垃圾图像分类系统的设计与实现

摘 要

我国是世界上垃圾最多的国家，近年来，我国愈来愈重视垃圾分类问题。然而绝大多数居民并不具备垃圾分类常识，同时也对掌握这一知识十分抵触。为了帮助居民快速正确的进行垃圾分类，本文提出了基于VGG16(Visual Geometry Group)卷积神经网络的垃圾图像分类系统。首先，采集数量可观的垃圾图像数据集，并且按一定比例将垃圾图像分为训练集，验证集和测试集；然后，在Windows操作系统下，以Python为主要编程语言对VGG16卷积神经网络中完成对数据集的训练，并且依次记录卷积神经网络在数据集中每次的训练过程中的准确率和损失率；最后，选择合适训练次数的卷积神经网络作为分类模型在数据集中进行测试，并且根据选好的分类网络构建出便于用户使用的界面。实验结果显示，基于VGG16的垃圾图像分类系统在本文自己构建的测试集中平均准确率达到了90%，并且在一定的背景干扰下仍能维持检测的稳定性，解决了居民们垃圾分类困难的问题。

装

订

线

关键词：深度学习；卷积神经网络；实时检测；垃圾图像分类

**Design and Implementation of Garbage Image Classification System**

**ABSTRACT**

China can produce enormous amounts of rubbish in recent years, which raises the governmental awareness of the rubbish sorting. However, most citizens lack awareness of rubbish sorting, meanwhile, they are unlikely to learn some knowledge of garbage sorting. This article suggests a garbage sorting method which is based on VGG16(Visual Geometry Group) to help the dwellers to sort garbage conveniently as well as correctly. First of all, collect a dataset with a number of rubbish images and divide it into train set, validation set and test set. In addition, on windows, practice VGG16 classification network mainly by python and recorded the train accuracy and train loss of the training progress at each training epoch. Last but not least, choose the most suitable network of any training time to test its performance, which can be used on programming the graphical interface for users. The experimental result shows that this neural network can achieve 90% of average accuracy on test set, it also can sustain the sorting stability for some images which exist ambient interference, reduce the difficulty of dwellers garbage classification.

**Key words：**Deep learning；Convolutional neural network；Real-time detection ；Garbage image classification

装

订

线

目 录

[1 引言 1](#_Toc38481395)

[1.1 研究背景及意义 1](#_Toc38481396)

[1.2 国内外发展现状 1](#_Toc38481397)

[1.2.1 基于卷积神经网络的图像分类的国内外研究现状 1](#_Toc38481398)

[1.2.2 基于纹理特征的图像分类的国内外研究现状 2](#_Toc38481399)

[1.3 垃圾图像分类拟采用的方法 2](#_Toc38481400)

[1.4 论文结构 3](#_Toc38481401)

[2 垃圾图像分类系统分析 4](#_Toc38481402)

[2.1 可行性分析 4](#_Toc38481403)

[2.2 需求分析 4](#_Toc38481404)

[3 卷积神经网络概述 6](#_Toc38481405)

[3.1 卷积神经网络 6](#_Toc38481406)

[3.2 VGGNet网络 10](#_Toc38481407)

[4 垃圾图像分类网络模块的设计与实现 14](#_Toc38481408)

[4.1 VGG16的优点: 14](#_Toc38481409)

[4.2 垃圾图像数据集的构建: 14](#_Toc38481410)

[4.3 基于VGG16的垃圾图像检测模型: 15](#_Toc38481411)

[4.3.1 数据集的导入与图像预处理 15](#_Toc38481412)

[4.3.2 迁移学习 16](#_Toc38481413)

[4.3.3 网络训练 17](#_Toc38481414)

[4.3.4 实验环境 20](#_Toc38481415)

[5 垃圾图像分类应用模块的的设计与实现 21](#_Toc38481416)

[5.1 导入模型 21](#_Toc38481417)

[5.2 对预测图片进行预处理 21](#_Toc38481418)

[5.3 预测图片 21](#_Toc38481419)

[5.4 垃圾图像识别操作步骤图 22](#_Toc38481420)

[5.5 垃圾图像报错 23](#_Toc38481421)

[6 垃圾图像分类系统实验结果 25](#_Toc38481422)

[7 总结与展望 28](#_Toc38481423)

[7.1 全文总结 28](#_Toc38481424)

[7.2 本文存在的不足与展望 28](#_Toc38481425)

[参考文献 30](#_Toc38481426)

[致 谢 31](#_Toc38481427)

2. 引言

1.1 研究背景及意义

随着经济的快速发展，人们的生活水平已经来到了一个新的高度，同时人们的垃圾产生量也日益增加，大量的垃圾由于得不到有效处理，造成了资源的浪费，同时也破坏了许多的城市的市容市貌，然而垃圾经过分类之后，利用过的垃圾可以转化为新的能源，也可以减少传统焚烧，填埋带来的环境问题，因此垃圾分类是当前处理垃圾的一个很好的选择。另一方面随着生活垃圾分类制度入法的到来，垃圾分类对于每一位公民显得更加重要。然而，现阶段居民普遍缺少垃圾分类常识，加之日益紧张的生活节奏垃圾分类实施的效果还有待加强。

21世纪的今天，我国已经进入智能化时代。图像识别，图像处理等技术在各行各业中都发挥了重大的作用，例如在无人零售领域[1]，通过图像识别技术使计算机根据商品图像自动识别商品，大大节省了人力，降低了成本，便利了人们的生活。针对当前人们缺乏垃圾分类常识这一问题，如果可以基于用户上传的垃圾图像实时监测出垃圾的分类种类，对于人们的生活便利，和城市的环保市容具有重大的意义。

1.2 国内外发展现状

垃圾图像分类在实际应用中具有很大的意义，但是现阶段，基于图像的垃圾分类研究不多，目前用于垃圾图像分类的方法有基于卷积神经网络[2]的图像分类，基于纹理[3]的图像分类等,针对这一问题国内外的许多专家学者对此进行了研究。

1.2.1 基于卷积神经网络的图像分类的国内外研究现状

利用卷积神经网络进行图像分类即使待识别的样本形状发生了变化，仍能完成高效稳定的识别任务，除此之外卷积神经网络通过依次分层对图像样本的拓扑结构展示的数据抽象，提取隐式的样本特征，使传统分类方式中的特征提取的数据量大，需要大量图像预处理操作，计算复杂度高的问题得到了规避。卷积神经网络也在近几年广泛用于各类视频检测[4],图像分类[5][6]等领域,如Picon等人[7]提出了三种不同的CNN（Convolutional Neural Networks）架构用于分类农作物疾病种类，实验基于10万张图像组成的五种农作物不同疾病阶段的数据集，采用上述的CNN架构将上下文非图像元数据（例如裁剪的信息）合并到基于图像的卷积神经网络中，降低了疾病分类的复杂性，实验结果显示，农作物病害分类网络通过嵌入向量级联的方式将背景信息整合在一起，取得了0.98的均衡精度，改进了以往的所有方法。

Mittal等人[8]通过利用全卷积网络的深度架构检测图像中的垃圾,该模型可以将用户上传的带有地理标签的图像进行垃圾区域的检测和分割，在引入的垃圾图像集GINI（Garbage In Images）中进行训练,平均准确率达到了87.69%。吴碧程等人[9]提出了一种基于卷积神经网络的智能垃圾分类系统，该系统将卷积神经网络移植到嵌入式设备树莓派，并且和机械分类结构相结合，在实际操作中嵌入式设备树莓派启动摄像头摄取废弃物样本，预处理之后，运用卷积神经网络获取垃圾图像分类信息，该方案易于实现，成本较低且达到了较高的识别准确率。余希鹏[10]提出了一种基于卷积神经网络的处理道路垃圾图像的方法，对于实时获取的道路图像，测出图像中的垃圾，再将同视角下获得的干净道路图像中的道路区域作为空间约束，实现了垃圾的检测，该方法具有良好的性能，在测试数据集中获得了96.79%的像素精度和93.53%的MeanIOU(Intersection over Union)。

1.2.2 基于纹理特征的图像分类的国内外研究现状

采用图像的纹理进行分类常常采用统计法，频谱法等方法[11]。Hoang[12]采用统计法来实现自动检测沥青路面剥落，他提出采用彩色通道统计特性提取的基于图像纹理的特征和灰度共生矩阵作为输入变量来表征路面状态，根据提取的一组特征将图像样本分为非解离和解离两类，并利用开发好的基于随机梯度下降逻辑回归方法的程序来进行沥青路面状况调查，该方法具有良好的性能。

频谱法是根据波动的频率分布得出波动源的类别的方法，该方法有效可靠，除了用于物理研究也可以用在工业生产中，例如高光谱成像系统，该系统采用高分辨率光谱相机，在探测物体间特征的同时能获得上百个波段带宽小于10nm 的光谱信息[13]因光谱分辨率高、图谱合一、可实现快速无损检测等特点现已广泛应用于农业、医学、遥感等领域。赵冬娥等人[14]通过建立识别分类模型在近红外波段(780 ~ 1000nm)提取了垃圾样本图像的特征波段，并利用SAM(Spectral Angle Mapper)光谱匹配算法和Fisher判别分析算法分别对测试样本ROI(Region Of Interest)区域内测试点分类实现了将图像中待分类的常见垃圾图像如所料瓶、一次性筷子等成功分类，结果表明，利用SAM判别方法在可回收垃圾的高光谱图像中实现检测垃圾图像与分类的准确度更高，达到了99.33%，而通过Fisher判别方法分析训练样本集得出判别函数式和判别准则，对测试样本点集分类，测试样本点集整体分类准确度为99%。

黄兴华等人[15]提出了一种结合了纹理特征融合和支持向量机的分类算法来提取道路图像中的垃圾，实验中首先利用这一改进算法提取垃圾图像纹理特征，之后根据样本图像的特征直方图对支持向量机进行训练分类，最后再对图像样本采用形态学处理提取垃圾的质心位置。实验结果表示除了可以实现稳定的识别率外，还可以准确提取垃圾的质心位置。

由以上可以看出（1）近年来用来进行图像分类的方法种类繁多，在垃圾图像分类方面基于卷积神经网络的方法和改进的基于纹理特征分类的方法应用更为广泛，两种方法也各有优劣（2）利用卷积神经网络进行垃圾图像识别具有很高的识别精度，基本都在85%以上，而且成本较低。（3）利用卷积神经网络进行图像分类需要一定的数据集，但是目前公共资源渠道缺少一个良好的垃圾图像集，因此需要自己建立一个类别精确，基数充分的图像集。本文主要针对居民生活中的垃圾图像进行检测，根据公开资源渠道，现阶段并没有相关的检测手段，因此具有重要的意义。

1.3 垃圾图像分类拟采用的方法

本文对居民生活中拍摄，获取的图像进行检测，将垃圾分类[16]为(1)可回收物,例如废纸，塑料，金属等生活常见的废弃物(2)其他垃圾，包括果壳，卫生纸等(3)厨余垃圾，该类别主要是厨房中剩余的剩菜剩饭和果皮等(4)有害垃圾，例如废弃电池，温度计等。在实现算法上，本文提出了一种基于VGG(Visual Geometry Group)网络的垃圾图像检测分类方法。软件开发上本文的垃圾图像分类系统主要采用Wxpython进行主要页面的开发。

1.4 论文结构

本文共分为七章:

第一章 引言：主要阐述垃圾图像检测系统的意义和背景；之后介绍目前国内外对于垃圾图像分类的研究现状；随后引出本文针对该问题采用的解决方案；最后介绍全文各章的主要内容。

第二章 垃圾图像分类系统分析：本章主要对垃圾图像分类系统进行可行性分析，以及进行需求分析，明确系统的功能需求。

第三章 卷积神经网络概述：本章主要介绍卷积神经网络的基本概念和组成结构；其次阐述经典的也是本文主要采用到的卷积神经网络模型VGGNet模型。

第四章 垃圾图像分类网络模块的设计与实现：本章主要介绍基于VGG16的垃圾图像检测技术的具体实施步骤，首先介绍选用VGG16进行实验的原因，优点；随后阐述本文中自己构建的数据集；最后阐述如何利用VGG16进行图片训练等。

第五章 垃圾图像分类应用模块的设计与实现：本章主要介绍如何利用训练好的网络模型进行图片识别;随后阐述垃圾图像检测模块的操作步骤图；最后介绍垃圾图像报错模块的具体实现图。

第六章 垃圾图像分类系统实验结果：本章主要介绍垃圾图像训练完之后的模型在不同数据集的详细表现；然后介绍垃圾图像存在背景干扰情况下的表现情况，确定系统检测性能的鲁棒性。

第七章 总结与展望：本章主要总结本文所做的具体研究，分析目前仍存在的一些不足，并对之后发展的趋势进行展望。

2 垃圾图像分类系统分析

2.1 可行性分析

该阶段，本文主要通过分析现有的图像分类方法来判断垃圾图像分类系统能否正常开发，开发之后是否具有实用性。下面本文将从经济，技术以及操作可行性三个方面来对垃圾图像分类系统进行可行性分析。

(1)经济可行性:经济可行性主要分析系统的投入和应用以及后续系统在实际应用的价值。

投入方面，本系统在开发过程中所需要的设备极少，使用支持摄像的只能手机和家用的计算机就可以开发。应用方面，用户只需要上传手机拍摄的图片就可以快速正确的得到上传图像所属的垃圾类别，在操作简单的同时，也保证了输出结果的正确性。在实际应用方法，垃圾图像分类系统的普及可以大大的减少用户垃圾分类的困难，便利居民的生活，因此具有一定的应用价值。

(2)技术可行性:技术可行性主要根据当下的算法研究以及软件应用这两个方面来分析垃圾图像分类系统在技术方面的可行性。  
算法研究方面，随着深度学习的兴起，目前应用于图像分类的算法种类繁多。其中神经网络在图像分类中具有广泛的应用，通过研究这些神经网络，从而选择出适合垃圾图像分类的神经网络。再根据修改训练网络的不同参数实现分类准确率较高的垃圾图像分类网络。软件方面，目前软件开发的种类众多，开发语言丰富具备软件开发的条件。因此根据以上的技术分析，垃圾图像分类系统可以很好的实现。

(3)操作可行性:操作可行性指垃圾图像分类系统在操作方面是否可以供用户良好的使用。

垃圾图像分类系统的界面简洁美观，操作简单，而且功能按钮明确，可以使用户很快的掌握系统的使用方法。

2.2 需求分析

本文开发的垃圾图像分类系统，通过分析用户在垃圾分类时遇到的困难，分析出用户具体的功能需求。

(1)系统功能需求分析

图2-1展示了垃圾图像分类系统中各个模块间的不同功能。



图2-1 垃圾图像分类系统功能分布图

(2)系统非功能性分析

易用性:要求系统界面简洁，操作方便，便于用户操作。

可靠性:能够根据用户上传的各种垃圾图像返回出相应的垃圾类别。

稳定性:系统能够长时间使用，而且保持垃圾图像检测的准确性。

扩展性:系统可以根据用户需求的增加，相应完善系统的相关功能。

3 卷积神经网络概述

随着深度学习的快速发展，很多领域都用到了卷积神经网络。卷积神经网络尤其在图像分类中得到了广泛的应用，随着一些典型的模型如LeNet[17],AlexNet[18]，GoogleNet[19],VGGNet的先后诞生发展，逐渐使卷积神经网络在图像识别中成为了主流解决方案。在这其中VGGNet虽然性能并非最优，但是结构相对简洁，实现规整，因此常常被用在多种图像分类任务中。

3.1 卷积神经网络

神经网络的灵感来源于动物的神经系统，人工的神经网络一般表现为互相之间连接的神经元，一个神经元通常由多个负责接收信息的树突和一个负责传递给其他神经元的轴突组成，轴突的末端一般还有多个与其他神经元树突连接的轴突末梢，1943年McCulloch和Pitts参考上述的生物神经元，提出了抽象的神经元模型，神经元模型如图3-1所示:



图3-1 神经元模型

其中，x1,x2,x3为输入，w1,w2,w3为神经元的权重，g是sgn激活函数的值，b为偏执权重，y的对应计算公式如3-1所示:

 (3-1)

如上文所述，多个神经元连接在一起便构成了神经网络，如图3-2是一个双层神经网络:



图3-2 两层神经网络

图2-2中，各层计算公式如公式2-2，2-3，2-4:

 (2-2)

 (2-3)

 (2-4)

两层神经网络通过两层的线性模型了数据之间的真实的非线性函数，因此也可以将神经网络认为是复杂函数的拟合。对于图3-2的两层神经网络，x1,x2,x3为输入层，z为输出层，其余为隐含层。

卷积神经网络与神经网络相比，最大的区别在于卷积神经网络在神经网络的输入层之前连接了卷积层，因此卷积层就变成了卷积神经网络的数据输入层[20]。一个卷积神经网络的架构如图3-3所示:

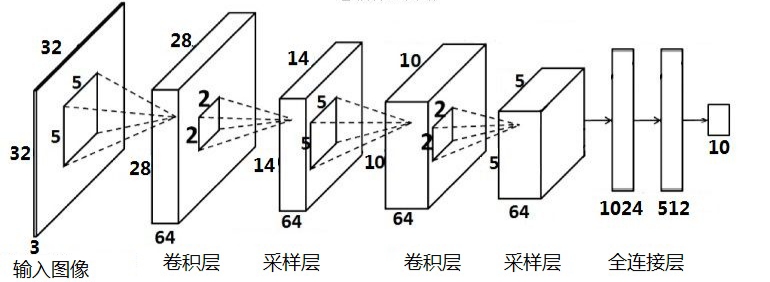


图3-3 卷积神经网络

一个卷积神经网络主要由卷积层，采样层/池化层，全连接层组成，下面详细介绍这三个主要网络层。

1. 卷积层

卷积层由若干个卷积核和偏移权值组成，卷积核和输入层的图像点积和累加可以得到一张特征图。卷积核用于提取图像的局部特征，一个卷积层可以由多个卷积核，每一个卷积核提取图像的局部特征不同。卷积层的卷积过程如图3-4所示:

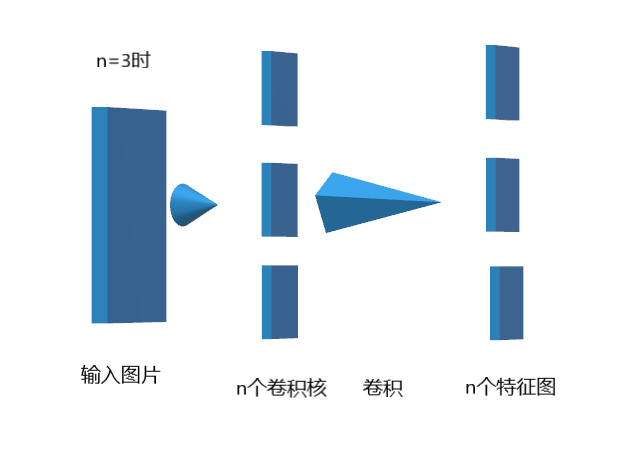


图3-4 卷积过程

图2-5中，对于一张宽w，高h，通道数为c的图片，该图片由c个w列h行的矩阵构成。输入矩阵的通道数=卷积核的通道数。一般来说，假设c=3时，卷积时会根据p(padding:填充,卷积核滑动到图像边缘时，图像边缘的像素填充个数),s(stride：步长，在原图片中水平方向和垂直方向每次的步进长度)从图片的左上角开始，卷积核的对应通道与输入图片进行通道点积累加和的操作，得到一张特征图，然后将3张特征图相对应的位置和偏移矩阵b的相应位置相加，结果即为该卷积核对应的特征图，三个卷积核，就有三个特征图。特征图宽度与高度计算公式如3-5，3-6:

 （3-5）

 (3-6)

(2)采样层/池化层

除了卷积层，卷积神经网络也常常使用采样层/池化层来缩减模型的大小，提升计算的速率，提升提取特征的鲁棒性(Robust)[21],对于一个4\*4的矩阵进行最大池化，最大池化的树池是一个2\*2矩阵，过程如图3-5所示:

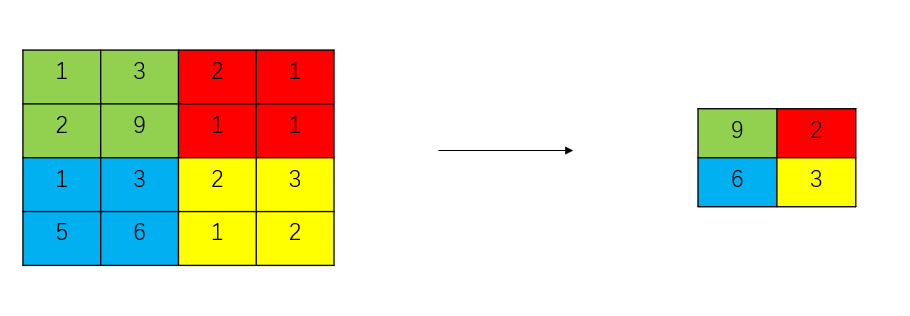
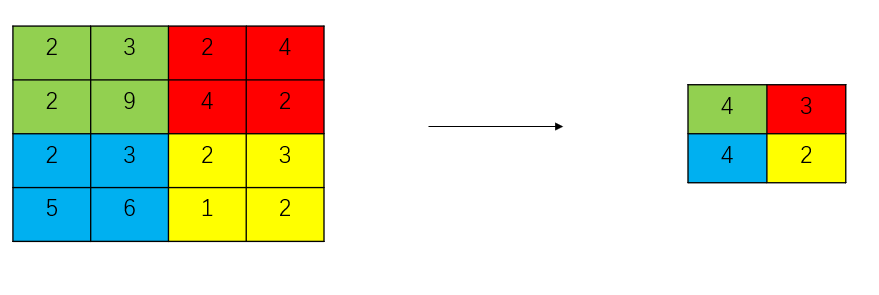


图 3-5 最大池化

最大池化过程相对简单，将4\*4的矩阵划分为不同的2\*2区域，在选择每个区域的最大元素值就构成了最大池化的2\*2结果图。其中应用的2\*2区域，步幅是2，这些就是最大池化的超参数。

除了最大池化还有一种平均池化，平均池化相对而言应用较少，运算过程中和最大池化的不同在于选取的元素为平均值而不是最大元素值，平均池化的过程如图3-6所示:

  
图3-6 平均池化

1. 全连接层

一般来说，全连接层常常出现在卷积神经网络的最后几层中，用于对之前的特征进行加权求和，从而起到分类的作用。该层的神经元和前一层每个神经元都有权重连接，但是相同层的神经元节点互相之间没有连接。

3.2 VGGNet网络

VGGNet网络，全称是Oxford Visual Geometry Group，由牛津大学主要负责机器学习的小组Robotics Research Group提出，在2014年ILSVRC竞赛中取得了第二名的成绩，该模型通过不断堆叠3\*3的小型卷积核和2\*2的最大池化层，成功构建了16-19层的卷积神经网络，现阶段被广泛的用于提取图像的特征。

VGGNet主验证了可以通过不断堆叠网络结构从而提升网络的性能。由于参数量多集中在最后的全连接层上，因此层数的增长并不会造成参数的爆炸。其中两个3\*3的卷积层等同于1个5\*5的卷积层。，如图3-7所示:

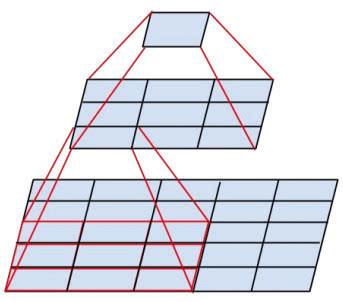


图3-7 两个3\*3的卷积层，类似于一个5\*5的卷积层

同理，三个3\*3的卷积层，可以相当于一个一个7\*7的卷积层，但是相对而言，三个3\*3的卷积层的参数量，只有一个7\*7卷积层的一半。除此之外三个3\*3的卷积层也可以多进行两次非线性操作，因此可以使三个3\*3得卷积层的特征学习能力更强。

下面我们重点介绍本文使用的VGG16的结构，结构如图3-8所示:

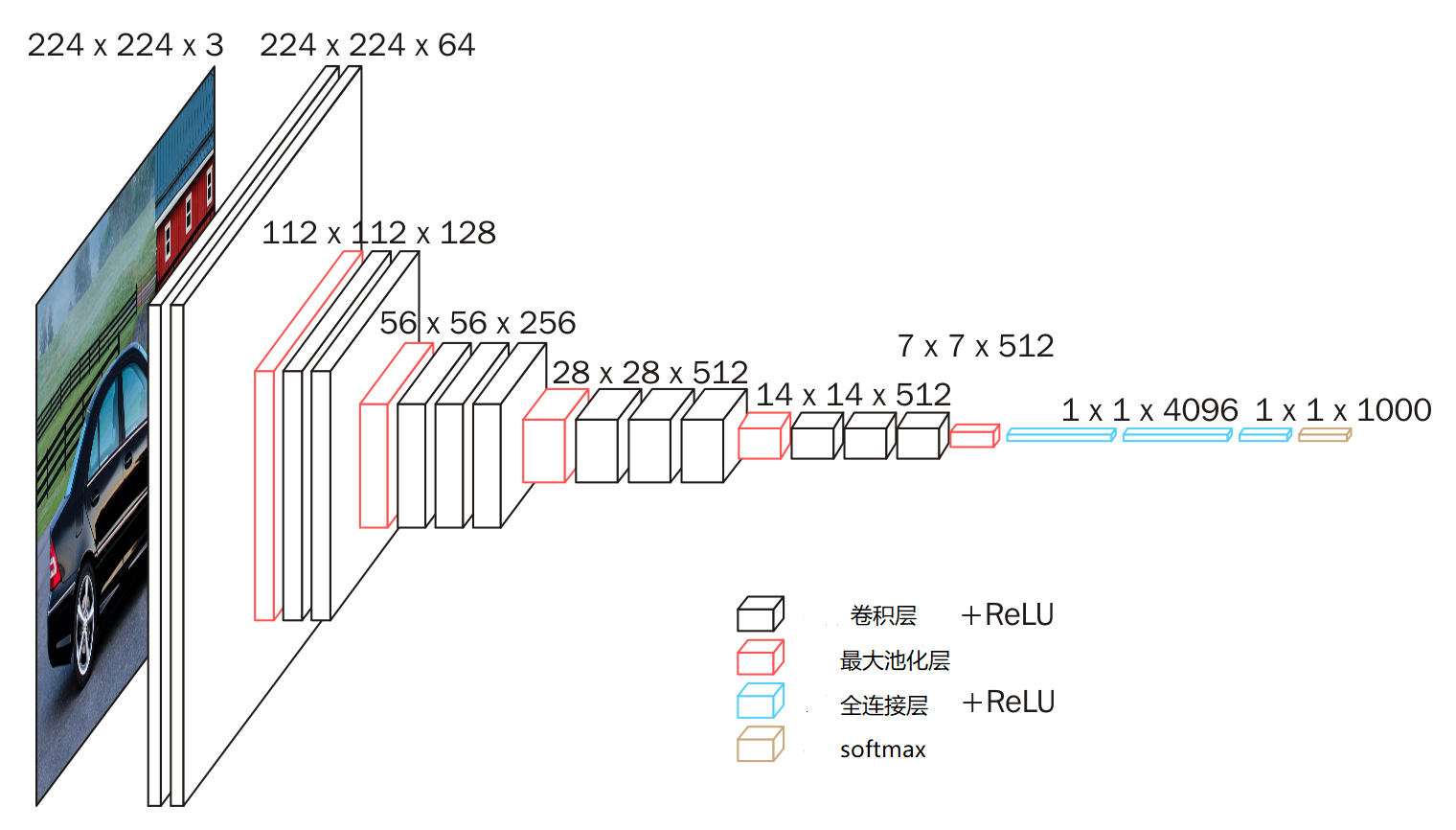


图3-8 VGG16结构图

其中，ReLU[22]（Rectified Linear Unit）线性整流函数，为神经网络中的激活函数，对于输入的正数输出原数，对于负数进行置零操作，如图3-9所示:

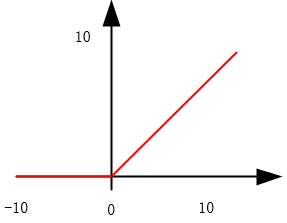


图3-9 ReLU

Softmax[23]是机器学习中一个很常见的函数，对于数组V中的第i个元素，该元素的Softmax值为公式3-7:

 (3-7)

也就是说某个元素的Softmax值为该元素的指数与所有元素指数和的比例，这样就方便了神经网络最后获取到更有可能正确的值.

表2-1为根据卷积核大小和卷积层数目不用的VGG网络的六种不同结构的各层具体参数:

表2-1 六种不同结构的VGG网络各层配置

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ConvNet Configuration | | | | | |
| A | A-LRN | B | C | D | E |
| 11 weight layers | 11 weight layers | 13 weight  layers | 16 weight  layers | 16 weight  layers | 19 weight  layers |
| Input (224 \* 224 RGB image) | | | | | |
| conv3-64 | conv3-64  **LRN** | conv3-64 **conv3-64** | conv3-64 conv3-64 | conv3-64 conv3-64 | conv3-64 conv3-64 |
| maxpool | | | | | |
| conv3-128 | conv3-128 | conv3-128  **conv3-128** | conv3-128 conv3-128 | conv3-128 conv3-128 | conv3-128 conv3-128 |
| maxpool | | | | | |
| conv3-256 conv3-256 | conv3-256 conv3-256 | conv3-256 conv3-256 | conv3-256 conv3-256 **conv1-256** | conv3-256 conv3-256 **conv3-256** | conv3-256 conv3-256 conv3-256 conv3-256 |
| maxpool | | | | | |
| conv3-512  conv3-512 | conv3-512  conv3-512 | conv3-512  conv3-512 | conv3-512  conv3-512  **conv1-512** | conv3-512  conv3-512 **conv3-512** | conv3-512  conv3-512 conv3-512  **conv3-512** |
| maxpool | | | | | |
| conv3-512  conv3-512 | conv3-512  conv3-512 | conv3-512  conv3-512 | conv3-512  conv3-512 **conv1-512** | conv3-512  conv3-512 **conv3-512** | conv3-512  conv3-512 conv3-512  conv3-512 |
| maxpool | | | | | |
| FC-4096 | | | | | |
| FC-4096 | | | | | |
| FC-1000 | | | | | |
| Soft-max | | | | | |

D为VGG16网络所采用的详细参数配置，根据表所示，VGG16一共包含了13个卷积层，3个全连接层，5个池化层，其中卷积层和全连接层由于具有权重系数，因此也可以成为权重层，总层数为卷积层和全连接层数的和16层。

总的来说，VGG16网络结构相对简单，而且在图像识别中应用广泛，可以作为很好的解决图像分类问题的手段。

4 垃圾图像分类网络模块的设计与实现

垃圾图像分类网络模块是垃圾图像分类系统中的核心部分。正如上文所述，卷积神经网络在图像分类中具有广泛的应用，因此本文决定经典的卷积神经网络VGG16来进行垃圾图像检测。

4.1 VGG16的优点:

VGG16和VGG19都是VGGNet中常用的网络，相对而言VGG19的卷积层比VGG16的卷积层要多3层，本质上并没有区别。虽然在性能上来说可能VGG19要略微出色一些，但是VGG16的结构更加轻量化，因此在训练上更加快速，容易些。因此在性能差别并不大的情况下，我们选择更能节省时间的前者，在实际应用中VGG16也更广泛些。

4.2 垃圾图像数据集的构建:

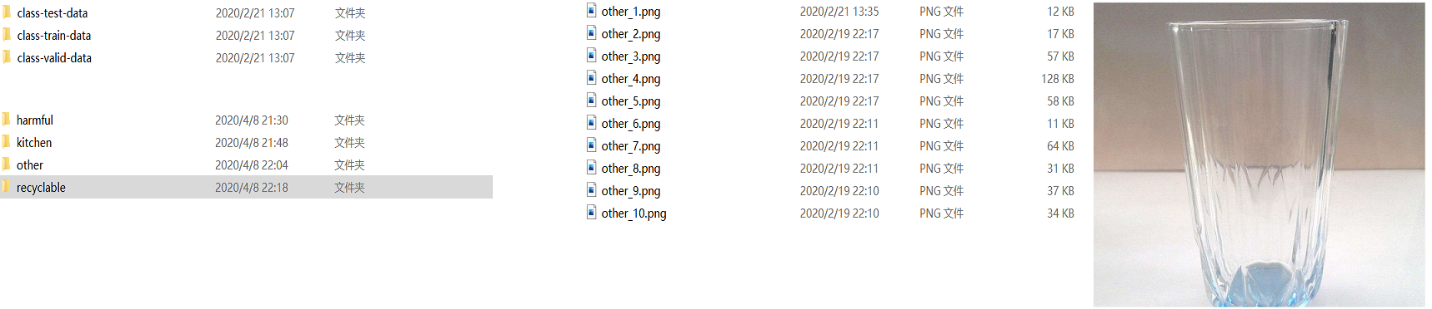
图像数据集在图像分类中尤为重要，一般来说，越大越精确的数据集，训练出来的神经网络准确率更高，目前在计算机视觉领域有几种比较常见的数据集，比如ImageNet[24]，CIFAR[25]（Canada Institude For Advanced Research），LFW[26]（Labeled Faces in the Wild）等，但是在垃圾图像方面还没有一个比较公开的，常用的数据集，因此本文构建了一个小型的垃圾图像数据集，图像集如图4-1：  


图4-1 垃圾图像数据集

(左边为数据集内的各个文件夹，最右侧为玻璃杯属于可回收垃圾)

该数据集由200张人为拍摄或者来源网络的垃圾图像组成，图像采集完全后，按照6:2:2的比例将图像集分为训练集,测试集，验证集。之后再按照类别将其分为harmful,other,recyclable,kitchen四类,最后再统一按照图片类别，将其重命名为类似harmful\_1的形式。

4.3 基于VGG16的垃圾图像检测模型:

本节在VGG16框架下，利用Python完成垃圾图像集的训练。

* + 1. 数据集的导入与图像预处理

(1)导入数据集

创建名为cls2.py的文件，首先导入训练图片，代码如下所示:

for d in os.listdir(traindir):  
 categories.append(d)  
 # 图片的数量  
 train\_imgs = os.listdir(traindir + d) //导入训练集图片  
 valid\_imgs = os.listdir(validdir + d)//导入验证集图片  
 test\_imgs = os.listdir(testdir + d)//导入训练集图片  
 n\_train.append(len(train\_imgs))//添加训练集图片数量  
 n\_valid.append(len(valid\_imgs))//添加验证集图片数量  
 n\_test.append(len(test\_imgs))//添加测试集图片数量  
 # 导入训练图片   
 for i in train\_imgs:  
 img\_categories.append(d)//获取类别  
 img = Image.open(traindir + d + '/' + i)//按照图片名字依次打开图片  
 img\_array = np.array(img)//

其中triandir，testdir为数据集路径，在本文中路径为: datadir = 'E:/class\_sortrubbish/',traindir = datadir + 'class-train-data/'

(2)图片预处理

将图像的尺寸转化为224\*224,改变图片的对比度，对图片用均值和标准差等进行标准化处理，代码如下:

transforms.Compose([ #Transfoms 是很常用的图片变换方式，可以通过compose将各个变换串联起来  
 transforms.Resize((224,224)), #将输入的图片转换成给定的尺寸的大小  
 transforms.RandomRotation(degrees=15), #按角度旋转图像。  
 transforms.ColorJitter(), #随机改变图片的亮度、对比度和饱和度。   
 # 标准化图像  
 transforms.ToTensor(), #将图片转成Tensor类型的  
 transforms.Normalize([0.485, 0.456, 0.406], #用均值和标准差对张量图像进行标准化处理  
 [0.229, 0.224, 0.225]) # Imagenet standards  
]),

(3)加载数据集的图像

代码如下 :  
'train':  
datasets.ImageFolder(root=traindir, transform=image\_transforms['train'])

4.3.2 迁移学习

迁移学习的主要思想是从相关领域迁移标注数据或者知识结构、完成或改进目标领域或任务的学习效果。首先我们导入已经在庞大数据集训练好的VGG16预训练模型，由于预训练模型中之前的层的参数都已经相对充足，无需训练，因此将分类层之前的层进行冻结操作，之后引入新定义的分类层，从而我们在训练时，只用训练自己的分类层即可，总的来说使用迁移学习加快了我们的训练时间，提升了网络的性能。

(1)导入预训练模型

model = models.vgg16()  
model.load\_state\_dict(torch.load('E:/class\_sortrubbish/vgg16-397923af.pth')) #model.load\_state\_dict(torch.load(PATH))恢复模型中的参数：使用这种方法，我们需要自己导入模型的结构信息。

预训练模型是一种深度学习架构，通过提前在一些其他的庞大的数据集训练过的初始化参数，再根据本文中的垃圾图像集进行学习，得到适合垃圾图像集的参数，本文用到的是从Pytorch官方下载的VGG16的预训练模型。因为在迁移学习中需要使用别人训练好的权重参数，冻结预测层之前的所有的权重参数，进行接下来的训练，因此需要冻结层的操作。

(2)冻结VGG16中分类层之前的层

for param in model.parameters(): #对原模型中的参数进行遍历操作

param.requires\_grad = False #将参数中的param.requires\_grad全部设置为False，这样对应的参数将不计算梯度，同时也就不会再进行梯度更新了

(3)定义新的分类层

之后，定义新的全连接层结构并重新赋值给model.classifier。（考虑到输出结果个数需修改全连接层）在完成了新的全连接层定义后，全连接层中的parma.requires\_grad参数会被默认重置为True，所以不需要再次遍历参数来进行解冻操作。

pytorch使用torch.nn.Sequential快速搭建神经网络：torch.nn.Sequential是一个顺序容器，模块将按照构造函数中传递的顺序添加到模块中。另外，也可以传入一个有序模块代码如下:  
model.classifier[6] = nn.Sequential(

nn.Linear(n\_inputs, 256),

nn.ReLU(),

nn.Dropout(0.2),

nn.Linear(256, n\_classes),

nn.LogSoftmax(dim=1)

* + 1. 网络训练

(1)损失函数与优化函数

损失函数是机器学习中很常见的一个函数，主要用来负责衡量预测模型的性能，表现预测和实际数据的差距程度，一般来说损失函数越小，模型的鲁棒性越好，本文损失函数选择的是负对然似然(negative log likelihood )，公式如4-1:

 (4-1)

在计算出模型的损失值后，需要利用损失值来进行模型参数的优化，本文的优化函数使用的是Adam(Adaptive Moment Estimation)[27]算法,Adam算法是深度学习中十分常见的一种优化算法，在模型训练优化的过程中，通过让每个参数获得自适应的学习率，达到优化质量和速度的双重提升。代码如下:

criterion = nn.NLLLoss() #Negative Log Liklihood(NLL) Loss:负对数似然损失函数。

optimizer = optim.Adam(model.parameters(), lr=0.0001,) #优化函数

(2)梯度

下面我们介绍一下下面过程中计算损失函数使用到的梯度的概念，梯度可以定义为一个函数的全部偏导数构成的向量，一般将函数f的梯度记作: *∇f*,计算公式如4-2:

（4-2）

梯度向量的方向即为函数值增长最快的方向，也就是说沿着梯度方向，函数值增长最快。机器学习中常用*γ*表示学习率(Learning rate),同时也在梯度下降法中表示步长，一般来说通过计算出目标函数的梯度，并且在其反方向更新完参数，此过程也就是达到了函数值减小最快的效果，也就是说在经过迭代之后，目标函数就可以很快的达到一个极小值，如果这个极小值也是全局最小值，那么使用梯度下降法就可以保证收敛到全局最优解。

(3)网络训练

定义训练函数,batchsize = 64,batchsize为一次训练提取的样本数;训练次数Epoch=30,代码如下所示:

#定义训练函数

def train(model, #模型

criterion, #损失函数

optimizer, #优化函数

train\_loader,

valid\_loader,

save\_file\_name,

max\_epochs\_stop=3,

n\_epochs=20, #训练次数总共是20次

print\_every=2):

Model.train() #启用 BatchNormalization 和 Dropout,否则一旦batchsize过小，会造成失真。

#开始在训练集循环训练每张图片

for ii, (data, target) in enumerate(train\_loader):

# Tensor转到gpu(图像处理器)

if train\_on\_gpu:

data, target = data.cuda(), target.cuda()

print("data:",data)

print("target:", target.data)

#梯度初始化为零，便于下面计算loss

optimizer.zero\_grad()

(4)计算损失率和准确率

Loss(损失函数)和acc(accuracy:准确率 )是评量一个模型好坏的基本指标，计算代码如下:

#梯度的loss

loss = criterion(output, target)

#反向传播

loss.backward()

# 更新梯度参数

optimizer.step()

#平均损失\*样本的数量计算训练的损失

train\_loss += loss.item() \* data.size(0)

#通过计算最大对数概率来计算精度

\_, pred = torch.max(output, dim=1)

correct\_tensor = pred.eq(target.data.view\_as(pred))

# 平均准确度\*样本的数量计算训练准确率

accuracy = torch.mean(correct\_tensor.type(torch.FloatTensor))

train\_acc += accuracy.item() \* data.size(0)

验证集样本同上，这里不再多做赘述，平均损失和平均准确率的代码如下:

# 平均损失=总的损失/数据集总长度

train\_loss = train\_loss / len(train\_loader.dataset)

valid\_loss = valid\_loss / len(valid\_loader.dataset)

# 平均准确率=总的准确率/数据集总长度

train\_acc = train\_acc / len(train\_loader.dataset)

valid\_acc = valid\_acc / len(valid\_loader.dataset)

(5)保存训练好的模型

代码如下:

def save\_checkpoint(model, path):

model\_name = path.split('-')[0]

assert (model\_name in ['vgg16'

]), "Path must have the correct model name"//若路径出错，及时报错

checkpoint = {

'class\_to\_idx': model.class\_to\_idx,

'idx\_to\_class': model.idx\_to\_class,

'epochs': model.epochs,

}//定义存储数据结构

# 提取最终分类器和状态字典

if model\_name == 'vgg16':

checkpoint['classifier'] = model.classifier

checkpoint['state\_dict'] = model.state\_dict()

# 添加优化器

checkpoint['optimizer'] = model.optimizer

checkpoint['optimizer\_state\_dict'] = model.optimizer.state\_dict()

# 保存到指定路径

torch.save(checkpoint, path)

4.3.4 实验环境

本文的操作系统采用的Windows10,64位操作系统;用到主要的编程语言是Python;训练过程中用到了pytorch,numpy,pandas,os等，主要使用NIVIDA GPU提升训练的效率,GPU的版本为 GeForce GTX 1050 Ti,计算机内存为 8G。

5 垃圾图像分类应用模块的的设计与实现

5.1 导入模型

首先，导入之前训练好的模型，名为vgg16-transfer-4.pth，路径为训练后导入的路径，具体代码如下:

checkpoint = torch.load(path, map\_location='cpu')

model = models.vgg16()

#读取分类器等具体信息

model.classifier = checkpoint['classifier']

model.load\_state\_dict(checkpoint['state\_dict'])

5.2 对预测图片进行预处理

#修改图片的尺寸，标准差，均值等，保持和训练时的图片一致

img = image.resize((224, 224))

img = np.array(img).transpose((2, 0, 1)) / 448

means = np.array([0.485, 0.456, 0.406]).reshape((3, 1, 1))

stds = np.array([0.229, 0.224, 0.225]).reshape((3, 1, 1))

5.3 预测图片

根据图片路径，读取图片，返回预测图片的类别。

#使模型开始预测，训练时为model.train()

model.eval()//避免输入数据会改变权值

out = model(img\_tensor)

ps = torch.exp(out)//获取网络中返回的类别标签指数

#求最可能性最大的类别

topk, topclass = ps.topk(topk, dim=1)

# 提取最终获得的类别信息

top\_classes = [

model.idx\_to\_class[class\_] for class\_ in topclass.cpu().numpy()[0]

]

5.4 垃圾图像识别操作步骤图

本模块的具体操作步骤图，如图5-1:



图 5-1 垃圾图像检测操作流程图

如上图所示，用户登录系统之后，可以从本地选择识别的图片，图片选择好之后再调用训练好的VGG16分类网络，图片在网络中经过处理后，返回最有可能类别的标签，之后再根据标签转换为最后的输出结果。(例如标签为harmful,返回有害垃圾)

5.5 垃圾图像报错

虽然算法取得了不错的识别精度，但是仍然可能存在误判的情况，针对这一问题，本文引用了实时报错功能，对于识别错误的图像用户只需要简单的步骤就可以上传给管理人员错误的图像，管理人员再定期将错误图像加入到垃圾图像集中进行训练，可以不断完善算法的识别精度，报错过程如下图5-2:  


图5-2 垃圾图像报错过程

具体实现代码如下:

(1)获得上传图片

def show\_message(self, word=""):

#word = "您要上传的是这张图片?"

file = open(self.FileName.GetValue())#获取垃圾图像文件路径

err\_img1 = str(self.FileName.GetValue())#将路径转化为字符串

err\_img1 = err\_img1.replace('\\', '/')#替换字符串中的’\\’

dlg = wx.MessageDialog(None, "您要上传的是这张图片："+err\_img1 +"?", u"报错", wx.YES\_NO | wx.ICON\_QUESTION)#确认用户的上传图片

if dlg.ShowModal() == wx.ID\_YES:

self.show\_message2(err\_img=err\_img1)

dlg.Destroy()

(2)获得正确垃圾图像种类

box = wx.TextEntryDialog(None, '请输入正确的垃圾种类', '报错', '可回收物')#对话框，用户输入正确垃圾种类

(3)发送到管理员的邮箱

s = smtplib.SMTP\_SSL("smtp.qq.com", 465) # 邮件服务器及端口号

s.login(msg\_from, passwd)#根据发送邮箱用户账号，密码登录

s.sendmail(msg\_from, msg\_to, msg.as\_string())#邮件发送

print("发送成功!")#发送成功后显示

6 垃圾图像分类系统实验结果

网络训练完之后，在训练集和验证集的准确率和损失率如图6-1,图6-2

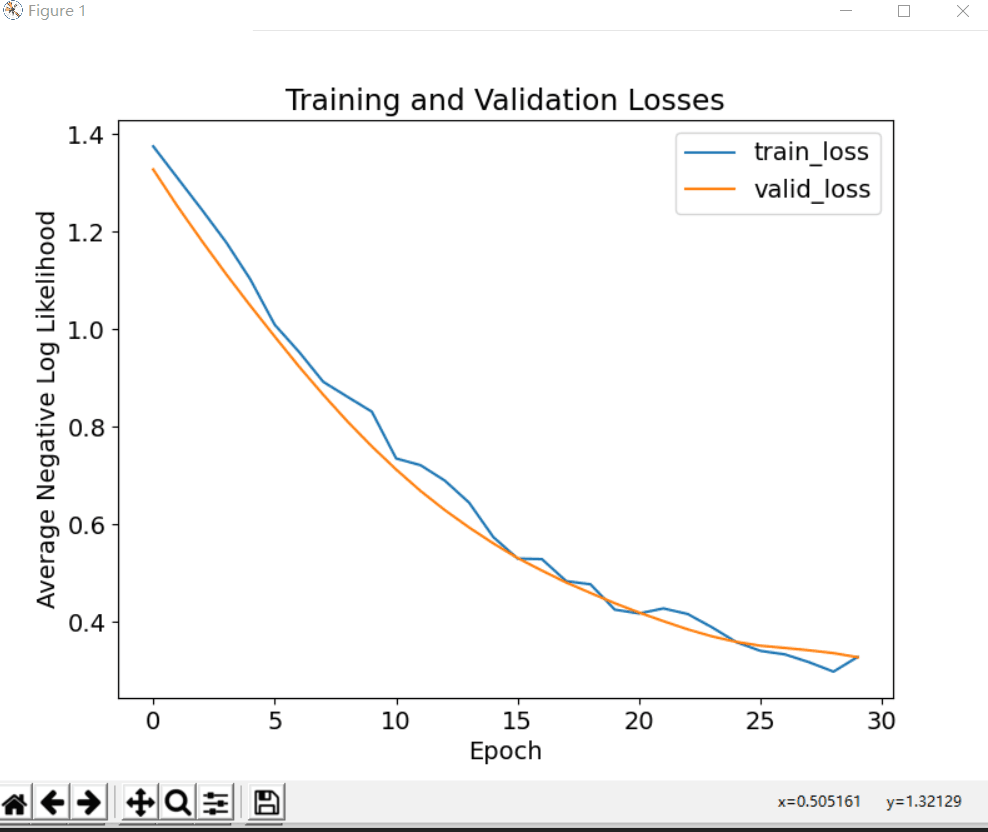


图6-1 训练集和验证集消耗率

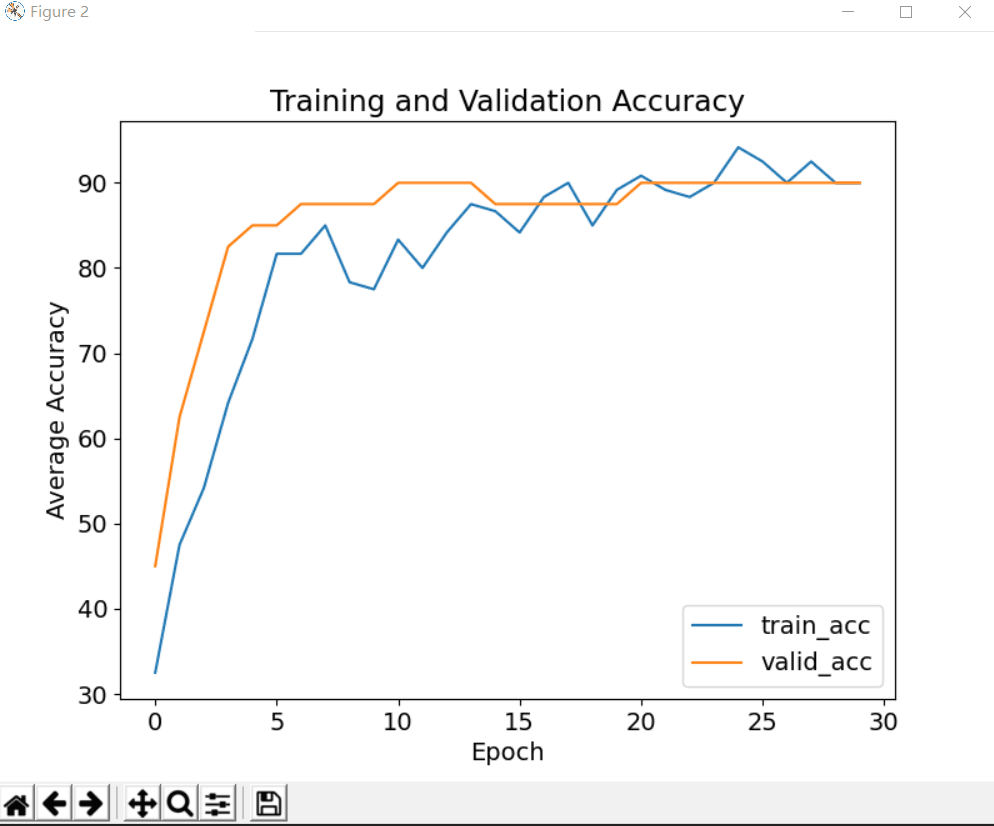


图6-2 训练集和验证集准确率

由图4-1，4-2可以发现，随着训练次数(Epoch)的上升，损失率在一直减少，说明在此过程中一直在学习，在接近30次的时候train\_loss略有回升，valid\_loss趋于稳定但是总体较小，且接近，结合train\_acc和valid\_acc的不断上升慢慢趋于90%，可以说明网络对于数据的训练是处于适度拟合的。

之后再利用训练好的网络模型在从未训练过的40张测试集上测试准确度,准确度如下图6-3所示:

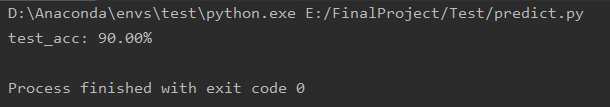


图6-3 测试集准确率

由上图可知，训练好的模型可以在40张测试集中取得90%的准确率，下图6-4为三张图像背景存在干扰的预测结果:



图6-4 干扰背景下的垃圾图像检测

上图中分别是骨头，药品，以及矿泉水瓶三张图像存在一定干扰情况下的检测结果，可以看到即使是在背景有所干扰的情况下，垃圾图像的识别结果依旧很准确。

从检测结果和测试集准确率可以得出以下几种结论:

(1)基于VGG16的垃圾图像检测平均准确率很高，在测试集，验证集都取得了高达90%的准确率。

(2)对于不同环境下的垃圾图像检测鲁棒性很高，面对背景干扰仍能正确的识别出垃圾种类。

7 总结与展望

7.1 全文总结

本文主要研究了垃圾图像的检测算法，基于当前垃圾分类的大力推行的背景，针对大多数居民对于垃圾种类的陌生，自主分类困难的问题，提出了一种基于VGG16网络的垃圾图像分类算法。根据用户实时上传的生活垃圾图像，首先将图像处理成训练网络匹配的格式，随后根据分类层给出的垃圾种类概率，判别出最有可能的垃圾类别，最后再根据不同用户的生活城市，返回出垃圾的正确类别。在本文自主构建的数据集上获得了很高的识别率，同时在一定的背景干扰下也取得了不错的效果。本文主要完成的工作如下:

(1)分析了目前垃圾分类面对的时代背景，概括了垃圾图像分类的研究意义，总结了目前国内外对于垃圾图像分类的研究现状。并且对于本文中使用的卷积神经网络进行了一定程度的概括。面对目前缺少垃圾图像数据集的问题，通过自己拍摄和网络搜索，搜集了一定数量的常见垃圾图像，构建了垃圾种类齐全的垃圾图像数据集。

(2)面对目前不同城市的垃圾种类繁多不统一的问题，收集了各个城市的垃圾分类材料，结合不同城市的不同政策，不同称呼，将垃圾种类根据不同的城市返回相应的结果，如厨余垃圾在上海与北京分别为湿垃圾和厨余垃圾。防止了用户面对垃圾返回结果困惑的问题。

(3)针对识别过程中可能会存在的问题，本文采取了用户可以实时报错的功能，面对分类错误的垃圾图像，用户可以实时上传图片，并将图片邮件给管理人员，管理人员定期查看识别错误的图像，将图像加入到训练集，不断提升算法的识别精度和效率。

总的来说，算法表现程度良好，在存在背景干扰的测试集中仍然取得了90%的准确率。

7.2 本文存在的不足与展望

虽然实验整体效果不错，但是本文仍存在一定的不足，还有一定的进步空间：

(1)本文的数据集种类划分，本文根据垃圾图像的种类，将不同的物体归并为同一种的垃圾，然而物体的繁多必然使垃圾图像识别相对困难，降低了垃圾图像的识别精度，因此如果将图像按照自身的种类进行分类，再根据识别的种类判断输入何种垃圾，相对而言更为科学，可以大大提升算法的识别准确度和效率。

(2)本文提出的垃圾图像算法只能针对一种垃圾图像进行识别提取，然而生活中有时需要识别的垃圾图像中的种类远不止一种，因此如何处理多种垃圾堆叠于一种图像的问题，仍然需要进一步的研究。

(3)本文采取的垃圾图像种类只选择了生活中最为常见的几种，然而有时面对一些冷门的垃圾用户更有可能会束手无措，因此构建垃圾种类更加丰富，垃圾图像更为繁多的垃圾图像集也是未来需要解决的问题。

装

订

线

参考文献

1. 梅啟成.基于深度学习的商品图像识别方法研究[D].广州：广东工业大学,2018.
2. 朱芸芸.基于卷积神经网络的图像分类方法研究[D]. 北京：北京交通大学, 2016.
3. 王晓东.纹理图像分类系统及其关键技术研究[D]. 南京：南京邮电大学, 2013.
4. Wang D, Tang J L, Zhu W J, *et al*. Dairy goat detection based on Faster R-CNN from surveillance video[J]. Computers and Electronics in Agriculture. 2018, 154.
5. Cheng, Phillip M, Malhi, *et al*. Transfer Learning with Convolutional Neural Networksfor Classification of Abdominal Ultrasound Images[J]. Journal of Digital Imaging. 2017, 30(2):234－243．
6. 罗浩伦, 冯泽霖, 冉钟南等. 基于VGG16网络的茶叶嫩芽自动检测研究[J]. 农业与技术. 2020, 40(01):15-17.
7. Picon A, Seitz M, Alvarez-Gila A, *et al*. Crop conditional convolutional neural networks for massive multi-crop plant disease classiﬁcation over cell phone acquired images taken on real ﬁeld conditions[J]. Computers and Electronics in Agriculture. 2019, 167: 1-10.
8. Mittal G, Yagink K B, Grag M, *et al*. SpotGarbage: Smartphone App to Detect Garbage UsingDeep Learning[C]. Acm International Joint Conference on Pervasive & Ubiquitous Computing. 2016:940-945.
9. 吴碧程, 邓祥恩, 张子憧等. 基于卷积神经网络的智能垃圾分类系统[J]. 物理实验. 2019, 39(11): 44-49
10. 余希鹏.基于卷积神经网络的固定视角道路垃圾检测[D].武汉：华中科技大学,2019.
11. 王龙. 图像纹理特征提取及分类研究[D]. 青岛：中国海洋大学, 2014.
12. Hoang N. Automatic detection of asphalt pavement raveling using image texture based feature extraction and stochastic gradient descent logistic regression[J]. Automation in Construction. 2019, 105: 1-12.
13. 刘立新, 李梦珠, 赵志刚等.高光谱成像技术在生物医学中的应用进展[J].中国激光. 2018,45(02):214-223.
14. 赵冬娥, 吴瑞, 赵宝国等. 高光谱成像的垃圾分类识别研究[J].光谱学与光谱分析. 2019, 39(03): 921-926.
15. 黄兴华, 叶军一, 熊杰. 基于纹理特征融合的道路垃圾图像识别及提取[J]. 计算机工程与设计. 2019, 40(11): 3212:-3218:3305.
16. Tong Y Q, Liu J F, Liu S Z. China is implementing “Garbage Classification” action[J]. Environmental Pollution. 2020, 259.
17. Oliveira F M C, Borin E. Partitioning Convolutional Neural Networks to Maximize the Inference Rate on Constrained IoT Devices[J]. Future Internet. 2019, 11(10)
18. Boedi R M, Banar N, Tobel J D, *et al*. Effect of Lower Third Molar Segmentations on Automated Tooth Development Staging using a Convolutional Neural Network[J]. Journal of Forensic Sciences. 2020, 65(2).
19. Teerakawanich N, Leelaruji C, Pichetjamroen A. Short term prediction of sun coverage using optical flow with GoogLeNet[J]. Energy Reports. 2020,6(Supl.2).
20. 吴正文.卷积神经网络在图像分类中的应用研究[D].成都：电子科技大学,2015.
21. 赵长安,贺风华．多变量鲁棒控制系统[M].哈尔滨：哈尔滨工业大学出版社, 2011
22. Chen Z, Ho P H. Global-connected network with generalized ReLU activation[J]. Pattern Recognition. 2019, 96.
23. 万磊, 佟鑫, 盛明伟等.Softmax分类器深度学习图像分类方法应用综述[J].导航与控制.2019, 18(06):1-9, 47.
24. Kornblith, Simon, Shlens J, *et al*. Do better imagenet models transfer better[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition . 2019
25. 陈泓佑,陈帆,和红杰等.基于样本特征解码约束的GANs[J/OL].自动化学报.1-14[2020-04-10].https://doi.org/10.16383/j.aas.c190496.
26. Sun J, Yang W, Gao R Q,Xue J H, *et al*. Inter-class angular margin loss for face recognition[J]. Signal Processing: Image Communication. 2020, 80.
27. Kingma D, Ba J. Adam：a method for stochastic optimization. Computer Science[J]. 2014

致 谢

四年的大学时光一晃而过，四年的时间收获到了很多，除了知识以外，学习到更多的是关于为人，关于品质，在这里想感谢很多人，感谢老师同学们的包容和教导才有了本科阶段的收获。

首先想感谢，我毕业设计项目的导师田学东教授，很荣幸在毕业设计选题阶段能选中田老师。还记得我第一次看到毕业设计的选题时，垃圾图像分类就使我眼前一亮，我觉得这个选题既能学习到图像处理，深度学习的知识也能在真正完成之后有所应用。之后在整个毕业设计的过程中，田老师严格的要求，一丝不苟的态度也让我学到了很多；在开题报告阶段，由于本人在听课时的心猿意马导致一开始的开题报告写的漏洞百出，幸好有田老师不厌其烦的一遍又一遍的指导修改，才能顺利完成开题报告。毕业设计项目的完成过程中田老师的许多建议也补充了自己在设计时没有想到的问题，丰富了系统的架构。在此，相对老师在整个过程中的辛勤付出，真诚的说一声谢谢！

其次我想感谢本次毕业设计同组的各位同学们，仍记得在我实验进行的过程中，本人心态多次出现问题，有几次甚至出现了放弃的念头，在此想感谢我的同学们耐心的在我遇到困难时给与的积极开导和帮助。虽然只和大家在一起完成毕业设计的时间才聊聊数月，但是我相信这段岁月将是我未来路上永远的回忆。

感谢我的家人，感谢我的父母对我一直以来的信任和支持，同时也给我了良好的生活环境，让我不用操心学习之外的琐碎杂事，感谢我的父母，我会在未来的道路上一直铭记父母的教诲，也会用自己未来的行动来回报父母对我无私的支持。

最后也感谢评审老师对于论文的耐心阅读和批评指导！