**毕业设计实验报告**

|  |
| --- |
| **实验内容** |
| **姓名：陈骋**  **毕业设计题目：基于深度学习的影像报告多模态检索**  **简介：作为医生诊断患者疾病的一个重要工具，影像报告中往往包含着影像、病史、影像学表现、影像学诊断等图像、文字、标签等多个模态的数据。通过深度学习实现影像报告多模态特征之间的跨模态检索，不仅可以帮助医生查阅类似案例，还可以为医生提供相应的病灶表现和诊断信息，帮助他们更加快速、更加准确的进行病理诊断。** |
| **实验记录** |
| **第一阶段的实验主要围绕毕业设计的前期准备**  **首先，通过开源医疗数据库open-i提供的接口，抓取了来自MedPix的CT、X光、核磁共振三种影像报告，并下载了印第安纳大学医院提供的Chest X-ray数据集。**  **其次，通过查阅相关领域文件初步了解了多模态检索的概念，并初步了解了可用于多模态检索的各种方法，例如生成对抗网络、受限制玻尔兹曼机、对应自编码器等**  **下一阶段将对数据进行进一步的处理，以满足后续实验对数据质量的阅读。同时学习并使用不同的多模态检索模型进行初步的实验，加深对各种模型的认识，同时选取更加适合此次毕业设计的模型。**  **时 间： 2020年 12月 31日** |
| **实验记录** |
| **第二阶段的实验主要围绕着多模态检索模型的学习与构建**  **通过相关文献的阅读，了解了几种用于多模态检索的深度学习网络，例如对应自编码器、跨模态生成对抗网络、对应首先玻尔兹曼机等。第二阶段的实验中，实现了自定义的对应自编码器。并在MNIST数据集上进行了实验并取得了较好成绩。**  **其中影像报告文本数据通过借助NLTK库进行分词、去掉停用词、提取词干等预处理，并使用Doc2Vec模型实现了影像报告文本的向量化。影像报告中的图像信息则预先通过图像增强处理后再读入内存中。**  **现阶段已经使用GPU服务器训练50轮得到模型，但通过输出的模型损失观察到模型的收敛速度较慢，且损失较高。因此下一阶段将围绕着模型的优化进行，将尝试引入其他的神经网络对图像数据和文本数据进行特征提取，构建更为优秀的深度学习网络。同时下一阶段也将开始搭建多模态检索的演示平台。**  **时 间： 2021年 1月 31日** |

|  |
| --- |
| **实验记录** |
| **第三阶段的实验主要围绕着多模态检索模型的优化与扩展，同时搭建了检索的图形界面。**  **第二阶段的实验构建了可以应用于多模态检索的对应自编码器，但通过在MNIST数据集上测试发现模型的准确率较差，并且在实验数据集上出现了数据间没有区分度的情况，且模型的训练损失下降较慢。**  **因此第三阶段的实验主要围绕模型参数的优化，在实验过程中尝试对模型的结构、训练轮次、batch\_size等模型参数进行调整，并尝试通过其他深度学习模型对图像数据进行特征提取，同时构建了对应全模态自编码器。**  **通过在MNIST数据集上实验发现现阶段对图像数据的特征提取效果不够理想，而对应全模态自编码器的效果较对应自编码器提升不大。**  **因此接下来的实验将全力围绕着各个模型的优化。**  **时 间： 2021年 2月 28日** |

|  |
| --- |
| **实验记录** |
| **最后阶段的实验搭建了对应卷积自编码器以及扩展的对应自编码器。**  **同时对每一个模型进行了Top-20%准确率的计算，通过比较不同模型之间的检索准确率分析和探究各模型的优缺点。**  **最后还实现了可视化本地检索系统的搭建。**  **时 间： 2021年 4月 28日** |