**深 圳 大 学 实 验 报 告**

|  |
| --- |
| **课程名称 脑与认知科学**  **项目名称 图像中的注意区域预测**  **学 院 计算机与软件学院**  **专 业 软件工程（腾班）**  **指导教师 钟圣华**  **报 告 人 黄亮铭 学号 2022155028**  **实验时间 2024年11月06日至2024年11月26日**  **提交时间 2024年11月26日** |

**教务处制**

# 一、实验目的与要求

1. 通过实验加深对注意过程的了解；

2. 掌握Matlab或Python等常用工具的基本使用方法；

3. 具备使用机器学习方法设计基本的图像注意区域预测的能力；

# 二、实验内容与方法

1. 请简述你对图像的注意预测的定义、难点的理解。

**定义**：图像注意预测是计算机视觉领域地图像分析技术。图像注意预测旨在模拟人类视觉系统（人类视觉系统会自动地识别出感兴趣地区域，并对感兴趣的区域进行处理，忽略不感兴趣地区域）的行为，通过智能算法预测人类视觉在图像中最可能关注地区域（又称显著性区域）。这些区域通常与图像的特定特征密切相关。

**难点**：**1）**传统注意预测算法在预测效果的提高上有局限性。因为其特征提取和特征学习都以图像本身空间特征作为基础，缺乏多语义和上下文等深度特征，所以相较于人眼仍然具有较大差距，很难检测到人眼注视信息包含的大量高级语义信息。此外，不同人的注意力机制存在一定差异，在大部分传统模型中未加入先验信息，处理相对困难。**2）**对于深度学习注意预测算法，卷积神经网络中具有下采样的操作，会逐渐降低特征分辨率，并在此过程中丢失不同尺度包含的特征信息。而且，深度学习算法在神经网络的设计上针对不同任务需要不断变更，算法比较复杂。此外，深度学习模型存在可解释性差的缺点。

# 三、实验结论或体会

1. 请叙述一种注意预测的常用评价标准和计算方法。

AUC是一种常用的注意预测评价标准，用于评估预测结果与真实注视点之间的一致性。它通过计算受试者工作特性曲线（ROC曲线）的面积来衡量注意预测结果的优劣。AUC值介于0到1之间，值越接近1表示预测的显著性图与真实注视点的匹配程度越高。例如，如果接近0.5表示模型效果与随机猜测相当，如果接近1.0表示模型的预测结果完全符合真实注视点。

**计算方法**：

1. 输入参数要求：①预测显著性图：表示模型预测的显著性概率分布，通常是连续值矩阵；②真实注视点图：表示人类注视的实际位置，通常是二值图（1表示注视点，0表示非注视点）。
2. 从预测显著性图中提取真实注视点位置的显著值作为正样本，从非注视点区域中随机采样显著值作为负样本。
3. 通过通过阈值遍历，将显著性图的像素值分类为正类或负类，计算对应的TPR和FPR。计算公式如下：

其中，是正样本预测为正类数量；是负样本预测为正类数量；是正样本预测为负类数量；是负样本预测为负类数量。

1. 根据ROC曲线（一种对同一信号刺激的反应在几种不同的判定标准下所得的结果，用来度量真阳和假阳）积分，得到AUC值，公式为：。

# 四、练习题（每一题都请详细回答，在系统上传代码）

1．请选择一种基于图像的注意预测算法，使用该算法进行图像注意区域的预测，并使用注意预测的常用评价标准进行评价。在这里请给出方法简介、原图像、预测结果、评价结果。进行标准数据集上的显著性检测之后，请自选一张生活照（景物或者人都可以），并使用选择的注意预测算法进行显著性预测，在实验报告中附上原图和显著性预测图。

参考链接：**<http://saliency.mit.edu/downloads.html>**

[**https://saliency.tuebingen.ai**](https://saliency.tuebingen.ai)

4.1MSI-Net算法简介

MSI-Net又称多尺度信息网络，是一种基于深度学习的显著性检测模型了。通过引入多尺度信息融合机制，其结合局部细节和全局上下文特征，来精准预测图像中的显著性区域。其核心设计包括多尺度特征提取模块以及融合策略，前者用于同时捕获图像的不同尺度信息，后者通过逐步整合不同尺度的特征图生成最终的显著性预测图。MSI-Net具有高效的特征表达能力，能够适应复杂的场景和多样化的显著性检测任务，在视觉显著性预测、对象分割等领域表现优异。与最先进的方法相比，该网络基于轻量级图像分类主干，因此为计算资源有限的应用提供了一个合适的选择，以估计复杂自然场景中的人类注视。

MSI-Net的作者也在GitHub的源码中放上了如下的对比图，可以看到，其预测效果表现比较好。

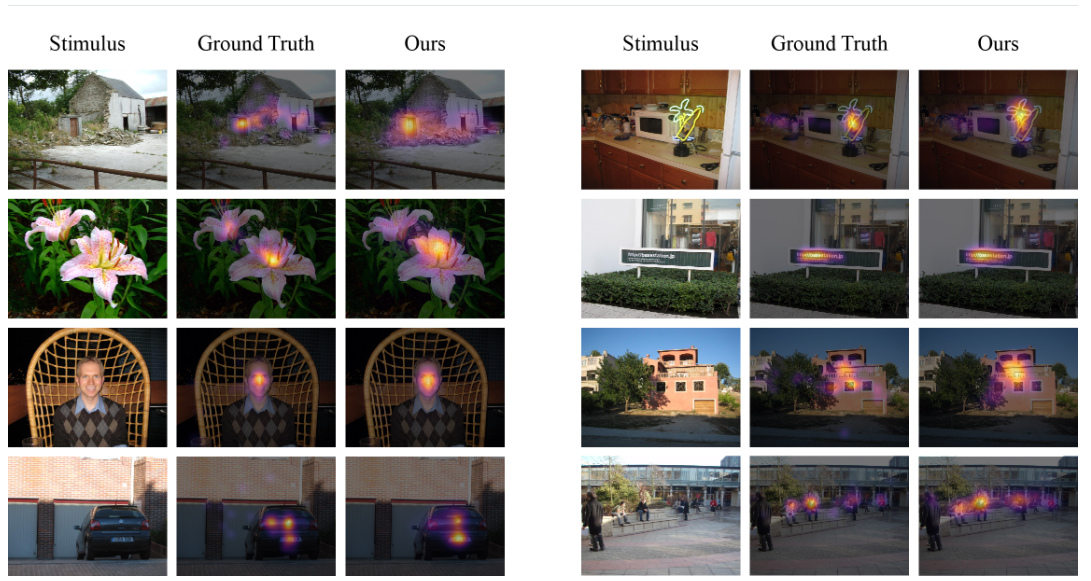


图1对比图

MSI-Net算法的主体架构为编码器-解码器架构，如图2所示。其架构包含一个具有不同扩张率的多个卷积层的模块，以平行捕捉多尺度特征。同时，我们将得到的特征信息与全局场景信息相结合，以准确预测视觉显著性。

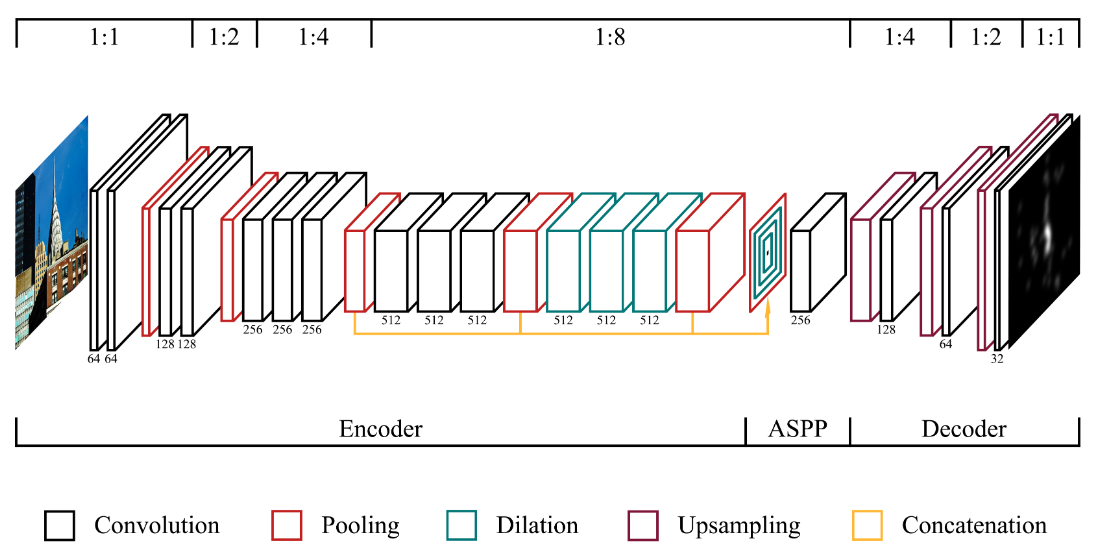


图2MSI-Net算法架构

4.2使用MSI-Net

MSI-Net的使用流程如下图所示。我将根据一下流程图分点介绍如何MSI-Net。

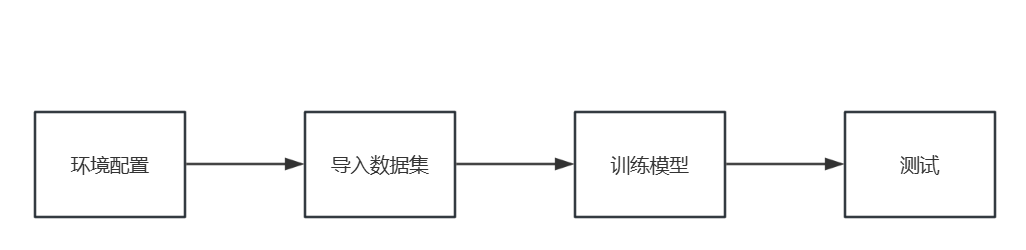


图3MSI-Net使用流程

1. **环境配置**：首先在GitHub上下载对应的模型源代码到本地，然后在命令行进入到对应的目录下输入命令创建环境。接下来进入该虚拟环境（），使用查看安装结果。观察下图，发现上述命令已经自动帮我们创建虚拟环境并安装模型所需要的包。



图3虚拟环境

但是，**此时输入命令直接运行模型不能成功**，会出现ImportError的错误。通过查阅资料发现是因为CUDA版本太新导致。但是，将CUDA降级会导致其他环境出现错误，因此只能寻求其他办法予以解决。进一步查询，发现可以使用TensorFlow的cpu版本进行训练。

综合上述分析，我的任务变更为下载cpu版本TensorFlow。

首先输入命令卸载gpu版本的TensorFlow，然后输入命令安装cpu版本TensorFlow。运行模型，出现函数不存在的错误，错误出现在h5py。通过查阅资料得知，这是因为conda自动安装的h5py的版本太高导致，因此输入命令和。再次运行模型，发现模型已经可以正确运行。

1. **导入数据集**：该模型支持多种数据集，如salicon、mit1003、cat2000和dutomron等多种数据集，默认数据集为salicon。本次实验我采用默认数据集salicon。模型导入数据集的方式可以是手动导入，也可以让模型在训练阶段自动导入。这里我选择让模型自动导入。
2. **训练模型**：在训练模型之前，我们需要在文件中修改字段为。在上述虚拟环境中输入命令即可开始训练，默认的训练轮数为10轮，该数据也可以在配置文件中进行修改。训练结束后，最佳的模型权重会保存在如下目录。
3. **测试**：使用3）中训练好的模型对自己的图片或者测试集的图片进行注意预测，命令为，其中对应的是用于训练时的数据集，对应的是需要测试的图片，最后得到的结果会存储在中。

在原始图像和预测结果中随机抽取两组，发现模型的预测效果表现比较好。

|  |  |
| --- | --- |
|  |  |
|  |  |

图4左：原始图像 右：显著性预测图像

**此外实验要求进行标准数据集上的显著性检测之后，请自选一张生活照（景物或者人都可以），并使用选择的注意预测算法进行显著性预测，在实验报告中附上原图和显著性预测图。**因此，我选择我在南京旅游景点夫子庙时拍摄的照片进行显著性预测。结果如下图所示，可以发现，该模型的预测效果还是比较好的。

|  |  |
| --- | --- |
|  |  |

图5左：原始图像 右：显著性预测图像

4.3评价MSI-Net

官方已经将评价体系集成为中的一个库。因此，我们可以使用命令安装对应的包，如下图所示。

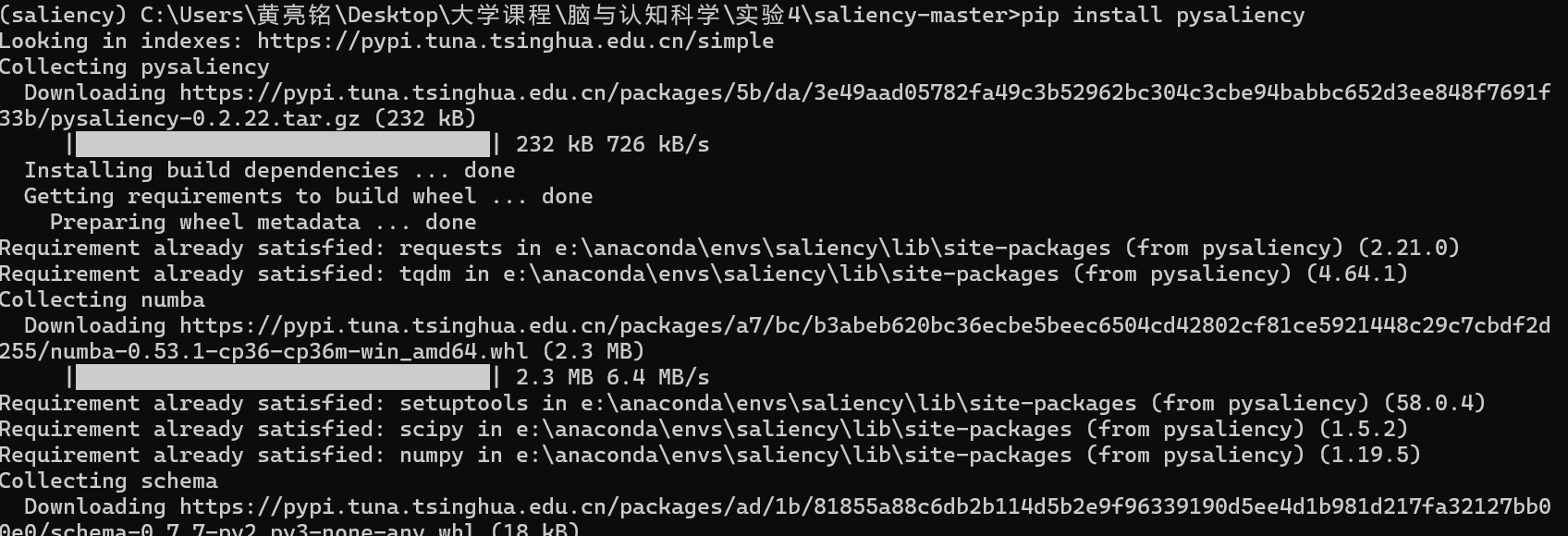


图6安装

然后根据官方文档进行代码的编写，如下图所示。

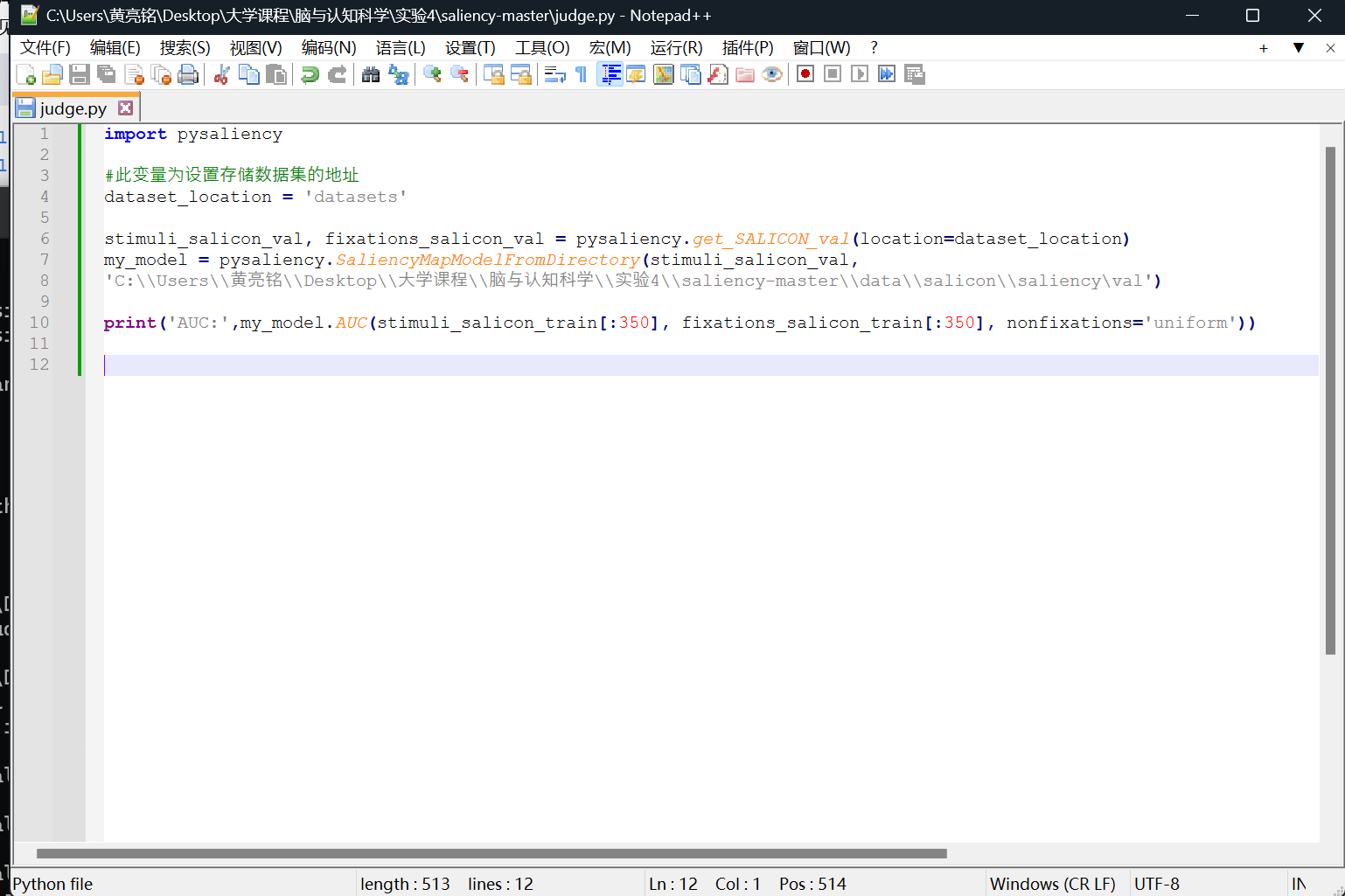
. 

图7评价代码

最终结果如下图所示。官网上的AUC的值为0.87，我们测试出来的值为0.82，两者相差不大，说明我们的模型的准确度是比较高的。

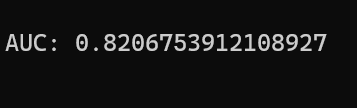


图8AUC值

4.4练习题总结

1. 通过本次实验，我了解了MSI-Net的原理和使用方法，并学习了使用MSI-Net对图片进行显著性预测的方法。
2. 在本次实验过程中，我在环境配置上遇到了较多的问题，而且问题的解决办法也比较繁琐，但是最终经过不懈努力将问题一一解决，模型得以正常地训练和测试。
3. 通过本次实验，我了解了如何使用这个官方库对模型进行评价，并且我实际使用了该库对自己训练的模型进行评价。

|  |
| --- |
| 指导教师批阅意见：  成绩评定：  指导教师签字：  年 月 日 |
| 备注： |

注：1、报告内的项目或内容设置，可根据实际情况加以调整和补充。