

FCCM: 基于非小细胞肺癌CT图像的医疗诊断与病情分析

组员：张弛 517021910658 程铭 517021910750

背景介绍

随着生物医学技术的发展，人们对各种复杂病情有了更为深入的了解与认识。随着CT(Computed Tomography电子计算机断层扫描)等技术的出现，医生能够较为直接地研究部分病灶的物理特征。但是通过CT图像对病人的病情进行分析，不仅需要医生极其丰富的行医经验，也需要消耗大量的时间与精力。

另一方面，随着计算机科学技术的发展，结合计算机技术辅助病情诊断与分析，得到了广泛的关注。近年来，深度卷积神经网络的提出，使计算机能够在图像中提取出丰富的特征信息，在物体分类任务、模式识别任务等等领域都表现出极大的潜力。结合深度卷积神经网络处理CT图等医疗图像，用于病灶的定位与分析，并对病情的发展进行预测，是目前生物医学的热门研究方向。

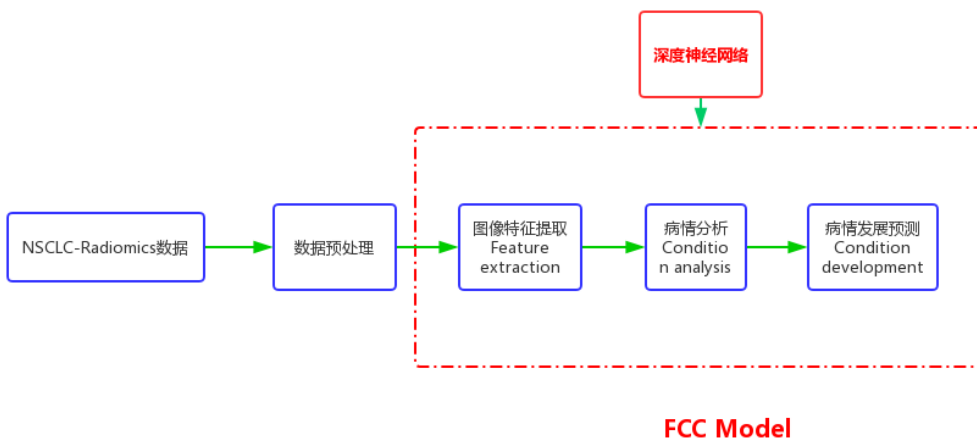
关键词：医疗图像诊断，深度卷积网络，病灶定位与分类，病情发展预测

研究内容及方向

研究内容（FCCM的建立）

我们小组的研究方向是，利用肺癌病人的CT图像与医生对于病人的病情分析等信息，建立“*图像特征提取 (Feature extraction) ——病情分析 (Condition analysis) ——未来发展预测 (Condition development)*”的医疗模型（FCC Model），能够实现通过向模型输入病人CT图像数据，从而对病人的病情进行准确评价，并对病情未来发展进行预测。

FCCM模型框图如下：



数据获取（NSCLC-Radiomics）

The Cancer Imaging Archive (TCIA)是癌症研究的医学图像的开放获取数据库。其中有种类丰富的各类癌症疾病的数据集，部分数据标注了患者结果，治疗细节，基因组学，病理学和专家分析等丰富信息。

我们采用了其中于近期更新(2019年11月8日更新)的非小细胞肺癌患者CT图像数据集，名为NSCLC-Radiomics，可以确定该数据集未被用于机器学习领域的模型训练(更新时间较近，距今仅一个月不到，且该网站会列出引用某数据集的论文列表，该数据集未被引用过)。

预期目标

我们的神经网络模型能够直接处理病人肺部的CT图像，提取其中的细节特征，对病人的患病种类（Histology）与病情程度（OverallStage）进行准确分类，同时对病人的未来预期存活时间（SurvivalTime）进行正确预测。

FCCM共包括两个子任务：预测和目标分类。

预测任务

对于指标“存活时间”而言，算法通过读入输入图像数据并提取相关特征，预测当所患病为当前情况时，该病人的存活时间。将预测值与该病人实际的存活时间相比较，便可评估算法精度。通过对算法的不断改进及优化，使得算法预测值与实际值较为接近。

分类任务

对于指标“患病状态”、“患病种类”而言，算法通过读入输入图像并提取相关特征，得出在患病为当前情况时，该病人的患病状态（I类、II类、III类）以及患病种类（肺腺癌、肺鳞状细胞癌、肺大细胞癌、非小细胞肺癌、不详）。将算法的输出类别与实际类别相比较，便可评估模型精确度。通过对算法及模型的改进和优化，使得算法的输出类别与真实类别相符。

多模型比较分析与新算法的提出

为更好地完成整个任务，我们决定同时建立多种模型、采用多种优化算法对图像数据进行处理和特征提取，并加以比较分析。通过对比分析各算法的性能优劣，得出相关结论。与此同时，我们通过模型输入输出之间的关系，可以从中得出医学领域（肺癌）相关知识结论，对肺癌这一病症进行更为细致的分析，亦有助于协助医生工作。

另外，为更好地提取图像特征并完成预测及分类任务，我们将使用深度卷积网络来学习图像的特征。考虑到医学图像与普通场景图像的不同，我们将整合、改进现有算法，提出更适合对医学图像（CT图）处理的全新算法，从而提高预测精度和分类精度。