

分类号 TP393.1  
U D C

密级 公开  
编号 10741



## 硕士学位论文

论文题目 基于半监督深度学习的非小细胞肺癌淋巴  
结转移诊断研究

研究生姓名: 吴祎璠

指导教师姓名、职称: 何江萍 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: 2024年5月31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：吴祥博 签字日期：2024.5.31

导师签名：何江华 签字日期：2024.5.31

## 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，          （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：吴祥博 签字日期：2024.5.31

导师签名：何江华 签字日期：2024.5.31

# **Diagnosis of Lymph Node Metastasis in Non-small Cell Lung Cancer Based on Semi-supervised Deep Learning**

**Candidate : Yifan Wu**

**Supervisor: Jiangping He**

## 摘 要

肺癌是一种全球范围内重要的癌症类型，其发病率和死亡率一直居高不下。尤其是非小细胞肺癌（Non-small Cell Lung Cancer, NSCLC）患者占肺癌病例的80%-85%。对于非小细胞肺癌患者来说，淋巴结转移在治疗方案制定中起着至关重要的作用。近年来，医学影像分析领域中深度学习技术的应用呈上升趋势，可以用于辅助淋巴结转移的分期工作。PET/CT作为一种多模式成像系统，结合了代谢和形态学信息，被广泛应用于分期评估。然而，获取和标注医学影像数据的成本较高，而且并不是所有患者都能接受被认为是分期金标准的病理检测。因此，影像诊断成为另一种选择，但其准确性受到多重因素的影响。目前已有一些研究提出了解决噪声标签问题的方法，但在医学领域中，对于联合噪声标签训练问题的研究较少。因此，本研究旨在运用半监督深度学习技术来帮助诊断患有非小细胞肺癌的患者是否存在淋巴结转移，为淋巴结转移的分期评估提供更有效的辅助工具。主要研究内容如下：

(1) 本文首先介绍了一种改进自 Mean Teachers 方法的师生模型，名为双学生师生模型（Dual Student with Teacher and Correction, DSTC），该模型包含两个学生网络和纠正网络，用于淋巴结转移预测。经过预处理的 PET/CT 肿瘤块作为输入，学生网络以病理和影像标签为目标，提取图像的深层次特征并预测淋巴结转移。在两个学生网络之间引入单向学习方法，以病理标签的指导下辅助影像标签预测淋巴结转移。教师网络的权重采用学生网络的指数移动平均值，通过一致性学习使得学生网络的预测与之相似。尽管放射科医生根据影像检查结果及相关疾病信息已经提供了诊断，为后续治疗和疗效评估提供了重要依据，但其准确度不如病理诊断。因此，在模型中加入了半监督策略纠正网络，以病理标签为指导，减少影像标签潜在的错误，纠正后的标签将用于学生网络的训练。该模型在具有 328 例影像标签的非小细胞肺癌患者的 PET/CT 数据集上进行了五折交叉验证，其中只有 112 例拥有病理标签。在测试集上，该模型实现了准确率为 0.88 和 AUC 为 0.8，优于放射科医生的诊断和其他方法的预测，证实了本研究的模型能够有效预测非小细胞肺癌淋巴结转移。

(2) 由于医生手动勾画肿瘤区域的成本较高，本文还使用了一种无需肿瘤分割的多尺度图像淋巴结转移诊断模型。该模型以 PET/CT 图像的二维正面最

大密度投影和三维半身图像作为输入，构建了一个多尺度、多模态的半监督深度学习模型，能够利用未带有病理标签的数据来生成伪标签并参与训练过程。该模型采用了两个基于 ResNet-18 的特征提取网络，用于提取二维和三维深度特征，每个特征提取网络包含 PET、CT 和堆叠三个分支。在堆叠的分支中引入了融合模块，以融合不同模态和上下文信息。最终将两个维度图像的特征进行拼接，用于预测淋巴结转移。实验结果表明，该模型在淋巴结转移分类研究中有效减少了对医生手动标注病灶区域的依赖，实现了分类准确率为 0.86。

淋巴结转移作为疾病诊断过程中的关键环节，其准确性直接关系到后续治疗方案的制定以及患者预后的评估。因此，本文在淋巴结转移诊断领域所展开的研究，有望为医生提供更准确、更可靠的诊断依据，有助于改善患者的治疗效果和生活质量。

**关键词：**半监督深度学习 非小细胞肺癌 淋巴结转移 PET/CT

## Abstract

Lung cancer is a significant type of cancer on a global scale, with high incidence and mortality rates. In particular, Non-small Cell Lung Cancer (NSCLC) patients account for 80%-85% of lung cancer cases. For NSCLC patients, lymph node metastasis plays a crucial role in treatment planning. In recent years, the application of deep learning techniques in the field of medical image analysis has been on the rise, providing assistance in staging lymph node metastasis. PET/CT, as a multimodal imaging system combining metabolic and morphological information, is widely used for staging evaluation. However, the cost of obtaining and annotating medical imaging data is high, and not all patients can undergo pathological testing, considered the gold standard for staging. Therefore, imaging diagnosis becomes an alternative, but its accuracy is influenced by multiple factors. While some studies have proposed methods to address noisy label issues, research on joint noisy label training problems in the medical field is limited. Therefore, this study aims to use semi-supervised deep learning approach to help diagnose whether patients with NSCLC have lymph node metastasis, and to provide a more effective auxiliary tool for the staging assessment of lymph node metastasis. The main research contents are as follows:

- (1) This paper first proposes a Dual Student with Teacher and

Correction (DSTC) model, a modification of the Mean Teachers method, with two student networks and a correction network for predicting lymph node metastasis. Pre-processed PET/CT tumor patches serve as input, with the student networks targeting pathological and imaging labels to extract deep features from images and predict lymph node metastasis. A unidirectional learning method is introduced between the two student networks, to assist in predicting lymph node metastasis based on imaging labels under the guidance of pathological labels. The weights of the teacher network are updated using the exponential moving average of the student networks' weights, ensuring similarity between the student networks through consistency learning. Despite radiologists providing diagnosis based on imaging results and relevant disease information, which serves as important evidence for subsequent treatment and efficacy assessment, its accuracy does not match that of pathological diagnosis. Therefore, a semi-supervised strategy correction network is added to the model, guided by pathological labels, to reduce potential errors in image labels. The corrected labels are then used for the training of the student network. The model is tested on a dataset of PET/CT images from 328 NSCLC patients, of which only 112 had pathological labels, through five-fold cross-validation. On the test set, the model achieves an accuracy of 0.88 and an AUC of 0.8, outperforming radiologists' diagnosis and other prediction methods, confirming the model's ability to effectively

predict lymph node metastasis in NSCLC.

(2) Due to the high cost of manual delineation of tumor regions by physicians, this paper also uses a multi-scale image lymph node metastasis diagnosis model without tumor segmentation. The model constructs a multi-scale, multi-modal semi-supervised deep learning model with two-dimensional frontal maximum density projections and three-dimensional half-body images of PET/CT images as inputs. It is able to utilize data without pathological labels to generate pseudo-labels and participate in the training process. The model employs two ResNet-18-based feature extraction networks for extracting two-dimensional and three-dimensional deep features. Each feature extraction network contains three branches of PET, CT, and stacking. A fusion module is introduced in the stacked branches to fuse different modal and contextual information. Finally, the features of the two dimensional images were spliced together to predict lymph node metastasis. The experimental results show that the model effectively reduces the dependence on doctors manually labeling the lesion area in the classification of lymph node metastasis, and achieves a classification accuracy of 0.86.

Lymph node metastasis is a key link in the process of disease diagnosis, and its accuracy is directly related to the formulation of subsequent treatment plans and the evaluation of patient prognosis.



Therefore, the research in the field of lymph node metastasis diagnosis is expected to provide doctors with more accurate and reliable diagnostic basis, which will help improve the treatment effect and quality of life of patients.

**Keywords:** Semi-supervised deep learning; Non-small cell lung cancer; Lymph node metastasis; PET/CT

# 目 录

<b>1 绪论</b> .....	1
1.1 研究背景与意义 .....	1
1.1.1 非小细胞肺癌 .....	1
1.1.2 PET/CT .....	2
1.1.3 非小细胞肺癌淋巴结转移诊断的意义 .....	3
1.2 研究现状 .....	4
1.2.1 监督学习方法 .....	4
1.2.2 半监督学习方法 .....	7
1.2.3 噪声标签学习方法 .....	8
1.3 研究内容与创新 .....	10
1.4 本文组织结构 .....	11
<b>2 相关理论</b> .....	13
2.1 卷积神经网络 .....	13
2.2 ResNet 网络 .....	14
2.3 半监督学习 .....	16
2.4 病理诊断和影像诊断 .....	17
2.5 本章小结 .....	17
<b>3 基于 DSTC 的半监督淋巴结转移诊断</b> .....	18
3.1 研究流程 .....	18
3.2 实验数据 .....	19
3.2.1 数据介绍 .....	19
3.2.2 数据预处理 .....	25
3.3 深度学习模型架构 .....	21
3.3.1 单向学习 .....	24
3.3.2 一致性学习 .....	25
3.4 总损失函数 .....	26
3.5 实验 .....	26

3.5.1 实验环境及训练策略 .....	26
3.5.2 评估指标 .....	26
3.5.3 实验结果分析 .....	28
3.6 讨论 .....	33
3.7 本章小结 .....	35
<b>4 基于多尺度多模态融合网络的无肿瘤分割淋巴结转移诊断 .....</b>	<b>35</b>
4.1 研究流程 .....	35
4.2 实验数据及预处理 .....	36
4.3 深度学习模型架构 .....	36
4.4 实验 .....	38
4.4.1 训练过程和实验环境 .....	38
4.4.2 实验结果分析 .....	39
4.5 讨论 .....	42
4.6 本章小结 .....	43
<b>5 总结与展望 .....</b>	<b>44</b>
5.1 总结 .....	44
5.2 展望 .....	45
<b>参考文献 .....</b>	<b>46</b>
<b>致谢 .....</b>	<b>51</b>

# 1 绪论

## 1.1 研究背景与意义

### 1.1.1 非小细胞肺癌

癌症是一种由细胞分裂过程中发生的病理生理改变引起的严重疾病，已经成为全球造成大量死亡的主要原因之一<sup>[1][2]</sup>。尤其是肺癌，在世界范围内是一种重要的癌症类型，给社会带来了严重负担。依据 2022 年全球癌症统计报告显示，肺癌患病比例达到了 11.4%，而致死率则占到 18%<sup>[3]</sup>。每年都有大量人被诊断为肺癌，并且它也是导致许多人死亡的主要原因之一。在我国，肺癌同样是常见的恶性肿瘤，根据国家癌症中心于 2024 年初发布的数据显示，2022 年我国肺癌的发病率和死亡率都排名第一<sup>[4]</sup>。

依据显微镜下肿瘤细胞的形态将肺癌区分，可以分为小细胞肺癌（Small Cell Lung Cancer, SCLC）与非小细胞肺癌（Non-small Cell Lung Cancer, NSCLC）两种类型。小细胞肺癌是一种高度侵袭性的肺癌，其特点包括细胞体积较小、核与胞浆比例高，生长迅速且早期就会转移到其他部位。由于其侵袭性和早期转移的特点，小细胞肺癌通常在诊断时已经扩散到其他部位，这对治疗策略的选择产生了影响。对于小细胞肺癌而言，其治疗方式主要采用化学治疗和放射治疗，而在疾病晚期且癌细胞已经播散的情形中，手术切除的疗效甚微。大部分肺癌患者（大约 80%至 85%）被诊断患有非小细胞型，这一类别在众多肺癌种类中占据主导地位<sup>[5]</sup>。非小细胞肺癌又可以进一步细分为多个亚型，包括腺癌、鳞癌和大细胞癌等。这些亚型在细胞学和组织学上有所不同，对于治疗和预后评估都具有重要性。

在非小细胞肺癌中，淋巴结转移是最常见的转移途径之一<sup>[6]</sup>。在早期未发生远处转移的患者中，发生区域淋巴结转移的比例为 13.0%-40.3%<sup>[7]</sup>。准确评估淋巴结转移对制定非小细胞肺癌患者的治疗计划至关重要<sup>[6]</sup>。评估淋巴结转移的方法包括 CT、PET/CT、淋巴结活检等，其中淋巴结活检被认为是诊断淋巴结转移的金标准。对于已确定存在淋巴结转移的患者，治疗策略通常包括化疗、放疗和可能的手术切除。淋巴结转移与患者预后密切相关，通常与较差的生存率和较高的疾病复发率相关。因此，准确评估淋巴结转移状态可以帮助医生制定个体化的治疗和随访计划，提高患者的治疗效果和生存率。

## 1.1.2 PET/CT

正电子发射断层扫描技术（Positron Emission Tomography, PET）是一种无创成像方式，通过注射放射性化合物和检测放射性示踪剂的分布，提供生理信息<sup>[8]</sup>。在 PET 扫描过程中，放射性同位素被标记到参与血流或代谢过程的化合物上，随后注射到患者体内。这些同位素在体内会发生衰变，释放出正电子，并与组织中的负电子相遇，产生两个方向相反、能量为 0.511MeV 的  $\gamma$  光子。这些光子沿着相反路径到达探测器，如果它们在短时间窗口内到达并符合时间重叠条件，脉冲处理器将发出信号以重建放射性同位素的分布。通过使用类似葡萄糖的示踪剂，PET 可以描述人体内葡萄糖摄取的增加情况，并准确确定癌症影响的区域。PET 成像已经从研究成像发展成为肿瘤学诊断和分期的标准组成部分，为医生提供了有关组织和器官功能及状态的临床有用信息。所提供的信息类型取决于显像剂和疾病，可以包括检测、分类、分期、预后、治疗计划、评估治疗反应和监测<sup>[8]</sup>。

电子计算机断层扫描（Computed Tomography, CT）是一种快速扫描并生成人体组织图像的技术。通过让 X 射线以不同角度通过患者的身体，并使用探测器检测辐射信号，然后根据信号的强弱来重建图像，以显示人体解剖结构的信息<sup>[9]</sup>。CT 图像通过显示不同组织或器官对 X 射线的吸收差异，以不同的灰度值表示。高吸收率的区域显示为白色，代表骨骼等密度较高的组织；低吸收率的区域显示为黑色，代表充满空气的肺部。CT 图像可以显示微小病变以及脑部、肺部、肝脏、胆囊等软组织器官，在神经影像学、肺部影像学、骨骼肌肉影像学、肿瘤学和急诊医学等多个医学领域得到广泛应用<sup>[10]</sup>。

CT 作为肿瘤成像的重要手段，虽然能够提供较好的肿瘤解剖结构信息，但缺乏显示生理差异的能力。与之相比，PET 在确定组织代谢活动方面具有独特的优势，但它需要更高分辨率的解剖信息的帮助，而这是它无法提供的。为了克服这一问题，PET/CT 技术作为一种非侵入性的多模态显像方式被引入。它将 PET 和 CT 图像相结合，充分利用了 CT 的解剖结构信息和 PET 的代谢信息。相比常规的影像学检查，如 B 超、CT 和 MRI，PET/CT 具有更高的灵敏度和准确性。PET/CT 能够探测到细胞代谢的生物学活性，从而对潜在的病变进行早期检测和诊断。此外，PET/CT 还能够提供全身范围内的信息，一次性完成对全身病灶的排查，这对于评估肿瘤的分期和转移情况非常重要。在肿瘤诊断和治疗指导方面，PET/CT 已经得到了广泛的应用，并被认为是目前诊断和指导治疗肿瘤的最佳手段之一。它能够提供更准确和全面的信息，帮助医

生做出更精确的诊断和分期、评估治疗效果，以及指导个体化的治疗策略等<sup>[11][12]</sup>。这种综合性的信息对于肿瘤患者的个体化治疗至关重要，可以提高治疗效果并减少患者的不适和副作用。

最大强度投影（Maximum Intensity Projection, MIP）是一种图像处理技术，用于将三维数据转化为突出显示最大数值的二维图像<sup>[13]</sup>。它通过在感兴趣的体积区域内发射平行光线，选择沿光线路径的最大数值来生成图像。如图 1.1 所示右边的 PET 最大强度投影，这种处理方法可以突出显示血管或其他高密度结构，提供清晰可见的视觉图像，帮助医生进行疾病诊断和分析。

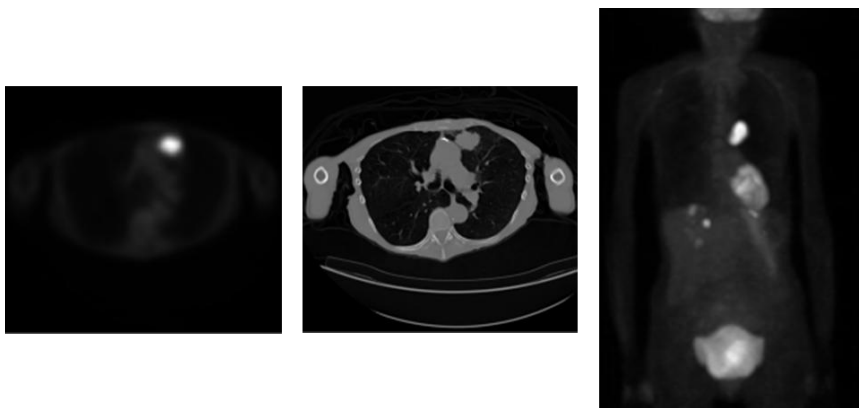


图 1.1 PET/CT 和 MIP 示例图。从左至右依次为 PET、CT 以及 PET 最大强度投影图像

### 1.1.3 非小细胞肺癌淋巴结转移诊断的意义

淋巴结转移在非小细胞肺癌的诊断中扮演着重要的角色，对预后评估、分期确定、治疗决策和预测都有重要影响。淋巴结转移是非小细胞肺癌预后不良的独立危险因素之一，其存在提示肿瘤已经扩散到淋巴系统，增加了治疗难度和预后风险。分期是确定肿瘤的严重程度和扩散程度的过程，可以帮助制定最佳的治疗方案，并预测患者的预后。淋巴结转移的存在与否以及淋巴结转移的数量和位置等因素都对分期结果产生影响。因此，在非小细胞肺癌的分期过程中，准确诊断淋巴结转移具有重要意义，并能为治疗方案的选择提供参考。在临床实践中，医生通常会使用多种方法来评估淋巴结转移，主要包括影像学检查和组织病理学检查等。

然而，目前淋巴结转移的诊断主要依赖于病理标签的监督学习。由于某些患者无法接受手术或肿瘤位置不适合进行病理检查，这导致他们无法获得病理标签。而影像标签虽然适用于所有患者，但其中一些可能存在准确性问题。医学影像的解读和判读是一个复杂的过程，受到医生经验、设备质量以及图像质量等多种因素的影响，可能会导致标

签的不准确性。为了解决病理标签缺失和噪声标签的问题，近年来一些研究已经开始探索针对这些情况下的医学图像分类工作。然而，大多数方法仅依赖于其中一种标签进行训练和预测，而未能充分利用两种标签的信息。因此，开发一种能够同时利用病理标签和影像标签的淋巴结转移诊断方法显得尤为重要。

## 1.2 研究现状

### 1.2.1 监督学习方法

影像组学是一种利用数字医学图像转化为可挖掘的高维数据，并通过分析这些数据来支持决策的技术。通过从感兴趣区域中提取出高维特征数据，这些特征可以包括形状、大小、纹理、强度等，并结合临床数据，然后可以使用机器学习方法或统计学方法对这些高维数据进行挖掘，以开发出能够改善诊断、预后和预测准确性的模型<sup>[14]</sup>。

Dai 等人<sup>[15]</sup>进行了一项回顾性研究，旨在通过结合 PET/CT 影像和临床数据对肺腺癌患者的纵隔淋巴结转移进行预测。研究共纳入两家机构的 320 例患者，其中包括 288 例内部病例和 32 例外部病例。医生首先确定了原发肺肿瘤感兴趣的区域，并提取了 190 个影像组学特征；接着采用曼-惠特尼 U 检验和 LASSO 算法对这些特征进行了降维处理，得到 10 个最优影像组学特征；最后采用逻辑回归建立了一个模型，以预测患者是否存在纵隔淋巴结转移。在内部测试集中，该方法表现良好，AUC 达到了 0.885，准确率为 86.2%；而在外部测试集上，该方法也取得了很高的性能，AUC 值达到了 0.948，准确率达到了 90.6%。研究结果表明，结合临床和 <sup>18</sup>F-FDG PET/CT 影像组学的综合分析，能够显著提高对肺腺癌患者纵隔淋巴结转移的预测能力。这一研究为肺腺癌的诊断和治疗提供了重要的参考依据。

Liu 等人<sup>[16]</sup>回顾性地分析了 180 名宫颈癌患者，旨在利用磁共振成像（Magnetic Resonance Imaging, MRI）提取影像组学特征来预测淋巴结转移情况。研究中首先采用软件对 MRI 图像进行手动三维分割的方法，主要是为了识别每个横截面上的淋巴结和肿瘤区域，并通过人工分割来进行准确的标注；然后从归一化的图像中提取了 122 个影像组学特征，并通过最小绝对收缩和选择算子（Least Absolute Shrinkage and Selection Operator, LASSO）对临床和影像组学特征进行筛选，从中筛选出关键预测指标；最后利用逻辑回归、多项式朴素贝叶斯、支持向量机、决策树等七种机器学习方法构建了预测

模型。研究表明,使用多项式朴素贝叶斯模型可以更好地预测宫颈癌患者术前淋巴结转移,其 AUC 为 0.745,准确度为 0.778。这项研究证实了将 MRI 图像的影像组学特征与患者的临床特征相结合的多项式朴素贝叶斯模型可以作为一种非侵入性工具,用于术前评估淋巴结转移,并证明了该模型的有效性和稳定性。

Zhao 等人<sup>[17]</sup>进行了一项回顾性分析,对 199 名乳腺癌患者的超声影像进行了评估,并结合外周血 T 细胞亚群分析来预测哨兵淋巴结的转移情况。首先由两名医生利用专业软件在超声影像上勾勒出感兴趣的区域,并提取了经过标准化处理后的图像组学特征;然后对产生的 864 个影像组学特征使用类间/类内相关系数进行一致性分析,再通过曼-惠特尼 U 检验和 LASSO 进行特征选择和降维,最终筛选出了 19 个特征;最后采用支持向量机、逻辑回归、朴素贝叶斯和分类决策树构建了单变量模型,并选取性能最优的模型来建立组合模型。在验证集上,他们使用逻辑回归构建的综合模型实现了 0.79 的 AUC。该研究构建了具有较强可重复性和临床意义的综合模型,能够用于预测乳腺癌中的哨兵淋巴结转移,并可以为与手术程序相关的临床决策提供有效指导。

Li 等人<sup>[18]</sup>对 570 例甲状腺乳头状癌患者进行了回顾性分析,通过使用超声图像提取的影像组学特征来预估颈部淋巴结转移的转移情况。首先通过软件手动勾勒出感兴趣区域;然后提取出 1561 个放射学特征,包括形状特征、强度特征和纹理特征,使用采用 T 检验或曼-惠特尼 U 检验进行特征选择,使用各种机器学习算法开发了放射组学模型;接着进行单变量和多变量逻辑回归来识别独立危险因素,并建立了临床模型;最后整合放射组学和临床模型创建组合诺谱图,取得了更好的 AUC 值。这项研究证明了超声放射组学技术为甲状腺乳头状癌患者颈部淋巴结转移的预测提供了一种定量、客观的方法,但临床指标仍不可替代。

深度学习是一种机器学习技术,它基于神经网络的概念,其核心思想是通过多个隐藏层来提取高级抽象特征,使原始输入数据能够转化为更具有高级表示的形式,通过反向传播算法进行权重调整,实现自动学习和模式识别的能力<sup>[19]</sup>。与影像组学相比,深度学习不需要手动提取定量特征,而是通过自动学习数据的高层次语义信息来提高特征学习效率。这样的特点使得深度学习在医学图像分析领域具有独特的优势,并在该领域取得了重要的应用和研究进展。

Wu 等人<sup>[20]</sup>提出了一项创新的研究,旨在使用深度多模态学习网络来预测原发性甲状腺癌患者的淋巴结转移情况。他们将临床上多个数据模态(如超声图



像和临床记录)整合到深度学习算法中,并采用了一系列复杂的网络结构和算法来实现该目标。该研究基于深度卷积神经网络和多层感知器来从超声图像和临床记录中提取抽象特征。通过融合和输入这些特征到长短期记忆网络(Long Short-Term Memory, LSTM)中,该研究可以学习多模态数据的上下文信息。最后,通过应用全连接层对 LSTM 网络的输出进行分类,以预测淋巴结转移的状态。研究结果显示,这种深度多模态学习网络在两个独立验证集中表现出了非常良好的性能,其 AUC 值达到了 0.973。这些结果对于指导临床医生在诊断原发性甲状腺癌淋巴结转移时具有重要意义。此外,该研究还有助于增进我们对深度多模态学习算法在预测过程中的工作原理和模式的理解。

Zhang 等人<sup>[21]</sup>对 523 名已经病理证实为局部晚期胃癌的患者进行了回顾性收集。他们利用预训练的卷积神经网络从患者的术前 CT 图像中提取深度学习特征,并将支持向量机作为分类器。研究评估了模型的性能,并与基于训练组开发的放射组学模型进行了比较。研究结果显示,基于 ResNet 提取的特征的最佳模型在测试组中表现更好,其 AUC 为 0.796,准确度为 75.2%,相较于放射组学模型的 AUC 为 0.704 和准确度为 61.8%,有显著的提高。这一研究为胃癌的预测和治疗决策提供了一种新的非侵入性方法,并具有潜在的临床应用前景。

Wan 等人<sup>[22]</sup>对 611 名接受术前 MRI 检查的 T1-2 期直肠癌患者进行了回顾性分析,研究了四种二维和三维残差网络(ResNet18, ResNet50, ResNet101 和 ResNet152)在 T2 加权图像上进行训练和测试,旨在识别具有淋巴结转移的患者。三名放射科医生独立评估 MRI 上的淋巴结状态,并将其结果与深度学习模型进行了比较。研究结果显示,八个深度学习模型在训练集中的 AUC 范围为 0.80-0.89,在验证集中的范围为 0.77-0.89。其中,基于三维网络结构的 ResNet101 模型在测试集中表现最佳,其 AUC 值为 0.79,远高于放射科医生的综合 AUC 值 0.54。这一研究结果表明,基于深度学习的模型可能成为预测 T1-2 期直肠癌患者淋巴结转移的有用工具,可以协助医生进行个体化的治疗决策。

Zhou 等人<sup>[23]</sup>在乳头状甲状腺癌颈部淋巴结转移预测方面建立了一个基于深度卷积神经网络的集成模型。该研究利用 ResNet-50 构建了一个集成深度学习模型,结合临床变量和术前超声图像(BMUS 和 CDFI)来预测患者是否出现

颈部淋巴结转移。实验证明，集成深度学习模型在测试集中取得了较高的预测性能，内部测试集的 AUC 为 0.86，外部测试集的 AUC 为 0.77。相对于所有放射科医生的平均结果，集成模型表现出更好的预测性能。该研究结果为临床实践提供了准确的医学决策依据，有助于实施个体化治疗。

Tian 等人<sup>[24]</sup>进行了一项回顾性研究，对来自六家医院的 1325 例实性浸润性肺腺癌患者进行了分析，开发并验证了一个结合放射组学和深度学习的融合模型，旨在预测跨多个中心的患者术前隐匿性淋巴结转移情况。研究建立了三种预测模型：放射组学模型（使用决策树和支持向量机）、深度学习模型（使用 ResNet-18, ResNet-34, ResNet-50, DenseNet-121 和 Swin Transformer）以及融合模型（将放射组学和预训练的 ResNet-18 特征通过串联技术整合）。融合模型在验证和测试集上表现出色，平均 AUC 为 0.754，优于其他模型。这项研究展示了放射组学-深度学习融合模型在预测隐匿性淋巴结转移方面的泛化能力，有望为跨多个中心的实性浸润性肺腺癌患者的个性化治疗提供帮助。

### 1.2.2 半监督学习方法

在医学应用中，深度学习面临着许多挑战之一是缺乏可靠的标签数据<sup>[25][26]</sup>。为了增加深度学习方法在医学领域的应用和成功，必须找到解决数据稀缺和不可靠标签的问题的方法。半监督学习可能是一个具有潜力的解决方案，它通过利用未标记的数据和更多的图像信息来提高模型性能，并减少对大量标记数据的需求。这一方法可以帮助我们充分利用现有的数据资源，并更好地开发深度学习在医学领域的潜力。

Su 等人<sup>[27]</sup>在他们的研究中提出了一种名为 Semi-HIC 的半监督深度学习方法，旨在对组织病理图像进行分类。通过引入半监督损失函数，将关联周期一致性损失和最大条件关联损失相结合，利用未标记的图像块来解决组织病理图像中类间相似性和类内变异性的问题，从而显著提高了分类性能。此外，他们还采用高效的网络结构和级联 Inception 块来学习丰富且有区分力的特征表示。该方法在 Bioimaging 2015 挑战数据集和 BACH 数据集上取得了良好的结果，这对于减轻病理学家的负担，提高组织病理图像的自动分析能力具有重要意义。

Liu 等人<sup>[28]</sup>提出了一种新的医学图像分类方法，该方法结合了半监督生成对抗网络（Generative Adversarial Network, GAN）和伪标记技术。通过这种方法，能够从未标记样本中提取有用的特征，并通过 K-means 聚类算法为这些样本生成伪标记，不仅增强了

图像特征对分类任务的影响，还有效地解决了类内差距不明显的问题，并提高了伪标记的准确性。经过实验证明，该方法在肺部疾病诊断的 X 射线图像分类和乳腺癌诊断的组织病理图像分类方面，相较于已有的方法，具有更优异的性能。值得一提的是，虽然该方法仅使用了少量的标记数据，但与传统的监督学习方法相比，其性能相当甚至更好。

Atha 等人<sup>[29]</sup>提出了一种半监督深度学习方法 SSBTCNet，用于脑肿瘤的分类识别。该方法采用一种结合自编码器和多层感知器的网络架构，通过同时训练这两个模型来提高脑部 MRI 的分类性能。此外，为了提高 SSBTCNet 的效率，研究中还引入了一种基于模糊逻辑的增强样本生成方法，并且利用数据增强技术改进了脑肿瘤和非肿瘤的分类方法。实验结果表明，该方法在多个不同的数据集上的准确率明显优于文献中使用的其他脑肿瘤分类方法。

Su 等人<sup>[30]</sup>提出了一种名为 Semi-LAC 的半监督方法，该方法采用局部增强技术对病理图像的局部区域进行随机增强，通过增加图像的多样性并避免混合不相关的区域，该方法能够提高图像分类的准确性。此外，作者还引入了方向一致性损失函数，以确保原始图像和增强图像的特征和预测结果上保持一致。该损失函数要求低置信度的预测或特征与高置信度的预测或特征对齐，进而提升分类网络对病理图像的判别能力和预测准确性。实验证明该方法在病理图像分类中有效，对于应对标注数据不足的问题具有重要的实际意义。

### 1.2.3 噪声标签学习方法

近年来，深度学习技术在众多行业实现了显著进展，但由于其对准确标注数据的依赖性，采集大规模高质量数据变得困难而耗时。噪声标签是指数据集中存在标签错误或标签不准确的情况，在深度学习算法中，噪声标签的存在可能会导致模型性能下降或训练过程不稳定。因此，一些研究人员致力于解决噪声标签的问题，并提出了各种方法来应对这一挑战<sup>[31]</sup>。

Li 等人<sup>[32]</sup>提出了一个名为 DivideMix 的创新框架，用于在存在噪声标签的情况下进行半监督学习。该框架通过对每个样本的损失分布建模成混合模型，从而将训练数据分为含有干净样本标注的集合和含有噪声样本的无标注集合两部分。然后，利用半监督学习方法在标注和无标注数据上同时训练两个不同的网络，以避免确认偏差的产生。为了进一步改进模型的性能，DivideMix 采用了两个关键策略。首先，每个网络使用另一个网络的数据集划分，以增加模型的鲁棒性和泛化能力。其次，在半监督训练阶段，对

标注样本进行标签共同细化,对无标注样本进行标签共同猜测。这两个策略的结合使得 DivideMix 能够更好地应对噪声标签和未标注样本,从而提高了半监督学习的性能。研究人员通过对多个基准数据集的实验结果进行评估,证明了 DivideMix 相较于现有方法在性能上的明显改进。这项创新的框架为解决噪声标签下的半监督学习问题提供了一种有效而可行的解决方案。

Yi 等人<sup>[33]</sup>提出了一个全新的端到端概率性纠正框架,专门解决标签噪声问题,该框架通过两个关键的方面来解决这个问题。首先,它将图像的标签建模为一个包含所有可能标签分布的概率分布,而不是仅仅是一个固定的分类值。这个改变使得我们能够统一地处理受噪声影响和无噪声的标签。其次,该框架通过在网络参数学习和标签学习中持续更新和维护标签分布,以纠正噪声。它以端到端的方式直接纠正噪声并更新标签分布,该框架首先使用噪声标签初始化标签分布,然后通过不断更新标签分布来逐步纠正噪声。最重要的是,网络的损失函数利用的是更新后的标签分布而不是原始的噪声标签。这种概率性纠正框架在合成数据和真实世界数据集上都取得了显著更好的性能。

Wei 等人<sup>[34]</sup>提出了一种名为 JoCoR 的鲁棒学习范式,用于解决深度学习中存在嘈杂标签的问题。该方法通过同时使用两个网络对同一批次的数据进行预测,并计算每个训练样例的共同正则化损失,然后选择损失较小的样例来同时更新两个网络的参数。通过这种联合损失进行训练的方式,两个网络会逐渐趋向相似,从而提高了鲁棒性。实验证明,JoCoR 在基准数据集如 MNIST 和 CIFAR-10 上对于处理嘈杂标签具有显著的改进效果。这种联合训练和共同正则化的方法为解决嘈杂标签问题提供了一种新的解决思路。

Zhu 等人<sup>[35]</sup>提出了一种新颖的鲁棒噪声标签纠正方法,该方法采用共同教学学习范式和基于一致性的策略,旨在降低噪声标签对结果的影响,并有效利用数据中的有用信息。该方法由两个关键模块组成,即自适应噪声率估计模块和基于一致性的噪声标签纠正模块。自适应噪声率估计模块的主要作用是计算数据集中的噪声率,从而可以辅助检测噪声标签的存在。然而,在临床应用领域,这个模块通常无法得到可靠的估计结果。因此,研究团队将重点放在基于一致性的噪声标签纠正模块上。基于一致性的噪声标签纠正模块的目标是检测和纠正数据集中的噪声标签,以减少其对分类结果的干扰,并能够充分利用数据中的有用信息。通过分析数据集中标签的一致性,该方法能够识别出潜在的噪声标签,并进行相应的纠正操作。这种方法在公共皮肤病变数据集和甲状腺超声图像数据集上进行了实验,证明了其能够适应不同类型的数据集,并有效地纠正噪声标

签。尤其在医学图像分类方面，该方法具有很高的应用潜力。

目前，虽然已经有一些针对噪声标签的分类方法取得了良好的结果，但是在医学图像领域，特别是在 PET/CT 图像上诊断非小细胞肺癌患者是否存在淋巴结转移的研究还非常有限。同时，对于病理标签和影像标签的共同学习也是一个待开拓的领域。因此，本文旨在探索使用半监督深度学习方法，结合病理标签和影像标签的联合学习，利用 PET/CT 图像进行非小细胞肺癌淋巴结转移的诊断。

### 1.3 研究内容与创新

本研究针对 PET/CT 非小细胞肺癌淋巴结转移诊断问题，首先提出了一种基于半监督深度学习的解决方案。本文使用了一个端到端的半监督深度学习模型（Dual Student with Teacher and Correction, DSTC），该模型能够从医生勾画的肿瘤金标准中提取 PET/CT 肿瘤块的深度学习特征，通过两个网络分别学习病理标签和影像标签，然后教师网络能够结合来自这两个领域的知识来输出淋巴结转移的分期结果。该模型通过引入单向学习方法，有效地利用了病理标签的深度学习信息来指导影像标签的深度学习。教师网络的模型权重由学生网络共享，并且通过一致性学习来确保教师和学生网络同样表现良好。此外，我们还引入了半监督纠正网络，它可以利用病理标签对影像标签进行细化，最后通过对比实验和消融实验验证了模型的有效性。

为了解决医生手动勾画肿瘤区域的时间和劳力成本高以及不一致性的问题，本文进一步使用了无肿瘤分割的方式，通过多尺度、多模态影像的半监督深度学习模型对淋巴结转移进行诊断。使用患者的原始半身 PET/CT 图像以及 PET/CT 的正面最大强度投影图像来构建模型，并采用半监督学习中的伪标签方法来训练没有病理标签的数据。在模型中，引入了两个特征提取网络，分别针对二维最大强度投影图像和三维 PET/CT 图像进行深度学习特征提取，并添加了三个分支来处理 PET 模态、CT 模态和两者的堆叠图像。同时，在 ResNet 网络中引入了融合模块，以更好地融合不同模态的信息和上下文信息。最后，将二维和三维的深度学习特征进行融合，用于非小细胞肺癌淋巴结转移的诊断，取得了与先前方法相近的预测效果。

本文的创新点在于：

(1) 在淋巴结转移的诊断任务中，研究采用了两个相互独立的学生网络。第一个学生网络专门针对病理标签进行训练，而第二个学生网络则专注于影像标签的学习。为了提升学习效果，研究引入了一种单向学习方法，使得第一个学生网络能够指导第二个学

生网络的学习过程。实验结果表明，影像标签和这种单向学习机制对淋巴结转移的诊断任务起到了积极的推动作用。为了进一步提高非小细胞肺癌淋巴结转移的诊断精度，研究加入了一种深层特征引导的半监督纠正网络。这个网络结合了深度学习特征和原始的影像标签作为输入，利用半监督学习算法纠正影像标签。经过纠正的标签随后被用于强化第二个学生网络的训练过程，从而提升了其诊断性能。

(2) 使用一种无需医生提供肿瘤区域标注、无需分割的深度学习方法，利用患者的三维 PET/CT 图像和二维最大密度投影来预测非小细胞肺癌患者是否存在淋巴结转移，并且在特征提取网络中加入特征融合模块。实验结果显示，使用多尺度图像模型比单尺度模型更优，融合模块也表现出更好的性能。

## 1.4 本文组织结构

第一章，绪论。在本章中分为三个部分，第一部分介绍了文章的研究背景与意义，说明预测非小细胞肺癌患者的淋巴结转移的重要性；第二部分将概述淋巴结转移预测和噪声标签学习的研究现状，包括基于监督学习方法、半监督学习方法和噪声标签学习方法的研究进展；第三部分，对本文的研究内容、创新点以及文章的结构进行了说明。

第二章，相关理论。本章旨在探讨涉及到的理论知识，主要包括深度学习中常用的卷积神经网络及其关键组成部分：卷积层、池化层、全连接层和激活函数，以及 ResNet 网络、半监督学习方法和病理诊断与影像诊断的相关理论基础。

第三章，基于 DSTC 的半监督淋巴结转移诊断。本章主要分为四个部分，第一部分，首先介绍了研究流程和实验数据预处理工作，其中通过医生描绘的肿瘤区域与原图进行对比，选取肿瘤区域，并使用线性插值方法标准化肿瘤块的大小。第二部分，详细介绍了深度学习模型的设计和搭建过程。具体可分为如下步骤：第一，以 ResNet-18 网络为基础建立深度学习模型，为了提高对非小细胞肺癌淋巴结转移的预测效果，考虑到影像标签的利用，本章采用两个学生网络来学习病理标签和影像标签，并通过单向学习方法实现病理对影像的指导。此外，还通过学生网络权重的指数移动平均值构建了教师网络，并与学生网络进行一致性学习。同时，引入半监督纠正网络，利用病理标签对影像标签进行进一步细化。第三部分，对实验结果进行了对比分析，主要包括两个方面的实验：首先，与其他不同方法进行对比实验，与两种半监督方法和四种噪声学习方法进行了对比，结果证明本章研究的模型具有更好的性能；其次，与医生诊断结果进行对比分析，通过混淆矩阵和 Kaplan-Meier 曲线的使用，来评估模型的诊断准确性；最后，

进行了深度学习模型的消融实验，用以验证本章所加入的单向学习方法和纠正网络的有效性。

第四章，基于多尺度多模态融合网络的无肿瘤分割淋巴结转移诊断。在第三章中，使用的深度学习方法需要医生首先勾画出肿瘤区域，这增加了医生的工作量。因此，获取有标注的数据集是一项非常具有挑战性的任务。面对这个问题，本章旨在探索在无需肿瘤区域标注的情况下进行患者淋巴结转移预测的方法。与前一章相同，本章的实验环境和评价标准保持一致。实验主要分为两个步骤，首先是数据预处理，本章的模型引入了 PET/CT 图像的二维正面最大密度投影和三维半身图像，以利用全身上下文特征。然后建立半监督深度学习模型。对于没有病理标签的数据，采用半监督学习中的伪标签方法进行训练。本章使用了两个特征提取网络 ResNet-18，用于提取二维和三维深度学习特征。每个特征提取网络都包含 PET、CT 和堆叠三个分支，并且在堆叠分支中引入融合模块，用以融合不同模态和上下文信息。最终将两个维度图像的特征拼接起来进行淋巴结转移预测。最后进行了对比消融实验，以证明多尺度多模态方法的有效性，并验证融合模块对模型预测能力的提升作用。

第五章，总结与展望。对本文内容进行了总结，并讨论了研究中存在的缺陷和不足之处，分析了未来研究中可以提升的方向。

## 2 相关理论

### 2.1 卷积神经网络

卷积神经网络 (Convolutional Neural Networks, CNN) 是当前深度学习领域中被广泛采用的一种模型<sup>[36]</sup>。它最初受到生物视觉系统神经机制的启发,旨在模拟人类大脑的视觉神经元对事物进行分析和处理的能力,使得计算机能够类似于人类对世界进行观察和理解。在 20 世纪 80 至 90 年代,研究者开始对卷积神经网络进行初步研究,时间延迟网络<sup>[37]</sup>和 LeNet<sup>[38]</sup>是当时最早应用的卷积神经网络模型<sup>[39]</sup>。随着 21 世纪的到来,深度学习理论的提出以及计算设备性能的改善,卷积神经网络迅速发展,并广泛应用于计算机视觉、自然语言处理等领域<sup>[36]</sup>。2012 年, Krizhevsky 等人<sup>[40]</sup>提出的 AlexNet 成为第一个成功应用深度卷积神经网络结构于大规模图像数据集的重要里程碑,也被视为深度学习发展史上的关键转折点。随后,基于 AlexNet 的网络模型如 VGGNet<sup>[41]</sup>、GoogLeNet<sup>[42]</sup>、ResNet<sup>[43]</sup>、DenseNet<sup>[44]</sup>等相继提出,进一步推动了卷积神经网络的发展。下面将对卷积神经网络的基本结构进行介绍。

#### (1) 卷积层

在卷积神经网络中,卷积层 (Convolutional layer) 是非常关键的组成部分,其主要作用是从输入数据中提取关键特征。卷积层内部包含了多个学习参数的滤波器,也称为卷积核。这些卷积核会在输入数据上滑动并进行卷积操作,得到特征图。不同的卷积核专门用于检测输入数据中的不同特征,例如边缘和纹理等。通过叠加多个卷积层,神经网络可以学习到更复杂和抽象的特征。卷积层的参数包括卷积核的大小、步长和填充,这些参数共同影响着卷积层输出特征图的大小<sup>[45]</sup>。卷积核的大小可以设定为小于输入图像大小的任意值,如果卷积核的大小较大,它可以提取出更复杂的输入特征<sup>[45]</sup>。卷积步长决定了卷积核在特征图上滑动时的间距。当步长为 1 时,卷积核会逐个扫描特征图的每个元素;而当步长为  $n$  时,它将跳过  $n-1$  个像素,在下一次操作时继续扫描<sup>[46]</sup>。填充是一种人为增加特征图大小的方法,用于抵消卷积操作中的尺寸缩小效应。

#### (2) 池化层

在卷积层之后,特征图会经过池化层 (Pooling layer) 进行特征选择和信息过滤。池化层的作用是通过输入特征图进行下采样,从而压缩特征维度。这样做有两个主要目的。一方面,可以简化网络的计算复杂度,降低计算成本,同时也可以在一定程度上减



少过拟合的风险。另一方面，池化层聚焦于提取特征图中的重要特征。池化层的计算方式与卷积层类似，通过调整池化大小、步长和填充来控制<sup>[45]</sup>。卷积神经网络中常用的池化层种类包括最大池化（Max pooling）、平均池化（Average pooling）和全局池化（Global pooling）等。在最大池化中，特征图被划分为多个池化区域，每个区域选取最大值作为输出；而在平均池化中，则计算每个区域的平均值作为输出。全局池化将整个特征图压缩为一个值，可以是全局平均或全局最大。通过池化操作，可以减小特征图的大小，简化网络计算复杂度。同时，池化层也能够提取出主要的特征信息，增强模型的表达能力和泛化能力。

### （3）全连接层

在卷积神经网络中，通常会在输出前添加一个或多个全连接层（Fully Connected Layer, FC）。这些全连接层的主要作用是将卷积层和池化层提取到的特征进行整合，从而得到最终的分类结果。全连接层中的每个神经元都与上一层的所有神经元相连接。不同于卷积层和池化层具有特征提取能力的特点，全连接层本身并不能提取特征，它试图利用已经提取到的高阶特征来完成学习任务。

### （4）激活函数

激活函数（Activation functions）在卷积神经网络中扮演着重要的角色。通过将输入加权求和后应用激活函数，可以帮助神经网络学习和理解复杂的非线性关系。激活函数的主要作用是改变数据之间的线性关系。如果网络只包含线性变换，多层网络可以转化为单层网络，失去其多层结构的优势。然而，引入激活函数赋予了网络实际意义，并增强了网络的能力。它增加了网络的容量，使其能够学习复杂的事物和数据，并实现输入输出之间的非线性映射。此外，激活函数还能对数据进行归一化，将其映射到特定范围内再传递给下一层。这样做的好处是限制数据的扩张，避免因数据过大而导致的溢出风险。常用的卷积神经网络激活函数包括 Sigmoid、Tanh 和 ReLU。

## 2.2 ResNet 网络

ResNet 是一种深度学习模型，由微软研究院提出，并在图像识别、目标检测和语义分割等任务中展现出卓越的性能。它的核心组成部分是多个残差块（Residual block），每个残差块内部包含了多个卷积层和批量归一化层。图 2.1 展示了一个深度为 18 层的 ResNet 网络结构。在这里，网络的深度指的是网络的权重层，包括了池化、激活和线性层，但不包括批量归一化层和池化层。通过引入残差块，ResNet 成功地解决了深层

网络训练过程中的梯度弥散和消失问题，使得网络更加易于训练和优化。它的引入对于深度学习的研究和应用有着重要的意义。

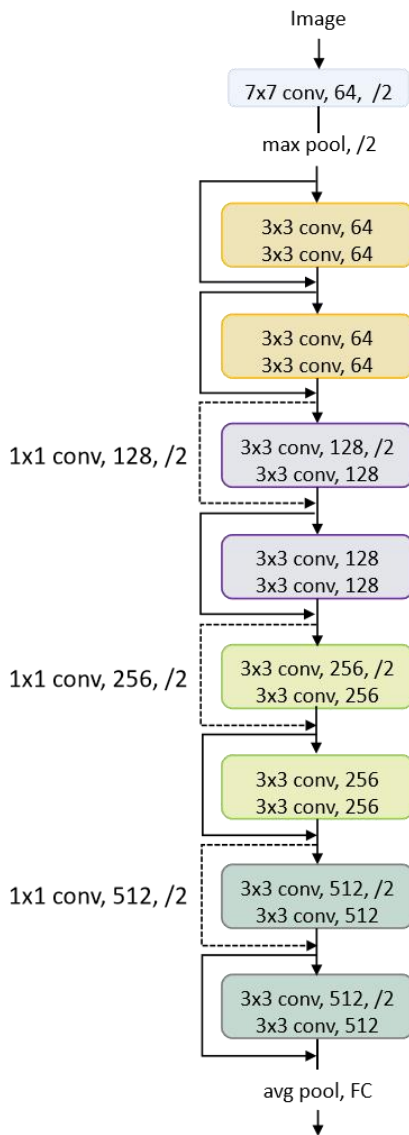


图 2.1 ResNet-18

为了解决深度神经网络训练中的梯度消失和表达能力限制问题，ResNet 提出了一种创新的解决方案——引入捷径连接 (Shortcut connection)。如图 2.2 所示，捷径连接的作用是将前一层的输出直接与后一层的输入相加，以保留前一层的重要特征信息。这样的设计使得网络能够更深入地堆叠层级，并实现跨层的信息传递，从而有效地缓解了训练困难和信息丢失的问题。

此外，ResNet 使用了一种称为全局平均池化层 (Global average pooling) 的技术，

取代了传统的全连接层，这样做有助于减少网络中的参数数量，并降低过拟合的风险。这种设计结构使得 ResNet 更加轻松地提取图像中的抽象特征，从而在各种计算机视觉任务中获得了突破性的性能。

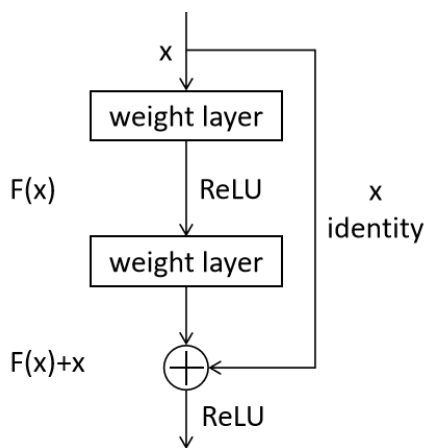


图 2.2 捷径连接

## 2.3 半监督学习

深度学习是由 Hinton 等研究者<sup>[19]</sup>于提出，是机器学习的一个新领域。传统的机器学习算法需要手动设计特征，而深度学习则能够利用数据自动学习特征表示。通过构建多层神经网络的层次化结构，并借助反向传播算法，深度学习可以有效地应对复杂的非线性关系和大规模数据<sup>[19]</sup>，取得了在计算机视觉、医疗诊断等领域的引人瞩目成就。

半监督深度学习是在深度学习在监督任务方面取得显著进展的基础上发展起来的。深度学习领域面临一个挑战，即需要大量的标签数据来训练模型。然而，标注数据的收集通常需要专业人士的参与，需要大量的时间和精力，这对实际应用造成了困扰。与此同时，我们可以轻松获取无标签数据。因此，半监督学习成为一种吸引人的研究方向，可以利用这些无标签数据来辅助有标签数据的学习。在某些高成本领域，例如医学图像分析和自然语言处理，半监督学习可以有效地解决数据稀缺的问题。在半监督深度学习中，可以充分利用无标签数据来提升模型性能。通过使用无标签数据，可以扩大训练数据集的规模，并更好地捕捉数据的潜在结构和规律，从而使模型更好地理解整个数据空间的特点。通过利用无标签数据的分布特征，模型的泛化能力得到提高，从而在数据获取成本较高或有限标签数据的情况下，取得更好的效果。

## 2.4 病理诊断和影像诊断

尽管肿瘤的生化、免疫和影像诊断得到了长足的发展，但病理学诊断仍然是确定肿瘤性质的主要依据。病理检查被视为诊断肿瘤最准确和可靠的方法，通过观察器官的肉眼改变、显微镜下观察组织结构和细胞病变特征，从而做出精确的疾病诊断。相比临床上根据病史、症状和体征等进行的分析性诊断（可能存在多个可能诊断）以及依赖各种影像技术所进行的诊断，病理学诊断更加客观和准确。病理诊断是对疾病最本质的判断，是诊断的“金标准”，也是对疾病的“终审判决”。尽管淋巴结活检是一种常用的检查方法，但仍然存在一些风险，比如术后出血和纵膈炎等并发症。此外，由于有些淋巴结与大血管位置相近，进行有创性检查可能不适合。因此，无创评估淋巴结转移可以弥补淋巴结活检的不足，影像学技术成为最佳选择。

影像学检查是临床上常用的辅助诊断手段，通过观察影像结果可以了解病变部位组织结构和血流灌注等情况的改变，进而对病变性质进行判断。在进行影像分析之前，需要详细了解患者的病史和其他相关检查资料，以便进行有针对性的观察和诊断。影像诊断是临床医学中扮演重要角色，并为后续的治疗和疗效评估提供必要依据。CT 是诊断淋巴结转移最常用的无创影像检查方法，通常通过 CT 检查测量淋巴结的最大短径来评估其是否增大，但有 28% 的增大淋巴结在病理检查上是阴性的，同时也有 20% 正常大小的淋巴结病理检查上显示为转移病变<sup>[47]</sup>。

因此，本文对影像标签进行细化，并与病理标签进行联合学习，通过综合应用病理学和影像学的手段，以提高淋巴结转移的诊断准确度。

## 2.5 本章小结

本章主要介绍了本文涉及的理论知识，包括卷积神经网络的基本构成、ResNet 的网络结构、半监督学习以及病理诊断和影像诊断。

### 3 基于 DSTC 的半监督淋巴结转移诊断

在本章的 PET/CT 非小细胞肺癌淋巴结转移诊断研究中，我们正在处理一个非传统的半监督分类问题。通常，半监督学习需要使用少量已标记数据和大量未标记数据。然而，在我们的研究中，我们采纳了 328 例非小细胞肺癌患者作为研究对象，所有患者都能够获取到影像标签，但只有 112 例患者拥有病理标签。本研究介绍了一种具有两个学生网络和纠正网络的师生模型（Dual Student with Teacher and Correction, DSTC），借助半监督学习的思想来利用所有 PET/CT 图像进行学习，对非小细胞肺癌患者的淋巴结转移进行分类。同时，我们设计了一种单向学习方法和纠正网络来有效利用病理标签和影像标签的信息。最后，我们进行了对比实验和消融实验，对实验结果进行了分析和讨论。

#### 3.1 研究流程

图 3.1 展示了本章实验的整体流程，分为三个主要步骤：（A）首先进行数据预处理，根据放射科医生勾画的肿瘤感兴趣区域，对 PET/CT 原始图像进行预处理，目的是得到肿瘤块；（B）然后，建立一个用于预测淋巴结转移的深度学习模型，该模型使用肿瘤块和影像标签作为输入。为了提升模型的性能并充分利用可用的信息，我们设计了一种单向学习方法和标签纠正方法。通过单向学习方法，病理标签和影像标签之间的信息关联得到了充分利用，而标签纠正方法则可以自动修正影像标签，从而提高模型训练的准确性和可靠性；（C）最后，对模型进行评价与分析，包括 AUC、准确度、敏感性、特异性等分类评估指标以及 Kaplan-Meier 曲线。

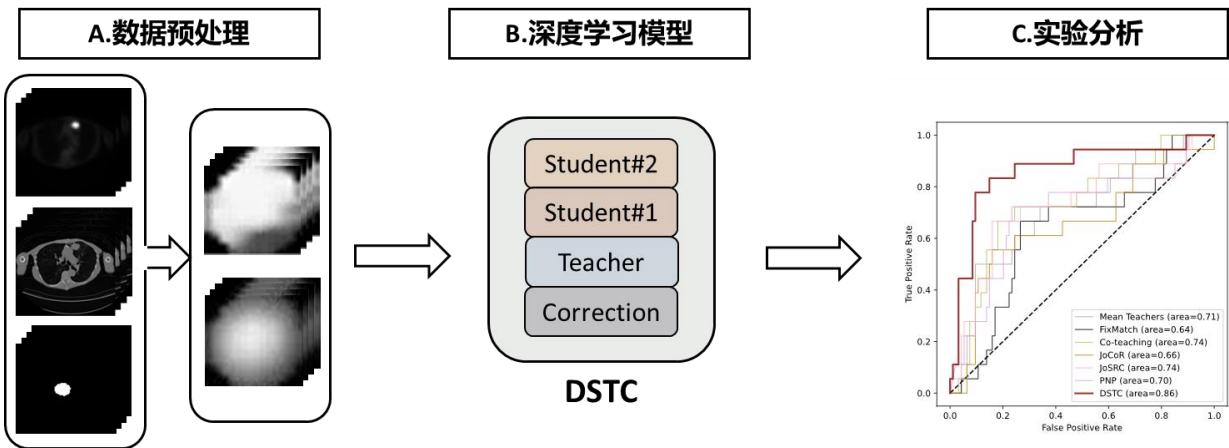


图 3.1 整体实验流程图

通过上述三个主要流程，本文旨在利用一种基于半监督深度学习的方法，用于预测淋巴结转移。该方法不仅考虑了病理标签和影像标签之间的信息关联，还通过单向学习和标签纠正方法提高了模型的预测能力和准确性。最后，通过全面的评价指标对模型进行评估和分析，从而为临床实践提供决策支持。

## 3.2 实验数据

### 3.2.1 数据介绍

本研究的实验数据使用了美国芝加哥大学医学院在 2004 年至 2014 年期间收集的非小细胞肺癌患者的 PET/CT 图像和临床资料。所有的数据都已通过相关机构审查委员会的批准，并且患者同意放弃知情权。

我们获得了 935 例有病理证实的非小细胞肺癌患者，其中没有发生脑转移或其他疾病。研究的主要目标是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诊断。为了达到这个目标，需要进一步筛选符合研究内容的患者。患者的纳入准则如下：

(1) 患者必须具有完整的 PET/CT 影像数据，这些影像数据包含全身信息，能够清晰地展示人体高代谢组织、器官和骨骼等；(2) 不能获得临床特征的患者将被排除；(3) 未发生远端转移的患者也将被纳入研究。经过筛选，最终共有 328 例非小细胞肺癌患者符合研究要求。其中，只有 112 例患者接受了淋巴结转移的病理检查，其余 216 例患者只有放射科医生提供的影像标签。在这 112 位具有病理标签的非小细胞肺癌患者中，有 18 例被确认患有淋巴结转移，而另外 94 例没有发生淋巴结转移。

本研究中的 243 例患者的 PET/CT 图像是在 2012 年 3 月 15 日之前使用 Reveal HD 成像扫描仪获得的，而 85 例患者的图像则是在该日期之后由 Simens mCT 成像扫描仪获得的<sup>[48]</sup>。所有的 PET/CT 图像都由两名经验丰富的放射科医生阅读使用 MIMvista 软件手动勾画出每位患者的肿瘤区域。

具体的患者数据统计如表 3.1 所示。在本研究中，共有 328 例非小细胞肺癌患者被纳入研究。这些患者的年龄是连续的变量，平均年龄为 68.7 岁，主要集中在 55 岁及以上。就种族而言，大部分患者为黑人（58.8%），其次是白人（39.0%）。另外，大多数患者曾经吸烟（55.8%）或者仍在吸烟（37.8%）。在治疗方面，41.5%的患者接受了手术，50.3%的患者不接受手术，还有 8.2%的患者未接受任何治疗。根据组织学亚型分类，大多数患者被诊断为腺癌（48.9%）和鳞状细胞癌（34.8%）。通过影像诊断发现，其中

180 例患者存在淋巴结转移，而 148 例则未出现转移情况。此外，有 216 例患者未接受病理检测，无法确定是否存在淋巴结转移。而在接受病理检测的 112 例患者中，仅有 18 例患者发生淋巴结转移。

表 3.1 数据特征统计

特征	统计值	特征	统计值
患者总数	328 (100%)	治疗方式, 个数 (%)	
年龄均值, 标准差	68.7 (10.7)	手术	136 (41.5%)
年龄阶段, 个数 (%)		未手术	165 (50.3%)
≤44	2 (0.6%)	未治疗	27 (8.2%)
45-54	39 (11.9%)	组织学亚型, 个数 (%)	
55-64	76 (23.2%)	腺癌	161 (48.9%)
≥65	211 (64.3%)	鳞状细胞癌	114 (34.8%)
性别, 个数 (%)		大细胞癌	12 (3.7%)
女	192 (58.5%)	非特指型非小细胞肺癌	34 (10.4%)
男	136 (41.5%)	其他	7 (2.2%)
人种, 个数 (%)		影像淋巴结转移诊断, 个数 (%)	
白	128 (39.0%)	转移	148 (45.1%)
黑	193 (58.8%)	未转移	180 (54.9%)
其他	7 (2.2%)	病理淋巴结转移诊断, 个数 (%)	
吸烟状态, 个数 (%)		转移	18 (5.5%)
从未	21 (6.4%)	未转移	94 (28.7%)
目前在吸	124 (37.8%)	无	216 (65.8%)
曾吸	183 (55.8%)		

### 3.2.2 数据预处理

为了消除不同的 PET/CT 扫描仪带来的差异，需要进行统一化预处理。具体流程如下所示：（1）将原始图像的 DICOM 格式转换为 nii 格式；（2）由于 PET 图像的灰

度值会受到显影剂剂量的影响，因此需要进行药物剂量校正以降低其影响。这种校正方式可以通过对像素值进行调整来实现，即基于患者体重归一化为 PET 代谢影像的标准摄取值 (SUV)；(3) 对 PET/CT 图像重新配准，并重采样到  $256 \times 256$  大小；(4) 医生通常会使用二值图像来标记肿瘤的位置，然而这些金标准图像并不包含患者肿瘤处的具体信息，无法直接作为模型的输入，因此需要从患者的 PET/CT 图像中获取感兴趣区域。使用肿瘤金标准图像作为模板，利用像素相乘的方法，将肿瘤以外的区域像素值设为 0，在 PET/CT 图像中创建感兴趣区域；(5) 沿感兴趣区域的边界将 PET/CT 图像裁剪成肿瘤块，并利用三线性插值将所有肿瘤块统一调整为  $64 \times 64 \times 64$  大小，最后进行最小最大值归一化。具体的数据预处理流程如图 3.2 所示。

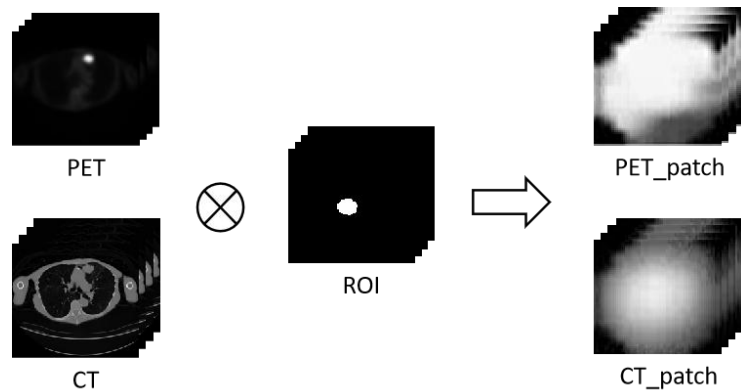


图 3.2 数据预处理流程图

### 3.3 深度学习模型架构

Mean Teachers<sup>[49]</sup> 是一种半监督学习算法，于 2018 年提出。该算法的核心思想是将模型分为教师和学生，通过教师模型生成学生学习的目标，并用学生模型的权重来生成教师模型的权重。具体来说，使用分类损失来比较学生模型的输出和标签，并使用一致性损失来比较学生模型的输出和教师模型的输出。在更新学生模型的权重之后，通过将学生模型的权重与教师模型的权重进行加权平均，使用指数移动平均 (Exponential Moving Average, EMA) 值来更新教师模型的权重。通过这种方式，教师模型可以获得学生模型在训练过程中的知识，并提供更准确的预测。然而，与传统的半监督学习问题不同，我们的研究中只有少数患者有病理标签，而所有人都有影像标签。因此，我们将 Mean Teachers 模型进行了扩展，引入了两个学生模型和一个纠正网络的师生模型



DSTC。通过使用两个学生网络，对病理标签和影像标签进行有效分离处理。同时，我们还引入了单向学习的方式，其中使用影像标签训练的学生模型可以从使用病理标签训练的学生模型中受益。教师网络仍然是通过指数移动平均值从两个学生网络中构建而成的，总结了学生网络在训练过程中获得的知识，并通过实现一致性学习来帮助学生网络更好地进行泛化和推理。为了进一步提高影像标签的质量，我们还构建了半监督纠正网络，利用病理标签对影像标签进行纠正，并在模型训练中重复使用这些纠正后的标签。通过这种方式，我们可以改善影像标签的准确性和可靠性。

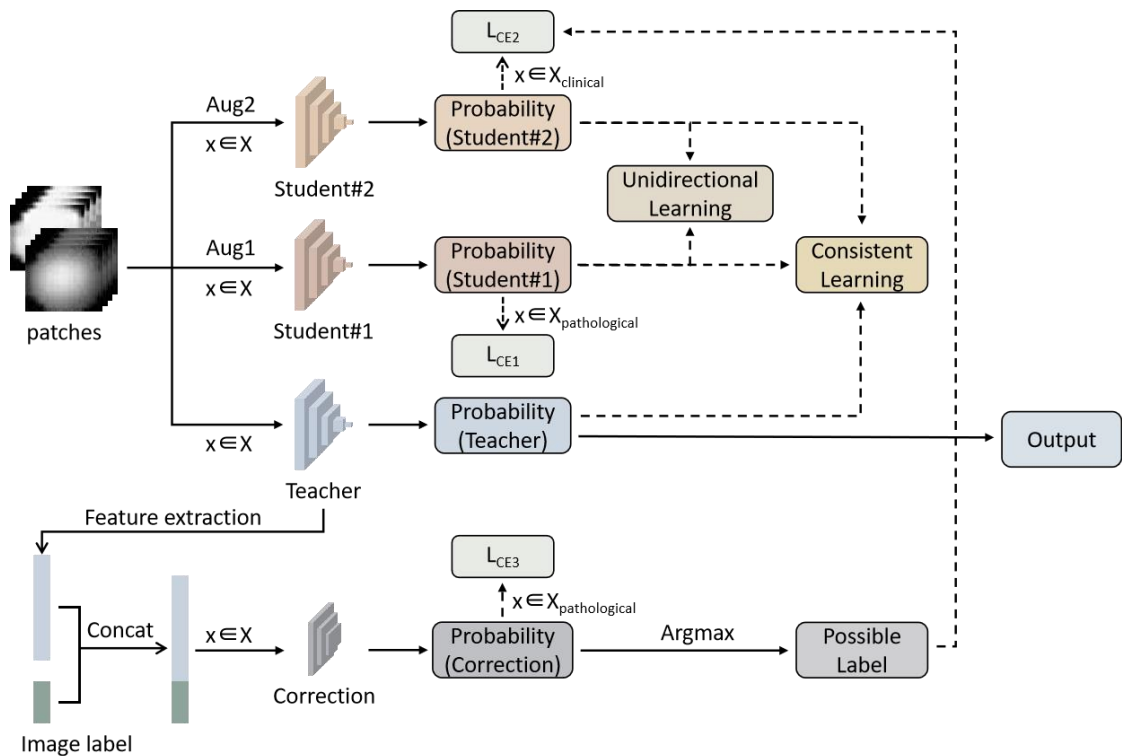


图 3.3 DSTC 模型架构

图 3.3 展示了深度学习模型 DSTC 的主要架构。DSTC 模型中包含四个网络，分别是两个学生网络、一个教师网络和一个纠正网络。教师网络和两个学生网络的结构是相同的，它们都采用 ResNet-18 作为特征提取器，已有相关研究证明参数更少的 ResNet-18 在图像分类上有更好的性能<sup>[24][35][50]</sup>。在完成提取特征后，通过引入两个全连接层对特征进行降维处理，最后采用 softmax 激活函数得到预测结果。纠正网络则由三个全连接层组成，在每个全连接层之后都连接有一个激活函数<sup>[51]</sup>。

在这个模型中，教师网络输入经过预处理后的肿瘤块，而学生网络#1 和学生网络#2

则分别输入经过模糊和翻转增强处理后的肿瘤块。随后,使用 ResNet 进行深度学习特征提取,并预测发生淋巴结转移的概率。在一批患者的输入数据中,当病理标签可用时,学生网络#1 将使用病理标签计算交叉熵损失 ( $L_{CE1}$ ) 进行训练,而学生网络#2 则通过单向学习和一致性学习的损失进行训练;当病理标签不可用时,学生网络#2 将使用影像标签计算交叉熵损失 ( $L_{CE2}$ ) 进行训练,而学生网络#1 则通过一致学习损失进行训练。具体的交叉熵损失  $L_{CE1}$  和  $L_{CE2}$  的计算方式见公式 (3.1) 和 (3.2):

$$L_{CE1} = \frac{1}{K} \sum_{i=0}^{C-1} \sum_{n=0}^{K-1} y_n(i) \log(p_n(i)) \quad (3.1)$$

$$L_{CE2} = \frac{1}{M-K} \sum_{i=0}^{C-1} \sum_{n=0}^{M-K-1} \hat{y}_n(i) \log(\hat{p}_n(i)) \quad (3.2)$$

其中,  $C$  为分类数(这里  $C=2$  为转移和未转移两类),  $K$  为有病理标签的患者数,  $M$  为患者总数;  $y$  为病理标签,  $\hat{y}$  为影像标签;  $p(i)$  和  $\hat{p}(i)$  表示由学生网络#1 和学生网络#2 产生的患者属于  $i$  类的预测概率。

单向学习是一种评估学生网络的损失的方法,其重点在于评估学生网络#2 通过学生网络#1 进行学习的情况。这种评估方法可以帮助我们了解学生网络#2 在具体任务中的表现和学习进展。而一致学习则是评估两个学生网络从教师网络中学习的情况,目的是增强学生网络和教师网络之间的一致性。单向学习和一致学习的细节将在后面小节中详细描述。

在上述两个学习方法中,教师网络起到了非常重要的作用。教师网络通过将两个学生网络的结果与自己相似,来确保学习的一致性。为了达到这个目的,教师网络的权重会通过指数移动平均方法从两个学生网络中进行更新。这样生成的教师网络具有更好的稳定性和可靠性,它综合了学生网络在训练过程中所获得的知识,并帮助学生网络更好地进行泛化和推理。因此,教师网络的存在不仅提高了整体模型架构的性能和泛化能力,还加快了训练收敛的速度。指数移动平均更新过程描述为:

$$\varphi_t = \alpha\varphi_{t-1} + (1 - \alpha)(\lambda_1\theta_t + \lambda_2\hat{\theta}_t) \quad (3.3)$$

其中,  $\alpha$  为设置为 0.99 的超参数,  $t$  为训练步长,  $\lambda_1$ 、 $\lambda_2$  为两个学生网络的权重占比,  $\theta$ 、 $\hat{\theta}$  为两个学生网络的权重,  $\varphi$  为教师网络的权重。

由于指导学生网络#2 学习的影像标签会存在一定的不准确性,为了解决这个问题,我们引入了半监督学习策略的校正网络。具体而言,我们首先从教师网络的特征提取器

中提取出最后卷积层的特征向量，然后将这些特征与对应患者的二值影像标签进行拼接，形成组合特征。将组合特征用作校正网络的输入，使用交叉熵损失  $L_{CE3}$ （如公式 3.4 所示）来衡量预测概率与病理标签之间的差异。通过最小化该损失函数，使得预测概率更加接近真实的病理标签，从而提高影像标签的准确性。为了选择最可靠的标签，我们在校正网络的输出上应用  $\text{argmax}$  运算，选取概率值最大的标签作为新的标签用于训练学生网络#2，从而提高其分类精度和泛化能力。

$$L_{CE3} = \frac{1}{K} \sum_{i=0}^{C-1} \sum_{n=0}^{K-1} y_n(i) \log(\bar{p}_n(i)) \quad (3.4)$$

其中， $\bar{p}$  是校正网络输出的预测概率。

### 3.3.1 单向学习

为了提高学生网络#2 对淋巴结转移的诊断准确性，我们引入了单向学习策略。单向学习策略的核心思想是让学生网络#2 从使用病理标签训练的学生网络#1 中进行学习，以促进病理和影像信息的整合。

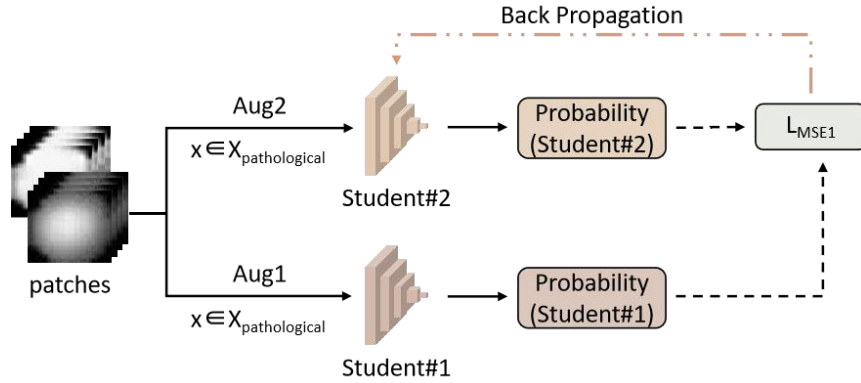


图 3.4 单向学习

单向学习的过程如图 3.4 所示，我们使用了一致性损失来度量学生网络#1 和学生网络#2 的预测概率之间的一致性，即  $L_{MSE1}$ 。通过最小化一致性损失  $L_{MSE1}$ ，我们能够让学生网络#2 从学生网络#1 中学习到更准确的淋巴结转移诊断知识。一致性损失  $L_{MSE1}$  的定义如下：

$$L_{MSE1} = \frac{1}{M-K} \sum_{i=0}^{C-1} \sum_{n=0}^{M-K-1} (p_n(i) - \hat{p}_n(i))^2 \quad (3.5)$$

这种单向学习策略的引入，使学生网络#2 能够更好地向病理标签贴近，从而提高其对淋巴结转移的诊断准确性。同时，通过整合病理和影像信息，学生网络#2 能够得到更全面、准确的预测结果，为医生提供更可靠的诊断依据。

### 3.3.2 一致性学习

图 3.5 展示了三个网络之间的一致性学习过程。

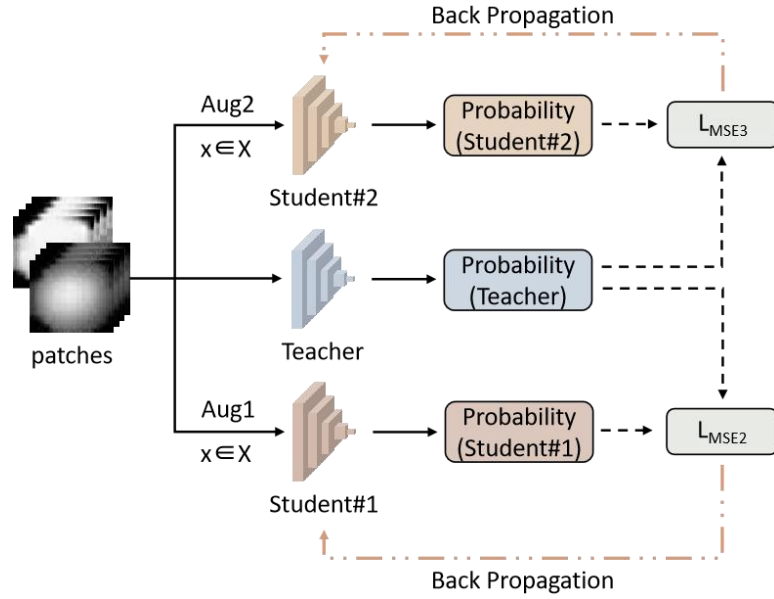


图 3.5 一致性学习

为了鼓励两个学生网络能够产生与教师网络相似的输出，引入了教师网络的输出作为学习目标，引导那两个学生网络进行学习。为了评估两个学生网络和教师网络输出之间的差异，我们使用了一致性损失  $L_{MSE2}$  和  $L_{MSE3}$  进行量化。这些损失的定义如下：

$$L_{MSE2} = \frac{1}{M} \sum_{i=0}^{C-1} \sum_{n=0}^{M-1} (p_n(i) - \tilde{p}_n(i))^2 \quad (3.6)$$

$$L_{MSE3} = \frac{1}{M} \sum_{i=0}^{C-1} \sum_{n=0}^{M-1} (\hat{p}_n(i) - \tilde{p}_n(i))^2 \quad (3.7)$$

其中， $\tilde{p}$  是教师网络输出的预测概率。

通过一致性学习，我们能够促使两个学生网络向教师网络学习，并增强它们的分类能力。在构建教师网络的过程中，我们采用了指数移动平均方法来计算两个学生网络的权重加权和。在迭代的过程中，这三种网络相互学习、相互影响，使得它们的分类能力

得到进一步增强。最终，我们使用教师网络进行实际测试，以验证模型性能。

### 3.4 总损失函数

本章的 DSTC 模型由四个部分组成：两个学生网络，一个教师网络和一个校正网络。教师网络是通过指数移动平均方法从两个学生网络构建而来，因此不需要使用损失函数；学生网络#1 的总损失由病理标签的交叉熵损失  $L_{CE1}$  和与教师网络的一致学习损失  $L_{MSE2}$  组成；学生网络#2 使用影像标签替代病理标签，其总损失包括无病理标签患者的交叉熵损失  $L_{CE2}$ ，有病理标签患者的单向学习损失  $L_{MSE1}$ ，以及与教师网络的一致学习损失  $L_{MSE3}$ 。此外，校正网络的损失仅包括病理标签的交叉熵损失  $L_{CE3}$ 。每个网络的总损失定义如下：

$$L_{student\#1} = L_{CE1} + \mu_1 L_{MSE2} \quad (3.8)$$

$$L_{student\#2} = L_{CE2} + \mu_2 L_{MSE1} + \mu_3 L_{MSE3} \quad (3.9)$$

$$L_{correction} = L_{CE3} \quad (3.10)$$

其中  $\mu_1$ 、 $\mu_2$ 、 $\mu_3$  为每个单项损失的权重，为了保证教师网络对两个学生网络的公平性，我们将  $\mu_1$  和  $\mu_3$  设置为相等的值。

## 3.5 实验

### 3.5.1 实验环境及训练策略

本章实验所使用的硬件环境为 NVIDIA GeForce 3090 显卡，内存容量为 24GB，软件环境是基于 Linux 18.04 操作系统，使用 Python 3.7 版本的 Pytorch 1.11 进行开发。

实验采用了五折交叉验证的方法进行训练。具体而言，我们将数据集随机平均划分为 5 个子集，在每次实验中，其中 4 个子集被用作训练集，而剩下的 1 个子集则作为测试集。通过循环测试不同的子集，我们得到了教师网络的五个测试结果，并将这五个测试结果的均值作为最终的实验结果。在实验过程中，我们使用 NAdam 优化器，并将学习率设置为 0.001，最终的分类标签则是通过应用 0.5 作为阈值来确定的。

### 3.5.2 评估指标

本文中所用到的模型评价指标主要有 ROC 曲线下面积 (AUC)、准确度 (Accuracy)、

敏感性 (Sensitivity)、特异性 (Specificity)、阳性预测值 (PPV)、阴性预测值 (NPV) 以及 Kaplan-Meier 曲线, 本节将对此进行详细介绍。

### (1) 混淆矩阵

混淆矩阵是一种广泛使用的评估分类方法的工具, 如表 3.2 所示, 它能直观地表达模型的性能, 展示真实与预测类别的情况。

表 3.2 混淆矩阵

真实类别	预测类别	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

在表 3.2 中, TP (True positive) 表示将正例预测为正例的样本数; FP (False positive) 为将反例预测为正例的样本数; FN (False negative) 为将正例预测为反例的样本数; TN (True negative) 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根据混淆矩阵中的要素表示出来。

### (2) 准确度 (Accuracy)、敏感性 (Sensitivity)、特异性 (Specificity)、阳性预测值 (PPV)、阴性预测值 (NPV)

Accuracy 反映了分类器对整体的判断能力, 计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (3.11)$$

Sensitivity 表示真实正类中被正确预测为正类的样本数占总的真实正类样本数的比例, 计算公式如下:

$$Sensitivity = \frac{TP}{TP + FN} \quad (3.12)$$

Specificity 表示真实负类中被正确预测为负类的样本数占总的真实负类样本数的比例, 计算公式如下:

$$Specificity = \frac{TN}{FP + TN} \quad (3.13)$$

PPV 表示真阳性的阳性预测比例, 计算公式如下:

$$PPV = \frac{TP}{TP + FP} \quad (3.14)$$

NPV 表示真阴性的阴性预测比例, 计算公式如下:

$$NPV = \frac{TN}{FN + TN} \quad (3.15)$$

### (3) ROC 曲线下面积 (AUC)

ROC<sup>[52]</sup> 曲线是根据一系列不同的二分类方式(分界值或决定阈),以真阳性 TPR 为纵坐标,假阳性率 FPR 为横坐标绘制的曲线。

$$TPR = \frac{TP}{TP + FN} \quad (3.16)$$

$$FPR = \frac{FP}{FP + TN} \quad (3.17)$$

AUC 是 ROC 曲线下的面积,反映了分类器对正负样本的分类能力。AUC 值越接近 1,分类器的性能越好<sup>[53]</sup>。

### (4) Kaplan-Meier 曲线

Kaplan-Meier 曲线<sup>[54]</sup>又称生存曲线,是一种经典的生存分析技术,主要分析单一因素对生存期的影响,用于绘制描述患者数据生存时间分布的生存曲线。生存曲线以生存时间为横轴,生存率为纵轴,绘制而成的连续型的阶梯形曲线,用以说明生存时间与生存率之间的关系。生存曲线显示了一段时间内患者的生存状态,对于评估不同变量对生存时间的影响非常有帮助。

## 3.5.3 实验结果分析

本研究进行了以下实验来证明该淋巴结转移诊断模型 DSTC 的有效性:(1)对比实验,我们与六种现有方法在淋巴结转移分类上进行了比较。(2)与医生诊断结果的具体比较,我们利用混淆矩阵和 Kaplan-Meier 生存曲线对两者进行比较分析;(3)消融实验,我们分析了本章的模型架构中每个加入的组件所带来的性能变化。

### (1) 对比实验

首先与六种现有方法在淋巴结转移分类上进行了比较。Mean Teachers<sup>[49]</sup> 和 FixMatch<sup>[55]</sup> 被提出用于半监督分类,而其余四种方法 Co-teaching<sup>[56]</sup>、JoCoR<sup>[35]</sup>、Jo-SRC<sup>[57]</sup> 和 PNP<sup>[58]</sup> 则侧重于处理标签中可能存在错误的标签噪声。Mean Teachers 方法中使用了两个模型,学生模型是主要的训练模型,教师模型则是一个平均了学生模型权重的模型,无标签样本通过使用教师模型生成的目标来训练学生模型;FixMatch 使用弱增强对无标记图像进行预测生成伪标签,并仅在产生高置信度预测时保留伪标签,然后使用强增强对同一无标记图像进行预测,并将预测结果与伪标签计算损失,以此训练模型;Co-teaching 使用双重网络框架来识别具有可靠标签的数据实例,并促进网络之间的知

识共享；JoCoR 采用联合训练的方法来减少两个网络之间的差异，解决深度学习中的噪声标签问题；Jo-SRC 使用每个样本的两种不同视图来预测其成为干净样本或超出分布的概率，并集成一致性正则化以提高模型的生成能力；PNP 通过使用两网络方法从噪声标签中实现鲁棒学习，一个网络负责预测类别标签，另一个网络预测标签中存在的噪声类型，使用联合损失函数进行更新。

表 3.3 对比实验结果

模型	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
Radiologists	0.79	0.47	0.86	0.45	0.89	-
Mean Teachers	0.82	0.50	0.88	0.55	0.90	0.71
FixMatch	0.70	0.63	0.71	0.33	0.92	0.65
Co-teaching	0.80	0.57	0.75	0.44	0.91	0.71
JoCoR	0.78	0.57	0.82	0.47	0.91	0.65
Jo-SRC	0.81	0.68	0.83	0.46	0.93	0.73
PNP	0.79	0.73	0.80	0.45	0.94	0.72
<b>DSTC</b>	<b>0.88</b>	<b>0.77</b>	<b>0.90</b>	<b>0.62</b>	<b>0.95</b>	<b>0.84</b>

如表 3.3 所示，本章的模型 DSTC 在各项指标上都达到了最高。放射科医生提供的影像标签是二值形式，不能用于计算 AUC。在表 3.3 中，放射科医生对所有患者诊断的准确度为 0.79，敏感性为 0.47，特异性为 0.86，PPV 为 0.45，NPV 为 0.89。PNP 在敏感性上位于第二名，以 0.04 的差距低于本章方法；Mean Teachers 在特异性上位于第二名，以 0.02 的差距低于 DSTC。上述方法的准确度分布在 0.70-0.82，敏感性在 0.50-0.73，AUC 在 0.65-0.73。本章的 DSTC 将准确度提高到 0.88，敏感性提高到 0.77，特异性提高到 0.90，PPV 提高到 0.62，NPV 提高到 0.95，AUC 提高到 0.84，显著优于的医生诊断结果以及其他六种方法。图 3.6 中的 ROC 曲线也对比了各方法之间的性能。



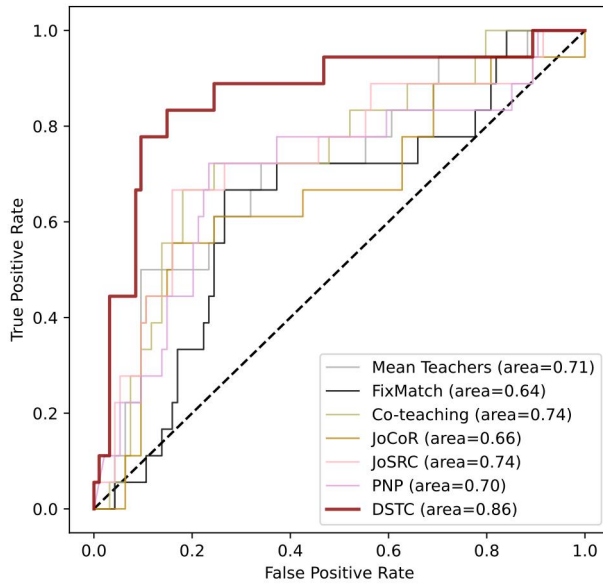


图 3.6 对比实验的 ROC 曲线

(2) 与医生诊断结果的具体比较

如图 3.7 所示，左边为医生诊断的混淆矩阵，右边为本章 DSTC 模型预测结果的混淆矩阵。

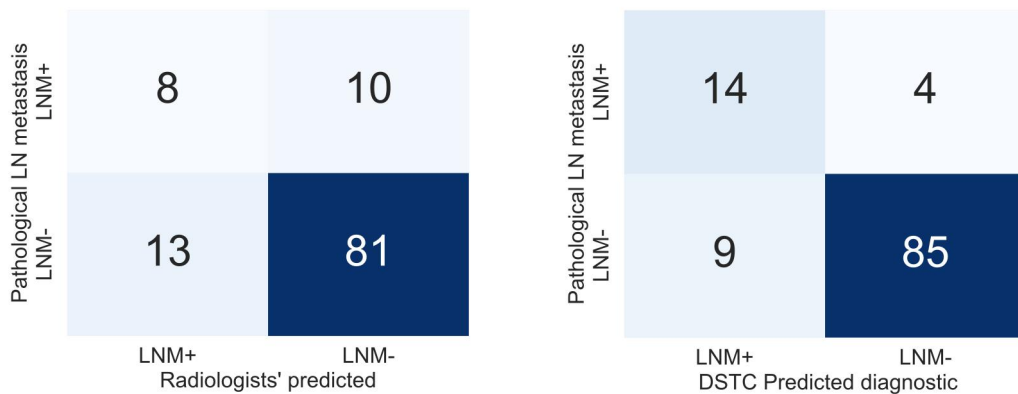


图 3.7 混淆矩阵

在研究对象中，有 18 例患者发生淋巴结转移，其中医生正确诊断了 8 例，错误诊断了 10 例；而在 94 例未发生转移的病例中，医生正确诊断了 81 例，错误诊断了 13 例。本章 DSTC 模型预测结果显示，在有转移的病例中正确预测了 14 例，错误预测了 4 例；在无转移的病例中，正确预测了 85 例，错误预测了 9 例。不论病例是否存在淋巴结转移，DSTC 模型的预测准确率均优于医生的诊断结果。

此外, 基于放射科医生提供的影像标签和本章所提方法的预测标签, 我们绘制了 328 例患者的 Kaplan-Meier 生存曲线, 如图 3.8 所示。

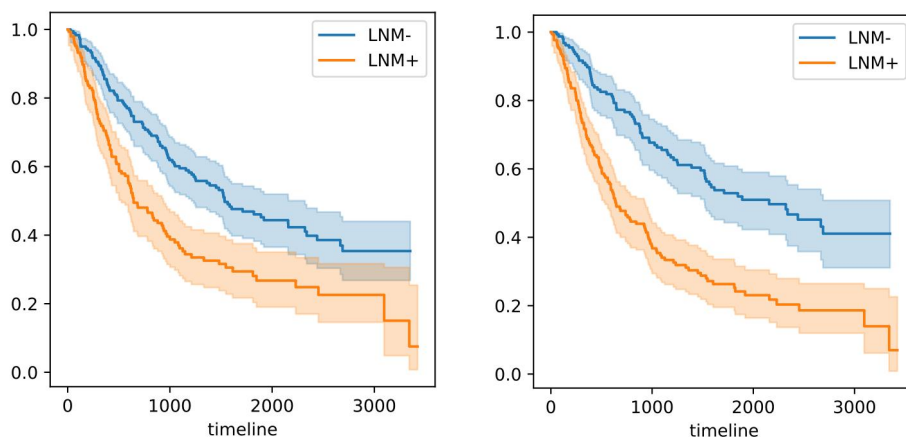


图 3.8 Kaplan-Meier 曲线

在图 3.8 中, 左边是根据影像标签绘制的 Kaplan-Meier 曲线, 右边是根据预测标签绘制的 Kaplan-Meier 曲线。结果显示, 两组曲线的  $p$  值均小于 0.005, 但使用预测标签的两组患者之间的差异更为显著, 且对比两边的 Kaplan-Meier 曲线, 无论是否淋巴结转移, 生存率都有所提升, 这表明该模型可以提供更有用的预测价值。

### (3) 消融实验

为了进一步验证本文方法核心模块的有效性, 进行了消融实验, 其结果如表 3.4 所示。设置 1: 完整的 DSTC 模型; 设置 2: 去掉纠正网络; 设置 3: 去掉单向学习; 设置 4: 同时去掉单向学习和纠正网络; 设置 5: 仅使用有病理标签的数据进行监督学习。

表 3.4 消融实验结果

模型	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
DSTC	<b>0.88</b>	<b>0.77</b>	<b>0.90</b>	<b>0.62</b>	<b>0.95</b>	<b>0.84</b>
无纠正网络	0.80	0.63	0.83	0.50	0.92	0.68
无单向学习	0.80	0.75	0.81	0.47	0.95	0.74
无单向学习和纠正网络	0.83	0.47	0.89	0.53	0.90	0.73
全监督	0.81	0.50	0.87	0.45	0.90	0.63

如表 3.4, 在设置 2 中把纠正网络移除后, 学生网络#2 使用原始的影像标签进行训练,

准确率下降了 0.08；在设置 3 中去除单向学习后，直接计算两个学生网络之间所有输出的一致性损失参与每个学生网络的损失回传，AUC 比完整的方法下降了 0.1；在设置 4，单向学习和纠正网络均被移除，敏感性比完整模型下降了 0.3，这是对对照组中下降最多的设置；在设置 5 中使用完全监督策略训练，即仅使用具有病理标签的 112 例患者进行训练，网络结构与完整方法中的教师网络一致，准确率下降了 0.07，敏感性下降了 0.27，AUC 下降了 0.21。这些指标得分证明了这些核心模块在 DSTC 模型中可以有效提升淋巴结转移诊断性能。图 4.9 中对应的 ROC 曲线展示了消融实验的结果。

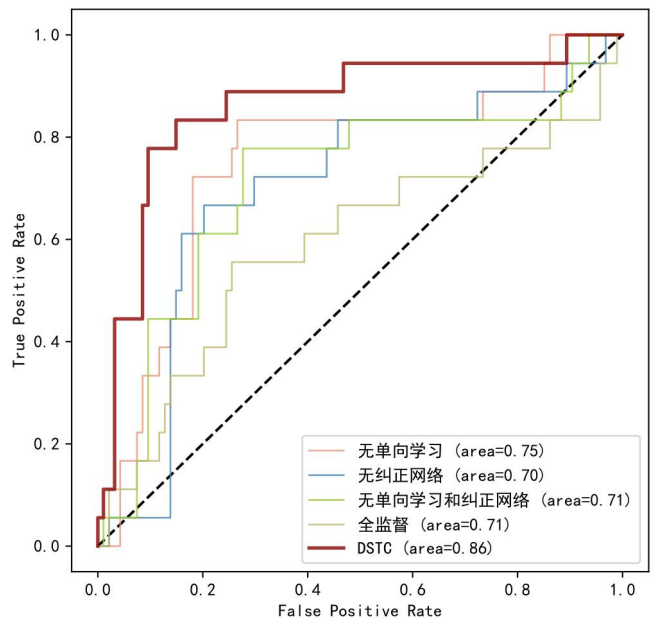


图 3.9 消融实验的 ROC 曲线

此外，我们进行了几个实验来确定 3.4 节中两个学生网络总损失函数的超参数  $\mu_1$  和  $\mu_2$  的选择，因为  $\mu_1$  和  $\mu_3$  是相等的，所以表中只列出了  $\mu_1$  的值。结果如表 3.4 所示，最终设置  $\mu_1$  和  $\mu_3$  为 0.8， $\mu_2$  为 0.6。

表 3.5 超参数的消融实验

$\mu_1$	$\mu_2$	Accuracy	Sensitivity
1	1	0.77	0.57
1	0.8	0.80	0.70
1	0.6	0.85	0.75
1	0.4	0.82	0.62

续表 3.5

$\mu_1$	$\mu_2$	Accuracy	Sensitivity
<b>0.8</b>	<b>0.6</b>	<b>0.88</b>	<b>0.84</b>
0.6	0.6	0.84	0.70

### 3.6 讨论

本研究介绍了一种基于半监督策略的深度学习模型 DSTC，对非小细胞肺癌患者的淋巴结转移诊断问题进行了改进。该模型成功地将 PET/CT 图像、病理标签和影像标签有效地结合起来，能够准确地区分患者是否存在淋巴结转移，并证明了这种学习方式的有效性。

当前，在研究淋巴结转移时，很多研究都使用病理标签作为学习目标，因为病理标签被认为是诊断的金标准。但是病理标签并不总是可获得的，这使得放射科医生提供的影像标签对一些患者来说是一个替代选择。值得注意的是，由于放射科医生的诊断能力或医学图像质量问题，影像标签的准确性可能并不高。对于这些噪声标签，目前医学领域主要采用标签平滑的方法来处理，或者直接忽略带有噪声的样本。然而，这种处理方式可能导致信息丢失或模型性能下降。

为了结合病理和影像标签来诊断淋巴结转移，我们在模型中使用两个学生网络以这两种类型标签为学习目标来提取图像特征，并且通过单向学习促进知识共享。此外在学生网络权重的基础上建立教师网络，鼓励学生网络进行一致性持续学习。在对比实验中，仅以病理标签或影像标签为学习目标的模型表现并不好，将全监督深度学习模型（仅使用具有病理标签的数据）、无纠正网络深度学习模型、无单向学习深度学习模型等性能进行比较。根据消融研究的结果，当不使用单向学习时，敏感性评分从 0.77 下降到 0.75。影像标签的潜在不准确性可能降低淋巴结转移诊断的准确性。因此，我们采用半监督策略，在病理标签的指导下纠正影像标签。在没有使用纠正网络的消融实验中，导致灵敏度评分从 0.85 下降到 0.63。这表明校正网络在提高 LN 转移诊断任务的准确性方面起着至关重要的作用。通过比较的结果可知，以病理标签学习结果指导影像标签学习的单向学习过程和半监督纠正影像标签的网络的加入，使模型的效果达到最优，可能是因为影像标签数据在学习过程中对信息的进一步补充，以及纠正网络对影像标签的细化，这都证明了每一个模块对于模型性能的提升。在对比医生诊断的混淆矩阵后，我们发现在有

淋巴结转移的情况下，本章实验的 DSTC 模型比医生多诊断正确了 6 例，这说明本方法能够更准确地诊断出患有淋巴结转移的患者。通过及时对可能存在转移风险的患者进行进一步检查，可以早期发现异常情况，并制定有效治疗方案，从而避免肿瘤的进一步扩散和转移，提高患者的生存率，改善其生活质量。

目前的研究具有一些局限性。首先，由于这项研究的回顾性性质，未来需要前瞻性研究来验证这个工作。其次，我们的工作是小样本，因此采用五折交叉验证来训练深度学习模型，未来的研究将利用来自多中心研究的外部验证集。最后，我们对于图像数据的处理还需要医生手动勾画肿瘤区域，未来将采用自动分割或无分割方法。

### 3.7 本章小结

在本章中，我们介绍了一种基于 DSTC 的半监督非小细胞肺癌患者淋巴结转移诊断模型，并详细介绍了实验的整体流程。我们对 DSTC 模型的架构以及每个核心模块进行了详细描述，并展示和讨论了实验结果。通过进行对比实验和消融实验，验证了 DSTC 模型的良好性能。该框架成功地将病理标签和影像标签进行联合训练，能够有效帮助放射科医生提高诊断准确性，并进一步改善患者的生存状况。

## 4 基于多尺度多模态融合网络的无肿瘤分割淋巴结转移诊断

在上一章所使用的深度学习模型，虽然针对非小细胞肺癌患者的淋巴结转移预研取得了超过放射科医生的实验结果，但无论是完全监督学习是第四章的半监督深度学习模型都依赖于大量有标注的数据集，即需要由医生手动勾画的肿瘤区域，并没有减少医生的工作量。因此本章采用了一种无需勾画肿瘤区域的自动预测非小细胞肺癌的淋巴结转移情况的方法，利用 2D MIP 图像和 3D PET/CT 图像作为输入，使用 ResNet 作为主干网络并引入了融合模块构建深度学习网络，通过半监督学习算法训练自动诊断淋巴结转移的深度多模态融合模型。

### 4.1 研究流程

为了更好的提取到患者肿瘤区域之外的全身特征，本章选用患者的全身 PET/CT 的 MIP 图像作为深度学习模型的输入，同时将半身 PET/CT 作为另一个输入以提供深层次高维特征。实验基于 ResNet 构建分类网络，引入融合模块，目的是从中间的特征层对多模态信息进行融合。实验的整体流程如图 4.1 所示。

首先是数据预处理，生成全身 MIP 图像和截取半身 3D PET/CT 图像；然后以 ResNet 为主干网络，并在中间插入特征融合模块，构建深度学习模型中的两个独立分支；再将两个网络输出的特征进行堆叠，输入全连接网络(FC)进行降维后，经过 softmax 激活函数得到预测概率；最后，使用同样的分类评估指标对模型进行评价与分析。

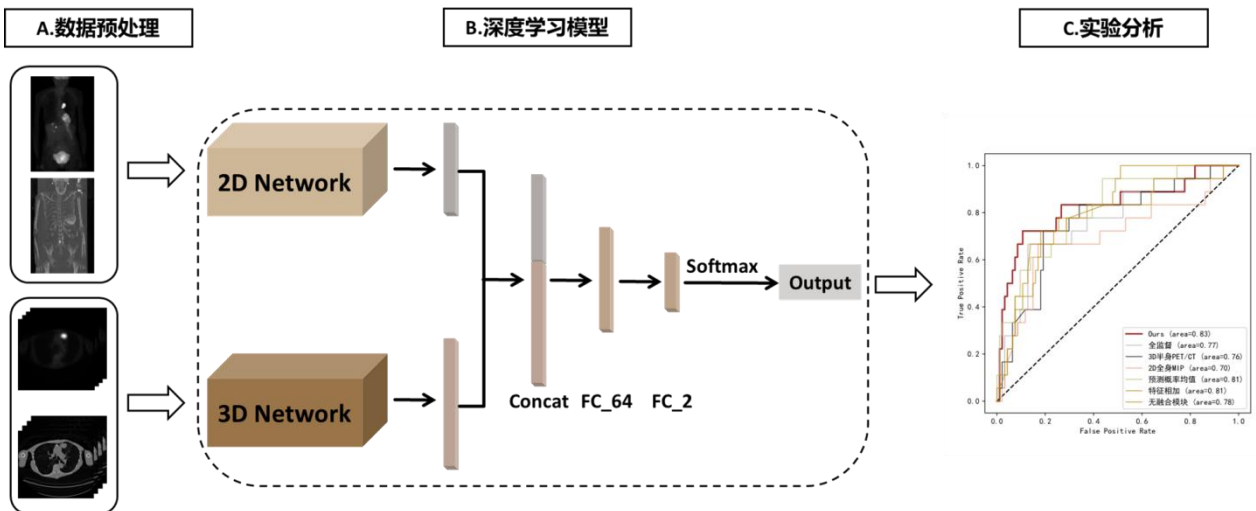


图 4.1 整体实验流程图

## 4.2 实验数据及预处理

本实验所用研究数据来源与第三章所用数据相同，具体相关介绍见本文 3.2.1 小结关于实验数据的介绍部分。

在利用多尺度图像对非小细胞肺癌患者进行淋巴结转移预测时，依据配准后的 PET/CT 图像，从正面分别得到两个视图的 MIP 图像，并将它们调整为  $192 \times 128$  大小进行归一化。另外，由于淋巴结转移通常只发生在患者的上半身，为了减少计算负担并提高效率，我们截取了半身 PET/CT 图像，调整为  $64 \times 64 \times 64$  大小进行了归一化处理。与上一部分的工作不同的是，在无肿瘤分割的非小细胞肺癌患者的淋巴结转移的研究中，本研究仅采纳了 112 例患者的病理标签，328 例患者的影像标签数据并未参与训练过程，因此在这一部分的研究中采用的是半监督方法，为没有病理标签的患者生成伪标签。

## 4.3 深度学习模型架构

在本研究中，我们包含了没有病理标签的患者。对于具有病理标签的数据，我们使用正常的监督学习方法进行训练，将输出结果与病理标签进行交叉熵损失计算。而对于其他没有病理标签的患者，我们采用半监督学习方法，参考图 4.2 来进行处理。首先，我们对原始的 2D 和 3D 图像进行特征提取。通过设定阈值为 0.5，我们可以生成伪标签，并利用这些预测结果。然后，我们对所有图像进行弱增强，并将这些增强后的图像与对应的伪标签一起进行交叉熵损失计算。

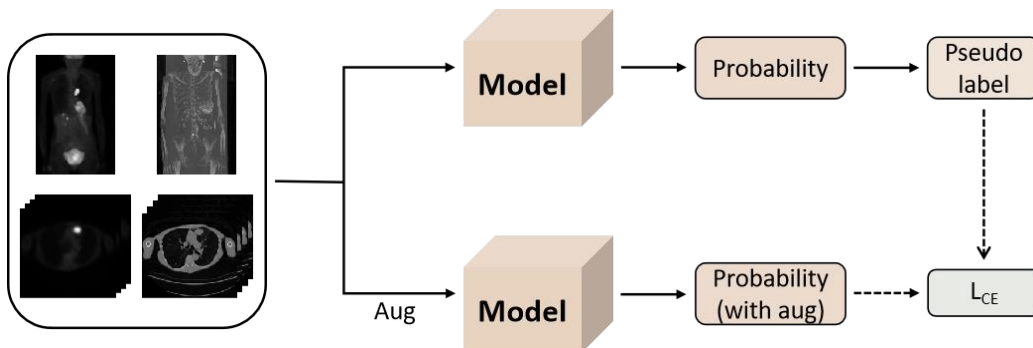


图 4.2 半监督学习过程

特征提取在无分割预测淋巴结转移任务中扮演着关键的角色。根据 Meng 等人<sup>[59]</sup>

的工作，他们证明了中间融合策略可以允许信息在多个尺度上进行融合，从而获得更好的性能。因此，我们在特征提取网络中引入了融合（Fusion）模块，以融合不同模态的信息和上下文信息。特征提取分为两个部分。第一部分是通 2D 网络对 MIP 图像进行特征提取，目的是利用到患者的全身特征。第二部分是通 3D 网络对半身 PET/CT 图像进行特征提取，使用半身图像可以提高运行速度和减少冗余信息。具体的特征提取网络结构如图 4.3 所示。

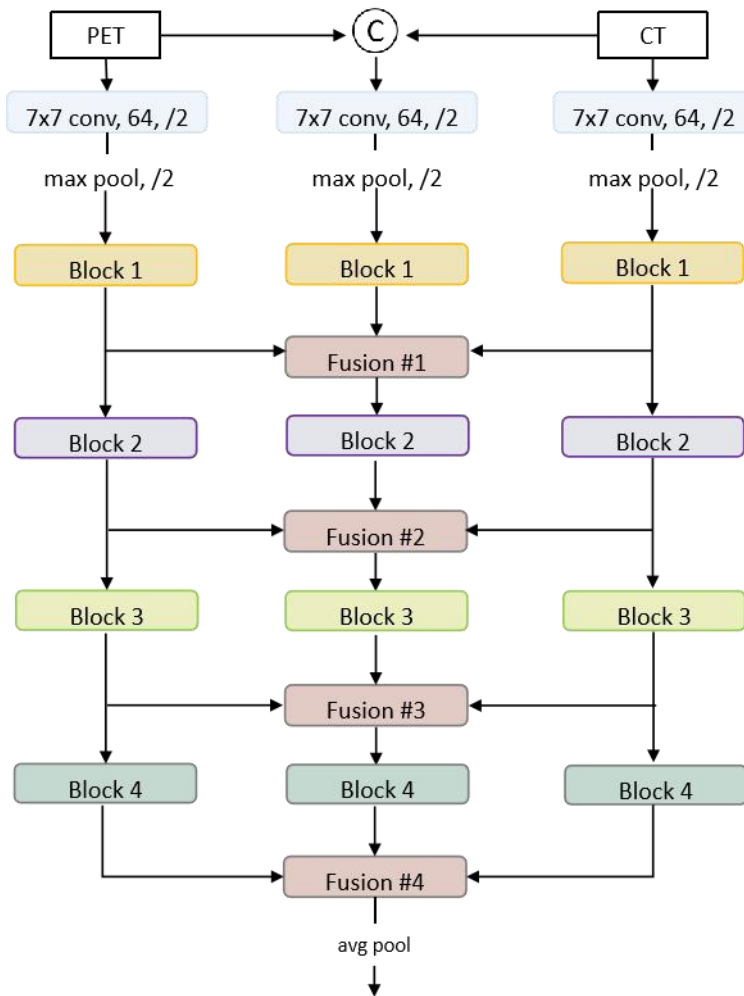


图 4.3 特征提取网络

特征提取网络包括三个分支用于特征提取和融合，左右两个分支命名为 PET 和 CT，分别从其半身图像或对应的 MIP 图像中提取特征，分支中包括四组 ResNet 卷积块（Block）；中间的分支名为 Concat，从堆叠的图像中提取特征，分支中包括四组卷积块和融合模块。首先通过一个卷积层、最大池化层和第一个卷积块对三个图像提取特征。然后使用融合模块将 PET 和 CT 特征与 Concat 的特征进行融合，再将融合模块输出



的特征传递到中间分支的下一组卷积块。上述过程重复四次，最后输出第四个融合模块的特征，进行后续全局平均池化、降维和 softmax，得到淋巴结转移预测结果。

如图 4.4，是融合模块的详细结构。 $F_{PET}^i$ 、 $F_{CT}^i$  和  $F_{Concat}^i$  是 PET、CT 和 Concat 分支的第  $i$  个卷积块输出的特征；先将  $F_{PET}^i$  和  $F_{CT}^i$  经过一层卷积，然后进行批处理归一化 (Batch Normalization, BN) 和 ReLU 激活函数，得到特征  $F_{PC}^i$ ；再将  $F_{PC}^i$  和  $F_{Concat}^i$  进行堆叠，重复上述卷积操作，得到最终的融合特征  $F_{fusion}^i$ 。

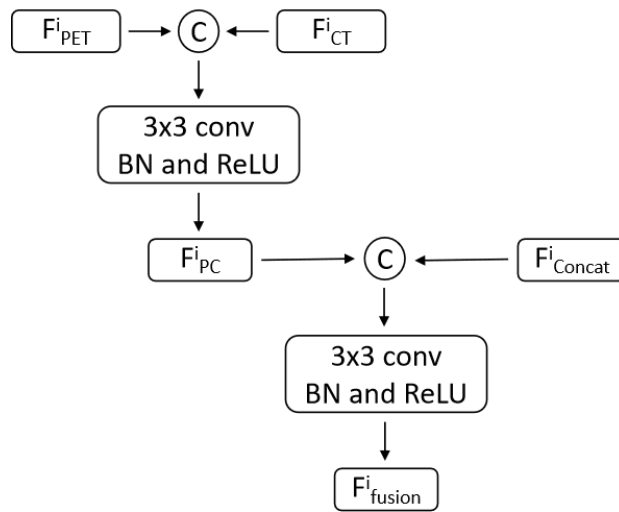


图 4.4 融合 (Fusion) 模块

通过上述步骤，完成对 2D 和 3D 图像的特征提取和融合，四次融合模块的应用实现了不同模态信息及上下文信息的融合。最后将整个特征提取网络输出的融合特征，经过两个全连接层和 softmax 函数，可以得到患者最终分类为转移或未转移。

## 4.4 实验

### 4.4.1 训练过程和实验环境

本研究涵盖了 328 例非小细胞肺癌患者，接受了病理检查，其中 18 例发生了淋巴结转移，94 例未发生淋巴结转移，其余的患者没有使用影像标签信息。在实验过程中，使用五折交叉验证的方法进行半监督训练，使用 NAdam 优化器，学习率设置为 0.001，以 0.5 作为阈值确定最终的分类标签。

在训练过程中，我们使用了交叉熵损失函数分别计算监督学习的损失和半监督学习

的算是，将预测的分类分数分别与病理标签和伪标签进行了比较。具体的损失函数可以写为：

$$L = L_{labeled} + \alpha L_{unlabeled} \tag{4.1}$$

$$L_{labeled} = \frac{1}{K} \sum_{i=0}^{C-1} \sum_{n=0}^{K-1} y_n(i) \log(p_n(i)) \tag{4.2}$$

$$L_{unlabeled} = \frac{1}{M-K} \sum_{i=0}^{C-1} \sum_{n=0}^{M-K-1} \bar{y}_n(i) \log(p_n(i)) \tag{4.3}$$

其中，C 为分类数(这里 C=2 为转移和未转移两类)，K 为有病理标签的患者数，M 为患者总数；y 为病理标签， $\bar{y}$  为伪标签；p 表示模型预测的患者属于 i 类的预测概率； $\alpha$  为损失的权重。

本章使用与第四章相同的实验环境，详细实验环境介绍见 3.5.1 节。

#### 4.4.2 实验结果分析

本研究中所用到的模型评价指标主要有 AUC 值、准确度 (Accuracy)、敏感性 (Sensitivity)、特异性 (Specificity)、阳性预测值 (PPV)、阴性预测值 (NPV)，详细介绍及公式见本文 3.5.2 小结。

(1) 与医生诊断结果的具体比较

对具备病理标签的 112 例患者，我们通过混淆矩阵与医生诊断进行对比。

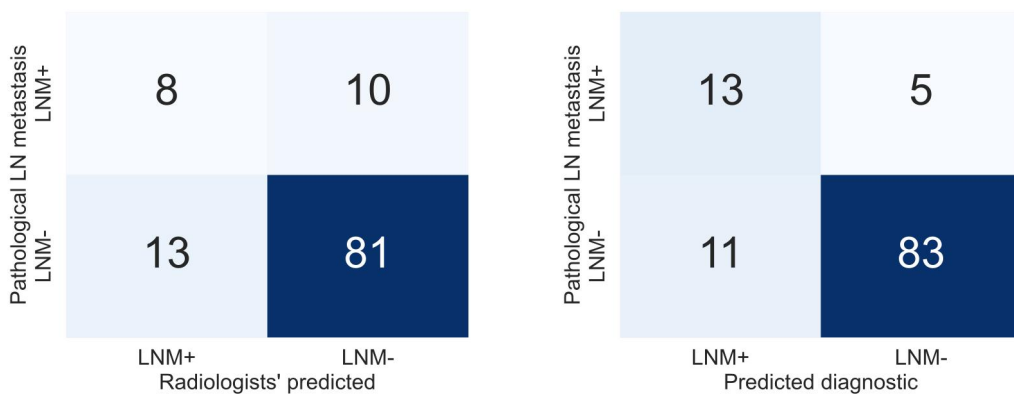


图 4.5 混淆矩阵

如图 4.5 所示，左边为医生诊断的混淆矩阵，右边为本章实验的混淆矩阵，实验结果表明，在处理淋巴结转移病例时，本章方法相比于医生诊断更为准确。在有转移的病

例中，本章方法成功预测了 13 例，仅有 5 例错误，而医生只成功诊断了其中的 8 例，错误了 10 例。同时，在未发生转移的病例中，本章方法也表现出色，成功预测了 83 例，错误 11 例，而医生只成功诊断了 81 例，错误了 13 例。综合来看，无论是有转移还是没有转移的情况下，本章方法都比医生的诊断更为可靠和准确。这表明我们的方法在淋巴结转移的诊断中具有潜在的临床应用前景。具体评估指标数值如表 4.1 所示。

表 4.1 与医生诊断的指标对比

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
Radiologists	0.79	0.47	0.86	0.45	0.89	-
<b>Ours</b>	<b>0.86</b>	<b>0.72</b>	<b>0.89</b>	<b>0.57</b>	<b>0.94</b>	<b>0.81</b>

### (2) 不同输入的消融实验

实验对比了不同模态作为输入的分类性能，在相同的特征提取网络，仅以 3D 半身 PET/CT 图像或 2D 全身 MIP 图像作为输入，结果展示在表 4.2 中。

在表 4.2 中，仅使用深度学习提取 3D 半身 PET/CT 特征得到的准确率和 AUC 分别为 0.78 和 0.80，在发生淋巴结转移的病例中虽然取得了提升，但是在未发生转移的病例中不如医生诊断；仅使用深度学习提取 2D 全身 MIP 特征得到的准确率和 AUC 分别为 0.80 和 0.71，虽然 AUC 有所提升，但是特异性没有超过医生诊断结果；将两种图像结合得到了 0.86 的准确率和 0.81 的 AUC，取得了较大的提升，这证明了将 3D PET/CT 图像特征与 2D MIP 图像特征结合的有效性。

表 4.2 不同输入的消融实验

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
3D 半身 PET/CT	0.78	0.62	0.81	0.41	0.91	0.80
2D 全身 MIP	0.80	0.67	0.82	0.43	0.93	0.71
<b>2D + 3D</b>	<b>0.86</b>	<b>0.72</b>	<b>0.89</b>	<b>0.57</b>	<b>0.94</b>	<b>0.81</b>

### (3) 不同学习方式的消融实验

在相同的特征提取网络下，仅以 112 例有病理标签的数据进行完全监督实验，准确度为 0.83，AUC 为 0.77，详见表 4.3 中数值，虽然预测结果比医生略好，但效果不如

半监督学习。

表 4.3 不同学习方式的消融实验

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
全监督	0.83	0.55	0.88	0.55	0.91	0.77
<b>半监督</b>	<b>0.86</b>	<b>0.72</b>	<b>0.89</b>	<b>0.57</b>	<b>0.94</b>	<b>0.81</b>

通过整合无标签的数据，半监督学习技术不仅能够提升数据样本的多样性，而且有助于构建更复杂的数据分布模型，能够最大限度地利用现有的数据资源，不仅提高了分类算法的准确性，还增强了其鲁棒性，使得模型能够更好地适应各种分类任务。

#### (4) 不同特征融合方式的消融实验

在该研究中，我们采用了一种堆叠的方式来融合 2D 和 3D 特征提取网络输出的特征向量。我们对比了三种融合策略：①将 2D 和 3D 特征提取网络输出的特征向量分别输入两个全连接层进行降维，然后对由 softmax 激活函数输出的预测概率取均值<sup>[60]</sup>；②将 2D 和 3D 特征提取网络输出的特征向量按通道维度进行相加<sup>[61]</sup>，并取平均值，然后再进行降维和分类；③去除特征提取网络中堆叠分支的融合模块，对三个分支的特征进行拼接并降维至 64，再将两个维度图像的特征进行融合、降维和分类。

表 4.4 不同特征融合方式的消融实验

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
预测概率均值	0.81	0.6	0.85	0.43	0.92	0.79
特征相加	0.78	0.50	0.83	0.41	0.90	0.81
<b>特征堆叠</b>	<b>0.86</b>	<b>0.72</b>	<b>0.89</b>	<b>0.57</b>	<b>0.94</b>	<b>0.81</b>
无融合模块	0.83	0.67	0.87	0.53	0.93	0.77

通过表 4.4 可以看出，直接对特征向量进行堆叠可以取得更好的结果，而融合模块也对模型的预测能力具有提升作用。

最后，在图 4.6 中展示了所有消融实验的 ROC 曲线，可以看到本章实验的模型 AUC 最高。

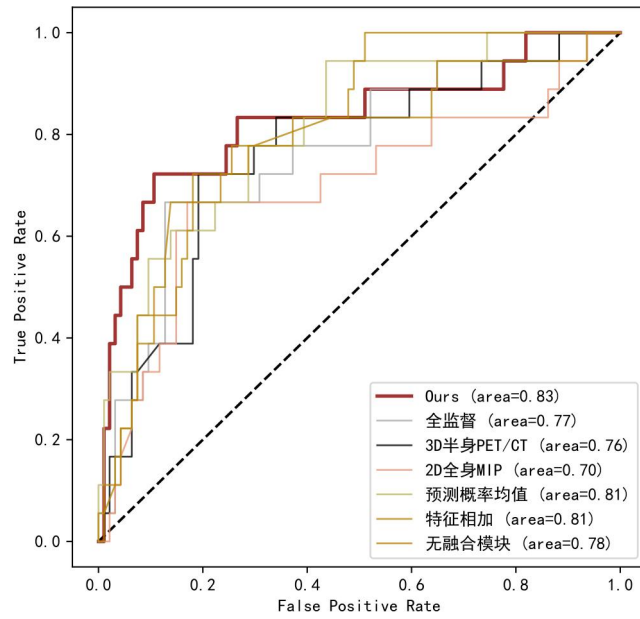


图 4.6 各消融实验 ROC 曲线

## 4.5 讨论

本研究开发了一种无肿瘤分割的、多尺度多模态的半监督深度学习模型，用于预测非小细胞肺癌患者的淋巴结转移状态。随着深度神经网络的不断发展，深度学习逐渐向多模态技术迈进。多模态学习指建立模型使机器从多模态中学习不同模态的信息，并实现模态之间的信息交流和转换。模态本质上是信息传递的通道，这些来自不同来源的数据在语义上相互关联，并且有时还能提供互补的信息。利用多模态数据不仅可以改善神经网络的性能，还能更好地提取特征，从而在更大范围内进行预测。因此，本研究将多模态深度学习模型应用于非小细胞肺癌患者的淋巴结转移状态的预测，无需进行肿瘤分割操作。

对于部分没有病理标签的数据，我们通过半监督中的伪标签方法进行训练，即使用原始图像的预测标签形成伪标签，与弱增强图像的预测进行交叉熵损失比较，与全监督实验相比，取得了很大提高。对于多模态数据，我们用两个网络分别对 2D 和 3D 影像进行特征学习，每个网络又分为三个分支对 PET、CT 以及堆叠图像提取深层特征。在堆叠图像的分支上，我们对 ResNet-18 的每个残差块之后加入了融合模块，以融合三个分支的特征，融合不同模态的信息和上下文信息。通过结合 2D、3D 影像，实验结果表明，综合 PET/CT 和 MIP 深度学习特征的淋巴结转移预测准确率比单独使用任何一种影像都有提高。

针对这两章实验的预测结果，模型在预测发生淋巴结转移的病例时，其表现明显超越了医生，而在预测未发生转移的病例时，其表现则与医生大致相当。这种显著的差异，我们认为，可能与数据集的分布特点、模型的训练策略以及医生和模型在诊断过程中的不同侧重点和思维模式有着紧密的关联。具体而言，由于发生淋巴结转移的病例在数据集中可能更为显著，或者它们本身具备某些易于辨识的特征，使得模型在训练过程中能够相对容易地捕捉并学习到这些关键信息。因此，在进行预测时，模型能够展现出较高的准确率和优越性。然而，在预测未发生淋巴结转移的病例时，模型的表现可能存在一定的局限。这可能是由于数据集中未发生转移病例的多样性和复杂性相对较低，导致模型难以充分学习到所有相关的特征和信息。此外，医学图像的解读和诊断往往需要丰富的医学知识和临床经验，而模型在这方面可能还有待提升和完善。

## 4.6 本章小结

在本章中，我们以非小细胞肺癌淋巴结转移诊断为研究对象，详细描述了一种利用多尺度、多模态图像数据在 ResNet-18 中加入融合模块的半监督深度学习方法，用于预测无肿瘤分割的淋巴结转移。

通过对比实验和消融实验，验证了本章深度学习模型的良好性能，该框架可以在无需分割肿瘤的情况下实现淋巴结转移的预测，从而帮助放射科医生提高诊断准确性，减少医生的工作量，具有较大的现实意义。

## 5 总结与展望

### 5.1 总结

肺癌是全球第二常见的癌症，并且是导致癌症相关死亡的主要原因。在非小细胞肺癌这一最常见的类型中，淋巴结转移在其诊断中具有重要的意义，对分期确定和治疗决策等方面都产生着重要影响。病理诊断是对疾病进行本质判断的“金标准”，然而，并非所有患者都能接受病理诊断。因此，有必要开发基于半监督深度学习的方法来辅助预测非小细胞肺癌患者的淋巴结转移情况。

本文研究的主要内容是基于 PET/CT 图像，使用半监督深度学习模型对非小细胞肺癌患者的淋巴结转移进行诊断。首先介绍了非小细胞肺癌的研究背景及意义，并从监督学习、半监督学习和噪声学习相关工作三个方面阐述研究现状；然后详细介绍了深度学习的理论基础，包括卷积神经网络的基本构成、ResNet、半监督深度学习以及影像诊断和病理诊断；最后，运用本文的半监督深度学习模型解决非小细胞肺癌患者淋巴结转移分类问题。具体的工作内容如下：

(1) 基于 DSTC 的半监督淋巴结转移诊断。为了充分利用患者的影像标签信息，我们使用两个学生网络分别以病理标签和影像标签作为目标，设计了单向学习模块让病理标签能够指导影像标签学生网络的学习。然后基于两个学生网络的权重的指数移动平均值构建教师网络，它总结了学生网络在训练过程中获得的知识，并通过一致性学习帮助学生网络更好地进行泛化和推理。由于影像标签会存在一定的不准确性，我们加入了以深度特征和影像标签为输入的校正网络，通过半监督学习策略让病理标签指导影像标签的细化，新的标签继续对学生网络训练。通过一系列的对比实验和消融实验，验证了 DSTC 模型的良好性能，证明了每一个模块对于模型性能的提升；无论是发生转移还是没有转移的情况下，该模型的预测都优于医生诊断。

(2) 基于多尺度多模态融合网络的无肿瘤分割淋巴结转移诊断。医生勾画肿瘤区域极大的消耗了人力、物力和财力，因而本章考虑在无需医生勾画肿瘤区域、无需自动分割的情况下，使用半监督深度学习的方法对非小细胞肺癌患者的淋巴结转移情况进行预测。对于研究数据中没有病理标签的患者，我们采用半监督学习中的伪标签方法，以原始图像的预测标签与弱增强图像的预测进行交叉熵损失比较；对于有病理标签的患者，依旧是利用病理标签进行完全监督学习。我们的输入采用 PET/CT 三维半身图像和二

维正面最大密度投影，分别构建二维、三维图像的特征提取网络，包括 PET、CT 和堆叠三个分支。为了融合不同模态和上下文信息，我们在堆叠分支中设计了融合模块，将三个分支的特征在网络中间多次融合。最后将三维深度学习特征和二维正面最大强度投影的深度学习特征进行融合，并通过多个全连接层进行分类，得到发生淋巴结转移的概率。整个诊断的过程在测试时不需要医生对肿瘤区域的勾画，实验结果也显示出了良好性能，取得了与上一模型相近的结果。通过消融实验，验证了多尺度图像的特征融合、融合模块的有效性，同时也展现出了半监督学习策略在医学领域的应用价值，充分利用有限的标记数据和大量的未标记数据，显著提升分类算法的准确性和鲁棒性。

## 5.2 展望

尽管本文研究在利用半监督深度学习进行非小细胞肺癌淋巴结转移诊断方面取得了不错的实验结果，但仍有许多方面值得进一步优化和探讨。

在多种标签利用方面。本文的第一个研究主要是在淋巴结转移的诊断中加入影像标签是否能够起到积极的作用，并且对影像标签进行了细化。尽管实验结果证明了这种多标签学习网络和纠正网络的有效性，但这种半监督的标签纠正方式仍然相对简单、直接。另外，在第二个无分割研究中，仅利用了半监督生成的伪标签，并未对影像标签进行处理。因此，在未来的研究中，需要改进多种标签数据的学习方法，开发更复杂、精细的标签纠正算法，以充分发挥不同标签的效果并提高对淋巴结转移的诊断准确性。

在肿瘤分割方面。本文的第二个研究在无需肿瘤分割的情况下进行淋巴结转移预测，并未针对如何改进网络的定位能力以更准确地定位肿瘤区域。因此，在接下来的研究中将探索一种新的方法，即将自动肿瘤分割作为上游任务，利用分割结果进行分类和生存分析等任务，通过这种任务之间的互相促进，我们可以实现更好的性能表现。

在淋巴结转移分期方面。本研究的主要目标是预测非小细胞肺癌患者是否会发生淋巴结转移，从而将其划分为二分类问题。然而，对于淋巴结转移的不同分期（N0、N1、N2 和 N3），对于治疗策略具有重要意义。后续工作将尝试进一步对发生淋巴结转移的患者进行具体的分类，这将有助于制定个性化的治疗计划并改善患者的预后。

最后，我们的研究数据仅限于一个医学院的数据来源，尽管我们成功地在淋巴结转移分类问题上，通过使用半监督深度学习模型提高了诊断准确度，但这并不能被视为真正的大数据。因此，为了进一步评估我们的模型在不同患者群体和成像方案中的有效性和普适性，在未来可以将研究扩展到多个医学中心的数据。



## 参考文献

- [1] Matthews H K, Bertoli C, De Bruin R A M. Cell cycle control in cancer [J]. *Nature reviews: molecular cell biology*, 2022, 23(1): 74-88.
- [2] Hanahan D. Hallmarks of Cancer: New Dimensions [J]. *Cancer Discovery*, 2022, 12(1): 31-46.
- [3] Chhikara B S, Parang K. Global cancer statistics 2022: the trends projection analysis [J]. *Chemical Biology Letters*, 2023, 10(1): 451.
- [4] Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022 [J]. *Journal of the National Cancer Center*, 2024.
- [5] Leiter A, Veluswamy R R, Wisnivesky J P. The global burden of lung cancer: current status and future trends [J]. *Nature Reviews Clinical Oncology*, 2023, 20(9): 624-639.
- [6] Lu Y, Ma T, Wang L, 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): 6.
- [7] Chen Y H, Chu S C, Wang L Y, et al. Prognostic Value of Combing Primary Tumor and Nodal Glycolytic-Volumetric Parameters of 18F-FDG PET in Patients with Non-Small Cell Lung Cancer and Regional Lymph Node Metastasis [J]. *Diagnostics*, 2021, 11(6): 1065.
- [8] Jose Vaquero 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.
- [9] Jung H. Basic Physical Principles and Clinical Applications of Computed Tomography [J]. *Progress in Medical Physics*, 2021, 32(1): 1-17.
- [10] Whatmough C, Lamb C R. Computed Tomography: Principles and Applications [J]. *Compendium on Continuing Education for the Practicing Veterinarian*, 2006, 28(11): 789-798.
- [11] Wu Y, Li P, Zhang H, et al. Diagnostic value of fluorine 18 fluorodeoxyglucose positron emission tomography/computed tomography for the detection of metastases in non-small-cell lung cancer patients [J]. *International Journal of Cancer*, 2013, 132(2): E37-E47.

- [12]Griffeth L K. Use of PET/CT Scanning in Cancer Patients: Technical and Practical Considerations [J]. Baylor University Medical Center Proceedings, 2005, 18(4): 321-330.
- [13]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.
- [14]Gillies R J, Kinahan P E, Hricak H. Radiomics: Images Are More than Pictures, They Are Data [J]. Radiology, 2016, 278(2): 563-577.
- [15]Dai M, Wang N, Zhao X, et al. Value of Presurgical 18F-FDG PET/CT Radiomics for Predicting Mediastinal Lymph Node Metastasis in Patients with Lung Adenocarcinoma [J]. Cancer Biotherapy and Radiopharmaceuticals, 2022.
- [16]Liu S, Zhou Y, Wang C. et al. Prediction of lymph node status in patients with early-stage cervical cancer based on radiomic features of magnetic resonance imaging (MRI) images [J]. BMC Medical Imaging, 2023, 23(1): 101
- [17]Zhao M, Zheng Y, Chu J, et al. Ultrasound-based radiomics combined with immune status to predict sentinel lymph node metastasis in primary breast cancer [J]. Scientific Reports, 2023, 13(1): 16918.
- [18]Li M H, Liu L, Feng L, et al. Prediction of cervical lymph node metastasis in solitary papillary thyroid carcinoma based on ultrasound radiomics analysis [J]. Frontiers in Oncology, 2024, 14.
- [19]LeCun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [20]Wu X, Li M, Cui X-W, et al. Deep multimodal learning for lymph node metastasis prediction of primary thyroid cancer [J]. Physics in medicine & biology, 2022,67(3): 035008.
- [21]Zhang A-Q, Zhao H-P, Li F, et al. Computed tomography-based deep-learning prediction of lymph node metastasis risk in locally advanced gastric cancer [J]. Frontiers in Oncology, 2022, 12.
- [22]Wan L, Hu J, Chen S, et al. Prediction of lymph node metastasis in stage T1-2 rectal cancers with MRI-based deep learning [J]. European Radiology, 2023, 33(5): 3638-3646.
- [23]Zhou L-Q, Zeng S-E, Xu J-W, et al. Deep learning predicts cervical lymph node metastasis in clinically node-negative papillary thyroid carcinoma [J]. Insights into Imaging, 2023,14(1): 222.

- [24] Tian W, Yan Q, Huang X, et al. Predicting occult lymph node metastasis in solid-predominantly invasive lung adenocarcinoma across multiple centers using radiomics-deep learning fusion model [J]. *Cancer Imaging*, 2024,24(1): 8.
- [25] Langlotz C P, Allen B, Erickson B J, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop [J]. *Radiology*, 2019, 291(3): 781-791.
- [26] Ching T, Himmelstein D S, Beaulieu-Jones B K, et al. Opportunities and obstacles for deep learning in biology and medicine [J]. *Journal of The Royal Society Interface*, 2018, 15(141): 20170387.
- [27] Su L, Liu Y, Wang M, et al. Semi-HIC: A novel semi-supervised deep learning method for histopathological image classification [J]. *Computers in Biology and Medicine*, 2021, 137: 104788.
- [28] Liu K, Ning X, Liu S. Medical Image Classification Based on Semi-Supervised Generative Adversarial Network and Pseudo-Labeling [J]. *Sensors*, 2022, 22(24): 9967.
- [29] Atha Z, Chaki J. SSBTCNet: Semi-Supervised Brain Tumor Classification Network [J]. *IEEE Access*, 2023, 11(1): 141485-141499.
- [30] Su L, Wang Z, Shi Y, et al. Local augmentation based consistency learning for semi-supervised pathology image classification [J]. *Computer Methods and Programs in Biomedicine*, 2023, 232: 107446.
- [31] Liang X, Liu X, Yao L. Review-A Survey of Learning from Noisy Labels [J]. *ECS Sensors Plus*, 2022, 1(2): 021401.
- [32] Li J, Socher R, Hoi S C H. DivideMix: Learning with Noisy Labels as Semi-supervised Learning[C]. *ICLR*, 2020.
- [33] Yi K, Wu J. Probabilistic End-To-End Noise Correction for Learning With Noisy Labels [C]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019: 7010-7018.
- [34] Wei H, Feng L, Chen X, et al. Combating Noisy Labels by Agreement: A Joint Training Method with Co-Regularization [C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020: 13723-13732.

- [35]Zhu M, Zhang L, Wang L, et al. Robust co-teaching learning with consistency-based noisy label correction for medical image classification [J]. *International Journal of Computer Assisted Radiology and Surgery*, 2023, 18(4): 675-683.
- [36]Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks [J]. *Pattern Recognition*, 2018, 77: 354-377.
- [37]Waibel A, Hanazawa T, Hinton G, et al. Phoneme recognition using time-delay neural networks [I]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1989, 37(3): 328-339.
- [38]Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [C]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [39]LeCun Y, Bengio Y. *Convolutional networks for images, speech, and time series* [M]. MIT Press, 1998: 255-258.
- [40]Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks [J]. *Advances in neural information processing systems*, 2012, 25(2): 84-90.
- [41]Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *Computer Science*, 2014.
- [42]Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015: 1-9.
- [43]He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 770-778.
- [44]Huang G, Liu Z, Maaten L V D, et al. Densely Connected Convolutional Networks [C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 2261-2269.
- [45]Goodfellow I, Bengio Y, Courville A. *Deep learning* [M]. The MIT Press, 2016.
- [46]Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning [J]. *arXiv: 1603.07285*, 2016.
- [47]Hua J, Li L, Liu L, et al. The diagnostic value of metabolic, morphological and heterogeneous parameters of 18F-FDG PET/CT in mediastinal lymph node metastasis of

- non-small cell lung cancer [J]. *Nuclear Medicine Communications*, 2021, 42(11): 1247-1253.
- [48] Zhang C, Liao C, Penney B C, et al. Relationship between Overall Survival of Patients with Non-Small Cell Lung Cancer and Whole-Body Metabolic Tumor Burden Seen on Postsurgical Fluorodeoxyglucose PET Images [J]. *Radiology*, 2015, 275(3): 862–869.
- [49] Tarvainen A, Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results [J]. *Advances in neural information processing systems*, 2018, 30.
- [50] 易婉晴, 郑雪怡, 张状等. 基于多期相 CT 图像的深度学习 ResNet18 和 ResNet50 模型诊断肾结核的效能比较[J]. *中国防痨杂志*, 2024, 46 (3) :288-293.
- [51] Zheng G, Awadallah A H, Dumais S T. Meta Label Correction for Noisy Label Learning [C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11053-11061.
- [52] Fawcett T. An introduction to ROC analysis [J]. *Pattern Recognition Letters*, 2005, 27(8): 861-874.
- [53] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2000.
- [54] Kaplan E L, Meier P. Nonparametric Estimation from Incomplete Observations [J]. *Journal of the American Statistical Association*, 1958, 53(282): 457–481.
- [55] Sohn K, Berthelot D, Li C L, et al. FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence [C]. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020: 13.
- [56] Han B, Yao Q, Yu, X, et al. Co-teaching: Robust Training of Deep Neural Networks with Extremely Noisy Labels [C]. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 2018: 11.
- [57] Yao Y, Sun Z, Zhang C, et al. Jo-SRC: A Contrastive Approach for Combating Noisy Labels [C]. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021: 5188-5197.
- [58] Sun Z, Shen F, Huang D, et al. PNP: Robust Learning from Noisy Labels by Probabilistic Noise Prediction [C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022: 5301-5310.

- [59] Meng M, Gu B, Fulham M J, et al. DeepMSS: Deep Multi-Modality Segmentation-to-Survival Learning for Survival Outcome Prediction from PET/CT Images [J]. CoRR abs/2305.09946, 2023.
- [60] Zhang J, Huang Y, Zhang Z, et al. Whole-Body Lesion Segmentation in 18F-FDG PET/CT [J]. arXiv: 2209.07851, 2022.
- [61] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.

## 致谢

光阴似箭，时间如流，三年的硕士研究生求学即将结束。在这段时间里，我拓宽了自己的视野，增长了见识，接触了以往未涉足的领域，感受到了广阔而深邃的世界。

首先，我要感谢我的导师何江萍教授。在他严格的要求下，我更加努力投入科研工作，养成了良好的科研习惯。在他的鼎力支持下，我接触到许多优秀人才，开阔了眼界。何老师对科研的一丝不苟和学术追求的精神深深地影响着我，让我深刻认识到科研的严谨性和持之以恒的重要性。我还要感谢实验室的丁晓阳老师、李强老师、韩金仓老师、杨佳澄老师、李子荣老师和李兵老师，他们为实验室的建设做出了巨大贡献，对我的科研工作提供了重要支持。

其次，我要表达对学校的感谢，感谢学校为我们提供了优质的学习资源和舒适的学习环境。特别是信息工程与人工智能学院的教师们，他们不仅在课堂上传授知识，解答我们的疑惑，还在培养我们的能力方面起到了至关重要的作用。

最后，我要感谢一直陪伴在我身边的父母和朋友。感谢我的父母给予了我无私的经济支持和关怀，正是因为他们的理解和鼓励，我才能够全身心地专注于研究并成功地完成了我的学业。感谢我的室友们，我们一起经历风风雨雨，相互支持和鼓励，这段时间让我们都不断成长，变得更加出色。感谢实验室的同学们，我们相互学习、共同成长，也是他们对我的科研和工作提出的建议，不断激励着我成长。

