

## 基于改进EfficientNetV2的有害垃圾图像分类方法

原慧琳, 刘军涛, 黄碧, 韩真, 冯宠

### 引用本文

原慧琳, 刘军涛, 黄碧, 韩真, 冯宠. 基于改进EfficientNetV2的有害垃圾图像分类方法[J]. 计算机科学, 2022, 49(11A): 211100100-5.

YUAN Hui-lin, LIU Jun-tao, HUANG Bi, HAN Zhen, FENG Chong. [Classification Method of Harmful Garbage Images Based on Improved EfficientNetV2](#) [J]. Computer Science, 2022, 49(11A): 211100100-5.

---

### 相似文献推荐 (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)

##### [基于机器学习的剩余使用寿命预测实证研究](#)

Empirical Research on Remaining Useful Life Prediction Based on Machine Learning  
计算机科学, 2022, 49(11A): 211100285-9. <https://doi.org/10.11896/jsjcx.211100285>

##### [基于差分进化算法的字符对抗验证码生成方法](#)

Adversarial Character CAPTCHA Generation Method Based on Differential Evolution Algorithm  
计算机科学, 2022, 49(11A): 211100074-5. <https://doi.org/10.11896/jsjcx.211100074>

##### [融合多层次视觉信息的人物交互动作识别](#)

Human-Object Interaction Recognition Integrating Multi-level Visual Features  
计算机科学, 2022, 49(11A): 220700012-8. <https://doi.org/10.11896/jsjcx.220700012>

##### [R-YOLOv5:自动切割的旋转的文本检测模型](#)

R-YOLOv5:Auto-cutting, Rotated Text Detection Model  
计算机科学, 2022, 49(11A): 210900185-6. <https://doi.org/10.11896/jsjcx.210900185>

##### [基于注意力机制的手写体数字识别](#)

Handwritten Digit Recognition Based on Attention Mechanism  
计算机科学, 2022, 49(11A): 211100009-5. <https://doi.org/10.11896/jsjcx.211100009>

# 基于改进 EfficientNetV2 的有害垃圾图像分类方法

原慧琳<sup>1</sup> 刘军涛<sup>2</sup> 黄碧<sup>2</sup> 韩真<sup>2</sup> 冯宠<sup>2</sup>

1 东北大学秦皇岛分校管理学院 河北 秦皇岛 066004

2 东北大学信息科学与工程学院 沈阳 110819

(1000289@nequ.edu.cn)

**摘要** 随着工业的迅速发展,垃圾产生的数量呈爆炸式上升,使得垃圾处理成为一个世界性的难题。我国政府对环境的关切也逐渐加深,不断推出了各种垃圾分类政策以及法律法规以监督市民进行垃圾分类。垃圾处理特别是有害垃圾处理,如电子废弃物等,如果处理不当,会对环境产生恶劣影响。有害垃圾图像数据具有数据质量低、图像不清晰的特点,采集来自不同设备的图像样本又有明显差异,因此有害垃圾图像处理面临巨大挑战,同时有害废弃物分类结果关系到环境污染问题,且目前产出的垃圾数量巨大,要求具有较高的处理速度和准确度。文中提出了一种基于卷积神经网络和注意力机制的垃圾图像分类方法。该方法不需要对输入的图像进行手工提取特征,通过深度学习模型框架,弥补传统图像处理算法的不足,实现对有害垃圾准确、高效的分类,可以较好地识别多种类型的有害垃圾。经实验验证,所提方法在 harmful-waste 数据集上的准确率达到 97.46%,相比其他算法模型,其模型训练时间更短,性能更优。利用深度学习的方法,部署自动化垃圾分类模型,对于环境保护有重要意义。

**关键词:** 神经网络;注意力机制;深度学习;垃圾分类;图像分类

中图分类号 TP391

## Classification Method of Harmful Garbage Images Based on Improved EfficientNetV2

YUAN Hui-lin<sup>1</sup>, LIU Jun-tao<sup>2</sup>, HUANG Bi<sup>2</sup>, HAN Zhen<sup>2</sup> and FENG Chong<sup>2</sup>

1 College of Management, Northeastern University at Qinhuangdao, Qinhuangdao, Hebei 066004, China

2 College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

**Abstract** With the rapid development of industry, the amount of waste generated has also exploded, making waste disposal a worldwide problem. The Chinese government's concern for the environment has gradually deepened, and various garbage classification policies and laws and regulations have been continuously introduced to supervise citizens' garbage classification. Garbage disposal, especially hazardous garbage such as electronic waste, if it is improperly handled, will result in bad influence. Hazardous garbage image data has the characteristics of low data quality and unclear images. The image samples collected from different devices have obvious differences. Therefore, the image processing of hazardous garbage faces huge challenges. At the same time, the classification results of hazardous waste are related to environmental pollution issues. The amount of waste produced is huge, requiring high processing speed and accuracy. This paper proposes a garbage image classification method based on convolutional neural network and hybrid attention mechanism. This method does not need to manually extract features from the input image. Through the deep learning model framework, it overcomes the shortcomings of traditional image processing algorithms, achieves accurate and efficient classification of hazardous waste, and can better identify multiple types of hazardous waste. Experiment shows that the proposed method has an accuracy rate of 97.47% on the harmful-waste data set, the model training time is shorter, and its performance is better than other algorithm models. Using deep learning methods to deploy automated garbage classification models is of great significance to environmental protection.

**Keywords** Convolutional neural network, Attentional mechanism, Deep learning, Garbage classification, Image classification

## 1 引言

近年来,随着城市化的发展和居民生活水平的提高,生活垃圾数量急剧增加,全球许多城市都面临着垃圾处理危机。垃圾分类回收任务是减少废物产生、缓解环境和改善整个国家经济的正确途径<sup>[1]</sup>。特别是电子废弃物等有害垃圾,由于

含有重金属等,需要精细分类以免影响环境。分类任务的效率和质量在很大程度上取决于分类原材料的纯度和准确性。然而,回收任务通常需要大量劳动力成本,计算机视觉和深度学习(Deep Learning, DL)技术有助于自动检测和分类回收任务的废物类型。DL是人工智能的一个分支,它可以解决视觉识别和分类问题,而卷积神经网络(Convolutional Neural Net-

基金项目:东北大学产学研战略合作项目(71971050)

This work was supported by the Northeastern University "Industry University Research Strategic Cooperation Project"(71971050).

通信作者:刘军涛(1822061948@qq.com)

work, CNN)是目前主要的图像识别方法的 DL 体系结构之一。

Yang 等<sup>[2]</sup>收集了总共 2527 张图像以创建废物图像数据集,即 TrashNet。使用支持向量机(Support Vector Machine, SVM)和支持向量机的 ResNet50 模型对垃圾网中的垃圾图像进行分类,两个模型的准确率分别为 63%和 87%。除垃圾网外,很少建立废物数据集,如 AquaTrash 和 VN 垃圾<sup>[3]</sup>,并且这些数据集存在一些缺点,例如特定环境中的废物量相对较少,而且数据集不是开源的。此外,该研究将 ResNext<sup>[4]</sup>作为 CNN 的基础模型,并结合多项改进,对无机废物和医疗废物进行自动分类。配备基于 CNN 的算法和数字传感器的垃圾分类设备可以实现 90%左右的垃圾分类精度。然而,尽管由 CNN 驱动的现有废物分类能够有效工作,但废物分类的准确性仍有待提高。且大型网络复杂度高,训练时间长、推理速度慢等特点导致其无法应用于在嵌入式设备中,难以落地实施。近年来,许多学者开始研究基于 MobileNet, ShuffleNet<sup>[5-6]</sup>等的轻量级网络,实现垃圾实时分类。

自开创性的 AlexNet<sup>[7]</sup>被提出以来,许多学者不断研究以进一步改善 CNN 的性能,提出了 GSop-Net, CBAM, SANet<sup>[8-10]</sup>等注意力模块。Squeeze-and-Excitation Networks (SENet)<sup>[11]</sup>是由自动驾驶公司 Momenta 在 2017 年公布的一种全新的图像识别结构,通过学习 channel 之间的相关性,筛选出了针对通道的注意力。将 SENet 嵌入到其他分类模型中,计算量有小幅提升,但是效果提升明显。CBAM 是集成通道和空间注意力机制的模块,可以将其嵌入到卷积神经网络,该模块从空间和通道两个方向生成注意力特征图。

本文研究以轻量级的卷积神经网络模型 EfficientNetV2<sup>[12]</sup>为主干网络,提出了一种基于改进的基于深度学习的有害垃圾分类方法。改进后的网络,浅层更新了基本模块,采取 CBAM 注意力机制,能够有效提取浅层的空间信息。对于深层更加抽象的特征信息,依然采用 SENet 抑制部分干扰。该方法无须人工提取特征,网络模型通过算法自动提取特征,弥补了传统图像分类的不足,可实现对垃圾图像的准确、高效分类。

## 2 相关工作

### 2.1 ResNet

ResNet<sup>[13]</sup>由微软实验室于 2015 年提出,致力于解决深度卷积神经网络模型退化的问题,在 2015 年取得了 ILS-VRC-2015 的冠军。ResNet 系列网络使用残差模块构建网络单元,这些单元包括一系列卷积层和池化层,并且提出 Batch Normalization 加速训练。在本文研究中,ResNet 采用  $224 \times 224$  大小的图像输入。

### 2.2 MobileNetV3

MobileNetV3<sup>[14]</sup>是谷歌团队于 2019 年提出的轻量级神经网络,该网络结合了 MobileNet 的深度可分离卷积、MobileNetV2<sup>[15-16]</sup>的 Inverted Residuals 和 Linear BottleNeck、SE 模块,利用神经结构搜索来搜索网络的配置和参数。MobileNetV3 提供了两种配置,分别适用于对资源要求不同的情况。在 ImageNet 分类任务中,MobileNetV3 相比 MobileNetV2 精度提高了大约 4.6%,耗费时间减少了 5%。本文研究采用 MobileNetV3-s 作为对比,输入图像尺寸为  $224 \times 224$ 。

### 2.3 EfficientNetV2

在之前的研究中,研究人员主要关注的是准确率和参数数量,但近年来,一些学者开始关注网络的训练速度和推理速度。2021 年,谷歌团队提出了 EfficientNetV2<sup>[17-18]</sup>,该网络在测试准确率和参数数量上都优于之前的网络,在 ImageNet 公开数据集测试网络的特征提取部分由浅层的 Fused-MBConv 模块和深层的 MBConv<sup>[19]</sup>模块组成。其结构如图 1 和图 2 所示。根据输入的尺寸不同,EfficientNet-v2 的 3 种配置分别为 EfficientNetV2-s, EfficientNetV2-m 和 EfficientNetV2-l,其输入图像尺寸依次递增。为了加快训练速度,本研究采用 s 配置进行实验,其训练输入尺寸为  $300 \times 300$ 。

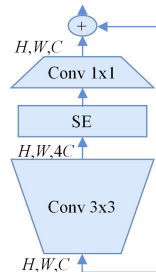


图 1 Fused-MBConv 模块结构

Fig. 1 Structure of Fused-MBConv

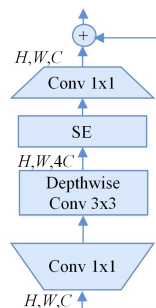


图 2 MBConv 模块结构

Fig. 2 Structure of MBConv

## 3 基于改进 EfficientNetV2 的有害垃圾图像分类方法

本研究以搭建低消耗高精度的有害垃圾分类模型为目的,以轻量级网络 EfficientNetV2 为主干网络。在网络浅层考虑到丰富的空间信息,使用通道和空间注意力机制,并更新了浅层模块。在训练过程中,采用迁移学习的方法,加快训练速度。

### 3.1 通道和空间注意力机制

CBAM<sup>[20]</sup>是集成通道和空间的注意力机制模块(见图 3),可以将其嵌入到卷积神经网络中,并且参数数量和计算量只有小幅提升。给定一个  $F \in \mathbb{R}^{C \times H \times W}$  的特征图作为输入,CBAM 模块会依次生成一个  $M_c \in \mathbb{R}^{C \times 1 \times 1}$  的通道注意力特征图和一个  $M_s \in \mathbb{R}^{1 \times H \times W}$  的空间注意力特征图。整体注意力模块的运算可概括为:

$$F = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

其中,  $\otimes$  代表点积运算,  $F'$  为原始输入的特征图  $F$  经过通道注意力子模块的输出,  $F''$  为  $F'$  经过空间注意力机制输出的注意力图。下面介绍每个注意模块的细节信息。

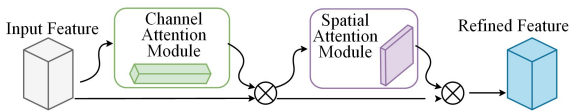


图3 通道和空间注意力机制

Fig. 3 Channel and attention mechanisms

在通道注意力子模块中(见图4),首先使用最大池化和平均池化操作聚合特征图的空间信息,生成两个不同的空间上下文描述: $F_{avg}^c$ 和 $F_{max}^c$ ,分别表示平均池化和最大池化得到的特征描述。然后,将这两个特征描述输入到多层感知机模型进行计算,得到相应的特征向量。两个特征向量融合后,在使用 sigmoid 函数处理,最终得到通道注意力特征图 $M_c$ 。通道注意力计算过程可总结如下:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (3)$$

其中, $\sigma$ 表示 sigmoid 函数, $MLP$ 为多层感知机前向计算函数, $W_0$ 和 $W_1$ 为两层线性层的权重, $F_{avg}^c$ 和 $F_{max}^c$ 分别为输入特征图 $F$ 经过平均池化和最大池化运算后的结果。

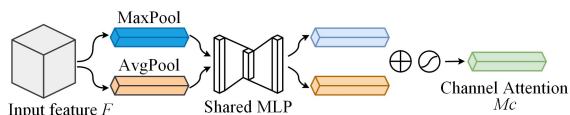


图4 通道注意力机制

Fig. 4 Channel attention mechanism

空间注意力机制如图5所示。首先利用平均池化和最大池化操作压缩通道信息,生成二个特征图: $F_{avg}^s \in \mathbb{R}^{1 \times H \times W}$ 和 $F_{max}^s \in \mathbb{R}^{1 \times H \times W}$ 。然后通过一个标准卷积层将它们连接并进行卷积运算,生成二维的空间注意力图。其计算过程如下:

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (4)$$

其中, $\sigma$ 表示 sigmoid 函数; $f^{7 \times 7}$ 表示卷积操作,其卷积核大小为7。

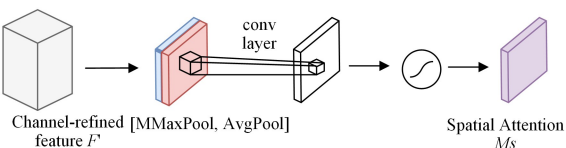


图5 空间注意力机制

Fig. 5 Spatial attention mechanism

### 3.2 C-MBConv 模块

原始的 EfficientNetV2 由浅层的 Fused-MBConv 和深层的 MBConv<sup>[21]</sup>组成。由于浅层特征图含有大量的空间信息,因此在本研究中浅层使用 GBAM 注意力机制,能够同时兼顾通道和空间方向上的细节特征。为此,我们将 GBAM 机制嵌入 Fused-MBConv 中,并将新的模块命名为 C-MBConv,详细结构如图6所示。

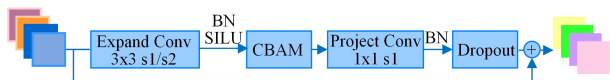


图6 C-MBConv 模块结构

Fig. 6 Structure of C-MBConv

### 3.3 基于混合注意力机制的 EfficientNetV2

使用 C-MBConv 模块取代 EfficientNetV2 原有的浅层

Fused-MBConv 模块,搭建出了基于混合注意力机制的 EfficientNetV2,即 Hybrid-Attention-EfficientNetV2 (HA-EfficientNetV2),其模型整体结构如图7所示。新的网络在浅层不仅考虑通道上的重要特征,而且能有效捕获到大量的空间特征信息。实验证明,使用混合注意力策略能够有效提高分类准确率,与其他主流网络的性能相比更有竞争力。

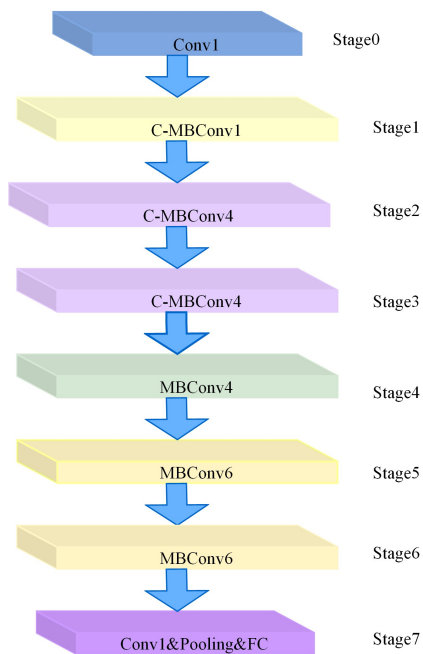


图7 HA-EfficientNetV2 结构

Fig. 7 Structure of HA-EfficientNetV2

### 3.4 迁移学习

迁移学习利用模型在相似任务中的学习权重,避免模型从零开始学习,能够使模型获得更优的初始化权重参数。本研究利用 EfficientNetV2 在 ImageNet 数据集上的预训练权重作为模型训练的初始化参数权重,能够显著提升模型训练速度。

使用迁移学习的方法为下载官方开源模型权重。对于改进模型的部分,由于结构已经发生改变,无法导入原模型的初始化参数,在导入权重时应跳过。

## 4 实验与结果

由于垃圾分类工作缺乏公开的标准数据集,本研究中采用课题组收集的数据集在不同网络上进行对比实验。经实验验证,本研究方法能够对有害垃圾实现快速、准确的分类。

### 4.1 数据集与实验环境

本文采用 Harmful-waste 数据集进行研究,该数据集是由本课题组通过网络爬虫和实地拍摄完成。数据集包括8种常见的有害垃圾图像,表1所列为每种有害垃圾的数量分布。从表中可以看出,数据集存在严重的数据量不平衡问题,干电池数量远远多于其他样本数量。为了解决数据集不平衡带来的问题,本文先对数据进行均衡化处理,并通过随即裁剪和翻转的方式对所有数据集进行简单的扩充。在扩充前将90%的数据划分出来充当训练集,其余数据用于测试。

实验平台配置如下:CPU为AMD R7 4800H,搭载NVIDIA 1650显卡,操作系统为Windows 10,python版本为3.7,使用的深度学习框架为pytorch 1.7。

表1 Harmful-wsate 数据集样本分布

Table 1 Sample distribution of Harmful-waste dataset

垃圾名称	数目/张	垃圾名称	数目/张
纽扣电池	220	杀虫剂	355
干电池	224	铅蓄电池	306
灯管	329	电池板	229
胶水	237	药片	210

## 4.2 实验结果与分析

本研究选取当下流行的多种神经网络进行实验对比,以证实本模型的有效性。对比实验模型包括 ResNet, ResNext, MobileNetV3 等,分别使用上述模型于本研究的自建数据集 Harmful-waste 中进行训练。在训练过程中,记录每一个训练周期的训练集损失和测试集准确率,以便及时掌握模型的训练情况,确保各个模型在收敛的状态下完成训练。

各个模型在 Harmful-waste 数据集上的训练结果曲线如图 8 和图 9 所示。图中,  $E$  为训练的迭代周期,  $A$  为测试准确率,  $L$  为训练损失。从训练曲线可知,虽然 EfficientNetV2 网络采用了轻量级的架构,但它在有害垃圾分类任务上仍然取得了优异成绩,训练结果优于 MobileNetV3, ResNet50, ResNext50<sup>[4]</sup> 等。本研究算法基于 EfficientNetV2 网络进行了改进,准确率分别比 ResNet101, EfficientNetV2 提高了 0.6% 和 1.2%。原因如下:1) EfficientNetV2 本身是优秀的轻量级网络,在训练速度、推理速度和测试准确率等方面优于之前的网络;2) 本文算法模型在网络浅层引入 CBAM 结构,在通道及空间维度上集成了注意力机制,使得重要的特征得到加强,而不重要的特征得到抑制。相比原始的 EfficientNetV2,网络在增加一定参数的前提下,能够捕获浅层更丰富的空间特征信息;3) 使用迁移学习的策略,利用模型在大规模数据集中的权重,在相似任务中能够更好地初始化网络参数,进而提升模型的性能。

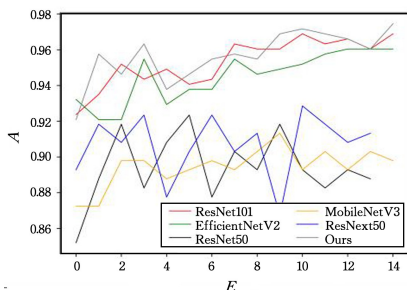


图 8 各模型在本研究测试集上的准确率变化曲线

Fig. 8 Accuracy curve of each model in test set

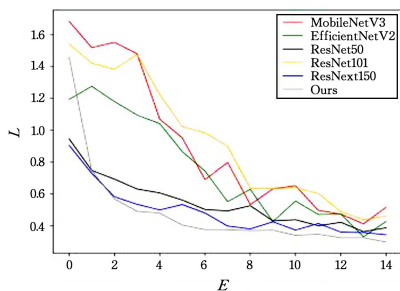


图 9 各模型在本研究测试集上的训练损失曲线

Fig. 9 Training loss curve of each model in test set

为了进一步分析本文研究算法在其他分类任务场景下的性能提升效果,添加公共数据集 Stanford Dogs Dataset 和

CUB\_200\_2011 进行对比实验。在测试集上的分类精度如表 2 所列。表中  $A_s$  为模型在数据集 Stanford Dogs Dataset 上的测试准确率,  $A_c$  为模型在 CUB\_200\_2011 数据集上的测试准确率。可以看出,本研究算法在 Stanford Dogs Dataset 和 CUB\_200\_2011 数据集上的测试准确率分别为 88.56% 和 87.05%, 仅次于 ResNet101 模型。EfficientNetV2 的准确率低于 ResNet101 的原因是 ResNet101 网络太深,模型复杂,在大型数据集上往往能取得更好的学习效果。但 ResNet101 网络体积过大,不能保证实时性,在计算资源有限的场景下难以部署。因此,本文研究算法模型同样适用于其他分类任务,在复杂的场景下仍能保持较高准确率。

表 2 各模型在公共数据集中的测试准确率

Table 2 Test accuracy of each model in common data set

models	$A_s$	$A_c$
ResNet50	78.65	77.58
ResNext50	80.49	80.58
ResNet101	89.32	88.04
MobileNetV3	81.13	66.81
EfficientNetV2	86.62	86.45
Ours	<b>88.56</b>	<b>87.05</b>

在有害垃圾分类任务中,各模型在本研究自建数据集上的测试实验结果如表 3 所列。表中  $T_i$  为单张图片推理时间,  $T_r$  为使用迁移学习后训练网络到收敛所需时间。可见本研究算法的准确率最高,训练时长和单张图片推理时间比原始网络有小幅增加,具有较好的实时效果。

表 3 各模型在本研究自建数据集训练结果对比。

Table 3 Comparison of training results of each model on self-built dataset in this paper

models	$A/\%$	$T_i/\text{ms}$	$T_r/\text{min}$
ResNet50	92.35	145	12.5
ResNext50	93.89	130	16.3
ResNet101	96.93	265	26.2
MobileNetV3	91.82	58	10.5
EfficientNetV2	96.32	65	13.5
Ours	<b>97.47</b>	<b>89</b>	<b>16.7</b>

将本文算法预测结果进行可视化测试,选取一组常见的真实场景下的有害垃圾图像,将图像传入模型中进行识别,结果如图 10 所示。



图 10 有害垃圾图像识别效果

Fig. 10 Harmful waste images recognition effect

图 10 中括号内数值是识别结果的置信度。可以看出,在不同环境下,模型能够准确地识别出有害垃圾种类。因此,本文研究提出的改进的 EfficientNetV2 模型是高精度、低消耗的网络模型。

**结束语** 针对传统图像分类的准确度不高,本文研究通过改进 EfficientNetV2 模型,结合迁移学习的技术,实现了对有害垃圾的快速、准确分类。相比以往流行的深度学习模型,本文研究中改进算法在 Harmful-waste 数据集上实现了最高的测试准确度,且训练时间和推理时间都较短。但实际应用中,有害垃圾的种类繁多,形式各种各样,因此建立更全面的数据集十分必要,这是下一步研究的方向之一。另一方面,部分有害废弃物外观特征较为相似,如果处理不当会对环境造成恶劣影响,细分多分类识别面临的挑战更加艰巨,研究也更有价值,这也是我们下一步需要进行的研究。

### 参 考 文 献

- [1] MA A, MAT B, SF A, et al. Artificial intelligence applications in solid waste management: A systematic research review [J]. Waste Management, 2020, 109: 231-246.
- [2] YANG M, THUNG G. Classification of trash for recyclability status[J]. CS229 Project Report, 2016, 2016(1): 3.
- [3] VO A H, LE H S, VO M T, et al. A Novel Framework for Trash Classification Using Deep Transfer Learning [J]. IEEE Access, 2019, 7: 178631-178639.
- [4] XIE S, GIRSHICK R, DOLLÁR P, et al. Aggregated Residual Transformations for Deep Neural Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1492-1500.
- [5] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [J]. arXiv: 1704. 04861, 2017.
- [6] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6848-6856.
- [7] 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.
- [8] GAO Z, XIE J, WANG Q, et al. Global Second - Order Pooling Convolutional Networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2020.
- [9] 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.
- [10] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks [C]//2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 2235-2239.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [12] TAN M, LE Q V. EfficientNetV2: Smaller Models and Faster Training[C]//International Conference on Machine Learning. PMLR, 2021: 10096-10106.
- [13] 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.
- [14] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2020.
- [15] GAVAI N R, JAKHADE Y A, TRIBHUVAN S A, et al. MobileNets for flower classification using TensorFlow[C]//2017 International Conference on Big Data, IoT and Data Science (BIGDATA). IEEE, 2017: 154-158.
- [16] 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 (CVPR). IEEE, 2018.
- [17] SUNIL C K, JAIDHAR C D, PATIL N. Cardamom Plant Disease Detection Approach Using EfficientNetV2 [J]. IEEE Access, 2021, 10: 789-804.
- [18] LIU D, WANG W, WU X, et al. EfficientNetv2 Model for Breast Cancer Histopathological Image Classification[C]//2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI). IEEE, 2022: 384-387.
- [19] KOUMPAROULIS A, POTAMIANOS G. Accurate and Resource-Efficient Lipreading with Efficientnetv2 and Transformers[C]//2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022: 8467-8471.
- [20] MA B, WANG X, ZHANG H, et al. CBAM-GAN: generative adversarial networks based on convolutional block attention module[C]//International Conference on Artificial Intelligence and Security. Cham: Springer, 2019: 227-236.
- [21] QU H, YANG J, SHEN M, et al. Fault diagnosis of rolling bearing under time-varying speed conditions based on EfficientNetv2 [J]. Measurement Science and Technology, 2022, 33(6): 065023.



**YUAN Hui-lin**, born in 1969, Ph.D, professor. Her main research interests include modeling and optimization of complex systems, information retrieval, and artificial intelligence.



**LIU Jun-tao**, born in 1996, postgraduate. His main research interests include machine learning and artificial intelligence.