



(12)发明专利申请

(10)申请公布号 CN 111144496 A

(43)申请公布日 2020.05.12

(21)申请号 201911383091.0

(22)申请日 2019.12.27

(71)申请人 齐齐哈尔大学

地址 161006 黑龙江省齐齐哈尔市建华区
文化大街42号

(72)发明人 石翠萍 谭聪 苗凤娟 刘文礼
王天毅

(74)专利代理机构 哈尔滨市松花江专利商标事
务所 23109

代理人 时起磊

(51)Int.Cl.

G06K 9/62(2006.01)

G06N 3/04(2006.01)

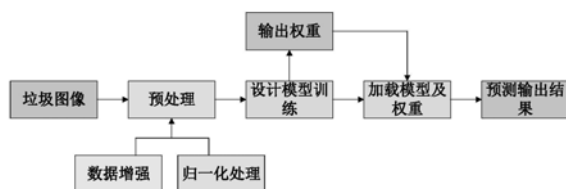
权利要求书2页 说明书14页 附图7页

(54)发明名称

一种基于混合卷积神经网络的垃圾分类方法

(57)摘要

一种基于混合卷积神经网络的垃圾分类方法,它属于垃圾分类回收技术领域。本发明解决现有方法对垃圾分类的精度低、且需要的训练时间长的问题。本发明的混合卷积神经网络模型灵活运用了卷积层、批标准化、最大池化层和全连接层,将BN批标准化用于每层卷积层和全连接层之后,进一步增强模型提取特征的能力,使各层作用充分发挥,进而得到较好的分类结果。利用BN层的正则化效果,适当的加入最大池化层对特征进行统计,减少特征维度,提高表征能力,且能够很好拟合,收敛速度快、参数量少、计算复杂度低,较传统卷积神经网络有明显优势。同时,模型采用SGDM+Nesterov的优化器,最终模型在图像上的分类准确率达到92.6%。本发明可以应用于生活垃圾分类。



1. 一种基于混合卷积神经网络的垃圾分类方法,其特征在于,该方法包括以下步骤:

步骤一、加载垃圾图像,并对垃圾图像进行预处理,获得预处理后的垃圾图像;

步骤二、构建混合卷积神经网络,将步骤一获得的预处理后垃圾图像输入混合卷积神经网络进行训练,以获得训练好的混合卷积神经网络;

所述混合卷积神经网络的结构为:

从混合卷积神经网络的输入端开始,混合卷积神经网络依次包括第一网络模块、第二网络模块、第三网络模块、第四网络模块、展平层、第一全连接层、第一全连接层后的批标准化层、第二全连接层、第二全连接层后的批标准化层以及Softmax分类器;

从所述第一网络模块的输入端开始,第一网络模块由两个基础模块以及一个最大池化层组成,其中:每个基础模块包括一个卷积层以及一个批标准化层;

其中:第二网络模块、第三网络模块和第四网络模块的结构均与第一网络模块的结构相同;

步骤三、将待分类的垃圾图像输入训练好的混合卷积神经网络,并加载训练好的混合卷积神经网络的权重,输出对待分类垃圾图像的分类结果。

2. 根据权利要求1所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述步骤一中,对垃圾图像进行预处理,获得预处理后的垃圾图像;其具体过程为:

通过对加载的垃圾图像进行数据增强和归一化处理,来获得预处理后的垃圾图像;

所述数据增强的方式包括对垃圾图像进行随机缩放、翻转、平移和旋转。

3. 根据权利要求1所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述第一网络模块的通道数为32;第二网络模块的通道数为64;第三网络模块的通道数为128;第四网络模块的通道数为256,第一全连接层的通道数为128,第二全连接层的通道数为64。

4. 根据权利要求3所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述混合卷积神经网络中每个卷积层均采用 3×3 的卷积核,且步幅都为1;每个最大池化层均采用 2×2 的过滤器,且步幅都为 2×2 。

5. 根据权利要求4所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述混合卷积神经网络采用的激活函数为Relu,优化器为SGDM+Nesterov,优化器的动量参数设置为0.9。

6. 根据权利要求5所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述将步骤一获得的预处理后垃圾图像输入混合卷积神经网络进行训练,以获得训练好的混合卷积神经网络,其具体过程为:

步骤1、设置初始学习率为0.01,在训练过程中添加监控保存损失函数值最低且精度值最高的模型机制;

步骤2、在当前学习率 lr 下,对混合卷积神经网络进行训练,当学习停滞时,若连续15次训练中模型的性能得不到提升,则学习率下降0.1,获得新的学习率 $lr \times 0.1$;

步骤3、重复步骤2的过程,新的混合卷积神经网络在新的学习率下继续训练;

步骤4、直至当前学习率对应的损失函数值开始增加、停滞改善时,则提前停止机制被激活,训练30次后停止训练,获得训练好的混合卷积神经网络。

7. 根据权利要求6所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所

述混合卷积神经网络中卷积层的每个卷积核提取出图像特征,即提取出特征图谱,提取出的特征图谱的具体计算公式如下:

$$y_j^l = \theta \left(\sum_{i=1}^{N_j^{l-1}} w_j \otimes x_i^{l-1} + b_j^l \right), j=1,2,\dots,M$$

其中, y_j^l 表示第j个特征图谱; w_j 表示对应的卷积核; x_i^{l-1} 表示上一层第i个特征图谱,即将上一层第i个特征图谱作为当前的输入; b_j^l 表示偏置量; N_j^{l-1} 表示每个特征图谱的特征数量; M 表示每个卷积层的特征图谱数; $\theta(\cdot)$ 为激活函数。

8.根据权利要求7所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述第一全连接层和第二全连接层的输出的表达式为:

$$h_{w,b}(x) = \theta \left(\sum_{i=1}^3 W_i x_i + b \right)$$

其中, $h_{w,b}(x)$ 表示全连接层的输出; x_i 表示上一层神经元的输出,即全连接层的输入; W_i 表示神经元之间连接的权值; b 表示偏置量, $\theta(\cdot)$ 表示激活函数。

9.根据权利要求8所述的一种基于混合卷积神经网络的垃圾分类方法,其特征在于,所述混合卷积神经网络的性能评估指标为精确率、回归率和F1-score值;

所述精确率P、召回率r以及F1-score值 F_1 的计算公式如下:

$$P = \frac{TP}{TP + FP}$$

$$r = \frac{TP}{TP + FN}$$

$$F_1 = \frac{2rP}{2 \times TP + FP + FN}$$

其中,TP表示将实际正样本数预测为正样本的个数,FN表示将实际正样本数预测为负样本的个数,FP表示将实际负样本预测为正样本的个数。

一种基于混合卷积神经网络的垃圾分类方法

技术领域

[0001] 本发明属于垃圾分类回收技术领域,具体涉及一种基于混合卷积神经网络的垃圾分类方法。

背景技术

[0002] 垃圾分类回收在日常生活中占有非常重要的地位,随着人们生活水平的提高,日常生活垃圾越来越多。以前垃圾分类都是通过人力进行的,随着人工智能的兴起,利用深度学习和其它智能技术对垃圾进行自动分类受到了普遍的欢迎。Lulea技术大学于1999年开展了一个项目,开发了一个使用机械形状标识符回收金属废料的系统。在贝叶斯计算框架中使用了SIFT和轮廓形状的特征,其系统基于Flickr材料数据库。JinqiangBai等人设计了一个新型拾垃圾的机器人,机器人能够利用深层神经网络进行垃圾识别,准确自主地检测垃圾,效果很显著。2016年,一种能够区分堆肥的自动垃圾被RaspberryPI回收利用,他们的系统是使用谷歌的TensorFlow开发的,该系统的缺点是只能区分堆肥材料。

[0003] ArtzaiPicon等人利用高光谱数据允许构建比标准彩色图颜色表示更精细的模型来采样有色金属垃圾的特性。提出的模糊光谱和空间分类器算法融合了光谱和空间特征,通过构造生物的光谱模糊集来降低高光谱数据的维数,实验结果表明,当光谱空间特征用于有色金属垃圾时分类率得到了很大的提升。2012年,Alex Krizhevsky等人使用基于卷积神经网络的AlexNet,在ImageNet竞赛图像分类任务中取得了最好成绩。此后,好的卷积神经网络相继被提出,能够较好地用于目标检测和分类。NoushinKarimian等人提出一种新的分类方法,利用磁感应光谱法对三种金属进行分类,可以构造出有效的分类器。S. Shylo等人利用毫米波成像技术加多个传感器来提供互补数据,从而提高了废纸和卡片的分类性能。RutqvistD等人使用自动机器学习来解决生活中智能垃圾管理系统的容器排空问题,利用现有的人工工程模型及其改进的传统机器学习算法,使用随机森林分类器达到最佳效果,还提高了回收容器排空时间的预测质量。Zhao Dong-e等人提出利用高光谱成像系统采集垃圾样品的近红外高光谱图像,通过去噪和反射率信息的黑白校正反演对采集的样本图像进行预处理,然后用主成分分析训练样本的感兴趣区域。结果证明,经过SAM的测试和分类对可回收垃圾高光谱图像处理可以得到更准确的结果。Zheng, JJ等人提出使用数理统计的方法来表述个体有限理性,并使用无标度网络的特定图结构来表征群体结构,利用Python进行仿真实验,该文对个体有限理性的表征,具有一定的理论价值,同时,对垃圾分类具有推广作用。Chu Y等人提出一种多层混合方式的深度学习系统,可自动分类城市公共区域内个人处理的垃圾,使用多层感知机器 (MLP) 方法整合图像特征和其它特征信息,得到很好的分类性能。Yusoff S H等人设计了一个自动分离金属的可回收生活垃圾并记录垃圾数据的系统。Zeng等人提出一种利用机载高光谱数据监测大面积垃圾分布的方法,提出一种新的高光谱图像分类网络——多尺度卷积神经网络,对高光谱图像数据的像素进行分类,生成二值垃圾分割图。该算法在大面积垃圾检测方面有良好的性能。Seok-BeomRoh等人利用混合技术构造了一种模糊径向基函数神经网络分类器,实验结果表明具有特征提取技

术的分类系统的分类性能好,能有效的回收垃圾。

[0004] Kennedy等人基于VGG-19为基础模型的迁移学习,在垃圾图像上的分类准确率为88.42%,很好的利用了VGG-19的提取特征的能力。传统的机器学习,需要标定大量的训练数据,会耗费大量的人力与物力,但迁移学习可以从现有的数据中迁移知识,用来帮助将来的学习,它能得到更短的训练时间,更快的收敛速度,更精准的权重参数。Adedeji等人采用50层残留网络预处理(ResNet-50)构建的卷积神经网络模型作为提取器,利用支持向量机(SVM)进行分类,在垃圾图像数据集上测试达到了87%的准确率。Chen Zhihong等人提出一种基于计算机视觉的垃圾自动分拣机器人抓取系统,为了实现目标物体的精确抓取,采用RPN和VGG-16模型进行物体识别和姿态估计。Stephen L等人利用MobilNet生成模型,在ImageNet大型视觉识别挑战赛上训练的模型中迁移学习,得到了87.2%的准确率。后期经过优化和量化,达到了89.34%的准确率,并成功运用在移动设备上。何凯明博士首次提出的残差网络,在2015年的ImageNet上大放光彩。但是模型的加深,学习能力也会出现“退化”现象,即模型层次加深时,错误率会提高。因此,该网络并不适合数据集少的垃圾分类。Ruiz V等人利用深度学习经典模型的优势,训练和比较不同的深度学习体系,以自动分类垃圾类型,其中最优的结合Inception-ResNet模型,在垃圾图像上达到了88.60%的准确率。Costa等人研究了不同类型的神经网络,将垃圾图像分为四类,其中采用的KNN、SVM、RF预训练模型方法得到的精度分别为88.0%、80.0%、85.0%。但就现有的研究工作来看,大多数的垃圾识别分类精度还比较低,且训练时间相对较长。

发明内容

[0005] 本发明的目的是为解决现有方法对垃圾分类的精度低、需要的训练时间长的问

题。

[0006] 本发明为解决上述技术问题采取的技术方案是:一种基于混合卷积神经网络的垃圾分类方法,该方法包括以下步骤:

[0007] 步骤一、加载垃圾图像,并对垃圾图像进行预处理,获得预处理后的垃圾图像;

[0008] 步骤二、构建混合卷积神经网络,将步骤一获得的预处理后垃圾图像输入混合卷积神经网络进行训练,以获得训练好的混合卷积神经网络;

[0009] 所述混合卷积神经网络的结构为:

[0010] 从混合卷积神经网络的输入端开始,混合卷积神经网络依次包括第一网络模块、第二网络模块、第三网络模块、第四网络模块、展平层、第一全连接层、第一全连接层后的批标准化层、第二全连接层、第二全连接层后的批标准化层以及Softmax分类器;

[0011] 从所述第一网络模块的输入端开始,第一网络模块由两个基础模块以及一个最大池化层组成,其中:每个基础模块包括一个卷积层以及一个批标准化层;

[0012] 其中:第二网络模块、第三网络模块和第四网络模块的结构均与第一网络模块的结构相同;

[0013] 步骤三、将待分类的垃圾图像输入训练好的混合卷积神经网络,并加载训练好的混合卷积神经网络的权重,输出对待分类垃圾图像的分类结果。

[0014] 本发明的有益效果是:本发明提出了一种基于混合卷积神经网络的垃圾分类方法,本发明的混合卷积神经网络模型灵活运用了卷积层、批标准化、最大池化层和全连接

层,使各层作用充分发挥,进而得到较好的垃圾分类结果。本发明大量使用BN批标准化,将其用于每层卷积层和全连接层之后,进一步增强了模型提取特征的能力。利用BN层的正则化效果,适当的加入最大池化层对特征进行统计,减少特征维度,提高了表征能力,且能够很好拟合,收敛速度快、参数量少、计算复杂度低,较传统卷积神经网络有明显的优势。实验中,首先对垃圾图像数据集进行预处理,以便混合卷积神经网络能够很好对其进行特征提取。其次,对模型的提出和改进进行了详细的说明,并对优化器的选择进行了对比,进而选择出适合模型的优化器。最后,通过多种指标对提出的垃圾分类模型进行评估,在垃圾图像上得到了高达92.6%的准确率。与现有的多种基于TrashNet数据集的垃圾分类方法进行对比,提出的方法能够得到最高的分类结果,且具有较低的计算复杂度,缩短了训练时间。

附图说明

- [0015] 图1是本发明的一种基于混合卷积神经网络的垃圾分类方法的流程图;
- [0016] 图2是数据增强后的垃圾图像;
- [0017] 图3为本发明采用的混合卷积神经网络的结构图;
- [0018] 图4为本发明的一个基础模块和一个网络模块的结构图;
- [0019] 图5为两个 3×3 的卷积核感受视野与1个 5×5 的卷积核感受视野的对比图;
- [0020] 图6为SGD与SGDM的梯度下降对比图;
- [0021] 图7为优化器Adam、SGD和SGDM+Nesterov的准确率对比图;
- [0022] 图8为TrashNet数据库的垃圾图像;
- [0023] 图9为本发明采用模型在训练过程中的训练精度曲线图;
- [0024] 图10为本发明采用模型在训练过程中的训练损失曲线图;
- [0025] 图11为本发明采用的混合卷积神经网络在测试集上输出的混淆矩阵图;
- [0026] 图12为可视化混合卷积神经网络的特征图。

具体实施方式

- [0027] 具体实施方式一:如图1所示,本实施方式所述的一种基于混合卷积神经网络的垃圾分类方法,该方法包括以下步骤:
- [0028] 步骤一、加载垃圾图像,并对垃圾图像进行预处理,获得预处理后的垃圾图像;
- [0029] 步骤二、构建混合卷积神经网络,将步骤一获得的预处理后垃圾图像输入混合卷积神经网络进行训练,以获得训练好的混合卷积神经网络;
- [0030] 所述混合卷积神经网络的结构为:
- [0031] 从混合卷积神经网络的输入端开始,混合卷积神经网络依次包括第一网络模块、第二网络模块、第三网络模块、第四网络模块、展平层、第一全连接层、第一全连接层后的批标准化层、第二全连接层、第二全连接层后的批标准化层以及Softmax分类器;
- [0032] 从所述第一网络模块的输入端开始,第一网络模块由两个基础模块以及一个最大池化层组成,其中:每个基础模块包括一个卷积层以及一个批标准化层;
- [0033] 其中:第二网络模块、第三网络模块和第四网络模块的结构均与第一网络模块的结构相同;
- [0034] 步骤三、将待分类的垃圾图像输入训练好的混合卷积神经网络,并加载训练好的

混合卷积神经网络的权重,输出对待分类垃圾图像的分类结果。

[0035] 具体实施方式二:本实施方式与具体实施方式一不同的是:所述步骤一中,对垃圾图像进行预处理,获得预处理后的垃圾图像;其具体过程为:

[0036] 通过对加载的垃圾图像进行数据增强和归一化处理,来获得预处理后的垃圾图像;

[0037] 所述数据增强的方式包括对垃圾图像进行随机缩放、翻转、平移和旋转。

[0038] 一般而言,比较成功的神经网络需要大量的参数,而使得这些参数可以正常工作需要大量的数据进行训练,然而实际情况中往往并没有那么多数据。考虑到垃圾图像数据库TrashNet的样本比较少,在垃圾图像预处理时对其进行了数据增强,以增加训练样本数。它不仅能增强训练的数据量,提高模型的泛化能力,还能增加噪声数据,提高模型的鲁棒性。如图2所示为数据增强后的垃圾图像。

[0039] 具体实施方式三:本实施方式与具体实施方式一不同的是:

[0040] 在每一层卷积层和全连接层后都加上批标准化,使模型的提取特征能力进一步增强,可以有效的避免梯度消失和梯度爆炸,减小模型的结构复杂度。

[0041] 利用卷积层来提取图像特征,BN层来提高网络的泛化能力、打乱训练的数据和加快模型的收敛速度。在训练时,BN是基于每一个小批量计算的,将训练时每一个批量数据对应的均值和方差记录下来,利用它们来计算整个训练集的均值和方差,其计算公式为:

$$[0042] \quad \mu_{\beta} = \frac{1}{m} \sum_{i=0}^m x_i$$

$$[0043] \quad \delta_{\beta}^2 = \frac{1}{m} (x_i - \mu_{\beta})^2$$

$$[0044] \quad E[x] \leftarrow E_{\beta}[\mu_{\beta}]$$

$$[0045] \quad Var[x] \leftarrow \frac{m}{m-1} E_{\beta}[\delta_{\beta}^2]$$

[0046] 其中,m指的是小批量尺寸, β 为一个批量大小为m的数据集,x为一层的输入。批标准化是对每个特征图进行的,即对每个特征图不同位置的批标准化采取同样的操作。假设特征图大小为 $p \times q$,则对该特征图的BN相当于对大小为 $m' = |\beta| = m \cdot pq$ 的特征批标准化。选择了BN可以有效避免梯度消失和爆炸,与参数初始值关系较小,且起到了正则化的效果。可以去掉Dropout和L2正则化,以减小模型的结构复杂程度。

[0047] 图3为本发明采用的混合卷积神经网络的结构图,表1为混合卷积神经网络模型结构和对应层的参数;

[0048] 表1

[0049]

卷积神经网络结构				
类型	过滤器	步幅	填充	输出
输入	-	-	-	64×64×3
Conv2D_1	3×3	1×1	Same	64×64×32
BN_1	-	-	-	64×64×32
Conv2D_2	3×3	1×1	Same	64×64×32

[0050]

BN_2	-	-	-	64×64×32
Max_pooling_1	2×2	2×2	-	32×32×32
Conv2D_3	3×3	1×1	Same	32×32×64
BN_3	-	-	-	32×32×64
Conv2D_4	3×3	1×1	Same	32×32×64
BN_4	-	-	-	32×32×64
Max_pooling_2	2×2	2×2	-	16×16×64
Conv2D_5	3×3	1×1	Same	16×16×128
BN_5	-	-	-	16×16×128
Conv2D_6	3×3	1×1	Same	16×16×128
BN_6	-	-	-	16×16×128
Max_pooling_3	2×2	2×2	-	8×8×128
Conv2D_7	3×3	1×1	Same	8×8×256
BN_7	-	-	-	8×8×256
Conv2D_8	3×3	1×1	Same	8×8×256
BN_8	-	-	-	8×8×256
Max_pooling_4	2×2	2×2	-	4×4×256
Flatten_1	-	-	-	1×1×4094
Full_connected_1	-	-	-	1×1×128
BN_9	-	-	-	1×1×128
Full_connected_2	-	-	-	1×1×64
BN_10	-	-	-	1×1×64
Softmax分类器	-	-	-	1×1×6
所有参数	1709926			

[0051]

训练参数	1707622
未训练参数	2304

[0052] 为了解决每次卷积操作时,图像角落边的像素被遗漏,在输出中利用较少,从而丢掉图像边缘位置的特征信息,对每层卷积层都使用0填充。最后输出网络的总参数为1709926万,参数量相对于深层的卷积神经网络来说,已经很小了。

[0053] 具体实施方式四:本实施方式与具体实施方式一不同的是:所述第一网络模块的通道数为32;第二网络模块的通道数为64;第三网络模块的通道数为128;第四网络模块的通道数为256,第一全连接层的通道数为128,第二全连接层的通道数为64。

[0054] 第一网络模块、第二网络模块、第三网络模块、第四网络模块的通道数分别采用32、64、128、256。保证当输出特征图尺寸减半时,输出特征图的通道数加倍,这样可保证相邻卷积层所包含的信息量不会相差太大。批标准化即批量数据,把数据分成小批量进行随机梯度下降,且在每批数据进行前向传递时对每一层都进行标准化处理。本发明采用一层卷积层和一层BN层作为基础模块,对卷积层的输出进行标准化处理。每两个基础模块后加一层最大池化层作为一个网络模块,卷积层后的激活函数采用“修正线性单元(ReLu)”。该激活函数计算复杂度低,不需要进行指数运算,只需要一个阈值即可得到激活值。且在 $x > 0$ 的范围不会出现梯度饱和、梯度消失现象。

[0055] 提出的基础模块和网络模块结构图如图4所示。利用这样的模块进行混合,每个网络模块的通道数采用32、64、128、256等,模块数以 2^n 进行增长。在四个网络模块进行混合后,利用一层展平层进行展平,将其用在卷积层和全连接层的过渡,用来将输入全连接层的数据压平。紧接着利用两层全连接层,全连接层的通道数本发明采用小通道数,分别为128和64,相比于大通道数减少了参数和计算量,最后采用Softmax经典分类算法进行分类。

[0056] 本发明通过改变网络模块的数量和通道数进行模型改进,根据改进前后模型的准确率对比,选择准确率最高的模型作为本发明的混合卷积神经网络模型。如表2所示,为提出混合卷积神经网络模型改进过程的指标记录表;

[0057] 表2.

	提出方法	结构	准确率	迭代一次耗时
[0058]	初始模型	一个网络模块	86.20%	114毫秒/步
	第一次模型改进	三个网络模块混合	87.20%	189毫秒/步
	第二次模型改进	三个网络模块和一个基础模块混合	89.70%	201毫秒/步
[0059]	第三次模型改进	四个网络模块混合	92.60%	223毫秒/步
	第四次模型改进	五个网络模块混合	88.50%	240毫秒/步

[0060] 改进是基于主体框架的基础上进行微调。本发明所有改进都是基于模块的混合数进行改进的,混合完成的网络所连接的展平层和全连接层参数都相同。在训练模型之前,首先对提出的初始模型进行训练,得到的准确率为86.2%。第一次模型改进采用三个网络模块进行混合,分别采用32、64、128个通道数,加载数据集训练后得到了87.2%的准确率,迭代的平均时间为189毫秒/步,训练得到的准确率不是很理想。

[0061] 第二次模型改进是在第一次模型改进的基础上加上一个通道数为256的基础模块,经过训练后,该模型得到了89.70%的准确率,迭代的平均时间为201毫秒/步。可见,分类准确率有所提升,但模型复杂度增加,训练迭代时间变长。有待进一步改进。

[0062] 考虑到模型稍浅、提取特征的能力还不足,第三次模型改进是在第一次模型改进的基础上加上一个通道数为256的网络模块,训练后得到了92.6%的准确率,迭代的平均时间为223毫秒/步。可见,在时间可接受的范围内,精度上有了明显的提高。

[0063] 在上述较好结果的基础上,进一步加深网络深度,看是否会随着模型的复杂度提高,导致错误率的提高。因此,在第三次模型改进的基础上加上通道数为384的网络模块,训练后得到了88.5%的准确率。很明显,随着模型复杂度的提高,错误率增大了。经过模型改进结果表明,由四个网络模块混合成的模型效果最佳,准确率达到92.6%,迭代一次的平均时间为233毫秒/步。因此,本发明采用四个网络模块混合成的模型作为垃圾分类模型。以上所有模型的训练使用的优化器都是SGDM+Nesterov。

[0064] 具体实施方式五:本实施方式与具体实施方式四不同的是:所述混合卷积神经网络中每个卷积层均采用 3×3 的卷积核,且步幅都为1;每个最大池化层均采用 2×2 的过滤器,且步幅都为 2×2 。

[0065] 结合深度学习和卷积神经网络的特点,本发明拟从复杂度低、参数少、计算量小的卷积神经网络入手,并采用 3×3 的小卷积核。在VGGNet中曾指出,两个 3×3 的卷积核拥有和1个 5×5 的卷积核相同的感受视野。卷积核感受视野比较图如图5所示。因此,参数量可进一步减少。例如,在卷积核的数量都为n时,1个 5×5 的卷积层中的参数量为 $25n$,而2个 3×3 的卷积层中的参数量才为 $18n$,差别较为明显。与此同时,2个 3×3 的卷积层拥有比1个 5×5 的卷积层更多的非线性变换,即 3×3 的卷积层可以使用两次非线性激活函数,而 5×5 的卷积层只使用一次,使卷积神经网络对图像的特征的学习能力更强。因此,使用 3×3 的卷积核,既可以保证感受视野,又可以减少卷积层的参数。

[0066] 具体实施方式六:本实施方式与具体实施方式五不同的是:所述混合卷积神经网络采用的激活函数为Relu,优化器为SGDM+Nesterov,优化器的动量参数设置为0.9。

[0067] SGDM+Nesterov是指在随机动量梯度下降法(SGDM)的基础上引入Nesterov动量,动量参数设置为0.9。

[0068] 优化器的选择对于深度学习的训练有着极其重要的作用,关系到训练能否快速收敛并且取得较高的准确率和召回率。常见的优化器有Adam、GradientDescent、Momentum等。本发明主要对Adam、随机梯度下降法(SGD)、SGD+Momentum+Nesterov在提出的模型上进行研究比较。优化器Adam集合了AdaGrad和RMSProp的优点,计算高效、方便实现,更新步长和梯度大小无关,只和alpha、beta_1、beta_2有关。优化器SGD的参数更新是针对每一个样本集,从一批训练样本中随机选取一个样本,SGD没有动量概念的时候公式为:

[0069]
$$m_t = g_t, V_t = I^2$$

[0070]
$$\eta_t = \alpha \cdot g_t$$

[0071] 其中,计算目标函数关于当前参数的梯度为

[0072]
$$g_t = \nabla f(w_t)$$

[0073] 根据历史梯度计算一阶动量和二阶动量,分别为

[0074]
$$m_t = \phi(g_1, g_2, \dots, g_t)$$

[0075]
$$V_t = \psi(g_1, g_2, \dots, g_t)$$

[0076] 计算当前时刻的下降梯度:

$$[0077] \quad \eta_t = \alpha \cdot \frac{m_t}{\sqrt{V_t}}$$

[0078] 根据下降梯度进行更新:

$$[0079] \quad w_{t+1} = w_t - \eta_t$$

[0080] 因为SGD的最大缺点就是下降速度慢,且存在持续震荡现象,停留在一个局部最优解。为了解决这一问题,可在SGD基础上加上一阶动量:

$$[0081] \quad m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$$

[0082] 对各个时刻梯度方向上的指数求移动平均值,大约等于最近 $\frac{1}{(1-\beta_1)}$ 个时刻的梯度

向量和平均值。于此同时为了解决SGD困在局部最优的问题,再在SGDM (Loshchilov I, Hutter F. Sgdr: Stochastic gradient descent with warm restarts [J]. arXiv preprint arXiv:1608.03983, 2016.) 和 (Zoph, Barret, and Quoc V. Le. "Neural architecture search with reinforcement learning." arXiv preprint arXiv:1611.01578 (2016)) 的基础上加上Nesterov, 梯度下降的方向是由累积动量决定的,加了Nesterov可以不看当前的梯度方向,而是跟着累积动量走,进一步判断梯度下降方向如何走。这时的下降方向为

$$[0083] \quad g_t = \nabla f \left(w_t - \frac{\alpha \cdot m_{t-1}}{\sqrt{V_{t-1}}} \right)$$

[0084] 然后,利用下一个点的梯度方向,与历史累计动量相结合,计算当前时刻的累积动量。

[0085] SGD在没有动量和有动量的情况下梯度下降的摆动过程如图6所示。为了使成本函数最优化,即从椭圆的边缘开始梯度下降到达中心的最小值点。它将会慢慢摆动到中心点,这样的上下摆动减慢了梯度下降的速度,这样就无法使用更大的学习率了,否则可能会偏离函数范围,但是希望上下摆动的幅度小且梯度下降的速度快,于是采用动量梯度下降法,它可以很好减少摆动幅度和加快下降速度。

[0086] 在相同条件下,Adam、SGD和SGDM+Nesterov在提出的最好模型下训练的比较结果如图7所示。这里展示的是训练过程中保存的权重模型的精度值。由图7可以看出,前期SGD的效果是最好的,随着训练次数的增加Adam和SGD逐渐趋于平稳,在训练的后期SGDM+Nesterov的效果最好。这也很好的证实了SGD能够达到全局最优解,而且训练的最佳精度也要高于其他优化算法,只是对学习率的设置要求比较高容易停在鞍点。Adam下降速度快,很容易跳过鞍点,无需干预学习率的设置,但是很容易在局部最小值处震荡,存在在特殊数据集下出现学习率突然上升现象,造成不收敛的情况,它集合和其它算法的优点也包含了它们的缺点。所以,Adam优化算法不一定适合本发明模型,于是本发明采用的是SGDM+Nesterov作为优化器。

[0087] 表3给出了三种优化器的具体准确率和迭代一次的平均耗时。可以明显看出SGDM+Nesterov优化器的准确度和迭代一次的平均耗时最佳。

[0088] 表3. 优化器Adam、SGD和SGDM准确度及耗时

	优化器	准确率	迭代一次耗时
[0089]	Adam	90.2%	235毫秒/步
	SGD	89.7%	225毫秒/步
	SGDM+Nesterov	92.6%	223毫秒/步

[0090] 具体实施方式七：本实施方式与具体实施方式六不同的是：所述将步骤一获得的预处理后垃圾图像输入混合卷积神经网络进行训练，以获得训练好的混合卷积神经网络，其具体过程为：

[0091] 步骤1、设置初始学习率为0.01，在训练过程中添加监控保存损失函数值最低且精度值最高的模型机制；

[0092] 步骤2、在当前学习率 lr 下，对混合卷积神经网络进行训练，当学习停滞时（即在学习过程中受到阻碍时），若连续15次训练中模型的性能得不到提升，则学习率下降0.1，获得新的学习率 $lr \times 0.1$ ；

[0093] 步骤3、重复步骤2的过程，新的混合卷积神经网络在新的学习率下继续训练；

[0094] 步骤4、直至当前学习率对应的损失函数值开始增加、停止改善时，则提前停止机制被激活，训练30次后停止训练，获得训练好的混合卷积神经网络。

[0095] 本实施方式训练过程中添加了监控保存损失函数值最低且精度值最高的模型机制，当第1次训练时，第1次训练得到的损失函数值为最低损失函数值和最高精度值，模型自动保存第1次训练获得的模型，当第2次训练时，若第2次训练得到的损失函数值比第1次训练得到的损失函数值低，且精度值比第1次训练得到的精度值高，则模型自动保存第2次训练得到的模型，若第2次训练得到的损失函数值不比第1次训练得到的损失函数值低，或第二次训练得到的损失值比第一次训练得到的损失值低，但第二次训练的精度值不比第一次训练的精度值高，则模型还是保存第1次训练得到的模型，以此类推，通过不断的训练过程，保存出来损失函数值最低精度值最高的模型。

[0096] 具体实施方式八：本实施方式与具体实施方式七不同的是：所述混合卷积神经网络中卷积层的每个卷积核提取出图像特征，即提取出特征图谱，提取出的特征图谱的具体计算公式如下：

$$[0097] \quad y_j^l = \theta \left(\sum_{i=1}^{N_j^{l-1}} w_j \otimes x_i^{l-1} + b_j^l \right), j=1, 2, \dots, M$$

[0098] 其中， y_j^l 表示第 j 个特征图谱； w_j 表示对应的卷积核； x_i^{l-1} 表示上一层第 i 个特征图谱，即将上一层第 i 个特征图谱作为当前的输入； b_j^l 表示偏置量； N_j^{l-1} 表示每个特征图谱的特征数量； M 表示每个卷积层的特征图谱数； $\theta(\cdot)$ 为激活函数，常见的有Tanh、Sigmoid、ReLU、Softmax等。

[0099] 具体实施方式九：本实施方式与具体实施方式八不同的是：所述第一全连接层和第二全连接层的输出的表达式为：

$$[0100] \quad h_{w,b}(x) = \theta(W_x^T x)$$

$$[0101] \quad h_{w,b}(x) = \theta \left(\sum_{i=1}^3 W_i x_i + b \right)$$

[0102] 其中, $h_{w,b}(x)$ 表示全连接层的输出; x_i 表示上一层神经元的输出, 即全连接层的输入; W_i 表示神经元之间连接的权值; b 表示偏置量, $\theta(\cdot)$ 表示激活函数。

[0103] 具体实施方式十: 本实施方式与具体实施方式九不同的是: 所述混合卷积神经网络的性能评估指标为精确率、召回率和F1-score值;

[0104] 所述精确率P、召回率r以及F1-score值 F_1 的计算公式如下:

$$[0105] \quad P = \frac{TP}{TP + FP}$$

$$[0106] \quad r = \frac{TP}{TP + FN}$$

$$[0107] \quad F_1 = \frac{2rP}{2 \times TP + FP + FN}$$

[0108] 其中, TP表示将实际正样本数预测为正样本的个数, FN表示将实际正样本数预测为负样本的个数, FP表示将实际负样本预测为正样本的个数。

[0109] 实验及结果分析

[0110] 该部分首先对TrashNet数据集进行图像预处理, 使模型能够更好的接受和提取图像的特征。其次, 利用训练得到的模型, 在TrashNet数据集上进行评估, 得到评估模型的数据指标。最后, 在相同数据集上, 将本发明模型与其他模型进行分类准确度和混淆矩阵的对比, 以证明本发明提出方法的有效性。在这部分中, 本发明模型指用四个网络模块进行混合得到的模型。

[0111] 1、数据集处理

[0112] 本发明采用的垃圾图像数据库是TrashNet数据库, 该数据集是由斯坦福大学的Mindy Yang和Gary Thung创建的包含六类RGB垃圾图像, 总共2527张图像, 其中纸箱: 403张, 玻璃: 501张, 金属: 410张, 纸: 594张, 塑料: 482张, 垃圾: 137张, 图像像素为 513×384 。该数据库的图像可视化如图8所示。本发明先对垃圾图像数据库进行数量上的统计, 进而合适的划分出训练集、验证集和测试集, 三者的比例划分为(8:2:1), 每一个数据集的类别数量如表4所示。

[0113] 表4. 训练集、验证集和测试集数量

	纸箱	玻璃	纸	金属	塑料	垃圾
[0114] 训练集数量	293	364	298	431	350	99
验证集数量	73	91	74	106	88	25
测试集数量	37	46	38	55	44	13

[0115] 如图2所示为数据增强后的垃圾图像。

[0116] 2、测试结果分析

[0117] 为了验证提出模型的有效性, 本发明在TrashNet数据集上利用提出模型进行了大量测试, 并采用多种指标对其进行评估。使用的输入模型图像为 $(64 \times 64 \times 3)$ 的大小, 以此

进一步减少模型的参数量,进而减少计算量,大大缩短训练迭代的时间。并加入了提前停止机制和学习率下降机制,本发明将耐心值设为30次,当发现损失值开始增加,停止改善时,提前停止机制被激活,则经过30次后自动停止训练。在当前学习率 lr 下,每当学习停滞时,连续15次训练中模型的性能得不到提升,则学习率下降0.1,获得新的学习率 $lr \times 0.1$ 。批次数设为32,该模型是基于Keras开发的,且训练在GeForce 940MX显卡上完成。本发明提出的模型加载垃圾分类图像训练后达到了92.6%的准确率。该模型整个训练过程的训练精度曲线和损失曲线如图9和图10所示,可以看到模型前期的验证精度和损失值波动很大,不停的震荡到后面趋于稳定,比较符合优化器SGDM的特点,得到了较好的效果。

[0118] 通过在TrashNet数据集上进行模型性能评估,分别给出了垃圾图像类别的精确率、回归率和F1-score值,数量表示类别的测试图像数量,如表5所示。表5还给出了模型分类的宏平均值、微平均值,以及权重的平均值,微平均是把所有类别都一次性考虑进来,计算类别预测的准确率;宏平均即是对每个类别分开考虑,单独计算每个类别的准确率,最后再进行算术平均得到该测试集的准确率。由表5结果可知,垃圾类别的召回率最低,纸箱类别和金属类别的召回率最高,其与训练图像的数量和特征有着极大的关联。在数据集中垃圾类别数量最少且特征和其它类别很相近,纸箱和金属的训练数量多且特征简单易学习。

[0119] 表5. 模型在测试集上测试的报告表

	类型	精确率	召回率	F1-值	数量
[0120]	纸箱	0.92	0.94	0.93	37
	玻璃	0.87	0.79	0.83	46
	金属	0.83	0.97	0.89	38
	纸	1.00	0.93	0.96	55
	塑料	0.88	0.93	0.90	44
	垃圾	1.00	0.77	0.87	13
[0121]	微平均值	0.91	0.91	0.91	233
	宏平均值	0.92	0.89	0.90	233
	权重平均值	0.91	0.91	0.91	233

[0122] 图11为本发明通过加载模型在TrashNet数据集上测试得到的混淆矩阵。它主要用于比较分类结果和实际预测结果,可以把分类结果的准确度显示在一个混淆矩阵中。与此同时,混淆矩阵能够更好的展示出分类算法性能,可以从另一个角度给出分类模型的表现以及错误类型。由图11可以看出,每一类别预测的正确率都集中在对角线上,六类预测正确的准确率都较高,表明了本发明模型具有较好的分类性能。

[0123] 如图12为提出方法得到的可视化卷积神经网络的特征图。这里给出了本发明模型的卷积层的特征图。由图12可见,本发明模型第一层卷积获得的特征还比较明显,特征数据与原始数据还很接近。随着层数越深,特征也变得越来越抽象,本发明模型的第四层卷积获得的特征相对于第一层卷积获得的特征更抽象,但人眼还可以判断出其特征与原图还有相似度。随着模型层数的进一步加深,第八层卷积获得的特征更加的抽象了,完全判断不出与

原图特征的相似性。这证实了卷积神经网络的工作过程。可以认为卷积神经网络通过对图像特征进行预处理,进行适量的数据降维,去掉无用成分,以便模型学到更多有用的特征。本发明模型采用的八层卷积得到的模型准确率最高,由本发明模型改进的过程(由表2可知)可知超过八层卷积层和少于八层卷积层的效果都不是最好的。因此,对于不同的数据集而言,并不是模型的复杂度越高提取特征的能力就越好,所得到的分类精度就越高。

[0124] 在相同数据集下,将本发明方法与其他基于深度学习的方法进行对比,Kennedy T等人利用基于VGG-19的迁移学习方法,探索在少量数据情况下利用预训练的大型网络达到了88.42%的准确率。Awe等人使用了一个微调更快的Faster R-CNN模型,对混杂的垃圾图像分类,分类准确率达到68.30%。Adedeji等人采用50层残留网络预处理(ResNet-50)构建的卷积神经网络模型作为提取器,利用支持向量机(SVM)进行分类,在垃圾图像数据集上达到了87%的准确率。Ruiz V等人利用深度学习经典模型的优势,使用垃圾图像数据集训练和比较了几种卷积神经网络的效果,最后结合Inception-ResNet模型达到了88.60%的最好准确率。Costa等人研究了不同类型的神经网络,将垃圾图像分为四类,其中采用的KNN、SVM、RF预训练模型方法得到的精度分别为88.0%、80.0%、85.0%。Rabano S L等人利用MobileNet生成的模型,测试精度达到了87.2%,并对其进行优化和量化后达到了89.34%的准确率。Satvilkar M等人使用卷积神经网络,使用多种方法并列排列,最佳方法分类精度为89.81%。Yang M等人使用带有支持向量机的比例不变特征变换的卷积神经网络(SVM),实验表明准确率达到63.0%。本发明提出的模型,在垃圾分类图像上达到的准确率是最优的。此外,本发明方法的模型参数量少,计算量少,迭代时间短。这意味着采用本发明方法进行垃圾分类,能够在较低的计算复杂度下,得到较高的垃圾分类正确率。

[0125] 利用本发明模型和文献(“Kennedy,Tom.”OscarNet:using transfer learning to classify disposable waste.”CS230 Report:Deep Learning.Stanford University,CA,Winter(2018)”)以及“Ruiz V,Sánchez Á,Vélez J F,et al.Automatic Image-Based Waste Classification[C]//International Work-Conference on the Interplay Between Natural and Artificial Computation.Springer,Cham,2019,11487:422-431.”)中的混淆矩阵对角线的最高准确率进行对比。

[0126] 将文献(“Kennedy,Tom.”OscarNet:using transfer learning to classify disposable waste.”CS230 Report:Deep Learning.Stanford University,CA,Winter(2018)”)作为文献1;

[0127] 将文献(“Ruiz V,Sánchez Á,Vélez J F,et al.Automatic Image-Based Waste Classification[C]//International Work-Conference on the Interplay Between Natural and Artificial Computation.Springer,Cham,2019,11487:422-431.”)作为文献2;

[0128] 从平均准确率来看,本发明模型平均准确率最高。与此同时,通过6个类别准确率的趋势来看,本发明模型更加平稳,波动较小,不会偏向任何一个类别。而且由表6中极差值可知,本发明模型极差值最小,对垃圾图像6个类别的预测均衡性好,较为集中,更能突出本发明模型的稳定性好,有良好的分类性能。

[0129] 表6.混淆矩阵对比表

	纸箱	玻璃	金属	纸	塑料	垃圾	平均值	极差
文献 1	0.949	0.752	0.90	0.752	0.401	0.406	0.694	0.548
文献 2 的 VGG-16	0.930	0.680	0.88	0.670	0.770	0.620	0.758	0.310
[0130] 文献 2 的 VGG-19	0.850	0.780	0.85	0.850	0.600	0.690	0.770	0.250
文献 2 的 Inception-ResNet	0.880	0.880	0.98	0.900	0.770	0.690	0.850	0.290
本发明	0.900	0.790	0.97	0.930	0.930	0.770	0.888	0.200

[0131] 本发明的上述算例仅为详细地说明本发明的计算模型和计算流程,而并非是对本发明的实施方式的限定。对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其它不同形式的变化或变动,这里无法对所有的实施方式予以穷举,凡是属于本发明的技术方案所引伸出的显而易见的变化或变动仍处于本发明的保护范围之列。

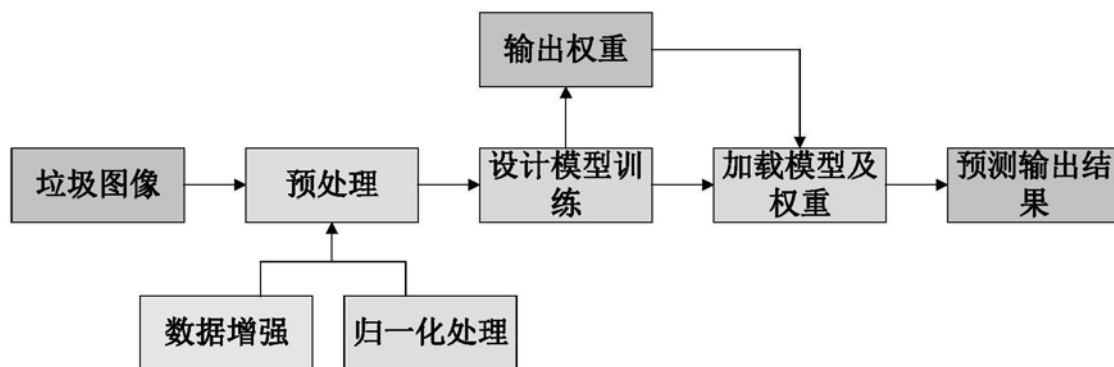


图1

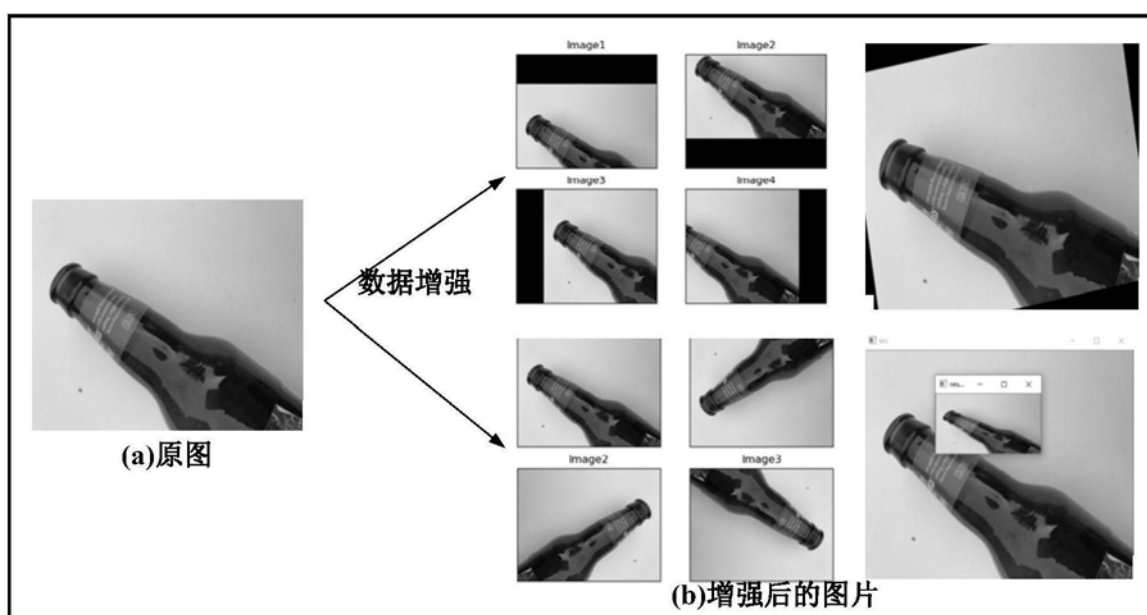


图2

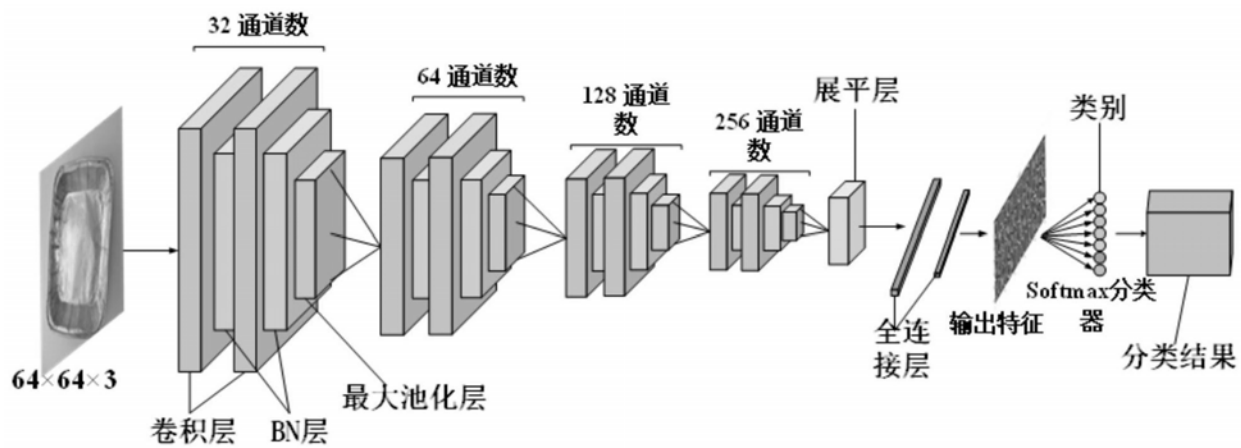


图3

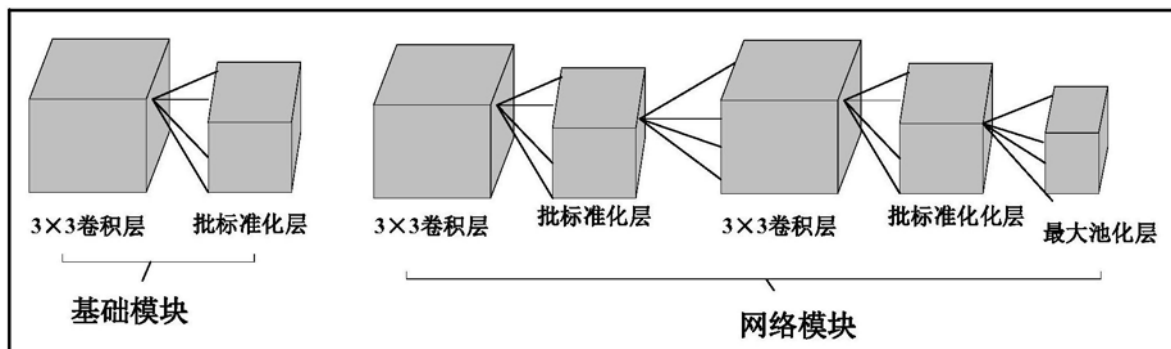


图4

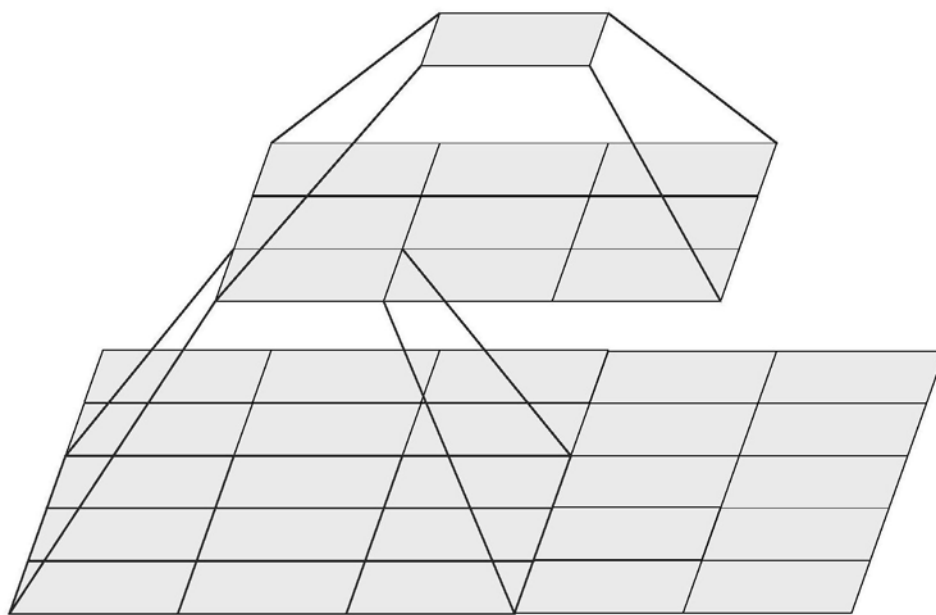


图5

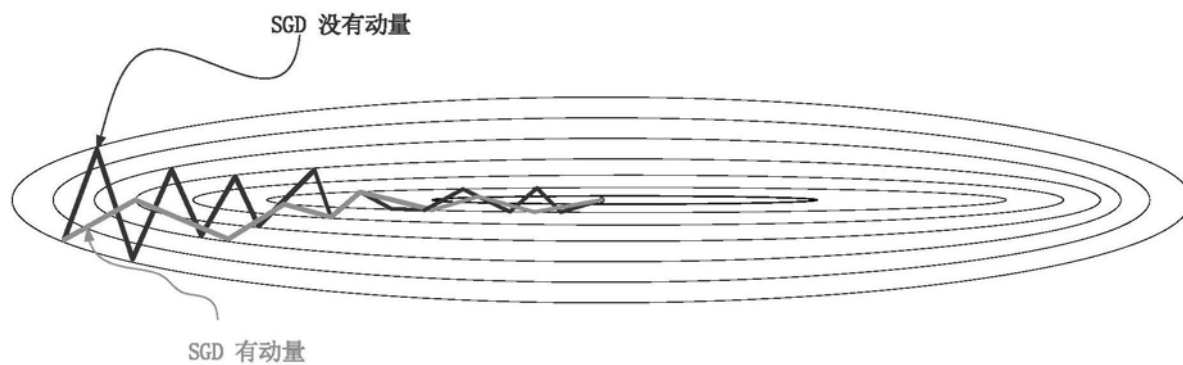


图6

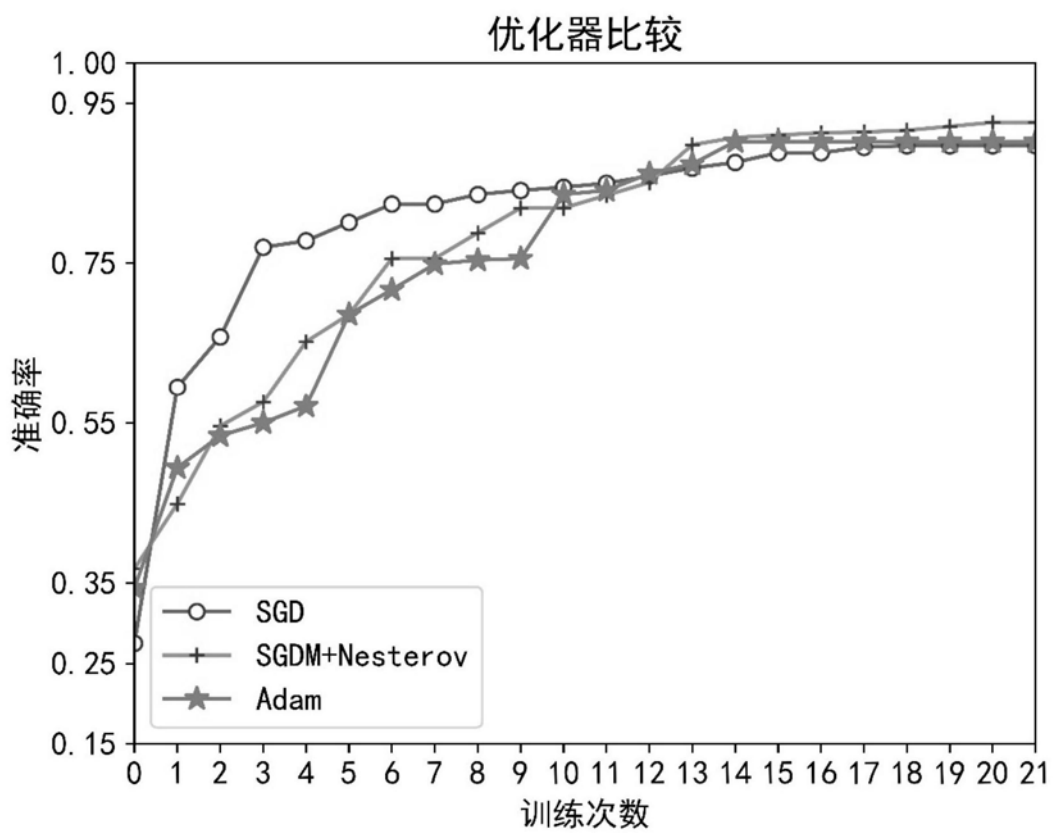


图7

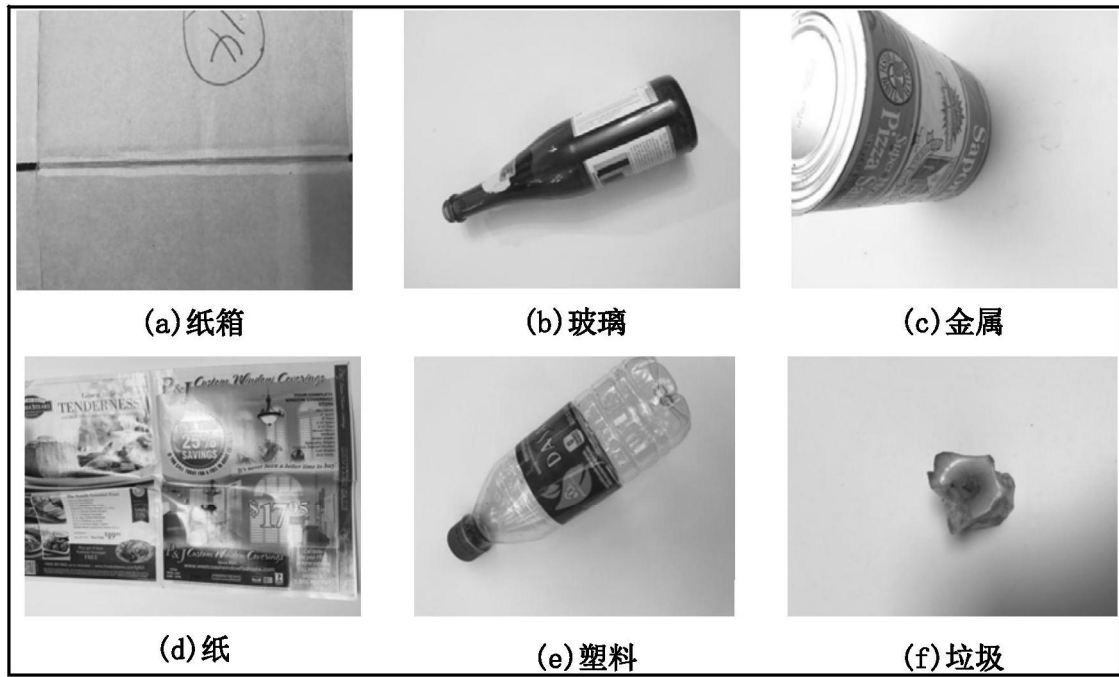


图8

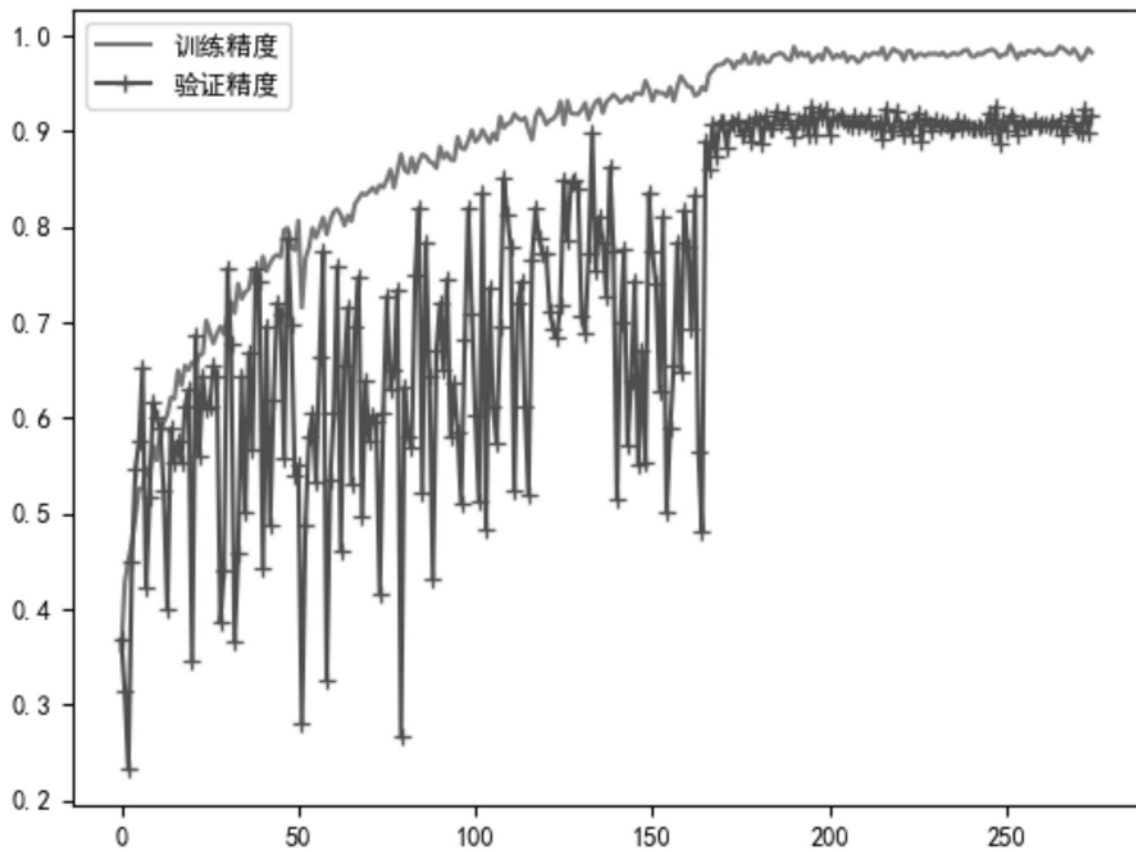


图9

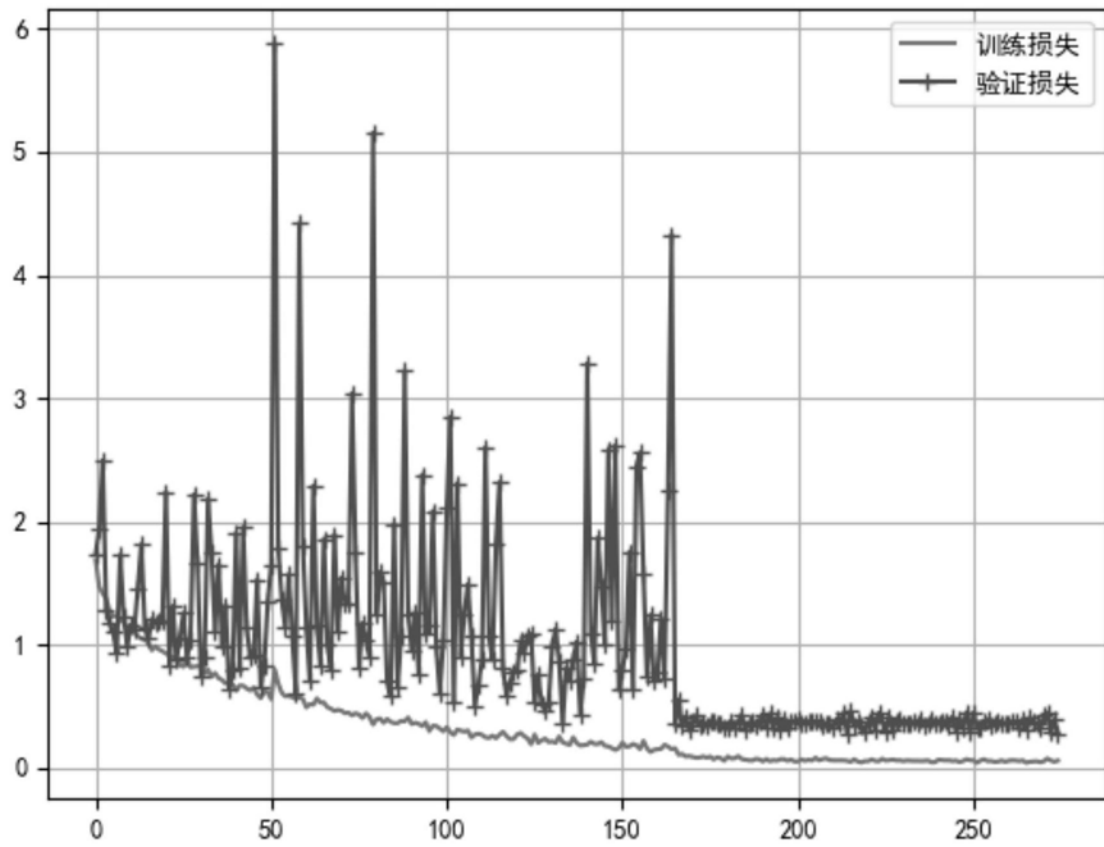


图10

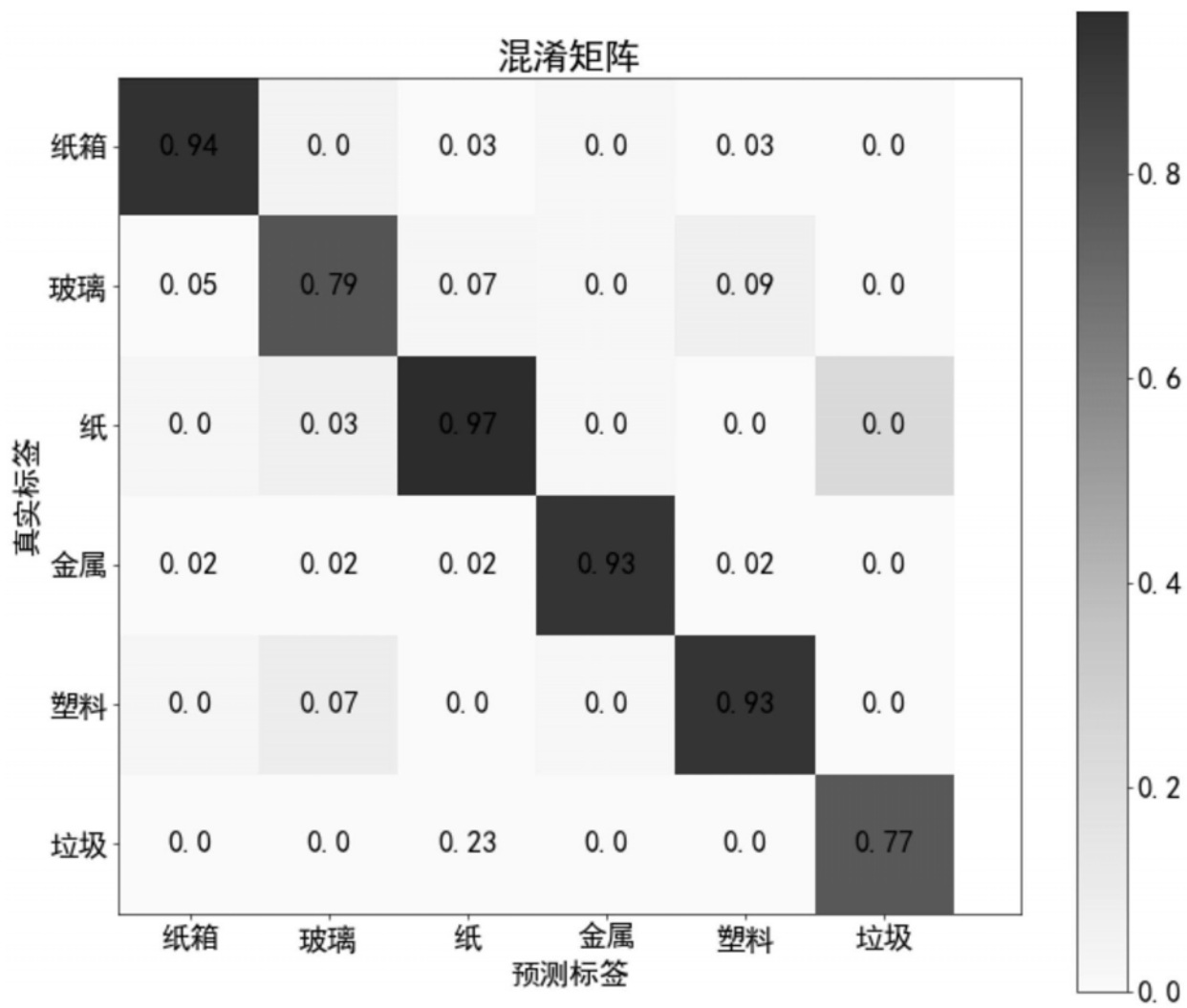


图11

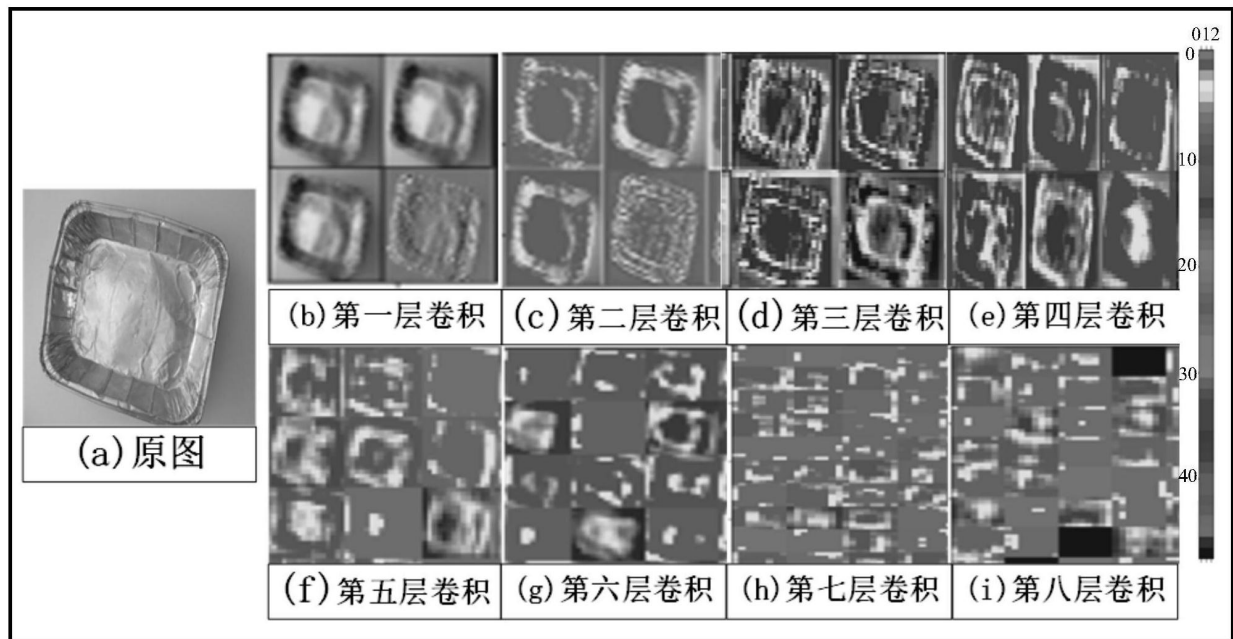


图12