第9期
2016年9月

Vol. 44 No. 9 Sep. 2016 ΙL

# 改进贝叶斯 ARTMAP 的 迁移学习遥感影像分类算法

## 韩敏,杨雪

(大连理工大学电子信息与电气工程学部,辽宁大连116023)

摘 要: 遥感分类旨在从图像光谱中提取资源环境监测可用的地理信息,然而基于模式分类的图像处理技术受 光谱漂移影响而缺乏历史样本重复利用的有效策略,制约着有限目标样本下遥感分类精度的提高.针对该问题,本文 构建了基于改进的贝叶斯 ARTMAP 神经网络的迁移学习遥感影像分类算法,通过提高谐振匹配性来抑制类别扩散, 利用节点的离散增量期望最大化参数更新策略,将历史遥感样本中的地物分类先验信息迁移到目标模型当中.实验结 果表明本文方法能有效利用历史遥感数据弥补缺少目标训练数据的不足,相比于其他样本利用策略大幅提高遥感影 像分类精度.

关键词:	遥感分	分类;	信息提取;	迁移学习;	ARTMA	ΑP;	神经网络		
中图分类号	<b>ት:</b> ገ	FP751;	TP181	文献标识	只码:	Α	文章编号:	0372-2112 (2016	) 09-2248-06
电子学报 U	JRL:	http://	/www.ejour	nal. org. cn			<b>DOI</b> : 10.39	69/j. issn. 0372-211	2.2016.09.033

## Transfer Learning Using Improved Bayesian ARTMAP for Remote Sensing Image Classification

### HAN Min, YANG Xue

(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning116023, China)

Abstract: Remote sensing classification aims at extracting available geographic information from image spectrum for resources and environment monitoring, but due to the spectral drift effect, the lack of effective strategies on historical sample reuse for image processing technology based on pattern classification restricts remote sensing classification accuracy with limited target samples. To solve this problem, this paper proposes a transfer learning algorithm for remote sensing classification using improved Bayesian ARTMAP neural network. More productive resonance matching is used to suppress the unattractive property of category proliferation, so that the incremental expectation maximization can be introduced to update parameters adaptively. The classification prior knowledge of the historical samples is transferred to the target model. The experimental results show that this method can effectively compensate for the lack of target training data by reusing the historical samples and significantly improve the accuracy of remote sensing image classification compared with other sample utilization strategy.

Key words: remote sensing classification; information extraction; transfer learning; ARTMAP; neural network

## 1 引言

难以快速采集充足而准确的样本数据已成为遥感 影像分类应用面临的最大问题.同时,现有算法缺乏重 复利用历史样本的有效策略,导致历史样本资源的浪 费<sup>[1]</sup>.如何将历史数据中的先验知识迁移到目标模型 当中,是提高遥感影像分类效率与精度的新的切入点. 传统的分类算法无法对历史数据展开有效应用的 主要原因是:历史样本光谱特性与待分类影像的光谱 特性不服从相同的概率统计分布(光谱漂移现象)<sup>[2]</sup>. 迁移学习是近年来提出解决域适应问题的机器学习理 论<sup>[3,4]</sup>,如基于 Adaboost 的迁移学习算法<sup>[5]</sup>、半监督的 域适应迁移成分分析方法<sup>[6]</sup>、支持向量机的迁移学习 理论<sup>[7,8]</sup>.但目前遥感影像方面的迁移学习研究成果还

收稿日期:2015-01-14;修回日期:2015-04-29;责任编辑:覃怀银

基金项目:国家重点基础研究发展规划(973 计划)项目(No. 2013CB430403);国家自然科学基金项目(No. 61374154)

较少<sup>[9]</sup>,如吴田军提出遥感影像对象级分类的迁移学 习方法<sup>[10]</sup>;以及 Liu 提出的基于案例推理时空信息重 用的遥感影像分类方法<sup>[11]</sup>.但是上述方法均需要大量 的遥感数据积累以及复杂的数据管理方案来建立系 统,如何直接地将历史遥感影像中的有效信息迁移到 目标模型中,从而降低影像分类对目标影像中有标记 样本的依赖仍需要进一步的研究.

本研究中以贝叶斯 ARTMAP<sup>[12]</sup>为具体模型来重点 探讨迁移学习在遥感影像分类中的问题.提出基于改 进的贝叶斯 ARTMAP 的迁移学习遥感影像分类算法 (Transfer Learning using Improved Bayesian ARTMAP,TL-IBAM).首先改进贝叶斯 ARTMAP 谐振机制,降低类别 扩散的影响,之后将历史遥感影像数据下得到的网络 参数作为迁移学习中的先验信息,在目标影像的样本 信息支撑下,以离散增量形式实现参数期望最大化,将 历史数据中的知识迁移到目标模型中,实现贝叶斯 AR-TMAP 网络的迁移学习.

## 2 贝叶斯 ARTMAP 神经网络

贝叶斯 ARTMAP<sup>[12]</sup> 由输入模块 ART<sub>a</sub>和输出模块 ART<sub>b</sub>组成,中间由映射域进行连接,其本身的特性使其 能够不断地学习新的模式,其训练算法可由图1 简要描述,参数说明见表1,具体论述参见文献[13].



₽	1	参	娈	븥	列	表
~		Y	×	ᆂ	~"	1×

<b>水</b> <sup>1</sup> 罗又里 <u>刘</u> 衣							
x:M 维输入向量	n <sub>j</sub> :激活次数	$ \boldsymbol{\Sigma}_j $ :节点超体积					
y:输出类标	W <sup>ab</sup> :映射矩阵	S <sub>max</sub> :警戒参数					
<b>μ</b> <sub>j</sub> :均值向量	w <sup>a</sup> :ARTa 节点 j	P <sub>min</sub> :谐振参数					
<b>∑</b> ;:协方差矩阵	$w_k^b$ :ARTb 节点 $k$	$N_a$ :节点个数					

贝叶斯 ARTMAP 神经网络能够根据数据的复杂程度自适应地调整网络结构,更适合于处理非平稳的遥感影像数据<sup>[14]</sup>.但是,如果获胜节点与输出类别不匹

配,便会过度调用匹配跟踪机制,产生大量的中间层节 点造成类别扩散现象.产生过多的类别节点会退化分 类精度<sup>[13]</sup>,同时过多的节点给模型调整带来巨大的负 担,不利于先验信息的迁移学习.因此,限制类别扩散 是保持 ARTMAP 网络模型稳定以及实现迁移学习的重 要前提.

## 3 基于改进贝叶斯 ARTMAP 的迁移学习

首先改进贝叶斯 ARTMAP 神经网络,在限制其类别扩散的基础上,利用离散增量形式实现期望最大化算法<sup>[15]</sup>,

2249

|

将历史数据中的先验信息迁移到目标模型中,建立基于改进贝叶斯 ARTMAP 神经网络的迁移学习算法.

## 3.1 改进谐振机制的贝叶斯 ARTMAP

贝叶斯 ARTMAP 节点竞争过程中基于最小错误率利 用后验概率<sup>[12]</sup>进行获胜节点的选择. 从获胜节点 $w_{J}^{a}$ 的测 试条件 $|\Sigma_{J}| \leq S_{max}$ 可以看出警戒测试防止某一类别过于强 大,贝叶斯推理已为衡量多个假设的置信度提供了定量的 方法,能够测试输入模式与类别模式间的联接强度,因此 调节警戒参数对提高预测准确度的作用较小. 在分布重叠 区出现错误预测时,调用匹配跟踪测试 $p(w_{K}^{b}|w_{J}^{a}) < P_{min}$ , 对警戒阈值进行调整 $S_{max} = |\Sigma_{J}| - \delta(0 \leq \delta < < |\Sigma_{J}|)$ ,仍无 法避免 ART 模块内部无监督的类别节点选择过程出现错 误结果,便会发生类别扩散.

本文从改进网络谐振机制的角度来简化模型:在 节点的竞争选择之前,考虑将第*j*个类别节点映射到输 出类别K:g(j) = K,其中 $g(\cdot)$ 为节点与类标间的映射 函数,当新样本x输入时,首先根据样本类标y,直接进 行节点谐振判断K = y,将与类标相对应的类别节点设 为局部激活节点 $\tilde{w}_j$ ,局部激活节点集合为 $\tilde{w} = \{w_j^a \mid j \in$  $g^{-1}(K)\}$ .之后从局部激活节点中进行获胜节点的选 择<sup>[12]</sup>.获胜节点能够确保训练过程中模块间的谐振,无 需匹配跟踪测试 $p(w_k^b \mid w_j^a) < P_{\min}$ 过程,之后直接对获 胜节点进行警戒测试 $|\Sigma_j| \leq S_{\max}$ ,重置未能通过警戒测 试的获胜节点,如果没有通过警戒测试的节点则创建 新的节点.根据顺序最大似然法对通过警戒测试的获

改进模块谐振判断后,能够确保获胜节点与样本 输出建立正确的关联,消除匹配跟踪带来的类别扩散, 从而降低节点个数,提高模型的精度,这是进一步迁移 学习的前提基础.

#### 3.2 改进型贝叶斯 ARTMAP 的迁移学习算法

将当前影像中获取样本数据按照类标进行划分, 设 $L_c = \{x | y = c\}$ 为属于类别c的全部样本构成的数据 集. 划分后的样本集分别对历史网络模型参数进行迁 移. 首先根据样本集 $L_c$ 所属类别按照贝叶斯 ARTMAP 神经网络改进的谐振机制进行节点激活: $K = c, \tilde{w} = \{w_j^a$  $l j \in g^{-1}(K)\}, g^{-1}(\cdot)$ 为类标与节点间的逆映射函数. 根据激活节点 $\tilde{w}_j$ 的描述参数 $\mu_j, \Sigma_j, n_j$ 计算出各激活节 点对于标记样本 $x \in L_c$ 的后验概率:

$$\varphi_{j} = p(\tilde{w}_{j} \mid \boldsymbol{x}) = \frac{p(\boldsymbol{x} \mid \boldsymbol{w}_{j})p(\boldsymbol{w}_{j})}{\sum_{i \in g^{-1}(K)} p(\boldsymbol{x} \mid \boldsymbol{w}_{i}^{a})p(\boldsymbol{w}_{i}^{a})}$$
$$= \frac{N(\boldsymbol{x} \mid \boldsymbol{\mu}_{j}, \boldsymbol{\Sigma}_{j})n_{j}}{\sum_{i \in g^{-1}(K)} N(\boldsymbol{x} \mid \boldsymbol{\mu}_{i}, \boldsymbol{\Sigma}_{i})n_{i}}$$
(1)

以及节点 w<sub>j</sub>在局部激活节点当中的先验概率:

$$\xi_j = p(\tilde{w}_j) = \frac{n_j}{\sum_{i \in \mathcal{V}} n_i}$$
(2)

以后验概率作为样本权值估算出每一个激活节点  $\tilde{w}_i$ 所对应的节点迁移参数 $\mu'_i$ 、 $\Sigma'_i$ 与迁移概率 $\xi'_i$ 、 $\varphi'_i$ :

$$\begin{cases} \boldsymbol{\mu}_{j}^{'} = \frac{1}{\sum_{\boldsymbol{x} \in L_{c}} p(\tilde{\boldsymbol{w}}_{j} + \boldsymbol{x})} \sum_{\boldsymbol{x} \in L_{c}} \boldsymbol{x} p(\tilde{\boldsymbol{w}}_{j} + \boldsymbol{x}) \\ \boldsymbol{\Sigma}_{j}^{'} = \frac{1}{\sum_{\boldsymbol{x} \in L_{c}} p(\tilde{\boldsymbol{w}}_{j} + \boldsymbol{x})} \sum_{\boldsymbol{x} \in L_{c}} p(\tilde{\boldsymbol{w}}_{j} + \boldsymbol{x}) (\boldsymbol{x} - \boldsymbol{\mu}_{j}^{'}) (\boldsymbol{x} - \boldsymbol{\mu}_{j}^{'})^{\mathrm{T}} \\ \boldsymbol{\xi}_{j}^{'} = \frac{1}{|L_{c}|} \sum_{\boldsymbol{x} \in L_{c}} p(\tilde{\boldsymbol{w}}_{j} + \boldsymbol{x}) \\ \boldsymbol{\varphi}_{j}^{'} = \frac{\boldsymbol{\xi}_{j}^{'} N(\boldsymbol{x} + \boldsymbol{\mu}_{j}^{'}, \boldsymbol{\Sigma}_{j}^{'})}{\sum_{i \in \boldsymbol{\xi}^{'} \mid N(\boldsymbol{x} + \boldsymbol{\mu}_{i}^{'}, \boldsymbol{\Sigma}_{i}^{'})} \end{cases}$$
(3)

式(3)中, $|L_{e}|$ 代表  $L_{e}$ 的样本个数. 之后按照离散增量 式期望最大化算法对激活节点的迁移参数 $\mu'_{j}$ 、 $\Sigma'_{j}$ 以及 迁移先验概率 $\xi'_{j}$ 进行更新,更新后的变量分别表示为  $\mu''_{i}$ 、 $\Sigma''_{i}$ 及 $\xi''_{i}$ :

$$\begin{cases} \xi_{j}^{'} = \xi_{j}^{'} + \frac{\varphi_{j}^{'} - \varphi_{j}}{|L_{c}|} \\ \mu_{j}^{'} = \mu_{j}^{'} + \frac{(\varphi_{j}^{'} - \varphi_{j})(\mathbf{x} - \mu_{j}^{'})}{|L_{c}|\xi_{j}^{*}} \\ (\mathbf{\Sigma}_{j}^{'})^{-1} = \frac{\xi_{j}^{'}}{\xi_{j}^{'}} \left[ (\mathbf{\Sigma}_{j}^{'})^{-1} - \frac{(\varphi_{j}^{'} - \varphi_{j})(\mathbf{\Sigma}_{j}^{'})^{-1}(\mathbf{x} - \mu_{j}^{'})(\mathbf{x} - \mu_{j}^{'})^{\mathrm{T}}(\mathbf{\Sigma}_{j}^{'})^{-1}}{|L_{c}|\xi_{j}^{*} + (\varphi_{j}^{'} - \varphi_{j})(\mathbf{x} - \mu_{j}^{'})^{\mathrm{T}}(\mathbf{\Sigma}_{j}^{'})^{-1}(\mathbf{x} - \mu_{j}^{'})} \right] \\ |\mathbf{\Sigma}_{j}^{c}| = \left(\frac{\xi_{j}^{'}}{\xi_{j}^{'}}\right)^{M} |\mathbf{\Sigma}_{j}^{'}| \left[ 1 + \frac{(\varphi_{j}^{'} - \varphi_{j})(\mathbf{x} - \mu_{j}^{'})^{\mathrm{T}}(\mathbf{\Sigma}_{j}^{'})^{-1}(\mathbf{x} - \mu_{j}^{'})}{|L_{c}|\xi_{j}^{*}} \right]$$

$$(4)$$

计算出与各节点相对应的迁移参数的对数似然值  $\ln P(L_e | \boldsymbol{\mu}^{''}, \boldsymbol{\Sigma}^{''}, \boldsymbol{\xi}^{''}) = \sum_{x \in L_e} \ln \{ \sum_{j \in g^{-1}(K)} \boldsymbol{\xi}_j^{''} N(\boldsymbol{x} + \boldsymbol{\mu}_j^{''}, \boldsymbol{\Sigma}_j^{''}) \}, 以$ 及各节点相对应的描述参数的对数似然值  $\ln P(L_e | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\xi}) = \sum_{x \in L_e} \ln \{ \sum_{j \in g^{-1}(K)} \boldsymbol{\xi}_j N(\boldsymbol{x} + \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \}, \text{式 中}, \boldsymbol{\mu}^{''} = [\boldsymbol{\mu}_{1,1}^{''}, \boldsymbol{\mu}_{2}^{''} \cdots]^{\mathrm{T}}, \boldsymbol{\mu} = [\boldsymbol{\mu}_{1,1}, \boldsymbol{\mu}_{2} \cdots]^{\mathrm{T}}, \boldsymbol{\Sigma}^{''} = [\boldsymbol{\Sigma}_{1,1}^{''}, \boldsymbol{\Sigma}_{2}^{''} \cdots]^{\mathrm{T}}, \boldsymbol{\Sigma}$   $= [\boldsymbol{\Sigma}_{1,1}, \boldsymbol{\Sigma}_{2} \cdots]^{\mathrm{T}}, \boldsymbol{\xi}^{''} = [\boldsymbol{\xi}_{1,1}^{''}, \boldsymbol{\xi}_{2}^{'''} \cdots]^{\mathrm{T}}, \boldsymbol{\xi}$   $= T_{1,2} + T_{1,2} + T_{2,2} + T_{2$ 

 $\ln P(L_{e}|\boldsymbol{\mu}^{"},\boldsymbol{\Sigma}^{"},\boldsymbol{\xi}^{"}) - \ln P(L_{e}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\xi})| \leq \varepsilon$ (5) 式中,  $\varepsilon$  为收敛阈值,  $|\cdot|$ 表示取绝对值. 当满足不等式 的收敛标准时,则停止对样本集  $L_{e}$ 的迁移训练; 否则将 激活节点的迁移参数传递给网络节点的描述参数: $\xi_{j} = \xi_{j}^{"};\boldsymbol{\mu}_{j} = \boldsymbol{\mu}_{j}^{"}; (\boldsymbol{\Sigma}_{j})^{-1} = (\boldsymbol{\Sigma}_{j}^{"})^{-1}; |\boldsymbol{\Sigma}_{j}| = |\boldsymbol{\Sigma}_{j}^{"}|.$ 利用新的 描述参数重复对样本集  $L_{e}$ 的迁移训练,直至收敛.

至此算法完成了网络的迁移学习过程. 网络的测试过程为基于贝叶斯判据<sup>[12]</sup>的类别归属确定.

## 4 仿真实验

此节首先改进贝叶斯 ARTMAP(IBAM) 在抑制类

第 9 期

别扩散以及提高分类精度方面的有效性,在此基础上证明基于改进贝叶斯 ARTMAP 的迁移学习算法(TL-IBAM)在利用先验信息进行遥感影像分类建模方面的实用性.

#### 4.1 改进贝叶斯 ARTMAP 分类实例

首先将使用标杆数据验证改进贝叶斯 ARTMAP(IBAM)的精度及稳定性,贝叶斯 ARTMAP(BAM)<sup>[12]</sup>、 模糊 ARTMAP(FAM)<sup>[16]</sup>、ART-EMAP(AEM)<sup>[17]</sup>作为比 较对象. 每组数据进行 10 次十折交叉检验,平均精度和 标准差见表 2.

数据集	IBAM		BAM		FAM		AEM	
	OA(%)	STD	OA(%)	STD	OA(%)	STD	OA(%)	STD
Vote	91.53	0.0416	91.47	0.0406	74.23	0.1532	73.12	0.1724
Tic-tac-toe	71.29	0.1402	69.69	0.0529	56.75	0.1009	63.04	0.0990
Lymph	81.86	0.1037	77.21	0.1063	63.14	0.1395	68.93	0.1364
Hepatitis	82.60	0.1072	80.87	0.0979	78.87	0.0976	80.47	0.1035
Breast-w	95.80	0.0234	95.75	0.0243	91.45	0.0682	92.59	0.0566
Balance-scale	90.02	0.0357	84.92	0.0507	62.84	0.0788	76.53	0.0570
Krkp	75.79	0.0208	63.11	0.0280	74.21	0.1080	67.24	0.1340

表2 UCI 数据仿真结果

从表 2 中可知,本文所提 IBAM 算法在实验数据上 取得最优结果,说明具有更高的分类精度,且通过比较 标准差可以看出 IBAM 算法的结果更加稳定,表现出更 突出的分类性能.

其次,验证 IBAM 的遥感数据分类能力,以及抑制 类别扩散的能力.利用大连 SPOT-5 遥感数据进行仿真 实验,影像分辨率为 2.5m. 在影像数据中选取 400 个数 据点通过目视解译赋予类标来构造样本集,进行 10 次 十折交叉检验来验证算法的精度.与 BAM<sup>[12]</sup>、FAM<sup>[16]</sup>、 AEM<sup>[17]</sup>、极限学习机(Extreme Learning Machine, ELM)、 决策树(Decision Tree, DT)以及 BP 神经网络进行比较. 其中 IBAM 与 BAM 的警戒参数均为  $S_{max} = 10^{-12}$ , FAM、 AEM、AMIC 的警戒参数分别为 0.96、0.90、0.75, BP 神 经网络的最大迭代次数为 1000.

#### 表3 各分类算法精度及节点个数

	IBAM(本文)	BAM	FAM	AEM	ELM	DT	BP
平均精度(%)	96.04	85.28	82.72	65.60	93.16	94.48	94.56
节点个数	9	22	53	42	50	-	10

从表3中可以看出,本文GIBAM 算法在精度方面 优于其余算法.与此同时,本文所改进的贝叶斯 ART-MAP 神经网络通过改变网络谐振机制,抑制匹配跟踪 带来的类别扩散,相比于同类 ARTMAPG 神经网络而言 节点个数更少,获得最佳的算法性能.

#### 4.2 迁移学习遥感影像分类实例

为验证所提 TL-IBAM 算法在遥感影像分类建模方面的实用性,首先利用 393 × 236 像素大小的 2003 年以及 2007 年大连 SPOT-5 遥感影像进行实验.两期影像空间分辨率均为 2.5m,包含蓝、绿、红、近红外四个波段.



图2 大连SPOT-5影像分类结果

以图 2(a) 作为历史影像,图 2(b) 作为当前解译的 目标影像.在图 2(a) 中获取了 1078 个有标记样本点. 从图 2(b) 中获取 234 个有标记样本点,不完备的训练 样本集无法对复杂的光谱特征建立有效的模型,因此 图 2(c) 中仅利用目标影像样本的分类结果并不能令人 满意,特别是对建筑的识别误分类较多.图 2(d) 和图 2 (e)分别给出了利用历史数据获得目标影像分类结果 以及混合历史数据与当前数据后获得的目标遥感影像 分类结果,可以看出由于两幅影像光谱特征的不一致 性,在缺乏历史数据的有效重用策略时,历史数据的引 入反而会降低分类的准确性.本文所提的迁移学习算 法 TL-IBAM 在少量有标记样本下辅以大量的历史数

据,获得 2007 年目标影像的最优分类结果(见图 2 (*f*)).同时给出分类精度的定量评价指标,图2(*b*)给出 精度评价参考点的位置及大小,表4给出了总体精度 以及 Kappa 系数.

通过表4以及图2的比较可以看出本文所提的迁 移学习遥感影像分类方法获得了最优的分类结果,不 仅分类准确度高,而且地物类型的完整性强.这主要是 因为本文所提迁移学习方法具有将历史模型中的知识 迁移到目标模型中的能力,充分利用了历史数据的有 效信息,降低了对当前影像中有标记样本的依赖,因此 获得更为准确与完整的分类结果.

其次,利用辽河口湿地1776×2177×7大小的TM 影像,验证TL-IBAM算法对大场景影像数据中的适用 性.本组试验中扩展了历史数据的采样范围,如图3(a) 所示,分别从三幅遥感影像不同区域采集735、1764、 1225个样本.以2009年该地区遥感影像为目标影像, 从中采集980个样本构成当前数据集,警戒参数设定为 S<sub>max</sub> = 10<sup>-5</sup>.首先以1995年影像样本作为历史数据进行 学习,之后在历史数据中依次加入1987年与2004年样 本数据,分别进行学习,比较不同历史数据下的分类结 果,遥感影像分类精度的定量指标结果如表5所示.

表 4 2007 年大连 SPOT-5 遥感影像分类精度

	当前数据 结果	历史数据 结果	混合数据 结果	迁移学习 结果
OA(%)	86.53	71.60	75.10	94.34
Kappa 系数	0. 8251	0. 6383	0. 6798	0. 9269

	当前数据结果	Ŀ	历史数据结界	艮	混合数据结果			迁移学习结果		
数据来源	2009	1995	+ 1987	+ 2004	1995	+ 1987	+ 2004	1995	+ 1987	+ 2004
节点个数	52	49	57	62	90	103	110	36	38	39
OA(%)	88.11	70.37	74.54	77.33	77.65	79.52	80.11	89.77	91.26	92.70
Kappa 系数	0.8453	0.6122	0.6657	0.7057	0.7139	0.7298	0.7427	0.8687	0.8932	0.9144

#### 表5 辽河口湿地 TM 遥感影像分类精度

图 3 给出了全部 3724 个样本作为历史数据,不同 样本利用策略下的分类结果.初始模型建立了 62 个节 点,在迁移过程中不断对节点参数进行更新,其中 | **Σ**<sub>j</sub> | >10<sup>2</sup>的节点发生退化,最后获得 39 个迁移节点.利用 对称 Kullback-Leibler (KL)散度<sup>[18]</sup> 衡量各节点的迁移 强度,水体、沼泽、水田、裸地、建筑类别中迁移强度最

大节点的 KL 散度分别为 68. 5999,227. 7071,48. 8644, 109. 0713,30. 5086,其中沼泽类迁移程度最大,建筑类 的迁移程度最低. TL-IBAM 算法建立了历史数据的重 用策略,较大程度上提高了大场景目标影像的分类精 度,其中增加历史数据的规模有助于建立完备的先验 信息,从而提高迁移学习的精度.



14		沿伴	小田	保地
	图3	辽河	口湿地TM影像	分类结果

表6	计算时间比较
----	--------

样本数量(个)	10 <sup>2</sup>	10 <sup>3</sup>	10 <sup>4</sup>	10 <sup>5</sup>
ELM 运行时间(s)	0.003878	0.027158	0.316916	13.909087
TL-IBAM 运行时间(s)	0.018322	0.062932	0. 541773	5.375229

TL-IBAM 算法分类过程中运行时间与样本规模 n 及节点个数 N 线性相关,以辽河口湿地影像数据为基础,比较不同数据量下的运行时间(见表 6),其中节点数为 100.实验环境为 Intel(R) Core(TM) i5 760 四核处理器,32 位操作系统,3.49GB 内存及 2.80GHz 主频.可以看出所提算法在数据量较大时仍有较快的计算速

度,就其分类效率上来讲,所提算法同样适用于大场景 遥感影像的分类.

### 5 结论

建筑

围绕利用有限样本数据建立遥感影像分类模型的 目标,本文在历史数据的辅助下,提出了一种基于改进 贝叶斯 ARTMAP 的迁移学习遥感影像分类算法,其中 改进的贝叶斯 ARTMAP 神经网络有效抑制类别扩散现 象的产生,为实现 ARTMAP 的迁移学习提供了有利条 件,而离散增量期望最大化算法将历史数据中的先验 知识作为迁移计算的起点,在目标影像的样本数据指

第 9 期

导下,调整贝叶斯 ARTMAP 节点描述参数,实现了目标 影像的迁移学习. 仿真实验证明本文所提算法有效利 用了模型先验信息,避免了遥感影像历史样本的浪费, 降低了对目标影像标记样本的依赖性,获得了较好的 分类性能.

## 参考文献

[1] 唐晓亮,韩敏. 基于改进型模糊 ARTMAP 网络的 CBR 遥感图像分类系统研究[J]. 中国图象图形学报,2007, 12(6):1033-1040.
TANG Xiao-liang,HAN Min. A study of CBR system for remote sensing classification based on modified fuzzy AR-

TMAP neural network [J]. Journal of Image and Graphics,2007,12(6):1033 – 1040. (in Chinese)

- [2] DAUMÉ H, MARCU D. Domain adaptation for statistical classifiers [J]. Journal of Artificial Intelligence Research, 2011,26:101 – 126.
- [3] PAN S J, YANG Q. A survey on transfer learning [J].
   IEEE Transactions on Knowledge and Data Engineering, 2010,22(10):1345 - 1359.
- [4] 王皓,高阳,陈兴国.强化学习中的迁移:方法和进展
  [J].电子学报,2008,36(12):39-43.
  WANG Hao, CHEN Xing-guo. Transfer of reinforcement learning: the state of the art[J]. Acta Electronica Sinica, 2008,36(12):39-43. (in Chinese)
- [5] 于重重,田蕊,谭励,等. 非平衡样本分类的集成迁移学 习算法 [J]. 电子学报,2012,40(7):1358-1363.
  YU Chong-chong, TIAN Rui, TAN Li and Xu-yan. Integrated transfer learning algorithmic for unbalanced samples classification [J]. Acta Electronica Sinica, 2012,40(7): 1358-1363. (in Chinese)
- [6] PAN S J,TSANG I W, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22(2):199 – 210.
- [7] 倪彤光,王士同,应文豪,邓赵红. 迁移组概率学习机
  [J]. 电子学报,2013,41(11):2207-2215.
  NI Tong-guang, WANG Shi-tong, YING Wen-hao, DENG Zhao-hong. Transfer group probabilities based learning machine [J]. Acta Electronica Sinica,2013,41(11):2207-2215. (in Chinese)
- [8] BRUZZONE L, MARCONCINI M. Toward the automatic updating of land-cover maps by a domain-adaptation SVM classifier and a circular validation strategy [J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47 (4):1108-1122.
- [9] ZHANG Y S, ZHENG X W, et al. Semi-supervised manifold learning based multigraph fusion for high-resolution remote sensing image classification [J]. IEEE Geoscience and Remote Sensing Letters, 2014, (11):464-468.

[10] 吴田军,骆剑承,夏列钢,等. 迁移学习支持下的遥感影像对象级分类样本自动选择方法[J]. 测绘学报,2014,43(9):908-916.

WU Tianjun,LUO Jiancheng,XIA Liegang, et al. An automatic sample collection method for object-oriented classification of remotely sensed imageries based on transfer learning [J]. Acta Geodaetica et Cartographica Sinica, 2014,43(9):908-916. (in Chinese)

- [11] LIU Y, LI X. Domain adaptation for land use classification: A spatio-temporal knowledge reusing method [J].
   ISPRS Journal of Photogrammetry and Remote Sensing, 2014,98;133 – 144.
- [12] VIGDOR B, LERNER B. The Bayesian ARTMAP [J]. IEEE Transactions on Neural Networks, 2007, 18 (6): 1628 – 1644.
- [13] TANG X, HAN M. Semi-supervised Bayesian ARTMAP [J]. Applied Intelligence, 2010, 33(3):302 - 317.
- [14] LEIVA-MURILLO J M, GÓMEZ-CHOVA L, et al. Multitask remote sensing data classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51 (1): 151 – 161.
- [15] NG S K, MCLACHLAN G J. On the choice of the number of blocks with the incremental EM algorithm for the fitting of normal mixtures [J]. Statistics and Computing, 2003,13(1):45-55.
- [16] CARPENTER G A, GROSSBERG S, et al. Fuzzy ART-MAP: A neural network architecture for incremental supervised learning of analog multidimensional maps [J]. IEEE Transactions on Neural Networks, 1992, 3(5):698-713.
- [17] CARPENTER G A, ROSS W D. ART-EMAP: A neural network architecture for object recognition by evidence accumulation [J]. IEEE Transactions on Neural Networks, 1995, 6(4):805 - 818.
- [18] COETZEE F M. Correcting the Kullback-Leibler distance for feature selection [J]. Pattern Recognition Letters, 2005,26(11):1675-1683.

#### 作者简介



**韩 敏(通信作者)** 女,1959 年 8 月出生 于吉林延吉.现为大连理工大学教授、博士生导 师、模糊信息处理与机器智能研究所副所长.从 事 3S 系统信息融合、神经网络理论及其应用、复 杂系统建模与预测的研究.

E-mail:minhan@dlut.edu.cn

**杨** 雪 女,1990年2月出生于河北秦皇岛.现为大连理工大学 电子信息与电气工程学部硕士研究生.主要研究方向为遥感影像信 息提取.