

基于改进残差网络的混凝土砂石骨料种类识别研究

曹庆园, 朱建鸿

引用本文

曹庆园, 朱建鸿. [基于改进残差网络的混凝土砂石骨料种类识别研究](#)[J]. 计算机科学, 2024, 51(11A): 231000082-6.

CAO Qingyuan, ZHU Jianhong. [Study on Identification of Concrete Sand and Gravel Aggregate Types Based on Improved Residual Network](#) [J]. Computer Science, 2024, 51(11A): 231000082-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[MB-ATMK:融合属性权重和时序元知识的多行为序列推荐模型](#)

MB-ATMK: Multi-behavior Sequential Recommendation Integrating Attribute Weights and Temporal Meta-knowledge

计算机科学, 2024, 51(11A): 231100047-9. <https://doi.org/10.11896/jsjcx.231100047>

[FCTNet:基于双城深度学习的公交车到站时间预测方法](#)

FCTNet: Bus Arrival Time Prediction Method Based on Dual Domain Deep Learning

计算机科学, 2024, 51(11A): 231000180-7. <https://doi.org/10.11896/jsjcx.231000180>

[基于季节分解的混合神经网络的时间序列预测](#)

Time Series Prediction of Hybrid Neural Networks Based on Seasonal Decomposition

计算机科学, 2024, 51(11A): 231200008-7. <https://doi.org/10.11896/jsjcx.231200008>

[基于相对位置编码转换器模块的深度步态识别网络](#)

Deep Gait Recognition Network Based on Relative Position Encoding Transformer

计算机科学, 2024, 51(11A): 240400064-6. <https://doi.org/10.11896/jsjcx.240400064>

[基于改进超像素采样的立体匹配网络](#)

Stereo Matching Network Based on Enhanced Superpixel Sampling

计算机科学, 2024, 51(11A): 231100005-7. <https://doi.org/10.11896/jsjcx.231100005>

基于改进残差网络的混凝土砂石骨料种类识别研究

曹庆园 朱建鸿

江南大学轻工过程先进控制教育部重点实验室 江苏 无锡 214122

(1249015804@qq.com)

摘要 为解决混凝土砂石骨料复杂种类识别精度低的问题,实现砂石骨料种类自动识别,提出了一种适用于混凝土砂石骨料种类识别的 CM-ResNet18 网络模型。首先采集骨料图像数据集,并采用数据增强方法增加样本以提升模型的鲁棒性;其次选择 ResNet18 模型作为主干网络,融合 CBAM 模块和 MHSA 模块,以增强模型的特征提取能力;最后增加 Dropout 函数以提高神经网络的泛化性能,同时在训练中引入迁移学习以加快网络收敛速度,并增大最后一层学习率,使其更好地适应训练数据并提高模型性能。实验结果表明,CM-ResNet18 模型在原材料识别中取得了高达 99.09% 的准确率。与其他网络模型 AlexNet, VGG19, EfficientNet, ResNet18, ResNet34 相比,CM-ResNet18 模型在识别准确率、精确率、召回率、F1-score 上均有提高,表明该方法在混凝土砂石骨料识别中具有较高的实用性和可行性。

关键词: 砂石骨料;数据增强;残差网络;注意力机制;迁移学习

中图分类号 TP391

Study on Identification of Concrete Sand and Gravel Aggregate Types Based on Improved Residual Network

CAO Qingyuan and ZHU Jianhong

Key Laboratory of Advanced Process Control for Light Industry, Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract In order to solve the problem of low accuracy of identification of complex types of concrete sand and gravel aggregates and realize the automatic identification of sand and gravel aggregate types, a CM-ResNet18 network model suitable for the identification of concrete sand and gravel aggregate types is proposed. Secondly, the ResNet18 model is selected as the backbone network, the CBAM module and the MHSA module are fused to enhance the model's ability to extract features, and then the Dropout function is added to improve the generalization performance of the neural network, and transfer learning is introduced into the training to accelerate the network convergence speed, and the last layer learning rate is increased to better adapt to the training data and improve the model performance. Experimental results show that the CM-ResNet18 model achieves an accuracy of up to 99.09% in the identification of raw materials. Compared with other network models AlexNet, VGG19, EfficientNet, ResNet18 and ResNet34, the CM-ResNet18 model has improved the recognition accuracy, precision, recall rate and F1-score, and the results show that the method has high practicability and feasibility in the identification of concrete sand and gravel aggregates.

Keywords Sand and gravel aggregate, Data augmentation, Residual network, Attention mechanisms, Transfer learning

1 引言

在混凝土自动化生产过程中,各砂石骨料需要通过旋转喂料机和机械臂进入不同的配料仓。为确保砂石骨料正确进入配料仓,各混凝土砂石骨料在入仓初期需要进行种类识别。目前在大部分企业混凝土生产过程中,主要依靠巡检员通过人工对入仓砂石骨料进行种类判定。然而由于砂石骨料种类繁多,人工识别耗时高,增加了混凝土加工企业的人工成本。因此,研究一种适应实际生产需求的砂石骨料种类自动识别方法十分必要。

目前,利用数字处理技术和机器视觉技术对砂石骨料进行识别的研究逐步展开^[1], Li^[2] 利用图像处理技术对混凝土粗骨料的长宽比系数进行提取,实现了对混凝土粗骨料的

检测。Liu 等^[3] 利用数字图像处理技术分割出岩石颗粒与孔隙,从而提取岩石颗粒的轮廓信息,这对砂石骨料识别具有借鉴意义。卷积神经网络(Convolutional Neural Network, CNN)因其具有权值共享和局部视野的特点,具有很强的鲁棒性,被广泛应用于图像识别^[4]。Ceng 等^[5] 通过 SqueezeNet 网络模型进行训练,在较短时间内实现了岩石薄片图像分类,并取得了较好的分类效果。Yang 等^[6] 通过大量砂石骨料物理属性数据,就砂石骨料品质问题展开了研究,提出了基于卷积神经网络的砂石骨料分类模型。Yu 等^[7] 提出了一种能够直接从破碎工序上识别骨料成分的算法模型,采用遗传算法(GA)与支持向量机(SVM)相结合的算法,对破碎、筛分后的人工砂石骨料物理特征量进行训练,建立的 GA-SVM 模型具有较强的分类效果。Yan 等^[8] 提出了一种基于卷积神经网络

基金项目:国家自然科学基金(61973139)

This work was supported by the National Natural Science Foundation of China(61973139).

通信作者:朱建鸿(zhu0012008@163.com)

的砂石骨料图像分类模型 CNN13, 该分类模型参考经典卷积神经网络 VGG16 (Visual Geometry Group 16) 模型进行网络结构的改进和参数优化, 实现对每个等级的砂石骨料进行分类, 且分类准确率较高。Tuan 等^[9] 开发了一种用于粗集料颗粒识别和形状分析的自动立体视觉检测系统 (SVIS), 将深度学习和立体视觉技术结合起来, 计算单元转换因子和每个粒子的厚度, 以方便粒子形状分析, 该系统弥补了图像处理技术的不足, 极大地辅助了现场骨料检测的决策过程。

上述方法介绍了利用图像处理技术和深度学习的方法实现对砂石骨料的识别, 但是这些方法仅能识别特征区别明显的砂石骨料种类, 在混凝土生产过程中, 各种砂石骨料中会存在材料化学属性一致, 但密度不同 (即归为不同种类砂石骨料) 的情况因其密度不同会导致生产的混凝土性能差异, 针对数据集中存在相同材料但不同密度的复杂种类场景, 其识别准确率亟待改善。为进一步提高混凝土砂石骨料的分类识别精度, 本文提出了对经典残差网络 (Residual Network18, ResNet18) 进行改进的新方法, 以 ResNet18 模型为基础模型, 融入 CBAM 和 MHSA 注意力机制, 增加 Dropout 函数, 在模型训练过程中引入迁移学习机制并增大最后一层学习率。实验结果表明, 该方法可以更好地适应砂石骨料图像的特点, 明显提高砂石骨料图像的识别准确率。

2 数据集构建

2.1 研究对象

本文以进入料仓前混凝土各砂石骨料为识别对象, 在混凝土生产企业收集共 9 种混凝土砂石骨料, 即 3.95 粗籽 (名称前面数字表示密度)、3.95 矿粉、4.3 粗籽、4.3 矿粉、4.4 铁精粉、5.0 铁精粉、细石、黄沙、青石料, 混凝土砂石骨料样品如图 1 所示。



图 1 混凝土砂石骨料样本

Fig. 1 Sample of concrete sand and gravel aggregate

2.2 图像采集与预处理

实验中采用的是海康机器人公司的 MV-CA004-10UC 进阶型面阵相机, 相机采用 Sony 的 IMX287 CMOS 芯片, 该工业相机拍摄的图像分辨率为 720 像素 \times 540 像素。使用海康机器人的工业相机进行混凝土砂石骨料各类别图像的采集, 各类混凝土砂石骨料图像如图 2 所示。

将采集到的图像按约 7:3 的比例划分为训练集和测试集, 其中训练集中的图像通过随机水平和垂直翻转图像、随机调整图像的亮度、对比度和饱和度以及添加随机噪声或模糊等十余种方法来扩充数据集以满足数据多样性, 经过数据处理后的砂石骨料数据集如表 1 所列。

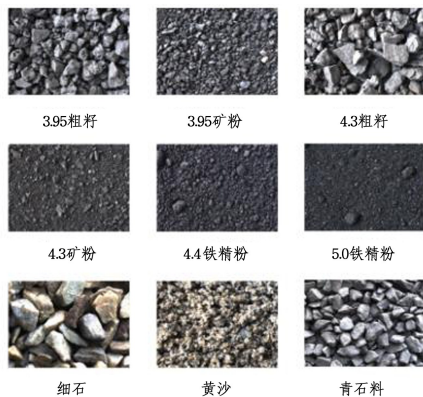


图 2 部分混凝土砂石骨料图像

Fig. 2 Partial image of concrete sand and gravel aggregate

表 1 数据集展示

Table 1 Dataset display

骨料类别	训练集	测试集	数据总数
3.95 粗籽	1155	252	1407
3.95 矿粉	1179	252	1431
4.3 粗籽	1144	252	1396
4.3 矿粉	1172	252	1424
4.4 铁精粉	1198	252	1450
5.0 铁精粉	1179	252	1431
细石	1160	252	1412
黄沙	1198	252	1450
青石料	1195	252	1447

3 建立识别模型

3.1 ResNet18 模型

为保证砂石骨料分类任务的高准确率和实时性, 选取网络层数适中、收敛速度快、能够平衡训练速度和网络深度的 ResNet18 作为主干网络, 在保证砂石骨料特征提取能力的同时使网络有较快的训练速度和推理速度。ResNet18 是由 He^[10] 等提出的一种深度神经网络模型。它的主要特点是引入了残差连接和深度残差学习。这两个关键概念的引入可以解决神经网络中的梯度消失和梯度爆炸问题。残差结构如图 3 所示。

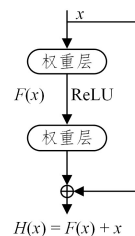


图 3 残差结构图

Fig. 3 Residual structure diagram

图 3 中, x 表示输入, $H(x)$ 表示输出, $F(x)$ 表示残差。传统的 CNN 尝试拟合输入 x 与输出 $H(x)$, 而残差网络结构则尝试学习 $F(x) + x$, 而不是直接学习输入与输出之间的潜在映射 $x \rightarrow H(x)$ 。残差单元在接收输入数据后, 如果一个浅层网络的准确率已经达到饱和状态, 则在其后面添加几个恒等映射层, 使得输出等于输入。通过增加网络的深度但不增加误差, 使得神经网络的层数可以超过之前的限制, 并提高了识别准确性。

3.2 CBAM 注意力机制

CBAM 注意力机制 (Convolutional Block Attention Module, CBAM) 由两个子模块组成: 通道注意力机制 (Channel Attention Module, CAM) 和空间注意力机制 (Spatial Attention Module, SAM)^[11]。CBAM 注意力机制结合了通道注意力

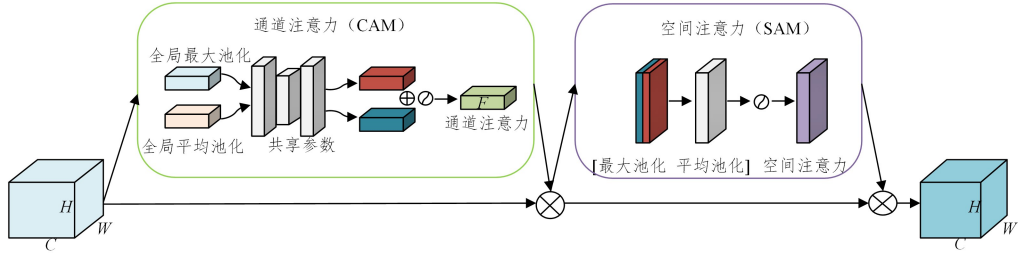


图 4 CBAM 注意力机制

Fig. 4 CBAM attention mechanism

CBAM 的模型结构如图 4 所示, 通道注意力 (CAM) 模块将输入的特征图 $F(H \times W \times C)$ 分别经过基于宽和高的全局最大池化和全局平均池化, 得到两个 $1 \times 1 \times C$ 的特征图, 接下来使用共享的隐藏多层感知器 (MLP) 对它们进行处理, 之后再送入两个全连接层, 最后使用 sigmoid 激活函数输出得到权重系数。将权重系数与输入特征图相乘加权到之前的通道上, 更新各通道的特征权重, 即式(1)中的 $M_c(F)$:

$$M_c(F) = \sigma\{W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))\} \quad (1)$$

其中, F_{avg}^c 为特征图经过全局平均池化的结果, F_{max}^c 为经过全局最大池化的结果, σ 指代 sigmoid 函数, 它是一种常用于深度学习中的激活函数, W_0 和 W_1 是共享感知机的训练权重。

空间注意力 (SAM) 模块将通道注意力模块输出特征图作为本模块的输入特征图。首先进行一个基于通道的全局最大池化和全局平均池化, 得到两个尺寸为 $H \times W \times 1$ 的特征图, 然后将这两个特征图基于通道做拼接操作。然后经过一个 7×7 卷积操作, 降维为 $H \times W \times 1$, 再经过 sigmoid 函数生成空间注意力特征, 即式(2)中的 $M_s(F)$:

$$M_s(F) = \sigma\{f^{7 \times 7}([F_{avg}^s; F_{max}^s])\} \quad (2)$$

其中, σ 指代 sigmoid 函数, f 代表卷积操作, $[]$ 代表通道拼接操作, F_{avg}^s 为特征图经过全局平均池化的结果, 接着将该向量和该模块的输入特征图做乘操作, 得到最终生成的特征。

3.3 多头自注意力机制

为了进一步有效地提取丰富的砂石骨料细节特征, 多头自注意力机制 (Multi-Head Attention, MHSA) 是一种通过并行计算多个独立的注意力头来捕捉输入序列中的关系和依赖性的神经网络模块^[14], 如图 5 所示, 其中每个头都会对输入序列进行注意力计算, 最后通过线性变换和拼接操作将各个注意力头的输出进行整合, 得到最终的表示。

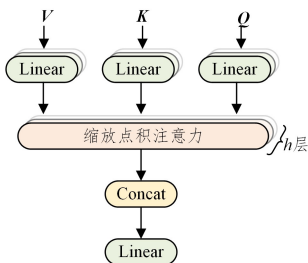


图 5 MHSA 注意力机制

Fig. 5 MHSA attention mechanism

和空间注意力, 能够同时关注特征图的通道维度和空间维度^[12]。在 ResNet18 骨干网络基础上, 引入了通道注意力与空间注意力模块, 以便在砂石骨料图像特定通道中和不同区域上放大有用信息, 抑制不相关信息, 这有助于网络更好地捕捉到砂石骨料图像中的重要特征^[13]。

多头自注意力的关键结构是 MHSA 自注意力层, MHSA 由多个独立的自注意力层组成。每个头都有独立的查询 Q 、键 K 和值 V 向量, 并且这些向量都会被映射到可学习权重矩阵 W_i^Q, W_i^K, W_i^V :

$$Q_i = XW_i^Q \quad (3)$$

$$K_i = XW_i^K \quad (4)$$

$$V_i = XW_i^V \quad (5)$$

这可将输入转换到不同的表示空间中, 且不同权重矩阵之间相互独立, 将 Q_i, K_i, V_i 输入到第 i 个并行头中, 并计算向量之间的点积相似度, 计算过程如式(6)所示:

$$Z_i = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i, i = 1, 2, \dots, h \quad (6)$$

将 Q_i, K_i, V_i 通过除以可学习参数 $\sqrt{d_k}$, 来保证点积结果范围不变, 并且减小梯度变化的影响, 更好地进行反向传播和优化。将得到的相似度通过 softmax 函数进行归一化, 从而得到每个头的注意力权重分布, 用于对值向量进行加权求和, 如式(7)所示:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(Z_0, Z_1, \dots, Z_h) W^o \quad (7)$$

式(7)中将所有自注意力层的输出合并, 并与权重矩阵 W^o 相乘, 得到最终的 MHSA 输出, MHSA 从多个角度提取特征之间的相关性, 得到了更加丰富全面的特征信息。

3.4 改进识别模型

本文提出了一种新的 CM-ResNet18 (CBAM and MHSA-ResNet18) 识别模型, 基于 ResNet18 的改进网络模型综合了 CBAM 和 MHSA 注意力机制, 这样的改进能够充分利用 CBAM 和 MHSA 的优势, 在图像分类任务中进一步提升性能, 使网络更具表达能力和泛化能力。改进的网络模型如图 6 所示, 图中虚线框表示在第三层卷积层后加入 MHSA 注意力机制、在第四层卷积层后加入 CBAM 和 MHSA 注意力机制, 有助于网络更好地捕捉到砂石骨料图像中的重要特征, 用以增强模型性能; 添加 Dropout 函数用于在训练过程中随机地将某些神经元的输出置为零, 从而增加模型对输入的鲁棒性, 使得模型对噪声或变化更具适应性, 减少过拟合; 同时在训练过程中增加最后一层学习率, 以加快对最后一层权重的更新, 使其适应训练数据并提高模型性能。将更多的注意力集中在对最后一层的权重进行调整, 从而更好地适应识别任务的特定需求。

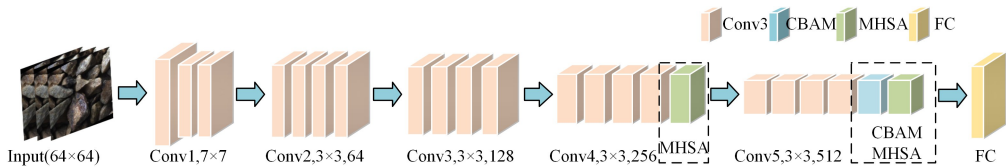


图 6 CM-ResNet18 模型

Fig. 6 CM-ResNet18 model

3.5 实验环境与评价指标

本文模型训练的实验硬件环境如下: Intel 酷睿 i5-11320H 处理器, 运行内存为 16 GB, CPU 频率为 3.2 GHz, 搭载 NVIDIA GeForce MX450 GPU。软件环境基于 Win11 系统, Anaconda 3 平台, 通过 Python 语言在 Pytorch 深度学习框架中搭建并运行网络模型, 卷积神经网络模型的训练效果与模型参数设置密切相关, 设置批大小(batch_size)为 128, 每次训练迭代 50 轮, 并使用 Adam(Adaptive Moment Estimation) 优化器。Adam 优化器具有占用内存小、应用简单、计算效率高等优点^[15]。初始学习率设置为 0.0001, 在图像进入网络训练之前进行预处理和归一化处理, 将其统一为 64 像素 * 64 像素大小。使用迁移学习^[16], 用 ResNet18 网络模型在 ImageNet(ILSVRC-2012)数据集上预训练出的权重来初始化当前模型权值, 加速网络收敛, 提高训练效率^[17]。选择准确率、识别时间作为网络模型的主要评价指标, 并结合混淆矩阵计算精确率、召回率、特异度和 F1 值来综合评价网络模型对混凝土各砂石骨料种类的识别能力。

4 研究结果与分析

4.1 有无迁移学习结果对比

本文应用迁移学习的方式为基于模型的迁移, 即重新利用模型中的参数。通过迁移学习用 ResNet18 网络模型在 ImageNet(ILSVRC-2012)数据集上预训练出权重和特征提取能力, 然后在预训练好的 ResNet18 网络中添加 CBAM 和 MHSAs 注意力机制, 通过在砂石骨料数据集上进行训练来调整模型以适应本文分类识别任务。通过一组有无迁移学习对比实验, 证实了迁移学习的优势。

实验结果如图 7 所示, x 轴代表训练次数, 左侧 y 轴代表准确率, 右侧 y 轴代表损失值。这 4 条折线中, 蓝色代表“有迁移学习准确率”, 绿色代表“有迁移学习损失值”, 红色代表“无迁移学习准确率”, 黄色代表“无迁移学习损失值”。

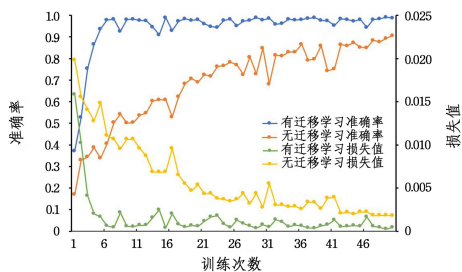


图 7 有无迁移学习的实验结果

Fig. 7 Experimental results with or without transfer learning

经过 50 轮迭代, 图中显示具有迁移学习的 CM-ResNet18 模型在准确率和损失值上始终优于无迁移学习的结果, 并且网络收敛速度相比无迁移学习模型明显加快, 训练效率得以提升。这表明具有迁移学习的 CM-ResNet18 网络

模型不但可以提高识别率和测试精度, 同时收敛速度也有明显提高, 无需从头开始训练新模型, 可以节省大量的时间和计算资源。

4.2 不同网络模型对比

为了验证本文模型改进的有效性, 将本文研究的 CM-ResNet18 网络模型与常见用于图像分类的 ResNet18、ResNet34、AlexNet^[18]、VGG19^[19] 和 EfficientNet^[20] 网络模型在混凝土砂石骨料数据集上的训练结果进行对比。训练前, 将各模型参数调整为一致, 在相同的训练环境下进行 50 次迭代。各个模型测试准确率、精确度、召回率、F1-score 和识别时间等性如表 2 所列。

表 2 不同网络模型的性能对比

Table 2 Performance comparison of different network models

模型	准确率/%	精确率/%	召回率/%	F1-score/%	时间/ms
AlexNet	90.71	91.36	90.68	90.80	14.2
VGG19	88.00	88.41	87.80	87.87	29.5
EfficientNet	97.83	97.98	97.85	97.82	19.1
ResNet18	96.50	96.71	96.50	96.52	15.3
ResNet34	97.47	97.57	97.60	97.44	18.0
CMResNet18	99.09	99.10	98.80	98.93	15.8

对比各网络模型的测试准确率发现, 各卷积网络模型在混凝土砂石骨料种类数据集上的最终识别准确率都在 88% 以上。CM-ResNet18 网络模型的测试准确率略高于 ResNet34 和 ResNet18 网络模型, 50 次迭代训练后 CM-ResNet18 网络模型的测试集识别准确率可达到 99.09%, 比经典 ResNet18 模型的测试准确率高 2.59%, 较更深层次的 ResNet34 网络模型高 1.62%, 比 EfficientNet 模型高了 1.26%。VGG19 模型的准确率仅为 88%, 比 CM-ResNet18 网络模型的识别准确率低 11.09%, AlexNet 模型最终的识别准确率仅为 90.71%, 比 CM-ResNet18 网络模型的识别准确率低 8.38%。综上, 与其他模型相比, CM-ResNet18 模型在混凝土砂石骨料数据集上的识别效果最佳。

得到上述最优性能的原因在于在 ResNet18 中加入 CBAM 和 MHSAs 机制可以带来以下的优势: MHSAs 注意力机制能够从不同的角度观察混凝土各个砂石骨料图像内容, 而 CBAM 注意力机制能够关注特征图中不同通道和空间位置的重要性。通过引入这些注意力机制, 模型可以从多个角度学习和融合信息, 提高了对混凝土砂石骨料图像的理解能力。这样的结合使得模型能够同时考虑到局部和全局信息, 提高了对图像的感知能力。

当对每个模型的识别时间进行分析时, 由表 2 可知 AlexNet 识别时间最短, 仅为 14.2 ms; VGG19 识别时间为 29.5 ms, 相对较长; EfficientNet 识别时间为 19.1 ms; ResNet18 识别时间为 15.3 ms, 接近于 AlexNet 的速度; ResNet34 识别时间为 18.0 ms, 相对较快; CM-ResNet18 识别时间为 15.8 ms, 速度快且性能优秀。这说明其性能提升的同时

并未大幅增加复杂度,体现了模型改进的有效性。综合来看,CM-ResNet18在这组数据中表现最优,拥有最高的准确率、精确度、召回率和F1-score值,并且在识别时间方面也表现良好。

4.3 消融实验

为了验证CM-ResNet18模型中CBAM和MHSA模块的有效性,在相同数据集上进行了消融实验,实验结果如表3所列。

表3 消融实验结果
Table 3 Ablation experiment results

模型	准确率	精确率	召回率	F1-score
ResNet18	96.50	96.71	96.50	96.52
ResNet18+CBAM	97.66	97.84	97.65	97.66
ResNet18+CBAM+MHSA (仅添加在第四层)	98.36	98.52	98.35	98.34
CM-ResNet18	99.09	99.10	98.80	98.93

随着各个模块的加入,模型的准确率逐渐提升,其中ResNet18+CBAM比ResNet18的分类准确率提高1.16%,表明CBAM添加到ResNet18中的空间和通道注意力机制可以帮助模型更好地理解输入图像,并提升其特征提取能力;ResNet18+CBAM+MHSA(仅添加在第四层)比仅添加CBAM机制的准确率提高了0.7%,表明MHSA通过全局视野,模型可以更好地把握图像的全局特征和关联,有助于提升模型的性能;最后,CM-ResNet18与ResNet18+CBAM+MHSA(仅添加在第四层)进行对比,准确率、精确率、召回率和F1-score分别有0.73%,0.58%,0.45%和0.59%的提升,这表明CM-ResNet18中多层的MHSA可以逐渐扩大感受野并捕捉到更复杂的图像信息,与CBAM机制结合从而增强了模型的特征提取能力和表达能力,使得准确率得以提高。

4.4 混凝土砂石骨料种类识别结果的混淆矩阵

混淆矩阵可以用于可视化地展示分类模型的性能,是评判模型结果的重要指标之一,常用于评判分类器模型^[21]。本文研究的CM-ResNet18网络模型使用测试集测试模型的效果,得到混淆矩阵如图8所示,预测值密集分布在对角线上,说明该模型的分类非常准确。观察预测错误的样本分布可知,个别3.95矿粉和4.3矿粉出现预测错误,可能是因为数据集中个别3.95矿粉和4.3矿粉的颜色和粉碎粒度等特性都趋于一致;以及个别4.4铁精粉和4.3矿粉的颜色极为相近。在特征提取过程中易混类特征差异极小,难以区分,因此容易出现识别错误。

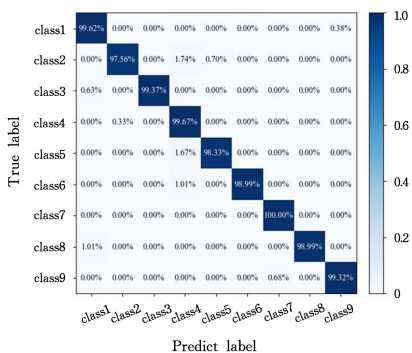


图8 混凝土砂石骨料测试集混淆矩阵

Fig. 8 Concrete sand aggregate test set confusion matrix

总体来说,本文提出的CM-ResNet18模型在采集的混凝土砂石骨料数据集上的分类性能优秀,能够较好地对混凝土砂石骨料种类进行区分,准确地对混凝土砂石骨料种类进行判别。

CM-ResNet18模型混凝土砂石骨料各种类识别结果如表4所列。

表4 CM-ResNet18模型混凝土砂石骨料种类识别结果
Table 4 Identification results of concrete sand and gravel aggregate by CM-ResNet18 model

种类	精确率	召回率	特异度	F1-score
3.95粗籽	98.13	99.62	99.79	98.87
3.95矿粉	99.64	97.56	99.95	98.59
4.3粗籽	100	99.37	100	99.68
4.3矿粉	95.82	99.67	99.44	97.70
4.4铁精粉	99.33	98.33	99.91	98.83
5.0铁精粉	100	98.99	100	99.49
细石	99.33	100	99.91	99.66
黄沙	100	98.99	100	99.49
青石料	99.66	99.32	99.95	99.49

由表4可知,本文模型的召回率、F1值、特异度均很高,9种混凝土砂石骨料分类特异度均在99%以上。以F1-score为标准对其分类性能排序的结果为4.3粗籽、细石、5.0铁精粉、黄沙、青石料、3.95粗籽、4.4铁精粉、3.95矿粉、3.95矿粉,其中3.95矿粉的识别效果略有不足。由表2可知,测试集的识别准确率达99.09%,精确率达99.10%,召回率达98.80%,F1-score达98.93%。这表明CM-ResNet18模型在混凝土砂石骨料种类识别任务上表现优异。

结束语 为有效解决混凝土原材料种类的识别问题,本文提出了CM-ResNet18模型,该模型选择ResNet18模型作为主干网络,融合CBAM模块和MHSA模块以增强模型对特征的提取能力,增加Dropout函数以提高神经网络的泛化性能。本文首先对混凝土砂石骨料原始数据集进行了充分的扩充,随后将CM-ResNet18模型与迁移学习策略相结合,使得模型能够在相对较少的训练迭代次数内达到较高性能水平,从而显著缩短训练时间。本文将训练得到的CM-ResNet18模型的识别结果与当前广泛使用的其他分类网络模型进行了全面对比。结果显示,CM-ResNet18网络模型对混凝土砂石骨料图像的识别率可达到99.09%,验证了CM-ResNet18网络模型的可行性。本文对于实现砂石骨料种类的自动识别具有重要的应用价值。通过快速、准确地识别混凝土砂石骨料的种类,可以提高混凝土生产的效率和质量控制水平。

参考文献

- [1] FANG H Y, YANG J B, HUANG W J, et al. A kind of detection device and method for aggregate morphological quality; China, Fujian; CN106969708A[P]. 2017-07-21.
- [2] LI X. Study on aspect ratio detection method of concrete coarse aggregate based on image processing[J]. Shanxi Architecture, 2018, 44(5): 103-105.
- [3] LIU C, XU Q G, SHI B, et al. Digital image recognition method and application of rock particle and pore system[J]. Chinese Journal of Geotechnical Engineering, 2018, 40(5): 925-931.
- [4] XIONG C Y, WANG T. Image recognition of building compo-

- nents based on convolutional neural network[J]. *Computer Science*, 2021, 48(S1):51-56.
- [5] CHENG G J, LI B, WAN X L, et al. Research on image classification of rock slices based on SqueezeNet convolutional neural network[J]. *Mineralogy and Petrology*, 2021, 41(4):94-101.
- [6] YANG S R, SU M, WANG H. Sand and gravel aggregate classification model based on convolutional neural network[J]. *Sichuan Cement*, 2019, 280(12):15-16.
- [7] YU L J, TONG X, SHEN G L, et al. Classification method of sand and gravel aggregate using GA-SVM model[J]. *Journal of Huaqiao University(Natural Science Edition)*, 2020, 41(2):137-141.
- [8] YAN R, LIAO J D, WU X Y, et al. Research on sand and gravel aggregate classification method based on convolutional neural network [J]. *Progress in Laser & Optoelectronics*, 2021, 58(20):211-218.
- [9] TUAN N M, KIM Y, LEE J Y, et al. Automatic stereo vision-based inspection system for particle shape analysis of coarse aggregates[J]. *Journal of Computing in Civil Engineering*, 2022, 36(2):04021034.
- [10] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[J]. *CoRR*, 2015, abs/1512.03385.
- [11] FENG K, CHASPARI T. A Review of Generalizable Transfer Learning in Automatic Emotion Recognition[J]. *Frontiers in Computer Science*, 2020, 2:9.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778.
- [13] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]// *Proceedings of the European Conference on Computer Vision(ECCV)*. 2018:3-19.
- [14] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. *arXiv:2010.11929*, 2020.
- [15] LYU C, HEYER P, GOOSSENS B, et al. An unsupervised transfer learning framework for visible-thermal pedestrian detection[J]. *Sensors*, 2022, 22(12):4416.
- [16] KINGMA D, BA J. Adam: a method for stochastic optimization [C]// *International Conference on Learning Representations (ICLR)*. 2015:1-15.
- [17] DU Z W, ZHOU H, LI C Y, et al. Review of small target detection algorithms for deep convolutional neural networks [J]. *Computer Science*, 2022, 49(12):205-218.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// *Proceedings of the 26th Conference on Neural Information Processing Systems(NIPS)*. Lake Tahoe, USA, 2012:1106-1114.
- [19] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition. [J]. *CoRR*, 2014, abs/1409.1556.
- [20] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[C]// *International Conference on Machine Learning*. PMLR, 2019:6105-6114.
- [21] BANAN A, NASIRI A, TAHERI-GARAVAND A. Deep learning-based appearance features extraction for automated carp species identification[J]. *Aquacultural Engineering*, 2020, 89:134-139.



CAO Qingyuan, born in 1998, master's candidate. His main research interests include machine vision and computer intelligent control system research.



ZHU Jianhong, born in 1964, associate professor, master tutor. His main research interests include Internet of Things and industrial automation intelligent control system.