

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 基于治疗前和治疗后的信息对局部晚期
非小细胞肺癌患者的生存分析研究

研究生姓名: 马亚楠

指导教师姓名、职称: 丁晓阳 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: 2023年6月6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： 马亚楠 签字日期： 2023年5月20日

导师签名： [Signature] 签字日期： 2023年5月20日

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： 马亚楠 签字日期： 2023年5月20日

导师签名： [Signature] 签字日期： 2023年5月20日

A Survival Analysis Study of Patients With Locally Advanced Non-Small Cell Lung Cancer Based on Pre- and Post-Treatment Information

Candidate : Ma Ya'nan

Supervisor: Ding Xiaoyang

摘 要

已经有研究^[1]证明患者接受治疗后肿瘤较高的最大标准化摄取值（Maximum Standardized Uptake Value, SUVmax）与局部晚期非小细胞肺癌患者较差的生存率相关，并且当肿瘤 SUVmax 取 5.0 的二值截取值时，其与患者的生存预后具有显著的相关性（ $p < 0.05$ ）。而在临床实践中，患者在接受治疗后并不会总是再次进行 PET/CT 检查。因此，基于患者接受治疗前的 FDG-PET 图像预测出患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 可能会对预测局部晚期非小细胞肺癌患者的总生存期有潜在的价值。并且由于患者治疗后的数据搜集比较困难，目前对非小细胞肺癌患者的生存分析大部分都是基于患者接受治疗前的一些信息，很少有同时基于患者治疗前和治疗后的信息进行的生存分析。

因此，本文基于患者接受治疗前的 FDG-PET 对患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 进行了预测，并评估了其对局部晚期非小细胞肺癌患者总生存期的影响；并进一步研究了患者接受治疗前和接受治疗后的临床信息对患者生存分析的影响。

基于治疗后 SUVmax 的预测对局部晚期非小细胞肺癌患者的生存分析研究。研究内容包括两个部分，分别是患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 的预测以及生存分析。首先是患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 的预测。其次，采用 Cox 单因素和多因素分析对患者的临床特征进行挑选，将预测出的结果以及临床特征同时输入到 Cox 比例风险模型（Cox Proportional Hazards Model^[2], Cox 模型）中对患者进行生存分析的预测。最终的实验结果表明利用患者接受治疗前的 FDG-PET 图像可以准确预测出患者接受治疗后的肿瘤 SUVmax 是否大于 5.0，并且预测出的结果可以提升患者生存分析的准确性。

基于临床信息对局部晚期非小细胞肺癌患者的生存分析研究。研究内容同样包括两个部分，临床特征的选择以及患者生存分析的预测。利用 Cox 单因素和多因素分析进行临床特征的选择，然后用 Cox 模型对患者生存分析进行预测。最终的实验结果显示，利用患者治疗前和治疗后的临床信息可以准确预测出患者

的生存率。

本文最终的实验结果显示，基于患者接受治疗前的 FDG-PET 预测出的患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 的结果可以显著提升患者生存分析的准确性，是局部晚期非小细胞肺癌患者预后的一个重要因素，对医生的临床决策具有一定的指导意义。基于临床特征对患者进行生存分析的研究可以更加深入地了解患者的病情和治疗效果，为医生制定更加个性化的治疗方案提供参考。这项研究对于提高局部晚期非小细胞肺癌患者的生存率和治疗效果具有重要的临床意义。

关键词：局部晚期非小细胞肺癌 FDG-PET 生存分析 CNN SUVmax 预测

Abstract

It has been demonstrated^[1] that the higher maximum standardized uptake value (SUVmax) of tumors after patient treatment is associated with poorer survival in patients with locally advanced non-small cell lung cancer and that when the tumor SUVmax takes a binary cutoff value of 5.0, it is significantly associated with survival prognosis of patients correlation ($p < 0.05$). In clinical practice, treated patients do not always undergo repeat PET/CT. Therefore, prediction of post-treatment tumor SUVmax greater than 5.0 based on pre-treatment FDG-PET may be potentially valuable in predicting overall survival in patients with locally advanced non-small cell lung cancer. Since it is difficult to collect post-treatment data, most of the current survival analyses of patients with non-small cell lung cancer are based on some pre-treatment information, and few studies have examined analyses based on both pre- and post-treatment information, this paper further investigates the impact of pre- and post-treatment clinical information on the survival analysis of patients.

Therefore, this article predicted whether the tumor post-treatment SUVmax was greater than 5.0 based on the patient's pre-treatment information, and evaluated its impact on the overall survival of locally advanced non-small cell lung cancer patients. Furthermore, the study investigated the impact of pre- and post-treatment clinical information on patient survival analysis.

This study focuses on the survival analysis of locally advanced non-small cell lung cancer patients based on the prediction of post-treatment SUVmax. The study consists of two parts: predicting whether the tumor SUVmax after treatment is greater than 5.0 and

survival analysis. First, a 3D convolutional neural network (3D CNN) was used to extract features from the pre-treatment whole-body FDG-PET to predict whether the tumor SUVmax after treatment is greater than 5.0. Second, Cox proportional hazards model^[2] (Cox Model) was used to conduct survival analysis by selecting clinical features with Cox univariate and multivariate analysis and combining them with the predicted results. The results indicate that using pre-treatment FDG-PET images can predict the tumor SUVmax after treatment relatively accurately and improve the accuracy of survival analysis.

This study focuses on the survival analysis of locally advanced non-small cell lung cancer patients based on clinical information. The study also consists of two parts: selecting clinical features and predicting patient survival. Cox univariate and multivariate analysis were used to select clinical features, and Cox Model was used to predict patient survival. The results indicate that using both pre- and post-treatment clinical information can accurately predict patient survival time, and the selected clinical features reflect the impact of clinical information on patient survival prognosis.

This article predicted the impact of tumor SUVmax after treatment on patient survival analysis and found that the prediction results based on pre-treatment FDG-PET can significantly improve the accuracy of patient survival analysis, which is an important factor in the prognosis of locally advanced non-small cell lung cancer patients and has certain guiding significance for clinical decision-making. The study of patient survival analysis based on pre- and post-treatment clinical features can provide reference for doctors to develop more personalized treatment plans, and deepen the understanding of patients' condition and treatment effectiveness. This research has important clinical significance for

improving the survival rate and treatment effectiveness of locally advanced non-small cell lung cancer patients.

Keywords: Locally Advanced Non-small Cell Lung Cancer; FDG-PET; Survival Analysis; CNN; SUVmax; Prediction

目 录

1 引 言	1
1.1 研究背景与意义	1
1.2 研究现状	2
1.2.1 基于分类的 3D CNN 研究现状	2
1.2.2 生存分析研究现状	5
1.3 本文主要研究内容及创新	9
1.3.1 本文主要研究内容	9
1.3.2 本文主要创新点	9
1.4 本文组织结构	10
1.5 本章小结	11
2 相关知识介绍	12
2.1 局部晚期非小细胞肺癌	12
2.2 PET/CT	13
2.3 基于分类的 CNN	14
2.4 生存分析	16
2.4.1 Kaplan-Meier 方法	17
2.4.2 Log-rank 检验	18
2.4.3 Cox 模型	18
2.5 评价指标	21
2.5.1 分类评价指标	21
2.5.2 生存分析评价指标	22
2.6 本章小结	23
3 基于治疗后 SUVmax 预测对局部晚期非小细胞肺癌患者的生存分析研究	24
3.1 研究方法	24

3.1.1 实验流程	24
3.1.2 实验数据介绍	25
3.1.3 实验数据处理	26
3.1.4 实验验证	30
3.1.5 对患者接受治疗后的肿瘤 SUVmax 预测的分类模型	33
3.1.6 基于预测出的治疗后 SUVmax 的生存分析	35
3.2 实验结果与分析	35
3.2.1 对患者接受治疗后的肿瘤 SUVmax 预测的分类实验结果	35
3.2.2 基于预测出的治疗后 SUVmax 的生存分析实验结果	37
3.2.3 讨论	37
3.3 本章小结	38
4 基于临床信息对局部晚期非小细胞肺癌患者的生存分析研究	40
4.1 研究方法	40
4.1.1 实验流程	40
4.1.2 实验数据介绍	41
4.1.3 实验数据处理	41
4.1.4 Cox 单因素和多因素分析的临床特征挑选	41
4.1.5 生存分析模型的训练过程	41
4.2 实验结果与分析	42
4.2.1 Cox 单因素和多因素分析的临床特征挑选结果	42
4.2.2 生存分析实验结果	45
4.2.3 讨论	45
4.3 本章小结	45
5 总结与展望	47
5.1 总结	47
5.2 展望	48
参考文献	50

致 谢..... 58

1 引言

1.1 研究背景与意义

癌症是全球主要的公共卫生问题之一，也是死亡的主要原因之一^[7]，其中，肺癌是世界上最常见的癌症之一。肺癌大致可以分为两类：小细胞肺癌、非小细胞肺癌。其中，小细胞肺癌患者大约占肺癌患者的 15%，其余的 85% 为非小细胞肺癌患者。据研究显示，确诊为非小细胞肺癌的患者中有超过大约三分之一（33-50%）的患者表现为局部晚期疾病。局部晚期非小细胞肺癌通常指的是 III 期的非小细胞肺癌^[8]。III 期非小细胞肺癌是一组异质性疾病，包括大肿块、纵隔淋巴结受累和主要纵隔结构的侵犯。

随着治疗技术的进步，虽然无法手术的 III 期非小细胞肺癌的预后得到了进一步的改善，但是患者的预后仍然很差^[9]。同步放化疗的一个挑战是难以在治疗后快速确定疾病状态^[1]。18-氟脱氧葡萄糖（^[18F]FDG）PET 结合计算机断层扫描（CT）作为 PET/CT 联合检查已成为接受放疗的非小细胞肺癌患者分期和检查的主要内容^[10]。PET 对非小细胞肺癌的分期是有效的，与传统的分期相比，对局部或远端转移性疾病具有更高的敏感性和特异性^[11]。SUV 是使用 18-氟脱氧葡萄糖（^[18F]FDG）放射线在 PET 上测量的，是肺癌管理中用于疾病检测、分期和评估治疗反应的一种标准成像方式^[12]。

在对局部晚期非小细胞肺癌患者进行生存分析的研究中，本文发现在治疗前 FDG-PET 测量得到的肿瘤 SUV_{max} 对患者的生存预后没有影响。van Elmpt W^[13] 等人指出，单独使用治疗前 FDG-PET 获得的肿瘤指标对患者的总生存期没有预后价值。此外，Machtay M^[1] 等人明确指出，患者接受治疗前的肿瘤 SUV_{max} 与生存率之间没有关系。

另一方面，已经证实，在放化疗后相对快速获得的^[18F]FDG-PET 可以预测长期预后^[1]。治疗后肿瘤 SUV_{max} 的价值可能表明了患者对治疗^[14]的反应。治疗后较高的肿瘤 SUV_{max} 与 III 期非小细胞肺癌患者较差的生存率相关。并在探索

分析中发现, 当对局部晚期非小细胞肺癌患者治疗后的肿瘤 SUVmax 使用 5.0 的二值截取值^[1] ($p < 0.05$) 时, 它与生存率显著相关。

但是, 目前治疗后肿瘤 SUVmax 在使用 5.0 二值截取值时对生存分析的影响尚不清楚。并且, 由于 PET/CT 检查非常昂贵, 只有少数患者会在治疗后再次进行 PET/CT 筛查。因此, 本文的研究重点是根据患者治疗前的 18F-FDG PET/CT 预测局部晚期非小细胞肺癌患者接受治疗后的肿瘤 SUVmax 是否大于 5.0, 并利用预测出的结果对其进行生存分析研究, 确定患者接受治疗后的肿瘤 SUVmax 使用 5.0 的二值截取值时对其总生存期的影响。并进一步研究了患者接受治疗前和接受治疗后的临床信息对患者生存分析的影响。

1.2 研究现状

1.2.1 基于分类的 3D CNN 研究现状

由于疾病检测是根据医学图像进行分析的, 而每次扫描所得到的医学图像包含数百万像素, 研究人员和临床医生了解这种扫描需要很长时间, 因此在临床上计算机技术被应用于患者疾病的诊断^[15]。随着科学技术的发展, 机器学习和其他深度学习的方法已经被用来提取医学图像中的有用信息来进行医学图像的分类^[16]。计算机辅助诊断 (CAD) 为放射科医生提供了一种无创的解决方案和一种客观的预测, 以协助医生进行疾病的诊断。现有的 CAD 方法可分为两类: 基于手工制作的特征的分类模型和基于自动特征提取的深度神经网络^[17]。第一类方法通常用来测量放射学特征, 例如, 病灶的大小、位置、形状、纹理, 并采用分类器来确定病灶的状态。但是, 收集和选择一个有用的特征子集来诊断病灶并不容易, 而且存在引入测量误差的风险。在第二类方法中, 基于深度 CNN 的模型可以自动从医学图像中学习到诊断病灶的特征。

CNN 通常被用来检测医学图像中的疾病状况。由于 CNN 具有数据驱动的特性, 因此它能够理解不同类别之间的细微差异, 而不是传统的基于规则的特征技术, 如小波和主成分分析 (PCA)^[18]。因此, 基于分类的 CNN 将对医学数据集

提供有效的帮助。CNN 已被广泛用于研究医学图像，使用 3D CNN 的优点是能够将整个三维医学图像序列作为一个整体进行分析，而传统的 2D CNN 则是一次处理一张图像。因此，3D CNN 可以从相邻的切片中提取到三维特征，但是，这些特征无法从单一的二维图像中检测出来。

Guida 等人^[19]利用 1100 名骨关节炎患者的膝关节核磁共振图像（Magnetic Resonance Imaging, MRI）以及 X 射线数据研究了 3D CNN 在膝关节骨关节炎诊断中的应用，分别对膝关节骨关节炎的严重程度（0-4 级）以及是否患有骨关节炎（严重程度 ≤ 1 和严重程度 ≥ 2 ）进行了分类。分类结果显示，与使用来自同一组病人的 X 射线图像训练的 CNN 模型相比，所提出的带有 MRI 的三维模型在五类分类（54.0%和 50.0%）和二类分类（83.0%和 77.0%）中都取得了更高的准确性。结果表明，与目前使用的 X 射线方法相比，应用 3D CNN 模型的 MRI 在提高膝关节骨关节炎的临床诊断准确性方面具有更大的潜力。

Suryakanth 等人^[20]利用多模态医学成像数据集使用 3D CNN ResNet 进行三维图像的分类。本文首先将数据集中的二维图像转化为三维图像数据，因为三维图像可以提供比二维图像数据更多的信息。然后，将三维 CT 和 MRI 进行融合，融合后的图像经过引导滤波过滤，以便对其进行进一步处理。之后对融合后的图像进行增强。最后，将融合后的图像放进 3D CNN ResNet 进行分类。最终本文所提供的方法达到了 98%的准确率。作者所设计的模式以一种有效的方式预测了疾病的分期。也说明了 3D CNN 在医学图像分类中的有效性。

Dey 等人^[19]利用 1010 例肺结节患者的 CT 数据，基于放射科医生所提供的注释从 CT 中获取对应的三维图像块作为网络的输入对肺结节进行分类。基于此，作者提出了四个双通道的 CNN，包括一个基本的 3D CNN、一个新型的多输出网络、一个 3D DenseNet 网络和一个具有多输出的 3D 增强 DenseNet 网络。四个网络在该数据集上的分类性能超过了大多数现有的方法。三维多输出 DenseNet 网络在端到端的肺结节诊断任务上达到了最优的分类精度（0.904），也进一步验证了 3D CNN 在医学图像分类上的优势。

Kruthika 等人^[21]使用来自阿尔茨海默病神经成像倡议（ADNI）的 MRI 数据对阿尔茨海默病（AD）进行了分类。作者提出了一种基于内容的图像检索系统

(CBIR)，该系统使用 3D CapsNets、3D CNN 和预训练的三维自编码器技术来进行早期的阿尔茨海默氏症的检测。3D CapsNets 可以在小数据集上快速学习，并且能够处理图像旋转和转换。使用 3D CapsNets 开发的 CBIR，可以准确的检测出早期的阿尔茨海默氏症，准确率高达 98.42%。研究表明，与现有的基于二维的方法相比，他们所提出的方法有更好的效果。

在参考文献^[22]中，Ker 等人使用的是来自新加坡国家神经科学研究所的 399 名患者的三维脑 CT 图像，大约包含了 12,000 张图像。作者开发了一种三层的浅层 3D CNN 进行脑出血的 2 分类（正常与特定的异常类）和 4 分类（正常、含有蛛网膜下腔出血（SAH）、沟内出血（IPH）、急性硬膜下出血（ASDH））。与 3D VGG 网络和 3D GoogLe 网络相比，该网络在较短的训练时间内获得了最好的结果。二分类的 F1-score 分别为 SAH: 0.927、IPH: 0.932、ASDH: 0.952、脑部多发性创伤出血（BPH）: 0.919。多分类的 F1-score 为: 0.938。与 3D VGG 网络和 3D GoogLe 网络相比，该网络可以在较短的训练时间内获得较好的结果。

Shaish 等人^[23]收集了 2017 年病理证实为恶性肿瘤的患者进行的 PET/CT 扫描，两名放射科医生每次从 PET/CT 检查的未增强 CT 部分中选择 1 到 10 个淋巴结对其进行勾画，并记录了淋巴结的 SUVmax。作者将淋巴结的 CT 块以及原发肿瘤的组织学类型作为 3D CNN 的输入，用来预测淋巴结的 SUVmax 是否大于 2.5。研究结果表明，CNN 能够准确地从未增强的 CT 图像和患者的肿瘤组织学亚型中预测淋巴结的 SUVmax。

Kang 等人^[24]对肺部图像数据库联盟和图像数据库资源倡议数据库（LIDC-IDRI）所提供的 CT 图像进行了二分类（良性和恶性）和三分类（良性、原发性恶性和转移性恶性）。因为 3D CNN 能够充分利用肺部结节的三维背景信息，并且，多视图策略已被证明有助于提高 2D CNN 在肺部结节分类中的性能。因此，作者使用链式结构和有向无环图结构的三维多视图 CNN（MV-CNN），包括 3D Inception 和 3D Inception-ResNet，来探索肺结节的分类。所有研究结果都是通过十折交叉验证得到的。结果显示，3D MV-CNN 的性能大大超过了 2D MV-CNN。说明了 3D 空间信息对分类任务的重要性。

由于目前还没有可以用来自动区分有无 AD 病理的语言缺陷患者，基于此，

Marco 等人^[25]建立了一个 CNN，基于 3D 脑 MRI 图像对 AD 病理与非 AD 病理的语言缺陷患者进行分类。并添加了一个迁移学习技术对患者的 3D 大脑 MRI 图像进行去噪，目的是用来提升模型分类的准确性。最后的结果显示，3D CNN 模型的准确性达到了 0.62 的准确率，在采用迁移学习技术对患者的图像进行去噪之后，模型的准确率达到 0.74。

以上的研究均表明了 3D CNN 可以对医学图像进行纹理分析，并从中提取到与分类任务相关的一些特征。一些研究证明，与 2D CNN 相比，3D CNN 可以提取到二维图像无法提供的三维空间信息，甚至在一些分类任务中，3D CNN 可以获得更好的分类效果。目前已有的关于医学图像分类的任务中，3D CNN 的输入通常需要有医生勾画的肿瘤轮廓或者肿瘤的大致位置。但是，这种方法需要医生手动对肿瘤进行标注，费时费力。因此，可以考虑采用阅片的方式，采用 3D CNN 通过对患者的医学影像进行特征提取，从而对其进行分类。并且，在已有的分类任务中，已经有研究成功的从患者的影像数据中准确预测出病灶的 SUVmax 是否大于 2.5，说明了患者的医学图像中提供了关于病灶 SUVmax 的相关信息。

1.2.2 生存分析研究现状

生存分析主要是用来分析感兴趣事件发生时间的方法。生存分析通常用于生物统计学，主要用来确定患者在接受治疗后的存活机会，例如，化疗后的癌症患者。患者的生存数据是被长期记录的，会记录患者的状态，如死亡、治愈或退出研究。此外，生存分析也可以应用于其他领域，分析顾客流失、囚犯被释放后再犯的可能性。生存分析中使用的一种方法是 Cox 模型，它可以将感兴趣的事件随着时间发生的风险进行量化。在整个分析过程中，生存将被认为是感兴趣的事件没有发生时的情况。此外，该模型可以用来分析一些特征对生存的影响，并可以确定一定时间后结果发生的可能性，以预测事件是否会发生。另一方面，Cox 模型还可以用来进行特征的研究，以确定该特征与事件是否存在较大的相关性^[26]。它已成为处理删失数据的标准方法，并已广泛应用于临床试验和其他生物医学研究。这种模式激发了许多相关工作。

Sisay 等人^[27]研究的数据来自埃塞俄比亚南部 2020 年 10 月 1 日至 2021 年 9 月 30 日确诊冠状病毒（COVID-19）的患者，共 413 名，其中有 194 例为严重病例（46.9%），有 77 例（39.6%）死亡。主要研究的是确诊 COVID-19 病例中发生重症 COVID-19 的时间和风险因素。作者采用 Kaplan-Meier 曲线图和 log-rank 检验来确定类别变量对 COVID-19 严重程度的影响。采用双变量和多变量 Cox 回归模型确定重症 COVID-19 的风险因素。最终确定重症 COVID-19 的风险因素的特征包括：年龄（>65）、咳嗽、胸痛、头痛、伴随性疾病、哮喘、入院时症状发作超过 5 天。最终通过研究得出结论：确诊 COVID-19 患者中有症状和有伴随性疾病的患者应该被密切监测。

Tian 等人^[28]研究的是关于头颈部癌症——下咽鳞状细胞癌（HSCC）的患者预后。作者采用的是 2010 年至 2016 年间被诊断为 HSCC 的 2198 名患者的临床数据，这些数据是从 Surveillance, Epidemiology, and End Results（SEER）数据库中检索得到的。文中利用 Cox 模型来评估每个变量与总生存期之间的关系。并进行了 Cox 多变量回归分析，将分析的结果用来制作列线图。预后列线图的准确性用校准曲线和 C-index 来评估。Cox 多变量回归分析显示，年龄、婚姻状况、种族、T 期、N 期、M 期、TNM 期、局部治疗和化疗都与患者总生存期相关。校准曲线显示预测的总生存期和实际总生存期之间具有很高的一致性。说明，列线图对 HSCC 患者的生存结果提供了相对准确和适用的预测。

Feroz 等人^[29]采用的是 2019 年至 2021 年在印度祈祷国卡马拉·尼赫鲁纪念癌症医院收集到的 176 名临床诊断为胆囊癌（GBC）的患者。基于所收集到的数据，作者评估了 GBC 患者的生存状况，并探讨其改善和预防的预后因素。作者采用 Kaplan-Meier 法分析了患者的生存率；生存率差异以及预后因素的分析均采用 log-rank 检验；采用 Cox 模型估计死亡率结果的风险比。结果显示患者的总中位生存时间为 5 个月，1 年、2 年和 3 年生存率分别为 24.4%、8.5%和 4.5%。在接受手术和放疗后，患者 3 年生存率分别提高到 19.5%和 35%。多变量分析结果显示，患者所在城市地区（HR = 1.568, p = 0.040）、胆结石（HR = 1.571, p = 0.049）、N 期（HR = 1.468, p = 0.029）和 M 期（HR = 2.289, p < 0.0001）是预后的独立风险因素，而手术（HR = 0.573, p = 0.030）是 GBC 预后的保护因素。

Molina-Fernández 等人^[30]的研究主要是为了根据疾病的阶段和治疗方法来评估喉癌（LC）患者的生存率，研究包括临床诊断为原发性 LC 的成年患者，诊断时患有 LC 和癌症转移的患者以及患有其他类型肿瘤的患者被排除在外。作者采用单变量和多变量分析来确定接受 LC 治疗和事件（死亡）发生的时间之间的关联。计算了总生存率、病因特异性生存率（CSS）和无病生存率（DFS）。结果显示，晚期肿瘤患者（III 期和 IV 期）的 LC 死亡风险几乎是初始肿瘤阶段（I 期和 II 期）的三倍；接受手术治疗的患者比按照器官保留方案（OPP）治疗的患者有更高的生存机会。最后得出结论，相对于单纯的放疗，手术治疗能改善初始 LC 患者五年后的 CSS 和 DFS。此外，在晚期 LC 患者中，手术治疗与补充性放疗相关，可提供更好的 CSS 和 DFS。

由于目前还没有针对泌尿生殖系统（GU）小细胞癌（SCC）的大型前瞻性研究或临床试验，大多数数据都是从肺部 SCC 推断出来的。Baralo 等人^[31]利用 SEER 数据库，使用 log-rank 检验分析了其发病率趋势、总生存率和癌症特异性生存率。作者使用 Cox 模型对变量进行了分析，分析结果显示，膀胱和前列腺的 SCC 是 GU SCC 最常见的类型，分别有 1836 和 606 例。在 2018 年，膀胱和前列腺 SCC 的发病率是 2010 年的两倍（ $P < 0.001$ ）。膀胱 SCC 患者的总生存期和癌症特异性生存期明显长于前列腺 SCC 患者（ $P < 0.0001$ ）。高龄、生长范围较广、淋巴结受累、无手术干预和存在转移灶的 SCC 膀胱患者的生存结果更差（ $P < 0.05$ ）。对于前列腺 SCC 患者，只有高龄是影响患者不良预后的风险因素（ $P < 0.05$ ）。

Vargas 等人^[32]主要研究了在同一机构治疗的高级别胶质瘤（HGG）患者在 10 年内的生存情况，主要分析了 HGG 相关因素的重要性以及在病情进展时进行挽救性手术的作用。研究所纳入的数据是 2008 年至 2017 年间在 Gregorio Marañón 医院（西班牙马德里）治疗的一系列世界卫生组织（WHO）III/IV 级胶质瘤患者，包含患者的临床、放射学和解剖学病理数据，最终纳入研究的患者有 133 名。文中的生存分析采用 Kaplan-Meier 曲线。每个变量对生存率的影响采用 log-rank 检验计算。采用 Cox 回归计算显示显著影响的变量的风险分层。最终的结果显示，大约三分之一的 HGG 患者在病情进展时可能有资格接受挽救手术。在这一分组

的患者中，抢救性手术与患者较长的生存期明显相关。

子宫内膜癌是印度第四种常见的妇科癌症，Kumar 等人^[33]评估了印度一家三级护理中心的子宫内膜癌复发模式以及患者的生存预后，同时也对影响子宫内膜癌复发的危险因素进行了分析。对 2014 年 1 月至 2021 年 11 月期间在三级护理中心接受治疗的子宫内膜癌患者进行了回顾性分析（ $n = 115$ ），主要采用的是 Kaplan-Meier 方法。研究结果显示，患者的中位总生存期为 47.7 个月，而中位无病生存期（DFS）为 50 个月。年龄超过 60 岁（ $p = 0.0092$ ），淋巴血管空间侵犯（LVSI）（ $p = 0.0001$ ），结节受累（ $p = 0.018$ ），FIGO 分期较高（ $p = 0.031$ ）和组织学等级较高（ $p = 0.00014$ ）是与复发明显相关的主要风险因素。年龄超过 60 岁、组织学等级较高、II 型组织学、存在 LVSI 和淋巴结受累是复发的危险因素。

Chernova 等人^[34]对 2015 年至 2019 年在新西伯利亚地区肿瘤医院接受诊断的 3007 名鳞状细胞肺癌和 3049 名肺腺癌患者的医疗档案进行了分析。该研究包括 4758 名男性和 1298 名女性。最终的结果显示，大多数（96%）非小细胞肺癌患者的年龄都在 50 岁以上。鳞状细胞肺癌和肺腺癌患者的 5 年生存率低于 20%。鳞状细胞肺癌患者的中位生存期为 443 天。在诊断为非小细胞肺癌时，67% 的患者为疾病的 III/IV 期。在 50 岁到 61 岁的非小细胞肺癌患者中，被诊断为疾病 III 期的患者生存期最长（10 至 15 年）。在对新西伯利亚地区鳞状细胞肺癌和肺腺癌患者生存情况的分析表明，疾病阶段和年龄对中位生存期有明显影响。这些都是非小细胞肺癌一般发病率和死亡率的特征。

以上研究所采用的生存分析方法包括：Kaplan-Meier 分析、Log-rank 检验、Cox 模型。虽然目前所存在的生存分析的方法除了以上的三种传统的方法外，还有基于深度学习的方法可以从医学图像中提取相关特征来对患者进行生存预测，但是，从关于生存分析的研究现状来看，传统方法仍然是做生存分析研究中常用的一种方法。与基于深度学习的生存分析方法比较，传统方法不仅可以预测患者的生存率，还能够从众多因素中判断出影响患者生存预后的因素。因此，传统方法仍然是做生存分析的主要选择。

1.3 本文主要研究内容及创新

1.3.1 本文主要研究内容

本文的主要研究内容包括两个部分，第一部分是基于治疗后的肿瘤 SUVmax 的预测对局部晚期非小细胞肺癌患者的生存分析研究。其主要的研究内容为：首先是在已有的局部晚期非小细胞肺癌患者数据集上进行了实验验证，证明了从治疗后 FDG-PET 上获取的肿瘤 SUVmax 在取 5.0 的二值截取值时，肿瘤 SUVmax 与生存率有显著的相关性，之后利用 Cox 单因素和多因素分析进一步证明了患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 是影响生存分析的有效因素。其次，利用 3D CNN 基于患者治疗前的全身 FDG-PET 对患者接受治疗后的 SUVmax 是否大于 5.0 进行了分类。最后将临床数据与预测出的治疗后 SUVmax 进行结合，采用 Cox 模型对患者进行生存分析。最终的结果显示，基于患者接受治疗前的 FDG-PET 对患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 的预测结果可以提升患者生存分析的准确性，说明基于治疗前 FDG-PET 预测出的治疗后 SUVmax 是否大于 5.0 是局部晚期非小细胞肺癌患者预后的一个重要因素，并且对医生的临床决策有一定的帮助。

第二部分是基于患者治疗前和治疗后的临床信息对局部晚期非小细胞肺癌患者的生存分析，首先是采用 Cox 单因素和多因素分析对患者的临床信息进行了挑选，其次采用 Cox 模型对患者进行生存预测，最终的结果显示，在测试集中，患者生存分析的准确性 C-index 达到了 0.70 (0.04)，从每一折挑选出患者的临床特征来看，治疗后的信息也对患者的生存分析产生了必要影响。基于治疗前和治疗后的临床信息对患者生存分析的研究可以更详细的了解患者的病情和治疗效果，为医生制定后续的治疗方案提供参考。

1.3.2 本文主要创新点

本文的创新点主要包括以下几个方面：

- 1、在对患者进行接受治疗后的肿瘤 SUVmax 是否大于 5.0 的预测时，能够

在没有患者接受治疗后信息的基础上，利用患者接受治疗前的全身 FDG-PET 影像数据，采用阅片的方式，准确预测出患者在接受治疗后肿瘤 SUVmax 是否大于 5.0。

2、本文在进行患者治疗后肿瘤 SUVmax 是否大于 5.0 的预测时，能够根据患者接受治疗前的 FDG-PET 准确预测出没有接受治疗后 FDG-PET 检查的患者在接受治疗后的肿瘤 SUVmax 是否大于 5.0，并且预测出的结果是局部晚期非小细胞肺癌患者预后的一个重要因素，能够在原有的基础上进一步提升患者生存分析的准确性。

3、在基于患者临床信息对局部晚期非小细胞肺癌患者进行生存分析时，不仅利用到了患者接受治疗前的临床信息，还应用了患者接受治疗后的临床特征对患者进行了生存分析。

1.4 本文组织结构

本文首先对研究内容的研究背景、研究意义以及研究现状进行了介绍，其次详细介绍了本文的两个工作，并在最后进行了总结与展望，本文的具体组织结构如下：

第一章是绪论部分。本章主要对本文两个工作的研究背景以及研究意义进行了介绍，然后对基于分类的 3D CNN 的研究现状以及生存分析的研究现状进行了简要阐述，之后介绍了本文的主要研究内容、创新点以及组织结构。

第二章主要是对文章中涉及到的一些概念和理论知识进行了介绍，其中包括局部晚期非小细胞肺癌的介绍、PET/CT 的介绍、基于分类的 CNN 的详细描述、生存分析的介绍、在进行生存分析实验时所用到的方法也进行了详细描述以及本文研究所涉及到的一些模型的评价指标的详细介绍。

第三章主要介绍的是本文的第一个工作：基于治疗后 SUVmax 预测对局部晚期非小细胞肺癌患者的生存分析研究。介绍了第一个工作的实验流程以及相关实验。首先对实验数据来源、实验所用到的患者影像数据以及临床数据的选择、患者临床特征的挑选数据预处理部分进行了详细介绍。然后对本研究的可行性进

行了实验验证，主要描述了实验验证的流程以及实验验证的结果。其次对本研究的实验流程和方法进行了详细介绍，包括实验所用到的模型，模型训练时的参数设置以及最终的实验结果与分析。

第四章主要介绍了本文的第二个工作：基于患者治疗前和治疗后的临床信息对局部晚期非小细胞肺癌患者的生存分析。首先是对本研究要用到的患者的临床数据进行了处理，然后对本研究在进行特征挑选以及生存分析的方法进行了描述，并且罗列出了五折交叉验证中每一折所挑选出来的患者的临床特征。最后就是对生存分析结果的描述与分析。

第五章是总结与展望部分，主要对本文所研究的两个内容进行了总结，以及对本文工作存在的不足之处进行了说明，有望在未来的工作中进行进一步的修正。

1.5 本章小结

本章主要介绍了本文的研究背景以及研究意义，并且对本文所涉及到的基于分类的 3D CNN 以及生存分析的研究现状进行了简要介绍，并且详细阐述了本文的主要研究内容以及创新点。

2 相关知识介绍

本文的主要内容是局部晚期非小细胞肺癌患者的生存分析研究，采用的方法包括 3D CNN 的分类模型以及 Cox 模型。因此，本章主要对本研究所涉及到的一些相关知识做了详细介绍。主要包括：局部晚期非小细胞肺癌、PET/CT、基于分类的 3D CNN、生存分析模型以及相关评价指标。

2.1 局部晚期非小细胞肺癌

癌症是世界范围内的一个主要公共卫生问题^{[7][35]}，并且是造成死亡的主要原因之一^{[7][35][36]}。据 2022 年全球癌症统计显示，全球有超过 1930 万例的癌症病例，大约有 1000 万人死于癌症^[37]。常见的癌症包括：乳腺癌、肺癌、淋巴癌、前列腺癌和脑癌等。其中乳腺癌是世界各地女性中最普遍的癌症类型，肺癌和前列腺癌是男性中最普遍的癌症类型。

原发性支气管肺癌（简称：肺癌）是常见的恶性肿瘤之一，超过一半的肺癌患者在被确诊后一年内死亡^[38]，是癌症死亡的主要原因^{[7][35]}。肺癌是男性和女性中的第二大主要癌症^[38]，也是第二种最常见的癌症。肺癌是在男性中约占 20%、在女性中约占 11%的恶性肿瘤。随着科技的进步以及医疗水平的提高，虽然在男性中肺癌的发病率逐渐下降，但是女性肺癌的发病率却快速提升^[39]。肺癌可以大致分为小细胞肺癌和非小细胞肺癌。其中，非小细胞肺癌患者约占所有肺癌病例的 85%^[40]。

局部晚期非小细胞肺癌通常是指 III 期的非小细胞肺癌^[8]。约 27% 的非小细胞肺癌患者在诊断时患有局部晚期疾病^[41]。尽管所有这些患者都被归类为肿瘤-结节-转移（TNM）分期系统的第三期^[42]，但这是一个相当不一样的疾病群体，包括大型肿瘤、侵入纵隔和胸壁的肿瘤，以及纵隔淋巴结转移伴有小的甚至无法检测的原发病灶的患者^[43]。由于临床表现的多样性，目前人们普遍认为 III 期非小细胞肺癌需要一种由胸外科、医学肿瘤学和放射治疗专家组成的多学科团队的多模式治疗方法^[44]。需要局部和全身治疗相结合，来实现疾病的彻底控制，以减少

局部和远处复发的风险。局部晚期非小细胞肺癌患者由于各种因素被认为其肿瘤无法切除，其中包括结节过大、多发性淋巴结、广泛的纵隔、气道或其他器官侵犯，和（或）原有的医疗合并症^[45]。

不可切除的局部晚期非小细胞肺癌是一组具有不同临床表现和预后的异质性疾病，如果用病理学特征定义，它主要包括 IIB 至 IIIC 期^[46]。同步化疗（cCRT）是全球无法手术患者的一种标准的治疗方法^[47]。基于此，本文所研究的患者主要包含的是接受 cCRT 的患者，从而评估无法手术的局部晚期非小细胞肺癌患者在接受 cCRT 后的生存预后。

2.2 PET/CT

正电子发射断层扫描（Positron emission tomography, PET）被越来越多地用于各种恶性肿瘤的诊断、分期和随访。它已经被研究用于评估各种肿瘤，包括孤立性肺结节、肺癌、淋巴瘤、黑色素瘤、乳腺癌和结肠直肠癌等。PET 能够在尚未显示出基于形态学标准的异常外观的器官中，在分子水平上显示出异常的代谢活动，并有助于区分恶性和良性病变以及对恶性肿瘤进行分期^[48]。与 PET 不同的是，计算机断层扫描（Computed tomography, CT）是依靠解剖学变化来诊断、分期和跟踪癌症^[46]。

氟脱氧葡萄糖（2-[F-18]fluoro-2-deoxy-d-glucose, 18F-FDG）PET/CT 是一种混合的功能和解剖成像方法，在肿瘤学中有大量的应用。18F-FDG PET/CT 提供的是一种无创的代谢和解剖成像方式。作为放射性同位素的氟化物-18，它的半衰期时间较短，因此可以在有限的剂量下成像。氟化物-18 被用来以化学方式取代葡萄糖上的一个羟基，由此产生的类似于葡萄糖的 FDG 被细胞吸收^[49]。

恶性肿瘤细胞的特征是增殖速度快、体积会变大、同时会伴有局部侵袭和远端转移^[50]，进而肿瘤细胞通常可以比正常组织吸收更多的葡萄糖，与葡萄糖不同的是，FDG 不会进一步的代谢，被困在代谢活跃的细胞中，因此 FDG 可以对肿瘤进行有效的成像。据统计，有超过 95% 的 PET 扫描是以 FDG 作为放射性载体进行的^[50]。在临床实践中，18F-FDG PET/CT 常用于多种癌症的评估和管理，包

括肿瘤诊断、分期、再分期、治疗监测和放疗计划。因此，利用 PET 进行的肿瘤成像技术最近在寻找新的、有效的治疗方法和理解分子生物标志物在癌症治疗过程中的作用方面变得尤为重要^[51]。使用 18F-FDG PET 在提供重要的肿瘤相关定性和定量代谢信息方面具有重要价值，对诊断和随访非常重要^[49]。并且，治疗前的 18F-FDG PET 已被深入研究，并在多种肿瘤疾病中显示出良好的预后价值^{[52][53]}。SUV 是在 PET 上测量的，因此，本文采用的是 FDG-PET 图像来预测患者的 SUV_{max}。

2.3 基于分类的 CNN

在深度学习中，CNN 是一种深度神经网络，它主要用于视觉图像，是人工神经网络（Artificial Neural Networks, ANN）的一个特殊架构，由 Yann LeCun 在 1998 年提出。CNN 的一个流行应用是图像分类。

CNN 由两部分组成：第一部分由卷积层和池化层组成，主要的作用是进行图像特征的提取。第二部分由全连接层和密集层组成，主要是对提取的特征进行若干非线性变换，并对图像进行分类。

对一张自然图像来说，人类肉眼看到的是一张彩色图像，而当把这张彩色图像输入计算机时，计算机接收到的图像是一个像素矩阵。在常见的 RGB 彩色模式下，如果一幅彩色图像的大小为 512×512 ，那么计算机接收到的像素矩阵的大小将会是 $512 \times 512 \times 3$ 。这里，512 代表矩阵的高度和宽度，3 代表 RGB（红、绿、蓝）三个颜色通道，其中每个颜色通道由一个单独的矩阵表示。本文中所用到的医学图像属于单通道的灰度图像，计算机读取到的是一个通道数为 1 的像素矩阵。

图像分类一般被定义为一项任务。在该任务中，图像将会被作为输入，然后使用特定的算法建立一个模型，最终输出图像类别的概率。在这个过程中，图像将会被标记为一个特定的类别。给定一张猫的图像，对人类来说，它的一些基本特征是耳朵、鼻子和尾巴等，但计算机是从像素层面上来理解这幅图像，所以计算机所识别出来的基础特征是它的轮廓和边界等因素。通过使用几个不同的卷积

层和池化层，计算机就能顺利地从图像中提取到一些像素层方面的特征。

CNN 模型中一般包括输入层、卷积层、池化层、全连接层、输出层、激活函数。其中输入层就是将输入的图像输入到 CNN 模型中，输入网络的图像大小可以根据实际需求进行调整；卷积层是由一个固定大小的滤波器组成，也是最为重要的一层，卷积的数学运算是在输入图像和滤波器之间进行的。这是一个从图像中提取大部分基本特征的阶段，如尖锐的边缘和曲线，因此该层也被称为特征提取层；之后进行池化操作，池化也被称为降采样，是为了缩小图像的空间体积；全连接层（Fully Connected Layer, FC）被放在 CNN 模型的最终分类输出之前。这些层用于在分类前对结果进行平移，涉及几个偏置、权重和神经元。在分类前附加一个 FC 层会产生一个 N 维向量，其中 N 是目标任务的类别种类；最后，输出层由类别标签组成，它主要是通过使用 one-hot 编码方式对类别进行编码。激活函数被用来决定神经网络的输出，是 CNN 模型的核心。它决定了一个特定的神经元是否应该被激活，是对输入信号进行的非线性变换。然后，这个转换后的输出被作为输入传递到下一层的神经元。常用的激活函数包括：Sigmoid、ReLU、Softmax 等。基于分类的 CNN 模型的基本架构如图 2.1 所示。具体来说就是，首先将需要处理的数据经过输入层输入到 CNN 中，然后经过卷积层和池化层对输入网络的数据进行特征提取，之后经过全连接层和输出层对数据进行分类。

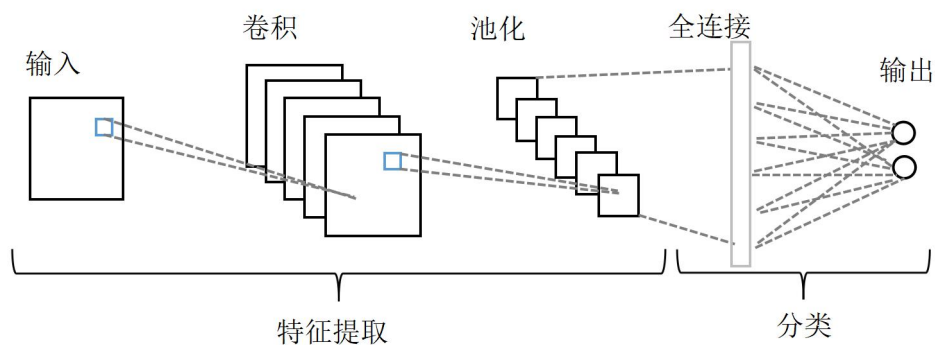


图 2.1 基于分类的 CNN 模型的基本框架

2.4 生存分析

生存分析是用于分析事件发生时间的一种方法。生存分析的生存时间通常指的是事件失败的时间、患者的生存时间或事件发生的时间。例如，肿瘤复发时间、在经过一些治疗干预后，直到患者死亡的时间、直到机器部件出现故障的时间等等。在进行分析时，事件如果均发生，则可以使用标准的回归分析方法。但是，事件发生的时间被限制为带有偏态分布的正分布；并且生存超过某一时间点的概率可能比事件发生的预期时间更有意义；并且在生存分析中用于回归的风险函数可以比线性回归更深入地了解其失效机制。

生存分析的生存时间包括以下特点：1、通常是连续的；2、可能某些事件的时间还不完全确定（例：对某些实验对象不清楚他们确切的生存时间，但是可能知道他们的生存时间至少等于一段时间）；3、到观察时间结束，事件还未发生，该事件就为截尾事件；4、生存时间总是大于等于 0 的。

当有一些关于受试者的事件发生的时间信息时，删失数据就会存在。但不知道事件发生的确切时间。删失可能发生的原因一般有三个：第一，受试者在研究结束前不会经历该事件；第二，受试者在研究期间失去随访；第三，受试者在中途退出了这项研究。由以上原因所产生的数据，均属于右删失数据。右删失数据包括以下几种情况：1、当研究设计在结束 C 年随访后发生。在这种情况下，每个在研究过程中没有观察到事件的受试者都在 C 年被删失；2、该研究设计在 C 年之后结束，但被删失的受试者并非都有相同的删失时间；3、当该研究有预先指定的事件数量时就结束研究。

生存率是肿瘤学研究中一个关键的衡量指标，它被视为事件发生的时间。最常见的结果是总生存期、无事件生存期、无病生存期、无复发生存期和无进展生存期^[54]。生存分析常用的方法包括 Kaplan-Meier 方法^[55]、对数秩检验（Log-rank 检验）以及 Cox 模型。Kaplan-Meier 方法是在 1958 年英国科学家 Kaplan 和 Meier 提出的，是基于概率乘法定理根据生存时间分布来估计生存率以及中位生存时间，适用于未分组的大样本或者小样本数据生存率的 Kaplan-Meier 分析。Log-rank 检验是用来比较多组生存曲线的一种非参数方法，一般是根据生存率及其标准差

来计算各组生存曲线之间的差异性。Cox 模型是一种允许删失数据存在的一种分析多个因素对生存时间影响的多因素分析方法。

2.4.1 Kaplan-Meier 方法

Kaplan-Meier 方法是一种基于概率乘法定理根据生存时间分布来估计生存率以及中位生存时间的方法，是由 Kaplan 和 Meier 提出的一种非参数估计生存率的最常用的方法。具体来说就是：可以从一组样本中的生存时间估计出在相同情况下在特定时间内存活的人口比例。其结果通常用生存曲线图表示，其中 X 轴表示随访时间，Y 轴表示存活人口的比例。

设 T 表示事件失败时间的随机变量，则 $S(t) = P(T \geq t)$ 为生存函数。则生存曲线的概率估计可以用以下公式 (2.1) 表示：

$$S_{t+1} = S_t \times \left(1 - D_{t+1} / N_{t+1}\right) \quad (2.1)$$

其中， S_t 表示患者在时间 t 存活的概率， D_{t+1} 表示从 t 到 $t+1$ 时刻实际发生的事件个数， N_{t+1} 表示 $t+1$ 时刻仍然存活的受试者的数量。在时间 $t+1$ 仍然存活的受试者的数量是通过在 t 时间存活的患者减去在 t 时刻死亡的患者，再减去在 t 时刻删失的数据得到的，即 $N_{t+1} = N_t - D_t - C_t$ ，其中 C_t 表示在 t 时刻删失的受试者的数量^[56]。

即使真实的生存函数必须是连续的，但是由估计得到的 Kaplan-Meier (KM) 曲线是一个阶跃函数。此外，“危险”表明患者在 t 时间之前没有经历过事件，在 t 时间之前或 t 时间没有删失。不同时间点的危险患者数量通常显示在 KM 图上。根据随后的事件和删失情况，这个数字只能保持不变或随着时间的推移而减少。根据曲线和风险数量，很容易确定生存随时间变化的模式。KM 方法中重要的假设是截尾事件的时间分布与真实生存时间无关。需要进行生存分析的另一个原因是，生存时间通常有一个偏态分布。不考虑生存时间的平均值/中值等删减信息的统计时，中位生存时间被认为是一个合适的估计值。其中，中位生存时间是指生存函数等于 50% 的时间。描述中位生存时间的时候通常带有 95% 的置信区间

(CI)，使研究人员有可能意识到感兴趣的人群一般能存活多久。

2.4.2 Log-rank 检验

除了估计整个人群的生存函数外，研究者可能对不同突变组的患者是否有不同的生存行为或男性和女性患者的生存行为是否不同感到好奇。比较组间差异的经典统计方法，如 2-样本 t 检验或方差分析，并不适用于生存分析，因为生存数据包括删失数据。Log-rank 检验是一种非参数假设检验，可用于比较不同样本的生存分布^[57]。

Log-rank 检验是用来比较多组生存曲线之间的差异性。基于 Log-rank 检验的推论是，比较整个生存曲线，而不仅仅是中位生存时间或某个特定时间点的生存概率。检验的无效假设是在任何时间点的生存曲线之间没有差异，而另一假设则是存在生存差异。Log-rank 检验实际上是一种特殊的分层卡方检验，因为它比较的是随访期间每个时间点观测到的事件数与预期的事件数（分层的变量是时间）。与其他统计检验类似，Log-rank 检验最终获得一个 p 值，该值与所选择的显著性水平进行比较，以评估治疗组在生存率方面是否有显著差异^[56]。P 值小于 0.05 通常被认为具有统计学意义，这表明有强有力的证据反对原假设。

2.4.3 Cox 模型

Cox 比例风险模型，也被称为 Cox 回归，比例风险回归，或相对风险回归，是一个研究失败时间的回归程序。在生存分析中，常用 Cox 模型来估计风险因素对生存结果的影响。Cox 模型准确描述了风险函数^[2]中的特征和风险之间的相互作用。给定含有 d 个特征的向量 $x \in \mathfrak{R}^d$ 和参数 $\beta \in \mathfrak{R}^d$ ，可以将风险函数定义为：

$$\lambda(t|x) = \lambda_0(t)e^{\beta^T x} \quad (2.2)$$

最初的风险函数（基线风险函数） $\lambda_0(t)$ 是半参数化的，并不需要由 Cox 模型指定。这将减少对 Cox 模型的限制。基线风险函数描述的是 $x = 0$ 时的风险，

即未合并特征时的风险。风险函数描述了基线风险与特定样本特征之间的关系，以量化特定时间的危险或风险。该模型只需要满足比例风险假设，即一个样本的风险与另一个样本^[58]的风险成正比。当比值不依赖于时间时，两个样本 x_1 和 x_2 满足这个假设，比值计算方法如下式 2.3 所示：

$$\frac{\lambda(t|x_1)}{\lambda(t|x_2)} = \frac{\lambda_0(t)e^{\beta^T x_1}}{\lambda_0(t)e^{\beta^T x_2}} = \frac{e^{\beta^T x_1}}{e^{\beta^T x_2}} \quad (2.3)$$

此外，训练数据的相对风险与平均风险的比值是由样本 x_k 与样本均值 \bar{x} 定义的，如式 2.4 所示：

$$\frac{\lambda(t|x_k)}{\lambda(t|\bar{x})} = \frac{\lambda_0(t)e^{\beta^T x_k}}{\lambda_0(t)e^{\beta^T \bar{x}}} = \frac{e^{\beta^T x_k}}{e^{\beta^T \bar{x}}} \quad (2.4)$$

（一）Cox 模型等同于 Poisson 回归

参数 β 可以通过最大化部分似然来找到。对于每个样本 i ，假设 x_i 是一个特征向量； T_i 为事件发生时的时间或截尾时间，如果感兴趣的事件没有发生在该时间段内，则是最后一次观察样本的时间； D_i 死亡指标，其中 1 表示事件发生的时间，0 表示事件删失的时间。Cox 模型的部分似然可以用广义线性模型 Poisson 回归的似然来拟合，因为两者的似然是成正比的^[59]。

Cox 模型可以用 Poisson 回归来解释。给定累积风险 $\Lambda(t)$ 以及样本 i ，参数 β 的估计可以通过将死亡指标 D_i 看作 Poisson 分布，其平均值 $\mu_i = \Lambda(t_i)e^{\eta_i}$ ，其中， $\eta = \beta^T x$ 。联结函数被修改为 $\beta^T x = \log(\mu_i) - \log(\Lambda(t_i))$ 。

（二）偏似然

为了可以用公式表示偏似然， f 个独立的失败时间按照递增被排序为 $t_1 < t_2 < \dots < t_f$ ， $j(i)$ 表示在时间 t_i 失败的样本的索引。当每个时间段最多只有一个样本失败时，Cox 模型的偏似然函数可以写为：

$$L(\beta) = \prod_{i=1}^f \frac{\lambda_0(t_i) e^{\beta^T x_{j(i)}}}{\sum_{j \in R_i} \lambda_0(t_i) e^{\beta^T x_j}} = \prod_{i=1}^f \frac{e^{\beta^T x_{j(i)}}}{\sum_{j \in R_i} e^{\beta^T x_j}} \quad (2.5)$$

其中，风险集 R_i 是死亡或删失时间发生在 t_i 或 $R_i = \{k | T_k \geq t_i\}$ 之后的样本索引集。这代表了在时间 t_i 有风险的样本中，一个样本在时间 t_i 发生失败的概率。由于基线风险 λ_0 被抵消了，所以半参数属性可以在这里表现出来。

但是，上面的部分可能性没有考虑到并列事件，所以概率没有那么准确。如果在时间 t_i 的死亡人数 d_i 大于 1，就会发生并列事件。Breslow 引入了一个不同的偏似然函数来处理并列问题^[60]。给定的 $I(i)$ 是样本在时间 t_i 或 $I(i) = \{k | D_k = 1 \text{ and } T_k = t_i\}$ 失败的集合的索引，偏似然函数可以重新定义如下所示：

$$L(\beta) = \prod_{i=1}^f \frac{e^{(\sum_{s \in I(i)} \beta^T x_s)}}{(\sum_{j \in R_i} e^{\beta^T x_j})^{d_i}} \quad (2.6)$$

(三) 最小化负偏对数似然

参数 β 可以通过最小化负偏对数似然 $\ell(\beta)$ 来得到，其定义如下：

$$\ell(\beta) = \log(L(\beta)) = \sum_{i=1}^f \log \frac{e^{(\sum_{s \in I(i)} \beta^T x_s)}}{(\sum_{j \in R_i} e^{\beta^T x_j})^{d_i}} = \sum_{i=1}^f [(\sum_{s \in I(i)} \beta^T x_s) - d_i \log \sum_{j \in R_i} e^{\beta^T x_j}] \quad (2.7)$$

带 ℓ_1 正则化的最小化负对数似然，在如下目标 $f(\beta)$ 的优化问题中实现：

$$f(\beta) = -\ell(\beta) + \lambda \|\beta\|_1$$

$$\min_{\beta} f(\beta) \quad (2.8)$$

(四) Cox 模型的生存函数

从 Cox 模型中得到的生存函数可以用于对样本生存的预测，因为它是样本在时间 t 后存活的概率。生存函数的定义为：

$$S(t) = \exp(-\Lambda(t)) \quad (2.9)$$

累积风险或危险函数 $\Lambda(t)$ 定义为：

$$\Lambda(t) = \int_0^t \lambda(s) ds \quad (2.10)$$

其中， $\lambda(t)$ 为风险函数，即在生存到 t 时间的情况下， t 时刻的瞬时死亡概率。它也可以被重写为：

$$\lambda(t) = -\frac{d}{dt} \log S(t) \quad (2.10)$$

对于一个给定的样本 x ，生存函数可以在 t 时刻被重写为：

$$\begin{aligned} S(t|x) &= \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(s|x) ds\right) \\ &= \exp\left(-\int_0^t \lambda_0(s) e^{\beta^T x} ds\right) \\ &= \exp\left(-e^{\beta^T x} \int_0^t \lambda_0(s) ds\right) \\ &= S_0(t) e^{\beta^T x} \end{aligned} \quad (2.11)$$

其中，基线累积风险为 $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$ ，基线生存函数为 $S_0(t) = e^{-\Lambda_0(t)}$ 。生存函数 $S(t|x)$ 或时间 t 后的生存概率由 $S_0(t) e^{\beta^T x}$ 定义^[58]。

2.5 评价指标

2.5.1 分类评价指标

对于二分类问题，常用的分类器评价指标包含混淆矩阵、准确率 (Accuracy)、精准率 (Precision)、召回率 (Recall)、F1-score。针对二分类问题，分类器预测出的结果只有两种情况，即正例 (Positive) 和反例 (Negative)，分别用 P 和 N 表示，如果分类器预测结果正确用 T 表示，预测出错用 F 表示。则将会出现四种结果，分别为 TP、TN、FP、FN。其中，TP 代表实际类别为正例，分类器同样预测为正例的情况；同样，TN 代表实际类别为反例，分类器同样预测为反例的情况；FP 代表实际类别为反例，分类器预测为正例的情况；FN 代表实际类别为正例，分类器预测为反例的情况。混淆矩阵如表 2.1 所示：

表 2.1 二分类混淆矩阵

		实际类别	
		1	0
预测类别	1	TP	FP
	0	FN	TN

准确率是正确分类的样本数与总样本数之间的比值，计算公式如下所示：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.12)$$

精准率指的是分类正确的正样本个数与分类器预测为正样本个数的比例，计算公式如下：

$$Precision = \frac{TP}{TP + FP} \quad (2.13)$$

召回率指的是分类正确的正样本个数与实际的样本个数的比值，计算公式如下：

$$Recall = \frac{TP}{TP + FN} \quad (2.14)$$

F1-score 为精确率和召回率的调和平均值，计算公式如下：

$$F1\text{-score} = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.15)$$

2.5.2 生存分析评价指标

C-index, C 指数, 即一致性指数 (index of concordance), 用来评价生存分析模型的准确结果。C-index 计算的是模型预测出来的结果与实际观测到的结果一致的概率, 取值范围在[0,1]之间, 计算方法如下:

$$C\text{-index} = \frac{1}{n} \sum_{i \in \{1, \dots, N\} | \delta_i = 1} \sum_{t_j > t_i} I[f_i > f_j] \quad (2.16)$$

其中, n 是可比较数据对的数量, $I[\cdot]$ 表示指示函数, t 指的是实际的时间观测值, f 表示模型预测出来的患者的风险概率。C-index 的数值越大, 表示模型的预测结

果越好^[61]。

2.6 本章小结

本章主要介绍的是本文中涉及到的一些概念及知识介绍,主要介绍了局部晚期非小细胞肺癌、FDG-PET/CT 的概念,并且对基于分类的 3D CNN 进行了简要介绍,然后详细介绍了在生存分析中用到的 Kaplan-Meier 方法、Log-rank 检验和 Cox 模型的原理,最后介绍了本文中所用到的关于分类以及生存分析的评价指标。

3 基于治疗后 SUVmax 预测对局部晚期非小细胞肺癌患者的生存分析研究

在基于治疗后肿瘤 SUVmax 的预测对局部晚期非小细胞肺癌患者的生存分析研究中, 包括 143 名同时包含治疗前和治疗后 FDG-PET/CT 图像的患者以及 40 名只包含治疗前 FDG-PET/CT 图像的患者。为了准确预测出患者在接受治疗后的肿瘤 SUVmax 是否大于 5.0, 本研究采用了 3D CNN 对患者的 FDG-PET 采用阅片的方式, 使 CNN 自动从患者的图像中提取到相关信息从而对其进行预测。为了进一步研究患者治疗后肿瘤 SUVmax 对患者生存预后的影响, 本文基于预测出的结果采用 Cox 模型对患者进行了生存分析的研究。

3.1 研究方法

本节首先对本实验的实验流程进行了介绍, 其次详细介绍了本实验的数据处理过程, 主要介绍了患者医学图像的选择、本实验所纳入数据的详细情况、临床特征的选择以及对患者 FDG-PET 图像的预处理, 之后主要对实验的可行性进行了验证, 然后主要对本实验所用到的 3D CNN 的模型以及网络的训练过程进行了详细介绍, 最后介绍了本实验的生存分析部分。

3.1.1 实验流程

本研究的主要目的是确定治疗后的肿瘤 SUVmax 与患者总生存期之间的关系, 并进一步确定, 利用 3D CNN 基于治疗前 FDG-PET 预测出的治疗后肿瘤 SUVmax 是否大于 5.0 的 0-1 标签是否可以提高患者生存分析的准确性。在本研究中, 本文所采用的方法包括两个部分: ①基于 3D CNN 使用治疗前 FDG-PET 来预测局部晚期非小细胞肺癌患者治疗后的 SUVmax 是否大于 5.0; ②评估第一部分预测出来的患者治疗后的肿瘤 SUVmax 是否大于 5.0 对局部晚期非小细胞肺癌患者的预后影响。具体的实验步骤如图 3.1 所示。

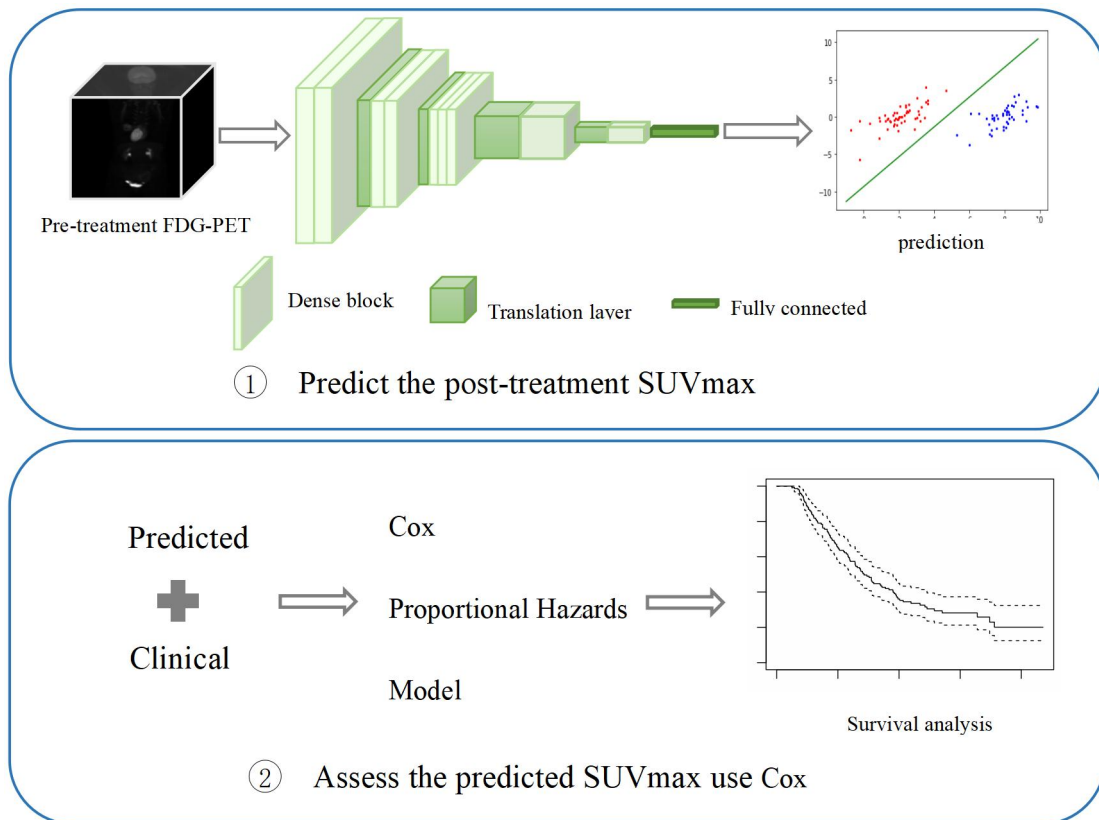


图 3.1 本文的研究流程图。① 一种分类器的网络结构，用于预测在治疗后没有 FDG-PET 的患者在治疗后的 SUVmax。② 通过生存分析的结果来评估预测的肿瘤 SUVmax 的准确度。

3.1.2 实验数据介绍

(一) 数据来源

本文所采用的实验数据为公开数据集，由美国放射学会核心实验室合作提供。患者数据来自美国放射影像网络 (ACRIN) 6668/放射治疗肿瘤组 (RTOG)/0235，进行的多中心临床试验。ACRIN 6668 多中心临床试验的目的是确定治疗后不久 FDG-PET 成像的 PET SUV 测量值是否是预测最终放化疗后长期临床结果(生存)的有用指标。

数据选取的是 18 岁以上的 III 期和无法手术的 II 期非小细胞肺癌患者^[9]。纳入研究的患者同时接受常规的以铂金为基础的化放疗，不做手术；并且允许放疗后巩固化疗。在治疗前进行了全身 FDG-PET 扫描。第二次治疗后全身 FDG-PET 扫描在放疗后约 14 周进行（辅助化疗后至少 4 周）。治疗前 FDG-PET 扫描是在

符合 ACRIN 标准的扫描仪上进行的。治疗后的 FDG-PET 扫描需要在治疗结束后 12-16 周内进行，使用与治疗前扫描相同的扫描仪^{[62][1]}。患者数据的采集是从 2005 年 6 月份开始到 2009 年 5 月结束。总共从 37 家机构中招募到了 250 名患者，其中有 8 名患者没有可供评估的治疗前 PET，最终该数据集中有 242 名患者，可在 TCIA^[63]中找到。

研究中使用的是用传统的现代设备/技术进行的 FDG-PET(有或没有 PET/CT)扫描。扫描前患者必须禁食 4 小时，并且在注射 FDG 前血糖水平低于 200 mg/dL。FDG 的剂量没有规定，推荐的剂量是 0.14 至 0.21 mCi/kg（约 10 至 20mCi）。发射扫描在注射 FDG 后 50 至 70 分钟开始，包括从上/中颈到股骨近端的身体。发射和透射扫描的采集时间符合制造商的建议^[1]。根据标准的临床实践，在治疗结束后对患者进行了至少 2 年的观察（或直到死亡）。允许进行非协议 PET 成像，但不强制要求。

（二）临床特征的初步筛选

本文所研究的数据中，患者的临床信息包括各种各样的表格，其中包括患者的资格审查表、初步评估表、PET 技术评估表、由核医学放射科医生填写的 PET 半定量评估表、随访表等与患者相关的一些信息。首先，在兰大二院的医生协助下从中挑选出与患者生存预后相关的一些临床信息；然后将患者的临床信息进行汇总，之后剔除掉一些重复的信息之后，最终得到初步处理好的 242 名患者的临床表格数据。

3.1.3 实验数据处理

该数据集总共有 250 名患者，共有 242 名非小细胞肺癌患者的影像数据，其中 239 名患者有 PDG-PET 图像数据，180 名患者有治疗前和治疗后的 FDG-PET 图像，156 名患者有治疗前和治疗后的 FDG-PET/CT 图像（训练数据）；59 名患者只有治疗前的 FDG-PET 图像，45 名患者有治疗前的 FDG-PET/CT 图像（测试数据）。然后将患者的 FDG-PET/CT 图像经过 3DSlicer^[64]处理得到 FDG-PET 图像对应的 SUV 图像，去除处理异常和 SUV 图像异常的数据以及临床分期非 II

B、IIIA、IIIB 的患者；对训练数据，去除缺少治疗前临床信息以及无治疗后 SUV 的患者，最终纳入 143 名患者；对测试数据，去除无生存期和生存状态的患者，最终纳入 40 名患者。关于患者群体的详细信息列于表 3.1。

表 3.1 本研究所使用数据的分布情况

患者临床特征	训练数据集 (N = 143)		测试数据集 (N = 40)	
	No.	%	No.	%
年龄				
中位数	65		63.5	
范围	37-82		37-81	
临床分期				
II B	5	3.5	3	7.5
III A	82	57.3	16	40
III B	56	39.2	21	52.5
其它	0	0	0	0
性别				
男性	95	66.4	26	65
女性	48	33.6	14	35
Performance status				
0 (fully active)	75	52.4	10	25
1 (ambulatory, capable of light work)	68	47.6	30	75
种族				
Asain	16	11.2	7	17.5
Black or African American	12	8.4	5	12.5
White	113	79	1	2.5
Native Hawaiian or other Pacific Islander	0	0	25	62.5
Unknown	5	3.5	3	7.5

续表 3.1

患者临床特征	训练数据集 (N = 143)		测试数据集 (N = 40)	
	No.	%	No.	%
化疗方案				
1 (Cisplatin)	2	1.4	1	2.5
2 (Carboplatin)	23	16	8	20
3 (Paclitaxel)	60	42	16	40
4 (Etoposide)	25	17.5	5	12.5
5 (Navelbine)	6	4.2	0	0
6 (Taxotere)	11	7.7	2	5
8 (Other, specify)	15	10.5	2	5
Data not available	1	0.7	6	15
Adjuvant (postradiotherapy)				
chemotherapy given				
yes	27	18.9	11	27.5
no	93	65	23	57.5
Data not available	23	16.1	6	15
Radiation dose, Gy				
<50	0	0	6	15
50-60	5	3.5	3	7.5
60-70	101	70.6	17	42.5
≥70	35	24.5	8	20
Data not available	2	1.4	6	15

由上述表 3.1 可知，本文所纳入的数据中，局部晚期非小细胞肺癌患者的年龄分布在 37 到 82 岁之间。在训练数据集中有 3.5% 的患者处于不可手术的 II 期，其余患者均处于 III 期；在测试数据中，有 7.5% 的患者处于不可手术的 II 期，其

余患者处于III期。并且训练数据集和测试数据集的男女比例以及是否在接受放疗后辅助化疗的比例相似，这说明了测试数据集的数据分布与训练数据集相似。

(一) 临床特征选择

本文所研究的工作是：基于局部晚期非小细胞肺癌患者治疗前的一些信息（包括临床特征以及影像数据）是否可以相对准确的预测出患者接受治疗后的肿瘤 SUVmax。因此，在进行临床信息挑选的时候，只能应用患者治疗前的信息，避免涉及到患者接受治疗后的一些数据，以免数据泄露。然后，利用训练数据集将所挑选出来的患者治疗前的一些临床信息用 Cox 单因素回归分析进行挑选，最终挑选出与生存分析相关的临床特征有六个。纳入的六个临床特征分别为：绝对粒细胞的化验结果、患者接受治疗前的总的代谢性肿瘤体积、患者接受治疗前的原发瘤的代谢性肿瘤体积、对侧位于纵隔内的淋巴结是否受影响、患者接受治疗前的所有病变的代谢性肿瘤体积、患者接受治疗前的原发瘤的总代谢量。它们在临床信息表格中对应的代码编号分别为：ile53、lee7、lee8、sse10、sse16、sse21。最后将六个临床特征在训练数据集上使用 Cox 多因素回归分析进行挑选，最终挑选出两个临床特征与患者的生存预后相关，分别为：绝对粒细胞的化验结果（ile53）、对侧位于纵隔内的淋巴结是否受影响（sse10）。关于临床特征 Cox 单因素、多因素分析的结果如表 3.2 所示。

表 3.2 单因素多因素分析中的临床信息对患者总生存期的预后意义

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
ile53	16.3 (2.1 - 129.2)	0.008	30.4 (3.7 - 247.3)	0.001
lee7	1.0 (1.001 - 1.005)	0.0003	-	-
lee8	1.0 (1.001 - 1.005)	0.002	-	-
sse10	1.0 (1.515 - 2.54)	0.008	1.8 (1.2 - 2.7)	0.004
sse16	1.0 (1.001 - 1.005)	0.010	-	-
sse21	1.0 (1.0 - 1.004)	0.028	-	-

（二）医学图像数据预处理

实验所采用的数据需要进一步的处理才能开始进行实验。对于患者的影像数据来说，因为患者的数据是从 37 家不同的机构中获得的，因此得到的患者的影像数据的层厚以及图像像素的范围是不一样的，因此需要考虑将患者的影像数据进行重采样以及对患者的影像数据进行归一化操作。首先将患者的 FDG-PET 图像进行重采样，调整患者进行 FDG-PET 检查时得到的图像的层厚，保证不同患者间的 FDG-PET 层厚相同。之后将患者的影像数据进行归一化操作，转换到 0 到 1 之间，这样不仅可以加速网络在训练时收敛的速度还可以减少医学图像在拍摄时由于光线原因造成的干扰。最后为了保证模型在训练时能够进行批处理操作，将图像进行了重采样操作，将图像的大小统一调整为 $128 \times 128 \times 128$ 。

3.1.4 实验验证

由于本文是在 Machtay 等人^[1]研究结果的基础上进行的进一步研究，并且在数据筛选过程中有一部分数据因为数据缺失等原因进行了剔除，导致本文的研究数据与 Machtay 等人^[1]所纳入的数据有所不同，为了保证实验能够正常进行，在开始进行研究之前对实验的可行性进行了验证。实验可行性验证是为了进一步验证本文所研究内容的合理性。

（一）实验验证流程

Machtay 等人^[1]在文章中得出结论，患者治疗后较高的肿瘤 SUV 与 III 期非小细胞肺癌较差的生存率相关，并发现治疗后肿瘤 SUV_{max} 在采用 5.0 的二值截取值时与患者的生存率有显著的相关性，基于此，本文在所纳入的数据集中进行了实验验证。具体是为了验证，在本文所纳入的数据集中，患者接受治疗后的肿瘤 SUV_{max} 在采用 5.0 的二值截取值时同样与患者的生存率有显著的相关性，并且患者接受治疗前的肿瘤 SUV_{max} 与其生存率的相关性并不显著。

实验验证的流程具体如下：首先，采用 Cox 单因素和多因素分析，验证患者治疗前和治疗后肿瘤 SUV_{max} 在采用 5.0 的二值截取值时是否为影响患者生存预后的关键因素；其次，以患者接受治疗前和治疗后的肿瘤 SUV_{max} 是否大于

5.0 为界，将患者分为两组，采用 Kaplan-Meier 方法和 Log-rank 检验来验证两组患者的生存曲线是否存在明显差异。若验证结果显示，治疗前患者肿瘤 SUVmax 不是影响患者生存预后的关键因素，且两组患者的生存曲线并不存在明显差异；而治疗后肿瘤 SUVmax 是影响患者生存预后的关键因素，且两组患者 Cox 单因素和多因素分析的生存曲线存在明显差异，则证明了本实验的可行性。

(二) 实验验证结果

关于治疗前和治疗后患者肿瘤 SUVmax 的 Cox 单因素和多因素分析结果如表 3.3 所示，结果显示患者治疗后 SUVmax 在采用 5.0 的二值截取值时，与患者的生存预后显著的相关性，其回归系数分别为 2.0 和 2.1，P 值分别为 0.0004 和 0.0003。

表 3.3 治疗前和治疗后 SUVmax 对患者总生存期的生存预后

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
治疗前 SUVmax	0.8 (0.1 - 6.1)	0.87	0.6 (0.1 - 4.5)	0.64
治疗后 SUVmax	2.0 (1.4 - 3.0)	0.0004	2.1 (1.4 - 3.1)	0.0003

在依据患者肿瘤 SUVmax 是否大于 5.0 将患者分为不同的两组时，采用 Kaplan-Meier 方法得到的分析结果如图 3.2、和图 3.3 所示。由于本文所纳入的数据中，只观察到一例治疗前肿瘤 SUVmax 小于 5.0 并且在随访期内死亡的患者，因此出现了如图 3.1 所示的情况。为了进一步确定患者治疗前的肿瘤 SUVmax 对患者的影响，依据 Machtay 等人^[1]的研究，本文进一步研究了肿瘤 SUVmax 是否大于 7.0 对两组不同患者生存曲线的影响，分析结果如图 3.4 所示。

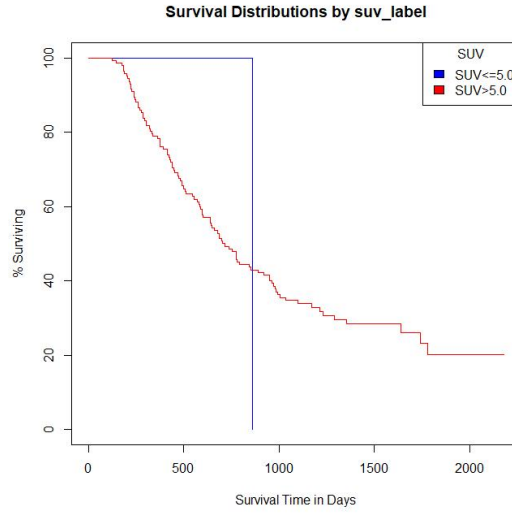


图 3.2 对治疗前肿瘤 SUV_{max} 使用 5.0 的二值截取值时的 KM 曲线

结果显示，治疗后的肿瘤 SUV_{max} 在采用 5.0 的二值截取值时可以将两组患者完全分开，且其 P 值为 0.0003，说明治疗后肿瘤 SUV_{max} 是否大于 5.0 是影响患者生存预后的一个关键因素。在对患者治疗前的肿瘤 SUV_{max} 采用 5.0 和 7.0 的二值截取值时，两组患者生存曲线之间的 P 值分别为 0.9 和 0.2，均与患者的生存预后无明显影响，也进一步验证了本研究可行性。

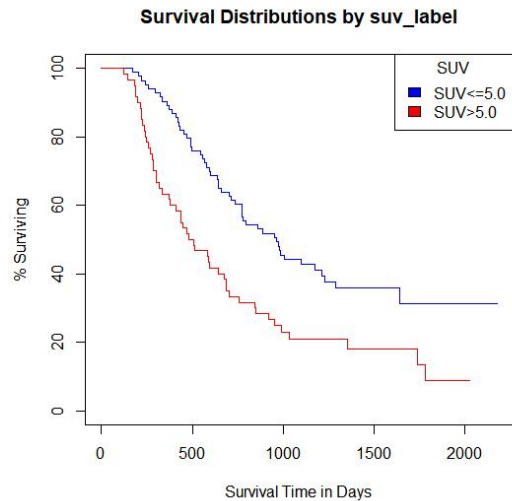


图 3.3 对治疗后肿瘤 SUV_{max} 使用 5.0 的二值截取值时的 KM 曲线

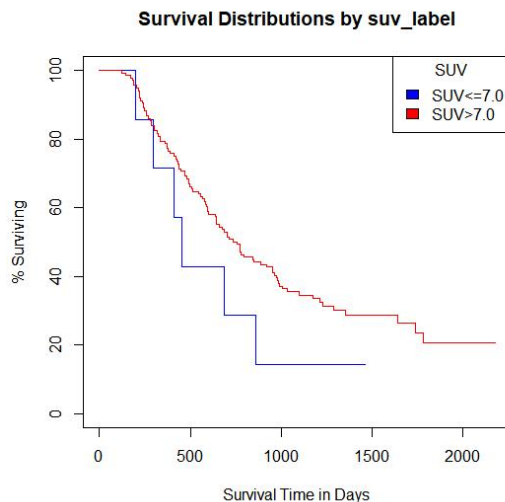


图 3.4 对治疗前肿瘤 SUVmax 使用 7.0 的二值截取值时的 KM 曲线

3.1.5 对患者接受治疗后的肿瘤 SUVmax 预测的分类模型

在采用 3D CNN 对患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 进行预测时，本文所采用的分类模型为 3D DenseNet169。3D DenseNet169 由三个模块组成，分别为：Dense block、Translation layer 以及最后的全连接层 Fully connected。在模型的每个 Dense block 中，每个卷积层的特征图大小保持一致，并可以在通道维度上进行连接，其具体结构如图 3.5 所示。Dense block 由批量归一化 Batch Normal 层，激活函数 Relu 层以及卷积核大小分别为 $1 \times 1 \times 1$ 和 $3 \times 3 \times 3$ 的卷积 Conv 层组成。模型中的 Translation layer 具体结构如图 3.6 所示，由批量归一化 Batch Normal 层，激活函数 Relu 层，卷积核大小别为 $1 \times 1 \times 1$ 的卷积 Conv 层以及最后的平局池化 AvgPool 层组成。在模型的最后跟了一个分类层，是由 Relu 激活函数、全局平均池化、以及全连接层构成的。将患者的三维图像输入网络中，最终经过分类层得到最后的分类结果。

Densenet 密集的连接方式，可以提升网络的训练速度，使得网络更加容易训练；并且由于 Densenet 网络的连接方式可以对特征进行重复利用，最后网络分类时使用的是其低级特征^[65]，网络的具体结构可在参考文献^[63]中找到。

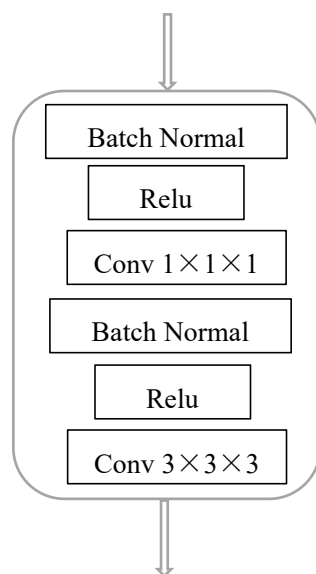


图 3.5 模型中的 Dense block 结构示意图

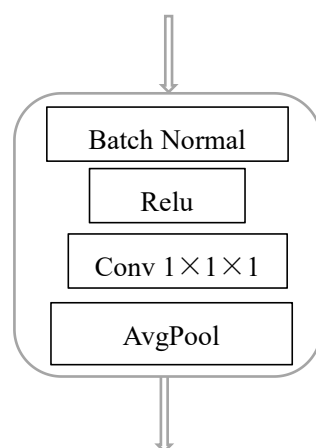


图 3.6 模型中的 Translation layer 结构示意图

根据本文所采用的数据的分布情况：数据中有 143 名患者有治疗前和治疗后的数据，40 名患者仅有接受治疗前的数据。因此在根据患者治疗前的 FDG-PET 来预测患者接受治疗后的肿瘤 SUVmax 时，只能根据同时拥有治疗前和治疗后数据的 143 名患者进行模型的训练、以及参数的选择。仅有治疗前数据的 40 名患者由于没有治疗后的数据，因此只能作为测试集，在测试集上进行模型预测结果的评价时，可以采用 Cox 单因素以及多因素分析的方法对分类模型预测出的结果好坏进行进一步的评价。

3.1.6 基于预测出的治疗后 SUVmax 的生存分析

在生存分析部分，采用 Cox 模型来对患者的生存预后进行分析。模型的训练是由包含 143 名患者的训练数据集进行的，模型的测试是由包含 40 名患者的测试数据进行的。整个生存分析包含两个实验，第一个实验是在只包含患者的两个临床数据上进行的，第二个实验是在第一个实验的基础上，在训练数据集上加上了患者治疗后的肿瘤 SUVmax 是否大于 5.0 的 0-1 标签，测试数据集加上的是由分类模型预测出的患者接受治疗后肿瘤 SUVmax 是否大于 5.0 的 0-1 标签。最后对两个生存分析的实验结果进行了显著性检验，用来判定两个生存分析实验结果的不同是不是由偶然性引起的。患者的生存分析采用的是 R 语言编程软件进行的。

3.2 实验结果与分析

本节主要对本实验两部分的实验结果进行了详细描述。主要介绍了 3D DenseNet169 网络的分类结果、训练过程中所用到的损失函数、模型的参数设置以及生存分析部分的相关实验结果介绍。

3.2.1 对患者接受治疗后的肿瘤 SUVmax 预测的分类实验结果

(一) 模型的损失函数

模型在进行训练过程中是通过最小化预测结果与真实值之间的差距对模型进一步的优化，采用交叉熵损失来衡量预测结果与真实值之间的差距。交叉熵损失函数的定义如下：

$$l(\Theta) = \frac{1}{n} \sum_{i=1}^n H(y^{(i)}, \hat{y}^{(i)}) \quad (3.1)$$

$$H(y^{(i)}, \hat{y}^{(i)}) = - \sum_{j=1}^q y_j^{(i)} \log \hat{y}_j^{(i)} \quad (3.2)$$

其中， Θ 表示模型参数； n 表示样本个数； q 表示类别数； $y^{(i)}$ 是一个 n 维向量，

表示真实的类别标签； $\hat{y}^{(i)}$ 同样是一个 n 维向量，表示预测出的类别概率^[66]。

（二）模型的参数设置

模型在训练过程中的轮训练次数，即 epoch 设置为 100，学习率(learning rate) 设置为 2×10^{-4} ，Batch size 大小设置为 6，优化器选择的是 Adam 算法来实现的，模型最后参数的选择是根据验证集来进行挑选的。

（三）实验设备

分类网络所采用的实验是在 Ubuntu 20.04 的操作系统上进行的，采用的是 Python3.9 的编程语言，在 Pytorch1.11.0 的深度学习环境下进行的实验，并使用 MONAI 进行实验参数调整以及模型的训练。实验设备带有 NVIDIA GeForce RTX3090 的显卡。

（四）模型在验证集上的分类结果

模型是在训练集上进行训练的，模型在训练时从训练集中的 143 名患者随机挑选其中的 70% 用来进行模型的训练，其余 30% 的患者作为验证集，用来对模型的参数进行选择。模型在验证集上的表现结果作为对分类模型的初步评价。分类模型在验证集上表现的结果如表 3.4 所示。实验结果显示，模型对验证集上的患者进行治疗后的 SUVmax 是否大于 5.0 的预测结果是比较准确的，模型的准确率达到 0.63，召回率达到 0.76。

表 3.4 分类模型在验证集上的实验结果

网络	准确率	精确率	召回率	F1_score	AUC
3D DenseNet169	0.63	0.52	0.76	0.62	0.64

（五）测试集上分类结果的评价

在测试集上对分类的结果好坏进行评价时，采用的是 Cox 单因素、多因素分析，具体结果如表 3.5 所示。结果显示，模型在测试集上预测出的患者治疗后的 SUVmax 与其总生存期也有显著的影响，在对其进行单因素和多因素分析时，其 P 值均小于 0.05，也进一步证明了模型预测结果的准确性。

表 3.5 预测的治疗后 SUVmax 对患者总生存期的影响

变量	单因素		多因素	
	HR (95% CI)	P	HR (95% CI)	P
预测的治疗后 SUVmax	2.25 (1.1 - 4.7)	0.034	2.6 (1.2 - 5.9)	0.017

3.2.2 基于预测出的治疗后 SUVmax 的生存分析实验结果

两个实验的生存分析结果如表 3.6 所示。在训练数据集上对患者进行生存分析时，加上患者治疗后 SUVmax 是否大于 5.0 的 0-1 标签后，患者生存分析的准确性得到了显著提高，C-index 由 0.59 提升到了 0.66，P 值为 0.002。在测试集上，加上分类模型预测出的患者治疗后肿瘤 SUVmax 是否大于 5.0 的 0-1 标签后，测试集生存分析结果的 C-index 从 0.59 提升到了 0.68，P 值为 0.013，同样小于 0.05。

表 3.6 生存分析实验结果 (C-index)

	训练集	测试集
Clinical	0.59	0.59
Clinical+SUVmax (0-1)	0.66	0.68
P	0.002	0.013

训练集患者生存分析结果的提升说明了患者接受治疗后的肿瘤 SUVmax 在采用 5.0 的二值截取时对患者生存预后的重要性；测试集上患者生存分析结果的提升不仅进一步验证了基于患者治疗前 FDG-PET 对患者接受治疗后肿瘤 SUVmax 是否大于 5.0 预测结果的准确性，还证明了基于治疗前 FDG-PET 预测出的患者治疗后肿瘤 SUVmax 是否大于 5.0 对患者生存分析的有效性。

3.2.3 讨论

表 3.4 中，患者治疗后的肿瘤 SUVmax 是否 5.0 的结果显示出了模型预测结

果的准确性，其中模型的召回率达到了 0.76。但是由于测试集没有患者接受治疗后的肿瘤 SUVmax 的数据，因此，进一步研究了预测出的结果对患者生存分析的影响，实验结果如表 3.5 所示。表中的 HR 表示风险概率，主要是用来评估预测出的结果对患者总生存期的影响大小，HR 的数值小于 1 表示该因素对患者的生存预后是有利影响，HR 的数值大于 1 表示该因素对患者的生存预后是不利影响。实验结果显示，从患者治疗后的 FDG-PET 上测量得到的患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 的 0-1 标签，对患者生存预后单因素和多因素影响的 HR 数值分别为 2.0 和 2.1；分类模型预测出的患者接受治疗后的肿瘤 SUVmax，对患者生存预后单因素和多因素影响的 HR 数值分别为 2.25 和 2.6，进一步说明了模型预测结果的准确性。表 3.6 中，生存分析的实验结果表明患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 是影响患者生存分析的重要因素，并且可以显著提升患者生存分析的结果，说明患者接受治疗后的肿瘤 SUVmax 是否大于 5.0 是影响患者生存预后的重要因素。本文的实验表明，基于患者治疗前的影像可以准确预测出患者接受治疗后的肿瘤 SUVmax 是否大于 5.0，并且预测出的结果对患者接受治疗后的生存预后具有重要指导意义。

3.3 本章小结

本章主要对基于治疗后 SUVmax 预测对局部晚期非小细胞肺癌的生存分析的工作进行了详细描述。首先是实验数据部分进行了介绍，包括总体实验数据的选择、实验所纳入数据的分布情况、患者在接受治疗前的临床特征进行挑选的详细信息以及在 CNN 中所用到的患者治疗前的医学影像数据 FDG-PET 的预处理步骤。其次对本文所研究的内容进行的可行性进行了验证，可行性验证是在基于 Machtay 等人^[1]工作的基础上进行的实验，具体采用了 Cox 单因素、多因素回归分析、Kaplan-Meier 方法和 Log-rank 检验来对患者治疗后 SUVmax 是否大于 5.0 对患者生存预后的影响进行了验证。之后是对本研究所用的方法部分进行了介绍。方法具体包含了两个部分，分别是对患者治疗后肿瘤 SUVmax 是否大于 5.0 的预测以及在第一部分结果的基础上对患者进行的生存分析。文中分别对两个部分的

实验进行了详细的介绍,在对患者治疗后肿瘤 SUVmax 是否大于 5.0 的预测部分详细介绍了本文所使用的分类预测模型 3D DenseNet169 的具体结构,以及在训练模型时所使用的模型参数设置、使用的损失函数。最后对模型的分类结果进行了描述。在基于预测出的患者治疗后 SUVmax 结果的基础上对患者进行生存分析的部分,主要详细介绍了生存分析的模型的训练过程以及最后得出的结果。最终的实验结果显示,在对没有治疗后信息的患者进行的治疗后患者肿瘤 SUVmax 预测的基础上进行的生存分析研究可以在一定程度上提升患者生存分析的准确性,并且对患者的生存预后具有显著性意义。

4 基于临床信息对局部晚期非小细胞肺癌患者的生存分析研究

在基于患者的临床信息对局部晚期非小细胞肺癌患者的生存分析研究中，是基于 138 名患者的治疗前和治疗后的临床信息对患者进行的研究。首先是对患者的临床信息进行挑选，挑选出与患者生产出分析相关的临床特征，其次再基于挑选出的特征对患者进行生存分析。

4.1 研究方法

4.1.1 实验流程

在对患者进行生存分析之前，需要挑选出与患者生存分析相关的若干临床特征，然后再进行生存分析的研究。在利用训练数据集进行临床特征选择时，采用的是 Cox 单因素和多因素回归分析的方法，从中挑选出与患者生存分析相关的因素 ($P < 0.05$)；然后在进行生存分析时也是采用的 Cox 模型对患者进行生存预测。具体的实验流程如图 4.1 所示。

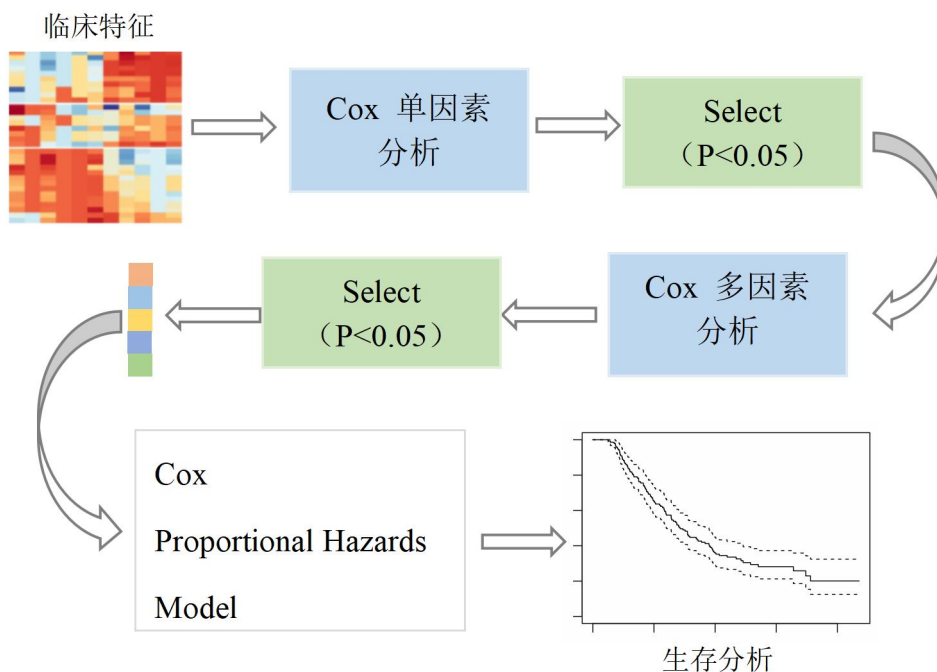


图 4.1 生存分析实验流程图

4.1.2 实验数据介绍

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3.1.2小结关于实验数据的介绍部分。

4.1.3 实验数据处理

在利用临床信息对局部晚期非小细胞肺癌患者进行生存分析时,采用的是患者治疗前和治疗后的临床信息。依据兰大二院的医生协助挑选出的临床特征对患者的临床信息进行进一步的处理,从中剔除临床信息空白的特征,以及只有一个特征的信息,以及缺少治疗前和治疗后临床信息的患者,最后挑选出 138 名患者。最终基于 138 名患者的临床信息对患者进行生存分析的研究。

与上一部分的工作不同的是,在基于临床信息对局部晚期非小细胞肺癌患者的生存分析的研究中,本研究所采用的 138 名患者的数据均含有治疗前和治疗后的一些特征,因此在这一部分的研究中采用的是五折交叉验证。在对临床信息进行特征选择的时候需要用到训练数据集,然后依据训练数据集挑选出的若干临床特征挑选出测试数据所含有的特征。

4.1.4 Cox 单因素和多因素分析的临床特征挑选

实验采用的是五折交叉验证的结果,因此每一折在进行临床特征挑选时的训练数据不同,最终挑选出的临床特征也不一样,每一折挑选出的临床特征如表 4.1 所示。每一折挑选出的临床特征个数不一样,分别是五个、三个、三个、四个、五个。表格中的字母表示的是实验挑选出的临床特征在患者临床表格中的相关编码。

4.1.5 生存分析模型的训练过程

Cox 比例风险回归模型的训练是由训练数据集完成的,然后在测试数据集上进行测试。在总共的 138 名患者中,有 40 名(29.0%)患者在随访过程中死亡。

生存分析的实验同样是在 Ubuntu 20.04 的操作系统上进行的，采用的是 R 语言编程软件，RStudio 对代码进行编译，Cox 单因素和多因素分析采用的是 R 语言的 Survival 包，Cox 比例风险回归模型采用的是 R 语言的 Survcomp 包对患者进行生存分析的。

表 4.1 每一折挑选出的临床特征 (P<0.05)

第一折	f1e8	ile5	ile40	sse40	tfe7
第二折	lee12	sse33	sse42		
第三折	ile40	lee10	sse16		
第四折	f1e8	ile40	lee11	sse33	
第五折	ile40	lee8	sse33	sse49	tfe10

4.2 实验结果与分析

本节主要对生存分析研究中临床特征挑选出的结果以及生存分析的实验结果进行了详细介绍。

4.2.1 Cox 单因素和多因素分析的临床特征挑选结果

关于每一折单因素和多因素分析挑选所得到的关于临床特征的具体分析结果如表 4.2-4.6 所示，表中变量列的字母表示在所挑选出的患者的临床特征在临床信息表格中对应的代码编号。

第一折的训练数据集利用 Cox 单因素、多因素分析总共挑选出五个临床特征，具体如表 4.2 所示，分别是：患者在随访过程中的肿瘤是否转移到新的部位(f1e8)、患者在接受治疗前的身体状况(ile5)、患者疾病的诊断性说明(ile40)、患者接受放射性治疗后转移性高代谢病变的数量(sse40)、患者终止化疗的原因(tfe7)。

第二折的训练数据集共挑选出三个临床特征，具体如表 4.3 所示，分别是：患者接受治疗后的高代谢病变的数量(lee12)、患者接受放射性治疗后对侧位

于纵隔内的淋巴结是否出现转移性疾病（sse33）、患者接受放射性治疗后原发瘤的 SUVmax（sse42）。

第三折的训练数据集共挑选出三个临床特征，具体如表 4.4 所示，分别是：患者疾病的诊断性说明（ile40）、患者接受治疗后的肿瘤 SUVmax（lee10）、患者接受放射性治疗前所有病变的总代谢性肿瘤体积（sse16）。

第四折的训练数据集共挑选出四个临床特征，具体如表 4.5 所示，分别是：患者在随访过程中的肿瘤是否转移到新的部位（fle8）、患者疾病的诊断性说明（ile40）、患者接受治疗后的肿瘤 SUV 峰值（lee11）、患者接受放射性治疗后对侧位于纵隔内的淋巴结是否出现转移性疾病（sse33）。

表 4.2 第一折单因素、多因素分析挑选得到的临床特征结果

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
fle8	0.04 (0.01 - 0.09)	<0.001	0.03 (0.02 - 0.6)	0.02
ile5	2.14 (1.15 - 4.01)	0.017	0.003 (0.001 - 0.006)	0.004
ile40	0.04 (0.01 - 0.27)	0.01	0.01 (0.001 - 0.1)	<0.001
sse40	1.0 (1 - 1.01)	0.02	0.01 (0.001 - 0.1)	0.04
tfe7	21.4 (0.65 - 184.5)	0.005	0.20 (0.012 - 0.3)	0.04

表 4.3 第二折单因素、多因素分析挑选得到的临床特征结果

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
lee12	1.34 (1.18 - 1.53)	<0.001	1.40 (1.212 - 1.620)	<0.001
sse33	3.55 (1.66 - 7.59)	0.001	3.83 (1.764 - 8.311)	<0.001
sse42	1.11 (1.03 - 1.2)	0.007	1.11 (1.031 - 1.187)	0.005

第五折的训练数据集共挑选出五个临床特征，具体如表 4.6 所示，分别是：患者疾病的诊断性说明（ile40）、患者接受治疗前原发肿瘤的代谢性肿瘤体积

(lee8)、患者接受放射性治疗后对侧位于纵隔内的淋巴结是否出现转移性疾病 (sse33)、患者接受放射性治疗后结节病变的总糖酵解活性 (sse49)、患者化疗时所采用的药剂种类 (tfe10)。

表 4.4 第三折单因素、多因素分析挑选得到的临床特征结果

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
ile40	0.04 (0.00 - 1.98)	0.005	0.04 (0.002 - 0.81)	0.035
lee10	1.12 (1.05 - 1.19)	<0.001	1.13 (1.06 - 1.20)	<0.001
sse16	1.0 (1 - 1.01)	0.019	1.004 (1.000775 - 1.0075)	0.016

表 4.5 第四折单因素、多因素分析挑选得到的临床特征结果

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
f1e8	0.07 (0.00 - 0.82)	0.034	0.03 (0.02 - 0.44)	0.01
ile40	0.07 (0.00 - 0.49)	0.008	0.05 (0.009 - 0.33)	0.001
lee11	1.11 (1.05 - 1.19)	0.001	1.10 (1.02 - 1.20)	0.02
sse33	3.86 (1.8 - 8.29)	0.001	3.65(1.59 - 8.36)	0.002

表 4.6 第五折单因素、多因素分析挑选得到的临床特征结果

变量	单因素		多因素	
	HR (CI 95%)	P	HR (CI 95%)	P
ile40	0.03 (0.02 - 0.09)	0.009	0.07 (0.0004 - 0.11)	<0.001
lee8	1.00 (1.0 - 1.01)	<0.001	1.008 (0.001 - 1.012)	<0.001
sse33	4.05 (1.83 - 8.95)	0.001	4.33 (0.001 - 11.21)	<0.001
sse49	1.00 (1.0 - 1.01)	0.010	1.007(0.001 - 1.01)	0.013
tfe10	0.05 (0.0 - 0.37)	0.014	0.05 (0.003 - 0.66)	0.02

4.2.2 生存分析实验结果

利用 Cox 模型对患者进行生存分析的结果如表 4.7 所示。在训练数据集上每一折患者生存分析的结果均在 0.7 以上，五折交叉验证的平均值为 0.74，标准差为 0.03；在测试数据集上，患者生存分析五折的 C-index 平均值为 0.70，标准差为 0.04，显示出比较稳定的结果。

表 4.7 生存分析的实验结果 (C-index)

	训练集	测试集
第一折	0.76	0.71
第二折	0.72	0.63
第三折	0.70	0.69
第四折	0.75	0.70
第五折	0.78	0.77
平均值 (标准差)	0.74 (0.03)	0.70 (0.04)

4.2.3 讨论

表 4.2-4.6 中，每一折 Cox 单因素和多因素临床特征挑选的实验结果表明，患者接受治疗后的信息同样对患者的生存预后提供了重要信息。表 4.7 中生存分析的实验结果表明，基于患者治疗前和治疗后的临床信息同样可以准确预测出患者的生存预后，并且可以更加深入地了解患者的病情和治疗效果，为医生制定更加个性化的治疗方案提供参考。这项研究对于提高局部晚期非小细胞肺癌患者的生存率和治疗效果具有重要的临床意义。

4.3 本章小结

本章主要对本文的第二个工作进行了详细介绍。首先主要介绍了本文的第二个工作所涉及到的数据，以及对数据进行的挑选部分，所涉及到的临床数据首先

是由兰大二院的医生进行筛选。然后将剩下的数据经由 Cox 单因素、和多因素分析，分五折对其进行挑选；每一折所挑选出的患者的临床特征及实验分析结果都进行了详细介绍。最后对实验的结果进行了展示和分析。

5 总结与展望

5.1 总结

本文是一项前瞻性工作研究，研究内容包括两个部分，第一部分主要是基于患者治疗后 SUVmax 预测对局部晚期非小细胞肺癌患者的生存分析，其具体流程如下：首先，由于本文所研究数据与 Machtay M^[1]等人相同，但纳入条件不同，因此需进一步在本文所纳入的数据集上进一步进行实验验证；实验验证结果显示在对治疗后肿瘤 SUVmax 采用 5.0 的二值截取值时对患者生存预后的影响是显著的 ($P < 0.05$)；得出结论，对治疗后肿瘤 SUVmax 使用 5.0 的二值截取值时对临床决策具有一定的价值；之后对其进行 Cox 单因素和多因素分析，结果显示，治疗后肿瘤 SUVmax 的二值化标签对患者的生存预后具有显著的影响。其次，使用患者接受治疗前的 FDG-PET 直接预测患者接受治疗后的 SUVmax 是否大于 5.0，采用的是 3D DenseNet169 对其进行预测，模型的训练和参数挑选由同时包含治疗前和治疗后 FDG-PET 的 143 名患者进行，然后在仅包含治疗前 FDG-PET 的 40 名患者上进行预测。最后采用上一步预测出的患者治疗后 SUVmax 的二值化标签对患者进行生存分析，结果显示，在训练集中加上治疗后 SUVmax 的二值化标签后，患者生存分析的准确性即 C-index 由 0.59 提升到了 0.66，且 P 值为 0.002，说明准确性的提升并不是由偶然性引起的，证明了加上患者治疗后 SUVmax 的二值化标签后对患者的生存分析有一定的帮助。在测试集中，加上预测出的治疗后 SUVmax 的二值化标签后患者生存分析的准确性由 0.59 提升到了 0.68，且 P 值为 0.013，证明了结果的提升并不是由偶然性引起的，也进一步说明了模型预测结果的准确性。最后得出结论：基于患者接受治疗前的 PET 预测出的患者接受治疗后的肿瘤 SUVmax 是否大于 5.0，对预测局部晚期非小细胞肺癌患者的总生存期具有一定的价值，且对临床决策具有一定的帮助。

在基于患者治疗前的 FDG-PET 的 SUV 图像对患者进行治疗后 SUVmax 是否大于 5.0 的预测时，采用的是患者的全身 PET 图像。在测试集上对模型预测出的结果进行评价时，单因素、多因素以及 Cox 回归分析均显示预测出的结果有

较高的准确性。说明了在没有医生勾画肿瘤金标准的基础上，对患者的影像数据采用阅片的形式，可以相对准确的预测出患者接受治疗后的肿瘤 SUVmax 是否大于 5.0，也在一定程度上减轻了医生的工作量，减少了医生勾画肿瘤金标准的时间。同时也进一步验证了，局部晚期非小细胞肺癌患者在进行了治疗前 FDG-PET/CT 筛查后，可以根据其相应的治疗前 FDG-PET 的 SUV 图像预测出其接受同步放化疗后的肿瘤 SUVmax 是否大于 5.0，从而预测出患者在接受治疗后的生存状况，对患者的生存预后具有重要指导性意义。

本文的第二部分工作是基于患者治疗前和治疗后的临床特征对患者进行生存分析。其具体流程如下：首先是基于 Cox 单因素多因素分析对患者的临床特征进行挑选，其次利用临床特征采用 Cox 回归分析对患者进行生存分析。因为本实验采用的是五折交叉验证，而且患者临床特征是基于训练数据集进行挑选的，所以，临床特征选择的这一过程要经过五次，最终挑选出五组不同的特征用于之后的生存分析。然后基于单因素和多因素分析挑选出来的五组特征对五折测试数据集的特征进行选择，最后基于挑选出的特征，用训练数据集对 Cox 模型进行训练，测试数据集用来进行测试。最终的结果显示，生存分析五折交叉验证结果 C-index 的平均值为 0.70，标准差为 0.04。结果表明，采用患者治疗前和治疗后的临床信息来对患者进行生存预后的效果比较稳定。

5.2 展望

虽然本文在基于患者治疗前和治疗后的信息对患者进行生存分析时取得了一些相对比较好的结果，但是，在本文的研究中还存在着一些不足之处，现主要对本文两个工作所存在的不足之处以及对未来的展望进行详细描述。

本文的第一个工作在进行研究时所采用的影像数据只有患者的 FDG-PET 图像，患者的 CT 图像同样可以提供患者的一些解剖结构信息对患者治疗后肿瘤 SUVmax 的预测进行一些辅助，但是由于数据配准的原因，并未采用患者 CT 部分的图像数据来对本文的研究任务进行辅助。在之后的工作中可以考虑对患者的 PET 和 CT 图像进行配准，从而充分利用患者的影像数据，使得分类网络可以提

取到更多相应的特征。此外，在进行预测时只采用了患者的影像数据，患者治疗前的临床数据并未应用到患者治疗后 SUVmax 的预测中，并且患者在接受治疗前的一些信息可能会影响到患者接受治疗后的临床状态以及恢复情况，因此，在之后的工作中加入患者的临床信息来进行患者治疗后 SUVmax 的预测也是将来需要考虑的一部分内容。

本文的第二个工作在对患者进行生存分析的研究时，只采用了患者接受治疗前和治疗后的临床数据集，采用的是传统的生存分析方法，并未利用患者的医学影像数据。在之后的工作中可以考虑进一步利用到患者治疗前和治疗后的影像数据，并采用深度学习基于端到端或者非端到端的方法，提取到患者的一些影像特征，从而进一步提升患者生存预后的准确性。但是在利用患者治疗前和治疗后的医学图像数据时，需要考虑到患者治疗前和治疗后 FDG-PET 图像的使用，以及配准问题，这将是本工作进一步研究的重点问题。

18F-FDG PET/CT 是在评估癌症患者时常用的一种成像方式^[54]。并且随着现代计算机科学技术的发展，患者的医学图像数据也将会在生存分析的研究中显示出越来越重要的作用，未来也必定会有更多关于患者生存分析方面的研究。

参考文献

- [1] Machtay, Mitchell & Duan, Fenghai & Siegel, Barry & Snyder, Bradley & Gorelick, Jeremy & Reddin, Janet & Munden, Reginald & Johnson, Douglas & Wilf, Larry & Denittis, Albert & Sherwin, Nancy & Cho, Kwan & Kim, Seok-Ki & Videtic, Gregory & Neumann, Donald & Komaki, Ritsuko & Macapinlac, Homer & Bradley, Jeffrey & Alavi, Abass. Prediction of survival by [18F] fluorodeoxyglucose positron emission tomography in patients with locally advanced non-small-cell lung cancer undergoing definitive chemoradiation therapy: results of the ACRIN 6668/RTOG 0235 trial [J]. *Journal of Clinical Oncology : official journal of the American Society of Clinical Oncology*, 2013, 31(30): 3823-3830.
- [2] D. R. Cox, Regression models and life-tables, *Journal of the Royal Statistical Society [J]. Series B : Methodological*, 1972, 34 (2): 187–220.
- [3] TA Gerds, MW Kattan, M Schumacher, and C Yu. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring [J]. *Statistics in Medicine*, 2013, 32 (13): 2173-2184.
- [4] Wolbers M, Blanche P, Koller MT, Witteman JC, Gerds TA. Concordance for prognostic models with competing risks [J]. *Biostatistics*, 2014, 15(3): 526-39.
- [5] Latouche A, Andersen PK, Rey G, Moreno-Betancur M. A Note on the Measurement of Socioeconomic Inequalities in Life Years Lost by Cause of Death[J]. *Epidemiology*, 2019, 30(4): 569-572.
- [6] Blanche P, Kattan MW, Gerds TA. The c-index is not proper for the evaluation of t -year predicted risks [J]. *Biostatistics*, 2019, 20(2): 347-357.
- [7] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022 [J]. *CA Cancer J Clin*. 2022, 72(1): 7-33.
- [8] Anna-Maria Shiarli, Priyanka Patel, Fiona McDonald. Reference Module in Biomedical Sciences[M]. Academic Press,2021, 810-824.

- [9] Mäurer, M., Käsmann, L., Fleischmann, D.F. et al. PET/CT-based adaptive radiotherapy of locally advanced non-small cell lung cancer in multicenter yDEGRO ARO 2017-01 cohort study [J]. 2022, 17(1): 29.
- [10] Grootjans W, de Geus-Oei LF, Bussink J. Image-guided adaptive radiotherapy in patients with locally advanced non-small cell lung cancer: the art of PET. *Q J Nucl Med Mol Imaging*. 2018 Dec;62(4):369-384. doi: 10.23736/S1824-4785.18.03084-4. Epub 2018 Jun 4. PMID: 29869486.
- [11] Truong MT, Viswanathan C, Erasmus JJ. Positron emission tomography/computed tomography in lung cancer staging, prognosis, and assessment of therapeutic response [J]. *J Thorac Imaging*, 2011, 26(2):132-46.
- [12] Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, Bruno DS, Chang JY, Chirieac LR, D'Amico TA, DeCamp M, Dilling TJ, Dowell J, Gettinger S, Grotz TE, Gubens MA, Hegde A, Lackner RP, Lanuti M, Lin J, Loo BW, Lovly CM, Maldonado F, Massarelli E, Morgensztern D, Ng T, Otterson GA, Pacheco JM, Patel SP, Riely GJ, Riess J, Schild SE, Shapiro TA, Singh AP, Stevenson J, Tam A, Tanvetyanon T, Yanagawa J, Yang SC, Yau E, Gregory K, Hughes M. Non-Small Cell Lung Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology [J]. *J Natl Compr Canc Netw*. 2022, 20(5): 497-530.
- [13] van Elmpt W, Ollers M, Dingemans AM, et al. Response assessment using 18F-FDG PET early in the course of radiotherapy correlates with survival in advanced-stage non-small cell lung cancer [J]. *J Nucl Med*, 2012, 53(10): 1514–20.
- [14] Brodin NP, Tomé WA, Abraham T, Ohri N. 18F-Fluorodeoxyglucose PET in Locally Advanced Non-small Cell Lung Cancer: From Predicting Outcomes to Guiding Therapy [J]. *PET Clin*, 2020, 15(1): 55-63.
- [15] Cheng, X., Zhang, L., & Zheng, Y. Deep similarity learning for multimodal medical images [J]. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2018, vol.6(3): 248-252.

- [16] Qayyum, A., Anwar, S. M., Awais, M., & Majid, M. Medical image retrieval using deep convolutional neural network. *Neurocomputing* [J]. *Neurocomputing*, 2017, vol.266: 8-20.
- [17] R. Dey, Z. Lu and Y. Hong. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) [M], Washington, DC, USA, 2018. 4, 774-778.
- [18] Tirupal, T., Mohan, B. C., & Kumar, S. S. Multimodal medical image fusion techniques – A review [J]. *Current Signal Transduction Therapy*, 2021, vol.16(2): 142-163.
- [19] Guida, Carmine & Shan, Juan & Zhang, Ming. Knee Osteoarthritis Classification Using 3D CNN and MRI [J]. *Applied Sciences*, 2021, vol.11(11): 5196.
- [20] Suryakanth, B. & Prasad, S. 3D CNN-Residual Neural Network Based Multimodal Medical Image Classification [J]. *WSEAS TRANSACTIONS ON BIOLOGY AND BIOMEDICINE*, 2022, 19: 204-214.
- [21] Kr, Kruthika & Rajeswari, & Maheshappa, H.D. CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis. *Informatics in Medicine Unlocked* [J]. *Informatics in Medicine Unlocked*, 2019, vol.16.
- [22] Ker J, Singh SP, Bai Y, Rao J, Lim T, Wang L. Image Thresholding Improves 3-Dimensional Convolutional Neural Network Diagnosis of Different Acute Brain Hemorrhages on Computed Tomography Scans [J]. *Sensors (Basel)*, 2019, 19 (9): 2167.
- [23] Shaish, Hiram & Mutasa, Simukayi & Makkar, Jasnit & Chang, Peter & Schwartz, Lawrence & Ahmed, Firas. Prediction of Lymph Node Maximum Standardized Uptake Value in Patients With Cancer Using a 3D Convolutional Neural Network: A Proof-of-Concept Study [J]. *American Journal of Roentgenology*, 2018, 212(2). 1-7.
- [24] Kang G, Liu K, Hou B, Zhang N. 3D multi-view convolutional neural networks for lung nodule classification [J]. *PLoS One*. 2017, 12 (11):e0188290.

- [25] Passa, Marco & Mandelli, Maria & Cobigo, Yann. Transfer learning in 3D-CNN increases classification of AD and no-AD in Primary Progressive Aphasia [J]. 2020.
- [26] Ko, Jessica. Solving the Cox Proportional Hazards Model and Its Applications [Z]. San Francisco, USA: EECS Department, University of California, Berkeley. May, 2017.
<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-110.html>
- [27] Sisay G, Mantefardo B, Beyene A. Time from symptom onset to severe COVID-19 and risk factors among patients in Southern Ethiopia: a survival analysis [J]. J Int Med Res. 2022, 50(8): 3000605221119366.
- [28] Tian, Shu & Li, Qin & Li, Ruichen & Chen, Xinyu & Tao, Zhonghua & Gong, Hongli & Wang, Xiaoshen & Hu, Xichun. Development and Validation of a Prognostic Nomogram for Hypopharyngeal Carcinoma [J]. Frontiers in Oncology. 2021,11:696952.
- [29] Feroz, Zainab & Gautam, Priyanka & Tiwari, Sonia & Shukla, Girish & Kumar, Munish. Survival analysis and prognostic factors of the carcinoma of gallbladder [J]. World Journal of Surgical Oncology, 2022, 20(1):403.
- [30] Molina-Fernández E, Palacios-García JM, Moreno-Luna R, Herrero-Salado T, Ventura-Díaz J, Sánchez-Gómez S, Vilches-Arenas Á. Survival Analysis in Patients with Laryngeal Cancer: A Retrospective Cohort Study [J]. Life (Basel). 2023, 13(2):295.
- [31] Baralo B, Schneider M, Baralo I. Survival analysis of small cell carcinomas of the genitourinary system [J]. Proc (Bayl Univ Med Cent). 2022, 36(1): 8-14.
- [32] Vargas López AJ, Fernández Carballal C, Valera Melé M, Rodríguez-Boto G. Survival analysis in high-grade glioma: The role of salvage surgery [J]. Neurologia (Engl Ed), 2023, 38(1):21-28.
- [33] Kumar, D. & Noushad, Navin & Viswanathan, Mega & Sharma, Ajay. Endometrial Carcinoma: Survival Analysis at a Tertiary Care Center in India [J].

- Indian Journal of Gynecologic Oncology.2022, 20(4):69.
- [34] Chernova, Galina & Kozlov, Vadim & Gulyaeva, Lyudmila. (2022). Survival analysis of patients with non-small cell lung cancer in Novosibirsk region from 2015 to 2019 [J]. Almanac of Clinical Medicine. 2022, 50 (1): 56-64.
- [35] Zheng, Rongshou & Zhang, Siwei & Zeng, Hongmei & Wang, Shaoming & Sun, Kexin & Chen, Ru & Li, Li & Wei, Wenqiang & he, Jie. Cancer incidence and mortality in China, 2016 [J]. Journal of the National Cancer Center, vol.2(1): 1-9.
- [36] Xia C, Dong X, Li H, Cao M, Sun D, He S, Yang F, Yan X, Zhang S, Li N, Chen W. Cancer statistics in China and United States, 2022: profiles, trends, and determinants [J]. Chin Med J (Engl). 2022, 135(5): 584-590.
- [37] Chhikara, Bhupender & Parang, Keykavous. Global Cancer Statistics 2022: The Trends Projection Analysis [J]. Chemical Biology Letters, 2022, vol.10(1): 451.
- [38] Y. Wu, J. Ma, X. Huang, S. H. Ling and S. Weidong Su. DeepMMSA: A Novel Multimodal Deep Learning Method for Non-small Cell Lung Cancer Survival Analysis [J]. 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Melbourne, Australia. 2021.10, 1468-1472.
- [39] SEER Training Modules, Lung Cancer. U. S. National Institutes of Health, National Cancer Institute [Z]. 2023/02/09. <<https://training.seer.cancer.gov/lung/intro/>>.
- [40] 中华医学会肿瘤学分会,中华医学会杂志社.中华医学会肿瘤学分会肺癌临床诊疗指南(2021 版)[J].中华肿瘤杂志,2021,43(06):591-621.
- [41] Morgensztern D, Ng SH, Gao F, et al. Trends in stage distribution for patients with non-small cell lung cancer: a National Cancer Database survey [J]. J Thorac Oncol 2010;5:29-33.
- [42] Brierley JD, Gospodarowicz MK, Wittekind C. UICC TNM classification of malignant tumours, 8th ed [M]. OxfordHoboken: Wiley, 2017, 106-12.
- [43] Huber RM, De Ruyscher D, Hoffmann H, et al. Interdisciplinary multimodality management of stage III non-small cell lung cancer [J]. Eur Respir Rev

- 2019, .28(152):190024.
- [44] Muriana, Piergiorgio & Perroni, Gianluca & Novellis, Pierluigi & Veronesi, Giulia. Robotic surgery for locally advanced non-small cell lung cancer [J]. Journal of Visualized Surgery, 2021, Vol.7: 26,
- [45] Prasad, Rahul & Williams, Terence. A narrative review of toxicity of chemoradiation and immunotherapy for unresectable, locally advanced non-small cell lung cancer [J]. Translational Lung Cancer Research. 2020, 9(5): 2040-2050.
- [46] F.C. Detterbeck, D.J. Boffa, A.W. Kim, L.T. Tanoue. The eighth edition lung cancer stage classification [J]. Chest, 2017, 151(1): 193-203.
- [47] Wang, Yimeng & Wang, Yao & Yu, Jinming & Meng, Xiangjiao. The treatment in patients with unresectable locally advanced non-small cell lung cancer: Explorations on hot issues [J]. Cancer Letters, 2022, 551: 215947.
- [48] Kapoor V, McCook BM, Torok FS. An introduction to PET-CT imaging [J]. Radiographics. 2004, 24(2):523-43.
- [49] Drzezo. Introduction to FDG PET/CT [Z]. Radiology Key. Jun 18, 2019. <https://radiologykey.com/introduction-to-fdg-petct/>.
- [50] Cotran RS, Kumar V, Collins T. Neoplasia. In:Cotran RS, Kumar V, Collins T, eds. Robbinspathologic basis of disease [J]. Philadelphia, Pa:Saunders, 1999, 260-327.
- [51] Devaki Shilpa Surasi, Pradeep Bhambhvani, Jon A. Baldwin, Samuel E. Almodovar, and Janis P. O ' Malley. 18F-FDG PET and PET/CT Patient Preparation: A Review of the Literature [J]. Journal of Nuclear Medicine Technology, 2014, 42 (1) 5-13.
- [52] Higashi K, Ueda Y, Arisaka Y, et al. 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resected non-small cell lung cancer [J]. J Nucl Med. 2002;43(1):39-45.
- [53] Tixier F, Hatt M, Le Rest CC, et al. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET

- [J]. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. 2012, 53(5):693-700.
- [54] Kuitunen, Ilari & Nikkilä, Atte & Ponkilainen, Ville & Uimonen, Mikko & Lohi, Olli. Survival Analysis and Cox Proportional Hazards Model Reporting in Pediatric Leukemia Studies-a Systematic Review [J]. *SN Comprehensive Clinical Medicine*. 5, 24 (2023).
- [55] Kaplan, E. & Meier, Paul. Nonparametric Estimation From Incomplete Observations [J]. *Journal of American Statistical Association*, 1958, 53:457-481.
- [56] Famularo, Simone & Bernasconi, Davide Paolo. Survival Analysis [M]. 2022.12, 89-108.
- [57] Su, Pei-Fang & Lin, Chou-Ching & Hung, Jo-Ying & Lee, Jung-Shun. (2022). The Proper Use and Reporting of Survival Analysis and Cox Regression [J]. *World Neurosurgery*. 161: 303-309.
- [58] Ko, Jessica. Solving the Cox Proportional Hazards Model and Its Applications[D]. San Francisco, USA: EECS Department, University of California, Berkeley. May, 2017.
<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-110.html>
- [59] P. McCullagh and J. A. Nelder. Generalized Linear Models [EB/OL]. 2019.1, <https://doi.org/10.1201/9780203753736>.
- [60] Norman Breslow. Covariance analysis of censored survival data [J]. *Biometrics*, 1974, 30(1): 89-99.
- [61] Yao, Jiawen & Zhu, Xinliang & Jonnagaddala, Jitendra & Hawkins, Nicholas & Huang, Junzhou. Whole Slide Images based Cancer Survival Prediction using Attention Guided Deep Multiple Instance Learning Networks [J]. *Medical Image Analysis*, 2020, 65:101789.
- [62] Kinahan, P., Muzi, M., Bialecki, B., Herman, B., & Coombs, L. Data from the ACRIN 6668 Trial NSCLC-FDG-PET [DB/OL]. The Cancer Imaging Archive, 2019, <https://doi.org/10.7937/tcia.2019.30ilqfcl>.

- [63] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository [J]. *Journal of Digital Imaging*, 2013, 26(6): 1045-1057.
- [64] Pieper, Steve & Halle, M. & Kikinis, Ron. 3D slicer [J]. *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2014, vol.1:632-635.
- [65] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks[J]. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, 2261-2269.
- [66] 阿斯顿·张, 李沐, 扎卡里·C.立顿, 亚历山大·J.斯莫拉. 动手学深度学习[M]. 北京: 人民邮电出版社: 2019.6, 40-41.

致 谢

时光匆匆，转瞬即逝。眨眼间，三年的硕士研究生生涯即将结束，在开始编写致谢之后，才深刻感觉到此刻的自己是真的要毕业了，即将迈出新的一步，踏入新的生活。回顾从确定被学校录取之后到现在时光，中间经历了许许多多，又不知从何说起。想起确定被学校录取的那天，当时只觉得我的考研之路没有白走、努力没有白费，但就是当时考研的选择让我能够有机会扩展自己的眼界、改变了自己的人生路。感谢当时坚持选择考研的自己，也感谢祖国以及母校对我的培养！

认真回想起三年的学习生涯，离不开我的导师丁晓阳教授的谆谆教诲和栽培。在我初入校园时是丁老师和蔼可亲的态度让我倍感亲切，削减了我从小以来对老师就有的莫名恐慌；在我遇到困难时，也是丁老师及时提供的帮助让我在学校时也会有种家的感觉；在我不知所措时，同样是丁老师给我指明了前进的方向。在学校里，丁老师给了我许多的指导和帮助，同样丁老师也以自己的行动向我们展示了他的努力和付出，让我受到了鼓舞，使我能够保持积极向上、不断进取的学习和生活态度。同时也感谢丁老师对我们实验室所提供的支持和帮助。由衷的感谢丁老师，祝愿丁老师身体健康，万事如意！

感谢我的另一位老师何江萍教授，在实验室的时光都是何老师时时刻刻陪伴着我们实验室的所有同学，同我们一起学习，一起进步。正是因为何老师，我们才能有在实验室进行学习和研究生活的机会；也是因为何老师，让我们在实验室的时光才不会那么枯燥；同样是何老师，在三年的研究生期间，我们才能够有外出开会，扩展眼界的机会。学习中，离不开何老师对我们的引导，在科研中遇到难题和瓶颈时，何老师总能为我提供一些新的思维角度；在生活中遇到问题时，何老师也总是给我们提供帮助。再次感谢何老师，祝愿何老师阖家欢乐，万事胜意！

同样也感谢从进入实验室开始就一直给我们提供帮助的李兵老师以及李子荣老师，李兵老师认真教学的态度和李子荣老师的认真指导让我收益颇多，祝愿两位李老师身体健康，心想事成！同时也要感谢实验室的师兄弟以及师姐师妹们，

感谢他们在我遇到困难时的认真解答以及提供的帮助，在与他们探讨问题以及共同进步的时光里，使我学到了许多。感谢研究生三年和我一块儿生活的舍友们，是她们在我迷茫和低落的时候给我鼓励，让我能够有坚持下去的信念！最后感谢我的父母，是他们让我能够在学校里无忧无虑的学习，是他们默默无闻的付出我才能够在学校里心无旁骛的完成我的学业！

以后的人生道路还很漫长，未知的困难和未来正在等着我们，但是追求和梦想会给予我们不断前进的动力，让我们能够在以后的道路上勇往直前！