

赵冀唯, 伊丕源. 面向矿物填图的典型高光谱图像分类方法对比研究[J]. 世界核地质科学, 2025, 42(2): 385-399. doi: 10.3969/j.issn.1672-0636.2025.02.013

ZHAO Jiwei, YI piyuan. Study on mineral identification and classification based on airborne hyperspectral imagery[J]. World Nuclear Geoscience, 2025, 42(2): 385-399(in Chinese).

面向矿物填图的典型高光谱图像分类方法对比研究

赵冀唯, 伊丕源

- 1 铀资源探采与核遥感全国重点实验室, 北京 100029
- 2 核工业北京地质研究院, 北京 100029
- 3 中核集团铀资源勘查与评价技术重点实验室, 北京 100029

摘要 高光谱遥感技术以其高光谱分辨率和丰富的光谱信息在诸多领域得到广泛应用。而高光谱图像分类是充分利用高光谱数据的关键技术之一, 旨在根据像元的光谱特征将其划分为不同的类别。在进行了调研与归纳高光谱图像分类技术的研究现状基础上, 对甘肃金昌以北的矿区进行试验, 主要从监督分类和非监督分类两个方面进行对比分析, 以光谱角法为例深入讨论了影响不同方法分类性能的关键因素。结果表明: 对于本实验区这类空间特征不足的区域, 监督分类方法的精度整体优于非监督分类, 其中最大似然法效果最佳, 同时也证明非监督分类并不适用于这类区域的矿物分类。

关键词 矿物填图; 高光谱遥感; 监督分类; 非监督分类

中图分类号 P627; P628 文献标志码 A 文章编号 1672-0636 (2025) 02-0385-15

Study on mineral identification and classification based on airborne hyperspectral imagery

ZHAO Jiwei, YI Piyuan

- 1 National Key Laboratory of Uranium Resources Exploration and Exploitation and Nuclear Remote Sensing, Beijing 100029, China
- 2 Beijing Research Institute of Uranium Geology, Beijing 100029, China
- 3 CNNC Key Laboratory of Uranium Resource Exploration and Evaluation Techniques, Beijing 100029, China

Abstract: Hyperspectral remote sensing technology has been widely used in many fields due to its high spectral resolution and rich spectral information. Object classification is one of the key techniques to fully use the hyperspectral data. Based on the investigation and summary of the research status of hyperspectral image classification technology, experiments were conducted in the mining area north of Jinchang, Gansu province. A comparative analysis was mainly carried out from two aspects: supervised classification and unsupervised classification. Taking the spectral Angle method as an example, the key factors affecting the classification

基金项目: 多模态卫星遥感数据接引设计及典型应用关键技术研究(编号: WDZC_2023_HDYY_101)资助

收稿日期: 2025-02-12 改回日期: 2025-03-15

第一作者: 赵冀唯, 男, 1995年生, 在读硕士研究生, 科研方向为遥感图像处理。E-mail: zhaojw1230@163.com

Supported by Multi-modal satellite remote sensing data access design and typical application key technology research (WDZC_2023_HDYY_101)

Received date: 2025-02-12 revised date: 2025-03-15

First author: ZHAO Jiwei, male, born in 1995, master's degree candidate, focusing on remote sensing image processing. E-mail: zhaojw1230@163.com

performance of different methods were deeply discussed. The results show that the accuracy of supervised classification method is better than that of unsupervised classification for the areas with insufficient spatial characteristics like the experimental area, in which the Maximum Likelihood Classification is the best, and it also proves that unsupervised classification is not suitable for mineral classification in the similar areas.

Keywords: mineral mapping; hyperspectral remote sensing; supervision classification; unsupervised classification

高光谱遥感技术能够同时获取目标的空间与光谱信息,即具有“图谱合一”的优势。尤其在光谱维度上,与传统的多光谱遥感相比,高光谱遥感能够获取地物在连续光谱上的精细光谱特征,为识别不同地物类别提供了更多信息^[1-5]。目前,高光谱遥感技术已在矿产勘探领域得到成熟的应用^[6-11],尤其是核地质系统自 2008 年在国内率先引进国际先进的 CASI/SASI 航空高光谱遥感测量系统^[12],航空高光谱遥感技术成为铀矿找矿领域新技术创新的重要方向^[13]。

高光谱图像分类是指利用算法根据像元的光谱特征将其分配到预定义的类别中,主要目的是识别和区分图像中不同类型的地表覆盖和材料,从而获取有关地球表面特征的详细信息。高光谱图像分类是充分挖掘和利用高光谱数据的关键技术之一,是高光谱遥感应用的重要基础^[14]。其分类方法大致可分为监督分类和非监督分类两大类^[15]。监督分类是基于已知类别标签的样本来优化分类器模型,并进而对未标签样本进行分类,代表性的方法有支持向量机^[16]、随机森林^[17]和神经网络^[18]等。非监督分类无需训练样本,直接根据像元间的相似性度量进行聚类,常见的

方法包括 k-means^[19]、ISODATA^[20]等。此外还出现了一些半监督分类方法,利用少量标记样本和大量未标记样本进行分类^[21-22]。结合具体应用需求选择和优化相应的方法,是提升高光谱遥感图像分类精度和应用效果的关键。

本文以亥姆寺东侧矿区为例,使用航空高光谱遥感影像为实验数据,应用部分典型的高光谱图像分类方法进行蚀变矿物填图分类实验,并对比分析了各种算法在典型矿物分类提取方面的效果,以期对相关研究提供参考。

1 实验数据

实验区域位于甘肃省金昌市以北,区域地表基本无植被覆盖,以裸土和矿石碎屑为主,主要矿物分布有高岭石、方解石、伊利石、白云石和第四纪堆积物等。图 1 为该区域航空高光谱影像,面积约为 3.51 km²,基于 SASI 航空成像光谱仪获取,波段范围介于 950~2 450 nm 之间,光谱分辨率为 15 nm,影像空间分辨率为 2 m。原始辐亮度数据共 101 波段,用 FLAASH 方法对研究区进行大气校正处理,剔除掉水汽噪声波段后剩余 89 个波段。通过野外查证分析后得到的样本数据表 1 所示,其光谱曲线如图 2 所示。

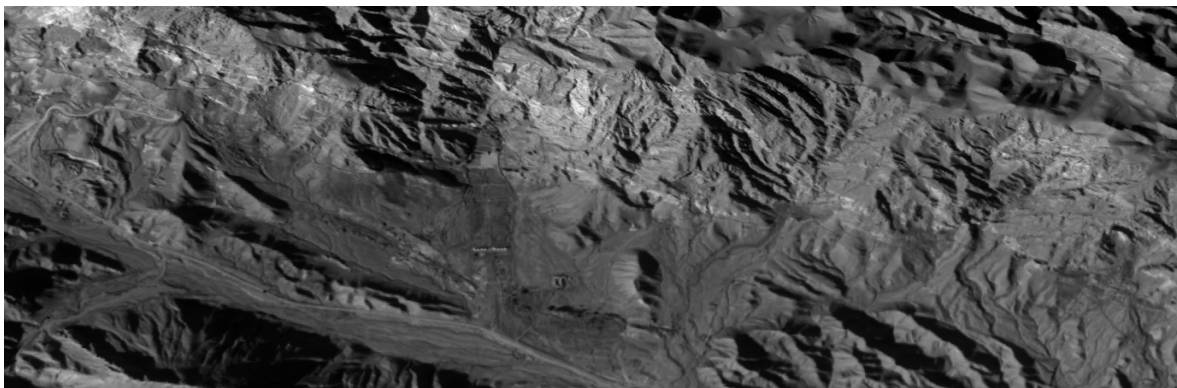


图 1 研究区航空高光谱遥感影像(RGB 组合为 980、965 和 950 nm)

Fig. 1 Airborne hyperspectral remote sensing images of the study area (RGB combination is 980, 965 and 950 nm)

表1 样本像元数

Table 1 Pixel numbers for different mineral sample

样本名称	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石
像元数	1 111	286	194	281	257	131	369

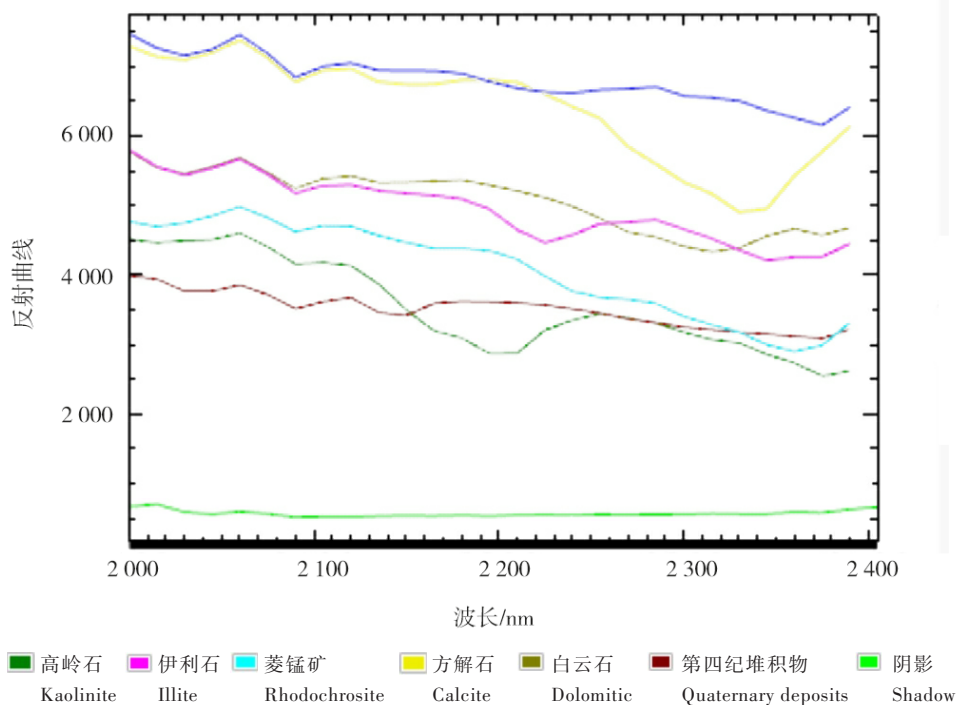


图2 样本光谱曲线

Fig. 2 Spectral curves of mineral samples

2 实验方法

分析总结高光谱遥感图像分类具体方法。主要分为监督分类和非监督分类两个大类,监督分类以光谱角分类、支持向量机分类、随机森林分类、最小距离分类、神经网络分类和最大似然分类为例,非监督分类以典型的k-means和ISODATA为例,对上述8种分类方法进行对比分析。在每个大类下面也对相应的算法进行了更加详细划分,具体的优缺点见表2。

2.1 监督分类方法

监督分类是基于已知分类的样本(通常称为训练集)来训练分类器,并将图像中的像素分配到预定义的类中^[23]。

1) 光谱角分类法(Spectral Angle Mapper, SAM)是将每个像素的光谱看作是高维空间中的一个向量,然后通过计算待分类像素与参考

光谱之间的光谱角进行分类。光谱角越小,说明待分类像素与参考光谱的相似度越高,反之则说明相似度越低。如图3所示,为计算像元光谱X与参考光谱Y间的夹角 α 。通过判定夹角 α 是否满足阈值条件来进行分类。

传统的SAM分类精度会受到混合其他物种的噪声光谱影响,在分类前进行纯净像元提取后的SAM未改进的光谱角分类的各项精度指标都高,苏岫^[24]等人在红树林分类中验证并取得良好的分类精度。SAM分类结果如图4所示,分类精度如表3所示。

2) 最小距离法(Minimum Distance to Mean, MD)是一种常用的高光谱图像分类方法,这种方法假设每个类别的光谱值是均匀分布的,并通过计算像元光谱与各类别光谱中心点的距离进行分类。

欧氏距离是最常用的计算距离的方法,但由于该方法只考虑了待分类样本与每个类别

表 2 高光谱遥感图像分类的方法优缺点

Table 2 Advantages and disadvantages of hyperspectral image classification methods

分类方法	优点	缺点	
监督分类	光谱角分类(SAM)	①计算简单, 易于实现和应用 ②对光谱纯度较高的像元分类效果好 ③对光谱维度和数据规模的适应性较好	①分类精度可能不及SVM等方法 ②对光谱混合像元的分类效果较差 ③容易受到光谱变异性和噪声的影响
	支持向量机分类(SVM)	①分类精度高, 特别是对光谱混合像元的分类效果好 ②能够处理高维特征数据 ③对小样本和非线性可分数据的分类性能好	①对参数敏感, 需要调节核函数等 ②训练过程复杂, 对大规模高光谱数据的计算效率较低 ③分类结果的物理解释性较差
	随机森林分类(RF)	①分类精度高, 对高维高光谱数据的分类效果好 ②对参数不敏感, 易于实现和应用 ③能够评估高光谱特征的重要性, 有助于波段选择	①模型复杂度较高, 对大规模高光谱数据的训练效率较低 ②对噪声和异常值的鲁棒性可能不如SVM ③分类结果的物理解释性一般
	最小距离分类(MD)	①计算简单, 易于实现和应用于高光谱数据 ②分类速度快, 适用于实时高光谱图像分类 ③对训练样本的依赖性小, 易于实现自动化分类	①分类精度相对较低, 特别是对光谱混合像元 ②对光谱变异性和噪声敏感, 容易出现错分和漏分 ③难以发现非线性和复杂的类别关系
	神经网络分类(NN)	①分类精度高, 特别是对光谱混合像元和复杂类别关系 ②能够自动学习和提取高光谱特征 ③具有较好的泛化能力和鲁棒性	①训练过程复杂, 需要大量的标注样本和计算资源 ②模型复杂度较高, 易出现过拟合, 泛化能力不确定
最大似然法分类(MLC)	①具有坚实的理论基础, 广泛应用于统计学 ②简单易用, 计算方法相对简单 ③可以利用训练数据自动确定分类模型的参数, 不需要手动设置	①对训练数据的质量和数量要求较高, 训练数据不足或质量不高可能影响分类效果 ②对总体分布的假设(如正态分布)可能不总是成立, 这可能影响分类的准确性	
非监督分类	k-means	①计算效率高, 适用于大规模高光谱数据 ②聚类结果易于解释, 可以与地物类别对应 ③可以利用先验知识, 如已知的地物类别数目	①需要预先指定聚类数k ②对光谱异常值和噪声敏感, 可能影响聚类精度 ③对异常值和噪声敏感
	ISODATA	①不需要预先指定聚类数, 可以自动确定 ②可以处理不同大小和形状的聚类 ③对噪声和异常值有一定的鲁棒性	①计算复杂度较高, 尤其是在数据量大时 ②参数选择(如聚类合并和分裂的阈值)对结果有较大影响 ③可能需要多次迭代以获得稳定结果

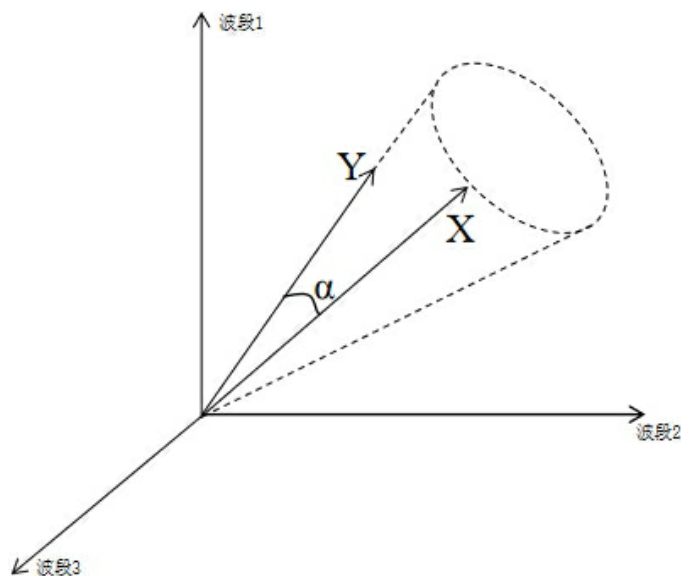


图 3 光谱角法示意图

Fig. 3 Schematic diagram of spectral angle method mapping

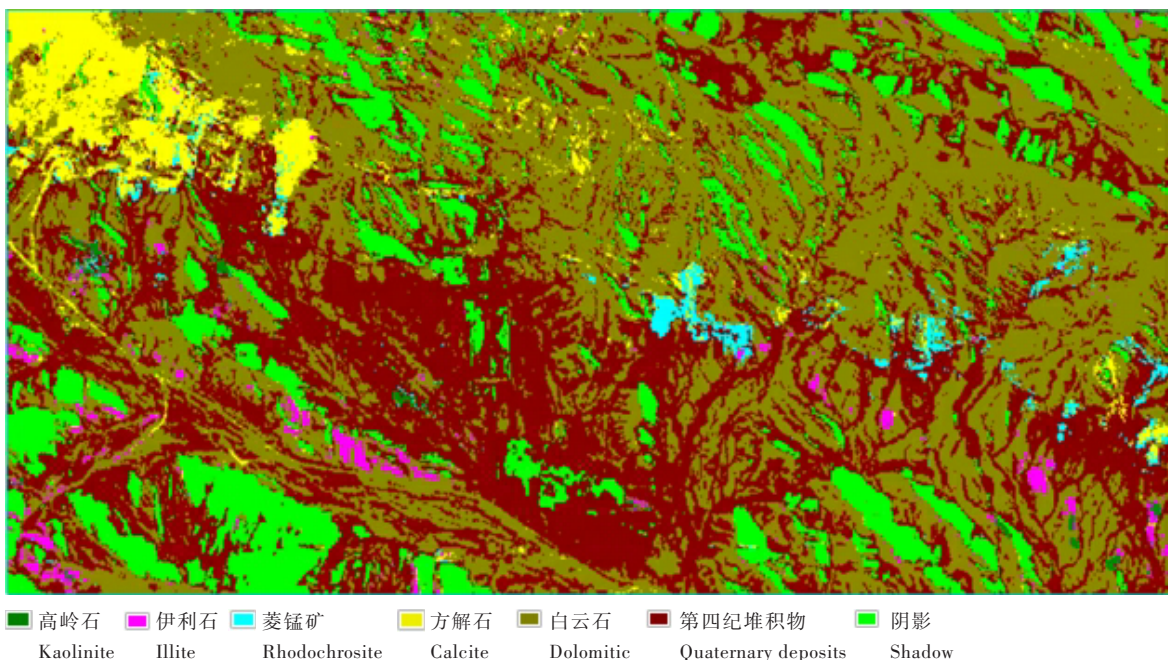


图 4 光谱角分类法分类结果

Fig. 4 Classification results by spectral angle mapping

表 3 光谱角分类法分类精度

Table 3 Classification accuracy of spectral angle mapping

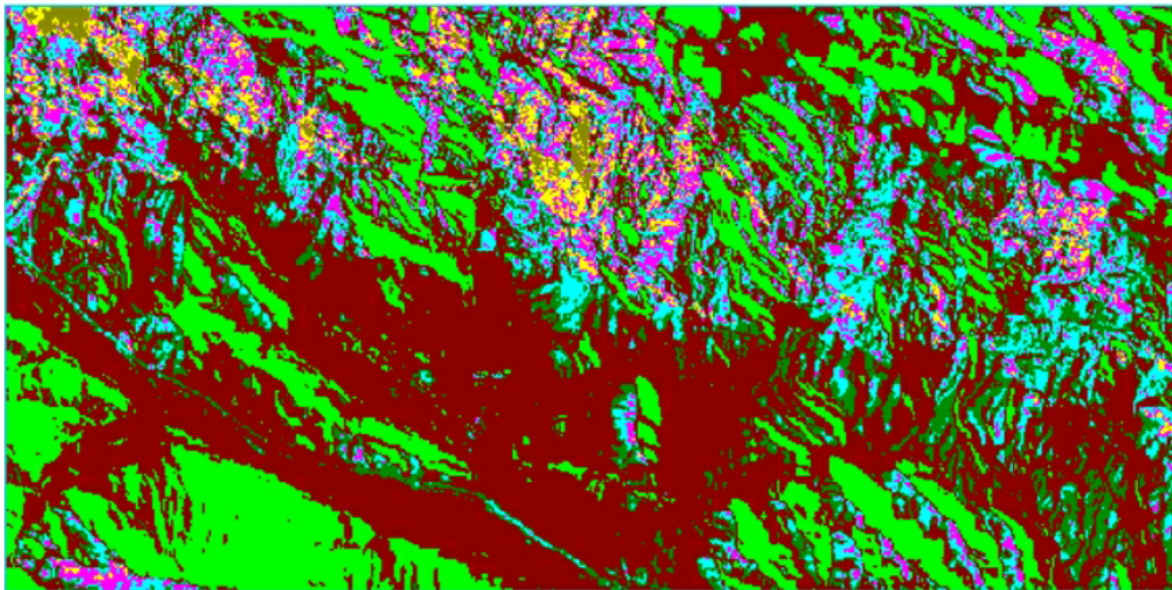
总体精度 = (2 264/2 629) 86.116 4%
Kappa 系数 = 0.817 5

类别	地表真实像元							总计
	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石	
阴影	1 107	20	0	0	0	0	2	1 127
第四纪堆积物	4	150	10	0	14	0	20	199
白云石	0	116	171	0	0	5	6	298
伊利石	0	0	0	202	5	1	17	225
菱锰矿	0	0	4	0	209	15	10	238
方解石	0	0	9	15	29	110	0	163
高岭石	0	0	0	64	0	0	314	379
总计	1 111	286	194	281	257	131	369	2 629

中心的距离,而没有考虑已知样本的分布,虽然分类速度很快,但准确率不高。此外,当特征维数很高时,MD的运算量和计算复杂度会明显增加,影响分类效率和准确性。因此,在应用MD时,需要充分考虑不同情况下的适用性和限制性,以确保其正确性和有效性。近年来,部分学者对其进行研究,但在高光谱领域研究较少,通常被用于和其他算法进行对比。MD分类结果如图5所示,分类精度如表4所示。

3) 支持向量机 (Support Vector Machine, SVM) 是一种强大的监督学习模型,常用于解决分类和回归问题。自从 Vapnik 和 Chervonenkis 于 1963 年提出以来, SVM 已经在多个领域显示出其卓越的性能,尤其是在图像分类和模式识别领域。在高光谱遥感图像分类的应用中, SVM 由于其在高维空间中的有效性和对小样本数据集的处理能力而变得尤为重要。

SVM 通过构建一个或多个超平面在高维或无限维空间中进行分类,从而有效解决这一



■ 高岭石 ■ 伊利石 ■ 菱锰矿 ■ 方解石 ■ 白云石 ■ 第四纪堆积物 ■ 阴影
Kaolinite Illite Rhodochrosite Calcite Dolomitic Quaternary deposits Shadow

图 5 最小距离法分类结果

Fig. 5 Classification result of minimum distance method

表 4 最小距离法分类精度

Table 4 Classification accuracy of minimum distance method

		总体精度 = (1 768/2 629) 67.249 9 % Kappa 系数 = 0.569 1						
		地表真实像元						
类别	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石	总计
阴影	1 111	14	0	0	0	0	5	1 130
第四纪堆积物	0	162	1	9	19	14	110	315
白云石	0	0	135	6	5	71	0	217
伊利石	0	50	15	173	24	3	79	344
菱锰矿	0	37	11	38	57	32	36	211
方解石	0	2	32	38	49	3	12	136
高岭石	0	21	0	17	103	8	127	276
总计	1 111	286	194	281	257	131	369	2 629

问题。在 SVM 的核心设计中,存在一个关键概念是最大化类间边界,即寻找一个最优的分类超平面,使得不同类别之间的间隔最大化。这种最大间隔原则是 SVM 优越泛化能力的根本所在^[25]。

近年来,随着机器学习技术的发展,SVM 已被应用于高光谱数据的特征选择和降维,进一步提高分类的准确性和效率。通过与其他机器学习技术如随机森林、深度学习模型等结合,SVM 在处理特别复杂的高光谱数据时展示

了更大的潜力和灵活性。

SVM 因其强大的分类能力、优秀的泛化性能以及对高维及小样本数据集的适应性,在高光谱遥感图像分类中扮演着至关重要的角色。未来,随着计算技术的进步和算法的优化,SVM 预计将在高光谱图像处理领域继续保持其重要地位。支持向量机法分类结果如图 6 所示,分类精度如表 5 所示。

4) 随机森林(Random Forest, RF)通过构建多个决策树来进行分类或回归预测,每棵树的

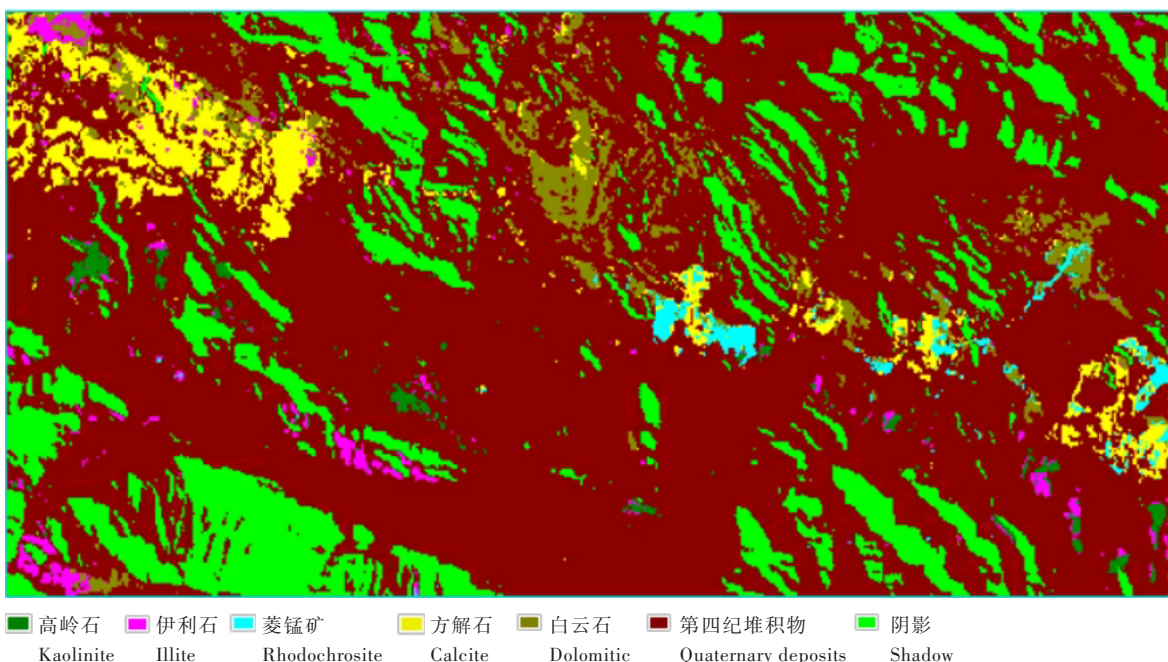


图 6 支持向量机法分类结果

Fig. 6 Classification results of support vector machine method

表 5 支持向量机法分类精度

Table 5 Classification accuracy of support vector machine method

总体精度 = (2 589/2 629) 98.478 5 %
Kappa 系数 = 0.980 0

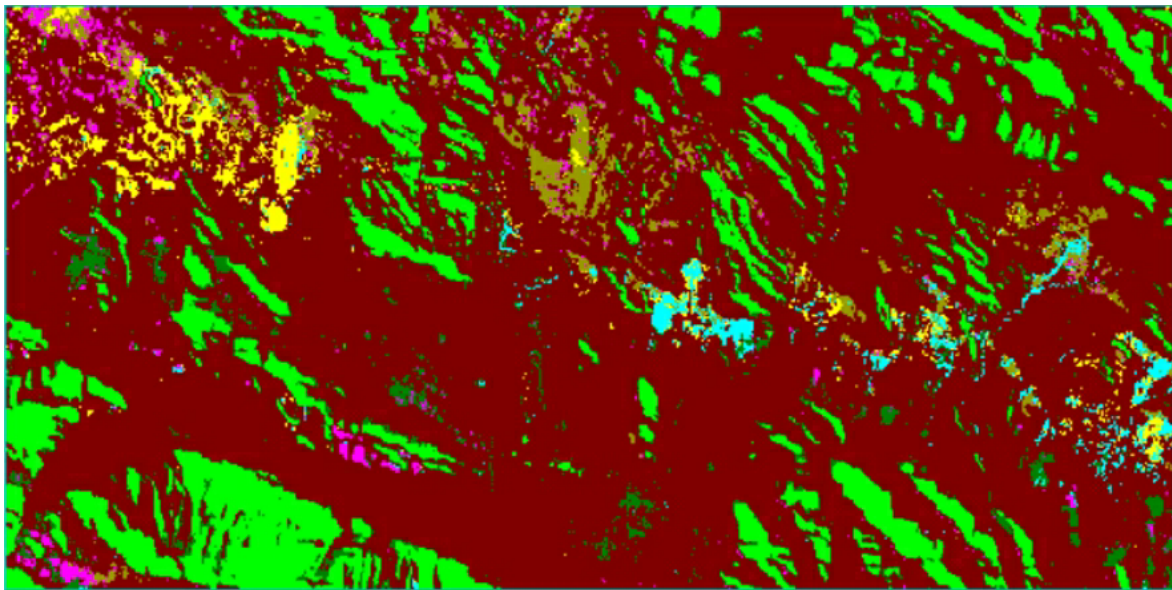
类别	地表真实像元							总计
	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石	
阴影	1 111	0	0	0	0	0	2	1 113
第四纪堆积物	0	283	0	0	0	1	1	285
白云石	0	3	189	5	0	0	0	197
伊利石	0	0	0	266	4	1	0	271
菱锰矿	0	0	0	0	253	8	0	261
方解石	0	0	5	0	0	121	0	126
高岭石	0	0	0	10	0	0	366	376
总计	1 111	286	194	281	257	131	369	2 629

构建过程中引入随机性,以提高模型的泛化能力^[26]。决策树是一种基础的机器学习算法,用于分类和回归任务。该算法生成的模型具有树状结构,并通过一连串的决策规则对数据进行递归分割,从而对目标变量的值进行预测^[27]。作为一种基于空间数据挖掘和知识发现的监督分类方法,决策树是一种有效的高光谱影像分类方法。尽管决策树在某些情况下可能需要较多的人工干预,并且可能不如其他一些算法效率高。

RF 对于波段数量的需求超过多光谱数据,

而高光谱数据的属性正好契合该技术的需求,两者的融合能够促进分类性能的提升,因此 RF 在高光谱遥感领域的分类问题上展现出极大的研究潜力。随机森林法对高维高光谱数据的分类效果好,对参数不敏感,易于实现和应用,但模型复杂度较高,对大规模高光谱数据的训练效率较低,分类结果的物理解释性一般。随机森林法结构简单,训练容易,具有很高的分类精度。RF 分类结果如图 7 所示,分类精度如表 6 所示。

5) 随着传感器的发展,高空间分辨率的高



■ 高岭石 ■ 伊利石 ■ 菱锰矿 ■ 方解石 ■ 白云石 ■ 第四纪堆积物 ■ 阴影
Kaolinite Illite Rhodochrosite Calcite Dolomitic Quaternary deposits Shadow

图 7 随机森林法分类结果

Fig. 7 Results of random forest classification

表 6 随机森林法分类精度

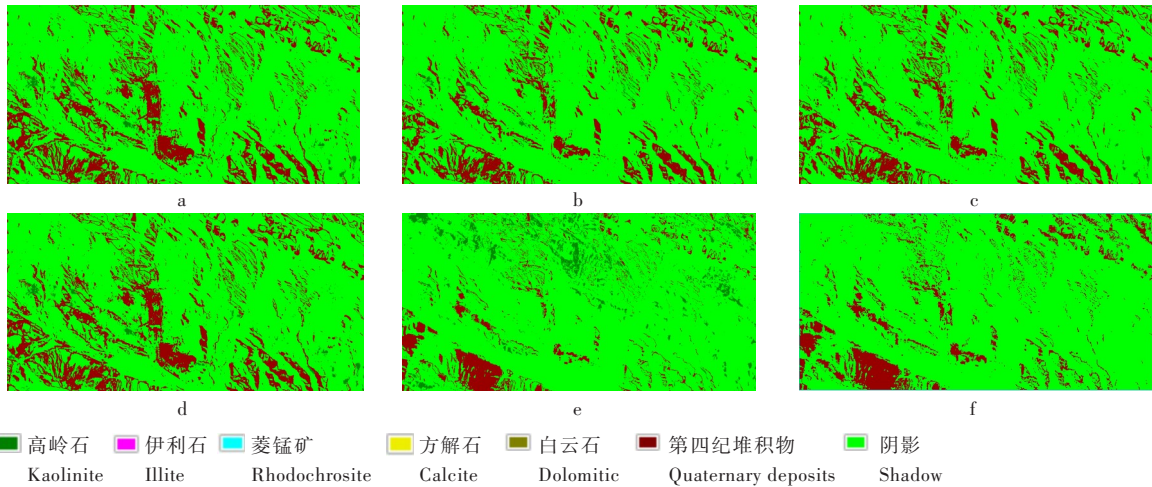
Table 6 Classification accuracy of random forest method

总体精度 = (2 625/2 629) 99.884 7 % Kappa 系数 = 0.997 6								
类别	地表真实像元							总计
	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石	
阴影	1 110	0	0	0	0	0	0	1 110
第四纪堆积物	1	286	0	0	0	0	0	287
白云石	0	0	194	1	0	0	0	195
伊利石	0	0	0	279	0	0	0	279
菱锰矿	0	0	0	0	257	0	0	257
方解石	0	0	0	0	0	131	0	131
高岭石	0	0	0	1	0	0	369	370
总计	1 111	286	194	281	257	131	369	2 629

光谱图像包含了丰富的空间信息,这些空间信息可以与光谱信息相结合从而提高分类精度。传统的机器学习分类方法如 SVM 和 RF,依赖于手工选取的特征,这在高光谱数据中可能不仅耗时而且难以捕捉到所有有用信息。相较之下,神经网络(Neural Networks, NN),尤其是卷积神经网络(Convolutional Neural Networks, CNN),能自动从原始数据中学习复杂的特征表示,显著提升了分类效果^[28]。

本文选择神经网络法基于前馈神经网络,是一种最简单的神经网络,各神经元分层排

列,每个神经元只与前一层的神经元相连。接收前一层的输出,并输出给下一层,各层间没有反馈。其是应用最广泛、发展最迅速的人工神经网络之一^[29]。图 8 展示神经网络法在不同阈值下的分类结果与分类精度。阈值参数决定了与活化节点级别相关的内部权重的贡献。它用于调节节点内部权重的变化,内部节点主要位于隐藏层。隐藏层是前馈神经网络中位于输入层和输出层之间的层次。一个前馈神经网络可以有一个或多个隐藏层,每个隐藏层包含多个节点,内部节点的主要功能是提



a—阈值为 0.9; b—阈值为 0.87; c—阈值为 0.85; d—阈值为 0.8; e—阈值为 0.75; f—阈值为 0.7。
a—The threshold is 0.9; b—The threshold is 0.87; c—The threshold is 0.85; d—The threshold is 0.8; e—The threshold is 0.75; f—The threshold is 0.7.

图 8 神经网络法分类结果

Fig. 8 Classification result of Neural Network method

表 7 神经网络法阈值为 0.85 时的分类精度

Table 7 Classification accuracy of Neural Network method with threshold of 0.85

总体精度 = (1 431/2 629) 54.431 3 % Kappa 系数 = 0.263 4								
地表真实像元								
类别	阴影	白云石	伊利石	菱锰矿	方解石	第四纪堆积物	高岭石	总计
阴影	1 090	194	276	257	131	157	160	2 262
白云石	0	0	0	0	0	0	0	0
伊利石	0	0	0	0	0	0	0	0
菱锰矿	0	0	0	0	0	0	0	0
方解石	0	0	0	0	0	0	0	0
第四纪堆积物	21	0	0	0	0	132	0	153
高岭石	0	0	5	0	0	0	209	214
总计	1 111	194	281	257	131	286	369	2 629

取和转换输入数据的特征。而节点权重则是指连接前一层节点到当前层节点的连接强度。这些权重是神经网络中的关键参数,决定了输入信号如何影响当前节点的输出。经过重复调整阈值,当阈值为 0.85 时,分类精度最高,分类精度如表 7 所示。

6) 最大似然法 (Maximum Likelihood Classification, MLC) 是一种广泛应用于高光谱遥感影像分类的监督分类方法。该方法通过估算待分类像元属于各个类别的概率,将像元分配给似然值最大的类别最大似然法在遥感分类领域因其理论基础扎实、分类效果较好而得到广泛应用。

MLC 考虑类别间的光谱特征分布和类别内部的协方差结构,能够有效区分光谱特征相似的类别,提高分类精度。并且该方法可以处理多波段、多类别的高光谱数据,适用于复杂的地物分类^[30-31]。但最大似然法需要计算协方差矩阵及其逆矩阵,对计算资源要求较高,处理大规模高光谱数据时计算复杂度较大^[32]。最大似然法分类结果如图 8 所示,分类精度如表 8 所示。

2.2 非监督分类

非监督分类是一种不依靠预先标记的训练样本来对图像数据中的像素进行分类。通过分析像素的光谱特征和它们之间的统计关

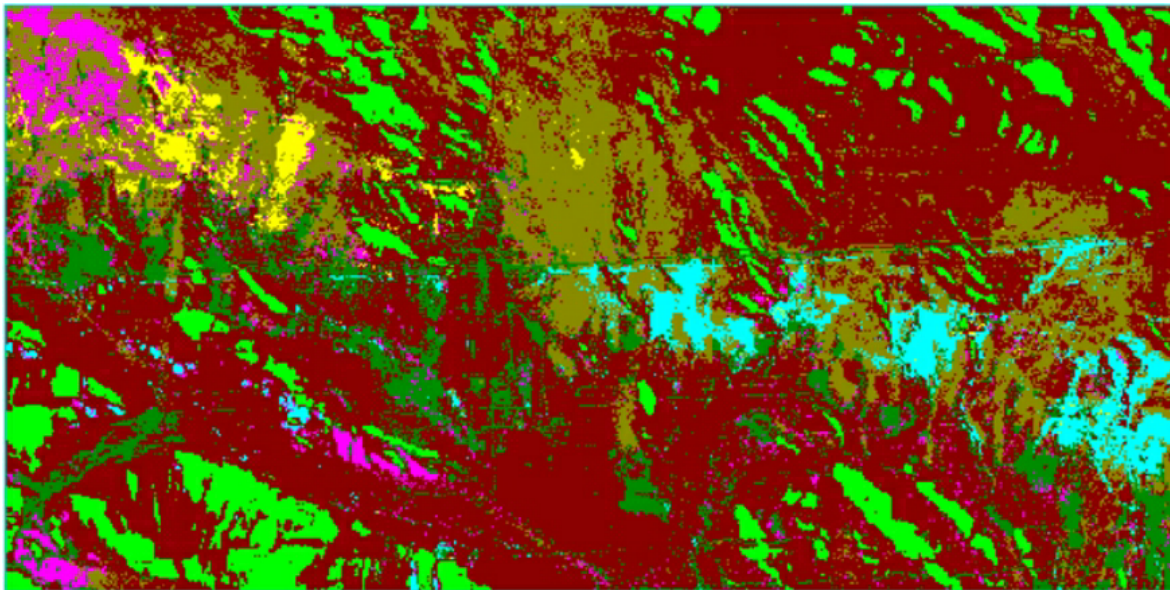


图 9 最大似然法分类结果

Fig. 9 Classification results by Maximum Likelihood

表 8 最大似然法分类精度

Table 8 Classification accuracy by Maximum Likelihood method

总体精度 = (2 628/2 629) 99.962 0 % Kappa 系数 = 0.999 5								
类别	地表真实像元							总计
	阴影	第四纪堆积物	白云石	伊利石	菱锰矿	方解石	高岭石	
阴影	1 111	0	0	0	0	0	0	1 111
第四纪堆积物	0	286	0	0	0	0	0	286
白云石	0	0	194	0	0	0	0	194
伊利石	0	0	0	280	0	0	0	280
菱锰矿	0	0	0	0	257	0	0	257
方解石	0	0	0	0	0	131	0	131
高岭石	0	0	0	1	0	0	369	370
总计	1 111	286	194	281	257	131	369	2 629

系来自动地将像素聚集成不同的类别或簇,然后按照一定规则自动根据像元光谱或空间等特征组成集群组。

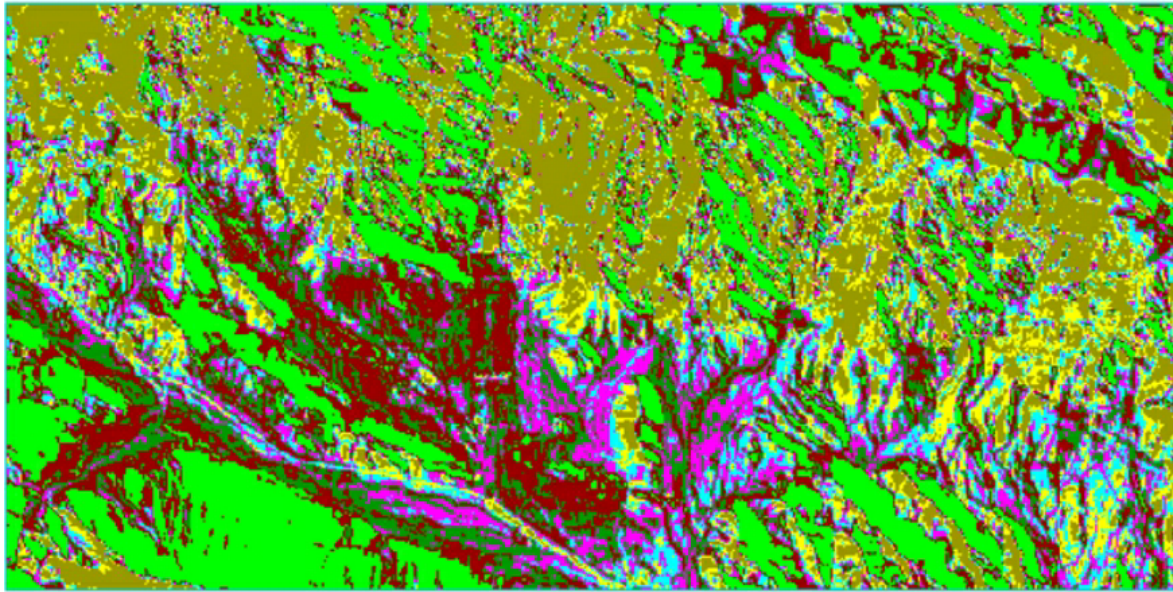
2.2.1 k-均值聚类

k-均值算法(k-means)在执行过程中需要计算每个点到各个簇中心的距离,这一步骤在数据量大时尤其耗时。为了提高效率,阴阳k-means被提出,它通过两个过滤器来识别和省略不必要的距离计算,从而显著提高算法的运行效率^[33]。此外,k-means对初始中心的选择

非常敏感,这会影响最终聚类的结果。为了解决这个问题,人们开发了更有效的初始化方法,如k-means++和基于密度的方法,这些方法能够增强聚类的稳定性和准确性^[34]。分类结果如图 10 所示,分类精度如表 9 所示。

2.2.2 ISODATA 聚类

ISODATA 聚类 (Iterative Self-Organizing Data Analysis Techniques Algorithm) 是 k-means 的一种扩展,其特点是在迭代过程中允许类别数量动态变化。该算法首先随机分配初始的



■ 高岭石 ■ 伊利石 ■ 菱锰矿 ■ 方解石 ■ 白云石 ■ 第四纪堆积物 ■ 阴影
 Kaolinite Illite Rhodochrosite Calcite Dolomitic Quaternary deposits Shadow

图 10 k-均值算法分类结果

Fig. 10 Classification result of k-means algorithm

表 9 k-均值算法分类精度

Table 9 Classification accuracy of k-means algorithm

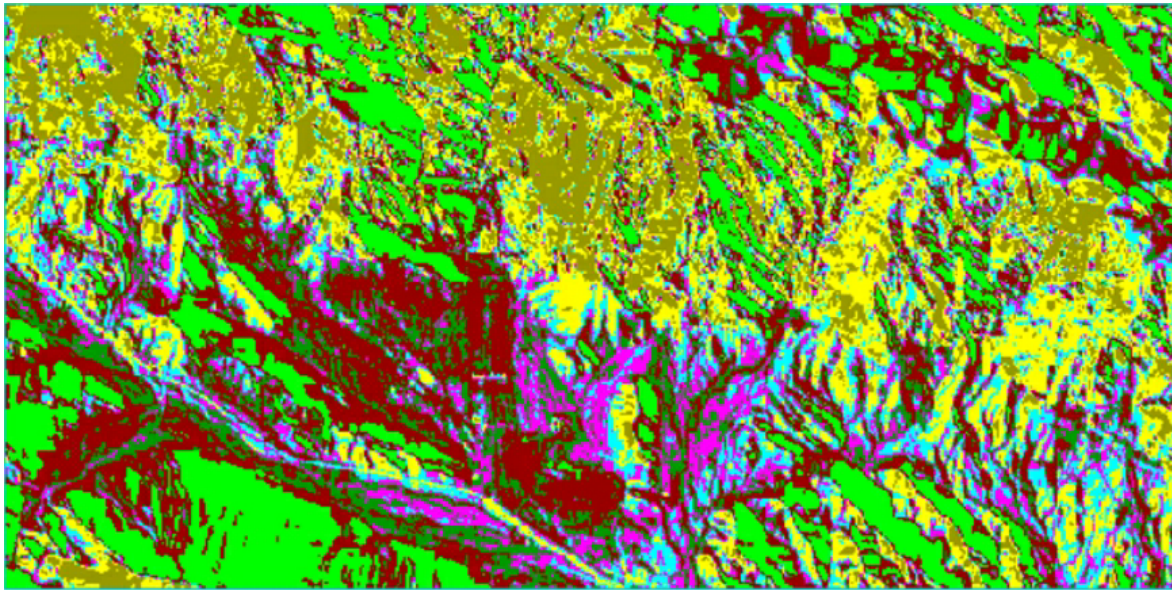
总体精度 = (1 521/2 629) 57.854 7 % Kappa 系数 = 0.454 5								
类别	地表真实像元							总计
	阴影	第四纪堆积物	高岭石	伊利石	菱锰矿	方解石	白云石	
阴影	1 111	13	4	0	0	0	0	1 128
第四纪堆积物	0	117	21	0	0	0	0	138
高岭石	0	28	39	0	0	1	0	68
伊利石	0	9	43	7	0	5	0	64
菱锰矿	0	13	54	3	48	10	1	129
方解石	0	17	59	13	81	7	1	178
白云石	0	89	149	258	128	108	192	924
总计	1 111	286	369	281	257	131	194	2 629

聚类中心,然后计算每个簇的均值和协方差。接下来,算法将每个数据点分配至最近的簇,并基于这些数据点更新簇的均值和协方差。这一过程持续迭代,直至连续迭代之间的变化满足预设的停止条件^[35]。这样的改动可以通过测量聚类均值从一次迭代到下一次迭代的变化距离来量化,也可以通过迭代之间变化的像素百分比来量化。它可以合并相似的簇或分裂过大的簇,提供更多的灵活性,更好地适应地表覆盖类型的多样性和复杂性。分类结果如图 11 所示,分类精度如表 10 所示。

3 讨 论

本文对大气校正后的高光谱数据采用 SAM、SVM、RF 等监督分类法和 k-means 等非监督分类法进行矿物分类,分类结果如表 11 所示。

通过地面真实感兴趣区与分类后的图像分别进行混淆矩阵分析评估分类性能。对比分析结果发现 MLC 和 RF 算法的分类性能较好,说明该算法可以处理高维度数据的分类计算;SVM 算法适用于特征数量较多情况下的数



■ 高岭石 ■ 伊利石 ■ 菱锰矿 ■ 方解石 ■ 白云石 ■ 第四纪堆积物 ■ 阴影
 Kaolinite Illite Rhodochrosite Calcite Dolomitic Quaternary deposits Shadow

图 11 ISODATA 聚类算法分类结果

Fig. 11 Classification result of ISODATA clustering algorithm

表 10 ISODATA 聚类算法分类精度

Table 10 Classification accuracy of ISODATA clustering algorithm

总体精度 = (1 567/2 629) 59.604 4 % Kappa 系数 = 0.479 7								
类别	地表真实像元							总计
	阴影	第四纪堆积物	高岭石	伊利石	菱锰矿	方解石	白云石	
阴影	1 111	0	2	0	0	0	0	1 113
第四纪堆积物	0	138	23	0	0	0	0	161
高岭石	0	20	36	0	0	1	0	57
伊利石	0	9	43	7	0	5	0	64
菱锰矿	0	17	64	3	62	10	1	157
方解石	0	43	123	108	94	26	6	400
白云石	0	59	78	163	101	89	187	677
总计	1 111	286	369	281	257	131	194	2 629

表 11 不同分类方法的分类结果对比

Table 11 Comparison of classification results of different classification methods

分类方法	SAM	MD	RF	SVM	NN	MLC	k-means	ISODATA
分类精度/%	44.009 1	67.249 9	99.884 7	98.478 5	54.431 9	99.962 0	57.854 7	59.604 4
Kappa 系数	0.405 7	0.569 1	0.997 6	0.980 0	0.263 4	0.999 5	0.454 5	0.479 7

表 12 第四纪堆积物与其他矿物的皮尔逊相关系数

Table 12 Pearson correlation coefficient between Quaternary deposits and other minerals

类别	高岭石	伊利石	菱锰矿	方解石	白云石
第四纪堆积物	0.887 2	0.902 4	0.865 2	0.846 5	0.927 1

据分类。这可能是导致SVM算法分类性能低于RF的主要原因;SAM算法分类精度低于MLC、SVM和RF,通过SAM结果的混淆矩阵中可以看出,第四纪堆积物和白云石之间错分较多。为验证两者的相关性,提取每个样本的光谱数据,通过皮尔逊相关系数对第四纪堆积物与其他样本进行比较分析,计算结果如表12所示。

通过计算可以看出第四纪堆积物与白云石的皮尔逊相关系数为0.927 1,两者相关性较高。SAM算法主要通过比较像元光谱与参考光谱之间的相似性来进行分类,当矿物光谱特征差异明显且稳定时,这类方法能够取得较好的分类效果;但如果矿物光谱特征相似或存在较多的同物异谱现象,其分类精度可能会下降。RF可以通过多波段组合的阈值分割区分细微差异,而SAM仅依赖整体光谱形状的相似性,这可能是分类精度不如RF的主要原因。

最小距离分类法假设每个类别的光谱值是均匀分布的,因此对于光谱值分布复杂的类别,这可能会显著影响分类性能;最大似然法对于样本的选择要求苛刻,要求样本能够准确代表各类别的光谱特征,样本中的混合像元可能导致分类精度下降;神经网络分类法基于研究区的分类性能较低。从不同的阈值对应的分类精度可以得出,当阈值设置为0.85时分类精度最高,为54.431 9%,与其他分类方法的精度仍有差距。

针对非监督分类不需要样本的情况,本实验设置分类结果为7类,通过分类结果地物光谱与样本光谱对比,同时在分类的结果与样本标签做叠加分析后,经过比对得到了样本和结果类的对应关系。

从分类结果可以看出,两种非监督分类方法只有对未分类区的分类结果高度一致。从影像上判断,未分类区是因地形效应所产生的阴影。从整体精度上来看,k-means方法分类精度为57.854 7%,ISODATA方法分类精度为59.604 4%,与监督分类方法的精度形成明显差异。非监督分类依赖数据自身统计特性,而遥感数据的复杂性、噪声干扰及算法局限性导致难以准确区分地物。

非监督分类无需预先提供样本数据,仅需设定预期的分类个数,即可自动完成分类。然而在矿区环境中,由于空间特征复杂且不明显,难以直接确定分类结果与已知样本之间的对应关系。因此,通常需要借助已知的地质图进行对比验证,但这种方法在实际应用中显得较为繁琐,且在没有现成地物参考图的情况下难以实现,这限制了非监督分类在矿区的实用性。

4 结 论

本研究基于航空高光谱影像数据,利用6种监督分类方法和2种非监督分类方法对研究区进行6种矿物分类提取工作,经对比分析,主要结论如下:

1)监督分类方法能实现矿物的分类,但不同算法的分类精确性有一定的差异。在基于本研究区的分类方法中,RF、SVM和MLC算法分类精度较高。

2)SAM的角度阈值设置,SVM、RF和NN的参数值都极大地影响着分类结果。

3)k-means和ISODATA分类精度不如监督分类的主要原因在于其分类过程缺乏先验知识约束,更容易受地形和噪声等因素的干扰,导致其分类结果的准确性和实用性受限。因此,在实际应用中,非监督分类可能更适合初步分类或探索性分析,而监督分类方法更适合需要高精度的应用场景。

参考文献

- 张川,易敏,童勤龙,等.深度卷积网络在航空高光谱岩性识别中的应用——以塔木素铀矿床北部地区为例[J].世界核地质科学,2024,41(1):33-46.
ZHANG Chuan, YI Min, TONG Qinlong, et al. Application of deep convolutional networks in airborne hyperspectral lithology identification: A case study of the northern Tamusu uranium deposit [J]. World Nuclear Geoscience, 2024, 41(1): 33-46 (in Chinese).
- 杨敏,傅炜舜,聂兴信,等.高光谱遥感技术在矿山地质环境调查中的应用[J].现代矿业,2024,40(1):48-52.
YANG Min, FU Weishun, NIE Xingxin, et al. Application of hyperspectral remote sensing technology

- in mine geological environment survey [J]. *Modern Mining*, 2024, 40(1): 48–52 (in Chinese).
- 3 韩晓青, 杨国防, 木红旭, 等. 新疆雪米斯坦铀成矿有利区高光谱遥感识别方法研究[J]. *世界核地质科学*, 2022, 39(2): 309–320.
HAN Xiaqing, YANG Guofang, MU Hongxu, et al. An identification method of hyperspectral remote sensing for the favorable uranium metallogenic area in Xuemisitan, Xinjiang[J]. *World Nuclear Geoscience*, 2022, 39(2): 309–320 (in Chinese).
 - 4 董秀军, 许强, 余金星, 等. 九寨沟核心景区多源遥感数据地质灾害解译初探[J]. *武汉大学学报(信息科学版)*, 2020, 45(3): 432–441.
DONG Xiujun, XU Qiang, SHE Jinxing, et al. Preliminary study on interpretation of geological hazards in Jiuzhaigou based on multi-source remote sensing data[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(3): 432–441 (in Chinese).
 - 5 LANDGREBE D. Hyperspectral image data analysis [J]. *IEEE Signal Processing Magazine*, 2002, 19(1): 17–28
 - 6 BIOUCAS-DIAS J M, PLAZA A, CAMPS-VALLS G. Hyperspectral remote sensing data analysis and future challenges [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2013, 1(2): 6–36.
 - 7 FAUVEL M, TARABALKA Y, BENEDIKTSSON J A, et al. Advances in spectral-spatial classification of hyperspectral images [J]. *Proceedings of the IEEE*, 2012, 101(3): 652–675.
 - 8 GHAMISI P, YOKOYA N, LI J, et al. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2017, 5(4): 37–78.
 - 9 黄海霞. 土地资源调研中卫星遥感技术的应用分析[J]. *电子元器件与信息技术*, 2022, 6(4): 18–21
HUANG Haixia. Application analysis of satellite technology in land remote sensing resources investigation [J]. *Electronic Components and Information Technology*, 2022, 6(4): 18–21 (in Chinese).
 - 10 许强, 郭晨, 董秀军, 地质灾害航空遥感技术应用现状及展望[J]. *测绘学报*, 2022, 51(10): 2020–2033.
XU Qiang, GUO Chen, DONG Xiujun. Application status and prospect of aerial remote sensing technology for geological hazards [J]. *Acta Geodaetica et Cartographica Sinica*, 2022, 51(10): 2020–2033 (in Chinese).
 - 11 许强, 朱星, 李为乐, 等. “天-空-地”协同滑坡监测技术进展[J]. *测绘学报*, 2022, 51(7): 1416–1436.
XU Qiang, ZHU Xing, LI Weile, et al. Technical progress of space-air-ground collaborative monitoring of landslide [J]. *Acta Geodaetica et Cartographica Sinica*, 2022, 51(7): 1416–1436 (in Chinese).
 - 12 叶发旺, 张川, 李瀚波, 等. “空天”高分辨率遥感技术及其在铀资源勘查中的应用进展与发展建议[J]. *铀矿地质*, 2021, 37(3): 313–328.
YE Fawang, ZHANG Chuan, LI Hanbo, et al. Strengthen the application of “aero-space” high resolution remote sensing technology, helping the fundamental potential investigation of uranium resources in China [J]. *Uranium Geology*, 2021, 37(3): 313–328 (in Chinese).
 - 13 叶发旺, 张杰林, 张川, 等. 航空高光谱遥感技术在铀矿找矿中的典型应用——以新疆雪米斯坦地区为例[J]. *世界核地质科学*, 2024, 41(2): 1672–0636.
YE Fawang, ZHANG Jielin, ZHANG Chuan, et al. Typical applications of airborne hyper-spectral remote sensing technology in uranium exploration: A case study of Xuemisitan area, Xinjiang [J]. *World Nuclear Geoscience*, 2024, 41(2): 233–249 (in Chinese).
 - 14 CAMPS-VALLS G, TUIA D, BRUZZONE L, et al. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods [J]. *IEEE Signal Processing Magazine*, 2013, 31(1): 45–54.
 - 15 HAUT J M, PAOLETTI M E, PLAZA J, et al. Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(11): 6440–6461.
 - 16 MELGANI F, BRUZZONE L. Classification of hyperspectral remote sensing images with support vector machines [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778–1790.
 - 17 HAM J, Chen Y, Crawford M M, et al. Investigation of the random forest framework for classification of hyperspectral data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(3): 492–501.
 - 18 CHEN Y, LIN Z, ZHAO X, et al. Deep learning-based classification of hyperspectral data [J]. *IEEE Journal of Selected topics in applied earth observations and*

- remote sensing, 2014, 7(6): 2094–2107.
- 19 RICHARDS J A. Analysis of remotely sensed data: The formative decades and the future[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3): 422–432.
- 20 ZHANG B, LI S, JIA X, et al. Adaptive Markov random field approach for classification of hyperspectral imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(5): 973–977.
- 21 DATTA A, GHOSH S, GHOSH A. Unsupervised band extraction for hyperspectral images using clustering and kernel principal component analysis[J]. International Journal of Remote Sensing, 2017, 38(3): 850–873.
- 22 LI J, BIOUCAS-DIAS J M, PLAZA A. Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(11): 4085–4098.
- 23 LILLESAND T, KIEFER R W, CHIPMAN J. Remote sensing and image interpretation[M]. Hoboken: John Wiley and Sons, 2015.
- 24 苏岫, 王祥, 宋德瑞, 等. 基于改进光谱角法的红树林高分遥感分类方法研究[J]. 海洋环境科学, 2021, 40(4): 639–646.
- SU Xiu, WANG Xiang, SONG Derui, et al. Research on high resolution remote sensing mangrove classification method based on improved spectral angle mapper[J]. Marine Environmental Science, 2021, 40(4): 639–646 (in Chinese).
- 25 VAPNIK V. The nature of statistical learning theory[M]. Berlin: Springer Science and Business Media, 2013.
- 26 BREIMAN L. Random forests[J]. Machine Learning, 2001, 45: 5–32.
- 27 QUINLAN J R. Induction of decision trees[J]. Machine Learning, 1986, 1: 81–106.
- 28 HAN J, ZHANG D, CHENG G, et al. Advanced deep-learning techniques for salient and category-specific object detection: A survey[J]. IEEE Signal Processing Magazine, 2018, 35(1): 84–100.
- 29 BEBIS G, GEORGIPOULOS M. Feed-forward neural networks[J]. IEEE Potentials, 1994, 13(4): 27–31.
- 30 ALAN J, JIA Xiuping. Remote sensing digital image analysis: An Introduction [M]. Berlin: Springer, 2006.
- 31 MATHER P M, KOCH M. Computer processing of remotely-sensed images [M]. Hoboken: John Wiley and Sons, 2022.
- 32 LILLESAND T, KIEFER R W, CHIPMAN J. Remote sensing and image interpretation[M]. Hoboken: John Wiley and Sons, 2015.
- 33 DING Y, ZHAO Y, SHEN X, et al. Yinyang k-means: A drop-in replacement of the classic k-means with consistent speedup[C]//International Conference on International Conference on Machine Learning, PMLR, 2015: 579–587.
- 34 ARTHUR D, VASSILVITSKII S. k-means⁺⁺: The advantages of careful seeding[C]//Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2007: 1027–1035.
- 35 BALL G H. ISODATA, a novel method of data analysis and pattern classification [J]. Stanford Research Institute, 1965: AD-699616.