**2020机器学习**

**印花疵点分类**

**组员：**

须子逸 20210240350

李帅

李锦程

朱柏霖

郭露

# 1 任务描述

基于机器学习模型对印花疵点数据集中标签为1、2、14的数据以及标签为1、2、4、14、20的数据分别做3分类及5分类任务；

利用ModelArts平台进行模型训练任务。

# 2 数据集介绍

# 3 数据预处理

3.1 裁剪

3.2 池化

3.3 差分

3.4 resize

# 4 模型介绍及实验

# 4.1 决策树

# 4.2 SVM

# 4.3 逻辑回归

# 4.4 简单CNN

# 4.5 ResNet-18

# 4.6 ResNet-50 on ModelArts

# 5 实验结果及分析

## 5.x 结果对比

三分类：

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 预处理方式 | | | | | | 分类模型 | | | | | | |
| 差分 | 最大池化 | 平均池化 | 缩放 | 填充 | 数据增强 | 决策树 | 逻辑回归 | SVM | CNN | ResNet-18 | ResNet-50 |
|  |  |  |  | √ |  |  |  |  |  |  |  |
| √ |  |  | √ |  |  |  |  |  |  |  |  |
| √ |  |  |  | √ |  |  |  |  |  |  |  |
| √ |  |  |  | √ | √ |  |  |  |  |  |  |
| √ | √ |  |  | √ | √ |  |  |  |  |  |  |
| √ |  | √ |  | √ | √ |  |  |  |  |  |  |

五分类：

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 预处理方式 | | | | | | 分类模型 | | | | | | |
| 差分 | 最大池化 | 平均池化 | 缩放 | 填充 | 数据增强 | 决策树 | 逻辑回归 | SVM | CNN | ResNet-18 | ResNet-50 |
|  |  |  |  | √ |  |  |  |  |  |  |  |
| √ |  |  | √ |  |  |  |  |  |  |  |  |
| √ |  |  |  | √ |  |  |  |  |  |  |  |
| √ |  |  |  | √ | √ |  |  |  |  |  |  |
| √ | √ |  |  | √ | √ |  |  |  |  |  |  |
| √ |  | √ |  | √ | √ |  |  |  |  |  |  |