

# 基于子区域多标签学习的露天煤矿区场景识别

赵银娣<sup>1,2</sup>, 卫虹宇<sup>1</sup>, 董霁红<sup>1</sup>, 董畅<sup>1</sup>

1. 中国矿业大学 环境与测绘学院, 徐州 221116;

2. 自然资源部退化及未利用土地整治工程重点实验室, 西安 710075

**摘要:** 露天煤矿开采对区域生态环境产生不利影响, 对其进行高效监管有利于矿区环境保护和可持续发展。随着遥感技术和人工智能的发展, 基于高分辨率遥感影像的露天煤矿区场景自动识别成为可能。本文针对单标签学习算法在场景子区域识别中识别率较低的问题, 将多标签学习策略和地理学第一定律相结合, 提出一种基于子区域多标签学习的露天煤矿区场景识别方法。为了区分露天煤矿区场景与其周边场景, 设置了6类矿区标签和7类非矿区标签, 对9768张场景子区域图像进行标注, 构建多标签数据集, 利用该数据集训练基于多标签学习的Inception-v3模型。场景识别时, 首先将一幅覆盖研究区的遥感影像划分为相同大小的子区域并进行多标签分类; 然后对含有矿区标签的子区域, 利用地理学第一定律对其矿区标签的相关性和完整性进行判定, 识别出属于露天煤矿区场景的子区域。胜利西露天煤矿区识别实验结果表明: 该方法提取的结果最接近真值, 显著高于单标签学习的识别精度; 其子区域多标签分类F1分数达到0.857, 与单标签学习中性能最好的ResNet50模型相比, 提高了8个百分点。本文提出的方法能够自动提取子区域内多类标签的有效特征, 提高露天煤矿区场景识别的精度, 其识别结果可为露天矿区开采管理提供数据支撑。

**关键词:** 高分辨率遥感影像, 露天煤矿区场景识别, 多标签学习, 场景子区域识别

**引用格式:** 赵银娣, 卫虹宇, 董霁红, 董畅. 2022. 基于子区域多标签学习的露天煤矿区场景识别. 遥感学报, 26(9): 1849-1858

Zhao Y D, Wei H Y, Dong J H and Dong C. 2022. Opencast coal mine scene recognition based on sub-region multi-label learning. National Remote Sensing Bulletin, 26(9): 1849-1858 [DOI: 10.11834/jrs.20210001]

## 1 引言

露天煤矿区场景识别是露天矿地质环境监管工作的重要内容之一。露天开采对生态环境的破坏极其严重, 容易导致地质灾害和环境污染等一系列问题, 利用高分辨率遥感影像识别露天煤矿区, 一方面可以为突发灾害提供决策依据, 另一方面有利于为矿区的土地管理和土地复垦规划提供相应帮助(胡振琪和谢宏全, 2005; 张凯翔, 2018)。

在遥感影像上识别露天煤矿区, 早期的方法主要依赖目视解译(毕如田等, 2008), 但是存在识别精度不高、工作效率较低等问题。针对这些问题, 国内外学者提出了许多方法, 主要分为基于像元的分类方法(程璐, 2017)和面向对象的

分类方法(祝振江, 2010)。这两类方法有效提高了露天煤矿区的识别精度和识别效率, 但也存在一些问题(张峰极, 2019)。基于像元的分类方法无法从整体上解译影像的空间、纹理等特征, 容易产生信息冗余, 面向对象的分类方法需要人为确定分类规则, 受人的主观因素影响较大。随着遥感影像分辨率的不断提高, 场景中包含的语义信息愈加丰富, 给露天煤矿区识别带来了更多挑战。

近年来, 场景识别技术(姜啸远, 2017)不断发展, 可以更好的提取露天煤矿区场景中包含的多类特征。起初, 研究者采用各种底层特征用于场景识别, 例如SIFT (Scale-Invariant Feature Transform)、HOG (Histogram of Oriented Gradient)、BoW (Bag of Words)等, 这些方法具有良好的逻辑

收稿日期: 2020-01-20; 预印本: 2020-07-24

基金项目: 自然资源部退化及未利用土地整治工程重点实验室开放基金(编号: SXDJ2019-4); 国家重点研发计划(编号: 2016YFC0501105); 中央高校基本科研业务费专项(编号: 2015XKMS050)

第一作者简介: 赵银娣, 研究方向为遥感数据处理、模式识别。E-mail: zhaoyd@cumt.edu.cn

通信作者简介: 卫虹宇, 研究方向为遥感目标检测、场景识别。E-mail: 925074644@qq.com

辑性,但是提取的特征泛化性不强,且在大量、多类场景信息的特征提取中效率较低(Zuo等,2014)。自AlexNet(Krizhevsky等,2012)使用卷积神经网络CNN(Convolutional Neural Network)大幅提高了图像分类的准确率,VGG(Simonyan和Zisserman,2015)、Inception(Szegedy等,2015;Ioffe和Szegedy,2015)、ResNet(He等,2016)等深度卷积网络相继提出,在特征提取(钱晓亮等,2018)、图像分类(张康等,2018)、场景识别(许凤晖等,2016)等领域取得了不错的成效。与传统方法相比,CNN提取的特征能够包含丰富的语义信息和结构信息,且对于平移、缩放等变形具有高度不变性,泛化能力强(杨鹏等,2018),因此本文采用场景识别技术,利用CNN模型提取的深层特征进行场景分类,从而识别出露天煤矿区。

露天煤矿区占地面积大且地物类别丰富,在高分辨率遥感影像上具有特征量大、特征多样的特点(张成业等,2022),对整个矿区场景进行特征学习并识别,给设备的计算能力提出了更高的要求。为了减少计算损耗,本文将场景划分为多个子区域,对露天煤矿区场景和非露天煤矿区场景的子区域进行特征学习,通过识别属于露天煤矿区场景的子区域从而识别出该场景。目前场景识别广泛使用单标签学习的策略(冯权洸等,2022),由于露天煤矿区场景子区域包含的地物在类别、位置、大小和形状等方面存在较大差异,并且不同地物间还存在遮挡问题,难以提取有效特征。

随着计算机视觉任务趋于多样化和复杂化,多标签学习(Tsoumakas和Katakis,2007;Zhang和Zhou,2014;陈智,2015)已逐渐成为机器学习领域的热点研究问题之一。不同于单标签学习中每个样本仅属于一个类别标签,多标签学习是通过学习有多个类别标签的数据来预测未知样本所属的标签集合。基于此,本文提出一种基于子区域多标签学习的露天煤矿区场景识别方法,通过学习露天煤矿区场景子区域中包含的多类标签的特征并进行多标签分类,再基于该场景区别于其他场景的标签进行分析,从而识别出所有属于该场景的子区域以构成整个场景。矿产资源开发要素(宋亚婷等,2016)主要包括采掘区、剥离区、边帮、排土场、废渣场、中转场地,这6类矿

区标签是露天煤矿区场景区别于非露天煤矿区场景的主要标签,另外,还设置了道路、水域、耕地、林地、草地、裸地、建筑物这7类非矿区标签与矿区标签进行区分,将露天煤矿区场景和非露天煤矿区场景的子区域都标注13类标签构成多标签数据集。鉴于Inception-v3(Szegedy等,2016)模型在ImageNet数据集上具有良好的特征提取能力,对Inception-v3模型进行改进,使其能够自动学习并提取多类标签的有效特征,训练该模型从而对子区域包含的多类标签进行预测。数据集中具有矿区标签的子区域即为露天煤矿区场景的子区域,但是在多标签预测时存在误差,需要分析子区域中的多标签预测结果对子区域进行判定,为此设计了一种场景子区域判定算法,结合地理学第一定律,根据矿区标签的标签相关性和标签完整性以判定该子区域是否属于露天煤矿区场景,识别出的子区域构成露天煤矿区场景。

## 2 方法原理

本文方法主要由基于多标签学习的卷积神经网络和场景子区域判定算法两部分组成。用基于多标签学习的卷积神经网络对数据集中子区域包含的多类标签进行学习,将输入影像划分为相同大小的子区域并进行多标签分类,采用场景子区域判定算法,根据分类结果判定子区域是否属于露天煤矿区场景,识别出的子区域构成露天煤矿区场景。

### 2.1 基于多标签学习的卷积神经网络

CNN在单标签学习任务中获得了优异的性能,然而,由于底层图像的复杂性和多标签训练数据的缺乏,如何更好地利用CNN处理多标签图像仍然是一个有待解决的问题(Shahriyar等,2018),本文对CNN模型进行改进以适应多类标签。

如图1所示,将CNN中经典模型之一的Inception-v3模型作为源模型,使用公开数据集ImageNet预训练模型对CNN模型的参数进行初始化,将露天煤矿区多标签数据集导入该模型,将最后一层全连接层的输出节点数调整为实验数据集的类别数,采用反向传播算法训练最后一层,通过参数迁移以得到更丰富的图像特征表示。为了使CNN模型可以学习多个语义标签并加快参数的更新速度,在CNN模型输出层采用Sigmoid函数作为输出层的激活函数(Kurata等,2016),并且

使用 Sigmoid 交叉熵损失函数代替单标签损失函数对模型进行训练优化。对于  $N$  个样本和  $K$  类标签, 损失函数  $J$  表达式如下:

$$J = -\frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K (y_n^k \ln(h(x_n)) + (1 - y_n^k) \ln(1 - h(x_n))) \quad (1)$$

其反向传播导数为

$$\frac{\partial J}{\partial x_n} = \frac{\partial J}{\partial h(x_n)} \frac{\partial h(x_n)}{\partial x_n} = \frac{1}{N} (h(x_n) - y_n^k) \quad (2)$$

式中,  $x_n$  表示第  $n$  个样本,  $y_n^k$  表示  $x_n$  对应标签集合中的第  $k$  个标签,  $h(x_n)$  表示 Sigmoid 函数, 其数学表达式为

$$h(x_n) = \frac{1}{1 + e^{-x_n}} \quad (3)$$

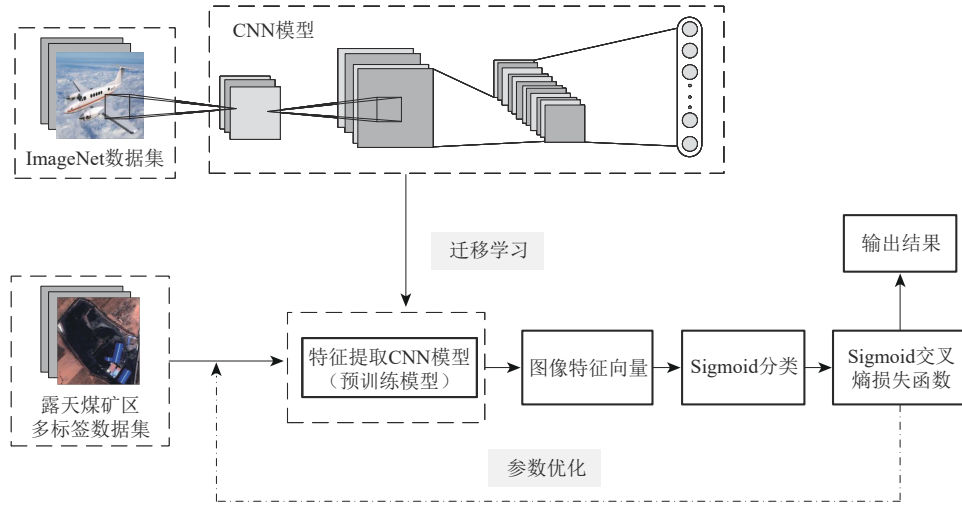


图1 基于多标签学习的卷积神经网络

Fig. 1 Convolutional neural network based on multi-label learning

## 2.2 场景子区域判定算法

设某类场景  $A$  共有  $p$  类核心标签, 多标签数据集共标注  $q$  ( $q > p$ ) 类标签, 为了判定每个子区域是否属于场景  $A$ , 本文设计了一种场景子区域判定算法, 图2为算法流程, 算法的主要思想如下:

(1) 获取相关标签集合。给定包含  $I$  个样本的多标签训练集  $D_{\text{train}} = \{(x_i, Y_i) | i = 1, 2, \dots, I\}$ , 其中  $Y_i = \{y_{i1}, y_{i2}, \dots, y_{iq}\}$  是样本  $x_i$  的  $q$  类标签集合。所有样本的标签集合  $L = \{l_1, l_2, \dots, l_q\}$  是一个  $q$  维的随机变量, 任意  $l_a$  和  $l_b$  的相关系数  $\rho_{ab}$  ( $a, b = 1, 2, \dots, q$ ) 计算公式为

$$\rho_{ab} = \frac{E((l_a - E(l_a))(l_b - E(l_b)))}{\sqrt{D(l_a)} \sqrt{D(l_b)}} \quad (4)$$

式中,  $D(l_a)$ 、 $D(l_b)$  分别为  $l_a$ 、 $l_b$  的方差,  $E(l_a)$ 、 $E(l_b)$  分别为  $l_a$ 、 $l_b$  的期望。

为了确定标签间的关联程度, 对于每类标签  $l_a$ , 取其所有相关系数 (除1以外) 的平均值作为阈值  $\theta$ , 仅考虑  $\rho_{ab}$  大于等于阈值  $\theta$  的标签, 因此, 标签  $l_a$  的相关标签集合为

$$\hat{L}_a = \{l_b | \rho_{ab} \geq \theta, b = 1, 2, \dots, q\} \quad (5)$$

(2) 获取子区域的多标签分类结果。将输入图像划分为相同大小的子区域, 得到包含  $T$  个子区域的多标签测试集  $D_{\text{test}} = \{\tilde{x}_t | t = 1, 2, \dots, T\}$ , 将其输入训练好的基于多标签学习的 CNN 模型进行  $q$  类多标签分类, 得到子区域  $\tilde{x}_t$  的预测标签  $\tilde{Y}_t = \{\tilde{y}_{t1}, \tilde{y}_{t2}, \dots, \tilde{y}_{tq}\}$ 。

(3) 根据核心标签的标签相关性对子区域进行第一次判定。根据地理学第一定律, 在空间上越靠近的事物或现象就越相似, 对每个子区域分析其是否存在核心标签以及子区域周边是否存在该核心标签的相关标签, 若是则提取该子区域。子区域  $\tilde{x}_t$  关于核心标签的预测标签  $\tilde{Y}_t = \{\hat{y}_{tj} | j = 1, 2, \dots, p\}$ , 对于每个标签  $\hat{y}_{tj}$ , 其相关标签集合为  $\hat{L}_{j\circ}$ 。子区域  $\tilde{x}_t$  的 8 个邻域集合为  $D_{\text{neighbor}} = \{\tilde{x}_{tm} | m = 1, 2, \dots, 8\}$ ,  $\tilde{x}_{tm}$  关于核心标签的预测标签  $\hat{Y}_{tm} = \{\hat{y}_{tm}^j | j = 1, 2, \dots, p\}$ 。假设  $\gamma_t$  为每个子区域  $\tilde{x}_t$  与其 8 个邻域的关于核心标签的相关标签总数, 计算公式如下:

$$\gamma_t = \sum_{j=1}^p \sum_{m=1}^8 \varphi(\tilde{x}_t) \quad (6)$$

$$\varphi(\tilde{x}_i) = \begin{cases} 0, & \hat{Y}_{im} \cap \hat{L}_j = \phi \\ 1, & \hat{Y}_{im} \cap \hat{L}_j \neq \phi \end{cases} \quad (7)$$

式中,  $\varphi(\tilde{x}_i)$ 为判断子区域的8个邻域是否存在第 $j$ 类核心标签的相关标签。根据每个测试样本 $\tilde{x}_i$ 的 $\gamma_i$ 值得到提取的子区域集合为 $D_A = \{\tilde{x}_i | \gamma_i > 0\}$ 。

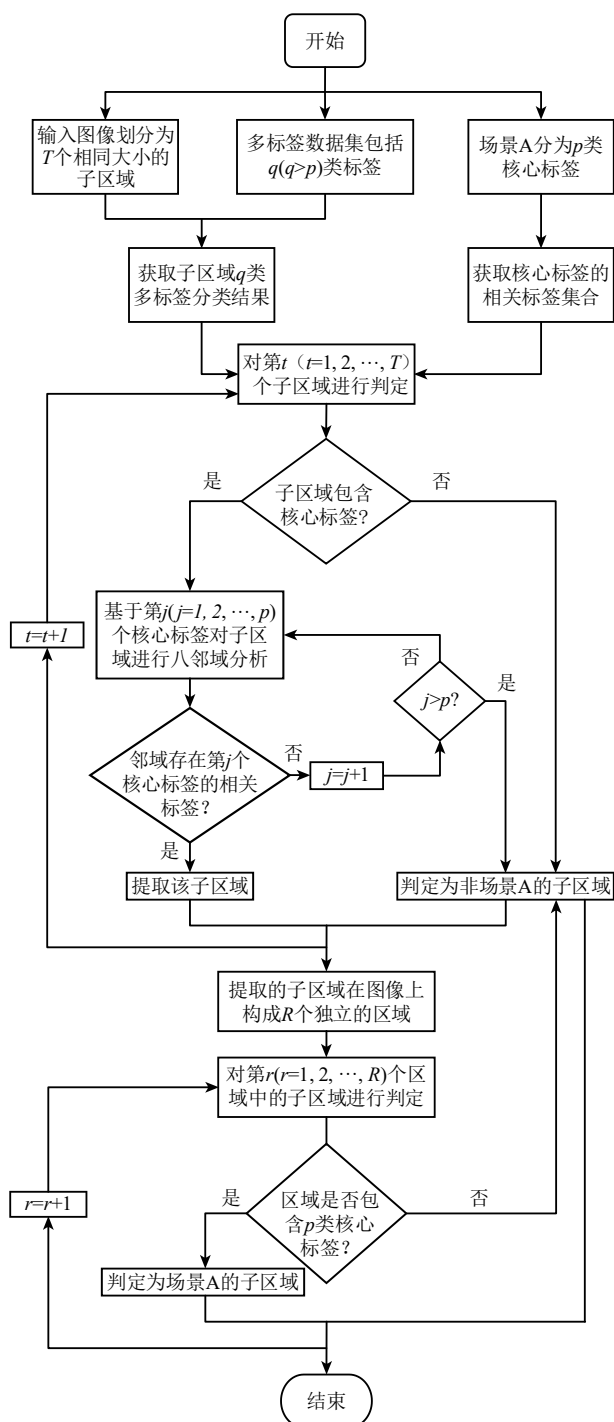


图2 场景子区域判定算法流程  
Fig. 2 Flowchart of the scene sub-region determination algorithm

(4) 根据核心标签的标签完整性对提取的子区域进行第二次判定。由于多标签分类存在误差, 需要过滤掉一部分有误的数据。将 $D_A$ 中相连的子区域 $\tilde{x}_i$ 分别合并到一个集合, 得到包含 $R$ 个独立区域的子区域集合 $D_R = \{D_r | r = 1, 2, \dots, R\}$ 。由于场景A包含 $p$ 类核心标签, 遍历集合 $D_R$ 中每个独立的区域 $D_r$ , 若 $p$ 类核心标签都存在, 则判定 $D_r$ 中的每个子区域属于场景A。

### 3 实验结果与分析

#### 3.1 研究区和数据集

##### 3.1.1 研究区概况

选择胜利煤田(马萧, 2011)的西露天煤矿区作为研究区, 该区域煤炭资源丰富, 露天煤矿分布比较集中。研究区所处位置为内蒙古高原中部, 大兴安岭西侧的北坡, 整体地形分布西北高、东南低, 东西两侧高差较大。该区域属中温带半干旱大陆性季风气候, 年均气温 $1.7\text{ }^\circ\text{C}$ 、降水量 $294.74\text{ mm}$ 。研究区大都是高台阶开采, 容易诱发坍塌、高空坠落等事故, 并且开采过程中产生的废渣会对地形地貌景观造成不同程度的破坏。研究区影像为谷歌影像, 如图3所示。



图3 研究区遥感影像图  
Fig. 3 Remote sensing image of the study area

##### 3.1.2 数据集

本文数据集可用在多标签学习和单标签学习中, 实验数据裁剪自谷歌影像, 数据集由露天煤矿区场景和非露天煤矿区场景的子区域构成, 共9768张图像。如图4所示, 若采用多标签学习, 共包含

13类标签，其中有6类矿区标签：采掘区、剥离区、边帮、排土场、废渣场、中转场地，有7类非矿区标签：道路、水域、耕地、林地、草地、裸地、建筑物。图4显示了每张图像包含的多类标签，图像由独热编码的方式进行标注。若采用单标签学

习，数据集分为露天煤矿区场景子区域和非露天煤矿区场景子区域这两类，其中具有矿区标签的数据属于露天煤矿区场景子区域。为了使一张图像涵盖多类标签，数据均裁剪为500×500像素大小的图像，每类图像数量不均衡，标签分布如图5所示。

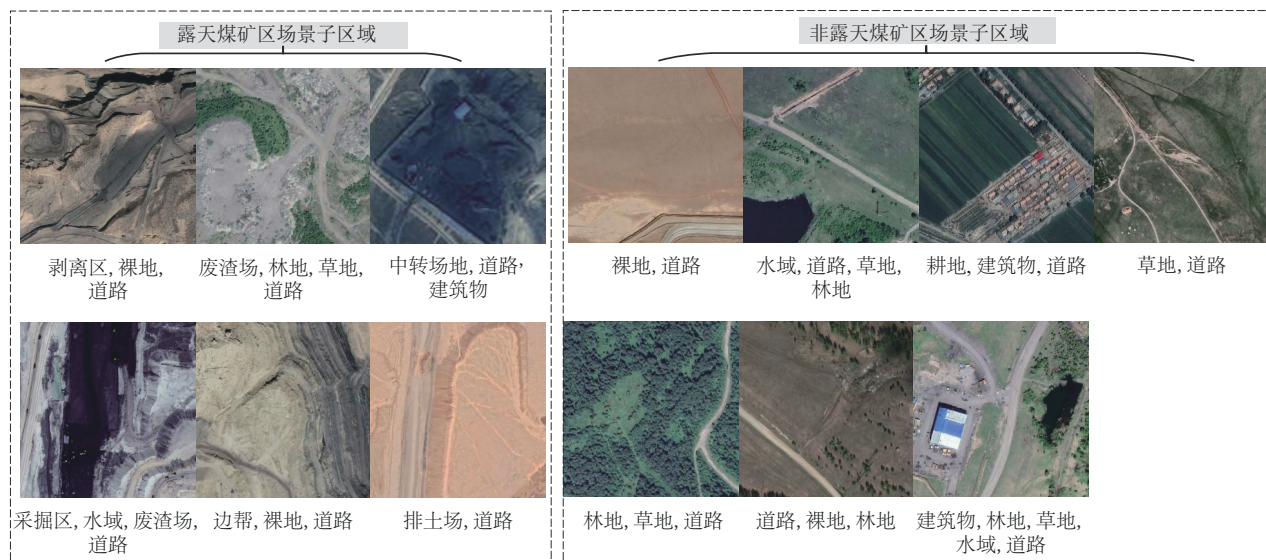


图4 多标签数据集示例

Fig. 4 Demonstration of the multi-label learning dataset

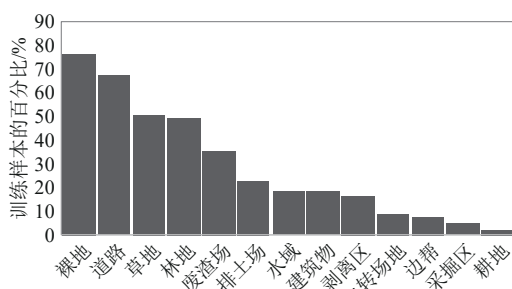


图5 标签分布图

Fig. 5 Distribution of labels

### 3.2 评价指标

本文使用的评价指标包括准确率 (Precision)、召回率 (Recall)、F1分数 (F1-Score) 和 F2分数 (F2-Score)，定义如下：

$$p = \frac{TP}{TP + FP} \quad (8)$$

$$r = \frac{TP}{TP + FN} \quad (9)$$

$$F_1 = (1 + \beta^2) \frac{pr}{\beta^2 p + r}, \beta = 1 \quad (10)$$

$$F_2 = (1 + \beta^2) \frac{pr}{\beta^2 p + r}, \beta = 2 \quad (11)$$

式中， $p$ 是准确率， $r$ 是召回率， $TP$ 为真正样本， $FP$ 为假正样本， $FN$ 为假负样本。

在子区域多标签分类中采用F2分数进行多标签阈值优化，采用准确率、召回率评估各标签的分类效果，以分析多标签分类对后续数据处理的影响。采用召回率、准确率和F1分数来评估露天煤矿区场景子区域的识别效果 (宋仁忠等, 2022)，F1分数的值若等于1，说明所有属于露天煤矿区场景的子区域被识别出来，即识别出整个露天煤矿区场景。

### 3.3 结果分析

#### 3.3.1 子区域多标签分类

将多标签数据集按8:2划分为训练集和测试集，在训练集上采用5折交叉验证法训练模型。实验采用I5处理器，运行内存为10 G，结合NVIDIA GTX 1060 (6 GB显存) GPU，在Ubuntu16.04系统上搭建CUDA9.0和CUDNN7.5的GPU运行环境，采用Pytorch深度学习框架训练模型。训练时先使用公开数据集ImageNet预训练模型对Inception-v3

模型参数进行初始化,再将最后一层全连接层的类别参数调整为实验数据集的类别数,将输出层的激活函数替换为Sigmoid函数。采用随机梯度下降优化器进行参数优化,动量为0.9,初始学习率为0.001,权重衰减系数为0.1(Dogo等,2016)。由于本实验需要将测试集的大部分类别预测出来以进行后续分析,每次训练结束后对每个类别基于F2分数进行阈值优化,搜索从0.01到1的100个潜在值作为每个标签的阈值,找到使F2分数最高的相应类别的阈值,根据阈值进行判定得到分类结果(Williams,2021)。

分析基于多标签学习的Inception-v3模型在测试集上生成的每类标签的准确率和召回率,如图6所示,大部分标签的召回率明显高于准确率,这是由于本文选择了专门针对F2分数优化的阈值,确保大部分标签分类正确以提高后续数据处理的精度。采掘区、中转场地和建筑物这三类标签的准确率和召回率相差较小,是因为这三类的特征明显,提取的特征表征能力强且在分类过程中不容易与其他类特征混淆。相比于矿区标签,大部分非矿区标签的准确率和召回率偏高,主要是本文数据集的标签分布不均衡,非矿区标签在数据集的占比较大造成的。

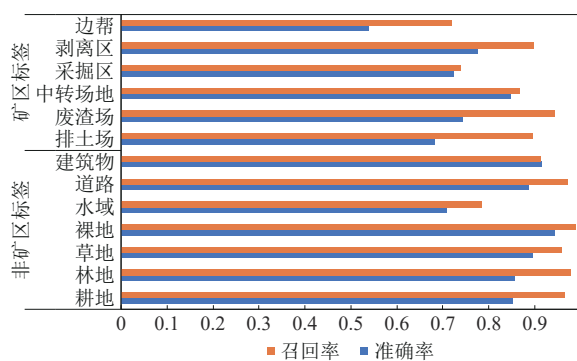


图6 子区域多标签分类结果

Fig. 6 Results of sub-region multi-label classification

将研究区遥感影像划分为相同大小的子区域,用训练的基于多标签学习的Inception-v3模型对子区域进行多标签分类,图7展示了部分子区域的多标签预测结果,黄色框代表错误预测的标签,蓝色框代表没有预测出来的标签。可以看到每个子区域包括多类标签,大部分标签分类正确,说明该多标签分类模型能从具有多类标签的样本中提取出有效的特征。对于分类有误的标签,模型对

采掘区、水域和废渣场容易混淆,对边帮容易误分为道路,由于本文方法主要基于矿区标签进行分析,矿区标签预测错误对场景子区域判定的精度有一定影响。

### 3.3.2 露天煤矿区场景识别

对研究区遥感影像的子区域进行识别,得到的结果如表1所示。表1中对比了本文方法和单标签学习算法对露天煤矿区场景子区域的识别结果。可以看到,单标签学习算法的最优F1分数为0.777,而本文子区域多标签分类F1分数达到0.857,其与单标签学习中性能最好的ResNet50模型相比,提高了8个百分点,说明该方法在露天煤矿区场景子区域识别中具有明显优势,识别能力更强。分析其原因,若采用单标签学习算法,把子区域分为露天煤矿区场景子区域和非露天煤矿区场景子区域,由于露天煤矿区场景子区域中包含非露天煤矿区场景子区域的一些地物,在特征提取中容易混淆,从而影响识别结果。而本文方法采用了多标签学习策略,能够学习到露天煤矿区场景子区域中区别于非露天煤矿区场景子区域的矿区标签,再结合矿区标签对子区域进行判定从而识别出露天煤矿区场景子区域。

此外,表1中本文方法的准确率达到最优精度,而召回率比准确率低了0.085,这主要是受子区域多标签分类的影响,结合图6进行分析,多标签分类模型对各类标签的召回率偏高,可以保证大部分标签预测正确,从而在露天煤矿区场景子区域识别中与非露天煤矿区场景子区域更好地区分,所以对露天煤矿区场景子区域识别的准确率较高,但是多标签分类中预测矿区标签的准确率不高,若没有预测出子区域包含的矿区标签,场景子区域判定算法就会将该子区域识别为非露天煤矿区场景子区域,使得召回率降低。

本文通过识别属于场景的子区域从而识别出整个场景,图8显示了研究区遥感影像的露天煤矿区场景识别结果,红色区域代表由识别出的子区域构成的露天煤矿区场景。将每种方法得到的识别图与真值图进行对比,单标签学习算法误将多个非露天煤矿区场景的子区域识别为露天煤矿区场景的子区域,而本文方法的识别结果接近真值,证明该方法在露天煤矿区场景识别中的有效性。

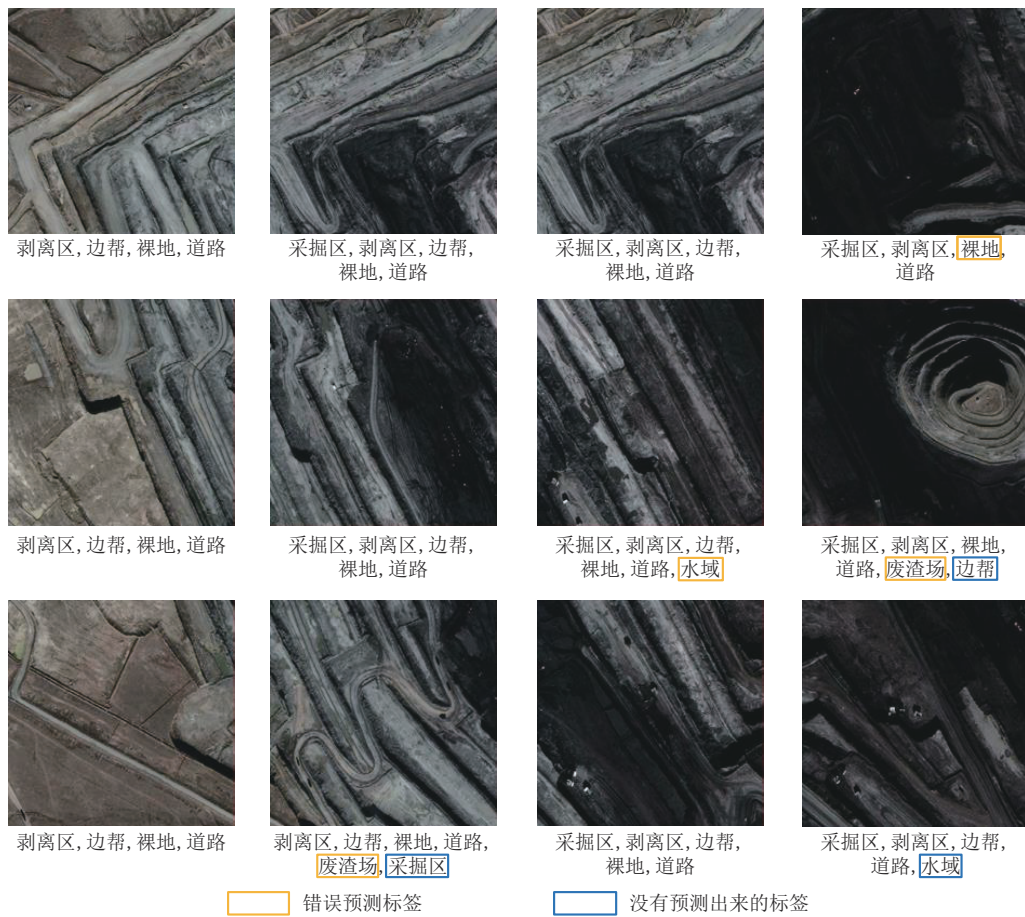


图7 研究区遥感影像的子区域多标签分类部分结果

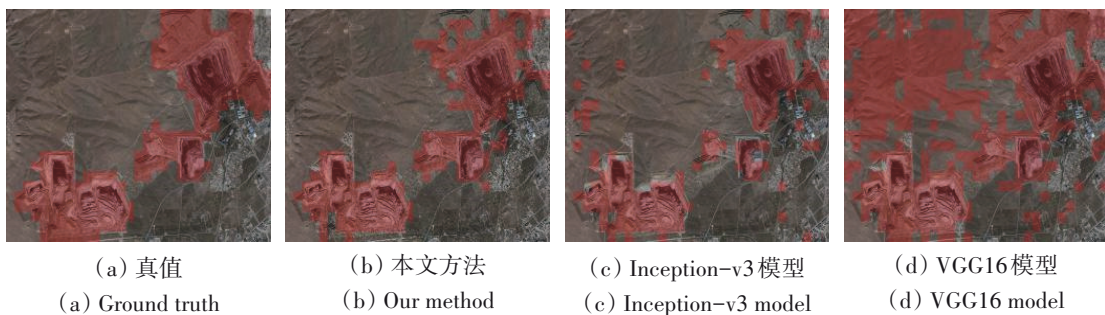
Fig. 7 Partial results of multi-label classification of remote sensing image sub-regions in the study area

表1 露天煤矿区场景子区域识别结果对比

Table 1 Comparison of scene sub-region recognition results in opencast coal mine areas

方法	学习策略	准确率	召回率	F1 分数
本文方法	多标签	<b>0.901</b>	0.816	<b>0.857</b>
ResNet50	单标签	0.734	0.825	0.777
Inception-v3	单标签	0.863	0.671	0.755
ResNet18	单标签	0.790	0.692	0.738
ResNet152	单标签	0.728	0.731	0.729
VGG16	单标签	0.495	<b>0.902</b>	0.665
VGG19	单标签	0.565	0.803	0.663

注:加粗文字表示各指标的最优精度。



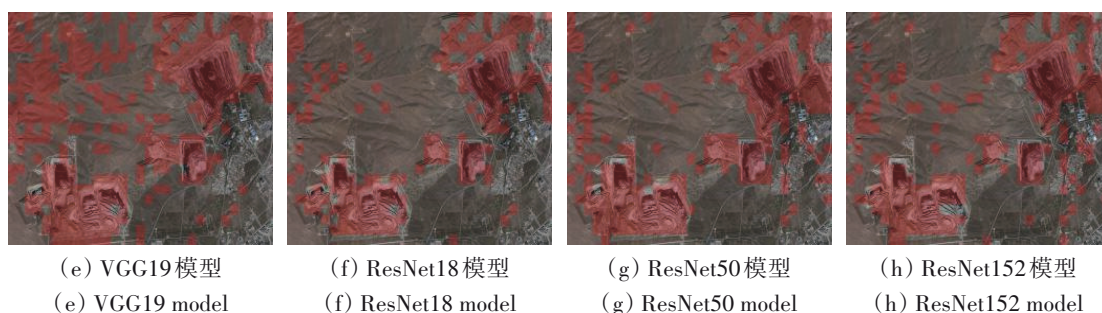


图8 露天煤矿区场景识别结果对比

Fig. 8 Comparison of scene recognition results in opencast coal mine areas

## 4 结论

本文将露天煤矿区场景划分为多个子区域进行特征学习和识别,针对单标签学习算法在场景子区域识别中识别率较低的问题,将多标签学习策略和地理学第一定律相结合,提出一种基于场景子区域多标签学习的场景识别算法。将场景划分为相同大小的子区域并标注多类标签,为了学习子区域内多类标签的特征,对Inception-v3模型进行改进使其适用于多类标签的数据。为了对子区域包含的多标签预测结果进行分析从而识别出属于露天煤矿区场景的子区域,设计了一种场景子区域判定算法,判定子区域是否属于露天煤矿区场景从而识别出露天煤矿区场景。实验结果表明:

(1) 基于多标签学习的Inception-v3模型能够自动学习并提取有效的多类标签的特征,但该模型对矿区标签的预测错误会导致露天煤矿场景子区域识别的召回率降低;

(2) 利用地理学第一定律对多标签分类结果进行分析从而识别露天煤矿区场景子区域,相比单标签学习算法的最优F1分数,本文方法的F1分数提高了8个百分点;

(3) 在研究区遥感影像上对露天煤矿区场景进行识别,本文方法的识别结果接近真值,证明该方法在露天煤矿区场景识别中的有效性。

本文方法是先进行多标签分类,再对分类结果进行分析以完成识别任务,后续还可以尝试算法优化以实现框架一体化;子区域多标签分类结果对露天煤矿场景子区域识别有一定影响,可以尝试其他算法提高多标签分类精度,以获得更好的场景识别效果;实验所用数据集标签分布不均衡,可以通过优化样本以提高模型精度。另外,

本文提出的方法也可以对研究区的多时相数据进行识别,以探究露天煤矿区范围的变化情况,为露天煤矿开采管理提供数据支撑。

## 参考文献(References)

- Bi R T, Bai Z K, Li H and Li W X. 2008. Land use changes in opencast mine based on RS and GIS technology. *Transactions of the CSAE*, 24(12): 201-204 (毕如田, 白中科, 李华, 李卫祥. 2008. 基于RS和GIS技术的露天矿区土地利用变化分析. *农业工程学报*, 24(12): 201-204 [DOI: 10.3321/j.issn:1002-6819.2008.12.042])
- Chen Z. 2015. Multi-label scene classification using convolutional neural network. Ji'nan: Shandong University (陈智. 2015. 基于卷积神经网络的多标签场景分类. 济南: 山东大学)
- Cheng L. 2017. Application of object-oriented combined SVM in information extraction of open-pit mine. Xining: Qinghai University (程璐. 2017. 面向对象结合支持向量机(SVM)在露天矿区信息提取中的应用研究. 西宁: 青海大学)
- Dogo E M, Afolabi O J, Nwulu N I, Twala B and Aigbavboa C O. 2018. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks//*Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems*. Institute of Electrical and Electronics Engineers Inc., United States [DOI: 10.1109/CTEMS.2018.8769211]
- Feng Q L, Chen B A, Li G Q, Yao X C, Gao B B and Zhang L C. 2022. A review for sample datasets of remote sensing imagery. *National Remote Sensing Bulletin*, 26(4): 589-605 (冯权泷, 陈泊安, 李国庆, 姚晓闯, 高秉博, 张连翀. 2022. 遥感影像样本数据集研究综述. *遥感学报*, 26(4): 589-605 [DOI: 10.11834/jrs.20221162])
- He K, Zhang X, Ren S and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hu Z Q and Xie H Q. 2005. Study on land use/cover change of coal



- mining area based on remote sensing images. *Journal of China Coal Society*, 30(1): 44-48 (胡振琪, 谢宏全. 2005. 基于遥感图像的煤矿区土地利用/覆盖变化. *煤炭学报*, 30(1): 44-48) [DOI: 10.3321/j.issn:0253-9993.2005.01.010]
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift//*Proceedings of the 32nd International Conference on Machine Learning*, PMLR 37: 448-456[<https://dl.acm.org/doi/10.5555/3045118.3045167>]
- Jiang X Y. 2017. Scene Recognition Based on Deep Learning. Nanjing: Nanjing University of Information Science & Technology (姜啸远. 2017. 基于深度学习的场景识别研究. 南京: 南京信息工程大学)
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25(2): 1097-1105 [DOI:10.1145/3065386]
- Kurata G, Xiang B and Zhou B. 2016. Improved neural network-based multi-label classification with better initialization leveraging label co-occurrence//*Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics [DOI: 10.18653/v1/n16-1063]
- Ma X. 2011. Ecological Risk Assessment of Vulnerable Mine Area — Taking Shengli East No.2 Open-pit as an Example. Beijing: China University of Geosciences (Beijing) (马箫. 2011. 脆弱性矿区生态风险评价——以胜利东二号露天矿为例. 北京: 中国地质大学(北京))
- Qian X L, Li J, Cheng G, Yao X W, Zhao S N, Chen Y B and Jiang L Y. 2018. Evaluation of the effect of feature extraction strategy on the performance of high-resolution remote sensing image scene classification. *Journal of Remote Sensing*, 22(5): 758-776 (钱晓亮, 李佳, 程焱, 姚西文, 赵素娜, 陈宜滨, 姜利英. 2018. 特征提取策略对高分辨率遥感图像场景分类性能影响的评估. *遥感学报*, 22(5): 758-776) [DOI: 10.11834/jrs.20188015]
- Shahriyar S A, Alam K M R, Roy S S and Morimoto Y. 2018. An approach for multi label image classification using single label convolutional neural network//*21st International Conference of Computer and Information Technology*. Institute of Electrical and Electronics Engineers Inc., United States[DOI: 10.1109/ICCITECHN.2018.8631970]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//*3rd International Conference on Learning Representations*[<https://arxiv.org/abs/1409.1556>]
- Song R Z, Zheng H Y, Wang D C, Shang Z, Wang Y J, Zhang C Y, Li Y. 2022. Classification of features in open-pit mining areas based on deep learning and high-resolution remote sensing images. *China Mining Magazine*, 31(7): 102-111 (宋仁忠, 郑慧玉, 王党朝, 尚志, 王兴娟, 张成业, 李军. 2022. 基于深度学习和高分辨率遥感影像的露天矿地物分类方法. *中国矿业*, 31(7): 102-111) [DOI: 10.12075/j.issn.1004-4051.2022.07.010]
- Song Y T, Jiang D, Huang Y H and Wan H W. 2016. Research on the priority of the land use types extraction of opencast mine area based on object-oriented classification. *Remote Sensing Technology Application*, 31(3): 572-579 (宋亚婷, 江东, 黄耀欢, 万华伟. 2016. 基于面向对象方法的露天煤矿用地类型提取优先级分析. *遥感技术与应用*, 31(3): 572-579) [DOI: 10.11873/j.issn.1004-0323.2016.3.0572]
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society [DOI: 10.1109/CVPR.2015.7298594]
- Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. 2016. Rethinking the inception architecture for computer vision//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society [DOI: 10.1109/CVPR.2016.308]
- Tsoumakas G and Katakis I. 2007. Multi-label classification: an overview. *International Journal of Data Warehousing and Mining*, 3(3): 1-13 [DOI: 10.4018/jdwm.2007070101]
- Williams C K I. 2021. The effect of class imbalance on precision-recall curves. *Neural Computation*, 33(4): 853-857 [DOI: 10.1162/neco\_a\_01362]
- Xu S H, Mu X D, Zhao P and Ma J. 2016. Scene classification of remote sensing image based on multi-scale feature and deep neural network. *Acta Geodaetica et Cartographica Sinica*, 45(7): 834-840 (许夙晖, 慕晓冬, 赵鹏, 马骥. 2016. 利用多尺度特征与深度神经网络对遥感影像进行场景分类. *测绘学报*, 45(7): 834-840) [DOI: 10.11947/j.AGCS.2016.20150623]
- Yang P, Cai Q Q, Sun H and Sun L H. 2018. Indoor scene recognition based on convolutional neural network. *Journal of Zhengzhou University (Natural Science Edition)*, 50(3): 73-77 (杨鹏, 蔡青青, 孙昊, 孙丽红. 2018. 基于卷积神经网络的室内场景识别. *郑州大学学报(理学版)*, 50(3): 73-77) [DOI: 10.13705/j.issn.1671-6841.2017260]
- Zhang C Y, Li F Y, Li J, Xing J H, Yang J Z, Guo J T and Du S H. 2022. Recognition of land use on open-pit coal mining area based on DeepLabv3+ and GF-2 high-resolution images. *Coal Geology & Exploration*, 50(6): 94-103 (张成业, 李飞跃, 李军, 邢江河, 杨金中, 郭俊廷, 杜宇航. 2022. 基于DeepLabv3+与GF-2高分辨率影像的露天煤矿区土地利用分类. *煤田地质与勘探*, 50(6): 94-103) [DOI: 10.12363/issn.1001-1986.22.01.0029]
- Zhang F J. 2019. Research on deep learning extraction method in open mining area based on multi-source remote sensing images. Hefei: Anhui University (张峰极. 2019. 多源遥感影像露天开采区深度学习提取方法研究. 合肥: 安徽大学)
- Zhang K, Hei B Q, Zhou Z and Li S Y. 2018. CNN with coefficient of variation-based dimensionality reduction for hyperspectral remote sensing images classification. *Journal of Remote Sensing*, 22(1):

- 87-96 (张康, 黑保琴, 周壮, 李盛阳. 2018. 变异系数降维的 CNN 高光谱遥感图像分类. 遥感学报, 22(1): 87-96) [DOI: 10.11834/jrs.20187075]
- Zhang K X. 2018. Research and demonstration application on interpretation sign system and information extraction method in open-pit mine. Wuhan: China University of Geosciences(Wuhan) (张凯翔. 2018. 露天矿地质环境解译标志体系和信息提取方法的研究及示范应用. 武汉: 中国地质大学(武汉))
- Zhang M and Zhou Z. 2014. A review on multi-label learning algorithms. IEEE Transactions on Knowledge and Data Engineering, 26(8): 1819-1837 [DOI: 10.1109/TKDE.2013.39]
- Zhu Z J. 2010. Study on mine area information extraction based on object-oriented high-resolution remote sensing image classification and its application. Beijing: China University of Geosciences (Beijing) (祝振江. 2010. 基于面向对象分类法的高分辨率遥感影像矿山信息提取应用研究. 北京: 中国地质大学(北京))
- Zuo Z, Wang G, Shuai B, Zhao L F, Yang Q Q and Jiang X D. 2014. Learning discriminative and shareable features for scene classification//European Conference on Computer Vision. Springer: 552-568 [DOI: 10.1007/978-3-319-10590-1\_36]

## Opencast coal mine scene recognition based on sub-region multi-label learning

ZHAO Yindi<sup>1,2</sup>, WEI Hongyu<sup>1</sup>, DONG Jihong<sup>1</sup>, DONG Chang<sup>1</sup>

1. School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China;

2. Key Laboratory of Degraded and Unused Land Consolidation Engineering, Ministry of Natural Resources, Xi'an 710075, China

**Abstract:** Opencast coal mining activities would lead to negative impacts on regional ecological environment, ensuring efficient monitoring and regulation of mining activities would promote environmental protection and sustainable development. With the development of remote sensing technology and artificial intelligence, there is great potential in automatically detecting opencast coal mine areas from high spatial resolution remote sensing imagery. Aiming at the problem of low recognition rate of scene sub-region recognition via single-label learning algorithm, this paper proposes an opencast coal mine scene recognition method by integrating multi-label learning and the first law of geography. In order to distinguish the opencast coal mine scenes from their surrounding different scenes, six categories of mining labels and seven categories of non-mining labels are set, and 9768 sub-region images are annotated to create a multi-label dataset. The Inception-v3 model is trained using the created dataset to perform multi-label classification. For scene recognition, firstly, the remote sensing image covering the study area is divided into non-overlapping sub-regions of the same size and the multi-label classification is carried out on the divided sub-regions. Then, inspired by the first law of geography, the sub-regions containing the mining labels are assigned to the coal mine scene type or not according to the correlation between the labels and the completeness of the labels. Finally, all the sub-regions judged as the coal mine scene type constitute the opencast coal mine scene recognition result from the high spatial resolution remote sensing image covering the study area. The experimental results show that the recognition result of Shengli west opencast coal mine areas obtained by the proposed method is much closer to the ground truth than the comparative methods based on single-label learning. The F1 score of the proposed method reaches 0.857 in the multi-label classification, with an increase of 8 percentage points compared to the ResNet50 single-label learning method which has the best performance in the compared single-label learning methods. The proposed method can automatically extract the effective features of multiple labels in sub-regions and improve the performance of opencast coal mine scene recognition, its recognition results can provide data support for opencast mining management.

**Key words:** high-resolution remote sensing image, opencast coal mine scene recognition, multi-label learning, scene sub-region recognition

**Supported by** Open Fund of Key Laboratory of Degraded and Unused Land Consolidation Engineering, Ministry of Natural Resources (No. SXDJ2019-4); National Key Research and Development Plan Project (No. 2016YFC0501105); Special Fund for Fundamental Research Funds of Central Universities (No. 2015XKMS050)