

基于 DBNMI 模型的海洋遥感影像自动标注方法

黄冬梅, 许琼琼, 杜艳玲, 贺 琪

(上海海洋大学信息学院, 上海 201306)

摘要:研究大规模海洋遥感影像管理的关键是缩小影像低层视觉特征与高层语义之间的鸿沟.针对海洋遥感影像中不同区域对语义相似性度量的贡献程度不同,提出一种基于深度信念网络多示例(deep belief networks multi-instance, DBNMI)的遥感影像语义自动标注模型.模型对初始输入遥感影像进行自适应分割,粗粒度划分海洋遥感影像背景区域和对象区域;对影像对象区域的低层视觉特征和高层语义概念间关系,利用深度信念网络模型进行自动建模;定量计算标注词间共现和对立的语义关系,改善图像标注结果.在公开遥感影像数据集上进行验证,实验表明所提出方法在标注精度上取得了较好效果.

关键词:深度信念网络;自适应分割;海洋遥感影像;图像标注

中图分类号:TP 391 **文献标识码:**A doi:10.3969/j.issn.0253-2778.2017.03.010

引用格式:黄冬梅,许琼琼,杜艳玲,等.基于 DBNMI 模型的海洋遥感影像自动标注方法[J].中国科学技术大学学报,2017,47(4):323-330.

HUANG Dongmei, XU Qionqiong, DU Yanling, et al. Ocean remote sensing image auto-annotation based on DBNMI model[J]. Journal of University of Science and Technology of China, 2017, 47(4): 323-330.

Ocean remote sensing image auto-annotation based on DBNMI model

HUANG Dongmei, XU Qionqiong, DU Yanling, HE Qi

(Department of Information and Technology, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Bridge the semantic gap between low-level visual feature and high-level semantic concepts has been the subject of intensive investigation on large scale remote sensing image management for years in order to improve the accuracy of automatic image annotation. An ocean remote sensing image auto-annotation method based on DBNMI Model was proposed for contributions of semantic similarity about different regions of ocean remote sensing images. Initial remote sensing images were adaptively segmented, ocean remote sensing images were divided into background and the object region by means of a coarse-grained method, the relationship between low-level visual feature and high-level semantics label of the object region was modeled automatically, using DBN model, and the co-occurrence relations and adversarial relations between semantic concepts for improving image annotation results were calculated. The proposed approach is evaluated on a public remote sensing image dataset. The experimental results show a satisfactory improvement on accuracy.

Key words: deep belief networks; adaptive segmentation; ocean remote sensing image; image annotation

收稿日期:2016-08-28;修回日期:2016-12-08

基金项目:国家重点基础研究发展计划(2012CB316206),国家自然科学基金(61272098,61402282)资助.

作者简介:黄冬梅,女,博士/教授,1964年生,研究方向:大数据和智能信息处理. E-mail:dmhuang@shou.edu.cn

通讯作者:贺琪,博士/副教授. E-mail:qihe@shou.edu.cn

0 引言

遥感影像语义自动标注是先让机器通过学习已知语义标签的影像示例,然后使用语义概念表示一幅未标注遥感影像的内容.随着多颗海洋遥感卫星的发射(如 SeaStar、GEOSAT、ERS- n 、HY-1 到 HY-3 等卫星),海洋遥感影像数量急剧增长.通过遥感影像语义标注,可以实现遥感影像数据分析、索引和检索等任务,高效利用遥感影像.人工标注的方法具有主观性及不精确性,费时费力,已不能满足当今“大数据”时代.机器学习等理论的发展,使得遥感影像语义自动标注成为一大研究热点.

众多学者研究遥感影像语义标注方法主要从以下方面着手:特征提取^[1-2]、分类模型^[3-5]、高层语义划分层次^[6]等.Bratanan 等^[1]提出一种将相同语义概念的像素组合到一起的映射方式,完成无语义低层特征到光谱语义索引映射再到人类理解的高层语义的映射.郑歆慰等^[2]提出一种空间约束多特征联合稀疏编码模型,实现遥感影像的自动标注.Liénou 等^[3]对每个概念给定一个训练集,利用最大似然方法给未知图像分配概率,基于潜在狄利克雷分配模型(latent Dirichlet allocation, LDA)模型进行学习.为了克服基于 LDA 模型的遥感图像标注方法不能对多光谱遥感影像进行标注的缺陷, Luo 等^[4]提出一种基于作者主题模型(author topic model, ATM)的遥感图像标注方法,通过颜色和形状特征描述符,对测试影像语义类的概率进行估计,进而分类.当区域类别多时,不同类区域低层特征相似,基于 ATM 的遥感图像标注方法分类准确度下降.为提高标注精度, Luo 等^[5]又提出一种基于作者-流派-主题模

型(author-genre topic model, AGTM)遥感影像标注方法,利用流派代表场景的标签,有效提高了标注精度.Chen 等^[6]在第一层利用多示例多标签学习(multi-instance multi-label learning, MIML)进行高分辨率遥感影像标注,第二层采用多尺度分层语义信息模拟示例在 MIML 的空间关系,第三层利用非参数概率模型合并可能的类.由于语义鸿沟的存在,影像语义的自动标注性能亟待提高.

近年来,深度学习在文本、图像和语音领域得到广泛应用,并取得很好的成效.Wei 等^[7]为解决多标签图片分类问题,提出一种灵活、深度的假说聚合型卷积神经网络模型(hypotheses-CNN-pooling, HCP),取得了较好的效果.杨阳等^[8]提出改进的深度神经网络模型,将图像类别信息作为标注信息,利用深度神经网络建立图像标注模型.吕启等^[9]提出基于深度信念网络(deep belief network, DBN)模型的遥感图像分类方法,在 SAR 图像数据的实验结果表明,分类结果比支持向量机(support vector machine, SVM)和传统神经网络(neural network, NN)方法的效果有显著提高.高常鑫等^[10]提出分层深度学习模型,对海量高分辨率遥感影像数据实现典型目标的检测.

现有工作对海洋遥感影像结构的复杂性考虑不多,一幅遥感影像的不同区域对语义相似性度量的贡献程度不同,海洋遥感影像中“对象区域”比“背景区域”的贡献要大很多.本文提出一种基于深度信念网络多示例(deep belief networks multi-instance, DBNMI)的遥感影像语义自动标注模型,主要工作如图 1 所示,分为三部分:①提出一种基于自适应多尺度分割方法,粗粒度划分海洋遥感影像背景区域

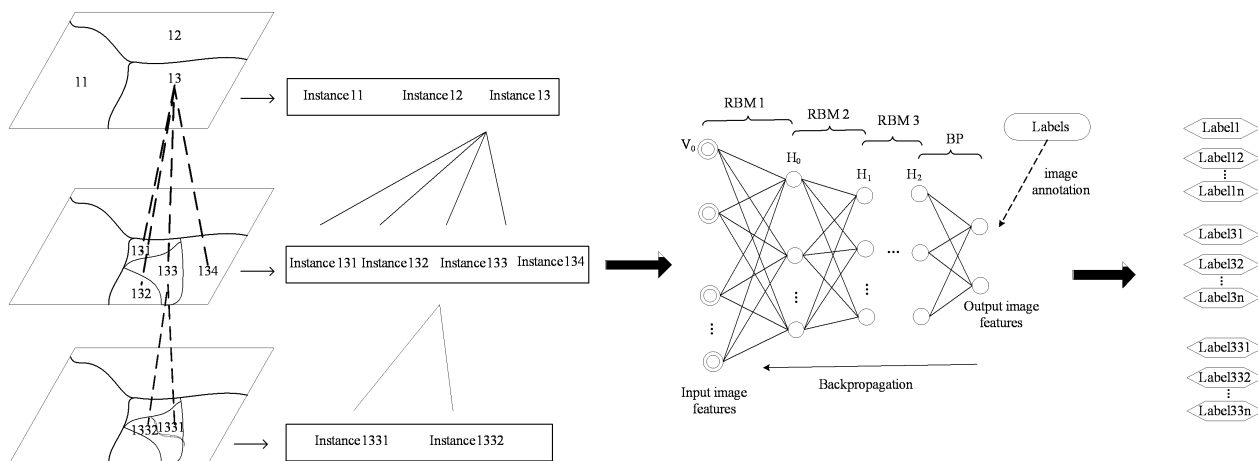


图 1 基于 DBNMI 模型影像标注过程

Fig.1 Process of image annotation based on DBNMI model

和对象区域,多示例表示海洋遥感影像;②利用深度信念网络模型对影像低层视觉特征和高层语义概念关系进行自动建模;③定量计算标注词间共现和对立的语义关系,改善图像标注结果。

1 基于自适应分割的遥感影像多示例构建

遥感影像空间结构及地理特征信息丰富^[11],而海洋遥感影像语义标注任务感兴趣的区域仅占影像很小一部分.在不同观测尺度下,感兴趣的对象和蕴含的语义概念也会不同,如一幅遥感影像在低分辨率下可以标注为“海岛”,但是在高分辨率下可以标注为“沙滩”和“森林”,因此本文对初始输入遥感影像进行处理,粗粒度划分海洋遥感影像背景区域和对象区域。

根据影像场景复杂性选择合适的分割尺度,简单场景选择大的尺度,复杂场景选择小的尺度,如图 2 所示.小波变换能提供影像在不同分辨率下的表达,实现影像按照不同尺度和位移的分解^[11].本文利用小波变换生成多分辨率影像的金字塔形式,自适应构建不同尺度的影像信息等级结构,完成遥感影像信息在不同尺度层间的传递。

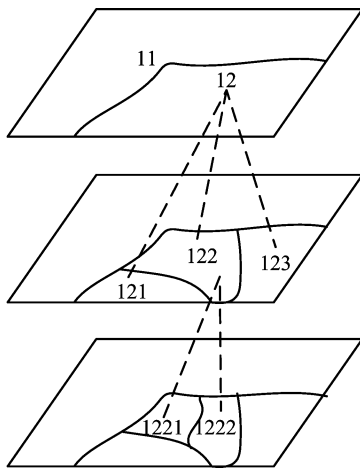


图 2 自适应多尺度分割

Fig.2 Adaptive multiscale segmentation

一幅海洋遥感影像概括的地理范围很大,利用 MIML 技术表示海洋遥感影像.定义一幅遥感影像用 (X_i, Y_i) 表示,其中 $X_i = \{x_i^1, x_i^2, \dots, x_i^m\}$ 表示示例空间, m 是示例个数, $Y_i = \{y_i^1, y_i^2, \dots, y_i^l\}$ 表示标签集, l 为标签个数;旨在从训练集中得到一个函数: $f_{MIML}: X^x \rightarrow 2^y$, 这里输入是一幅遥感影像产生的示例空间 X_i , 输出是该幅遥感影像对应的标签

集 Y_i .

每幅影像在不同空间尺度中作为一个包 (bag), 相关的分割区域视为示例 (instances)^[12], 如图 3 所示.其中第 i 层用包 $B_i = \{I_{i1}, I_{i2}, \dots, I_{im}\}$ 表示.设置阈值 t , 计算背景区域与对象区域相似度 s , 若 s 大于预设阈值 t , 第 i 层包需要继续划分; 小于阈值 t 则不需要。

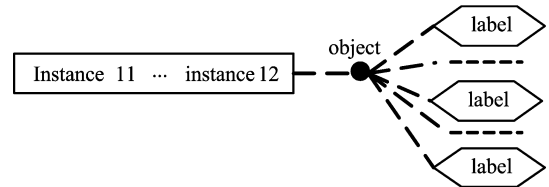


图 3 多示例多标签学习

Fig.3 Multi-instance multi-label learning

2 基于 DBN 模型的对象区域建模

研究大规模海洋遥感影像管理的关键是缩小影像低层视觉特征与高层语义之间的鸿沟.DBN 模型将输入数据逐层特征变换,转换成适合标注的理想特征,提高标注精度。

深度信念网络包括多层受限玻尔兹曼机 (restricted Boltzmann machines, RBM) 网络和一层反向传播 (back propagation, BP) 网络^[13].受限玻尔兹曼机作为深度信念网络模型的核心,由可见层和隐含层两层神经元组成, $v \in \{0, 1\}^I$ 表示可见层节点的状态, I 为可见层节点数目; $h \in \{0, 1\}^J$ 表示隐含层节点的状态, J 为隐含层节点数目。

当给定可见层节点的状态时,各隐含层节点的激活状态之间是相互独立的,即

$$P(h | v) = \prod_{j=1}^J P(h_j | v) \tag{1}$$

给定隐含层节点的状态时,各可见层节点的激活状态之间是相互独立的,即

$$P(v | h) = \prod_{i=1}^I P(v_i | h) \tag{2}$$

对遥感影像对象区域进行特征提取,将提取出的特征作为受限玻尔兹曼机的可见层输入, $v \in \{0, 1\}^I$ 二元变量,假设每个神经元遵循伯努利分布,定义该模型的能量函数如下:

$$E(v, h | \theta) = - \sum_i \sum_j w_{ij} v_i h_j - \sum_i b_i v_i - \sum_j a_j h_j \tag{3}$$

式中, $\theta = \{a, b, W\}$ 是模型的参数.可得到 (v, h) 的联合概率分布为

$$P(v, h) = e^{-E(v, h)} / Z \tag{4}$$

式中, Z 为归一化常数. 利用迭代的方式进行训练, 旨在学习出参数 $\theta = \{a, b, W\}$ 的值.

利用训练集影像, 将粗粒度划分的影像在不同空间尺度的包分别作为受限玻尔兹曼机的输入数据; 进行逐层特征变换, 将样本的特征表示从原空间转换到新特征空间, 获得从低层特征到高层语义表达最佳的权值, 利用反向传播网络对网络进行调优, 实现数据从低层到高层的特征提取, 形成适合标注的理想特征; 对不同包进行标注, 最终不同尺度层的标注词, 均作为该输入影像的标注词, 如图 4 所示.

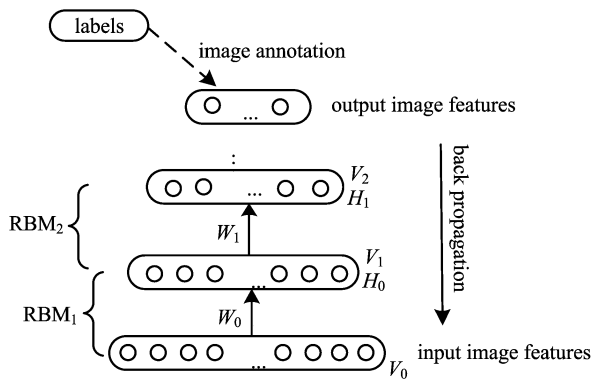


图 4 基于 DBN 模型影像自动标注

Fig.4 Image auto-annotation based on DBN model

3 标注词语义相关性度量

利用学习的模型自动完成新影像语义的标注, 遥感影像标注词间存在不同语义层次关系, 如相近、对立、包容等. 一副遥感影像被标注为“轮船”、“海水”等词, 则“港口”作为该影像标注词的概率也会提升.

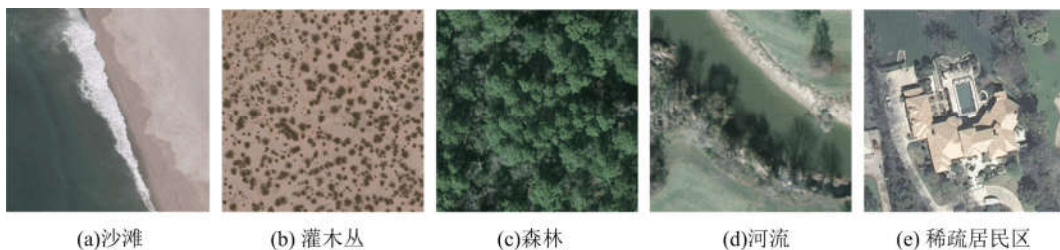


图 5 五个类的遥感影像示例

Fig.5 Training images from five semantic classes

采用 10 次交叉验证方法评估隐含层节点数的影响, 即将全部影像样本数据平均分成 10 份, 每次使用其中的 9 份用于训练, 剩余 1 份用于测试, 保证每个类的影像至少出现一次. 利用误差、精度验证影像标注方法的性能. 假设以某一标注词 ω 作为查询, 对已标注好的测试集影像数据进行检索, 假设标

MIML 中若包中至少有一个正例, 则该包被标记为正 (positive), 若包中所有示例都是反例, 则该包被标记为反 (negative). 这种简单的表达方式不足以描述标注词的语义层次和关系, 本文考虑标注词间的共现、对立关系, 定量计算标注词间的相似性.

定义 3.1 标注词间的共现关系可表示为: 当 $F(B_i, L) = 1$ 时,

$$\forall l \in L, \text{count}(B_i, l) \geq 1 \quad (5)$$

式中, B_i 为第 i 个包, L 代表语义概念集. 表示包 B_i 中含有标签 l , 则该包可以选择该标签所属语义概念集 L 中的其他标签.

定义 3.2 标注词间的对立关系可表示为: 当 $F(B_i, L_1, L_2) = 1$ 时,

$$\forall l_1 \in L_1, \forall l_2 \in L_2, \text{count}(B_i, l_1) \geq 1, \Delta(B_i, l_2) = 0 \quad (6)$$

式中, 语义概念集 L_1, L_2 属于对立关系, 包 B_i 中含有标签 l_1 , 则该包可以选择该标签所属语义概念集 L_1 中的其他标签, 但不能选择语义概念集 L_2 中的标签.

4 实验结果与分析

4.1 实验设置与评价

实验采用公开遥感数据集^[14]进行海洋遥感影像标注, 影像大小为 256×256 像素. 由于海洋遥感影像语义标注任务感兴趣的区域仅占影像很小一部分, 故选择数据集中海洋遥感影像代表性的 5 个类 (沙滩、灌木丛、森林、河流、稀疏居民区), 如图 5 所示.

注正确的影像数为 $|W_c|$, 标注错误的影像数 $|W_e|$, 可检索到的所有影像数为 $|W_M|$, 则标注的误差为

$$\text{error} = \frac{|W_e|}{|W_M|} \quad (7)$$

精度为

$$\text{precision} = \frac{|W_c|}{|W_M|} \quad (8)$$

为了验证提出方法的有效性,并与其他方法进行比较,设计了 3 项任务:

(I) 检验 DBNMI 方法结构对海洋遥感影像语义自动标注的影响;

(II) 验证 DBNMI 适合影像语义自动标注的程度;

(III) 验证语义相关性度量方式的有效性.

针对任务(I),设计了 4 组不同的实验进行影像标注.参数设置如下:隐含层层数分别取 1,2,3,4 层,每一隐含层的节点数从 100 取到 1 000,取值间隔为 100.

针对任务(II),设计了 4 组不同的实验.分别基于 SVM、NN、DBN 和 DBNMI 标注方法进行对比实验.SVM 于 1995 年首次提出,对高维度数据的标注有较好效果.本文采用高斯径向核(Gaussian RBF)函数,利用 LIBSVM 软件包^[15]进行基于 SVM 影像自动标注实验,基于 NN、DBN 方法均使用 DeepLearnToolbox 中代码,设置参数与 DBNMI

类似.

针对任务(III),设计了 4 组实验.采用 4 种方法:不考虑标注词间共现和对立关系进行的影像标注方法(not-DBNMI, nt-DBNMI),仅考虑标注词间共现关系的标注方法(co-occurrence tag-DBNMI, ct-DBNMI)仅考虑标注词间对立关系的标注方法(opposite tag-DBNMI, ot-DBNMI)以及本文 DBNMI 方法进行对比.实验中所有算法在 Matlab2013a 平台运行.

4.2 结果分析

4.2.1 模型参数设置对该方法的影响

表 1 是在隐含层层数分别取 1,2,3,4 层,每层节点数相同情况下,不同隐含层数不同隐含层节点数对误差的影响.

从表 1 可以看出,当隐含层数为 1、2 层时,节点数为 300 误差最小为 0.018.当隐含层数为 3 层时,隐含层节点数为 200 误差最小为 0.015.当隐含层数为 4 层时,隐含层节点数为 100 误差最小为 0.069.

表 1 DBNMI 参数对标注误差的影响

Tab.1 Effect of different parameters of DBNMI'error

隐含层数	节点数									
	100	200	300	400	500	600	700	800	900	1000
1	0.035	0.022	0.018	0.048	0.053	0.102	0.105	0.114	0.133	0.159
2	0.049	0.046	0.025	0.076	0.087	0.101	0.152	0.154	0.154	0.156
3	0.041	0.015	0.046	0.051	0.093	0.101	0.101	0.128	0.152	0.152
4	0.069	0.247	0.074	0.101	0.101	0.105	0.105	0.109	0.152	0.156

为研究隐含层及节点数目对误差的影响,对表 2 进行图示,如图 6 所示.

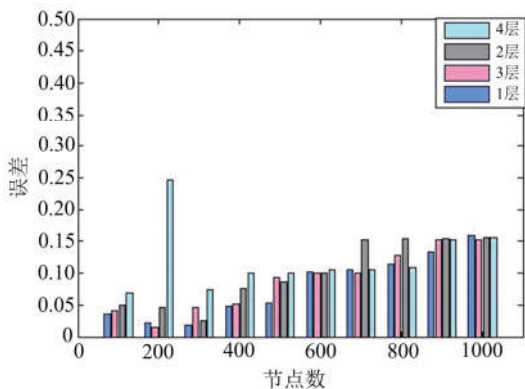


图 6 不同隐含层及节点数目的误差对比

Fig.6 Error comparison of different layers and nodes

由图 6 可知,随着隐含层节点数的增加,误差先减小后增大.主要是因为特征越多,训练影像数据在

每个特征上会稀疏,降低了 DBN 的泛化性能,导致过拟合现象.比较突出地,隐含层数为 4 层,节点数为 200 时,误差达到最高,可能造成的原因是隐含层层数增加导致反向梯度下降算法的误差增大,此后隐含层节点数的增加误差率都是不准确的.

说明本文海洋遥感影像语义自动标注方法中隐含层数最佳为 3 层,隐含层节点数最佳为 200.

4.2.2 相关方法的比较与分析

表 2 是 4 种方法的标注对比实验.相比于 SVM 标注方法,基于 DBNMI 的标注方法精度提高了 5%.与 SVM 针对每一个标注词设计分类器相比,深度学习逐层转换低层特征形成更加适合标注高层的特征,做法更加合理.

基于 DBNMI 的标注方法标注精度均优于 NN、DBN 的标注方法,分别提高 2%,3%.主要是因为本文考虑到了海洋遥感影像的复杂结构,自适应

分割海洋遥感影像,对标注性能有更好的效果.

表 2 不同标注方法的精度比较

Tab.2 Comparison of different methods

方法	SVM	NN	DBN	DBNMI
精度	0.73	0.76	0.75	0.78

在研究海洋遥感影像标注时,面对大规模海洋遥感影像数据,基于 DBNMI 的标注方法针对遥感影像数据特征规律进行无偏差估计,将特征学习融入到建立模型的过程中,具有海量遥感影像数据的标注优势.

4.2.3 标注词相关性度量检验

表 3 给出了是否考虑标注词间对立和共现关系的 DBNMI 方法对影像的标注结果.相比于 nt-DBNMI 方法,DBNMI 的标注方法精度提高了 1.5%.

表 3 标注词相关性度量有效性验证

Tab.3 Validation of label correlation measurement

方法	nt-DBNMI	ct-DBNMI	ot-DBNMI	DBNMI
精度	0.765	0.774	0.772	0.780

与 ct-DBNMI、ot-DBNMI 标注方法相比,DBNMI 标注方法标注精度都略有增加,分别提高 0.6%,0.8%.这说明本文提出的标注词改善方法,考虑了标注词的共现关系和对立关系,有效提高了影像语义标注的性能.

5 结论

遥感影像数据呈几何级增长,高效管理遥感影像成为研究热点.本文提出了一种基于深度信念网络多示例的遥感影像语义自动标注模型,以提高标注精度.实验表明,本文方法考虑了海洋遥感影像的复杂结构,对海洋遥感影像进行自适应分割,提升了海洋遥感影像语义标注效果.将来的工作拟在以下几个方面展开:①将本文提出的方法在更大规模的影像数据集上测试,以验证方法的普适性,进一步测试需要更大规模的实验平台;②图像分割算法、低层视觉特征之间的权重等对标注精度的影响也值得深究.

参考文献(References)

- [1] BRATASANU D, NEDELCU I, DATCU M. Bridging the semantic gap for satellite image annotation and automatic mapping applications[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2011, 4(1): 193-204.
- [2] 郑歆慰,胡岩峰,孙显,等.基于空间约束多特征联合稀疏编码的遥感图像标注方法研究[J].电子与信息学报,2014,36(8):1891-1898.
- [3] LIÉNOU M, MAÏTRE H, DATCU M. Semantic annotation of satellite images using latent dirichlet allocation[J]. Geoscience and Remote Sensing Letters, 2010, 7(1): 28-32.
- [4] LUO W, LI H L, LIU G H. Automatic annotation of multispectral satellite images using author - topic model[J]. Geoscience and Remote Sensing Letters, 2012, 9(4): 634-638.
- [5] LUO W, LI H L, LIU G H, et al. Semantic annotation of satellite images using author - genre - topic model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(2): 1356-1368.
- [6] CHEN K M, JIAN P, ZHOU Z X, et al. Semantic annotation of high-resolution remote sensing images via Gaussian process multi-instance multilabel learning[J]. Geoscience and Remote Sensing Letters, 2013, 10(6): 1285-1289.
- [7] WEI Y C, XIA W, HUANG J S, et al. CNN: Single-label to multi-label [J]. Computer Science, 2014; arXiv:1406.5726.
- [8] 杨阳,张文生.基于深度学习的图像自动标注算法[J].数据采集与处理,2015,30(1):88-98.
- [9] 吕启,窦勇,牛新,等.基于 DBN 模型的遥感图像分类[J].计算机研究与发展,2014,51(9):1911-1918.
- [10] 高常鑫,桑农.基于深度学习的高分辨率遥感影像目标检测[J].测绘通报,2014,(S1):108-111.
- [11] 黄志坚,黎湘,徐帆江.基于视觉复杂度的自适应尺度遥感影像分割[J].电子与信息学报,2013,35(8):1786-1792.
- [12] 葛永,吴秀清,洪日昌.基于多示例学习的遥感图像检索[J].中国科学技术大学学报,2009,39(2):132-136.
- [13] GE Y, WU X Q, HONG R C. Remote sensing image retrieval based on multiple instance learning [J]. Journal of University of Science and Technology of China, 2009, 39(2): 132-136.
- [14] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [15] YANG Y, NEWSAM S. Bag-of-visual-words and spatial extensions for land-use classification [C]// Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. San Jose, USA: ACM Press, 2010: 270-279.
- [16] CHANG C C, LIN C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 27: 1-27:27.