

分类号 TP391.1  
U D C

密级  
编号 10741

兰州财经大学

LANZHOU UNIVERSITY OF FINANCE AND ECONOMICS

## 硕士学位论文

论文题目 基于深度学习的非小细胞肺癌淋巴结转移  
辅助诊断研究

研究生姓名: 李彤

指导教师姓名、职称: 何江萍 副教授

学科、专业名称: 管理科学与工程

研究方向: 信息管理与信息系统

提交日期: 2022年5月29日

## 独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 李彤 签字日期： 2022.5.29

导师签名： 何以萍 签字日期： 2022.5.29

## 关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定， 同意（选择“同意”/“不同意”）以下事项：

1. 学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2. 学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入 CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名： 李彤 签字日期： 2022.5.29

导师签名： 何以萍 签字日期： 2022.5.29

# Deep Learning-Based Computer-Aided Diagnosis of Lymph Node Metastasis in Non-Small Cell Lung Cancer

**Candidate: Tong Li**

**Supervisor: Jiangping He**

## 摘要

肺癌作为对生命威胁最大的癌症之一，发病率和死亡率一直以来都居高不下。而非小细胞肺癌占全部肺癌患病人数的 85%。一般来说一旦发生了肿瘤转移，患者的生存期和预后也会变得更差。通过计算机辅助诊断是否发生淋巴结转移，可以以非侵入的方法计算出淋巴结发生转移的几率，进一步地给放射科医师提供必要信息从而提高诊断准确性，帮助改善病人存活期和预后。正电子发射断层扫描-计算机断层扫描（PET-CT）图像被认为是评估、分期、诊断肺癌的首选成像方式。影像组学是指从医学图像中提取和分析高维可挖掘数据，是癌症风险评估的重要预后工具。因此，本文从 PET-CT 图像中获取影像组学的信息，并用深度学习方法对非小细胞肺癌患者是否发生淋巴结转移做出了辅助诊断。主要研究如下：

（1）基于深度学习的非小细胞肺癌淋巴结转移辅助诊断。本研究提出了一种端到端的深度学习体系结构，融合传统影像组学特征、深度学习特征、临床特征以及医生诊断进行淋巴结转移辅助诊断。首先从患者的 PET 图像中提取传统的影像组学特征，再用曼惠特尼 U 检验将存在显著差异的特征剔除；其次，分别利用三维卷积神经网络和二维卷积神经网络提取三维原发肿瘤和二维正面半身最大强度投影的深度学习特征；最后将传统影像组学特征和深度学习特征进行融合，再加入临床特征和医生诊断结果，对淋巴结是否发生转移做出诊断。本文在 121 例非小细胞肺癌患者的 PET-CT 影像和临床数据上进行研究。在测试集上该模型达到了 0.82 的 AUC 和 0.86 的精确度，明显优于医生诊断（AUC:0.61，精确度:0.80），具有更好的性能，证实了融合多种特征能以非侵入的方式辅助诊断是否发生非小细胞肺癌淋巴结转移。

（2）两阶段的非小细胞肺癌淋巴结转移辅助诊断。本研究提出了一种两阶段的深度学习架构，通过分割任务使网络关注感兴趣区域，自动学习分类特征，消除测试阶段模型对医生勾画的金标准的依赖。首先在第一阶段使用一个分割网络对正面半身最大强度投影图像进行肿瘤分割，大致定位肿瘤区域，作为第二阶段三维肿瘤特征提取时的测试集处理依据；其次在第二阶段中利用三维分割-分类模型和二维分割-分类模型对三维肿瘤和二维正面半身最大强度投影提取深度学习特征；最后将两种深度学习特征融合，进行淋巴结转移辅助诊断。在测试集上该模型达到了 0.74 的 AUC 和 0.84 的精确度，证实了在测试集中不使用医生勾画的肿瘤金标准也能达到不错的性能。

**关键字：**非小细胞肺癌 PET-CT 图像 深度学习 淋巴结转移 辅助诊断

## Abstract

As one of the most threatening cancers to life, lung cancer has always had high morbidity and mortality. Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancers. In general, once tumor metastasis has occurred, the survival and prognosis of patients will also become worse. Computer-aided diagnosis of whether lymph node metastasis (LNM) occurs, the probability of LNM can be calculated in a non-invasive way, further providing radiologists with the necessary information to improve diagnostic accuracy and help improve patient survival and prognosis. Positron Emission Tomography -Computed Tomography (PET-CT) imaging is considered the imaging method of choice for assessing, staging, and diagnosing lung cancer. Radiomics refers to the extraction and analysis of high-dimensional mineable data from medical images, which is an important prognostic tool for cancer risk assessment. Therefore, in this paper, the radiomics features are extracted from PET-CT images and deep learning methods are used to assist in the diagnosis of whether LNM occur in patients with NSCLC. The main studies are as follows:

(1) Computer-aided diagnosis of LNM of NSCLC based on deep learning. In this study, an end-to-end deep learning architecture is proposed, which integrates traditional imaging radiomics features, deep learning features, clinical features, and physician diagnosis for LNM computer-aided diagnosis. Firstly, the traditional radiomics features of the patient are extracted from the PET images of the patient, and then the features with significant differences in the image

scanner are screened out by the Mann Whitney U test; secondly, the 3D convolutional neural network (CNN) and the 2D CNN are used to extract the deep learning features of the 3D primary tumor and the maximum intensity projection (MIP) of the 2D frontal bust, respectively; finally, the traditional radiomics features and deep learning features are fused, and then the clinical features and doctor's diagnosis results are added to diagnose whether the lymph nodes have metastasized. This paper conducts PET-CT imaging and clinical data from 121 patients with NSCLC. On the test set, the model achieved an accuracy of 0.82 AUC and 0.86, which was significantly better than the doctor's diagnosis (AUC: 0.61, precision: 0.80), and confirmed that the fusion of multiple features can assist in the diagnosis of NSCLC lymph node metastasis in a non-invasive manner.

(2) Two-stage computer-aided diagnosis of LNM in NSCLC. This study proposes a two-stage deep learning architecture that makes the network pay attention to the region of interest by splitting the task, automatically learning classification characteristics, and eliminating the dependence of the test phase model on the gold standard outlined by the doctor. Firstly, in the first stage, a segmentation network is used to segment the tumor of the frontal bust MIP image, roughly localizing the tumor area, which is used as the basis for the test set processing in the second stage of 3D tumor feature extraction; secondly, in the second stage, the 3D segmentation-classification model and the 2D segmentation-classification model are used to extract deep learning

characteristics from the 3D tumor and the 2D frontal bust MIP ; finally, the two deep learning features are fused to assist in the diagnosis of LNM. The model achieved an AUC of 0.74 and an accuracy of 0.84 on the test set, confirming that good performance could be achieved without the gold standard of tumors outlined by the doctor in the test set.

**Keywords:** Non-small cell lung cancer; PET-CT images; Deep learning; Lymph node metastasis; Computer-aided diagnosis

# 目 录

<b>1 绪 论</b> .....	<b>1</b>
1.1 研究背景与意义 .....	2
1.1.1 非小细胞肺癌 .....	2
1.1.2 PET-CT 介绍 .....	3
1.1.3 非小细胞肺癌淋巴结转移辅助诊断的临床价值 .....	5
1.2 研究现状 .....	6
1.2.1 传统的影像组学方法 .....	6
1.2.2 基于深度学习的方法 .....	8
1.3 主要研究内容及创新 .....	10
1.4 本文组织结构 .....	11
<b>2 相关理论</b> .....	<b>13</b>
2.1 卷积神经网络 .....	13
2.2 分类问题评价指标 .....	14
2.3 本章小结 .....	15
<b>3 实验数据</b> .....	<b>16</b>
3.1 数据来源与选择 .....	16
3.2 临床特征处理 .....	17
3.3 本章小结 .....	17
<b>4 非小细胞肺癌淋巴结转移辅助诊断</b> .....	<b>18</b>
4.1 方法 .....	18
4.1.1 研究流程 .....	18
4.1.2 传统影像组学特征提取 .....	19
4.1.3 深度学习数据预处理 .....	20
4.1.4 深度学习特征提取 .....	21
4.1.5 特征融合与分类 .....	22
4.2 实验结果与分析 .....	22
4.2.1 训练过程 .....	22

4.2.2 实验环境.....	23
4.2.3 淋巴结转移辅助诊断结果.....	23
4.2.4 讨论.....	27
4.3 本章小结.....	28
<b>5 两阶段非小细胞肺癌淋巴结转移辅助诊断.....</b>	<b>29</b>
5.1 方法.....	29
5.1.1 研究流程.....	29
5.1.2 肿瘤分割.....	30
5.1.3 以分割任务为引导的辅助诊断.....	32
5.2 实验结果与分析.....	35
5.2.1 训练过程.....	35
5.2.2 实验环境.....	36
5.2.3 两阶段淋巴结转移辅助诊断结果.....	36
5.2.4 讨论.....	40
5.3 本章小结.....	40
<b>6 总结与展望.....</b>	<b>41</b>
6.1 总结.....	41
6.2 展望.....	42
<b>参考文献.....</b>	<b>44</b>
<b>致 谢.....</b>	<b>50</b>

# 1 绪论

医学影像学<sup>[1]</sup>能够快速把人体内部的组织器官从结构、密度方面用影像表现出来,医生可根据影像中器官结构和密度的特点得到诊断信息,进而对人体的健康状况进行评估。医学影像学的出现使医疗诊断的形式逐渐从传统方式转变为非侵入的方式。传统的临床诊断方式是活组织检查,但活检意味着用侵入手段获取病变组织,这对病人的身体会造成极大的危害,还有引入并发症的风险<sup>[2]</sup>。而医学影像学则通过代谢信息或辐射检测捕获肿瘤特征,以非侵入的形式帮助医生进行诊断治疗<sup>[3]</sup>。此外,医学影像学很容易与现代的科学技术相结合,如计算机视觉、模式识别等,形成新的科学研究方向,这在未来的医疗诊断中将发挥更大的作用。

计算机辅助诊断(Computer-aided Diagnosis, CAD)<sup>[4]</sup>以计算机为辅助工具,根据身体的生理特点和生化特性,利用影像学、图像处理等技术帮助医生寻找病灶或预测疾病,提升诊断的准确度。计算机强大的数据运算和数据处理能力得到各行各业的青睐,尤其是在医学影像学的方面。在医学影像学中,计算机技术可以提取出影像中病变的结构,接着对提取到的病变特征进行量化,最后对量化特征加以处理,实现疾病的辅助诊断。但CAD只是一种辅佐医生的手段,并不能完全代替医生进行直接的医学判断。

CAD目前已广泛应用到医学影像学中,用于协作医生提高诊断准确性、减少漏诊、提高诊断效率。这是因为在传统的医疗诊断中,每个医生的知识水平各有不同,诊断时容易受到主观意识的影响;其次,不同的医生对影像的诊断有不同的结果;另外,人工诊断很容易遗漏细节的变化,导致结果存在差异。而计算机能客观地按照相同的准则去诊断影像,能够弥补人工诊断的不足<sup>[5]</sup>。

近年来,随着深度学习的飞速发展,人工神经网络被广泛用于智能技术中。它能够模仿大脑神经元进行数学处理,和大脑一样能自主学习、预测事物、记忆规律等,因此可以用于CAD。在各种医学分类问题中,深度学习方法展现出了比传统计算方式更准确的效果<sup>[6]</sup>,是最先进的技术之一。基于深度学习的CAD系统能够分析并处理大量医学数据,从影像中获取肿瘤的高阶纹理特征,也能在自动学习的过程中区分出肿瘤的异质性,进行精准诊断。CAD目前在临床应用中并不能直接代替医生,还面临着许多挑战,但在未来或将成为主流的医学图像计算方式,在智慧医疗中崭露头角,帮助医生提高诊断效率。

## 1.1 研究背景与意义

### 1.1.1 非小细胞肺癌

全球有约六分之一的死亡是癌症造成的<sup>[7]</sup>，每年因癌症死亡的人数高达一千万人左右，且发展中国家的死亡人数占总死亡人数的绝大比例。癌症中最常见的是乳腺癌、肺癌、结肠和直肠癌，其中因为癌症而死亡的最常见原因是肺癌。

肺癌作为一种起源于肺部的恶性肿瘤癌症，一直以来严重影响着人体的生命健康。最近几年，无论是在高收入水平还是中低收入水平国家，死于肺癌的病例报道明显增多，在报道中，男性患上肺癌的概率和因肺癌而死亡的概率高于女性。《2020 全球癌症报告》<sup>[8]</sup>显示，肺癌的发病率位居全球第二，死亡率在全球第一，并且由于人口老龄化、不健康习惯等因素的影响，预计到 2040 年，人类在癌症上的负担还将会提升一半，且新发癌症病例会到新的峰值。在 2020 年统计的癌症病例中，中国患癌人数在所有国家中位于前列，占到了总病例的 23.7%，在死亡病例中，中国占 30%。中国国家癌症中心的数据<sup>[9]</sup>显示，男性和女性患恶性肿瘤最多的类型仍然是肺癌，且肺癌病患的死亡率也是居于所有恶性肿瘤首位。

小细胞肺癌（Small Cell Lung Cancer, SCLC）和非小细胞肺癌（Non-Small Cell Lung Cancer, NSCLC）是肺癌中常用的两种分类，依据细胞外形和生存形式可以将两者区分。小细胞肺癌含有体积较小的细胞，通常始于胸部中部的支气管，只占肺癌患病人数的小部分；而非小细胞肺癌的细胞更大，也更常见，患病人数占比高达 85%。非小细胞肺癌<sup>[10]</sup>的主要亚型有腺癌、鳞状细胞癌和大细胞癌三种。腺癌通常始于肺部的外部区域，是最主要的非小细胞肺癌类型。腺癌在非小细胞肺癌占比达到了 40%，在肺癌中占比达到了 30%<sup>[11]</sup>。鳞状细胞癌通常起源于支气管等区域，但现在也有很多病例被诊断为在肺的外围<sup>[12]</sup>，占非小细胞肺癌的 30%左右。大细胞癌由于难以诊断只能通过排除其他可能性进行确定，可以发生在肺部的任何地方，但比较罕见，仅占非小细胞肺癌患者的 5%至 10%<sup>[13]</sup>。医生可以根据肺癌患者所属的亚型制定相应治疗方案。目前存在不同的方法来诊断肺癌，包括影像学检查、活组织检查、肺功能测试等。

癌症的主要特点是肿瘤生长和发展的速度较快，会出现全身的转移和扩散，癌症可能会从原发肿瘤开始快速地向周围组织扩散，也可能通过淋巴管或血液发生转移，还可能从胸腹、脊髓腔发生腔道种植，其中淋巴管和血液转移最为常见。在非小细胞癌中，

淋巴结转移作为最主要的和最常见的转移方式<sup>[14]</sup>，对分期和预后会产生重大影响，也会累及病人的身体状态。发生肿瘤转移另一大后果就是会导致患者出现多种并发症，进一步的降低患者的生存时间。因此，准确评估病人是否发生淋巴结转移能在早期给医生参考，进而有效治疗，有着重要的意义<sup>[15-16]</sup>。

### 1.1.2 PET-CT 介绍

正电子发射断层扫描技术（Positron Emission Tomography, PET）<sup>[17]</sup>是一种灵敏度较高的定量功能成像技术，通过无创的方式给病人注射放射性示踪剂，然后检测辐射并重建示踪剂分布来提供生理信息。具体的显像过程为：将发射正电子的放射性同位素标记到参与血流或代谢过程的化合物上，其中氟-18 是最常用的放射性同位素，然后将其注射到患者体内；放射性同位素会在患者体内发生衰变，当衰变到稳定状态时，发射的正电子前进 1-2 毫米后与组织的负电子结合，得到能量同为 0.511MeV 但互为反向的两个 $\gamma$ 光子；两个 $\gamma$ 光子经过体内的相反路径最终达到探测器，如果两者到达时间处于短期重合时间窗口（通常为 1-10ns）内，脉冲处理器将发射方波到工作站中，以准确地重建放射性同位素的分布。大量的光子探测器能沿着人体相关器官和组织收集多角度的放射性信息，通过这些信息能得到人体不同部位的冠状、矢状或横截面影像。

PET 可以通过使用葡萄糖类似物作为示踪剂描述人体内葡萄糖摄取增加的情况，恶性细胞比正常组织需要更多的葡萄糖来满足能量需求，因而可以确定癌症的影响区域<sup>[18]</sup>。此外，一些对葡萄糖需求量或存储量较大的器官（心脏、膀胱、肾脏等）、棕色脂肪以及炎症等也能在 PET 上呈现出高代谢信号。PET 可以通过代谢信息成像并进行定量分析，以此得到不同角度的可辅助医生诊断的截面影像；更微观的，可以从分子的层面得到人体内代谢物的活动情况。目前，PET 影像已经在临床诊断、指导治疗各种疾病中显示出独特的优越性，如肺癌、冠心病、结直肠癌、脑部疾病等。

电子计算机断层扫描（Computed Tomography, CT）<sup>[19]</sup>是一种可以快速对人体组织扫描成像的技术，该技术首先对人体穿过一系列不同角度的 X 射线，然后检测辐射并重建图像来提供人体的解剖结构信息。具体的显像过程为：在 CT 扫描仪机架内让 X 射线以不同的角度穿过病人的身体，其中一些 X 射线被身体吸收，另一些则被病人另一侧的探测器检测到，由于人体不同的组织或器官密度不同，对 X 射线的吸收和透过率不同，探测器接收到的信号强弱也会不同；将探测器得到的信号根据衰减情况进行计算从而获得

体素的 X 射线衰减或吸收值,根据得到的数值进行矩阵化运算以字符形式保存到计算机中,最终依据数值的不同对应不同的灰度值从而构成 CT 图像。

在 CT 图像中,人体器官和组织对 X 射线的吸收差异体现在灰度的深浅上,对于吸收率高的部位来说,图像上所呈现的为白色,代表部位为组织密度较高的骨骼;对于吸收率低部位来说,图像上所呈现的为黑色,代表部位为充满空气的肺部。此外,CT 图像可以在成像上呈现体内的细小病变,也能很好地显示出脑、肺、肝、胆等软组织器官。CT 通过人体对 X 射线的吸收状况成像并进行定量分析,可以重建出多角度的断层影像,查看器官和病变的关系,研究人体生理结构。目前,CT 影像已经在癌症定位、确定癌症类型、癌症分期等方面做出了卓越的贡献,为脑部疾病、胸部疾病、中枢神经系统疾病的治疗与诊断提供了较高的价值<sup>[20]</sup>。

虽然 CT 图像分辨率高,能准确描绘解剖细节,划分病变与周围组织的边界,但缺乏显示生理学上的关键差异的能力;而 PET 能描绘组织代谢的生理学信息,准确定位代谢异常的部位,但缺乏高分辨率的解剖结构信息。因此 PET-CT 应运而生,它对 PET、CT 进行了图像融合,有效结合了 CT 的解剖结构信息和 PET 的代谢信息。PET-CT 以超高的灵敏度提高了病灶清晰度,目前在疾病的诊断和定位方面已经得到了广泛的应用,在诊断方面,PET-CT 可用于疾病的复发诊断和肺癌检查等;在定位方面,PET-CT 可用于脑肿瘤精确定位、癫痫定位等。此外,在术后疗效随访和预后评估等方面 PET-CT 也发挥了重大的作用<sup>[21-24]</sup>。图 1.1 为 PET-CT 示例图,从左至右依次为 PET 图像、CT 图像、PET 最大强度投影图像,其中绿色十字表示肿瘤:

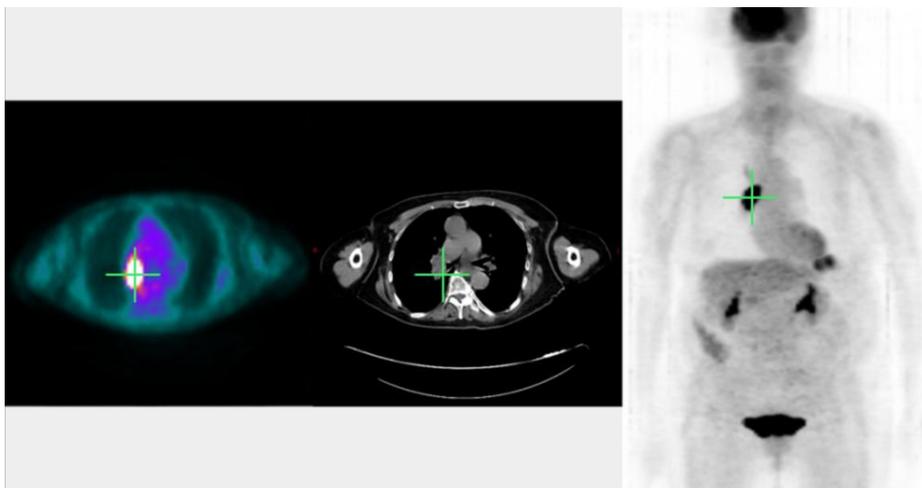


图 1.1 PET-CT 示例图

最大强度投影 (Maximum Intensity Projection, MIP) 技术是一种用于三维数据的可视化方法, 该技术以平行光线的方式将具有最大强度的体素投影到二维可视化平面, 通过描述从某一角度看到的体素的最大强度值, 捕获图像中包含的器官或组织的结构。该方法很好地识别了由高信号强度定义的结构, 可以用于疾病诊断<sup>[25]</sup>。例如在图 1.1 右边的 PET 正面最大强度投影图像中, 可以直观地看出人体的大脑、心脏、肾脏、膀胱比其他组织具有更高的强度值, 这是由于这些器官的代谢作用更强。此外, 由于恶性肿瘤需要更多的葡萄糖来满足能量需求, 代谢作用会强于周围组织, 所以也会显现出高亮度的特征。

### 1.1.3 非小细胞肺癌淋巴结转移辅助诊断的临床价值

非小细胞肺癌作为肺癌中最常见的一种, 是医疗人员和研究人员关注的重点。非小细胞肺癌的治疗和预后取决于多种因素, 包括肿瘤大小和等级、淋巴结转移状态、患者的临床表现等。其中, 肿瘤转移的程度被认为是非小细胞肺癌生存的最可靠预测指标<sup>[26]</sup>。因此, 准确预测、诊断非小细胞肺癌淋巴结转移对病人的治疗和预后十分重要。

肿瘤转移意味着生存期下降、复发率升高, 对患者生存状态的评估十分重要。非小细胞肺癌常常伴随着淋巴结转移或远端转移, 淋巴结转移作为最平常见的一种转移方式, 是相对具有较好预后的一种肿瘤转移。如果在肿瘤没有发生进一步扩散的情况下准确诊断, 就可以选择不同的治疗方案, 例如出现肺门淋巴结转移时可以选择手术得到根治性治疗, 而出现纵膈淋巴结转移时可以结合手术、化疗等多种治疗方式来延长患者生存时间<sup>[27]</sup>。目前最准确的识别方法是活组织检查, 这种方式对患者的身体会造成较大的危害, 还有引入并发症的风险, 所以大多数患者选择通过影像诊断来确定治疗方案。虽然影像学诊断不具有侵袭性, 但不同医生可能产生不同的诊断结果, 并且逐一分析影像需要耗费医生大量的时间、精力, 准确性还有待提高。因此, 可以通过计算机对影像进行客观分析, 进而对肿瘤转移进行诊断, 根据肿瘤转移程度指导医生制定合理的应对方案, 提高肺癌患者的生存率<sup>[16]</sup>。

影像组学特征可以实现对肿瘤的异质性、解剖学特征、形态学特征量化描述, 还能表示出肿瘤潜在的分子特性, 使其以非侵入的方式预测癌症、辅助诊断。越来越多的研究表明, 利用深度学习分析影像组学特征, 进行淋巴结转移辅助诊断, 其临床和研究意义是十分重要的<sup>[28-30]</sup>。

## 1.2 研究现状

目前,计算机辅助诊断已经广泛应用到发达国家的医疗系统中,发展中国家也正在快速普及这一技术。大数据和深度学习的发展推动着计算机辅助诊断的进步,医学影像领域也依赖于此技术。在肿瘤转移的预测和诊断方面,也涌现出了大量的研究和算法。这里根据使用方法的不同将利用影像组学进行肿瘤转移预测或诊断的研究分为传统的影像组学方法和基于深度学习的方法两大类。

### 1.2.1 传统的影像组学方法

与医生对医学图像直接做出解释并诊断的方法不同,影像组学是计算机辅助诊断系统的延伸,它通过从影像中提取定量特征,挖掘数据并定量分析进而实现临床决策支持。具体来说是通过图像特征提供肿瘤表型和微环境的信息,使用分析工具区分肿瘤在生理学上的高度异质性,从而提高医生的诊断准确性和对患者的预后可靠性,指导治疗方案选择。传统的影像组学方法通常需要提取影像的一阶、二阶及高阶统计量特征,并与患者数据相结合,通过机器学习方法或其他统计学工具进行数据挖掘,开发提高诊断准确性的预测或诊断模型<sup>[31-32]</sup>。

Li<sup>[33]</sup>等人根据超声图像特征预测了 126 例甲状腺癌患者的淋巴结转移情况,并确定了其临床价值。首先在超声图像中利用软件绘制感兴趣区域,在感兴趣区域中提出 1079 个高通量影像组学特征;将这些高通量影像组学特征经过特征共线操作剔除不重要的特征,最终筛选出 690 个特征,再利用假设检验法对得到的影像组学特征进行降维,进一步得到 91 个特征;最后在多种机器学习方法的对比中选择装袋算法(Bagging)来建立预测模型。实验结果表明该模型具有较好的效果(训练集AUC: 0.759; 验证集 AUC: 0.803)。此外,对超声图像特征与淋巴结转移预测的关系给出实验分析,例如不规则形状特征( $P=0.025$ )和微钙化特征( $P=0.004$ )与是否发生淋巴结转移存在显著差异,其 AUC 分别为 0.591 和 0.629。该研究证明了传统的影像组学方法评估超声特征预测甲状腺癌淋巴结转移的可行性,在诊断和预后中具有重要医学价值。

Ji<sup>[34]</sup>等人对 103 例肝内胆管癌患者进行了回顾性分析,根据 CT 图像提取影像组学特征进行淋巴结转移预测。首先使用软件手动分割原发肿瘤形成感兴趣区域,在此区域中提取了 105 个特征,其中包含了形状、纹理和一阶特征;然后使用 LASSO 进行特征

选择,得到非冗余且稳健的 8 个与淋巴结转移状态相关的特征;最后基于多变量 Logistic 回归分析建立一个结合影像组学特征和独立风险因素的模型。该模型在验证集上的 AUC 值为 0.8921,在 CT 报告的淋巴结转移阴性亚组中 AUC 为 0.9224。此外,基于影像组学的列线图取得了比临床预测模型、淋巴结形态学更好的性能。该研究证明了传统的影像组学模型能有效诊断肝内胆管癌淋巴结转移,为医学临床决策提供支持。

Zhang<sup>[35]</sup>等人对 61 例乳头状甲状腺癌患者进行了回顾性分析,根据磁共振图像(MRI)提取影像组学特征,利用这些影像组学特征为淋巴结转移预测提供更多信息。首先对肿瘤手动分割得到感兴趣区域,并采取一些预处理方式来增强类别差异,提高纹理对比度;接着用计算机提取图像中共计 396 个特征,提取的特征包括强度直方图、形状特征、灰度共生矩阵(GLCM)等总共五种类型;最后使用曼惠特尼 U 检验来选择特征,并构建三个预测模型(随机森林(RF)、支持向量机(SVM)、广义线性模型(GLM))进行评估。该方法在随机森林模型中取得了最佳的性能,达到了 0.85 的 AUC 和 0.87 的准确度。此外,实验结果证明了 T2WI 图像中的影像组学特征比 T2WI-FS 图像更能准确预测淋巴结转移情况。

Liu<sup>[36]</sup>等人根据动态对比增强磁共振成像(DCE-MRI)对 62 例乳腺癌患者进行了前哨淋巴结转移预测。首先在 DCE-MRI 上手动确定肿瘤轮廓得到 62 个感兴趣体积,接着在这些手动确定的感兴趣体积中自动获取 149 个影像组学特征;然后用选择 K 最佳和 LASSO 算法选择特征并降维,得到 6 个最佳特征用于预测;最后构建三个分类器(SVM、XGboost、逻辑回归(LR))验证所选特征在预测前哨淋巴结转移中的有效性。在训练集中 XGboost 模型取得了最优的 AUC (0.92) 和准确度 (0.82),而在验证集中 SVM 取得了最好的结果, AUC 为 0.85,准确度也为 0.85,并且均方误差也最低 (0.26)。该研究证明了利用传统的影像组学方法,能够在 DCE-MRI 上无创地预测乳腺癌患者的前哨淋巴结转移,从而辅助临床诊断,具有现实意义。

根据肿瘤在发展过程中形态或性质的变化可将非小细胞肺癌进行分期,分期对于治疗方案的选择、放疗靶区的划定都具有重要意义。然而,由于非小细胞肺癌在早期的形状特征不够明显,在患者身上体现的症状也很少,医生很难把握好对病人的治疗程度,如果诊断不够精确,容易对患者造成更大的伤害。为了在早期及时有效治疗非小细胞肺癌患者,如何快速精确地预测出患者体内淋巴结转移的几率意义重大。Cong<sup>[37]</sup>等人分析了 649 例 IA 期非小细胞肺癌患者的 CT 图像,对其进行了淋巴结转移预测,研究了影

像组学分析对肿瘤诊断的意义。首先在医生的帮助下用软件勾画出原发肿瘤的兴趣区域，利用工具提取 396 个包括直方图特征、GLCM 特征、灰度大小区域矩阵 (GLSZM) 特征在内的 6 种影像组学特征；然后采用曼惠特尼 U 检验和方差分析进行特征降维，并使用 LASSO 算法从提取到的特征中选择最稳健的 7 个影像组学特征；最后建立随机森林模型预测淋巴结转移。最终结合影像组学特征和临床特征的模型在训练集中 AUC 达到 0.911，在测试集中 AUC 达到 0.860，取得最优结果。该研究证明了影像组学特征能够很好地区分肿瘤异质性，且与临床特征相结合能提高临床模型的性能，可以更好地对非小细胞肺癌患者是否发生淋巴结转移进行预测，帮助医生选择治疗方案。

此外，还有很多利用传统影像组学方法对各种癌症类型的肿瘤进行转移预测和诊断的研究<sup>[38-41]</sup>。然而，传统的影像组学方法繁琐且耗时，它们通常需要经过手动分割、特征提取和选择，比较分类器性能等流程，并且在这些过程中可能存在较差的重复性和对参数的敏感性，无法适应特定的临床问题。

### 1.2.2 基于深度学习的方法

深度学习<sup>[42]</sup>是一种强大的特征学习、表示方法，它可以提取图像的高层次语义信息，自动地学习与任务相关的特征，从而提高特征学习效率，在独特的优势下已经迅速成为了分析医学图像的首选方式。与传统的影像组学方法提取的手工定量特征不同，深度学习通常提取医学图像中与特定临床任务有关的复杂信息。在早期，深度学习通常用于医学图像的分类和检测，并且取得了不错的结果，有时甚至能达到人类专家的水平<sup>[43-44]</sup>，这具有十分重要的临床意义。最近的研究表明，基于深度学习的影像组学方法还可以从医学图像中提取与预后相关的信息，并在涉及多种成像模式和应用领域的肿瘤转移预测、诊断任务中具有良好的前景。

Lee 等人<sup>[45]</sup>在 CT 图像上利用深度学习方法，对 202 例甲状腺癌患者是否发生颈部淋巴结转移进行了辅助诊断。首先由放射科医生手动识别 CT 图像上颈部淋巴结的病变，在病变周围取一个矩形区域作为深度学习模型的输入；然后建立八种深度卷积神经网络模型来区分未转移和转移淋巴结，包括 VGG16、VGG19、Exception、DensaNet121、ResNet 等模型；最后生成未转移和转移淋巴结的注意力热图，来可视化模型重点学习的区域。八种深度学习算法的 AUC 均达到 0.9 以上，其中 ResNet50 模型的性能最好，AUC 达到 0.953，准确度达到 0.904。此外，注意力热图显示出模型能从恶性肿瘤区域提取特征，

并将其作为淋巴结转移状态的分类基础，提高了可解释性。该研究证明了基于深度学习的计算机辅助诊断系统用于甲状腺癌患者颈部淋巴结转移的有效性，能进行准确分类帮助医生提高诊断效率。

Hu 等人<sup>[46]</sup>对 222 例胃癌患者做了回顾性分析，利用深度学习方法对显微扫描仪下的病例图像进行淋巴结转移诊断。该方法包括两个部分：淋巴结定量和转移癌识别。对于淋巴结定量，首先由经验丰富的病理学家标注感兴趣区域，然后结合目标检测中的 Faster RCNN 模型和图像分割中的 DeepLab 模型在感兴趣区域中分割淋巴结；对于转移癌识别，首先对图像进行处理防止内存占用过大，然后结合深度神经网络 Exception 和 DenseNet-121 提取与肿瘤转移相关的特征。在测试集上，淋巴结定量过程达到了 0.971 的准确度，转移癌识别过程取得了 0.935 的阳性预测值和 0.979 的阴性预测值。该研究设计出一个无创的计算机辅助诊断系统，在无需勾画肿瘤边界的情况下能准确识别出转移淋巴结和未转移淋巴结，大大减轻了医生工作量，提升了诊断精度，具有较高的临床应用价值。

Wu 等人<sup>[47]</sup>根据 MRI 图像，利用深度学习模型对 894 例早期宫颈癌患者是否发生了淋巴结转移进行了计算机辅助诊断。首先由放射科医生用紧贴着肿瘤边界的矩形边框来定位肿瘤，形成肿瘤感兴趣区域，还将矩形边框扩大 5 个像素来添加肿瘤周围信息，形成包含瘤周的感兴趣区域；然后建立一个类似 ResNet 的端到端的深度学习模型来预测淋巴结转移状态，还在分类前的全连接层中加入临床特征（年龄、FIGO 分期和 MRI-LN 状态）建立一个混合模型来提高模型性能。此外，还比较了不同的 MRI 序列在预测淋巴结转移问题上的效果。在验证集中，同时使用瘤内和瘤周区域的模型表现更好，并且在对比增强 T1 加权成像（CET1WI）序列中达到最好效果，AUC 为 0.844，准确度为 74.47%。在混合模型中这个组合的性能得到了进一步提升，AUC 达到了 0.933，准确度达到了 87.94%。该研究证明了肿瘤内部特征、肿瘤周围特征以及临床特征都能提升淋巴结转移识别模型的性能，并且 CET1WI 序列在宫颈癌淋巴结转移预测中能提供更大的价值，在临床辅助诊断方面很有前景。

Zhao 等人<sup>[48]</sup>用三维卷积神经网络对 51 例 T1 期肺腺癌患者的 CT 图像进行分析，准确地预测了淋巴结的转移情况。首先根据手动标注的肿瘤中心剪裁 CT 图像，得到包含肿瘤及肿瘤周围的感兴趣区域；然后将感兴趣区域放入一个类似 DenseNet 模型的三维卷积神经网络（DensePriNet）中自动提取与淋巴结转移相关的特征；接着将网络提取的

图像特征与临床特征（影像学特征、吸烟状态、癌胚抗原（CEA）水平等）连接；最后通过全连接层输出发生淋巴结转移的概率。该模型取得了 0.926 的 AUC 和 87.6% 的准确度，显著高于传统的影像组学方法（AUC:0.891，准确度：90.2%）和医生诊断（AUC:0.723，准确度：83.2%）。该研究开发了一个融合 CT 图像特征和先验临床特征的三维神经网络，对肺腺癌患者是否发生淋巴结转移进行自动诊断，在诊断精度上超越了医生，能够促进精准医疗的应用。

对于早期的非小细胞肺癌患者，医生通常使用直接手术的方法进行治疗，但由于早期症状不明显，这些患者在治疗过程中会被发现发生了纵膈淋巴结（N2）转移<sup>[49]</sup>。对这些患者来说，直接进行手术可能会造成不可逆转的伤害，因而不是治疗的最佳方案。所以，Zhong 等人<sup>[50]</sup>回顾性地分析了 2663 例临床 I 期非小细胞肺癌患者的 CT 图像，利用深度学习提高患者 N2 转移的预测准确性。首先在 CT 图像中截取一个包含肿瘤在内的方框形成感兴趣区域；然后建立一个 ResNet152 模型对感兴趣区域发生 N2 转移的概率进行预测；最后在内部测试集、外部测试集、前瞻性测试集中验证该方法的有效性。在内部测试集中，该方法的 AUC 为 0.82，准确度为 77%；在外部测试集中，该方法的 AUC 为 0.81，准确度为 81%；在前瞻性测试集中，该方法的 AUC 为 0.81，准确度为 78%。此外还探索了模型学习到的深度学习特征对预后的价值。该研究证明了深度学习特征可以用于非小细胞肺癌患者 N2 转移的准确预测，能够在早期提升医生的诊断精度，帮助制定治疗方案，具有较大的临床意义。

此外，在不同模态的成像和不同癌症类型中还有许多使用深度学习对肿瘤进行转移预测、诊断的研究<sup>[51-54]</sup>。但是，用 PET-CT 图像上的特征预测或诊断非小细胞肺癌患者是否出现淋巴结转移预案中，目前的研究和方法还存在可探究的空间。因此本文主要利用深度学习方法，基于 PET-CT 图像对非小细胞肺癌的淋巴结转移辅助诊断展开研究和探讨。

### 1.3 主要研究内容及创新

本文从 PET-CT 图像中获取非小细胞肺癌患者的影像特征，使用深度学习方法对非小细胞肺癌患者的淋巴结转移状况进行了辅助诊断。首先充分利用医生勾画的肿瘤金标准，在 PET 图像中提取传统的影像组学特征并加以处理，利用卷积神经网络提取原发肿瘤的三维深度学习特征，加入正面半身最大强度投影的二维深度学习特征，最后融合临

床特征和医生诊断，对非小细胞肺癌淋巴结转移进行判断。为了减少医生勾画肿瘤的工作量，提升工作效率，进一步使用两阶段、以分割任务为引导的方式消除测试阶段对肿瘤金标准的依赖。具体来说，在第一阶段利用正面半身最大强度投影图像对肿瘤进行分割，得到大致位置并作为第二阶段的测试集选取依据；在第二阶段，用分割-分类模型对三维肿瘤和二维正面半身最大强度投影进行特征提取，使模型能在学习时关注肿瘤区域，并在测试时不需要肿瘤金标准；最终融合两种深度学习特征，对非小细胞肺癌淋巴结转移进行辅助诊断。

本文的创新点在于：

(1) 在非小细胞肺癌淋巴结转移辅助诊断任务中，利用最大强度投影图像提取与肿瘤转移相关的特征，并且实验证实了加入最大强度投影图像提取到的特征能提升模型性能，这说明了全局信息对淋巴结转移诊断任务有促进作用，也将为其他研究带来新思路。

(2) 用深度学习的方法融合了传统的影像组学特征、三维原发性肿瘤深度学习特征、二维半身最大强度投影深度学习特征、临床特征以及医生诊断结果进行辅助诊断，并且实验证明了该融合能有效提升模型对非小细胞肺癌淋巴结转移诊断的性能，说明了多种数据融合的有效性，能为其他问题的解决提供借鉴。

(3) 提出一种两阶段的深度学习方法。该方法引入分割任务，不仅能对肿瘤进行大致定位，还能在特征学习中使网络关注在肿瘤区域上。这将消除测试阶段对医生勾画的肿瘤金标准的依赖，大大减少医生的工作量，节约医生时间，为下一步快速准确订制医疗预案打下坚实基础。

(4) 所提出的两个淋巴结转移辅助诊断模型的性能都要优于医生的判断，尤其是在肿瘤发生转移的情况中，模型诊断错误的病例个数少于医生的诊断。这有助于临床诊断中医生对病人身体状况的判断，也能帮助医生进行治疗决策、评估预后，有效提升非小细胞肺癌人群生活质量、降低死亡率。

## 1.4 本文组织结构

本文基于深度学习方法，利用 PET-CT 图像，对非小细胞肺癌患者是否会出现淋巴结转移进行了辅助诊断，具体的文章组织结构如下：

第一章：绪论。本章主要对非小细胞肺癌淋巴结转移辅助诊断的研究背景和意义进

行介绍，然后对现有的肿瘤转移预测、诊断的研究进行详细的阐述，最后对本文的内容、创新点以及文章的结构安排进行了说明。

第二章：相关理论。本章主要对涉及的理论知识进行介绍，包括卷积神经网络的基本结构以及分类问题中的评价指标。

第三章：实验数据。本章主要对研究数据的来源和选择过程进行了详细介绍，并列出了本文所用数据的具体分布。

第四章：非小细胞肺癌淋巴结转移辅助诊断。本章主要对淋巴结转移辅助诊断的研究流程、具体方法以及实验结果进行了详细的介绍。

第五章：两阶段非小细胞肺癌淋巴结转移辅助诊断。本章主要讲述两阶段的淋巴结转移诊断研究流程，以分割任务为引导的特征提取方式以及实验的结果。

第六章：总结与展望。对本文内容进行总结，并讨论了研究中存在的缺陷与不足之处，分析了未来工作中可以提升的方向。

## 2 相关理论

本文是基于深度学习的非小细胞肺癌淋巴结转移辅助诊断研究，主要使用的是深度学习中的卷积神经网络搭建模型进行计算机辅助诊断的，因此本章对卷积神经网络中最主要的部件及其作用进行介绍。由于本文探究的问题本质上属于分类问题，所以本章对分类问题中的评价指标（准确度、精准率、召回率、AUC 等）也进行了简要阐述。

### 2.1 卷积神经网络

卷积神经网络（Convolutional Neural Networks, CNN）是一种十分常用的深度学习模型<sup>[55]</sup>，它模仿生物的视觉神经元对事物进行分析、处理，使计算机能像人脑一样观察并理解世界。卷积神经网络是一种深层次的前馈神经网络<sup>[56]</sup>，它拥有超强的特征学习能力，能挖掘数据中反映本质的特征从而进行分类或可视化。20 世纪 80 年代，一些对卷积神经网络的探索为其应用到不同领域打下了基石，例如时间延迟网络<sup>[57]</sup>、平移不变人工神经网络<sup>[58]</sup>以及 LeNet<sup>[59]</sup>，它们被用于语音识别、医学图像处理等方面。随着深度学习理论被广泛提倡，更多的复杂卷积神经网络也涌现出来，用于解决金融、医疗、城市等问题。下面对卷积神经网络的基本结构进行介绍。

#### （1）卷积层

卷积神经网络在接收到一维、二维或多维数组后利用隐含层进行特征学习，其中卷积层是隐含层中最重要的提取特征的组件。卷积层内部包含多个卷积核，每个卷积核中的元素是一个神经元，能让网络在传播和迭代的过程中进行计算。每个神经元与前一层对应位置的相邻区域连接，这个区域就是卷积核的大小，通常是  $1 \times 1$ 、 $3 \times 3$ 、 $5 \times 5$  等，在进行计算时在卷积核大小的矩阵中做元素相乘并求和，再加入偏置量，完成一次卷积操作。卷积操作通过不断地运算从低级特征中提取高级语义信息，它的输出便是特征图。此外，卷积核的步长是影响特征图大小的重要参数。

在上述的过程中，可以发现卷积操作实际上就是矩阵相乘，因此输出永远是输入的线性组合，这种简单的学习无法满足复杂任务的需要。而现实中需要解决的问题往往是难以理解的、非线性的，因此在卷积层中引入激活函数来提高模型对非线性问题的表达能力，使模型应用到广泛的任务中。常用的激活函数有 Sigmoid<sup>[60]</sup>函数、Tanh 函数和 ReLU<sup>[61]</sup>等。

### (2) 池化层

在经过卷积层提取特征后，需要周期性地插入池化层来选择特征，去除冗余信息。池化层是一种下采样，它能降低特征维度，在保持特征不变的同时缩小输出特征图的尺寸，可以减少参数和计算量。此外，它能像激活函数一样实现非线性操作，使网络满足复杂问题的需要，还能降低过拟合的概率，扩大感受野。常见的池化层包括最大池化（Max Pooling）、平均池化（Average Pooling）和全局池化（Global Pooling）<sup>[62]</sup>等。最大池化是将输入按照步长划分为多个池化区域，对每个区域输出最大值；平均池化则是计算每个区域的平均值作为此区域的输出值；而全局池化则是将每个特征图输出为一个值，包括全局平均池化和全局最大池化。

### (3) 全连接层

在卷积神经网络输出前通常连接全连接层，对提取到的特征进行非线性组合，从而进行分类。大多数时候一个全连接层不能很好地解决非线性问题，因此常常使用两个或两个以上的全连接层来增强特征表达能力。具体来说，根据全连接层衔接的对象不同，可以将全连接操作表示为全局卷积或  $1 \times 1$  的卷积。可以看出全连接层实际上就是对特征进行映射，得到样本分类结果。

## 2.2 分类问题评价指标

在分类问题中，常用一些评价指标来评估分类器的分类性能，例如准确度（Accuracy）、精准率（Precision）、召回率（Recall）等。在医学领域的分类中，常出现数据不平衡的问题，因此需要结合多种评价指标进行综合评价。

### (1) 混淆矩阵

混淆矩阵<sup>[63]</sup>能直观地表达模型的分类性能，展示了真实与预测类别的情况。二分类混淆矩阵如表 2.1 所示：

表 2.1 二分类混淆矩阵

真实类别	预测类别	
	1	-1
1	TP	FN
-1	FP	TN

其中 TP 表示将正例预测为正例的样本数；FP 为将反例预测为正例的样本数；FN 为将正例预测为反例的样本数；TN 为将反例预测为反例的样本数。其他常用的分类评价指标也可以用混淆矩阵中的数表示出来。

(2) 准确度 (Accuracy)、精准率 (Precision)、召回率 (Recall)

Accuracy 反映了分类器对整体的判断能力，计算公式如下：

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (2.1)$$

Precision 表示预测为正例的样本中有多少是真正的正例，计算公式如下：

$$Precision = \frac{TP}{TP + FP} \quad (2.2)$$

Recall 表示所有真正的正例样本中有多少被分类器预测正确，计算公式如下：

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

(3) 受试者工作特征 (ROC) 曲线下的面积 (AUC)

无论是准确度、精准率还是召回率，都依赖于阈值得到具体的数值，阈值选取的不同会导致这些指标结果不同，所以用 ROC 曲线综合表现不同阈值下的预测情况。以 FPR (False Positive Rate) 为横坐标，TPR (True Positive Rate) 为纵坐标，便可以画出 ROC 曲线，ROC 曲线是一系列阈值 (FPR, TPR) 下的数值点的连线。其中：

$$TPR = \frac{TP}{TP + FN} \quad (2.4)$$

$$FPR = \frac{FP}{FP + TN} \quad (2.5)$$

AUC 则是 ROC 曲线下的面积，反映了分类器对正负样本的分类能力，尤其是能在样本不平衡的情况下进行合理评估，被广泛用于各种医学分类问题中。AUC 的取值在 0 到 1 之间，超过 0.5 则表示分类器比随机猜测的结果好，越接近 1 分类器性能越好<sup>[64]</sup>。

## 2.3 本章小结

本章简要地介绍了本文涉及的理论知识，包括卷积神经网络中最主要的部件及其作用以及针对本文的医学应用问题采用的各种分类指标。

## 3 实验数据

### 3.1 数据来源与选择

本文回顾性地分析了 2004 年至 2014 年在美国芝加哥大学医学院接受诊疗的非小细胞肺癌患者的 PET-CT 图像和临床资料。首先我们获得了 935 例有病理证实的、没有发生脑转移或其他疾病的非小细胞肺癌患者的 PET-CT 影像，这些影像包含了从头到脚的全身信息，并且能清晰地反映出人体高代谢的组织、器官及骨骼等。其次，本文主要使用深度学习方法，对非小细胞肺癌患者是否发生淋巴结转移进行辅助诊断，针对此任务需要进一步筛除不符合研究内容的的数据。筛除条件包括：（1）发生了远端转移的患者；（2）没有经过病理学检查（活体组织检查）确定是否发生淋巴结转移的患者；（3）无法获取临床特征（年龄、吸烟史等）的患者。最终在 935 例非小细胞肺癌患者中剩余 121 例患者符合研究条件。这 121 例非小细胞肺癌患者有 19 例患者经过活组织检查确定发生了淋巴结转移，另 102 例没有发生淋巴结转移。

由于医疗中心在 2012 年 3 月 15 日更换过一次扫描设备，所以 121 例非小细胞肺癌患者的 PET-CT 影像中有 100 例是 Reveal HD PET-CT 成像仪扫描的结果，21 例是 Siemens mCT 成像仪扫描的结果。不同的成像仪扫描的影像有不同的参数，例如切片的层厚、影像的分辨率以及像素对应的空间体素的大小等都存在两种不同的值，在应用到具体任务时可以通过预处理进行统一。

对每个患者都有经验丰富的放射科医生手动勾勒其体内的肿瘤，这些肿瘤是医生使用专业软件（MIMvista5.1.2）在 PET 和 CT 上分别勾画的，勾画的结果被称为肿瘤金标准（Gold Standard）。此外，由于一些患者的 CT 影像存在问题，并且考虑到 PET 图像具有较高的敏感性，可以单独用于医学任务分析，所以本文只使用患者的 PET 图像和医生勾画的 PET 肿瘤金标准作为数据输入进行辅助诊断。具体的患者数据统计如表 3.1 所示：

表 3.1 淋巴结转移辅助诊断的患者数据统计

	总计	淋巴结未转移	淋巴结转移
患者数目	<b>121</b>	102	19
年龄			
最小值	42.7	42.7	54
最大值	87	87	85
均值	66	67	65
吸烟史			
曾有吸烟史	64	55	9
吸烟	41	34	7
无吸烟史	16	13	3
Laterality Description			
Left	56	50	6
Right	65	52	13

由表 3.1 可知，121 例非小细胞肺癌患者中淋巴结发生转移的人数比例为 15.7%，是未发生转移人数的约五分之一。此外，这 121 例患者都是中老年人，并且曾有吸烟史或正在吸烟的人数远远多于不吸烟人数。Laterality Description 这个特征用于表示原发肿瘤位于左肺还是右肺，位于左右肺的患者比例相近。

### 3.2 临床特征处理

临床特征是患者身体数据的数字表达，通常年龄、临床 TNM 分期、吸烟史、组织学类型等特征被认为是重要的评估预后的指标。本文选择年龄、吸烟史以及 Laterality Description 作为非小细胞肺癌淋巴结转移辅助诊断的临床特征，其中年龄为数值型数据，可直接进行归一化处理，其他特征为分类数据，需要用数值表示后再进行归一化。例如在 Laterality Description 特征中，用 0 表示原发肿瘤位于左肺，1 表示原发肿瘤位于右肺。

### 3.3 本章小结

本章具体介绍了所用的 PET-CT 影像的来源、淋巴结转移辅助诊断研究中使用数据的筛选过程以及参与研究的患者的统计数据。此外还描述了对所用的临床特征（年龄、吸烟史、Laterality Description）的处理。

## 4 非小细胞肺癌淋巴结转移辅助诊断

在基于深度学习的 PET-CT 非小细胞肺癌淋巴结转移辅助诊断研究中，共包括 121 例非小细胞肺癌患者，其中 19 例患者发生了淋巴结转移，102 例患者未发生淋巴结转移。为了准确识别出淋巴结转移与未转移患者，本研究旨在充分挖掘可学习特征，从中自动捕获与淋巴结转移最相关的信息。一方面，从患者的 PET 图像中提取传统的影像组学特征，包括基于强度、纹理和形状的 833 个特征；另一方面，利用深度学习从患者的 PET 图像中提取特征，包括从原发肿瘤和正面半身最大强度投影中提取的深度学习特征。传统影像组学特征和深度学习特征都为淋巴结是否发生转移提供了可学习依据，临床特征能提供一些与肿瘤转移相关的统计量，医生诊断结果也能为模型带来先验知识。因此本研究提出了一个端到端的深度学习架构，该架构可以融合深度学习图像特征、传统影像组学特征、临床特征以及医生诊断结果，对非小细胞肺癌患者是否发生淋巴结转移进行辅助诊断。

### 4.1 方法

#### 4.1.1 研究流程

本研究的主要目的是充分挖掘有助于淋巴结转移辅助诊断的可学习特征，进而提升诊断的准确率，帮助临床医生提高工作效率，减少误诊或漏诊的风险。本文首先使用传统方法提取影像组学特征，并且初步利用机器学习方法进行淋巴结转移诊断。但在传统方法处理的过程中存在较差的重复性，模型对参数也十分敏感，诊断的性能往往效果不佳。因此，在利用传统影像组学特征的基础上，构建了一个融合深度学习特征、临床特征以及医生诊断的框架进行淋巴结转移辅助诊断，该框架包括四步流程：传统影像组学特征提取、深度学习数据预处理、深度学习特征提取、特征融合与分类。研究的具体流程如图 4.1 所示。

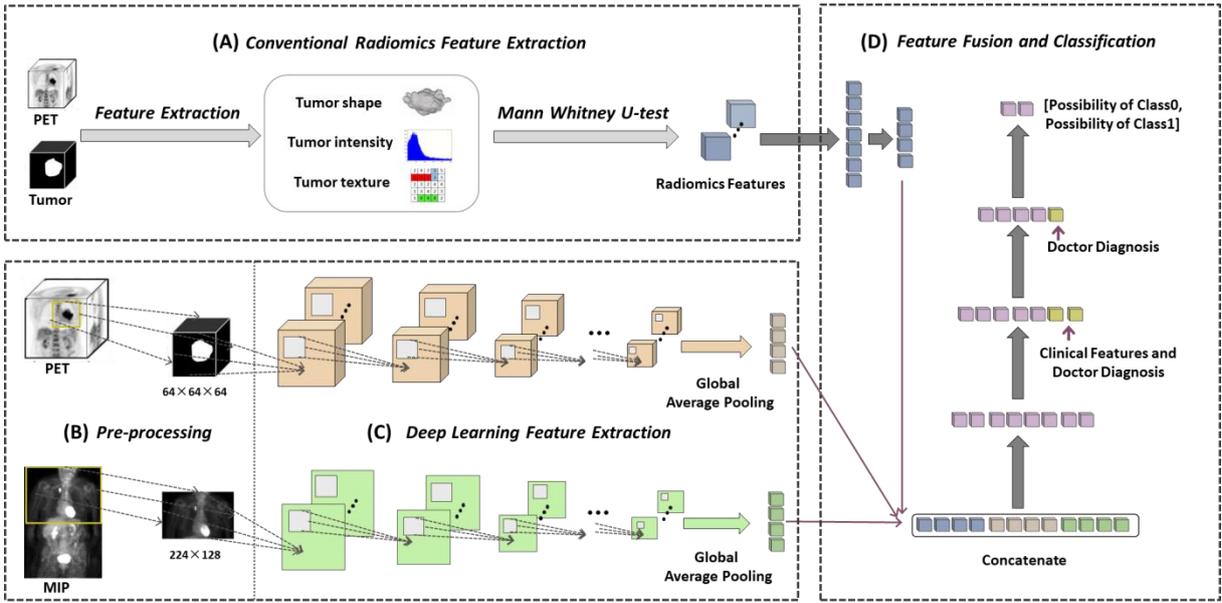


图 4.1 非小细胞肺癌淋巴结转移辅助诊断流程图

### 4.1.2 传统影像组学特征提取

传统影像组学特征提取及处理流程如图 4.1 (A) 所示。首先进行影像组学特征提取。Pyradiomics<sup>[65]</sup>是一个常用于在医学图像中提取影像组学特征的 Python 工具包。具体来说，用医生手动勾画的 PET 金标准图像描述原发肿瘤的位置信息，Pyradiomics 根据金标准获取 PET 图像中的一阶、二阶或高阶特征。总共提取出了 833 个影像组学特征，并且这些特征可以分为两类：（1）基于原始 PET 图像；（2）基于小波变换的图像。基于原始 PET 图像共提取了 105 个特征，包括肿瘤强度特征 18 个、形状特征 14 个以及纹理特征 73 个。其中肿瘤强度特征反映了单个体素的强度分布，不考虑体素之间的空间关系；形状特征描述了肿瘤的几何属性；纹理特征由不同矩阵计算体素的空间排列，反映了肿瘤的异质性，矩阵包括灰度共现矩阵（GLCM）、灰度相关矩阵（GLDM）、灰度游程矩阵（GLRLM）、灰度大小区域矩阵（GLSZM）、邻近灰度区域差异矩阵（NGTDM）五个类型<sup>[66]</sup>。此外，基于小波变换的图像共提取了 728 个特征，包括肿瘤强度特征 144 个、纹理特征 584 个。小波特征属于高阶特征，是通过对图像使用滤波操作、变换图像域来提取特征的，本实验使用的是 Coiflet<sup>[67]</sup>滤波器对图像进行了变换。但是传统的影像组学特征存在不稳定的问题，如果使用全部特征进行分析和诊断，不仅会增加模型诊断的难度和计算成本，还会导致模型性能无法进一步提升。因此在输入深度学习模型之前，

需要进行特征的筛除。

其次进行影像组学特征筛除。本研究使用的 PET 图像是由两种不同的 PET-CT 扫描仪进行成像的，不同设备的切片层厚、分辨率、像素空间等参数不同，成像的大小也不同，导致提取的影像组学特征存在差异，会影响分类器学习的过程，因此使用曼惠特尼 U 检验 (Mann Whitney U-test)<sup>[68]</sup>进行特征筛除。曼惠特尼 U 检验可以在本研究中用作剔除在两种扫描仪上差别显著的特征。具体来说，对两个不同扫描机器中的同一特征进行曼惠特尼 U 检验计算 U 值，U 值大于 0.05 则认为该特征在两种设备中没有显著差别，因此移除 U 值小于等于 0.05 的特征，保留其他服从同一分布的特征，以减轻不同设备对模型训练造成的影响。由于本文采用的是五折交叉验证，所以每一折训练数据保留下来的特征数量不同，测试数据则选择与训练数据相同的特征。尽管曼惠特尼 U 检验在一定程度上减轻了特征分布不统一的问题，但剔除过的影像组学特征仍然是高维度的，这会导致模型无法高效学习，因此还需要进行特征降维。

最后进行影像组学特征降维。全连接层经常被用于分类问题，它能够将学习到的特征表示映射到样本的标记空间，而多个全连接层的使用不仅能够增强模型非线性表达的能力，还能降低特征维度，将特征压缩为具有分类意义的向量。因此，对维度较高的传统影像组学特征，本文使用两个全连接层将其降至 64 个，与深度学习特征的维度匹配。最终得到的 64 个传统影像组学特征用于与深度学习特征融合，并在深度学习模型中进行分类，提高计算机辅助诊断的准确率。

#### 4.1.3 深度学习数据预处理

深度学习网络模型的图像输入分为两部分，如图 4.1 (B) 所示。一部分是利用原发肿瘤金标准剪裁 PET 图像生成的三维体。原发肿瘤的类型是医生选择合适治疗方案的依据，对原发肿瘤的识别与鉴定可使病人减少重复的检测、检查、影像扫描、活检等，并且原发肿瘤的状态极大影响了是否发生淋巴结转移的可能性。因此本研究以原发肿瘤作为深度学习模型的一个输入，利用卷积神经网络学习特征，对病人发生淋巴结转移的概率进行预测。在对 PET 图像预处理时，将原始 PET 图像与原发肿瘤金标准图像进行像素相乘形成感兴趣区域（即原发肿瘤以外的区域像素值为 0），再沿感兴趣区域边界裁剪成体素块，最后通过三线性插值将所有肿瘤图像调整为  $64 \times 64 \times 64$  大小，输入网络时进行图像归一化。

另一部分是经过预处理的正面半身最大强度投影图像。在 PET 图像中使用最大强度投影能有效观察到肿瘤或其他高代谢组织的位置和分布。而本文的研究对象是非小细胞肺癌患者的淋巴结转移与否，仅会发生在患者的上半身，所以只将正面半身最大强度投影图像作为深度学习模型的另一个输入，进行淋巴结转移辅助诊断。在对 PET 图像预处理时，将原始全身三维 PET 图像从正面进行最大强度投影，由于不同病人的 PET 切片层厚可能不同，所以通过双线性插值将图像调整为  $1\text{mm} \times 1\text{mm}$  截取半身图像，最后将所有半身图像调整为  $224 \times 128$  大小，输入网络时进行图像归一化。

#### 4.1.4 深度学习特征提取

深度学习特征提取网络包括两个部分，如图 4.1 (C) 所示。第一部分通过三维卷积神经网络提取原发肿瘤中与淋巴结转移相关的图像特征表示。本研究医疗数据的数量比较少，为了使网络参数减少、减轻网络训练的负担，所以选择已在自然图像分类中显示出优异性能的 VGG 网络<sup>[69]</sup>来提取特征。由于本研究中的原发肿瘤 PET 图像是三维的，因此采用了三维 VGG 网络结构。提取特征的过程可分为多个层次，每一层由两个三维卷积层组成，其中的卷积核大小为  $3 \times 3 \times 3$ ，个数保持不变。在两个层次之间，采用大小为  $2 \times 2 \times 2$  和步幅为 2 的最大池化进行下采样，降低特征映射的分辨率。每次下采样操作后，特征的大小将减少一半。此外，为了提高特征表示能力，每次下采样时特征通道数将增加一倍。采用批处理归一化 (Batch Normalization, BN) 层减小内部协方差偏移，ReLU 激活函数减小梯度消失的概率。本文共采取三次下采样操作，且在最后添加 dropout 层以消除过拟合，最终使用全局平均池化 (Global Average Pooling) 层得到 64 通道的三维原发瘤图像特征向量，用以之后的特征融合与分类。

第二部分通过二维卷积神经网络提取正面半身最大强度投影中与淋巴结转移相关的图像特征表示。为了扩大数据量，并且充分考虑全局信息，本文在第二部分中利用二维 VGG 网络结构提取半身最大强度投影图像特征。提取特征的过程同样可分为多个层次，每一层由两个二维卷积层组成，其中的卷积核大小为  $3 \times 3$ ，个数保持不变。在两个层次之间，采用大小为  $2 \times 2$  和步幅为 2 的最大池化进行下采样，降低特征映射的分辨率。同样采取三次下采样操作，并采用批处理归一化层、ReLU 函数、dropout 层以及全局平均池化层，最终得到 64 通道的二维正面半身最大强度投影图像特征向量，也用作之后的特征融合与分类。

原发性肿瘤图像提供了局部的、与肿瘤密切相关的深层次高维特征，半身正面最大强度投影图像提供了全局的、描述肿瘤位置与分布的低层次特征，因此本研究探究了将两者共同放入深度网络模型中训练是否能进一步提升模型分类的准确率，验证多种信息的输入是否有助于数据量扩大和相互的学习与促进。

#### 4.1.5 特征融合与分类

最后，将提取到的传统影像组学特征、深度学习特征、以及临床特征和医生诊断进行融合，得到分类结果，如图 4.1 (D) 所示。为了充分整合多源特征，将 64 维的传统影像组学特征向量、64 维的三维原发性肿瘤特征向量以及 64 维的二维正面最大强度投影特征向量用 Concatenate 操作进行连接，形成一个 192 维的特征向量，再通过全连接层进行特征融合与降维。部分临床特征能够为医生或计算机诊断提供重要依据，并且医生诊断也可以作为先验信息让网络学习得更好，所以在淋巴结转移分类前加入临床特征和医生诊断来提高模型准确度。为了保证临床特征向量(包括医生诊断，长度为 4)不被图像特征表示所淹没，当图像特征经过全连接层降至 16 维时与其连接，形成长度为 20 的输出向量，再通过全连接层进行融合。由于医生诊断较临床特征更具代表性和权威性，对网络的学习表征能力可能有更大的影响，所以本研究在输出分类结果前再一次将医生诊断向量与通道数为 4 的特征向量进行连接。最后，使用 softmax 激活函数输出淋巴结转移风险概率，使用 0.5 的概率阈值确定最终的二分类结果。

### 4.2 实验结果与分析

#### 4.2.1 训练过程

本研究共包括 121 例非小细胞肺癌患者，其中发生淋巴结转移的患者 19 例，未发生淋巴结转移的患者 102 例。在深度学习任务中遇到数据量较小的情况，常采用五折交叉验证来训练并测试模型，保证实验结果的泛化性。具体来说，在实验过程中，将数据平均分成 5 份，其中 4 份参与训练 1 份进行测试。这个过程重复五次，最终结果为五个测试集的综合结果。考虑到本文数据的类别极其不平衡，所以采用分层采样法划分数据集，确保训练集，测试集中各类别样本的比例与原始数据集的样本比例相同。比如原数据中，未发生转移和发生转移两类数据比例接近 5:1，通过分层采样切分的每个训练集

和测试集中两类的比例也约为 5:1。

为了进一步减轻数据不平衡带来的影响，本文在训练过程中采用带权重的二元交叉熵作为损失函数，即给少数类分配较大的权重，多数类分配较小的权重。通过最小化损失函数确定最优模型参数，最终用于测试。具体的损失函数可以写为：

$$\text{Loss}(X, y) = \frac{1}{N} \sum_i - [w_{pos} \cdot y_i \log p(Y=1|X) + w_{neg} \cdot (1 - y_i) \log p(Y=0|X)] \quad (4.1)$$

$$w_{pos} = \left( \frac{1}{pos} \times total \right) / 2 \quad (4.2)$$

$$w_{neg} = \left( \frac{1}{neg} \times total \right) / 2 \quad (4.3)$$

其中， $pos$ 为发生转移（正例）个数， $neg$ 为未发生转移（负例）个数， $total$ 为总病例个数。 $y_i$ 是淋巴结是否发生转移的金标准， $p$ 是网络输出的概率预测。

#### 4.2.2 实验环境

网络是使用基于 Python 3.8 版本的 TensorFlow2 实现的。本研究使用的 batch size 大小为 5，且利用 Adagrad 优化器来训练模型。为了防止过拟合，在 25 个 epoch 后早停了训练。该模型在 Ubuntu 18.04 系统上，使用 NVIDIA RTX2080Ti GPU 进行训练，并有 11 GB 显存可用。

#### 4.2.3 淋巴结转移辅助诊断结果

本研究进行了以下实验来证明该淋巴结转移辅助诊断模型的有效性：（1）本文方法与传统影像组学方法、医生诊断的整体性能比较；（2）消融实验，即分析本文提出的方法中每个新加入的组件带来的性能变化；（3）与医生诊断结果具体比较，即利用混淆矩阵展示模型的具体分类结果，并与医生诊断的混淆矩阵进行对比。这里使用最常用的分类评价指标进行说明，并绘制了受试者工作特性（ROC）曲线来直观地展示模型性能。

##### （1）与传统影像组学方法和医生诊断的性能比较

这里的传统方法是指使用 Xgboost 分类器对手工制作的特征进行分类并实现诊断。Xgboost 是机器学习中一种重要的集成思想，通过不断拟合决策树预测的残差来提升分

类结果，在各种领域的分类任务中展现了优异性能。这里的分类特征包括经过曼惠特尼 U 检验的传统影像组学特征、临床特征和医生诊断。模型性能如表 4.1 所示。在表 4.1 中，放射科医生对所有病例评估得出的 AUC 为 0.61，Precision 为 0.64，Recall 为 0.65，Accuracy 为 0.80。传统放射组学方法得出的 AUC 为 0.65，Precision 为 0.63，Recall 为 0.66，Accuracy 为 0.79，AUC 比医生诊断有所提高，但其他指标均与医生评估相近。本文的方法结合多种可学习特征，将 AUC 提高到 0.82，Precision 提高到 0.74，Recall 提高到 0.75，Accuracy 提高到 0.86，显著优于的医生诊断结果与传统的机器学习方法对手工特征的分类结果，性能最好。图 4.2 中对应的 ROC 曲线直观地展示了医生诊断、传统方法和本文方法的性能。

表 4.1 与传统影像组学方法和医生诊断对比实验结果

Model	Precision	Recall	Accuracy	AUC
Doctor	0.64	0.65	0.80	0.61
Xgboost based radiomics+clinic+doctor	0.63	0.66	0.79	0.65
<b>Ours</b>	<b>0.74</b>	<b>0.75</b>	<b>0.86</b>	<b>0.82</b>

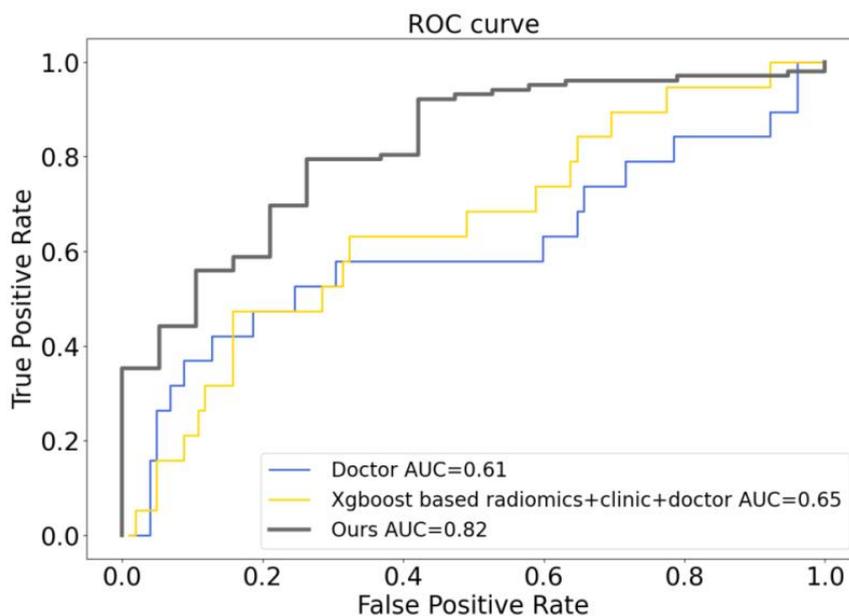


图 4.2 各方法的 ROC 曲线

## (2) 消融实验

为了验证本文方法每个步骤的有效性, 评估每个组件的性能, 进行了消融实验, 如表 4.2 所示。

表 4.2 淋巴结转移辅助诊断模型的消融实验

Model	Precision	Recall	Accuracy	AUC
3D deep learning	0.58	0.60	0.76	0.58
2D deep learning	0.60	0.66	0.71	0.65
3D+2D deep learning	0.60	0.64	0.75	0.72
3D+2D deep learning+radiomics(without U-test)	0.66	0.68	0.82	0.72
3D+2D deep learning+radiomics(with U-test)	0.69	0.69	0.83	0.78
3D+2D deep learning+radiomics+clinic	0.70	0.69	0.84	0.72
3D+2D deep learning+radiomics+doctor	0.71	0.74	0.84	0.73
3D+2D deep learning+ radiomics+clinic+doctor(front)	0.68	<b>0.76</b>	0.81	0.79
3D+2D deep learning+ radiomics+clinic+doctor(back)	0.67	0.70	0.82	0.80
<b>3D+2Ddeep learning+ radiomics+clinic+doctor(twice)</b>	<b>0.74</b>	0.75	<b>0.86</b>	<b>0.82</b>

在表 4.2 中, 仅使用深度学习提取三维原发肿瘤图像特征或二维正面半身最大强度投影图像特征进行淋巴结辅助诊断得到的 AUC 相对较低, 分别为 0.58 和 0.65。将二者结合得到 0.72 的 AUC 和 0.75 的 Accuracy, 虽然其他指标没有太大的提升, 但 AUC 超越了 0.7, 取得了重大突破, 这证明了将原发肿瘤图像特征与最大强度投影图像特征结合的有效性。加入传统的影像组学特征后, 虽然 AUC 没有进一步提升, 但其他指标均得到了大幅改善。如果使用的传统影像组学特征经过了曼惠特尼 U 检验, AUC 将提高到 0.78, Accuracy 将提高到 0.83, 可见曼惠特尼 U 检验和传统影像组学特征能提高模型性能。此外, 消融实验还将四种情况进行对比, 分别是: 只加入临床特征、只加入医

生诊断、同时加入临床特征与医生诊断、临床特征加在前一个全连接层而医生诊断加在最后一个全连接层。这四种情况的指标均有不同程度的波动和提升，但都没有达到最优性能，而是将医生诊断与临床特征同时加入前一个全连接层，再在最后一个全连接层加入一次医生诊断达到最好结果。最终本研究的方法，即结合三维原发肿瘤和二维正面半身最大强度投影的深度学习图像特征、经过曼惠特尼 U 检验的传统影像组学特征、临床特征以及两次医生诊断达到了最优的效果。该方法 AUC 为 0.82，Accuracy 为 0.86，Precision 和 Recall 分别为 0.74 和 0.75。图 4.3 中对应的 ROC 曲线展示了本文方法消融实验的结果。

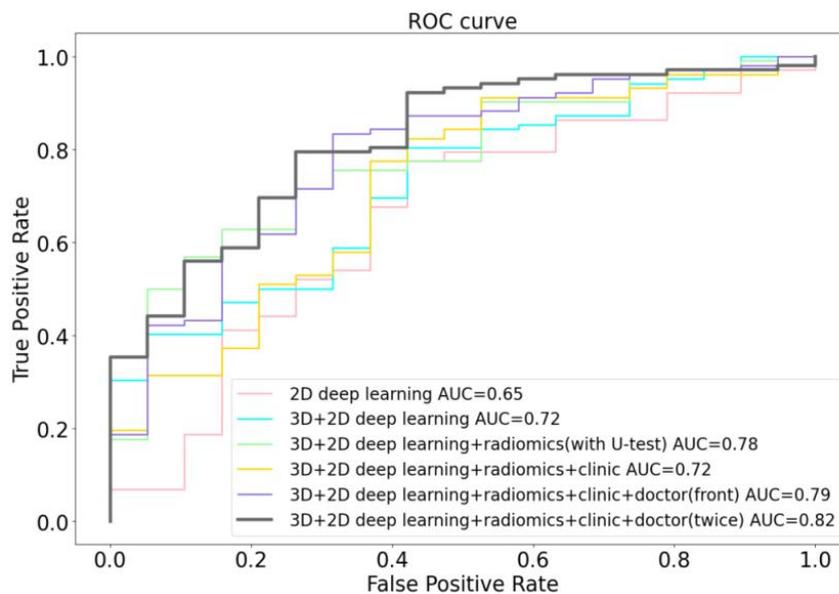


图 4.3 本文方法消融实验的 ROC 曲线

### (3) 与医生诊断比较

绘制混淆矩阵来进一步说明本文方法的临床意义，如图 4.4 所示。图 4.4 左边为本实验混淆矩阵，在发生淋巴结转移的病例中，本文方法预测正确 11 例，错误 8 例；在未发生转移的病例中，本文方法预测正确 93 例，错误 9 例。图 4.4 右边为医生诊断的混淆矩阵，在发生淋巴结转移的病例中，医生诊断正确 8 例，错误 11 例；在未发生转移的病例中，医生诊断正确 89 例，错误 13 例。可见，无论在有转移还是没有转移的情况下，本方法结果都优于医生诊断。

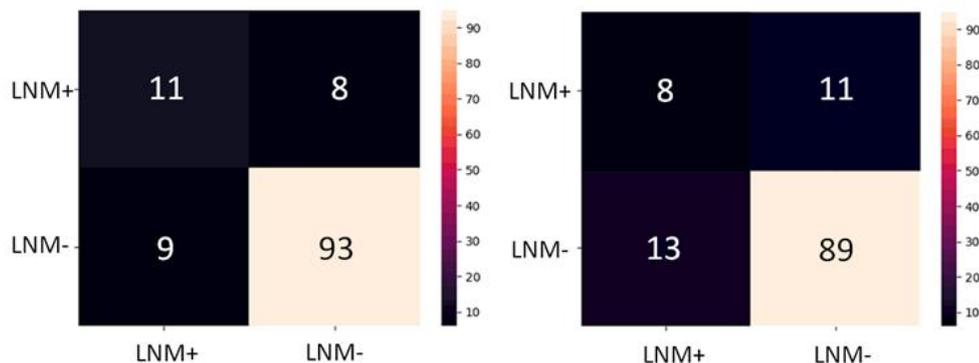


图 4.4 混淆矩阵

#### 4.2.4 讨论

表 4.1 和表 4.2 的结果表明，本文的方法具有更高的分类精度和 AUC。表 4.1 中，传统影像组学方法虽然能达到医生诊断的水平，但对参数和手工制作特征的依赖可能会限制其泛化性。卷积神经网络可以以端到端的方式实现较高的精度，因为深度学习方法可以自动挖掘与淋巴结转移相关的深层特征表示，降低对固定特征的依赖。表 4.2 显示，三维原发肿瘤特征与二维半身最大强度投影特征的融合能有效提升深度学习模型的性能，这表明局部精细特征和全局宏观特征能在诊断是否发生淋巴结转移的过程中产生共同学习和促进的作用。此外，加入经过曼惠特尼 U 检验的传统影像组学特征和先验临床特征（包括医生诊断）后，模型性能也有进一步的提升，证明了将多源输入融合进深度学习模型中能提升准确率，也说明了这些特征都是与淋巴结转移诊断相关的特征。最后，在与医生诊断的混淆矩阵的比较中，在有转移的情况中，本文的方法比医生少诊断错 3 例，这意味着本方法能比医生发现更多的淋巴结转移病人。如果能够及时地对存在转移潜在风险的患者进一步检查，就能及早发现异常，并且提前制定治疗方案，避免肿瘤发生进一步转移和扩散，从而提升患者的预后，改善病人的生活质量。

本研究提出的方法进一步改进了深度学习模型，从原发肿瘤捕获三维局部图像特征，从半身正面最大强度投影图像捕获二维全局图像特征，进一步与传统的手工特征和既往临床特征融合，并加入医生诊断，克服了标注训练数据不足的问题，减轻了对参数和手工特征的依赖，提高了模型的泛化性。

### 4.3 本章小结

在非小细胞肺癌淋巴结转移辅助诊断一章中，详细描述了淋巴结转移辅助诊断的研究流程、具体方法和实验结果。本章提出了一种新的基于深度学习的框架，用于非小细胞肺癌淋巴结转移辅助诊断。该框架可以成功地将先验临床知识、医生诊断、传统的手工特征和医学图像特征融合到深度学习模型中。通过这种端到端设计，所提出的框架实现了比专业医生或传统影像组学方法更好的分类性能，能帮助放射科医师提高诊断准确性，进而改善患者的生存状态。

## 5 两阶段非小细胞肺癌淋巴结转移辅助诊断

在上一章基于深度学习的端到端非小细胞肺癌淋巴结转移辅助诊断模型中，训练和测试过程均利用了医生勾画的肿瘤金标准及医生诊断结果，也就是说该方法是完全依赖医生判断的。虽然模型效果有较大的提升，但仅能起到辅助医生诊断的作用，并没有减少医生的工作量。为了能在保证模型性能没有较大损失的情况下减轻医生的负担，本研究提出一个两阶段的非小细胞肺癌淋巴结转移辅助诊断框架。该框架可以在第一阶段大致确定肿瘤区域来减少计算负担，在第二阶段利用分割任务消除测试时对医生注释的依赖，让计算机自动学习肿瘤区域及区域内特征，对非小细胞肺癌患者是否发生淋巴结转移进行辅助诊断，减轻医生工作量。

### 5.1 方法

#### 5.1.1 研究流程

本研究的主要目的是利用分割任务定位肿瘤区域，使辅助诊断模型在测试阶段不需要医生勾画的金标准，在保证诊断准确率的同时减少医生工作量，帮助临床医生提高工作效率。本文首先模仿医生诊断的思想，即直接将正面半身最大强度投影输入网络中学习，进行淋巴结转移诊断，但这种方式引入了较多无关信息，干扰了网络对于肿瘤转移相关特征的学习，表现出的性能不佳。因此本文采用两阶段的辅助诊断框架，在第一阶段，训练一个分割模型，在半身最大强度投影中大致确定肿瘤位置与范围，避免在第二阶段引入过多信息并减少计算机需要的算力；在第二阶段，训练两个以分割任务为引导的特征提取器，一个用三维分割-分类模型来提取三维肿瘤的与淋巴结转移相关的特征，另一个用二维分割-分类模型提取半身最大强度投影与淋巴结转移相关的特征，最终进行特征融合与分类。第二阶段中的测试图像均来自第一阶段的分割结果。研究的具体流程图如图 5.1 所示。

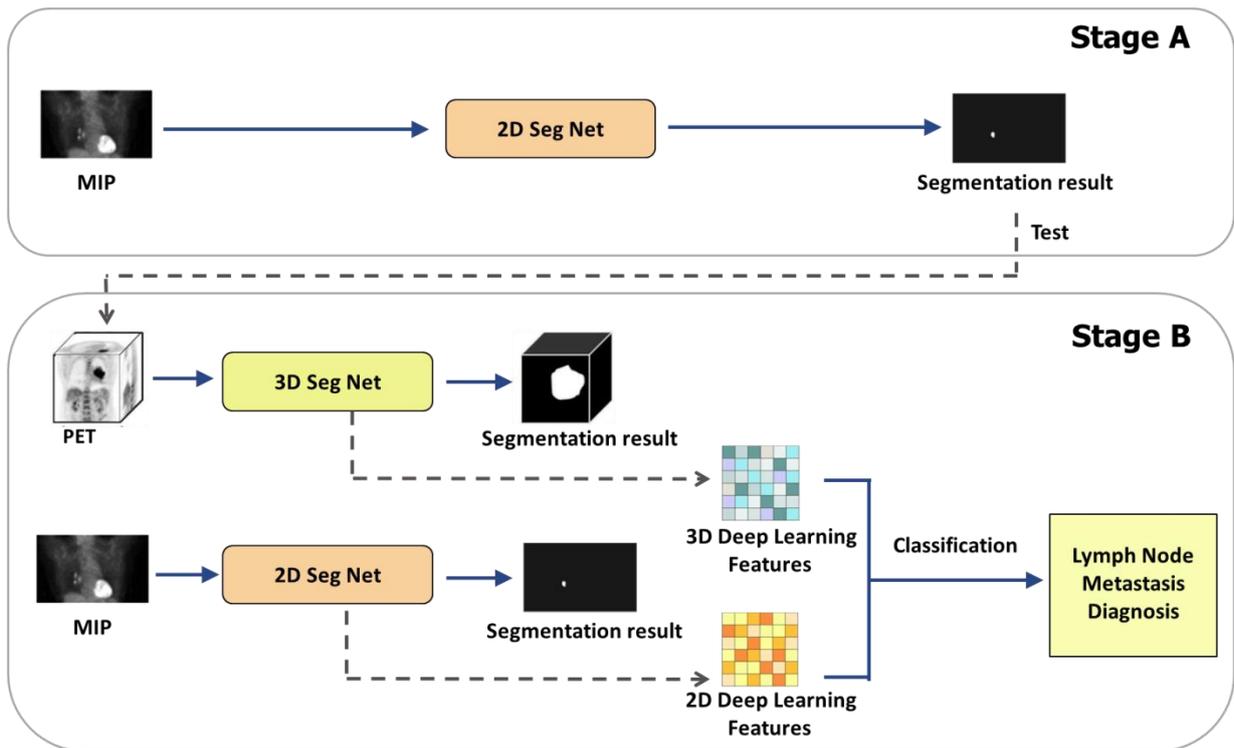


图 5.1 两阶段非小细胞肺癌淋巴结转移辅助诊断流程图

### 5.1.2 肿瘤分割

第一阶段是在正面半身最大强度投影图像中分割肿瘤，如图 5.1（Stage A）所示。Unet<sup>[70]</sup>是一种最常用的分割模型，它简单、高效、容易构建、可以从小数据集中训练，已广泛应用在医疗影像领域。因此本文利用 Unet 模型作为第一阶段分割网络的骨干，进行特征提取以及像素级别的分类。图 5.1(Stage A)中的 2D Seg Net 细节如图 5.2 所示。

在训练过程中，分割网络的输入为  $256 \times 128$  大小的正面半身最大强度投影图像。分割网络分为编码器和解码器两个部分。编码器用于特征的提取，可分为多个层次，每一层由两个二维卷积层组成，其中的卷积核大小为  $3 \times 3$ ，个数保持不变。在两个层次之间，采用大小为  $2 \times 2$  和步幅为 2 的最大池化进行下采样来降低特征映射的分辨率。每次下采样操作后，特征图的大小将减少一半，特征通道数将增加一倍。由于数据量小且分割任务相对简单，本研究共使用了三次下采样进行特征提取。解码器用于将特征映射大小恢复为原始图像大小，也可分为多个层次，每一层由两个二维卷积层组成，其中的卷积核大小为  $3 \times 3$ ，个数保持不变。与上采样过程相反，为了增加特征图的大小，使用

二维反卷积操作进行上采样。每次上采样操作后，特征映射的大小将增加一倍，特征通道数将减少一半。最后的特征映射分别表示对应像素被分类为背景和前景的概率，进而通过 softmax 函数获得肿瘤分割结果。

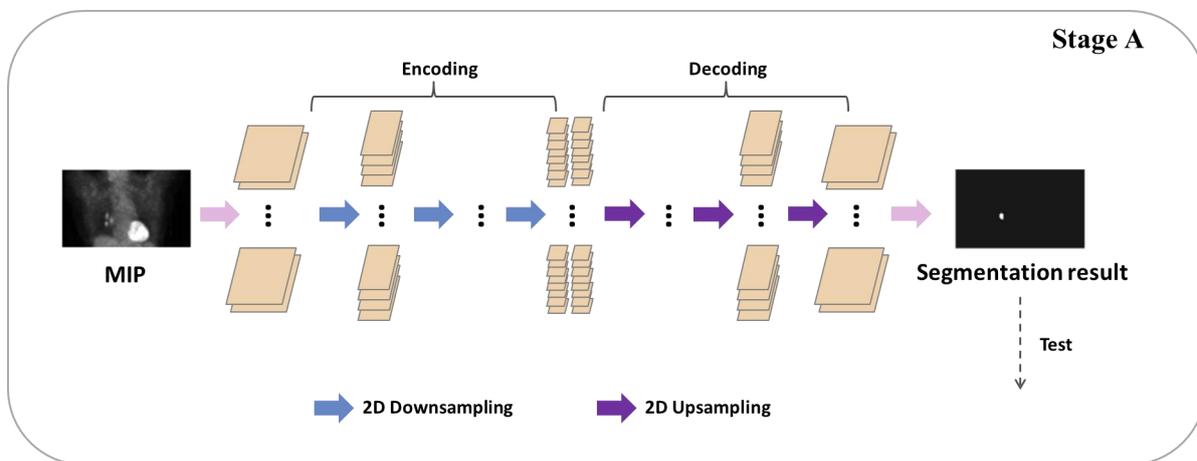


图 5.2 第一阶段肿瘤分割模型

这一阶段主要是为了大致定位肿瘤区域，使第二阶段的网络输入范围缩小，排除多余的干扰信息，并减轻网络运算的负担。这一阶段的分割网络在训练时需要医生勾画的肿瘤金标准来最小化损失函数进行优化，在测试时则不需要。测试得到的肿瘤分割结果将用于本研究的第二阶段。

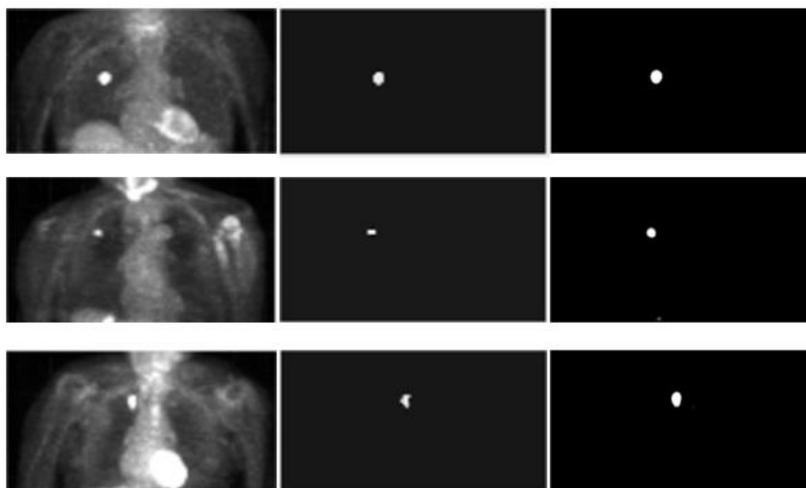


图 5.3 第一阶段分割结果

第一阶段肿瘤分割的部分结果如图 5.3 所示，第一列为三位患者的 PET 正面半身最大强度投影图像，第二列为医生勾画的肿瘤金标准的正面投影，第三列为模型对肿瘤的分割结果。该分割结果以一定的规则形成第二阶段肿瘤的测试集输入，具体规则见下一小节的描述。

### 5.1.3 以分割任务为引导的辅助诊断

第二阶段是以分割任务为引导的、在测试时不需要金标准图像的淋巴结转移辅助诊断，如图 5.1 (Stage B) 所示。在这一阶段，包括三个部分：(1)三维肿瘤特征提取；(2)二维正面半身最大强度投影特征提取；(3)特征融合与分类。具体模型如图 5.4 所示。

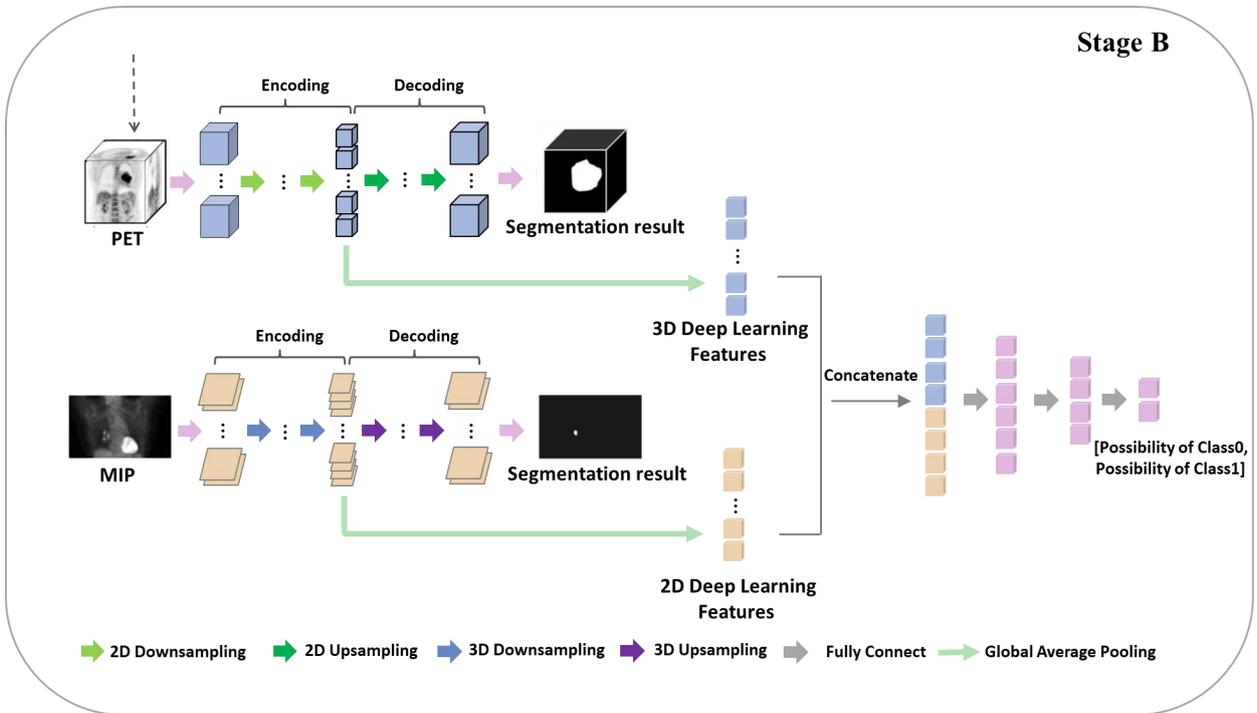


图 5.4 第二阶段以分割任务为引导的淋巴结转移辅助诊断模型

#### (1) 三维肿瘤特征提取

首先进行数据预处理。对于训练数据，利用医生勾画的肿瘤金标准剪裁原始 PET 三维图像，即在肿瘤金标准图像中，以肿瘤质心为中心，向上向下各选取 16 张切片（共 32 张切片），取 PET 图像对应位置的切片作为网络输入的三维体，如图 5.5 所示（左图为医生勾画的金标准，右图为 PET 图像）。选取相同数量的切片一是为了使图像输入尺

寸一致，二是为了增大在测试集上做相同操作时的误差容忍度，尽可能将整个肿瘤及周边区域输入网络中进行学习。因此，三维肿瘤训练数据的大小为  $128 \times 128 \times 32$ 。对于测试数据，将第一阶段生成的肿瘤结果作为金标准，用以剪裁原始 PET 图像。与训练数据的处理一样，从肿瘤中心上下选取 16 张切片进行裁剪，形成  $128 \times 128 \times 32$  大小的三维数据体，此过程不依赖医生的手动注释。对所有输入数据均做了归一化处理。

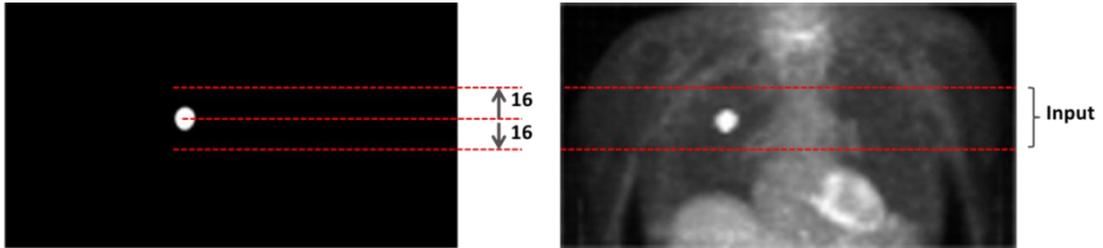


图 5.5 三维肿瘤数据输入处理方式示意图

其次进行特征提取。由于输入图像是  $128 \times 128 \times 32$  大小的立方体，所以采用三维的 Unet 结构进行特征提取。在编码过程中，使用多个层次的下采样进行特征提取，每一层由两个三维卷积层组成，其中的卷积核大小为  $3 \times 3 \times 3$ ，个数保持不变。在两个层次之间，采用大小为  $2 \times 2 \times 2$  和步幅为 2 的最大池化进行下采样来降低特征映射的分辨率。每次下采样操作后，特征映射的大小将减少一半，特征通道数将增加一倍，最终得到 128 个特征映射。这 128 个特征映射经过一个三维的全局平均池化层，形成一个 128 维的特征向量，用于是否发生非小细胞肺癌淋巴结转移的分类任务中。在这里，由于输入是包含肿瘤、肿瘤周围甚至许多无关区域的三维体，这将引入许多无关信息干扰网络的特征学习，因此使用一个分割任务（多个上采样）来引导网络关注重要区域（例如肿瘤）。在解码过程中，使用多个不同层次的转置卷积层进行上采样。这些不同层次的上采样块与下采样块相似，由两个三维卷积层组成，不同的是与上采样过程相反，使用三维反卷积操作增加特征图的大小，每次特征映射的大小增加一倍，特征通道数减少一半。同时，通过水平连接将每个层次的下采样特征转发给上采样块，这种方式可以收集用于肿瘤轮廓预测的细粒度细节，否则这些细节将在下采样路径中丢失。反过来，肿瘤区域通过反向传播过程在下采样过程中得到强调。最终上采样过程得到 2 通道的特征映射，分别表示对应像素被分类为背景和前景的概率，通过 softmax 函数获得肿瘤分割结果。

在训练时，分割网络需要医生勾画的肿瘤金标准来最小化损失函数进行优化。在测

试时，使用第一阶段得到的肿瘤大致区域作为输入，不需要具体的肿瘤金标准。分割任务引导了学习过程在训练和测试阶段关注语义上重要的区域，分类任务可用于是否发生淋巴结转移的预测，整体的三维分割-分类模型可对肿瘤中与淋巴结转移相关的特征进行有效提取。

### （2）二维正面半身最大强度投影特征提取

首先进行数据预处理。训练数据和测试数据都是将原始全身三维 PET 图像从正面进行最大强度投影，然后通过双线性插值将图像调整为  $1\text{mm} \times 1\text{mm}$  截取半身图像，最后将所有半身图像调整为  $224 \times 128$  大小，此过程不依赖医生的手动注释。对所有输入数据均做了归一化处理来提高图像质量。

其次进行特征提取。由于输入图像是  $224 \times 128$  大小的二维切片，所以采用二维的 Unet 结构进行特征提取。在编码过程中，使用多个层次的下采样进行特征提取，每一层由两个二维卷积层组成，其中的卷积核大小为  $3 \times 3$ ，个数保持不变。在两个层次之间，采用大小为  $2 \times 2$  和步幅为 2 的最大池化进行下采样来降低特征映射的分辨率。每次下采样操作后，特征映射的大小将减少一半，特征通道数将增加一倍，最终得到 128 个特征映射。这 128 个特征映射经过一个二维的全局平均池化层，形成一个 128 维的特征向量，用以是否发生非小细胞肺癌淋巴结转移的分类任务中。与三维肿瘤特征提取过程一样，为了使网络关注肿瘤区域，引入一个分割任务（多个上采样）。使用多个不同层次的转置卷积层进行上采样，每一层由两个二维卷积层组成。为了增加特征图的大小，使用二维反卷积操作进行上采样，每次特征映射的大小增加一倍，特征通道数减少一半。同时，通过水平连接将每个层次的下采样特征转发给上采样块，防止细节在下采样路径中丢失。反过来，肿瘤区域通过反向传播过程在下采样过程中得到强调。最终上采样过程得到 2 通道的特征映射，分别表示对应正面半身最大强度投影图像中像素被分类为背景和前景的概率，通过 softmax 函数获得肿瘤分割掩膜。

在训练时，分割网络需要医生勾画的肿瘤金标准来最小化损失函数进行优化，而测试时不需要。分割任务引导了学习过程在训练和测试阶段关注正面半身最大强度投影上重要的区域，即肿瘤区域，分类任务可用于是否发生淋巴结转移的预测，整体的二维分割-分类模型可对正面半身最大强度投影中与淋巴结转移相关的特征进行有效提取。

### （3）特征融合与分类

在三维肿瘤的特征提取中，网络提取了局部的、深层次的与肿瘤内部纹理或结构有

关的特征；而在二维正面半身最大强度投影特征提取中，网络提取了全局的、宏观的与肿瘤位置或分布有关的特征。二者分别提取到了一个 128 维的特征向量。为了充分整合三维和二维特征，将 128 维的三维肿瘤特征向量、128 维的二维正面半身最大强度投影特征向量用 Concatenate 操作进行连接，形成一个 256 维的特征向量，再通过多个全连接层进行特征融合与降维。最后，使用 softmax 激活函数输出淋巴结转移风险概率。

## 5.2 实验结果与分析

### 5.2.1 训练过程

本研究共包括 121 例非小细胞肺癌患者，其中发生淋巴结转移的患者 19 例，未发生淋巴结转移的患者 102 例。在第一阶段和第二阶段的实验过程中，均采用 5 折交叉验证的方法进行训练与测试。

在第一阶段中，训练过程仅包括分割任务，因此采用分割任务中常用的像素级交叉熵损失将预测的分割输出与肿瘤金标准进行比较。通过最小化损失函数确定最优模型参数，最终用于测试。具体的损失函数可以写为：

$$L_1 = \frac{1}{N} \sum_i - \left( y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{y}^{(i)}) \right) \quad (5.1)$$

$y^{(i)}$  表示样本的标签，即肿瘤金标准中像素值为 1 的是正类，像素值为 0 的是负类。 $\hat{y}^{(i)}$  为预测值的类别，即预测的分割结果中像素值为 1 的是正类，像素值为 0 的是负类。

在第二阶段中，训练过程包括两个任务，一个是三维肿瘤特征提取的分割-分类任务，一个是二维正面半身最大强度投影特征提取的分割-分类任务，最终二者提取的特征进行融合与分类，因此整个阶段共有三个损失函数来确定最优化模型的参数。具体的损失函数可以写为：

$$L_2 = L_{seg\_3D} + L_{seg\_2D} + L_{class} \quad (5.2)$$

$$L_{seg\_3D} = \frac{1}{N} \sum_i - \left( y_{3D}^{(i)} \log \widehat{y}_{3D}^{(i)} + (1 - y_{3D}^{(i)}) \log (1 - \widehat{y}_{3D}^{(i)}) \right) \quad (5.3)$$

$$L_{seg\_2D} = \frac{1}{N} \sum_i - \left( y_{2D}^{(i)} \log \widehat{y}_{2D}^{(i)} + (1 - y_{2D}^{(i)}) \log (1 - \widehat{y}_{2D}^{(i)}) \right) \quad (5.4)$$

$$L_{class} = \frac{1}{N} \sum_i - \left( y_{cls}^{(i)} \log \widehat{y_{cls}}^{(i)} + (1 - y_{cls}^{(i)}) \log (1 - \widehat{y_{cls}}^{(i)}) \right) \quad (5.5)$$

$L_{seg\_3D}$ 表示三维肿瘤分割任务的损失函数， $y_{3D}^{(i)}$ 是肿瘤金标准的像素级类别， $\widehat{y_{3D}}^{(i)}$ 是预测得到的分割结果的像素级类别，1为正类，0为负类； $L_{seg\_2D}$ 表示二维正面半身最大强度投影分割任务的损失函数， $y_{2D}^{(i)}$ 是最大强度投影中肿瘤金标准的像素级类别， $\widehat{y_{2D}}^{(i)}$ 是预测得到的分割结果的像素级类别，1为正类，0为负类； $L_{class}$ 表示淋巴结转移诊断的分类任务的损失函数， $y_{cls}$ 是发生淋巴结转移的标签， $\widehat{y_{2D}}^{(i)}$ 是预测得到标签，1为正类（发生淋巴结转移），0为负类（未发生淋巴结转移）。

### 5.2.2 实验环境

网络均是使用基于 Python 3.8 版本的 TensorFlow2 实现的。第一阶段使用的 batch size 大小为 5，且利用 Adam 优化器来训练模型，在 85 个 epoch 后早停了训练；第二阶段使用的设置一样，但在 25 个 epoch 后早停了训练。该模型在 Ubuntu 18.04 系统上，使用 NVIDIA RTX2080Ti GPU 进行训练，并有 11 GB 显存可用。

### 5.2.3 两阶段淋巴结转移辅助诊断结果

为了证明本研究提出的方法的有效性，进行了以下实验：（1）本文方法与传统的影像组学方法、医生诊断的整体性能比较；（2）消融实验，即所提出的两个阶段和分割任务的有效性；（3）与医生诊断结果比较，即利用混淆矩阵展示模型性能，和医生诊断作具体比较。

#### （1）与传统影像组学方法和医生诊断的性能比较

这里采用的传统方法是使用 Xgboost 分类器对手工制作的特征进行分类和预测。分类特征是经过曼惠特尼 U 检验的传统的影像组学特征、临床特征和医生诊断的融合特征。模型性能如表 5.1 所示。在表 5.1 中，放射科医生对所有病例评估得出的 AUC 为 0.61，Precision 为 0.64，Recall 为 0.65，Accuracy 为 0.80。传统的影像组学方法得出的 AUC 为 0.65，Precision 为 0.63，Recall 为 0.66，Accuracy 为 0.79，AUC 比医生诊断有所提高，但其他指标均与医生评估相近。本文使用两阶段法，在第一阶段大致分割和定位肿

瘤区域，用于第二阶段中三维肿瘤特征提取的输入，这一阶段能准确定位出 85%左右的肿瘤区域，只有 15%左右的肿瘤较难分割；在第二阶段以分割任务为引导提取三维肿瘤特征和二维正面半身最大强度投影特征，最后进行融合与分类。本文方法将 AUC 提高到 0.74，Precision 提高到 0.70，Recall 提高到 0.69，Accuracy 提高到 0.84，各项指标均优于医生诊断与传统方法，性能最好。图 5.6 中对应的 ROC 曲线提供了医生诊断、传统方法和本文方法的结果。

表 5.1 与传统影像组学方法和医生诊断对比实验结果

Model	Precision	Recall	Accuracy	AUC
Doctor	0.64	0.65	0.80	0.61
Xgboost based radiomics+clinic+doctor	0.63	0.66	0.79	0.65
<b>Ours</b>	<b>0.70</b>	<b>0.69</b>	<b>0.84</b>	<b>0.74</b>

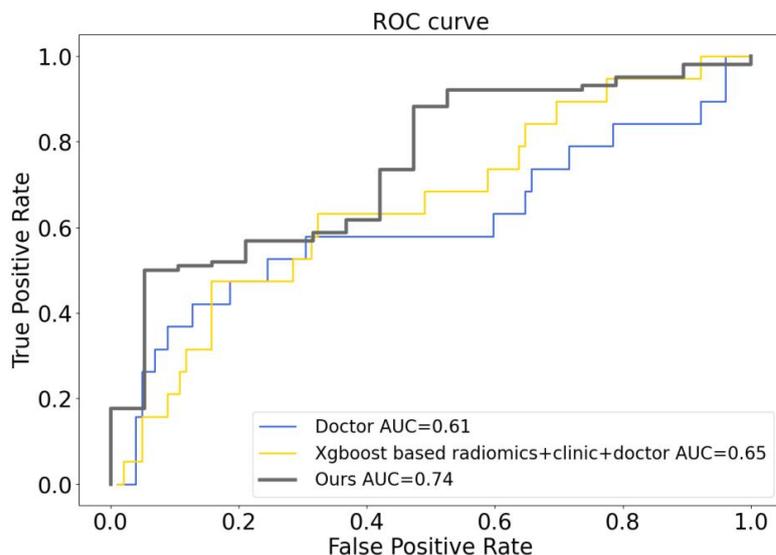


图 5.6 各方法的 ROC 曲线

## (2) 消融实验

为了验证本文方法两个阶段的有效性，进行了消融实验，如表 5.2 所示。如果没有第一阶段对肿瘤的大致分割与定位，则第二阶段中无法确定三维肿瘤进行特征提取的区域，所以只有对二维半身最大强度投影进行特征提取并预测的步骤。这一步骤包括两个

实验：（1）不使用医生注释的金标准，只输入最大强度投影图像让计算机自动学习发生淋巴结转移的概率，即单纯的分类任务。该过程取得的 AUC 相对较低，为 0.61，Precision、Recall、Accuracy 分别为 0.58、0.62、0.72；（2）仅在训练时使用医生注释的金标准，测试时不使用，即分割-分类任务。该过程取得的实验结果相较第一个实验的评价指标有所提升，AUC 为 0.65，Precision 为 0.63，Recall 为 0.64，Accuracy 为 0.79。证明了分割任务对分类任务有促进作用，能提高整体性能。

如果只对三维肿瘤进行特征提取，即利用第一阶段对肿瘤分割，大致进行定位，用于第二阶段中三维肿瘤特征提取的测试集选取，然后在第二阶段中只保留三维分割-分类模型对肿瘤进行特征提取，不对半身最大强度投影图像提取特征，得到的 AUC 为 0.70，Precision 为 0.59，Recall 为 0.66，Accuracy 为 0.69。

最终在加入第一阶段的情况下融合三维肿瘤和二维最大强度投影特征。此时对三维肿瘤的特征提取和二维正面半身最大强度投影的特征提取均采用分割任务为引导的方式，即使用分割-分类模型。最终取得的 AUC 为 0.74，Precision 为 0.70，Recall 为 0.69，Accuracy 为 0.84，性能最好。证明了两阶段、分割任务引导的有效性，也说明了三维局部的肿瘤特征与二维全局的最大强度投影特征的融合能提升诊断结果。图 5.7 中对应的 ROC 曲线展示了本文方法主要消融实验的结果。

表 5.2 两阶段淋巴转移辅助诊断模型的消融实验

Model	Precision	Recall	Accuracy	AUC
2D MIP	0.58	0.62	0.72	0.61
2D MIP with segmentation	0.63	0.64	0.79	0.65
3D tumor with segmentation	0.59	0.66	0.69	0.70
<b>Two-stage with segmentation(2D MIP+3D tumor)</b>	<b>0.70</b>	<b>0.69</b>	<b>0.84</b>	<b>0.74</b>

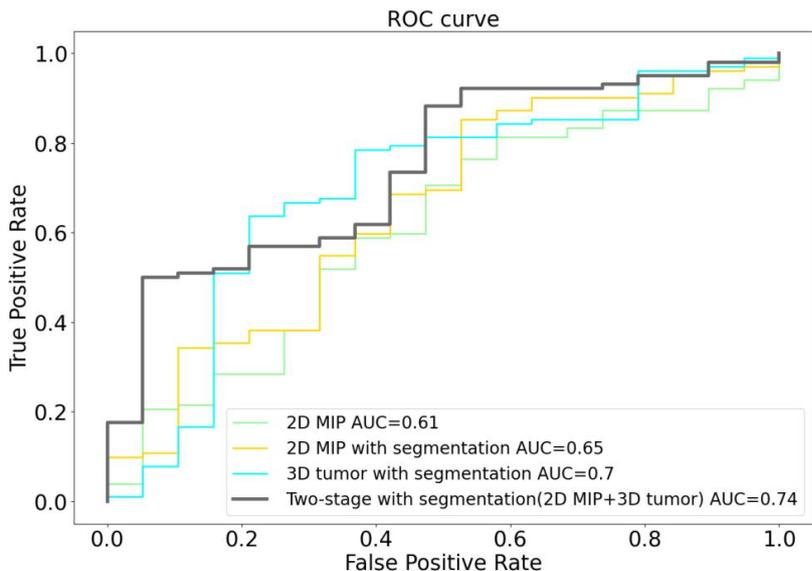


图 5.7 本文方法消融实验的 ROC 曲线

(3) 与医生诊断比较

绘制混淆矩阵来进一步说明本研究的临床意义，如图 5.8 所示。图 5.8 左边为本实验混淆矩阵，在发生淋巴结转移的病例中，本文方法预测正确 9 例，错误 10 例；在未发生转移的病例中，本文方法预测正确 93 例，错误 9 例。图 5.8 右边为医生诊断的混淆矩阵，在发生淋巴结转移的病例中，医生诊断正确 8 例，错误 11 例；在未发生转移的病例中，医生诊断正确 89 例，错误 13 例。可见，无论在有转移还是没有转移的情况下，本方法结果都优于医生诊断。

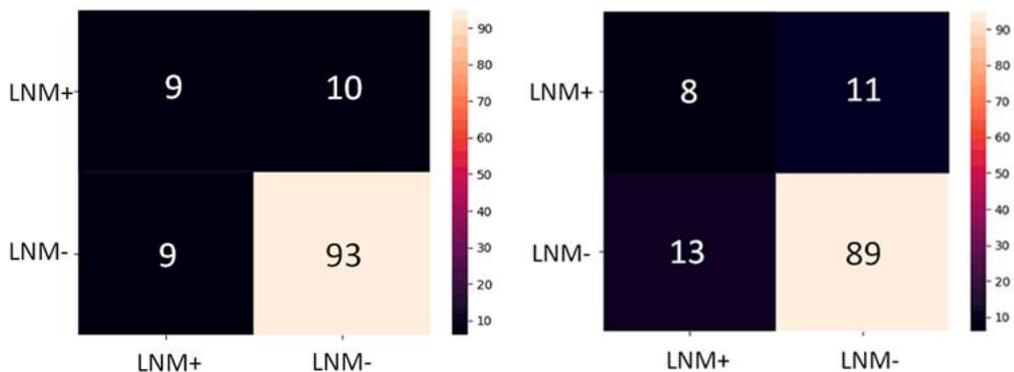


图 5.8 混淆矩阵

## 5.2.4 讨论

表 5.1 和表 5.2 的结果表明，本文的方法能在测试阶段不使用医生注释的金标准的情况下取得较好的结果。表 5.1 中，传统影像组学方法利用机器学习方法根据手工特征进行分类，手工特征是完全根据医生勾画的金标准提取出来的，虽然能达到医生诊断的水平，但对手工制作特征的依赖和对医生勾画的依赖会影响模型的泛化性能，并且不能减轻医生的工作量。而本研究使用的深度学习方法可以实现较高的精度，降低对固定特征的依赖，并且两阶段、以分割任务为引导的学习能在测试过程中不使用医生勾画的金标准，这将具有更大的现实意义。表 5.2 中，以分割任务为引导的学习方式比完全不使用医生勾画金标准的方式性能更好，这表明利用分割任务将网络的注意力放在肿瘤区域比直接对影像进行分析的效果好，也说明了肿瘤信息能为预测是否发生非小细胞肺癌淋巴结转移提供更多的判断依据。两阶段的方式能在测试阶段不使用医生勾画金标准的情况下提取三维肿瘤特征，最终与二维正面半身最大强度投影特征融合，提升了深度学习模型的性能。这也表明了局部精细特征和全局宏观特征能在诊断是否发生淋巴结转移的过程中产生共同学习和促进的作用。最后，在与医生诊断的混淆矩阵的比较中，在有转移的情况中，本文的方法比医生少诊断错 1 例，这一结果也具有十分重要的临床价值，能帮助医生提前选择合适的化疗方案，帮助提升患者生存率。

## 5.3 本章小结

在两阶段非小细胞肺癌淋巴结转移辅助诊断一章中，详细描述了利用两阶段的方式、引入分割任务进行淋巴结转移辅助诊断的研究流程、具体方法和实验结果。本章提出了一个两阶段的深度学习的框架，用于非小细胞肺癌淋巴结转移辅助诊断。该框架可以成功地将局部三维肿瘤特征和全局二维半身正面最大强度投影特征融合到深度模型中进行分类，并且这种自动化过程消除了测试阶段对精确的肿瘤勾画的依赖，同时获得了不错的整体性能。此外，实验结果还证明了肿瘤区域的特征对识别是否发生淋巴结转移有着重要作用。这种两阶段的设计能够减轻对参数和手工特征的依赖，提高模型的泛化性，减少医生的工作量，具有较大的现实意义。

## 6 总结与展望

### 6.1 总结

医学影像学能以无创的方式得到疾病的病理学特征，可以帮助医生进行临床诊断或决策，减少病人治疗的成本。而计算机辅助诊断能帮助医生寻找病灶、预测疾病发生的概率、分类病灶等。让计算机充分挖掘医学影像的特征，能打破传统方法的局限性，解决医生难以解决的问题。因此本文根据非小细胞肺癌患者的 PET-CT 图像提取特征，利用深度学习的方法对非小细胞肺癌患者的淋巴结转移进行了辅助诊断。

在非小细胞肺癌的淋巴结转移辅助诊断中，为了充分挖掘影像的可学习特征，我们提取了传统的影像组学特征、原发肿瘤的深度学习特征、正面半身最大强度投影的深度学习特征，并与临床特征和医生诊断进行了融合。具体来说，通过 Pyradiomics 提取原发肿瘤的传统影像组学特征，共 833 个特征，特征具有高维性并且由于设备不同，有些特征不能很好地帮助模型进行诊断。所以通过曼惠特尼 U 检验，剔除在两种设备中存在显著差异的特征，再用多个全连接层进行特征降维。此外，利用三维卷积神经网络提取原发肿瘤的内部高阶特征、二维卷积神经网络提取正面半身最大强度投影的全局特征，再将二者与传统影像组学特征相连。最后，用多个全连接层进行特征融合与降维，在输出结果前连接临床特征和医生诊断，得到发生淋巴结转移的概率。该模型的结果显示融合多种特征在非小细胞肺癌淋巴结转移辅助诊断方面具有良好的表现，并且无论是发生转移还是没有发生转移的情况下，模型的性能都优于医生的诊断。

为了进一步减轻医生工作量，减少勾画肿瘤金标准的时间，让计算机辅助诊断提供更大的便利，本文开发了一个两阶段、以分割任务为引导的模型，消除在测试阶段对肿瘤金标准的依赖，同时保证分类的准确性。具体来说，在第一阶段中对 PET 图像的正面半身最大强度投影进行肿瘤分割，对肿瘤大致定位，用于第二阶段中测试数据的处理。在第二阶段中，利用两个分割-分类模型对三维肿瘤区域和二维正面半身最大强度投影提取特征，其中分割任务能让网络关注在感兴趣区域（肿瘤区域）上，让网络自动学习与淋巴结转移相关的分类特征，此时三维肿瘤测试集通过第一阶段的预测结果进行确定。最后，连接三维肿瘤的深度学习特征和二维正面半身最大强度投影的深度学习特征，通过多个全连接层进行融合与分类，得到发生淋巴结转移的概率。整个辅助诊断的过程在测试时不需要医生对肿瘤区域的勾画，这能大大节省医生的时间，提升工作效率。该

模型的结果也显示出利用两阶段的方式能在测试不使用肿瘤金标准的情况下取得良好的表现，甚至能超越医生的诊断结果，这具有更大的现实意义，并且能为后续的工作提供更广的思路。

## 6.2 展望

尽管本文在利用深度学习进行非小细胞肺癌淋巴结转移辅助诊断方面取得了一些成果，但是仍然有很多方面有待进一步优化和探讨，主要有以下几个方面：

(1) 在医学影像的利用方面。本文的研究数据虽然是非小细胞肺癌患者的 PET-CT 图像，但由于部分 CT 影像存在问题，所以只使用了 PET 图像。CT 图像能够提供与 PET 不同的解剖结构信息，并且也有研究证明了在 CT 图像上进行淋巴结转移预测的可靠性。因此在之后的研究中可以利用 CT 影像，提取更多的与淋巴结转移相关的特征。

(2) 在多种数据的融合方面。本文的第一个研究主要在于探究不同的特征是否能对淋巴结转移的诊断起到积极作用，虽然结果证明了各种特征，包括传统影像组学特征、局部深度学习特征、全局深度学习特征以及临床特征对非小细胞肺癌淋巴结转移辅助诊断的有效性，但将各种特征的融合方式还是比较简单、直接，在之后的研究中可以对多种数据的融合方式进行改进，最大化发挥不同特征的效用。

(3) 在肿瘤的分割定位方面。本文的第二个研究主要目的是消除测试阶段对肿瘤金标准的依赖，因此设计了两阶段的方式，并且在第一阶段中对肿瘤进行大致分割定位。虽然第一阶段能准确地找到 85% 左右的肿瘤，但由于该分割任务是在二维半身最大强度投影中进行的，所以数据量较小、分割的准确性也有提高的空间。在之后的研究中可以改进肿瘤分割的输入或方式，提高肿瘤分割的准确性，减少第二阶段的选择误差和运算成本，进而提升模型的诊断性能。

(4) 数据不平衡问题。本文使用的 121 例非小细胞肺癌患者中只有 19 例发生了淋巴结转移，这是极其不平衡的。虽然在医学数据中经常存在数据不平衡的问题，但如果能通过一些策略或方法使模型能在数据不平衡时依然判断准确，则具有更大的意义。本文虽然利用带权重的交叉熵损失、数据的分层采样等方式减轻了数据不平衡的问题，但这些策略只能缓解一定程度的问题，因此在后续的研究中不但要进一步考虑数据不平衡的策略，还要从根本上缓解极度不平衡的问题，即寻找更多的可研究数据或扩充现有数据集。

最后，本文的研究思路首先是在训练和测试均使用医生勾画金标准，主要依靠金标准提取肿瘤与淋巴结转移相关的特征；其次进一步消除测试阶段对肿瘤金标准的依赖，通过两阶段、以分割任务为引导的方式使模型能在测试时自动寻找与肿瘤相关的特征，这一过程并不以损失过多的模型精度为代价。在后续的研究中，可以考虑在保证模型性能的情况下，训练和测试均不使用医生勾画的肿瘤金标准，让计算机自动学习并判断与淋巴结转移相关的特征，这将从真正意义上解放医生的双手，不用费时费力且重复地做勾画工作，具有更大的现实意义。

影像学与计算机技术的结合已在医学领域显现出十分重要的价值，提取的特征不但能反映出肿瘤的关键性病理信息，还能进一步进行组织学分类、生存期预测以及患者预后评估等。随着深度学习的迅速发展，越来越多关于计算机辅助诊断的研究涌现出来，可待研究和探讨的方法与思路也越来越多，未来必然还会有更多医工交叉的工作，为医生、患者带去更大的便利。

## 参考文献

- [1] 祝蕾,朱坤福.医学影像学的发展趋势探究[J].影像研究与医学应用,2018,2(02):84-85.
- [2] Doroshow J H, Kummar S. Translational research in oncology—10 years of progress and future prospects[J]. Nature reviews Clinical oncology, 2014, 11(11): 649-662.
- [3] Hatt M, Tixier F, Visvikis D, et al. Radiomics in PET/CT: more than meets the eye?[J]. Journal of Nuclear Medicine, 2017, 58(3): 365-366.
- [4] 刘延彤,李茂星,叶建军.计算机辅助诊断在医学影像诊断中的新进展[J].西北国防医学杂志,2006(06):450-451.DOI:10.16021/j.cnki.1007-8622.2006.06.025.
- [5] Nishikawa R M. Computer-aided detection and diagnosis[M]//Digital Mammography. Springer, Berlin, Heidelberg, 2010: 85-106.
- [6] Cheng J Z, Ni D, Chou Y H, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans[J]. Scientific reports, 2016, 6(1): 1-13.
- [7] Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer.[Internet][J]. 2021.
- [8] Wild C P, Weiderpass E, Stewart B W. World cancer report 2020[J]. Lyon: International Agency for Research on Cancer, 2020, 199: 512.
- [9] 国家肿瘤质控中心 . 2019 年中国癌症报告 [EB/OL]. (2019-03-19)[2021-02-28].<http://www.china-rt.cn/special/846.html>.
- [10] Clark S B, Alsubait S. Non small cell lung cancer[J]. 2020.
- [11] Zappa C, Mousa S A. Non-small cell lung cancer: current treatment and future advances[J]. Translational lung cancer research, 2016, 5(3): 288.
- [12] Funai K, Yokose T, Ishii G, et al. Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung[J]. The American journal of surgical pathology, 2003, 27(7): 978-984.
- [13] Brambilla E, Travis W D, Colby T V, et al. The new World Health Organization classification of lung tumours[J]. European respiratory journal, 2001, 18(6): 1059-1068.
- [14] Yun L U, Teng M A, Lei W, et al. Advances in Lymph Node Metastasis and Lymph Node Dissection in Early Non-small Cell Lung Cancer[J]. Chinese Journal of Lung Cancer,

- 2019, 22(8).
- [15] Kim M, Kim J K, Lee Y, et al. Diagnosis of lymph node metastasis in uterine cervical cancer: usefulness of computer-aided diagnosis with comprehensive evaluation of MR images and clinical findings[J]. *Acta Radiologica*, 2011, 52(10): 1175-1183.
- [16] Ha E J, Baek J H. Applications of machine learning and deep learning to thyroid imaging: where do we stand?[J]. *Ultrasonography*, 2021, 40(1): 23.
- [17] Vaquero J J, Kinahan P. Positron emission tomography: current challenges and opportunities for technological advances in clinical and preclinical imaging systems[J]. *Annual review of biomedical engineering*, 2015, 17: 385-414.
- [18] Griffeth L K. Use of PET/CT scanning in cancer patients: technical and practical considerations[C]//*Baylor University Medical Center Proceedings*. Taylor & Francis, 2005, 18(4): 321-330.
- [19] Jung H. Basic Physical Principles and Clinical Applications of Computed Tomography[J]. *Progress in Medical Physics*, 2021, 32(1): 1-17.
- [20] Whatmough C, Lamb C R. Computed tomography: principles and applications[J]. *Compendium*, 2006, 28(11): 789-800.
- [21] Tian M, Watanabe Y, Kang K W, et al. International consensus on the use of [18F]-FDG PET/CT in pediatric patients affected by epilepsy[J]. *European journal of nuclear medicine and molecular imaging*, 2021, 48(12): 3827-3834.
- [22] Humbert O, Bourg V, Mondot L, et al. 18F-DOPA PET/CT in brain tumors: impact on multidisciplinary brain tumor board decisions[J]. *European journal of nuclear medicine and molecular imaging*, 2019, 46(3): 558-568.
- [23] Borja A J, Hancin E C, Zhang V, et al. Potential of PET/CT in assessing dementias with emphasis on cerebrovascular disorders[J]. *European Journal of Nuclear Medicine and Molecular Imaging*, 2020, 47(11): 2493-2498.
- [24] Groheux D, Quere G, Blanc E, et al. FDG PET-CT for solitary pulmonary nodule and lung cancer: literature review[J]. *Diagnostic and interventional imaging*, 2016, 97(10): 1003-1017.
- [25] Prokop M, Shin H O, Schanz A, et al. Use of maximum intensity projections in CT

- angiography: a basic review[J]. *Radiographics*, 1997, 17(2): 433-451.
- [26]Beyaz F, Verhoeven R L J, Schuurbiens O C J, et al. Occult lymph node metastases in clinical N0/N1 NSCLC; a single center in-depth analysis[J]. *Lung Cancer*, 2020, 150: 186-194.
- [27]Katsumata S, Aokage K, Ishii G, et al. Prognostic impact of the number of metastatic lymph nodes on the eighth edition of the TNM classification of NSCLC[J]. *Journal of Thoracic Oncology*, 2019, 14(8): 1408-1418.
- [28]Takahashi Y, Takashima S, Hakucho T, et al. Diagnosis of regional node metastases in lung cancer with computer-aided 3D measurement of the volume and CT-attenuation values of lymph nodes[J]. *Academic Radiology*, 2013, 20(6): 740-745.
- [29]Bedrikovetski S, Dudi-Venkata N N, Maicas G, et al. Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: a systematic review and meta-analysis[J]. *Artificial Intelligence in Medicine*, 2021, 113: 102022.
- [30]Zhao J, Wang H, Zhang Y, et al. Deep learning radiomics model related with genomics phenotypes for lymph node metastasis prediction in colorectal cancer[J]. *Radiotherapy and Oncology*, 2022, 167: 195-202.
- [31]Gillies R J, Kinahan P E, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2): 563-577.
- [32]Xu X, Zhang H L, Liu Q P, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma[J]. *Journal of hepatology*, 2019, 70(6): 1133-1144.
- [33]Li F, Pan D, He Y, et al. Using ultrasound features and radiomics analysis to predict lymph node metastasis in patients with thyroid cancer[J]. *BMC surgery*, 2020, 20(1): 1-6.
- [34]Ji G W, Zhu F P, Zhang Y D, et al. A radiomics approach to predict lymph node metastasis and clinical outcome of intrahepatic cholangiocarcinoma[J]. *European radiology*, 2019, 29(7): 3725-3735.
- [35]Zhang H, Hu S, Wang X, et al. Prediction of Cervical Lymph Node Metastasis Using MRI Radiomics Approach in Papillary Thyroid Carcinoma: A Feasibility Study[J].

- Technology in Cancer Research & Treatment, 2020, 19: 1533033820969451.
- [36]Liu J, Sun D, Chen L, et al. Radiomics analysis of dynamic contrast-enhanced magnetic resonance imaging for the prediction of sentinel lymph node metastasis in breast cancer[J]. *Frontiers in Oncology*, 2019: 980.
- [37]Cong M, Feng H, Ren J L, et al. Development of a predictive radiomics model for lymph node metastases in pre-surgical CT-based stage IA non-small cell lung cancer[J]. *Lung Cancer*, 2020, 139: 73-79.
- [38]Sha X, Gong G, Qiu Q, et al. Discrimination of mediastinal metastatic lymph nodes in NSCLC based on radiomic features in different phases of CT imaging[J]. *BMC medical imaging*, 2020, 20(1): 1-8.
- [39]Qu J, Shen C, Qin J, et al. The MR radiomic signature can predict preoperative lymph node metastasis in patients with esophageal cancer[J]. *European radiology*, 2019, 29(2): 906-914.
- [40]Park V Y, Han K, Kim H J, et al. Radiomics signature for prediction of lateral lymph node metastasis in conventional papillary thyroid carcinoma[J]. *PloS one*, 2020, 15(1): e0227315.
- [41]Zhao B, Zhu H T, Li X T, et al. Predicting Lymph Node Metastasis Using Computed Tomography Radiomics Analysis in Patients With Resectable Esophageal Squamous Cell Carcinoma[J]. *Journal of Computer Assisted Tomography*, 2021, 45(2): 323-329.
- [42]LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *nature*, 2015, 521(7553): 436-444.
- [43]Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *nature*, 2017, 542(7639): 115-118.
- [44]Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. *Jama*, 2016, 316(22): 2402-2410.
- [45]Lee J H, Ha E J, Kim J H. Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT[J]. *European radiology*, 2019, 29(10): 5452-5457.
- [46]Hu Y, Su F, Dong K, et al. Deep learning system for lymph node quantification and

- metastatic cancer identification from whole-slide pathology images[J]. *Gastric Cancer*, 2021, 24(4): 868-877.
- [47]Wu Q, Wang S, Zhang S, et al. Development of a deep learning model to identify lymph node metastasis on magnetic resonance Imaging in Patients With Cervical Cancer[J]. *JAMA network open*, 2020, 3(7): e2011625-e2011625.
- [48]Zhao X, Wang X, Xia W, et al. A cross-modal 3D deep learning for accurate lymph node metastasis prediction in clinical stage T1 lung adenocarcinoma[J]. *Lung Cancer*, 2020, 145: 10-17.
- [49]Gao S J, Kim A W, Puchalski J T, et al. Indications for invasive mediastinal staging in patients with early non-small cell lung cancer staged with PET-CT[J]. *Lung Cancer*, 2017, 109: 36-41.
- [50]Zhong Y, She Y, Deng J, et al. Deep learning for prediction of N2 metastasis and survival for clinical stage I non-small cell lung cancer[J]. *Radiology*, 2022, 302(1): 200-211.
- [51]Zhou Z, Chen L, Sher D, et al. Predicting lymph node metastasis in head and neck cancer by combining many-objective radiomics and 3-dimensional convolutional neural network through evidential reasoning[C]//2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018: 1-4.
- [52]Dong D, Fang M J, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study[J]. *Annals of Oncology*, 2020, 31(7): 912-920.
- [53]Liu X, Zhang D, Liu Z, et al. Deep learning radiomics-based prediction of distant metastasis in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy: a multicentre study[J]. *EBioMedicine*, 2021, 69: 103442.
- [54]Zhang L, Wu X, Liu J, et al. MRI - Based Deep - Learning Model for Distant Metastasis - Free Survival in Locoregionally Advanced Nasopharyngeal Carcinoma[J]. *Journal of Magnetic Resonance Imaging*, 2021, 53(1): 167-178.
- [55]Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks[J]. *Pattern Recognition*, 2018, 77: 354-377.
- [56]Goodfellow I, Bengio Y, Courville A. *Deep learning*[M]. MIT press, 2016.

- [57]Waibel A, Hanazawa T, Hinton G, et al. Phoneme recognition using time-delay neural networks[J]. IEEE transactions on acoustics, speech, and signal processing, 1989, 37(3): 328-339.
- [58]Zhang W, Tanida J, Itoh K, et al. Shift-invariant pattern recognition neural network and its optical architecture[C]//Proceedings of annual conference of the Japan Society of Applied Physics. 1988: 2147-2151.
- [59]LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4): 541-551.
- [60]Rafferty J, Shellito P, Hyman N H, et al. Practice parameters for sigmoid diverticulitis[J]. Diseases of the colon & rectum, 2006, 49(7): 939-944.
- [61]Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[C]//Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2011: 315-323.
- [62]Lin M, Chen Q, Yan S. Network in network[J]. arXiv preprint arXiv:1312.4400, 2013.
- [63]Fawcett T. An introduction to ROC analysis[J]. Pattern recognition letters, 2006, 27(8): 861-874.
- [64]周志华. 机器学习[M].北京: 清华大学出版社, 2000:30-35.
- [65]Van Griethuysen J J M, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype[J]. Cancer research, 2017, 77(21): e104-e107.
- [66]Haralick R M, Shanmugam K, Dinstein I H. Textural features for image classification[J]. IEEE Transactions on systems, man, and cybernetics, 1973 (6): 610-621.
- [67]Salazar Á, Lorduy H. Approach to wavelet multiresolution analysis using Coiflets and a two-wave mixing arrangement[J]. Optics communications, 2008, 281(11): 3091-3098.
- [68]Sundjaja J H, Shrestha R, Krishan K. McNemar And Mann-Whitney U Tests[J]. 2020.
- [69]Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [70]Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

## 致 谢

三年的时光一晃而过，我的研究生生活至此也告一段落了。初入校园的欣喜、挑灯夜读的艰辛、休闲时刻的欢愉、独自科研的苦闷，每一桩每一件都仿佛昨日发生，让我历历在目。这三年的生活经历让我褪去了幼稚轻率的心态，在一点一滴的沉淀中提升自己，变得沉稳和宽阔；让我在不停地学习与感受中收获知识、友情，明白了智慧是行动和精神的武装，也明白了真挚的友情有多么弥足珍贵；让我开拓了眼界、增长了见识，接触到了不曾涉足的领域，也体会到了大千世界的包容与深奥。回望这令我挥洒心血和汗水的实验室，这一下课就人满为患的食堂，这给予我温暖和安全感的寝室，这虽然不大却承载了我们所有回忆的校园，我的心绪便再一次被拨动，我的感激之情也油然而生。

首先我要感谢我的导师何江萍副教授，在他的严格要求下，我才能约束自己，让自己全身心投入在科研中；在他的关心鼓励下，我才能突破自己，在每次想要退缩的时候坚守初心，砥砺前行；在他的鼎力支持下，我才能开阔视野，去外面的世界接触优秀的人才和力量；在他的辛勤耕耘下，实验室才能不断地发展，我们的科研团队才能不断地壮大。何老师对待科研的一丝不苟和对待工作的兢兢业业，无一不时时刻刻影响着我，让我能不自觉地提高对自己的要求，追求工作和生活的尽善尽美。我还要感谢实验室的其他老师，感谢韩金仓老师对我的生活和学习的关心，让我在迷茫的时刻找到方向；感谢李兵老师对我们学习上的教导，让我对不理解的知识有了更好的掌握；感谢丁晓阳老师对我们实验室的支持，让我们的科研团队有信心做出更多的研究成果。

其次我要感谢我的学校，为我们提供了很好的资源和工作场地，包括信息工程学院的任课老师们。他们在那三尺讲台上为我们传授无尽的知识，解答我们的疑惑，培养我们的能力，让我受益匪浅。他们精益求精的态度，将成为我未来道路中的信仰；他们孜孜不倦的品质，将成为我学习和模仿的表率；他们豁达开朗的心态，将成为我遇到问题时的指引。我深深地感谢老师们对我的帮助和影响，我相信在未来的日子依然能用到老师们教的知识 and 道理。

最后我要感谢一直陪伴在我身边的家人、朋友。感谢我的室友张杉杉、王文瑞、宋文琴，是她们在我最彷徨的时候给予我陪伴与帮助；感谢我的师兄师姐、师弟师妹们，是他们对我的科研和工作提出建议，不断鞭策我成长；感谢每一个鼓励过我、关心过我的朋友，让我感受到世界的温暖；感谢我的父母、家人，给我最坚实的力量和最强大的后盾，让我在前进的道路上能无所顾忌、勇往直前！