

土地利用分类粒子群优化概率神经网络半监督算法

王春阳^{1,2} 汤子梦¹ 吴喜芳¹ 李长春¹ 张合兵¹

(1. 河南理工大学测绘与国土信息工程学院, 焦作 454003; 2. 河南理工大学计算机科学与技术学院, 焦作 454003)

摘要: 针对以往土地利用监测大都采用监督分类算法,成本较高、错分漏分严重且受人为因素影响较大的问题,提出了一种粒子群优化概率神经网络的半监督分类算法。该算法通过粒子群优化算法优化分类器的参数,提高分类器的精度,运用香农熵选择高置信度的样本扩展初始训练样本集,将大量未标记的样本扩展到训练样本集中,减少了初始标签样本的数量,节约了成本,并与随机森林法、最大似然法、概率神经网络算法进行对比分析,总体精度较其他算法提高了1.25~6.57个百分点,Kappa系数达到0.8以上。对新乡市1996年、2004年、2013年、2020年的遥感影像进行土地分类,结果表明1996—2020年间新乡市的建设用地以中部地区新乡县为中心不断扩张,耕地面积也在不断增加,其他用地面积不断减少,沿黄绿地面积不断增加;土地流转方面耕地转建设用地最为明显,本研究为新乡市进一步合理开发土地资源提供了理论依据。

关键词: 土地利用分类; 半监督算法; 粒子群优化; 概率神经网络; 香农熵; 转移矩阵

中图分类号: TP79 文献标识码: A 文章编号: 1000-1298(2022)02-0167-10 OSID: 

Semi-supervised Land Use Classification Based on Particle Swarm Optimization Probabilistic Neural Network

WANG Chunyang^{1,2} TANG Zimeng¹ WU Xifang¹ LI Changchun¹ ZHANG Hebing¹

(1. School of Surveying and Mapping and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China
2. College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China)

Abstract: Aiming at the problem that most of the land use monitoring in the past uses supervised classification algorithms, which have high costs, wrong leakage points, and greatly affected by human factors, a semi-supervised classification algorithm was proposed for particle swarm optimization probability neural networks, which improved the classification accuracy. The algorithm optimized the parameters of the classifier through the particle swarm optimization algorithm, improved the accuracy of the classifier, and Shannon entropy was used to select high-confidence samples to expand the initial training sample set, a large number of unlabeled samples were expanded to the training sample set, the number of initial label samples were reduced, costs were saved, and it was compared and analyzed with random forest, maximum likelihood method, and probabilistic neural network algorithm, the classification accuracy was improved by 1.25 ~ 6.57 percentage points compared with that of other algorithms, and the Kappa coefficient reached more than 0.8. Through the land classification of the remote sensing images of Xinxiang City in 1996, 2004, 2013 and 2020, the results showed that the construction land of Xinxiang City from 1996 to 2020 was continuously expanded in Xinxiang County in the central region, and the cultivated land area was also increased, and the area of other land used was decreased, and the area along the Yellow green area was increased; the land circulation was the most obvious for the conversion of cultivated land to construction land. The research results provided a certain reference for the further rational development of land resources in Xinxiang City.

Key words: land use classification; semi-supervised algorithm; particle swarm optimization; probabilistic neural network; Shannon entropy; transfer matrix

收稿日期: 2021-11-07 修回日期: 2021-11-27

基金项目: 国家自然科学基金项目(41871333)、河南省科技攻关项目(222102110038、222102210131)和河南理工大学博士基金项目(B2021-19)

作者简介: 王春阳(1979—),男,讲师,博士,主要从事机器学习遥感应用研究,E-mail: wcy@hpu.edu.cn

通信作者: 吴喜芳(1987—),女,讲师,博士,主要从事遥感数据处理及应用研究,E-mail: wxf2020@hpu.edu.cn

0 引言

土地资源是人们赖以生存和发展的基础,随着全球环境气候不断变化,温室效应日益加剧,人类发展过程中对土地的利用以及引起的土地覆盖变化被认为是全球环境变化的重要组成部分^[1]。土地利用分类图是研究土地变化的前提,为生态服务和制定土地政策提供了依据^[2]。伴随着计算机技术的发展,土地利用/覆盖变化的研究方法日新月异,传统分类方法成本高,工程量大,且效果不佳^[3],机器学习如人工神经网络^[4]、支持向量机^[5]、决策树^[6]等是通过学习来获取经验并改进自身性能的方法,受人为因素影响较小,分类精度较高,被广泛应用在遥感影像的分类中^[7]。

机器学习的发展可以追溯到 1990 年,WARREN 等^[8]首次提出了神经网络结构模型,为机器学习的提出和发展奠定了基础,随后 MITCHELL 等^[9]提出了机器学习的概念。运用机器学习算法对遥感影像进行分类,既提高了分类精度,又节约了成本,因此成为学者们研究的热点^[10-13]。在许多数据挖掘和机器学习的任务中,存在大量未标记的样本和有限的标记样本^[14],马永建等^[15]将植被指数 NDVI 值和纹理特征灰度共生矩阵能量值作为特征波段,基于 U-Net 模型实现了对 GF-1 号遥感影像荒漠地区耕地的提取,相比于传统方法,分类精度得到了提高;王笑影等^[16]基于 R 语言和随机森林对 Landsat 8 OLI 影像进行土地覆盖的监督分类研究,该方法简单实用但仅用于中小区域土地的快速分类的情景;陈磊士等^[17]以 Landsat 8 和 Sentinel-1A 融合后的数据为实验数据,分别采用卷积神经网络和 BP 神经网络算法分类,结果发现卷积神经网络具有较好的分类结果;赵亚杰等^[18]运用随机森林算法对 Sentinel-2 和 Landsat 影像进行土地利用分类,实现了对地物的精细提取;陈义菁等^[19]运用支持向量机算法实现了对 Landsat 影像的分类,对探究影响土地利用分类的变化提供了依据。大多数学者提出的机器学习算法虽然有较好的分类精度,但样本需求量大、成本高,受人为因素影响干扰较大。半监督学习具有仅使用少量标记样本就可以达到分类效果的特点,成为近年来研究的热点话题。半监督机器学习算法适用于小样本类型,它是同时使用标记和未标记的样本来拟合模型。在某些情况下,添加未标记的数据可能使模型的准确性提高,在其他情况下可能会使准确率降低^[20]。

为了最大化有限使用在半监督机器学习算法中可用的监督数据,本文提出基于粒子群算法优化的

半监督概率神经网络框架,将其运用在中分辨率的 Landsat 影像中,获得新乡市土地利用分类图。以新乡市 1996 年、2004 年、2013 年、2020 年 4 期遥感影像为例,解译出不同时期的土地利用动态,通过监测近 14 年内的土地利用/覆盖变化,分析新乡市 1996—2020 年的土地利用转移矩阵,为新乡市人地关系研究和土地资源管理提供参考^[21],以期实现新乡市各种土地利用资源与经济增长、社会稳定、生态和谐共同发展。

1 研究方法和数据来源

1.1 研究区概况

新乡市地理位置示意图如图 1 所示,新乡市位于河南省的北部,地处北纬 34°53′~35°50′,东经 113°23′~114°59′,南临黄河,与郑州市、开封市隔河相望;北依太行山,与鹤壁市、安阳市毗邻;西连焦作市,与晋东南接壤;东接濮阳市,与鲁西相连,是国家重要的综合交通枢纽,也是豫北的经济、交通、教育、商贸物流中心;年平均气温 14℃^[22],7 月最热,平均气温 27.3℃;1 月最冷,平均气温 0.2℃,年平均降水量 656.3 mm,6—9 月降水量最大,且多暴雨。新乡市自然条件和土地耕作条件好,是中国粮棉生产区、国家优质小麦生产基地和河南省畜牧生产加工基地,是农业大市^[23]。截至 2020 年 11 月,新乡市的常住人口 625.29 万人,全市的农林牧业产值、粮食产量、居民消费水平大幅度提高。

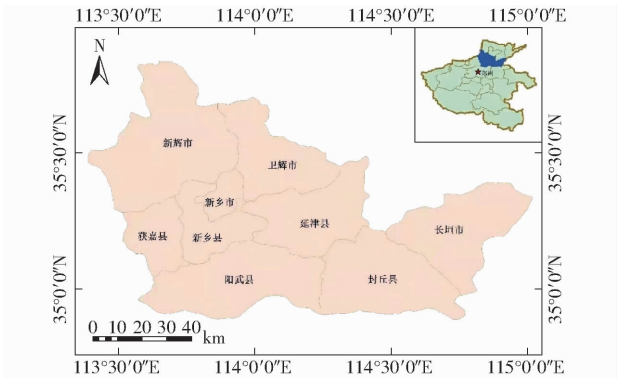


图 1 新乡市地理位置示意图

Fig. 1 Geographical location map of Xinxiang City

1.2 数据来源与预处理

根据新乡市的地理位置特点及获取遥感影像需满足的实验需求,本研究所使用的 4 期影像数据均来源于美国地质调查局 (USGS) 网站 (<https://earthexplorer.usgs.gov/>),由于受到 Landsat 数据影像品质的影响,无法下载到同一个月份的影像。为了保证实验结果的可靠性,选取了时间上相对较接近的影像数据(6、7、8 月)。选取 1996 年、2004 年、2013 年、2020 年的 6 月或 7 月或 8 月的影像,所选

的影像云量均在 5% 以下,该时期的影像品质较高便于土地利用/覆盖变化的研究。新乡市面积较大,覆盖研究区需要由行列号为 123 - 36、124 - 35、124 - 36 的 3 景影像拼接而成,空间分辨率为 30 m,本次研究使用的数据分别为:1996 年 7 月获取的 Landsat TM 数据,2004 年 6 月获取的 Landsat TM 数据,2013 年 7 月获取的 Landsat OLI 数据和 2020 年 8 月获取的 Landsat OLI 数据。

首先结合《全国遥感监测覆盖分类体系》建立的参考框架,综合考虑新乡市土地利用类型的特点,将研究区土地利用/覆盖类型确定为耕地、林地、水体、建设用地、草地、其他用地 6 类^[24],遥感影像的解译如表 1 所示。具体的实验步骤为:首先使用 ENVI 5.3 软件对遥感影像进行辐射定标、大气校正、裁剪、拼接等预处理工作^[25],使遥感影像的图像更加清晰、更加突出物体的标识;其次基于预处理后的遥感影像使用 ENVI 5.3 软件进行 6 类样本点的选取,将样本点和原始影像均转换为算法所需的 .mat 格式;然后基于 Matlab 2019a 平台运用粒子群优化概率神经网络的半监督算法、概率神经网络半监督算法进行分类,进行对比的随机森林法和最大似然法,则采用 ENVI 5.3 软件完成分类,并用 ArcGIS 软件制作新乡市的土地利用分类图;最后对新乡市 1996—2020 年土地利用的时空变化进行分析。实验流程如图 2 所示。

表 1 新乡市土地利用分类及遥感解译特征	
Tab.1 Land use classification and remote sensing interpretation signs in Xinxiang City	
土地利用类别	影像特征
建设用地	以亮白色为主,影像分布比较零散
耕地	绿色、暗绿色为主,形状较为规则,分布比较集中
林地	分布在起伏较大的山体附近,形状比较粗糙,边界模糊
草地	绿色、暗绿色为主,零星分布
水体	几何特征明显,形状局部平直,深蓝色、绿色为基调(黄河显示为黄色调)
其他用地	分布在起伏较大的地区,暗绿色为主,零星分布

2 实验方法

针对遥感影像分类中需要大量标记样本和分类精度较低的问题,提出了一种以半监督为主体,将粒子群优化与概率神经网络框架相融合的一种分类算法。首先,利用粒子群优化算法生成最优参数;其次,利用上述最优参数的概率神经网络分类器,根据标注的训练样本计算每个像素的类别概率,预测每个像素的类别;然后利用香农熵筛选高置信度的无

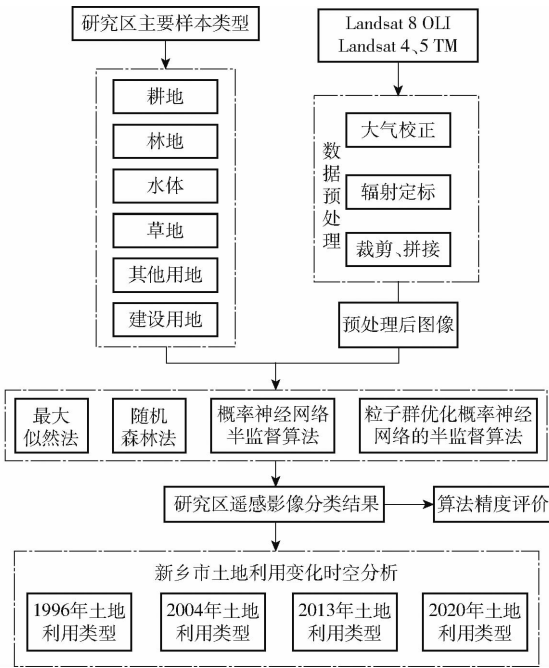


图 2 实验流程图
Fig.2 Experimental flow chart

标记样本,并将其添加到原始训练集中进行新一轮的预测,自我训练使用模型自己对未标记数据的预测结果,将其添加到已标记的数据集中,选择置信度高的预测样本将其添加到已标记的数据集中,不断地重新训练这个模型,直到没有更多可信的预测结果为止。这种算法可以满足训练样本较少的情况,将大量无标记的样本转换为有标记的样本,同时使用标记样本和未标记样本数据来拟合出能够真实反映类别信息的分类器,分类成本较低,分类精度不断提高。本研究首先选取样本总数的 5% 作为训练样本加入到训练样本集中,依次增加样本数的 10% 进行迭代,算法流程如图 3 所示。

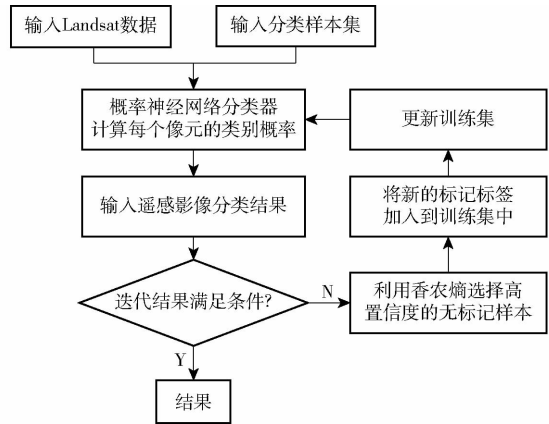


图 3 实验算法流程图
Fig.3 Flow chart of experimental algorithm

2.1 粒子群优化的概率神经网络

2.1.1 概率神经网络分类器

概率神经网络 (Probabilistic neural network,

PNN)是 SPECHT^[26]在 1988 年提出的一种神经网络模型。这种模型是以指数替代神经网络常用的 S 型激活函数,进而构造出能够计算非线性判别边界的概率神经网络,该判定界面接近于贝叶斯的最佳判定面。概率神经网络主要用于分类和识别领域,其中在分类方面应用最广泛,这种网络模型已经被广泛应用在非线性滤波、模式分类、联想记忆和概率密度估计中^[27]。这种模型分类错误率较低,风险比较小,但将其应用到 Landsat 遥感影像方面的研究较少。本研究将概率神经网络模型应用在遥感影像中,并对遥感影像数据中的地物信息进行分类,运用粒子群优化算法选取最优分类参数,使分类精度达到理想的结果。

概率神经网络的结构模型大致分为输入层、样本层、求和层和竞争层。在输入层中,网络计算输入向量与所有训练样本向量之间的距离;样本层里神经元的数量是训练样本的数量;求和层是将样本层的输出按类相加,相当于 n 个加法器;竞争层的神经元数量为 1,最后的结果也是由竞争层输出,输出的结果只有 1 个 1,其他的结果都是 0。

2.1.2 粒子群优化算法

粒子群优化算法具有快捷、收敛速度较快等优势^[28],但作为一种仿生算法,粒子群优化算法数学理论基础尚不完善,因而需要在理论上进一步研究和讨论。

在粒子群算法中,每一个粒子群代表一个可能的解,整个种群通过竞争和协作实现多维空间中对最优解的探索^[29]。在 D 维空间中,每个粒子被视作一个解,由当前位置 $X = (x_{i1}, x_{i2}, \cdots, x_{iD})$ 和当前速度 $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$ 组成,对于全局粒子群算法,粒子群 X_i 依据其历史最优值 $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$ 和全局最优值 $P_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$ 迭代更新。每个粒子更新位置和速度可以表示为

$$v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{id}^t) \quad (1)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^t \quad (2)$$

- 式中 t ——当前进化的次数
 $c_1、c_2$ ——正加速度常数
 $r_1、r_2$ —— $[0,1]$ 之间的随机数
 ω ——惯性权重
 p_{gd}^t ——迭代 t 次后全局最优值
 p_{id}^t ——迭代 t 次后个体极值
 $x_{id}^t、x_{id}^{t+1}$ ——迭代 $t、t+1$ 次后粒子位置
 $v_{id}^t、v_{id}^{t+1}$ ——迭代 $t、t+1$ 次后粒子速度

2.1.3 粒子群优化概率神经网络过程

概率神经网络的相关结构和对应神经元的数目由训练样本数据的数量和类别确定,网络模型的性能

能取决于网络的参数,而粒子群算法计算简单、收敛速度快,通过迭代获得最优解,可有效选取概率神经网络的最优参数。首先随机初始化粒子的位置和速度,构建网络,通过适应度函数 $f(x_i)$ 计算各个粒子的适应度,从而选取出个体的局部最优位置和全局最优位置,然后进行迭代,将得出的每个粒子的适应度与局部最佳位置的适应度相比,若当前适应度更大,则用当前粒子的位置作为全局最佳位置,不断进行迭代,更新个体的极值和群体极值。若满足条件则迭代终止,输出最优解;若未满足条件,则不断重复上述步骤直到满足条件为止。本文使用粒子群优化算法对概率神经网络的 spread 参数进行优化,利用粒子群的粒子间协同合作与知识共享对 spread 速度和位置进行准确定位,找到最佳的输入值,有效提高了概率神经网络分类模型的性能,从而最终能够得到最佳的土地分类结果,流程如图 4 所示。

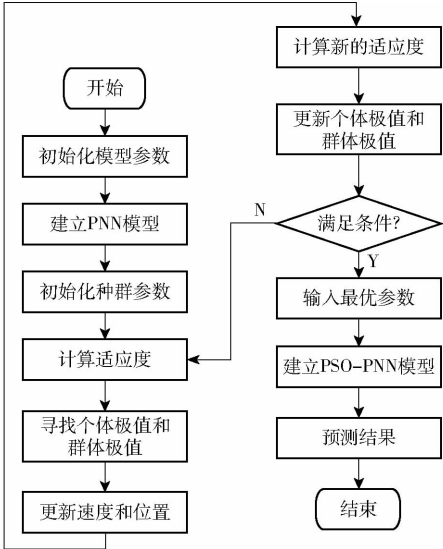


图 4 粒子群优化概率神经网络流程图
Fig.4 Flow chart of probabilistic neural network for particle swarm optimization

2.2 自训练半监督机器学习算法

YAROSKY 首先提出了自训练半监督算法^[30],该算法是先用有标记的样本去训练一个分类器,本文所采用的分类器是概率神经网络,然后利用分类器对无标记的样本进行概率预测,根据获得的置信度,选择一些高置信度且无标记的样本加入到原始训练集中作为新的训练样本集,重复这个过程直到完成给定的迭代次数。本文算法在多次迭代的分类精度中选择了最大精度,避免迭代次数过多出现过度拟合现象。若有原始已标记样本集 $L = \{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$, 无标记样本集 $U = \{(x_{m+1}), (x_{m+2}), \cdots, (x_n)\}$, 迭代的次数为 K , 算法流程为:

- (1)利用已标记样本集 L 训练得到模型 p 。
- (2)利用分类器 p 对无标记样本集 U 进行标记,选择置信度较高的 m 个样本,记作 L' ,并将其从
无标记样本中删除。
- (3)利用 $L \cup L'$ 对选择的分类器进行反复训练。
- (4)重复步骤(2)和步骤(3),直到迭代条件满足时为止。

3 实验结果分析

3.1 分类算法结果分析

选取新乡市 1996 年、2004 年、2013 年和 2020 年的 Landsat 影像作为实验数据,运用粒子群优化概率神经网络的半监督算法分类时,选取 5% 的训练样本,迭代 3 次每次增加总样本的 10% 扩充训练样本集,最后训练样本集总数为总样本的 35%,剩下的用来检验分类精度。新乡市各年份各类别输入的样本总数和训练样本总数如表 2 ~ 5 所示。获得各年份的土地利用分类图,分类总体精度达到 97%,Kappa 系数均达到 0.8。为了对比算法的分类精度,与随机森林、最大似然法和概率神经网络半监督算法进行了对比实验,各算法的分类结果如图 5 所示。

表 2 新乡市 1996 年总样本、训练样本和验证样本数

Tab.2 Total samples, training samples and verified samples in Xinxiang City in 1996

类别标签	土地利用类型	样本总数	训练样本数	验证样本数
1	水体	13 630	4 771	8 859
2	耕地	105 732	37 006	68 726
3	林地	39 781	13 923	25 858
4	草地	2 296	804	1 492
5	建设用地	53 785	18 825	34 960
6	其他用地	12 001	4 200	7 801

表 3 新乡市 2004 年总样本、训练样本和验证样本数

Tab.3 Total samples, training samples and verified samples in Xinxiang City in 2004

类别标签	土地利用类型	样本总数	训练样本数	验证样本数
1	水体	36 598	12 809	23 789
2	耕地	268 042	93 815	174 227
3	林地	59 365	20 778	38 587
4	草地	4 110	1 439	2 671
5	建设用地	149 415	52 295	97 120
6	其他用地	53 789	18 826	34 963

由图 5 可知,随机森林法和最大似然法在进行林地与其他用地的分类时,错分、漏分现象较为严重,较为明显的地方在图中已用红色区域圈出,概率神经网络算法分类效果较其他 2 种算法好。在较难区分的林地和草地中,本文算法可将 2 种地物区分

表 4 新乡市 2013 年总样本、训练样本和验证样本数

Tab.4 Total samples, training samples and verified samples in Xinxiang City in 2013

类别标签	土地利用类型	样本总数	训练样本数	验证样本数
1	水体	32 734	11 457	21 277
2	耕地	101 067	35 373	65 694
3	林地	1 526	534	992
4	草地	1 156	405	751
5	建设用地	89 796	31 429	58 367
6	其他用地	28 395	9 938	18 457

表 5 新乡市 2020 年总样本、训练样本和验证样本数

Tab.5 Total samples, training samples and verified samples in Xinxiang City in 2020

类别标签	土地利用类型	样本总数	训练样本数	验证样本数
1	水体	13 364	4 677	8 687
2	耕地	44 181	15 463	28 718
3	林地	16 267	5 693	10 574
4	草地	1 251	438	813
5	建设用地	50 382	17 634	32 748
6	其他用地	4 719	1 652	3 067

出来,而随机森林法和最大似然法将 2 种地物归为 1 种地物来划分,概率神经网络算法将 1996 年的部分林地划分为草地,如图 5b 所示,最大似然法也未能将 2020 年林地准确区分出来,如图 5o 所示;在进行水体分类时,最大似然法和概率神经网络算法未能将 2004 年的水体准确区分出来,如图 5f、5g 所示;对于耕地和建筑物分类效果不能满足要求;对其他用地进行分类时,2013 年分类错分最为明显,如图 5k、5l 所示,2 种算法将新乡市西北部的区域几乎都认为是其他用地,分类效果较差。在精度评定方面本文选取总体精度和 Kappa 系数 2 个评价指标,从表 6 可以看出,和其他 3 种算法相比,本研究提出的半监督算法精度提高了 1.25 ~ 6.57 个百分点,Kappa 系数也达到 0.8 以上,有效提高了分类总体精度,弥补了其他算法错分漏分的情况。

除此之外,本研究提出的算法因融合了半监督,所以适应训练样本少的情况,为了使实验结果对比更清晰,选择了一个小区域进行对比实验,各年份所选的区域如图 6 所示。分别选取了样本数的 5%、10%,每次迭代过程中都会新加入 5% 的训练样本到测试集中作为训练样本集,在对比分析不同样本数的分类精度时发现,当选取训练样本数为总数的 5% 时,1996 年、2004 年、2013 年、2020 年分类精度依次为 95.15%、95.93%、98.13%、95.57%;当选取样本数为 10% 时,分类精度依次为 94.08%、95.45%、97.62%、95.97%。2 种样本数量的分类精度均保持在理想状态,对比可以发现当选取的样

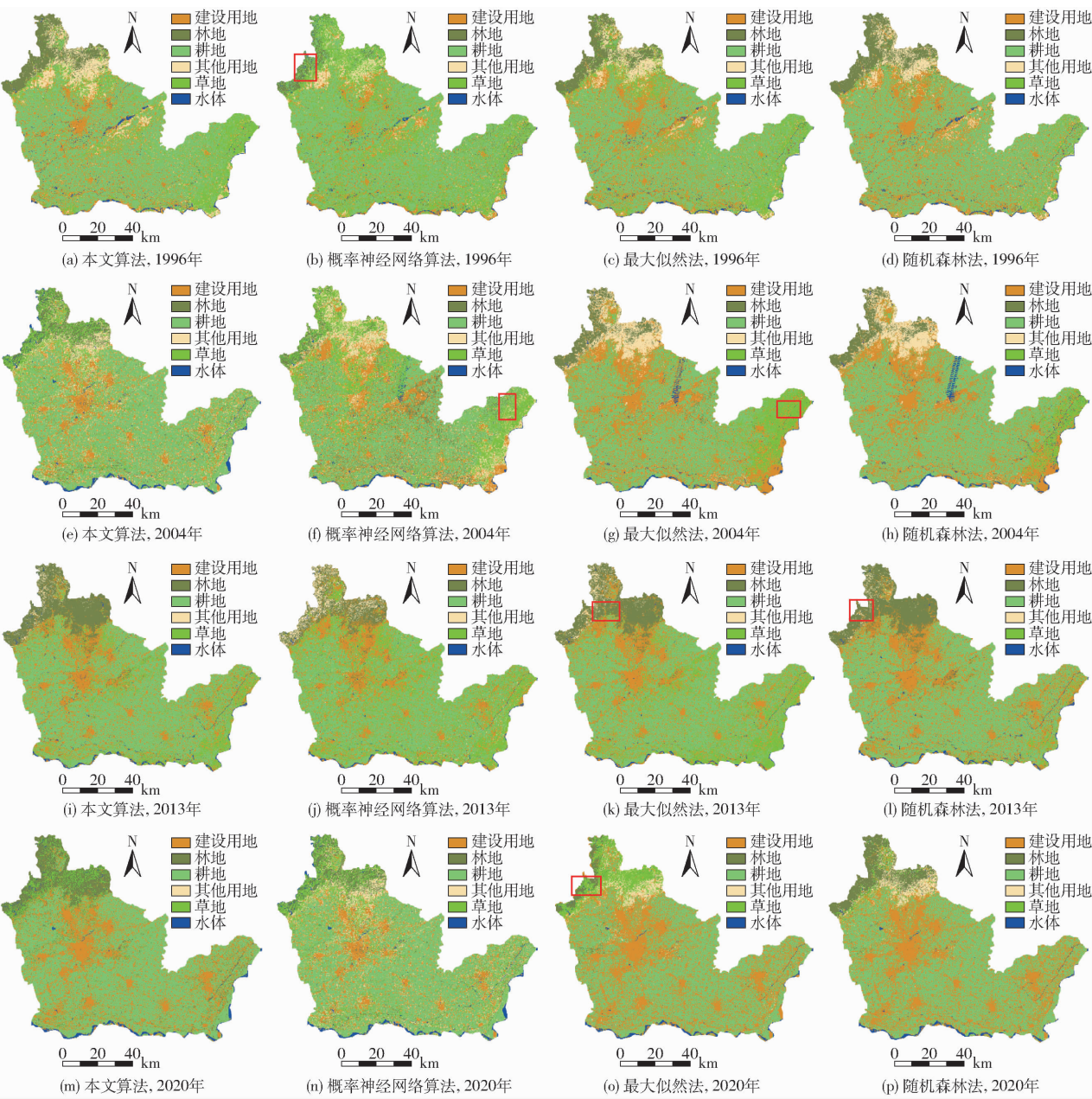


图 5 各算法土地利用分类图

Fig. 5 Land use classification diagrams of each algorithm

表 6 新乡市土地利用/覆盖变化总体精度评价

Tab.6 Overall accuracy evaluation of land use/coverage change in Xinxiang City

年份	分类方法	总体精度/%	Kappa 系数
1996 年	随机森林法	89.56	0.76
	最大似然法	90.12	0.73
	概率神经网络算法	93.65	0.75
	本文算法	95.15	0.81
2004 年	随机森林法	91.58	0.78
	最大似然法	90.36	0.76
	概率神经网络算法	92.98	0.81
	本文算法	95.93	0.86
2013 年	随机森林法	91.56	0.83
	最大似然法	92.32	0.79
	概率神经网络算法	94.72	0.84
	本文算法	98.13	0.87
2020 年	随机森林法	94.32	0.82
	最大似然法	93.68	0.81
	概率神经网络算法	93.51	0.83
	本文算法	95.57	0.87

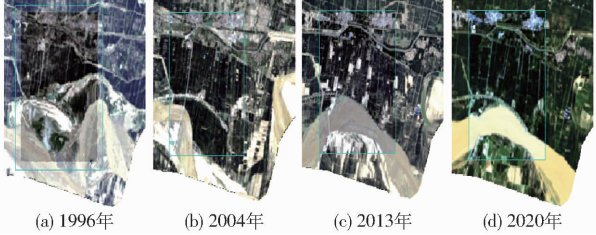


图 6 各年份研究区位置图

Fig. 6 Location map of study area for each year

本数量为 5% 时,各地物分类的精度也很高,和样本数量为 10% 在分类效果和分类精度上几乎没有差别,在较难区分的草地和林地,分类效果也较好,因此当训练样本较少时,本研究提出的算法仍然满足精度要求,各年份的土地利用分类如图 7 所示。

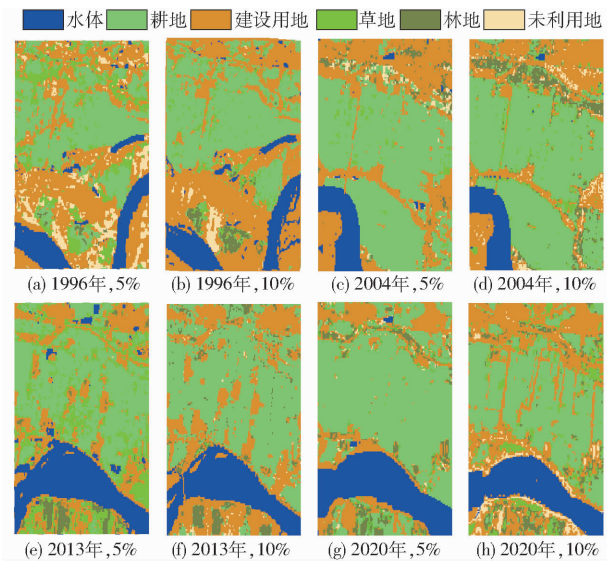


图 7 不同样本数的土地利用分类图

Fig. 7 Land use classification maps for different sample numbers

3.2 土地利用变化时空特征

3.2.1 土地利用时间变化特征

利用本文方法获得新乡市 4 年的土地利用分类图,然后利用 ArcGIS 平台的空间分析和统计功能,根据土地利用分类图,获得每一类的土地利用面积,分析土地利用变化的原因。新乡市各土地利用类型面积如表 7 所示。

从表 7 可以看出,在 1996—2020 年期间,耕地面积总体呈增加趋势,从 1996 年的 4 377. 86 km² 增加到 4 810. 95 km²,增加了 9. 9%;建设用地面积从 1996 年到 2020 年增加了近 38%;其他用地的面积相对减少,由 1996 年的 729. 40 km² 减少到 354. 16 km²,面积减少了 52%;草地面积由 1996 年的 974. 30 km²

表 7 新乡市各年份土地利用类型面积						
Tab. 7 Area statistics of various land use types in Xinxiang City						
年份	耕地	林地	草地	水体	建设 用地	其他 用地
1996 年	4 376. 93	611. 16	974. 30	134. 39	1 426. 34	729. 40
2004 年	4 209. 06	621. 39	624. 39	168. 09	2 181. 67	449. 69
2013 年	4 464. 77	865. 01	835. 69	124. 22	2 096. 03	520. 71
2020 年	4 810. 24	528. 76	465. 21	130. 70	1 963. 45	354. 16

减少到 465. 21 km²,减少了 52%;水体的面积几乎没有大的波动。由此看出,近 14 年来由于新乡市经济飞速发展,社会基础设施的不断增加、生活水平的不断改善、人口的不断增多导致对住房和工业的需求增多,从而导致建筑面积的不断扩张,其他用地面积不断减少。除此之外,由于人口数量的急剧增加,导致对粮食新的需求不断增加,耕地面积不断增加,草地面积急剧减少。

土地利用转移矩阵是定量研究土地利用类型间相互转换的数量和方向的主要方法,它能够具体反映土地利用变化的结构特征和各类型间的转移方向。以新乡市 1996 年和 2020 年的土地利用数据为基础,利用 ArcGIS 软件进行叠加分析,获得土地利用转移数据,并利用 Excel 制作转移矩阵,如表 8 所示。结果表明,新乡市土地利用的主要变化趋势为耕地和建设用地面积在增加,其中耕地转换为建设用地的面积最大。具体的流向为:草地主要流转为耕地和建设用地;耕地主要流转为建设用地和其他用地;建设用地主要流转为耕地;林地主要流转为草地;水体部分被开发为建设用地,部分作为耕地;其他用地转换为耕地、草地和建设用地。

表 8 1996—2020 年新乡市土地利用转移矩阵							
Tab. 8 Land use transfer matrix of Xinxiang City from 1996 to 2020							
1996 年	2020 年						总计
	草地	耕地	建设用地	林地	水体	其他用地	
草地	51. 30	550. 72	275. 68	32. 86	14. 86	48. 88	974. 30
耕地	66. 82	3 515. 60	657. 63	47. 84	19. 96	69. 08	4 376. 93
建设用地	42. 84	463. 53	824. 01	22. 40	38. 15	35. 41	1 426. 34
林地	80. 26	28. 61	14. 31	350. 89	11. 34	125. 75	611. 16
水体	5. 34	46. 21	39. 02	4. 67	35. 00	4. 15	134. 39
其他用地	218. 65	205. 57	152. 80	70. 10	11. 39	70. 89	729. 40
总计	465. 21	4 810. 24	1 963. 45	528. 76	130. 70	354. 16	8 252. 52

3.2.2 土地利用空间变化特征

由于耕地和建设用地是新乡市主要的 2 种土地利用类型,所以选取这 2 种土地利用类型来分析新乡市土地利用空间变化特征。1996—2020 年间,新乡市土地利用变化总体上表现为耕地的增加和建设用地的扩张。新乡市 1996—2020 年建设用地和耕

地空间分布如图 8 所示。

耕地是新乡市占比较大的土地利用类型,主要分布在地势比较平坦的东西部地区和南部的临黄河地区,西北部辉县山地较多、中部地区人口较多建筑用地分布较为密集,因此这 2 个区域耕地面积较少。总体来看,1996—2020 年新乡市耕地面积总体呈增

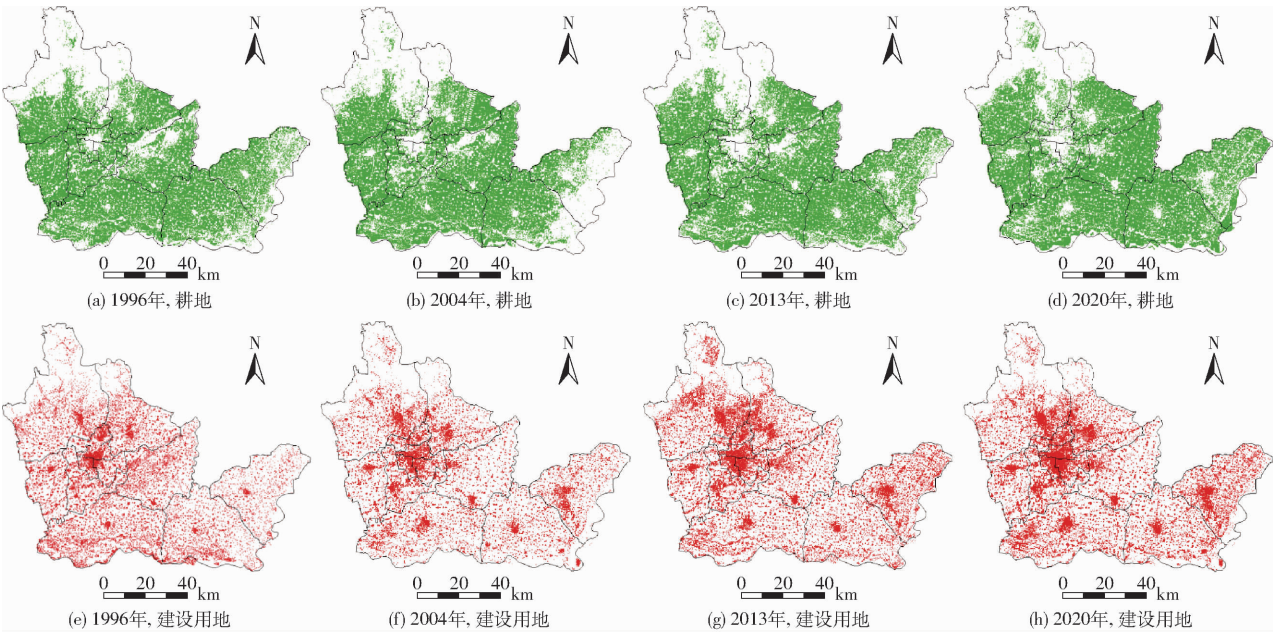


图 8 新乡市 1996—2020 年耕地、建设用地空间分布

Fig. 8 Spatial distribution maps of cultivated land and construction land in Xinxiang City from 1996 to 2020

加趋势,增加了 433.21 km²。从不同时期来看,1996—2004 年耕地面积减少,如新乡市东南部地区长垣市。2004—2020 年耕地面积有所增加,因为 2005 年新乡市进行了行政区划的调整,对农田实施了保护措施,耕地面积减少的趋势得到了改善。此外,国土资源部颁布的《国土资源部办公厅关于印发市县乡及土地利用总体规划编制指导意见的通知》和河南省颁布的《批转省国土资源厅关于做好土地利用总体规划修编工作指导意见的通知》这些政策,推进新乡市进行复垦和开发,这也是新乡市耕地面积不断增加的原因之一。

新乡市的建设用地主要集中在中心市区新乡县,并不断向外扩张。辉县市西北部紧邻太行山脉,山地丘陵面积较多,因此建设用地分布较少但却是林地的主要分布区域;新乡市南部的延津县和封丘县的建设用地变化不太明显;新乡市东南部的长垣市是重要的工业区,近年来建设用地面积也在不断增加。此外,新乡市是河南省中原地区的重要工业基地,工业生产总值较高,工业厂房的建设也在不断加快,进而导致建设用地增加。此外,人口的增长、城市化进程的加快、城市扩张均成为新乡市建设用地面积不断增加的原因。

4 结论

(1)提出了一种基于粒子群优化概率神经网络的半监督融合算法,该算法使用粒子群优化概率神

经网络分类器,使得分类器的性能更加优越,还融合了半监督适合小样本的优势。该算法在分类时可以使用少量的标记样本,解决了传统分类方法对标记样本数量要求高、分类精度低、成本高、错分漏分严重的问题,相比于其他 3 种传统算法,分类精度提高了 1.25~6.57 个百分点。由此得出,本研究提出的机器学习算法应用在 Landsat 影像的复杂地形区土地利用/土地覆盖分类中有速度快、准确率高、外界因素影响较小等优势,表现出了较强的适应性,运用机器学习算法获得土地利用/覆盖信息,并制定合理的政策成为未来研究的热点。

(2)对新乡市近年来土地利用现状进行分析,有利于解决新乡市目前存在的土地问题并为未来的土地利用决策提供依据。新乡市 1996—2020 年间耕地面积和建设用地面积分别增加了 433.21 km²和 537.11 km²;其他用地面积和草地面积均呈现出减少的趋势,水体面积变化小,14 年间仅减少了 3.45 km²;在土地流转方面,草地流转为耕地和建设用地的面积最大,变化最明显,水体的转换情况不是很突出,其他用地多转为草地和建设用地,耕地多转为建设用地。1996—2020 年新乡市土地利用变化主要受到城市化水平的提高,经济的增长以及国家的方针和地方的政策等多方面的共同影响。研究揭示了新乡市的近 14 年的土地利用变化特征,为该地区进一步合理开发和利用土地提供理论依据。

参 考 文 献

[1] KAVZOGLU T, MATHER P M. The use of backpropagating artificial neural networks in land cover classification [J].

- International Journal of Remote Sensing, 2003, 24(23): 4907–4938.
- [2] 韩文霆, 郭聪聪, 张立元, 等. 基于无人机遥感的灌区土地利用与覆被分类方法[J]. 农业机械学报, 2016, 47(11): 270–277.
HAN Wenting, GUO Congcong, ZHANG Liyuan, et al. Classification method of land cover and irrigated farm land use based on UAV remote sensing in irrigation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(11): 270–277. (in Chinese)
- [3] 刘文萍, 赵磊, 周焱, 等. 基于深度学习的无人机土地覆盖图像分割方法[J]. 农业机械学报, 2020, 51(2): 228–236.
LIU Wenping, ZHAO Lei, ZHOU Yan, et al. Deep learning based UAV landcover image segmentation method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 228–236. (in Chinese)
- [4] 乔平林, 张继贤, 卢秀山, 等. 基于神经网络的区域生态环境分类方法研究[J]. 地理与地理信息科学, 2004(2): 97–99.
QIAO Pinglin, ZHANG Jixian, LU Xiushan, et al. An artificially neural network method for regionally ecological classification[J]. Geography and Geo-Information Science, 2004(2): 97–99. (in Chinese)
- [5] 徐凯健, 田庆久, 杨闫君, 等. 遥感土地覆被分类的空间尺度响应研究[J]. 地球信息科学学报, 2018, 20(2): 246–253.
XU Kaijian, TIAN Qingjiu, YANG Yanjun, et al. Response of spatial scale for land cover classification of remote sensing[J]. Journal of Geo-information Science, 2018, 20(2): 246–253. (in Chinese)
- [6] 杨雪峰, 王雪梅. 基于决策树的多角度遥感影像分类[J]. 地球信息科学学报, 2016, 18(3): 416–422.
YANG Xuefeng, WANG Xuemei. Classification of MISR multi-angle imagery based on decision tree classifier[J]. Journal of Geo-information Science, 2016, 18(3): 416–422. (in Chinese)
- [7] 谷晓天. 基于机器学习的湟水流域土地利用/土地覆被分类研究[D]. 西宁: 青海师范大学, 2018.
GU Xiaotian. Research on land use/land cover classification based on machine learning in the Huangshui river basin[D]. Xining: Qinghai Normal University, 2018. (in Chinese)
- [8] WARREN S M, WALTER P. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biology, 1990, 52(1–2): 99–115.
- [9] MITCHELL T, BUCHANAN B, DEJONG G, et al. Machine learning[J]. Annual Review of Computer Science, 1990, 4(1): 417–433.
- [10] 李朝锋, 杨茂龙, 许磊, 等. 概率神经网络与 BP 网络模型在遥感图像分类中的对比研究[J]. 国土资源遥感, 2004(4): 11–13, 18.
LI Chaofeng, YANG Maolong, XU Lei, et al. A comparative study of probabilistic neural network and BP networks for remote sensing image classification[J]. Remote Sensing for Land & Resources, 2004(4): 11–13, 18. (in Chinese)
- [11] 喻昕玥, 吴良才. 概率神经网络在 ETM+ 影像分类中的应用[J]. 矿山测量, 2014(6): 30–32.
YU Xinyue, WU Liangcai. Application of probabilistic neural networks in ETM+ image classification[J]. Mine Surveying, 2014(6): 30–32. (in Chinese)
- [12] 胡剑策, 吴国平. 基于 K-L 变换的 BP 神经网络遥感图像分类[J]. 测绘科学, 2009, 34(3): 137–139.
HU Jiance, WU Guoping. BP neural network based on principle component analysis in multi-spectral remote sensing images classification[J]. Science of Surveying and Mapping, 2009, 34(3): 137–139. (in Chinese)
- [13] 李海洋, 范文义. 基于概率神经网络的遥感图像分类 MATLAB 实现[J]. 东北林业大学学报, 2008(6): 55–56, 62.
LI Haiyang, FAN Wenyi. MATLAB realization of sensing image classification based on probabilistic neural networks[J]. Journal of Northeast Forestry University, 2008(6): 55–56, 62. (in Chinese)
- [14] BASU S, BANERGEI A, MOONEY R J. Active semi-supervision for pairwise constrained clustering[C]//Proceedings of the 2004 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2004: 333–344.
- [15] 马永建, 汪传建, 赵庆展, 等. 基于 GF-1 遥感影像的荒漠区耕地分类与提取方法[J]. 石河子大学学报(自然科学版), 2021, 39(3): 383–390.
MA Yongjian, WANG Chuanjian, ZHAO Qingzhan, et al. Study on extraction of cultivated land in desert area with GF-1 remote sensing image on U-Net model[J]. Journal of Shihezi University(Natural Science), 2021, 39(3): 383–390. (in Chinese)
- [16] 王笑影, 周玉科, 温日红. 基于 Landsat-8 影像和随机森林方法的土地分类研究[J]. 测绘与空间地理信息, 2020, 43(11): 1–3.
WANG Xiaoying, ZHOU Yuke, WEN Rihong. Study on land cover classification using Landsat-8 image and random forest method[J]. Geomatics & Spatial Information Technology, 2020, 43(11): 1–3. (in Chinese)
- [17] 陈磊士, 赵俊三, 李易, 等. 基于机器学习的多源遥感影像融合土地利用分类研究[J]. 西南师范大学学报(自然科学版), 2018, 43(10): 103–111.
CHEN Leishi, ZHAO Junsan, LI Yi, et al. On land use classification by means of machine learning based on multi-source remote sensing image fusion[J]. Journal of Southwest China Normal University(Natural Science Edition), 2018, 43(10): 103–111. (in Chinese)
- [18] 赵亚杰, 王立辉, 孔祥兵, 等. 基于 Sentinel-2 和 Landsat8 OLI 数据融合的土地利用分类研究[J]. 福建农林大学学报(自然科学版), 2020, 49(2): 248–255.
ZHAO Yajie, WANG Lihui, KONG Xiangbing, et al. Land use classification based on data fusion of Sentinel-2 and Landsat8 OLI[J]. Journal of Fujian Agriculture and Forestry University(Natural Science Edition), 2020, 49(2): 248–255. (in Chinese)
- [19] 陈义菁, 张尹祯, 陈明珠, 等. 基于遥感影像数据的土地利用分类变化研究——以重庆市北碚区为例[J]. 科学技术创新, 2020(22): 191–192.
CHEN Yijing, ZHANG Yinzheng, CHEN Mingzhu, et al. A study on land use classification based on remote sensing data: a case study of Beibei District, Chongqing City[J]. Science and Technology Innovation, 2020(22): 191–192. (in Chinese)

- [20] 解银朋,周庆博,何金栋,等. 基于自训练的半监督 SQL 注入检测算法[J]. 陕西师范大学学报(自然科学版), 2021, 49(1): 37-43.
XIE Yinpeng, ZHOU Qingbo, HE Jindong, et al. Semi-supervised SQL injection detection based-on self-training[J]. Journal of Shaanxi Normal University(Natural Science Edition), 2021, 49(1): 37-43. (in Chinese)
- [21] 胡苏李扬,李辉,顾延生,等. 基于高分辨率遥感影像的神农架大九湖湿地土地利用类型变化及其驱动力分析——来自长时间尺度多源遥感信息的约束[J]. 国土资源遥感, 2021, 33(1): 221-230.
HUSU Liyang, LI Hui, GU Yansheng, et al. An analysis of land use changes and driving forces of Dajihu wetland in Shennongjia based on high resolution remote sensing image: constraints from the multi-source and long-term remote sensing information[J]. Remote Sensing for Land & Resources, 2021, 33(1): 221-230 (in Chinese)
- [22] 校园齐,程启先,余向洋,等. 基于灰色关联度的新乡市大气环境质量评价[J]. 环境与发展, 2019, 31(10): 9-10.
JIAO Yuanqi, CHENG Qixian, YU Xiangyang, et al. Evaluation of atmospheric environmental quality in Xinxiang City based on grey relational degree[J]. Environment and Development, 2019, 31(10): 9-10. (in Chinese)
- [23] 冯鑫. 新乡市农业产业集群发展问题及对策研究[D]. 新乡:河南师范大学, 2018.
FENG Xin. Research on the problems and countermeasures of the development of agricultural industrial cluster in Xinxiang City [D]. Xinxiang: Henan Normal University, 2018. (in Chinese)
- [24] 史利江,王圣云,姚晓军,等. 1994—2006 年上海市土地利用时空变化特征及驱动力分析[J]. 长江流域资源与环境, 2012, 21(12): 1468-1479.
SHI Lijiang, WANG Shengyun, YAO Xiaojun, et al. Spatial and temporal variation characteristics of land use and its driving force in Shanghai City from 1994 to 2006[J]. Resources and Environment in the Yangtze Basin, 2012, 21(12): 1468-1479. (in Chinese)
- [25] 刘雨,刘玉振. 城市土地利用结构变化及其驱动力研究——以河南省开封市为例[J]. 国土资源科技管理, 2011, 28(1): 67-73.
LIU Yu, LIU Yuzhen. Driving forces and evolution urban land use structure: a case study of Kaifeng City, Henan Province [J]. Scientific and Technological Management of Land and Resources, 2011, 28(1): 67-73. (in Chinese)
- [26] SPECHT D F. Probabilistic neural networks[J]. Neural Networks, 1990, 3(1): 109-118.
- [27] 巴桑,张正健,刘志红,等. 基于概率神经网络的遥感影像分类方法[J]. 高原山地气象研究, 2011, 31(3): 26-29.
BA Sang, ZHANG Zhengjian, LIU Zhihong, et al. Remote sensing image classification based on probabilistic neural network [J]. Plateau and Mountain Meteorology Research, 2011, 31(3): 26-29. (in Chinese)
- [28] 邹心遥,陈敬伟,姚若河. 采用粒子群优化的 SVM 算法在数据分类中的应用[J]. 华侨大学学报(自然科学版), 2016, 37(2): 171-174.
ZOU Xinyao, CHEN Jingwei, YAO Ruohu. Application of SVM algorithm based on particle group optimization in data classification[J]. Journal of Huaqiao University (Natural Science), 2016, 37(2): 171-174. (in Chinese)
- [29] 朱腾,余洁,谢东海,等. 粒子群优化算法在全极化 SAR 影像非监督分类中的应用[J]. 测绘科学技术学报, 2014, 31(1): 57-61.
ZHU Teng, YU Jie, XIE Donghai, et al. Particle swarm optimization in POLSAR image unsupervised classification[J]. Journal of Geomatics Science and Technology, 2014, 31(1): 57-61. (in Chinese)
- [30] SHI J F, LIU F, LIN Y H, et al. Polarimetric SAR image classification based on deep learning and hierarchical semantic model[J]. Acta Automatica Sinica, 2017, 43(2): 215-226.

(上接第 157 页)

- [22] 司乔瑞,袁寿其,CHERDIEU P,等. 径向导叶内流动分离特性的试验研究和控制[J]. 实验流体力学, 2015, 29(6): 48-53, 78.
SI Qiaorui, YUAN Shouqi, CHERDIEU P, et al. Experimental investigation on the characteristics and controlling of flow separation inside the centrifugal diffuser[J]. Journal of Experiments in Fluid Mechanics, 2015, 29(6): 48-53, 78. (in Chinese)
- [23] Rotodynamic pumps-hydraulic performance acceptance tests-Grade1, 2 and 3[S]. 2016.
- [24] 张德胜,吴苏青,施卫东,等. 不同湍流模型在轴流泵叶顶泄涡模拟中的应用与验证[J]. 农业工程学报, 2013, 29(13): 46-53.
ZHANG Desheng, WU Suqing, SHI Weidong, et al. Application and experiment of different turbulence models for simulating tip leakage vortex in axial flow pump[J]. Transactions of the CSAE, 2013, 29(13): 46-53. (in Chinese)
- [25] 张克危. 流体机械原理[M]. 北京:机械工业出版社, 2012.
- [26] ZHOU Y, ZHANG H, CHEN B. Influence of double-inlet design on the flow-head characteristics of axial-flow pump[J]. Journal of Hydrodynamics, 2021, 33(4): 763-772.