

基于特征优化的 SAR 图像水华识别方法



毋 琳^{1,2,3} 白 澜^{1,3,4} 孙梦伟^{1,3,4} 郭拯危^{1,3,4}

1 河南大学计算机与信息工程学院 河南 开封 475004

2 河南大学环境与规划学院 河南 开封 475004

3 河南省智能技术与应用工程技术研究中心 河南 开封 475004

4 河南省大数据分析与处理重点实验室 河南 开封 475004

(henuwl@henu.edu.cn)

摘要 内陆湖泊水华现象的频繁爆发,严重影响着地表水环境安全,严重阻碍了我国的生态文明建设。充分发挥合成孔径雷达(Synthetic Aperture Radar,SAR)遥感技术全天时、全天候的优势,可实现大尺度、周期性的水华识别与监测工作,对于地表水生态环境的保护与监管具有重大的现实意义。立足于SAR遥感目标识别技术的研究与应用,文中提出了一种基于特征优化的水华识别方法。该方法基于对水华SAR图像特征的深入分析与提取,应用ReliefF特征优化算法对全部的22个水华特征进行筛选与优化,得到包含10个特征的最优特征子集,并以反向传播(Back Propagation,BP)神经网络为分类识别器完成了多组对比实验,水华识别总体精度最高达81.39%,较优化之前提升了19.38%。实验结果表明,使用最优特征集不仅可以大幅降低算法复杂度,还可以有效地提升水华总体识别精度,具有进一步推广的实用价值。

关键词:合成孔径雷达图像;特征优化;水华识别;ReliefF算法;水生态

中图法分类号 TP751

Algal Bloom Discrimination Method Using SAR Image Based on Feature Optimization Algorithm

WU Lin^{1,2,3}, BAI Lan^{1,3,4}, SUN Meng-wei^{1,3,4} and GOU Zheng-wei^{1,3,4}

1 College of Computer and Information Engineering, Henan University, Kaifeng, Henan 475004, China

2 College of Environment and Planning, Henan University, Kaifeng, Henan 475004, China

3 Henan Engineering Research Center of Intelligent Technology and Application, Henan University, Kaifeng, Henan 475004, China

4 Key Laboratory of Analysis and Processing on Big Data of Henan Province, Henan University, Kaifeng, Henan 475004, China

Abstract The frequent outbreak of algal bloom in inland lakes has seriously affected the safety of surface water environment, and has brought great obstacles to the construction of ecological civilization in China. Taking full advantage of SAR(Synthetic Aperture Radar) remote sensing technologies, large-scale and periodic algal bloom discrimination and monitoring can be realized. It is of great practical significance for the protection and supervision of water environment. Based on the research and application of SAR remote sensing target recognition technology, this paper proposes an algal bloom discrimination method with feature optimization. After the in-depth analysis and extraction of algal bloom image features, the ReliefF algorithm is used to obtain the optimal feature set, which consists of 10 features from all 22 algal bloom features. And then, the BP (Back Propagation) neural network is as the classifier of this discrimination method to carry out a number of comparative experiments. The overall accuracy of the proposed method is 81.39%, which is 19.38% higher than that before optimization. The experimental results show that the optimal feature set can not only greatly reduce the algorithm complexity, but also effectively improve the discrimination accuracy of algal bloom, which has practical value for further promotion.

Keywords Synthetic aperture radar image, Feature optimization, Algal bloom discrimination, ReliefF algorithm, Water ecology

到稿日期:2020-08-21 返修日期:2020-10-30 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61871175);河南省高等学校重点科研项目(19A420005,21A520004);河南省科技攻关计划项目(202102210175,212102210093,212102210101);自然资源部国土卫星遥感应用重点实验室经费资助项目(KLSMNR-202102)

This work was supported by the National Natural Science Foundation of China(61871175), College Key Research Project of Henan Province(19A420005,21A520004), Plan of Science and Technology of Henan Province(202102210175,212102210093,212102210101) and Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China(KLSMNR-202102).

通信作者:郭拯危(gzw@henu.edu.cn)

1 引言

水华现象是水体富营养化的典型特征。当水体受到废水污染时,氮磷成分增加,在一定气象条件下藻类会大量繁殖,聚集形成水华^[1-2]。水华会迅速消耗水体中的氧气,导致水生生物缺氧死亡,且藻类会产生大量的毒素^[3],严重破坏水生态环境,对水生态安全造成了巨大影响。同时,水体水质恶化会进一步加剧我国水资源短缺的问题,不利于顺利实施可持续发展战略^[4]。近年来,水华问题已经成为我国水生态保护的重点关注内容之一,对水华的监测与预警也成为了保护水环境安全的一项重要工作。

近些年,卫星遥感技术在地表环境监测领域发挥了越来越重要的作用^[5-6]。光学遥感和微波遥感都是常用的卫星遥感技术,其中光学遥感应用较广泛,但易受云、雨、雾等恶劣天气以及光照的影响^[7]。星载合成孔径雷达(Synthetic Aperture Radar,SAR)是一种主动微波遥感技术,通过发射电磁波,获取地物反射的后向散射信息进行成像,具有全天时、全天候的对地观测能力,可以有效弥补光学遥感的不足^[8-9]。

国内外在SAR图像水华识别研究领域起步较晚。2009年,文献[10]发现ASAR图像中的暗斑区与MODIS图像中的水华区高度一致。在后续的水华识别研究中,SAR图像与光学遥感图像相结合成为了较为常用的方法。2013年,文献[11]基于ASAR图像和MODIS图像对太湖蓝藻水华进行提取,验证了SAR图像可以用于水华的监测。2014年,文献[12]结合ASAR图像和MERIS光学图像,对欧洲Cruonian湖中蓝藻水华不同阶段叶绿素浓度和后向散射系数关系进行了反演。2015年,文献[13]利用多时相SAR图像和光学遥感图像进行水华识别,得到67.74%的总体识别精度。2018年,文献[14]提出了一种基于支持向量机(Support Vector Machines,SVM)的水华识别方法,该方法摆脱了对光学图像的依赖,仅依靠Sentinel-1A SAR图像就实现了水华识别,总体识别精度达到了74.32%。

虽然文献[14]在水华识别过程中摆脱了对光学图像的依赖,将该领域研究推向了一个新的阶段。但是其水华识别过程仅使用了9个SAR图像特征,且并未评估各特征值对于识别的贡献,存在特征提取不完全、无法抑制消极特征影响的问题,因此,最终水华总体识别精度不高。近些年在与水华识别相似的SAR图像溢油识别的研究中,特征选择算法也常用于特征的筛选与优化,并获得了较好的识别精度和提升效果^[15-16]。

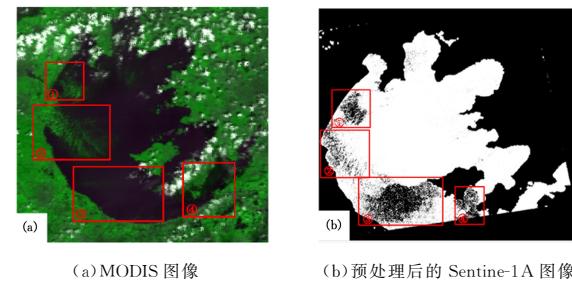
针对这一问题,本文结合ReliefF特征选择算法,提出了一种基于特征优化的SAR图像水华识别方法,并应用BP神经网络作为识别核心开展验证实验,得到了最高81.39%的总体识别精度。该实验结果表明,最优特征集的使用可以有效提升SAR图像水华识别的效果。

2 SAR图像水华识别方法

2.1 水华的SAR图像特点

湖库水华爆发期间,藻类快速死亡、腐败、分解,聚集在水面上形成大片绿色黏质物,对水面波纹具有较大的平滑作用。

SAR图像的后向散射情况与地物目标表面粗糙度有着直接的关系,水华减弱了水面 Bragg 散射,使后向散射回波降低^[17],呈现为较明显的水华图斑。图1为2017年9月12日太湖区域的同步MODIS图像与预处理后的Sentinel-1A图像。对照图1(a)可知,图1(b)中区域①和区域②为水华图斑,而区域③和区域④在视觉效果上与水华图斑相似,但却不是水华引起的,且该区域非太湖的水华爆发区域^[18-19],这些类似水华的图斑是由低风速(小于3.2m/s)等原因造成的^[14,20]。因此,有效地从SAR图像中区分水华与非水华图斑是完成水华识别的关键。



(a) MODIS 图像

(b) 预处理后的 Sentinel-1A 图像

图1 2017年9月12日太湖区域遥感图像

Fig. 1 Remote sensing image of Taihu Lake on September 12, 2017

2.2 SAR水华识别方法流程

基于特征的SAR图像水华识别方法一般包括SAR图像预处理、图斑分割、特征提取、水华识别4个阶段,其处理流程如图2所示。图斑分割是利用图像分割等技术将SAR图像中疑似水华的图斑完整提取的过程。经过该步骤的处理,将去除图斑周边的陆地、水体的影响,仅保留图斑区域的SAR图像信息。同时,该步骤还需结合相关的光学图像与气象水文信息判别每个图斑的性质,为获取到的每个图斑添加水华与非水华标签,形成图斑集,用于后续的识别处理^[21]。特征提取是水华识别的关键步骤,提取到的特征质量会直接影响最后的识别效果,本文后续将详细阐述本文的水华特征提取与优化过程。水华识别阶段一般由分类识别器通过训练学习调参后,实现水华识别。

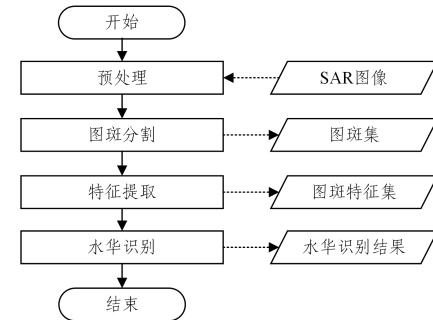


图2 SAR图像水华识别的一般流程

Fig. 2 Flow chart of algal bloom discrimination using SAR image

2.3 SAR图斑提取

本文应用改进的K-means算法^[14]对提取到的图斑SAR进行粗分割,进一步应用区域生长算法去除分割结果中可能存在的边缘噪声和零碎干扰,并使用形态学对图像进行处理,可得到只保留目标信息的图斑结果。图斑分割过程如图3所

示,其中,图 3(a)为 SAR 图斑原始图像,图 3(b)为改进 K-means 算法处理后的二值图,图 3(c)为区域生长算法处理后的二值图,图 3(d)为最终获取的图斑结果。

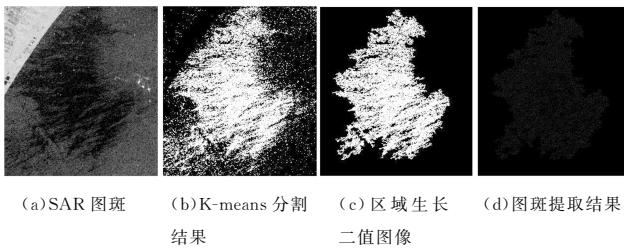
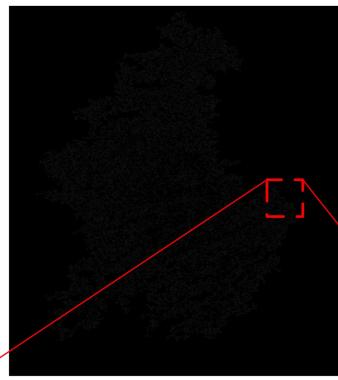


图 3 SAR 图像的图斑提取过程

Fig. 3 Region extraction process of SAR image

分割后的图斑会保持 SAR 图像目标区域的原始灰度信息,而其他区域灰度值置 0。图 3(d)中的局部区域灰度矩阵如图 4 所示。



8	0	5	2	6	4	8	12	5	5	13	9	0	3	13	16	15	0	0	16	20	17	14	15	1	0
11	13	15	11	12	4	17	16	6	7	8	7	5	4	5	8	13	5	1	2	13	9	16	17	16	4
4	6	4	10	6	0	15	3	0	5	2	4	7	4	1	5	12	0	0	0	7	20	11	6	10	11
18	16	9	16	5	2	16	1	0	2	4	8	7	3	9	16	16	18	5	4	0	20	3	12	20	14
15	14	8	7	0	2	3	0	0	4	7	11	6	0	5	11	4	11	4	10	4	18	7	13	11	13
2	3	4	2	1	12	1	0	0	0	2	9	10	4	8	12	6	8	7	7	13	2	15	18	14	11
0	1	0	4	10	19	2	0	1	15	5	2	13	18	15	14	8	13	20	7	6	7	6	18	12	7
4	2	4	0	3	0	0	0	3	4	10	10	6	5	12	17	11	0	17	7	1	3	6	14	4	5
2	0	3	0	2	0	4	4	14	3	9	9	2	4	13	12	3	0	16	6	1	4	6	11	5	7
2	3	0	2	0	12	7	6	3	7	7	11	14	17	13	3	1	2	13	11	17	21	8	0	0	0
11	18	6	0	0	9	1	0	0	0	0	12	18	17	6	0	4	21	15	2	0	2	0	0	5	0
11	15	8	0	1	0	0	0	4	1	4	13	13	19	19	5	12	0	0	0	0	0	0	2	0	4
3	5	6	0	6	3	5	0	3	0	0	0	0	8	16	0	4	1	0	1	2	0	0	4	0	0
0	8	8	9	3	12	7	4	0	7	3	4	0	10	14	15	18	3	13	19	4	0	0	0	0	0
10	11	0	10	16	13	9	0	1	0	2	0	7	9	10	4	14	17	9	15	0	4	3	0	6	0
12	18	7	11	9	10	8	7	13	2	2	0	0	6	4	5	12	14	16	24	2	0	0	0	0	4
13	5	6	21	15	16	9	12	15	1	0	1	0	9	0	1	3	0	1	13	5	0	0	6	0	0
3	13	9	12	3	19	17	21	16	3	1	4	4	7	0	0	0	7	4	11	18	4	0	13	2	5
10	11	14	16	3	17	13	21	19	14	11	7	8	0	0	0	2	0	0	3	15	1	1	22	15	10
18	4	16	17	8	16	8	14	14	17	14	10	15	2	3	0	1	7	20	22	19	0	2	17	13	3
1	16	17	1	0	17	16	10	13	7	9	14	9	6	3	0	2	16	12	6	0	2	7	11	9	
0	0	19	17	20	23	14	13	17	12	0	2	3	7	5	10	2	13	16	2	1	0	1	0	5	14
16	0	19	14	13	16	9	13	18	5	13	22	17	14	10	1	3	9	13	22	8	0	0	3	0	0
1	17	18	0	1	18	20	16	9	9	10	11	4	3	4	0	2	8	12	16	9	4	2	0	0	3
0	0	11	5	0	1	1	8	15	12	10	10	1	0	1	3	9	9	8	6	4	11	6	0	4	14
1	0	12	11	2	0	0	0	2	9	10	16	8	0	2	8	15	3	1	0	8	10	6	15	9	
4	0	5	7	1	1	4	1	0	4	3	12	8	0	2	8	16	14	4	6	4	4	13	17	17	8
8	2	10	18	11	8	15	17	18	5	0	3	4	0	6	9	9	4	12	8	1	10	14	7	1	
3	6	5	9	1	0	10	6	2	11	0	5	9	6	11	12	8	8	8	12	4	0	3	3	8	

图 4 图斑局部区域灰度矩阵

Fig. 4 Gray matrix of SAR image region

3 水华特征提取与优化

3.1 水华特征提取

水华图斑与非水华图斑在视觉效果上虽然十分相似,因水华本身的特征及水流等原因,水华图斑一般为带状形态,且水华区域内部图像纹理较为细腻、平滑,而边缘则有明显的跃迁特征。文献[14]采用人工提取水华图斑特征的方式,共获取了如表 1 所列的 9 个水华图斑特征,并对水华与非水华进行了识别。

表 1 文献[14]中使用的水华图斑特征

Table 1 Algal bloom features of SAR image region used in reference [14]

几何特征	灰度特征	纹理特征
面积 S	复杂度 COM	背景后向 散射均值与 梯度 $MRTB$
对比度 $MRTB$	方差之比 $MVRB$	边缘 梯度 GOE
		角二 矩阵 Asm
		对比度 Con
		熵 Ent
		倒数 差分矩 Rdm

本文继续对水华形成机理、研究区区域特征,以及 SAR 图像特性等多方面进行深入分析和研究,新增 7 个 Hu 不变矩的几何特征和 6 个纹理特征,如表 2 所列。其中, η_{pq} 为归一化中心距; x_{ij} 为图斑中像素点在水平和垂直方向上的梯度; $p(i,j)$ 表示图像 (i,j) 处的灰度值。

表 2 本文新增的水华图斑特征

Table 2 New algal bloom features of SAR image region used in this paper

特征类型	特征名称	特征公式
几何特征	$\phi_1 = (\eta_{20} + \eta_{02})$	
	$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$	
	$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$	
	$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$	
	$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + 3\eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$	
	$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$	
	$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$	
纹理特征	均值 $Mean$	$Mean = \sum_{i=0}^{255} \sum_{j=0}^{255} p(i,j) * i$
	方差 Var	$Var = \sum_{i=0}^{255} \sum_{j=0}^{255} p(i,j) * (i - Mean)^2$
	标准差 Std	$Std = \sqrt{Var}$
	相关性 Cor	$Cor = \frac{\sum_{i=0}^{255} \sum_{j=0}^{255} (i - Mean) * (j - Mean) * p(i,j)}{Var}$
	同质性 Hom	$Hom = \frac{\sum_{i=0}^{255} \sum_{j=0}^{255} p(i,j) * \frac{1}{1+(i+j)^2}}$
相异性 Dis	$Dis = \sum_{i=0}^{255} \sum_{j=0}^{255} p(i,j) * i - j $	

3.2 ReliefF 特征选择算法

特征优化是机器学习领域中的一种预处理技术,其目的是从原始特征集中去除低相关性或负相关性的特征,形成较低维的最优特征子集,避免高维数据降低算法效率、无关特征影响算法精度等,从而提高机器学习的性能^[22]。

1994 年,Kononenko 首次提出了 ReliefF 算法^[23],以特征权重表征每个特征与类别的相关性,通过多次迭代更新,最终依据权重值的大小选出最优特征集,其流程如图 5 所示。

从图斑样本集中随机选择一个样本 R ,在 R 的同类样本中寻找 k 个近邻样本 H ,在 R 的不同类样本中寻找 k 个近邻样本 U ,依据式(1)迭代更新样本 R 的每个特征的权重值 $W(*)$,直至所有样本计算完毕,以从大到小的顺序输出所有特征权重结果。

$$W(A) = W(A) - \frac{\sum_{j=1}^k diff(A, R, H_j)}{mk} + \sum_{C \in class(R)} \left[\frac{P(C)}{1-P(class(R))} \cdot \frac{\sum_{j=1}^k diff(A, R, U_j(C))}{mk} \right]$$

其中, $W(A)$ 表示特征 A 的权重; m 表示迭代次数; $class(R)$ 表示样本 R 的所属类别; C 表示样本 R 的不同类别; $P(*)$ 表示类别概率; $diff(A, R, H_j)$ 为样本 R 和样本 H_j 在特征 A 上的距离差值函数, 其距离差值的计算式如式(2)所示。

$$diff(A, R, H_j) = \begin{cases} \frac{|R[A] - H_j[A]|}{\max(A) - \min(A)}, & A \text{ 连续} \\ 0, & A \text{ 离散, 且 } R[A] = H_j[A] \\ 1, & A_r \text{ 离散, 且 } R[A] \neq H_j[A] \end{cases} \quad (2)$$

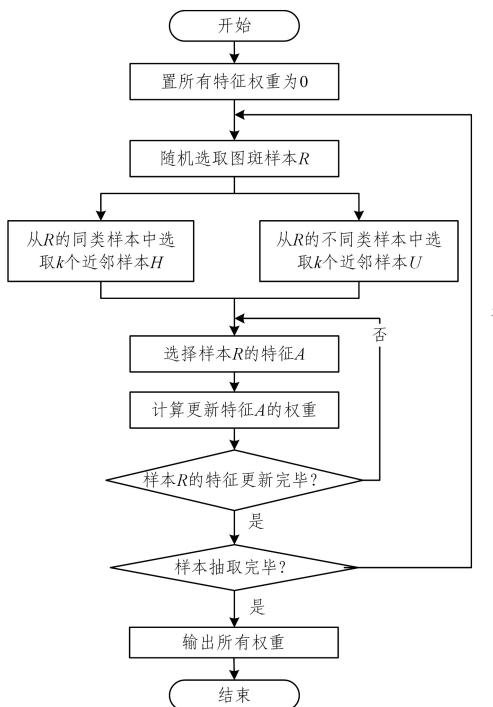


图 5 ReliefF 算法流程图

Fig. 5 Flow chart of ReliefF algorithm

表 3 特征选择的权重结果

Table 3 Weight result of feature selection

特征	Ent	Rdm	Cor	Con	Asm	ϕ_1	ϕ_2	Dis	ϕ_3	Mean
权重	0.12166	0.10145	0.09896	0.08599	0.05121	0.04359	0.03741	0.0299	0.02733	0.02539

由早期研究可知, 纹理特征是单极化单波段 SAR 图像识别中的重要特征。如表 3 所列, 最优特征集中共包含 10 个特征, 其中纹理特征有 7 个, 且相关性排名前 5 的 Ent, Rdm, Cor, Con, Asm 特征也都是表征纹理信息的。这也验证了纹理特征对于 SAR 图像目标识别的重要性。

4 实验与分析

4.1 研究区域

太湖($30^{\circ}56' - 31^{\circ}33' N, 119^{\circ}52' - 120^{\circ}37' E$)位于江苏省无锡市, 长江三角洲南部, 是我国五大淡水湖之一。太湖水域面积 $2338 km^2$, 平均水深 1.9 m, 其水质情况与苏州、湖州、无锡、宜兴市等周边城市的农业灌溉、水产养殖等生产活动密切相关。自 20 世纪 80 年代后期, 太湖水体开始出现富营养化现象, 90 年代富营养化状态逐年加剧, 重度富营养化状态由局部水体逐渐扩展至主湖区^[24-25]。2007 年太湖爆发了严重的蓝藻水华污染事件, 对当地居民的日常生活造成了巨大的

3.3 水华特征优化选择

水华图斑的 22 个特征全部应用于水华识别, 无疑会增加数据维度, 使算法的运行效率降低, 同时对识别精度的提升并不一定有较好的效果。只有将特征集合中的消极特征去除, 应用更加有效的表征特征进行水华识别, 才有可能更好地完成识别。

本文针对由 SAR 图像获取到的水华与类似水华图斑样本集合, 应用 ReliefF 算法, 获取 22 个特征的权重排序, 并设置适当的阈值, 从而完成特征的优化选择。

算法 1 ReliefF 特征优化算法

输入: 样本集合 D
 输出: 特征权重集合 $W = \{W_A | A=1, 2, \dots, n\}$
 1. /* 水华 SAR 图像特征优化方案 */
 2. 初始化: $m \leftarrow \text{Random}(0 \sim 72)$
 /* m 为迭代次数 */;
 $k \leftarrow \text{Random}(0 \sim 22)$
 /* k 为近邻数目 */;
 $n \leftarrow 22$ /* n 为特征数目 */;
 $W \leftarrow 0$ /* W 为特征权重 */;
 3. For $i=1$ to m do
 4. 从 D 中随机选择水华样本 R
 5. 在 D 中选择 k 个与 R 同类的近邻水华样本
 $H = \{H_j | j=1, 2, \dots, k\}$
 在 D 中选择 k 个与 R 不同类的近邻水华样本
 $U = \{U_j | j=1, 2, \dots, k\}$
 6. For $t=1$ to n do
 依据式(1)更新 W_t 的权重
 7. 特征权重 $W = \{W_A | A=1, 2, \dots, n\}$ 排序输出。

通过多次实验验证, 当迭代次数设置为 15, 近邻数目设置为 10, 阈值设置为 0.02 时, 可以获得对于水华识别较为有利的包含 10 个特征的最优特征集, 其中选中的特征及其权重如表 3 所列。

影响^[26]。2008 年以后, 国家与地方政府对太湖水华展开了大力度治理, 但水华爆发现象仍时有发生。因此, 本文选取太湖作为研究区域, 区域位置及其 SAR 图像如图 6 所示。

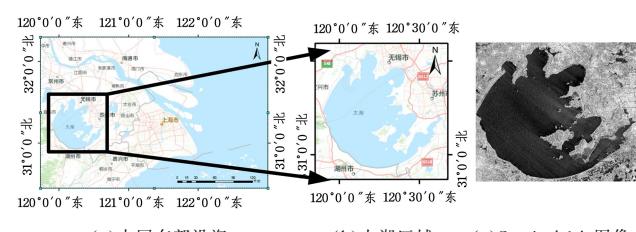


图 6 太湖研究区域位置及其 SAR 图像

Fig. 6 Location of Taihu Lake and its SAR image

4.2 SAR 图像数据

Sentinel-1A 卫星重访周期 12 天, 其干涉宽幅地距多视数据产品的幅宽为 $250 km$ 、分辨率为 $5 m \times 20 m$, 能够为水华识别研究提供长时间序列的遥感数据。实验采集了 2015 年

5月2日至2019年4月5日期间,覆盖太湖区域的27景Sentinel-1A SAR图像,具体的数据采集时间如表4所列。

表4 太湖Sentinel-1A SAR图像数据采集时间

Table 4 Data acquisition time of Sentinel-1A SAR image in Taihu Lake

区域	年份	图像时间
太湖	2015	05/02/2015, 05/21/2015, 08/18/2015, 09/11/2015, 10/05/2015
	2016	07/14/2016, 07/26/2016, 08/07/2016, 09/17/2016, 10/24/2016
	2017	08/07/2017, 08/19/2017, 08/31/2017, 09/12/2017, 09/24/2017, 10/06/2017, 10/30/2017
	2018	04/16/2018, 04/28/2018, 05/22/2018, 06/27/2018, 07/21/2018, 08/02/2018, 08/14/2018, 08/26/2018
	2019	03/18/2019, 04/05/2019

4.3 实验设置

Sentinel-1A SAR图像的预处理主要包括辐射校正、精致Lee滤波、基于距离-多普勒算子的几何校正,以及图像拉伸等步骤。通过该预处理操作,可以较大程度上保证SAR图像灰度值与目标地物回波强度的一致性,同时使水面信息更加突出,易于后续处理。

实验使用的水华与非水华图斑集合均由采集到的27景Sentinel-1A图像获取,经过对原始SAR图像裁剪、分割处理,并与同时期光学图像和气象水文资料对比确认,共获得水华图斑72个,非水华图斑57个。

水华识别分类器采用BP神经网络,设置其学习率为0.01。将所有样本的最优特征集采用留一法分为训练集和测试集,训练集用于BP网络的训练以及权值与阈值调优,测试集用于模型测试和识别结果输出。

4.4 结果分析

为验证特征优化对水华识别精度的有效提升性能,分别采用文献[14]的9个特征、未优化选择的22个全部特征,以及经过优化选择的10个特征,使用相同结构的BP神经网络和数据集展开验证实验。实验结果采用水华识别精度、非水华识别精度、总体识别精度,以及Kappa系数进行评价。

由表5可知,验证实验1使用了与文献[14]相同的特征,仅将识别分类器由SVM替换为BP网络,水华识别精度、非水华识别精度,以及总体识别精度均有一定的提升,其中水华识别精度提升了近8%,Kappa系数也增大了0.11。这说明采用BP网络可以较为有效地完成水华识别任务。

表5 水华识别结果的精度对比

Table 5 Accuracy comparison of algal bloom discrimination experimental results

特征数目	识别分类器	水华识别精度/%	非水华识别精度/%	总体识别精度/%	Kappa系数
文献[14]	SVM	74.36	74.29	74.32	0.48
验证实验1	BP网络	82.05	77.14	79.73	0.59
验证实验2	BP网络	59.13	71.43	62.01	0.24
本文方法	BP网络	82.43	80.00	81.39	0.72

在验证实验2中,仍然使用与验证实验1相同的BP网络,但将提取到的22个水华特征全部用于分类识别,得到的各项指标结果不升反降,总体识别精度下降至62.01%,Kappa系数仅有0.24。这一结果表明,仅增加特征数目,并不一

定对分类识别精度的提升具有积极作用,当特征集中出现与识别目标负相关的消极特征时,会导致识别效果下降。

在验证实验2的基础上,使用最优特征集进行水华识别,此时虽然特征数目与验证实验1相差不多,但由于采用了与水华目标高相关性的特征,各项识别结果的精度较前几种方法均有了明显提升,其中水华识别精度为82.43%,非水华识别精度为80.00%,总体识别精度达到了81.39%,Kappa系数为0.72。

上述验证实验表明,在使用人工特征识别方法开展SAR图像目标识别时,所用特征数目与识别结果精度并不存在直接相关关系,只有适当选择与识别目标相关性强的特征,才可以更好地达到预期目标。使用特征选择算法进行特征优化,不但可以降低数据处理的复杂度,提升识别算法的执行效率,还能够有效提升算法性能,获取更高的识别精度。

结束语 本文提出了一种基于特征优化的SAR图像水华识别方法,该方法针对水华的图像特点,采用ReliefF算法对其全部的22个特征进行优化选择,获取最优特征集,并以BP神经网络为核心分类器完成水华识别任务。本文使用太湖区域SAR图像数据集开展验证实验,结果表明,使用最优特征集可以较好地提高水华识别精度,在后续研究中可以从特征算法和分类器两个方面入手,进一步改善基于SAR图像的水华识别方法。

参 考 文 献

- [1] TAO F. Air-water CO₂ flux in an algae bloom year for Lake Hongfeng, Southwest China: implications for the carbon cycle of global inland waters[J]. Acta Geochimica, 2017, 36(4): 658-666.
- [2] KATRI R, JOSÉ E C B, MARIA H K V. Identifying Multiple Stressors that Influence Eutrophication in a Finnish Agricultural River[J]. Science of the Total Environment, 2019, 658: 1278-1292.
- [3] XIE P. Cyanobacteria bloom and its secondary damage[J]. Journal of Hydroecology, 2015, 36(4): 1-13.
- [4] GUO L E. Doing battle with the green monster of Taihu Lake [J]. Science, 2007, 317(5842): 1166.
- [5] GUO H D, ZHANG L. 60 years of radar remote sensing: Fourstage development. Journal of Remote Sensing[J]. Journal of Remote Sensing, 2019, 23(6): 1023-1035.
- [6] SUN W W, YANG G, CHEN C, et al. Development status and literature analysis of China's earth observation remote sensing satellites[J]. Journal of Remote Sensing, 2020, 24(5): 4-25.
- [7] CHEN L F, LIU Y Z, ZHANG P, et al. Road Extraction Algorithm of Multi-feature High-resolution SAR Image Based on Multi-Path RefineNet[J]. Computer Science, 2020, 47(3): 156-161.
- [8] WANG H, CHEN X, TIAN S Z, et al. SAR Image Recognition Based on Few-shot Learning [J]. Computer Science, 2020, 47(5): 124-128.
- [9] PENG J X, SU Y Q, XUE X R. SAR Image Feature Retrieval Method Based on Deep Learning and Synchronic Matrix[J].

- Computer Science,2019,46(6):196-199.
- [10] WANG J,WANG S X,YAN F L,et al. Algal bloom range extraction based on ASAR and MODIS images[J]. Remote Sensing Information,2009,3:54-57.
- [11] LIU J,WANG W C. Cyanobacteria bloom extraction based on ASAR and MODIS remote sensing data[J]. Journal of Anhui Agricultural Sciences,2013,41(2):889-892.
- [12] BRESCIANI M,ADAMO M,DE CAROLIS G,et al. Monitoring blooms and surface accumulation of cyanobacteria in the Curoian Lagoon by combining MERIS and ASAR data[J]. Remote Sensing of Environment,2014,146:124-135.
- [13] WANG G L. Active and passive remote sensing monitoring of cyanobacteria blooms in inland waters[D]. Shanghai:East China Normal University,2015.
- [14] WU L,WANG L,MIN L,et al. Discrimination of Algal-Bloom Using Spaceborne SAR Observations of Great Lakes in China [J]. Remote Sensing,2018,10(5):767.
- [15] MERA D,BOLON-CANEDO V,COTOS J M,et al. On the Use of Feature Selection to Improve the Detection of Sea Oil Spills in SAR Images[J]. Computers and Geosciences,2017,100: 166-178.
- [16] RAEISI A,AKBARIZADEH G,MAHMOUDI A. Combined Method of an Efficient Cuckoo Search Algorithm and Nonnegative Matrix Factorization of Different Zernike Moment Features for Discrimination between Oil Spills and Lookalikes in SAR Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2018,11(11):4193-4205.
- [17] JOHANNESSEN J A,SCHUCHMAN R A,DIGRANCES G,et al. Coastal Ocean Fronts and Eddies Imaged with ERS-1 Synthetic Aperture Radar[J]. Journal of Geophysical Research,1996,101:6651-6667.
- [18] KONG F X,GAO G. Reflection on the formation mechanism of cyanobacteria bloom in large shallow eutrophication lakes[J]. Acta Ecologica Sinica,2005(3):589-595.
- [19] KONG F X,MA R H,GAO J F,et al. Theory and Practice of prevention,prediction and early warning of cyanobacteria bloom in Taihu Lake[J]. Journal of Lake Sciences,2009(3):10-24.
- [20] BAI X H,HU W P,HU Z X,et al. Calculation of wind drift volume of sheet floating algal bloom in Meiliang Bay,Taihu Lake in summer 2004[J]. Environmental Science,2005,26(6):57-60.
- [21] ZHOU Y Y,CHEUNG J H,LIU T,et al. Review of Road Extraction for High-resolution SAR Images[J]. Computer Science,2020,47(1):124-135.
- [22] GUYON I M,GUNN S R,NIKRAVESH M,et al. Feature Extraction: Foundations and Applications[J]. Studies in Fuzziness & Soft Computing,2006,205(12):68-84.
- [23] KONONENKO I. Estimating attributes:Analysis and extensions of RELIEF[C]// European Conference on Machine Learning on Machine Learning. Springer,1994:171-182.
- [24] DUAN H,MA R,XU X,et al. Two-Decade Reconstruction of Algal Blooms in China's Lake Taihu[J]. Environmental Science & Technology,2012,43(10):3522-3528.
- [25] CHEN Y W,QIN B Q,TEUBNER K,et al. Long-term dynamics of phytoplankton assemblages: Microcystis-domination in Lake Taihu,a large shallow lake in China[J]. Journal of Plankton Research,2003,25(1):445-453.
- [26] YANG L Y,YANG X Y,REN L M,et al. Outbreak mechanism and control measures of cyanobacteria bloom in Taihu Lake[J]. Journal of Lake Sciences,2019,31(1):18-27.



WU Lin, born in 1978, associate professor, master supervisor. Her main research interests include SAR image processing techniques, and SAR image application of water environment.



GUO Zheng-wei, born in 1963, professor, master supervisor. Her main research interests include SAR image processing techniques, and SAR image application of ecological environment.