

基于改进Resnet18的垃圾分类收运监管方法研究

何彦弘¹, 徐怡宁², 傅嘉琪¹, 陈书航¹, 李俊峰¹

(1.浙江理工大学信息科学与工程学院, 浙江 杭州 310018;

2.浙江理工大学计算机科学与技术学院, 浙江 杭州 310018)

✉2019330301193@mails.zstu.edu.cn; 2019329621218@mails.zstu.edu.cn; 2019330301190@mails.zstu.edu.cn;

2019330361008@mails.zstu.edu.cn; ljf2003@zstu.edu.cn



摘要: 为解决垃圾分类收运过程中, 由于错误收运导致的混装率高、垃圾分类效果下降的问题, 文章提出了一种基于改进Resnet18的垃圾分类收运监管方法。首先, 在Resnet18残差结构的始端引入SE-Net通道注意力机制模块, 可以有效提升网络的特征提取能力; 其次, 采用一种基于空洞卷积的多尺度感受野融合模块, 使网络能够融合不同尺度的特征信息。实验结果表明, 残差结构始端和晚期融合的改进方案效果更佳, 加权 $F1$ 值较原Resnet18分别提升了9.26%和7.36%。改进后的模型加权的 $F1$ 值达97.27%, 较原Resnet18分别提升了10.26%, 并且对不同颜色垃圾桶识别的准确率更高。此外, 改进后的Resnet18的每秒检测帧数(FPS)达到65.03, 可以满足监管实时性的要求。同时, 采用数据增强和天气模拟的方法处理数据集, 使模型能适应多种环境, 提升了模型的鲁棒性。

关键词: Resnet18; 垃圾分类收运监管; SE-Net; 多尺度感受野融合; 扩张卷积

中图分类号: TP391.4 **文献标识码:** A

Waste Classified Collection and Transportation Supervision Approach based on Improved Resnet18

HE Yanhong¹, XU Yining², FU Jiaqi¹, CHEN Shuhang¹, LI Junfeng¹

(1.School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China;

2.School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

✉2019330301193@mails.zstu.edu.cn; 2019329621218@mails.zstu.edu.cn; 2019330301190@mails.zstu.edu.cn;

2019330361008@mails.zstu.edu.cn; ljf2003@zstu.edu.cn

Abstract: In order to solve the problem of high mixed loading rate and poor waste classification effect caused by incorrect collection and transportation in the process of waste classification and transportation, this paper proposes a waste classification and transportation supervision approach based on improved Resnet18. Firstly, SE-Net (Squeeze-and-Excitation Networks) channel attention mechanism module is introduced to the beginning part of residue structure of the Resnet18, which can effectively improve the network feature extraction capability. Secondly, a multi-scale receptive field fusion module based on dilation convolution is adopted to enable the network to feature information of different scales. The experimental results show that the improved model of infusion of the beginning part and the late part of the residual structure performs better, and the weighted $F1$ value increases by 9.26% and 7.36% respectively compared with the original Resnet18. The weighted $F1$ value of the improved model reaches 97.27%, which is 10.26% higher than that of the original Resnet18, and the recognition accuracy of different color trash cans is higher. In addition, the number of frames per second (FPS) of the improved Resnet18 reaches 65.03, which can meet the requirements of real-time monitoring. At the same time, data argumentation and weather simulation methods are used to process the dataset, so that the proposed model can adapt to a variety of environment and improve the robustness of the model.

Keywords: Resnet18; waste classification and transportation supervision; SE-Net; multi-scale receptive field fusion; dilated convolution

1 引言(Introduction)

随着经济的发展、城市化进程的推进及人们生活水平的提高, 垃圾产生量迅速增加, 从而增加了城市垃圾收集、运

输及垃圾处理系统的压力^[1]。2020 年我国城市固废量达310.9 百万吨, 同比增长1.5%。预计2022 年我国城市固废处理量将达到349.3百万吨, 大量的垃圾严重阻碍了城市发展与人们的

健康生活,因此研究如何更好地进行垃圾的分类回收对解决上述问题具有重要意义。垃圾分类是指按一定规定或标准将垃圾分类储存、投放和运输从而转化成公共资源的一系列活动的总称。垃圾分类可以分为三个步骤:前端分类、中端清运和后端回收利用。

近年来,机器学习迅速发展,在很多领域得到广泛应用,很多学者提出了采用机器学习方法辅助垃圾分类^[2-4],FENG等^[5]提出了一种基于Inception-v3迁移学习模型的办公室垃圾智能分类方法,WANG等^[6]提出了一种基于机器视觉的垃圾分类和回收系统,通过机器人对垃圾进行智能识别,根据垃圾类别进行分类回收。但是,上述研究的应用场合分别为办公室垃圾和散落的建筑垃圾,不适用于产生量更大的生活垃圾的分类和回收,在这个领域具有一定的局限性;赵珊等^[7]提出了基于MobileNetV2网络和隐式特征金字塔网络(IPFN)改进的单步多框目标检测器(SSD)垃圾实时分类检测,提高了对小目标的检测能力。

同时,刘南杉等^[8]提出了利用人工智能虚拟现实等技术对垃圾分类进行整体规划检测;CHUNG等^[9]提出了一种智能分类和环境监测系统实现垃圾自动分类和环境监测;YU等^[10]提出了利用人工智能技术,提高方程式废物处理的准确性和效率;CHEN等^[11]提出了一种基于“互联网+”的城市垃圾分类回收与运输体系框架,给出了垃圾分类运输系统智能化和信息化的解决方案,但并未提出如何对垃圾分类运输过程形成有效监管以提升垃圾分类的效果。

上述研究为应用计算机技术实现垃圾分类提供了重要的参考,但上述研究的重点均在垃圾分类前端,而在将机器视觉技术运用于垃圾的清运过程的监管方面有所欠缺。通过走访某市的环境卫生和生固废处置保障中心得知,垃圾收集和运输过程中混装率极高,存在“前端分类,后端混装”的现象。清运混装垃圾会挫伤市民实施垃圾分类的积极性,增大垃圾后续处理的难度,同时会影响垃圾分类效果,因此对该过程进行有效的监管是非常有必要的。清运过程会受到作业环境、时间等多种因素的影响,因此对其进行监管的难度大。若依赖于人工监控,工作量大且效率低。如果利用垃圾清运车上安装的工业摄像头对垃圾清运图像进行实时采集,并应用图像处理等技术进行分析,用计算机替代人工进行实时监控,就能有效提高对清运过程的监管效率,减少不规范清运和垃圾混装现象的发生,从垃圾分类的中端提升垃圾分类效果。结合深度卷积网络的相关研究^[12-15],本文提出一种基于改进Resnet18的垃圾桶颜色分类与匹配算法,解决垃圾清运过程中的监管问题。首先,通过在Resnet18添加SE-Net通道注意力机制模块对模型实现改进,并提出了一种多尺度感受野融合模块,添加在Resnet18中构成晚期融合用于提升网络性能;其次,提取垃圾车作业期间的监控视频的关键帧构建垃圾桶分类数据集;最后,利用改进后的Resnet18对上述

数据集进行训练,从而实现对垃圾清运的有效监管。

2 垃圾清运监管问题分析(Problem analysis of waste clearing and transportation supervision)

经过调研和访谈发现,垃圾清运过程主要存在以下问题。

(1)装车过程中挂载的垃圾桶数量超量。按照垃圾清运作业规定,垃圾车作业时,每次只能将两个垃圾桶挂在车上,并将桶内垃圾倒入清运车内,这是为了保证垃圾桶的垃圾不掉落地面上,保证作业场地的清洁;保障收运工人的作业安全,以免产生安全隐患。但在实际作业现场,作业人员为了提升工作效率而挂载三个垃圾桶,存在违规操作行为。

(2)垃圾桶盖未完全打开。在垃圾清运作业时,为了确保桶内的垃圾能够被清运干净,要求将垃圾桶盖完全打开后再进行作业。但在实际作业现场,作业人员为了方便而没有人工打开部分闭合的垃圾桶盖,从而影响了垃圾清运的效果。

(3)垃圾桶颜色(对应垃圾类别)与垃圾清运车收运垃圾类别不匹配。为了实现垃圾的分类清运,通常会将分类后的不同类型垃圾装在对应该颜色的垃圾桶中并采用不同的垃圾清运车运送。这就要求垃圾桶的颜色(即垃圾的类别)要与垃圾清运车收运的垃圾类型相匹配。但是实际收运过程中,常有不匹配的垃圾桶被错误清运的问题,导致垃圾的混装混运,降低了垃圾分类效果。

针对上述问题,本文提出了一种垃圾清运监管解决方案(图1),可实现对垃圾分类清运过程中的有效监管,具体可以分成两个部分实施:一是使用目标检测网络实现对垃圾桶及其状态的识别;二是使用分类网络对检测到的垃圾桶的类型进行判断。



图1 垃圾清运监管解决方案

Fig.1 Solution to waste clearing and transportation supervision

本文就“垃圾桶颜色(对应垃圾类别)与垃圾清运车收运垃圾类别不匹配”的问题,提出了一种基于改进Resnet18的解决方案,该方法可以对目标检测网络检测到的垃圾桶目标加以识别分类。按照不同的收运垃圾类型,本文对垃圾桶做出以下分类,分别为蓝色垃圾桶(可回收垃圾)、绿色垃圾桶(厨余垃圾)、灰色垃圾桶(其他垃圾)、红色垃圾桶(有害垃圾)、黄

色垃圾桶(其他垃圾)和未知颜色垃圾桶(未知垃圾),由于在清晨或夜晚等作业条件下的光线不足和垃圾袋反光问题,以及部分垃圾过大对垃圾桶造成遮挡,造成难以识别的问题,因此添加了未知颜色这一类别)。

3 Resnet18及其改进方法(Resnet18 and its improving method)

3.1 Resnet18

深度残差网络是由HE等^[16]于2016年提出的,残差网络结构如图2所示。该研究证明了引入跳层连接能有效解决梯度弥散导致的模型难以收敛和深度神经网络退化两大问题,并使计算轻量化。

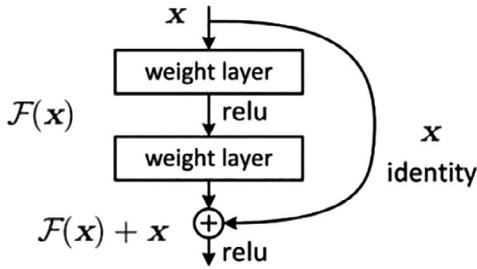
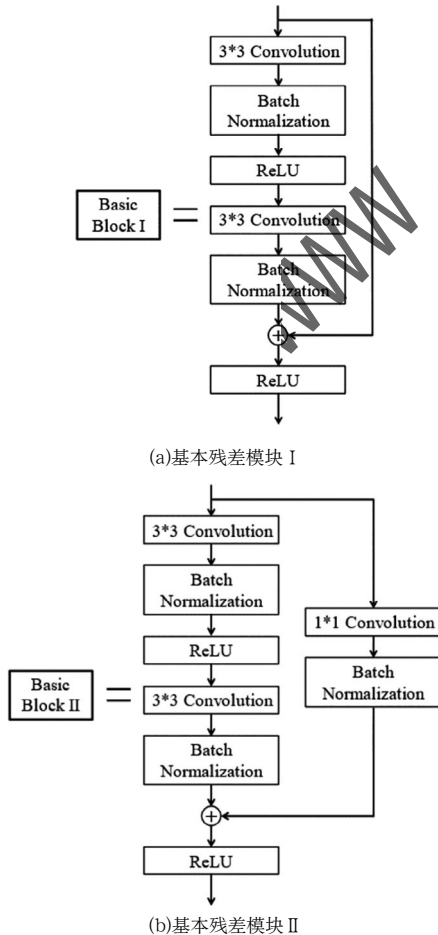


图2 残差网络基本结构

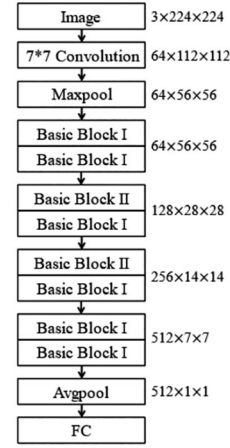
Fig.2 Basic structure of Resnet network

Resnet18是一种基于上述残差结构构造的网络,其模型如图3所示。Resnet18的残差模块中首先进行一次 3×3 卷积,其次进行批标准化和ReLU激活,再次进行一次 3×3 卷积和批标准化,最后与输入特征图相加后进入ReLU激活。残差模块II在跳层连接中加入了 1×1 卷积进行下采样。



(a)基本残差模块 I

(b)基本残差模块 II



(c)整体网络结构

图3 Resnet18模型

Fig.3 Resnet18 model

3.2 SE-Net注意力机制

SE-Net(Squeeze-and-Excitation Networks)是由HU等^[17]提出的一种通道注意力机制,其主要原理是根据全局信息学习通道特征权重,对通道信息重标定从而增大有效通道特征的权重,减小无效或效果较小的通道特征的权重,提升模型特征提取能力,其结构如图4所示。

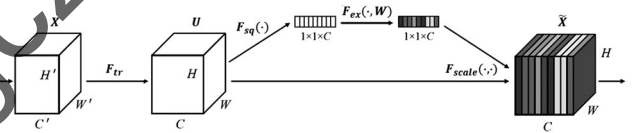
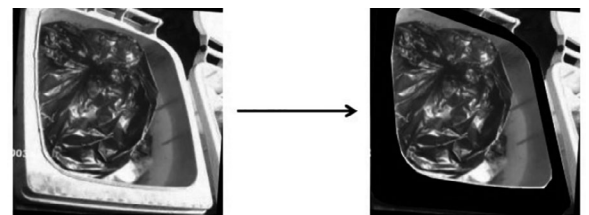


图4 SE-Net 结构

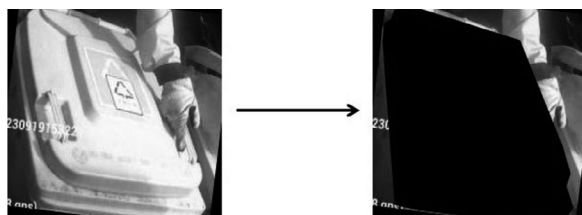
Fig.4 SE-Net structure

该算法可以分为三个部分,即通道特征压缩(Squeeze)、通道特征激发(Excitation)和通道权重重标定(Scale)。首先Squeeze操作通过全局平均池化操作将 C 通道 $H \times W$ 的特征图压缩成 C 通道 1×1 的特征图;然后Excitation操作对上一步结果进行全连接,得到 C/r 维的向量后进行ReLU激活,再对该结果进行一次全连接,将 C/r 维转换成 C 维向量,并通过Sigmoid激活,使其值分布在0—1;最后将通道权重加权到原始特征图上,完成对原始特征的重标定。

结合本文分类任务和图像数据特征,从图5中可以直观地看到,对垃圾桶分类识别较为关键的信息是闭盖垃圾桶的垃圾桶盖和开盖垃圾桶的垃圾桶边缘部分。通过SE-Net学习通道权重,理论上能抑制通道权值较小的特征,提升模型对通道的敏感性和模型的特征提取能力。



(a)开盖垃圾桶



(b)闭盖垃圾桶

图5 垃圾桶颜色识别关键信息(黑色部分)

Fig.5 Key information of trash can color

SE-Net常与Resnet残差网络按图6中的方法结合,但要注意在Resnet残差结构的不同位置加入SE-Net,可能会产生不同的效果。本文拟提出三种方案,分别将其插入在Resnet残差结构的始端(Pre-SE)、末端(Post-SE)和残差连接(Res-SE)处,其结构如图7所示。

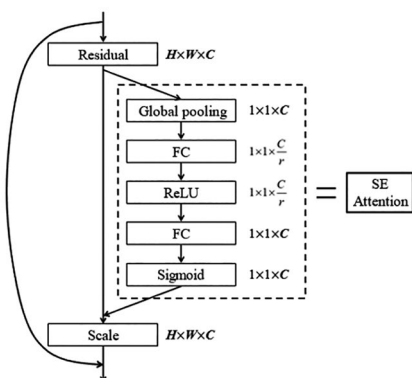
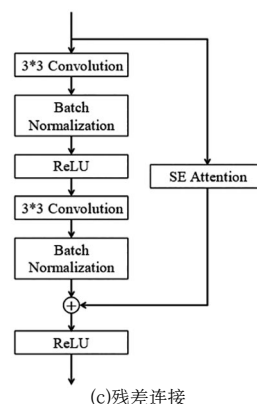


图6 Resnet与SE-Net结合方法

Fig.6 Combination of Resnet and SE-Net



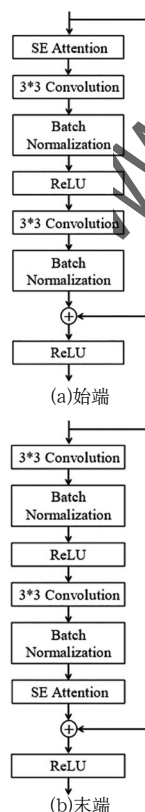
(c)残差连接

图7 三种SE-Net与Resnet结合方法

Fig.7 Three approaches to combining SE-Net and Resnet

3.3 多尺度感受野融合残差结构

Inception结构^[18]通过输入不同尺度的特征图预测不同大小的目标,受此启发,若对输入特征图进行多尺度融合,就能够增加模型对特征信息的感受野。结合本文图像数据特点(图8),可以直观地感受到多尺度融合的效果。本文数据集可以按照垃圾桶状态分为两大类,即开盖垃圾桶(Open Garbage Can)和闭盖垃圾桶(Close Garbage Can);对于开盖垃圾桶的颜色分类任务,最重要的信息位于垃圾桶图像的四周,而垃圾桶内的颜色对该任务影响甚微;而对于闭盖垃圾桶,其图像主体为垃圾桶颜色。进行多尺度融合后可以直观地发现(图8),开盖垃圾桶的特征信息增加了,而对于闭盖垃圾桶,多尺度融合导致其特征信息的减少是有限的。综上所述,理论上多尺度融合有利于提升本文数据集特征提取能力。



(a)始端

(b)末端



多尺度融合

(a)开盖垃圾桶



多尺度融合

(b)闭盖垃圾桶

图8 多尺度变换后的图像

Fig.8 Image of multi-scaled transformation

结合上述分析与残差网络的思想,拟提出一种针对本文数据集的多尺度感受野融合残差模块(Multi-Scaled Receptive Field Fusion Residual Module, MSRM),其结构如图9所示。

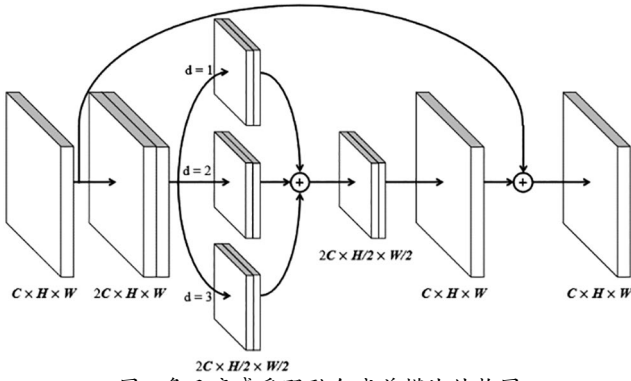


图9 多尺度感受野融合残差模块结构图

Fig.9 Residual module structure of multi-scaled receptive field fusion

上述多尺度感受野融合模块首先进行 1×1 卷积将特征图变换为 $2C \times H \times W$ ，然后采用空洞率不同的空洞卷积对特征图进行下采样和融合，最后利用 3×3 转置卷积进行上采样，将特征图的尺度变换为 $C \times H \times W$ 并与输入进行残差融合，最终得到输出特征图。需要注意的是，每次卷积后都进行了批标准化和ReLU激活。

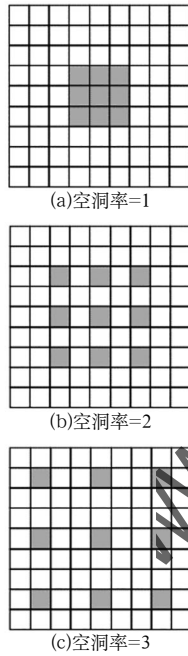


图10 不同空洞率的空洞卷积感受野比较

Fig.10 Receptive field comparison of the dilation convolution with different dilation rates

空洞卷积^[19]可以在保持卷积核大小不变的前提下扩大感受野，从图11中可以直观地感受到使用空洞率不同的空洞卷积能够提取到更大范围的信息，使得融合后的特征图具有更丰富的语义信息。

考虑到在中期和晚期融合的过程中，特征图自身的通道数较多，上述模块可能对网络性能有较大的影响，但普通的标准卷积可以由深度卷积和点卷积两个部分替代，因此使用深度可分离卷积^[20]重构上述模块中计算量较大的卷积，从而提升网络计算效率。

将上述模块插入Resnet18的不同位置分别构成早期融合

(Early-Fusion, EF)、中期融合(Mid-Fusion, MF)和晚期融合(Late-Fusion, LF)，具体如图11所示。

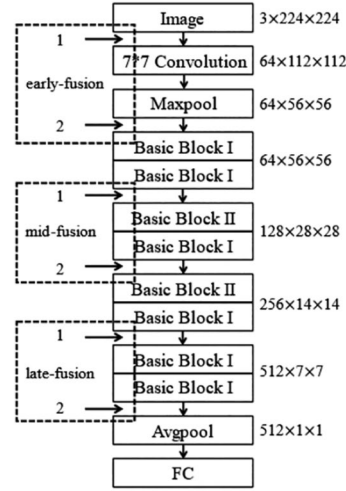


图11 不同位置的MSRM

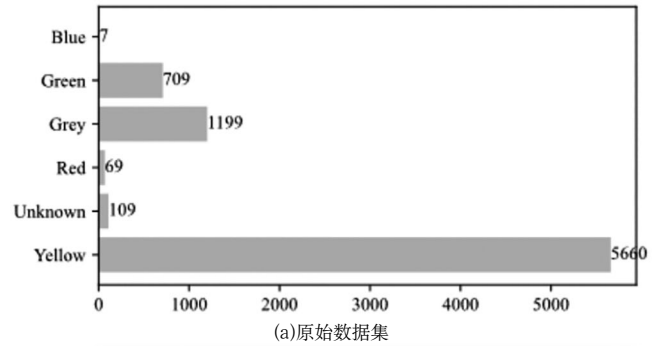
Fig.11 MSRM of different positions

4 实验与分析(Experiment and analysis)

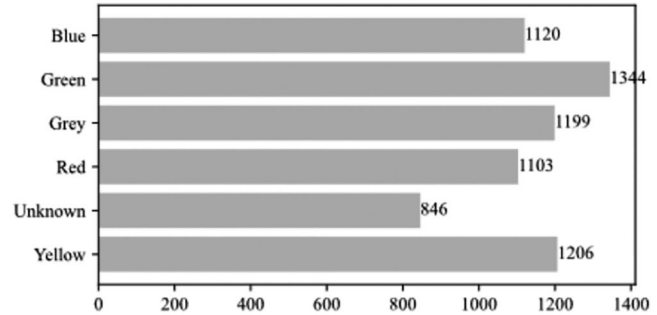
4.1 数据集

本文研究使用的图像来源于某市垃圾清运车搭载的工业摄像机拍摄的视频，通过提取视频关键帧和裁剪的方法构建数据集。原始图像共7,753张，具体分为6类，即黄色、灰色、红色、绿色、蓝色和未知颜色的垃圾桶。

由于不同颜色的垃圾桶使用的频率存在显著差异，导致数据集中样本分布失衡，原始数据集分布如图12(a)所示。为了使数据集分布均衡，通过人工重新拍摄和数据增强扩增数据集。最终数据集共包含6,818张图像，数据分布如图12(b)所示，统一缩放为 224×224 像素，并按照7:2:1的比例划分训练集、验证集和测试集，分别为4,773张、1,362张和683张。



(a)原始数据集



(b)调整后数据集

图12 数据集数据分布

Fig.12 Data distribution of the dataset

为了扩增数据集以提高模型的泛化能力，使其适应垃圾收运作业环境的复杂性，采用亮度调整、 $0^{\circ}-20^{\circ}$ 随机旋转、水平垂直翻转、高斯模糊和GridMask方法^[21]对数据进行增强。同时，本文采用开放计算机视觉库(OpenCV)模拟降雨、降雪和大雾等天气，提高模型在不同天气条件下分类的准确性。部分数据增强方法如图13所示，天气模拟如图14所示。



图13 数据集增强

Fig.13 Dataset augmentation

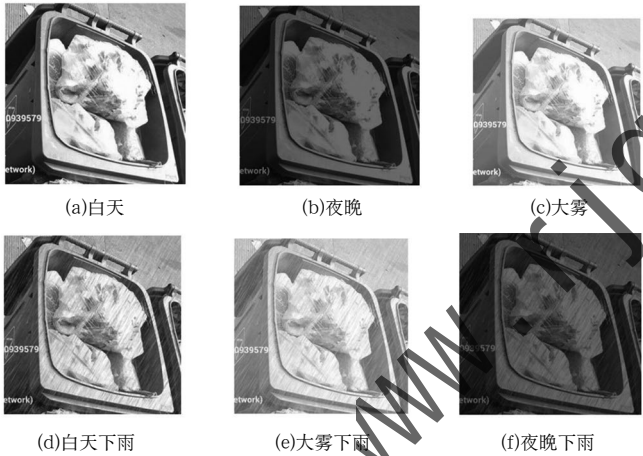


图14 模拟天气

Fig.14 Weather simulation

4.2 实验环境

本实验的硬件环境：处理器为Intel(R) Core(TM) i5-9300H CPU @ 2.40 GHz，内存为8 GB，显卡为Nvidia GTX 1650，操作系统为64 位，磁盘大小1 TB。

软件环境：操作系统为Windows 10，深度学习框架为Pytorch 1.10.1，开发环境为PyCharm，开发语言为Python，包管理器和环境管理器为Anaconda，GPU加速库为CUDA 11.3和cuDNN 8.2.1。

4.3 超参数设置

4.3.1 学习率

本文实验中，学习率采用具有重启机制的余弦退火调整策略^[22](Cosine Annealing with Warmup Restart)，该方法使学习率在一个周期内按照余弦函数随着迭代次数的增加而发生改变，同时通过热重启在设定的Epoch(一个

Epoch相当于使用训练集中的全部样本训练一次)增大学习率跳出局部最优解。通过预训练调整初始学习率和余弦退火策略的相关参数值，实验所用的学习率与迭代次数的关系如图15所示。

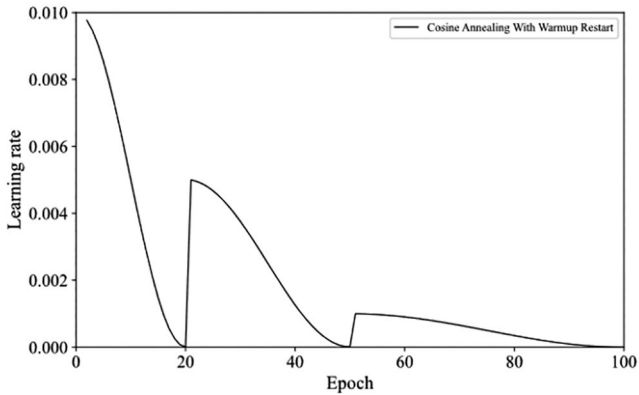


图15 余弦退火调整策略

Fig.15 Cosine annealing adjustment strategy

4.3.2 其他超参数

本实验输入的图像尺寸为 224×224 像素，训练100 Epoch，采用随机梯度下降(SGD)策略进行训练，初始学习率设为0.01，动量设置为0.9，衰减因子设置为0.000484，batch设置为16。

4.4 实验结果与分析

4.4.1 SE-Net实验

根据SE-Net通道注意力机制插入在Resnet18残差结构的不同位置，分别构成始端(Pre-SE)、末端(Post-SE)和残差连接(Res-SE)处的通道注意力机制。对三种不同位置的网络按照上述超参数设置进行实验，得到实验结果如表1所示。

表1 SE-Net实验具体结果

Tab.1 Specific results of SE-Net experiment

垃圾桶类别	模型	查准率/%	查全率/%	垃圾桶类别	模型	查准率/%	查全率/%
蓝色	Resnet18	96.46	93.17	红色	Resnet18	93.90	84.75
	Pre-SE	100.00	100.00		Pre-SE	99.06	92.95
	Post-SE	99.49	100.00		Post-SE	97.65	99.52
	Res-SE	98.48	100.00		Res-SE	98.59	95.02
绿色	Resnet18	93.22	96.46	未知颜色	Resnet18	50.00	77.88
	Pre-SE	96.95	98.62		Pre-SE	85.23	97.59
	Post-SE	98.64	99.32		Post-SE	85.80	92.07
	Res-SE	94.58	97.89		Res-SE	92.05	90.50
灰色	Resnet18	92.21	88.02	黄色	Resnet18	94.38	83.63
	Pre-SE	97.84	95.76		Pre-SE	97.59	95.67
	Post-SE	98.70	88.72		Post-SE	94.78	97.52
	Res-SE	97.84	93.00		Res-SE	93.98	97.91

考虑到不同颜色的垃圾桶在本研究的对象城市中的使用频率不同，因此对不同颜色的垃圾桶的查准率(查全率)进行加权作为该模型的加权平均查全(准)率。根据本文“4.1”中提

到的数据集，将原始数据集中不同颜色的垃圾桶的出现频率作为权值，得到以下加权算法：

$$\begin{aligned} \text{Weighted(Precision*Recall*)} = & 0.01 \times \text{Blue} + 0.09 \times \text{Green} + \\ & 0.16 \times \text{Grey} + 0.01 \times \text{Red} + \\ & 0.02 \times \text{Unknown} + 0.71 \times \text{Yellow} \end{aligned}$$

表2 SE-Net实验结果

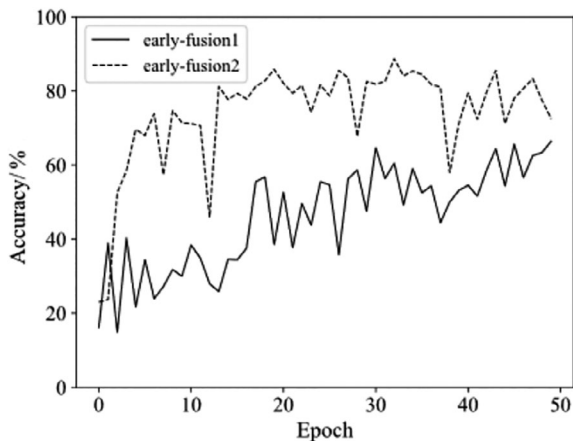
Tab.2 Results of SE-Net experiment

模型	查全率/加权查全率/%	查准率/加权查准率/%	F1值/加权F1值/%	模型大小/kB	FPS
Resnet18	93.06/86.69	85.48/87.32	89.11/87.01	43,771	77.92
Pre-SE	97.36/96.11	95.96/96.42	96.66/96.27	43,927	77.52
Post-SE	95.65/95.84	96.21/96.19	95.93/96.02	43,927	76.29
Res-SE	94.70/95.92	96.97/95.72	95.82/95.82	43,927	74.42

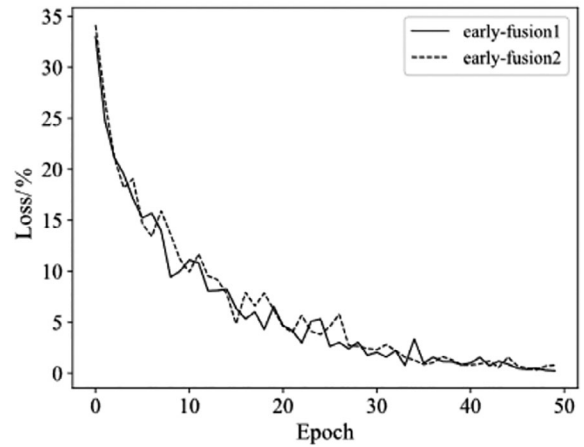
通过表2中的数据发现，加入SE-Net通道注意力机制后对模型的大小和检测速度影响较小，但能够在一定程度上提升网络的性能，从各个颜色分类的性能指标来看，对网络性能的提升较为全面。与Resnet18相比，是对未知颜色的垃圾桶有了更高的检测性能。Pre-SE、Post-SE和Res-SE分别较未改进的Resnet18在加权F1值上分别提升了9.26%、9.01%和8.81%。同时，加入SE-Net对模型的大小和每秒识别帧数(FPS)指标并没有太大的影响，该指标仅小幅度降低，但是综合上述指标对三种位置的SE-Net进行比较，显然在残差结构的始端(Pre-SE)加入SE-Net通道注意力机制对模型的性能有更大的提升。因此，本文选择Pre-SE作为一种改进方案。

4.4.2 MSRM实验

在Resnet网络结构的不同位置加入MSRM模块，分别形成早期融合、中期融合和晚期融合。对三种不同位置的网络按照上述超参数设置进行实验，实验结果表明，对于早期和中期融合，随着训练集Loss的降低，验证集上的精确率难以提升，如图16和图17所示，早期融合和中期融合会导致模型泛化能力恶化。同时，上述实验也发现MSRM加入在网络中的位置越深，对模型的影响逐渐从恶化转变为提升，晚期融合能在一定程度上提升网络的性能，尤其体现在对黄色垃圾桶和未知颜色垃圾桶的识别上。

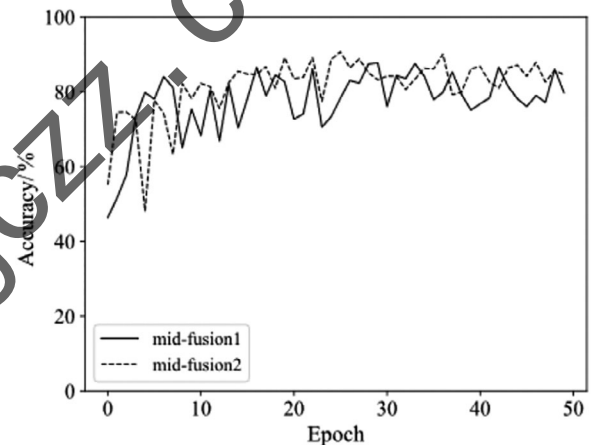


(a)准确率变化

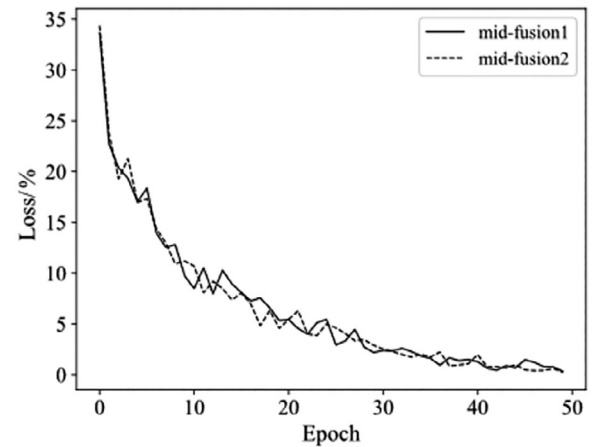


(b)损失值变化

图16 早期融合位置1和早期融合位置2训练准确率和损失
Fig.16 Training accuracy and loss of early-fusion at position 1 and 2



(a)准确率变化



(b)损失值变化

图17 中期融合位置1和位置2训练准确率和损失
Fig.17 Training accuracy and loss of mid-fusion at position 1 and 2

对网络性能有较明显效果的晚期融合进行进一步的数据分析，实验结果如表3和表4所示。

表3 MSRM 实验具体结果

Tab.3 Specific results of MSRM experiment

垃圾桶类别	模型	查准率/%	查全率/%	垃圾桶类别	模型	查准率/%	查全率/%
蓝色	Resnet18	96.46	93.17	红色	Resnet18	93.90	84.75
	LF-1	99.49	99.49		LF-1	86.85	98.93
	LF-2	99.49	100.00		LF-2	86.67	98.45
绿色	Resnet18	93.22	92.21	未知颜色	Resnet18	50.00	77.88
	LF-1	99.32	87.20		LF-1	89.77	90.29
	LF-2	95.25	97.23		LF-2	92.05	78.26
灰色	Resnet18	92.21	88.02	黄色	Resnet18	94.38	83.63
	LF-1	90.91	95.89		LF-1	96.79	97.57
	LF-2	85.28	96.57		LF-2	97.19	89.30

表4 MSRM 实验结果

Tab.4 Results of MSRM experiment

模型	查全率/加权查全率/%	查准率/加权查准率/%	F1值/加权F1值/%	模型大小/kB	FPS
Resnet18	93.06/86.69	85.48/87.32	89.11/87.01	43,771	77.92
EF-1	—	—	—	43,937	—
EF-2	—	—	—	45,868	—
MF-1	—	—	—	45,868	—
MF-2	—	—	—	47,629	—
LF-1	95.86/93.86	96.26/94.90	96.06/94.37	52,053	65.79
LF-2	94.95/93.16	91.15/93.30	93.02/93.23	76,710	73.12

对表3和表4中的数据进行分析，在具体分类上，MSRM模块对黄色垃圾桶和未知颜色垃圾桶识别的准确率有较明显的提升，LF-1和LF-2对于未知颜色垃圾桶的查全率分别提升了12.41%和0.38%；对黄色垃圾桶的查全率分别提升了13.94%和5.67%，但是对绿色、灰色和红色垃圾桶的识别的准确率均有不同程度的降低。从整体的实验结果来看，LF-1的加权查全率、加权查准率和加权F1值较Resnet18分别提升了7.17%、7.58%和7.36%，LF-2较Resnet18分别提升了6.47%、5.98%和6.22%。

但是，随着MSRM模块加入的位置越深，Resnet18加入节点的特征图通道数也随之增加，从而导致模型变大和检测速度降低。相比Resnet18，LF-1和LF-2的模型大小分别增加了18.92%和75.25%，检测速度也有所下降。

综上所述，实验结果表明，MSRM模块以晚期融合的方式加入网络的末端对网络性能有所提升，同时综合各种因素，本研究选择晚期融合在LF-1位置将MSRM模型加入网络中。

4.4.3 消融实验

结合实验“4.4.1”和“4.4.2”，本文改进后的Resne18网络如图18所示。

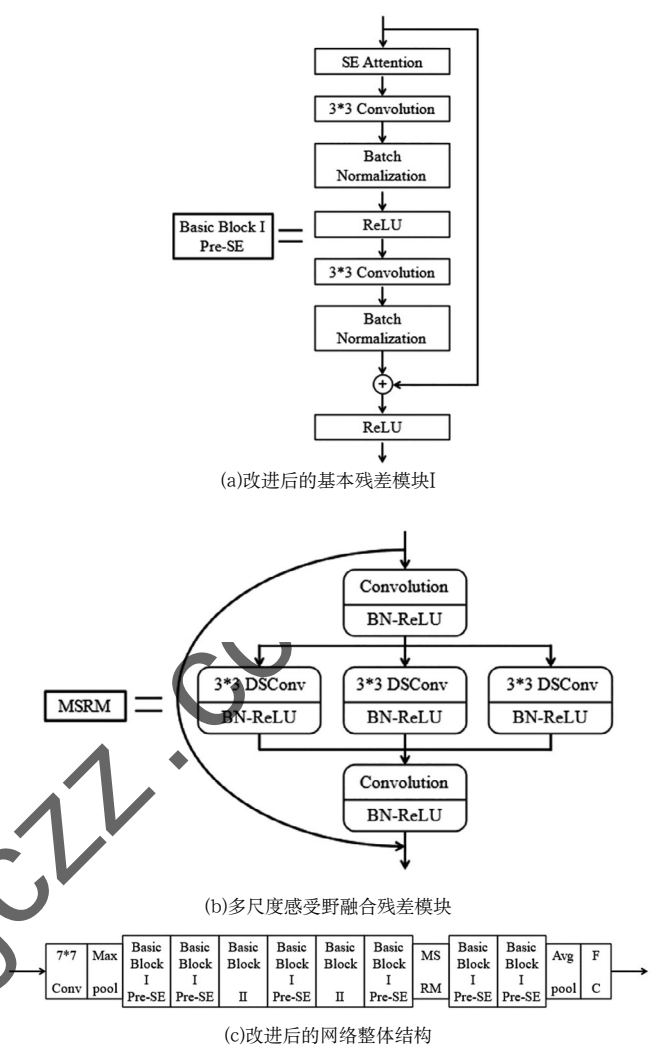


图18 改进的Resnet18

Fig.18 Improved Resnet18

对其进行消融实验以验证其效果，实验结果如表5和表6所示。

表5 消融实验具体结果

Tab.5 Specific results of ablation experiment

垃圾桶类别	模型	查准率/%	查全率/%	垃圾桶类别	模型	查准率/%	查全率/%
蓝色	Resnet18	96.46	93.17	红色	Resnet18	93.90	84.75
	Pre-SE	100.00	100.00		Pre-SE	99.06	92.95
	MSRM	99.49	99.49		MSRM	86.85	98.93
绿色	Both	100.00	100.00	未知颜色	Both	95.31	100.00
	Resnet18	93.22	96.46		Resnet18	50.00	77.88
	Pre-SE	96.95	98.62		Pre-SE	85.23	97.59
灰色	MSRM	99.32	87.20	黄色	MSRM	89.77	90.29
	Both	96.95	97.95		Both	94.89	94.35
	Resnet18	92.21	88.02		Resnet18	94.38	83.63
	Pre-SE	97.84	95.76		Pre-SE	97.59	95.67
	MSRM	90.91	95.89		MSRM	96.79	97.57
	Both	98.27	96.19		Both	97.99	95.31

表6 消融实验结果

Tab.6 Results of ablation experiment						
Pre-SE	LF1-MSRM	查全率/加权查全率/%	查准率/加权查准率/%	F1值/加权F1值/%	模型大小/kB	FPS
×	×	93.06/86.69	85.48/87.32	89.11/87.01	43,771	77.92
○	×	97.36/96.11	95.96/96.42	96.66/96.27	43,927	77.52
×	○	95.86/93.86	96.26/94.90	96.06/94.37	52,053	65.79
○	○	97.87/97.23	95.76/97.30	96.81/97.27	52,229	65.03

研究人员通过消融实验结果，发现Pre-SE和LF1-MSRM都能提升网络的性能，Pre-SE对网络整体性能都有所提升，LF1-MSRM的部分性能较Resnet18有所下滑，例如对红色垃圾桶查准率下滑7.05%，对灰色垃圾桶查准率下降1.3%，但对于未知颜色和黄色垃圾桶的查准率有较大幅度的提升。将上述两种模块叠加使用后，使网络性能有了较大的提升，尤其是对于未知颜色的垃圾桶，较单独的Pre-SE和LF1-MSRM查准率分别提升了9.66%和5.12%，但这是以牺牲一定的对其他颜色的检测性能为代价的，例如针对红色垃圾桶的查准率，改进后的Resnet18较单独使用Pre-SE改进的网络下降了3.75%。但从总体来看，改进后的Resnet18的性能是得到提升的，加权查全率、加权查准率和加权F1值较单独使用Pre-SE改进的网络分别提升了1.12%、0.88%和1%，较单独使用LF1-MSRM改进的网络分别提升了3.37%、2.4%和2.9%。

因此，改进后的网络的检测性能更加均衡，在牺牲少量速度的前提下，较未改进的Resnet18性能有了较大幅度的提升。

4.4.4 对比实验

将改进后的Resnet18与其他网络对比，得到结果如表7和表8所示。

表7 对比实验具体结果

Tab.7 Comparison of specific experiment results						
垃圾桶类别	模型	查准率/%	查全率/%	垃圾桶类别	模型	查准率/%
蓝色	Resnet18	96.46	93.17	蓝色	Resnet18	93.90
	Resnet34	93.43	93.43		Resnet34	87.32
	Resnet50	100.00	96.12		Resnet50	97.18
	Resnet101	94.95	93.53		Resnet101	90.14
	本文模型	100.00	100.00		本文模型	95.31
	Resnet18	93.22	96.46		Resnet18	50.00
绿色	Resnet34	90.85	96.75	绿色	Resnet34	64.77
	Resnet50	96.95	96.30		Resnet50	71.59
	Resnet101	93.90	96.18		Resnet101	73.86
	本文模型	96.95	97.95		本文模型	94.89
	Resnet18	92.21	88.02		Resnet18	94.38
	Resnet34	90.48	84.27		Resnet34	88.76
灰色	Resnet50	90.91	93.33	灰色	Resnet50	93.98
	Resnet101	90.04	90.83		Resnet101	92.77
	本文模型	98.27	96.19		本文模型	97.99

表8 对比实验结果

Tab.8 Comparison of experiment results					
模型	查全率/加权查全率/%	查准率/加权查准率/%	F1值/加权F1值/%	模型大小/kB	FPS
Resnet18	93.06/86.69	85.48/87.32	89.11/87.01	43,771	77.92
Resnet34	88.77/85.93	88.63/85.91	88.70/85.92	137,329	69.27
Resnet50	93.40/91.77	92.07/92.08	92.73/91.92	225,871	53.04
Resnet101	92.05/89.28	90.11/89.24	91.07/89.26	300,657	46.13
本文模型	97.87/97.23	95.76/97.30	96.81/97.27	52,229	65.03

从以上研究结果中可以看到，Resnet系列随着网络深度的加深，性能并没有得到较好的提升，检测速度降低明显。改进后的Resnet18的检测速度略低于Resnet34，但网络性能远超同类网络。

5 结论(Conclusion)

本文主要对Resnet18在垃圾分类收运和监管方面的应用展开了研究，并针对本文建立的垃圾桶数据集的特点，对Resnet18进行改进。通过上述实验，得了如下结论。

(1)在网络中加入SE-Net通道注意力机制能在几乎不影响检测速度的前提下提升模型性能，通道注意力机制提升了模型对通道的敏感性，提高了模型的特征提取能力。其中，将SE-Net加入Resnet残差结构的始端相比与加入末端和残差连接处具有更好效果。

(2)本文提出了一种多尺度感受野融合的残差模块，在垃圾桶数据集上有较为突出的表现，尤其是能够大幅度提升对未知颜色垃圾桶识别的准确率。同时，上述模块若加入在Resnet18的不同位置，则对模型性能有不同的影响，早期融合会导致模型性能恶化，使得训练过程难以收敛，晚期融合虽然在一定程度上降低了网络的检测速度，但是能够有效提升网络的性能。

(3)使用上述两种模块同时对Resnet18进行改进，通过消融实验表明，同时加入两种模块能够使网络性能更加均衡，并能有效提升网络对垃圾桶分类的正确率。

图像识别方法与深度学习应用于垃圾收运管理有着很大的研究空间和应用前景。下一步的研究方向是如何将目标检测网络和分类网络结合，或是集中嵌入智慧城市管理系统中或嵌入垃圾清运车车载设备上，从而实现对垃圾分类清运的实时监管和反馈，最终提高垃圾分类效果，实现更有效的资源再利用。

参考文献(References)

[1] CHEN S S, HUANG J L, XIAO T T, et al. Carbon emissions under different domestic waste treatment modes induced by garbage classification: case study in pilot communities in Shanghai, China[J]. Science of the Total Environment, 2020, 717:137193.

[2] CHEN Z C, YANG J, CHEN L F, et al. Garbage classification

- system based on improved ShuffleNet v2[J]. Resources, Conservation and Recycling, 2022, 178:106090.
- [3] LIU F C, XU H, QI M, et al. Depth-wise separable convolution attention module for garbage image classification[J]. Sustainability, 2022, 14(5):3099.
- [4] TANYA G, RAKSHIT J, DEVARSHI M, et al. A deep learning approach based hardware solution to categorise garbage in environment[J]. Complex & Intelligent Systems, 2022, 8(2):1129–1152.
- [5] FENG J W, TANG X Y. Office garbage intelligent classification based on inception-v3 transfer learning model[J]. Journal of Physics Conference Series, 2020, 1487(1):012008.
- [6] WANG Z, LI H, YANG X. Vision-based robotic system for on-site construction and demolition waste sorting and recycling[J]. Journal of Building Engineering, 2020, 32:101769.
- [7] 赵珊, 刘子路, 郑爱玲, 等. 基于MobileNetV2和IFPN改进的SSD垃圾实时分类检测方法[J]. 计算机应用, 2022, 42(S1):106–111.
- [8] 刘南杉, 裴云强, 蒋皓, 等. 基于VD-MobileNet网络的WebAR生活垃圾分类信息可视化方法[J]. 图学学报, 2022, 43(04):667–676.
- [9] CHUNG C Y, PENG I T, YEH J C. Intelligent classification and environment monitoring system combined with lora wireless transmission technology[C]// IEEE. 2018 International Symposium on Computer, Consumer and Control (IS3C). Los Alamitos, CA, USA: IEEE Computer Society, 2018:173–176.
- [10] YU K H, ZHANG Y, LI D, et al. Environmental planning based on reduce, reuse, recycle and recover using artificial intelligence[J]. Environmental Impact Assessment Review, 2021, 86:106492.
- [11] CHEN C, ZHOU M, HU P, et al. Research on the "Internet+" urban classified garbage recovery and transportation system framework[C]// EEEP. 2017 6th International Conference on Energy and Environmental Protection(ICEEP 2017). Paris, France: Atlantis Press, 2017:681–690.
- [12] LAGREE A, MOHEBPOUR M, METI N, et al. A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks[J]. Scientific Reports, 2021, 11(1):8025.
- [13] SU N, LIN Z S, YOU W L, et al. RMGCS: Real-time multimodal garbage classification system for recyclability[J]. Journal of Intelligent & Fuzzy System, 2022, 42(4):3963–3973.
- [14] MAHMUD T, RAHMAN M A, FATTAH S A. CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization[J]. Computers in Biology and Medicine, 2020, 122:219–229.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4):834–848.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[J/OL]. (2015–12–10)[2022–04–08]. <https://arxiv.org/abs/1512.03385>.
- [17] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42(8):2011–2023.
- [18] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[J/OL]. (2015–12–11)[2022–04–08]. <https://arxiv.org/abs/1512.00567>.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[J/OL]. (2016–04–30)[2022–04–08]. <https://arxiv.org/abs/1511.07122>.
- [20] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[J/OL]. (2017–04–04)[2022–04–08]. <https://arxiv.org/abs/1610.02357>.
- [21] CHEN P, LIU S, ZHAO H, et al. GridMask data augmentation[J/OL]. (2020–01–14)[2022–04–08]. <https://arxiv.org/abs/2001.04086>.
- [22] LOSCHILLOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts[J/OL]. (2017–05–03)[2022–04–08]. <https://arxiv.org/abs/1608.03983>.

作者简介:

何彦弘(2001–), 男, 本科生. 研究领域: 机器学习, 自动化技术.

徐怡宁(2000–), 女, 本科生. 研究领域: 计算机技术, 机器学习.

傅嘉琪(2001–), 男, 本科生. 研究领域: 电气工程, 自动化技术, 机器学习.

陈书航(2002–), 男, 本科生. 研究领域: 机器学习, 自动化技术.

李俊峰(1978–), 男, 博士, 副教授. 研究领域: 机器学习, 智能信息处理.