

DOI:10.19651/j.cnki.emt.2211744

基于改进 MobileNetV3 烧结断面火焰图像识别^{*}

梁秀满 安金铭 曹晓华 曾凯 王福斌

(华北理工大学电气工程学院 唐山 063210)

摘要: 烧结机尾断面火焰图像蕴含大量与烧结终点相关的特征信息,充分利用烧结火焰图像特征信息进行在线判断烧结终点状态,具有可行性及工程实际意义。针对烧结机尾断面火焰图像特征信息难以提取、识别精度低以及难以满足实时性等问题,提出一种基于改进的 MobileNetV3 烧结断面火焰图像识别算法。以 MobileNetV3 作为烧结终点火焰状态特征信息提取的基础模型,引入注意力机制;改进通道注意力结构,减少特征损失提高识别精度;引入空间注意力机制,设计双分支通道空间注意力模块精确捕捉了红火区在烧结断面火焰图像中的位置和内容信息;引入数据增强和余弦退火学习率来提高模型的泛化能力,并采用冻结训练策略加速模型收敛。在烧结火焰数据集上的实验表明,该算法能够充分利用烧结火焰图像中的特征信息,识别准确率达到 97.54%,较改进前提高了 6.41%。

关键词: 双分支通道空间注意力模块; MobileNetV3; 烧结火焰图像; 余弦退火

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Flame image recognition of sintering section based on improved MobileNetV3

Liang Xiuman An Jinming Cao Xiaohua Zeng Kai Wang Fubin

(College of Electrical Engineering, North China University of Science and Technology, Tangshan 063210, China)

Abstract: The flame image of the sintering machine tail section contains many features related to the sintering endpoint. It is feasible and practical to make full use of the feature information of the sintering flame image to judge the state of the sintering endpoint online. Aiming at the problems of difficult extraction of flame image feature information of sintering machine tail section, low recognition accuracy, and difficulty meeting real-time requirements, an improved MobileNetV3 sintering section flame image recognition algorithm is proposed. MobileNetV3 is taken as the basic model for feature extraction of flame state at the sintering endpoint, and the attention mechanism is introduced; it improves the attention structure of the SE channel to solve the problem of weak resolution of features extracted from the original model; The introduction of Spatial Attention (SA) mechanism and the design of Two Branch Channel Spatial Attention (TBCSA) module accurately capture the position and content information of the red fire zone in the flame image of the sintering section; The data enhancement and cosine annealing learning rate are introduced to improve the generalization ability of the model, and the freezing training strategy is used to accelerate the model convergence. The experiment on the sintering flame data set shows that the algorithm can fully use the feature information in the sintering flame image, and the recognition accuracy reaches 97.54%, which is 6.41 percentage points higher than before.

Keywords: two-branch channel spatial attention module; MobileNetV3; sintered flame image; cosine annealing

0 引言

烧结的产品质量与烧结工况息息相关,在工业生产中,烧结终点是决定烧结产品质量的重要因素。目前的烧结终点判定主要依赖于人工看火,由于受看火工经验、疲劳程度等主观因素的影响,导致看火结果不准确^[1]。烧结断面火

焰图像蕴含许多与烧结终点相关的特征信息^[2],因此利用烧结断面火焰图像的特征信息可以判断烧结终点状态。烧结断面火焰图像即烧结工艺过程中料层的燃烧状态,分为欠烧、过烧以及正常燃烧图像,准确识别 3 种烧结断面火焰图像对优化烧结过程、提高炼铁产量和节约燃料资源具有重要工程意义。

收稿日期:2022-10-19

* 基金项目:河北省自然科学基金(E2021209037)、河北省省属高等学校基本科研业务费研究项目(JYG2020004)资助

Wang 等^[3]提出遗传算法优化支持向量机参数的烧结终点预测,Liu 等^[4]根据烧结过程特点提出基于决策树算法和决策规格的烧结终点预测系统。李屹等^[5]将烧结火焰图像红火区、气孔区平均灰度和断面红火区厚度等特征信息,输入到 BP 神经网络对烧结图像进行烧结终点识别。杨春雨等^[6]提出利用 BP 神经网络构建红火层面积、重心和亮度等特征信息判断烧结终点。王福斌等^[7]提出分割烧结火焰图像红火区,对红火区长轴、短轴和离心率等特征信息进行聚类,实现烧结火焰图像分类。上述方法主要依赖于传统机器学习算法模型,需要结合先验知识和人工提取烧结终点相关的特征变量,受限于主观因素,难以实现高可靠和实时判断烧结终点。

相比传统的机器学习方法,深度学习已逐步应用在钢铁产业,包括:烧结火焰图像分类、板坯表面质量检测和铁水包扒渣识别等^[1]。深度学习在图像分类方面取得了飞速的发展,AlexNet^[8]、GoogLeNet^[9]、VGGNet^[10]和 ResNet^[11]等经典 CNN 架构,从分类精度、鲁棒性和泛化性等方面优于传统算法,广泛应用于钢铁生产领域。朱永红等^[12]在卷积神经网络模型嵌入 SE 模块,提高网络对关键特征的注意力,然而该模型只注重通道维度注意力,忽略了火焰空间特征信息对识别的影响。徐超^[13]利用免疫遗传算法全局寻优能力对神经网络模型进行改进,加快网络收敛速度,但是由于遗传算法编程的复杂性,需要编码解码一系列操作,以至于很难解决大规模计算量等问题。基于上述问题,本文提出改进的 MobileNetV3 轻量化网络来实现烧结断面火焰图像识别。MobileNetV3^[14]网络将通道注意力应用于视觉领域,有效提高一些图像的识别精度,但对于烧结断面火焰图像,红火区的空间位置信息对于烧结终点状态识别起到至关重要的作用,因此,结合烧结火焰图像特点,在 MobileNetV3 的基础上引入空间注意力,并与改进的 SE 通道注意力结合提出双分支通道空间注意力模块,使网络模型更加关注红火区的重要通道的特性映射以及所在空间位置的关键特征信息;利用数据增强、余弦退火算法和冻结训练,提高网络的泛化性并加快模型收敛。

1 实验数据

本研究试验自建火焰数据集,采集的图像来源于唐钢 320 m² 烧结机的烧结机尾断面火焰图像,数据集中的类别可分为正常烧结、欠烧和过烧 3 种状态,如图 1 所示,根据烧结料层自上而下的燃烧特点,从烧结机尾断面火焰图像上可以判断出烧结的状态。正常烧结状态所有矿料都进行了燃烧,红火区达到料层底部,背景区域位于料层断面的上方;欠烧状态红火区位于料层断面的上方或者台车的上方到台车的 1/2 处,红火区下方为未燃烧的混合料;过烧状态料层烧透,火焰呈窄条状,红火区上方为燃烧结束的背景区。实验数据为 22 675 张烧结断面火焰图像,数据集各个类别分布情况如表 1 所示。

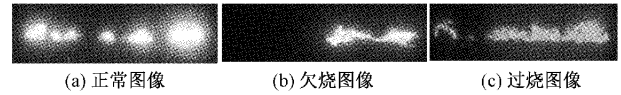


图 1 烧结断面 3 种火焰状态图像

表 1 数据集各类别分类情况

标签	类别	数量
0	正常图像	11 338
1	欠烧图像	4 535
2	过烧图像	6 802

2 烧结火焰识别模型

烧结断面火焰图像的分类主要依赖于红火区的特征信息,然而,烧结现场工艺复杂,受烟气、粉尘影响,采集到的断面火焰图像存在大面积的背景干扰因素,不利于网络模型辨识烧结火焰状态类别。考虑到今后便于部署的工业性要求,本研究选用了识别精度高,并且易于部署的轻量级网络 MobileNetV3 为基准网络。为使该模型适用于烧结断面火焰图像的识别,提高烧结终点状态识别精度,本研究改进 SE(squeeze and excitation)通道注意力,引入空间注意力模块与改进的 SE 通道注意力并行设计双分支通道空间注意力模块(two-branch channel spatial attention, TBCSA),对 MobileNetV3 模型进一步优化改进,据此将研究模型命名为 TBCSA-MobileNetV3,网络模型如图 2 所示。网络模型架构主要由瓶颈结构(blockneck, bneck)结构和尾部结构(Last stage)组成。bneck 结构的主要作用为火焰图像重要红火区的特征提取,由 11 个 bneck 层堆叠而成,将原 MobileNetV3 模型的 SE 注意力替换为 TBCSA 注意力。MobileNetV3 对 Last stage 结构进行了精简,移除较复杂的瓶颈层连接,降低了模型参数量。

2.1 通道注意力模块

通道注意力模块(channel attention block, CAB)为了研究特征映射通道之间的关系,提出改进的 SE 通道注意力机制,如图 3 所示。MobileNetV3 的 SE 通道注意力仅采用平均池化压缩信息如图 3(a)所示,受文献[15]的启发,平均池化方法可以学习到目标物体的程度信息,最大池化能够学习到物体的判别性特征,同时利用最大池化层和平均池化层相结合的方式可以更加精准的推断通道注意力;此外,SE 通道注意力利用全连接层对特征维度进行降维,降维对特征图的信息损失较大,受到 ECA-Net^[16]的启发,利用权重共享的 1D 卷积避免了重要特征信息的损失,有效实现了最大池化与平均池化跨通道交互,如图 3(b)所示,首先通过平均池化和最大池化来聚合特征图的空间信息压缩特征映射为 $1 \times 1 \times C$,然后利用快速 1D 卷积代替全连接层计算通道的重要特征信息的相关性,最后采用 Hard-sigmoid 获得输出特征图。整体运算过程如下:

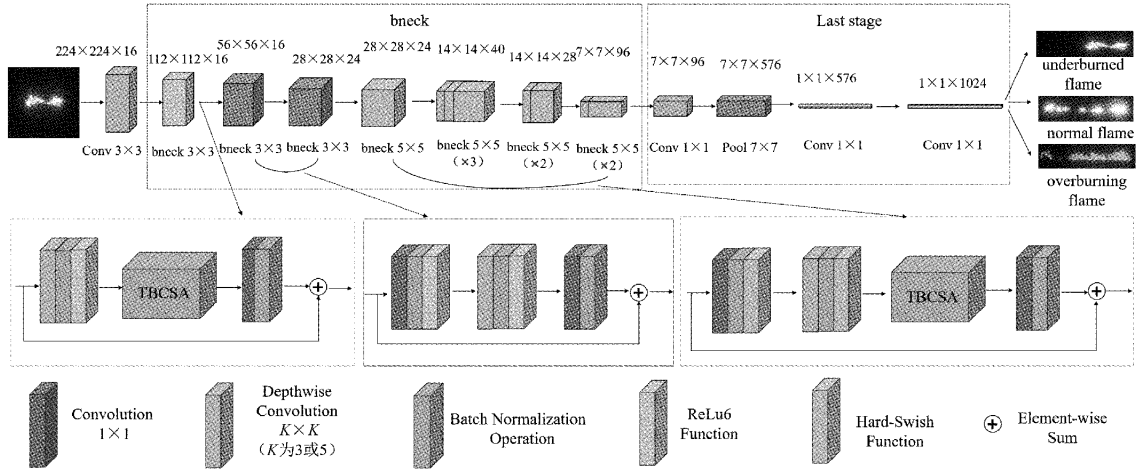


图 2 TBCSA-MobileNetV3 网络结构示意图

$$M_c = \sigma(C_1 D_k (AvgPool(F) + MaxPool(F))) \quad (1)$$

式中： F 表示输入特征图， σ 表示激活函数， $C_1 D_k$ 表示 1D 卷积。

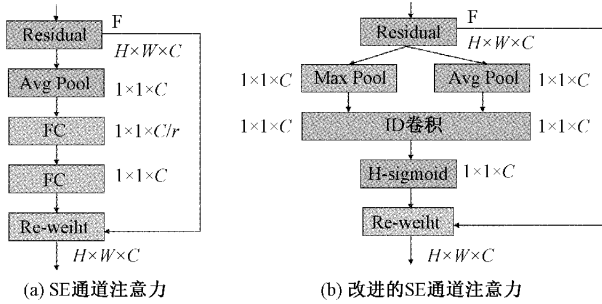


图 3 通道注意力模块

2.2 空间注意力模块

空间注意力模块 (spatial attention block, SAB) 为了研究特征映射空间之间的关系, 采用 $GE^{[17]}$ (Gather-excite) 聚集和激发空间注意力机制, GE 定义了 Gather 算子 ξ_G 和 Excite 算子 ξ_E 。基于 GE 简洁有效的特点, SA 注意力采用无参数的做法, 如图 4 所示。用全局平均池化操作作为聚合算子 ξ_G , 从每个特征图中的每个空间邻域提取出特征, 聚合给定空间范围内的神经元响应, 压缩特征映射为 $1 \times 1 \times C$ 。激发算子 ξ_E 为最近邻插值法, 通过 ξ_G 的值和输入值, 来调整特征图的重要程度, 产生一个与原始输入相同维度的新张量, 最后与输入特征进行点积运算得到输出特征图 $H \times W \times C$ 。

2.3 TBCSA 双分支通道空间注意力模块

为了同时关注烧结断面火焰图像红火区的空间域和通道域信息, 提高烧结状态识别的精度, 受 $DANet^{[18]}$ 网络的启发, 本文提出 TBCSA 双分支通道空间注意力, 包括 CAB 通道注意力模块和 SAB 空间注意力模块, 如图 5 所示。TBCSA 注意力将 CAB 和 SAB 模块并行设计来捕捉空间和通道维度的全局特征依赖关系, 给定一张烧结断面

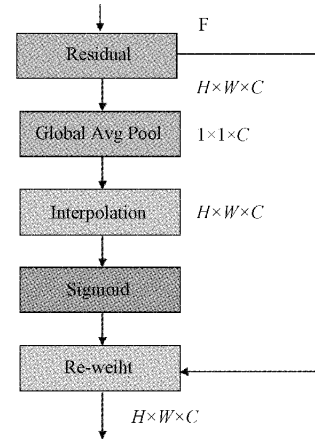


图 4 空间注意力模块

火焰图像, TBCSA 不仅能学习通道特征矩阵纹理、颜色和亮度等有价值的信息, 还能关注火焰红火区在输入特征映射的位置信息, 降低图像背景区域的关注度, 并且将两个注意模块的输出特征相加, 以进一步改进特征表示, 有助于更精确的断面火焰图像分类结果。

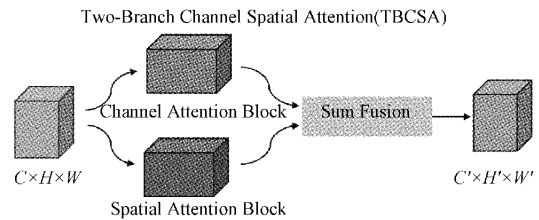


图 5 双分支通道空间注意力模块

3 网络训练策略

3.1 数据增强

烧结断面火焰图像具有多样性、旋转不变性、对称性等特点, 因此可对火焰数据集进行数据增强 (data augmentation, DA), 增加样本的多样性, 提高模型的泛化

能力,减少过拟合。文中采取多种数据增强方式,包括垂直翻转、水平翻转、旋转、随机饱和度、亮度和对比度调整以及这几种方式的随机组合,部分样本的数据增强示例如图 6 所示。

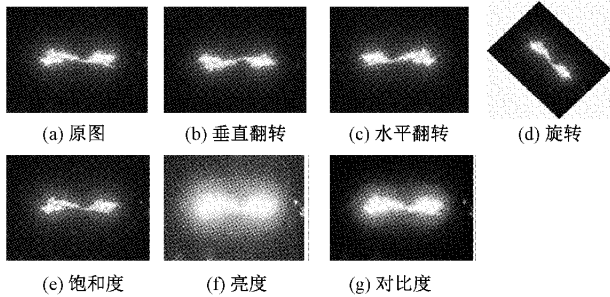


图 6 数据增强示例

3.2 学习率调整策略

学习率控制了基于损失梯度调整神经网络权值的速度,使损失函数收敛到最小值,为了避免损失函数陷入局部最小值情况,引入了余弦退火^[19](cosine annealing, CA)的学习率下降方式。余弦退火的原理如下:

$$\eta_i = \eta_{\min}^i + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i) \left(1 + \cos\left(\frac{T_{cur}}{T_i} \pi\right) \right) \quad (2)$$

其中, η_{\max}^i 和 η_{\min}^i 分别表示学习率的最大值和最小值, T_{cur} 则表示当前执行了多少个 epoch, T_i 表示第 i 次启动总的 epoch 数。

3.3 冻结训练

冻结训练(freeze training, FT)是利用迁移学习思想对网络模型进行微调,迁移学习迁移学习主要是指通过学习某一项任务达到比较好的效果后,再去学习目标任务,从而实现学习能力的继承或者更加快速的进入目标任务的学习。冻结部分不参与权重更新,可以将更多资源放在未冻结部分的网络参数,使得时间和资源利用都能得到较大

改善。

4 实 验

4.1 实验设置

本次实验环境硬件训练平台为:32 G,CPU:Intel(R) Xeon(R)E7,GPU:12 G 显存 NVIDIA-RTX3060Ti;软件平台为:操作系统 Ubuntu20.04、深度学习框架 PyTorch1.8.0 实现网络模型训练、编程语言为 Anaconda3、并行计算架构 CUDA11 和 GPU 加速库 CUDNN8.2.1 等。

4.2 消融实验

为了验证本文算法提出的双分支通道空间注意力和一系列训练策略对火焰图像识别精度的影响,本研究在烧结断面火焰图像数据集上进行消融实验,以 MobileNetV3 为基准网络,分别加入 2.1 节介绍的通道注意力模块,空间注意力模块以及双分支通道空间注意力来验证注意力机制对火焰识别精度的影响,在加入双分支注意力网络的基础上采用 2.2 节介绍的数据增强、余弦退火和冻结训练来评估文中提出的模型训练策略的有效性。数据增强为垂直翻转、水平翻转、随机旋转、随机饱和度、亮度和对比度调整以及这几种方式的随机组合。余弦退火使用 pytorch 导入的 CosineAnnealingWarmRestarts 包,参数设置为 $T_0=5$, $T_{mult}=2$ 。冻结训练采用 3 种方案:1)只冻结 bneck;2)只冻结 Last stage;3)bneck 和 Last stage 轮流冻结,以 5 个 epoch 为一个阶段,第一个 epoch 为 bneck 冻结,只训练 Last stage;第 2~5 个 epoch 为 Last stage 冻结,只训练 bneck。消融实验结果如表 2 所示。训练模型时,模型在验证集上的损失变化曲线和准确率变化曲线如图 7 所示,为验证注意力和训练策略的有效性,仅展示了本文提出的 TBCSA 注意力以及加入组合训练策略最终的损失值和准确率变化曲线。

表 2 TBCSA-MobileNetV3 模型消融试验结果

Model	CAB	SAB	TBCSA	DA	CA	FT		Acc/%	Paras/M
						bneck	Last stage		
	—	—	—	—	—	—	—	91.13	2.54
	√	—	—	—	—	—	—	92.15	2.49
	—	√	—	—	—	—	—	91.96	2.38
	—	—	√	—	—	—	—	94.69	2.59
MobileNetV3	—	—	√	√	—	—	—	95.81	2.59
	—	—	√	√	√	—	—	96.37	2.59
	—	—	√	√	√	√	—	96.74	2.59
	—	—	√	√	√	—	√	96.85	2.59
	—	—	√	√	√	√	√	97.54	2.59

注:CAB代表改进的SE通道注意力,SAB代表空间注意力,TBCSA代表双分支通道空间注意力,DA代表数据增强,CA代表余弦退火学习率,FT代表冻结训练策略,“—”为改进 MobileNetV3 时不使用该策略,“√”表示在 MobileNetV3 时使用该策略。

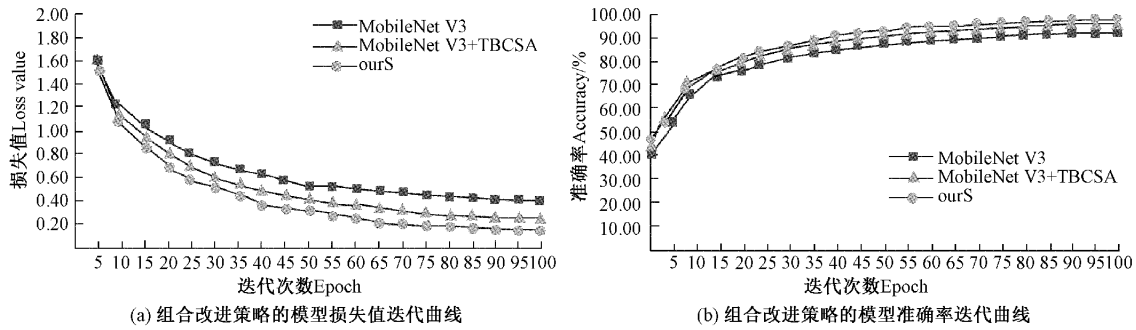


图 7 组合改进策略的模型迭代曲线

通过对比表 2 加入注意力模块后的网络模型对断面火焰图像分类的准确性,可以得出文中提出的通道注意力模块,模型参数量下降为 2.49 M,火焰识别精度提高 1.02%。引入空间注意力模块,火焰识别准确率提高 0.83%。采用本文提出的双分支通道注意力,火焰识别精度提高 3.56 个百分点,模型参数量仅为 2.59 M。综上,与单一通道注意力模块或空间注意力模块相比,混合注意力模块更加有效地提升模型的分类效果。

通过对比一系列训练策略得到的数据可知,采用数据增强的图像预处理策略,模型识别准确率提高了 1.12%,这说明数据增强的训练策略可以增加训练数据,提高模型泛化能力;利用余弦退火调整学习率,使模型准确率提升 0.56%;对比冻结训练的结果,可知方案 1)模型准确率提高 0.37%,方案 2)模型准确率提高 0.48%,方案 3)模型准确率提高 1.17%,所以模型采用冻结交替训练。综上,同时使用数据增强、余弦退火和冻结训练组合训练策略,模型的准确性整体提升了 2.85%。

从图 7 可以看出,加入 TBCSA 注意力,在验证集上的损失值降低,准确率提升,同时采用 TBCSA 注意力以及组合训练策略,损失值更小,识别率更高,表明改进后的模型具有良好的泛化能力。综合上述,本文提出的模型在验证集上的损失值进一步下降,识别准确率进一步上升,取得了良好的表现,证明了本文对 MobileNetV3 一系列改进的有效性。

4.3 不同算法模型的对比分析

为了验证算法的有效性和先进性,使用 AlexNet^[8]、GoogLeNet^[9]、VGGNet^[10]、ResNet^[11]、MobileNet^[14]、EfficientNet^[20] 网络模型以及 SE-Inception-ResNet-V2^[12]、GA-ResNet^[13] 主流网络模型进行对比实验,由表 3 可知,相比经典网络 AlexNet、GoogLeNet 和 VGGNet,其参数量仅是这些网络模型参数的 4.24%、37.05%、1.87%。与经典分类网络(除 MobileNet 和 EfficientNet 外的网络)相比,本实验提出的网络模型取得了较好的火焰识别准确率,并与轻量级网络 MobileNetV2、MobileNetV3 和 EfficientNet 相比火焰识别精度提高 6.66%、6.41% 和 5.49%,模型参数量为 2.59 M,浮点运算量为 0.07 B。这

表明本文采用的组合训练策略在不增加网络模型复杂度的前提下明显提高网络模型分类准确率。同时对比该领域现有的网络模型,相较于 SE-Inception-ResNet-V2^[12] 和 GA-ResNet^[13] 网络模型,识别准确率分别提高 5.18% 和 6.19%。这表明本文提出的双分支通道空间注意力网络模型烧结火焰识别效果优于现有的烧结火焰识别模型,取得了与目前主流方法更优的识别准确率。

表 3 各网络模型的性能对比

模型	模型参数量/ M	浮点运算量 FLOPs/B	准确率/ %
AlexNet ^[8]	61.1	0.72	87.56
GoogLeNet ^[9]	6.99	1.58	86.39
VGGNet ^[10]	138.36	1.05	80.15
ResNet-18 ^[11]	11.69	1.82	89.06
ResNet-50 ^[11]	25.56	4.14	90.70
ResNet-101 ^[11]	44.55	7.87	91.78
EfficientNet ^[20]	5.29	0.42	92.05
MobileNetV2 ^[14]	3.50	0.38	90.88
MobileNetV3 ^[14]	2.54	0.06	91.13
SE-Inception-ResNet-V2 ^[12]	20.38	3.75	92.36
GA-ResNet ^[13]	23.57	4.54	91.35
本文	2.59	0.07	97.54

5 结 论

本研究提出了一种基于改进 MobileNetV3 烧结断面火焰图像识别算法,在 MobileNetV3 模型结构中加入 TBCSA 双分支通道空间注意力模块,不仅能关注火焰的判别性特征信息,还能重点捕捉红火区在图像中的空间位置信息,提高模型的烧结状态识别精度。为了进一步提升模型的性能,采用数据增强、余弦退火和冻结训练的组合训练策略来训练网络模型。该模型在烧结火焰图像数据集上的识别准确率达到 97.54%,相较于一些现有的方法对火焰图像进行识别,识别精度取得了令人满意的结果。

本研究用图像识别在线判断烧结终点,对减轻工人劳动强度、提高系统自动化程度、实现钢铁工业智能制造等具有重大工业应用价值,强化深度学习在钢铁工业的实际应用。

参考文献

- [1] 李江昀,杨志方,郑俊锋,等.深度学习技术在钢铁工业中的应用[J].钢铁,2021,56(9):7.
- [2] 王福斌,刘贺飞,王蕊,等.烧结断面火焰图像多核 Boosting 显著性检测[J].计算机辅助设计与图形学学报,2021,33(9):1466-1474.
- [3] WANG D D, YANG K, HE ZH J, et al. Application research based on GA-FWA in prediction of sintering burning through point [C]. Proceedings of 2018 International Conference on Computer, Communications and Mechatronics Engineering (CCME 2018), 2018:394-401.
- [4] LIU S, LYU Q, LIU X, et al. A prediction system of burn through point based on gradient boosting decision tree and decision rules[J]. ISIJ International, 2019, 59(12):2156-2164.
- [5] 李屹,王莹.基于机器视觉的烧结终点识别方法研究[J].冶金自动化,2012,36(6):10-15.
- [6] 杨春雨,宋宝宇,杨东晓,等.鞍钢烧结生产过程综合智能控制系统[J].冶金自动化,2011,35(6):1-6,20.
- [7] 王福斌,王蕊,武晨.基于改进聚类算法的烧结火焰图像分类[J].激光与光电子学进展,2022,59(2):449-456.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015: 1-9.
- [10] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014, DOI: 10.48550/arXiv.1409.1556.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2016.
- [12] 朱永红,付瑶,李选亮,等.基于改进的卷积神经网络陶瓷梭式窑火焰图像分类方法[J].陶瓷学报,2022,43(2):302-309.
- [13] 徐超.基于深度卷积神经网络的窑内工况燃煤火焰图像识别算法研究[D].长沙:湖南大学,2019.
- [14] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314-1324.
- [15] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany, 2018:3-19.
- [16] WANG Q, WU B, ZHU P, et al. ECANet: Efficient channel attention for deep convolutional neural networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2020: 11531-11539.
- [17] JIE H, LI S, ALBANIE S, et al. Gather-excite: Exploiting feature context in convolutional neural networks[C]. NeurIPS, 2018.
- [18] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA: IEEE, 2019: 3146-3154.
- [19] LOSHCHILOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts [J]. ArXiv Preprint, 2016, ArXiv:1608.03983.
- [20] TAN M, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks [C]. 36th International Conference on Machine Learning (ICML 2019). Los Angeles, USA: Proceedings of Machine Learning Research, 2019: 6105-6114.

作者简介

梁秀满,副教授,硕士生导师,主要从事检测技术与智能装置方面的教学与研究工作。

E-mail:13230857556@126.com