

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 PET/CT 影像的非小细胞肺癌组织学  
亚型分类研究

研究生姓名: 张杉杉

指导教师姓名、职称: 韩金仓 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： 张树彤 签字日期： 2022.5.29

导师签名： 韩志包 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： 张树彤 签字日期： 2022.5.29

导师签名： 韩志包 签字日期： 2022.5.29

# **Histological subtype classification of non-small cell lung cancer based on PET/CT image**

**Candidate : Shanshan Zhang**

**Supervisor: Jincang Han**

## 摘要

肺癌是一种恶性肿瘤，它起源于人体支气管黏膜上皮或肺泡上皮。当前，肺癌的发病率仍然居高不下，且是全球癌症死亡的主要原因之一。临床上通常根据癌细胞的类型不同将原发性肺癌分为小细胞肺癌和非小细胞肺癌，非小细胞肺癌占据了原发性肺癌的80%以上，是原发性肺癌的主要构成部分。组织学亚型是非小细胞肺癌患者治疗方案制定的重要因素，术前区分非小细胞肺癌患者的组织学亚型，对于非小细胞肺癌患者的个性化诊疗有着重要意义。

为了在术前更加准确、无侵入的区分非小细胞肺癌患者的组织学亚型，本文研究在包含 151 例组织学亚型为鳞状细胞癌或腺癌的非小细胞肺癌患者数据集上，建立了用于术前无创区分非小细胞肺癌组织学亚型的融合分类模型和全自动分类模型。在非小细胞肺癌组织学亚型融合分类模型中，本研究结合了传统影像组学获取的手工特征、深度学习方法构建的卷积神经网络获取的深度学习特征、非小细胞肺癌患者临床信息，使用机器学习分类器作为最终分类来区分患者组织学亚型。在这一模型中，本研究通过一系列对比实验证实融合模型在非小细胞肺癌组织学亚型分类任务较单独的影像组学和深度学习方法有着更好的性能（AUC 值 0.834，准确率=0.755、精确率=0.752、召回率=0.736 和 F1 值=0.740）。在非小细胞肺癌组织学亚型全自动分类模型中，本研究克服了非小细胞肺癌组织学亚型分类研究中对于医生手动标注的患者病灶区域的依赖性，构建了三种全自动分类模型，通过实验比较选择基于患者 PET/CT 影像最大强度投影的 2D 深度学习方法作为最佳全自动分类模型，并在这一全自动分类模型中得到了 0.768 的分类准确度。

本文研究表明，传统影像组学和深度学习都能区分非小细胞肺癌患者的组织学亚型，其中融合传统影像组学获取的手工特征、深度学习方法构建的卷积神经网络获取的深度学习特征、非小细胞肺癌患者临床信息，使用机器学习分类器作为最终分类手段构建的融合分类模型较其他两种方法有着更大的优势。其次，脱离对医生手动标注的患者病灶区域的依赖，直接从患者医学影像入手判断其组织学亚型的非小细胞肺癌组织学亚型全自动分类方法是可行且有效的。

**关键字：**非小细胞肺癌 分类 影像组学 深度学习 机器学习 组织学亚型

## Abstract

Lung cancer is a malignant tumor. Currently, the incidence of lung cancer remains high and is one of the leading causes of cancer deaths worldwide. In clinical practice, non-small cell lung cancer accounts for more than 80% of primary lung cancers and is the major component of primary lung cancer. Histological subtype is an important factor in the development of treatment plans for patients with non-small cell lung cancer, and differentiating the histological subtype of non-small cell lung cancer patients before surgery is important for personalized treatment of non-small cell lung cancer patients.

A combined model and fully automated classification model for automatic and noninvasive distinguishing the histological subtypes of non-small cell lung cancer patients was established by Radiomics and deep learning methods on a dataset containing 151 non-small cell lung cancer patients whose histological subtypes were pathologically confirmed as squamous cell carcinoma and adenocarcinoma, in order to distinguish the histological subtypes of non-small cell lung cancer patients more accurately and noninvasively before surgery. The combined model of non-small cell lung cancer histological subtype classification was constructed by combining handcrafted features obtained from traditional Radiomics, deep learning features obtained from convolutional neural networks constructed by deep learning methods, and clinical information of non-small cell lung cancer patients using a machine learning classifier as the final classification tools. Using a series of comparative experiments, this study

confirmed that this combined model performed better for the non-small cell lung cancer histological subtype classification task (AUC= 0.834, accuracy= 0.755, precision= 0.752, recall= 0.736, and F1 value= 0.740) and outperformed the results obtained using either traditional imaging histology or convolutional neural network models alone.

In the fully automated classification model for non-small cell lung cancer histological subtypes, this study overcame the dependence on physician manually labeled patient lesion areas in non-small cell lung cancer histological subtype classification studies, constructed three fully automated classification models, and selected the 2D deep learning method based on the maximum intensity projection of patient PET/CT images as the best fully automated classification model by experimental comparison, and in this fully automated classification model, the classification accuracy of 0.768 was obtained in this fully automated classification model.

This article shows that both conventional Radiomics method and deep learning techniques can distinguish histological subtypes of non-small cell lung cancer patients, in which the combined model built by fusing manual features obtained from conventional Radiomics method, deep learning features obtained from convolutional neural networks constructed by deep learning methods, and clinical information of non-small cell lung cancer patients using machine learning classifiers as the final classification tools has more advantages than the other two methods. Furthermore, it is feasible and effective for non-small cell

lung cancer histologic subtypes to be classified directly from the patient's medical images, without relying on the physician's manual labeling of the patient's lesion area.

**Keywords:** Non-small cell lung cancer; Classification; Radiomics; Deep learning; machine learning; Histological subtypes

# 目 录

|                            |    |
|----------------------------|----|
| 1 绪 论.....                 | 1  |
| 1.1 研究背景与意义.....           | 1  |
| 1.2 非小细胞肺癌组织学亚型分类研究现状..... | 2  |
| 1.3 本文的研究内容与创新.....        | 5  |
| 1.4 本文组织结构.....            | 6  |
| 2 相关理论介绍.....              | 7  |
| 2.1 影像组学及其特征介绍.....        | 7  |
| 2.2 深度学习介绍.....            | 8  |
| 2.3 特征选择算法与机器学习分类器介绍.....  | 9  |
| 2.3.1 特征选择算法.....          | 9  |
| 2.3.2 机器学习分类器.....         | 10 |
| 2.4 本章小结.....              | 11 |
| 3 PET/CT 数据介绍.....         | 12 |
| 3.1 PET/CT 医学影像原理简要介绍..... | 12 |
| 3.2 本实验数据来源及肿瘤分割的原则介绍..... | 13 |
| 3.3 数据统计.....              | 14 |
| 3.4 PET/CT 图像预处理.....      | 14 |
| 3.5 本章小结.....              | 15 |
| 4 非小细胞肺癌组织学亚型融合分类方法.....   | 16 |
| 4.1 融合分类方法总体结构.....        | 16 |
| 4.2 影像组学特征提取与选择.....       | 17 |
| 4.3 深度学习特征提取.....          | 18 |
| 4.4 特征融合与分类模型构建.....       | 20 |
| 4.5 融合分类模型及实验结果.....       | 21 |
| 4.5.1 传统影像组学模型实验结果.....    | 21 |
| 4.5.2 深度学习分类模型实验结果.....    | 22 |
| 4.5.3 融合分类模型实验结果.....      | 23 |

|                                   |           |
|-----------------------------------|-----------|
| 4.6 本章小结 .....                    | 26        |
| <b>5 非小细胞肺癌组织学亚型全自动分类方法 .....</b> | <b>27</b> |
| 5.1 全自动分类模型总体结构 .....             | 27        |
| 5.2 全自动分类模型数据预处理 .....            | 28        |
| 5.3 PET/CT 最大强度投影图特征提取 .....      | 28        |
| 5.4 PET/CT 特征融合及分类 .....          | 29        |
| 5.5 全自动分类模型实验结果 .....             | 30        |
| 5.5.1 机器学习模型实验结果 .....            | 30        |
| 5.5.2 3D 深度学习分类模型实验结果 .....       | 31        |
| 5.5.3 2D 深度学习分类模型实验结果 .....       | 31        |
| 5.6 本章小结 .....                    | 33        |
| <b>6 总结与展望 .....</b>              | <b>34</b> |
| <b>参考文献 .....</b>                 | <b>38</b> |
| <b>致 谢 .....</b>                  | <b>43</b> |

# 1 绪论

## 1.1 研究背景与意义

当前来看，肺癌仍是癌症死亡的主要原因，占全球癌症死亡人数的 18.0%<sup>[1]</sup>。肺癌是一种具有明显的病理特征、临床行为和高度异质性的恶性上皮肿瘤<sup>[2]</sup>。根据癌细胞在显微镜下显示的特征，临床上将原发性肺癌分类为非小细胞癌和小细胞癌。这两类肺癌中前者占据了原发性肺癌的 80%，是肺癌最主要的类型。通过组织学亚型临床上能够对非小细胞肺癌进行区分，在这其中，鳞状细胞癌（SCC）和腺癌（ADC）是其最常见的两种组织学亚型。通常来说，两者共占据非小细胞肺癌的 58%<sup>[3]</sup>。这两种亚型的肿瘤在人体中的生长部位有所区分，腺癌主要表现为腺上皮的肿瘤，鳞状细胞癌则表现为鳞状上皮的肿瘤。同时，它们的区别还在于这两者的形状和位置，腺癌质地比较硬，呈圆形或是卵圆形并具有不完整性，通常发生在肺部周围；鳞状细胞癌细胞则具有角质化结构，通常位于肺部的更中央<sup>[4]</sup>。由于鳞状细胞癌和腺癌的治疗方案和用药方面存在较大差异，例如在用药方面，培美曲塞是 IV 期肺腺癌的首选治疗方案，而贝伐单抗不适合用于鳞癌<sup>[13]</sup>；在治疗方案选择上腺癌对于化疗较为敏感，而鳞状细胞癌则对化疗敏感度较低，更适合手术。因此组织学亚型成为肺癌治疗决策的重要因子<sup>[5]</sup>。

在非小细胞肺癌组织学亚型的区分通常将病理诊断作为区分鳞状细胞癌和腺癌的金标准。在当前的临床实践中，最常用的非小细胞肺癌组织学亚型鉴别方法是通过活检对肿瘤组织进行组织病理学分析。然而，在肿瘤组织的采样中，肿瘤的异质性、某些病灶位置穿刺无法触及容易导致采样错误或采样失败。另一方面，活检诊断程序对患者构成较大伤害，有较高感染的可能性。这些缺点要求自动化、无创性的组织学亚型检测方法的出现。

正电子发射断层扫描和计算机断层扫描(PET/CT)仪利用 18F-氟脱氧葡萄糖(18F-FDG)标记物和人体代谢成像，使 PET/CT 成为肺癌临床诊断的重要成像方式之一<sup>[6,7]</sup>，但放射科医生单纯根据 PET/CT 图像区分腺癌和鳞状细胞癌较为困难。当前，影像组学作为一种被广泛认可的特征计算方法，利用一定的计算规则从医学影像中定量地提取特征来表征肿瘤信息，这一方法能够提包括肿瘤内部的特征表现，引起了临床肿瘤学家的广泛关注<sup>[8]</sup>。影像组学方法提取到的特征包括了半定量特征和定量特征。Aerts<sup>[9]</sup>等人在临床上全面应用影像组学方法，这一研究在一个包含了 1019 例非小细胞肺癌患者

(788 例) 和头颈癌患者的大型数据集证实了影像组学在表征肿瘤异质性上的潜力。422 例非小细胞肺癌患者的 CT 影像被用作训练集提取了多于 400 个包括强度、形状、纹理、小波的手工影像组学特征。研究使用提取到的手工影像组学特征通过无监督聚类方法揭示了影像组学特征与非小细胞肺癌患者 TNM 分期、组织学亚型的关联性。同时, 研究还使用生存分析方法证实了影像组学特征的预后价值。这对医学影像分析研究有着极大的借鉴意义。

近年来, 深度学习, 即使用深度神经网络来学习图像特征的方法在医学图像分析方面取得了巨大的成功, 激发了基于深度学习的影像组学的发展。深度学习技术通常采用一个深度神经网络以端到端的方式从医学图像读取信息, 一方面, 这能够消除传统影像组学对耗时的手工特征提取的依赖; 另一方面, 基于深度学习的影像组学通过网络自动学习相关和稳健的特征, 而不需要人工干预<sup>[10]</sup>。换言之, 传统影像组学特征虽然可以达到上百上千个, 但这些特征都是通过人为制定的规则计算的, 受限于当下的医学成像知识。基于深度学习的影像组学克服了传统影像组学的这些缺点, 一定程度能够消除手工特征带来的人为偏见, 并有可能发现可能会被人工定义所忽略的特征。

组织学亚型在非小细胞肺癌患者治疗中的重要性和非小细胞肺癌组织学亚型病理检测的高风险、低效益等弊端使得自动、无创、成本效益高的辨别方法被需要。传统影像组学和深度学习技术在医学影像分析中的成功应用使自动、无创、成本效益高的方法成为可能。这种自动、无创的区分非小细胞肺癌组织学亚型方法能够大大降低病理诊断的成本和诊断程序的风险性。

## 1.2 非小细胞肺癌组织学亚型分类研究现状

传统影像组学在反映生物异质性上有着很大潜力, 可以通过非小细胞肺癌患者不同模态的医学影像获取病灶特征, 为非小细胞肺癌组织学亚型的诊断提供额外的有用信息<sup>[11]</sup>。此外, 影像组学方法在医学影像分析工作是相对经济有效的, 并且这一方法已经应用于肿瘤诊断、分期和治疗指导的常规临床实践。它基于特定规则提取大量的定量影像特征, 能够提供详细、全面的肿瘤表型特征在非小细胞肺癌组织学亚型自动分类研究问题上被广泛使用<sup>[12]</sup>。

Wu 等人<sup>[13]</sup>使用传统影像组学方法来区分肿瘤组织学亚型 (腺癌和鳞状细胞癌)。研究使用了来自两个不同医学中心的共 350 名非小细胞肺癌患者 CT 影像。其中, 包含了 198 位非小细胞肺癌患者的数据集 1 作为训练集, 剩余来自另一医学中心的包含 152

位非小细胞肺癌患者的数据集 2 作为测试数据集。这一研究所有影像数据病灶区域由医生进行手动勾画，在经过预处理的 CT 图像分割出的肿瘤体积后，研究使用影像组学方法在患者 CT 影像病灶区域部分提取了 440 个手工影像组学特征。这些手工影像组学特征包括了肿瘤的形状特征、强度特征以及纹理特征不同角度的量化。使用 ROC 曲线下面积 AUC 值进行各个特征与非小细胞肺癌患者组织学亚型相关性的单变量分析，分析揭示了在所有手工影像组学特征中有 53 个手工特征与非小细胞肺癌组织学亚型的显著相关。多变量分析中，这一研究使用了两阶段特征选择法，在第一阶段计算特征相关性，将特征数从 440 维减少到 67 维。特征选择的第二阶段则使用 24 种常用且计算效率高的特征选择方法增量选择了 5-45 个特征。将选出的特征在三个分类器中训练，并在测试集中测试各分类器的分类性能。最终在朴素贝叶斯分类器上得到了 0.72 的最优 AUC 值。这一工作证实了影像组学在非小细胞肺癌组织学亚型分类任务上的潜力，突出了无创和成本效益高的影像组学方法在肺癌治疗中的巨大潜力。Zhu X 等<sup>[14]</sup>回顾性研究了 129 例非小细胞肺癌患者，并从手动标记的肿瘤区域中提取了 485 个手工影像组学特征。通过使用逻辑回归方法，从 485 个手工影像组学中选择了五种特征来构建模型，选择出的 5 种手工特征在测试集中达到 0.893 的 AUC 值。Kirienko, Cozzi 等人<sup>[15]</sup>在 534 例肺癌患者的 PET/CT 数据集中区分了原发性肺癌和转移病灶，同时对原发肿瘤的组织学亚型（腺癌和鳞状细胞癌）进行鉴别。这一研究分别从 PET 影像和 CT 影像中使用影像组学方法提取了纹理特征，使用线性判别分析方法对患者病灶区域特征进行分析，通过 100 次重复实验证明了从 PET 影像和 CT 影像中提取的手工影像组学特征在区分原发肿瘤和转移病灶以及原发型肺癌肿瘤的组织学亚型的鉴别中显示出巨大的潜力。Ferreira Junior 等人<sup>[16]</sup>在组织学亚型的识别上也获得了类似的结果，证实了非侵入性和成本效益高的影像组学在肺癌组织学亚型分类中的巨大潜力。Ren, C 等人<sup>[17]</sup>基于 PET/CT 影像和机器学习方法来区分非小细胞肺癌（NSCLC）中的腺癌和鳞状细胞癌。这一研究在 315 名非小细胞肺癌患者数据中，开发了四个独立的预测模型，使用临床因素-肿瘤标志物，PET 影像提取的手工特征，CT 影像提取的手工特征及其组合来区分腺癌和鳞状细胞癌。使用最小绝对收缩和选择算子（LASSO）方法进行选择相关特征，构建了一个包含临床因素，肿瘤标志物和从 PET 影像和 CT 影像中提取的手工影像组学特征的组合模型并在非小细胞肺癌患者数据中得到出色的表现。Yang Fengchang<sup>[18]</sup>等收集了来自不同中心的三个非小细胞肺癌患者 CT 影像的数据集。研究过程中，使每个单独的数据集都有机会作为训练集，剩余两个数据集则充当测试集。实验中，研究使用手动标注的病灶区域和相应的

CT 影像中提取了总共 788 个手工影像组学特征。使用最小冗余最大相关特征选择 (mRMR)，顺序正向选择 (SFS)，最小绝对收缩和选择算子 (Lasso) 三种较为常用的特征来选择与当前学习任务相关的特征。根据所选特征对逻辑回归 (LR)，支持向量机 (SVM) 和随机森林 (RF) 这三种广泛使用的分类模型进行了独立评估，并在合并的数据集上得到 0.78 的 AUC 值，显示出相对较好的预测性能。

虽然传统机器学习与影像组学特征相结合的方法取得了一系列不错的结果，但是影像组学的特征多为明确设计或手工制作，虽然这些特征的数量可以达到数万个，但这仍受限于当前的医学成像知识，可能不能完全表征图像的异质性，从而限制影像组学模型的潜力。

随着神经网络这一强大的计算机视觉方法在医学上得到应用，从深度网络中提取图像深层特征成为可能。基于深度学习的影像组学能够通过神经网络自动地从影像数据中提取特征这一特性，以及它在医学影像分析中显示出的巨大应用潜力。越来越多学者将其应用于医学图像分析领域。这些由神经网络提取到的影像的深层特征相比于手工影像组学特征包含了更多的医学图像的抽象信息，这为进一步提高肿瘤组织学亚型鉴别的准确性带来了可能。Han 等人<sup>[19]</sup>在 1419 位非小细胞肺癌患者的 PET/CT 影像。使用共 688 个包括一阶统计特征和基于矩阵描述的纹理特征对十种特征选择方法，十种机器学习分类器和 VGG16 深度学习算法在非小细胞肺癌组织学亚型鉴别上的性能，构建区分鳞状细胞癌和腺癌的最佳分类模型。结果显示 VGG16 模型优于所有影像组学与机器学习相结合的方法。Chaunzwa 等人<sup>[20]</sup>在 157 例非小细胞肺癌患者的 CT 影像中对比了传统影像组学方法和深度学习方法，神经网络以 0.751 的 AUC 值成为区分腺癌和鳞状细胞癌的最佳分类模型。这一结论与 Han 等人的研究得出相似的结论。Atsushi Teramoto 等人<sup>[21]</sup>基于深度卷积神经网络方法开发了一种用于肺癌显微图像的自动分类方案。分类网络包含了三个卷积、三个池化和两个完全连接。构建好的网络被用来区分三种类型癌症 (腺癌，鳞状细胞癌和小细胞癌)，并使用三重交叉验证评估得到了大约 71% 的准确性，这一结果与放射科医师的准确性基本一致。

在以上研究中，我们注意到一方面使用深度学习方法进行非小细胞肺癌组织学亚型分类的工作多缺少与影像组学特征的结合，且已有研究表明，当与机器学习分类器(如支持向量机(SVM)和随机森林分类器(RF))一起使用时，深度学习派生的特征图可能在分类任务中优于原始深度学习模型<sup>[22-24]</sup>；另一方面，现有研究仍然依赖于医生对于病灶区域的标注，这一工作十分的耗时耗力。根据这一非小细胞肺癌组织学亚型分类任务的研

究现状，文章构建了用于在非小细胞肺癌患者影像中区分腺癌和鳞状细胞癌的融合分类模型与全自动分类模型。

### 1.3 本文的研究内容与创新

基于当前对于非小细胞肺癌组织学亚型分析研究的现状，本研究在非小细胞肺癌组织学亚型分类任务中提出融合传统影像组学和基于深度学习的影像组学的非小细胞肺癌组织学亚型融合分类方法（依赖于医生对病灶区域的标注）以及脱离医生标注病灶区域的组织学亚型全自动分类方法两个模块。融合方法将深度学习模型得到的深层抽象信息、传统影像组学得到的有效手工特征以及患者临床信息相结合放入机器学习分类器中，弥补了手工特征中缺少的深层抽象信息，进一步提升非小细胞肺癌组织学亚型分类的准确性。非小细胞肺癌组织学亚型全自动分类则脱离医生对于病灶区域的标记，通过构建 2D 卷积神经网络直接从非小细胞肺癌患者原始 PET/CT 影像中判断非小细胞肺癌患者的组织学亚型。

两个模块的工作为自动、无创、成本效益高的非小细胞肺癌组织学亚型分类提供方法上的指导，融合传统影像组学和基于深度学习的影像组学对非小细胞肺癌组织学亚型分类为现有组织学亚型分类工作提供了新的思路。非小细胞肺癌组织学亚型全自动分类模型则进一步减轻医生在勾画病灶区域中所消耗的时间成本，使得组织学亚型的鉴别更加高效。

本文的创新点在于：

（1）在非小细胞肺癌组织学亚型融合分类模型中，不同于一般将传统影像组学手工特征融合到深度学习中的融合方式，本文研究将深度学习特征和传统影像组学手工特征使用堆叠方式连接放入机器学习分类器中进行分类任务，进一步提升了非小细胞肺癌组织学亚型分类任务的性能。

（2）非小细胞肺癌融合分类模型为分类任务提供了一个灵活的框架。本研究大量的实验表明了包含传统影像组学手工特征、深度学习特征以及临床特征的融合特征在机器学习分类器中有着最优表现，这一框架的使用或许能使得部分分类任务的性能得到进一步的提升。

（3）非小细胞肺癌组织学亚型全自动分类模型中，本文使用了患者最大值投影的全身 PET/CT 影像作为网络的输入。鉴于非小细胞肺癌中鳞状细胞癌和腺癌在人体中的位置通常有一定区别这一医学常识，患者最大值投影的全身图像能够提供更加全局的信

息，有助于网络识别病灶区域的位置信息，这能够为其他研究提供借鉴。

(4) 非小细胞肺癌组织学亚型全自动分类模型能够绕过医生对患者病灶区域标记这一费时耗力的工作来进行组织学亚型的分类，证明了不依赖手动、半自动或自动标注的病灶位置的全自动分类模型能够区分非小细胞肺癌患者的组织学亚型，相较于先前非小细胞肺癌组织学亚型分类研究工作是一大突破。

## 1.4 本文组织结构

本文的主要章节布局如下：

第一章总的介绍了肺癌这一恶性疾病，简述了肺癌的类型以及非小细胞肺癌组织学亚型分型。介绍了传统影像组学和深度学习技术在医学影像分析中的现状及特点，同时展示了两种方法在非小细胞肺癌组织学亚型分类任务中的一些应用。

第二章主要介绍了传统影像组学特征提取过程及其特征大致类型和其基本定义。深度学习方面，介绍了深度学习技术原理和其发展现状。同时，对于本文所用到的特征选择算法和机器学习分类器也做了简单的描述。

第三章详细介绍了 PET/CT 影像的原理。同时，这一章节详细介绍了本文所用数据的来源和数据的筛选原则。同时根据数据中的基本信息对本文研究所选用的数据样本做了简要的数据分类统计，并展示了其基本构成。另外，本章还简要介绍了本研究构建网络前对于影像数据所做的预处理操作。

第四章主要介绍基于 PET/CT 影像的非小细胞肺癌组织学亚型融合分类模型的各部分构成、特征提取与筛选、分类模型性能及与相关方法的比较。这一工作将传统影像组学选取的手工影像组学特征、卷积神经网络获取的深度学习特征以及患者临床信息融合到机器学习分类器中得到最终分类。

第五章介绍了克服医生手动标注的病灶区域的非小细胞肺癌组织学亚型全自动分类模型。使用融合分类模型中效果较好的卷积神经网络构建了 2D 深度学习网络，使用非小细胞肺癌组织学患者最大值投影全身影像作为模型的输入。成功区分了非小细胞肺癌组织学亚型的。

第六章是对本文所有工作的总结，详细讨论了文章所述的两项工作的贡献与其不足，并提供了可能的解决方法。

本文的最后对关心和帮助本论文创作的老师、同学、朋友及家人表示最衷心的感谢。

## 2 相关理论介绍

### 2.1 影像组学及其特征介绍

影像组学是一个从医学图像中提取和分析定量数据的过程，它根据固定的计算规则将影像感兴趣区域部分信息进行量化，得到高维的定量特征。其应用于医学图像分析的工作流程包括四个部分，分别为：图像采集、图像预处理、感兴趣区域分割、影像组学特征提取<sup>[25]</sup>。

图像采集是指原始医学图像的收集，影像组学依赖于患者原始医学影像，医学影像有多种模态之分，例如：超声影像、电子计算机断层扫描（CT）、正电子发射型计算机断层显像（PET）、核磁共振影像（MRI）和 PET/CT 等。其中 CT 影像能够提供患者解剖结构信息，可以看到较为明显的人体骨骼架构。PET 影像虽然缺乏解剖特征，但它可以提供患者体内的代谢和功能信息。PET/CT 影像则将 PET 影像与 CT 影像融为一体，同时提供了上述两种信息。在本研究中，来自于美国芝加哥大学医学院的 151 位非小细胞肺癌 PET/CT 影像被使用做非小细胞肺癌组织学亚型分类研究。

通常来说，在图像采集过程中，所使用的参数和设备的可变性容易导致医学影像分析工作中的误差。因此，图像预处理是影像组学特征提取中极为重要的环节<sup>[26]</sup>。感兴趣区域分割同样是影像组学工作流程的重要步骤，它通常包括手动的、半自动的或全自动的，这一工作精确度极大一部分决定了要提取的手工影像组学特征的质量。本研究在非小细胞肺癌组织学亚型融合分类模型中采用手动的感兴趣区域分割方法，影像中的病灶区域定位由有经验的医师进行勾画。

特征提取是感兴趣区域分割后的下一步，同时也是影像组学的核心，它对有用信息进行量化，从而帮助表征正常和异常的放射图像。影像组学特征由根据既定的计算方式自动计算图像特征。这些手工定量特征一般可分为一阶强度特征、形状特征、纹理特征。其中，一阶强度特征顾名思义描述了图像感兴趣区域内的各个单独体素强度分布，且这一特征的描述是基于单像素或单体素的。形状特征描述了金标准肿瘤的几何结构，它包括了基于 2D 和 3D 影像的形状特征。各个特征均由多个统计量进行描述，例如一阶强度特征使用偏度、峰度、均匀度等来描述。

纹理特征能够更好地描述相邻像素间的关系，且这一特征有能最大的信息量。在影像组学方法中，纹理特征对于像素间相邻关系的描述及特征提取通常基于不同的描述矩

阵，这些矩阵有灰度共生矩阵（GLCM）、灰度级大小区域矩阵（GLSZM）、灰度游程矩阵（GLRLM）、相邻灰度差矩阵（NGTDM）、灰度依赖矩阵（GLDM）。

在医学影像中，灰度共生矩阵（GLCM）基于影像各像素之间的空间分布蕴含着影像纹理信息这一假定，描述了患者病灶区域的二阶联合概率函数。灰度级大小区域矩阵（GLSZM）量化了患者病灶区域具有相同灰度级的相邻互连像素或体素数量计数的大小。这一矩阵越平坦则描述了更均匀的纹理。灰度游程长度矩阵（GLRLM）量化了患者病灶区域具有相同灰度值的连续像素的像素数长度。相邻灰度差矩阵（NGTDM）量化了一个灰度值与其在距离内与其邻居的平均灰度值之间的差异总和。灰度依赖矩阵（GLDM）则 量化影像中心像素或体素与其邻域之间的依赖关系<sup>[27]</sup>。

## 2.2 深度学习介绍

深度学习是机器学习的一个子集，它以在人脑中发现的信息处理模式为灵感建立起来。当前深度学习已逐渐成为机器学习领域中应用最广泛的计算方法。当前，深度学习已经在多项复杂的认知任务上取得了出色的成绩，有些甚至超过了人类表现<sup>[28]</sup>。深度学习是源自于传统的神经网络，但其性能要优于传统神经网络，它同时采用转换和图形技术来建立多层学习模型，在整个深度学习算法中以自动方式实现特征的提取<sup>[29]</sup>。深度学习具有多层数据表示架构，每层算法都对输入的数据提供不同的解释，即数据在不同层级的特征表示<sup>[30]</sup>。

卷积神经网络是深度学习领域中最著名也最常用的算法，相较于之前的深度学习方法，它有三个主要优点，这三个优点是等效表示、稀疏交互和参数共享<sup>[31]</sup>。与传统的全连接（Fully Connected）网络不同，卷积神经网络中的共享权重和局部连接用于充分利用图像信号等二维输入的数据结构。这一操作使用极少的参数，既简化了网络训练过程又加快了网络运行速度。用于图像分类的卷积神经网络架构通常包括卷积层、池化层、激活函数和全连接层。

卷积层（Convolutional Layer）是卷积神经网络十分重要的部分，它由一组卷积滤波器（内核）组成。滤波器可以由离散数字或值的网格来描述，网格中每个值称为核权重。在卷积神经网络训练过程开始时，随机数被分配用作内核权重的初始化。在训练中需要不断调整这些权重，从而使网络能提取重要特征。卷积层的具体运算使用一个  $8 \times 8$  的灰度图像作为输入来举例，以  $3 \times 3$  随机权重初始化内核。内核在整个  $8 \times 8$  的灰度图像上水平和垂直滑动，每滑动一次计算确定输入图像和内核之间的点积，即对应值相乘

然后相加。重复整个过程，直到不能再滑动为止。内核在图像中滑动计算出的点积值表示输出的特征图。其中内核每次滑动的距离称之为步长。

池化层（Pooling Layer）通过卷积操作生成的特征图的二次采样。池化层缩小了大尺寸的特征图以创建更小的特征图。池化阶段的每个步骤中都保留了大部分主要信息（或特征）。池化层的计算与卷积操作类似，在执行池化操作之前，步长和内核都初始分配大小。池化包括平均池化、最小池化、最大池化、全局平均池（GAP）和全局最大池化等。最熟悉和最常用的池化方法是最大池化、最小池化和全局平均池化。以最大池化为例， $2 \times 2$ 的内核在特征图上移动，每次取重叠部分四个值中的最大值，重复整个过程，直到不能再滑动为止。内核在图像中滑动计算出的点积值表示输出的特征图。

将输入映射到输出是激活函数（Activation Function）的核心功能。输入值是通过计算神经元输入的加权和及其偏差来确定的。在卷积神经网络架构中的所有具有权重的层，例如 FC 层和卷积层之后使用非线性激活函数。激活层的这种非线性性能意味着输入到输出的映射将是非线性的，这些层使卷积神经网络能够学习更复杂的事物。同时，激活函数一个非常重要的特征是具有区分能力。常用于卷积神经网络的激活函数有 Sigmoid 激活函数、Tanh 激活函数、ReLU 激活函数等。

全连接层（Fully Connected Layer, FC）通常位于每个卷积神经网络架构的末尾。在全连接层内部，每个神经元都连接到前一层的所有神经元，即所谓的全连接方法。它常充当卷积神经网络中的分类器。全连接层的输入来自于最后一个池化层或卷积层，它遵循传统多层感知器神经网络的基本方法，其输出代表最终的卷积神经网络的输出。当下最常用的卷积神经网络有：AlexNet、VGG、GoogLeNet、ResNet、DenseNet、MobileNet 等。

## 2.3 特征选择算法与机器学习分类器介绍

### 2.3.1 特征选择算法

特征集是用来描述当前学习任务属性的集合，这一集合通常包含了大量信息，通常我们称对当前学习任务有用的属性为相关特征，对于当前学习任务无用的属性称为无关特征。特征选择就是从给定的特征集合中筛选出相关特征的过程<sup>[42]</sup>。对于影像组学方法来说特征选择非常重要，因为使用影像组学方法提取出的特征是成百上千的，这其中包含了相当一部分的无关特征。这些无关特征一方面容易在学习任务造成“维数灾难”问

题,另一方面会增加学习任务的难度,降低计算效率。常用的特征选择算法可以分为过滤式、包裹式以及嵌入式。过滤法特征选择代表方法有卡方验证、互信息法和方差阈值法。包裹式方法则以向前搜索、向后搜索和递归特征消除法为代表的。L1 正则方法和基于树的随机森林特征选择法则属于嵌入式。

过滤法中的方差阈值特征选择法 (Variance) 是一个十分简单且较为基本的特征选择方法。这一方法根据各个特征计算出方差,根据计算出的方差值大小,基于方差越大的特征对于当前学习任务有效性越高这一理论,规定一个特定阈值删除小于这一阈值的特征<sup>[37]</sup>。互信息法 (Mutual Information, MI) 则基于各个特征的互信息值选择特征。其中,互信息值定义为一个事件的出现对于另一个事件的出现所贡献的信息量。在使用互信息法选择特征是,特征的互信息值越大,则该特征能够对当前学习任务提供更有用的信息。因此,使用互信息法进行特征选择时通常是选取互信息值最大的若干个特征<sup>[38]</sup>。

递归特征消除法<sup>[39]</sup> (Recursive feature elimination, RFE) 作为包裹式的特征选择方法,是一种贪婪的优化算法。这一特征选择方法基于一个基模型进行反复训练,在每轮训练后消除特征重要性较低的特征,保留特征重要性较高的特征。在下一轮训练中,递归特征消除法使用上一次训练中未被选中的特征来进行下一次的模型构建和特征筛选。这一过程反复进行,直到所有特征都耗尽为止。最后递归特征消除法根据保留特征的顺序对其进行排名,选出最佳的特征子集。

最小绝对收缩和选择算子方法<sup>[40]</sup> (Least absolute shrinkage and selection operator, LASSO) 和随机森林方法 (Random Forest, RF) 属于嵌入式。LASSO 方法是一种采用了 L1 正则化的线性回归方法,这一方法通过压缩特征的系数使其趋近于 0 来达到特征选择的目的。随机森林特征选择法是通过度量特征重要性,来选择出重要性较高的特征。这一方法中的特征重要性使用袋外数据误差评估<sup>[41]</sup>。

### 2.3.2 机器学习分类器

在分类任务中,构造的分类模型称之为分类器。分类器能够基于已有标签的训练数据学习分类规则从而将未知标签的数据赋予标签。机器学习中常用的分类器有: K 近邻分类、支持向量机、决策树算法等。

K 最近邻 (K-nearest neighbor classification, kNN) 分类器顾名思义,这一分类器是未知标签的新数据与已有标签的训练集数据的每一对应特征进行比较,将与新数据最近的训练数据标签赋予新数据的一种分类算法<sup>[42]</sup>。决策树算法 (Decision Tree, DT) 是

从树的根节点出发，对数据的每一个特征进行判断，根据判断结果，将数据分配到其子节点中<sup>[42]</sup>。随机森林（Random Forest Classifiers, RF）分类器是指利用多棵决策树对样本进行训练并预测的一种算法，它的最终输出的类别是由个别决策树输出的类别进行投票来决定的<sup>[42]</sup>。自适应提升算法（Adaptive Boosting）是一种有监督的集成学习。这一算法是根据训练数据集构建多个弱分类器，这多个弱分类器根据一定规则器构成最终的分类型模型<sup>[42]</sup>。支持向量机（Support Vector Mechine, SVM）分类器是寻找空间中能将数据分离的具有最大间隔的超平面的一种分类算法。它可以分为线性可分模型、线性支持向量机、以及非线性支持向量机<sup>[42]</sup>。梯度提升算法（Gradient Boosting, GDBT）通过对弱预测模型（比如决策树）的集成产生预测模型<sup>[43]</sup>。它以分步的方式构建模型把所有树的结论累加起来做最终结论，在这一分类器中，每一棵树所学的是之前所有树得到的结论和的残差。逻辑回归（Logistic regression, LR）分类器的本质是极大似然估计，过程简单来说就是对于一个回归或者分类问题，建立代价函数，然后通过优化方法迭代求解出最优的模型参数，然后测试验证求解出的模型的好坏<sup>[42]</sup>。

## 2.4 本章小结

本章节对影像组学方法应用于医学影像分析的流程和影像组学特征的构成进行简要介绍，在影像组学之后简单介绍了深度学习及其发展现状。同时，对文中的影像组学特征挑选用到的特征选择算法和融合模型构建所用的机器学习分类器进行简单介绍。

## 3 PET/CT 数据介绍

### 3.1 PET/CT 医学影像原理简要介绍

PET 影像的全称为正电子发射型计算机断层显像技术 (Positron Emission Computed Tomography, PET), 是一种能够显示患者体内代谢过程的核成像技术, 它依赖于放射性核素产生的信号。PET 影像所使用的放射性核素主要有碳、氟、氧和氮的同位素 ( $^{11}\text{C}$ 、 $^{13}\text{N}$ 、 $^{15}\text{O}$ 、 $^{18}\text{F}$ ) 等, 这些放射核素也称为示踪剂。在 PET 影像的拍摄中, 通过注射将上述放射性核素注入被拍摄者体内, 这些放射性核素进入人体后会参与人体正常代谢, 在被拍摄者体内产生衰变。在放射性核素衰变过程中会发生质子、中子的相互转化, 转化过程中释放的正电子与组织中的负电子相互作用形成湮没辐射。PET 检测器接收到湮没辐射产生的两个等能量反方向的光子, 并从中推定出正电子发射点位于两探测头的连线上。再经由 360 度排列的多组配对探头, 获得探头对连线上的一维信息。最后, 经过信息中心点反射和适当的数学处理形成 PET 影像。在 PET 影像中, 代谢率高的组织或者病灶区域会呈现高代谢的亮信号, 代谢率低的部分则呈现暗信号。PET 影像这一特点使临床中和研究上能够较为容易地获取到患者病灶区域代谢情况, 对于临床治疗和医学影像研究有极大的辅助作用。但 PET 影像并不能展示被拍摄者的解剖结构信息, 对于人体各个器官只有十分模糊的显示, 并不足以判断病灶区域相对于人体位置。

CT(Computed Tomography)影像, 全称为电子计算机断层扫描, 它的本质是衰减系数成像。在 CT 影像的拍摄原理是利用 X 射线束对于被拍摄者人体进行扫描, 在这束 X 射线穿透人体过程中, 射线束的一部分能量被人体吸收, 射出的 X 射线产生了衰减。电子计算机断层扫描仪将射出的 X 射线准直后由探测器接收, 模数转换器将探测器接收到的大量信息转换为数字并由计算机处理计算出 X 射线在断面上单位体积的吸收值。这些计算好的吸收值 (CT 值) 采用矩阵形式表征, 通过数模转换器转换为不同灰度等级并显示为被拍摄者的 CT 影像。由于人体不同密度组织对于 X 射线存在不同的吸收量, 通常来说, 人体组织密度越高的地方对于 X 射线吸收量越大。相反地, 在人体组织密度较低的位置, 对于 X 射线的吸收也越少。因此, CT 影像在对于人体骨骼结构表征上有着较大优势。

正如前文对于 PET 影像和 CT 影像的介绍, 这两种影像各有长短。同时 PET 影像在临床诊断中需要患者解剖结构信息, 电子计算机断层扫描是最为常见能够较为清晰显

示人体解剖结构的成像方式。尽管当前影像配准方法有了长足发展，但是对于并非在同一时间拍摄的影像的配准仍然不够完善。PET/CT 影像是将正电子发射断层扫描仪(PET)和 X 射线计算机断层扫描 (CT) 相结合到一台拍摄仪器的核医学技术。这一拍摄技术能够在近乎相同的时间拍摄出患者的 PET 影像和 CT 影像，因此，这一技术可以同时提供患者两方面信息，解决了 PET 影像对于解剖结构信息的需求，在临床医学中有极大的应用价值。到目前为止，PET/CT 影像是惟一可以在活体上显示生物分子代谢、受体及神经介质活动的新型影像技术。在临床上 PET/CT 核医学成像技术已经被用于多种疾病的诊断与个性化诊疗方案确定，例如肿瘤良恶性的判断和分期等<sup>[35]</sup>。同时，PET/CT 核医学成像技术更有着灵敏度高、特异性高、全身显像、安全性好等特点，这对于接受治疗的患者来说也极为重要。然而，PET/CT 核医学成像技术由于其昂贵的价格在临床使用上仍然十分困难。另一方面，用于 PET/CT 影像拍摄的放射性药物的生产和运输成本也是 PET/CT 核医学成像技术的一大阻碍<sup>[36]</sup>。因此，基于 PET/CT 医学影像的分析仍然有十分重要的价值。图 3.1 是本文研究所用数据的一个 PET/CT 影像的实例图：

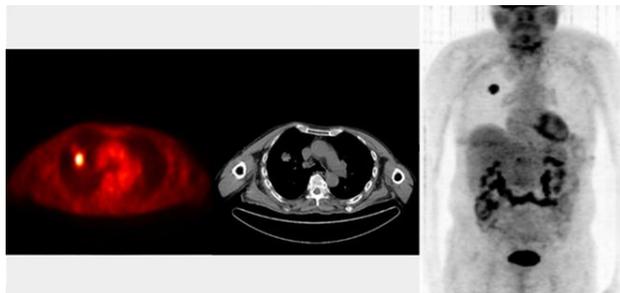


图 3.1 PET-CT 实例图 (从左到右依次为 PET 图像、CT 图像和 PET-CT 全身图像)

### 3.2 本实验数据来源及肿瘤分割的原则介绍

本次非小细胞肺癌组织亚型分类研究的实验数据来自于美国芝加哥大学医学院。研究对于 2004 年至 2014 年间芝加哥大学医学院接受治疗的、所有经病理证实为非小细胞肺癌 (NSCLC) 的患者病历进行回顾性审查，将所有符合以下标准的患者纳入研究：(1) 患者具有完整的 PET/CT 影像数据；(2) 患者没有明显双原发恶性肿瘤；(3) 患者在进行 PET/CT 拍摄前无脑转移迹象；(4) 有明确的病理诊断为腺癌或鳞状细胞癌；(5) 患者影像显示其未发生淋巴转移或远端转移。经过筛选的数据共包括 151 例病理检测为腺癌或鳞状细胞癌的非小细胞肺癌患者。

本次研究使用这 151 位非小细胞肺癌患者 PET/CT 影像数据及病理检测的组织学亚型标准进行非小细胞肺癌患者组织学亚型的分类研究。

### 3.3 数据统计

研究使用的 151 例非小细胞肺癌患者平均年龄为 68.7，男女性比例接近于 2:3。其中 88 例患者组织学亚型诊断为腺癌，63 例患者组织学亚型为鳞状细胞癌。表 3.1 展示了所选取病人的详细特征，这些特征包括患者年龄平均值、性别分布、组织学亚型分布、吸烟史以及影像最大 SUV 值。

表 3.1 非小细胞肺癌患者数据集的特征

| 特征                | 统计          |
|-------------------|-------------|
| 汇总                | 151 (100%)  |
| 年龄：平均值(方差)        | 68.7 (11.3) |
| 性别                |             |
| 女性：N (%)          | 87 (57.6)   |
| 男性：N (%)          | 64 (42.4)   |
| 组织学亚型             |             |
| 腺癌 (ADC)：N (%)    | 88 (58.3)   |
| 鳞状细胞癌 (SCC)：N (%) | 63 (41.7)   |
| 吸烟史               |             |
| 从不：N (%)          | 13 (8.6)    |
| 吸烟：N (%)          | 58 (38.4)   |
| 曾经：N (%)          | 80 (53.0)   |
| 最大 SUV 值：平均值(方差)  | 9.23 (7.6)  |

### 3.4 PET/CT 图像预处理

研究对于 151 例非小细胞肺癌患者 PET/CT 影像预处理流程包括三部分，分别是人工肿瘤区域分割、PET 代谢图像基于体重归一化为标准摄取值 (Standardized Uptake Value, SUV)、感兴趣区域 (Region of Interest, ROI) 获取。(1) 人工肿瘤区域分割：患者 PET/CT 影像中的肿瘤区域分割由两名具备资格且经验在十年以上的放射科医师进行。对于明显有争议的具体病灶，经放射科医师内部讨论解决，以实现最终的分割结果。

(2) PET 代谢影像基于体重归一化：由于 PET 影像灰度值受到显像剂剂量的影响较大，需要通过给药剂剂量校正，对像素值进行校正降低其影响，即基于患者体重归一化为

PET 代谢影像的标准摄取值。(3) 感兴趣区域 (Region of Interest, ROI) 获取: 医生对于病灶区域的标注产生的二值图不包含患者病灶信息, 不能作为模型的输入, 因此需要从患者原始影像中获取感兴趣区域, 具体为以病灶标签为模板, 从患者原始影像中扣取对应位置的影像信息, 达到为病灶区域赋予原始信息表征的目的。经过预处理后的影像用于非小细胞肺癌组织学亚型分类任务模型的构建与分析。

### 3.5 本章小结

本章节对本文中用到的数据进行介绍, 从所使用的 PET/CT 医学影像原理介绍、实验数据来源和肿瘤分割原则、用于本实验数据筛选原则对整体数据及进行介绍。同时对于筛选出的应用于本研究的 151 例非小细胞肺癌患者数据集做了数据统计, 展示了其基本数据构成。最后介绍了影像数据应用于研究前的预处理方法, 为后文模型的构建打下铺垫。

## 4 非小细胞肺癌组织学亚型融合分类方法

### 4.1 融合分类方法总体结构

当前对于非小细胞肺癌组织学亚型的分类任务中，传统影像组学和深度学习技术都有着不少应用。然而，对于这两种方法相结合的融合分类研究还比较少，且在当前的融合方法中，研究者多将传统影像组学方法获取的手工特征融合入深度学习的框架中。本研究提出一种不同的融合分类方法，这一用于非小细胞肺癌分类任务的融合方法将深度学习、影像组学以及患者临床信息相结合，它将深度学习模型得到的深层抽象信息、手工影像组学特征以及患者临床信息相结合放入机器学习分类器中，弥补了手工特征中缺少的深层抽象信息，进一步提升非小细胞肺癌组织学亚型分类的准确性，以期在肺癌筛查中辅助放射科医生对非小细胞肺癌患者的组织学亚型做出预判和指导，为临床治疗方式的选择和预后判断提供依据。图 4.1 显示了本研究的实验流程：

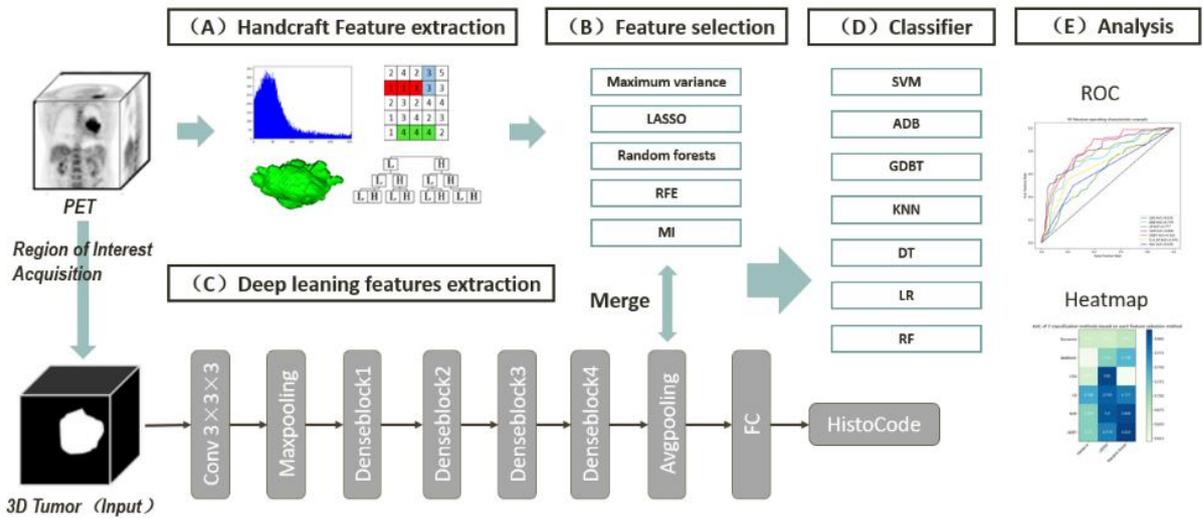


图 4.1 融合分类方法结构图

融合模型包括了五个部分，分别是（A）手工影像组学特征提取；（B）手工影像组学特征的选择；（C）深度学习特征的提取；（D）分类器构建；（E）模型分类性能的评估。手工影像组学特征提取部分研究使用传统影像组学方法获取了包括一阶、形状、纹理和基于小波滤波器的共 787 个特征；在特征选择部分五种特征选择方法对高维手工

特征进行挑选，得到相关特征集合；深度学习特征则由深度学习网络进行提取；分类器以融合了手工影像组学特征、深度学习特征和患者临床信息的特征集为基础构建；模型评估部分使用分类任务常用的指标量化，并借助热图进行展示。

融合模型的构建中，研究以传统影像组学方法选出高维影像组学特征中的最优相关特征子集，通过常用卷积神经网络分类性能的对比确定最优深度特征提取网络。因此，在本章中，研究展示了传统影像组学方法、深度学习技术在 151 位非小细胞肺癌组织学亚型分类任务上的性能以构建最优的融合分类模型。

## 4.2 影像组学特征提取与选择

目前影像组学已有成熟的软件可以使用，本研究使用了 python 环境下的 Pyradiomics 开源包，从医学影像中提取影像组学特征。使用 Pyradiomics 从每个患者病灶区域中提取 787 个手工影像组学特征，这些特征包含了以下四种类型：(1) 3D 形状特征(特征数量 = 13)；(2) 一阶强度特征(特征数量 = 18)；(3) 纹理特征(特征数量 = 68)；(4) 小波特征(特征数量 = 688)。表 4.1 显示了从 PET 图像获得的特征类型和数量。

表 4.1 影像组学特征统计

| 特征            | 统计  |
|---------------|-----|
| 汇总：数量 (N)     | 787 |
| 形状特征：N        | 13  |
| 强度特征：N        | 18  |
| 纹理特征：N        | 68  |
| 灰度共生矩阵：N      | 22  |
| 灰度游程矩阵：N      | 16  |
| 灰度大小区域矩阵：N    | 16  |
| 灰度依赖矩阵：N      | 14  |
| 小波特征：N        | 688 |
| 小波-强度特征：N     | 144 |
| 小波-灰度共生矩阵：N   | 176 |
| 小波-灰度游程矩阵：N   | 128 |
| 小波-灰度大小区域矩阵 N | 128 |
| 小波-灰度依赖矩阵：N   | 112 |

利用影像组学提取出的特征形成了高维的特征空间，这些高维特征的直接使用极有可能导致过拟合，破坏预测模型的泛化能力。且由于影像组学方法提取的手工特征大多

是冗余的，通常在使用手工影像组学特征是需要减少高度相关的特征以避免共线性。另一方面，高维特征降低了分类方法的计算效率与速度。基于上述因素，特征选择在影像组学方法中不可或缺，本文通过常用的五种特征选择算法来降低特征空间的维数，达到特征选择的目的。特征选择流程在图 4.2 中展示。

特征选择过程包括了数据归一化，特征选择算法的应用两部分，最终得到维数较低、冗余度较低的有效影像组学特征。使用患者 PET 影像和医生手动标记的病灶区域在 Pyradiomics 开源包上提取的 787 维影像组学特征量纲上参差不齐，归一化被用来将特征限定在一定的范围内量纲不一对整体数据集带来的不良影响。

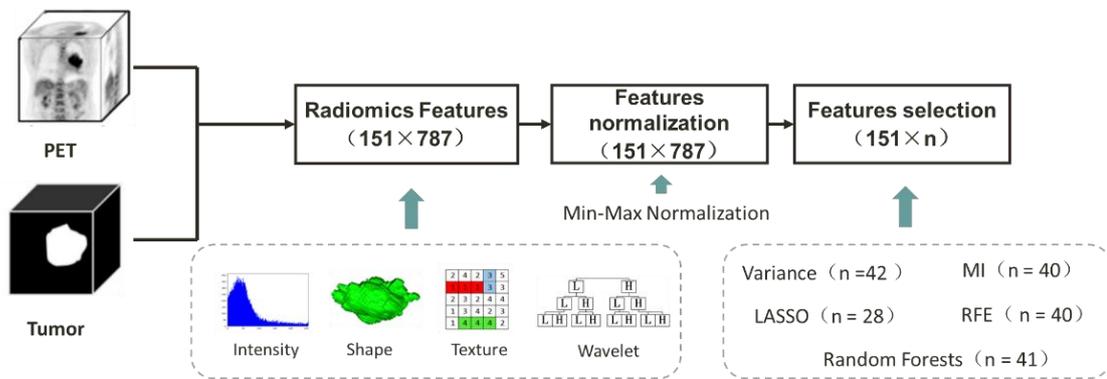


图 4.2 影像组学特征选择流程图

归一化后的影像组学特征的维数并没有改变，为了达到降低维数避免“维数灾难”的目的，研究选取了 5 种特征选择方法：方差选择法（Variance）、互信息法（Mutual Information, MI）、递归特征消除法（Recursive feature elimination, RFE）、最小绝对收缩和选择算子（Least absolute shrinkage and selection operator, LASSO）、随机森林（Random Forest, RF）特征选择法。经过特征选择后的影像组学特征维数较低，用于最终融合分类模型的构建。

### 4.3 深度学习特征提取

考虑到深度学习技术对图像深层抽象特征挖掘的有效性，研究基于稠密网络（DenseNet）构建了本文的深度学习特征提取网络。稠密网络理论基础来源于 ResNet 等非直连网络，主要思路是层与层之间的稠密连接（Dense-block），这一稠密块使每一层的输入都是之前所有层输出的拼接。稠密网络的每个稠密块（Dense-block）中每层的

输出特征图样大小一致可以直接连接,不同块间的特征大小不同,通过过渡层(Transition layer)中包含的池化层(pooling layer)实现下采样<sup>[32]</sup>。具体结构可以参照图 4.3。

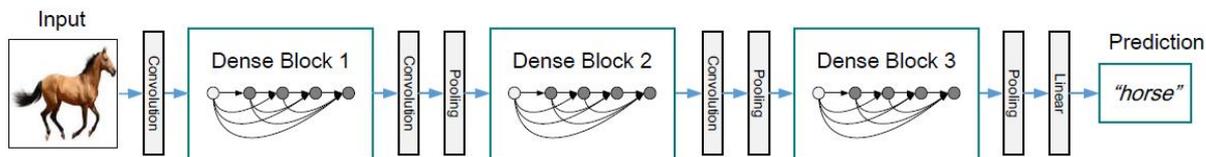


图 4.3 稠密网络示意图<sup>[32]</sup>

稠密网络充分利用输入层影像和每一层特征图的信息,它所具有的稠密块使得它能够更好地联系上下文信息,获取更加有效的图像表征。这对于小数据集的医学图像分析研究非常友好。因此,文章通过将稠密网络与其他两种常用于组织学亚型分类任务的网络模型进行对比后选择稠密网络作为最终深度抽象特征提取网络。本研究使用简化后的稠密网络,在模型下端的平均池化层输出 51 维的深度学习特征用于非小细胞肺癌组织学亚型分类融合模型中特征融合部分。

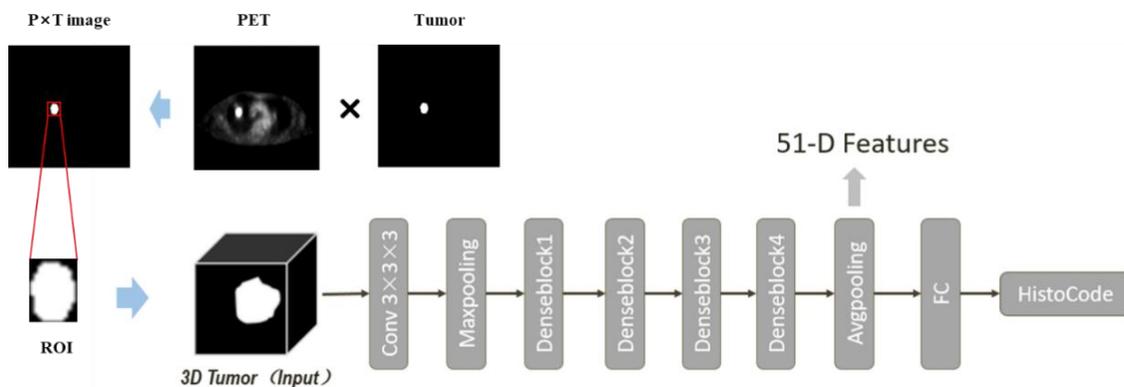


图 4.4 深度学习特征提取流程图

数据集中所包含的医生手动勾画的病灶标签为二值图,难以提供患者病灶区域的有效特征。因此,研究使用肿瘤轮廓图在校正后的 PET 图像中获取病灶区域。处理后的病灶区域作为深度学习特征提取网络—稠密网络的输入层。深度学习特征提取的流程在图 4.4 中展示。

患者病灶区域经由有经验的医师进行手动标注,得到了肿瘤图像,这一图像为二值图。在肿瘤图中,亮的部分表示患者的肿瘤区域,其像素值为 1,周围暗部为背景区域,

像素值为 0，研究使用乘法操作将患者校正后的 PET 影像与肿瘤图像相乘，得到的包含了患者原始 PET 影像信息的肿瘤区域（ROI）。含有信息的肿瘤区域三维影像经过尺寸归一化到  $64 \times 64 \times 64$  像素大小作为深度学习特征提取网络的输入层。经过核大小为  $3 \times 3 \times 3$  的卷积层、最大池化层、4 个稠密块在平均池化层中获取了深层的深度学习特征。

#### 4.4 特征融合与分类模型构建

经过特征选择的影像组学特征、深度特征提取网络获取的深度学习特征以及 5 种包括年龄、性别、种族、吸烟史、最大 SUV 值的临床特征进行融合，得到最终用于非小细胞肺癌组织学亚型分类的特征向量。这一特征向量既弥补了手工影像组学特征不能覆盖到的方面，同时为深度特征提取的深层信息弥补了低阶信息，临床特征则从临床角度诠释了影像信息。研究使用拼接的方式对手工影像组学特征、深层深度学习特征和 5 个量化好的临床信息进行融合，简单来说就是将三个一维向量拼接为一个一维长向量。拼接完成的特征向量放入分类器中得到非小细胞肺癌患者最终组织学亚型的分类。图 4.5 描述了特征融合与分类过程。

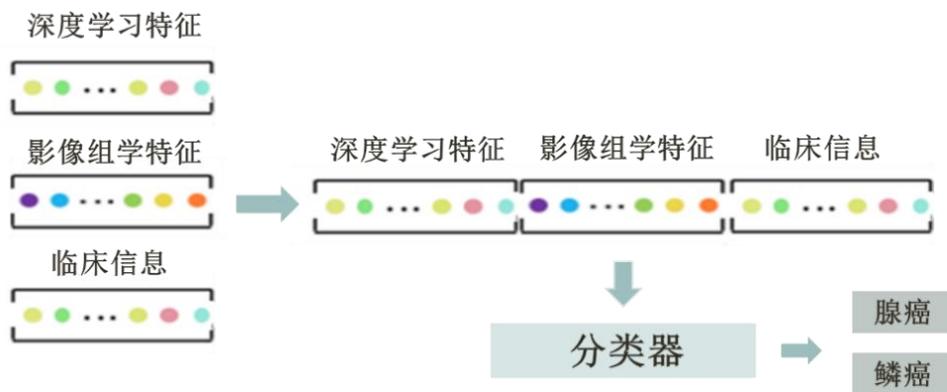


图 4.5 特征融合与分类

K-最近邻分类器 (K-nearest neighbor classification, kNN)、随机森林分类器 (Random Forest Classifiers, RF)、自适应提升算法 (Adaptive Boosting, ADB)、支持向量机 (Support Vector Mechine, SVM)、梯度提升算法 (Gradient Boosting, GDBT)、逻辑回归 (Logistic regression, LR) 和决策树算法 (Decision Tree, DT) 这 7 种较常使用的分类器被用来区分非小细胞肺癌患者的组织学亚型类型。对于每种特征选择方法和分类方法的交叉组合、不同特征的融合和分类方法的交叉组合均使用 5 折交叉验证得到最终分类结果。

## 4.5 融合分类模型及实验结果

融合模型整体构建如图 4.1 所示，使用手工影像组学制作的特征、深度学习特征以及患者临床信息构建机器学习分类器，对 151 例非小细胞肺癌患者的组织学亚型进行判别分类。

在模型结果对比中融合模型的各个部分均作为其对比方法进行结果展示。模型分类对比包括了传统影像组学模型（Conventional Radiomics Model, CRM），深度学习模型（Deep Learning Model, DLM），融合模型（Combined Model, CM）三个独立模型。传统影像组学模型是指使用手工影像组学特征与机器学习分类器，例如决策树、支持向量机等对非小细胞肺癌患者组织学亚型进行判别。其中传统影像组学模型中的手工影像组学特征来源于研究选取的 5 种特征选择方法选出的特征。深度学习模型是指利用深度神经网络，例如 VGG、稠密网络等从图像中直接学习信息并进行分类。

融合模型则是传统影像组学模型与深度学习模型的特征级融合，模型中的手工影像组学特征来源于传统影像组学模型中表现较好的特征选择方法。同样地，深度学习特征来源于用于比较的三种深度学习模型中表现最好的稠密网络。手工影像组学特征、深度学习特征和临床信息所构成的一维特征向量作为分类器的输入从而获得非小细胞肺癌患者组织学亚型分类结果。

模型实验结果的评估通过计算受试者工作特征(Receiver Operating Characteristic, ROC)曲线下面积(Area Under Curve, AUC)、准确度(accuracy)、精确度(precision)、召回率(recall)和 F1 值 (F1-score) 来量化。同时，研究将传统影像组学模型、深度学习模型和非小细胞肺癌组织学亚型分类融合模型进行比较，证实了文章阐述的非小细胞肺癌组织学亚型分类融合模型的性能优于传统影像组学模型和深度学习模型。实验将数据集划分成 5 份，每次实验使用其中 4 份数据作为训练数据集，剩下的 1 份数据作为测试数据集，共进行 5 次实验，使每份数据都充当过测试数据，5 次实验结果的平均性能作为最后的结果呈现。AUC 高于 0.60 的模型在类似研究中通常被认为具有预测性。

### 4.5.1 传统影像组学模型实验结果

使用传统影像组学方法构建的模型采用了 5 种特征选择算法和 7 种机器学习分类器。图 4.6 中反映了每个交叉组合的 AUC 值。总体而言，方差选择法选择的手工影像组学特征在 7 种分类器上的表现相对较差。相反，随机森林特征选择算法、递归特征消

除法和最小绝对收缩和选择算子特征选择算法表现出了相对较好的性能，综合看来以随机森林特征选择方法在各分类器中取得的效果最佳。

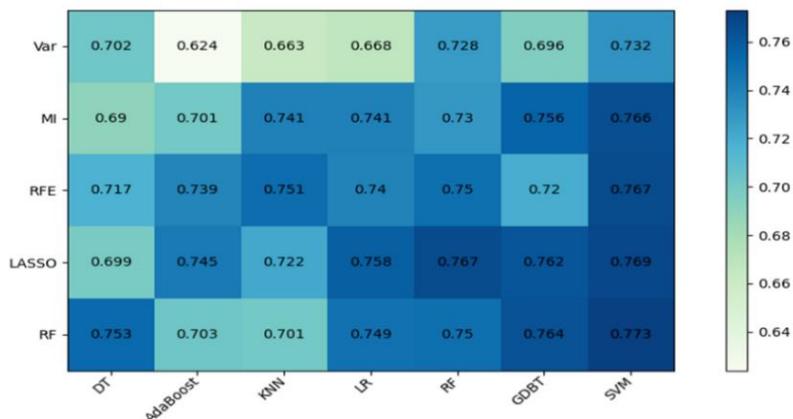


图 4.6 传统影像组学方法实验结果

研究使用的 7 种分类算法中，支持向量机（SVM）分类算法和随机森林（RF）分类算法的分类效果相较于其他分类器有较好效果。交叉组合 AUC 值前 5 名分别是随机森林（RF）特征选择算法+支持向量机（SVM）分类器、LASSO 特征选择方法+支持向量机（SVM）分类器、递归特征消除法（RFE）+支持向量机（SVM）分类器、LASSO 特征选择算法 + 随机森林（RF）分类器以及互信息法（MI）+支持向量机（SVM）分类器。其中，随机森林（RF）特征选择算法+支持向量机（SVM）分类器的 AUC 值、准确度(accuracy)、精确度(precision)、召回率(recall)和 F1 值（F1-score）分别为 0.773、0.695、0.696、0.664、0.666。

基于各个分类器及特征选择算法选出的相关特征集在传统影像组学模型中的表现，随机森林特征选择算法选出的手工影像组学特征和分类效果较好的支持向量机分类器被应用于融合分类模型中。

#### 4.5.2 深度学习分类模型实验结果

在深度学习模型中，研究使用了非小细胞肺癌组织学亚型分类任务常用的三种深度学习网络：VGG、残差网络和稠密网络对本研究的数据进行分类。通过使用相同的输入数据对于三种深度学习技术在 151 位非小细胞肺癌患者 PET/CT 影像上的分类性能的比较，稠密网络相较于其他两种网络有着更好的分类效果，在本研究数据集中得到 0.698

的 AUC 值, 同时, 网络的准确度(accuracy)、精确度(precision)、召回率(recall)和 F1 值 (F1-score) 分别为 0.690, 0.681, 0.683 和 0.682, 均优于其他两种网络。三种深度学习网络模型的性能在表 4.2 展示。

表 4.2 三种常用深度学习分类模型性能比较

| 深度学习网络   | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|----------|--------------|--------------|--------------|--------------|--------------|
| VGG      | 0.645        | 0.632        | 0.612        | 0.611        | 0.684        |
| ResNet   | 0.658        | 0.645        | 0.639        | 0.641        | 0.681        |
| DenseNet | <b>0.690</b> | <b>0.681</b> | <b>0.683</b> | <b>0.682</b> | <b>0.698</b> |

基于深度学习分类模型的性能, 在最终的融合分类模型中, 研究使用分类效果最好的稠密网络作为深度学习特征提取网络获得非小细胞肺癌组织学亚型分类融合分类模型所需的深度学习特征。

### 4.5.3 融合分类模型实验结果

在上述实验中, 研究获得了传统影像组学分类方法最优分类组合和深度学习分类方法的最优模型。因此, 在非小细胞肺癌组织学亚型融合分类模型中, 使用在传统影像组学分类模型中得到最优的手工影像组学特征集、在深度学习分类模型中确定最优的深度特征提取器 (DenseNet) 获得的最优深度学习特征集以及五个量化好的临床信息构建了非小细胞肺癌组织学亚型融合分类模型。其中, 模型比较分析中分类器使用了支持向量机算法。

最终的非小细胞肺癌组织学亚型融合分类模型与深度学习分类模型、传统影像组学分类模型进行比较。比较结果在表 4.3 中展示, 非小细胞肺癌组织学亚型融合分类模型显示出最好的性能。

表 4.3 非小细胞肺癌组织学亚型融合分类模型性能比较

| 模型            | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|---------------|--------------|--------------|--------------|--------------|--------------|
| 深度学习-DenseNet | 0.690        | 0.681        | 0.683        | 0.682        | 0.698        |
| 传统影像组学-SVM    | 0.695        | 0.696        | 0.664        | 0.666        | 0.773        |
| 融合分类模型        | <b>0.716</b> | <b>0.708</b> | <b>0.698</b> | <b>0.701</b> | <b>0.816</b> |

然而, 上述比较不能展示使用机器学习分类器作为最终分类手段是有效的, 因此研究比较了深度学习分类方法、将手工影像组学特征融入深度学习分类模型直接进行分类

的深度学习+影像组学方法、非小细胞肺癌组织学亚型融合分类方法的分类性能。其中，传统影像组学特征是由随机森林（RF）特征选择算法选出的 PET 手工影像组学特征，三种模型性能展示在表 4.4 中。深度学习与影像组学结合的模型得到了 0.737 的 AUC 值，相比于深度学习模型在各个指标上都有部分提升。将深度学习特征与手工影像组学特征相结合放入机器学习分类器中（以支持向量机为例）得到了 0.816 的 AUC 值。各个指标相较于前面两种分类模型有较大的提升。这一比较一方面显示了使用机器学习分类器作为最后的分类手段能够提升非小细胞肺癌组织学亚型分类任务的性能；另一方面，比较结果也展示了传统影像组学手工特征与深度学习特征的融合是有效的。

表 4.4 融合分类模型与不使用机器学习分类器模型性能比较

| 模型            | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|---------------|--------------|--------------|--------------|--------------|--------------|
| 深度学习-DenseNet | 0.690        | 0.681        | 0.683        | 0.682        | 0.698        |
| 深度学习+传统影像组学   | 0.697        | 0.687        | 0.681        | 0.683        | 0.737        |
| <b>融合分类模型</b> | <b>0.716</b> | <b>0.708</b> | <b>0.698</b> | <b>0.701</b> | <b>0.816</b> |

为了更充分的比较，降低结果出现的随机因素，研究还比较了三种深度学习模型加入机器学习分类器与直接分类模型的性能比较，实验所用机器学习分类器为支持向量机，直接将深度学习网络提取到的特征放入分类器中得到分类结果。实验结果展示在表 4.5 中。可以看到，加入机器学习分类器后，三种卷积神经网络方法性能都得到了一定程度的提升，其中以 VGG 模型新能提升最大，在各个指标上的提升分别为 0.045、0.049、0.071、0.017 和 0.0097。

表 4.5 深度学习分类模型与加入分类器的深度学习分类模型性能比较

| 模型                    | Accuracy     | Precision    | Recall       | F1-score     | AUC          |
|-----------------------|--------------|--------------|--------------|--------------|--------------|
| VGG                   | 0.645        | 0.632        | 0.612        | 0.611        | 0.684        |
| <b>VGG + 分类器</b>      | <b>0.690</b> | <b>0.681</b> | <b>0.683</b> | <b>0.682</b> | <b>0.781</b> |
| ResNet                | 0.658        | 0.645        | 0.639        | 0.641        | 0.681        |
| <b>ResNet + 分类器</b>   | <b>0.671</b> | <b>0.663</b> | <b>0.663</b> | <b>0.663</b> | <b>0.706</b> |
| DenseNet              | 0.690        | 0.681        | 0.683        | 0.682        | 0.698        |
| <b>DenseNet + 分类器</b> | <b>0.697</b> | <b>0.690</b> | <b>0.693</b> | <b>0.691</b> | <b>0.700</b> |

同时，为了比较五种特征选择方法选出的影像组学特征与深度学习特征融合后的性能，研究还比较了使用不同特征选择算法选择出的手工影像组学特征对分类模型性能的影响。选用的模型为 DenseNet 和支持向量机（SVM）的组合其中五种特征选择方法是：

方差选择法 (Variance)、互信息法 (MI)、递归特征消除法 (RFE)、LASSO、随机森林 (RF)。在互信息法、LASSO 和随机森林算法上取得了较好的分类性能。AUC 值分别为 0.802、0.805 和 0.816。

鉴于以上实验结果,手工影像组学特征、深度学习特征融合的有效性,研究在最终的非小细胞肺癌组织学亚型融合分类模型中添加了 5 种量化好的临床信息和非小细胞肺癌患者 CT 影像特征以检验其分类性能。其中,临床信息使用归一化方法统一其量纲。从 CT 影像中获取的影像组学特征数量与特征种类与 PET 影像中获取的相一致,并使用同样的特征选择算法选出相关特征集合。

表 4.6 展示了使用不同特征组的分类模型在非小细胞肺癌组织学亚型分类上的性能,这些特征组合包括: PET 手工影像组学特征+深度学习特征、PET 手工影像组学特征+临床信息+深度学习特征以及 PET 手工影像组学特征+CT 手工影像组学特征+临床信息+深度学习特征三种组合。实验表明,加入了从 CT 影像中获取的手工影像组学特征与临床信息模型性能能够得到部分提升,其中使用 PET 影像组学特征与临床特征组与深度学习特征融合得到了最高的 AUC 值,为 0.828。

表 4.6 融合分类模型中使用不同特征组的分类模型性能比较

| 模型 (SVM)                | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|-------------------------|--------------|--------------|--------------|--------------|--------------|
| PET 影像组学特征+深度学习特征       | 0.716        | 0.708        | 0.698        | 0.701        | 0.816        |
| PET/CT 影像组学+临床信息+深度学习特征 | 0.729        | 0.722        | <b>0.711</b> | <b>0.714</b> | 0.817        |
| PET 影像组学特征+临床信息+深度学习特征  | <b>0.729</b> | <b>0.723</b> | 0.709        | 0.713        | <b>0.828</b> |

为了进一步比较 PET 影像组学特征、CT 影像组学特征以及临床信息在非小细胞肺癌组织学亚型融合分类模型中的表现,我们使用随机森林作为特征选择方法,支持向量机 (SVM) 作为分类器对比了不同特征及其交叉组合在非小细胞肺癌组织学亚型融合分类模型中的表现,具体结果展示在表 4.7 中。这一比较显示了 PET 影像组学特征、CT 影像组学特征、临床特征及深度学习特征构成的特征组在非小细胞组织学亚型融合分类模型中取得较好效果。

在最终的细胞肺癌组织学亚型分类融合模型中,研究共使用了 7 种机器学习分类器比较了七种分类算法在不同特征集下的性能,图 4.7 显示了这些特征选择算法或特征组合在不同分类器中的性能,以 AUC 值构成的热图展示性能优越性,颜色越深即模型 AUC 值越高。

表 4.7 融合分类模型中使用不同特征组的分类模型性能比较

| 特征/特征组                     | 准确率          | 精确率          | 召回率          | F1 值         | AUC          |
|----------------------------|--------------|--------------|--------------|--------------|--------------|
| CT 影像组学特征+深度学习特征           | 0.697        | 0.703        | 0.672        | 0.673        | 0.745        |
| PET 影像组学特征+深度学习特征          | 0.716        | 0.708        | 0.698        | 0.701        | 0.816        |
| 临床特征+深度学习特征                | 0.690        | 0.693        | 0.653        | 0.653        | 0.788        |
| PET/CT 影像组学特征+深度学习特征       | 0.710        | 0.702        | 0.688        | 0.691        | 0.775        |
| CT 影像组学特征+临床特征+深度学习特征      | 0.665        | 0.652        | 0.642        | 0.644        | 0.745        |
| PET 影像组学特征+临床特征+深度学习特征     | <b>0.729</b> | <b>0.723</b> | 0.709        | 0.713        | <b>0.828</b> |
| PET /CT 影像组学特征+临床特征+深度学习特征 | <b>0.729</b> | 0.722        | <b>0.711</b> | <b>0.714</b> | 0.817        |

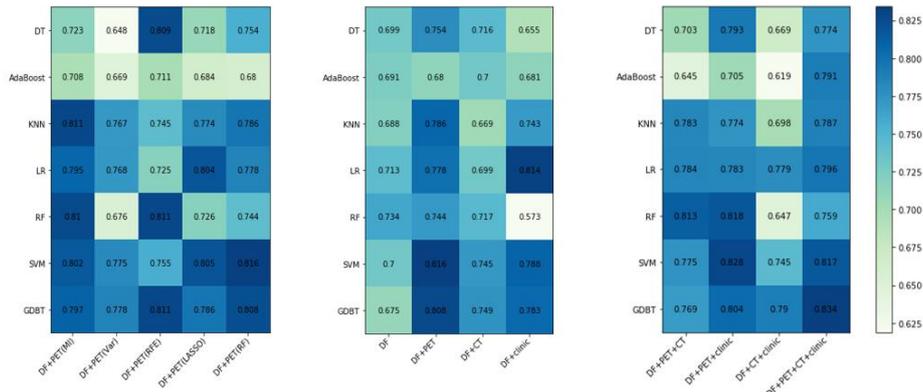


图 4.7 七个分类器的 AUC 值热图 (DF 表示深度学习模型提取的深度特征)

在各个交叉组合得到的 AUC 值构成的热图中，支持向量机（SVM）和梯度提升分类算法（GDBT）在每个特征集上的表现都优于其他分类器。其中，梯度提升分类算法（GDBT）在由 PET 影像组学特征、CT 影像组学特征、临床特征和深度学习特征相结合的特征集上获得了最优 AUC 值为 0.834，准确率、精确率、召回率和 F1 值分别为 0.755、0.752、0.736 和 0.740。

## 4.6 本章小结

本章节具体阐述了本文提取的基于强度、形状、纹理和基于小波的影像组学特征的具体提取和特征选择的过程和非小细胞肺癌组织学亚型融合分类模型中深度学习特征的提取流程。通过传统影像组学、深度学习技术的各种对比实验确定用于融合模型的最优手工特征集、最优深度学习特征提取器以及用于分类的机器学习分类器。最后详细描述非小细胞肺癌组织学亚型融合分类模型的结构和分类性能。并将其与不同方法进行比较，证明了融合模型的有效性。

## 5 非小细胞肺癌组织学亚型全自动分类方法

### 5.1 全自动分类模型总体结构

当前对于非小细胞肺癌组织学亚型研究中，大多依赖于医生对于病灶区域的标记。然而，这一工作是枯燥、费时、成本高的。因此，本研究提出一种端到端的非小细胞组织学亚型全自动分类方法，这一方法脱离了医生对于病灶区域标记，直接从非小细胞肺癌患者的医学影像中识别其组织学亚型，既避免了穿刺手术对于患者的伤害，同时解放医生双手，是一种成本低收益高的自动化分类方法。在这部分工作中，文章构建了机器学习分类模型、3D 深度学习分类模型以及 2D 深度学习分类模型，将三种全自动分类模型性能进行比较选取了最佳分类模型，即 2D 深度学习模型。这一最佳分类模型包括了三个部分，分别是 (A) PET 最大强度投影图特征提取；(B) CT 最大强度投影图特征提取；(C) 特征加权融合及分类。图 5.1 显示了最佳全自动分类模型结构：

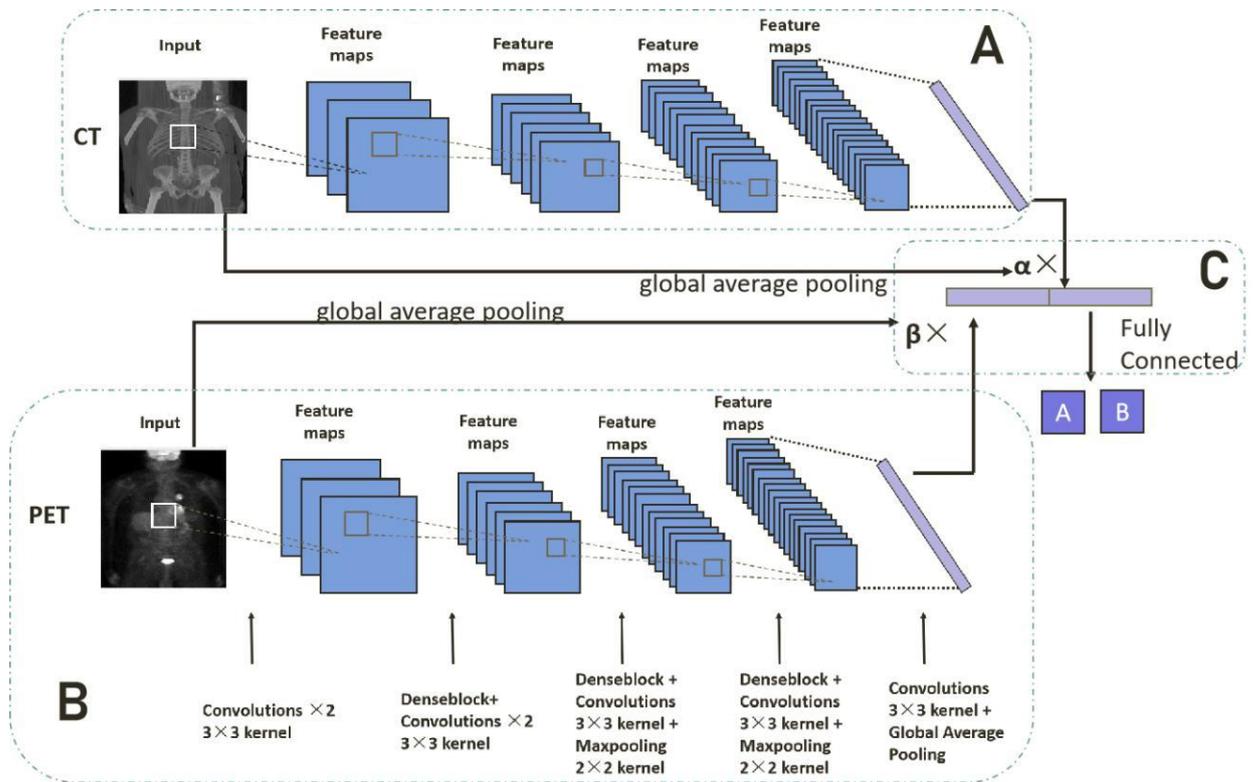


图 5.1 非小细胞肺癌组织学亚型全自动分类模型总体结构

在 PET/CT 影像特征提取部分使用 PET/CT 最大强度投影图作为输入层，2D 卷积神经网络作为特征提取网络得到 PET/CT 影像特征；特征融合部分则借助全局平均池化层赋予 PET/CT 特征权重，使用加权融合方式进行特征融合；分类部分则使用全连接层得到最终结果。

## 5.2 全自动分类模型数据预处理

在原始数据集中，151 位非小细胞肺癌患者原始 PET/CT 影像是 3 维立体影像，然而病灶区域只占据患者 3D 影像非常小的一部分。因此，在使用卷积神经网络在影像中区分非小细胞肺癌组织学亚型时需要对影像数据做进一步的处理，研究使用最大强度投影将 3D 影像转换为患者 2D 全身图像的最大强度投影（Maximal Intensity Projection, MIP）。

最大强度投影是一种简单的数据处理方法，它是按照患者影像某一视角使用透视法获得二维图像。具体来说，当光纤束穿过患者 3D 影像，影像中强度最大的像素值被保留投影到二维平面，形成了最大强度投影图。研究使用患者正面最大强度投影图作为深度学习网络的输入，最大强度投影示例图如 5.2 图所示。



图 5.2 非小细胞肺癌患者 2D 最大强度投影（从左至右分别为 PET、CT 的最大强度投影）

## 5.3 PET/CT 最大强度投影图特征提取

经过预处理后的 2D 最大强度投影被作为分类模型的输入层。PET 影像与 CT 影像所用特征提取网络一致，包括了卷积层、稠密块、最大池化层和全局平均池化层。以 PET 最大强度投影特征提取网络为例，具体结构如图 5.4 所示。

输入层 PET/CT 最大强度投影图大小为  $256 \times 256$ ，最大池化层步长设置为  $2 \times 2$ 。卷积层使用  $3 \times 3$  核大小的 2D 卷积，每一层卷积后衔接一个批正则化 (Batch Normalization, BN) 层，用于数据的归一化。经过卷积层、池化层、稠密块等操作网络最后使用全局平均池化将前一层得到的每张特征图平均为一个值，作为网络提取到的深度学习特征。在网络中，PET 最大强度投影全身图像和 CT 最大强度投影全身图像分别提取 128 维深度学习特征，共得到 256 维的 PET/CT 影像的深度学习特征。

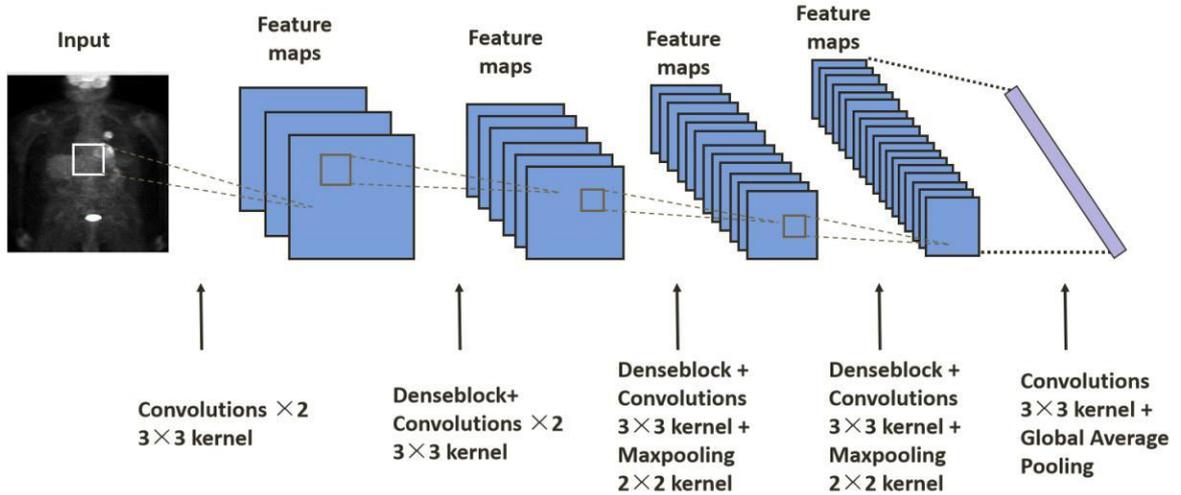


图 5.3 组织学亚型全自动分类模型网络结构

## 5.4 PET/CT 特征融合及分类

在临床中，CT 影像提供了患者的解剖结构信息，PET 则提供病变的代谢、功能等信息。因此，这从 PET 影像和 CT 影像所提取到的特征有各自不同的表征，对于网络正确区分组织学亚型也有不同作用。另一方面，在前文非小细胞肺癌组织学亚型融合分类模型中，患者 CT 影像被证实对于组织学亚型分类的贡献度较小。因此，为了区分 PET 影像与 CT 影像在网络中的作用，这一工作在非小细胞在 PET 影像特征和 CT 影像特征的融合上，使用了一种简单的加权融合方式，具体融合方法在图 5.4 中展示。

在特征提取网络中，研究使用 PET/CT 最大强度投影得到了 PET/CT 特征。同时，PET/CT 最大强度投影图经过全局平均池化层得到加权参数  $\alpha$  和  $\beta$ ，将  $\alpha$  与 CT 特征向量、 $\beta$  与 PET 特征向量分别相乘得到加权后的两组特征向量，最后使用连接操作将 PET 加权特征向量与 CT 加权特征向量相连得到融合后的特征向量并用于后续的分类。模型最终的分类由全连接层承担，在特征提取网络中研究得到了 PET/CT 影像加权融合后特

征向量,这一特征向量被放入网络最后的全连接层以得到全自动分类模型对于非小细胞肺癌患者组织学亚型最终的判断结果。

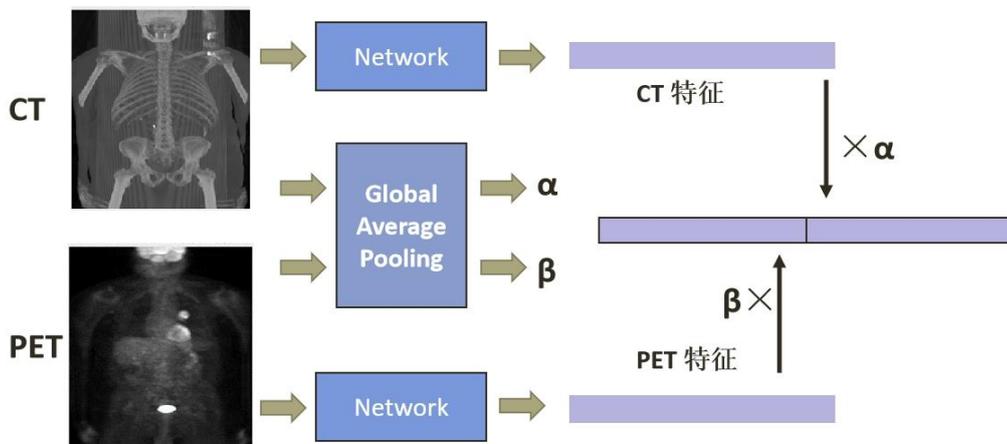


图 5.4 组织学亚型阅片分类模型特征融合方法示意图

## 5.5 全自动分类模型实验结果

在全自动分类模型实验结果的比较中,我们展示了机器学习方法、3D 深度学习模型与 2D 深度学习模型的性能,同时在深度学习方法中对比了使用不同模态图像的分类性能。模型实验结果的评估通过计算受试者工作特征(Receiver Operating Characteristic, ROC)曲线下面积(Area Under Curve, AUC)值、准确度(accuracy)、精确度(precision)、召回率(recall)和 F1 值(F1-score)来量化。

### 5.5.1 机器学习模型实验结果

在机器学习方法中,不依赖于医生对于病灶区域的标记也就意味着根据病灶区域得到的手工影像组学特征不能用于全自动分类任务。因此,研究以患者临床特征构建机器学习分类器来判断非小细胞肺癌患者的组织学亚型。由于患者临床信息维数很小,仅包含了患者的年龄、性别、种族、吸烟史、和最大 SUV 值五维信息,信息冗余度不高,因此,特征选择这一流程可以被省略。研究使用融合模型中比较的七种机器学习分类器,以五个临床特征作为输入来判断非小细胞肺癌患者组织学亚型,具体结果在表 5.1 中展示。

这一实验结果展示了使用患者临床因素和机器学习方法判断非小细胞肺癌患者组织学亚型是可行的,在七种机器学习分类器中,支持向量机和随机森林显示出较好的分类性能,但这一分类结果并不理想,七种分类器取得的准确度最优值仅为 0.682。

表 5.1 使用不同机器学习分类器模型的性能

| 分类器  | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|------|--------------|--------------|--------------|--------------|--------------|
| KNN  | 0.649        | 0.637        | 0.629        | 0.630        | 0.719        |
| ADB  | 0.656        | 0.654        | 0.644        | 0.644        | 0.693        |
| RF   | <b>0.675</b> | 0.665        | <b>0.656</b> | <b>0.658</b> | <b>0.750</b> |
| DT   | 0.576        | 0.571        | 0.573        | 0.571        | 0.654        |
| LR   | 0.649        | <b>0.721</b> | 0.586        | 0.546        | 0.747        |
| GDBT | 0.623        | 0.610        | 0.608        | 0.609        | 0.711        |
| SVM  | <b>0.682</b> | <b>0.754</b> | <b>0.653</b> | <b>0.654</b> | <b>0.752</b> |

### 5.5.2 3D 深度学习分类模型实验结果

使用深度学习方法构建的 3D 全自动分类模型的总体模型结构与 2D 模型一致，由步长  $2 \times 2 \times 2$  的最大池化层、核大小为  $3 \times 3 \times 3$  的 3D 卷积层和 3D 稠密块构成。在输入部分，由于患者病灶区域相对于整个影像是极小一部分，使用患者原始 3D 影像作为输入不仅会拉低网络计算和运行速度，还容易使网络受过较多冗余信息的干扰。因此研究不使用患者全身 3D 影像作为输入，而是取患者上半身影像作为 3D 深度学习分类模型的输入层。在 3D 深度学习分类模型中，实验对比了不同模态作为输入的分类性能，展示在表 5.2 中。

表 5.2 不同模态输入的非小细胞肺癌组织学亚型全自动分类模型性能比较

| 输入              | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|-----------------|--------------|--------------|--------------|--------------|--------------|
| CT 半身 3D 影像     | 0.600        | 0.616        | 0.610        | 0.596        | 0.634        |
| PET 半身 3D 影像    | 0.674        | 0.674        | 0.640        | 0.636        | 0.693        |
| PET/CT 半身 3D 影像 | <b>0.714</b> | <b>0.760</b> | <b>0.664</b> | <b>0.660</b> | <b>0.706</b> |

实验显示，以 PET/CT 半身 3D 影像作为 3D 网络的输入时，3D 模型得到了最佳的组织学亚型分类性能，且整体分类性能优于了机器学习分类模型中最佳分类器的分类性能。

### 5.5.3 2D 深度学习分类模型实验结果

端到端的 2D 深度学习全自动分类模型由融合特征经过全连接层得到组织学亚型分类结果，在模型中，研究比较了输入层为 PET 最大强度投影图、CT 最大强度投影图和 PET/CT 最大强度投影图的模型分类性能。将具体实验结果展示在表 5.3 中展示。

表 5.3 不同模态输入的非小细胞肺癌组织学亚型全自动分类模型性能比较

| 输入            | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|---------------|--------------|--------------|--------------|--------------|--------------|
| CT 最大强度投影     | 0.640        | 0.644        | 0.622        | 0.616        | 0.600        |
| PET 最大强度投影    | 0.692        | 0.724        | 0.692        | 0.672        | 0.737        |
| PET/CT 最大强度投影 | <b>0.728</b> | <b>0.754</b> | <b>0.692</b> | <b>0.688</b> | <b>0.759</b> |

根据分类结果, 2D 深度学习全自动分类模型在 PET/CT 最大强度投影图作为输入表现出最优性能。其中, 以 CT 最大强度投影图上表现最差, 而在 PET 最大强度投影图上的结果已经能够与将传统影像组学获取的手工影像组学特征融入到深度学习模型中的分类性能基本持平。同时, 2D 深度学习模型分类性能优于机器学习和 3D 深度学习分类模型分类性能, 成为最佳全自动分类模型。表 5.4 显示了三种全自动模型比较。

表 5.4 不同非小细胞肺癌组织学亚型全自动分类模型性能比较

| 模型      | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|---------|--------------|--------------|--------------|--------------|--------------|
| SVM     | 0.682        | 0.754        | 0.653        | 0.654        | 0.752        |
| 3D 深度学习 | 0.714        | <b>0.760</b> | 0.664        | 0.660        | 0.706        |
| 2D 深度学习 | <b>0.728</b> | 0.754        | <b>0.692</b> | <b>0.688</b> | <b>0.759</b> |

为了证实加权融合方式的有效性, 实验比较了加入权重参数的融合方式与不加入权重的融合方式在模型中的分类性能。其中, 不加权重参数即对网络提取到的 PET 影像特征和 CT 影像特征直接进行拼接。实验结果表明, 加入权重参数能够使得 PET/CT 特征更好的融合, 从而达到更好的分类效果。融合方式比较结果见表 5.5。

表 5.5 不同特征融合方式分类性能比较

| 融合方式  | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|-------|--------------|--------------|--------------|--------------|--------------|
| 非加权融合 | 0.678        | 0.692        | 0.634        | 0.618        | 0.754        |
| 加权融合  | <b>0.728</b> | <b>0.754</b> | <b>0.692</b> | <b>0.688</b> | <b>0.759</b> |

这一比较证明了加入权重的融合方式是有效的。同时, 我们看到在非加权融合方式中使用 PET/CT 最大强度投影的分类模型分类性能不如单独使用 PET 最大强度投影的分类结果, 原因可能是 CT 最大值投影对于非小细胞肺癌组织学亚型分类提供的信息作用较小, 这使得单独使用 PET 最大强度投影图分类性能优于非加权融合模型的性能。

患者临床特征对于非小细胞肺癌患者组织学亚型判别中的作用在融合模型中已经被证实，因此，研究还在最佳全自动分类模型中融入患者临床特征，这些特征包括了患者年龄、性别、种族、吸烟史以及最大 SUV 值。加入了临床信息的全自动分类模型与原始模型相比较，结果在表 5.6 中展示。加入临床特征的组织学亚型自动分类模型在 CT 最大强度投影图、PET 最大强度投影图以及 PET/CT 投影图作为输入时均有超越原始的全自动分类模型的性能。最终全自动分类模型在 PET/CT 最大值投影图+临床信息中端到端的非小细胞全自动分类模型得到最优性能，AUC 值、准确率、精确率、召回率和 F1 值分别为 0.793、0.768、0.790、0.760 和 0.750。

表 5.6 加入临床特征后全自动分类模型性能比较

| 输入                | 准确度          | 精确度          | 召回率          | F1 值         | AUC          |
|-------------------|--------------|--------------|--------------|--------------|--------------|
| CT 影像             | 0.640        | 0.644        | 0.622        | 0.616        | 0.600        |
| CT 影像+临床特征        | <b>0.660</b> | <b>0.654</b> | <b>0.628</b> | <b>0.614</b> | <b>0.652</b> |
| PET 影像            | 0.692        | 0.724        | 0.692        | 0.672        | 0.737        |
| PET 影像+临床特征       | <b>0.706</b> | <b>0.736</b> | <b>0.722</b> | <b>0.698</b> | <b>0.757</b> |
| PET 影像+CT 影像      | 0.728        | 0.754        | 0.692        | 0.688        | 0.759        |
| PET 影像+CT 影像+临床特征 | <b>0.768</b> | <b>0.790</b> | <b>0.760</b> | <b>0.750</b> | <b>0.793</b> |

这一工作不依赖于医生手动标注的病灶区域，且不经肿瘤分割过程，直接从患者影像识别非小细胞肺癌组织学亚型，这相较于先前非小细胞肺癌组织学亚型分类研究工作是一大突破。另一方面，非小细胞肺癌组织学亚型全自动分类模型效果基本与融合模型相当，这表明不依赖于病灶区域标记的全自动非小细胞肺癌组织学亚型分类是可行且有效的。同时，非小细胞全自动分类模型中创造性地使用了患者 PET/CT 最大值投影 2D 全身图像，这是相较于以往对于非小细胞肺癌组织学亚型分类研究的一大创新之处。

## 5.6 本章小结

在非小细胞肺癌组织学亚型全自动分类模型这一章节中，文章构建了三种分类模型以选择最佳的全自动分类模型来区分非小细胞肺癌患者的组织学亚型。通过对比以 2D 深度学习模型作为最终的全自动分类模型。针对这一最佳的全自动分类模型，章节展示了其总体结构，模型构建过程以及模型的性能。这一工作证明了不依赖手动标注的病灶位置的全自动分类模型能够区分非小细胞肺癌患者的组织学亚型。

## 6 总结与展望

本研究使用 18f-氟脱氧葡萄糖(18FFDG)标记物的正电子发射断层扫描和计算机断层扫描(PET/CT)影像在 151 位非小细胞肺癌患者数据集中自动区分组织学亚型(腺癌和鳞状细胞癌)。研究目标是无创预测非小细胞肺癌组织学亚型。研究最终成功构建并验证了一个包含临床信息、手工影像组学特征以及深度学习特征的非小细胞肺癌组织学亚型融合分类模型和一个脱离医生手动标记病灶区域的非小细胞肺癌患者组织学亚型全自动分类模型。两个模型均在非小细胞肺癌组织学亚型的自动区分上有良好的效果,能够在区分非小细胞肺癌组织学亚型这一临床上的重要因素提供帮助。

在非小细胞肺癌中,肿瘤组织学亚型不同导致非小细胞肺癌患者不同的临床策略和预后<sup>[33][34]</sup>。研究发现影像组学和自然图像识别任务中有效的卷积神经网络均可以用来区分非小细胞肺癌中最常见的组织学亚型。随着影像组学和神经网络在医学成像中显示出的巨大的应用潜力。一系列研究对比了影像组学和神经网络在非小细胞肺癌组织学亚型分类上的性能。Chaunzwa<sup>[20]</sup>等人与 Han<sup>[19]</sup>等人在非小细胞肺癌组织学亚型分类上都对比了传统影像组学与卷积神经网络分类模型的性能,证实了神经网络在区分非小细胞肺癌组织学亚型上的潜力。本研究提出将传统影像组学所提取的手工影像组学特征和卷积神经网络提取的深度学习特征相结合来进一步提升非小细胞肺癌组织学亚型的分类准确度的融合分类方法。在 151 位接受了 PET/CT 检查的非小细胞肺癌患者数据集中证实了将传统影像组学得到的手工影像组学特征融合到卷积神经网络中能够得到了优于单独使用深度学习算法或是传统影像组学方法的分类性能,这在一定程度上证明了手工影像组学特征和深度学习特征具有互补性。同时,这一结果直接证明了结合深度学习特征与手工影像组学特征在区分非小细胞肺癌患者组织学亚型任务中是有效的,且分类性能优于单独使用深度学习或传统影像组学特征。

同时,本文将卷积神经网络提取到的深度学习特征放入机器学习分类器中,得到了优于原始深度学习模型的性能,这与 Zeju Li 等人<sup>[22]</sup>和 Stephen Notley 等人<sup>[23]</sup>的研究有着相似结论,表明了使用神经网络作为一种手段来提取特征以供机器学习模型对某些数据集进行分类可能是一种合理的做法。基于这一论断,研究将构建并验证了一个包含临床因素、手工影像组学特征以及深度学习特征的非小细胞肺癌分类融合模型,不同于将传统影像组学得到的手工特征融入深度学习模型,研究使用堆叠的融合方法得到了包含源自传统影像组学的手工特征、源自深度学习方法构建的卷积神经网络提取的深度学习

习特征和五个量化好的临床信息的一维特征向量，利用机器学习分类器作为分类手段对融合后的一维特征进行分类。实验结果显示，在本研究 151 例非小细胞肺癌患者组织学亚型影像数据集中以机器学习分类器作为最终分类手段结果优于将传统影像组学提取到的特征融入到深度学习构建的卷积神经网络中。这为非小细胞肺癌组织学亚型分类研究提供了一条新的思路。

另一方面，这一非小细胞肺癌融合分类模型是一个灵活的框架，它的各个部分都可以非常灵活的替换。在传统影像组学部分随着技术不断进步手工影像组学特征，可以根据更优秀的特征提取方法来获取更具有代表性的人工影像组学特征。同时，特征选择方法可以根据不同数据选择最合适的特征选择方法进行，这可以得到更具有数据针对性的人工影像组学特征集；在深度学习模型建立部分，同样地，不同研究可以根据研究数据的特点来选择神经网络结构，以获取最具有表征性的深度学习特征集；当然，在最后的分类器选择也仍然是可替换的，随着研究的不断深入，我们相信更多的、更有效的技术也会应运而生，因此，在分类器选择上随着时间的增长，可选择性也会增加。因此，研究所提出的非小细胞肺癌组织学亚型融合分类模型的灵活性决定了它能够随着新技术的出现而更新，具有更强的生命力。

在前述的现有研究中，对于非小细胞肺癌组织学亚型分类无论是传统影像组学方法、深度学习方法大多依赖于医生手动标记的病灶区域，而医生对于病灶区域的标记过程十分枯燥耗时，且这一标注通常存在主观性，且需要有经验的医生进行。因此，文章基于当前非小细胞肺癌组织学亚型研究的另一个空缺开展了另一分类工作——非小细胞肺癌组织学亚型全自动分类模型。这一工作目的在于减少非小细胞肺癌组织学亚型分类任务对于病灶轮廓标记的依赖性，同时解放医生双手，能够为医生节省时间做更有意义的事情。在全自动分类模型中，文章构建了三种全自动分类方法，分别是机器学习方法，3D 深度学习方法及 2D 深度学习方法，这三种方法均不依赖于肿瘤的分割，在经过处理的患者影像中直接进行分类任务。经过实验对比最终确定了使用加权融合的 2D 深度学习模型作为最佳全自动分类模型。在此基础上，研究还将患者临床信息加入最佳的全自动分类模型中，得到了与融合分类模型相当的性能，证实了加权融合的方式是可行且有用的。这一工作一大亮点是我们使用了患者最大值投影的全身 PET/CT 影像作为网络的输入。鉴于非小细胞肺癌中鳞状细胞癌和腺癌在人体中的位置通常有一定区别这一医学常识，患者最大值投影的全身图像能够提供更加全局的信息，有助于网络识别病灶区域的位置信息。同时，2D 网络相较于 3D 网络能够节省更多计算量，运行速度

也有所提高。这一工作证明了不依赖于患者病灶区域标记的非小细胞肺癌组织学亚型分类任务是可行的且有效的，这为以后非小细胞肺癌组织学亚型亦或是其他疾病的研究提供了新的方向。

虽然文章的两个工作在非小细胞肺癌组织学亚型的无创分类上都获得了不错的结果，但目前的研究仍具有局限性。首先，文章研究所使用的 151 位非小细胞肺癌患者影像仍然是一个样本数量较小的数据集，这对于建立更好的组织学亚型预测模型有一定的阻碍，且较少的样本数量十分容易导致构建的模型不够稳定。但同时我们注意到，由于医学影像拍摄成本高、收集整理更加不易，且完整的数据集通常要耗费医生大量时间成本对其进行标注。因此在医学影像的分析研究中，这问题是普遍存在的，解决这一问题一方面需要从多个中心收集样本进行模型训练，以提高模型的稳定性和泛化能力，但这一解决方法所消耗的时间成本十分巨大，因此，更加经济的解决方案或许在于建立高质量的医学影像公开库；另一方面这一局限性要求我们不断推进新的技术，使深度学习等方法有更适合于小数据集研究的模型。第二，在非小细胞肺癌组织学亚型融合分类模型研究中，随机森林特征选择方法所选出的手工影像组学特征在 151 位非小细胞肺癌患者影像构成的数据集上表现出最好的性能，但我们并不能确信随机森林特征选择算法选择出的手工影像组学特征是最优的。同样地，深度学习特征的提取和使用上也具有上述不足，这就需要对更多的特征选择算法进行大量实验从而保证所选择出的手工影像组学特征或是深度学习特征对非小细胞肺癌患者影像信息有最优的表征性。第三，非小细胞肺癌组织学亚型融合分类模型在训练阶段仍需要使用医生勾勒的肿瘤轮廓图，这仍旧是一项耗时费力的工作。虽然全自动的非小细胞肺癌组织学亚型分类模型已经克服了这一缺点，且全自动分类模型能够取得与融合分类模型相当的性能。但是从两个模型的 AUC 值来看，全自动分类模型与融合分类模型还有一定差距。因此，在克服对病灶区域人工标注的依赖这一问题上还有很长的路要走。最后，研究所建立的模型均使用特征级的融合方式，且融合方式简单。因此，需要更加行之有效的方法来进行特征融合，且图像级融合方式和决策级融合方式应该被考虑。

综上所述，传统影像组学方法和深度学习技术都能够区分非小细胞肺癌的组织学亚型，即腺癌和鳞状细胞癌。融合包含源自传统影像组学的手工特征、源自深度学习方法构建的卷积神经网络提取的深度学习特征和五个量化好的临床信息的一维特征向量使用机器学习分类器作为分类手段的非小细胞肺癌融合分类模型取得了相对较好的性能。同时，研究也表明了脱离医生对病灶区域的手动标注直接从患者医学影像辨别是可行且

有效的。这为非小细胞肺癌组织学亚型分类研究提供了新的方向。故此，本文针对非小细胞肺癌组织学亚型分类提出的两种分类模型作为有前途的智能辅助诊断工具，一方面能够进一步降低癌症诊断的成本，另一方面能够以无创方式在肺癌筛查中辅助放射科医生对非小细胞肺癌患者的组织学亚型的诊断做出预判和指导，为临床治疗方式的选择和预后判断提供依据，促进肺癌患者的个性化治疗。

## 参考文献

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*. 2021 Feb 4. doi: 10.3322/caac.21660. Epub ahead of print. PMID: 33538338.
- [2] Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases[J]. *Nature reviews. Cancer*. 2014 Aug;14(8):535-546. DOI: 10.1038/nrc3775.
- [3] Constanzo J, Wei L, Tseng HH, El Naqa I. Radiomics in precision medicine for lung cancer[J]. *Transl Lung Cancer Res* 2017 Dec;6(6):635-647.
- [4] Schuurbiens OC, Meijer TW, Kaanders JH, Looijen-Salamon MG, de Geus-Oei LF, van der Drift MA, et al. Glucose metabolism in NSCLC is histology-specific and diverges the prognostic potential of 18FDG-PET for adenocarcinoma and squamous cell carcinoma[J]. *J Thorac Oncol* (2014) 9:1485–93.
- [5] Inamura K. Lung cancer: Understanding its molecular pathology and the 2015 WHO classification[J]. *Front Oncol* (2017) 7:193. doi: 10.3389/fonc.2017.00193
- [6] Takeuchi S, Khiewvan B, Fox PS, Swisher SG, Rohren EM, Bassett RL, et al. Impact of initial PET/CT staging in terms of clinical stage, management plan, and prognosis in 592 patients with non-small-cell lung cancer[J]. *Eur J Nucl Med Mol Imaging*. 2014;41(5):906–14.
- [7] Liu B, Dong J, Xu C, Zuo C, Le J, Guan Y, et al. Accuracy of 18F-FDG PET/CT for lymph node staging in non-small-cell lung cancers[J]. *Chin Med J*. 2009;122(15):1749–54.
- [8] Bingxin Gu and Mingyuan Meng and Lei Bi and Jinman Kim and David Dagan Feng and Shaoli Song, Prediction of 5-year Progression-Free Survival in Advanced Nasopharyngeal Carcinoma with Pretreatment PET/CT using Multi-Modality Deep Learning-based Radiomics[J]. 2021
- [9] H.J. Aerts, E.R. Velazquez, et al., Decoding Tumour Phenotype by Noninvasive Imaging Using a Quantitative Radiomics Approach[J]. *Nature Communications*, vol. 5, 2014.
- [10] A. Hosny, H.J. Aerts, R.H. Mak, Handcrafted versus deep learning radiomics for prediction of cancer therapy response[J]. *Lancet Digit Health*, 1 (2019) e106-e107

- [11] Chaunzwa, T.L., Hosny, A., Xu, Y. et al. Deep learning classification of lung cancer histology using CT images[J]. *Sci Rep* 11, 5471 (2021).
- [12] Cufer T, Ovcaricek T, O'Brien ME. Systemic therapy of advanced non-small cell lung cancer: major-developments of the last 5-years[J]. *Eur J Cancer* (2013) 49:1216–25.
- [13] Wu W, Parmar C, Grossmann P, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology[J].*Front Oncol* 2016;6:71.
- [14] Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer[J]. *Eur Radiol* (2018) 28:2772–8.
- [15] Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata A, Chiti A, Sollini M. Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions[J]. *Eur J Nucl Med Mol Imaging*. (2018)45:1649-1660.
- [16] Ferreira JR, Junior, Koenigkam-Santos M, Cipriano FEG, et al. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases[J]. *Comput Methods Programs Biomed* 2018;159:23-30.
- [17] Ren C, Zhang J, Qi M, Zhang J, Zhang Y, Song S, Sun Y, Cheng J. Machine learning based on clinico-biological features integrated 18F-FDG PET/CT radiomics for distinguishing squamous cell carcinoma from adenocarcinoma of lung[J]. *Eur J Nucl Med Mol Imaging*. (2020) 10.
- [18] Yang Fengchang, Chen Wei, Wei Haifeng, Zhang Xianru, Yuan Shuanghu, Qiao Xu, Chen Yen-Wei. Machine Learning for Histologic Subtype Classification of Non-Small Cell Lung Cancer: A Retrospective Multicenter Radiomics Study[J]. *Frontiers in Oncology*. (2021) 10: 3051-8.
- [19] Han Y, Ma Y, Wu Z, Zhang F, Zheng D, Liu X, Tao L, Liang Z, Yang Z, Li X, Huang J, Guo X. Histologic subtype classification of non-small cell lung cancer using PET/CT images[J]. *Eur J Nucl Med Mol Imaging* (2021)48:350–360.
- [20] Chaunzwa TL, Christiani DC, Lanuti M, Shafer A, Diao N, Mak RH, et al. Using deeplearning radiomics to predict lung cancer histology[J]. *J Clin Oncol* (2018) 36:8545–5.
- [21] Atsushi Teramoto, Tetsuya Tsukamoto, Yuka Kiriya, Hiroshi Fujita. Automated Classification of Lung Cancer Types from Cytological Images Using Deep Convolutional

- Neural Networks[J]. *BioMed Research International*. (2017)7: 6.
- [22] Li, Z., Wang, Y., Yu, J., Guo, Y. & Cao, W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma[J]. *Sci. Rep.* 7, 5467 (2017).
- [23] Notley, S., & Magdon-Ismail, M. Examining the use of neural networks for feature extraction: a comparative analysis using deep learning, support vector machines, and k-nearest neighbor classifiers[J]. *arXiv [csLG]* (2018). <http://arxiv.org/abs/1805.02294>
- [24] Setiono, R. & Liu, H. Feature extraction via Neural networks. In *Feature extraction, construction and selection: a data mining perspective*[J]. Springer, Boston, MA, 1998, 191–204.
- [25] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S. Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nat Rev Clin Oncol*. 2017 Dec;14(12):749-762. doi: 10.1038/nrclinonc.2017.141. Epub 2017 Oct 4. PMID: 28975929.
- [26] Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics[J]. *Radiol Med*. 2021 Oct;126(10):1296-1311. doi: 10.1007/s11547-021-01389-x. Epub 2021 Jul 2. PMID: 34213702; PMCID: PMC8520512.
- [27] Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillon-Robin, J. C., Pieper, S., Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic Phenotype[J]. *Cancer Research*, 77(21), e104–e107.
- [28] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. *J Big Data* 8, 53 (2021). <https://doi.org/10.1186/s40537-021-00444-8>
- [29] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*. 2015;521(7553):436–44.
- [30] Zhang Z, Cui P, Zhu W. Deep learning on graphs: a survey[J]. *IEEE Trans Knowl Data Eng*. 2020.
- [31] Goodfellow I, Bengio Y, Courville A, Bengio Y. *Deep learning*, vol. 1. Cambridge: MIT press; 2016.

- [32] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks[J]. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–08.
- [33] Yuan CZ, Tao XT, Zheng DF, Pan YJ, Ye T, Hu H, et al. The lymph node status and histologic subtypes influenced the effect of postoperative radiotherapy on patients with N2 positive IIIA nonsmall cell lung cancer[J]. *J Surg Oncol*. 2019;119(3):379–87..
- [34] McAleese J, Taylor A, Walls GM, Hanna GG. Differential relapse patterns for non-small cell lung cancer subtypes adenocarcinoma and squamous cell carcinoma: implications for radiation oncology[J].*Clin Oncol*. 2019;31(10):711–9.
- [35] Shao H, Ma X, Gao Y, Wang J, Wu J, Wang B, et al. Comparison of the diagnostic efficiency for local recurrence of rectal cancer using CT, MRI, PET and PET-CT: A systematic review protocol[J]. *Medicine (Baltimore)*. 2018;97(48):e12900.
- [36] Bertagna F, Albano D, Giovanella L, Giubbini R, Treglia G. F18-choline/C11-choline PET/CT thyroid incidentalomas[J]. *Endocrine*. 2019;64(2):203-208
- [37] Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection[J]. *J. Mach. Learn. Res.* 3, null (3/1/2003), 1157–1182.
- [38] Mohamed Bannasar, Yulia Hicks, and Rossitza Setchi. 2015. Feature selection using Joint Mutual Information Maximisation[J]. *Expert Syst. Appl.*42, 22 (December 2015), 8520–8532. DOI:<https://doi.org/10.1016/j.eswa.2015.07.007>
- [39] Gholami, B., Norton, I., Tannenbaum, A. R., & Agar, N. Y. (2012). Recursive feature elimination for brain tumor classification using desorption electrospray ionization mass spectrometry imaging[J]. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 2012, 5258–5261.
- [40] Ellis DE, Hubbard RA, Willis AW, Zuppa AF, Zaoutis TE, Hennessy S. Comparing LASSO and random forest models for predicting neurological dysfunction among fluoroquinolone users[J]. *Pharmacoepidemiol Drug Saf*. 2021 Dec 8. doi: 10.1002/pds.5391. Epub ahead of print. PMID: 34881470.
- [41] Speiser JL, Miller ME, Tooze J, Ip E. A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling[J]. *Expert Syst Appl*. 2019 Nov

15;134:93-101. doi: 10.1016/j.eswa.2019.05.028. Epub 2019 May 23. PMID: 32968335; PMCID: PMC7508310.

[42] 周志华. 机器学习[M].北京:清华大学出版社, 2000:30-35.

[43] Jerome H. Friedman .Greedy function approximation: A gradient boosting machine[J]. The Annals of Statistics, Ann. Statist. 29(5), 1189-1232, (October 2001)

## 致 谢

我们被无法预知却又早有预料的光阴推着结束了三年学生生涯，回想三年的故事，它像一本镌刻在记忆里的小说，模糊而又无比清晰。我想要用最华丽的辞藻去形容它，却发现简单平静是对这段日子最好的诠释。三年日复一日地从宿舍到实验室，从实验室到食堂这短短的路程简单而又平静。然而，每当我细细扒开它的纹路，无数闪着金光的记忆扑向我。这些记忆里，有谆谆教诲的恩师，有温暖调皮的室友，有良金美玉的师兄师姐，有青涩可爱的师弟师妹。这些点滴的记忆偷偷地推着我长大，给予我无限的面对挫折的勇气。在此，仅以最朴素的言词向成长路上陪伴和鼓励我的老师同学致以最诚挚的感谢。

饮水思源，不忘师恩。至此论文完成之际，向我的恩师韩金仓教授表达我深深的崇敬与感激。三年学术生活离不开您的谆谆教诲，您深厚的学术修养在科研道路上为我亮起一盏盏明灯，每当我失去自信变得迷茫之时您都能为我指明方向。同时还要感谢何江萍老师在科研与学习方面为我提供的帮助，您兢兢业业的科研态度，豁达大方的为人之道令我倍感温馨。您教给我们的为人处世的道理至今令我受益匪浅。另外，还要感谢丁晓阳老师、李兵老师、李子荣老师及信息工程学院的其他老师在学术上的倾囊相授。在这里我诚挚地向各位老师鞠躬致谢！三年研究生生涯，所获点滴足以让我在诸多方面都有所成长，心存感激。

陪伴是最长情的告白，三年时光，感谢最最亲爱的室友王文瑞、李彤、宋文琴的陪伴与鼓励。三年来我们无数次勾肩搭背的走在去食堂的路上；无数次一起奔跑在上课铃前时光中；无数次抱怨时间太少学业太重；无数次将彼此从迷茫中拉回来……这所有的清楚而又模糊的记忆在这本关于兰财的故事中占据着大量的篇幅，像是一缕缕丝线牵扯不断。每当我抚摸这些篇章，仿佛能感受到和煦的春风、炎炎的盛夏、金黄的秋叶和皑皑的冬日。同时，我也十分感谢在这个兰财能遇到这样可爱的你们，我们彼此相伴，彼此鼓励，在时光中一起成长成最好的样子。

翻开故事的下一篇章，是各位良金美玉师兄师姐、青涩可爱师弟师妹。师兄师姐的身影仿佛一帆旗帜，告诉我正确的航向；他们的话语仿佛一缕清风，送来我渴求已久的知识。师弟师妹则陪伴我走过最酷暑的盛夏，大家一起拼搏努力的样子是安静校园中的一道美丽的风景线。

故事的终章，是一路上支持、鼓励与关怀我的家人，他们永远是我坚强的后盾，激励我克服困难，不断向前。