

作业 3 & 作业 4

本次作业的两个任务记两次作业成绩。提交一个 ipynb 文件，在 notebook 中请明确标注两个任务的对应部分。

数据集下载：<https://disk.pku.edu.cn/link/AADD15DECA910E48BF9F9F145A609DB664>

数据集描述和参考代码：<https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>

任务 1. 利用图卷积神经网络或 Vision Transformer，实现肿瘤的分类。

任务 2. 利用 UNET 或其他分割工具，实现图像中的肿瘤区域识别。

在每项任务中，均要求：

- 1) 简介数据，导入并清洗数据(3 分)
- 2) 简介模型，pytorch 代码实现（4 分）
- 3) 合理评估模型。对任务一，比较至少一种基准模型（如经典机器学习，MLP 等）（3 分）
- 4) 可以参考提供的资料，但要规范引用，不能抄袭。
- 5) 附加题（总评最多+2 分，由主讲教师评定）：文献（链接如下）指出了该数据集存在的错漏之处。请总结文献内容，更正错误，并比较修改前后模型的表现。

<https://www.sciencedirect.com/science/article/pii/S2352340923003669?via%3Dihub>