

DOI:10.16136/j.joel.2023.01.0123

煤矸的轻量级智能分选网络

王天奇¹, 贾晓芬^{1,2*}, 杜圣杰¹, 郭永存², 黄友锐¹, 赵佰亭¹

(1. 安徽理工大学 电气与信息工程学院, 安徽 淮南 232001; 2. 安徽理工大学 省部共建深部煤矿采动响应与灾害防控国家重点实验室, 安徽 淮南 232001)

摘要:针对现有煤矸分选方法存在模型复杂、实时性差、特征易丢失等问题,构建了一种轻量化煤矸分选网络 GC-ResNet18。GC-ResNet18 利用幽灵卷积(ghost convolution, GC)线性生成 ghost 映射的特性,剔除煤和矸石相似性特征的冗余信息。借助 Softpool 的下采样激活映射,保留、凸显煤和矸石的特征信息并去除冗余参数,防止过拟合现象。引入 GC 自注意力机制,融合 SENet 的轻量化和 NLNet 长距离信息全局捕获的优势,使网络记忆、放大煤矸图像间的细微差异特征,提升煤矸图像的识别准确率。实验结果表明,GC 降低了 46.6% 的参数量,GC 自注意力机制在 CIFAR10、CIFAR100 上分别提升 1.44%、2.32% 的准确率,而 Softpool 池化在上述两个数据集中分别提升了 0.22%、0.17%。通过对比实验,全面改进后的 GC-ResNet18 网络在训练效率和分类精度上优于 VGG19-S-GDCNN、SBP-VGG-16 等模型,在 CIFAR10 和 CIFAR100 数据集中的分类精度与同规模的网络相比均达到最优的 94.07% 和 74.95%,并最终在自建煤矸数据集上达到了 97.2% 的分类准确率。

关键词:煤矸分选; 神经网络; 幽灵卷积(GC); 高效池化层; 自注意机制**中图分类号:**TD94;TD391.9 **文献标识码:**A **文章编号:**1005-0086(2023)01-0019-07

Lightweight intelligent separation network for coal and gangue

WANG Tianqi¹, JIA Xiaofen^{1,2*}, DU Shengjie¹, GUO Yongcun², HUANG Yourui¹, ZHAO Baiting¹

(1. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan, Anhui 232001, China; 2. State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines, Anhui University of Science and Technology, Huainan, Anhui 232001, China)

Abstract: Aiming at the problems of complex model, poor real-time performance and easy loss of characteristics in the existing coal gangue separation methods, a lightweight coal gangue separation network GC-resnet18 is constructed. GC-resnet18 uses ghost convolution (GC) to linearly generate ghost map to eliminate redundant information of similarity characteristics of coal and gangue. With the help of Softpool's down sampling activation mapping, the characteristic information of coal and gangue is retained and highlighted, and redundant parameters are removed to prevent over fitting. GC self-attention mechanism is introduced to integrate the advantages of lightweight of SENet and global capture of long-distance information of NLNet, so as to make the network remember and enlarge the subtle difference characteristics between coal gangue images, and improve the accuracy of coal gangue image recognition accuracy. The experimental results show that GC reduces the amount of parameters by 46.6%, and the accuracy of GC self-attention mechanism is improved by 1.44% and 2.32% on CIFAR10 and CIFAR100 respectively, while Softpool pooling is improved by 0.22% and 0.17% on the above two data sets respectively. Through comparative experiments, the comprehensively improved GC-ResNet18 network is better than VGG19-S-GDCNN, SBP-VGG-16 and other models in training efficiency and classification accuracy. The

* E-mail: jxfzbt2008@163.com

收稿日期: 2022-03-03 修订日期: 2022-03-25

基金项目: 国家自然科学基金面上项目(52174141)、安徽省自然科学基金面上项目(2108085ME158)、安徽省重点研究与开放计划(202104a07020005)和安徽高校协同创新项目(GXXT-2020-54)资助项目

classification accuracy in CIFAR10 and CIFAR100 data sets is the best 94.07% and 74.95% compared with the network of the same scale, and finally reaches 97.2% classification accuracy in self-built coal gangue data sets.

Key words: coal gangue sorting; neural network; ghost convolution (GC); high efficient pooling layer; self-attention mechanism

0 引 言

煤炭是世界上最重要的化石能源之一,开采过程中夹杂的矸石,不仅影响标煤燃烧热值,还会污染环境^[1],因此煤矸分选工作意义重大。然而,常用的重介质法和跳汰法,工艺复杂、设备成本高,水资源浪费严重;γ射线法等则污染环境且危害人体健康;而基于图像识别的煤矸分选方法,因具有效率高、成本低、无污染的特点而受到广大学者的关注。

卷积神经网络(convolutional neural network, CNN)可以自动学习图像的特征信息,已被广泛应用于煤矸图像的目标识别领域。如 PU 等^[2]通过引入著名的 VGG16 深度 CNN 模型和迁移学习技术识别出了煤与煤矸石,大大节省了人力成本; HONG^[3]等使用 AlexNet 并结合 SVM 算法实现了煤和矸石的高准确率识别; HUANG 等^[4]在 SqueezeNet 网络基础上结合深度可分离卷积构建了 SqueezeNet-DS 轻量级网络模型,实现了煤与矸石的快速识别。上述方法虽在煤矸石识别工作中取得了不错的效果,但是模型结构复杂,识别效率低,鲁棒性差,在复杂的实际分选环境下,结果差强人意。恶劣的实际分选现场环境、高速的运输皮带,都对煤矸分选模型的网络体积、复杂度和运行效率提出了更高要求,而现有的大多模型因自身可观的计算量等问题无法满足煤矸分选的实时性需要,因此,如何加快模型效率是目前煤矸分选领域亟待解决的问题之一。

为了提高煤矸石的分类识别精度和速度,实现高效率、高精度、低消耗、低人力的识别效果,提出一种轻量级 CNN 识别算法 GC-ResNet18,它是侧重于提取煤矸灰度和纹理特征信息差异的轻量级神经网络,对实现煤矸石精准分选工作具有重要意义。

1 轻量化煤矸识别网络 GC-ResNet18

GC-ResNet18 以 ResNet18^[5]为骨干网络,网络依次由初级特征提取模块(primary feature extraction module)、残差模块(residual module)和图像分类模块(image classification module)构成。其结构见图 1。初级特征提取模块中使用高效池化层 Soft-

pool^[6]替换常规池化层,更好地保证输入特征的完整性。残差模块中使用幽灵卷积(ghost convolution, GC)替换 GCR 块中的所有常规卷积层,同时在 GCR 块中引入 GC 自注意机制,以加快网络计算效率并提高对煤矸特征的识别能力。图像分类模块接收残差模块提取的有效特征,并利用 softmax 分类层得到对煤和矸石的分类结果。

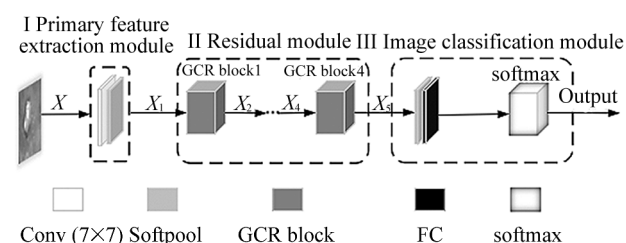


图 1 GC-ResNet18 网络结构图

Fig. 1 Diagram of GC-ResNet18 network structure

1.1 初级特征提取模块

初级特征提取模块旨在保留并凸显煤和矸石的灰度、纹理特征信息,而广泛使用的最大池化和平均池化两种操作又是丢失特征信息的主要环节,这将对 CNN 识别煤和矸石产生不利的影响。Softpool^[7]是一种新型高效池化层,是基于 softmax 加权方法来保留输入基本属性的同时放大更大强度的池化操作。笔者将其加入到初级特征提取模块中的普通卷积层之后,凭借其优异的池化效果达到网络的轻量化目的。Softpool 操作包括前向和反向两个过程,具体操作过程见图 2。

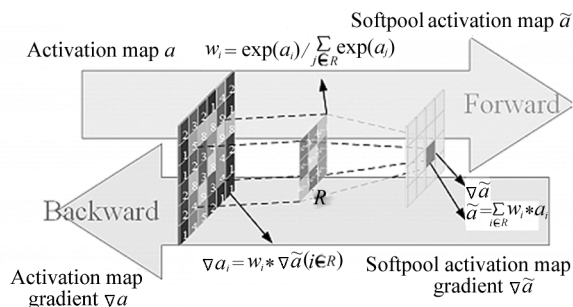


图 2 Softpool 操作过程

Fig. 2 Softpool operations

假设初始输入 X 经过 $\text{Conv}(7 \times 7)$ 卷积层后得到激活映射 a 。正向操作中, Forward 内核使用每个激活的指数 softmax 值作为权重来计算区域 R 的加权和, 梯度也使用相同的权重, 用 Backward 内核表示, 激活梯度与计算得到的 softmax 权重成正比, w_i 和 $\tilde{a}$ 计算方式为:

$$w_i = \frac{e^{a_i}}{\sum_{j \in R} e^{a_j}}$$

$$\tilde{a} = \sum_{j \in R} w_i * a_i. \quad (1)$$

反向操作中, 梯度更新与前向传递期间计算的权重成比例, 先对 $\tilde{a}$ 进行求导得到 $\nabla \tilde{a}$, 再由前向操作中得到的 w_i 计算得出 ∇a , 计算方式为:

$$\nabla a_i = w_i * \nabla \tilde{a} (i \in R), \quad (2)$$

式中, $\tilde{a}$ 表示 a 加权和, ∇a 表示 a 的梯度值, $\nabla \tilde{a}$ 表示 $\tilde{a}$ 的梯度值, R 表示激活区域内的最大近似值 ($i, j \in R$), w_i 表示区域 R 的权重。

与常规池化操作不同, Softpool 为区域 R 内的每个正激活分配一个最小的非零权值, 其权重与相应的激活值一起参与非线性变换, 其中较高的激活比较低的激活占更多的主导地位, 因为大多数池化操作都是在高维特征空间中执行, 突出较高的激活的影响比简单地选择最大值或平均值是一种更平衡的方法。该池化方法可以平衡 AvgPool 和 MaxPool 的不利影响, 同时兼有两者的有利属性。所以, 选择 Softpool 去除冗余参数特征在防止过拟合以及凸显重要特征方面的效果更优。此外, Softpool 还在节省计算资源和内存开销方面也有着更为出色的表现, 这对在实际工作环境中的嵌入煤矸识别模型提供了一定的优势。

1.2 残差模块

残差模块对煤矸的有效特征完成进一步筛选, 加强网络模型对关键特征的学习能力。该模块由 4 个依次相连的 GCR block 组成, 它是在 ResNet18^[5] 的基础残差块 (basic block) 中引入 GC 和 GC block 得到的, GCRblock 的结构见图 3, 它的卷积层全部

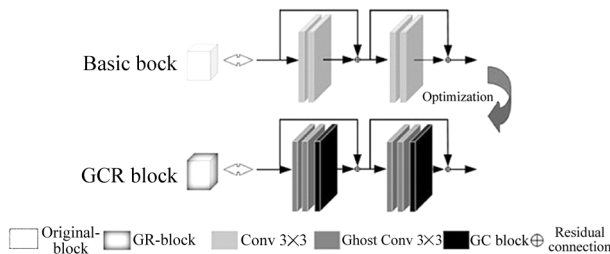


图 3 GCR block 实现过程

Fig. 3 GCR block implementation process

是 GC, 且在每个残差连接中都插入了 GC block 自注意模块。

1.2.1 GC

在进行煤矸石识别的过程中, CNN 产生的计算量大多来自于卷积运算过程, 因此, 从卷积层方面入手是实现模型轻量化的方法之一。HAN 等^[8] 提出一个仅通过少量卷积计算就能生成大量特征图的卷积操作——GC。

传统卷积提取图像特征会导致细节信息的遗漏, 且煤矸石分选对细节特征要求严格, 而 GC 作为典型代表, 其内部线性运算能够强化输入数据的区分能力, 以及削弱不相关因素。因此借助 GC, 并利用煤和矸石相似性之间的冗余信息, 线性地生成 ghost 映射, 可以实现对同类图像特征的快速、全面的学习。在每个独立的残差模块中, 利用不同大小的卷积核逐步提取特征信息, 利用可变步长卷积实现图像通道信息的维数增减, 前后层的信息流通过残差连接进行监督和加强。

引入的 GC 与常规卷积的区别在于采用线性变换代替常规卷积中的大部分卷积, 大大减少由卷积过程产生的计算量, 从而达到网络轻量化效果, 与此同时还加快了网络训练速度, 提升了分类识别效率。

1.2.2 GC 自注意力机制

现有的煤矸分选网络是通过学习煤和矸石之间的差异性特征, 注意力机制作为能有效提升模型性能的潜在手段, 可以直接有效地提升轻量型神经网络的性能且不给网络增加额外的负荷, 进而实现更精准的煤矸识别效果。常见的 CBAM^[9]、SE^[10]、ECA 等^[6] 注意力机制的原理多是基于图像特征在空间和通道方面的进一步利用, 但它们的作用范围也只是关注与之邻近区域的特征信息, 将其应用到煤矸等差异特征不大的识别任务中, 效果并不明显。GC^[11] 自注意机制是一种能够将全局上下文信息融合到通道中实现网络性能提升的方法, 它将 SENet^[10] 的轻量化特点和 NLNet^[12] 捕获长距离信息全局上下依赖关系的优势结合起来, 进而实现比单一网络更好的性能提升。

GC block 结构见图 4, 它能够使网络记忆并放大煤和矸石之间的细微差异特征并提醒网络在特征提取过程中重点学习这种差异特征, 以此作为煤矸的主要判断依据, 对于提升煤矸识别准确率来说极为有效。

图 4 中, C 表示输入图像通道数, H 、 W 分别表示输入图像高度和宽度, r 表示瓶颈比率, C/r 表示

瓶颈的隐藏表示维, W_k 、 W_{v1} 、 W_{v2} 分别表示不同位置的线性变换矩阵, $\otimes$ 表示矩阵乘法, $\oplus$ 表示加法操作。

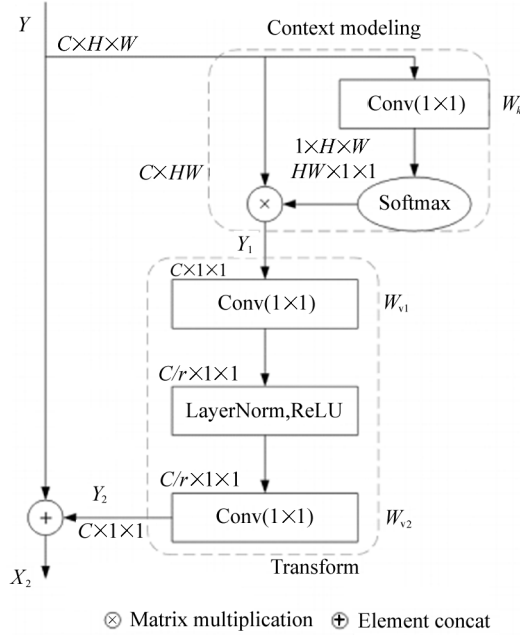


图4 GC block 结构图

Fig. 4 Diagram of GC block structure

GC block 是一种全局上下文建模构件, 内部包含上下文建模模块(Context Modeling)和信息转换模块(Transform)。其中 Context Modeling 模块用来捕获输入图像全局特征通道信息间的相关依赖性, 进而实现通道注意力机制的效果, 假设 GCR block1 中最后的 GC block 层输入为 Y , 则 Context Modeling 模块内的计算过程为:

$$Y_1 = Y \otimes (W_k Y)_{\text{softmax}}, \quad (3)$$

式中, W_k 表示 Context Modeling 模块内卷积层对应的线性变换矩阵, softmax 表示执行相应函数映射, Y_1 表示输入 Y 经过 Context Modeling 模块后得到的结果, 至此, 全局通道间依赖性信息捕获完成, GC block 实现自注意效果。随后将 Y_1 送入 Transform 模块中去, 使用两个 1×1 卷积层完成特征信息的瓶颈转换, 其目的是模仿 SE block 中瓶颈结构来有效减少参数量, 具体实现过程为:

$$Y_2 = W_{v2} (LN(W_{v1} Y_1))_{\text{ReLU}}, \quad (4)$$

式中, W_{v1} 、 W_{v2} 分别表示 Transform 模块内部不同位置的线性变换矩阵, LN 表示对层归一化, ReLU 表示执行相应激活函数操作, Y_2 为 Y_1 经过 Transform 模块计算后的输出结果, 因此完整的 GC block 模块实现过程为:

$$X_2 = Y + W_{v2} (LN(W_{v1} (Y \otimes (W_k Y)_{\text{softmax}})))_{\text{ReLU}}, \quad (5)$$

式中, X_2 为输入特征 Y 经过整个 GC block 后得到的结果, 同时也是 GCR block1 最终的输出。

至此, GC block 在有效降低参数量的同时实现全局通道信息间的自我关注机制。通过在残差连接中引入 GC 自注意模块, 实现了煤矸的高精度、高效率分类识别。

1.3 图像分类模块

进行煤矸分类识别的过程首先是将煤和矸石的图像输入到 CNN 模型中, 然后通过卷积层得到煤和矸石图像的轮廓、灰度、纹理等相关特征图, 通过池化层对特征图进行降维处理, 其次通过图像分类模块中的全连接层整合卷积层和池化层中学习到的特征信息, 得到煤和矸石的特征向量, 最终通过该模块的 softmax 分类层得到煤和矸石的种类。

该模块保留了 ResNet18 网络的原有结构, 包括 Softpool 池化层、全连接层 FC 和 softmax 分类器。在接收初级特征提取模块和 GC-ResNet18 残差模块中传入的特征信息时, 全连接层能够保持较大的模型容量, 从而保证模型表示能力的迁移, 而 softmax 分类器具有更好的解释性, 让取阈值等操作更加顺理成章, 输出为概率, 最终概率的最大值即为识别目标。

2 煤矸数据集建立

为了更好地实现实际应用环境下煤与矸石的有效分选, 对煤和矸石在实际场景下进行了充分的图像采集和数据集制作。煤和矸石如图 5 所示。采集煤矸石图像时保持光源大小一致, 光照强度不变, 减少光照对煤矸石图像的影响。由于煤矸石的光照、矿石位置与矿石形状随机性较强, 且图像纹理构成较为简单, 因此, 提取更为有效的特征尤为重要。煤与矸石在纹理和光泽上都有差异, 煤在图像中表现为有光泽的岩石, 且随着变质程度的增高显示出各种类型的光泽。矸石一般不具有光泽的成分, 所以其块状结构的图像表现一般, 为无光泽的固态岩石。因而, 提取煤和矸石图像的纹理特征用于煤矸分选存在可行性。此外, 煤的图像更黑, 而矸石的图像更加灰白, 这是煤和矸石最直接、最显著的外在特征。因此, 通过提取煤和矸石的灰度特征也同样可行。

煤矸数据集采集原始图像 931 幅, 大小均为 3016×4032 , 其中原煤图像 449 张, 矸石图像 482

张。为了防止训练过程出现过拟合现象,后期通过旋转、裁剪及亮度调节等方式实现数据集增强,最终得到煤矸数据集共包含 3 724 张 224×224 大小的图像,其中原煤图像 1 796 幅,矸石图像 1 928 幅。随后使用 Labelme 工具完成全部图像的标注工作。其中,训练集:测试集=4:1。

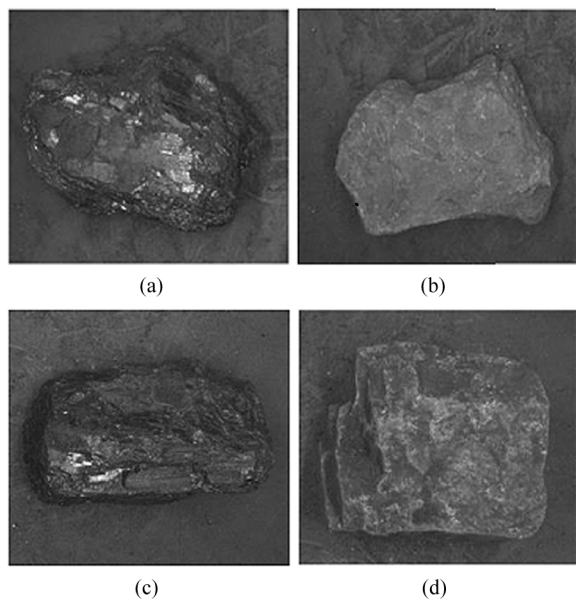


图5 煤和矸石:(a)(c) 煤;(b)(d) 矸石

Fig. 5 Coal and gangue:(a)(c) Coal;(b)(d) Gangue

3 实验结果及分析

实验在 Pytorch 1.5.0、NVIDIA GTX 2080Ti 的 PC 上开展,训练 GC-ResNet18 网络时,迭代轮数设置为 120 个 Epoch,学习率初始化为 0.1,每训练 30 个 Epoch 学习率衰减为原来的十分之一,采用随机梯度下降+动量法(SGD+Moument)训练。

实验在 CIFAR10、CIFAR100 及自制的煤矸数据集上进行。数据集 CIFAR10 和 CIFAR100 均是 32×32 的彩色图像,分别包含 10、100 个类别。使用参数量(Params, P)、浮点计算量(FLOPs, F)和测试集上的准确率(Acc, A)作为模型的评价指标,前两个指标的值越低越好,最后一个指标的值越高越好。

3.1 消融实验

GC-ResNet18 网络是以 ResNet18 为骨干网络,通过替换 GC、添加 GC 自注意力机制和优化池化层实现优化网络,针对上述 3 个方面的改进策略进行实验来测试其具体效果。

1) GC 替换

首先对 GC 的替换效果进行测试,测试方案为:
a. 在 ResNet18 网络上使用 CIFAR10、CIFAR100 标

准数据集进行训练。b. 在用 GC 替换 ResNet18 网络中所有常规卷积的改进网络上使用 CIFAR10、CIFAR100 标准数据集进行训练。测试结果见表 1。

由表 1 可见:方案 b 相较于方案 a 在模型参数量方面大幅降低,降幅约为 46.6%,而准确率方面则略微下降。综合来看,使用 GC 替换常规卷积给模型带来的性价比更高,因此,采用 GC 来优化 ResNet 是可行的。

表 1 GC 替换实验结果

Tab. 1 Experimental results of GC substitution

Result	CIFAR-10			CIFAR-100		
	P/M	F/G	$A/\%$	P/M	F/G	$A/\%$
a	0.238 5	0.037	92.97	0.238 5	0.037	73.68
b	0.127 3	0.022	92.53	0.127 3	0.022	73.12

2) 添加 GC 自注意机制

随后,在上一步实验基础上向网络中添加 GC 自注意机制并进行对比实验,实验方案为:c. 使用不添加注意力机制的网络在 CIFAR10、CIFAR100 标准数据集上进行测试。d. 使用添加了 GC 自注意力机制的网络进重复 c 中的实验。测试结果见表 2。

表 2 GC 自注意力测试结果

Tab. 2 Test results of GC self-attention mechanism

Result	CIFAR-10			CIFAR-100		
	P/M	F/G	$A/\%$	P/M	F/G	$A/\%$
c	0.127 3	0.022	92.53	0.127 3	0.022	73.12
d	0.127 1	0.022	93.86	0.127 1	0.022	74.82

由表 2 可见:添加了 GC 自注意力机制的方案 d 相较于未添加的方案 c 在参数量上略微降低,浮点计算量不变,但方案 d 在两种数据集上的测试准确率均有所提高,其中 CIFAR10 上提升 1.44%,CIFAR100 上提升 2.32%。由此可见,添加 GC 自注意力机制来提升模型识别准确率是可行的。

3) Softpool 池化层替换

进一步地,在确定 GC 自注意力机制添加策略有效性的基础上,设计不同实验方案来验证 Softpool 高效池化层的有效性,实验方案为:e. 在未引入 Softpool 池化层的网络上使用 CIFAR10、CIFAR100 标准数据集进行测试。f. 引入 Softpool 池化层后重复 e 中的实验。实验结果见表 3。

由表 3 可见:引入 Softpool 池化层后的方案 f 未引起参数量的变化,但在浮点计算量上有所降低,准确率方面则比方案 e 略微上升,其中在 CIFAR10 上

提升 0.22%, CIFAR100 上提升 0.17%。综合来看, 引入 Softpool 池化层在改善模型复杂度和提升准确率两个方面都有着一定的优化作用, 因此采用 Softpool 池化层优化网络是可行的。

表 3 Softpool 池化层替换实验结果

Tab. 3 Experiment result of Softpool pooling layer replacement

Result	CIFAR-10			CIFAR-100		
	P/M	F/G	A/%	P/M	F/G	A/%
e	0.127 1	0.022	93.86	0.127 1	0.022	74.82
f	0.127 1	0.020	94.07	0.127 1	0.020	74.95

综上, GC 替换可以使网络轻量化, GC 自注意力机制能有效提升模型准确率, Softpool 池化层在降低模型复杂度和提升模型性能上均有促进作用, 由此确定 GC-ResNet18 网络的结构见表 4。

表 4 GC-ResNet18 网络结构 (Input: 224×224)

Tab. 4 GC-ResNet18 network architecture (Input: 224×224)

Layer	Output	Layer parameters	GC block
Conv1	112×112	7×7, 64, stride=2	—
Ghost Conv2_x	56×56	3×3 SoftPool, stride=2	—
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	√
Ghost Conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	√
Ghost Conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	√
Ghost Conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	√
FC	1×1	SoftPool, FC, Softmax	—
FLOPs		0.98 G	—

3.2 对比实验

确定 GC-ResNet18 网络的结构后, 在两个经典数据集上开展对比实验, 以验证分类识别效果。在数据集 CIFAR10 上测试, 并和文献[8]、[13]、[14-17]中的方法对比, 结果见表 5, 粗体表示最好结果, 加下划线表示第二好的结果。GC-ResNet18 的 Params 和 FLOPs 均是最低值, Acc 最高为 94.07%。ResNet18 的 Params 指标和 Ghost-ResNet-56 的 FLOPs 指标分别排名第二, 但 Acc 仅为 92.9% 和 92.7%。GC-ResNet18 比 ResNet18 和 Ghost-ResNet-56 分别提高了 1.17% 和 1.37% 的准确率。

在 CIFAR100 上测试, 并和文献[18-22]中的方法对比, 结果见表 6, 由表 6 可见, DWConvXSepConv 获得了最好的 FLOPs 指标, GC-ResNet18 获得了最好的 Params 和最高的 Acc, 准确率达到 74.95%, 比 DWConvXSepConv 提高了 0.95%。

综合表 5—6 可知, GC-ResNet18 网络在 CI-

FAR10、CIFAR100 两个公开的数据集的分类识别精度均获得了最高值。因此, 该网络在训练效率和分类识别精度相较于其他 CNN 模型具有明显的竞争力。

表 5 与其他模型在 CIFAR-10 上的性能比较

Tab. 5 Performance comparison with other models on CIFAR-10

Model	P/M	F/M	A/%
ResNet18	<u>0.24</u>	182	92.9
VGG-16 ^[13]	15.0	313	93.6
Ghost-ResNet-56 ^[8]	0.43	<u>63</u>	92.7
R-MnasNet ^[14]	3.0	—	91.3
Dagger(VGGNet) ^[15]	2.7	119	<u>93.9</u>
L1-ResNet-56 ^[16]	0.73	91	92.5
SBP-VGG-16 ^[17]	—	136	92.5
GC-ResNet18(our)	0.127	20.48	94.07

表 6 与其他模型在 CIFAR-100 上的性能比较

Tab. 6 Performance comparison with other models on CIFAR-100

Model	P/M	F/M	A/%
ResNet18	<u>0.24</u>	182	72.76
ShuffleNet ^[18]	0.91	161	69.0
MobileNetV2 ^[19]	1.18	158	68.1
VGG16-half ^[20]	5.40	225	68.8
DWConvXSepConv ^[21]	1.59	16.8	<u>74.0</u>
VGG19-S-GD ^[22]	3.20	161	73.6
GC-ResNet18(our)	0.127	<u>20.48</u>	74.95

3.3 煤矸识别效果验证

GC-ResNet18 是针对煤矸分选应用搭建的 CNN 模型, 它在模型复杂程度和识别准确率方面具有优越性, 与其他主流网络模型相比, GC-ResNet18 性能处于上等水平, 将其应用到煤矸数据集反复训练, 实验结果取平均值, 如表 7 所示。

表 7 GC-ResNet18 在煤矸数据集上的表现

Tab. 7 Performance of GC-ResNet 18 on coal and gangue data sets

Model	Model size/M	A/%
ResNet18	11.69	96.7
DenseNet121	7.05	68.68
GhostNet	5.18	94.5
MobileNetV2	5.63	92.5
GC-ResNet18	6.23	97.2

由表 7 可知, GC-ResNet18 网络在煤矸分选数据集上表现出色, 模型大小仅为 6.23 M, 准确率达到 97.2%, 其体积小、精度高的特点对于实际应用场景的嵌入安装及精准实现煤矸分离作业具有十分重要的意义。

4 结 论

文中提出的一种具有潜在应用前景的轻量化煤矸分选网络 GC-ResNet18。通过引入 GC 和 Softpool 池化,有效去除了冗余参数和重叠信息,降低了模型复杂度。同时,能够保证对煤矸的灰度、纹理等特征的有效提取,提高网络模型的学习和分类能力,从而高精度、高效率地完成煤矸分选工作。GC 自注意力机制使得网络更加轻量化的同时,显著提高了分类效率和精度。本文在数据样本的扩充和减少图像采集中的噪声干扰等方面还有提升空间。同时,针对煤和矸石的不同类型的图像进行分选时,如何选取合适的图像特征作为标准达到预期分选效果,也是将来研究的重要方向。

参考文献:

- [1] XIE H P, WANG J H, WANG G F, et al. New idea of coal revolution and development of coal science and technology[J]. Journal of China Coal Society, 2018, 43(5): 1187-1197.
谢和平, 王金华, 王国法, 等. 煤炭革命新理念与煤炭科技发展构想[J]. 煤炭学报, 2018, 43(5): 1187-1197.
- [2] PU Y, APEL D B, SZMIGIEL A, et al. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning[J]. Energies, 2019, 12(9): 1735.
- [3] HONG H C. Research on coal gangue sorting algorithm based on machine vision[D]. Quanzhou: Huaqiao University, 2018.
洪惠超. 基于机器视觉的煤矸石分选算法的研究[D]. 泉州: 华侨大学, 2018.
- [4] HUANG M M. Study on coal gangue image recognition method based on convolutional neural network [D]. Huainan: Anhui University of Science and Technology, 2020.
黄曼曼. 基于卷积神经网络的煤矸石图像识别方法研究[D]. 淮南: 安徽理工大学, 2020.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [6] 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), June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 11531-11539.
- [7] STERGIOU A, POPPE R, KALLIATAKIS G. Refining activation downsampling with Softpool[EB/OL]. (2021-01-02)[2022-03-03]. <http://arxiv.org/abs/2101.00440>.
- [8] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 1577-1586.
- [9] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]. Cham: Springer, 2018: 3-19.
- [10] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8): 2011-2023.
- [11] CAO Y, XU J, LIN S, et al. GCNET: non-local networks meet squeeze-excitation networks and beyond [C]//IEEE/CVF International Conference on Computer Vision Workshops, October 27-28, 2019, Seoul, Korea (South). New York: IEEE, 2019: 1971-1980.
- [12] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7794-7803.
- [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2022-03-03]. <http://arxiv.org/abs/1409.1556>.
- [14] SHAH P, EL-SHARKAWY M. R-MnasNet: reduced MnasNet for computer vision [C]//2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), September 09-12, 2020, Vancouver, BC, Canada. New York: IEEE, 2020: 1-5.
- [15] SU X, YOU S, HUANG T. Data agnostic filter gating for efficient deep networks[EB/OL]. (2020-10-28) [2022-03-03]. <http://arxiv.org/abs/2010.15041>.
- [16] LIU Z, SUN M, ZHOU T. Rethinking the value of network pruning[EB/OL]. (2018-10-11) [2022-03-03]. <http://arxiv.org/abs/810.05270>.
- [17] HE Y, ZHANG X, SUN J. Channel pruning for accelerating very deep neural networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 1398-1406.
- [18] ZHANG X, ZHOU X, LIN M. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 6848-6856.
- [19] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 4510-4520.
- [20] SHEN C, WANG X, YIN Y, et al. Progressive network grafting for few-shot knowledge distillation [EB/OL]. (2020-12-09) [2022-03-03]. <http://arxiv.org/abs/2012.04915>.
- [21] CHEN J, LU Z, XUE J H. XSepConv: extremely separated convolution [EB/OL]. (2020-02-27) [2022-03-03]. <http://arxiv.org/abs/2002.12046>.
- [22] LIU Y, WENTZLAFF D, KUNG S Y. Rethinking class-discrimination based CNN channel pruning[EB/OL]. (2020-04-29) [2022-03-03]. <http://arxiv.org/abs/2004.14492>.

作者简介:

贾晓芬 (1978—), 女, 博士, 教授, 硕士生导师, 主要从事人工智能、机器学习方面的研究。