

基于神经网络的垃圾分类

张嘉浩，代洋飞，朱博医

浙江大学

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

1 背景介绍

2 算法尝试与求解结果

3 比较分析与结论

4 参考文献

选题背景与文献综述

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

选题背景：垃圾分类，一般是指按一定规定或标准将垃圾分类储存、分类投放和分类搬运，从而转变成公共资源的一系列活动的总称。自今年 7 月 1 日起，上海市将正式实施《上海市生活垃圾管理条例》。垃圾分类，看似是微不足道的“小事”，实则关系到 13 亿多人生活环境的改善，理应大力提倡。

- **早期**，学者们只能借助经典的图像分类算法完成垃圾图像分类任务。
例如：吴健等利用颜色和纹理特征，初步完成了废物垃圾识别。
传统算法的鲁棒性较差，处理方式复杂，所需时间较长，无法达到实时的效果。
- **卷积神经网络 (Convolution Neural Network, CNN) 飞速发展：**
2012 年，AlexNet 取得了 ImageNet 图像分类竞赛的冠军，标志着深度学习的崛起。
- **随后几年**，GoogleNet、VGGNet、ResNet 等算法提升了图像分类的精度，并成功应用于人脸识别、车辆检测等多个领域。

选题背景与文献综述

基于神经网络的垃圾分类

张嘉浩，代洋洋，
朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

- **斯坦福大学的 Yang 等**建立了 TrashNet Dataset 公开数据集, 包含 6 个类别, 共计 2527 张图片.
- **Ozkaya 等**搭建神经网络 (本文称之为 TrashNet) 并进行参数微调, 数据集 TrashNet Dataset 上准确率:97.86%
目前这一数据集上最佳分类网络
- **非公开数据集方面, Mittal 等**自制了 2561 张的垃圾图片数据集 GINI, 使用 GarbNet 模型, 准确率:87.69%.
- **郑海龙等**用 SVM 方法进行了建筑垃圾分类方面的研究
- **向伟等**使用分类网络 CaffeNet, 调整参数, 适用于水面垃圾分类, 自制的 1500 张图片数据集上识别率: 95.75%.
- **2019 年**, 华为举办垃圾图像分类竞赛

概述

基于神经网络
的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

学习框架：主要利用深度学习框架 Pytorch 和 MindSpore。

对于垃圾分类问题，我们小组主要进行了**四轮模型的调试**，具体过程如下。

第一轮尝试：DNN

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

我们首先搭建了简单的一个三层全连接神经网络（DNN）

```
# flatten inputs
inputs = Input(shape=input_shape)
dnn = Flatten()(inputs)

# fully connected layer
dnn = Dense(6)(dnn)
dnn = BatchNormalization(axis=-1)(dnn)
dnn = Activation('sigmoid')(dnn)
dnn = Dropout(0.25)(dnn)
```

全连接层并不适合这么大的图像信息，因此即便我们对该神经网络进行长达一个小时的训练，测试集准确度很低，最高仅达到 **0.3** 左右。

第二轮尝试：CNN——网络结构

基于神经网络的
垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

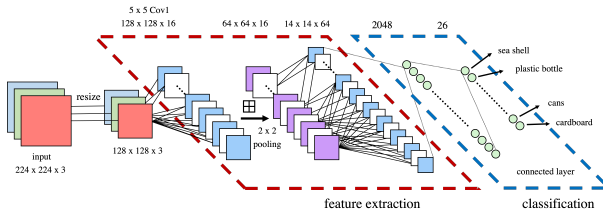
比较分析与结
论

参考文献

卷积神经网络通常被用于处理多阵列形式的数。该类网络的背后有四个关键思想，它们利用了自然信号的特性，分别是：**本地连接、共享权重、池化和多层的使用。**

该神经网络由数据输入层 (Input layer)，卷积计算层 (CONV layer)，ReLU 激励层 (ReLU layer)，池化层 (Pooling layer)，全连接层 (FC layer) 组成。

我们自己搭建了一个五层的卷积神经网络 (CNN)，其结构如下图所示：



图：卷积神经网络结构

第二轮尝试：CNN——网络结构

基于神经网络
的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

以第一层为例：

```
# conv1: Conv2d -> BN -> ReLU -> MaxPool
self.conv1 = nn.Sequential(
    nn.Conv2d(in_channels=3, out_channels=16, kernel_size=3,
              stride=1, padding=1),
    nn.Conv2d(in_channels=16, out_channels=16, kernel_size=3,
              stride=1, padding=1),
    nn.BatchNorm2d(16),
    nn.ReLU(),
    nn.MaxPool2d(kernel_size=2, stride=1),
)
```

第二轮尝试：CNN——图片变换

基于神经网络
的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

我们对图片进行了一系列变换，如旋转、翻转、灰度化以增强稳定度：

```
transform = transforms.Compose([
    #size transformation
    transforms.Resize((128,128)),
    #rotation
    transforms.RandomRotation((30,30)),
    #vertical flip
    transforms.RandomVerticalFlip(0.1),
    #Grayscale
    transforms.RandomGrayscale(0.1),
    #transform tensor
    transforms.ToTensor(),
    transforms.Normalize((0.5,0.5,0.5),(0.5,0.5,0.5))))
```

第二轮尝试：CNN——调整超参

基于神经网络
的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

在训练过程中主要调整以下参数

```
# #hyper parameter
batch_size = 32
num_epochs = 50
lr = 0.0001
num_classes = 25
image_size = 128 #### 64
```

第二轮尝试：CNN——训练结果

基于神经网络的
垃圾分类

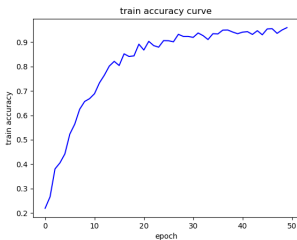
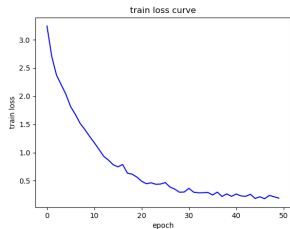
张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求
解结果

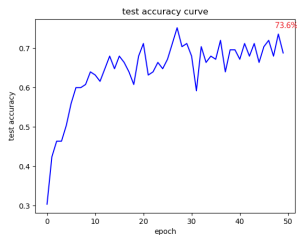
比较分析与结
论

参考文献



设置 $\text{epoch} = 50$ 进行训练. 虽然训练集损失值逐渐下降, 但是测试集表现不好, 正确率在 0.65 左右波动, 最后正确率在 70% 左右, 模型过拟合。这主要是由于训练集样本数量过少。

解决办法：数据增强



第三轮尝试： MobileNetV2

基于神经网络的垃圾分类

张嘉浩，代洋飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

此次尝试采用了 mo 平台提供的利用 MindSpore 搭建 MobileNetV2 网络模型的基础上进行参数修改。

网络简介：

在现实场景下，诸如移动设备、嵌入式设备、自动驾驶等等，计算能力会受到限制，MobileNet 由此提出。相较于传统网络，它有两个特点来提高运算速度：

- 使用深度可分离卷积 (depthwise separable convolutions) 替代传统卷积。
- 引入了两个收缩超参数 (shrinking hyperparameters)：宽度乘子 (widthmultiplier) 和分辨率乘子 (resolutionmultiplier)

第三轮尝试： MobileNetV2

基于神经网络的垃圾分类

张嘉浩，代洋洋，朱博医

背景介绍

算法尝试与求解结果

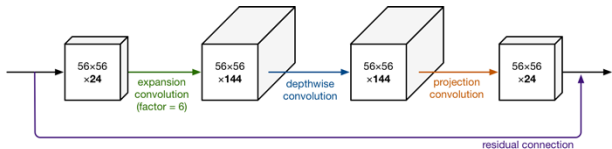
比较分析与结论

参考文献

MobileNetV2 主要引入了两个改动： Linear Bottleneck 和 Inverted Residual Blocks。

- Inverted Residual Blocks: 类似于沙漏和梭子形态的区别。
- Linear Bottlenecks: 对于低维度空间处理，论文中把 ReLU 近似为线性转换。

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	-



图： MobileNetV2 网络结构

第三轮尝试： MobileNetV2——数据准备

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

将脚本、预训练模型的 Checkpoint 和数据集组织为如下形式：

```
├─ main.ipynb # 入口Jupyter Notebook文件
|
├─ src_mindspore
|   ├─ dataset.py
|   ├─ mobilenetv2.py
|   └─ mobilenetv2-200_1067_gpu_cpu.ckpt
|
├─ results/mobilenetv2.mindir # 待生成的MindSpore0.5.0模型文件
|
├─ train_main.py # 将 main.ipynb Notebook 训练模型代码转化为py文件
|
└─ datasets/5fbdf571c06d3433df85ac65-momodel/garbage_26x100/ # 数据集
    ├─ train/
    ├─ val/
    └─ label.txt
```

图：文件组织形式

第三轮尝试： MobileNetV2——参数配置

配置后续训练、验证、推理用到的参数。可以调整以下超参以提高模型训练后的验证精度，如下所示：

```
# Train super parameters
config = EasyDict({
    # the dimension of the output layer
    "num_classes": 26,
    # mean, max, Head part of the pooling method
    "reduction": 'mean',
    "image_height": 224,
    "image_width": 224,
    # In view of the performance of the CPU
    "batch_size": 24,
    "eval_batch_size": 10,
    "epochs": 4,
    "lr_max": 0.01,
    "decay_type": 'constant',
    "momentum": 0.8,
    "weight_decay": 3.0,
    # ... Ingore other parameters
})
```

第三轮尝试： MobileNetV2——训练策略

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

一般情况下，模型训练时采用静态学习率，如 0.01。随着训练步数的增加，模型逐渐趋于收敛，对权重参数的更新幅度应该逐渐降低，以减小模型训练后期的抖动。所以，模型训练时可以采用动态下降的学习率，常见的学习率下降策略有：

- polynomial decay/square decay;
- cosine decay;
- exponential decay;
- stage decay.

这里实现 cosine decay 和 square decay 下降策略。我们对网络的调整主要是以下几个参数：

- “reduction”，部分池化方式
- “batch size”，批尺寸
- “lr rate”，最大学习率

第三轮尝试： MobileNetV2——训练结果

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

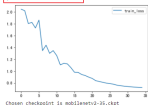
算法尝试与求解结果

比较分析与结论

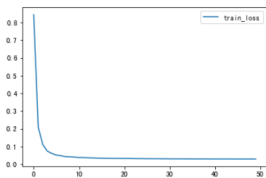
参考文献

前后两次训练结果对比。前者波动剧烈，后者较为平滑。改变的主要是：batch_size，训练 50 轮，最终得到准确率为 93.9%，损失曲线如下。

```
epoch: 26, time cost: 4.453917727813721, avg_loss: 0.7699189124397561
epoch: 27, time cost: 5.492337438876823, avg_loss: 0.7689247873373523
epoch: 28, time cost: 5.913162888188184, avg_loss: 0.7688935588888888
epoch: 29, time cost: 4.787590839842053, avg_loss: 0.75523089335504
epoch: 30, time cost: 4.27372084993745, avg_loss: 0.745558344535528
epoch: 31, time cost: 4.58389802482089, avg_loss: 0.748938444535528
epoch: 32, time cost: 4.38342648982344, avg_loss: 0.7358187433441743
epoch: 33, time cost: 4.758888888888889, avg_loss: 0.734830908259576
epoch: 34, time cost: 5.3598074775811, avg_loss: 0.726853246779882
[warn] [16] [38726:18876838618888,404890845] [02-11-18 21:02:33.282.182] [mindspore/
validating_processer.cc:12] be performed with dataset non-slice mode.
epoch: 35, time cost: 4.5337156898035648, avg_loss: 0.7242450812220674
[warn] [16] [38726:18876838618888,404890845] [02-11-18 21:02:33.282.182] [mindspore/
validating_processer.cc:12] be performed with dataset non-slice mode.
[acc': 0.9372896167799999, loss': 0.8682786779999999]
```



```
{'loss': 0.22210491988131315, 'acc': 0.9397435897435897}
```



第三轮尝试： MobileNetV2——预测结果

基于神经网络的垃圾分类

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

在所给的验证集中，大多数种类垃圾可以全部识别正确，个别垃圾种类会识别错误一张，预测的结果如下图所示：

```
加载模型路径: ./results/Ckpt_mobilenetv2/mobilenetv2-20.ckpt
./datasets/5fbd5f71c06d3433df85ac65-momodel/garbage_26x100/val/00_01/00040.jpg Hats
./datasets/5fbd5f71c06d3433df85ac65-momodel/garbage_26x100/val/00_01/00055.jpg Hats
./datasets/5fbd5f71c06d3433df85ac65-momodel/garbage_26x100/val/00_01/00037.jpg Hats
./datasets/5fbd5f71c06d3433df85ac65-momodel/garbage_26x100/val/00_01/00010.jpg Paint bucket
./datasets/5fbd5f71c06d3433df85ac65-momodel/garbage_26x100/val/00_01/00064.jpg Hats
```

图：预测结果：帽子

```

+加载模型路径: ./results/ckpt_mobilenetv2/mobilenetv2-20.ckpt
./datasets/5fdbf571c06d3433df85ac65-momodel/garbage_26x100/val/00_02/00055.jpg Newspaper
./datasets/5fdbf571c06d3433df85ac65-momodel/garbage_26x100/val/00_02/00094.jpg Newspaper
./datasets/5fdbf571c06d3433df85ac65-momodel/garbage_26x100/val/00_02/00007.jpg Newspaper
./datasets/5fdbf571c06d3433df85ac65-momodel/garbage_26x100/val/00_02/00059.jpg Newspaper
./datasets/5fdbf571c06d3433df85ac65-momodel/garbage_26x100/val/00_02/00074.jpg Newspaper

```

图：预测结果：报纸

第四轮尝试：Swim-Transformer

基于神经网络的
垃圾分类

张嘉浩，代洋
飞，朱博医

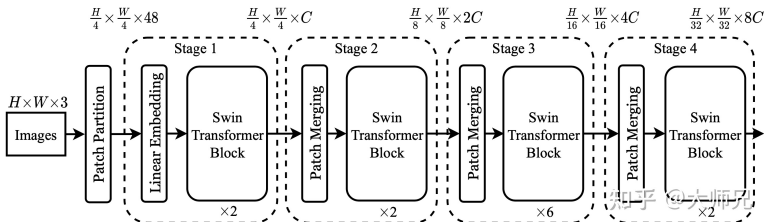
背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

Swin Transformer 是一种通用型的 Transformer 网络结构，在在分类，检测，分割任务上都取得了 state-of-the-art（最好）的效果。
它的网络结构如下所示：



图：Swin-transformer 网络结构

第四轮尝试：Swim-Transformer——训练成果

基于神经网络的垃圾分类

张嘉浩，代洋
飞，朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

我们将自己的训练集进行重新排列，以适用于 Swim-transformer 网络，经过了 80+ 次的训练，网络在训练集和验证集的准确率都接近或达到了 0.9 左右。我们对训练完的网络进行测试，在网上找了一张红帽子图片，进行预测，得到预测结果如下：



(a) pic1.

```
(base) PS C:\Users\David\Desktop>swim_transformer.pytrain -gpredict.py
class: Baseball prob: 0.00774
class: Basketball prob: 0.00774
class: Bicycle prob: 0.00774
class: Boat prob: 0.00774
class: Car prob: 0.00774
class: Cardboard prob: 0.00774
class: Ceramic Dood prob: 0.00774
class: Dirty Cloth prob: 0.00774
class: Disposable Chopsticks prob: 0.00774
class: Eggshell prob: 0.00774
class: Fluorescent Lamp prob: 0.00774
class: Glass Bottle prob: 0.00774
class: Glasses prob: 0.00774
class: Hats prob: 0.971
class: Lighter prob: 0.00774
class: Petalware prob: 0.00774
class: Newspaper prob: 0.00774
class: Old Mirror prob: 0.00774
class: Paper Bowl prob: 0.00774
class: Paint Bucket prob: 0.00774
class: Paper prob: 0.00774
class: Plastic Bottle prob: 0.00774
class: Seashell prob: 0.00774
class: Rubber Capsule prob: 0.00774
class: Toothbrush prob: 0.00774
class: Vegetable Leaf prob: 0.00774
(base) PS C:\Users\David\Desktop>
```

(b) pic2.

图：Swim-T 红帽子预测结果

比较分析与结论

基于神经网络
的垃圾分类

张嘉浩, 代洋
飞, 朱博医

背景介绍

算法尝试与求
解结果

比较分析与结
论

参考文献

在此次实验中我们实验的网络主要是 **CNN** 和 **MobileNet**。

- 之所以选择 CNN, 是因为卷积网络的精髓就是适合处理结构化数据, 而该数据在跨区域上依然有关联, 常被用于图像分析。
- 在本次实验中, 我们认为我们自己搭建的 CNN 的准确率不太理想的原因是因为网络模型还不是最优, 并且样本数量不足。
- 而在 MobileNet 网络中, 虽然最后准确率可以稳定在 93% 以上, 并且预测效果也算比较理想。
- 但是我们认为 MobileNet 的可调度很小, 我们也不能很清楚的了解到他其中的具体网络结构。

参考文献 I

基于神经网络的垃圾分类

张嘉浩, 代洋
飞, 朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

- [1] 吕思敏. 以史为鉴, 开启垃圾分类新时代, 城乡建设, 2020(3): 30-32.
- [2] Lowe DG, Distinctive image features from scale-invariant keypoints, in *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [3] Harri C, Stephens M, A combined corner and edge detector in *Proceedings of the 4th Alvey Vision Conference*, Manchester, UK. 1988. 207-217.
- [4] Zhang XK, Wang Y, Gou MR, et al, Efficient temporal sequence comparison and classification using gram matrix embeddings on a riemannian manifold in *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA. 2016. 4498-4507.
- [5] Ozkaya U, Seyfi L. Fine-tuning models comparisons on garbage classification for recyclability, *arXiv: 1908.04393*, 2019.
- [6] Mittal G, Yagnik KB, Garg M, et al. SpotGarbage: Smartphone app to detect garbage using deep learning, *Proceedings of 2016 ACM International Joint Conference*, Heidelberg, Germany. 2016. 940-945.
- [7] Kingma DP, Ba J. Adam. A method for stochastic optimization, *arXiv: 1412.6980*, 2017

参考文献 II

基于神经网络的垃圾分类

张嘉浩, 代洋
飞, 朱博医

背景介绍

算法尝试与求解结果

比较分析与结论

参考文献

- [8] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems*, Lake Tahoe, NV, USA. 2012. 1106–1114.
- [9] Harri C, Stephens M, A combined corner and edge detector, *Proceedings of the 4th Alvey Vision Conference*, Manchester, UK. 1988. 207–217.
- [10] Vapnik V. Statistical Learning Theory, *New York: Wiley*, 1998. 401–492.
- [11] 吴健, 陈豪, 方武. 基于计算机视觉的废物垃圾分析与识别研究, *信息技术与信息化*, 2016(10): 81-83
- [12] 向伟, 史晋芳, 刘桂华, 等. 改进 CaffeNet 模型在水面垃圾识别中的应用, *传感器与微系统*, 2019, 38(8): 150-152, 156.
- [13] Kingma DP, Ba J. Adam, A method for stochastic optimization, *arXiv: 1412.6980*, 2017.