

DOI: 10.13382/j.jemi.B2306473

# 基于轻量化网络和知识蒸馏的回转窑工况识别\*

王欣 邓章俊 秦斌

(湖南工业大学电气与信息工程学院 株洲 412007)

**摘要:**回转窑烧成带图像蕴含丰富火焰信息,准确识别燃烧状态是回转窑优化控制的前提,基于卷积神经网络的方法可以快速识别回转窑工况,提高回转窑自动化水平,但存在网络规模大,所需计算资源多的问题,为此提出了一种基于轻量化网络和知识蒸馏的回转窑工况识别方法,在网络卷积层后引入协方差池化层改进教师模型和学生模型,以改进的轻量化网络 MobilenetV2 作为学生模型的主干网络,以改进的 Resnet50 作为教师模型的主干网络,通过构建混合蒸馏损失函数,将教师模型蕴含的丰富分类标签信息迁移到学生模型中,并将蒸馏训练得到的学生模型作为回转窑工况识别模型,以提高网络对高相似火焰图像的识别精度。实验结果表明,经蒸馏后的学生模型总体识别准确率相较于原始模型提高了 3.33%,对测试集中 3 种工况的识别率分别达到了 93%、99%、90%,准确率和模型规模均优于目前其他主流网络模型,满足实际生产中实时、低成本要求。

**关键词:** 回转窑;火焰图像;工况识别;卷积神经网络;轻量化网络;知识蒸馏

中图分类号: TP391.4;TH69

文献标识码: A

国家标准学科分类代码: 520.2040

## Working condition recognition based on lightweight network and knowledge distillation for rotary kilns

Wang Xin Deng Zhangjun Qin Bin

(College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412007, China)

**Abstract:** The firing zone images of rotary kilns contain rich flame information and accurate combustion state recognition are the premise of optimal control for the rotary kilns. The working conditions can be quickly recognized, and the automation level of the rotary kilns can be improved through the convolutional neural network-based methods, but there are problems with large network size and high computational resources required. Therefore, a working condition recognition method based on lightweight network and knowledge distillation was proposed in this paper. The teacher model and the student model were improved by introducing the collaborator differential layer after the convolutional layer of the network. The improved lightweight network MobilenetV2 was used as the backbone network of the student model while the improved Resnet50 was used as the backbone network of the teacher model. The rich classification label information contained in the teacher model was transferred to the student model by constructing the mixed distillation loss function and the student model obtained by distillation training was used as the working condition recognition model to improve the recognition accuracy of the high similar rotary kiln flame images. The experimental results show that the overall recognition accuracy of the improved student model is increased by 3.33% compared with the original model, and the recognition accuracy of the three working conditions in the test set reaches 93%, 99%, 90% respectively. The accuracy and network size are better than those of other mainstream networks, and the requirements of real time and low cost in actual production process are met.

**Keywords:** rotary kiln; flame image; working condition recognition; convolutional neural network; lightweight network; knowledge distillation

## 0 引言

回转窑是水泥熟料煅烧的关键设备,在生产过程中,准确识别回转窑燃烧状态,对于提高熟料质量合格率,保证回转窑安全运行,降低企业生产成本具有重要意义<sup>[1]</sup>。

窑内煅烧区温度是表征熟料在窑内燃烧状态的主要参考指标,在实际生产过程中,一般通过控制送煤量、风门开度和窑转速等受控变量调节窑内温度<sup>[2]</sup>,然而受回转窑运行环境和设备本身特点等诸多因素限制,一般的测温系统(例如红外辐射系统,热电偶系统和色度计系统)难以准确测量窑内煅烧区温度。目前,对于烧结燃烧状态的识别大部分仍然由人工完成,通过窑头火孔直接观测燃烧情况或通过窑内摄像头间接观察判断,受窑内烟雾粉尘等干扰量影响,容易忽略一些异常火焰特征信息。

回转窑烧成带火焰图像蕴含了丰富的状态信息<sup>[3]</sup>,通过烧成带图像可以识别回转窑的燃烧状态,为解决水泥熟料煅烧过程中的检测问题,提高熟料煅烧质量提供了有效手段<sup>[4]</sup>。Guo 等<sup>[5]</sup>提出了一种基于奇异值分解(singular value decomposition, SVD)和支持向量机(support vector machine, SVM)的回转窑燃烧状态识别方法,使用基于奇异值分解的特征作为输入, SVM 作为分类器识别工况。该方法具有较低时间复杂度,可以实现回转窑工况实时识别,但是火焰图像存在许多冗余信息,仅靠传统机器学习方法不能准确提取火焰特征信息,因此后续有学者将研究重点放在回转窑火焰图像分割上,通过最大类间方差法(Otsu)<sup>[6]</sup>等算法将目标区域从原始火焰图像背景中分离出来,然后再进行特征提取。文献<sup>[7]</sup>提出了一种基于 Otsu-K-means 和 SVM 的水泥回转窑燃烧状态识别方法,首先通过提出的 Otsu-K-means 方法有效分割火焰图像的重要区域,然后再提取 10 个特征参数作为状态识别特征输入到 SVM 进行分类。该方法能更高效地提取火焰特征信息,不足之处在于对于不同工况的回转窑,如果火焰区域不能准确分割,将影响下一步的特征提取,进而影响识别精度。

近年来,深度学习在图像识别领域取得了突破性的进展,不少学者将深度学习应用于回转窑工况识别。Li 等<sup>[8]</sup>使用卷积神经网络进行特征迁移自动识别燃烧状态,与传统机器学习方法相比,该方法无需对火焰图像进行分割或特征参数构建等预处理操作,能够获得更准确、更可靠的自动分类效果,缺点则在于其特征迁移参数未针对回转窑火焰图像优化。栾庆磊等<sup>[9]</sup>基于深度迁移学习方法,通过构建动态卷积神经网络获取火焰图像的特征空间,并将特征向量输入至随机配置网络分类器(stochastic configuration network, SCN)识别燃烧状态,其

优势在于能针对火焰图像动态提取多层次特征空间。Hu 等<sup>[10]</sup>提出了基于迁移学习和注意力机制的回转窑工况识别方法,通过空间注意力机制改善模型对高相似度回转窑火焰图像的识别。由于基于静态图像的识别方法容易受到回转窑复杂工况下各种粉尘导致的噪声影响,文献<sup>[11]</sup>提出了一种基于循环神经网络(recurrent neural network, RNN)和回声状态网络(echo state network, ESN)的识别与预测方法,其优势在于能利用视频图像序列所蕴含的时空特征获得更好的工况识别效果,但其在工况预测问题上仍需进一步改进。

基于深度学习的回转窑工况识别方法多依赖于高性能的复杂网络模型,其参数量大,所需计算资源多,不能很好满足实际生产中对工况识别低延迟、低成本的要求。为了解决传统卷积神经网络模型难以在低算力设备上运行问题,轻量化的网络模型应运而生<sup>[12-14]</sup>,它具有网络参数量小、计算延迟低的特点,但识别精度与复杂模型相比还有待提高。

知识蒸馏<sup>[15-16]</sup>作为一种新兴的模型压缩方法,可以有效改善轻量化网络性能。它是一种教师-学生(Teacher-Student)训练结构,能以轻微的性能损失为代价将复杂模型的知识迁移到简单模型中。

综上所述,受轻量化模型与知识蒸馏算法启发,论文基于轻量化网络并结合知识蒸馏算法,提出了一种针对水泥回转窑的工况识别方法。首先,选取先进的复杂网络作为教师模型并训练好,然后通过构建蒸馏损失函数训练轻量化的学生模型,将教师模型学习到的知识蒸馏传递到学生模型中。为了进一步提高模型对高相似度火焰图像的识别精度,论文在模型中引入了融合二阶信息的协方差池化层。论文所提方法的主要贡献如下:

1) 改进了网络的池化层。在特征提取阶段,在网络卷积特征层后引入具有二阶信息的协方差池化方法,替换基于全局最大池化或全局平均池化的一阶池化方法,提高了网络对相似火焰图像的识别精度。

2) 应用知识蒸馏算法,通过监督学习方式,从复杂且需要大量计算资源的高精度模型中提取有用特征,辅助训练所占内存小、反应速度快的轻量化模型,保留轻量化模型优势的同时提高其识别精度,满足了实际生产中实时、低成本识别工况要求。

3) 实验结果表明,经蒸馏后的学生模型总体识别准确率相较于原始模型提高了 3.33%,对测试集中 3 种工况的识别率分别达到了 93%、99%、90%,准确率和模型规模均优于目前其他主流网络模型。

## 1 基于轻量化网络和知识蒸馏的回转窑工况识别

从模式识别的角度来看,所提回转窑燃烧状态识别

方法是基于已知状态的火焰图像样本特征建立模型,并通过模型识别未知状态下回转窑火焰图像。建立的模型分成两个部分,即教师模型和学生模型。教师模型采用高性能、复杂模型,目的是尽可能多的从样本数据中提取特征;学生模型采用轻量化模型,目的是尽可能满足低成本算力设备的部署要求。回转窑工况识别模型的工作流程如图1所示,主要步骤如下:

1)改进教师模型与学生模型,通过在卷积神经网络卷积层后添加协方差池化层(图1中的MPN-COV layer)

改善原网络对于高相似回转窑火焰图像的识别准确率。

2)将训练集输入 Resnet50,经迭代训练得到教师模型。此时,教师模型的网络参数不再发生变化。

3)蒸馏训练。学生模型通过构建的蒸馏损失函数进行训练并更新参数。其中,蒸馏损失函数由教师模型输出的概率分布信息和学生模型随迭代次数不断更新并输出的概率分布信息构成。

4)回转窑工况识别,由经蒸馏训练后的学生模型完成。

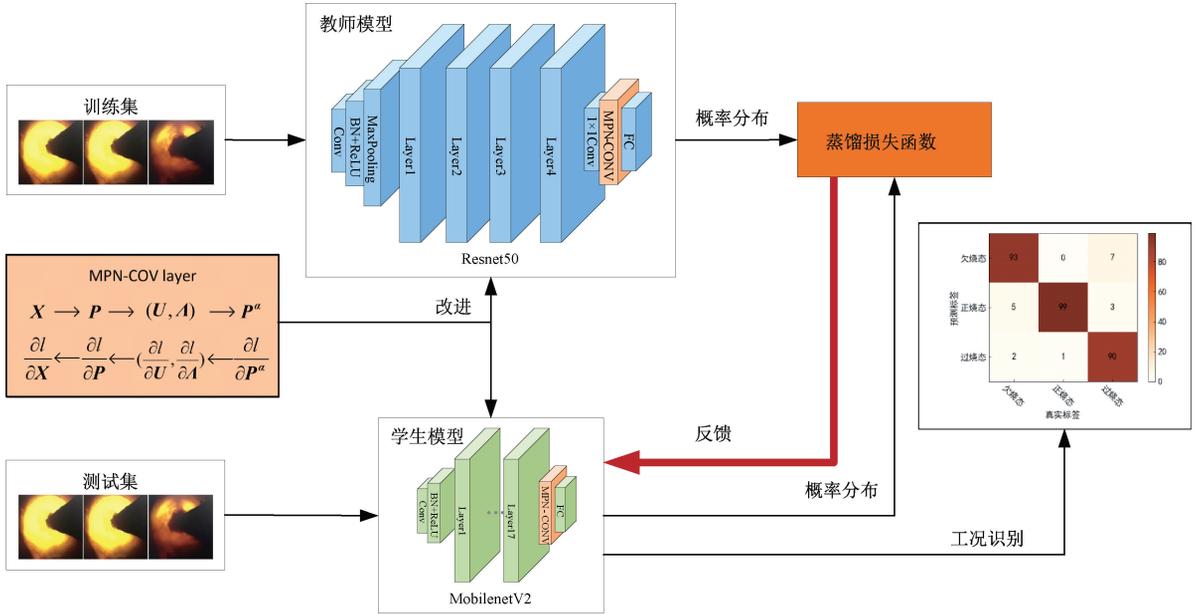


图1 回转窑工况识别模型的工作流程

Fig. 1 Workflow of rotary kiln working condition recognition model

1.1 改进的池化方法

回转窑的火焰图像识别属于类间差异非常小的小样本分类任务,回转窑火焰属于同一基础类别,其3种不同的燃烧状态属于更加细致的子类。

二阶信息(second-order information)已被证实能在细粒度图像分类上取得更好的效果<sup>[17-18]</sup>,因此论文将二阶信息引入回转窑火焰识别网络模型中,用来改进教师模型和学生模型的识别性能。

协方差池(covariance pooling)作为一种二阶池策略,已应用于包括图像分割和分类<sup>[19-20]</sup>的多个计算机视觉任务。首先,与传统池化方法不同,协方差池考虑了二阶统计量(即协方差),因此可以获得更紧凑和更具区分性的表示。如图2所示,协方差矩阵P的非对角项表示两个不同特征图之间的协方差,对角项表示每个特征图的方差,能有效融合不同的特征映射。

矩阵幂归一化协方差法 MPN-COV(matrix power normalized covariance)<sup>[19,21]</sup>,是一种端到端的协方差池化

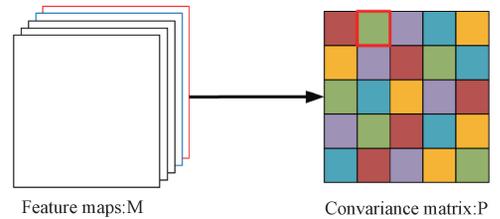


图2 协方差池化

Fig. 2 Covariance pooling

方法,可以通过梯度反向传播不断更新参数训练网络模型。如图3所示,MPN-COV作为中间层,插入到最后的卷积特征层和全连接层之间,替换传统的一阶池化层。

对于网络的正向传播,设最后的卷积特征层为  $M \in \mathbb{R}^{s \times s \times d}$ ,将M降维得到  $X \in \mathbb{R}^{d \times N}$ ,X为卷积神经网络的最后一层特征图矩阵,其中  $N = s^2$ ,d为特征图的通道数。然后由X计算协方差矩阵P,并对矩阵P进行本征分解,得到正交矩阵U和对角矩阵A,其中P矩阵的幂可

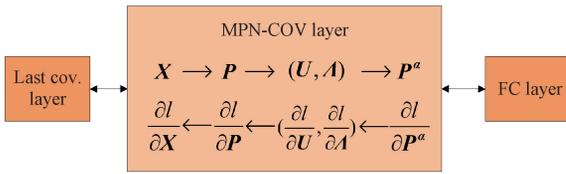


图 3 矩阵幂归一化协方差法的正向传播和反向传播

Fig. 3 Forward and backward propagations of ConvNets with MPN-COV

以转换为计算对角矩阵  $\mathbf{A}$  的幂,如式(1)、(2)所示:

$$\mathbf{P}^\alpha = \mathbf{U}\mathbf{F}(\mathbf{A})\mathbf{U}^T \quad (1)$$

$$\mathbf{F}(\mathbf{A}) = \text{diag}(\lambda_1^\alpha, \lambda_2^\alpha, \dots, \lambda_d^\alpha) \quad (2)$$

将  $\mathbf{P}^\alpha$  输入到全连接层,相对的,反向传播以相反的顺序依次计算相关的偏导数。

正向传播过程中的协方差矩阵计算公式:

$$\mathbf{P} = \frac{1}{N} \sum_{i=1}^N (x_i - u)(x_i - u)^T \in \mathbb{R}^{d \times d} \quad (3)$$

其中,  $[x_1, x_2, \dots, x_N] \in \mathbb{R}^{d \times N}$  为最后一层卷积层的特征图矩阵,  $u = \frac{1}{N} \sum_{i=1}^N x_i \in \mathbb{R}^d$ ,  $u$  即为每一通道特征图

的均值。 $\mathbf{P}$  和  $\mathbf{P}^\alpha$  均为对称矩阵,因此输入到全连接层的特征为  $\frac{d(d+1)}{2}$  个。

### 1.2 教师模型与学生模型的构建

教师模型的主干网络采用 Resnet50<sup>[22]</sup>,是目前先进的深层网络代表。其核心思想是使用的残差结构,不仅解决了传统网络的梯度消失问题,而且网络层数的增加使得特征提取能力更强,分类效果更好。学生模型的主干网络采用 MobilenetV2<sup>[23]</sup>,作为一种基于人工设计方法的轻量化网络模型,其核心在于使用了深度可分离卷积,具有模型参数少、计算速度快的优点,能够实现回转窑实时工况识别。

深度可分离卷积分为两步:逐通道卷积(depthwise convolution)和逐点卷积(pointwise convolution)。逐通道卷积即一个通道只被一个卷积核卷积,并且只在二维平面进行,由于卷积核数量与特征图的通道数相同,因此无法扩展特征图,也无法融合不同通道在相同空间上的特征信息,为了解决这两个问题,需要逐点卷积作进一步处理。逐点卷积的与常规卷积相似,通过一个  $1 \times 1$  的卷积核来融合通道之间的特征信息,通过设置卷积核数量扩展特征图,图 4 展示了逐通道卷积与逐点卷积的运算过程。

式(4)为传统卷积计算量公式,式(5)为深度可分离卷的参数量公式:

$$P_1 = D_k \times D_k \times M \times N \quad (4)$$

$$P_2 = D_k \times D_k \times M + M \times N \times 1 \quad (5)$$

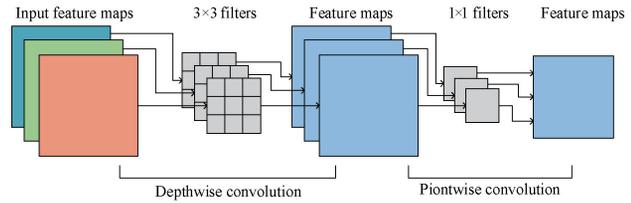


图 4 深度可分离卷积的计算过程

Fig. 4 The calculation of deep-separable convolution

式中:  $M$  表示输入通道数量,  $N$  表示输出通道数量,  $D_k$  表示输出通道卷积核大小。由于深度可分离卷积分为两步,因此总参数量为逐通道卷积参数量与逐点卷积参数量之和,相比传统卷积大大减小。

改进的 Resnet50 网络如图 5 所示,其输入是  $224 \times 224 \times 3$  的火焰图像,首先经过标准卷积层和批归一化,并经过 ReLU 激活层和最大池化层,然后依次经过 4 个残差模块得到  $7 \times 7 \times 2048$  特征输出,  $7 \times 7 \times 2048$  是一个小样本高纬度特征,如果直接用来计算协方差会非常困难且不可靠,因此在输入到 MPN-COV 层之前设置  $1 \times 1 \times 256$  的卷积层降维,最终得到  $256 \times (256+1)/2$  的特征输入到全连接层中。

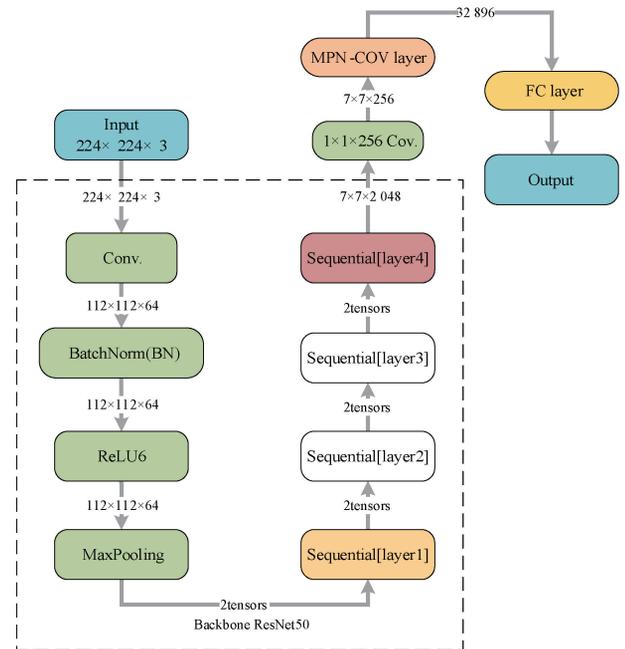


图 5 改进的 Resnet50 网络

Fig. 5 Improved Resnet50 network

改进的 MobilenetV2 网络如图 6 所示,同样输入  $224 \times 224 \times 3$  火焰图像,经过标准卷积层,执行批归一化(batch normalization, BN)和 ReLU6 激活,再经过 17 个采用深度可分离卷积的反向残差层,将得到的特征输入 MPN-COV 层,经全连接层获得输出。为减小模型参数

量,这里去除了原网络最后的  $1 \times 1 \times 1$  280 卷积层,因此输入到 MPN-COV 层的是  $7 \times 7 \times 320$  特征。

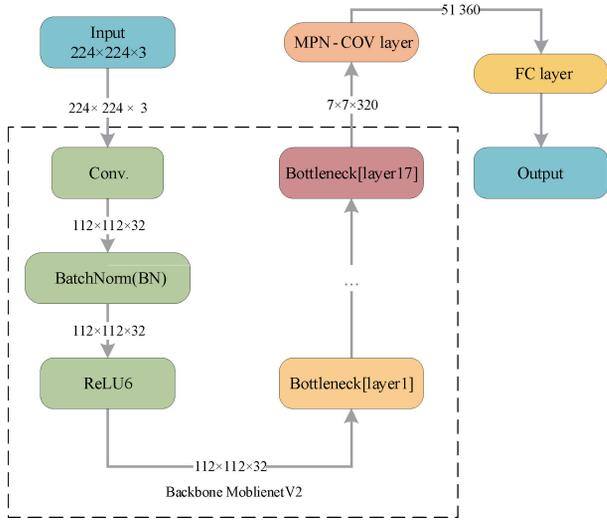


图 6 改进的 MobilenetV2 网络

Fig. 6 Improved MobilenetV2 network

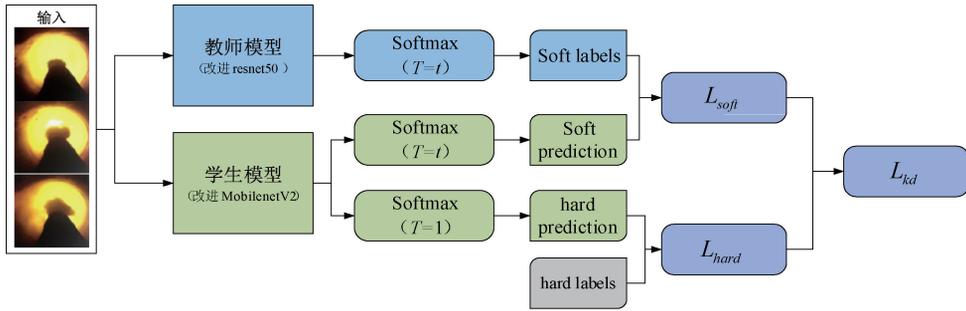


图 7 基于 Softmax 输出层的知识蒸馏方法结构

Fig. 7 Structure of knowledge distillation method based on Softmax layer

从学生模型与教师模型所学的数据集上分析,两者之间的关系可以表述为:对于给定的一个回转窑火焰图像数据集  $D = (X, Y)$ ,  $X$  表示输入数据,  $Y$  表示对应的硬标签,  $Y$  中每个样本的标签  $y \in \{0, 1\}$ , 其中  $y = 1$  表示对应正确的类别标签, 对其余两类则为 0。在整个蒸馏训练中, 教师模型  $T_w$  在数据集  $D$  上进行训练学习。而学生模型  $S_w$  则以教师模型生成的软标签为优化目标, 在新的标记数据集  $\tilde{D} = \alpha(X, T_w(X)) + (a - 1)D$  训练学习, 其中  $T_w(X)$  为教师模型软标签,  $\alpha$  为软标签比例系数。

在回转窑的工况识别中, 教师模型所输出的“软标签”包含经  $T$  平滑后的 3 类工况概率分布信息, 学生模型能从教师模型的概率分布信息中学习到更具类别区分度的信息。

因此, 将 1.2 节中改进的 Resnet50 和改进的 MobilenetV2 分别作为教师模型和学生模型, 通过教师模型输出的软标签指导学生模型进行参数更新, 使学

### 1.3 基于知识蒸馏的回转窑工况识别方法

知识蒸馏 (knowledge distillation) 是模型压缩的常用方法, 其通过模型参数量多、学习性能较强的教师模型输出监督信息, 辅助训练轻量化的学生模型, 以达到更好的性能和精度。其中教师模型输出的监督信息称之为知识, 学生模型学习迁移来自教师模型的监督信息过程称之为蒸馏。

回转窑工况识别采用基于 Softmax 输出层的知识蒸馏方法, 其基本思想为: 学生模型不再只从数据集的“硬标签”中学习, 同时也从教师模型输出的“软标签”中学习。

从损失函数上分析, 蒸馏训练的学生模型的损失函数  $L_{kd}$  由  $L_{soft}$  和  $L_{hard}$  两部分组成, 分别对应“软标签”和“硬标签”学习, 训练流程如图 7 所示。其中  $T$  为蒸馏温度, 用来平滑 Softmax 所输出的概率分布, 它可以使学生模型在训练过程中关注“硬标签”中所没有的负标签信息, 并从经过蒸馏温度  $T$  软化后的标签中学习同类内距离更加紧密的概率分布信息。

生模型尽可能达到教师模型的识别精度, 利用基于知识蒸馏的回转窑工况识别方法进行模型训练的具体步骤为:

1) 将带标签的回转窑火焰图像输入教师模型得到各类工况的概率分布, 并通过蒸馏温度  $T$  对概率分布信息进行平滑处理, 得到教师模型的软标签 (soft labels)  $q_{SLi}$ , 计算公式为:

$$q_{SLi} = \frac{e^{x_{Ti}/T}}{\sum_{i=1}^n e^{x_{Ti}/T}} \quad (6)$$

式中:  $x_{Ti}$  为教师模型最后输出的某一类工况的 logits,  $n$  为工况类别数。

2) 将带标签的回转窑火焰图像输入学生模型得到各类工况的概率分布, 同样通过蒸馏温度  $T$  对概率分布信息进行平滑处理, 得到学生模型的软预测 (soft predictions)  $q_{SPLi}$ , 计算公式为:

$$q_{SPi} = \frac{e^{x_{Si}/T}}{\sum_{i=1}^n e^{x_{Si}/T}} \quad (7)$$

式中:  $x_{Si}$  为学生模型最后输出的某一类工况的 logits。

3) 将回转窑火焰图像输入学生模型, 直接由 softmax 函数得到学生模型的硬预测 (hard predictions)  $q_{HPi}$ , 计算公式为:

$$q_{HPi} = \frac{e^{x_{Si}}}{\sum_{i=1}^n e^{x_{Si}}} \quad (8)$$

4) 由式(6)和(7)得到的软标签和软预测进行计算得到软损失函数值, 计算公式为:

$$L_{soft} = - \sum_{i=1}^C q_{SLi} \log(q_{SPi}) \quad (9)$$

式中:  $C$  为工况类别数。

5) 由式(8)得到的学生模型硬预测和硬标签进行计算得到硬损失函数值, 计算公式为:

$$L_{hard} = - \sum_{i=1}^C y_i \log(q_{HPi}) \quad (10)$$

6) 将式(9)和(10)的两种损失函数通过加权得到混合损失函数  $L_{hd}$ , 计算公式为:

$$L_{hd} = (1 - \alpha)L_{hard} + \alpha T^2 L_{soft} \quad (11)$$

式中: 温度  $T$  和系数  $\alpha$  是需要自主设置的超参数, 将影响知识蒸馏的效果。

## 2 实验结果与分析

### 2.1 数据集描述

回转窑烧成状态可以分为“正烧态”、“欠烧态”和“过烧态”3种工况, 其中“欠烧态”和“过烧态”是异常烧成状态。其示例图如图8所示。

在“欠烧态”下, 物料燃烧不充分, 火焰亮度较低且发光面积小; 在“过烧态”下, 物料在窑内剧烈燃烧, 火焰亮度高且发光面积大, 火焰中常带有过曝的亮斑; 在“正烧态”下, 物料燃烧充分, 火焰亮度均匀。当回转窑处于过烧或欠烧状态时, 窑内的烧结产物会出现烧结过度或不完全烧结的情况, 都不符合质量要求。

实验数据集通过某水泥回转窑的窑内工业摄像机采集, 各个类别样本数量均衡, 每类各 500 张, 总计 1 500 张, 标签全部由专家进行标定。实验取 80% 作为训练样本, 取 20% 作为测试样本。

### 2.2 实验设计与设置

实验具体分为 3 个部分, 第 1 部分是对协方差池化方法进行有效性验证, 实验主要对比不同池化方法在回转窑工况识别上的效果, 方法包括传统的一阶池化方法全局最大池化 (global max pooling, GMP) 和全局平均池化



图 8 回转窑火焰的 3 种燃烧态

Fig. 8 Three combustion states of rotary kiln flames

(global average pooling, GAP) 以及二阶池化方法 BCP<sup>[24]</sup> 和 B-CNN<sup>[18]</sup>, 以总体准确率和类准确率进行量化分析。

第 2 部分是对知识蒸馏算法进行有效性验证, 实验主要对比蒸馏后的学生模型同未蒸馏的学生模型以及教师模型在准确率和混淆矩阵上的差异。

第 3 部分对比其他经典模型以及规格相近的轻量化模型, 从模型识别准确率、精准度和模型规格上验证所提方法的优越性。

模型训练在一台配备 Intel i5 12400 处理器, 32 GB RAM 以及 NVIDIA RTX3060 的 Ubuntu PC 上进行。实验使用 Pytorch 架构, SGD 优化器, 并使用交叉熵损失函数。学习率设置为 0.02, 批大小为 60, 动量设置为 0.9, 训练 100 次。

由于各个类别样本数量均衡, 实验将总体准确率 (overall accuracy, OA) 和精准度 (Precision) 作为评价指标, 计算方法如式(12)和(13)所示:

$$OA = \frac{\text{num}(test_{pred} = test_{true})}{\text{mun}(test)} \times 100\% \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

其中,  $test_{pred}$  表示测试集图像由模型判断得到的分类结果,  $test_{true}$  表示测试集图像的标注类别,  $\text{num}(test_{pred} = test_{true})$  表示识别正确的数量,  $\text{mun}(test)$  表示测试集的样本总数。TP 表示被正确预测为某一类的数量, FP 表示被错误预测为这一类的数量。在本文中, 欠烧态、正烧态和过烧态的精准度分别为  $P_u$ 、 $P_n$  和  $P_o$ , 平均精准度为  $P_a = \frac{P_u + P_n + P_o}{3} \times 100\%$

### 2.3 实验结果与分析

#### 1) 改进池化方法的实验分析

图9为Resnet50采用各池化方法的测试集准确率变化曲线。可以看出,二阶池化方法的MPN-COV、CBP和B-CNN均优于一阶池化方法的GMP和GAP,。进一步分析曲线,MPN-COV不仅准确率最高,曲线的震荡程度也最小,没有过拟合现象。

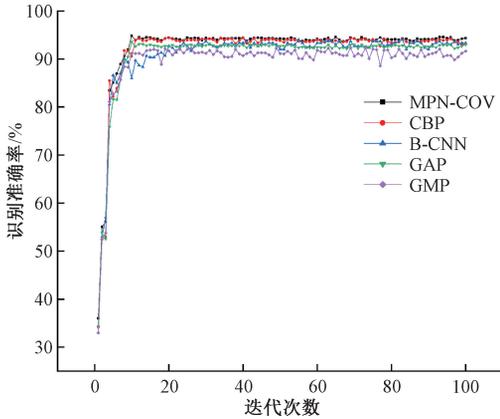


图9 各池化方法在 Resnet50 上的识别准确率曲线

Fig.9 Recognition accuracy curve of each pooling method on Resnet50

表1为各池化方法在Resnet50上的类别识别准确率和OA。其中,采用MPN-COV池化方法的总体准确率达到了94.67%,相较于Resnet50默认的GAP池化方法提高了1.33%。进一步分析可知,各池化方法对于正烧态的识别准确率都十分理想,而对于欠烧态和过烧态两种不正常的烧成状态,采用MPN-COV池化方法的Resnet50表现出了更优异的性能,对比Resnet50默认的GAP池化方法分别提升了2%和1%,说明MPN-COV这种二阶池化方法可以更好的融合卷积层的特征映射,提高了模型对高相似火焰的识别能力。

表1 各池化方法在 Resnet50 上的识别准确率

Table 1 The recognition accuracy of each pooling method on Resnet50 (%)

池化方法	欠烧态	正烧态	过烧态	OA
MPN-COV	94	99	91	94.67
CBP	93	98	91	94.00
B-CNN	92	98	91	93.67
GAP	92	98	90	93.33
GMP	90	97	88	91.67

表2为MPN-COV池化方法在MobilenetV2上的实验结果,其总体识别准确率相较于MobilenetV2默认的GAP池化方法提升了0.67%,欠烧态和过烧态的识别准确率均有提升。

2) 知识蒸馏的实验分析

为验证知识蒸馏算法的有效性,将知识蒸馏与协方

表2 各池化方法在 MobilenetV2 上的识别准确率

Table 2 The recognition accuracy of each pooling method on MobilenetV2 (%)

池化方法	欠烧态	正烧态	过烧态	OA
MPN-COV	90	96	88	91.33
GAP	89	96	87	90.67
GMP	89	95	86	90.00

差池化层结合,改进后的学生模型称为 MobilenetV2-MPNCOV-KD,改进后的教师模型称为 Resnet50-MPNCOV,式(11)中超参数 $\alpha$ 设置为0.8, $T$ 设置为8。

图10展示了所提模型的测试集准确率变化曲线,分析图10和表3可以得到,经教师模型蒸馏训练后,改进的学生模型较原模型总体准确率提高了3.33%,3种工况的识别准确率分别提升了4%、3%和3%,识别准确率接近于教师模型。

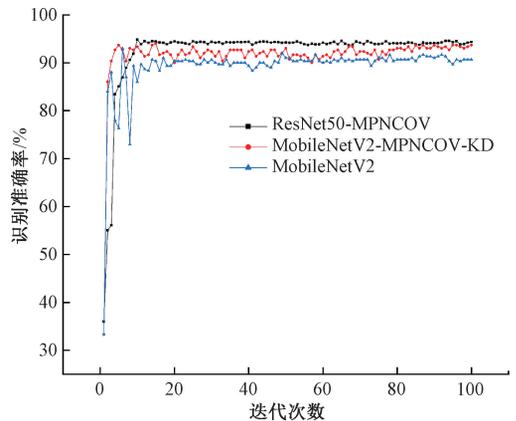


图10 对比蒸馏效果的识别准确率曲线

Fig.10 The recognition accuracy curves with distillation

表3 教师模型和学生模型的识别准确率

Table 3 The recognition accuracy of teacher model and student model (%)

所用模型	欠烧态	正烧态	过烧态	OA
Resnet50-MPNCOV (Teacher model)	94	99	91	94.67
MobilenetV2-MPNCOV-KD (Student model)	93	99	90	94.00
MobilenetV2	89	96	87	90.67

图11(b)展示了所提方法的混淆矩阵,同图11(a)和(c)对比可知,未经蒸馏的 MobilenetV2 模型在欠烧态和过烧态两种状态下的识别精确率上表现不佳,具体来说,欠烧态更容易误识别成过烧态,而过烧态则更容易误识别成欠烧态。经过蒸馏后,学生模型对这两种工况状态下的识别准确率分别提升了4%与3%。值得注意的是,教师模型和学生模型在混淆矩阵上具有相似性,两者

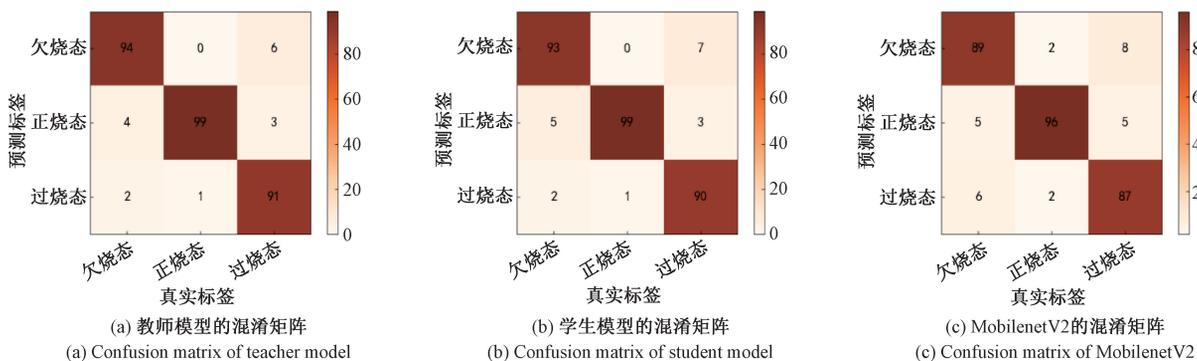


图 11 蒸馏实验的混淆矩阵

Fig. 11 Confusion matrix for knowledge distillation experiments

的差异在于学生模型多将一张欠烧态识别成了正烧态以及多将一张过烧态识别成了欠烧态,这说明通过知识蒸馏算法,教师模型能向学生模型传递有效的类别概率分布信息,指导学生模型向教师模型的分类结果优化,以获得接近教师模型的识别性能。

为直观地体现学生模型在蒸馏前后的性能提升,列举了几组较难识别的异常状态的回转窑火焰图像,如图 12 所示。

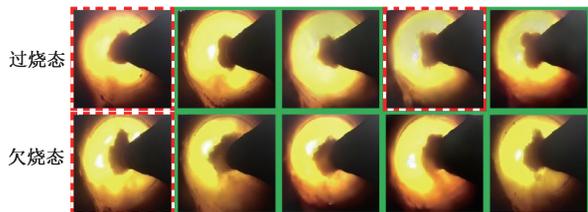


图 12 学生模型较难识别的回转窑火焰图像

Fig. 12 Rotary kiln flame images difficult to be recognized by the student model

在图 12 中,所列举图像均为学生模型在未经蒸馏时识别错误的图像,实线框的表示在经蒸馏后识别正确的图像,虚线框的表示仍识别错误的图像。具体分析这些图像,可以发现在过烧态和欠烧态两种异常状态下,两种状态的判别特征有许多交叉的地方。所列举的欠烧态图像均有少量的过曝亮斑,而过烧态图像也有类似欠烧态因粉尘多导致的画面有些“脏”的现象,这导致模型难以分辨它们。

在较为相似的类上,合适的蒸馏温度  $T$  对教师标签的软化实际上扩大了类之间的中心距离,意味着它使两个类别更容易区分;在有较大差异的类上,它只是稍微拉近了类之间的距离,不会从根本上改变所有类之间的关系分布<sup>[25]</sup>。具体到回转窑的工况识别,教师模型的软标签可以使学生模型更容易区分过烧态和欠烧态,而对于正烧态的识别只是略微降低预测概率值,不会影响最终

识别结果。因此,可以从图 12 中看到学生模型在经蒸馏训练后提高了对这些高相似火焰图像的识别精度。

此外,为进一步探究蒸馏温度  $T$  对蒸馏学习效果的影响,在不同蒸馏温度  $T$  上进行了实验,结果如表 4 所示。可以看出,随着温度  $T$  的逐渐上升,模型的识别准确率呈现先上升后下降的趋势,  $T = 8$  时准确率最高。  $T = 16$  时,模型的总体准确率相较于未经蒸馏时的 91.33% 反而下降了 0.33%,在欠烧态下降了 2%,这是由于过高的蒸馏温度使教师模型的负标签更为凸显,导致负标签与正标签过于接近,进而导致学生模型学习到了劣化后的软标签,这种劣化在预测概率值相对不高的不正常烧成状态下(对比正烧态)会更加严重,因此模型对于不正常烧成状态的识别准确率有所下降。

表 4 不同蒸馏温度  $T$  下的识别性能

Table 4 Recognition performance under different distillation temperature  $T$  (%)

$T$	欠烧态	正烧态	过烧态	OA
2	91	97	89	92.33
4	91	98	90	93.00
8	93	99	90	94.00
12	91	97	90	92.67
16	88	97	88	91.00

3) 对比实验

为评估模型性能,本文将提出的 MobilenetV2-MPNCOV-KD 模型与 Alexnet<sup>[26]</sup>、VGG16<sup>[27]</sup>、Resnet34、Resnet50、ShufflenetV2<sup>[28]</sup> 以及 MobileViT<sup>[29]</sup> 模型进行性能比较,实验结果如表 5 所示。由表 5 可知,所提改进模型效果最佳,总体准确率达到 94.00%,平均精准度达到 94.09%,超过了先进的 Resnet50 模型,但对比模型规格,所提模型的计算量和参数量仅是 Resnet50 的 8.47% 和 9.46%,轻量化的优势明显。对比同属于轻量化模型的 ShufflenetV2,所提模型的准确率和平均精准度均提高

了3%,两者在计算量和参数量上处于同一量级。ShufflenetV2在计算量上有优势,为本文所提模型的42.86%,参数量则与所提模型相当。MobileViT是一种结合了Mobilenet与视觉Transformer的轻量化网络,在ImageNet-1k数据集上有着优秀的性能。对比所提模型,两者在计算量与参数量上处于同一量级,但MobileViT在识别准确度和平均精确度上分别降低了1.67%和1.71%,这是由于Transformer擅长表征全局的语义特征,

不太适合回转窑工况识别这种更注重局部特征表征的细粒度分类任务,再加上Transformer结构对训练集的数据量需求较高,因此未能达到较好的识别效果。

综上所述,论文所提出的回转窑工况识别方法,能在确保模型轻量化的基础上,实现超过复杂模型的识别性能,对比规格处于同一量级的其他轻量化模型也具有性能优势,能够较好地满足回转窑工况识别任务。

表5 与其他模型的对比实验

Table 5 Comparison experiment with other models

所用模型	OA/%	$P_a$ /%	欠烧态 $P_u$	正烧态 $P_n$	过烧态 $P_o$	计算量 FLOPs/G	参数量 parameters/M
MobilenetV2-MPNCOV-KD(本文)	94.00	94.09	0.930 0	0.925 2	0.967 7	0.35	2.23
Resnet50	93.33	93.37	0.929 3	0.924 5	0.947 4	4.13	23.51
Resnet34	91.33	91.43	0.927 8	0.888 9	0.926 3	3.68	21.29
VGG16	89.67	89.68	0.887 8	0.897 2	0.905 3	15.47	134.27
Alexnet	86.33	86.44	0.865 9	0.846 8	0.880 4	0.71	57.02
ShufflenetV2	91.00	91.09	0.900 0	0.897 2	0.935 5	0.15	2.32
MobileViT	92.33	92.38	0.918 4	0.916 7	0.936 2	0.70	2.30

### 3 结 论

针对回转窑工况识别困难问题,利用回转窑火焰图像蕴含的丰富特征信息,提出了一种基于轻量化网络和知识蒸馏的回转窑工况识别方法。首先,将二阶的协方差池化方法引入到教师模型和学生模型中,以提高二者对高相似火焰图像的识别能力。然后通过知识蒸馏算法,将教师模型中蕴含的分类信息传递给学生模型,使学生模型不断对教师模型的分类结果作训练优化。实验结果表明,所提方法不仅在识别性能上优于目前的主流模型,同时保持了轻量化模型规格,能以较低硬件成本满足回转窑工况识别任务。下一步工作将考虑不平衡样本的分类任务,提升模型对于异常工况的识别性能。

### 参考文献

[1] ZHANG X G, ZHANG L, CHEN H, et al. Prediction of coal feeding during sintering in a rotary kiln based on statistical learning in the phase space [J]. ISA Transactions, 2018, 83:248-260.

[2] CHEN H, YAN T T, ZHANG X G. Burning condition recognition of rotary kiln based on spatiotemporal features of flame video[J]. Energy, 2020, 211:118656.

[3] 蔡国源, 牛玉广, 刘雪菲, 等. 基于图像卷积变分自编码的电站锅炉燃烧稳定性评价方法[J]. 仪器仪表学报, 2022, 43(3): 210-220.

CAI G Y, NIU Y G, LIU X F, et al. Combustion stability judgment of power plant boiler based on image

convolutional variational auto-encoder [J]. Chinese Journal of Scientific Instrument, 2022, 43 (3): 210-220.

[4] RAHUL S, CAMY J, SHYAM K, et al. Versatile image processing technique for fuel science: A review [J]. Science of the Total Environment, 2021, 780:146469.

[5] GUO S Y, SHENG Y X, CHAI L. SVD-based burning state recognition in rotary kiln using machine learning[C]. IEEE Conference on Industrial Electronics and Applications, 2017: 154-158.

[6] 潘峰, 孙红霞. 基于蝗虫算法的图像多阈值分割方法[J]. 电子测量与仪器学报, 2019, 33(1): 149-155.

PAN F, SUN H X. Algorithm for image segmentation based on grasshopper optimization algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(1): 149-155.

[7] ZHANG R F, LU S Z, YU H L, et al. Recognition method of cement rotary kiln burning state based on Otsu-Kmeans flame image segmentation and SVM [J]. International Journal for Light and Electron Optics, 2021, 243(2):167418.

[8] LI T, TAO L, PENG T, et al. Rotary kiln combustion state recognition based on convolutional neural network [J]. Journal of Physics: Conference Series, 2020, 1575(1): 012030.

[9] 栾庆磊, 陈克琼. 基于深度迁移学习的回转窑燃烧状况智能感知[J]. 重庆大学学报, 2019, 42(9):85-92.

LUAN Q L, CHEN K Q. Intelligent cognition of rotary

- kiln burning state based on deep transfer learning [J]. *Journal of Chongqing University*, 2019, 42(9): 85-92.
- [10] HU Y C, ZHENG W H, WANG X, et al. Working condition recognition based on transfer learning and attention mechanism for a rotary kiln [J]. *Entropy*, 2022, 24(9): 1186.
- [11] 张茜, 马文科, 周晓杰. 基于回声状态网络的回转窑工况识别方法[J]. *控制工程*, 2021, 28(1): 142-148. ZHANG Q, MA W K, ZHOU X J. Recognition method of rotary kiln working condition based on echo state network [J]. *Control Engineering of China*, 2021, 28(1): 142-148.
- [12] 王伟彤, 李长峰, 朱华波, 等. 轻量化 CenterNet 网络的二维条码定位算法[J]. *电子测量与仪器学报*, 2022, 36(5): 128-135. WAN W T, LI CH F, ZHU H B, et al. Two-dimensional barcode positioning algorithm of lightweight CenterNet network [J]. *Journal of Electronic Measurement and Instrumentation*, 2022, 36(5): 128-135.
- [13] 梁秀满, 邵彭娟, 刘振东, 等. 自适应特征融合的轻量级交通标志检测方法[J]. *电子测量技术*, 2022, 45(23): 107-112. LIANG X M, SHAO P J, LIU ZH D, et al. A lightweight traffic sign detection method based on adaptive feature fusion [J]. *Electronic Measurement Technology*, 2022, 45(23): 107-112.
- [14] 李仁鹰, 钱慧芳, 郭佳豪, 等. 基于 M-YOLOv4 模型的轻量化目标检测算法[J]. *国外电子测量技术*, 2022, 41(4): 15-21. LI R Y, QIAN H F, GUO J H, et al. Lightweight target detection algorithm based on M-YOLOv4 model [J]. *Electronic Measurement Technology*, 2022, 41(4): 15-21.
- [15] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [J]. *Computer Science*, 2015, 14(7): 38-39.
- [16] 顾佼佼, 翟一琛, 姬嗣愚, 等. 基于 BERT 和知识蒸馏的航空维修领域命名实体识别[J]. *电子测量技术*, 2023, 46(3): 19-24. GU J J, ZHAI Y CH, JI S Y, et al. Aviation maintenance text named entity recognition based on BERT and knowledge distillation [J]. *Electronic Measurement Technology*, 2023, 46(3): 19-24.
- [17] IONESCU C, VANTZOS O, SMINCHESCU C. Training deep networks with structured layers by matrix backpropagation [C]. *International Conference on Computer Vision*, 2015, 9: 2965-2973.
- [18] RAO D S, CH R B, KIRAN V S, et al. Plant disease classification using deep Bilinear CNN [J]. *Intelligent Automation and Soft Computing*, 2021, 30(1): 167-176.
- [19] LI P H, XIE J T, WANG Q L, et al. Is Second-order information helpful for large-scale visual recognition? [C]. *IEEE International Conference on Computer Vision*, 2017: 2089-2097.
- [20] HE N J, FANG L Y, LI S T, et al. Remote sensing scene classification using multilayer stacked covariance pooling [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(12): 6899-6910.
- [21] LI P H, XIE J T, WANG Q L, et al. Towards faster training of global covariance pooling networks by iterative matrix square root normalization [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 947-955.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [23] SANDLER M, HOWARD A, ZHU M, et al. MobileNet V2: Inverted residuals and linear bottlenecks [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 4510-4520.
- [24] CHETOUANI A, TREUILLET S, EXBRAYAT M, et al. Classification of engraved pottery sherds mixing deep-learning features by compact bilinear pooling [J]. *Pattern Recognition Letters*, 2020, 131: 1-7.
- [25] SHEN Z Q, LIU Z C, XUE D J, et al. Is label smoothing truly incompatible with knowledge distillation: An empirical study [C]. *International Conference on Learning Representations*, 2021.
- [26] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [27] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]. *International Conference on Learning Representations*, 2015: 1-14.
- [28] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design [C]. *European Conference on Computer Vision*, 2018, 122-128.
- [29] MEHTA S, RASTEGARI M. MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer [C]. *International Conference on Learning Representations*, 2022.

## 作者简介



王欣,2004 年于中南林业科技大学获硕士学位,2010 年于中南大学获博士学位,现为湖南工业大学教授,硕士生导师,主要研究方向为复杂系统建模、智能控制和优化、人工智能在各种工程问题中的应用。

E-mail: wangxin97p@163.com

**Wang Xin** received her M. S. degree in 2004 from Central South University of Forestry & Technology and Ph. D. degree in 2010 from Central South University, respectively. She is now the professor and master supervisor at the Hunan University of Technology. Her main research interests include the complex system modeling, intelligent control and optimization, and the application of artificial intelligence to resolving various engineering problems.



邓章俊,现为湖南工业大学硕士研究生,主要研究方向为图像处理、机器视觉。

E-mail: 1290329541@qq.com

**Deng Zhangjun** is a M. Sc. candidate of Hunan University of Technology. His main research interests include image processing and machine vision.



秦斌(通信作者),1988 年于中南工业大学(现中南大学)获硕士学位,2006 年于中南大学获博士学位。现为湖南工业大学教授,博士生导师,主要研究方向为复杂工业过程建模与智能控制、机器学习与模式识别、风力发电、系统优化与调度。

E-mail: qinbin99p@163.com

**Qin Bin**(Corresponding author) received his M. Sc. degree in 1988 and Ph. D. degree in 2006 from Central South University, respectively. He is now a professor and doctoral supervisor at Hunan University of Technology. His main research interests include complex industrial process modeling and intelligent control, machine learning and pattern recognition, wind power generation, system optimization and scheduling.