**《图像分类系统需求分析说明书》**

**项 目 名 称：** 图片分类系统

**成 员 名 单： 王迪 马鲜艳 范广宝 陈晨**

**导 师： 张曙**

**工 程 领 域： 图像分类**

**研 究 方 向： 物体的分类与检测**

**中国科学技术大学软件学院**

**目录**

[一. 引言 3](#_Toc530949331)

[1.1 目的 3](#_Toc530949332)

[1.2 背景 3](#_Toc530949333)

[二. 需求概述 4](#_Toc530949334)

[2.1 目标 4](#_Toc530949335)

[2.2 研究内容 4](#_Toc530949336)

[2.3 用例分析 5](#_Toc530949337)

[三. 功能需求 8](#_Toc530949338)

[3.1 系统流程结构图 8](#_Toc530949339)

[3.2 功能模块图 8](#_Toc530949340)

[3.3 功能模块需求说明 9](#_Toc530949341)

[3.3.1 数据预处理模块 9](#_Toc530949342)

[3.3.2 图像分类及检测模块 9](#_Toc530949343)

[3.3.3 可视化界面展示模块 10](#_Toc530949344)

[四. 非功能需求 11](#_Toc530949345)

[4.1 性能需求 11](#_Toc530949346)

[4.2 运行环境需求 11](#_Toc530949347)

[4.3 可维护性 11](#_Toc530949348)

一. 引言

1.1 目的

本项目的主要目的是实现图像的分类，拟采用深度学习的算法对图片进行分类与检测，为了提高图像分类与检测的准确率，在mask Faster R-CNN和SSD等众多图像分类算法模型中，尝试多种优化算法，激活函数，通过对比各种算法模型的优缺点，最终得出识别准确率较高的模型。

经过建模、图像数据预处理、深度学习实现图像检测分类、算法优化等多个步骤，最终实现对测试图片结构的可视化展示。

1.2 背景

图像分类是作为图像视觉信息的主要表达方式，人们通过获取图片，经过筛选分类，最终可以实现快速的信息分类，提高了生产生活效率。目前，图像物体分类与检测已经成为计算机视觉领域的一个主要研究方向，它主要是一种利用计算机对图像进行处理、分析和理解，从而实现不同模式图像物体的分类。

随着大数据的发展，深度学习已经广泛应用于各个领域，2012年Hinton引入CNN解决ImageNet问题并取得巨大成功以来，ImageNet成为深度学习理论创新和技术突破的引擎，在计算机视觉领域掀起了一股研究热潮。这种基于深度学习的模型，其基本思想是通过有监督或者无监督的方式学习层次化的特征表达，来对物体进行从底层到高层的描述。

在大数据驱动的深度学习背景下，通过低层信号到高层特征的函数映射，来建立学习数据内部隐含关系的逻辑层次模型，以模仿人脑的视觉认知推理过程，从而使学习的特征具有更强的泛化能力和表达能力。因此基于深度模型架构通用性和统一性的视觉识别框架是当前计算机视觉领域的发展趋势，也是目前深度学习理论在视觉领域取得更深层次技术突破的一个切入点。二. 需求概述

2.1 目标

利用已有的对物体图像分类与检测的算法，进行分析后加以改进，提高算法效率，实现后进行可视化处理。使用开源框架TensorFlow进行图像的检测与分类，对于输入的图像可以进行分析，对图像中物体进行检测、分类及标识。利用SSD模型统一目标检测过程，SSD模型对原有模型进行相应的替换修改，能够做到实现比原模型更好的准确度与速度。

图像分类系统预计实现输入图像的检测分类与标识，并将结果以可视化形式展示。

2.2 研究内容

本项目的主要研究内容：

1. 图像数据集的预处理；

主要是在建模前对训练数据进行处理，本项目图像数据集的预处理主要包括对图像进行采集、图像格式进行转换、对采集的而图像进行标注、实心图像集的分类。

1. 模型的构建及训练；

本项目是基于TensorFlow框架，调用物体检测的多个模型，如SSD模型、R\_CNN模型、Faster R\_CNN模型，Mask R\_CNN模型等，将处理好的训练集数据放入模型中进行训练，训练过程中观察模型参数变化，待模型的损失函数收敛后，生成自己的模型文件。

1. 模型的对比及优化；

利用验证集对多个模型进行测试验证，通过调整模型中参数达到优化效果。

1. 物体分类及检测；

输入测试数据通过训练好的模型对输入的数据进行分类识别，最终时间图像的分类与检测。

1. 物体的形态识别；

在实现图像分类与检测的同时，可以对图像的体态进行识别，从而实现不同物体的体态识别。

1. 图像检测结果的可视化展示。

通过可视化界面实现用户测试图像的输入，最终将识别结果显示。

2.3 用例分析

图像分类系统主要面向个人用户，管理员用户作为系统后台人员参与系统维护。系统用例图如图2-1所示。

个人用户用例包括输入图片、查看图片识别结果。

管理员用户用例包括输入图片数据集、查看图片识别结果、对比不同模型识别结果、调用模型、训练模型、评估模型。

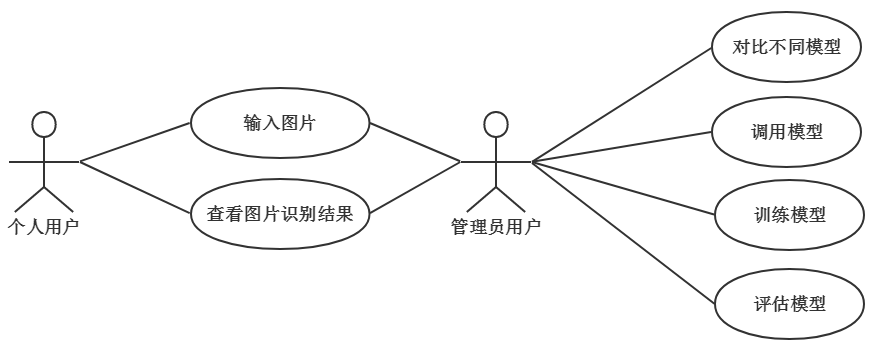


图 2-1 系统用例图

用例1：输入图片

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 1 | 名称 | 输入图片 |
| 简单描述 | 个人用户将需要测试的图片输入到系统中进行测试 | | |
| 参与者 | 个人用户 | | |
| 前置条件 | 个人用户已经获得测试图片数据 | | |
| 后置条件 | 个人用户将测试图片输入后，可以输出测试结果 | | |
| 基本流程 | 1. 个人用户获得可以测试的图片数据 2. 个人用户输入测试图片数据 3. 图片分类系统对图片进行分检测，最终输出测试分类结果。 | | |
| 特殊需求 | 用户输入的测试图片包含多个测试对象 | | |

用例2：查看图片识别结果

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 2 | 名称 | 查看图片识别结果 |
| 简单描述 | 管理员用户将测试的图片识别分类结果反馈给用户 | | |
| 参与者 | 管理员用户、个人用户 | | |
| 前置条件 | 训练数据集、建立训练模型，对测试图像进行训练 | | |
| 后置条件 | 将图像识别结果反馈给个人用户 | | |
| 基本流程 | 1. 管理员用户获取测试数据的识别结果 2. 管理员用户将测试识别结果反馈给个人用户 | | |
| 特殊需求 | 图像识别结果应该具有实时性 | | |

用例3：训练模型

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 3 | 名称 | 训练模型 |
| 简单描述 | 通过数据集对模型进行训练 | | |
| 参与者 | 管理员用户 | | |
| 前置条件 | 预处理完的图像数据集 | | |
| 后置条件 | 训练出不同的评估模型 | | |
| 基本流程 | 1. 管理员用户获取图像数据集 2. 对图像数据集进行预处理 3. 对于不同的训练模型分别对图像数据集进行训练 4. 生成不同的图像评估模型 | | |
| 特殊需求 | 尽量提高模型的识别效率 | | |

用例4：调用模型

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 4 | 名称 | 调用模型 |
| 简单描述 | 管理员用户调用已经建立的模型 | | |
| 参与者 | 管理员用户 | | |
| 前置条件 | 需要测试的图像数据 | | |
| 后置条件 | 训练出不同的测试评估模型 | | |
| 基本流程 | 1. 输入要训练的图像数据 2. 调用模型实现图像数据的训练 | | |
| 特殊需求 | 尽量提高模型的识别效率 | | |

用例5：对比不同的模型

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 5 | 名称 | 对比不同的模型 |
| 简单描述 | 对训练得到的不同的模型进行对比 | | |
| 参与者 | 管理员用户 | | |
| 前置条件 | 训练出不同的评估模型 | | |
| 后置条件 | 获得不同模型的识别准确率，选取准去率最高的评估模型 | | |
| 基本流程 | 1. 管理员用户对不同的评估模型进行测试 2. 输入测试图像数据 3. 对比不同的模型对图像识别分类的准确率 4. 选取准确率最高的评估模型 | | |
| 特殊需求 | 尽量提高模型的识别效率 | | |

用例6：评估模型

|  |  |  |  |
| --- | --- | --- | --- |
| ID | 6 | 名称 | 评估模型 |
| 简单描述 | 对选取的准确率最高的模型进行测试评估 | | |
| 参与者 | 管理员用户 | | |
| 前置条件 | 选取识别准确率最高的模型 | | |
| 后置条件 | 对选取的模型进行进一步评估 | | |
| 基本流程 | 1. 管理员获取准去率最高的模型 2. 输入图像测试数据 3. 获取最终的测试结构 | | |
| 特殊需求 | 尽量提高模型的识别效率 | | |

三. 功能需求

3.1 系统流程结构图

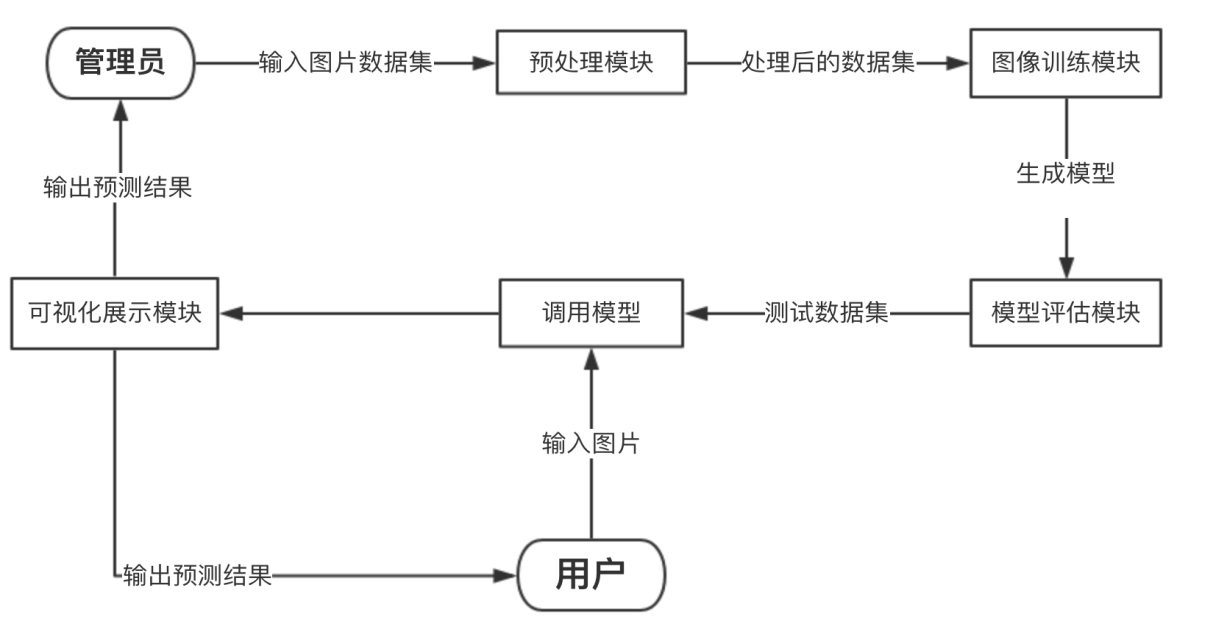


图3-1 系统流程结构图

3.2 功能模块图

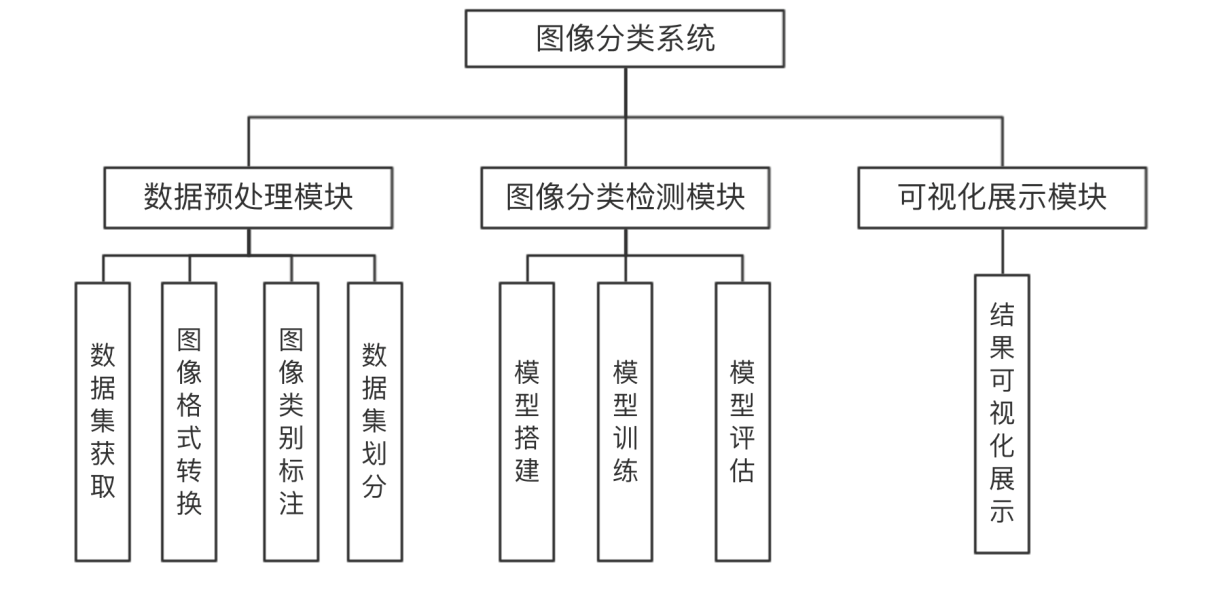


图3.2 系统功能模块图

3.3 功能模块需求说明

**3.3.1 数据预处理模块**

数据预处理模块主要实现输入数据的标准化。因为系统所处理图像有统一的格式要求，输入的图像应经过预处理，达到可以被系统处理的要求。主要功能有：数据集获取、图像格式转换、数据集划分及图像类别标注。

**（1）数据集获取**

从网络上获取合适的数据集，对数据集进行统一命名以便后续工作。

**（2）格式转换**

TensorFlow使用官方推荐的数据读取标准格式：tfrecord。tfrecord是一种同时存放图像数据和标签的二进制数据格式，有利于使用内存及数据的使用。将获得的数据集保存为tfrecord格式。

**（3）图像类别标注**

对图像进行分类标注。使用labelimage标注工具，用矩形框进行标注，标注出图像中物体的位置及名称。

**（4）数据集的划分**

模块调用TensorFlow中目标检测模型，使用训练集训练模型，使用验证集检验样本，计算均方误差进行各模型的比较，从中选出误差最小的模型为结果。最后用测试集数据对模型进行评估。

**3.3.2 图像分类及检测模块**

这个模块是分类系统的核心模块。主要通过调用谷歌开源TensorFlow框架中物体检测模型，通过训练集中的训练数据让模型进行学习；通过使用TensorBoard工具观察再利用使用验证集为检验样本，然后计算预测均方误差，比较各个模型的预测均方误差，选择预测均方误差最小的拟合模型为选择模型；最后利用测试集对模型进行测试并检验模型的泛化能力。

**（1）模型搭建**

配置好TensorFlow环境，利用pip安装好所需的包，为调用目标检测模型、使用训练集对被选中的模型进行训练做好准备。

**（2）模型训练**

将已被处理的数据集输入模型进行训练，训练结束后得到训练后的模型。

**（3）模型的优化**

利用验证集对模型进行检验，调整参数进行模型优化。

**（4）模型的评估**

使用测试集对模型进行测试，主要测试检测准确性、速度等。

**3.3.3 可视化界面展示模块**

该模块对结果进行可视化展示。

四. 非功能需求

4.1 性能需求

**（1） 响应时间**

图像分类系统应尽可能实现端到端的实时查询。用户输入图片，应该在短暂的时间内输出检测结果。

**（2）结果的准确性和合理性**

图像分类系统对用户输入的图像中物体的位置和分类识别的准确率应该至少与市面上系统准确率相当，这要求我们必须对模型的算法和参数进行优化。

**（3）灵活性**

系统要有好的接口，便于以后的扩充，以适应信息源的增加或者减少。

4.2 运行环境需求

**开发平台：**OSX 内存16G；Windows7 内存8G。

**语言：**Python

**开发工具：**Anaconda、Jupyter Notebook

**技术平台：**TensorFlow

**技术API：**Object-Detection API、Machine Learning API

4.3 可维护性

图像分类系统编码应遵循规范，便于维护。