

基于多源遥感数据的植被精细分类方法对比

李雨阳¹, 刘舫^{2*}, 胡文君², 李芹², 冯予泽¹, 杨长保¹, 陈天逸¹

1. 吉林大学 地球探测科学与技术学院, 长春 130026;

2. 云南省地质环境监测院, 昆明 650216

摘要: 针对东北虎豹国家公国内独特的温带森林植被, 开展了关于多源遥感影像覆盖分类的研究, 旨在利用现代遥感技术与先进的机器学习方法, 对复杂的植被类型进行精细分类, 从而更好地理解该区域的生态环境状况, 并为东北虎豹国家公园的保护工作提供科学依据。首先采用多尺度图像分割技术, 将原始遥感影像分解成一系列具有相似纹理和颜色特征的小区域, 有助于减少数据冗余并突出植被的不同特征; 接着对这些分割后的图像进行特征优化, 提取能够有效区分不同植被类型的特征参数; 随后将这些优化后的特征参数输入至3种不同的机器学习分类器中进行训练和测试, 包括支持向量机、C5.0决策树以及随机森林。实验结果显示, 在分类精度方面, 随机森林精度最高, 达91.33%, C5.0决策树精度达到89.59%, Kappa系数分别为0.90和0.88, 而支持向量机的分类精度较低, 仅为58.96%, Kappa系数为0.54。这一结果表明, 对于植被覆盖类型区分以及其他地物的识别, 随机森林算法具有显著优势, 不仅能够在复杂背景下准确区分不同的植被类型, 还能有效识别其他地物, 如道路、建筑物等。

关键词: 多源遥感影像; 植被覆盖分类; 支持向量机; 决策树; 随机森林

Comparison on vegetation fine classification methods based on multi-source remote sensing data

LI Yuyang¹, LIU Fang^{2*}, HU Wenjun², LI Qin², FENG Yuze¹, YANG Changbao¹, CHEN Tianyi¹

1. College of Geo-Exploration Science and Technology, Jilin University, Changchun 130026, China;

2. Yunnan Provincial Institute of Geological Environmental Monitoring, Kunming 650216, China

Abstract: Aiming to the unique temperate forest vegetation within the Northeast Tiger and Leopard National Park, a study on multi-source remote sensing image coverage classification has been conducted, which aims to utilize modern remote sensing techniques and advanced machine learning methods to perform fine classification of

收稿日期/Received: 2024-08-31; 修订日期/Revised: 2024-11-09; 出版日期/Published: 2025-05-25

基金项目: 国家自然科学基金项目 (42130805、42074154)

第一作者: 李雨阳 (2000—), 男, 硕士研究生, 主要从事遥感和地理信息系统的研究应用。E-mail: yuyangl22@mails.jlu.edu.cn

* 通信作者: 刘舫 (1986—), 女, 高级工程师, 主要从事遥感地质、矿山生态修复研究。E-mail: fangzhouyn@163.com

© Editorial Office of World Geology. This is an open access article under the CC BY-NC-ND 4.0 license.

complex vegetation types, thereby better understanding the ecological conditions of the region and providing scientific basis for the conservation efforts in the Northeast Tiger and Leopard National Park. Firstly, the authors employed multi-scale image segmentation technology to decompose the original remote sensing images into a series of small regions with similar texture and color features, which helps reduce data redundancy and highlight different characteristics of vegetation. Subsequently, these segmented images has been optimized by extracting feature parameters that effectively distinguish different types of vegetation. Then, the authors input these optimized feature parameters into three different machine learning classifiers for training and testing, including support vector machine (SVM), C5.0 decision tree and random forest. The experimental results show that, in terms of classification accuracy, random forest achieving the highest accuracy of 91.33%, and C5.0 decision tree reaching 89.59%, with the Kappa coefficients of 0.90 and 0.88, respectively. In comparison, the classification accuracy of support vector machine was lower, at only 58.96%, and the Kappa coefficient is 0.54. These results indicate that, for distinguishing vegetation cover types and identifying other land features, the random forest algorithm has significant advantages. It not only accurately distinguishes different vegetation types in complex backgrounds, but also effectively identifies other features such as roads and buildings.

Keywords: multi-source remote sensing image; vegetation cover classification; support vector machine; decision tree; random forest

0 引言

植被是地表覆盖的植物群落总称,是生态系统中至关重要的组成部分,包括林地、灌丛以及农作物等^[1]。植被不仅具有保护水源、防止土壤流失和沙漠化等功能,还能够降低气温、净化空气,并对气候起到调节作用^[2]。因此,植被覆盖相关研究一直是遥感应用研究的热点问题,受到了国内外学者的广泛关注^[3]。

近年来,随着卫星传感器技术的不断进步,高分辨率遥感数据的应用逐渐成为分类研究中的主流。高分辨率遥感影像的应用使得植被分类精度得以提高。光学图像具有丰富的光谱信息,可以利用可见光到红外波段的植被反射光谱曲线来区分不同类型的森林^[4]。然而,光学遥感存在监测时间受限和图像识别难度大的问题,因为图像中可能存在“同物异谱”和“同谱异物”的现象,导致部分地物信息无法准确提取。相比之下,合成孔径雷达具有全天候、全天时成像能力,并且具有较强的穿透力和高质量的影像,提供了更丰富有效的数据。因此,如何充分利用光学和雷达遥感的优势,结合光学卫星影像的光谱信息和微波卫星回波信息,成为解决遥感图像植被识别的有效途径^[5-7]。如Haywood等^[8]利用增强回归树模型和高质量地面调查数据,在机器学习框架内融合 Sentinel-2 和 Sentinel-1 数

据,进而可靠地估计地上森林生物量;Kaplan等^[9]将 SAR 影像与光学影像融合,有效地监测湿地和湿地植被的生物特征,提高了湿地分类精度;Kussul 等^[10]利用多时相 Landsat8 与合成孔径雷达(synthetic aperture radar, SAR)影像对乌克兰农作物进行分类,其结果精度超过 90%,为广泛的商业农业应用开辟了新的可能性;郭交等^[11]采用最大似然法和支持向量机两种方法,融合哨兵 1 号雷达影像和哨兵 2 号光学影像,提高了农作物的分类精度;刘戈等^[12]以哨兵 2 号遥感影像为数据源,设计了基于特征优选的卷积神经网络(convolutional neural network, CNN)分类方法,实现了对农作物的精细分类;陈媛媛等^[13]利用 Sentinel-1A 和 Sentinel-2A 数据进行基于面向对象的决策树方法对丽水市莲都区进行分类和精度评价,为高植被覆盖度区域的信息提取提供了一定的参考。

遥感影像分类的算法众多,但并没有特定的分类方法较其他算法有较大的精度优势。不同的数据源与地物类型适用不同的分类算法,因此,针对遥感图像植被分类中选择不同分类器的比较也是遥感影像分类研究中的重要一环。

东北虎豹国家公园作为重要的温带森林植物物种保护区,具有极高的研究价值。本研究将哨兵一号合成孔径雷达影像与哨兵二号光学影像组合成多源遥感影像,旨在解决光学影像监测时间限制和光

谱混淆问题, 提升植被覆盖分类的准确性和稳定性。首先, 确定了最优的多源组合遥感影像尺度, 并采用了面向对象的影像多尺度分割技术, 优化了影像特征提取过程。最后, 通过对比分析支持向量机、决策树和随机森林等多种机器学习分类器在精细分类中的表现, 为该地区遥感影像植被覆盖分类提供了可靠的参考和实用指导。

1 研究区概况

研究区位于东北虎豹国家公园的一部分, 地处吉林、黑龙江两省交界的老爷岭南部区域 ($130^{\circ}41'38.283'' \sim 131^{\circ}07'38.215''E$, $43^{\circ}20'43.475'' \sim 43^{\circ}38'49.433''N$) (图 1)。东部起始于吉林省珲春林业局青龙台林场, 与俄罗斯边境地区接壤, 西部至吉林大兴沟林业局岭东林场, 最北为黑龙江三道林场, 南至珲春敬信林场^[14]。虎豹公园地处中俄朝三国交界的连接地带, 地貌复杂多样, 主要以丘陵峡谷为特征, 同时还分布有其他地貌类型, 如盆地、台地及平原等。虎豹公园的海拔均在 1 500 m 以下, 最高山峰老爷岭高度为 1 447 m, 主要山体高度在 1 000 m 以下, 山脉连绵起伏, 沟壑纵横, 整体地势从中部向四周逐渐下降, 南北两端为海拔较低的山谷和山地, 受上游水系侵蚀的影响, 存在许多熔岩峡谷。虎豹公园山体坡面的侵蚀程度较弱。

东北虎豹国家公园位于亚洲温带针阔叶混交林生态系统的中心地带, 森林生态系统是该公园的主体。该地区森林面积广阔, 主要植被类型是温带针阔叶混交林。此外, 矿产资源也非常丰富, 已知的金属及非金属矿物高达几十种, 煤炭、油页岩、石灰石、黄金、花岗岩、大理石、玄武岩、铁矿石、钨、钼等矿产储量丰富, 已有 30 多种矿产被开发。东北虎豹国家公园是野生动植物分布最丰富的地区之一, 选择该研究区域可以最大程度地保证研究结果的代表性。

2 数据处理

2.1 影像数据获取

哨兵 (Sentinel) 系列卫星是欧洲哥白尼计划的专用系列卫星。

Sentinel-1 卫星作为高分辨率合成孔径雷达卫星, 自身携带的 C 频段合成孔径雷达 (SAR) 可

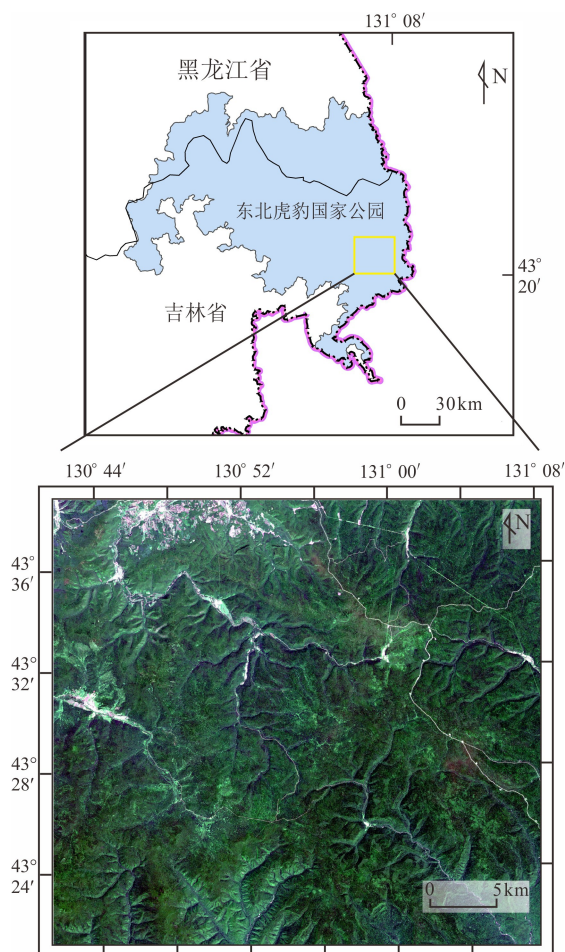


图 1 研究区示位置示意图

Fig. 1 Overview of study area

昼夜成像, 具有全天候的成像能力, 属于主动微波遥感卫星^[15]。Sentinel-1 卫星采用太阳同步轨道, 卫星轨道高度 693 km, 倾角 98.18° , 轨道周期 99 min, 重访周期 12 d。实验研究选择了采集时间为 2020 年 9 月 15 号哨兵 1 号双极化 VV-VH 合成孔径雷达影像, 影像成像模式为干涉宽幅 (interference width, IW), 分辨率为 10 m。

Sentinel-2 卫星是一颗高分辨率多光谱成像卫星, 自身携带一枚多光谱成像仪 (multispectral imager, MSI), 卫星具有高分辨率与高重访的特性, 主要用于全球陆地观测, 包括陆地植被、土壤和水资源、内陆水道和沿海地区以及发布相关的应急信息^[16]。Sentinel-2 卫星采用太阳同步轨道, 轨道高度 786 km, 倾角 98.50° , 轨道重访周期为 5 d。全球陆地表面成像数据每 10 d 更新一次。每个轨道周期的平均观测时间为 16.3 min, 峰值为

31.0 min。卫星上的多光谱成像仪工作波段为可见光、近红外及短波红外，覆盖 13 个光谱波段，部分光谱波段属性见表 1。多光谱图像的幅宽为 290 km，地面分辨率分别为 10 m、20 m、60 m。实验研究选择采集于 2020 年 8 月 25 号哨兵 2 号光学遥感影像，影像包含 13 个波段，选择其中 7 个波段（波段：2、3、4、5、6、7、8：蓝、绿、红、红边、近红外）进行植被分类，并将应用于分类的波段分辨率重采样到 10 m。影像云量覆盖为 0%。

表 1 Sentinel-2 传感器光谱波段属性
Table 1 Spectral band attributes of Sentinel-2 sensor

波段	描述	中心波长/ μm	分辨率/m
Band 1	超蓝（沿海和气溶胶）	0.443	60
Band 2	蓝色	0.490	10
Band 3	绿色	0.560	10
Band 4	红色	0.665	10
Band 5	红边波段	0.705	20
Band 6	红边波段	0.740	20
Band 7	红边波段	0.783	20
Band 8	近红外波段（宽）	0.842	10
Band 8A	近红外波段（窄）	0.865	20
Band 9	水蒸气波段	0.945	60
Band 10	短波红外波段	1.375	60
Band 11	短波红外波段	1.610	20
Band 12	短波红外波段	2.190	20

2.2 采样点的获取与选择

根据东北虎豹公园研究区遥感影像的目视解译成果以及实地取样考察结果，精心选择了检验点和样本点（图 2），将研究区划分为 10 类地物，包括阔叶林、针叶林、稀疏阔叶林、针阔混交林、采石场采矿场、旱地、湿地、城镇住宅用地、水体和道路。

2.3 数据预处理

Sentinel-1 影像数据预处理主要包括数据输入、辐射定标、多视合成、滤波和地形校正等步骤。使用 SNAP 软件对研究区域的双极化数据进行了预处理。

对于 Sentinel-2 L1C 遥感影像的大气校正，采用了 Sen2Cor 工具。Sen2Cor 是由欧洲航天局（European Space Agency, ESA）官方提供的工具集，用于生成和格式化 Sentinel-2 Level 2A 产品。

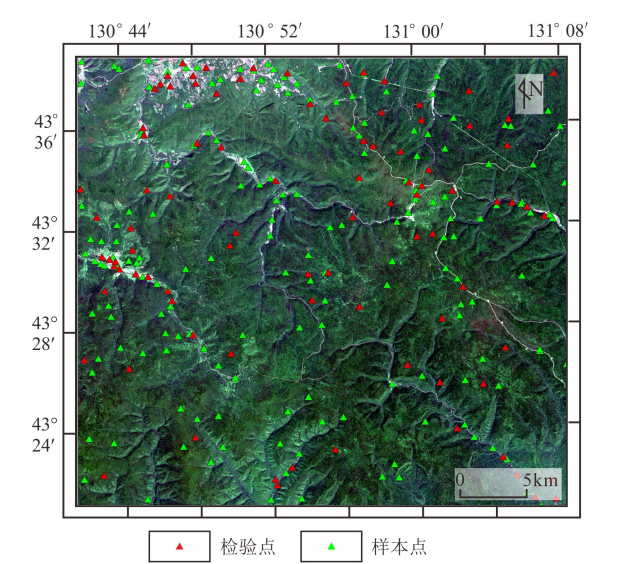


图 2 检验点与样本点分布
Fig. 2 Distribution of test points and sample points

3 方法原理

3.1 多尺度分割

在传统的光学遥感中，遥感影像的分析主要以像元为主。然而，随着图像分辨率的不断提高，影像信息变得越来越丰富。除了自身的光谱特征外，影像地物还包括独特的几何特征和纹理特征。与此同时，影像内部的光谱差异也日益显著，简单的单元像素属性已无法满足应用的需求。因此，研究人员引入了边缘检测和区域块提取等思想，发展出了面向对象的影像分割方法^[17]。目前常用的面向对象影像分割算法包括多尺度分割、均值漂移分割和分水岭变换分割等。

多尺度分割方法采用了模糊子集理论，在指定的感兴趣区域目标对应的尺度下，结合光谱特征、位置形状特征等，将影像分割成连结的同质性较高、异质性较低的影像区域^[18]。

多尺度分割采用了分形网络演化算法（fractal net evolution approach, FNEA）（基于异质性最小区域合并），该算法采用自底向上的增长区域方法。异质性主要表现为对象的光谱和形状异质性，其增加必须低于某一阈值。

具体而言，FNEA 算法以像元层为基础，选择合适的尺度阈值作为合并的结束条件。合并准则依据光谱和形状特征加权的异质性参数评价。首先，

从像元开始, 将特征相似的相邻区域合并形成次级对象的区域。然后, 基于异质性最小原则, 合并特征相似的次级区域。在每次合并时, 都会判断合并后区域的异质性是否超过尺度阈值。如果超过了阈值, 则不进行合并; 如果小于阈值, 则继续合并次级区域。这个过程循环进行, 直到达到合并条件的尺度阈值, 最终完成多尺度的分割^[19]。

3.2 多源影像特征优化

在多源遥感影像面向对象影像分类的过程中, 为了克服仅依赖光谱特征的局限性, 笔者选择利用丰富的雷达影像和光学影像特征信息进行植被分类。分割对象的特征值基于对象内部像元的特征和关系进行计算, 包括图 3 所展示的光谱特征、几何特征、纹理特征和植被指数。然而, 随着对影像特征的不完善和应用, 过多的特征信息可能导致分类过程中的信息冗余和无关特征的出现, 进而影响植被分类的准确性, 甚至引发维数灾难^[20]。因此, 对多源影像的特征信息需要进行优化筛选和降维, 以确定最佳的分类特征组合。

本研究基于哨兵 1 号和哨兵 2 号波段组合影像, 根据 VV 极化波段、VH 极化波段、蓝色波段、绿色波段、红色波段、红边波段和近红外波段的波段特征进行特征提取, 并构建面向对象的特征空间。为评估特征的重要性, 采用随机森林特征选择算法进行特征重要性评价。

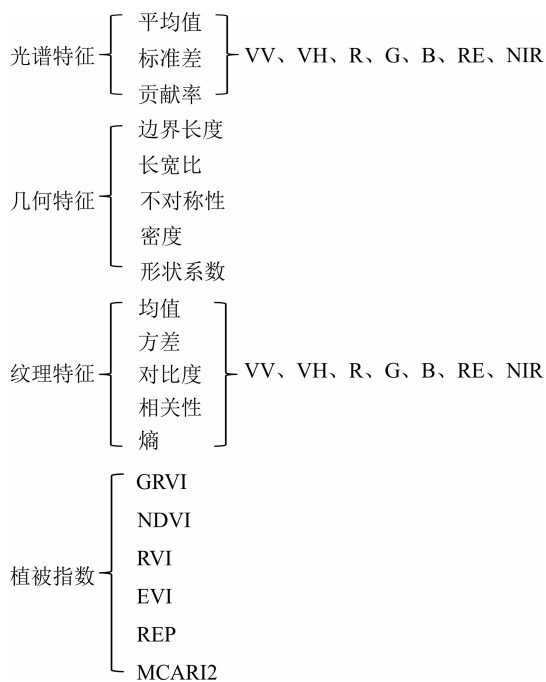


图 3 未筛选特征属性空间

Fig. 3 Feature attribute space before screening

随机森林是一种包裹式特征筛选方法, 使用一组决策树进行标签预测, 并计算每个特征的重要性。如图 4 所示, 每棵决策树都使用不同的训练数据集和特征子集进行训练, 并将其预测结果综合起来得到最终的预测结果。

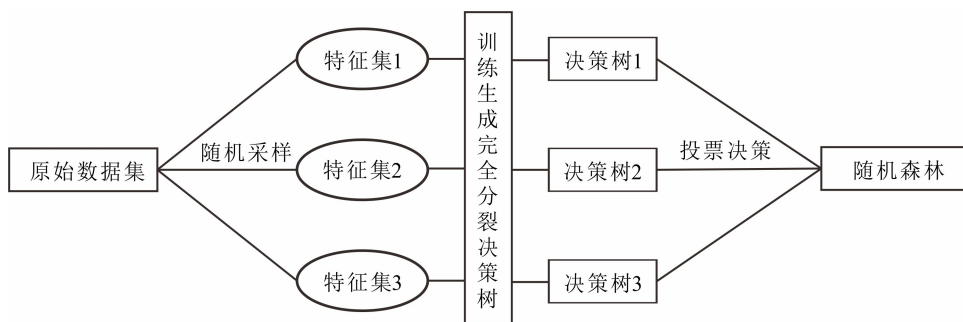


图 4 随机森林模型

Fig. 4 Random forest model

3.3 分类方法

分类算法通常包括面向对象的支持向量机、C5.0 决策树和随机森林。在对遥感影像进行植被分类后, 必须对其进行精度评价。精度评价是影像分类中不可或缺的一部分, 可以有效地分析和评估

不同分类方法, 了解其优缺点, 以提高植被地物分类的准确性。目前, 最常用的方法是混淆矩阵法, 并使用 Kappa 系数评估分类结果的准确性。

混淆矩阵是评估分类器性能的常用工具, 用于表示预测结果与真实标签之间的关系。混淆矩阵是

一个 n 行 n 列的方阵, 其中行值表示实际地物类别在遥感影像中的数量, 而列值表示分类结果中对应地物类别的数量。主对角线上的样本数量越多, 说明分类的准确性越高。制图精度是指某一地物类型中正确分类的样本数量除以参考数据中该类型样本的总数 (按列方向计算), 反映了该地物类型在地图上正确分类的概率。用户精度是指某一地物类型中正确分类的样本数量除以分类样本中该类型样本的总数 (按行方向计算), 反映了该地物类型被正确分类的概率。总体分类精度是指主对角线上所有正确分类的样本数量与所有样本数量的比值, 反映了分类地物类别与实际地物类别一致的概率。

Kappa 系数用于评价整体混淆矩阵, 是衡量分类结果一致性和准确性的重要指标。Kappa 系数的取值范围为 0~1, 数值越接近 1, 说明分类结果的准确性越高。

3.3.1 SVM 支持向量机

支持向量机 (support vector machine, SVM) 是一种常用于解决分类和回归问题的监督学习模型, 既适用于二分类问题也适用于多分类问题。SVM 的目标是在给定训练数据的情况下, 构建一个能够对新数据进行准确分类的模型^[21]。

SVM 的主要思想是要找到一个能够将两类数据分开的超平面, 并使得两类数据点到超平面的距离尽可能远, 这个距离被称为间隔。间隔最大化的目标是使得这个间隔尽可能大。在间隔最大化的过程中, 只有部分数据点对确定超平面起着关键作用, 被称为支持向量, 这些支持向量决定了超平面的位置。SVM 的一个重要特性是可以使用核函数将输入空间映射到更高维的特征空间, 从而使得非线性问题变成线性可分或近似可分的问题。常用的核函数包括线性核、多项式核、径向基函数核等。

支持向量机具有良好的泛化性能, 特别是在处理高维数据时表现出色。支持向量机能够有效地防止过拟合, 这意味着它不会过度依赖于训练数据, 而是能够很好地适应新的、未见过的数据, 这使得 SVM 成为处理复杂数据集的一种强大工具。支持向量机的另一个优点是其对于异常值的鲁棒性。如图 5 所示, 由于 SVM 的决策是基于支持向量, 即距离超平面最近的一些数据点, 因此它对于那些远离决策边界的数据点不敏感, 所以 SVM 在处理包含噪声的数据时表现出色。

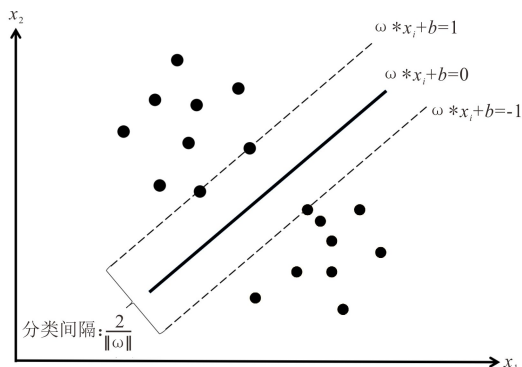


图 5 SVM 超平面与间隔

Fig. 5 SVM hyperplane and margin

3.3.2 C5.0 决策树算法

决策树是一种基于树状结构的监督学习算法, 常用于解决分类和回归问题。它通过对数据特征的选择和判断, 从根节点开始将数据分成越来越小的组, 直到数据被完全分离为一个叶子节点。每个内部节点代表一个特征判断, 而每个叶子节点代表一个类的预测^[22]。

C5.0 是决策树算法中的经典之作, 是对 ID3 和 C4.5 算法的改进和扩展。由 Quinlan^[23] 于 1993 年提出, C5.0 在 C4.5 的基础上做了重要改进。该算法采用了启发式方法来构建决策树, 旨在生成更小、更快、更准确的模型^[24]。在构建过程中, C5.0 首先对数据进行预处理, 包括处理缺失值和异常值等。这一预处理的目的是确保决策树的质量和性能, 并提升模型的泛化能力。C5.0 利用信息增益、信息增益比等指标来选择最优特征进行节点划分, 同时引入了悲观剪枝技术, 使得特征选择更加准确。

该算法采用自顶向下的贪婪策略, 递归地在每个节点上选择最优的特征进行划分, 直到满足停止条件。在节点划分过程中, C5.0 采用了增强的剪枝策略以防止过拟合。完成决策树的构建和剪枝后, C5.0 可以用于分类和预测任务^[25]。

C5.0 在特征选择、剪枝等方面做出了改进, 以提高模型的准确性和泛化能力。作为一种强大且灵活的机器学习算法, C5.0 在实际应用中被广泛使用。

3.3.3 随机森林算法

随机森林 (random forest) 是一种集成学习方

法, 旨在通过构建多个决策树并将它们组合起来进行预测, 从而提高整体模型的性能和稳定性。决策树作为随机森林的基石, 采用树形结构进行分类, 根据特征的条件进行决策。在随机森林中, 每个决策树都充当基学习器, 独立地从训练数据中学习并做出预测。

在构建每棵决策树的过程中, 随机森林采用了特征的随机选择, 它从所有特征中随机选取一部分作为候选特征, 而不是使用全部特征, 以增加模型的多样性, 减少决策树之间的相关性, 从而提高整体模型的泛化能力。同时, 随机森林通过随机采样的方式生成每棵决策树的训练数据集。通常情况下, 采用自助采样, 即从原始数据集中有放回地抽取与原始数据集大小相同的样本作为训练数据。这种方式使得每棵决策树都具有略有差异的训练数据, 增加了模型的多样性。

针对每个随机选择的特征和随机采样的训练数据, 构建一棵决策树。决策树的构建过程通常采用递归的方式, 通过对特征进行划分, 使得每个子节点中的样本尽可能属于同一类别或具有相似的特征。在每个节点上, 通过选择最优的特征和划分方式来建立决策树。当进行预测时, 随机森林中的每棵决策树都会对样本进行分类或回归预测。对于分类问题, 通常采用投票的方式, 即将每棵决策树的分类结果进行统计, 选择得票最多的类别作为最终的预测结果^[26]。

随机森林能够有效地减少过拟合现象, 并且通常具有很好的泛化能力。由于随机森林中的每棵决策树都是独立训练的, 因此它们可以并行生成, 从而有效提高了训练速度。同时, 随机森林也适用于处理高维数据和具有复杂特征关系的数据集。

4 结果分析

4.1 多尺度分割结果

在进行多尺度分割时, 本文采用逐渐增加分割尺度的方法, 从小到大对每个尺度进行分割, 以寻找最优尺度。这一策略旨在提高分割的精度和准确性, 同时有助于从图像中提取更多的有用信息。对于地物而言, 最优分割尺度应当使得分割结果的多边形能够清晰地显示不同种类地物的边界, 避免产生过于破碎或模糊的边界。

这种多尺度分割的方法有助于克服单一尺度下可能出现的局限性, 使得分割更加全面和准确。通过逐步增加尺度, 可以更好地捕捉地物的细节信息, 提高分割结果的质量。图 6 为最终分割结果, 呈现出清晰而准确的地物边界, 验证了本文采用的多尺度分割策略的有效性。

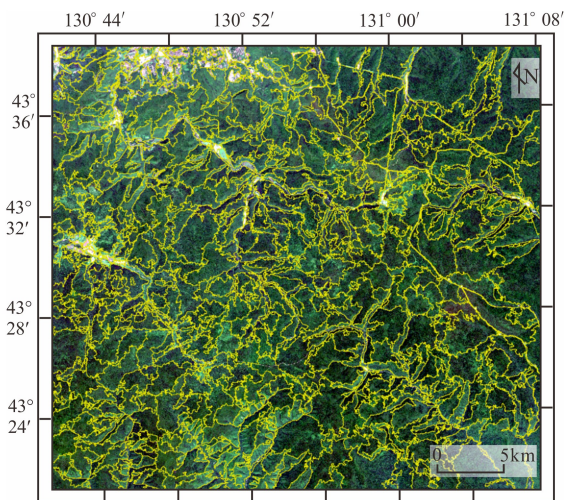


图 6 多尺度分割结果

Fig. 6 Multi-scale segmentation results

4.2 多源影像特征优化结果

实验的特征来自多源遥感组合影像, 这些影像包含多个波段, 因此用于分类的特征数量较多。根据其重要性, 利用随机森林算法对 89 种特征进行排序, 首先从高到低列出前 30 种重要性的特征如图 7 所示。

结果显示, 光学影像的光谱特征的均值及标准差在重要性评分中占据较高的位置, 其中红绿蓝波段的均值被评定为最重要的。在前 30 种特征中, 灰度共生矩阵纹理特征数量占比最大, 特别是 VV 极化后向散射系数纹理均值特征在重要性评价中表现较高, 表明在传统光学影像分类中引入 SAR 影像波段可以显著提升影像的分类精度。此外, 光学影像波段的纹理特征和植被专题指数的重要性排名也相当靠前, 说明在提取高分辨率光学影像的纹理特征和专题指数后, 对不同植被类型和土地利用种类的描述更为详细, 从而有效提高了影像的分类精度。最终选取的特征参数如表 2 所示。

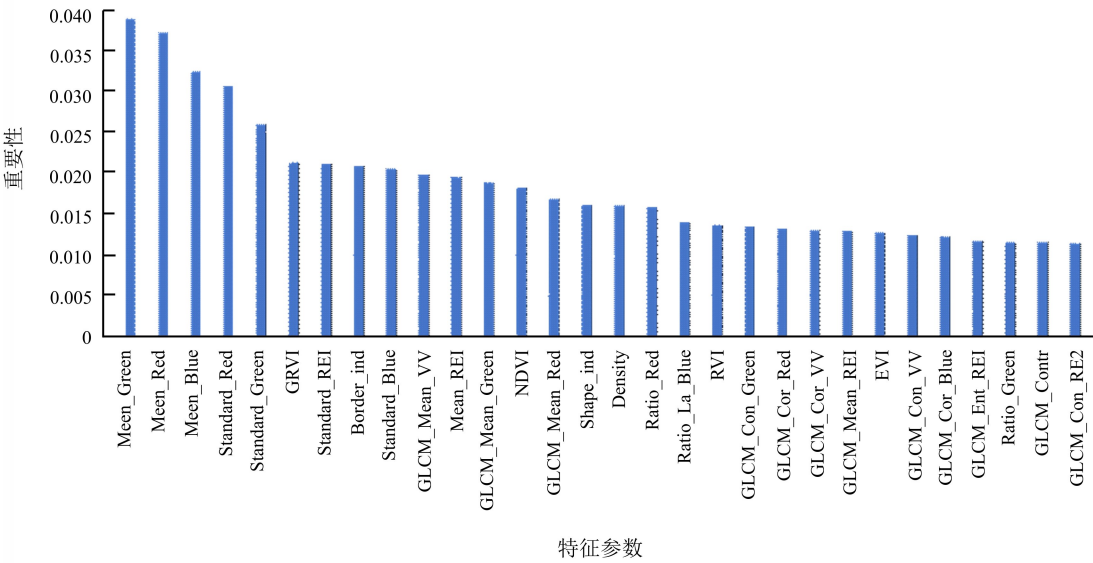


图 7 组合影像特征重要性排序

Fig. 7 Importance ranking of composite image features

表 2 特征参量
Table 2 Feature parameters

特征种类	特征参量
光谱特征	Mean _ Green、Mean _ Red、Mean _ Blue、Standard _
	Red、Standard _ Green、Standard _ REI、Standard _ Blue、Mean _ REI、Ratio _ Red、Ratio _ Blue
几何特征	Border _ index、Shape _ index、Density
纹理特征	GLCM _ Mean _ VV、GLCM _ Mean _ Green、GLCM _
	Mean _ Red、GLCM _ Con _ Green、GLCM _ Cor _ Red、GLCM _ Cor _ VV、GLCM _ Mean _ REI
植被指数	GRVI、NDVI、RVI

4.3 多源遥感组合影像分类结果

为了研究不同分类器对多源组合影像的分类效果，分别采用了支持向量机算法、决策树 C5.0 算法以及随机森林算法，并结合筛选优化的特征变量进行了植被及其他地物类型的分类。具体的分类结果见图 8 ~ 10。

从图 8 中支持向量机算法的分类结果可以看出，整体上植被覆盖主要为阔叶林，针叶林主要分布在东南部区域，而针阔混交林则集中在西北区域。湿地覆盖面积相对较广，主要分布在中部地区。然而，对具体地物类型的分布分析显示，支持向量机分类结果存在一定的误判现象。例如，将研究区中部植被覆盖地区误判为湿地或水体，同时东

北方向的道路被错误地归类为采矿用地。因此，支持向量机算法的分类结果显示出准确度还有待提升。

对于图 9 中决策树算法的分类结果来说，整体上植被分布主要为西北部的阔叶林和稀疏阔叶林，而东南部则主要为针叶林和针阔混交林。具体地物类型的分布分析显示，城市住宅用地多分布在阔叶林附近，而旱地一般则环绕在城镇住宅用地周围，湿地则多集中在水体附近。虽然决策树 C5.0 算法的整体分类效果较好，但仍存在部分区域对象的误分类现象，例如将一些道路误判为采矿用地。这可能是由于采矿用地多分布于道路两侧，容易被算法划分为同一对象，从而导致误分类现象。

图 10 中随机森林算法的分类结果显示，整体上植被分布与决策树算法相似，西北部为主要的阔叶林和稀疏阔叶林，东南部为主要的针叶林和针阔混交林。对其他地物类型的分析显示，旱地一般环绕在城镇住宅用地周围，而湿地多集中在水体附近。与决策树算法相比，随机森林算法的湿地分类结果占地面积较小且更为合理。总体来看，随机森林算法的植被分类效果优于前两种算法，尤其是在针叶林和阔叶林的区分方面表现更好。

按照混淆矩阵的计算方法，分别计算了面向对象支持向量机分类、决策树 C5.0 算法分类以及随

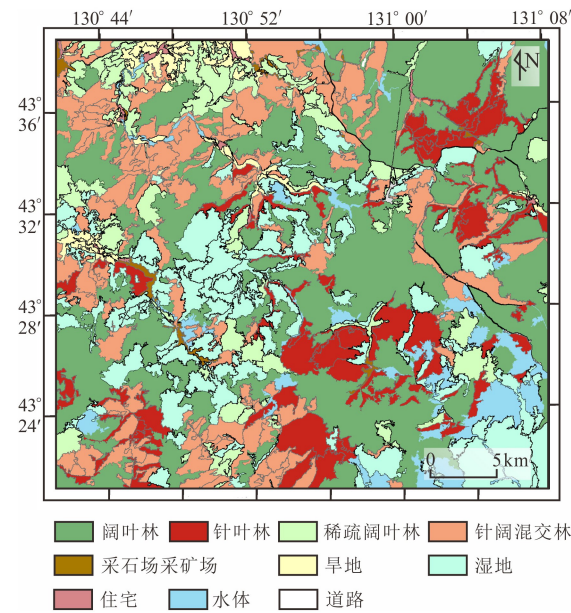


图 8 支持向量机算法面向对象分类结果
Fig. 8 Object-oriented classification results of SVM algorithm

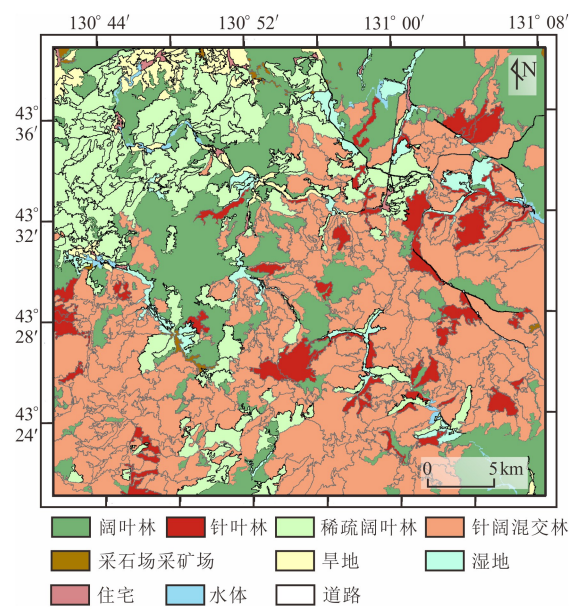


图 9 决策树算法面向对象分类结果
Fig. 9 Object-oriented classification results of decision tree algorithm

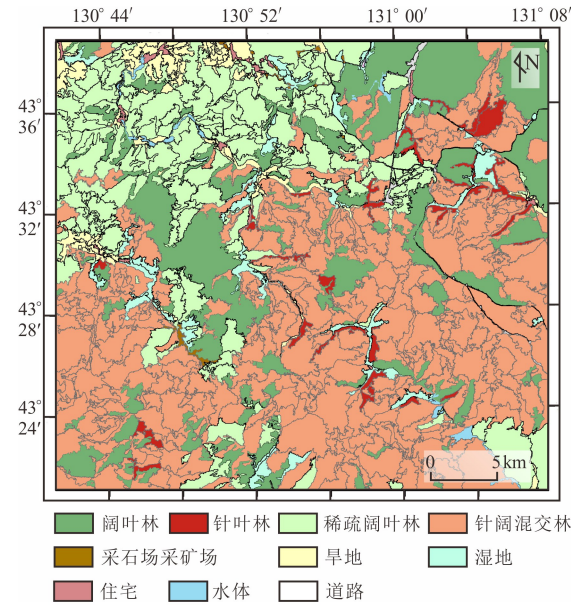


图 10 随机森林算法面向对象分类结果
Fig. 10 Object-oriented classification results of random forest algorithm

面向对象的 C5.0 决策树算法, 最低的是基于面向对象的支持向量机算法。其中, 随机森林分类的总体精度为 91.33%, Kappa 系数为 0.90; C5.0 决策树分类的总体精度为 89.59%, Kappa 系数为 0.88; 支持向量机分类的总体精度为 58.96%, Kappa 系数为 0.54 (表 3)。

其中支持向量机的性能高度依赖于核函数及其参数的选择。在研究中, 尽管进行了参数优化, 但仍可能存在参数选择不当的情况, 这可能会影响 SVM 的分类效果。与其他两种方法 (例如随机森林或深度学习方法) 相比, 支持向量机在处理非线性数据时可能存在一定的劣势, 其他方法可能更适合于捕捉复杂的数据模式和非线性关系, 在实验中表现更好。

通过对具体地物类型的制图精度和用户精度数据进行对比分析, 发现随机森林算法在采矿用地类型和湿地类型的精度低于 C5.0 决策树, 但在其他地物类型上的精度均高于其他两种算法分类, 尤其是对针叶林和阔叶林的分类效果明显优于其他算法, 制图精度分别达到了 95.87% 和 89.30%。综合来看, 采用 3 种分类器对东北虎园研究区进行面向对象植被覆盖分类时, 随机森林算法的分类效果最佳, 精度最高。

机森林算法分类的制图精度 (producer accuracy, PA)、用户精度 (user accuracy, UA)、总体分类精度 (overall accuracy, OA) 和 Kappa 系数, 并对 3 种分类算法的精度进行了比较评价。综合 3 种算法的分类结果和精度评价表可以看出: 基于面向对象的随机森林算法具有最高的分类精度, 其次是基于

表 3 3 种分类算法分类精度评价表

Table 3 Classification accuracy evaluation of three algorithms

地物类型	SVM		决策树		随机森林	
阔叶林	26.67	44.43	76.19	88.89	89.30	88.90
针叶林	62.58	83.33	88.89	80.10	95.87	97.80
稀疏阔叶林	53.84	36.84	89.10	99.50	98.60	94.11
针阔混交林	39.13	20.45	90.19	90.85	95.00	91.30
采石场采矿场	64.28	90.00	88.89	80.00	73.50	78.60
旱地	90.91	95.23	91.30	95.30	95.21	96.80
湿地	21.43	21.43	91.67	79.10	77.10	91.32
城镇住宅用地	98.65	86.67	87.50	93.00	96.80	76.47
水体	71.43	90.91	80.65	82.30	80.25	89.90
道路	92.86	92.87	83.62	85.71	87.50	87.50
PA/UA	PA	UA	PA	UA	PA	UA
OA/%	58.96		89.59		91.33	
Kappa	0.54		0.88		0.90	

5 结论

(1) 为提升面向对象分类结果的准确度，采用随机森林优化筛选降维，对从多源遥感数据不同波段提取的特征信息进行处理，最终将 89 种特征变量优化缩减到 23 种。

(2) 在进行基于多源遥感组合影像的面向对象植被精细分类时，通过比较随机森林、C5.0 决策树和支持向量机 3 种分类算法的表现，发现随机森林算法在分类精度上表现最为突出，达到了 91.33%，有效地实现了不同植被覆盖类型的区分及其他地物类型的识别。

参考文献:

[1] BROWN L R, MAGAGULA I P, BARRETT A S. A vegetation classification and description of Telperion Nature Reserve, Mpumalanga, South Africa [J]. Vegetation Classification and Survey, 2022, 3: 199–219.

[2] DÍAZVILLA M V E, CRISTIANO P M, EASDALE M, et al. Archetypal classification of vegetation dynamics of a humid subtropical forest region from North–East Argentina [J]. Remote Sensing Applications: Society and Environment, 2023, 30: 100966.

[3] HE Y Y. Research on the extraction technology of Earth remote sensing images based on computer BP neural network [C] //2022 IEEE 2nd International Conference on

Data Science and Computer Application (ICDSCA), Dalian, 2022: 1020–1024.

[4] SHEN Z, MIAO J, WANG J J, et al. Evaluating feature selection methods and machine learning algorithms for mapping mangrove forests using optical and synthetic aperture radar data [J]. Remote Sensing, 2023, 15 (23): 5621.

[5] 雷俊杰, 杨武年, 李红, 等. 哨兵光学及 SAR 卫星影像协同分类研究 [J]. 现代电子技术, 2022, 45 (2): 135–139.

LEI J J, YANG W N, LI H, et al. Research on collaborative classification of sentinel optical and SAR satellite images [J]. Modern Electronic Technology, 2022, 45 (2): 135–139.

[6] GAŠPAROVIĆ M, KLOBUČAR D. Mapping floods in low-land forest using Sentinel-1 and Sentinel-2 data and an object-based approach [J]. Forests, 2021, 12 (5): 553.

[7] 李林. 极化 SAR 和光学影像的土地覆盖协同分类 [D]. 南京: 东南大学, 2022.

LI L. Synergistic classification of land cover using polarimetric SAR and optical imagery [D]. Nanjing: Southeast University, 2022.

[8] HAYWOOD A, STONE C, JONES S. The potential of sentinel satellites for large area above ground forest biomass mapping [C] //IGARSS 2018 – 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018: 9030–9033.

- [9] KAPLAN G, AVDAN U. Sentinel-1 and Sentinel-2 data fusion for mapping and monitoring wetlands [M]. Beijing: Preprints, 2018: 729-734.
- [10] KUSSUL N, LEMOINE G, GALLEGO F J, et al. Parcel-based crop classification in Ukraine using Landsat-8 data and Sentinel-1A data [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9 (6): 2500-2508.
- [11] 郭交, 朱琳, 靳标. 基于 Sentinel-1 和 Sentinel-2 数据融合的农作物分类 [J]. 农业机械学报, 2018, 49 (4): 192-198.
GUO J, ZHU L, JIN B. Crop classification based on fusion of Sentinel-1 and Sentinel-2 data [J]. Transactions of the Chinese Society of Agricultural Machinery, 2018, 49 (4): 192-198.
- [12] 刘戈, 姜小光, 唐伯惠. 特征优选与卷积神经网络在农作物精细分类中的应用研究 [J]. 地球信息科学学报, 2021, 23 (6): 1071-1081.
LIU G, JIANG X G, TANG B H. Research on feature selection and convolutional neural networks for precision crop classification [J]. Journal of Earth Information Science, 2021, 23 (6): 1071-1081.
- [13] 陈媛媛, 雷鸣, 王泽远, 等. 基于 Sentinel 卫星影像的土地利用类型提取: 以丽水市莲都区为例 [J]. 森林工程, 2022, 38 (2): 54-61.
CHEN Y Y, LEI M, WANG Z Y, et al. Extraction of land use types based on sentinel satellite images: a case study of Liandu District, Lishui City [J]. Forest Engineering, 2022, 38 (2): 54-61.
- [14] 魏文丽. 基于遥感与 GIS 的野生东北虎豹生境适宜性研究 [D]. 长春: 吉林大学, 2019.
WEI W L. Study on habitat suitability of wild Northeast Tiger and Leopard based on remote sensing and GIS [D]. Changchun: Jilin University, 2019.
- [15] HRISHIKESH K, TAJDARUL H S, FALK A, et al. Space-time evolution of land subsidence in the National Capital Region of India using ALOS-1 and Sentinel-1 SAR data: evidence for groundwater over exploitation [J]. Journal of Hydrology, 2022, 605: 127329.
- [16] NASIR F L, ROSHANAK D, ANDREW K S, et al. Integrating semi-supervised learning with an expert system for vegetation cover classification using Sentinel-2 and Rapid Eye data [J]. Remote Sensing, 2022, 14 (15): 3605.
- [17] ZHANG R J, SHEN Y L. Multi-scale expression of coastal landform in remote sensing images considering texture features [J]. Remote Sensing, 2024, 16: 3862.
- [18] YE W H, ZHANG W, LEI W M, et al. Remote sensing image instance segmentation network with transformer and multi-scale feature representation [J]. Expert Systems with Applications, 2023, 234: 121007.
- [19] 孙坤, 鲁铁定. 顾及多尺度分割参数的 FNEA 面向对象分类 [J]. 测绘通报, 2018 (3): 43-48.
SUN K, LU T D. FNEA object-oriented classification considering multi-scale segmentation parameters [J]. Bulletin of Surveying and Mapping, 2018 (3): 43-48.
- [20] 张杲. 面向对象的光学遥感影像特征提取与分类研究 [D]. 郑州: 河南大学, 2021.
ZHANG G. Research on object oriented optical remote sensing image feature extraction and classification [D]. Zhengzhou: Henan University, 2021.
- [21] 张怡. 基于贝叶斯算法优化支持向量机的遥感分类研究 [D]. 上海: 华东师范大学, 2022.
ZHANG Y. Research on remote sensing classification based on Bayesian algorithm optimization support vector machine [D]. Shanghai: East China Normal University, 2022.
- [22] 许伟. 基于遥感影像与 C5.0 数据挖掘算法的土壤制图研究 [D]. 武汉: 华中农业大学, 2019.
XU W. Research on soil mapping based on remote sensing images and C5.0 data mining algorithm [D]. Wuhan: Huazhong Agricultural University, 2019.
- [23] QUINLAN J R. C4.5: programs for machine learning [M]. San Mateo, Calif: Morgan Kaufmann Publishers, 1993: 1-302.
- [24] RAMO R, GARCÍA M, RODRÍGUEZ D, et al. A data mining approach for global burned area mapping [J]. International Journal of Applied Earth Observation and Geoinformation, 2018, 73: 39-51.
- [25] SIKNUN G P, SITANGANG I S. Web-based classification application for forest fire data using the Shiny framework and the C5.0 algorithm [J]. Procedia Environmental Sciences, 2016, 33: 332-339.
- [26] AMINI S, SABER M, RABIEI-DASTJERDI H, et al. Urban land use and land cover change analysis using random forest classification of landsat time series [J]. Remote Sensing, 2022, 14: 2654.