

基于深度学习的垃圾分类系统设计与实现

摘要

自打从工业以来，机械逐渐代替手工，人类的生产力水平得以指数级提升。生产力水平提升的同时也使得垃圾产生的数量在飞速上涨。在国内越繁华的城市，人均垃圾产量越是大大超出了国内平均水平。大量的垃圾危害了人们的身体健康，也侵蚀着人类的活动区域面积。如何处理垃圾、怎样做到环保高效，这在我国以及全世界范围内都是一个亟待面对的棘手问题。但随着深度学习技术的发展，借助 python 技术使自动分拣垃圾具有了实现的可能性。

本系统旨在利用计算机高速的处理计算能力，采用现在流行的深度学习方法，构建神经网络，通过摄像头捕捉图片信息，对捕捉到的垃圾图像进行分类，同时进行及时的数据反馈和操作。本文的主要工作是以深度学习的理论和技术作为基础，并且针对我国垃圾分类现状以及垃圾处理所面临的多种问题，将改进后的网络模型应用到垃圾分类处理中，通过学习大量垃圾图片的特征得到模型参数并进行垃圾检测。尝试将深度学习运用到垃圾分类上，实现一个自动分类识别视频监控并标注 Python 系统。同时使用当下流行的小程序技术，将传感器采集到的数据通过 ESP8266 模块发送到服务端，构建起一个垃圾信息实时查看系统，便于工作人员随时查看垃圾箱里的垃圾信息并且及时对这些垃圾信息进行处理。

关键词：垃圾分类；深度学习；Python；OpenCV

1. 绪论

1.1 选题背景及研究意义

垃圾，可以被定义为无用的东西，也可以认为是错误放置的社会资源。垃圾与人类的生活息息相关。垃圾的产生总是不可避免的，但是对垃圾进行处理是可以做到的。随着时代的发展和科技的进步，我们可以有多种不同的手

段处理垃圾。在古代我国，人们曾经用填埋、焚烧方式处理垃圾。但那个时候由于垃圾产生量较少，用焚烧处理的方式并未造成太多的生态问题。自打从工业以来，机械逐渐代替手工，人类的生产力水平得以指数级提升。生产力水平提升的同时也使得垃圾产生的数量在飞速上涨。据不完全统计，仅 2018 年，我国垃圾的总产量达到 2.2801 亿吨，人均垃圾产量达到了 0.44 吨。在国内越繁华的城市，像北京、上海、广州、深圳，人均垃圾产量越是大大超出了国内平均水平。大量的垃圾危害了人们的身体健康，也侵蚀着人类的活动区域面积。世界上有一些城市已经被垃圾包围，甚至形成垃圾城[1]。

所以，如何处理垃圾、怎样做到环保高效，这在我国以及全世界范围内都是一个亟待面对的棘手问题。

目前我国在垃圾处理的问题上没有明确的法律规定和设置惩罚制度，只能依靠人们的自觉来完成。而且公民从小在垃圾处理、垃圾分类的教育宣传不够，更加剧了后期垃圾处理人员对垃圾分拣处理难度。

合理地垃圾进行分类是最终处理垃圾之前的关键步骤。

公民将垃圾按类别投放进不同的垃圾箱之后，垃圾被统一运送到垃圾处理厂。目前国内对垃圾进行处理的方式大多都是采用人力的方式，由专门的工作人员对垃圾进行分类。这种处理方式效率太过于低下，跟不上社会高速发展的节奏。与此同时也满足不了对海量垃圾的处理需求。而且工作人员长时间工作在垃圾处理的场所，对其身体健康非常不利。

深度学习技术在我国已经是日渐成熟，但是建立在深度学习基础上对垃圾进行分类处理，目前还属于刚起步阶段。由于垃圾的形状不一以及颜色变化较大，不容易手动提取类别特征，算法的识别精度不高，反应的速度也不能满足生产现场嵌入式设备的实时性要求。本文的工作目标是将深度学习方法用于垃圾图像识别的领域，开发垃圾分类中准确率高、速度快的深度学习算法，以利于设备快速分拣垃圾，提高效率，并降低人工成本。所以，本课题有较大的理论意义与应用价值。

1.2 研究现状和发展趋势

深度学习是机器学习里面的一部分。所谓深度学习，主要是从零到一仿照人脑思维的运作方式构建出来的“机器脑袋”。深度学习这个概念在 1940 年前后被提出，经过多年的发展，深度学习的模型越来越多元化，准确性也在不断提升。我国在深度学习方面起步较晚，但在某些领域已经取得重大的成功。当今社会已经逐渐走向“智能化”，而深度学习作为其中一部分，在未来的发展中将会扮演着重大的角色。

1.3 项目总体目标

针对垃圾日产量大的严峻形势以及垃圾分类难的现状，本文旨在利用计算机高速的处理计算能力，采用现在流行的深度学习方法，构建神经网络，通过摄像头捕捉图片信息，对捕捉到的垃圾图像进行分类，同时进行及时的数据反馈和操作。本文的主要工作是以深度学习的理论和技术作为基础，并且针对我国垃圾分类现状以及垃圾处理所面临的多种问题，将改进后的网络模型应用到垃圾分类处理中，通过学习大量垃圾图片的特征得到模型参数并进行垃圾检测。尝试将深度学习运用到垃圾分类上，实现一个自动分类识别视频监控并标注 Python 系统。同时以 arduino 为 MCU，通过 ESP8266 Wifi 模块将信息发送到服务端，使用当下流行的小程序技术，构建一个垃圾信息实时查看系统，便于工作人员随时查看垃圾箱里的垃圾信息并且及时对这些垃圾信息进行处理。

1.4 本文结构说明

本文的主要研究内容为基于机器视觉的垃圾分类系统和实时数据显示，开发出端到端的视觉目标检测模型，可以作为垃圾分拣机器人的前端视觉传感器，也可以用作单独的垃圾图像识别系统。同时使用硬件和通信技术，利用第三方平台实现数据实时显示。

第一章主要介绍了我国垃圾现状以及深度学习未来发展趋势，同时描述了本项目的总体目标。

第二章介绍了本文所用到的一些技术。大体上可分为硬件技术、软件技术。硬件技术主要包括本文使用到的各种硬件的使用原理、电路图信息、数据发送流程、控制过程等。软件技术主要包括两方面。一方面是深度学习所对应的有关基础理论，包括但不限于 tensorflow、Python 编程语言、OpenCV 库；另一方面是小程序开发以及腾讯第三方提供的服务器支持。

第三章是系统的需求分析和可行性分析。

第四章对模型进行实验。

第五章是垃圾分类系统的总体设计，阐述了系统的框架、功能等。

第六章是通过图片将系统运行的结果显示出来。

第七章主要是对系统进行测试。

第八章是对全文进行总结以及对自己不足的改进与期望。

2. 相关技术介绍

2.1 硬件技术

2.1.1 通讯模块

通讯模块采用的是 ESP8266[6]，该模块的功耗比较低。在编写硬件代码的过程中，可以设置该模块的工作模式。在本文中，采用的是 STA 模式。让该模块连接到互联网，实现信息的交互。

2.1.2 温湿度采集模块

温湿度采集模块采用的是数字温湿度传感器 DHT11，这是一种价格实惠且功耗低的温湿度传感器 [7]。该传感器

由一个电阻式感湿元件和一个 NTC 测温元件组成。它主要适合在本文中对温湿度和精确性要求较低的情况。

DHT11 的数据格式列表为：8bit 湿度整数数据+ 8bit 湿度小数数据+ 8bit 温度整数数据+ 8bit 湿度小数数据+ 8bit 校验和。

DHT11 通讯过程如图 2-1 所示。

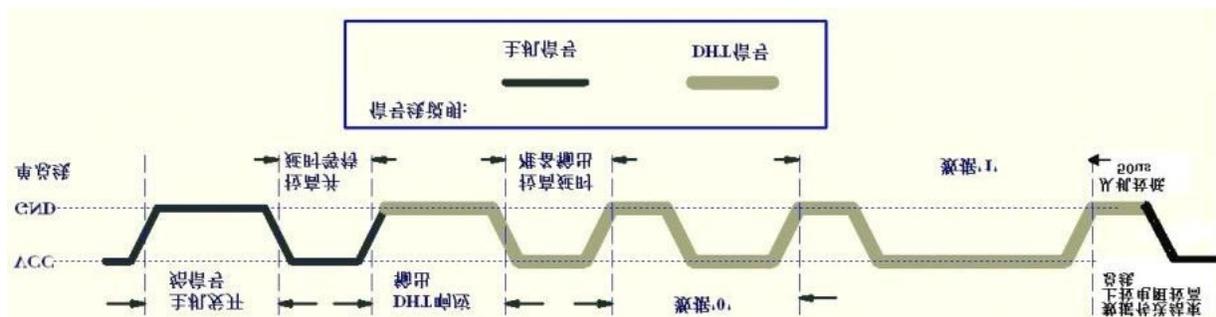


图 2-1 DHT11 通信过程

2.1.3 超声波模块

超声波模块采用的是 HC-SR04 型的超声波测距模块[8]。超声波模块由三个部分组成，包括了超声波发射器、接收器以及控制电路。如图 2-2 所示。



图 2-2 超声波模块图

2.1.4 舵机模块

舵机，又称伺服马达。可通过编写代码，命令舵机转动特定的角度，进而实现在本文中开启或关闭垃圾桶盖的功能。

2.2 软件技术

2.2.1 卷积神经网络

机器学习就是让计算机模拟人类的学习行为的一门科学。机器学习分为判别式模型和生成模型。

目前主要研究的卷积神经网络 (Convolutional Neural Networks, CNN) 在深度学习中模拟的是人的大脑的神

经元所组成的神经网络。卷积神经网络一般可以划分为：输入 (Input)、卷积层 (Convolutional layer,

Conv)、激活层 (Rectified linear unit, ReLU)、池化层 (Pool)、全连通层 (Fully connected layer, FC)

[10]。

首先需要了解卷积层的特点。滤波器 (Filter) 或者卷积核 (Kernel) 是指一个方阵。让方阵在输入图像的矩

阵上从左到右、从上到下进行移动，滤波器走过的区域就是感知区 (Receptive field)。在这个过程中，将滤波器和感知区进行运算就可以得到一个输出值，这个值也就是卷积值，而一次平移的操作就是一次卷积操作。不同的滤波器对不同的图形特征有不同的反应。比如下图就是两个分别对垂直特征和水平特征反应的滤波器。由此可知，一个好的滤波器可以检查输入图像是否有需要的特征。

传统的 CNN 中，卷积层之后常常伴有激活层和池化层，主要作用是提供更好的特征以及进行特征的压缩得到主要特征和连接所有特征得到输出值。卷积神经网络模型图如图 2-4 所示。

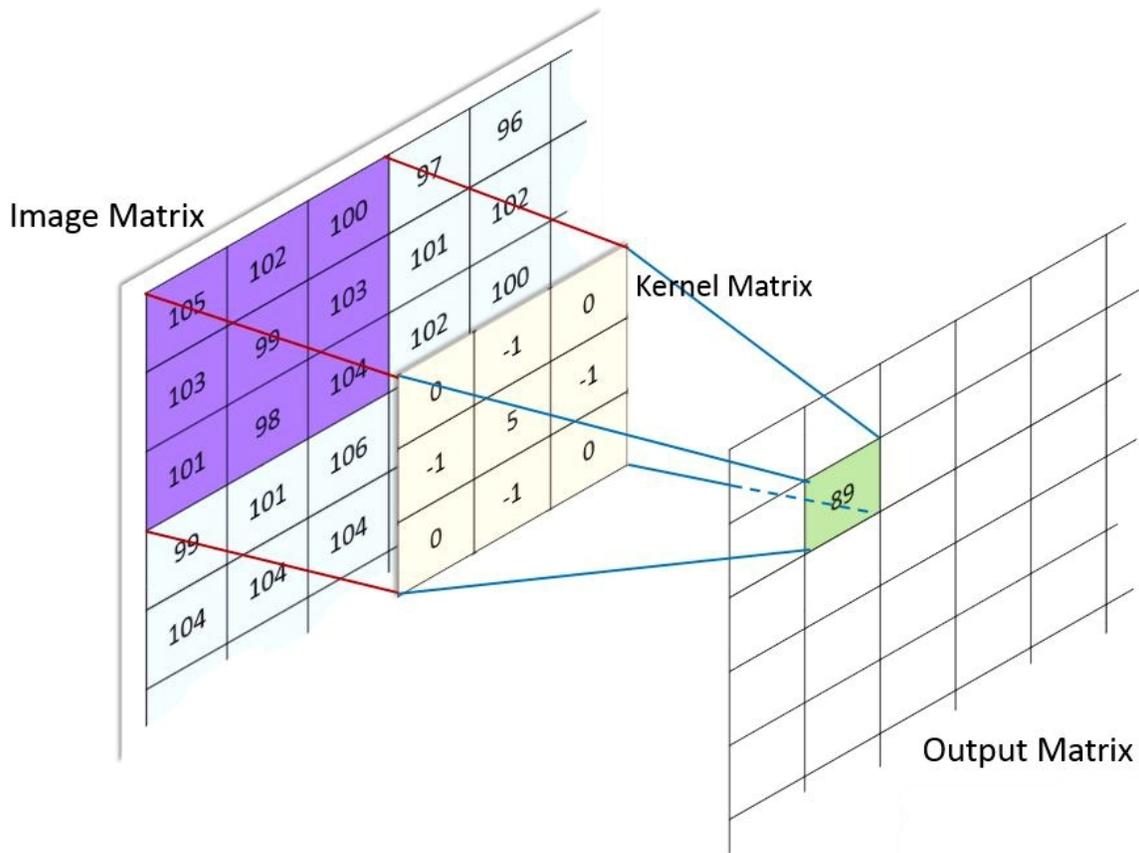


图 2-4 卷积神经网络模型

2.2.2 OpenCV 库

OpenCV 是一个第三方库，支持在多平台中使用。主要用于计算机视觉方面，同时提供多种编程语言的接口。

2.2.3 微信小程序

微信小程序在用户的操作上与现有的 app 的使用操作大致相同，但是微信小程序与 app 相比最大的好处就是不需要用户在使用的时候去下载软件，方便用户的使用。同时由于微信具有普遍性，微信小程序也有更开阔的市场。同时，腾讯官方提供开发者平台和详细的技术文档，使得开发微信小程序的门槛降低[12]。

2.2.4 巴法物联网云平台

巴法物联网云平台是基于一定的通信协议，保证数据在低带宽、不可靠的网络中传输有效的数据搭建起来的一个与物联网感知层交互的云平台。采用了订阅/发布的模式。在本系统中，微信小程序端订阅了 temp 的主题，于是就可以接收到这个主题里面的消息[13]。

3. 系统分析

3.1 需求分析

3.1.1 功能需求

(1) 硬件功能需求

本文的硬件功能包括：温湿度监测功能、舵机在一定条件下转动功能、超声波探测功能、信息传输功能。

系统功能列表如表 3-1 所示。

表 3-1 系统功能表

功能标题	模块名称	系统功能
温湿度监测功能	温湿度传感器 DHT11	检测垃圾桶内温度情况
舵机转动功能	舵机模块	自动/手动开启垃圾桶盖
超声波探测功能	超声波模块	探测垃圾桶内垃圾的存储情况
信息传输功能	ESP8266	用于传输采集到的信息

(2) 软件功能需求

本文的软件功能包括两个方面。一方面是智能识别出待检测垃圾的信息功能；另一部分是通过微信小程序将传感器收集到的信息在小程序端展示。

3.1.2 运行和开发环境

本系统所需的硬件开发环境包括一台最好拥有 GPU 处理器的笔记本电脑、一块 arduino 开发板、ESP8266 Wifi 模块、舵机模块若干、超声波模块若干、温湿度传感器模块若干。

所需的软件包括 Arduino IDE、Pycharm、Anaconda、微信开发者工具，运行和开发环境的要求不高，都可以实现，运行和开发环境配置如表 3-2 所示。

表 3-2 开发环境配置表

类别	标准配置
计算机硬件	Nvidia 1050Ti, 内存 8G, i5 处理器
硬件	arduino 开发板*1、ESP8266 Wifi 模块、舵机模块若干、超声波模块若干、温湿度传感器模块若干
软件	Arduino IDE、Pycharm、Anaconda、微信开发者工具

3.2 可行性分析

3.2.1 技术可行性

在硬件方面，arduino、ESP8266 Wifi 模块使用起来都是比较方便，操作比较简单，功能也很丰富。温湿度传感器、超声波传感器、舵机模块都是经常被使用的模块，具有广泛性和准确性。

在软件方面，之前对硬件编码有一定的基础，操作起来不难。在微信小程序方面，借助微信官方提供的第三方介绍文档进行学习研究后，实践起来容易上手。至于深度学习上，我觉得是本系统里面最难也是最核心的一部分。通过深度学习从而实现垃圾分类在理论上是可行的。但是由于自己之前对其了解的不深，所以重新学习花了不少时间。借助各大学习网站，对 tensorflow、OpenCV 有了更加深刻的认识和更多的掌握。

3.2.2 经济可行性

本文的消费主要是花在购买硬件设备上，包括超声波模块、DHT11 温湿度传感器、舵机、ESP8266 Wifi 模块等

。但是这些硬件的价格都比较低。目前的话，最需要的是依赖一台拥有 GPU 处理器的电脑，因为该电脑的消费是比较大的。但是后期可以将电脑部署到云端或者购买运算能力强的服务器，实现多个地方共享一台服务器，从而大大的减轻经济负担。通过垃圾分类，将大大的提升垃圾处理的速度，解放劳动力，从而创造出更多价值。

4. 模型效果实验

4.1 数据集介绍及预处理

Cifar10 数据集共有 6 万张图片，总计 10 个类别。其中训练集 50000，测试集 10000。官方的 Cifar10 把训练集分为 5 个批次，每个批次中有包含全 10 个类别的共 10000 张图片，且各个类别的数目是不同的。本文在训练集每个批次中随机取出 10%，共 5000 张图片作为验证数据集。Cifar10 原始数据集大小为 32×32 ，对于具有一定深度的 VGG-16 模型，容易造成特征丢失，本文在加载数据时采用双线性插值法将数据集 resize 成 64×64 的大小，此外还对训练、验证、测试集都做了归一化数据预处理。处理后的数据集信息如表 4-1 所示：表 4-1 Cifar10 数据集划分

数据名称	数据类别	训练集	验证集	测试集	图像大小	存储方式
Cifar10	10	45000	5000	10000	64×64	RGB

4.2 模型对照实验

实验采用经典的 VGG-16 作为主干网络，考虑到图像较小和网络较深，在原始的 VGG-16 上做了一些简单的调整，将第五个 block 删除，第四个 block 合并为一次卷积操作。为减少参数量便于训练，第一个全连接层设为 500 个输出节点，删除第二个全连接层。将精简后的 VGG-16 作为基础网络。本文下面所提到的基础网络都是精简后的 VGG-16。

4.3 聚类分组归一化效果验证

首先对本文提出的新的归一化方法——聚类分组归一化进行效果实验[15-16]，将基础网络和批归一化作为对照 [14]，即在基础网络下依次进行无归一化、聚类归一化、批归一化模型训，并将结果进行对比。由于计算机的算力限制，本次对照实验进行了 19 次循环迭代，即 epoch=19，batch_size 选择为 32。在输出方面每 2 个 steps 输出一次训练的 accuracy 和 loss，每 1 个 epoch 输出一次平均精度、平均损失且进行一次验证集测试。实验结果如图 4-1 所示。

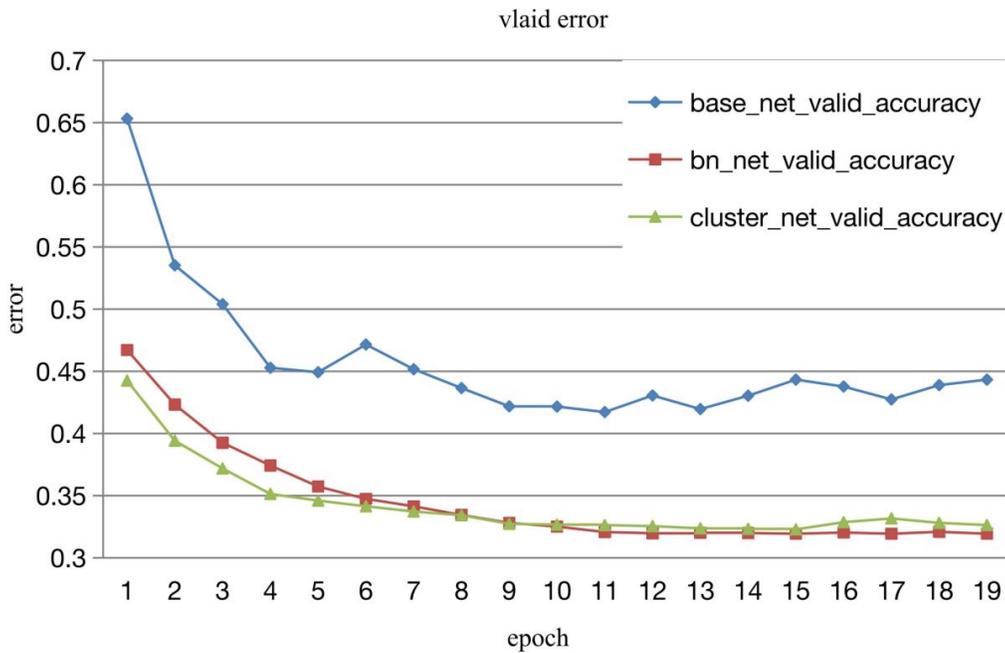


图 4-1 Cifar10 验证集错误率

图 4-3 是三种模型在每个 epoch 迭代后验证集上的错误率，由图可知与基础网络（base）相比，加入聚类归一化的网络模型（cluster）在第一个 epoch 结束后在验证集上就达到了 44.2%的错误率，此后的所有 epoch 都有着明显低于基础网络的错误率，实验证明聚类归一化确实能够起到加速模型训练, 提高模型准确率的作用。聚类归一化模型与批归一化模型相比在前 9 个 epoch 上有着更小的 error，即在前 9 次迭代聚类归一化模型要优于批归一化模型，能够使网络更快的收敛。

三种模型在训练数据集上每个 epoch 迭代的损失如表 4-2 所示：表 4-2 三种模型训练损失

1	1.9079	1.4492	1.4115
2	1.4686	1.1654	1.1030

Epoch	Base net train loss	Bn net train loss	Cluster net train loss
3	1.3056	1.0294	0.9702
4	1.2037	0.9339	0.8837
5	1.1341	0.8586	0.8173
6	1.0851	0.7967	0.7628
7	1.0391	0.7444	0.7172
8	1.0154	0.6940	0.6772
9	0.9664	0.6513	0.6382
10	0.9445	0.6094	0.5994
11	0.8955	0.5672	0.5656
12	0.8550	0.5329	0.5339

13	0.8255	0.4970	0.5014			
14	0.7785	0.4604	0.4710			
15	0.7554	0.4260	0.4389			
16	0.7217	0.3928	0.4083			
17	0.6978	0.3601	0.3794			
18	0.6745	0.3293	0.3517			
19	0.6534	0.2987	0.3284	VAR	0.0986	0.0946 0.0796

根据上表的训练损失值绘制折线图如图 4-2 所示。

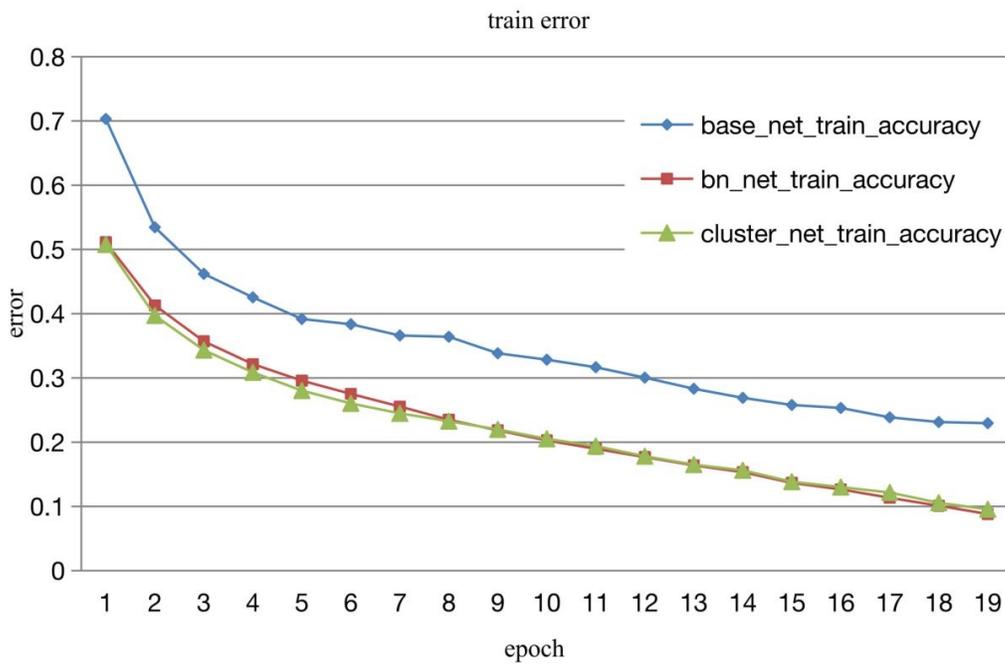


图 4-2 三种模型迭代损失

从图 4-2 中可以看出聚类归一化模型的损失与批归一化模型损失相比更加平滑（方差更小），因此对加入聚类归一化的模型可以考虑采用更大的学习率加快模型训练。

4.4 注意力机制效果验证

实验同样采用对照的方式进行，将基础网络与 Cbam 结合作为实验组，基础网络作为对照组，epoch=19, batch_size=32。输出方面的设置与聚类分组归一化效果验证的设置相同[17]。实验结果如图 4-3 所示。

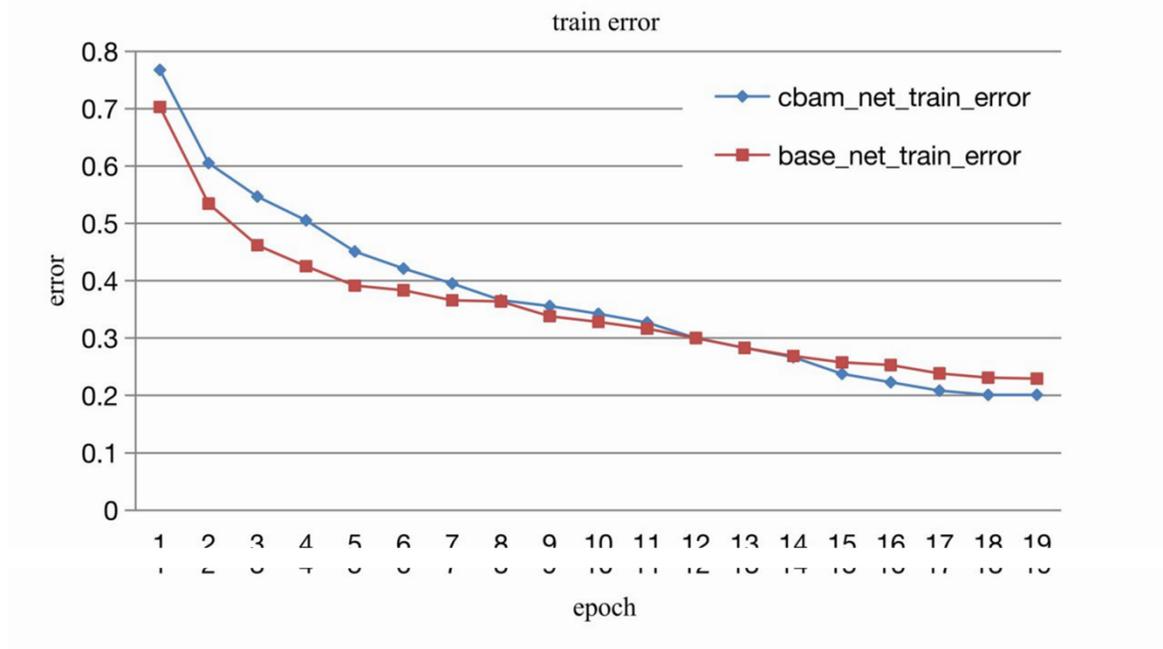


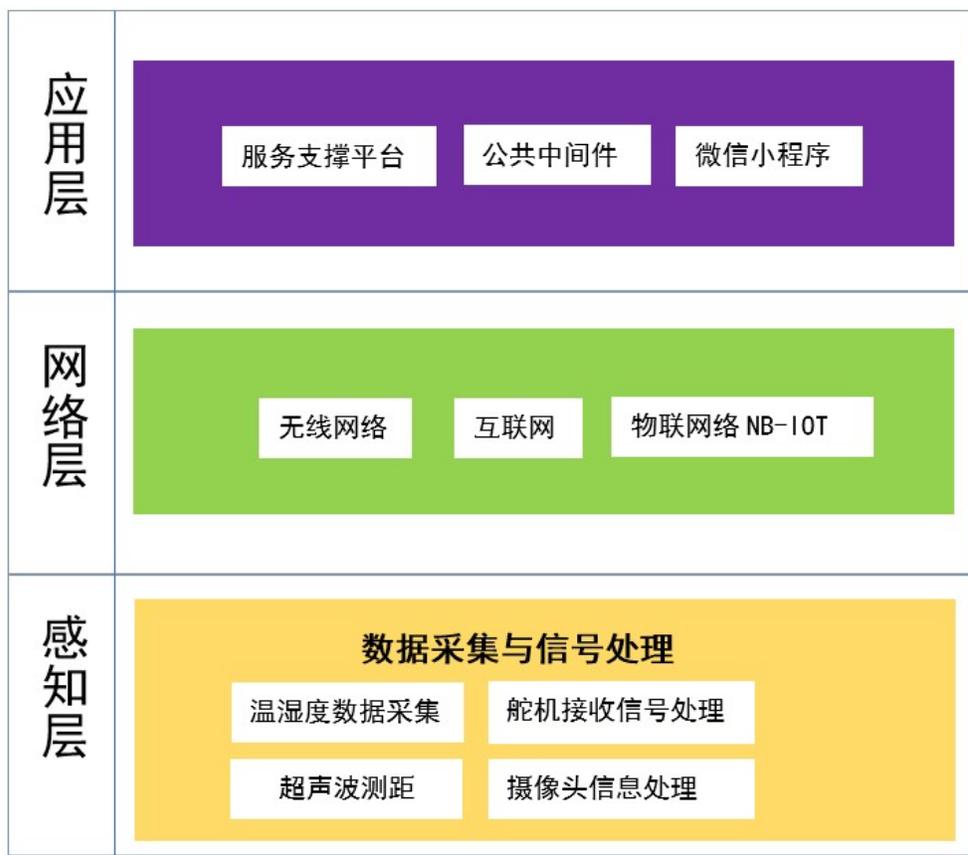
图 4-3 Cbam 训练误差

由图 4-3 可以看出网络模型加入注意力机制，在第 12 次迭代时，有着与基础网络相同的错误率，继续进行迭代，加入注意力机制的网络的错误率比原始网络要小，实验证明 Cbam 确实能起到提高网络训练精度的目的。

5. 系统设计

5.1 总体框架

本文采用的体系架构为物联网的三层架构，即感知层、网络层和应用层。本系统总体架构图如图 5-1

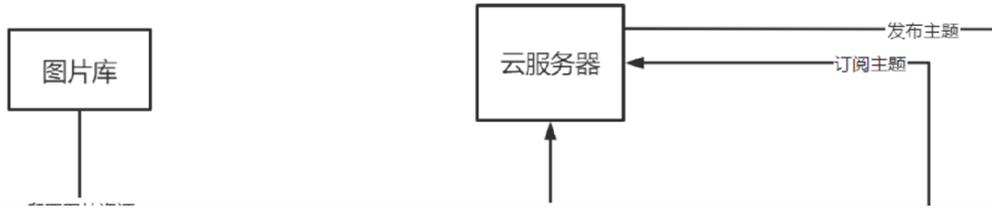


所示。

图 5-1 系统总计架构

在本系统中，学习大量垃圾图片的特征得到模型参数并进行垃圾检测。将深度学习运用到垃圾分类上，实现一个自动分类识别视频监控并标注 Python 系统。同时以 arduino 为 MCU，通过 ESP8266 Wifi 模块将信息发送到服务端，使用当下流行的小程序技术，构建一个垃圾信息实时查看系统，便于工作人员随时查看垃圾箱里的垃圾信息并且及时对这些垃圾信息进行处理。

在本系统中，先是通过爬虫获取相关的图片构建图片库，再将这个库通过 PC 端已有的神经网络进行训练。通过摄像头捕捉亟待识别的图片，与已有的图片进行比对，发出控制信号到单片机端。单片机对不同信号做出不同反馈。同时按照设定的时间频率将温湿度、距离的数据发送到云端。微信小程序端通过订阅云端的数据，将数据信息显示到小程序界面上。本系统信息流向图如图 5-2 所示。



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/867142044031006116>