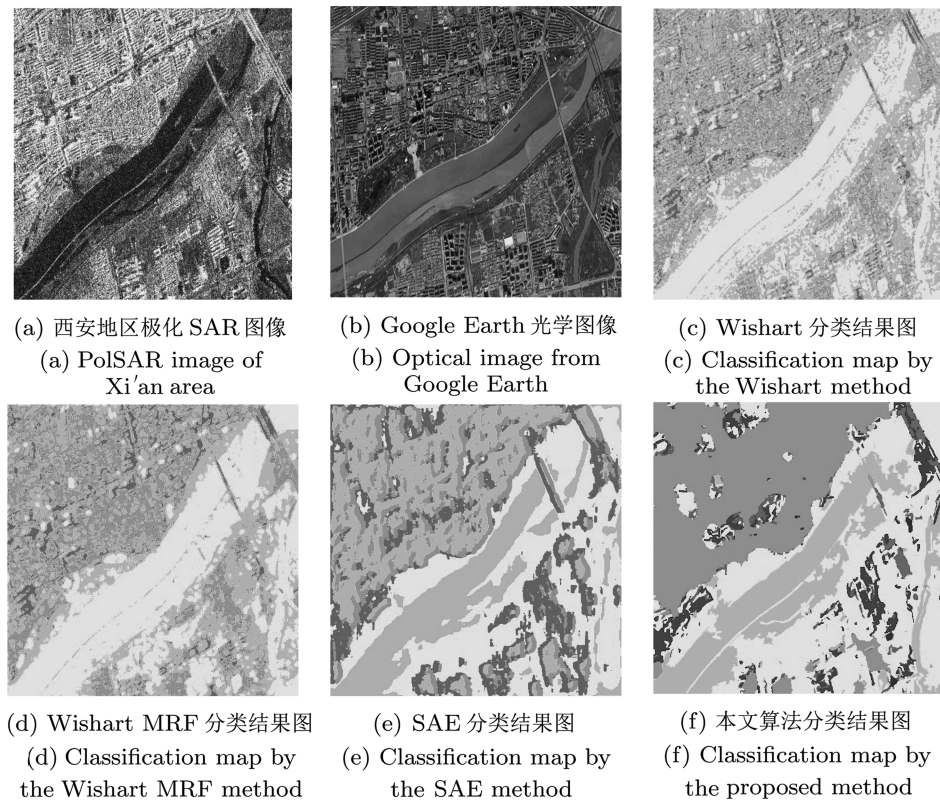


图 10 西安地区极化 SAR 图像分类结果图

Fig. 10 Classification maps of synthetic PolSAR image

包含地物结构; 反之, 应选取较大的图像块。

由于合成极化 SAR 图像 (图 7(a)) 包含标准图, 我们选择它来对图像块大小进行分析. 不同图像块大小对分类精度的影响如图 11 所示. 图像块大小分别从  $3 \times 3$  到  $31 \times 31$  进行实验. 从图 11 中可以看出, 本文算法在图像块大小取  $5 \times 5 \sim 13 \times 13$  之间都能取到稳定的较好的分类结果, 说明该算法对像块大小选取较为鲁棒. 然而, 当图像块太大或者

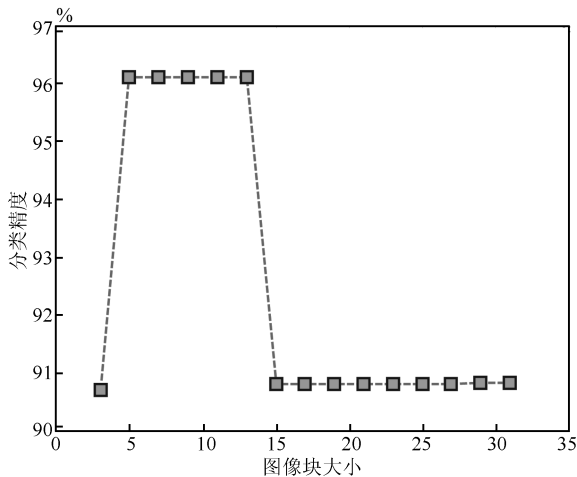


图 11 图像块大小对分类精度的影响

Fig. 11 Effect of block size on classification accuracy

图太小时, 不能得到好的分类结果, 那是因为太小的图像块难以包含图像结构, 而太大会包含多类结构, 使类间的特征差异性减少.

另外, 网络层数的设置也是用户根据经验选取. 不同层数网络对分类精度的影响的实验结果如图 12 所示,  $3 \sim 6$  层的网络结构分别用来进行比较. 从图 12 中可以看出, 图 7(a) 是较为简单的图像, 本文算法在不同网络层都能得到稳定的分类结果, 对网络层数较为鲁棒.

## 5 结束语

本文提出了一种新的无监督的极化 SAR 图像分类算法, 该算法结合了深度学习和层次语义模型的优势, 根据层次语义模型, 将图像首先划分为聚集、结构和匀质区域. 对聚集区域, 通过深度自编码模型学习地物结构特征, 将不同聚集地物区分开来. 对匀质和结构区域, 分别采用不同的策略进行分割和分类. 实验结果证明, 与传统的基于 Wishart MRF 方法和深度自编码方法相比, 该算法能够得到区域一致性好且边界保持的分类结果. 由于本文算法对聚集区域内部的线目标容易丢失, 依赖于层次语义模型的参数, 在后续工作中, 我们将加入相应的语义规则, 对重要目标进行保留. 此外, 如何自适应

地选择深度学习的参数也是我们下一步的工作。

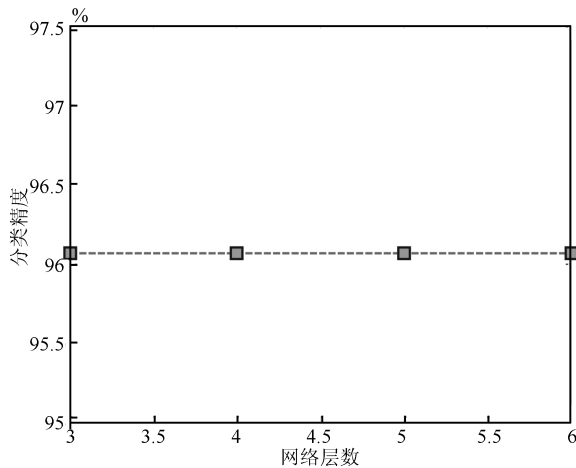


图 12 不同层数网络对分类精度的影响

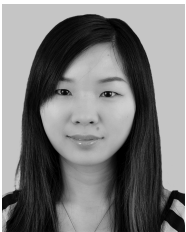
Fig. 12 Effect of network level on classification accuracy

## References

- 1 Cloude S R, Pottier E. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 1997, **35**(1): 68–78
- 2 Zhao L W, Zhou X G, Jiang Y M, Kuang G Y. Iterative classification of polarimetric SAR image based on the Freeman decomposition and scattering entropy. In: Proceedings of the 1st Asian and Pacific Conference on Synthetic Aperture Radar, APSAR 2007. Huangshan, China: IEEE, 2007. 473–476
- 3 Lee J S, Grunes M R. Classification of multi-look polarimetric SAR data based on complex Wishart distribution. In: Proceedings of the National Telesystems Conference. Washington D. C., USA: IEEE, 1992. 7/21–7/24
- 4 Beaulieu J M, Touzi R. Segmentation of textured polarimetric SAR scenes by likelihood approximation. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, **42**(10): 2063–2072
- 5 Shan Z L, Wang C, Zhang H, Wu F. Change detection in urban areas with high resolution SAR images using second kind statistics based G0 distribution. In: Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Honolulu, HI: IEEE, 2010. 4600–4603
- 6 Bombrun L, Vasile G, Gay M, Totir F. Hierarchical segmentation of polarimetric SAR images using heterogeneous clutter models. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, **49**(2): 726–737
- 7 Lee J S, Grunes M R, Ainsworth T L, Du L J, Schuler D L, Cloude S R. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, **37**(5): 2249–2258
- 8 Zhang B, Ma G R, Zhang Z, Qin Q Q. Region-based classification by combining MS segmentation and MRF for POL-SAR images. *Journal of Systems Engineering and Electronics*, 2013, **24**(3): 400–409
- 9 Ersahin K, Cumming I, Yedlin M. Classification of polarimetric SAR data using spectral graph partitioning. In: Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing. Denver, USA: IEEE, 2006. 1756–1759
- 10 Ersahin K, Cumming I G, Ward R K. Segmentation and classification of polarimetric SAR data using spectral graph partitioning. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, **48**(1): 164–174
- 11 Niu X, Ban Y F. An adaptive contextual SEM algorithm for urban land cover mapping using multitemporal high-resolution polarimetric SAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, **5**(4): 1129–1139
- 12 Liu F, Shi J F, Jiao L C, Liu H Y, Yang S Y, Wu J, Hao H X, Yuan J L. Hierarchical semantic model and scattering mechanism based PolSAR image classification. *Pattern Recognition*, 2016, **59**: 325–342
- 13 Bengio Y. Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, 2009, **2**(1): 1–127
- 14 Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1798–1828
- 15 Vincent P. A connection between score matching and denoising autoencoders. *Neural Computation*, 2011, **23**(7): 1661–1674
- 16 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25. Lake Tahoe, Nevada, USA: Curran Associates, Inc., 2012. 1097–1105
- 17 Salakhutdinov R, Hinton G. An efficient learning procedure for deep Boltzmann machines. *Neural Computation*, 2012, **24**(8): 1967–2006
- 18 Yildirim S, Cemgil T A, Aktar M, Ozakin Y, Ertuzun A. A Bayesian deconvolution approach for receiver function analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, **48**(12): 4151–4163
- 19 Lee H, Ekanadham C, Ng A Y. Sparse deep belief net model for visual area V2. In: Advances in Neural Information Processing Systems 20. Cambridge, MA: MIT Press, 2008. 873–880
- 20 Hyvärinen A, Hoyer P O. A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research*, 2001, **41**(18): 2413–2423
- 21 Ito M, Komatsu H. Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *The Journal of Neuroscience*, 2004, **24**(13): 3313–3324



- 22 Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P A. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 2010, **11**: 3371–3408
- 23 Marr D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W. H. Freeman and Company, 1982.
- 24 Guo C E, Zhu S C, Wu Y N. Primal sketch: integrating structure and texture. *Computer Vision and Image Understanding*, 2007, **106**(1): 5–19
- 25 Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603–619
- 26 Tirilly P, Claveau V, Gros P. Language modeling for bag-of-visual words image categorization. In: Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval. New York, USA: ACM, 2008. 249–258
- 27 Ning J F, Zhang L, Zhang D, Wu C K. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 2010, **43**(2): 445–456
- 28 Feng J L, Cao Z J, Pi Y M. Polarimetric contextual classification of PolSAR images using sparse representation and superpixels. *Remote Sensing*, 2014, **6**(8): 7158–7181
- 29 Rignot E, Chellappa R. Segmentation of polarimetric synthetic aperture radar data. *IEEE Transactions on Image Processing*, 1992, **1**(3): 281–300



石俊飞 西安理工大学计算机信息与工程学院讲师。2016 年获得西安电子科技大学计算机学院博士学位。2009 年获得河南师范大学计算机学院学士学位。主要研究方向为极化 SAR 图像分类, 语义模型, 计算机视觉。本文通信作者。  
E-mail: shijunfei3@126.com

(**SHI Jun-Fei** Lecturer at the School of Computer Science and Technology, Xi'an University of Technology. She received her Ph. D. degree from the School of Computer Science and Technology, Xidian University in 2016. She received her bachelor degree from the

School of Computer Science and Technology, Henan Normal University in 2009. Her research interest covers polarimetric SAR image classification, semantic model, and computer vision. Corresponding author of this paper.)



刘芳 西安电子科技大学计算机学院教授。1984 年获得西安交通大学计算机科学与技术专业学士学位, 1995 年获得西安电子科技大学计算机学院硕士学位。主要研究方向包括图像和信号处理, SAR 图像处理, 多尺度几何分析, 学习理论和算法, 数据挖掘。  
E-mail: f63liu@163.com

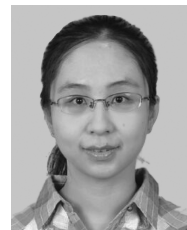
(**LIU Fang** Professor at the School of Computer Science, Xidian University and Technology. She received her bachelor degree in computer science and technology from Xi'an Jiaotong University in 1984 and master degree in computer science and technology from Xidian University in 1995. Her research interest covers synthetic aperture radar image processing, multiscale geometry analysis, learning theory and algorithms, and data mining.)



林耀海 福建农林大学计算机与信息学院讲师。主要研究方向有图像处理, 智能信号处理。E-mail: lyh953@qq.com

(**LIN Yao-Hai** Lecturer at the School of Computer Science and Technology, Fujian Agriculture and Forest University. His research interest covers image processing and intelligent signal

processing.)



刘璐 西安理工大学计算机信息与工程学院讲师。2015 年获得西安电子科技大学电子工程学院博士学位。主要研究方向为极化 SAR 图像分类。  
E-mail: liulu0613@163.com

(**LIU Lu** Lecturer at the School of Computer Science and Technology, Xi'an University of Technology. She received her Ph. D. degree from the School of Electronic Engineering, Xidian University in 2015. Her main research interest is polarimetric SAR image classification.)