

# 基于CBAM-EfficientNet-B0的 铁谱图像磨损类型识别 算法

刘胜慧

天津工业大学，控制科学与工程学院，天津

收稿日期：2022年6月23日；录用日期：2022年7月4日；发布日期：2022年7月14日

## 摘要

铁谱图像磨损类型识别是分析机械设备磨损故障的重要方法。针对磨粒数据集样本少且不同磨损类型在质地、形状和颜色上的差异微小导致的分类精度低的问题，提出一种基于改进EfficientNet网络的铁谱图像磨损类型识别算法。本文选择EfficientNet-B0作为磨损类型识别的基本模型，将CBAM注意力模块融合到EfficientNet-B0中，构建CBAM-EfficientNet-B0，从而提高磨粒的聚焦能力和信息表达能力。本文构建了五种磨损类型的磨粒图像数据集。在测试数据集上测试了CBAM-EfficientNet-B0的磨损类型识别能力。实验结果表明，本文提出的磨损类型识别算法CBAM-EfficientNet-B0准确率为92.55%，较改进前EfficientNet-B0算法的准确率提升了2.51%，提高了机械设备磨损状态识别精度和效率。将CBAM-EfficientNet-B0与MobilenetV3、Resnet50、VGG16和ViT分类模型进行比较，实验结果表明CBAM-EfficientNet-B0的精确率、召回率和准确率均高于对比实验中的其他方法。该研究为设备的状态维修和故障诊断提供了新的技术选择。

## 关键词

铁谱分析，磨损类型识别，EfficientNet，深度学习

## A Wear Type Recognition Algorithm for Ferrography Images Based on CBAM-EfficientNet-B0

Shenghui Liu

School of Control Science and Engineering, Tiangong University, Tianjin

Received: Jun. 23<sup>rd</sup>, 2022; accepted: Jul. 4<sup>th</sup>, 2022; published: Jul. 14<sup>th</sup>, 2022

文章引用：刘胜慧. 基于 CBAM-EfficientNet-B0 的铁谱图像磨损类型识别算法[J]. 图像与信号处理, 2022, 11(3): 101-112. DOI: 10.12677/jisp.2022.113012

## Abstract

Ferrographic image wear type identification is an important method to analyze the wear failure of mechanical equipment. Aiming at the problem of low classification accuracy caused by the small number of samples in the abrasive particle dataset and the small differences in texture, shape and color of different wear types, a wear type recognition algorithm based on improved EfficientNet network was proposed. In this paper, EfficientNet-B0 is selected as the basic model for wear type recognition, and the CBAM attention module is integrated into EfficientNet-B0 to construct CBAM-EfficientNet-B0, thereby improving the focusing ability and information expression ability of abrasive particles. In this paper, a dataset of abrasive grain images for five types of wear is constructed. The wear type recognition ability of CBAM-EfficientNet-B0 is tested on the test dataset. The experimental results show that the accuracy of the wear type identification algorithm CBAM-EfficientNet-B0 proposed in this paper is 92.55%, which is 2.51% higher than that of the EfficientNet-B0 algorithm before the improvement, which improves the accuracy and efficiency of mechanical equipment wear state identification. Comparing CBAM-EfficientNet-B0 with MobileNetV3, Resnet50, VGG16 and ViT classification models, the experimental results show that the precision, recall and accuracy of CBAM-EfficientNet-B0 are higher than other methods in the comparative experiments. This research provides new technical options for condition maintenance and fault diagnosis of equipment.

## Keywords

Ferrographic Analysis, Wear Type Identification, EfficientNet, Deep Learning

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着科学技术的进步,现代工业正朝着智能化、自动化的方向迅速发展。高度自动化的制造体系对工业设备的稳定性提出了更高的要求,因此,故障诊断技术在现代制造业中显得尤为重要。研究发现,机械磨损是造成设备失效的主要因素,大约有80%的机械故障是由磨损引起的[1]。在设备高强度运转过程中,磨损会加剧机械零部件的损耗,降低设备的工作效率、精度,影响设备正常运行,甚至会发生灾难性事故[2]。因此,通过技术手段对磨损进行有效地、实时地监测对于维持设备稳定、高效地运转有着重要的实际意义。

铁谱分析技术设备故障诊断中的重要技术手段。铁谱图像中的磨粒形态与分布能够反映设备的磨损状态。铁谱分析技术可以通过数字图像处理技术对磨粒进行识别,从而判断设备的磨损状态。到目前为止,研究人员在建立磨粒自动分类系统方面做了大量的工作。K. Jardine 人利用面积、周长和伸长参数来区分磨损碎屑,结果表明,这些形态特征能有效地分辨某些种类的磨损碎屑,但是有几种大于20微米的异常碎片无法准确识别[3]。Li Q 等人采用了基于磨粒形状检测的支持向量机算法[4]。Peng Z 等人提出了一种称为 FECNN 的卷积神经网络模型,用于识别包括切削、球形、疲劳和滑动颗粒在内的单个磨损颗粒的图像[5]。Fengguang Jia 等人提出了一种用于铁谱磨粒图像分析的新型识别网络,以 DenseNet121 为基础网络,采用迁移学习方法构建 DCNN 模型(FWDNet),获得了90.15%的准确率[6]。

不同的磨损机理产生的磨粒各异,因而可以通过磨粒分析来识别引起磨损故障的磨损机理。与润滑油分析有关的磨损机理可以分为以下几类:粘着磨损、磨料磨损、疲劳磨损和腐蚀磨损[7]。本文构建了新的分类器来识别上述磨损类型及正常磨损。为了提高磨损类型识别的精度,本文采用 EfficientNet-B0 [8] 作为磨损类型识别的基本模型,将卷积注意力模块(Convolutional Block Attention Module, CBAM) [9]融合到 EfficientNet-B0 中,构建 CBAM-EfficientNet-B0,从而提高网络模型磨粒信息的表达能力。这是 EfficientNet 和 CBAM 模块在磨损类型识别的首次应用。

## 2. 本文方法

### 2.1. EfficientNet

EfficientNet 是一个通过缩放三个参数(即深度、宽度和分辨率)而开发的高效的分类网络。网络有不同的版本,从 B0 到 B7,逐渐增加了子块的数量。本文选择 EfficientNet-B0 作为特征提取器。在该网络中,输入图像分辨率为  $224 \times 224$ , EfficientNet 是基于 MBConv 模块建立的网络。MBConv 模块的示意图如图 1 所示。EfficientNet-B0 的网络结构图如图 2 所示。

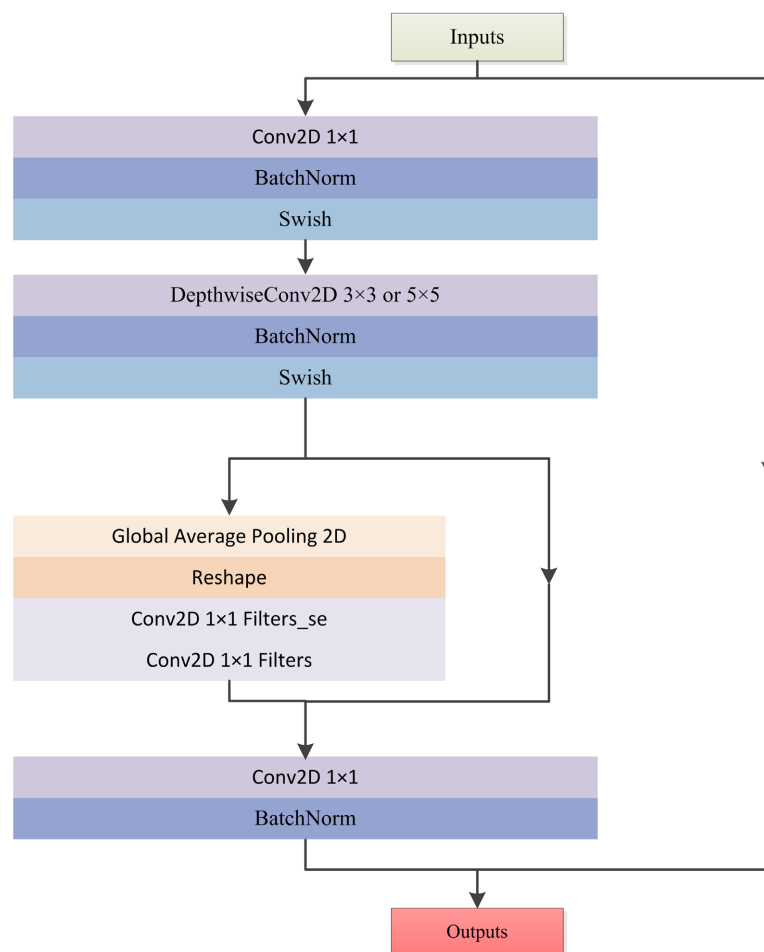


Figure 1. Schematic diagram of MBConv module

图 1. MBConv 模块示意图

如图 1 所示, MBConv 模块由左右两个部分组成, 左边是该模块的主干部分, 右边是该模块的残差

边。首先对输入特征层进行卷积核大小为  $1 \times 1$  的卷积操作以实现升维，然后进行标准化和激活函数。接着进行卷积核大小为  $3 \times 3$  或  $5 \times 5$  的深度可分离卷积，之后进行标准化和激活函数。再进行添加注意力机制，注意力机制部分也包括两个部分，左边是 SE (Squeeze Excitation) 注意力模块，右边是需要添加注意力的特征层。在 SE 模块中，首先对输入特征层进行全局平均池化，然后利用缩减比对神经元数量进行缩减，其宽度或神经元数量由缩减比决定。缩减比越高，瓶颈中的神经元数量越少，反之亦然。接着进行  $1 \times 1$  卷积降维和  $1 \times 1$  卷积升维，此时获得的特征层通道数和输入特征层通道数相等，可以代表注意力机制输入的特征层每个通道的权重，最后进行相乘操作实现了注意力机制的施加。最后通过一个卷积核大小为  $1 \times 1$  的卷积进行降维操作，标准化操作后增加残差边得到最后的输出特征层。

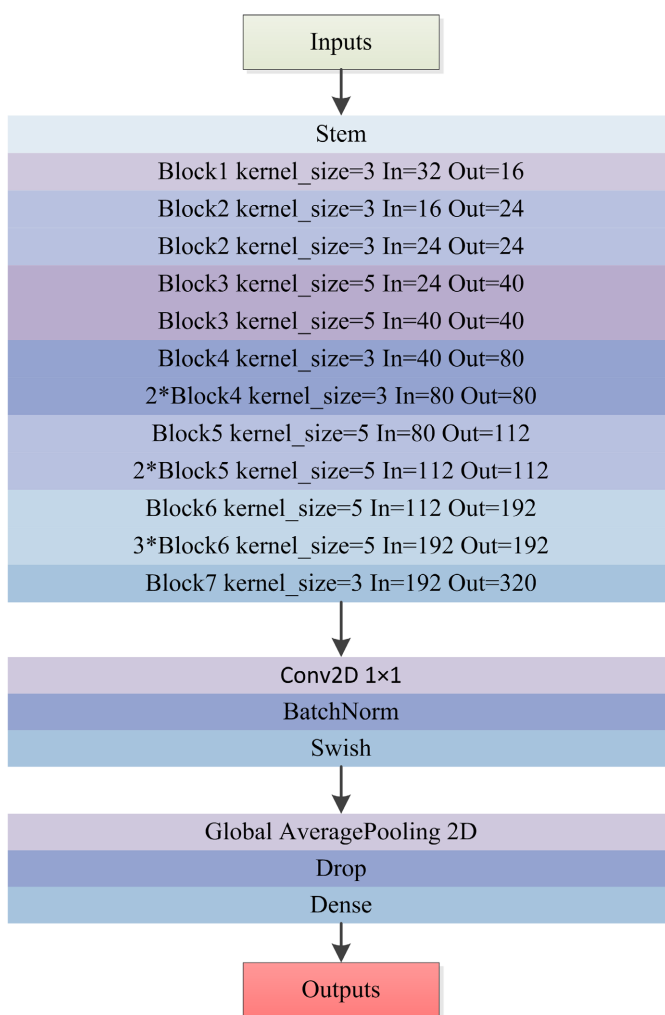


Figure 2. EfficientNet-B0 network structure diagram

图 2. EfficientNet-B0 网络结构图

EfficientNet 由 1 个 Stem、16 个 MBConv 模块、Conv2D、GlobalAveragePooling2D 和 Dense 组成。Stem 用于进行特征的初步提取，由卷积、标准化和激活函数构成。该网络的核心是 16 个 MBConv 模块。MBConv 模块是 EfficientNet 特有的特征提取结构，在模块堆叠的过程中完成高效的特征提取。Conv2D + GlobalAveragePooling2D + Dense 是该网络的分类头，全连接后用于最终的分类。

## 2.2. 融合 CBAM 模块的 EfficientNet-B0 网络

CBAM (Convolutional Block Attention Module)是一种新的卷积注意力模块,创新地提出了通道注意力模块(Channel Attention Module, CAM)与空间注意力模块(Spatial Attention Module, SAM)融合的注意力机制。该注意力模块对于前馈卷积神经网络来说简单又有效,同时又具有轻量化、通用性强的优点,可以无缝集成到任何卷积神经网络中以从中间的“输入特征图”中获得后续的“精炼特征图”,从而提高网络的分类精度的效果。CBAM 的结构如图 3 所示。

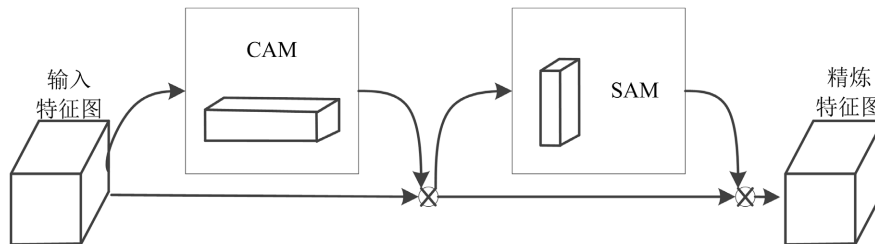


Figure 3. CBAM network structure diagram

图 3. CBAM 网络结构图

CAM 的网络结构图如图 4 所示。CAM 较 SE 做了修改,不是通过全局平均池化将特征进行下采样,而是将输入特征图分解为 2 个后续的维度向量,其中一个向量由全局平均池化层生成,另一个向量由全局最大池化层生成。全局平均池化层主要用于聚合空间信息,而全局最大池化层以图像中对象边缘的形式保留了更丰富的上下文信息,从而得到更精细的通道注意力。简而言之,平均池化具有平滑效果,而最大池化具有更锐利的效果,但更精确地保留了对象的自然边缘。研究表明,使用全局平均池化层和最大池化比仅使用平均池化得到的结果更好。然后分别经过共享多层感知机(Multilayer Perceptron, MLP)。将 MLP 输出的特征进行基于逐元素的加和操作,再经过 Sigmoid 激活操作,生成最终的通道注意力特征图。将该通道注意力特征图和输入特征图做逐元素的乘法操作,生成空间注意力模块需要的输入特征。

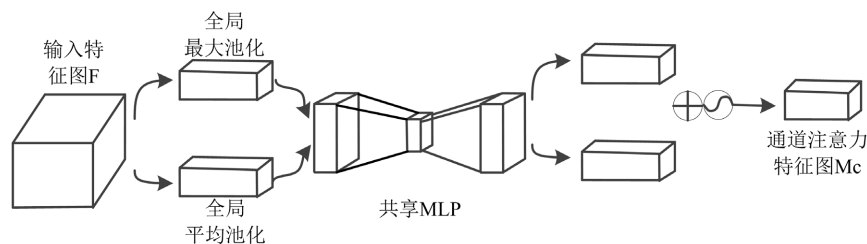


Figure 4. CAM network structure diagram

图 4. CAM 网络结构图

SAM 的网络结构图如图 5 所示。空间注意模块由三部分组成,第一部分为通道池化层,输入层被分解成两个通道即最大池化和平均池化,然后将这两个通道汇集作为下一层的输入。因此,这个卷积层是一个空间维度保持卷积。第二部分是批量标准化(BatchNorm, BN)层,用于对卷积的输出进行归一化和缩放。第三部分是 Sigmoid 激活层。Sigmoid 是一种概率激活,它将所有值映射到 0 到 1 之间的范围内。然后使用简单的元素乘积将这个空间注意掩码应用于输入张量中的所有特征图。

原 EfficientNet-B0 网络中虽包含通道注意力机制 SE,但 CBAM 中的 CAM 较 SE 做了改进,同时原网络未考虑空间信息注意力,而 CBAM 中包含了这一机制,将 CBAM 引入到 EfficientNet-B0 网络可以

进一步提升网络的特征提取能力从而使网络的识别精度更高。

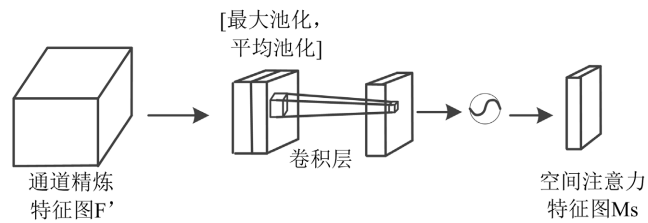


Figure 5. SAM network structure diagram  
图 5. SAM 网络结构图

改进后的 EfficientNet-B0 将 CBAM 融入到了 EfficientNet-B0 中, 并将其命名为 CBAM-EfficientNet-B0。CBAM-EfficientNet-B0 的结构图如图 6 所示。

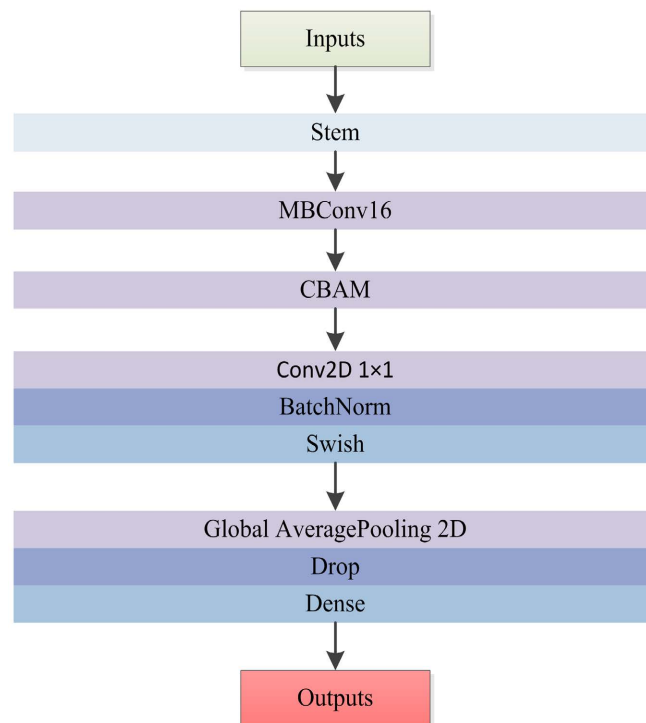


Figure 6. CBAM-EfficientNet-B0 network structure diagram  
图 6. CBAM-EfficientNet-B0 网络结构图

如图 6 所示, CBAM-EfficientNet-B0 主要由 EfficientNet-B0 和 CBAM 模块组成。在 CBAM-EfficientNet-B0 中, EfficientNet-B0 模型用于提取磨粒特征, 将 CBAM 模块插入在 16 个 MBCConv 模块之后用于将提取到的特征进行细化。当输入磨粒图像的特征图  $F$  输入到 CBAM 模块时, CBAM 模块首先将特征图  $F$  发送到通道注意力模块进行处理, 得到输入磨粒图像的通道注意力特征图  $M_c$ 。然后, 将  $M_c$  与  $F$  相乘得到空间注意力模块所需的精炼磨粒特征图  $F'$ 。然后, 将  $F'$  送入空间注意力模块, 生成磨粒图像的空间注意力特征图  $M_s$ ,  $M_s$  乘以  $F'$  生成 CBAM 模块的输出磨粒图像特征图。该模型输入为图像分辨率为  $224 \times 224$  的三通道图像, 识别结果作为输出。



### 3. 实验过程及结果分析

#### 3.1. 数据集

本文创建了包括正常磨损、粘着磨损、磨料磨损、疲劳磨损和腐蚀磨损的五类磨损类型的磨粒图像数据集，一共包含 3076 张图像，按照 9:1 的比例划分训练集和测试集。磨粒数据集具有类间差异小、类内差异大的特点。各类磨损图像样本如图 7 所示。磨损类型标签的数据分布如表 1 所示。

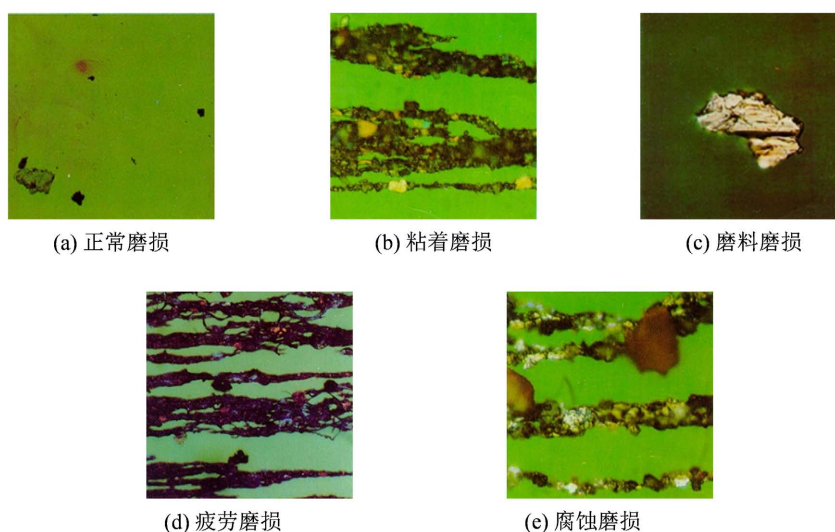


Figure 7. Sample images of five wear types  
图 7. 五种磨损类型样本图像

Table 1. Wear type label data distribution table

表 1. 磨损类型标签数据分布表

磨损类型	正常磨损	粘着磨损	磨料磨损	疲劳磨损	腐蚀磨损	总数
训练集	594	510	390	570	730	2794
测试集	58	56	41	53	74	282

#### 3.2. 实验环境

本文算法运行系统为 Windows10, 采用深度学习开源框架 PyTorch, 程序使用 Python 语言进行编写, 并使用 GPU 处理器加快模型的训练速度, 本文所使用的计算机配置如表 2 所示。在此硬件环境下, 基于 PyTorch 框架实现 CBAM-EfficientNet-B0 网络, 算法运行的软件环境如表 3 所示。

Table 2. Computer configuration used in this article

表 2. 本文所使用的计算机配置

名称	参数
设备名称	LAPTOP-3UQ8BH10
CPU 型号	Intel i7-10870H
内存容量	16GB
硬盘容量	1T

Continued

显卡芯片	NVIDIA GeForce RTX 2070 with Max-Q Design
显存容量	16GB
核心/线程数	八核心/十六线程

**Table 3.** Software operating environment  
**表 3.** 软件运行环境

名称	版本
操作系统	Windows10
显卡驱动	Cuda10
Python 模块管理工具	Anaconda3
深度学习框架	PyTorch1.6
PyTorch 工具包	Torch Vision
图像处理库	OpenCV2
编程语言	Python3.7

### 3.3. 评价指标

为了评价提出的磨损类型识别模型 CBAM-EfficientNet-B0 的分类性能, 本文采用准确率(Accuracy)、混淆矩阵(Confusion matrix)、精确率(Precision)、召回率(Recall)和 F1 分数(F1-score)来作为评价标准[10]。

对于图像分类模型, 识别准确率是非常重要的评价指标。识别准确率为全部分类正确的样本数除以总的样本数量的概率值, 计算表达式如公式(1)所示。

$$\text{Accuracy} = \frac{P_c}{P_{ALL}} \times 100\% \quad (1)$$

其中,  $P_c$  表示被正确预测的样本数量,  $P_{ALL}$  表示测试集的总样本数量。

混淆矩阵的作用是帮助分析每个类别的每个类别的误分类情况。混淆矩阵大小为(n, n)的方阵, 其中 n 表示类别的数量。该矩阵的行和列分别表示预测值和真实值。

精确率表示测试集中被预测为正样本的数量中被正确预测的概率。召回率表示测试样本中的正样本被正确预测的概率。精确率和召回率的计算表达式如公式(2)和公式(3)所示。

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

其中 TP、FP 和 FN 分别代表真阳性、假阳性和假阴性。真阳性表示识别类型为阳性, 预测为阳性; 假阳性表示识别类型为阴性, 预测为阳性; 假阴性表示识别类型为阳性, 预测为阴性。精度和召回率是相互独立的, 精度高意味着误识别率低, 召回率高意味着漏识别率低。

F1 分数为精确率和召回率的调和平均数, 表示了图像识别模型的精确程度, 该指标综合考虑了分类模型的性能, 其表达式如公式(4)所示。

$$F1 = \frac{2\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$



### 3.4. 实验设置及结果分析

网络的训练流程如图 8 所示。首先对输入图像进行预处理，使其分辨率置为  $224 \times 224$ ，并结合 ImageNet 预训练模型对系统初始化。训练过程预设为 200 轮 epoch，采用余弦退火[11]衰减策略进行学习率的调整。余弦退火算法促进学习率周期变化。设置最大学习率为  $1e-3$ ，最小学习率为  $1e-6$ 。前 100 轮 epoch 的 Batchsize 设置为 32，后 100 轮 epoch 的 Batchsize 设置为 16。为进一步提升模型的性能，试验时采用迁移学习和在线数据增强策略，并采用随机梯度下降优化器[12]来优化损失函数。

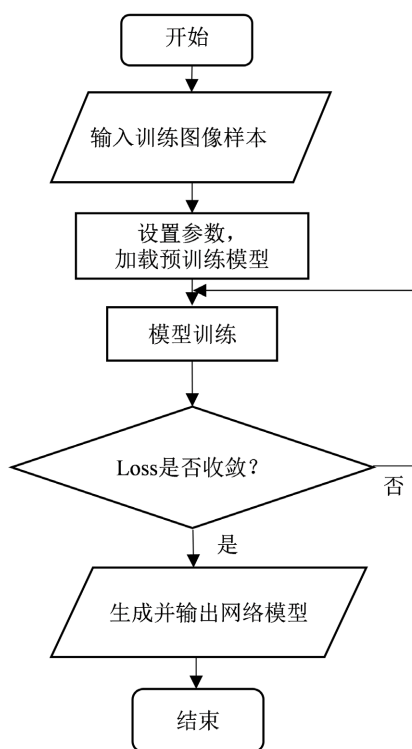


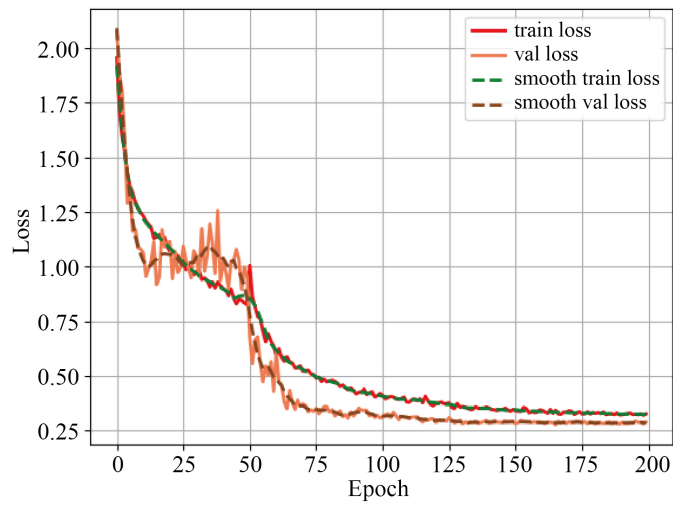
Figure 8. Network training flowchart  
图 8. 网络训练流程图

模型训练的损失函数曲线如图 9 所示。从图 9 可以看出，本文提出的算法训练过程中 Loss 值逐渐平稳，最后达到收敛。

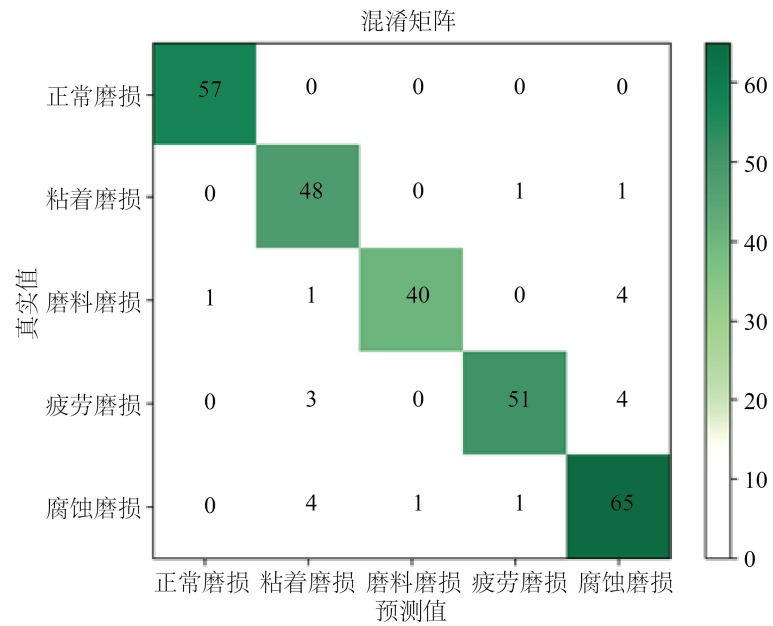
接下来使用磨损类型测试数据集图像评估 CBAM-EfficientNet-B0 模型的识别性能。用于在测试数据集中识别磨损类型的 CBAM-EfficientNet-B0 的混淆矩阵如图 10 所示。

从混淆矩阵可以清晰地看出各类磨损情况，磨损类型的误识别主要发生在粘着磨损和腐蚀磨损之间。这是由于磨粒数据集中粘着磨损和腐蚀磨损之间的差异极小且细微。对 CBAM-EfficientNet-B0 在测试集上的定量评估结果如表 4 所示。

由表可知，提出的 CBAM-EfficientNet-B0 对测试数据集中的磨损类型识别取得了很好的效果。从单个类型识别的角度来看，CBAM-EfficientNet-B0 对 5 种磨损类型的识别精确率高于 86.96%，召回率高于 85.71%，F1 值高于 90.57%。识别效果最好的磨损类型为正常磨损，CBAM-EfficientNet-B0 对正常磨损识别的精确率、召回率和 F1 值分别为 100%、98.28% 和 99.13%。从总体品种识别的角度来看，我们发现 CBAM-EfficientNet-B0 对测试数据集识别的总体准确率为 92.55%，平均精确率为 92.49%，平均召回率为 93.12%，平均 F1 值为 92.64%，表明 CBAM-EfficientNet-B0 具有较稳定的识别所有磨损类型的能力。



**Figure 9.** Loss curve of model training  
**图 9.** 模型训练的 Loss 曲线



**Figure 10.** Confusion matrix of CBAM-EfficientNet-B0  
**图 10.** CBAM-EfficientNet-B0 的混淆矩阵

**Table 4.** CBAM-EfficientNet-B0 quantitative evaluation result  
**表 4.** CBAM-EfficientNet-B0 定量评估结果

磨损类型	精确率/%	召回率/%	F1 值/%	准确率/%
正常磨损	100.00	98.28	99.13	
粘着磨损	96.00	85.71	90.57	
磨料磨损	86.96	97.56	91.95	92.55
疲劳磨损	87.93	96.23	91.89	
腐蚀磨损	91.55	87.84	89.66	

为了更好地评估 CBAM-EfficientNet-B0 模型的识别性能,在磨粒测试数据集上将 CBAM-EfficientNet-B0 模型与其他经典分类网络 MobilenetV3、Resnet50、VGG16、ViT 和 EfficientNet-B0 进行了比较。本文平均精确度、平均召回率、准确率作为评价指标,使用不同模型的定量评价结果如表 5 所示。

**Table 5.** Quantitative evaluation results of different models  
**表 5.** 不同模型的定量评价结果

模型种类	平均精确率/%	平均召回率/%	准确率/%
MobilenetV3	82.21	79.84	80.42
Resnet50	83.90	83.08	83.22
VGG16	80.43	80.94	80.07
ViT	81.89	80.91	81.12
EfficientNet-B0	89.29	89.97	90.04
<b>CBAM-EfficientNet-B0</b>	<b>92.49</b>	<b>93.12</b>	<b>92.55</b>

从表 5 中可以清楚地看出, EfficientNet-B0 与 MobilenetV3 [13]、Resnet50 [14]、VGG16 [15]和 ViT [16]相比, EfficientNet-B0 对磨损类型的平均精确度、平均召回率、准确率都是最高的,说明 EfficientNet-B0 对磨损类型的识别能力更强。EfficientNet-B0 对磨损类型的总体识别准确率为 90.04%,比 MobilenetV3、Resnet50、VGG16、ViT 分别提高了 9.62%、6.82%、9.97%和 8.92%。本文算法基于 EfficientNet-B0 进行改进,改进后较改进前的平均精确度、平均召回率、准确率均有所提高,分别提升了 3.20%、3.15%和 2.51%。这是因为在网络中加入 CBAM 注意力机制起到了显著提高磨损类型识别的作用。

#### 4. 结论

铁谱磨粒图像的智能识别和分类能够反映磨损状态,为设备的状态维修和故障诊断提供依据。在本研究中,数据集包含五种磨损类型的磨粒图像,包括正常磨损、粘着磨损、磨料磨损、疲劳磨损和腐蚀磨损。本文利用基于深度学习的计算机视觉技术,提出了磨损类型识别模型 CBAM-EfficientNet-B0 对五种磨损类型进行识别。经实验测试,本文提出的 CBAM-EfficientNet-B0 能够准确识别 5 种磨损类型。在测试数据集上与对比实验中使用的其他模型相比, CBAM-EfficientNet-B0 在评价指标上具有明显的优势。

该研究为深度学习技术在设备的状态维修和故障诊断中的应用提供了新的有效策略和理论基础,为磨损类型的自动识别提供了技术支持。建议今后的工作可以进一步丰富磨损类型数据集,比如疲劳磨损又分为点蚀、剥落和剥层三种表现形式,细分种类为精准故障诊断研究提供充分的数据基础。在未来的研究中,将优化 CBAM-EfficientNet-B0 模型的速度,并尝试在移动端进行模型部署。

#### 参考文献

- [1] 关浩坚, 贺石中, 李秋秋, 杨智宏, 覃楚东, 何伟楚. 卷积神经网络在装备磨损颗粒识别中的研究综述[J]. 摩擦学学报, 2022, 42(2): 426-445.
- [2] Wang, S., Wu, T., Wang, K. and Sarkodie-Gyan, T. (2020) Ferrograph Analysis with Improved Particle Segmentation and Classification Methods. *Journal of Computing and Information Science in Engineering*, **20**, Article ID: 021001. <https://doi.org/10.1115/1.4045291>
- [3] Jardine, K., Lin, D. and Banjevic, D. (2006) A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance. *Mechanical Systems and Signal Processing*, **20**, 1483-1510. <https://doi.org/10.1016/j.ymsp.2005.09.012>
- [4] Li, Q., Zhao, T., Zhang, L., Sun, W. and Zhao, X. (2017) Ferrography Wear Particles Image Recognition Based on Ex-

- tre Learning Machine. *Journal of Electrical and Computer Engineering*, **2017**, Article ID: 3451358. <https://doi.org/10.1155/2017/3451358>
- [5] Peng, P. and Wang, J. (2019) FECNN: A Promising Model for Wear Particle Recognition. *Wear*, **432-433**, Article ID: 202968 <https://doi.org/10.1016/j.wear.2019.202968>
- [6] Jia, F., Wei, H., Sun, H., Song, L. and Yu, F. (2022) An Object Detection Network for Wear Debris Recognition in Ferrography Images. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, **44**, Article No. 67. <https://doi.org/10.1007/s40430-022-03375-4>
- [7] 向进. 改进分水岭和灰靶理论在铁谱图像分析中的研究[D]: [硕士学位论文]. 南京: 南京航空航天大学, 2016.
- [8] Tan, M. and Le, Q. (2019) Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, Vol. 97, Long Beach, 9-15 June 2019, 6105-6114.
- [9] Woo, S., Park, J., Lee, J.Y. and Kweon, I.S. (2018) Cbam: Convolutional Block Attention Module. *Proceedings of the European Conference on Computer Vision*, Munich, 8-14 September 2018, 3-19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- [10] Lever, J., Krzywinski, M. and Altman, N. (2016) Classification Evaluation: It Is Important to Understand Both What a Classification Metric Expresses and What It Hides. *Nature Methods*, **13**, 603-605. <https://doi.org/10.1038/nmeth.3945>
- [11] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., *et al.* (2016) SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision*, Amsterdam, 11-14 October 2016, 21-37. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
- [12] Goyal, P., Dollár, P., Girshick, R., Noordhuis, P., Wesolowski, L., Kyrola, A., *et al.* (2017) Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. arXiv preprint arXiv: 1706.02677.
- [13] Howard, A., Sandler, M., Chu, G., Wang, W., Chen, L.-C., Tan, M., *et al.* (2019) Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/ICCV.2019.00140>
- [14] Mukti, I.Z. and Biswas, D. (2019) Transfer Learning Based Plant Diseases Detection Using ResNet50. *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, 20-22 December 2019, 1-6. <https://doi.org/10.1109/EICT48899.2019.9068805>
- [15] Qassim, H., Verma, A. and Feinzimer, D. (2018) Compressed Residual-VGG16 CNN Model for Big Data Places Image Recognition. *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, 8-10 January 2018, 169-175. <https://doi.org/10.1109/CCWC.2018.8301729>
- [16] Kolesnikov, A., Dosovitskiy, A., Weissenborn, D., Weissenborn, D., Zhai, X., Unterthiner, T., *et al.* (2021) An Image Is Worth  $16 \times 16$  Words: Transformers for Image Recognition at Scale. arXiv:2010.11929.