

遥感影像分类后处理的地统计方法

姚娜¹ 林宗坚² 张景雄¹ 任超锋¹

(1 武汉大学遥感信息工程学院,武汉市珞喻路129号,430079)

(2 中国测绘科学研究院,北京市北太平路16号,100039)

摘要:采用地统计学中带局部均值的简单克里格方法和协同克里格方法,利用基于训练样本的指示数据(位置和类别)和基于光谱初分类的类别概率数据的空间结构信息,对未知点位的类别发生概率进行预测,从而修正初分类结果。实验结果表明,两种方法所获得的精度相较初分类的精度均有明显提高,这种充分利用训练样本信息改善分类结果的策略不局限于特定的初始分类器。

关键词:分类;信息;地统计学;克里格;Arif指数;精度

中图分类号:P237.4

基于遥感影像的土地覆盖制图的实质是影像分类^[1]。因此,分类器是制约影像分类精度、影响成图质量的一个关键因素。依据信息论,分类器可以视为一个通过挖掘输入特征向量(如光谱特征等)所含信息来减少初始不确定性的系统。该系统的理想情形是能够完全消除初始不确定性,即类别最终得以正确确定^[2]。实际上,该理想情形至少受限于以下两个因素:①输入特征向量,如类别相对于给定特征向量的可分性;②分类器性能,如分类器对于含噪数据的鲁棒性^[2]。

因此,对于遥感影像监督分类,即使训练样本的特征向量信息充足,分类精度也会因分类器(信息处理通道)对输入特征向量信息的非充分学习(信息损失)而低于其潜在水平。这进而导致训练样本在分类前的类别指示向量与分类后的后验概率(或隶属度)向量之间存在残差。

传统的像元级分类常常忽略了邻近像元所蕴含的空间依赖性信息^[3-4]。基于地统计学的遥感影像分类,正是利用这种空间信息来改善初始分类结果,实现输入向量信息与空间信息的融合。

本文通过对残差进行变异函数建模,并分别采用地统计学中的带局部均值的简单克里格方法与协同克里格方法对该残差作空间预测,最后用逐像元上的预测残差补偿分类器所预测的后验概率。

1 克里格方法

输入分类器的特征向量在数学上可以抽象为一个连续型随机变量,该变量在地统计学中被称为区域化变量。变异函数正是反映研究区域内区域化变量相关性及其空间结构的定量模型^[3-5]。

假设区域化变量 $Z(x)$ 为待研究的某种属性,则克里格方法是利用 $Z(x)$ 的原始数据及其变异函数的结构特点,对未采样点的区域化变量的取值进行线性无偏最优估计。它实质是一簇广义最小二乘回归算法^[4]:

$$\hat{Z}(x) - m(x) = \sum_{a=1}^{n(x)} \lambda_a(x) [z(x_a) - m(x_a)] \quad (1)$$

1.1 带局部均值的简单克里格方法

当 $m(x)$ 已知且随点位 x 变化(即 $m(x)$ 非平稳)时,式(1)便转化为带局部均值的简单克里格^[6]。

若某研究区包含 K 种土地覆盖类别, $Z(x)$ 应替换为某未知点位隶属于类别 $k(k \in K)$ 的布尔指示向量 $I_k(x)$ 。 $I_k(x)$ 的指示转换公式为:

$$I(x; k) = \begin{cases} 1, & \text{类别为 } k \\ 0, & \text{其他} \end{cases} \quad (2)$$

此时,式(1)可写为:

$$\hat{I}_k(x) = \sum_{a=1}^{n(x)} \lambda_a(x) [i_k(x_a) - p_k(x_a)] + p_k(x) \quad (3)$$

式中, $i_k(x_a)$ 为训练样本点 x_a 的类别指示向量; $p_k(x)$ 和 $p_k(x_a)$ 分别表示未知点 x 和训练样本点 x_a 经分类器预测所得的后验概率或隶属度。依据式(3)计算所得的 $\hat{I}_k(x)$ 是未知点 x 隶属于类别 k 的概率估值。当类别数为 K 时, 式(3)需递归 K 次, 再对 K 个概率估值进行归一化处理, 即可得到点 x 关于 K 个类别归属的指示向量 $I_K(x)$ 。

式(4)的目的是将训练样本的指示向量与经分类器预测的概率向量相结合, 利用训练样本残差的线性组合来预测未知点位的残差, 进而修正分类器直接预测的类别概率估值。

1.2 协同克里格方法

在地统计学中, 实现信息融合的另一常见方法是协同克里格。当预测研究域区内 K 种土地覆盖类别的空间分布时, 假设 $I_k(x)$ 为主变量, $P_k(x)$ 为次变量, 则 $I_k(x)$ 的协同克里格估计为:

$$\hat{I}_k(x) = m(x) + \sum_{a=1}^{n_a(x)} \lambda_a(x) [i_k(x_a) - m(x_a)] + \sum_{\beta=1}^{n_\beta(x)} \lambda_\beta(x) [p_k(x_\beta) - m(x_\beta)] \quad (4)$$

1.3 精度上限——Arif 指数

若能够估算检验样本特征向量的潜在最高分类精度, 则其与分类器预测所得精度之差, 便是对分类过程中信息损失的一种定量度量。

Arif 指数正是一种对潜在最高分类精度进行估算的指标。其最大特点是: 独立于特定的分类器, 直接预测输入特征向量的潜在最大分类精度。当类别完全可分时, Arif 指数为零, 随着类别混淆情况的加重, Arif 指数逐渐接近其最大值 1。

Arif 指数的定义式为^[7]:

$$\text{Arif Index} = (N - \sum_{k=1}^N S(k)) / N \quad (6)$$

该指数与最高精度之间的关系为^[7]:

$$\text{Highest Accuracy} = 100 - \text{Arif Index} \times (100 - \text{Accuracy}_{\text{lower_bound}}) \quad (7)$$

其中, 最低精度是指数量最多的类别占总体的百分比。

2 数据集描述

实验所用的土地覆盖类型数据来源于美国第

二代全国土地覆盖数据集 (national land cover dataset 2001, NLCD2001)。它将美国全境划分为 65 个制图区域, 含 16 种土地覆盖类型; 包含 Landsat5 TM 和 Landsat7 ETM⁺ 影像, 影像均采用 Albers 等面积投影, 重采样至 30 m 分辨率^[8]。

NLCD2001 所用的遥感影像数据源均来自美国地质调查局地球资源观测与科学中心 (USGS EROS Center)^[8], 本文下载了一景 Landsat ETM⁺ 多光谱影像, 并将该影像数据与土地覆盖数据配准后, 选取了北纬 $47^{\circ}41' \sim 47^{\circ}54'$, 西经 $109^{\circ}01' \sim 109^{\circ}21'$ 的研究区, 像幅大小为 500×500 个像元。该区域的土地覆盖类别综合为开放水体、森林、草地/灌木、裸地/沙地、农田和湿地等 6 类。

鉴于研究区影像逐像元的土地覆盖类别均已知, 本文随机抽取了 1 500 个训练样本, 其余像元作为检验样本, 训练样本与检验样本容量比极小。

3 实验与分析

由于本文方法独立于分类过程, 仅需分类器提供后验概率向量, 因此可选分类器范畴较大。本文采用了在算法收敛性、训练速度与分类精度等方面都具有较高性能的支持向量机 (support vector machine, SVM)^[9], 通过与统计学软件接口的 LIBSVM 软件包, 解决 $K(K > 2)$ 类土地覆盖分类问题^[10]。

SVM 所得的总体精度为 73.79%。而对检验样本集的输入特征向量进行 Arif 指数的估计结果为 0.411, 对应的潜在最高分类精度为 79.47%。该值反映了分类器对于输入特征向量信息的损失, 也说明了分类精度有待提高的空间。

表 1 比较了 SVM 的分类结果与两种克里格方法的修正结果。两种方法的总体精度和 Kappa 系数较直接分类均有了明显提高。前者分别提高了 2% 与 5%, 后者由 0.58 分别提高到 0.63 和 0.65。

此外, 带局部均值的简单克里格方法 (下文简称 SK) 的修正效果整体优于协同克里格方法。其可能原因是后者的主、次变量的趋势面通过全局范围内的空间域平滑所得, 而 SK 的趋势面是局域化至像元级的, 因而其残差意义更为准确。而且 SK 大大减少了变异函数模型的数量与建模时间, 因此本文更为推崇 SK 方法。

以 SK 为例, 随机显示了 15 个真实类别为农田, 而 SVM 分类器对其预测类别为其他 (非农

田)的检验样本。图 1 比较分析了 SVM 所预测的后验概率与 SK 修正所得的后验概率的数量关系(为了更加清晰地显示修正前后概率值的波动幅度,SK 对 SVM 预测的后验概率进行残差补偿后,其修正值未经归一化处理)。图中横坐标表示检验样本数量,纵坐标代表后验概率。折线所串联的●为经由 SVM 分类器预测后,检验样本隶属于农田的后验概率;▲为经由 SVM 分类器预测后,检验样本隶属于其农田的最大后验概率。显然,图 1(a)~1(b)反映了 SVM 未能作出正确的预测。为经 SK 进行残差补偿后,检验样本隶属于农田的后验概率;■为 SK 对原本隶属于非农田类别的最大后验概率进行残差补偿后的修正值。

图 1(a)~1(c)反映了 SK 的总体效果,即通过全局的残差补偿,将训练样本所提供的类别空间分布信息内插至全局,在提高像元隶属于真实类别的后验概率的同时,降低其隶属于错误类别的后验概率。这种效果对于协同克里格方法同样成立。以图 1(a)为例,真实类别为农田的像元,经 SVM 预测后,其隶属于农田的概率低于非农田的概率,导致了漏分。而经 SK 修正后,这些像元隶属于农田的概率增大;相应地,其隶属于非农田的概率降低,有效地修正了分类结果。这一点从表 1 中农田类别生产者精度的提高中也可以看出。在图 1(b)中,错误类别后验概率的减小与正确类别后验概率的增加,不足以扭转分类器直接

预测的结果,但 SK 修正效果仍然可见。

图 1(d)中,SK 对 SVM 原本正确预测的类别进行了错误的修改。这主要源于以下两个原因:① 特征向量本身的易混淆性导致分类后隶属于农田的后验概率优势微弱;② 局部样本数量稀少且分布不均导致局域残差空间变异模拟的欠准确及空间内插过程的过机械,即空间域残差补偿的欠准确使得其所造成的概率扰动会轻易抵消甚至倒置原本微弱的概率优势。同理,结合表 1,相较 SVM 直接预测所得的生产者精度或用户精度,克里格方法修正后所得的相应精度变化,并不反映该方法对某些类别的特殊适应性。而且图 1(d)所代表的检验样本占全体农田检验样本的百分比仅为 1.1%,相较于图 1(a)中正确修正的 58%的样本,SK 的修正效果值得推崇。

为了进一步验证本文方法的有效性,笔者还将 SK 应用于另一景含 17 种土地覆盖的 TM 影像,包括光谱波段、高程等 11 个输入特征。Arif 指数估算的潜在最高分类精度为 72.22%,而采用广义线性模型分类所得的总体精度和 Kappa 系数分别为 65.55%和 0.62。经残差补偿后,两者分别提升至 75.45%和 0.73,效果显著。值得说明的是,Airf 指数所估计的最高精度 72.22%与修正精度 75.45%并不矛盾,因为前者的估计完全依赖于输入特征向量,而未考虑空间信息的引入。

表 1 SVM 分类器与两种克里格方法的精度评价指标/(%)

Tab. 1 Accuracy Assessment Indexes of SVM Classification and Two Kriging Methods

方法类别	SVM 分类		协同克里格		带局部均值的简单克里格	
	生产者精度	用户精度	生产者精度	用户精度	生产者精度	用户精度
开放水体	37.29	53.14	24.41	48.65	31.86	50.81
森林	78.95	84.17	76.52	81.59	77.95	83.12
草地/灌木	57.36	66.33	58.86	64.91	61.29	64.47
裸地/沙地	91.96	74.51	86.25	82.76	87.53	81.82
农田	15.92	59.75	73.28	76.11	72.80	73.93
湿地	10.32	41.80	35.98	20.51	27.16	54.83
总体精度	73.79		75.77		77.02	
Kappa 系数	0.58		0.63		0.65	

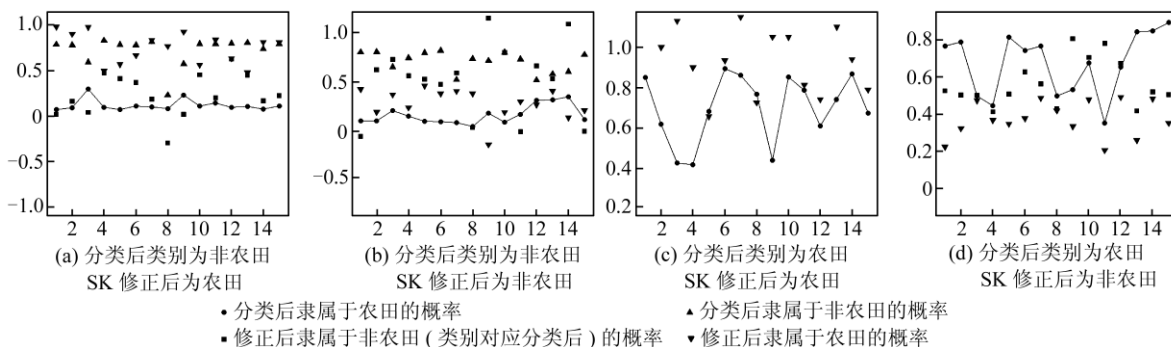


图 1 SVM 分类器与 SK 修正的后验概率增减情况(以农田为例)

Fig. 1 Thea Posterior Probabilities of the SVM Classifier and the SK Method (e. g., Farmland)

4 结 语

本文方法独立于分类过程,可以视作一种分类后处理方法。它旨在补偿分类器对输入特征向量信息的客观损失,通过充分利用训练样本已知的类别归属信息与分类所得的类别概率估计之间的差异,并依据这种差异的空间结构信息,将其内插至全局,进而弥补了分类器对样本信息挖掘的不足。此外,这种全局残差补偿的方法在原理上独立于应用所涉及的土地覆盖类型,但其应用效果,即分类精度的提高幅度,会受到样本容量、分布等的影响。基于可控或理想条件下的模拟数据的验证、基于可分性的特征向量的提取等科学问题仍值得进一步探讨。

参 考 文 献

- [1] Foody G M. Status of Land Cover Classification Accuracy Assessment[J]. *Remote Sensing of Environment*, 2002, 80:185-201
- [2] Battiti R. Using Mutual Information for Selecting Features in Supervised Neural Net Learning [J]. *IEEE Transactions on Neural Networks*, 1995, 5 (4): 537-551
- [3] Goovaerts P. Geostatistical Incorporation of Spatial Coordinates into Supervised Classification of Hyperspectral Data[J]. *Geographical Systems*, 2002, 4: 9-111

- [4] 张景雄, Goodchild M. 野外空间采样的渐进式策略 [J]. *武汉大学学报·信息科学版*, 2008, 33(5): 441-445
- [5] 张景雄. 空间信息的尺度、不确定性与融合[M]. 武汉:武汉大学出版社,2008
- [6] Goovaerts P. *Geostatistics for Natural Resources Evaluation* [M]. New York: Oxford University Press, 1997
- [7] Arif M, Afsar F A, Akram M U, et al. Arif Index for Predicting the Classification Accuracy of Features and its Application in Heart Beat Classification Problem[C]. *Advances in Knowledge Discovery and Data Mining: Proceedings of the 13th Pacific-Asia Conference*, Thailand, 2009
- [8] Homer C, Dewitz J, Fry J, et al. Completion of the 2001 National Land Cover Database for the Conterminous United States[J]. *Photogrammetric Engineering and Remote Sensing*, 2007, 73(4): 337-341
- [9] 万幼川, 黄俊. 几何和图论特征对高分辨率遥感影像土地利用分类的影响[J]. *武汉大学学报·信息科学版*, 2009, 34(7): 794-797
- [10] Meyer D. Support Vector Machines-The Interface to Libsvm in Package e1071[J/OL]. <http://cran.r-project.org/web/packages/e1071/vignettes/svm-doc.pdf>, 2009

第一作者简介:姚娜, 博士生, 研究方向为低空无人飞行器遥感系统、遥感影像的信息量与不确定性分析。
E-mail:nayao@foxmail.com

Geostatistical Approaches to Post-classification of Remote Sensing Image

YAO Na² LIN Zongjian¹ ZHANG Jingxiong² REN Chaofeng²

(1 School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Chinese Academy of Surveying and Mapping, 16 Beitaping Road, Beijing 100039, China)

Abstract: This paper explores two methods pertaining to geostatistics, i. e. , simple kriging with local mean and cokriging, to predict class occurrences based on training samples' indicator transforms (location and classes) and spectrally derived class probabilities, thus calibrating the a posterior class probability vectors derived from initial spectral classification. The results showed that classification accuracy is significantly increased by these two methods for utilizing spatial information contained in training samples and initial spectral classification, compared with those obtainable with spectral classification. Moreover, the proposed methods constitute a valuable strategy for making fuller use of information residing in training data for improving spectrally derived classification, which is independent of the specific classifiers initially adopted for image classification.

Key words: classification; information; geostatistics; Kriging; Arif index; accuracy