

多时相遥感影像样本迁移模型与地表覆盖智能分类

杜培军^{1,2}, 林 聪^{1,2}, 陈 宇³, 王 欣⁴, 张 伟^{1,2}, 郭山川^{1,2}

(1. 南京大学 地理与海洋科学学院, 江苏 南京 210023; 2. 自然资源部国土卫星遥感应用重点实验室, 江苏 南京 210023; 3. 中国矿业大学 环境与测绘学院, 江苏 徐州 221116; 4. 成都理工大学 地质灾害防治与地质环境保护国家重点实验室, 四川 成都 610059)

摘要: 利用时间序列遥感影像重建过去几十年的地表覆盖是实现时空多维地理场景感知与动态建模的基础, 但存档历史遥感影像分类面临样本选择难、多时相影像协同解译水平低的问题。研究提出了一种基于已有土地覆盖产品与对应遥感影像中几何及属性特征约束的样本时空迁移方法, 将迁移获得的训练样本嵌入多时相地表覆盖分类框架, 获得多期地表覆盖分类结果, 实现历史时期地表环境的智能感知与动态制图。太湖流域多时相分类的结果表明, 无监督样本迁移方法可以充分利用先验几何约束和光谱属性, 从参考地表覆盖产品中快速获得可靠的训练样本, 多时相分类精度均高于 89%, 满足大区域多时相地表覆盖的分类需求, 为地理环境演变建模提供了有效支持。

关键词: 遥感影像分类; 地表覆盖; 迁移学习; 聚类分析; 变化检测

中图分类号: P237

文献标志码: A

Training Sample Transfer Learning from Multi-temporal Remote Sensing Images for Dynamic and Intelligent Land Cover Classification

DU Peijun^{1,2}, LIN Cong^{1,2}, CHEN Yu³, WANG Xin⁴, ZHANG Wei^{1,2}, GUO Shanchuan^{1,2}

(1. School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China; 2. Key Laboratory for Land Satellite Remote Sensing Application of Ministry of Natural Resources, Nanjing 210023, China; 3. School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China; 4. State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China)

Abstract: Reconstructing historical land cover dynamics by time series remote sensing images is the basis of geographic scene sensing and modeling. However,

classification of historical archived remote sensing images is quite difficult due to the limited training samples and low interpretation capability for multi-temporal images. A spatio-temporal training sample transfer method from existing land cover products and corresponding remote sensing images was proposed based on geometric and attribute constraints. The transferred training samples were then embedded into a multi-temporal land cover classification framework to solve the problem of high sample labeling cost in multi-temporal image classification. The multi-temporal classification results in the Taihu Basin show that the proposed unsupervised sample transfer method can make full use of the prior geometric constraints and spectral properties of the products and images, and the classification accuracy in large scale area is over 89%. The results demonstrate that the proposed method is effective for land cover updating and geographic environment evolution modelling.

Key words: remote sensing image classification; land cover; transfer learning; clustering analysis; change detection

对地观测卫星提供了多谱段、多分辨率、多时相、海量的准实时及存档遥感影像, 具备了大数据种类多、体量大、动态多变、高价值以及冗余模糊的“5V”特征^[1]。虽然对地观测数据在体量、时效性、分辨率等方面提升迅速, 但遥感数据处理、分析以及进一步感知、认知的能力尚显缺乏, 影像信息智能提取与理解成为亟需解决的关键技术问题^[2-4]。因此, 突破遥感信息解译的传统模式, 采用知识构建多属性的样本库, 借助先进机器学习方法的支持, 从而高效利用海量遥感影像数据, 将遥感数据转化为知识, 深化遥感在各个行业内的应用, 是遥感信息科技发展

收稿日期: 2022-05-02

基金项目: 国家自然科学基金(41631176)

第一作者: 杜培军(1975—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为城市遥感, 遥感信息智能处理与地学分析。E-mail: peijun@nju.edu.cn



论文
拓展
介绍

的重要方向^[1,5]。

地表覆盖是地表各种物体的类型、状态、特征与自然属性的综合体^[6-7]。遥感影像是地表覆盖分类的重要数据源^[8]。在过去的几十年内,地表覆盖遥感解译在数据、策略、模型、计算能力、应用等方面取得了重大的进展^[9]。机器学习方法已成为地表覆盖分类制图的重要技术手段^[10]。

国内外许多研究团队致力于大尺度地表覆盖产品的研究,得到了一系列不同空间分辨率的全球尺度地表覆盖产品。低分辨率数据覆盖产品如:全球 5km 分辨率的 34 年连续年度地表覆盖产品 GLASS-GLC^[11],全球 1 km 空间分辨率的 IGBP DISCover 地表覆盖产品^[12],马里兰大学全球 1km 覆盖数据集^[13],全球 1km 覆盖数据库 GLC2000^[14],MODIS 全球地表覆盖类型产品等^[15-16]。随着 Landsat 系列以及 Sentinel-2 卫星影像的广泛应用,30m 及更高分辨率地表覆盖产品的生产成为新的研究重点。Gong 等^[17]制作了 2010 年全球 30m 地表覆盖产品 FROM-GLC 2010,进一步更新生产了 FROM-GLC 2015 与 FROM-GLC 2017,此后采用 2014—2015 年 Landsat 8 全球影像标记生成包含了约 14 万样本点的训练样本集,对 Sentinel-2 数据进行分类得到了全球首套 10m 空间分辨率的地表覆盖产品 FROM-GLC10^[18]。国家基础地理信息中心于 2014 年发布了 GlobeLand30 V2000 与 GlobeLand30 V2010,目前 GlobeLand30 V2020 已经可以公开获取^[19-20]。Zhang 等^[21]以已有全球地表覆盖产品为基础,自动生成时空光谱库,生产了 30m 分辨率全球精细地表覆盖产品 GLC_FCS30。

全球或大尺度的地表覆盖产品为地表覆盖分类研究提供了丰富的数据集。特别是对于基于监督算法的地表覆盖分类研究,通过对已有地表覆盖产品的充分利用,可以挖掘潜在的训练样本标签或分类规则,解决当前地表覆盖监督分类选取高质量训练样本费事费力的问题。在这一背景下,以已有地表覆盖产品为参考,充分利用产品中有效的知识,自动完成地表覆盖制图具有重要的实用价值。因此,探索样本迁移方法在地表覆盖动态制图中的作用,推进遥感影像的智能化解译,支持大数据时代下的遥感影像数据处理、分析、理解与应用,对以地表覆盖为基础的相关科学研究具有重要的意义。

迁移学习的主要思想是通过采用已有知识对不同但是相关的领域问题进行求解^[21]为多源多时相的地表覆盖分类提供了更加低成本的解决方法,在遥

感领域得到了广泛的关注与研究^[22-29]。领域适应(Domain Adaptation, DA)是一类重要的迁移学习方法。遥感影像的领域适应问题是将不同区域的影像或同一区域不同时相的影像考虑为源领域与目标领域两个部分,利用源领域中的知识训练分类器,解决目标领域的影像分类问题^[30]。领域适应在遥感影像迁移学习中可以被划分为四种类型,分别为不变特征选择法^[31-32](DA by selecting invariant features)、数据分布适应法^[33-34](DA by adapting data distribution)、分类器适应法^[35-36](DA by adapting the classifier)与主动学习法(DA by active learning)^[37-38]。上述的迁移学习算法往往比较复杂,难以在大尺度地表覆盖制图中应用。而基于变化检测的样本迁移方法高效、简便、鲁棒性强,具备在大尺度地表覆盖分类中广泛应用的潜力。

相比其他迁移学习方法,基于变化检测的样本迁移方法无需对源领域与目标领域之间的特征或分类器进行调整,通过确定双时相/多时相影像之间的不变区域来降低目标领域影像与源领域影像之间的统计分布差异。基于变化检测的样本迁移方法在降低领域之间的分布差异上有较好的鲁棒性,广泛应用于大尺度地表覆盖分类研究,如联合变化检测与分类的地表覆盖自动更新方法^[39]、变化检测驱动的迁移学习(Change Detection driven Transfer Learning, CDTL)方法^[40]、面向对象的变化检测与源领域样本迁移^[41]、时间序列影像变化检测与不变区域参考地表覆盖标签迁移^[42]、源领域影像分类后验概率高置信度样本标签选择与迁移^[43]、联合深度学习与变化检测的样本迁移方法^[44]等,以上方法均在应用场景中取得了理想的样本迁移与地表覆盖更新制图效果。但是这类样本迁移方法大多基于准确的源领域训练样本,当源领域训练样本不够准确时,例如采用已有地表覆盖产品作为类别标签,则缺乏降低不确定性的策略,目标领域影像将难以获得理想的地表覆盖分类结果。

综上,对于大尺度区域的多时相地表覆盖分类,缺乏快速、自动化的样本迁移方法。直接采用基于变化检测的样本迁移算法缺乏对源领域训练样本不确定性的考虑。虽然综合多源地表覆盖产品可以提供更加可靠的训练样本^[45-46],或是依赖时间序列的多源遥感信息来提高含噪声样本集下的地表覆盖分类精度^[47-48],但是这些方法依赖过多的输入数据,导致模型相对复杂,对大尺度区域或缺乏多源参考数据的历史存档数据的地表覆盖分类应用并不完全适

用。因此,本文提出了一种快速的、轻量级的样本迁移方法,依赖更少的数据输入从地表覆盖产品中自动获得高质量训练样本,与基于变化检测的多时相分类相结合,自动获得高精度多时相地表覆盖制图截图。研究以太湖流域为试验区,提出一种几何与光谱属性约束下的无监督样本迁移算法,实现利用单景影像有效降低地表覆盖产品的不确定性,研究基于产品和多时相影像的样本迁移模型与自动分类框架,以期为长时间序列地理环境演变分析提供支持。

1 研究区和数据

1.1 研究区

研究区域包含太湖流域及周边快速城镇化地区,如图1所示。太湖流域位于中国东部地区,处于长三角城市群的核心位置,包括了上海市、江苏南部、浙江北部以及部分安徽地区,总面积约37 000 km²,流域整体以平原地貌为主。太湖流域是长三角城市群重要组成部分,经济水平发达,城镇化程度高,人类活动剧烈,地表覆盖变化快速。

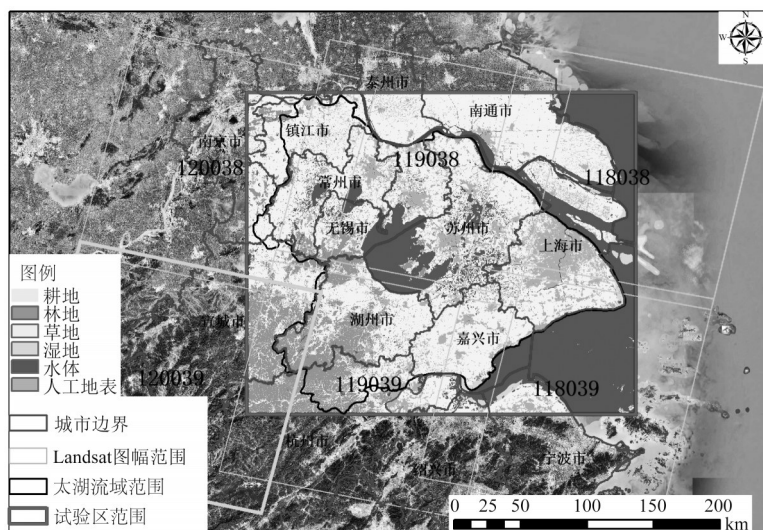


图1 试验区范围

Fig.1 Region of study area

根据地表覆盖产品 GlobeLand30 的分类统计结果,研究区主要地表覆盖类型为人工地表、耕地、水体、林地、草地和湿地,其他地类占比均低于0.03%。夜间灯光遥感数据通过观测城镇地区的人造光源信息来反映当地的人类活动情况和城镇化水平^[49]。分析研究区1992—2013年间的多期DMSP-OLS夜间灯光遥感数据(图2)可知,该区域在1992年就具备较好的城镇化基础,主要的大城市已经具备一定城镇化规模。1992—2007,研究区城镇化进展迅速,主要表现为大城市的持续扩张以及小城市的迅速城镇化,使得太湖流域内的城市群有明显一体化的趋势。至2010年,太湖流域及其周边地区已经具有较高的城镇化水平,2010—2013年期间,研究区的城镇化仍在进一步发展,城区范围仍然有较为明显的扩展。综上,研究区地表覆盖类型复杂,地表覆盖持续长时间地快速变化,是典型的快速城镇化地区。

1.2 数据

选用的地表覆盖产品为GlobeLand30全球地表

覆盖产品,该产品基于POK的遥感影像制图技术制作,研究采用了V2000以及V2010两个年份的产品。GlobeLand30产品分类的影像主要包括Landsat的TM、ETM+以及OLI传感器的多光谱数据以及中国环境减灾卫星(HJ-1)的多光谱数据,空间分辨率均为30m。GlobeLand30产品组织开展了大量第三方的精度评价工作,评价得到V2010产品的总体精度为83.50%,Kappa系数为0.78^[50]。

研究区为包含太湖流域空间范围的外接多边形,试验区大小为9 752×8 074,共计78 737 648个像元,需要6幅Landsat影像覆盖,图幅号分别为120/038、119/038、118/038、120/039、119/039与118/039(图1)。各个年份的影像采用当年Landsat年度观测序列中值合成,试验中选取的影像年份为1990年、1995年、2000年、2005年、2010年与2015年共计6期中值合成数据。经过投影转换与裁切等预处理,获得试验区的多时相影像数据。

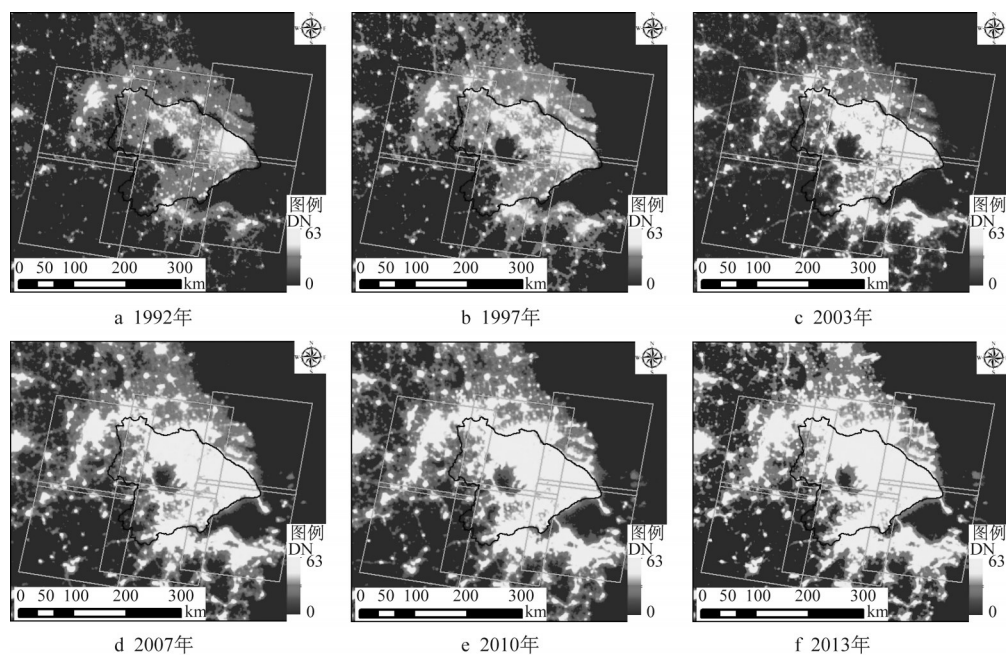


图2 研究区多年DMSP-OLS夜间灯光数据年度合成数据

Fig.2 Multi-temporal DMSP-OLS night light data in the study area

2 多时相样本迁移模型与分类方法

2.1 方法原理与流程

高精度的地表覆盖产品中存在部分可利用的类别标签,结合同一年份或相近时相的遥感影像生成训练样本集,重复利用这些训练样本,可以减少甚至完全避免手动标记新的训练样本。地表覆盖产品可以视为土地斑块与其对应类别的集合,提供了类别信息与代表地表覆盖空间连续性的土地斑块信息。

地表覆盖产品中的斑块提供了一种有效的几何约束,可以作为迁移与优化原始地表覆盖信息的一种先验局部空间单元。考虑产品中的斑块信息作为先验知识,提出一种针对地表覆盖产品的样本优化迁移模型,在无需手动设置任何参数与选择新训练样本的前提下,利用单时相影像快速准确地从地表覆盖产品中获取有价值的可利用样本标签,替代传统遥感影像监督分类中的手动选取样本的环节。将提出的样本迁移模型嵌入多时相影像分类的算法中,快速获得高精度多时相地表覆盖制图结果。

该方法主要包括三个步骤:

(1)以地表覆盖产品斑块为局部单元,迁移样本标签:从斑块单元的几何约束确定地表覆盖不确定性分析的计算单元,通过挖掘影像光谱属性信息有效降低地表覆盖产品的不确定性;

(2)以类别信息为全局单元,优化样本标签:考

虑不同地物类型在特定特征空间内的分布,从全局出发将不同地物类型分别构建高斯混合模型(Gaussian Mixture Model, GMM),通过求解GMM的过程去除错分斑块以及步骤(1)带来的错误标签;

(3)变化检测与集成学习分类器协同获取多时相地表覆盖分类结果:通过变化检测技术确定源影像与目标遥感影像之间的不变区域,形成多时相目标影像的训练样本集,获得多时相地表覆盖分类结果。

为了方便描述方法与公式部分涉及的变量,对输入的数据统一定义:令 $X=(X_1, X_2, \dots, X_q, \dots, X_Q)$ 代表对同一区域重复观测 Q 次的Landsat多时相多光谱影像集,其中 $X_q \in R^{B \times d}$ 代表由 B 个波段和 d 个像素组成的获取于 t_q 时间的Landsat多光谱影像;令 $M=\{\Omega, p\}$ 代表 t_q 时相的地表覆盖产品,其中 $\Omega=\{\omega_i\}_{i=1}^I$ 代表 I 种地表覆盖类别的集合,而 $p=\{p_j\}_{j=1}^J$ 代表 J 个土地斑块组成的集合,其中每个斑块具有唯一的地表覆盖类型。

2.2 局部单元样本迁移方法

原始地表覆盖产品中的多边形信息难以直接作为几何约束,主要原因有以下两点:

(1)由于影像分类的椒盐效应,产品中存在一些细小的多边形,不适合作为局部单元开展聚类算法;

(2)由于分类错误、后处理、以及地物真实分布,

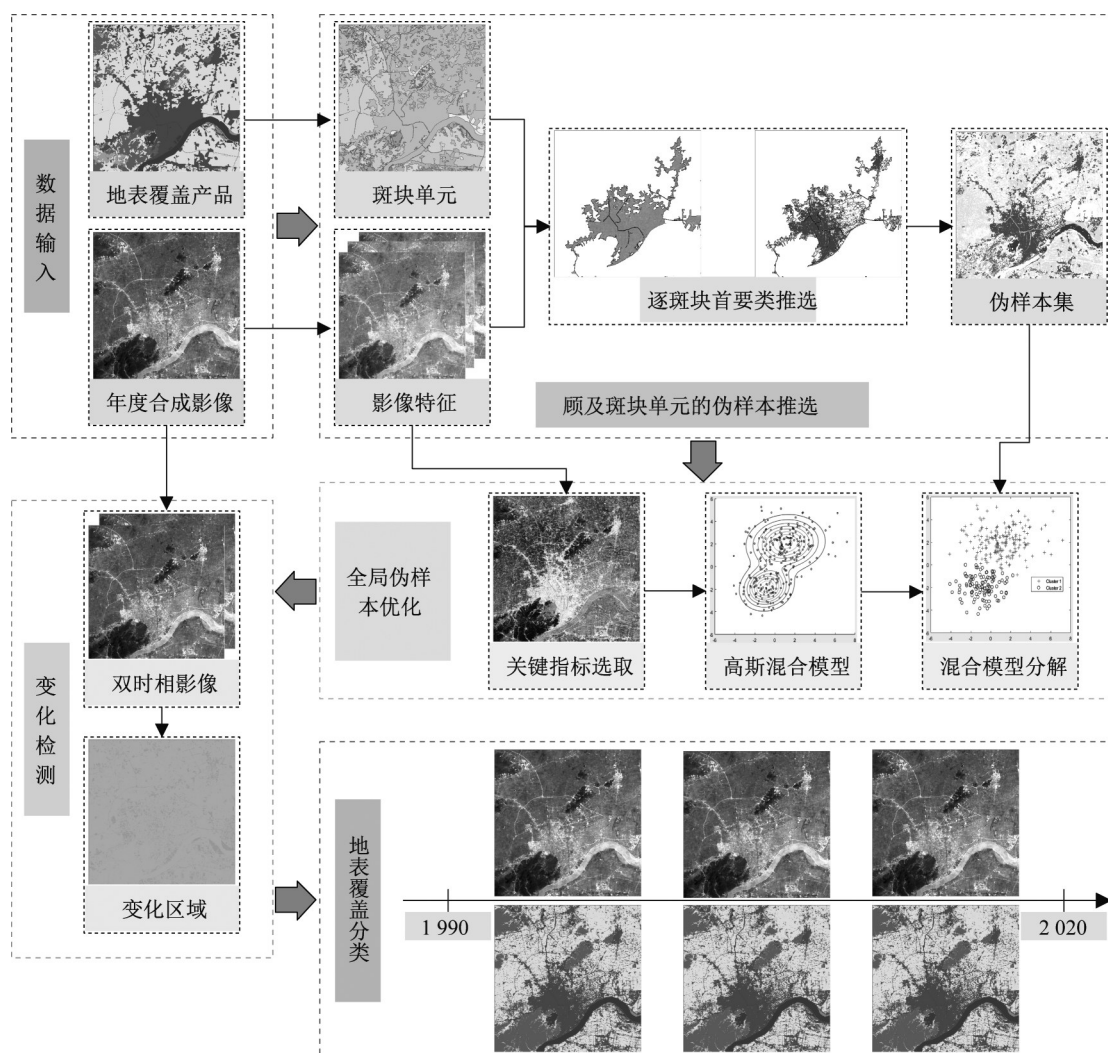


图3 基于无监督样本迁移方法的多时相地表覆盖分类流程

Fig.3 Flowchart of multi-temporal land cover classification based on unsupervised sample transfer

产品中存在由于斑块联合构成大多边形,使得局部单元的聚类算法难以收敛,属于错误的几何先验知识。基于此,采用形态学开运算预处理原始产品,使得细小的多边形可以被去除,斑块联合构成的多边形可以被分解为有效的斑块单元,形成有效的几何约束。

为了有效约束斑块单元内的错误地表覆盖信息,采用集成多种光谱指数特征的归一化光谱向量^[51] (normalized difference spectral vector, NDSV) 作为聚类的特征输入。该特征对于聚类算法有以下优势:

(1) NDSV 计算得到的光谱向量特征是归一化且全局连续的,适合直接作为无监督模型的输入特征;

(2) NDSV 的每一维特征信息相互关联,有助

于分析复杂地表覆盖环境下的各类地物的分布,增益聚类算法的相似性度量过程;

(3) 该方法完整计算了所有的波段组合,保障了在聚类分析过程中有效约束斑块单元内的错误地表覆盖信息。对于第 d 个像元而言,原始光谱特征为 $\mathbf{x}_d^B = (x_d^{b_1}, x_d^{b_2}, \dots, x_d^{b_B})$, 相应的 NDSV 基于式(1)计算为

$$\mathbf{y}_d = \begin{bmatrix} f(x_d^{b_1}, x_d^{b_2}) \\ f(x_d^{b_1}, x_d^{b_3}) \\ \vdots \\ f(x_d^{b_1}, x_d^{b_B}) \\ f(x_d^{b_2}, x_d^{b_3}) \\ \vdots \\ f(x_d^{b_{B-1}}, x_d^{b_B}) \end{bmatrix} \quad (1)$$

式中: y_d 代表第 d 个像元的光谱特征向量; $x_d^{b_B}$ 为 b_B 波段的反射率值; $f(x_d^{b_{B-1}}, x_d^{b_B})$ 为 b_{B-1} 波段与 b_B 波段的归一化差值。将计算得到的NDSV特征与地形特征(高程、坡度)组合, 形成属性约束特征集。

将 K -means 作为基础的聚类算法对 $p = \{p_j\}_{j=1}^J$ 逐个实施聚类分析, 以第 j 个斑块 p_j 为例, 其中 p_j 由 N 个像元组成, 则 p_j 对应的多光谱数据为 $X_{p_j} = \{x_1^B, x_2^B, \dots, x_n^B, \dots, x_N^B\}$, 对应的属性特征集为 $Y_{p_j} = \{y_1, y_2, \dots, y_n, \dots, y_N\}$ 。假设 p_j 对应的类别标签为 ω_u , 为了从 p_j 中分离出与 ω_u 正确关联的子集, 将 Y_{p_j} 通过 K -means 方法聚类划分为 K_j 个簇 $\{C_j^1, C_j^2, \dots, C_j^{K_j}\}$, K -means 通过最小化平方误差完成对簇的划分, 即

$$E = \sum_{k=1}^{K_j} \sum_{y_n \in C_j^k} \|y_n - \mu_k\|_2^2 \quad (2)$$

式中: E 为平方误差; K_j 为预期划分的簇数; y_n 为 Y_{p_j} 中 n 的第 n 个特征向量; C_j^k 为第 k 个簇, μ_k 为 C_j^k 的均值向量。

K_j 是上述过程中唯一需要输入的变量, 且最优的聚类簇数可以更好地划分 Y_{p_j} 。通过计算 Calinski-Harabasz (C-H) 指数来寻优每个局部单元内的最优簇数, C-H 指数通过方差比准则 (variance ratio criterion, VRC) 来评价聚类效果的好坏, 在聚类结果的基础上, 计算总体簇间方差 σ_B (overall between-cluster variance) 与总体簇内方差 σ_w (overall within-cluster variance), 通过式 (3) 计算方差比:

$$V_{K_j} = \frac{\sigma_B}{\sigma_w} \times \frac{(N - K_j)}{(K_j - 1)} \quad (3)$$

式中: N 为 Y_{p_j} 的特征向量的数目, K_j 为聚类簇数, V_{K_j} 为簇数为 K_j 下的方差比结果。通过定义簇数范围, 逐个计算 VRC 结果, 将 VRC 最大值对应的簇数作为当前局部单元下的聚类簇数。

在几何与属性约束下, 将每个局部单元内的像素集合划分为多个簇, 将占比最多的簇保留并继承原始地表覆盖产品的类别标签^[48], 迁移得到伪样本集。之所以称为伪样本集, 是考虑到地表覆盖产品几乎不可能保证每个斑块都分类正确, 因此当前的样本集中存在一定数量的错误, 需要进一步优化。

2.3 全局样本优化方法

局部单元样本迁移方法从地表覆盖产品 $M =$

$\{\Omega, p\}$ 中获取了伪训练样本集 $D_{pseudo} = \{d_{p_1}, d_{p_2}, \dots, d_{p_j}, \dots, d_{p_j}\}$ 。为了尽量剔除伪样本集中错误的样本, 获得一个优化后的训练样本集 D ($D \in D_{pseudo}$), 提出了一种基于高斯混合模型的全局样本优化方法。从全局影像特征出发, 构建高斯混合分布, 从统计分布角度约束伪样本集中的错误样本。采用高斯混合模型分解的手段将伪样本集划分, 保留正确分布, 获得目标训练样本集, 自动完成样本优化的过程。

对 D_{pseudo} 按照对应类别标签 ω_i 进行分解, 在分类体系 $\Omega = \{\omega_i\}_{i=1}^I$ 下将伪样本表达为不同地类伪样本集的集合 $D_{pseudo} = \{D_{\omega_1}, D_{\omega_2}, \dots, D_{\omega_i}, \dots, D_{\omega_I}\}$ 。 D_{ω_i} 为类别 ω_i 对应的伪样本集, 可以视为 ω_i 类与非 ω_i 类的两个高斯分布的混合, 对 D_{ω_i} 构建如下高斯混合模型:

$$p_m(y) = \alpha_1 p(y|\mu_1, \Sigma_1) + \alpha_2 p(y|\mu_2, \Sigma_2) \quad (4)$$

其中 $p_m(\cdot)$ 为概率密度函数, $\alpha_1, \mu_1, \Sigma_1$ 分别为第一个高斯分布的混合系数、均值向量与协方差向量, $\alpha_2, \mu_2, \Sigma_2$ 分别为第二个高斯分布的混合系数、均值向量与协方差向量。

采用迭代期望最大化 (EM) 计算高斯混合模型, 假设训练集 $D_{\omega_i} = \{d_1^{\omega_i}, d_2^{\omega_i}, \dots, d_{N_i}^{\omega_i}\}$ 由 N_i 个样本组成, 令 GM_1 与 GM_2 分别代表生成 D_{ω_i} 训练集的高斯混合成分, 则对于第 n_i ($n_i \leq N_i$) 个样本而言, 有隶属于 GM_j ($j = 1, 2$) 的后验概率 $\gamma_{(n_i, j)}$ 。EM 算法计算高斯混合模型分为两个部分, 首先执行 E 步, 通过当前的混合成分参数 $\{(\alpha_j, \mu_j, \Sigma_j) | j = 1, 2\}$, 计算 $d_{n_i}^{\omega_i}$ 属于 GM_1 与 GM_2 的后验概率 $\gamma_{(n_i, 1)}$ 与 $\gamma_{(n_i, 2)}$; 然后将 E 步的后验概率作为权重, 采用最大似然法估算新的混合参数 α_j , 均值向量 μ_j 以及协方差向量 Σ_j , 这部分为算法的 M 步。通过 E 步和 M 步的不断迭代更新, 直到 M 步似然函数的变化率 ($\Delta LL(\cdot)$) 小于设定阈值。

选择 NDVI、MNDWI 以及地形数据中的坡度特征作为关键特征, 完成对地表覆盖主要类别水体、人工地表、林地、草地以及耕地的全局优化, 具体优化流程如图 4 所示。

2.4 多时相地表覆盖分类

通过变化矢量分析 (change vector analysis, CVA) 结合大津法 Otsu 的变化检测方法确定不变区域, 传递不变区域的样本标签, 获得待分类时相的训练样本。采用随机森林 (random forest, RF) 作为分

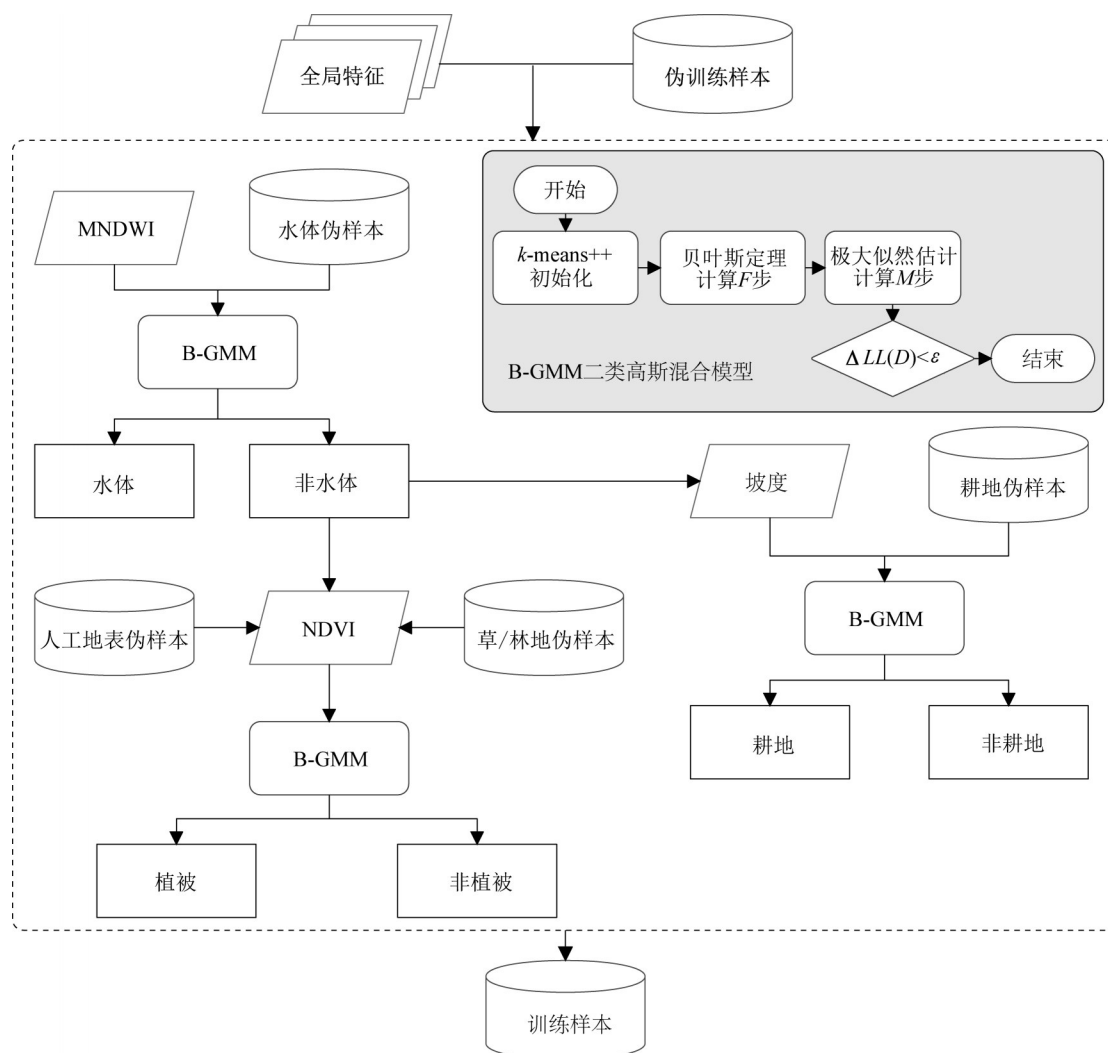


图4 基于高斯混合模型的全局伪样本优化流程

Fig.4 Global pseudo-sample optimization process based on Gaussian mixture model

类器完成多时相地表覆盖分类制图。由于研究中用于分类的训练样本并非手动选择,导致最终用于多时相分类的样本中存在少量错误,另一方面,由于RF中的基分类器互相之间相关性较弱,各自的错误预测也应该是几乎不相关的,因此对RF通过对多个基分类器的集成学习可以提高最终分类的结果。其他的类似的集成学习方法也可以替代RF完成本研究的分类任务。

3 结果与分析

对GlobeLand30 V2000与2000年中值合成影像、GlobeLand30 V2010与2010年中值合成影像分别执行无监督样本迁移方法,得到2000年与2010年的训练样本。通过变化检测方法,将不变区域的样本进行传递,基于2000年的训练样本得到1990年与

1995年的训练样本,基于2010年训练样本的得到2005年与2015年的训练样本。采用随机森林作为分类器,只采用影像原始的多光谱数据与地形特征(高程、坡度)参与分类,将6个时相的训练样本输入分类器,得到6个时相的地表覆盖分类结果,如图5所示。

为了进一步验证地表覆盖分类结果的可靠性,定量评价各期地表覆盖分类结果,基于GEE平台的TimeSync+(TimeSync-Plus)工具结合Landsat影像时序信息来标记验证样本。在随机生成样本点的基础上,在TimeSync+中获得多期验证样本点数据,将样本点叠加在影像上进行目视解译的验证、删除与增选,在2010与2015年选择Google Earth高分影像辅助上述过程,获得多期验证样本点数据,对地表覆盖分类结果进行精度评价。

表1为1990—2015年共6期的地表覆盖分类精

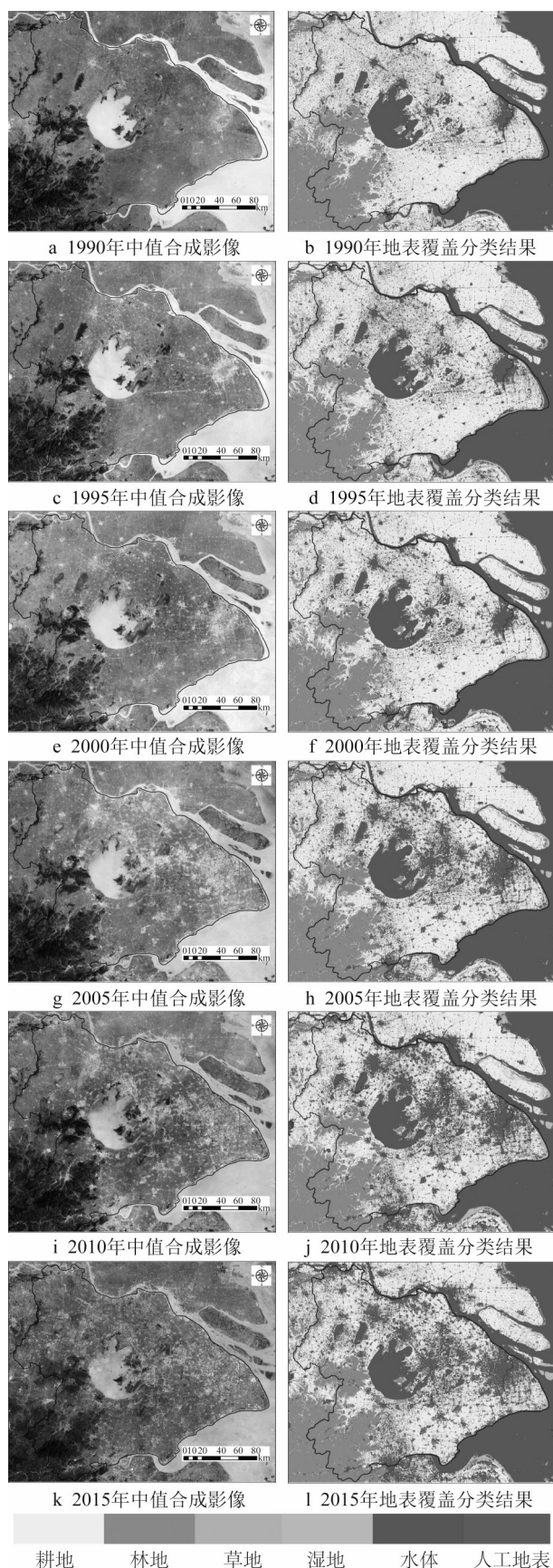


图5 太湖流域1990—2015年多时相地表覆盖分类结果
Fig.5 Multi-temporal land cover classification results in the Taihu Basin from 1990 to 2015

度评价结果。可以看出,2000年与2010年的总体精度超过91%,表明样本迁移方法可以有效替代手动标记样本的过程,获得高质量的训练样本。在只采用原始多光谱特征与地形特征的基础上,其他年份的地表覆盖分类精度也都在90%左右,因此提出的样本迁移方法与多时相分类技术可以自动生成可靠的地表覆盖分类结果。

为了进一步验证比较提出方法的可靠性,选择2000年的分类结果与GlobeLand30 V2000进行目视比对,选择2010年的分类结果与FROM-GLC 2010进行对比,选择2015年的分类结果与FROM-GLC 2015以及GLC_FCS30的2015年产品进行对比。

从图6对比结果来看,提出的无监督样本迁移模型支持的地表覆盖分类结果在空间细节上对比原始产品GlobeLand30有显著提升,更加符合影像中地物真实分布的状况。由于采用基于面向对象的分类方法以及人工编辑后处理,导致GlobeLand30产品虽然有较好的评价精度结果,但是实际上没有对30m影像中的地物进行空间上的精准制图,提出的无监督样本迁移方法基于GlobeLand30,在保持较高的分类精度的前提下,获得了细节更加丰富的地表覆盖分类结果。

从图7对比结果来看,本研究获得的2010年地表覆盖产品分类结果明显优于FROM-GLC 2010。FROM-GLC 2010存在将大量耕地与人工地表错分为裸地的情况,而事实上研究区只存在极少量的裸地。同时,FROM-GLC 2010也将大量的林地错分为了耕地,而提出的方法基本上正确分类了这些林地像元。值得注意的是,FROM-GLC 2010是较早发布的全球地表覆盖产品,因此在地表覆盖情况比较复杂的太湖流域地区精度存在一定不足。

从图8对比结果来看,提出的方法得到了较好的2015年的地表覆盖分类,与影像中的地物类别信息基本一致。GLC_FCS30在研究区内也取得较好的地表覆盖制图结果,但是相比于本研究的分类结果,明显存在将水体错分为草地、湿地,以及将耕地错分人工地表的情况。FROM-GLC 2015在人工地表与水体上取得了非常好的分类制图结果,但是存在大量将耕地错分为林地与草地的情况。总体而言,2015年的地表覆盖分类在研究区内优于GLC_FCS30以及FROM-GLC 2015。

表 1 太湖流域 1990—2015 年年地表覆盖分类精度评价结果

Tab.1 Accuracy assessment results of the Taihu Basin from 1990 to 2015

年份	1990	1995	2000	2005	2010	2015
OA/%	89.87	90.20	91.58	89.50	91.52	90.12
Kappa	0.867 4	0.871 7	0.890 7	0.864 6	0.890 3	0.871 7
样本数量	1 570	1 520	1 401	1 571	1 828	1 973

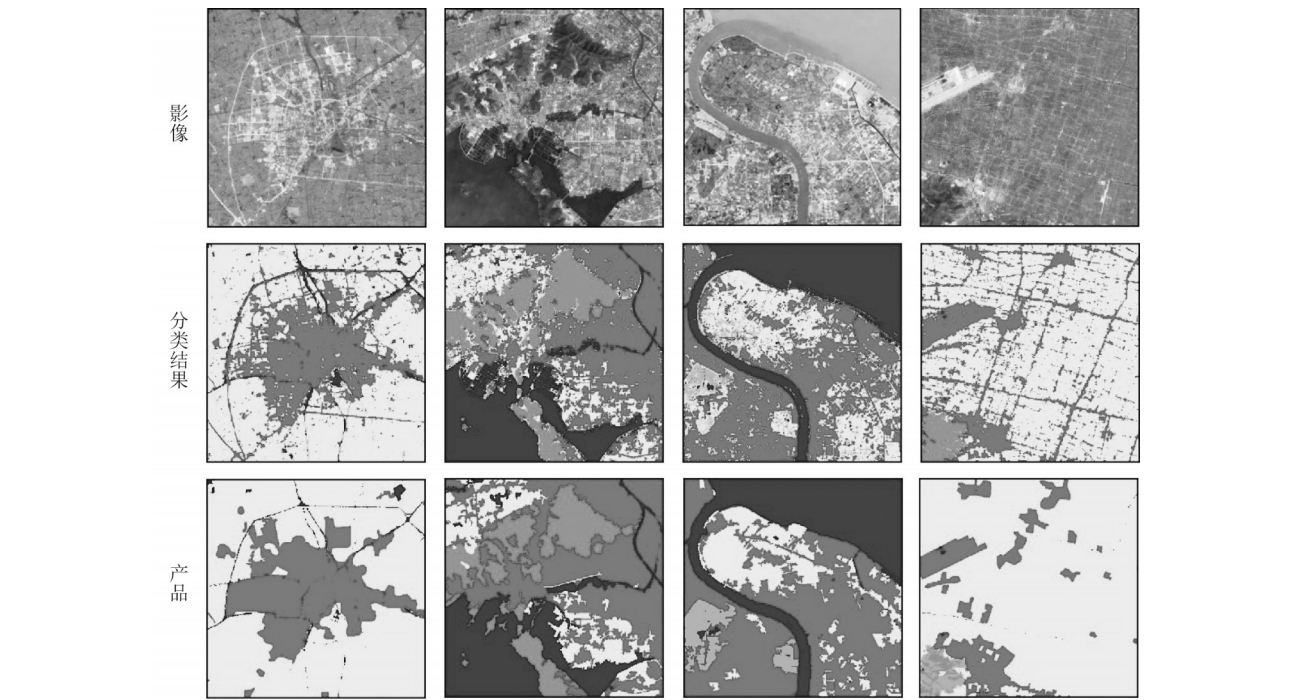


图 6 2000 年地表覆盖分类结果与 GlobeLand30 V2000 的对比

Fig.6 Comparison of land cover classification results in 2000 with GlobeLand30 V2000

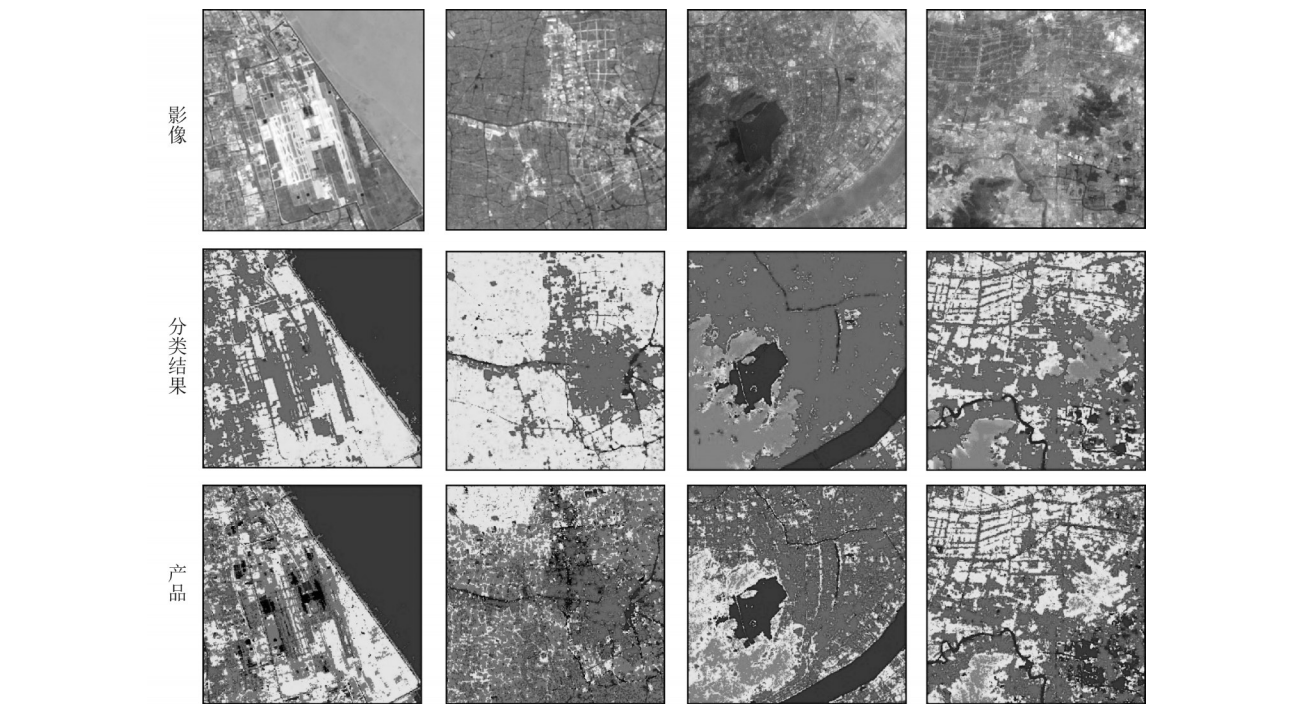


图 7 2010 年地表覆盖分类结果与 FROM-GLC 2010 的对比

Fig.7 Comparison of land cover classification results in 2010 with FROM-GLC 2010

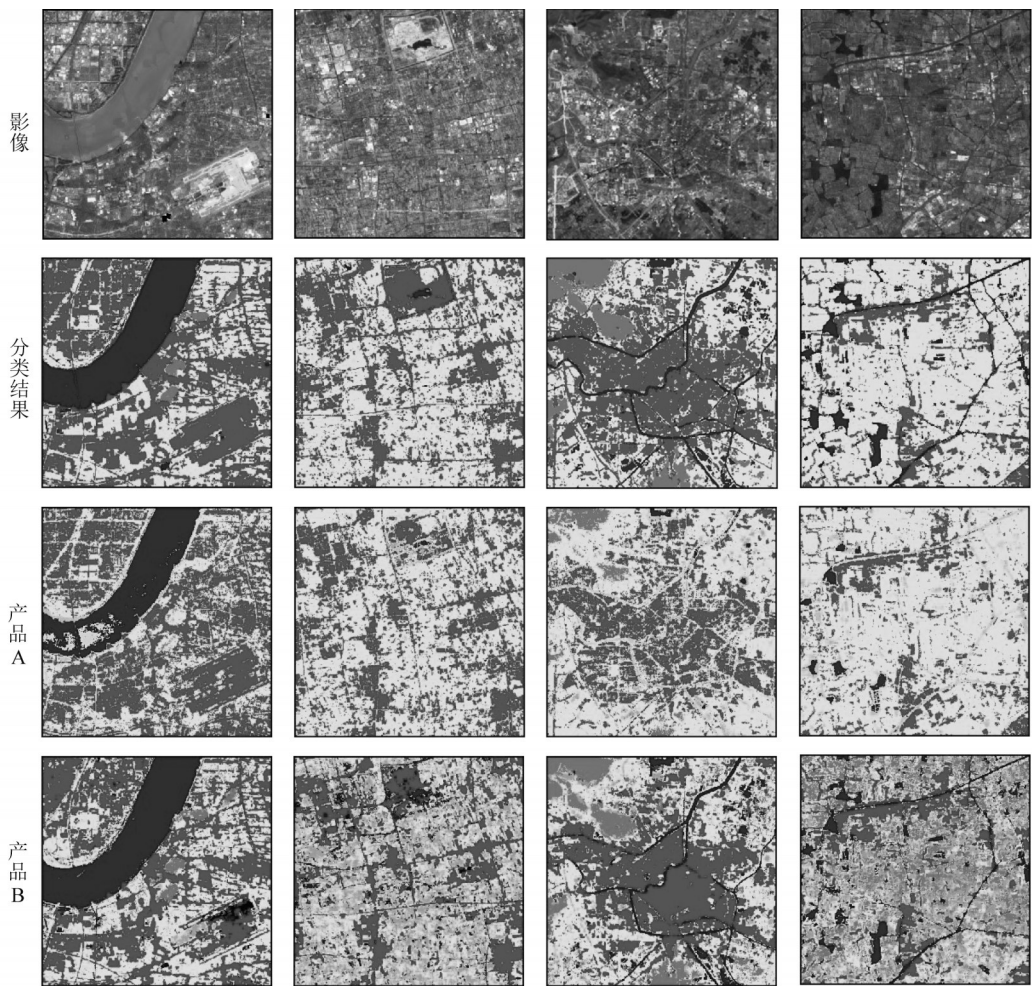


图8 2015年地表覆盖分类结果与GLC_FCS30(产品A)以及FROM-GLC 2015(产品B)的对比

Fig.8 Comparison of land cover classification results in 2015 with GLC_FCS30 and FROM-GLC 2015

4 结论

针对历史时期地表覆盖分类中对训练样本数量与质量的需求,研究提出了一种几何与属性约束下的无监督样本迁移模型和分类框架,得到的太湖流域1990—2015年的多期地表覆盖分类结果精度均优于89%。因此,无监督样本迁移方法充分利用了已有土地覆盖产品的几何约束和遥感影像的光谱特征,从局部斑块尺度和全局样本分布上对地表覆盖产品中隐含的信息进行了优选,可以形成高质量的训练样本集,在多分类器集成系统支持下获得高精度多时相土地覆盖分类结果,是一种轻量级、可靠、快速的样本迁移模型。

未来将进一步改进时间序列变化检测方法,降低样本在时序传递造成的误差,为长时间序列地表覆盖动态制图和地理环境时空感知提供可靠的信息支持。

作者贡献声明:

杜培军:负责论文总体设计,提出需求与思路,完成论文的前言与结论部分,统筹论文写作与修改。

林聪:负责方法的代码编写与实现,完成论文方法与试验部分的写作。

陈宇:负责论文初稿的整体修改,负责论文全过程的格式、排版,完成了研究区介绍部分。

王欣:协助完成了方法的代码编写与实现,修改了论文的方法与试验部分。

张伟:负责论文中数据的预处理部分。

郭山川:在论文完稿过程中提出了大量的修改建议。

参考文献:

- [1] 张兵. 遥感大数据时代与智能信息提取[J]. 武汉大学学报(信息科学版), 2018, 43(12): 108.
ZHANG Bing. Remotely sensed big data era and intelligent information extraction [J]. Geomatics and Information Science of Wuhan University, 2018, 43(12): 108.
- [2] 李德仁. 论21世纪遥感与GIS的发展[J]. 武汉大学学报(信

- 息科学版), 2003(2): 127.
- LI Deren. Towards the development of remote sensing and GIS in the 21st century [J]. *Geomatics and Information Science of Wuhan University*, 2003(2): 127.
- [3] 李德仁, 童庆禧, 李荣兴, 等. 高分辨率对地观测的若干前沿科学问题[J]. *中国科学: 地球科学*, 2012 (6): 15.
- LI Deren, TONG Qingxi, LI Rongxing, *et al.* Current issues in high-resolution earth observation technology [J]. *Science China Earth Sciences*, 2012 (6): 15.
- [4] 李德仁, 王密, 沈欣, 等. 从对地观测卫星到对地观测脑[J]. *武汉大学学报(信息科学版)*, 2017, 42: 143.
- LI Deren, WANG Mi, SHEN Xin, *et al.* From earth observation satellite to earth observation brain [J]. *Geomatics and Information Science of Wuhan University*, 2017, 42: 143.
- [5] 李德仁, 张良培, 夏桂松. 遥感大数据自动分析与数据挖掘[J]. *测绘学报*, 2014, 43(12): 1211.
- LI Deren, ZHANG Liangpei, XIA Guisong. Automatic analysis and mining of remote sensing big data [J]. *Journal of Geodesy and Geoinformation Science*, 2014, 43(12): 1211.
- [6] FOLEY J A. Global consequences of land use [J]. *Science*, 2005, 309(5734): 570.
- [7] 陈军, 陈晋, 宫鹏, 等. 全球地表覆盖高分辨率遥感制图[J]. *地理信息世界*, 2011, 9(2): 12.
- CHEN Jun, CHEN Jin, GONG Peng, *et al.* Higher resolution global land cover mapping [J]. *Geomatics World*, 2011, 9 (2): 12.
- [8] TOWNSHEND J R G. Land cover [J]. *International Journal of Remote Sensing*, 2007, 13(6/7): 1319.
- [9] WULDER M A, COOPS N C, ROY D P, *et al.* Land cover 2.0 [J]. *International Journal of Remote Sensing*, 2018, 39 (12): 4254.
- [10] DU P, BAI X, TAN K, *et al.* Advances of four machine learning methods for spatial data handling: a review [J]. *Journal of Geovisualization and Spatial Analysis*, 2020, 4(1): 13.
- [11] LIU H, GONG P, WANG J, *et al.* Annual dynamics of global land cover and its long-term changes from 1982 to 2015 [J]. *Earth System Science Data*, 2020, 12(2): 1217.
- [12] LOVELAND T R, REED B C, BROWN J F, *et al.* Development of a global land cover characteristics database and IGBP discover from 1 km AVHRR data [J]. *International Journal of Remote Sensing*, 2010, 21(6-7): 1303.
- [13] HANSEN M C, DEFRIES R S, TOWNSHEND J R G, *et al.* Global land cover classification at 1 km spatial resolution using a classification tree approach [J]. *International Journal of Remote Sensing*, 2010, 21(6-7): 1331.
- [14] BARTHOLOMÉ E, BELWARD A S. GLC2000: a new approach to global land cover mapping from earth observation data [J]. *International Journal of Remote Sensing*, 2007, 26 (9): 1959.
- [15] FRIEDL M A, SULLA-MENASHE D, TAN B, *et al.* MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets [J]. *Remote Sensing of Environment*, 2010, 114(1): 168.
- [16] SULLA-MENASHE D, GRAY J M, ABERCROMBIE S P, *et al.* Hierarchical mapping of annual global land cover 2001 to present: the MODIS collection 6 land cover product [J]. *Remote Sensing of Environment*, 2019, 222: 183.
- [17] GONG P, WANG J, YU L, *et al.* Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data [J]. *International Journal of Remote Sensing*, 2012, 34(7): 2607.
- [18] GONG P, LIU H, ZHANG M, *et al.* Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017 [J]. *Science Bulletin*, 2019, 64(6): 370.
- [19] CHEN J, BAN Y, LI S. China: open access to earth land-cover map [J]. *Nature*, 2015, 514(7523): 434.
- [20] CHEN J, CHEN J, LIAO A, *et al.* Global land cover mapping at 30m resolution: a POK-based operational approach [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 103: 7.
- [21] ZHANG X, LIU L, CHEN X, *et al.* GLC_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery [J]. *Earth System Science Data*, 2021, 13: 2753.
- [22] 庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. *软件学报*, 2015, 26(1): 26.
- ZHUANG Fuzhen, LUO Ping, HE Qing, *et al.* Survey on transfer learning research [J]. *Journal of Software*, 2015, 26 (1): 26.
- [23] Persello C, Bruzzone L. Active learning for domain adaptation in the supervised classification of remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(11): 4468.
- [24] MATASCI G, VOLPI M, KANEVSKI M, *et al.* Semisupervised transfer component analysis for domain adaptation in remote sensing image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53 (7): 3550.
- [25] SAMAT A, GAMBA P, ABUDUWAILI J, *et al.* Geodesic flow kernel support vector machine for hyperspectral image classification by unsupervised subspace feature transfer [J]. *Remote Sensing*, 2016, 8(3): 234.
- [26] LI X, ZHANG L, DU B, *et al.* Iterative reweighting heterogeneous transfer learning framework for supervised remote sensing image classification [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(5): 2022.
- [27] YANG H L, CRAWFORD M M. Spectral and spatial proximity-based manifold alignment for multitemporal hyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(1): 51.
- [28] TUIA D, MARCOS D, CAMPS V G. Multi-temporal and multi-source remote sensing image classification by nonlinear

- relative normalization [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 120: 1.
- [29] TUIA D, VOLPI M, TROLLIET M, *et al.* Semisupervised manifold alignment of multimodal remote sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(12): 7708.
- [30] TUIA D, PERSELLO C, BRUZZONE L. Domain adaptation for the classification of remote sensing data: an overview of recent advances [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2016, 4(2): 41.
- [31] 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.
- [32] PERSELLO C, BRUZZONE L. Kernel-based domain-invariant feature selection in hyperspectral images for transfer learning [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(5): 2615.
- [33] PETITJEAN F, KETTERLIN A, GANÇARSKI P. A global averaging method for dynamic time warping, with applications to clustering [J]. *Pattern Recognition*, 2011, 44(3): 678.
- [34] SUN H, LIU S, ZHOU S, *et al.* Unsupervised cross-view semantic transfer for remote sensing image classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(1): 13.
- [35] BRUZZONE L, CHI M, MARCONCINI M. A novel transductive SVM for semisupervised classification of remote-sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(11): 3363.
- [36] CHI M, BRUZZONE L. Semisupervised classification of hyperspectral images by SVMs optimized in the Primal [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(6): 1870.
- [37] DEMIR B, BOVOLO F, BRUZZONE L. Detection of land-cover transitions in multitemporal remote sensing images with active-learning-based compound classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(5): 1930.
- [38] TUIA D, PASOLLI E, EMERY W J. Using active learning to adapt remote sensing image classifiers [J]. *Remote Sensing of Environment*, 2011, 115(9): 2232.
- [39] CHEN X, CHEN J, SHI Y, *et al.* An automated approach for updating land cover maps based on integrated change detection and classification methods [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012, 71: 86.
- [40] DEMIR B, BOVOLO F, BRUZZONE L. Updating land-cover maps by classification of image time series: a novel change-detection-driven transfer learning approach [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(1): 300.
- [41] YU W, ZHOU W, QIAN Y, *et al.* A new approach for land cover classification and change analysis: integrating backdating and an object-based method [J]. *Remote Sensing of Environment*, 2016, 177: 37.
- [42] XU Y, YU L, ZHAO F R, *et al.* Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa [J]. *Remote Sensing of Environment*, 2018, 218: 13.
- [43] CROWSON M, HAGENSIIEKER R, WASKE B. Mapping land cover change in northern Brazil with limited training data [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 78: 202.
- [44] WU T, LUO J, ZHOU Y N, *et al.* Geo-object-based land cover map update for high-spatial-resolution remote sensing images via change detection and label transfer [J]. *Remote Sensing*, 2020, 12(1): 174.
- [45] 黄亚博, 廖顺宝. 多源数据的土地覆被样本自动提取[J]. *遥感学报*, 2017, 21(5): 757.
- HUANG Yabo, LIAO Shunbao. Automatic collection for land cover classification based on multisource datasets [J]. *Journal of Remote Sensing*, 2017, 21(5): 757.
- [46] XU G, CHEN B. Generating a series of land covers by assimilating the existing land cover maps [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 147: 206.
- [47] ZHUO L, SHI Q, TAO H, *et al.* An improved temporal mixture analysis unmixing method for estimating impervious surface area based on MODIS and DMSP-OLS data [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 142: 64.
- [48] PARIS C, BRUZZONE L, FERNANDEZ P D. A novel approach to the unsupervised update of land-cover maps by classification of time series of multispectral images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(7): 4259.
- [49] 余柏蒴, 王丛笑, 宫文康, 等. 夜间灯光遥感与城市问题研究: 数据、方法、应用和展望 [J]. *遥感学报*, 2021, 25(1): 342.
- YU Bailang, WANG Congxiao, GONG Wenkang, *et al.* Nighttime light remote sensing and urban studies: data, methods, applications, and prospects [J]. *Journal of Geodesy and Geoinformation Science*, 2021, 25(1): 342.
- [50] XIE H, TONG X, MENG W, *et al.* A multilevel stratified spatial sampling approach for the quality assessment of remote-sensing-derived products [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(10): 4699.
- [51] TRIANNI G, ANGIULI E. Urban mapping in landsat images based on normalized difference spectral vector [J]. *IEEE Geoscience and Remote Sensing Letters*, 2013, 11(3): 661.