

多光谱波段筛选的煤矸石快速定位

来文豪, 周孟然*, 王锦国, 胡天羽, 孔茜茜, 胡锋, 卞凯, 朱梓伟

安徽理工大学电气与信息工程学院, 安徽 淮南 232000

摘要 煤矿的现代化建设需要实现煤矸石的智能分选。设计了一个基于多光谱图像的煤矸石检测模型,其目的是快速、准确地为智能分选提供煤矸石位置。首先,采集煤和煤矸石的多光谱数据,利用最佳指数因子理论从多光谱图像的 25 个波段中选择 3 个构成伪 RGB 图像。其次,设计一个轻量的检测模型,记为 YOLO mg。最后,基于伪 RGB 图像将设计的模型用于煤矸石检测研究,并通过非极大值抑制联合置信度阈值过滤冗余边界框。实验结果显示,YOLO mg 在 204 pixel \times 204 pixel 的输入分辨率下,实现了煤矸石 91.91% 的平均检测精度,并且 115 张光谱图像的检测耗时仅为 1.225 s。所设计的轻量检测模型能实现煤矸石最快速的识别和定位,这对煤矸石的智能分离操作具有重要意义。

关键词 光谱学; 多光谱图像; 最佳指数因子; 波段选择; YOLO mg; 识别和定位

中图分类号 O433.5

文献标志码 A

doi: 10.3788/CJL202148.1611001

1 引言

煤作为人类最主要的一次能源之一,每年都被大量开采^[1]。开采过程的伴生物煤矸石在原煤中的质量占比超 10%^[2],它会降低煤的燃烧效率,加剧环境污染。因此需将煤矸石从原煤中分离,该过程被称为“洗煤”^[3-4]。煤炭行业作为我国重要的传统能源行业,其智能化建设直接关系到我国国民经济和社会智能化的发展进程^[5]。发展现代煤矸石分离技术对煤矿的智能化建设具有重要的推动作用。首先通过机器学习算法识别煤矸石,然后利用机械手^[6]或气动装置^[7]将煤矸石从煤中分拣出,该方法被称为煤矸石智能分离法^[8]。传统方法是利用射线识别煤矸石^[9-10],成本高且会产生辐射。随着计算机技术的发展,基于成像技术的煤矸石识别已得到越来越多的关注^[11-12]。例如,Dou 等^[13]提取了煤矸石图像的 19 个特征,并利用支持向量机算法识别了最优特征,构造了最优分类器;Lai 等^[14]为了克服 RGB 成像在煤矸石识别中易受可见光干扰这一不足,提出了基于多光谱成像技术的煤矸石识别,并研究了不同波段的煤矸石识别特性。通过识别只能判

断图像中有没有煤矸石,而无法对煤矸石进行定位。因此,基于成像的煤矸石智能分离仍需进一步研究。

目标检测是识别和定位的联合^[15-16],随着深度学习理论的发展和计算机硬件的进步,其已取得了巨大成果^[17]。以 YOLO(You Only Look Once)^[18]和 SSD(Single Shot Multi-Box Detector)^[19]为代表的定位模型,在速度上有着巨大优势^[20]。人们又相继公布了 YOLO 和 SSD 的改进版本,分别记为 v2^[21]、v3^[22]、v4^[23]以及 M2Det^[24]。当前的目标检测模型都是面向大规模视觉挑战赛而设计的,模型较为复杂,这使得它们能在各种复杂任务中都有出色的表现,但在面向具体任务时,往往又有改进空间。例如,Tian 等^[24]针对同一模型不能适应不同生长阶段苹果的检测问题,通过改进 YOLO v3 实现了果园苹果的实时检测。针对煤矸石检测,Li 等^[25]提出煤矸石区域建议网络(Coal and Ganguue Regional Proposal Networks, CG-RPN)模型,该模型可实现煤矸石定位,但是检测速度有待提升;Lü 等^[26]在 YOLO v3 基础上提出了一种用于煤矸石检测的级联网络,但模型较为复杂。本课题组通过改进 YOLO v4 研究了煤矸石检测,实现了煤矸石的

收稿日期: 2020-12-17; 修回日期: 2021-01-11; 录用日期: 2021-02-18

基金项目: 国家重点研发计划(2018YFC0604503)、安徽理工大学 2019 年博士创新基金项目(2019CX1005)、安徽省博士后科研资助项目(2019B350)、安徽理工大学高层次引进人才科研启动基金(2021yjrc02)

通信作者: *mrzhou8521@163.com

高精度识别和定位,并且在一定程度上减少了检测时间^[27]。但改进的首要目的是追求更高的检测精度,检测速度有待进一步提升。煤矸石的定位速度直接关系到煤矿的生产效率,针对此问题,基于 YOLO v4 架构,本文设计了一个煤矸石检测模型,旨在实现煤矸石更快的识别和定位。

基于多光谱成像技术检测煤矸石,可降低环境中可见光的干扰。由于多光谱图像包含几十个波段,每个波段都是一张二维的光谱图像,将全部的多光谱数据用于煤矸石检测,不仅极大地限制检测速度,且冗余信息还会对检测造成不可预知的干扰。文献^[27]选择的组合波段尽管实现了煤矸石的准确定位,但缺乏理论依据。最佳指数因子(Optimum Index Factor, OIF)在高光谱波段选择中扮演重要角色^[28],如 Acharya 等^[29]将 OIF 用于卫星图像的最佳组合波段选择。综上所述,OIF 可用于多光谱组合波段的选择。

针对多光谱图像分辨率低以及检测目标仅有煤和煤矸石两类,本文设计的检测模型的主干网络采用相对较小的输入分辨率以及相对较窄和浅的网络宽度。此外,为了追求更快的检测速度,对模型颈部进行了轻量化的设计。面向具体的生产需求,本文设计了一个轻量检测模型,记为 YOLO mg,其目的是实现煤矸石更快速和高效的识别和定位,为煤矸

石的智能分离操作提供坐标信息。此外,本文的轻量检测模型的设计也为其他领域的快速目标检测模型的设计提供了参考。

2 实验材料与相关理论

2.1 实验设备及数据采集

多光谱成像技术就是把入射的全波段或宽波段的光信号分成若干个窄波段的光束,然后把它们分别成像在相应的探测器上,从而获得不同光谱波段的图像。实验所用光谱成像是由德国 XIMEA 生产的 xiQ 系列低功耗超紧凑型面阵相机,面阵相机的分辨率为 2048 pixel×1088 pixel,光谱的通道数一共有 5×5 个,共有 25 个波段,波段 1 到波段 25 对应的光谱波长分别为 891,900,882,872,959,798,811,786,773,682,748,762,736,722,697,932,939,924,914,954,851,863,841,829,946 nm。每个波段的光谱图像的分辨率为 409 pixel×216 pixel。采集的煤矸石多光谱数据如图 1 所示。

图 1 是煤和煤矸石的影像图,可以看出,煤和煤矸石的多光谱影像的差异较大。图 2 是不同波段的光谱数据,可以看出,在不同波长下,煤和煤矸石的光谱强度存在差异。实验一共采集 640 组多光谱图像,采集数据的实验划分如表 1 所示。

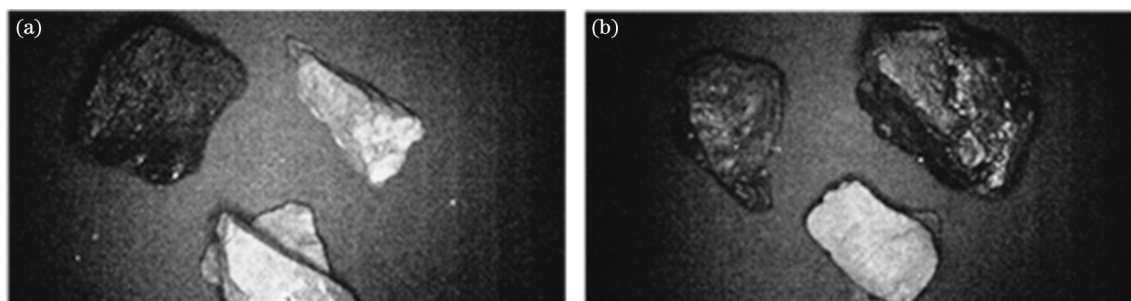


图 1 煤和煤矸石多光谱数据。(a)一块煤;(b)两块煤

Fig. 1 Multispectral data of coal and coal gangue. (a) One piece of coal; (b) two pieces of coal

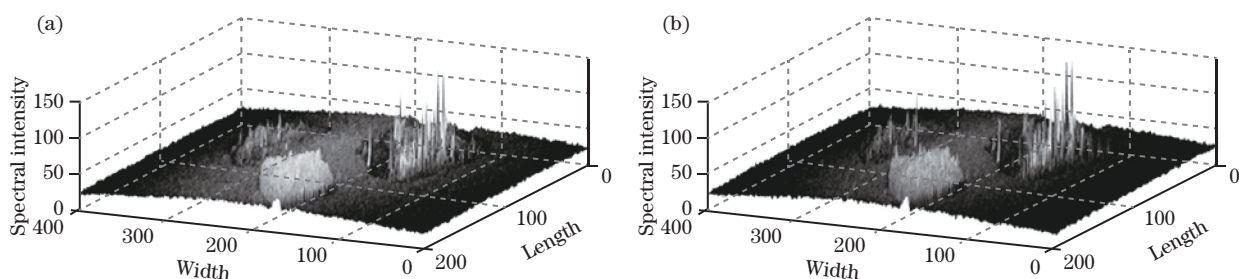


图 2 不同波段的光谱数据。(a)最小波长的光谱数据;(b)最大波长的光谱数据

Fig. 2 Spectral data at different bands. (a) Spectral data for minimum wavelength; (b) spectral data for maximum wavelength

表 1 煤矸石检测研究的数据划分

Table1 Data division for research of coal gangue detection

Sample	Training set		Validation set		Test set		Total
	1 lump	2 lumps	1 lump	2 lumps	1 lump	2 lumps	
Coal	140	55	15	10	20	10	250
Gangue	140	55	15	10	20	10	250
Mix	-	70	-	15	-	15	100
Total	280	180	30	35	40	35	600

表 1 中的“1 lump”代表一张光谱图像中只有一块煤或煤矸石，“Mix”代表一张光谱图像中既包含煤也有煤矸石。训练数据和验证数据分别为 460 组和 65 组。测试集不仅包含表 1 中展示的 75 组，还包括 10 组三块煤、10 组三块煤矸石、两块煤一块煤矸石以及一块煤两块煤矸石的光谱图像各 10 张，一共 115 张。

实验使用的是个人台式计算机，处理器为 Intel (R) Core (TM) i7-9700K CPU (32GB RAM)，显卡使用的是 NVIDIA © RTX 2070，显存为 8GB。

2.2 OIF 组合波段选择理论

实验使用的多光谱成像仪有 25 个波段，意味着每组多光谱图像包含 25 张单通道的图像，将这些数据全部用于煤矸石检测，大量的冗余信息不仅影响检测精度，而且会极大降低检测速度。OIF 综合考虑单波段图像的信息量及各波段间的相关性，计算简单，易于实现，在高光谱图像波段的选择方面得到了广泛的应用。对于给定波段的光谱图像，OIF 的计算公式为

$$O = \sum_{k_1=1}^n s_{k_1} / \sum_{k_1=1}^n \sum_{k_2=1}^n |\rho_{k_1 k_2}|, \quad (1)$$

式中： O 为 OIF； s_{k_1} 为组合波段中第 k_1 个波段的光谱图像的标准差； $\rho_{k_1 k_2}$ 为组合波段中第 k_1 个波段和第 k_2 个波段的光谱图像的相关系数； n 为选择的波段数，文中被设置为 3。OIF 方法联合标准差和相关系数用于波段选择时，既使用了图像的信息量，也使用到了图像间的相关性。第 k_1 个波段的标准差以及第 k_1 个波段与第 k_2 个波段的光谱图像的相关系数的计算式分别为

$$\rho_{k_1 k_2} = \frac{\sum_i \sum_j (A_{ij} - \bar{A})(A'_{ij} - \bar{A}')}{\sqrt{\sum_i \sum_j (A_{ij} - \bar{A})^2 \times \sum_i \sum_j (A'_{ij} - \bar{A}')^2}}, \quad (2)$$

$$s_{k_1} = \frac{1}{N \times M} \sqrt{\sum_{i=1}^N \sum_{j=1}^M (A_{ij} - \bar{A}')^2}, \quad (3)$$

式中： A_{ij} 和 A'_{ij} 分别是第 k_1 个波段和第 k_2 个波段

的光谱图像的第 i 行第 j 列的像素值； \bar{A} 和 \bar{A}' 分别是它们的均值； N 和 M 是图像的长和宽的像素点数。相关性系数越小， k_1 和 k_2 的差异相对就越大。

2.3 检测模型的评价指标

目标检测是图像中物体识别和定位的组合，仅使用准确率是无法评价一个目标检测模型的性能好坏。在目标检测领域，精确率、召回率、平均精度 (Average Precision, AP) 和均均精度 (mean Average Precision, mAP) 是使用最广泛的评价指标。当预测结果为正例时，记为 Positive；当预测结果为反例时，记为 Negative；当预测结果与真实值相同时，记为 True；当预测结果与真实值相反时，记为 False。精确率、召回率和平均精度的计算式分别为

$$P = T / (T + F), \quad (4)$$

$$R = T / (T + F'), \quad (5)$$

$$A = \int P dR, \quad (6)$$

式中： P 为精度，代表在预测结果为正样本的数据中，实际正样本数所占的比例； T 为正确预测的正样本数量； F 为错误预测为正样本的样本数量； F' 为错误预测为负样本的样本数量； R 为召回率，代表模型预测的所有样本中正确预测样本的占比； A 为平均精度。

3 目标检测理论及模型设计

YOLO 是一种将传统目标检测问题巧妙转化为回归问题的深度学习模型。相比于其先前版本和 RCNN^[30] 等其他模型，YOLO v4 的一个显著优点就是允许用户在一般性能的图像处理器 (Graphics Processing Unit, GPU) 上训练模型。因此，参考 YOLO v4 的框架，本文设计了一个轻量的煤矸石多光谱图像定位模型，记为 YOLO mg。

3.1 目标检测及其相关理论

YOLO v4 模型框架可分为三个部分，分别为用于特征学习的主干网络 (backbone)、标签预测和边框回归的头部 (head) 以及将 backbone 与 head 相连的颈部 (neck)。YOLO v4 的检测原理如图 3 所示。

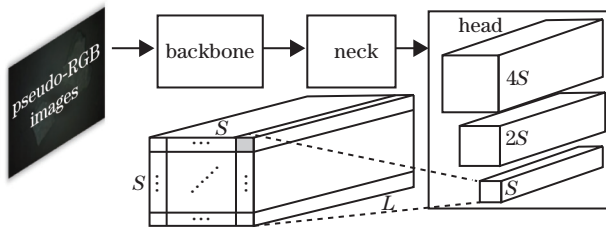


图 3 YOLO v4 模型简图

Fig. 3 Simplified diagram of YOLO v4 model

其头部采用三种刻度的网格, 网格的分辨率为 $S \times S$ 、 $2S \times 2S$ 和 $4S \times 4S$, 在模型确定的情况下, S 取决于图像的输入分辨率, 在 $416 \text{ pixel} \times 416 \text{ pixel}$ 分辨率下, S 的值为 13. L 为每个网格的长度, $L = (5 + C) \times B$, 其中 C 为检测目标的类别数, B 为锚框数, YOLO v3 和 YOLO v4 的 B 取值为 3.

所谓的交并比 (IoU), 就是预测边界框与实际边界框的相交面积除以相并的面积。在目标检测中, 一般情况下, 当 IoU 大于等于 0.5 就认为目标被准确定位。

在实际检测中, 同一目标往往会被预测出多个边界框。非极大值抑制 (Non-Maximum Suppression, NMS)^[31] 常被用于过滤冗余的预测边界框, 所谓非极大值抑制, 就是当两个预测边界框的 IoU 大于给定值时, 保留得分高的边界框。在应用中, NMS 的 IoU 的取值一般为 0.3~0.5, 本文设为 0.4.

3.2 检测模型的评价指标

YOLO mg 的网络架构分为三部分, 分别为主干网络、颈部和头部, 网络连接方式如图 4 所示。

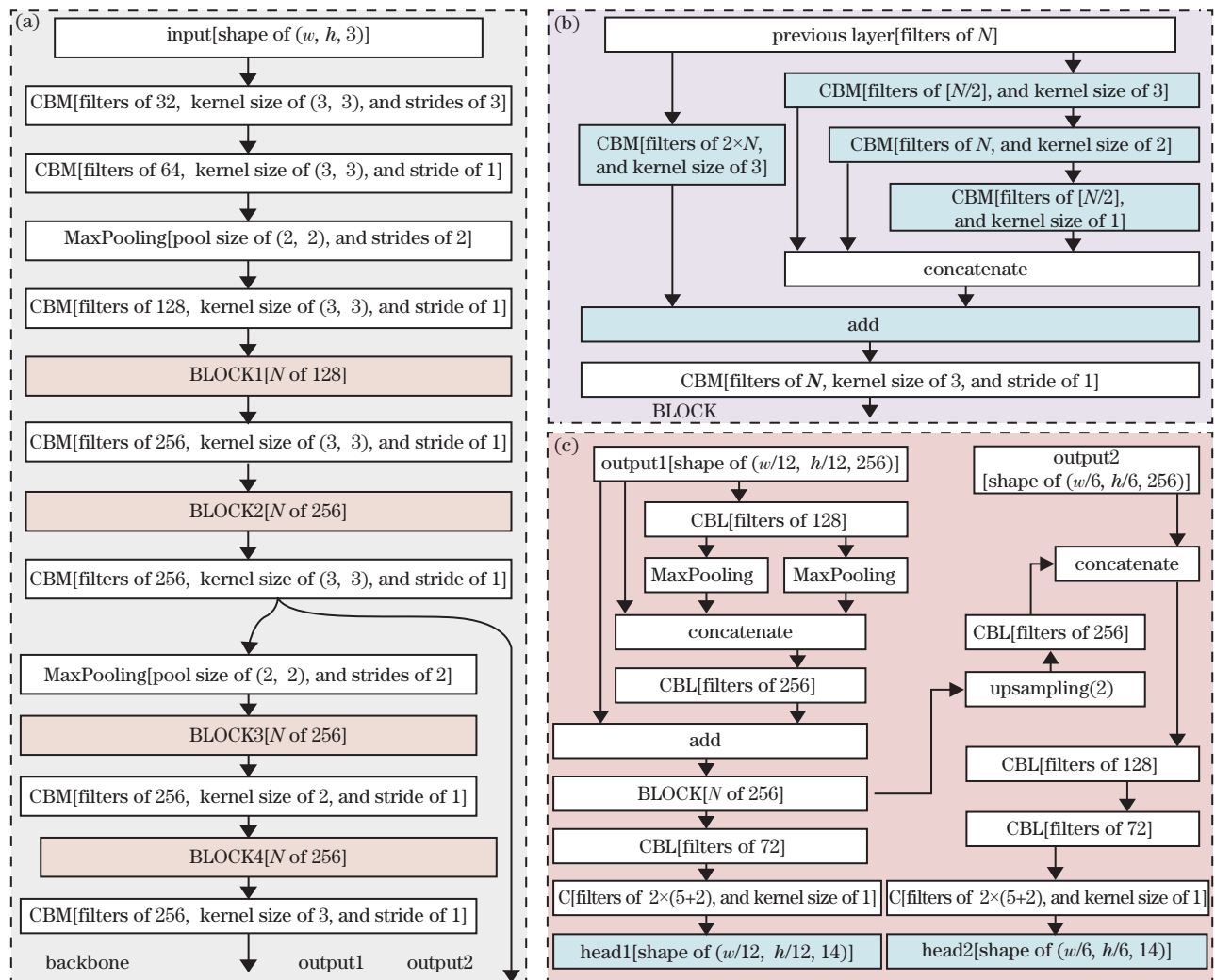


图 4 YOLO mg 的网络结构。(a) 主干; (b) 头部; (c) 颈部

Fig. 4 Network structure of YOLO mg. (a) Backbone; (b) head; (c) neck

在图 4(a) 中, “CBM” 代表卷积、批归一化和 Mish 激活, w 和 h 分别为输入图像的宽和高, “filters” 为卷积核数, “strides” 为卷积步长, N 为与之相连的前一

层特征图数, $[N/2]$ 为 N 除以 2 取整, “shape” 为输出的特征图的尺度。在图 4(b) 中, BLOCK 是主干网络中的 BLOCK1、BLOCK2、BLOCK3 和 BLOCK4 的网

网络结构。所设计的 BLOCK 使用 5 个卷积层,但其网络最大深度只用 4 层,因此主干网络的最大深度为 23 层。BLOCK 前三层卷积操作的核大小分别为 3×3 、 2×2 和 1×1 。BLOCK 块的输入输出通道数相同,即特征图数不变。

由于多光谱图像分辨率相对较小,为了进一步加快模型的检测速度,所设计的轻量主干网络的输入采用相对较小的分辨率。此外,主干网络的第一卷积层采用步长为 3 的卷积操作,它可以快速减小特征图尺寸。以 $204 \text{ pixel} \times 204 \text{ pixel}$ 分辨率为例,特征图像分辨率经历一次卷积操作就被降至为 $68 \text{ pixel} \times 68 \text{ pixel}$ 。Hasanpour 等^[32]指出,具有简单 Max-Pooling 架构的卷积神经网络(Convolutional Neural Networks, CNN)模型是优于跨步卷积模型。主干网的下采样除第一个采用多步长卷积操作外,其余均采用最大池化法,共有两个最大池化层,池化核为 3×3 ,池化窗口滑动步长为 2。

图 4(c) 是检测模型的 neck,“output1”和“output2”对应主干网络的输出,“BLOCK”与主干网络中设计的 BLOCK1 结构相同,“CBL”代表依次执行卷积、批归一化和 LeakyReLU 激活操作,其中卷积操作的核大小为 3×3 ,滑动步长为 1,“C”代表仅采用卷积操作。当输入分辨率为 $204 \text{ pixel} \times 204 \text{ pixel}$ 时,neck 的输出矩阵维度分别为 $17 \times 17 \times 14$ 和 $34 \times 34 \times 14$,其中,17 和 34 是检测模型头部的网格刻度,其大小由输入图像分辨率决定,14 是网格长度,取决于检测的目标类别数和锚框数。

4 基于波段组合的煤矸石检测研究

4.1 组合波段的选择

实验使用的多光谱成像仪有 25 个波段,为了实现煤矸石的快速检测,从 25 个波段中选择 3 个构成一个伪 RGB 图像以用于煤矸石检测,如图 3 所示的

伪 RGB 图像。在 25 个波段中选出 3 个组合,共有 2300 种选择,将 OIF 用于组合波段选择,OIF 值最大的 10 个组合如表 2 所示。

表 2 OIF 值最大的 10 个组合

Rank	OIF	Combination band
1	11.138	[7, 12, 23]
2	11.095	[7, 9, 23]
3	11.085	[7, 9, 12]
4	11.073	[7, 12, 21]
5	11.064	[7, 9, 21]
6	11.034	[7, 21, 23]
7	11.030	[7, 12, 13]
8	11.030	[7, 13, 21]
9	11.028	[7, 13, 23]
10	11.005	[7, 12, 24]

表 2 所示的最大的 10 个 OIF 值是基于所有光谱图像的相关性系数的平均值和标准差的平均值计算得到的。从表 2 中可知,波段 7、12 和 23 的 OIF 值最大,为 11.138,因此煤矸石检测的组合波段的选择为[7, 12, 23]。

4.2 检测模型的训练

本文设计的 YOLO mg 的主干网络与 YOLO 或其改进有较大差异,因此,训练时不使用预训练权重。首先冻结主干网络,使用小学习率训练 20 个 Epoch,接着使用相对大的学习率训练 40 个 Epoch,最后再调整学习率训练 40 个 Epoch。在模型训练的整个过程中,监测验证集的损失,以实时调整学习率,使模型得以稳定的训练。优化函数始终为 Adam,学习率衰减因子设为 0.5,验证集的损失出现增大趋势就降低学习率。三个训练过程的初始学习率分别设置为 0.00025、0.0005 和 0.0001,训练批大小分别设为 4、2 和 2。模型训练过程的曲线如图 5 所示,其中 L 为损失, L_e 为学习率。

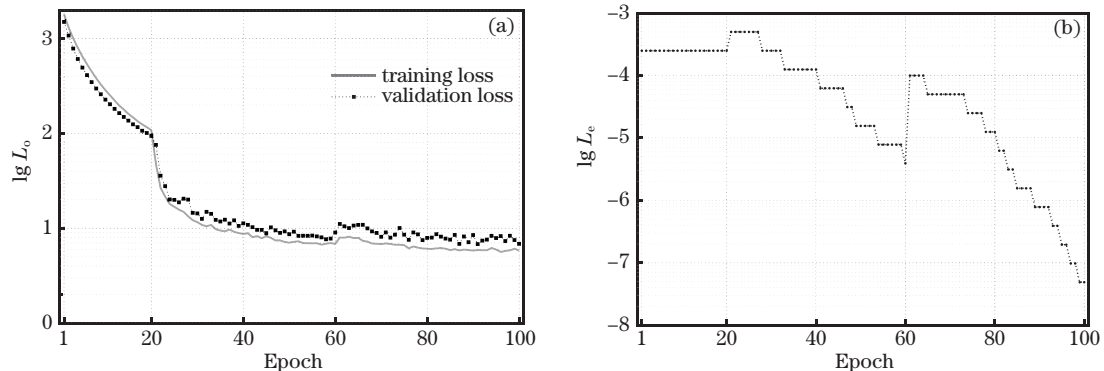


图 5 训练过程曲线。(a) 损失; (b) 学习率

Fig. 5 Curves for training process. (a) Loss; (b) learning rate

图 5(a)是训练过程的损失曲线,图 5(b)是学习率变化曲线。在前 20 个训练 Epoch 中,模型的损失保持稳定的下降,学习率没发生衰减。第一段训练结束时,模型的训练集和验证集的损失从 1805.40 和 1500.60 分别下降至 107.92 和 94.22。在第二段训练中,验证集的损失下降到 20 左右开始发生波动,此时学习率发生衰减。大约经历 50 个 Epoch 训练后,验证集的损失下降减缓,稳定在 8 左右,此时的学习率已降到 0.0000039。为了防止模型陷入局部最优,调整学习率为 0.0001,再训练 40 个 Epoch。第三段训练开始时,模型的损失稍微增大,之后下降,模型的验证集的损失最小为 6.81。

由图 5 可知,本文设计的煤矸石检测模型能被稳定训练,且没有发生过学习现象。验证集的损失

从初始的 1500.60 下降到 6.81 的最小值,意味着其学习能力较强。

4.3 煤矸石识别和定位研究

在训练过程中监测验证集的损失,并根据验证集的损失值保存模型的权值。在实验中,从第 40 个 Epoch 开始,将每次验证集损失降低时的权值保存,共保存 13 个,分别是第 42、43、44、45、50、52、56、57、58、76、79、87 和 91 个 Epoch 保存的模型。将测试集数据用于全部保存的模型测试,以第 91 个训练 Epoch 保存的模型为例,煤和煤矸石的召回率-精度的曲线如图 6 所示。

图 6 所示召回率-精度曲线是利用非极大值抑制过滤冗余预测边界框后的结果,其中置信度得分小于 0.25 的边界框被过滤,实线与横轴的面积是煤或煤矸石的 AP。不同迭代次数下的测试结果如表 3 所示。

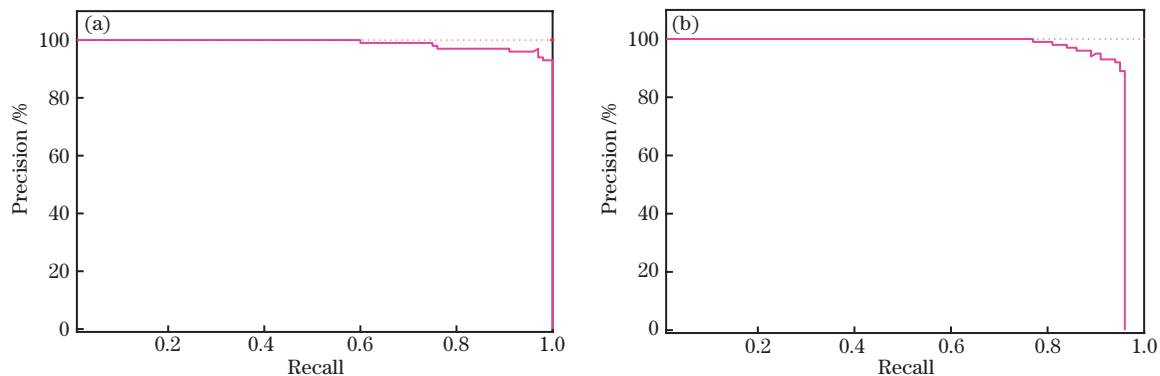


图 6 召回率-精度曲线。(a) 煤; (b) 煤矸石

Fig. 6 Precision versus recall. (a) Coal; (b) coal gangue

表 3 模型的检测结果

Table 3 Detection results of model

Epoch	Validation loss	Original detection result				Result after non-maximum suppression			
		A / %	T / Total	F	mAP / %	A / %	T / Total	F	mAP / %
42	10.27	91.80	108/116	17	92.73	78.70	92/116	3	83.46
43	9.62	93.98	110/116	19	94.37	89.08	104/116	3	89.24
44	9.61	90.35	107/116	18	93.95	85.71	101/116	5	90.98
45	8.88	92.93	111/116	46	94.40	88.84	105/116	12	88.57
50	8.65	93.56	109/116	12	94.71	87.80	102/116	2	91.84
52	8.32	95.13	111/116	18	96.70	90.39	105/116	2	93.94
56	8.21	94.98	111/116	23	96.84	88.70	103/116	1	93.34
57	8.02	95.01	111/116	21	96.92	88.67	103/116	1	93.75
58	7.65	95.00	111/116	25	96.99	89.48	104/116	3	93.85
76	7.55	96.63	113/116	24	97.81	91.97	107/116	6	93.90
79	7.50	94.97	111/116	24	97.06	91.91	107/116	4	94.76
87	6.84	94.84	111/116	16	96.84	90.33	105/116	3	94.58
91	6.81	94.94	111/116	18	96.96	90.35	105/116	4	94.26

在表 3 中,“Original detection result”是检测结果没有经过 F 抑制计算得的结果,“Result after non-maximum suppression”是检测结果通过非极大值抑制,且置信度阈值设为 0.25, IoU 的值设置为 0.4。本文研究的是基于多光谱成像的煤矸石快速

分离,因此更关注煤矸石的检测精度,所以表 3 中的 A 和 T 是煤矸石的检测结果。

表 3 中第 42~58 个 Epoch 模型是第二段训练保存而得的,第 76、79、87 和 91 个 Epoch 模型是第三段训练保存而得的。以原始检测结果为例,在第

二段训练中,随着训练次数的增加,煤和煤矸石的检测 mAP 呈现增长趋势;此外,在 50 个训练 Epoch 前,保存的模型的检测结果波动相对较大,在 50 个训练 Epoch 后,模型的检测结果相对稳定。在第三段训练中,第 76 个训练 Epoch 所保存的模型的 mAP 最大,为 97.81%,且随着训练的进行,mAP 有降低趋势。对比第二段和第三段保存的模型的测试结果,可以发现,尽管在第三段训练中,模型的损失减少并不显著,但第三段训练保存模型测试的结果优于第二段训练保存模型的测试结果,尤其是对

煤矸石的检测结果。

采用非极大值过滤冗余预测边界框,并通过置信度阈值过滤得分相对低的预测边界框。经过滤后,煤矸石最大检测精度为 91.97%,煤和煤矸石的 mAP 为 93.90%,此时对应的 F 仅有 6 个。对比抑制前后的 F ,尽管煤矸石的 A 有所下降,但是 F 的减少相对更多。例如对于第 76 个 Epoch 保存的模型,抑制 F 后,其 F 减少 18 个,而 T 仅减少 4 个。 F 抑制后,第 79 个 Epoch 保存的模型的 mAP 最大,检测结果对比如图 7 所示。

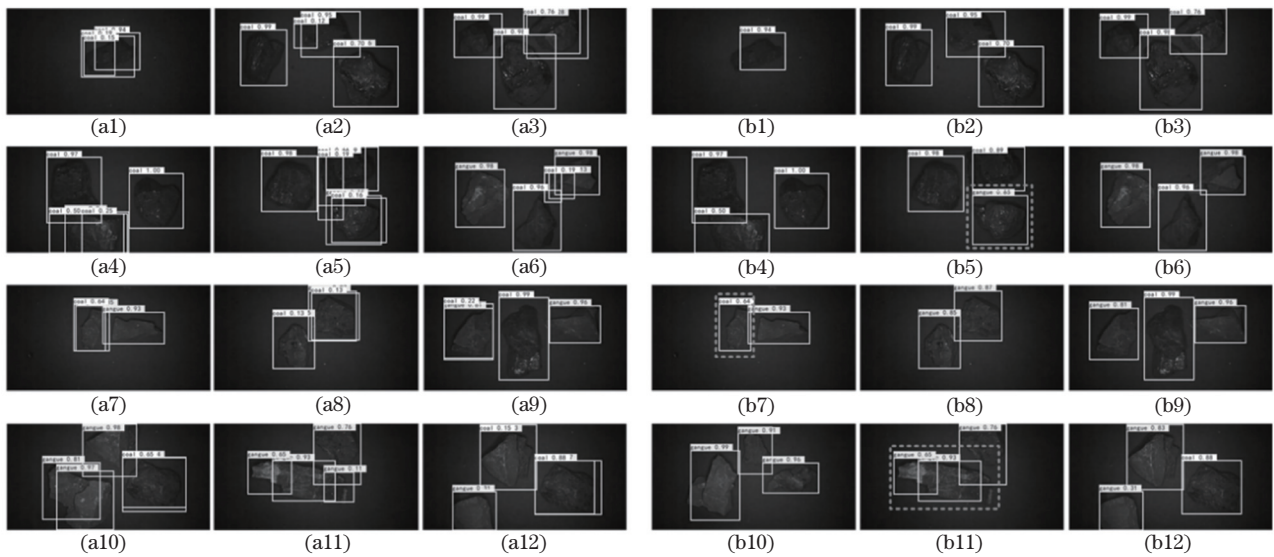


图 7 煤和煤矸石的识别和定位结果。(a1)~(a12) 冗余边界框没有被抑制;(b1)~(b12) 冗余边界框被抑制

Fig.7 Recognition and location results of coal and coal gangue. (a1)~(a12) Redundant bounding boxes without suppression; (b1)~(b12) redundant bounding boxes with suppression

图 7(a1)~(a12)是从测试集中随机选择的同一目标被预测出多个边界框的光谱图像,图 7(b1)~(b12)是图 7(a1)~(a12)过滤冗余预测边界框后的光谱图像,其中虚线圈出的是被错误检测的目标。从图

7(a1)~(a12)中可以发现,在过滤冗余边界框前,同一目标可能会被预测出许多边界框,即被重复定位,过滤冗余边界框后,目标的重定位被有效减少。图 7(b1)~(b12)中煤和煤矸石的坐标信息如表 4 所示。

表 4 煤和煤矸石的坐标信息

Table 4 Coordinate information of coal or coal gangue

Number	Coordinate and score for bounding box 1						Coordinate and score for bounding box 2						Coordinate and score for bounding box 3					
	L	S	x_1	y_1	x_2	y_2	L	S	x_1	y_1	x_2	y_2	L	S	x_1	y_1	x_2	y_2
1	0	0.94	49	176	125	267	-	-	-	-	-	-	-	-	-	-	-	-
2	0	0.99	43	52	157	145	0	0.95	6	173	99	292	0	0.7	77	238	200	368
3	0	0.99	11	61	101	158	0	0.98	54	141	206	266	0	0.76	0	201	93	315
4	0	1.00	54	245	166	354	0	0.97	21	80	155	190	0	0.5	136	85	216	235
5	1	0.85	99	224	199	335	0	0.98	19	95	132	207	0	0.89	0	224	90	329
6	1	0.98	46	63	164	162	1	0.98	19	262	98	352	0	0.96	88	178	211	275
7	1	0.93	56	192	121	316	0	0.64	40	134	135	198	-	-	-	-	-	-
8	1	0.87	13	189	115	285	1	0.85	64	118	171	202	-	-	-	-	-	-
9	1	0.96	43	251	119	354	1	0.81	48	40	153	139	0	0.99	26	150	194	250
10	1	0.99	54	77	196	176	1	0.96	79	221	141	334	1	0.91	6	172	102	241
11	1	0.93	75	117	157	243	1	0.76	0	199	124	294	1	0.65	69	68	143	156
12	1	0.83	1	115	134	229	1	0.31	137	60	216	148	0	0.88	74	225	183	345

在表 4 中,“L”代表标签,“0”代表煤,“1”代表煤矸石,“S”代表得分。 (x_1, y_1) 、 (x_2, y_2) 分别是边界框的左上和右下坐标。此外,表 4 中的斜体加粗是错误定位的样本。

4.4 与经典检测模型的对比研究

YOLO mg 是基于 YOLO v4 架构设计的,因此将 YOLO v4 也用于煤和煤矸石的检测,此外还

将经典的检测模型 YOLO v3 以及 M2Det 用于煤和煤矸石的识别和定位,最后对比结果。训练经典模型时,加载预训练权值,因此共训练 40 个 Epoch。YOLO v4 的输入分辨率为 416 pixel \times 416 pixel、512 pixel \times 512 pixel 和 608 pixel \times 608 pixel, M2Det 和 YOLO v3 的输入分辨率为 320 pixel \times 320 pixel。不同模型的检测结果如表 5 所示。

表 5 各模型的测试结果

Table 5 Detection results of each model

Model	Coal			Coal Gangue			mAP /%	Detection time /s
	AP /%	TP/Total	FP	AP /%	TP/Total	FP		
YOLO mg-204	97.61	113/115	7	91.91	107/116	4	94.76	1.255
YOLO mg-300	98.27	114/115	7	92.62	108/116	3	95.45	1.642
YOLO v4-416	98.43	115/115	10	89.94	105/116	4	94.19	4.386
YOLO v4-512	97.39	113/115	7	91.37	107/116	7	94.38	5.533
YOLO v4-608	96.36	111/115	2	95.51	112/116	8	95.94	7.871
YOLO v3-320	91.48	108/115	9	87.63	106/116	6	89.56	2.832
M2Det 320	98.97	114/115	8	90.80	106/116	5	94.89	4.212

YOLO mg-204 和 YOLO mg-300 代表本文设计的检测模型,且输入分辨率分别为 204 pixel \times 204 pixel 和 300 pixel \times 300 pixel,其他模型同理。检测时间是 115 张光谱图像的检测总耗时,并且测试 5 次,结果取平均值。

就检测精度而言,YOLO v3 对煤矸石的检测精度为 87.63%,最小。输入分辨率为 608 pixel \times 608 pixel 的 YOLO v4 对煤矸石的检测精度最大,但是检测速度相对最慢,测试集共用时 7.781 s。无论是采用 204 pixel \times 204 pixel 输入,还是采用 300 pixel \times 300 pixel 输入,YOLO mg 对煤矸石的检测精度优于 v3-320、v4-416 和 M2Det 320。此外,YOLO mg 的 115 张图像检测时间最少,仅为 1.255 s。

5 结 论

将多光谱成像技术用于煤矸石的智能分离研究,通过 OIF 选取最佳组合波段,并设计了一个检测模型以实现煤矸石的快速识别和定位。首先,将多光谱成像技术用于煤矸石的识别和定位,利用 OIF 理论,从 25 个波段中选择 3 个最佳组合波段用于煤矸石的检测研究。其次,设计一个快速检测模型,在不使用预训练权值情况下,通过设置三种学习率实现模型的快速、稳定训练,在 204 pixel \times 204 pixel 输入分辨率下,通过非极大值抑制和置信度阈值过滤冗余预测边界框,实现煤矸石 91.91% 的平均检测精度。最后,把经典的目标检测模型用于煤矸石检测研究。结果显示,利用所设计的

YOLO mg 检测 115 张光谱图像的时间仅为 1.255 s,在所对比模型中,耗时最少,达到设计目的。

参 考 文 献

- [1] BP. Book review: BP statistical review of world energy [J]. Energy Exploration & Exploitation, 1989, 7(6): 468.
- [2] Li J Y, Wang J M. Comprehensive utilization and environmental risks of coal gangue: a review [J]. Journal of Cleaner Production, 2019, 239: 117946.
- [3] Gui X H, Liu J T, Cao Y J, et al. Coal preparation technology: status and development in China [J]. Energy & Environment, 2015, 26(6/7): 997-1013.
- [4] Chelgani S C, Parian M, Parapari P S, et al. A comparative study on the effects of dry and wet grinding on mineral flotation separation-a review [J]. Journal of Materials Research and Technology, 2019, 8(5): 5004-5011.
- [5] Wang G F, Wang H, Ren H W, et al. 2025 scenarios and development path of intelligent coal mine [J]. Journal of China Coal Society, 2018, 43(2): 295-305.
王国法,王虹,任怀伟,等.智慧煤矿 2025 情景目标和发展路径 [J]. 煤炭学报, 2018, 43(2): 295-305.
- [6] Shang D Y, Wang Y W, Yang Z Y, et al. Study on comprehensive calibration and image sieving for coal-gangue separation parallel robot [J]. Applied Sciences, 2020, 10(20): 7059.
- [7] Zheng K H, Du C L, Li J P, et al. Underground pneumatic separation of coal and gangue with large

- size (≥ 50 mm) in green mining based on the machine vision system [J]. Powder Technology, 2015, 278: 223-233.
- [8] Liang X G, Li Y F, Li Y, et al. The intelligent dry cleaning technology: study, application and developing trend [J]. Coal Preparation Technology, 2019(1): 92-96, 102.
梁兴国, 李云峰, 李燕, 等. 智能干选技术研究应用及发展趋势[J]. 选煤技术, 2019(1): 92-96, 102.
- [9] Guo Y C, Yu Z S, Lu Y C. Research on photoelectric intelligent separation technology of coal and gangue based on NP-FSVM with the PSO algorithm [J]. Coal Science and Technology, 2019, 47(4): 13-19.
郭永存, 于中山, 卢熠昌. 基于 PSO 优化 NP-FSVM 的煤矸光电智能分选技术研究[J]. 煤炭科学技术, 2019, 47(4): 13-19.
- [10] Zhao Y D, He X M. Recognition of coal and gangue based on X-ray [J]. Applied Mechanics and Materials, 2013, 275/276/277: 2350-2353.
- [11] Cao X G, Li Y, Wang P, et al. Research status of coal-gangue identification method and its prospect [J]. Industry and Mine Automation, 2020, 46(1): 38-43.
曹现刚, 李莹, 王鹏, 等. 煤矸石识别方法研究现状与展望[J]. 工矿自动化, 2020, 46(1): 38-43.
- [12] Rao Z Y, Wu J T, Li M. Coal-gangue image classification method [J]. Industry and Mine Automation, 2020, 46(3): 69-73.
饶中钰, 吴景涛, 李明. 煤矸石图像分类方法[J]. 工矿自动化, 2020, 46(3): 69-73.
- [13] Dou D Y, Wu W Z, Yang J G, et al. Classification of coal and gangue under multiple surface conditions via machine vision and relief-SVM [J]. Powder Technology, 2019, 356: 1024-1028.
- [14] Lai W H, Zhou M R, Hu F, et al. A study of multispectral technology and two-dimension autoencoder for coal and gangue recognition [J]. IEEE Access, 2020, 8: 61834-61843.
- [15] Andreopoulos A, Tsotsos J K. 50 years of object recognition: directions forward[J]. Computer Vision and Image Understanding, 2013, 117(8): 827-891.
- [16] Sun J, Guo D B, Yang T T, et al. Real-time object detection based on improved YOLOv3 network [J]. Laser & Optoelectronics Progress, 2020, 57(22): 221505.
孙佳, 郭大波, 杨甜甜, 等. 基于改进的 YOLOv3 网络的实时目标检测[J]. 激光与光电子学进展, 2020, 57(22): 221505.
- [17] Li C Y, Yao J M, Lin Z X, et al. Object detection method based on improved YOLO lightweight network [J]. Laser & Optoelectronics Progress, 2020, 57(14): 141003.
李成跃, 姚剑敏, 林志贤, 等. 基于改进 YOLO 轻量化网络的目标检测方法[J]. 激光与光电子学进展, 2020, 57(14): 141003.
- [18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [19] Deepa R, Tamilselvan E, Abrar E S, et al. Comparison of Yolo, SSD, faster RCNN for real time tennis ball tracking for action decision networks[C]//2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), April 4-6, 2019, Sathyamangalam, India. New York: IEEE Press, 2019: 1-4.
- [20] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [21] Zhang X X, Zhu X. Moving vehicle detection in aerial infrared image sequences via fast image registration and improved YOLOv3 network [J]. International Journal of Remote Sensing, 2020, 41(11): 4312-4335.
- [22] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-05-23)[2020-12-15]. <https://arxiv.org/abs/2004.10934>.
- [23] Zhao Q J, Sheng T, Wang Y T, et al. M2Det: a single-shot object detector based on multi-level feature pyramid network [EB/OL]. (2018-11-12)[2020-12-15]. <https://arxiv.org/abs/1811.04533v3>.
- [24] Tian Y N, Yang G D, Wang Z, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417-426.
- [25] Li D J, Zhang Z X, Xu Z H, et al. An image-based hierarchical deep learning framework for coal and gangue detection [J]. IEEE Access, 2019, 7: 184686-184699.
- [26] Lü Z, Wang W D, Xu Z Q, et al. Cascade network for detection of coal and gangue in the production context [J]. Powder Technology, 2021, 377: 361-371.
- [27] Lai W H, Zhou M R, Hu F, et al. Coal gangue detection based on multi-spectral imaging and improved YOLOv4[J]. Acta Optica Sinica, 2020, 40(24): 2411001.
来文豪, 周孟然, 胡锋, 等. 基于多光谱成像和改进 YOLOv4 的煤矸石检测[J]. 光学学报, 2020, 40

- (24): 2411001.
- [28] Guo T, Hua W S, Liu X, et al. Rapid hyperspectral band selection approach based on clustering and optimal index algorithm [J]. *Optical Technique*, 2016, 42(6): 496-500.
郭彤, 华文深, 刘恂, 等. 基于聚类 and 最佳指数的快速高光谱波段选择方法 [J]. *光学技术*, 2016, 42(6): 496-500.
- [29] Acharya T, Yang I, Lee D. Land cover classification of imagery from landsat operational land imager based on optimum index factor [J]. *Sensors & Materials*, 2018, 30(8): 1753-1764.
- [30] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [31] Neubeck A, van Gool L. Efficient non-maximum suppression [C] // 18th International Conference on Pattern Recognition (ICPR'06), August 20-24, 2006, Hong Kong, China. New York: IEEE Press, 2006: 850-855.
- [32] Hasanpour S, Rouhani M, Fayyaz M, et al. Towards principled design of deep convolutional networks: introducing SimpNet [EB/OL]. (2018-02-17) [2020-12-15]. <https://arxiv.org/abs/1802.06205>.

Fast Location of Coal Gangue Based on Multispectral Band Selection

Lai Wenhao, Zhou Mengran*, Wang Jinguo, Hu Tianyu, Kong Xixi,
Hu Feng, Bian Kai, Zhu Ziwei

*School of Electrical and Information Engineering, Anhui University of Science and Technology,
Huainan, Anhui 232000, China*

Abstract

Objective As one of the most important primary energy resources, coal needs to be mined in large quantities every year. In this process, the mass proportion of biological coal gangue in raw coal is more than 10%. It needs to be separated from raw coal, which is called coal washing. As an important traditional energy field in China, the intelligent construction of coal industry is directly related to the process of the economic and social intelligence. Therefore, it is of great significance for the intelligent construction of coal mines to develop modern coal gangue separation technologies. Coal gangue is identified by machine learning algorithm, and then manipulators or pneumatic devices are used for the separation of coal gangue from coal, which is called intelligent separation of coal gangue. At present, coal gangue recognition based on an imaging technology has attracted more and more attention, but the recognition can only realize whether there is coal gangue, but cannot locate the coal gangue. In addition, the locating speed of coal gangue is directly related to the production efficiency of coal mines. This paper designs a coal gangue detection model based on the framework of YOLO v4, which aims to realize the faster identification and location of coal gangue, and provides the position information of coal gangue more quickly for the intelligent sorting system. In order to reduce the interference of environmental visible light, coal gangue is identified and located based on the multispectral imaging technology. In order to reduce redundant information and further improve detection speed, three pseudo RGB images are selected from 25 bands of multispectral images by using the optimal index factor theory.

Methods Firstly, the collected multispectral data of coal and gangue in the laboratory and the selected three bands from 25 bands of multispectral images were used to form the pseudo RGB images by using the best index factor theory. The training data, validation data, and test data were 460, 65, and 115 spectral images, respectively. Second, a lightweight coal gangue detection model was designed based on multispectral imaging, which is named YOLO mg. The maximum depth of its backbone network is 23 convolution layers. In order to achieve stable and fast training of YOLO mg without using the pre-training weights, three learning rates, which are 0.00025, 0.0005 and 0.0001, are used, respectively. Finally, the model YOLO mg was applied to the study of coal gangue detection based on the pseudo RGB images, and the redundant bounding box was filtered by the non-maximum suppression combined with confidence threshold.

Results and Discussions The rapid identification and location of coal gangue based on multispectral imaging are

studied. The maximum depth of the backbone network of YOLO mg is 23 layers (Fig. 4). OIF is used for multispectral band selection, the maximum value is 11.138, and the corresponding combination band selection is [7, 12, 23] (Table 2). The combination band is used for the designed YOLO mg training. Three learning rates are set and trained for 100 epochs. The convergence speed of the designed model is fast, and the verification set loss decreases from the initial 1500.60 to the minimum of 6.81. In addition, compared with training loss and verification loss, YOLO mg is not over learning (Fig. 5 and Table 3). The redundant prediction boundary box is filtered by the non-maximum suppression combined with confidence threshold. Under the input resolution of 204 pixel \times 204 pixel, the average accuracy of coal gangue detection is 91.91%, and there are only four false positive (FP), which effectively reduces the wrong positioning of coal gangue in the prediction results (Table 4 and Fig. 7). Compared with the advanced YOLO v4, M2Det, and YOLO v3, the YOLO mg designed here maintains an excellent detection accuracy, and the positioning speed of coal gangue is the fastest. The 115 combined band spectral images only take 1.255 s (Table 5).

Conclusions The intelligent coal washing technology is an important part of the realization of coal mine modernization construction. The multispectral imaging technology is applied to the study of intelligent separation of coal gangue. The optimal combination band is selected by OIF, and a detection model is designed to realize the location of coal gangue. Firstly, the multispectral imaging technology is used to identify and locate coal gangue. Three optimal combination bands are selected from 25 bands for coal gangue detection research by the OIF theory, which can effectively reduce the redundant multispectral information. Second, a fast detection model is designed. Without using the pre-training weights, the model can be trained quickly and stably by setting three learning rates. Finally, the detection time of 115 spectral images by YOLO mg designed in this paper is only 1.255 s, which is the least time-consuming in the comparison models, so that the rapid detection of coal gangue is realized and the design purpose is achieved. The lightweight detection model of YOLO mg designed here aims at the specific scientific problems of coal gangue localization based on multispectral imaging and provides a guidance for the lightweight model design of positioning problems in other fields.

Key words spectroscopy; multispectral image; optimal index factor; band selection; YOLO mg; identification and location

OCIS codes 300.6170; 110.4234; 100.4996; 150.1135