

DOI:10.19651/j.cnki.emt.2209535

# 轻量级注意力 X 射线矿石检测方法\*

杨文龙<sup>1,2,3</sup> 郭明钰<sup>1</sup>

(1. 江西理工大学电气工程与自动化学院 赣州 341000; 2. 江钨集团赣州有色冶金研究所有限公司 赣州 341000;  
3. 赣州有色冶金研究所有色金属矿冶装备工业设计中心 赣州 341000)

**摘要:** 针对缺乏矿石数据集和矿石分类识别模型等因素,自建以 X 射线照射成像的矿石图像为数据集,并以 MobileNet V2 为主网络,提出基于改进 MobileNet V2 轻量级矿石分类模型算法。首先,通过调整扩展因子和宽度因子大幅减少模型参数量,实现模型轻量化的目的;其次,通过在部分倒残差模块和原模型分类器中嵌入高效通道注意力机制,并将剩余倒残差模块替换为含深度空洞卷积的并行特征提取网络,以增强模型特征信息提取能力,提升模型识别准确率;最后,使用迁移学习的训练方式初始化权重,加速模型训练。经过改进,该算法矿石识别准确率提升至 96.720%,对比 VGG16、GoogleNet、Xception、ShuffleNet 和 MobileNet V2 在准确率和矿石检测速度都获得了提升。综合而言,相比本文实验中其他算法而言,改进算法针对矿石的识别性能具有更佳表现。

**关键词:** 深度学习;X 射线矿石图像分类;MobileNet V2;有效通道注意力机制;并行特征提取网络;迁移学习

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.20

## Lightweight attention parallel X-ray ore detection algorithm

Yang Wenlong<sup>1,2,3</sup> Guo Mingyu<sup>1</sup>

(1. School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China;  
2. Ganzhou Nonferrous Metallurgy Research Institute, Ganzhou 341000, China; 3. Nonferrous Metal Metallurgical  
Research Institute of Ganzhou Mining and Metallurgy Equipment Work Design Center, Ganzhou 341000, China)

**Abstract:** In view of the lack of ore data set and ore classification and recognition model, a lightweight ore classification model algorithm based on improved mobilenet V2 is proposed, which takes the ore image imaged by X-ray irradiation as the data set and mobilenet V2 as the main network. Firstly, by adjusting the expansion factor and width factor, the amount of model parameters is greatly reduced to realize the purpose of model lightweight; Secondly, the efficient channel attention mechanism is embedded in some inverse residual modules and the original model classifier, and the residual inverse residual module is replaced by a parallel feature extraction network with deep hole convolution, so as to enhance the ability of model feature information extraction and improve the accuracy of model recognition; Finally, the training method of transfer learning is used to initialize the weight and accelerate the model training. After improvement, the ore recognition accuracy of the algorithm is improved to 96.720%. Compared with vgg16, googlenet, xception, shufflenet and mobilenet V2, the accuracy and ore detection speed have been improved. In general, compared with other algorithms in this experiment, the improved algorithm has better performance for ore recognition.

**Keywords:** deep learning; X-ray ore image classification; MobileNet V2; efficient channel attention; parallel feature extraction network; transfer learning

## 0 引言

随着长时间的矿产资源开采,各种矿石资源日益贫化,矿石品味越来越低,提升矿石资源的有效利用率极为迫切。

矿石的分选也由人工手选发展到如今的重选、图像选矿和 X 射线选矿<sup>[1]</sup>。

近年来,随着计算机视觉类技术的应用越来越广泛,不乏有大量研究人员将视觉技术应用至矿山中,提升矿石分

收稿日期:2022-04-07

\* 基金项目:江西省普通高校校级研究生创新专项资金项目(XY2021-S113)资助

选效率。Deo 等<sup>[2]</sup>提出了一种基于卷积神经网络的图像处理技术来检测材料堆积,以达到回收物料的目的。Firla 等<sup>[3]</sup>利用图像处理算法以直方图的形式展现矿石粒度分布大小。Lian 等<sup>[4]</sup>利用图像处理技术在线采集矿石粒度参数。邓田等<sup>[5]</sup>使用改进的 PSO-Faster R-CNN 算法对矿石进行识别,速度和精度都得到一定的提升,但算法在处理高维度复杂特征的能力欠缺,且比较依赖硬件资源。文大鹏等<sup>[6]</sup>利用激光诱导击穿光谱技术结合 PCA-PSO-SVM 算法对 12 类矿石进行识别分类,并且获得可观的准确率,但算法并未对实际的 X 射线矿石图像进行分类识别,泛化性不足。高瑜璋等<sup>[7]</sup>使用机器视觉技术对提取的图像和声音特征进行融合和分类,模型在复杂的环境下也能有效的检测出矿用输送带的撕裂状态。

在实际应用中,许多检测模型只能部署在高算力的设备中运行,为使网络模型可以部署在一些小型算力的边缘设备上,深度学习模型的发展向更轻量化、更精准化持续进步。由 2014 年的 GoogleNet<sup>[8]</sup>,2015 年的 VGG<sup>[9]</sup>、ResNet<sup>[10]</sup>等主流大型网络向轻量化方向优化网络,2016 年提出的 SqueezeNet<sup>[11]</sup>沿用 AlexNet<sup>[12]</sup>网络结构,依靠将  $3 \times 3$  的卷积替换为  $1 \times 1$  的卷积进行压缩,使其参数量降低至等性能的 AlexNet 的 2.14%。自此 ShuffleNet<sup>[13]</sup>、Xception<sup>[14]</sup>、MobileNet<sup>[15]</sup>、DenseNet<sup>[16]</sup>等更多性能优越的轻量级网络被提出。轻量级网络的模型小、速度快等特性使其可以部署在经济型边缘设备上(如 Jetson nano)中,如 Zhao 等<sup>[17]</sup>将轻量级网络模型部署在两个 NVIDIA 的 Jetson nano 板上,推理速度达到 35 FPS,实现大豆全表面

的实时识别。Knapheide 等<sup>[18]</sup>为部署于 FPGA 上的轻量级神经网络设计硬件加速器,实现了每帧 32 mJ 的能效。

轻量化的实现使得检测模型对硬件资源的需求降低,节约了许多计算成本。本文将机器视觉技术应用在矿石检测中,致力于开发一种轻量化的矿石分类网络模型,以 Mobilenet V2<sup>[19]</sup>为主网络,对其做出以下几点改进:1)为使模型更轻量化,通过适应性调整网络的宽度因子  $\alpha$  和扩展因子  $t$  来获取模型参数数量的降低;2)在模型提取图像特征的过程中,嵌入高效相关性通道注意力模块,加强有用特征的提取和无用特征的抑制;3)基于深度可分离空洞卷积,设计并行多尺度特征提取网络,增加模型的宽度和感受野;4)针对模型的训练,本文采用多重迁移学习训练网络,初始化网络权重,加快网络的训练速度。改进后,本文算法对比其他主流算法将具有更优的表现,并在保证准确率和检测速率的前提下模型更轻量化,可以部署至一些算力不足的边缘设备中进行检测,为今后数字化矿山建设节约成本。

### 1 MobileNet V2 网络改进

以设计高精度、高效率 and 低内存消耗的黑钨矿石分选算法为目的,建立以 Mobilenet V2 网络为主干,通过调整扩展因子  $t$  来降低模型参数量,在倒残差模块中加入高效一维卷积相关性通道模块和并行多尺度特征提取模块进行改进,提升网络对矿石的适应性识别能力。模型训练时使用迁移训练初始化权重。改进后模型总体结构如图 1 所示,其中步长为 1 时使用 Bottleneck1 模块,步长为 2 时使用 Bottleneck2 模块。

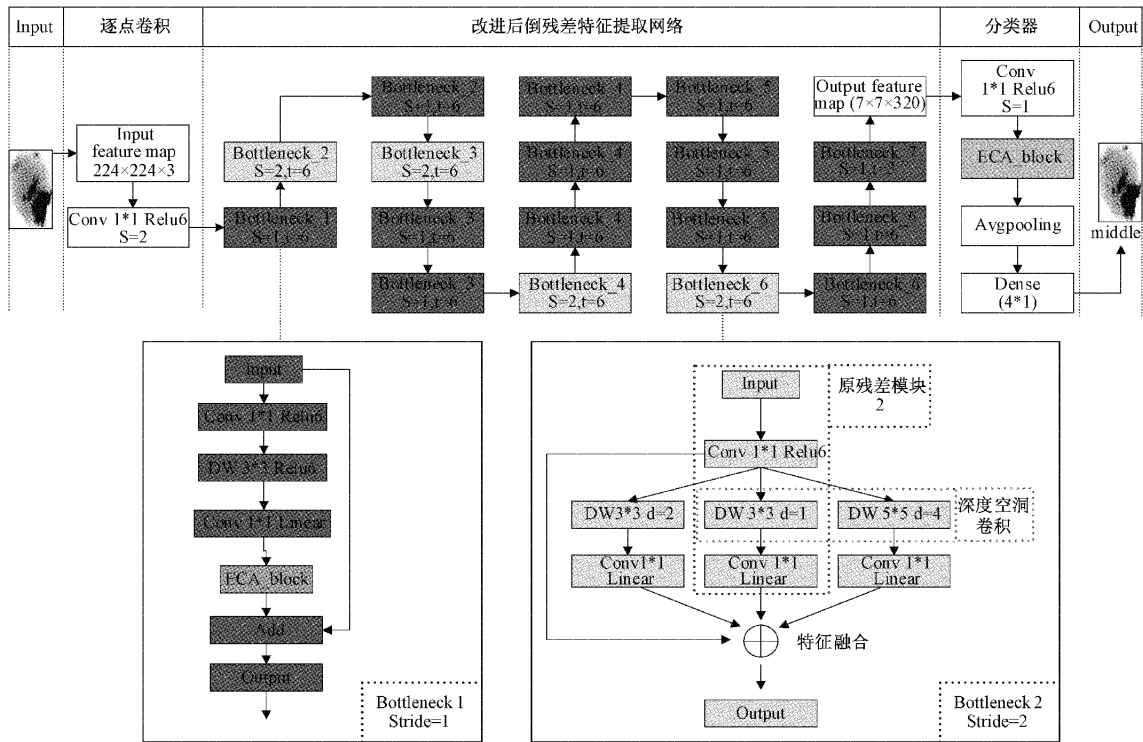


图 1 改进后模型总体结构

1.1 MobileNet V2 简介

Mobilenet V2 网络是一种轻量级的神经网络,它在 Mobilenet V1 的基础上进行了改进,其主要的优点是:1)沿用 V1 的优点,为降低模型计算量和参数量,使用深度可分离卷积代替普通卷积;2)模型主要结构由反向残差瓶颈层组成,既加深了网络结构深度,也增强了特征的表达能力;3)由于 relu6 非线性激活函数在低精度计算中更具鲁棒性,所以用其代替普通的 relu 激活函数。并在逐点卷积中添加线性瓶颈层,减少低维特征信息的丢失。

反向残差瓶颈层(bottleneck residual block)的主要任务是对输入图像进行特征提取。首先逐点卷积(pointwise convolution, PW)对输入的图像进行升维操作,然后输入至卷积核为  $3 \times 3$  的深度卷积(depthwise convolution, DW)进行高维特征提取,最后经一维卷积进行降维输出。通过此改进,可以增加网络的深度,在不增加网络过多运算量和参数量的同时,更好地提取有用信息。倒残差结构如图 2 所示。

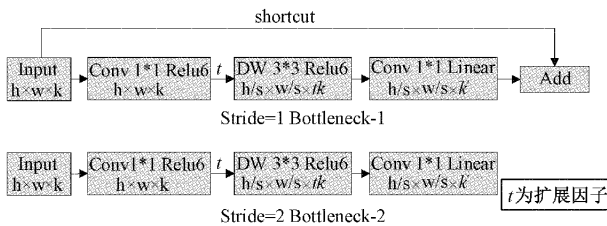


图 2 倒残差模块结构

由图 2 可以看出,只有当步长为 1 时,才有残差连接,且只有同时满足输入特征矩阵与输出特征矩阵深度相同时,残差连接才有作用,而此时倒残差模块的输出便等于输入加上主分支上的输出。

倒残差结构由深度可分离卷积组成,而深度可分离卷积由深度卷积单元和逐点卷积单元组成。逐点卷积就是卷积核为  $1 \times 1$  的普通卷积,其作用便是使用卷积核进行通道间的像素融合。而深度卷积与普通卷积不同,负责卷积运算。在标准的卷积运算中,如果输入为  $h_i \times w_i \times d_i$ ,卷积核  $K \in R^{k^2 \times d_i \times d_j}$ ,卷积后输出为  $h_j \times w_j \times d_j$ ,则消耗的计算量为:

$$FLOPs = h_i \times w_i \times d_i \times k^2 \times d_j \quad (1)$$

参数量为:

$$P = k^2 \times d_i \times d_j \quad (2)$$

如果进行深度可分离卷积运算,则计算量为:

$$FLOPs_{DW} = h_i \times w_i \times d_i (k^2 + d_j) \quad (3)$$

参数量为:

$$P_{DW} = k^2 \times d_i \times \alpha + d_i \times d_j \times \alpha^2 \quad (4)$$

其中,  $h_i, w_i$  为输入特征图的高度和宽度,  $k$  为卷积核大小,  $\alpha$  表示超参数宽度因子(width multiplier),默认为 1。

由上可知深度可分离卷积和普通卷积计算量的比值为:

$$K_{FLOPs} = \frac{FLOPs_{DW}}{FLOPs} = \frac{1}{d_i} + \frac{1}{k^2} \quad (5)$$

参数量的比值为:

$$K_P = \frac{P_{DW}}{P} = \frac{1}{d_j} + \frac{1}{k^2} \quad (6)$$

其中,  $d_i$  为输入通道数,  $d_j$  为输出通道数。可以看出,当标准卷积的卷积核与深度可分离卷积同为 3 时,使用深度可分离卷积可以使参数量和运算量减少约 8/9,而只损失少量的精度。深度可分离卷积过程如图 3 所示。

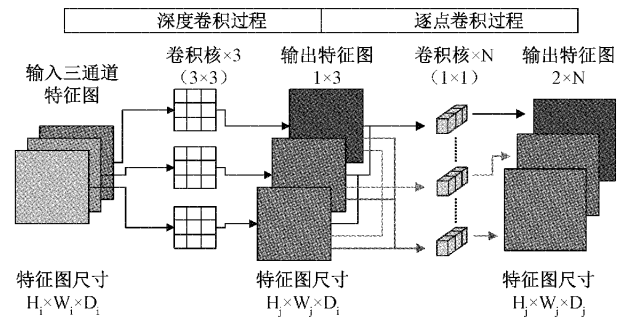


图 3 深度可分离卷积过程

1.2 超参数调整

本文研究中,希望通过调整网络参数对其进行一定的降参,在 Mobilenet V2 结构体系中,可以使用输入图像分辨率和宽度因子(alpha)作为可调超参数,当输入图像分辨率在 96~224 之间,宽度因子可在 0.35、0.5、0.75、1.0 间可调,由深度可分离卷积参数量计算公式可知,当调整  $\alpha < 1$  时,模型参数量获得相应降低。在 MobileNet V1 中通过调整宽度因子得到的性能如表 1 所示。

可以看出,经过调整之后,模型参数确实获得大幅降低,但模型准确率也随之大幅下降,而模型准确率下降可以通过网络优化来获得提升。

表 1 不同宽度因子 V1 性能

模型	准确率/%	计算量/M	参数量/M
MobilenetV1(1.0)	70.6	569	4.2
MobilenetV1(0.75)	68.4	325	2.6
MobilenetV1(0.5)	63.7	149	1.3

为防止改进对特征信息提取性能的影响,本文调整 MobileNet V2 网络结构中最后一层倒残差层扩展因子  $t=3$ ,同时设置超参数  $\alpha=0.75$  以轻量化矿石分类模型。

1.3 ECA 模块

在 MobilenetV2 网络结构中,模型的特征提取任务通常是由倒残差模块来执行的,为了避免网络提取过多的无用信息,多关注于有用特征信息,本文在倒残差模块中添加高效注意力机制(efficient channel attention, ECA-net)<sup>[20]</sup>,在特征通道不进行降维的情况下,有效捕获跨通道交互的信息,增加模型对有用特征信息提取能力。

通道注意力机制模块 ECA net 是对 SE net<sup>[21]</sup>进行改进的轻量、即插即用的注意力机制模块。其特点便是引入后模块附加的参数较少,计算量可以忽略不计,同时可以带来显著的性能提升。如图 4 所示,输入通过全局平均池化(global average pooling,GAP)获得的聚合特征,ECA 模块执行卷积核大小为  $k$  的快速一维卷积来生成通道权重,其中  $k$  通过通道维度  $C$  的映射自适应地确定。

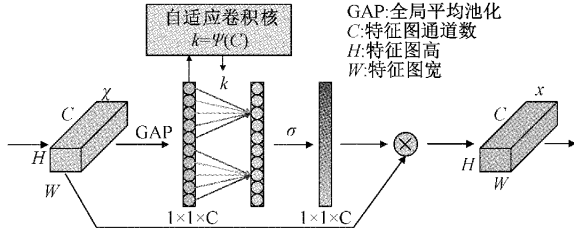


图 4 ECA 模块示意图

通道权重计算公式为:

$$\omega_i = \sigma\left(\sum_{j=1}^k \omega^j y_i^j\right), y_i^j \in \Omega_i^k \quad (7)$$

其中,  $\sigma$  表示 Sigmoid 函数,  $\Omega_i^k$  为  $y_i$  的  $k$  个相邻通道集合。 $\omega^j$  为全通道共享的权重参数。

最后的通道注意力映射公式为:

$$C = \phi(k) = 2^{\gamma \times k^b} \quad (8)$$

其中,  $C$  为通道维数,  $k$  为卷积核大小。在给定的通道维数  $C$  并将  $\gamma, b$  设置为 2 和 1 的情况下,卷积核可以根据式(9)自适应确定。可以看出,高维通道具有更长的相互作用距离,而低维通道通过非线性映射具有更短的相互作用距离。(  $|t|_{odd}$  表示  $t$  的最邻近奇数)

$$k = \phi(C) = \left\lfloor \frac{\log(C) + b}{\gamma} \right\rfloor_{odd} \quad (9)$$

具体来说,ECA 模块通过每个通道与其  $n$  个相邻通道相互作用,同时避免渠道维度缩减,以其轻量级的方式捕捉跨通道互动信息,记录有效的通道注意力。本文将 ECA 模块串联至倒残差模块(当 stride=1)的第二个逐点卷积后。具体位置如图 5 所示。

残差连接(当步长stride=1且输出特征通道数等于输入特征通道数时才有残差连接)

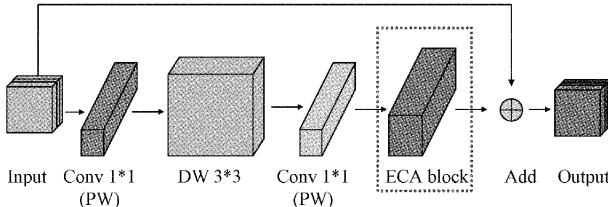


图 5 改进后倒残差模块 1

在 MobileNetV2 网络顶部分类器中,最后一层倒残差模块输出后经卷积、平均池化和全连接层后输出分类类别。为使分类器在对提取到的特征图具有更好的分类效果,提升网络对小特征的矿石图像的识别能力,本文在卷

积层后加入 ECA 模块,具体参数如表 2 所示。

表 2 嵌入注意力机制的分类器结构

输入	操作	输出
320×7×7	Conv2d 1×1 ReLU6	1 280×7×7
1 280×7×7	ECA_layer	1 280×7×7
1 280×7×7	Avgpooling	1 280×1×1
1 280×1×1	Dense	4×1

### 1.4 并行特征提取网络

并行特征提取网络设计是对 stride=2 倒残差模块的改进,改进后的倒残差网络结构如图 6 所示,为使网络所提取的特征信息更广,本文将深度卷积与空洞卷积相结合,设计由多个并列且独立的子网络组成的倒残差模块,其保留了原始的 PW 与 DW 串联结构,并并联了两个深度空洞卷积与 PW 卷积串联的特征金字塔网络(feature pyramid network)<sup>[22]</sup>,这样做增加了网络的宽度,可以在降低特征图分辨率损失的同时增大感受野。

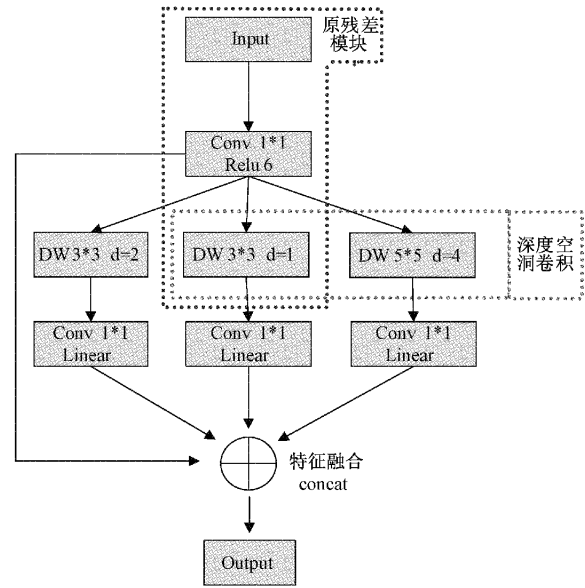


图 6 改进倒残差模块 2

感受野计算公式为:

$$n = k + (k - 1) \times (d - 1) \quad (10)$$

其中,  $k$  为卷积核大小,  $d$  为空洞率。

由式(10)可知,深度空洞卷积在选取不同空洞率的情况下具有不同的感受野。虽然增大  $d$  可以增加网络的感受野,但过大的空洞率会产生过多且无用的权重,影响网络的特征提取,所以本文中取  $d_1 = 2, d_2 = 4$ 。

### 1.5 迁移学习

在生活中,有许多利用已有的知识来学习新知识的例子,而在深度学习中的迁移学习便是如此,其核心便是学习两者之间的相关性。利用 Mobilenet V2 已在 ImgNet 数据集中训练好的权重文件,使模型不需要从零开始训练,

缩短模型总体训练时间,同时获得适合本数据集的初始训练权重文件。

本文迁移学习训练方案如下:下载官方的预训练权重文件;以 MobileNet V2 为主干建立模型导入权重文件并进行训练;首先冻结主干模型,对改进的分类器结构训练若干个批次,然后解冻主干模型,将训练好分类器的权重文件和官方权重文件一起导入至模型中进行训练,最后保存模型训练后的权重。

## 2 实验与分析

本次实验平台主要配置如下: Intel Core i7-9700 CPU, 32 GB RAM 3.00 GHz, NVIDIA-SMI 3080 GPU, Windows 10 专业版操作系统, Python3.8 语言, 解释器 PyCharm Community Edition 2020.3.3, Pytorch1.7.1 深度学习框架, CUDA 11.4 加速训练。本文算法设置的超参数如表 3 所示。

表 3 算法超参数设置表

优化器	损失率	损失函数	轮次	批次大小
Adam	0.001	Crossentropy	200	16

### 2.1 X射线矿石图像数据集

在现有的深度学习数据集中,暂时还没有一个公开的具有权威性的矿石数据集。本文以钨矿石为样本目标,构建了一个由 X 射线透射的矿石图像组成的数据集。

本次数据集的采集在位于赣南的某矿山企业选矿厂完成,为提升数据集的泛化能力,在 X 射线选矿机运行于不同射线照射环境下对矿石进行取图,获得不同环境下的 X 射线矿石数据集。由于金属矿石相比废石在原子序数上有较大差距,钨矿石更难穿透,而废石更容易穿透,所以原矿石经射线照射后,由钨矿和废石之间吸收 X 射线的差异性可以对两者进行区分。如图 7 所示为原矿石经 X 射线照射后的图像,可以看出,以 X 射线透射后所取的矿石图像特征分布明显、且成像后呈黑白图样,精矿部分由于 X 射线难以穿透呈深黑色,废石部分由于容易穿透呈浅灰色。由此组成的数据集具有特征分布明显的特点,在深度学习中利用此特点可以更好的提取不同矿粒的含矿量分布特征,且减少模型在训练时的信息吞吐量,提升学习效率和准确率。

在初步收集矿石射线图像后,将所有图像经矿山专业人士进行判定挑选,弃除存在争议的矿石图像,整个数据集共有 11 926 张图像,将矿石图像 1:1 分为含矿(6 046 张)和废石(5 880 张)。定义矿石在整颗矿粒的分布含量不同可分为四类做模型训练:高含矿量(high)、中含矿量(middle)、低含矿量(less)和不含矿(none)。在本文的模型训练中,将数据集按 8:1:1 划分为训练、验证和测试集。每类所含的图片总量和划分后数据集如表 4 所示。

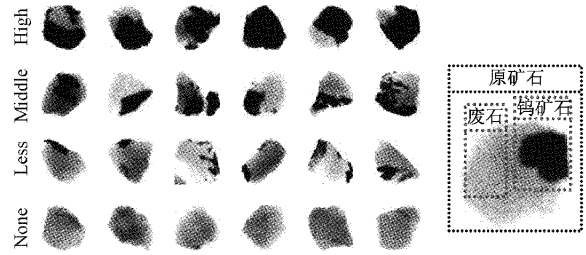


图 7 本数据集四类矿石图样

表 4 自建矿石数据集划分表

类别名称	图片数量/张	类别名称	图片数量/张
high	2 219	less	1 868
middle	1 959	none	5 880
训练集	9 545	验证集	1 193
测试集	1 188	总量	11 926

### 2.2 改进后实验指标对比

#### 1) 实验设计和结果

为证明本文中算法的各个改进方案对原算法具有增益效果,特设计消融实验对本文算法进行分析比较。在本次实验中,以模型在测试集的准确率、模型参数量和模型训练时间作为评价指标,对模型做 50 次迭代训练,其他参数设置和表 3 相同。具体实验变量控制方案如下:

实验 1:初始 MobileNet V2 网络。

实验 2:调整原网络参数矩阵,将最后一层倒残差层扩展因子设为 3,宽度因子设为 0.75。

实验 3:分类器中添加 ECA-net 注意力机制模块。

实验 4:分类器中添加 SE-net 注意力机制模块,与实验 3 作对比分析。

实验 5:当 stride=1 时,在倒残差模块中添加 ECA 注意力机制模块。同时分类器中也添加注意力机制模块。

实验 6:当 stride=2 时,将倒残差模块替换为并行特征提取网路结构。

实验 7:使用本文所有改进方案的模型。

如表 5 是本次消融实验的实验结果。

表 5 消融实验结果

实验序号	准确率/%	参数量/M	训练时间/s
1	94.164	2.229	2 472
2	93.996	<b>1.182</b>	<b>1 472</b>
3	94.676	2.229	2 519
4	94.362	3.320	2 946
5	95.434	2.314	2 493
6	94.776	3.821	3 487
7	<b>96.490</b>	1.362	1 684

#### 2) 实验结果分析

由实验 1 和实验 2 对比可以看出,虽然将扩展因子和

宽度因子调整之后模型准确率下降了 0.168%，但模型参数量却获得了 36.9% 降幅，此损失完全可以由其他改进弥补；

由实验 1 和实验 3 对比可知，由本文改进的分类器使模型准确率获得 0.512% 的提升，同时模型参数量并无增加；

实验 4 是为了对比两种注意力机制模块的扩展实验，由实验 1 和实验 4 对比可以看出，SE 注意力机制模块也能使模型准确率提升 0.198%，但相较于 ECA 模块，添加此模块模型参数量增加了 48.9%，且对模型准确率的增益效果也不如 ECA 模块；

由实验 1 和实验 5 对比可以看出，在 stride=1 时的每一层倒残差模块中添加轻量级注意力机制模块(ECA)使模型准确率提升了 1.270%，且模型参数只增加了 3.8%；

由实验 1 和实验 6 对比可以看出，将 stride=2 时的每一层倒残差模块替换为本文中的并行特征提取网络，使模型准确率提升了 0.612%，由于此并行结构由三个卷积核不同的分支构成，增加了模型宽度和加大感受野的同时，也会增加模型的参数量，所以参数量也有相应的提高。

实验 7 是本文改进后的算法实验，和原网络对比可以发现，经过本文的改进后，本文算法在自建矿石数据集的测试准确率为 96.490%，相比 MobileNet V2 提升 2.326%，同时算法参数量也比原算法更轻量化，为后期模型部署在边缘设备提供可行性。本次消融实验模型训练过程如图 8 所示。

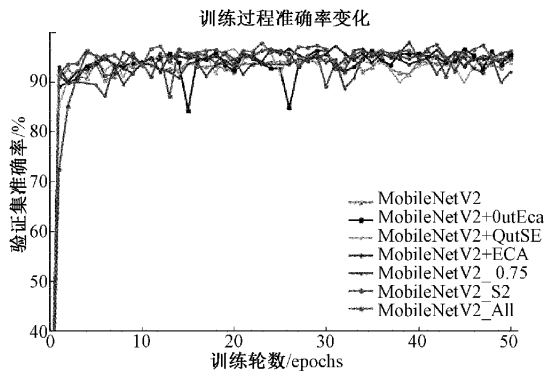
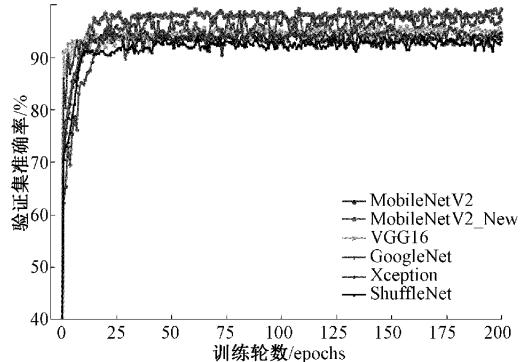


图 8 消融实验训练过程示意图

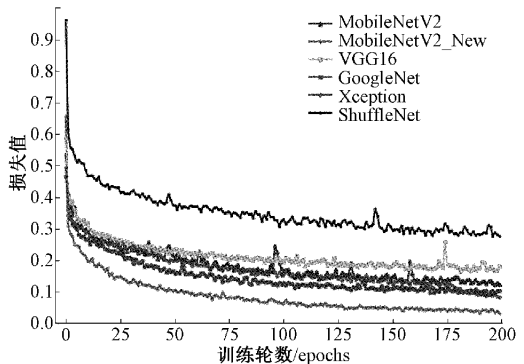
### 2.3 其他算法实验对比与分析

#### 1) 实验设计与结果

为了证明本算法改进相比其他算法具有一定优越性，本次实验在同一环境下针对 MobileNet V2、Vgg19、GoogleNet、Xception 和 ShuffleNet 五种流行算法与本文改进算法进行对比实验。以各算法在测试集中的测试准确率(Acc)、同一条件下检测矿石图片所用时间(T)、模型参数量和保存模型大小作为评价指标。训练过程可视化如图 9 所示，可以看出，使用迁移学习的训练方法使得模型刚开始就有很好的训练效果。



(a) 训练过程准确率变化



(b) 训练过程损失值变化

图 9 训练过程可视化示意图

将训练完的模型在测试集上测试，分步测试各算法，识别测试集中 1 188 张矿石图片，测试过程可视化热力图如图 10 所示，也可称测试结果平均混淆矩阵。其中纵坐标表示矿石真实类别标签，横坐标表示算法预测矿石标签，右侧渐变线则表示算法预测准确率，准确率越高颜色越深，反之越浅，对角线表示算法对各类别预测的准确率。

#### 2) 实验结果分析

由图 10(a)可以看出，在未对 MobileNet V2 算法进行改进时，虽然算法准确率为 94.203%，但针对 less 这一类别来说，却只有 70% 准确率，其中有 3% 被识别为高含量矿，27% 被识别为中含量矿，1% 识别为不含矿。这可能是由于少含量矿石图像和中含量矿石图像在含矿特征信息上极为相似，算法丢失微小特征信息，混淆对有用和无用信息的提取而导致的。由图 10(b)可以看出，当在算法中加入 ECA 注意力机制和并行特征提取网络后，算法增强了对特征信息的提取能力，less 类别准确率提升为 89%，其中 10% 被识别为高含量矿，1% 被识别为不含矿。显然本文的改进对算法具有一定的增益效果。

各模型识别性能如表 6 所示。对比实验结果可以发现，与 VGG16、GoogleNet 和 Xception 相对大型的网络对比，本文算法在准确率、检测时间、模型参数量和模型大小都优于这三种算法，在准确率上分别提升 2.69%、2.94% 和 0.67%。由于本文算法属于轻量级算法，而后者属于大型网络，模型深度相对较深，更为复杂，所以本文算法在测

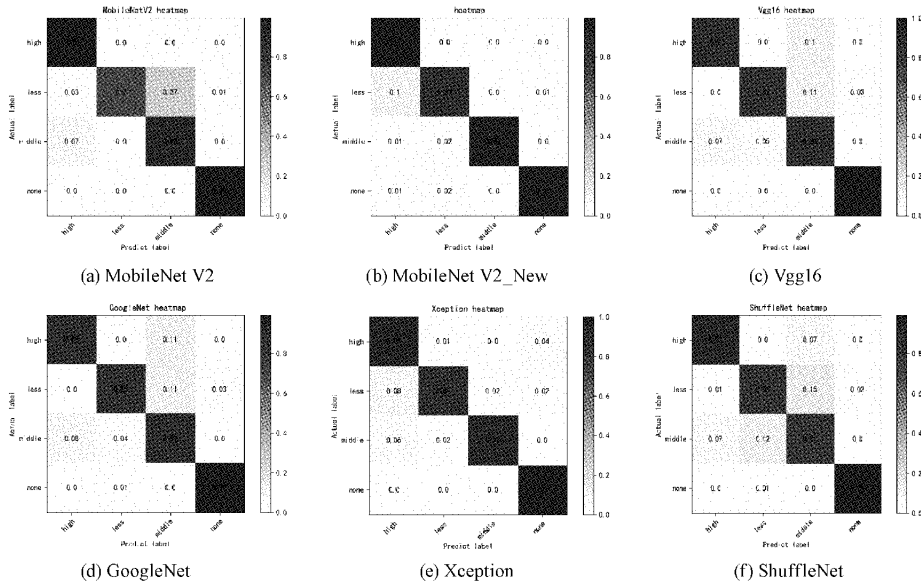


图 10 各算法测试结果可视化热力图

表 6 各算法实验结果

模型	准确率/%	时间/ms	P/M	参数量/M
本文算法	<b>96.720</b>	27	1.362	5.4
MobileNetV2	94.203	46	2.229	8.5
VGG16	94.037	128	14.717	168
GoogleNet	93.780	86	10.316	39.35
Xception	96.051	98	20.870	239
ShuffleNet	92.724	30	<b>1.258</b>	<b>4.8</b>

试时间、参数量和模型大小方面显然会有更优表现,且在  
在实际应用中本算法可以部署于算力不足的边缘设备中。  
与同样是轻量级网络 ShuffleNet 和 MobileNet V2 相比,  
本文算法在准确率上分别高于后者 3.996% 和 2.517%。  
虽然在模型参数量和模型大小上 ShuffleNet 网络优于本  
文算法,但其在检测时间上却比本文算法慢 3 ms,参数量  
只比本文算法少 0.104 M,模型大小也比本文算法小  
0.6 M。由于本模型算法的检测目标是针对矿石分类而  
言,而矿产资源日渐贫化的今天,矿石分类的准确率显然

是本文的首要评价指标。

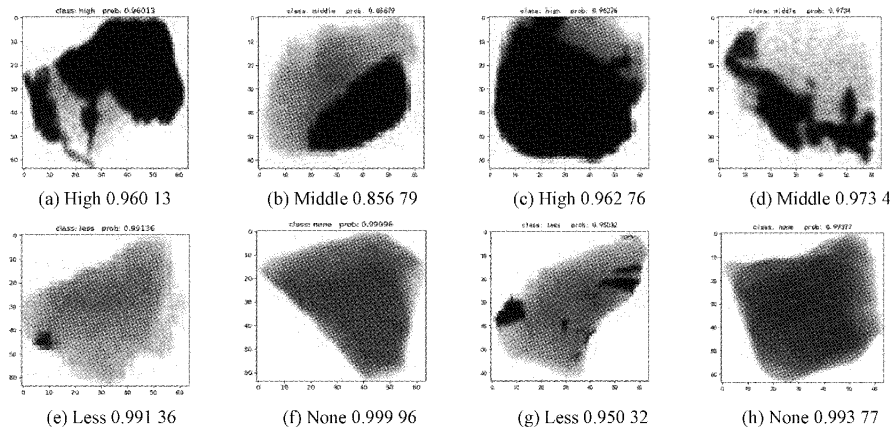
所以综合对比几种模型在矿石数据上的表现可知,本  
文算法在对 X 射线矿石检测的准确率优于其他算法,虽然  
在其他评价指标上表现不是最佳,但与最佳相比只有轻微  
差距。综合考虑下,本文算法具有最佳表现,为部署在边  
缘设备上应用提供可行性。

### 2.4 识别测试

为证明本文算法在部署于算力不足的边缘设备中具  
有良好的表现,将本文中矿石分选模型部署于英伟达  
Jetson nano developer kit(B01)中,对同一文件夹中存放的  
50 张 X 射线矿石图像进行识别准确率测试和识别速度测  
试,打印并记录模型识别结果。

#### 1) 识别准确率测试

在识别的结果中随机抽选三组矿石图样如图 11 所  
示,其中图样下方为矿石正确标签和算法检测时打印出  
预测为正确标签的准确率。可以看出,模型对矿石分选  
的准确率具有较好的表现,满足矿石分选的准确率要  
求。



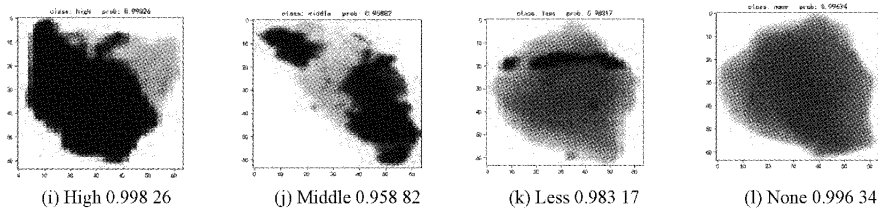


图 11 本文算法矿石识别结果图

2) 矿石识别速度测试

在矿石分选系统中,为保证分选速率,模型识别矿石后将识别信息传递到分选执行系统进行下一步分选执行动作,这就对模型的识别速度有较高的实时性要求,在保证识别准确的前提下执行机构获得识别信息并及时进行分选操作。紧凑而正确的识别并执行分选操作是提升分选速率的关键。本次识别速度测试对比其他算法,识别结果如图 12 所示。可以看出,本文算法在识别速度方面具有较好的表现,可以满足矿石分选系统的实时性分选需求。

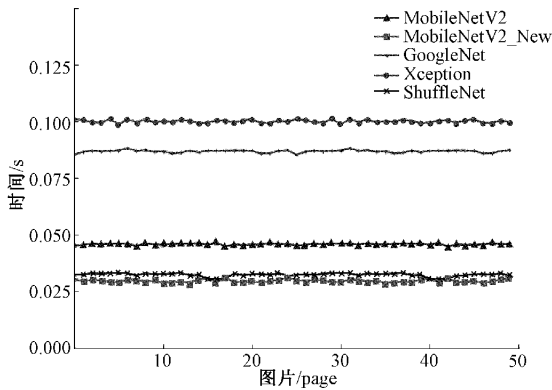


图 12 Jetson nano(4 GB) 识别速度

3 结 论

以 MobileNet V2 为主干模型应用于矿石检测,对其进行算法改进。对于网络更轻量化,本文通过调整扩展因子和宽度因子使得模型参数量获得大幅降低,有利于模型部署在算力不足的边缘设备中;对于提升网络的识别性能,本文通过对主干中的残差模块改进,将步长为 1 的残差模块替换为本文嵌入 ECA 注意力机制的倒残差模块,将步长为 2 的倒残差模块替换为本文的并行特征提取网络,该特征提取网络使得模型更宽,具有更大的感受野。有利于模型具有更强的有用特征信息提取能力,对本文自建 X 射线矿石数据集中部分矿石特征分辨不明显这一特性,具有很好的适应性;对于分类器识别分类而言,本文在一维卷积层后嵌入 ECA 模块,对分类信息聚合提取有用信息,加强分类性能。综合以上改进,本文改进算法在自建矿石数据集上的识别能力具有很好的表现,为今后矿山智能化建设打下基础。

本文改进算法主要针对于矿石的识别,在矿石分类识别上具有一定的性能提升,但针对目前全球矿产资源日渐贫化这一大环境,96.720%的准确率还是会有部分矿石未被识别到,造成资源浪费,显然在识别准确率方面还需要更高的提升空间;本文算法对矿石的识别只是属于矿石分选的第一部分:检测,为使后续分选执行系统更好的分选矿石,对矿石的大小、在传送带中的相对位置检测同样重要;在数据集方面,本文只针对于经 X 射线照射后的黑钨矿石数据集,而全球矿石种类繁多,完善现有各类矿石数据集也是机器视觉技术广泛应用于矿山的关键一步。后期将需要对这些方面进行下一步研究。

参考文献

[1] 葛丁, 梁殿印. 基于 X 射线透射的矿石品位检测方法研究[J]. 有色金属:选矿部分, 2019(4): 87-93.

[2] DEO A J, BEHERA S K, DAS D P. Detection of material build-up during online size distribution analysis of iron ore green pellets using MobilenetV2 [C]. 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA), 2021, DOI: 10.1109/IRIA53009.2021.9588756.

[3] FIRLA M, LIPNICKI P, LEWANDOWSKI D. Image processing algorithm for the assessment of the ore fragmentation size distribution [C]. 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, 2019: 505-512.

[4] LIANCHENG M A, ZHANG Y, SONG G, et al. Ore granularity detection and analysis system based on image processing [C]. 2019 Chinese Control And Decision Conference (CCDC), 2019: 359-366.

[5] 邓田, 余翼. 基于 PSO-Faster R-CNN 改进算法的矿石识别分类研究[J]. 矿业研究与开发, 2021, 41(2): 178-182.

[6] 文大鹏, 梁西银, 苏茂根, 等. 激光诱导击穿光谱技术结合 PCA-PSO-SVM 对矿石分类识别[J]. 激光与光电子学进展, 2021, 58(23): 9.

[7] 高瑜璋, 乔铁柱, 车剑. 矿用输送带纵向撕裂视听融合检测方法[J/OL]. 电子测量技术: 1-6. [2022-05-17]. <http://gfffgd97c66c3752f4337hvcfukuoc6bco6cpp>.



- [8] SZEGEDY C, WEI L, JIA Y, et al. Going deeper with convolutions [C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2015:1-12.
- [9] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014:1-14.
- [10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. 2016 the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016:770-778.
- [11] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5MB model size [J]. ArXiv Preprint, 2016, ArXiv:1602.07360.
- [12] TECHNICOLOR T, RELATED S, TECHNICOLOR T, et al. ImageNet classification with deep convolutional neural networks[C]. International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc, 2012: 1097-1105.
- [13] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018: 6848-6856.
- [14] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 1251-1258.
- [15] HOWARD A, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii: IEEE, 2017: 1704-1712.
- [16] HUANG G, LIU Z, LAURENS V, et al. Densely connected convolutional networks[J]. IEEE Computer Society, 2016, DOI:org/10.48550/arXiv.1608.06993.
- [17] ZHAO G, QUAN L, LI H, et al. Real-time recognition system of soybean seed full-surface defects based on deep learning-sciencedirect [J]. Computers and Electronics in Agriculture, 2021, DOI: org/10.1016/j.compag.2021.106230.
- [18] KNAPHEIDE J, STABERNACK B, KUHNKE M. A high throughput MobileNetV2 FPGA implementation based on a flexible architecture for depthwise separable convolution[C]. 2020 30th International Conference on Field-Programmable Logic and Applications(FPL), 2020.
- [19] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
- [20] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2020, DOI:10.1109/CVPR42600.2020.011155.
- [21] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, DOI: org/10.48550/arXiv.1709.01507.
- [22] SEUNG W K, HYONG K K, JEE Y S, et al. Parallel feature pyramid network for object detection [J]. Springer, Cham, 2018:239-256.

#### 作者简介

杨文龙,高级工程师,硕士生导师,主要从事数字矿山、智能工厂建设和智能选矿装备研发。

E-mail: YWL6582@163.com

郭明钰(通信作者),硕士研究生,主要从事计算机视觉方面的研究。

E-mail: 1065420173@qq.com