

SVM 在多源遥感图像分类中的应用研究

何灵敏¹⁾ 沈掌泉¹⁾ 孔繁胜¹⁾ 刘震科²⁾

¹⁾(浙江大学人工智能研究所, 杭州 310027) ²⁾(中国石油化工股份公司中南分公司研究院, 长沙 410007)

摘要 在利用遥感图像进行土地利用/覆盖分类过程中,可采用以下两种途径来提高分类精度:一是通过增加有利于分类的数据源,引入地理辅助数据和归一化植被指数(NDVI)来进行多源信息融合;二是选择更好的分类方法,例如支持向量机(SVM)学习方法,由于该方法克服了最大似然法和神经网络的弱点,非常适合高维、复杂的小样本多源数据的分类。为了提高多源遥感图像分类的精度,还研究了支持向量机在遥感图像分类中模型的选择,包括多类模型和核函数的选择。分类结果表明,支持向量机比传统的分类方法具有更高的精度,尤其是基于径向基核函数和一对多类方法的支持向量机模型更适合多源遥感图像分类,因此,基于支持向量机的多源土地利用/覆盖分类能大大提高分类精度。

关键词 支持向量机 分类 多源数据 模型选择

中图法分类号: TP751.1 文献标识码: A 文章编号: 1006-8961(2007)04-0648-07

Study on Multi-source Remote Sensing Images Classification with SVM

HE Ling-min¹⁾, SHEN Zhang-quan¹⁾, KONG Fan-sheng¹⁾, LIU Zhen-ke²⁾

¹⁾(Artificial Intelligence Institute, Zhejiang University, Hangzhou 310027)

²⁾(Central-South Petroleum Bureau, China Petroleum and Chemical Corporation, Changsha 410007)

Abstract There are two ways to improve the performance of land cover classification with remote sensing images. The first way is to apply new data source including GIS data and normalized difference vegetation index(NDVI) to multi-source information fusion. The second one is to use methods with higher accuracy. Support vector machines(SVM) overcome the defects of maximum-likelihood and neural networks classifiers. SVMs are suitable to process complex data of high dimension and small number of training data. In this paper, selection of SVM models including kernel functions and multi-class methods is studied in order to improve the accuracy of multi-source remote sensing images classification. Experimental results show that the SVMs have higher accuracy than other traditional classifiers for the classification of multi-source remote sensing data. The SVM with a RBF kernel function and One-against-one multi-class method is the best classifier in this study. SVM methods could greatly improve the multi-source land cover classification.

Keywords support vector machines(SVM), classification, multi-source data, model selection

1 引言

土地在人与环境的关系中占有非常重要的作用,特别是精确实时的土地利用信息可以应用到土地资源规划、环境变化研究等等。近年来,遥感在土地利用/覆盖分析中得到许多成功的应用,然而,许

多问题如分类精度等问题仍然存在^[1]。

由于一些地物具有非常相似的光谱特性,因此使用单一来源的遥感数据可能很难识别。由于地表地物的分布与地貌密切相关,而地貌又是由地形决定的,因此地理辅助数据能够帮助识别这些地物^[2-4],本文利用增强型专题绘图仪(enhanced thematic mapper, ETM)图像数据,结合地理辅助数

基金项目:国家自然科学基金资助项目(40201021)

收稿日期:2005-12-19;改回日期:2006-04-19

第一作者简介:何灵敏(1974~),男,1999年毕业于浙江大学地球科学系获硕士学位,现为该校计算机科学与技术专业博士研究生。研究方向为人工智能、模式识别和遥感图像处理等。E-mail:lemon_hill@163.com

据、植被指数来进行信息融合,由于大大丰富了遥感图像的物地信息,从而为利用遥感图像进行土地利用/覆盖分类提供了信息保证。

传统学习方法的基础是统计学,它是采用了所谓经验风险最小化 (empirical risk minimization, ERM) 准则,但其前提是要有足够多的样本,传统方法不仅当样本数目有限时难以取得理想的效果,而且这些学习方法本身还有各自的缺点,如最大似然法,它是假设待分类的输入数据服从高斯分布,但是有些数据,如地理辅助数据却不符合这个要求。虽然神经网络和决策树分类器的分类精度一般高于最大似然法^[5-7],然而它们都不是稳定的分类器。

支持向量机 (support vector machines, SVM) 是 Vapnik 等人在 1995 提出的一种学习方法^[8,9],由于它是建立在统计学习理论 (statistical learning theory, SLT) 的 VC (vapnik-chervonenkis) 维理论和结构风险最小化 (structural risk minimization, SRM) 准则基础上的,因此能较好地解决小样本、非线性、高维数据和局部极小等实际问题,因其易用、稳定和具有相对较高的精度而得到广泛的应用。

SVM 在遥感图像分类中的应用,国内外专家做了不少研究^[10-15],使用的图像数据包括陆地卫星 ETM 图像、SPOT 图像、超光谱图像和 SAR 图像等,而对多源遥感图像分类的应用则相对较少,其研究的重点大多集中在 SVM 和传统分类方法的分类效果的比较上,而对算法的模型选择和计算复杂度等则关注较少。由于 SVM 是一种通用的两类学习方法,因此在具体的应用中如何选择合适的模型,包括多类方法和核函数的选择是值得研究的问题,另外,分类算法的计算时间也是一个重要的考虑因素。本文利用多类 SVM 来处理多源数据,并通过比较各种多类方法和使用不同核函数情况下的分类精度和计算时间,提出了一个最佳的模型。

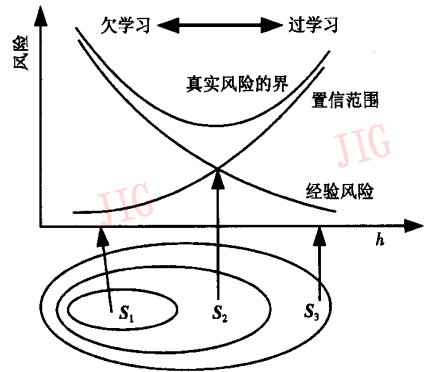
2 SVM

2.1 结构风险最小化准则

传统的学习方法采用了所谓经验风险最小化准则,即通过求经验风险最小值来近似代替期望风险最小值,事实上,由于这种代替并没有经过充分的理论论证,而只是直观上合理的想当然的做法,所以经验风险最小化准则在样本有限时是不合理的。

Vapnik 等人提出了一种在有限样本情况下的

较完善的机器学习理论体系——统计学习理论,它被认为是目前针对小样本统计估计和预测学习的最佳理论。这种统计学习理论提出了一种新的策略,即先把指示函数集^[8]构造为一个函数子集序列,并使各个子集按照 VC 维^[16]的大小排列;然后在每个子集中寻找最小经验风险,在子集间折中考虑经验风险和置信范围,以取得实际风险的最小,这种思想被称为“结构风险最小化准则”(如图 1 所示)。



函数子集: $S_1 \subset S_2 \subset S_3$, VC 维: $h_1 < h_2 < h_3$

图 1 结构风险最小化示意图

Fig. 1 Structure risk minimization

基于 SVM 的分类方法就是这种思想的具体实现,即先通过设计函数集的某种结构,使每个子集中都能取得最小的经验风险,然后通过选取适当的子集使置信范围最小,这样在这个子集中使经验风险最小的函数就是用于分类的最优判别函数。

2.2 两类 SVM 分类器

两类 SVM 的工作机理可概括为:通过寻找一个分类超平面来使得训练样本中的两类样本点能被分开,并且距离该平面尽可能地远(如图 2 所示), H 是最优分类面, H_1 和 H_2 之间的距离 m 称为分类间隔 (margin),最优分类面就是要求分类面不仅能将两类正确分开,而且能使分类间隔最大, H_1, H_2 上的

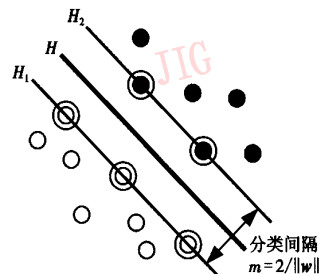


图 2 线性可分条件下的最优分类面

Fig. 2 Optimum separation plane

训练样本称作支持向量,而对线性不可分的问题,则可通过核函数将低维输入空间的数据映射到高维空间,以便将原低维空间的线性不可分问题转化为高维空间上的线性可分问题。

给定一个样本集 $T = \{(x_1; y_1), (x_2; y_2), \dots, (x_l; y_l)\}$, 其中, $x_i \in \mathbf{R}^n$ 是输入向量, $y_i \in \{1; -1\}$ 是两类问题的标签, SVM 方法就是为了寻找一个两类之间的最优分类面 $w \cdot \Phi(x) + b = 0$, 其中, w 是此分类面权重向量, $b \in \mathbf{R}$ 是偏置, Φ 表示把样本集 T 映射到一个高维的特征空间。

SVM 寻找最优分类面的过程等价于求解一个凸二次优化问题(式(1)):

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s. t. } & y_i [w \cdot \Phi(x_i) + b] \geq 1 - \xi_i \\ & \xi_i \geq 0 \quad i = 1, 2, \dots, l \end{aligned} \quad (1)$$

其中, C 是惩罚系数, 用来控制误差 ξ (松弛变量) 边界的平衡。使用拉格朗日优化方法和对偶原理来解问题(式(1)), 即得到式(2):

$$\begin{aligned} \min \quad & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{j=1}^l \alpha_j \\ \text{s. t. } & \sum_{i=1}^l y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C \quad i = 1, 2, \dots, l \end{aligned} \quad (2)$$

其中, $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ 是核函数, 可以用 Chunking 算法^[17], Osuna 算法^[18], SMO 算法^[19] 等来解式(2), 就可得到以下具有最大间隔的最优分类面:

$$\sum_{i=1}^l \alpha_i y_i K(x, x_i) + b = 0 \quad (3)$$

最后得到的决策函数是

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i K(x, x_i) + b \right) \quad (4)$$

SVM 方法有以下几个主要优点:(1)专门针对有限样本情况,其目标是得到现有信息下的最优解,而不仅仅是样本数趋于无穷大时的最优值;(2)从理论上说,算法得到的将是全局最优解,这就解决了在神经网络方法中无法避免的局部极值问题;(3)算法将实际问题通过非线性变换转换到高维特征空间,在高维空间中通过构造线性判别函数来替换原空间中的非线性判别函数,这就巧妙地解决了维数问题,其算法复杂度与样本维数无关。

2.3 核函数

SVM 需要利用核函数 $K(x_i, x_j) = \Phi(x_i) \cdot$

$\Phi(x_j)$ 来实现从原来空间到特征空间的映射,只要满足 Mercer 条件的对称函数即可作为核函数,常用的核函数有:

(1) 多项式核函数:

$$K(x_i, x_j) = (\gamma x_i \cdot x_j + b)^d \quad d = 1, 2, \dots \quad (5)$$

(2) 径向基(radius basis function, RBF)核函数:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (6)$$

(3) Sigmoid 核函数:

$$K(x_i, x_j) = \tanh(\gamma(x_i \cdot x_j) + b) \quad (7)$$

其中, γ, b, d 是核函数参数,但究竟选择哪个核函数至今没有理论上的指导,只能凭经验进行选择。

2.4 多类 SVM

SVM 本身是用来解决两类问题的,当前有两大类多类方法:一是集成许多两类问题,已经有比较成熟的算法,如一对一方法(one-against-one, OAO)、一对多方法(one-against-all, OAA)以及有向无环图(directed-acyclic-graph, DAG)支持向量机等;另一种是在优化公式中直接考虑多类问题,其缺点是算法选择的目标函数过于复杂而难以实现,而且计算量很大,所以本文使用前者来进行试验。

一对多方法是使用最早和使用最广泛的多类策略^[20]。对于一个 $N(N > 2)$ 类问题,由 N 个两类 SVM 分类器组成,每个两类分类器用来区分一个类和剩余的类,最后得到的决策结果是具有最大决策值的类获胜。由于是一类对其余的类,因此很可能发生显著的样本数不均衡现象,从而导致出现一定的误差。

一对一方法包括 $N(N-1)/2$ 个两类 SVM 分类器。每一个训练好的分类器用来区分两个类别^[21],所有的 $N(N-1)/2$ 个两类分类器采用投票法来进行决策。

有向无环图 SVM 是 Platt 等人提出^[22],其训练过程和 OAO 方法一样,一共训练出 $N(N-1)/2$ 个两类 SVM 分类器。不同的是,在测试阶段, DAG 方法使用一个从根开始的二分有向无环图,它有 $N(N-1)/2$ 个内部节点和 N 个叶子节点。例如给定一个测试样本,从根节点开始,采用二分决策,即根据其输出值转到左节点或者右节点,一直到叶子节点,最后由叶子节点来确定最后的归属。

3 分类

多源信息融合和分类要经过以下的步骤:(1)对

各种来源的数据进行预处理,包括把地理辅助数据由矢量转为栅格,以及对遥感图像进行各种校正等;(2)对不同来源的数据进行精确地配准;(3)使用合适的分类器进行分类;(4)进行评价和进一步处理。

3.1 试验区 and 数据

本试验区位于浙江省龙游县,遥感数据采用的是 Landsat ETM 图像,试验区大小为 512×512 pixels (见图 3(a))。6 个波段的 TM 数据分别为:TM-1 到 TM-5 及 TM-7,图像生成于 2002 年 3 月 7 日。增加的地理辅助数据包括数字高程模型(digital elevation model, DEM)、坡向和坡度 3 个数据通道。

由于植被叶绿素在波长 $0.69 \mu\text{m}$ 处具有强吸收特征,因此通过红外与近红外波段反射率的比值或线性组合可实现对植被信息的表达,归一化植被指数(normalized difference vegetation index, NDVI)被广泛应用于植被覆盖的定量研究^[23],归一化植被指数使得具有植被覆盖的地表和没有植被覆盖的地表更能区分开,归一化植被指数的计算公式为

$$NDVI = \frac{TM4 - TM3}{TM4 + TM3} \quad (8)$$

另外,采集数据时根据实地调查获得的土地利用图来指导取样。根据对土地利用图和 TM 图像的分析,把试验区的地物分为 8 类,经采样,共获得两组样本集,每组样本集包括训练样本和测试样本(参见表 1)。

表 1 研究区各地物类别的样本

Tab.1 Samples of land cover categories in the study site

类别 标号	地物类别	样本集 1		样本集 2	
		训练 样本数	测试 样本数	训练 样本数	测试 样本数
C1	水体	60	60	60	60
C2	建筑物	60	60	60	66
C3	水田	60	60	80	88
C4	裸露土壤	60	60	60	61
C5	果园	60	60	62	62
C6	茶园和桑园	60	60	68	60
C7	林地	60	60	80	81
C8	其他	60	60	77	65
总数		480	480	547	543

3.2 分类过程

训练前先把原始数据都归一化到 $(-1, +1)$,接着利用训练样本训练各种类型的 SVM 分类器,训

练时,先利用交叉验证来确定 SVM 的 C 和 γ 等参数;然后使用测试本来验证所训练的分类器,并根据分类精度和计算时间选择合适的模型;最后用所选择的 SVM 分类器对整个试验区进行分类。为了比较 SVM 分类器的分类效果,因此,本文也利用最大似然法(maximum likelihood classification, MLC)、神经网络分类法(BP neural network, BPNN)进行分类。

试验时运行的软硬件环境:CPU 是 AMD 公司的 Athlon™ XP 1800+, 512M 内存,操作系统是 Microsoft Windows XP Professional SP2。

3.3 试验结果及分析

为验证分类效果,分别从分类精度和计算时间两方面对不同的分类器模型进行分析比较,试验结果参见表 2 和表 3,表中 C 是惩罚参数, b, d, γ 是核函数参数。

从分类精度方面来看,首先,考虑不同的分类方法,从表 2 可以看出,绝大多数 SVM 分类器的总体精度比最大似然法和神经网络分类法高出许多,这说明 SVM 是相对稳定和高效的分类器,具有很好的泛化能力;其次,考虑 SVM 的各个不同模型,除去样本集 1 的基于 Sigmoid 核函数和 OAA 多类方法的模型(精度是 73.96%),其他的各种 SVM 模型分类精度相差不大,基本上都在 76%~78%(样本集 1)和 79%~82%(样本集 2)之间。相比较而言,径向基核函数的分类精度略高于多项式和 Sigmoid 核函数的分类精度,分别在 77% 和 81% 以上,且其各种多类方法分类精度的稳定性也比其他核函数略好。比较 3 种多类方法,OAA 多类方法的分类精度一般高于 OAA 方法和 DAG 方法,尤其是基于径向基的 OAO 分类器,其分类精度是所有 SVM 分类模型中最高的。总体上说,OAA 方法的稳定性不如 OAO 方法和 DAG 方法,其原因是 OAA 方法中各个两类分类器的数据不均衡的缘故。所以从精度和算法稳定性方面来说,径向基核函数和 OAO 方法是最好的选择。

从计算时间上来考虑,因为遥感数据往往是海量数据,所以分类算法的计算时间很重要。表 3 给出了各种 SVM 模型的支持向量数目和训练、测试时间,其最后一列是对整个 512×512 pixels 大小的试验区进行分类所用的时间(每种方法均试验 3 次,取其平均值)。从表 3 可知,对比各种核函数,Sigmoid 核函数的训练时间少于多项式和径向基核

表 2 多源数据的测试精度

Tab. 2 Test accuracy in percentage for the classification methods applied to multi-source data

分类方法		样本集 1			样本集 2		
核函数	多类方法	参数	Kappa 系数 (%)	总体精度 (%)	参数	Kappa 系数 (%)	总体精度 (%)
	MLC		64.05	68.54		69.44	73.53
	BPNN	3 层	71.23	74.87	3 层	72.89	77.35
	OAO	$C=2^5, \gamma=2^{-4}, b=0.5$	74.29	77.50	$C=2^6, \gamma=2^{-4}, b=0.5$	78.95	81.37
Sigmoid	OAA	$C=2^4, \gamma=2^{-4}, b=0$	70.24	73.96	$C=2^8, \gamma=2^{-4}, b=0.5$	78.14	80.15
	DAG	$C=2^5, \gamma=2^{-4}, b=0.5$	74.29	77.50	$C=2^6, \gamma=2^{-4}, b=0.5$	77.38	79.62
	OAO	$C=2^8, \gamma=2^{-3}, b=0.5, d=4$	74.52	77.71	$C=2^9, \gamma=2^{-2}, b=0.5, d=3$	78.95	81.37
多项式	OAA	$C=2^7, \gamma=2^{-3}, b=0, d=3$	72.62	76.04	$C=2^7, \gamma=2^{-2}, b=0, d=4$	78.14	80.15
	DAG	$C=2^8, \gamma=2^{-3}, b=0.5, d=4$	72.86	76.25	$C=2^9, \gamma=2^{-2}, b=0.5, d=3$	78.95	81.37
	OAO	$C=2^7, \gamma=2^{-3}$	75.24	78.33	$C=2^9, \gamma=2^{-5}$	79.31	82.14
径向基	OAA	$C=2^5, \gamma=2^{-3}$	74.05	77.29	$C=2^9, \gamma=2^{-6}$	78.56	81.77
	DAG	$C=2^7, \gamma=2^{-4}$	73.81	77.08	$C=2^9, \gamma=2^{-5}$	78.46	81.77

表 3 各种 SVM 分类器训练测试所需的时间

Tab. 3 Training and test time with different SVM classifiers applied to multi-source data

SVM 模型		样本集 1				样本集 2			
核函数	多类方法	支持向量个数	训练时间 (ms)	测试样本分类 (ms)	整个试验区分类 (s)	支持向量个数	训练时间 (ms)	测试样本分类 (ms)	整个试验区分类 (s)
	OAO	245	90	100	53	251	100	110	55
Sigmoid	OAA	398	431	120	66	413	310	140	67
	DAG	245	90	80	46	251	100	100	54
多项式	OAO	199	170	80	45	225	220	85	49
	OAA	275	360	90	50	274	900	95	55
	DAG	199	170	70	43	225	220	80	47
径向基	OAO	198	180	70	39	207	251	85	39
	OAA	252	411	80	40	253	912	90	41
	DAG	202	120	70	38	207	250	80	39

函数,然而对于更重要的测试时间,却多于多项式和径向基函数,而径向基和多项式相比则相差不大。比较各种多类方法,OAA 方法的训练和测试时间都是最多的,尤其训练时间是 OAO 方法和 DAG 方法的 2~4 倍,DAG 方法的测试时间最少,OAO 方法居中,但和 DAG 方法相差不大。分类器的测试时间主要与支持向量的数目有关,OAO 和 DAG 方法虽然所含的两类分类器数目要比 OAA 方法多得多,但是由于每个两类分类器的支持向量数目少,总的支持向量数目比 OAA 方法少,因此总体分类时间反而比 OAA 方法少。

从上面的分析可以得出以下认识:在核函数的选

择上,一方面径向基核函数无论从分类精度和分类时间上来看,都是最合适的选择;另一方面,径向基核函数的参数最少,只有一个 γ ,这样在训练过程中要确定的 SVM 参数就只有 C 和 γ ,比 Sigmoid 核函数(2 个核参数)和多项式核函数(3 个核参数)更易使用。而 OAO 多类方法虽然在分类时间上略大于 DAG 方法,但是其在精度上要略好于 DAG 方法,综合考虑,OAO 方法应该是首选,因此,基于径向基核函数的 OAO 多类支持向量机分类器是最好的选择。

为了分析各地物类别的分类精度,表 4 给出了测试样本的误差矩阵(基于径向基核函数和 OAO 多类方法),从表 4 可以看出,果园(C5)、茶园和桑

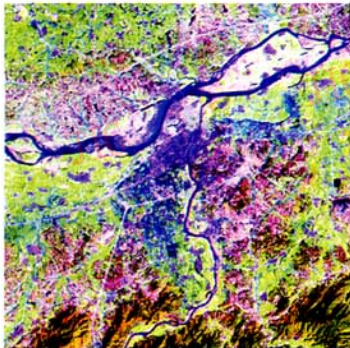
表 4 基于径向基核函数和 OAO 多类方法的 SVM 分类的误差矩阵

Tab.4 Error matrix of classification based on SVM with RBF kernel and OAO multi-class method applied to multi-source data

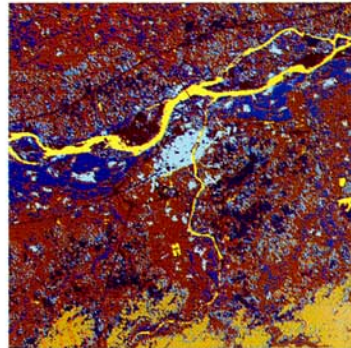
实际	分类								总计	
	C1	C2	C3	C4	C5	C6	C7	C8		
样本集 1	C1	60	0	0	0	0	1	0	0	61
	C2	0	48	0	0	2	0	1	0	51
	C3	0	0	54	0	1	1	0	0	56
	C4	0	1	1	60	0	0	0	0	62
	C5	0	0	1	0	34	28	15	1	79
	C6	0	3	4	0	11	26	6	2	52
	C7	0	8	0	0	9	3	38	1	59
	C8	0	0	0	0	3	1	0	56	60
	总计	60	60	60	60	60	60	60	60	480
样本集 2	C1	60	0	0	0	0	1	0	0	61
	C2	0	59	0	0	0	0	1	0	60
	C3	0	0	81	1	1	2	0	0	85
	C4	0	1	2	59	0	5	0	0	67
	C5	0	0	2	0	39	25	12	0	78
	C6	0	2	2	1	12	27	8	1	52
	C7	0	4	1	0	9	0	59	2	76
	C8	0	0	0	0	1	0	1	62	64
	总计	60	66	88	61	62	60	81	65	543

园(C6)、林地(C7)的分类正确率相对较低,精度分别在 60%、45% 和 65% 左右,采用其他类型的分类器(如 BPNN、MLC)精度更低。通过分析这三类地物在各通道(特征)的像元灰度值发现,它们有如下特征:(1)是它们在 TM 各波段的灰度值非常接近,即光谱特征非常相似,因为这些地物的地表大部分都是绿色木本植物;(2)在 DEM、坡向和坡度这三个数据通道,以上三类地物的灰度值差异性相对大些,即这些地物的分布和地形有密切关系。另外还对单源数据(仅 TM 图像数据)和多源数据的分类结果进行了比较,试验结果显示,多源数据对这三类地物样本的分类结果比单源数据要高 4% 左右,这就是为什么增加数据源能提高分类精度的原因。而要进一步提高分类精度,最有效的方法是采用更高分辨率的遥感图像。

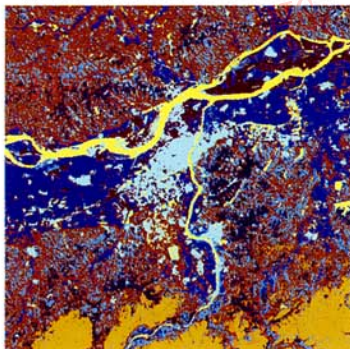
图 3(b)、图 3(c)、图 3(d)是整个试验区的分类结果图像(经伪彩色处理),参考实际的土地利用信息发现,基于 SVM 的多源遥感图像的分类结果(图 3(c)、图 3(d))更加合理。



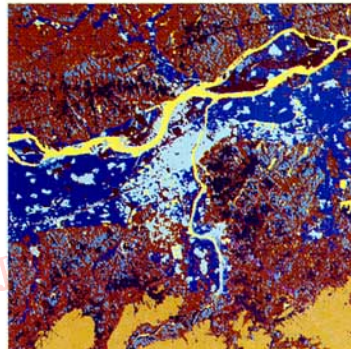
(a) 试验区的遥感图像(TM543 真彩色合成)



(b) 最大似然分类结果



(c) SVM 分类结果
(训练样本是样本集 1,RBF 核,OAO)



(d) SVM 分类结果
(训练样本是样本集 2,RBF 核,OAO)

图 3 试验区彩色图像和分类结果((b),(c),(d)是伪彩色图像)

Fig. 3 The classification results of different classifiers

4 结 论

本文利用支持向量机对多源遥感图像进行了分类研究,并从分类精度、计算时间和算法稳定性等角度比较了各种支持向量机模型,其中模型的选择包括核函数和多类方法的选择。试验结果表明,总体上支持向量机具有较高的分类精度,基于径向基核函数和一对一多类方法的支持向量机模型是最合适的模型;同时实验结果也表明,遥感数据加上其他辅助数据能提高分类精度。总之,基于多类支持向量机的多源遥感土地利用/覆盖分类能取得较好的分类效果。

参考文献 (References)

- 1 Foody G M. Status of land cover classification accuracy assessment [J]. *Remote Sensing of Environment*, 2002, **80**(1): 185 ~ 201.
- 2 Eiumnroh A, Shresta R P. Can DEM enhance the digital image classification [A]. In: *Proceedings of Asian Conference on Remote Sensing* [C], Kuala Lumpur, Malaysia, 1997.
- 3 Bektas F. *Remote Sensing and Geographic Information Integration: A Case Study; Bozcaada & Gokceada Island* [D]. Msc Thesis, Institution of Science and Technology, Istanbul Technical University, Turkey, 2003.
- 4 Harris Paul M, Ventura Stephen J. The integration of geographic data with remotely sensed imagery to improve classification in an urban area [J]. *Photogrammetric Engineering and Remote Sensing*, 1995, **61**(8): 993 ~ 998.
- 5 Kanellopoulos I, Wilkinson G G. Strategies and best practice for neural network image classification [J]. *International Journal of Remote Sensing*, 1997, **18**(4): 711 ~ 725.
- 6 Liu C M, Zhang L J, Davis C J, *et al.* Comparison of neural networks and statistical methods in classification of ecological habitats using FIA data [J]. *Forest Science*, 2003, **49**(4): 619 ~ 631.
- 7 Friedl M A, Brodley C E. Decision tree classification of land cover from remotely sensed data [J]. *Remote Sensing of Environment*, 1997, **61**(3): 399 ~ 409.
- 8 Vapnik V N. *The Nature of Statistical Learning Theory* [M]. New York: Springer-Verlag, 1995.
- 9 Zhang Xue-gong. Introduction to statistical learning theory and support vector machines [J]. *Acta Automatica Sinica*, 2000, **26**(1): 32 ~ 42. [张学工. 关于统计学习理论与支持向量机 [J]. *自动化学报*, 2000, **26**(1): 32 ~ 42.]
- 10 Roli F, Fumera G. Support vector machines for remote-sensing image classification [A]. In: *Proceedings of EOS/SPIE International Society for Optical Engineering* [C], Barcelona, Spanish, 2001, **4170**: 160 ~ 166.
- 11 Luo Jian-cheng, Zhou Cheng-hu, Leung Yee, *et al.* Support vector machine for spatial feature extraction and classification of remotely sensed imagery [J]. *Journal of Remote Sensing*, 2002, **6**(1): 50 ~ 55. [骆剑承, 周成虎, 梁怡等. 支撑向量机及其遥感影像空间特征提取和分类的应用研究 [J]. *遥感学报*, 2002, **6**(1): 50 ~ 55.]
- 12 Zhang J, Zhang Y, Zhou T. Classification of hyperspectral data using support vector machine [A]. In: *Proceedings of IEEE International Conference on Image Processing* [C], Thessaloniki, Greece, 2001: 882 ~ 885.
- 13 Hermes L, Friauff D, Puzicha J, *et al.* Support vector machines for land usage classification in Landsat TM imagery [A]. In: *Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium* [C], Hanburg, Germany, 1999: 348 ~ 350.
- 14 He L M, Kong F S, Sheng Z Q. Multiclass SVM based land cover classification with multisource data [A]. In: *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics* [C], Guangzhou, China, 2005: 3541 ~ 3545.
- 15 Huang C, Davis L S, Townshend J R G. An assessment of support vector machines for land cover classification [J]. *International Journal of Remote Sensing*, 2002, **23**(4): 725 ~ 749.
- 16 Vapnik V N, Levin E, Le C Y. Measuring the VC-dimension of a learning machine [J]. *Neural Computation*, 1994, **8**(5): 51 ~ 76.
- 17 Vapnik V N. *Estimation of Dependence Based on Empirical Data* [M]. New York: SpringerVerlag, 1982.
- 18 Osuna E, Freund R. An improved training algorithm for training support vector machines [A]. In: *Proceedings of IEEE Workshop on Neural Networks for Signal Processing* [C], New York, 1997, **1**: 252 ~ 256.
- 19 Platt J. *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines* [M]. Cambridge, MA, USA: MIT Press, 1998: 169 ~ 182.
- 20 Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines [J]. *IEEE Transactions of Neural Networks*, 2002, **13**(2): 415 ~ 425.
- 21 Ulrich H G Krefiel. Pairwise classification and support vector machines [A]. In: Scholkopf B, Burges C J C, Smola A J eds. *Advances in Kernel Methods: Support Vector Learning* [C], Cambridge, MA, USA: MIT Press, 1999, 255 ~ 268.
- 22 Platt J, Cristianini N, Shawe-Taylor J. Large margin dags for multiclass classification [J]. *Advances in Neural Information Processing Systems*, 2000, **12**: 547 ~ 553.
- 23 Du Ming-yi, Wu Wen-bo, Guo Da-zhi. Research on multi-source based classification geographic information of desertification [J]. *Journal of Image and Graphics*, 2002, **7A**(7): 740 ~ 743. [杜明义, 武文波, 郭达志. 多源地学信息在土地荒漠化遥感分类中的应用研究 [J]. *中国图象图形学报*, 2002, **7A**(7): 740 ~ 743.]