



南京大學

研究生毕业论文 (申请硕士学位)

论文题目 基于 GAN 生成病灶的三维医疗图像扩增技术
作者姓名 侯韵晗
专业学位类别(领域) 电子信息(软件工程领域)
研究方向 软件工程
指导教师 陈振宇 教授

2022 年 5 月 25 日

学 号： MF20320064

论文答辩日期： 2022 年 5 月 20 日

指导教师： (签字)

3D Medical Image Augmentation Technology Based on GAN Generating Lesions

by

Yunhan Hou

Supervised by

Professor Zhenyu Chen

A dissertation submitted to
the graduate school of Nanjing University
in partial fulfilment of the requirements for the degree of
MASTER OF ENGINEERING
in
Software Engineering



Software Institute
Nanjing University

May 25, 2022

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目： 基于 GAN 生成病灶的三维医疗图像扩增技术

工程硕士（软件工程领域） 专业 2020 级硕士生姓名：侯韵晗
指导教师（姓名、职称）： 陈振宇 教授

摘 要

近年来，深度学习在图像领域的应用越来越广泛。特别是在医疗领域，随着国家对“智慧医疗”扶持力度的增大以及医疗机构对计算机辅助诊断（CAD）技术需求的扩大，深度学习模型被大量应用于病灶识别和诊断决策。研究证明，深度学习在医疗目标检测、分割、诊断等方面都具有出色表现。但与其他领域的深度学习网络相同，这些网络的准确度也依赖于样本数据集的质量。目前，医疗影像数据集规模小、质量参差不齐，医疗诊断模型准确度不理想。隐私保护、数据安全等因素也加剧了医疗数据集的共享难度。因此需要探索其他途径来获取医疗数据样本。

为解决样本匮乏的问题，研究者们提出通过数据扩增来增加样本数量。随着技术的发展，越来越多的医疗目标检测开始选用三维图像作为样本输入。相对于二维医疗图像，三维医疗图像能更好地结合人体冠状面、矢状面和横断面的三维信息。但数据类型的升级也给数据扩增带来了新的难题，针对三维图像的扩增方法亟待研究。

针对数据集不足的问题，本文使用 GAN（3D-DCGAN）进行病灶样本生成，实现数据扩增。本文选取肺结节作为研究对象，从 LIDC-IDRI 数据集中获取原始的胸部 CT 影像。医疗数据具有其独特的信息存储格式，所以针对原始数据集进行信息提取和影像预处理也是本文工作的关键一环。本技术通过结节信息处理、样本获取、卷积生成对抗网络生成病灶等流程实现了肺结节病灶样本的生成。同时，本文使用基于模型训练准确度对比的方法对生成的病灶进行了质量评估验证，证实了本文所提出的技术能够有效的扩充数据集，改善样本分布不均衡的问题，提升训练模型的准确度。

关键词： 三维医疗图像生成；数据扩增；卷积生成对抗网络；肺结节检测

南京大学研究生毕业论文英文摘要首页用纸

THESIS: 3D Medical Image Augmentation Technology Based on
GAN Generating Lesions

SPECIALIZATION: Software Engineering

POSTGRADUATE: Yunhan Hou

MENTOR: Professor Zhenyu Chen

Abstract

In recent years, deep learning has been applied more and more widely in the field of image recognition. Especially in the medical field, a large number of CAD(Computer-Aided Diagnosis) systems have been developed and put into practice due to the country's strong support for WITMED (Wise Information Technology of Med) and the strong demand of medical institutions for computer-aided Diagnosis technology. Studies have proved that deep learning has excellent performance in medical object detection, segmentation, diagnosis, and other tasks. However, the accuracy of these deep learning networks also depends on the quality and quantity of training sample data sets. In the medical field, privacy protection, data security and other problems aggravate the difficulty of data sharing to a certain extent. The detection accuracy of many CAD systems needs to be improved. Therefore, other ways to obtain medical data samples need to be explored.

To solve the problem of sample shortage, the researchers propose to increase the sample size through data augmentation. With the development of technology, more and more medical targets (such as lung nodules) detection is no longer based on 2D images, but using 3D images as input. Compared with 2D medical image, the 3D medical image can better combine 3D information of the coronal plane, sagittal plane and cross section of the human body. The upgrade of data types also brings new problems to data augmentation, and it is urgent to study the augmentation methods for 3D images.

To solve the problem of insufficient medical data set, 3D-DCGAN is used in this paper to generate lesion samples and realize data augmentation. Lung nodules are selected as the research object, and the LIDC-IDRI dataset is selected as the original data

set. Because of the unique information storage format of medical data, it is necessary to extract information and preprocess images from the original data set. In general, the project achieved the generation of lung nodular lesion samples through nodule information processing, sample acquisition, and lesion generation (using 3D-DCGAN). At the same time, the quality of generated lesions was evaluated based on the method of model training accuracy comparison, which confirmed that the technology proposed in this paper can effectively augment the data set, solve the problem of unbalanced sample distribution, and improve the accuracy of the training model.

keywords: 3D Medical Image Generation, Data Augmentation, DCGAN, Lung Nodule Detection

目 录

目 录	v
插图清单	vii
附表清单	ix
第一章 绪论	1
1.1 研究背景与意义	1
1.2 国内外研究现状	3
1.3 研究目标及主要研究工作	5
1.4 本章小结	6
第二章 相关理论综述	7
2.1 生成对抗网络技术	7
2.1.1 简介	7
2.1.2 算法原理	8
2.1.3 医疗领域的 GAN 生成	12
2.2 激活函数	14
2.3 损失函数	16
2.4 通用图像扩增技术	18
2.5 本章小结	19
第三章 病灶数据分析	21
3.1 医疗病灶分析	21
3.2 影像数据分析	25
3.3 本章小结	28
第四章 技术关键算法	29
4.1 结节信息提取算法	32
4.2 样本获取算法	36
4.3 通用图像扩增算法	41
4.3.1 旋转	42
4.3.2 翻转	44

4.3.3 亮度	45
4.3.4 噪声	46
4.4 3D-DCGAN 生成结节算法	50
4.4.1 卷积生成对抗网络技术	50
4.4.2 判别器实现算法	52
4.4.3 生成器实现算法	53
4.4.4 判别器、生成器的博弈	54
4.4.5 损失函数	55
4.5 本章小结	56
第五章 生成病灶及质量评估实验	59
5.1 3D-DCGAN 网络训练	59
5.1.1 实验环境	59
5.1.2 模型训练与病灶生成	60
5.2 质量检测手段	63
5.3 质量验证实验	65
5.3.1 研究问题	65
5.3.2 实验设计	65
5.3.3 评估指标	67
5.3.4 实验结果	68
5.4 本章小结	71
第六章 总结与展望	73
致 谢	75
参考文献	77
简历与科研成果	85

插图清单

2-1 GAN 结构原理示意图	7
2-2 GAN 基本思路示意图	8
2-3 Sigmoid 函数	14
2-4 Tanh 函数	15
2-5 ReLU 函数	16
2-6 Leaky ReLU 函数	16
2-7 通用图像扩增技术	19
3-1 2020 年全球癌症发生率与死亡率占比。A: 全性别癌症发生率与 死亡率。B: 男性癌症发生率与死亡率。C: 女性癌症发生率与 死亡率	23
3-2 肺结节样式	24
3-3 胸部检查影像	25
3-4 LIDC-IDRI 的 xml 文件结构	28
4-1 本文技术结构图	29
4-2 某检查序列的文件夹示例图	33
4-3 3000611 检查中的多个结节	33
4-4 xml 文件详情	34
4-5 结节信息读取代码	36
4-6 人体基本面	37
4-7 Dicom 坐标系	38
4-8 三维重建实现代码	40
4-9 获取的原始样本	41
4-10 绕坐标轴旋转	42
4-11 旋转实现代码	43
4-12 肺结节样本旋转变换	44
4-13 翻转实现代码	45

4-14 肺结节样本翻转变换	45
4-15 亮度变换实现代码	46
4-16 肺结节样本亮度变换	46
4-17 肺结节样本添加高斯噪声	47
4-18 肺结节样本添加椒盐噪声	48
4-19 肺结节样本添加泊松噪声	48
4-20 噪声实现代码	49
4-21 针对肺结节生成网络的判别器	55
4-22 针对肺结节生成网络的生成器	55
4-23 网络结构	57
5-1 DCGAN 网络训练-判别器损失函数	60
5-2 DCGAN 网络训练-生成器损失函数	61
5-3 使用模型生成肺结节	61
5-4 生成结节-7500 次迭代	62
5-5 生成结节-15000 次迭代	62
5-6 生成结节-20000 次迭代	63
5-7 混淆矩阵	67
5-8 训练 3D-CNN-Acc 曲线对比	69
5-9 训练 3D-CNN-Loss 曲线对比	70
5-10 训练 3D-CNN-ROC 曲线对比	70

附表清单

3-1 LIDC-IDRI 数据集基本信息.....	26
3-2 dicom 文件重要数据列表.....	27
4-1 结节信息	35
5-1 实验环境	59
5-2 DCGAN 模型相关参数	60
5-3 质量度量指标	64
5-4 实验数据集设计	66
5-5 验证实验参数	66
5-6 验证实验结果	69

第一章 绪论

1.1 研究背景与意义

近年来，人工智能领域发展迅猛，众多深度学习模型的准确度与鲁棒性都有着出色表现，因此也在生活生产的各个领域得到了广泛应用。同时，广泛的实际应用也相应的启发和促进着领域化模型的算法研究，研究者们有针对性地对于不同领域和不同场景提出了一系列的深度学习网络 [1]。但在实际应用当中，深度学习模型存在的不足也被一一暴露出来。其中尤为突出的问题是，一些高精度领域往往对模型的质量与性能有着较为苛刻的要求，但模型受算法、样本等一系列客观因素的影响，往往难以满足这样的高精度要求，深度学习模型的质量提升遭遇瓶颈。因此在一些情况下，人工智能不但没有如愿起到辅助作用为人们带来便利，反而增加了人力物力的重复消耗，甚至会引导使用者走向错误的方向，得到与事实相反的结论。

医疗行业是一个典型的重度依赖计算机辅助，同时又对辅助系统的精度要求较高的领域。事实上，基于深度学习的医疗影像分析近年来也一直以来也是国家关注的重点研究领域。在 2021 年，我国印发了《公立医院高质量发展促进行动（2021-2025 年）》文件，提出重点建设“三位一体”智慧医院，利用互联网、人工智能和大数据等信息化技术，加强医疗健康与互联网的融合程度，促进医疗现代化进程的进展。可见当前医疗行业进行信息智能化发展的紧迫性。

在医疗领域，从智能提取病灶特征，到病灶分割分类，再到疾病预测诊断，医疗领域基于人工智能诊断方法的升级带来了越来越智能的应用。医疗领域的图像识别应用大致可以分为四类 [2]：组织分割，例如血管分割、脑区划分；病灶识别，其中包含简单的区域识别以及精确的轮廓识别；病灶分类，对于识别出的病灶进行分类；病症综合诊断，根据病灶特点，例如：大小、形态、生长速度、数量、位置等多方位特征进行病症诊断。病症综合诊断这些高级智能化功能的准确度与可靠性严重依赖于病灶分割与分类的质量，因此保证这些基础功能的准确与可靠非常重要。

虽然目前深度学习已经能够在通用图像识别与分类方面达到较高的准确度，但是，在医学领域，医学图像识别和分类的精准度仍然不尽如人意，模型质量的提升遭遇瓶颈。究其原因，主要是缺乏高质量的训练数据集 [3]，而深度学习往往又要求具有充足且高质量的训练数据。研究证明，深度学习模型质量与数据及质量密切相关，只有保证模型用充足的、高质量的数据进行学习，才能够得到质量高的优秀模型 [4]。相较于通用领域，医疗领域由于受到多方面客观因素的影响，其满足标准的数据更加匮乏，数据缺乏问题更加严重 [5]。首先，由于隐私伦理的原因，医学图像的传播必须要考虑到患者的隐私保护问题，必须以尊重患者意愿为前提。在一系列标准的约束下，临床数据共享困难。其次，基层医院信息化程度不同，设备先进性不同，许多医院共享渠道不畅通，信息技术不支持，导致有效样本数量少，即使具有有效样本，共享效率也非常低，或是数据信息闭塞，缺少渠道进行共享。并且，医学图像要成为能够进行训练的样本，必须经过数据标注。普通行业的日常图片的标注可以通过众包等方式来方便的完成，而医学图片的解读对标注者的专业性要求非常高，标注工作只能依靠经验丰富的专家来完成。显然，满足要求的专家数量并不多。即使能够抽调一些专家进行标注，也将耗费大量的人力物力财力，这些成本也将成为降低医疗机构安排标注活动积极性的因素，最终形成恶性循环。因此，在众多因素的加持下，具有极强领域特殊性的医疗图像样本面临着数量少质量低的问题。现有医疗数据集不能满足模型训练的需要，影响了深度学习在医疗图像分析上的应用。因此，医疗领域中有效样本缺少的问题亟待解决。所以，本文借鉴了通用领域的样本数据处理方法，集合医疗领域图像的特性，有针对性地对医疗图像进行数据扩增。

数据扩增是深度学习解决样本缺少问题的常用方法。它是通过轻微变换现有数据或创建新的合成图像，来得到足够应用于模型训练的数据，从而增加训练样本数据量。数据扩增的应用领域有图像扩增 [6]、文本扩增 [7]、雷达点云扩增 [8] 等等。由于不同的数据类型具有不同的特性，所以需要使用不同的扩增策略对它们进行数据变换。具体到图像扩增领域，一般可以通过仿射变换和弹性变换来进行图像变异。仿射变换包含旋转、裁剪等不改变像素值的扩增方法。弹性变换包含亮度偏移、对比度偏移等不改变像素点位置的扩增方法。同时，也有研究者提出通过合成变换，例如通过随机擦除、特征融合来生成更多样化的样本。研究者通过上述方法对二维医疗图像进行变化，证实了通过数据扩增扩充数据集对于二维医疗图像识别模型的训练有显著帮助。

随着医疗行业的发展，医护人员对计算机辅助诊断系统的需求也发生了改变。临幊上常使用电子计算机断层扫描（CT）来进行部位检查。CT 扫描可以理解为用 X 射线束对人体某部位进行成像，用一个个横截面来展示器官的情况，每个横截面之间保持一定厚度。但是，医生分析病情往往需要将矢状位、冠状位、横断面三个视角的图像结合在一起观看，常用的二维图像显然不够直观，于是三维图像展示应运而生。目前医疗机构所配备的可视化仪器基本都支持三维图像展示。因此，三维医疗图像的分割与分类也逐步进入研究者的视野，越来越多的三维图像识别模型被设计与应用 [9]。但它们同样在准确率提高方面存在瓶颈，缺乏足够的高质量三维图像样本。因此，本文着眼于三维医疗图像的数据扩增，以期增加训练样本的数量，提升模型准确度与鲁棒性。

1.2 国内外研究现状

近年来，深度学习在医疗图像识别和疾病诊断中的应用愈加广泛 [10]。其中，卷积神经网络（CNN）表现格外突出。1989 年，Lecun 等 [11, 12] 首次提出卷积神经网络，但由于当时用于训练的样本数量少且质量低，卷积网络在实际应用中表现平平。2012 年，Krizhevsky 等 [7] 提出的卷积神经网络模型 AlexNet 引起了研究者们的广泛关注。研究者开始关注卷积神经网络在图像领域的优秀表现。在医疗领域也不例外，例如 Cernazanu-Glavan 等 [13] 在 2013 年使用 CNN 实现了自动从 X-Ray 图像中分割出骨骼部分，实验表明该模型具有较高的准确率。Su 等 [14] 在 2015 年提出使用快速扫描卷积神经网络（FCNN）实现乳腺癌的区域分割，去除了原始 CNN 中的冗余计算，大大提高了分割速度。

与一般的常见图像不同的是，医疗图像中的检测目标往往较小，因此对于精度的要求就更高。小目标检测是医疗图像中重要的研究问题，比如肺结节检测和视网膜眼底病变检测。2018 年，Benzamin 等 [15] 使用卷积神经网络进行视网膜眼底病变的检测，文中使用滑窗方法（Sliding Window）将 16×16 的图像块输入一个 8 层的 CNN 网络进行二分类。肺结节检测同样也是小目标的检测，但与视网膜眼底病变不同，肺结节检测越来越多的从二维图像分析变为三维图像分析，且肺结节检测依靠灰度图像（CT/X-Ray）。

在肺结节检测研究中，Setio 等 [16, 17] 提出了多视图卷积网络（ConvNets）。该网络使用三个候选探测器识别结节，之后通过多个 ConvNets 组合减少假阳

性，ConvNets 网络处理提取的二维图像块，然后将结果输入到多种融合方法中。结果表明，该多视角卷积网络能够达到 90.1% 的检测灵敏度，鲁棒性高。随着医疗领域需求和卷积网络技术的进步，3D-CNN 应运而生并应用在医疗图像检测中，Hamidian 等 [18] 提出 DeepMed 网络，这是一种 3D-CNN 网络，它由卷积层-池化层-卷积层-池化层-卷积层-全连接层-全连接层构成，它使用 LIDC-IDRI 数据集进行训练和验证，实验证明该网络灵敏度能够达到 91%。Huang 等 [19] 也在 3D-CNN 肺结节识别方面进行了灵敏度和性能的优化。

与在通用领域一样，深度学习在医疗影像分析领域准确度的提高受到了样本数量不足的限制。医疗领域由于隐私等约束，其样本共享更加困难，所以数据扩增成为了至关重要的技术手段。Ahmet Haydar Ornek 等 [20] 使用 CNN 对新生儿热图像进行分类。他们对原始图像进行整体的图像变换，比如：旋转、翻转、缩放、锐化、亮度变化等。实验证明，通过图像变换扩增数据集，使用扩增数据集训练 CNN 网络，网络的准确率能够提升 26.29%，特异性提升了 29.09%。这也证实了，通用领域中应用的种种图像变换手段，应用在医疗领域在一定程度上同样有效。

2021 年，Hou 等开发了 TauMed[21] 系统实现对医疗二维图像的扩增。鉴于病灶区域中包含关键的特征信息，该文不仅实现了整体图像变换，更提出了病灶与非病灶区域分别处理的思路，实现了基于语义的局部图像变换，即对病灶区域进行扩张、删除、复制等变换，和对非病灶区域进行噪声添加、对比度变换等变换。实验表明，使用 TauMed 进行扩增，能够将 ResNet-50 网络的识别准确度由 96.76% 提升到 98.35%，效果优于常用的图像扩增工具库。

Aria Pezeshk 等 [22, 23] 实现了一种图像融合工具，能够将病灶区域插入一般图像，以此来实现数据扩增。实验证明，使用该工具能够生成大量图像，使用扩增数据集对分类器进行训练，可以有效解决数据缺乏的问题，提高分类器的准确度。这项研究也同样说明生成的病灶能够有合理方法与无病灶图像进行结合，侧面证明了通过病灶生成工作扩充样本进行模型训练的可行性。

Bermudez 等 [24] 在 2018 年使用使用生成对抗网络 (GAN) 实现无监督的脑部 MRI 合成。该方法仅需要很少的数据集，即可得到较多的扩充数据。让两名领域专家不知情地评估真实图像和合成图像，发现两种图像的评分相似，即专家很难判断真假图像。这项研究证明了 GAN 在使用小数据集合成医疗图像上的优秀表现，为本文进行三维医疗图像病灶生成提供了前提。

1.3 研究目标及主要研究工作

深度卷积神经网络在医疗领域应用广泛且表现优秀。例如，已有研究表明深度学习可以有效在肺部 CT[25] 和 X 光 [26] 中识别出 COVID-19。但这些模型仍然存在准确率增长瓶颈，其中最关键的诱因是训练样本规模小、质量低。为了解决医疗图像缺乏的问题，本文试图通过数据扩增手段生成高质量图像用于模型训练。另外，随着深度学习技术的快速发展和医疗领域需求逐步扩大，三维医疗图像识别技术广泛应用，模型可以识别出三维影像中的病灶辅助专家进行医疗诊断，但三维医疗图像扩增技术发展迟缓。所以本文着眼于三维医疗图像进行扩增技术的研究。

目前，已有研究针对二维医疗图像提出了语法整体变换和病灶分离的语义局部变换。实验结果表明，这两种扩增思路均能生成有效样本，使用扩增后数据集对模型进行训练，能够切实提升模型识别准确度，提高模型鲁棒性，扩增效果优于目前常用的扩增方法库。

针对三维医疗图像本文首先考虑依然沿用上述扩增思路。但与二维图像不同的是，三维图像需要多考虑 z 轴维度的影响。首先，在使用一些仿射扩增方法对三维图像进行变换时，二维图像变换中微小的误差会被再次放大。举一个直观的例子，在对病灶进行旋转、平移时，移动后会造成原位置像素空缺，需要利用算法进行填补。从二维到三维的转换，算法需要进行升级，原有误差将会扩大。所以，虽然三维仿射变换能够变化的角度方向增加，但需要考虑的图像质量问题也更加棘手，工作量增加，而造成的误差对于高敏感度的医疗图像来说无疑雪上加霜。其次，现有工作已经针对二维图像提出了例如病灶复制、病灶插入的合成扩增方法。但一直以来，针对病灶插入的研究都存在诸多问题。人体器官、组织、病灶的生长发展具有一定的规律，病灶相对于组织器官的位置往往也具有规律，而这种生长规律是深层次的、是细微的，难以用简单的算法进行模拟。因此，在对图像进行病灶插入等变换时，角度、位置的确定一直是一个难题。而在三维图像的处理中，这种难题被进一步的扩大，维度的增加为位置的确定增加新的考虑因素。所以，即使原有扩增手段可以应用在三维医疗图像变换上，也很有可能会有较大的误差，为了减少误差达到医疗领域标准，难免消耗大量人力物力。

在这样的背景下，生成式对抗网络的优势显露出来。生成式对抗网络（GAN）是一种深度学习模型，包含生成模型和判别模型两部分。这两部分通

过相互博弈不断进行学习和优化 [27]。生成对抗网络的优势在于它能够学习到图像中的细小的深度的特征，适合生成变化较小的小物体。在医疗图像中，常见的病灶、结节、细胞均属于上述类型的小物体。原本很难设定的算法与规律都可以通过 GAN 来自动发掘与学习。另外，GAN 相较于常见扩增手段具有更高的灵活性。因为通过人工手段去设定图像变化的参数很难覆盖所有情况，而 GAN 可以很好地解决这个问题。GAN 能够生成更加多样的数据，全面扩充数据集，从而提升深度学习网络的泛化能力。同时，受现实原因影响，医疗图像数据集通常存在数据类型分布不均衡的情况 [28]。对于那些非常罕见的疾病，相关病例少、检查影像少，从而导致模型训练样本少。通过生成对抗网络就可以合成这些不常见的医疗图像。

本文的研究目标就是挖掘生成对抗网络在三维医疗图像扩增方面的潜能，实现无条件的样本生成，提供便捷、有效的扩增技术。本文使用 GAN 中的 3D-DCGAN 网络生成病灶扩增数据集，并对扩增后数据集进行质量评估，以此来检验该扩增方法的效果，验证 3D-DCGAN 在三维医疗图像扩增上的优势。本文将 3D-DCGAN 生成的数据用于 3D-CNN 病灶判别模型的训练，根据训练后模型的准确率等评估指标对数据的质量和可靠性进行评估，即对本文方法的有效性进行评估。

1.4 本章小结

本章是文章绪论，首先针对深度学习领域样本缺乏问题进行介绍，尤其是在医疗领域，数据集数量少质量低的问题更加严重。训练样本的缺乏直接导致了医疗诊断模型的质量问题，许多医疗系统难以达到高精度的要求，急需针对医疗领域进行图像数据扩增方法的研究。同时，本章介绍了医疗图像识别以及医疗图像数据扩增技术的研究现状，解释了本文采用生成对抗网络实现三维医疗图像扩增的原因。最后介绍了本文的研究目标和主要研究工作，即，使用生成对抗网络实现三维医疗图像病灶的扩增。

第二章 相关理论综述

2.1 生成对抗网络技术

2.1.1 简介

生成对抗网络（Generative Adversarial Network，GAN）[29] 是 Ian J. Goodfellow 等人在 2014 年提出的一种深度学习网络。在刚被提出时，它被用做是一种非监督式的学习方法。非监督学习是机器学习中的一类方法，该类方法模型输入的训练样本是没有被标记的，学习时其样本分类的结果真伪未知，同时也未受到监督式增强，网络本身将自动从样本中探寻规则。

生成对抗网络由生成器（简称 G）、判别器（简称 D）两部分组成。生成器用来探究和学习数据的特征和分布规则，从而根据规则生成逼真的假数据。生成的数据同真实数据一同输入判别器，判别器鉴别假样本和真实数据。二者通过交互，不断博弈，调整二者相关参数。在这个过程中，生成器试图找到更加真实的数据分布，生成更加逼真的数据，而判别器会逐渐难以区分生成器生成的数据和真实的样本数据。GAN 结构如图 2-1 所示。

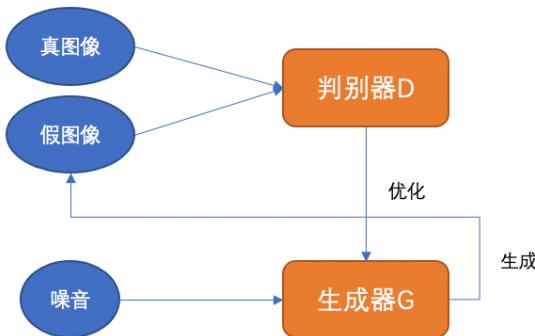


图 2-1: GAN 结构原理示意图

生成对抗网络中生成器（G）和判别器（D）的博弈可以用下列公式描述：

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(D(x))] + \mathbb{E}_{z \sim p_z(z)}[1 - \log(D(G(z)))] \quad (2-1)$$

其中， z 代表噪声， $p_z(z)$ 代表噪声分布， x 代表真实数据， $p_{\text{data}}(x)$ 代表真实数据分布。表达式中 $\log D(x)$ 表示判别器对真实样本的判断， $\log(1 - D(G(z)))$ 表示对生成的样本的判断。判别器尽力将真假样本进行区分，这个过程可以表示为最大化 $\mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(D(x))] + \mathbb{E}_{z \sim p_z(z)}[1 - \log(D(G(z)))]$ 。生成器尽力找到数据分布规律，这个过程可以表示为最小化 $\mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$ ，实现中往往将 $\mathbb{E}_{z \sim p_z(z)}[\log D(G(z))]$ 最大化。

图 2-2 是 GAN 的基本思路解释图。生成式对抗网络通过变更判别器分布（图中蓝色虚线）和生成器分布（图中绿色实线）进行训练，判别器（D，蓝色）试图对真实样本（黑色）与生成器生成样本（绿色）进行判别。图中下方的直线代表噪声（ z ）的采样域，上方直线代表真实数据（ x ）的定义域。箭头代表映射 $x = G(z)$ 。

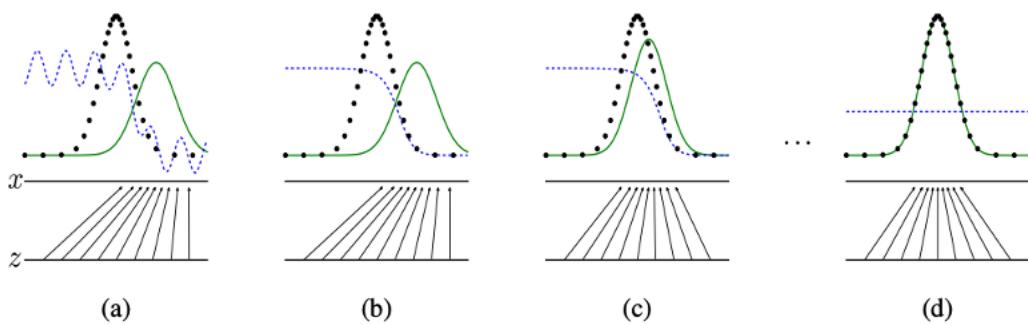


图 2-2: GAN 基本思路示意图

(a) 初始状态 (b) 固定生成器，优化判别器 (c) 固定判别器，优化生成器 (d) 博弈后，生成数据的分布近似真实样本，判别器难以判断真伪

2.1.2 算法原理

已知生成对抗网络是分为生成器（G）和判别器（D）两部分的。生成器用来生成逼真数据，判别器用来对生成图片进行真假判断，两部分相互优化，并更新参数，最终使得生成器的输出接近真实图像。本节将对 GAN 的算法原理进行描述。

GAN 的判别器主要用于判断样本的真伪，本质可以看作进行二分类任务，

模型使用 Sigmoid 将输出标量映射到 0~1 之间，用来表示判别器对样本真伪的判断，该值越接近 1 代表判别器认为它更大概率为真实数据。因此，可以将其建模为伯努利分布。伯努利分布（Bernoulli distribution）也称 0~1 分布，属于离散型的概率分布。伯努利试验有成功和失败两种情况。成功时，有随机变量为 1；失败时，有随机变量取值为 0。记其成功概率为 $p(0 \leq p \leq 1)$ ，失败概率为 $q = 1 - p$ 。伯努利分布中，对于随机变量 y ，有 $y = 0$ 和 $y = 1$ 两种状态，满足分布： $y = 1$ 的概率为 p ， $y = 0$ 的概率为 $1 - p$ 。即， $p(y) = p^y(1-p)^{1-y}$, $0 < p < 1$ 。对于 N 个样本，有 $\{y_1, y_2, \dots, y_N\}$ ($y_i \in \{0, 1\}, i \in [1, N]$)，对 p 进行最大似然估计。最大似然估计（Maximum Likelihood Estimation，简称 MLE）是用于参数估计的一种方法。详细地说，如果有随机样本满足某种概率分布，但有一些参数是未知的。那么进行多次试验后，根据结果能够对参数进行一个估值。如果某个参数值能够使样本出现的概率最大，那么就将这个参数当作估计的真实值，这就是最大似然估计的方法。公式 2-2 表示对 p 进行最大似然估计。

$$\begin{aligned}
 p^* &= \arg \max_p \mathcal{L}(p) \\
 &= \arg \max_p \log L(p) \\
 &= \arg \max_p \log \prod_{i=1}^N p(y_i) = \arg \max_p \log \prod_{i=1}^N p^{y_i}(1-p)^{1-y_i} \\
 &= \arg \max_p \left[\sum_{i=1}^N y_i \log p(y_i) + (1 - y_i) \log (1 - p(y_i)) \right] \quad (2-2) \\
 &= \arg \max_p \frac{1}{N} \left[\sum_{i=1}^N y_i \log p(y_i) + (1 - y_i) \log (1 - p(y_i)) \right] \\
 &= \arg \max_p \left[\frac{1}{N} \sum_{y_i=1} \log p(y_i) + \frac{1}{N} \sum_{y_i=0} \log (1 - p(y_i)) \right] \\
 &\approx \arg \max_p \mathbb{E}_y [\log p(y)] + \mathbb{E}_y [\log(1 - p(y))]
 \end{aligned}$$

其中， $\mathbb{E}_{x \sim p(x)}[f(x)] \approx \frac{1}{N} \sum_{i=1}^N f(x_i)$, $x_i \sim p(x)$ ，将样本函数值的平均值近似写作随机变量函数的数学期望。

具体到样本真伪的判断，对于样本集 x ，有函数 $y = D(x; \phi)$ 代表某数据为“真”的概率，输出值范围 $[0, 1]$ 。与上述伯努利分布类似，该数据分布律可以表示为： $y = 1$ 的概率为 $p(y | x; \phi)$ ， $y=0$ 的概率为 $1 - p(y | x; \phi)$ 。通过对该分布进行最大似然估计，即利用获取到的 (x, y) 的样本值，便可以对 ϕ 进行估计。

即对 ϕ 进行最大似然估计求解最优参数 ϕ^* , 公式 2-3 是这一过程的表达式。

$$\begin{aligned}
 \phi^* &= \arg \max_{\phi} \mathcal{L}(\phi) \\
 &= \arg \max_{\phi} \log L(\phi) \\
 &= \arg \max_{\phi} \log \prod_{i=1}^N p(y_i | x_i; \phi) \\
 &= \arg \max_{\phi} \log \prod_{i=1}^N p(y_i | x_i; \phi)^{y_i} (1 - p(y_i | x_i; \phi))^{1-y_i} \\
 &= \arg \max_{\phi} \left[\sum_{i=1}^N y_i \log p(y_i | x_i; \phi) + (1 - y_i) \log (1 - p(y_i | x_i; \phi)) \right] \quad (2-3) \\
 &= \arg \max_{\phi} \frac{1}{N} \left[\sum_{i=1}^N y_i \log p(y_i | x_i; \phi) + (1 - y_i) \log (1 - p(y_i | x_i; \phi)) \right] \\
 &= \arg \max_{\phi} \left[\frac{1}{N} \sum_{x_i \in x^r} \log p(y_i | x_i; \phi) + \frac{1}{N} \sum_{x_i \in x^g} \log (1 - p(y_i | x_i; \phi)) \right] \\
 &\approx \arg \max_{\phi} \mathbb{E}_{x \sim p_r(x)} [\log p(y | x; \phi)] + \mathbb{E}_{x \sim p_g(x)} [\log (1 - p(y | x; \phi))]
 \end{aligned}$$

其中, x^r 代表真实样本, x^g 代表假样本, $p_r(x)$ 代表真实样本分布, $p_g(x)$ 代表假样本分布。

在生成对抗网络中, 函数 $y = D(x; \phi)$ 其实代表了判别器的功能, 对输入地样本数据进行辨别, 输出样本为真概率。对判别器的训练也可以理解为对参数进行最大似然估计, 得到最优参数 ϕ^* , 从而进行准确的判断。

在训练之后, GAN 的判别器已经可以较为准确地分辨真实数据和生成器生成的数据。这时, 需要固定判别器, 对生成器进行优化。

与判别器类似, 生成器的功能也可以使用函数来描述。生成器是通过噪声来生成与真实图像相似的假图像, 这一过程可以表示为 $x = G(z; \theta)$, 其中 z 代表一些随机噪声, 是同 ϕ 类似的未知参数。该分布可以表示为 $x \sim p(x | z; \theta)$ 。生成器的生成样本将会输入到判别器中进行判断, 我们希望判别器将这些样本均判断为负, 认为他们均是假样本。那么对于 N 个样本, 有 $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ ($y_i = 0, i \in [1, N]$), 对 θ 进行最小似然估计可以表示为公式 2-4。

$$\begin{aligned}
\theta^* &= \arg \min_{\theta} \mathcal{L}(\theta) = \arg \min_{\theta} \log L(\theta) \\
&= \arg \min_{\theta} \log L(\theta) = \log \prod_{i=1}^N p(y_i | x_i; \phi^*) \\
&= \arg \min_{\theta} \log \prod_{i=1}^N p(y_i | x_i; \phi^*)^{y_i} (1 - p(y_i | x_i; \phi^*))^{1-y_i} \\
&= \arg \min_{\theta} \sum_{i=1}^N y_i \log p(y_i | x_i; \phi^*) + (1 - y_i) \log (1 - p(y_i | x_i; \phi^*)) \quad (2-4) \\
&= \arg \min_{\theta} \sum_{i=1}^N \log (1 - p(y_i | x_i; \phi^*)) \\
&= \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N \log (1 - p(y_i | x_i; \phi^*)) \\
&= \arg \min_{\theta} \mathbb{E}_{x \sim p(x|z;\theta), z \sim p(z)} [\log (1 - p(y | x; \phi^*))]
\end{aligned}$$

在生成对抗网络中，函数 $x = G(z; \theta)$ 其实就是生成器的功能，用于生成所需的数据。在网络的训练中，生成器部分的训练也可以理解为是对参数进行最小似然估计，得到最优参数 θ^* ，从而对模型进行优化。

所以，能够得到生成器和判别器的优化公式：

$$\begin{cases} D^* = \arg \max_D \mathbb{E}_{x \sim p_r(x)} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log (1 - D(G(z)))] \\ G^* = \arg \min_G \mathbb{E}_{z \sim p(z)} [\log (1 - D^*(G(z)))] \end{cases} \quad (2-5)$$

即， $G^* = \arg \min_G \max_D \mathbb{E}_{x \sim p_r(x)} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log (1 - D(G(z)))]$

具体算法可以表示为：

1. 判别器优化：使用真实样本数据和虚假样本数据训练判别器，调整判别器参数
2. 生成器优化：使生成器输出的数据迷惑上一步中优化得到的判别器
3. 重复以上步骤直到判别器难以识别图像真伪，生成器输出数据质量较高

2.1.3 医疗领域的 GAN 生成

自 GAN 提出以来，其在计算机视觉领域的应用层出不穷，在图像处理方面的应用效果得到证实，在图像生成方面的效果尤为显著。

虽然生成对抗网络在刚提出时，是一种无监督的生成模型，但在其在半监督学习、完全监督学习方面也被认证具有应用意义。在图像生成方面，根据生成器和判别器数量与结构的不同，可以分为直接方法，分层方法和迭代方法。直接方法中，模型包含一个生成器和判别器，没有分支，效果良好。而分层方法则主要通过两个生成器和判别器，将图像分成两部分生成，在进行结合。迭代方法则使用结构相似的生成器通过迭代逐步生成更细粒度的图像 [30]。GAN 能够通过无监督学习方法，从训练样本中学习到图像的深层特征。训练得到的生成器能够生成非常逼真的图像内容 [31]。同时，GAN 不仅生成的图像与真实图像非常相似，且具有较高的清晰度，能满足大多数场合的使用需求。在图像转换领域，使用生成对抗网络可以将某一类型图像甚至文字转换为另一类型图像不仅包含有监督的学习方法例如：Pix2Pix[32]，也有无监督的学习方法，例如：CycleGAN[33]。图像转换在医学图像转换、人脸表情转换等方面已有应用。

具体到医学图像领域，医疗图像识别领域的深度学习网络层出不穷，许多模型的识别准确度已经能够媲美甚至超过专家的诊断。但与通用领域的深度学习模型一样，模型的准确度除了受到本身算法质量的影响之外，还依赖于用于模型训练的样本数量与质量。训练样本质量低下往往会导致模型准确率不高。没有优秀的学习教材，模型很难学习到真实的图像规律。受到伦理、隐私等约束的限制，医疗图像的获取与分享困难重重，所以医疗图像识别模型训练往往缺少有效样本，样本数量不足，质量参差不齐。因此，在医疗等原始样本匮乏的领域内，进行图像扩增是非常有必要。另外，进行医疗图像识别模型的训练需要使用带有标签的样本作为数据集。在通用领域，为了获取标签，可以采用众包等手段进行标注。而在医疗领域，样本的标注需要大量的领域知识和实践经验，往往只有行业专家能够进行此项工作。但这样的人力物力又是极为稀缺的，如果想要培养能够进行标注的人员更是难上加难，而图像扩增正能解决这样的问题。同时，对于医疗图像来说，原始样本的分布往往是极不均衡的，即大多数的样本都是正常的不含病灶的。而大多数模型的关注点就在于病灶部分，所以有效的样本数量不平衡。目前通常采用随机采样解决样本不均衡的问题。但上采样可能会得到较为近似的图像数据，下采样会导致一定程度上的信

息缺失 [34]。那么对于含病灶的样本进行变异扩增就显得尤为重要。GAN 能够有效解决样本类别不均衡的问题，为模型训练提供更加合理的训练数据集。

在分析了医疗图像扩增的必要性后，需要有针对性地选择合适的医疗图像扩增手段。对于常见图像而言，对于图像整体进行仿射扩增和弹性扩增就能够大大增强模型准确度。所谓仿射扩增，就是对于图像的像素点进行位置偏移，而像素值不进行改动，例如：旋转、翻转、反射、裁剪等等。弹性扩增，就是不改变像素点位置，对于像素值本身进行有规则的变化，例如：对比度偏移、亮度偏移、噪声、卷积、雾化等等。但这样的扩增效果又是有限的。比如，医疗图像中病灶往往是研究的关键，不合理的整体变化会导致病灶区域的特征丢失，因此也有研究提出了病灶与非病灶区域分开进行变换的扩增思路，并通过实验证明了局部扩增方法的有效性。[21]

但由二维图像上升到三维图像，一些变异算法的误差会被放大，例如病灶缩小、放大等变换很容易违背病灶生长规律。医疗图像本身特征复杂，人体组织生长变异规则复杂，病灶生成规则难以轻易探究。因此，在进行三维医疗图像扩增时，GAN 便成为了一种不错的选择。GAN 能够自动学习图像特征，避免了人为变动图像造成的误差。生成器和判别器自动进行博弈并优化，能够输出真实的三维医疗图像。

在医学图像处理领域，GAN 的应用先例也很多。比如，有研究使用 GAN 形成原始图像与去噪图像的映射来消除医疗图像中的噪声，避免了 CNN 等方法中使图像质量下降的问题 [35]。由于操作失误、环境影响或患者体态问题，一些医疗影像在采样时会出现部分信息丢失的情况，而这些信息的缺失往往会影响医生对病情的判断。使用 GAN 可以建立信息缺失图像与正常图像的映射来对图像进行重建。以上几个方面的应用都体现了生成对抗网络在医疗影像细节特征学习上的优势和医疗影像生成方面的优势。另外，GAN 也被应用在了医疗图像分割识别当中，例如，使用 GAN 进行脑、眼、胸等部位图像的识别分析。

同通用领域一样，生成对抗网络在医疗领域也有图像生成和图像转换方面的应用。不论是最开始的无监督生成，还是之后的条件生成对抗网络即 CGAN，各类生成对抗网络都被证实能够生成大量高质量的医疗影像样本。

GAN 在挖掘深层次微小特征上具有独特的优势，因此选用它来进行三维病灶图像生成非常契合，即依次优化判别器与生成器，使模型到学习病灶的结构分布规律。卷积生成对抗网络（DCGAN）是在 GAN 的基础上使用卷积神经网

络代替了原始 GAN 中的多层感知机，特征学习能力更强 [36]。DCGAN 能学习到更深层次的特征，生成的图像质量更佳，满足医疗影像的高精度需求，所以本文使用三维 DCGAN 技术实现病灶生成。另外，使用生成对抗网络技术生成病灶能够有效缓解大部分医疗数据集存在的样本分布不均的问题。

2.2 激活函数

上小节简要介绍了生成对抗网络的算法原理。包括生成对抗网络在内的众多神经网络中，有一些非常重要的函数影响着网络性能，比如激活函数、损失函数。在神经网络中，一般会使用激活函数引入非线性，强化学习能力。在本文使用的 DCGAN 中，其生成器的输出层使用 Tanh 作为激活函数，其他层使用 ReLU 作为激活函数。而判别器中输出层使用 Sigmoid 作为激活函数，其他所有层都使用 Leaky ReLU 作为激活函数。

Sigmoid 函数是卷积网络中常用的激活函数，也称作 Logistic 函数。它对输出范围进行控制，将输入映射到 $(0,1)$ 中，其函数表达式为：

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2-6)$$

但 sigmoid 函数的输出是非 0 均值的，因此向下层神经元传入的输入也为非 0 均值的，从而会引起收敛缓慢。并且，Sigmoid 函数导数的范围在 $[0,0.25]$ 中，在网络反向传播时，梯度值会不断变为原来的 0.25 倍，随着层数的增加梯度值会越来越接近 0，从而造成梯度消失，导致难以训练。其函数及导数图像如图 2-3 所示。

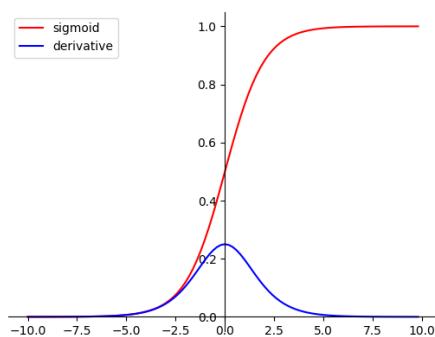


图 2-3: Sigmoid 函数

Tanh (Hyperbolic Tangent) 函数是双曲正切函数，它和 Sigmoid 函数一样对输出进行了限制，将输入映射到 (-1,1) 中。所以 Tanh 函数能够改善梯度消失问题，解决了 Sigmoid 函数因均值非 0 出现的收敛缓慢的问题 [37]。其函数表达式为：

$$\begin{aligned} f(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ &= \frac{2}{1 + e^{-2x}} - 1 \end{aligned} \quad (2-7)$$

其函数及导数图像如图 2-4 所示。

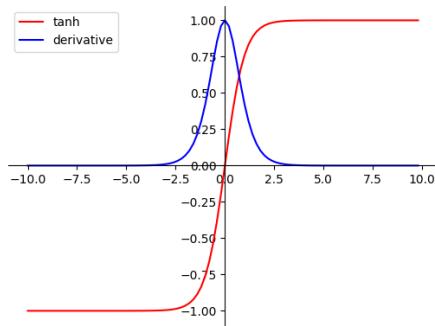


图 2-4: Tanh 函数

ReLU (Rectified Linear Unit) 函数 [38] 是修正线性单元函数，函数表达式为：

$$f(x) = \max(0, x) \quad (2-8)$$

ReLU 函数是非线性函数。在 $x > 0$ 的情况下，它的导数为 1，所以它能够解决上述网络的梯度消失问题。在 $x < 0$ 的情况下，它将函数值置为 0 去除噪音。ReLU 函数公式简单。使用它作为激活函数，不仅能够简化运算，还能较好地解决梯度消失问题，加快收敛。其函数及导数图像如图 2-5 所示。

由于 ReLU 函数中将 $x < 0$ 时的函数值置为 0，这样可能会导致一些关键特征被忽略，所以一些函数对 ReLU 函数 $x < 0$ 的部分进行了改动，比如 Leaky ReLU 函数 [39]。Leaky ReLU 的函数表达式为：

$$f(x) = \max(\alpha x, x) \quad (2-9)$$

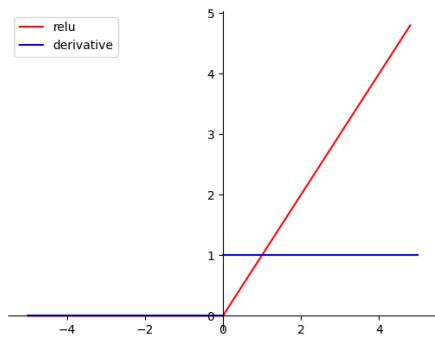


图 2-5: ReLU 函数

其函数及导数图像如图 2-6 所示

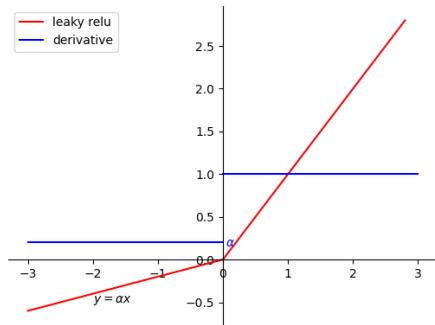


图 2-6: Leaky ReLU 函数

2.3 损失函数

除了上节介绍的激活函数，损失函数也是神经网络中的一个重要函数。生成对抗网络中的关键环节就是利用生成器和判别器的博弈不断优化，使得生成的假数据与真实数据近似。而在深度学习中，往往是通过定义损失函数来度量两个数据分布的近似程度。在深度学习领域中，通常使用损失函数来衡量模型在进行分类预测时不准确的程度，即预测与真实的差距。损失函数计算得到的值越小说明该模型的鲁棒性就越优秀。

“熵”也是损失函数中的一个重要概念。一般使用“熵”表示系统内在的

混乱程度，表示事件包含的信息量，又称为“信息熵”，可以表示为：

$$S(p_1, p_2, \dots, p_n) = -K \sum_{i=1}^n p_i \log_2 p_i \quad (2-10)$$

其中， S 表示熵， i 表示概率空间中所有可能的样本， p_i 表示该样本的出现几率， K 是和单位选取相关的任意常数。

而想要衡量与随机变量 x 相关的两个独立事件概率分布间的差异，通常会使用相对熵来表示。相对熵（relative entropy）也被称为 KL 散度（Kullback-Leibler divergence），等于两个数据分布的信息熵的差值。因此，可以使用相对熵对两个分布之间的距离进行度量。两个随机分布越相似，那么分布的相对熵就越小；分布间的差距越大，相对熵越大。随着分布的接近，它们间的相对熵逐渐趋近于 0。如果设 $P(x), Q(x)$ 是随机变量 X 上的两个概率分布，如果随机变量是连续型的，则相对熵的定义为：

$$\text{KL}(P||Q) = \int P(x) \log \frac{P(x)}{Q(x)} dx \quad (2-11)$$

除了使用相对熵来度量差距，研究者还在此基础上提出了交叉熵 [40] 的概念：

$$\begin{aligned} D_{KL}(A||B) &= -S(A) + H(A, B) \\ &= \sum_i P_A(x_i) \log \left(\frac{P_A(x_i)}{P_B(x_i)} \right) \\ &= \sum_i P_A(x_i) \log(P_A(x_i)) - P_A(x_i) \log(P_B(x_i)) \end{aligned} \quad (2-12)$$

$$S(A) = - \sum_i P_A(x_i) \log P_A(x_i) \quad (2-13)$$

$$H(A, B) = - \sum_i P_A(x_i) \log(P_B(x_i)) \quad (2-14)$$

其中， $D_{KL}(A||B)$ 代表 A 与 B 的相对熵， $S(A)$ 代表 A 的熵， $H(A, B)$ 代表 A 与 B 的交叉熵。当 $S(A)$ 为常量， $D_{KL}(A||B)$ 与 $H(A, B)$ 等价。

在模型训练中，因为训练数据是确定的，所以其熵也是确定值，因此使用交叉熵与 KL 散度效果一致。为了减少计算成本，通常使用交叉熵作为损失函数进行度量。本文模型也是选用了交叉熵作为损失函数。

2.4 通用图像扩增技术

扩增手段已经广泛应用在深度学习的数据预处理过程中，实现了对训练样本的扩增，提升了模型训练的效果。在图像扩增领域，按照不同维度扩增方法有多种分类方式，充分了解这些方法对于本文在验证实验中选择合适的扩增手段作为对照是十分必要的。如果在这一部分选择使用不合理的扩增手段，很可能因破坏样本可靠性导致模型训练效果受到影响，无法作为对照组进行实验，从而导致对GAN生成样本的质量评估出现较大误差。

根据一次扩增变换中使用到的样本数量为标准，图像扩增可以分为单样本扩增和多样本扩增。只使用到一个样本的变换方法就称为单样本扩增，使用多个样本则称为多样本扩增。广义的来讲，本文使用GAN生成数据也属于多样本扩增。由于一些常用的多样本扩增（SamplePairing[41]、mixup[42]等）虽然在性能上提升效果可观，但可解释性不强，为了保证关键病灶部分不受到破坏，这里本文选用单样本扩增作为对照的扩增手段。

根据进行扩增的图像变异范围不同，图像扩增可以分为针对图像本身的整体扩增和针对病灶（或非病灶）部分的局部扩增。以整个图像范围的像素为变换目标的操作称为整体扩增方法，将病灶与非病灶部分区域分开处理的操作称为局部扩增方法。本文希望使用对抗生成网络生成样本也是希望避免在三维图像中采取局部扩增方法时产生较大误差。从出发点来说，本文技术与局部扩增方法相似，都是关注病灶本身的特征。本文技术希望通过训练后的生成器模型生成病灶，使得扩增得到的数据尽量不要受到非病灶区域中一些组织的负面影响。从提升模型训练的角度来说，通过整体变异方法进行图像扩增，并使用扩增数据集训练判别模型，能够增强模型的鲁棒性。受到仪器配置、采集环境等各种客观因素的影响，医疗影像的质量参差不齐。使用整体扩增方法能在一定程度上降低这些客观因素对于识别模型的影响。综合以上两点，在验证实验中，本文选择使用整体扩增方法作为验证实验中的对照扩增方法。

根据对像素点的不同变异方式，可以粗略将图像扩增分为仿射扩增和弹性扩增。仿射扩增就是不改变像素点的值，只改变像素点的位置，例如：旋转、翻转等等。弹性扩增就是不改变像素点的位置，但回对像素点的值进行一定变化，例如：对比度偏移、亮度偏移、模糊卷积等等。这些变异方法已被大量实验证实能够产生大量有益于模型训练的样本，能够稳定提高模型的准确度与鲁棒性。图2-7展示了常见的变异方法。

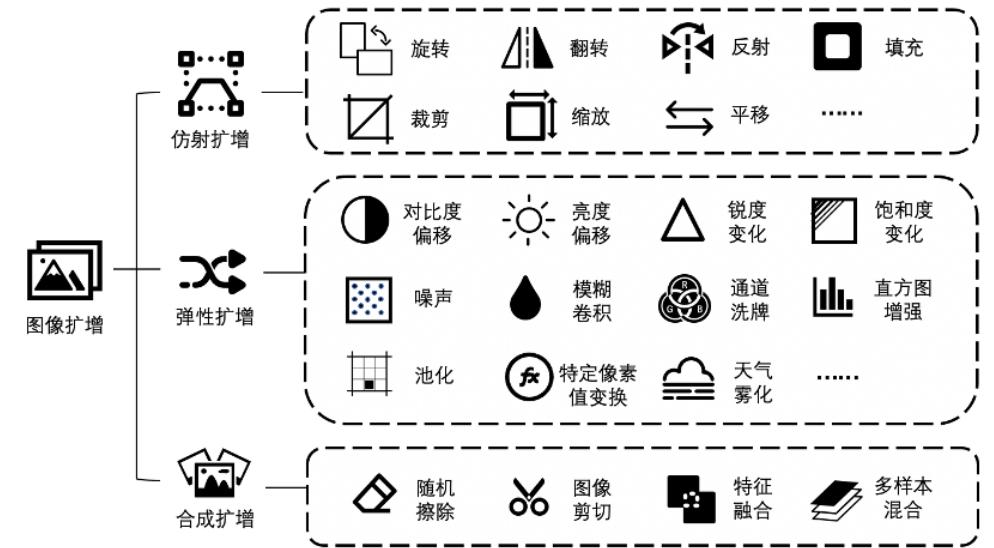


图 2-7: 通用图像扩增技术

本文的验证实验中选用了一些通用图像扩增技术作为对照技术，将使用这些技术生成的数据集作为对照训练集，与生成对抗网络生成的数据集进行对比，分别用于病灶判断模型的训练，通过对比训练结果来证实本文技术的有效性。

2.5 本章小结

本章主要介绍了本文技术使用到的相关理论知识，包括生成对抗网络技术、深度学习相关函数以及常用的图像扩增技术。针对生成对抗网络技术，通过阐述其原理与结构，结合相关研究，分析了其用于病灶生成的可行性。同时，本文在验证实验当中使用了通用图像扩增技术作为对照技术，将其生成样本作为对照数据集，用以验证本文技术。本章对相关理论的阐述为之后关键技术的研究建立了理论基石。

第三章 病灶数据分析

医疗影像具有多样性的特点。首先，人体有多种的组织器官。其次，每种组织器官都有其不同的病症与表征。并且，不同器官又另有多类可观测的影像类型。如果要确定本文实验所需的数据集，必须要从目前深度学习图像识别在医疗领域的发展与最新需求入手，结合医疗图像本身的特征进行分析。

3.1 医疗病灶分析

在确定数据集之前，首先需要明确选用的具体组织器官类型、对应病症及病灶的类型，以及对应深度学习图像识别模型的类型。人体的器官、组织、病灶本身具有多样性。深度学习模型能够针对不同部位能够辅助诊断多种疾病 [43]。例如针对心脏，可以辅助诊断先天性心脏病、动脉硬化和冠心病；针对大脑，分析诊断自闭症、阿尔茨海默病；针对眼部，通过血管分割识别早期糖尿病病人视网膜病变；针对肺部，发现肺结节预防早期肺癌、诊断肺炎等等。结合生成对抗网络对于深度细小特征的优势，实验目标将从微小病灶中考虑。

事实上，即使确定了器官及病灶类型，用于诊断该疾病的影像也种类多样，具有不同的影像类型。常见医学影像类型有 MRI 磁共振成像、CT 影像、PET 影像、X-ray 影像等等 [44]。一般来说，深度学习模型的分析都只针对特定器官的特定影像进行分析，因此，在进行医疗图像数据扩增的实验设计与数据集选定时，也是针对特定病灶的特定影像类型进行扩增，并根据该病灶或组织的特性及影像特点选择扩增方法。

目前国际常用医疗图像数据集类型丰富，例如：CHD Atlas 数据集包含先天性心脏病人心脏 MRI 核磁共振成像、ABIDE 数据集包含自闭症患者脑部 MRI 核磁共振成像、ADNI 包含阿尔茨海默氏病人脑部 MRI 核磁共振成像以及 PET 影像、OASIS 包含阿尔茨海默病人脑部 MRI 核磁共振成像、DRIVE 数据集包含早期糖尿病病人视网膜 PNG 图像、LIDC 数据集包含 DICOM 格式的肺部 CT 图像、LUNA16 数据集包含 MHD 格式的肺部 CT 图像。除此之外，还存在一些综合型医疗影像数据库。

在综合分析医疗领域需求和当前常用公开数据集资源后，本文在小型组织病灶中选定肺部结节为实验病灶。具体分析如下。

癌症是全世界大多数国家和地区最重要的居民死亡原因。据杂志 CA CANCER J CLIN 发布的《全球癌症统计 2020 年版》（Global Cancer Statistics 2020）统计，全球范围内 2020 年新增癌症病例 1930 万例（不包括 1810 万例非黑色素瘤皮肤癌），较 2018 年的 1810 万例有所上升。癌症死亡近 1000 万例（不包括 990 万例非黑色素瘤皮肤癌），超过 2018 年的 960 万。在全球中的 112 个国家或地区，因癌症死亡是其居民排名前二的死亡原因。另有 23 个国家和地区，癌症在其所有居民死亡原因中排名前四。

肺癌是全球第二常见的癌症。如图所示 3-1。据统计，2020 年新增肺癌病例 220 万例，约占诊断癌症的十分之一（11.4%）。对于男性而言，肺癌的发病率在所有癌症中排名第一，对于女性而言，肺癌的发病率排名第三。在 2020 年之前，肺癌一直是全球发生率最高的癌症。统计表明，虽然 2020 年度女性乳腺癌（230 万例新发病例，11.7%）已超过肺癌（11.4%）成为最常见的癌症，但肺癌仍然是在全球范围内造成死亡人数最多的癌症。

肺癌是导致死亡人数最多的癌症。据统计，2020 年新增肺癌死亡病例 180 万例，仍然是癌症死亡的主要原因，占死亡人数的 18.0%。对于男性而言，肺癌的死亡率在所有癌症中排名第一，对于女性而言，肺癌的死亡率排名第二，仅次于乳腺癌。总体而言，男性的肺癌发病率和死亡率大约是女性的 2 倍 [45]，统计图像如 3-1 所示。

在中国，癌症造成的影响更加不容小觑。2020 年的报告显示，癌症已经成为中国人排名第一的死亡原因。同时，肺癌始终威胁着中国居民的身体健康。在中国，男性的肺癌发病率在所有癌症中保持最高，死亡率也继续排名第一，死亡占比进一步提升到了 21.5%。同时，截止到 2020 年，中国女性发病率最高的癌症类型也是肺癌，死亡率同样占据第一。到 2020 年，中国居民肺癌的发病数占到了全球的 37%，死亡数占到全球的 39.8%，肺癌防治情况不容乐观 [45]。

研究表明，肺癌患者在早期诊断出并进行治疗，其五年生存率能够达到 70%，但是如果患者在晚期才诊断出肺癌，其五年生存率不足 15% [46]。在对 2010 年 ~2014 年间确诊的肺癌患者进行统计时，肺癌患者确诊后的 5 年存活率仅有 10% ~20%。肺癌症状通常在癌症晚期才会出现，包括持续咳嗽、痰中带血、胸痛、声音嘶哑、呼吸急促恶化，以及反复发作的肺炎或支气管炎等。

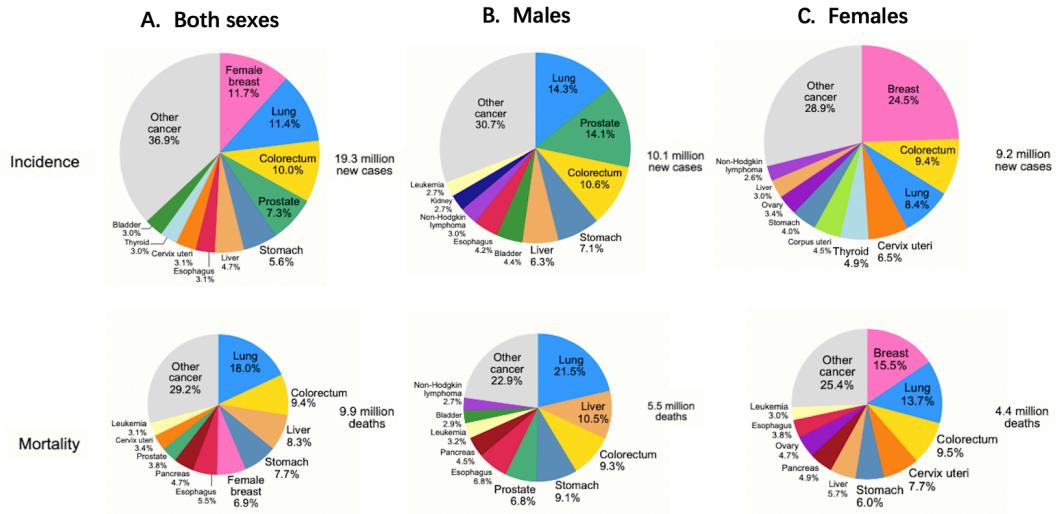


图 3-1: 2020 年全球癌症发生率与死亡率占比。A: 全性别癌症发生率与死亡率。B: 男性癌症发生率与死亡率。C: 女性癌症发生率与死亡率

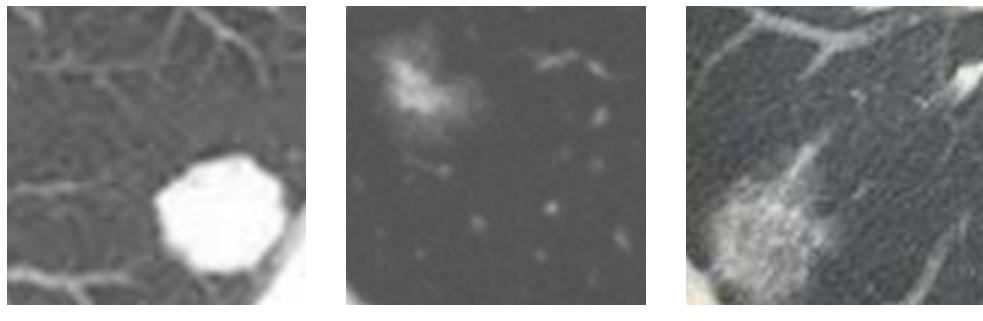
而在早期，患者很难通过外部表征来收到预警。随着医学研究和医疗手段的进步，肺结节筛查通常被认为是早期肺癌筛查的有效手段。肺结节的产生是肺部病变的表现，也是肺癌的早期表现。研究表明 [47–51]，通过对高危人群进行低剂量计算机断层扫描（LDCT）筛查有助于早期诊断癌症，从而提升成功治疗的可能性。低剂量螺旋 CT 辐射小，且与常规剂量的效果非常接近 [52]。医生通过观察肺部计算机断层扫描影像中肺结节的大小、形态、数目、位置、生长速率等表现评估结节情况，进行肺结节识别和定性，从而有效识别早期肺癌。最近，荷兰-比利时肺癌筛查试验报告称，试验通过 CT 容积扫描、引入生长率评估，在 10 年的随访中发现男性肺癌死亡率降低了 24%，女性肺癌死亡率降低了 33% [51]。美国癌症协会（ACS）建议 50 至 80 岁、曾有 20 包年吸烟史或在过去 15 年内戒烟的健康成年人按照美国预防服务工作组的指导意见每年进行低剂量计算机断层扫描检查 [53]。

我国也开展了一系列肺癌早诊早治项目，以提高居民肺癌筛查意识，建立高质量的肺癌筛查体系，规范医疗机构筛查技术手段。我国对于早期肺癌的诊断同样以肺结节的识别与定性作为方法。在我国进行的多种筛查项目中，专家发现，在对 CT 影像进行筛查时，假阳性结节的错判概率较高，结节识别准确度有待提升 [47]。

可以说，肺结节是肺癌早期确诊的重要识别依据 [54]。传统的肺结节识别主要是通过专业医生对病人的 CT 影像进行人工查阅来实现，但人工方法在临

床应用上暴露出诸多问题，比如人眼判断容易忽略掉一些小型的肺结节，判断的准确度有所欠缺。所以，随着智能医疗的提出，深度学习在医疗领域的应用蓬勃发展，越来越多的机构引入计算机来进行疾病发现与诊断。具体来说，越来越多的机构开始使用深度学习解决病灶识别问题，在肺癌检测方面就是通过深度学习进行肺结节的检测，并结合影像可视化技术，辅助医生进行诊断。引入计算机辅助检测不仅提高了检测的准确率，也缩短了识别耗时和人力成本。

肺结节 [47] 也被称为肺部斑点，它是肺中直径小于或等于 30 毫米的病灶。在医学影像中，肺结节显现为高密度的阴影，呈类圆形或不规则形，被肺实质包围。根据肺部上病灶数量的不同，肺结节可以分为单发性结节和多发性结节，如果肺上只诊断出单个斑点，则被称为单发性肺结节。对于多发性肺结节进行评估，往往参考其中最具威胁型的结节的性质。根据 NELSON 临床试验和 PanCan 临床研究数据发现，结节数量在一到四枚间，多发性肺结节患者的患癌几率随结节数量增大而增加，但超出范围后，患癌几率随结节数量增大而降低。总体而言，多发性结节患者癌症恶性率低于单结节患者。根据结节的密度不同，可以将其分为实性结节、纯磨玻璃结节和部分实性结节（混合型结节）。部分实性结节、纯磨玻璃结节、实性结节，其为恶性的概率依次升高。



(a) 实性结节

(b) 混合型结节

(c) 纯磨玻璃结节

图 3-2: 肺结节样式

目前对于肺结节的推荐检测方法是对肺部影像图片进行三维重建，立体直观地观察结节的数量与形态，并使用 CAD 辅助医生进行结节识别。如果仅依靠专家人眼评估二维图像，小结节的识别率不高，且对于结节形态的观察不直观。并且，使用二维图像判断结节性质，一般是通过计算结节最大直径与最小直径的平均值，通过该均值大小对结节进行定性，忽略结节的纵向长度。使用 CAD 进行辅助识别，则可以通过三维建模并计算结节容积。容积倍增时间是有效的结节定性指标之一，将其作为指标进行判断，数据更加准确，定性效果更优秀，也能够提升小结节的识别率。

临床胸部检查的方式一般有胸部 CT 检测和胸部 X 光检查，他们都是利用 X 射线穿过不同密度组织具有不同衰减强度的特性来实现的。CT 检查是按一定层间隔对肺部进行横断面成像 [55]，X 光通常是对胸部整体进行一个平面整体成像。就检测效果而言，胸部 CT 精确度高，而 X 光片成像容易受到其他组织的干扰，造成肺结节被骨头等部位遮挡 [56]。所以对于肺结节的检查通常还是依靠胸部 CT 进行。低剂量螺旋 CT 是世界公认的早期肺癌筛查手段，被证实能够明显提高肺癌检测率和患者生存率。它使用低辐射剂量扫描，对人体部位进行拍摄成像，精度较高。进行肺结节识别的深度学习网络通常也是以胸部 CT 作为输入样本的。因此，本文进行肺结节影像生成时同样选用胸部 CT 为样本。以此结论为基础，需要进一步在现有医疗数据集中选择合适的胸部 CT 数据集。

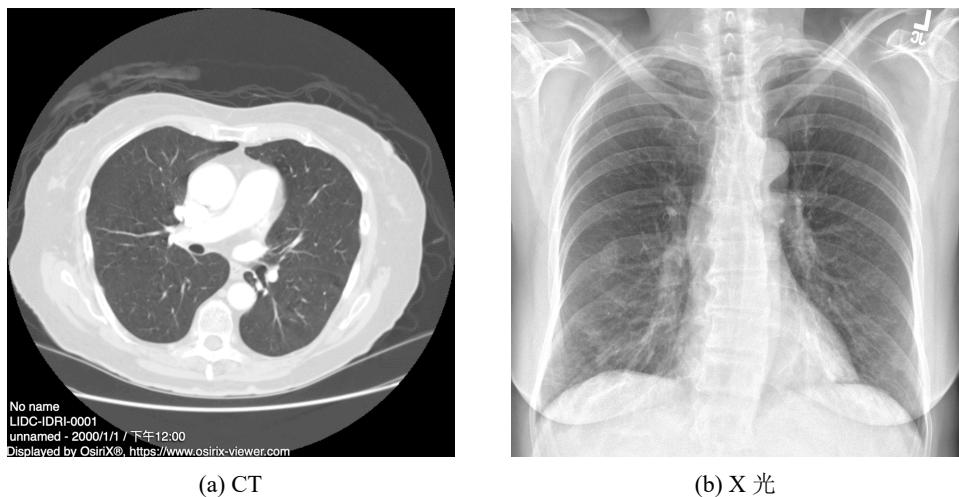


图 3-3: 胸部检查影像

3.2 影像数据分析

近年来，国内外有不少机构发布了共享的肺部影像数据集，比如国内的 Tianchi，国外的 LUNA16、Kaggle、LIDC-IDRI 等等。根据数据集中肺部影像存储格式的不同，可以将其分为两类。Tianchi 和 LUNA16 使用 mhd 格式存储影像数据，数据集通过 mhd 和 raw 文件来对病例的检查影像资料进行格式化保存，而 Kaggle 和 LIDC-IDRI 使用 dcm 格式存储影像数据。

LIDC-IDRI 数据集^①是最大的公用肺结节数据集。它提供了 1018 个病例的

^①<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>

胸部医疗影像，并提供了对结节的标记注释。该数据集项目由美国国家癌症研究所发起，旨在为肺癌检测与诊断的计算机辅助诊断系统提供有效的数据样本 [57]。表 3-1 中展示了 LIDC-IDRI 数据集的基本信息情况。

表 3-1: LIDC-IDRI 数据集基本信息

数据集信息	
影像格式	CT（计算机断层扫描） SEG（DICOM 分割） CR（计算机 X 线成像） DX（数字化 X 线摄影）
患者人数	1018
研究总数	1308
序列总数	1398
影像总数	244,617
影像大小 (GB)	125

该数据集通过 dicom 文件来保存肺部影像。每个 dicom 文件可以转换为一个二维图像 [58]，代表肺部某层的横截面扫描。所以，一次检测的影像由多个 dicom 文件组成，代表多层的肺部横截面扫描。数据集通过 xml 文件来保存病灶的详细信息，其中包括四位经验丰富的专业胸部放射科医师对肺结节进行的标注注释。每位专家分别对检查的肺部影像进行识别和注释。结节注释类型有三种：“nodule > or =3 mm”，“nodule <3 mm” 和 “non-nodule > or =3mm”。在最后，每位专家独立审查自己及其他三名专家的标记，得出最终结论。

LIDC-IDRI 数据集中，每次检查的记录都被保存在一个文件夹中，该文件夹下有数个 dicom 文件和一个 xml 文件。虽然每个数据集都有其特殊的存储形式，但由于其存储内容相似、信息类型相似，所以本文提出的信息处理方法同样可以应用于其他类型的胸部 CT 影像数据集以及其他医疗影像数据集。

在该数据集中，一次病例检查的文件夹结构为：LIDC-IDRI-XXXX / Study Instance UID / Series Instance UID / *.dcm *.xml
 XXXX : 病例序号，0000~1018；
 Study Instance UID : 某病例的检查编号；
 Series Instance UID : 某检查的序列编号；
 *.dcm *.xml: 文件夹中包含多个 dcm 格式的文件和一个 xml 格式的文件

LIDC-IDRI 中采用 dicom 格式保存影像。医疗数位影像传输协定（DICOM）是用于医学图像处理、存储和传输的常用标准协定。如果多个医疗仪器都能适配 dicom 格式，那他们之间能够通过传输 dicom 格式的文件来传递医疗影像信息。具体在 LIDC-IDRI 数据集中，一个 dicom 文件存储的影像代表了某病人在某次检测中在某一横切面上的肺部成像。数据集中一个文件夹包含了一次检查中的全部影像信息，即多个 dicom 文件和一个 xml 文件。将多个 dicom 文件组合，就代表了一次检测当中的所有横截面成像，据此可以进行三维重建。LIDC-IDRI 数据集中，每个 dicom 文件中保存的 CT 横截面切片都是长 512，宽 512 的。除此之外，dicom 文件还存储了多种重要元数据，如表 3-2 所示。

表 3-2: dicom 文件重要数据列表

数据属性	作用
OP Instance UID	每张 dcm 切片具有的独一无二的编号
Patient IDs	病例的 ID 号
Slice Thickness	该次检测使用的切片厚度。在检测中，由于器械或设置的不同，会采用不同厚度的横截面切片。但在一次检测中，所有切片间的厚度是保持一致的。
Study Instance UI	某病例的检查编号
Image Position (Patient)	该切片图像的原点（最左上角的点）对应的空间坐标 (x,y,z)，单位：毫米
Slice Location	该横截面在空间 z 坐标轴上的相对位置。单位：毫米。一般情况下，使用 Slice Location 或 Image Position 中的 z 坐标均可以定位切片位置。
Pixel Spacing	像素间隔
Rescale Slope, Rescale Intercept	使用这两个值可以将一些图像的像素转换为 HU 值（CT 值）
Pixel Data	影像数据

由上面的介绍可以知道，三维图像重建可已通过将检查序列的文件夹中所有 dicom 文件进行组合来实现，每个 dicom 文件都包含 z 轴坐标信息（ImagePositionPatient[2] 或 Slice Location），根据 z 轴坐标进行排序并堆叠即可得到三维影像数据。

xml 文件记录的是专家对于该病例一次检查图像的所有病灶标注，其信息格式与一般 xml 文件相同，使用树状节点记录关键信息。其中重要的节点结构数据图 3-4 所示。

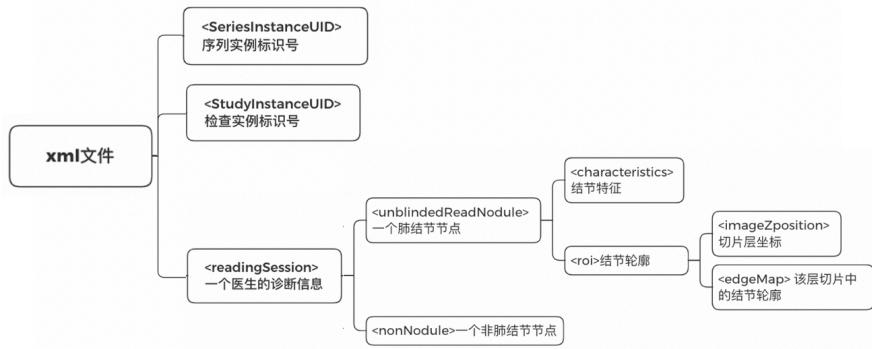


图 3-4: LIDC-IDRI 的 xml 文件结构

<SeriesInstanceUID>: 检查的序列编号。

<StudyInstanceUID>: 病例的检查编号。

<readingSession>: 专家信息标示区域。该节点下的信息代表了某专家对该此检测图像中各类型结节的标注，包括结节轮廓和结节特征属性。

<unblindedReadNodule>: 专家标注出的一个肺结节的信息。

<nonNodule>: 专家标注出的一个非肺结节的信息。

<characteristics>: 结节特征属性

<roi>: 结节轮廓

<imageZposition>: 该切片所在的 z 轴位置

<edgeMap>: 该切片层中的结节轮廓

3.3 本章小结

本章主要介绍了本文扩增病灶的选择以及医疗影像数据集的选择。从深度学习技术在医疗领域的应用出发，结合医学领域需求，本章确定了肺部为本文研究的组织器官类型，确定了肺癌为本文研究的目标病症，确定了肺结节为本文研究的目标病灶，确定了胸部 CT 为本文处理的影像类型。随后，本章介绍了本文所选数据集 LIDC-IDRI 及该数据集的特点，为医疗数据预处理流程提供信息支撑。

第四章 技术关键算法

本技术的代码实现可以分为三个模块：原始数据处理、生成对抗网络生成病灶、质量验证实验。其中，前两个部分属于本技术的核心内容，后一个部分属于为了验证技术成果所进行的验证实验，具体的技术实现结构如图 4-1 所示。

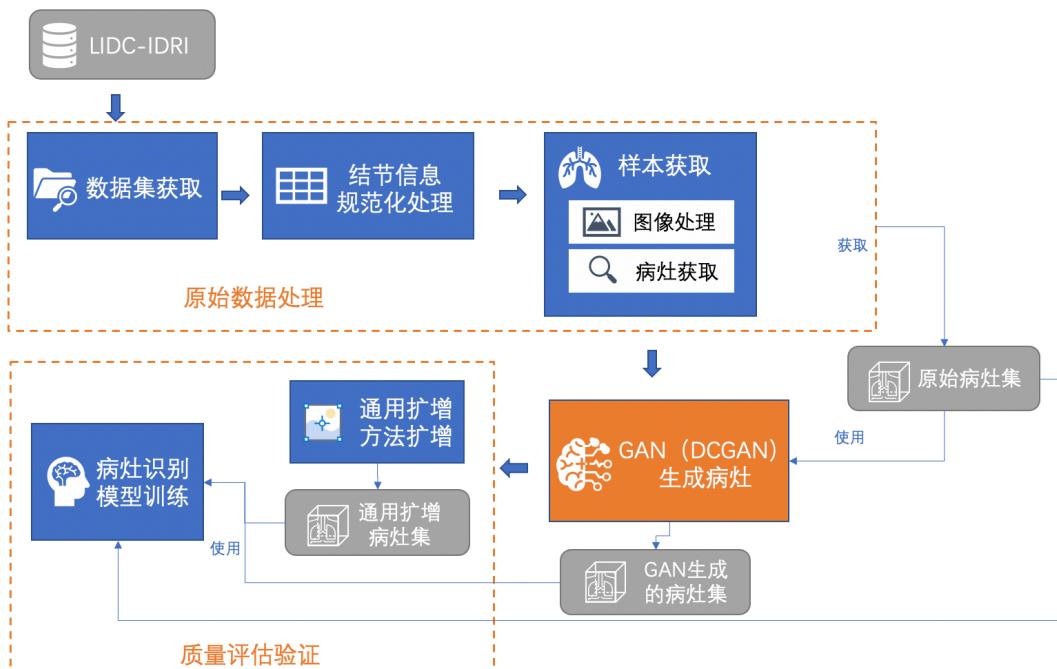


图 4-1: 本文技术结构图

项目中的每个模块中都包含了一到多个关键技术。比如，在将医疗图像用于模型训练前，需要对于原始影像进行处理。不仅需要对数据集进行获取，更重要的是进行结节信息的提取和规范化处理，同时根据提取到的信息进行图像的转换和样本集的生成。生成对抗网络（深度卷积生成对抗网络）是本技术的核心算法，是整个技术能够实现的理论依据，了解其理论及内涵，结合领域知识进行应用是重中之重。同时，为了对技术进行验证和对扩增后的数据集进行质量检验，项目设计了相应的验证实验，其中需要应用到通用图像扩增算法作为对照技术。使用该技术生成的图像作为对照进行模型训练实现质量评估。下

一章还会对验证实验进行更加详细的介绍。本章则针对上述提到的结节信息提取算法、样本获取算法、通用图像扩增算法和生成对抗网络生成结节算法进行详细描述。

为了实现对肺结节数据集的扩增，生成三维的肺结节图像，本文进行了一系列的工作。首先，本文介绍了医疗图像数据集发展现状，描述了目前已有的肺结节数据集，并根据项目需要选择了适合的数据集 LIDC-IDRI。不同的数据集往往拥有不同的数据格式和存储形式，自成一套系统。对于医疗图像数据集而言，也是如此。不同的研究机构提供的数据影像资源，格式、信息具有一定差异性，所以在选定数据集后对于其详情的深入了解与分析必不可少。针对于该数据集，本文详细介绍了其影像的存储格式，信息的描述方式以及元数据的表示含义等等。获取数据集后得到的数据是未经过任何处理的，使用这样的数据进行训练会非常的复杂，需要反复的进行转换和读取操作。另外，图像数据集往往是巨大的，在用于图像变换、深度学习等方面时，其文件大小往往会影响许多操作的可能性，降低操作的效率，耗费大量的时间成本。因此，为了保证较高的效率，就要求尽量优化对数据集的处理流程，提前进行规范化的信息处理。并且，对于医疗图像而言，为了展现更加细微的特征，图像往往需要包含更多的细节，其文件大小更加庞大。任何降低其清晰度的操作，都有可能影响图像所体现的内容，影响判断。与此同时，本次研究使用的是三维图像，相对于常规的二维医疗图像而言，其包含的信息更多、范围更广，更需要在图像集大小的约束下，尽量减少工作量，减少重复的不必要的转换提取工作。所以，本文提前对数据集中的结节影像及信息进行获取，以便之后能够方便的进行训练样本验证样本测试样本的生成，相应算法在本章将进行描述。

经过获取结节信息的流程后，能够生成一组较为规整的、具有一定规则的结节信息数据。但如果要满足模型训练的需要，仅仅如此还不够，还需要进一步结合数据集影像进行结节的提取和转换。这一过程的进行就需要依赖上一步所生成的结节信息。深度学习的训练就是通过使用训练数据集对模型进行训练，使得网络学习到相应参数，从而拥有“智能”。数据质量的好坏决定了模型训练的质量，影响其所具有的能力。这里数据质量不仅代表原始数据集的数据的质量，更代表着研究者进行数据预处理流程的质量。深度学习模型训练通常需要提前将数据划分为训练集、验证集和测试集。训练集是用于网络模型的训练调试过程，验证集用于在训练的过程中验证模型训练效果，测试集用于测试评估模型的最终能力。优秀的样本集设定划分，能够一定程度上提升模型的

泛化能力，避免欠拟合和过拟合的发生。在医疗影像数据集的处理中更需要注意样本集的合理划分，因为它的数据分布与通用图像数据集不同，负样本的数量远多于正样本。与此同时，医疗影像的保存形式与通用图像数据集有较大差异，文件内部往往包含更多注释信息。并且如果要想能够将其用于模型训练，进行下一步的操作，还需要将医疗图像进行像素化。总之，从数据集到可用的样本病灶集，仍然需要进行一系列的处理工作，本章将会对于该过程算法进行介绍。

上一章进行了 GAN 原理的分析以及相关领域成果的介绍，分析了通过 GAN 生成病灶的可能性。而目前，虽然通过对抗生成网络进行二维图像生成的实例较多，但在三维图像生成以及三维医疗图像生成上，GAN 技术还没有普遍应用。本章将会针对三维 GAN 技术，结合病灶生成、肺结节生成的领域知识，对生成对抗网络用于结节生成的算法进行描述。

对于本文所研究技术——生成对抗网络生成三维医疗图像病灶，为了验证技术的可靠性和模型的准确度，本文设计了验证实验。通过对照实验的方法将本文技术与通用扩增方法进行对比，分别使用本文技术和通用扩增方法生成扩增数据集，然后通过组合的方式生成训练样本集，对肺结节识别模型进行训练，对比不同样本集下训练模型的准确度，从而验证方法的有效性。本文的验证实验，不仅能够证明生成对抗网络生成数据集的方法对于模型训练有益，更能证明其在普通方法基础上仍然对训练有益，验证实验的目的着重放在生成对抗网络对于普通方法的补充作用。具体的数据质量验证实验设计将会在下一章进行介绍。但在这之前，有必要对通用图像扩增方法进行了解和选择。图像扩增通用方法种类多样，为了使实验达到更加优秀的质量检验效果，必须要根据方法的分类，结合医疗领域图像特点选用合适的扩增方法，减少前期图像处理带来的图像质量损失。并且，上升到三维图像，其扩增手段应用的可能性与有效性也需要进一步的斟酌。如果使用通用图像扩增手段得到的样本质量较低，对质量评估实验的可靠性将会造成较大影响，从而难以合理验证本文技术的有效性。

上文针对本文技术中的四个关键过程进行了简要论述。本章将针对项目中这几个关键过程所使用的思路、算法以及实现进行详细的介绍。

4.1 结节信息提取算法

结节信息提取的主要工作是获取数据集文件中的结节信息、并规范化保存，为生成实验样本做准备。其目的是通过信息提取和存储来简化后续样本生成的工作，提高效率，降低工作量，也便于对于数据和信息的有效管理。因此，构建结节信息提取算法必须要从模型训练所需的信息和数据集能提供的信息入手，进行分析与设计。

首先，训练样本集的组成要求原始数据集能够提供正负样本。负样本的获取相较于正样本较为简单，因为在整个 CT 图像中绝大部分的区域都属于没有病灶（结节）的，在这些非病灶区域中进行随机截取即可。而正样本则要求区域中含有病灶（结节）部分，且在原始图像中非病灶区域远大于病灶区域。因此，不论是生成正样本还是负样本都要依靠结节的位置信息。LIDC-IDRI 数据集在 xml 文件中提供了每位病例每次检测中的结节位置标注信息。所以，结节信息提取的关键任务就是提取数据集中结节所在的详细位置并规范化保存。

LIDC-IDRI 数据集的文件保存结构如下：LIDC-IDRI-XXXX / Study Instance UID / Series Instance UID / *.dcm *.xml
XXXX : 病例序号, 0000~1018;
Study Instance UID : 某病例的检查编号;
Series Instance UID : 某检查的序列编号;
*.dcm *.xml: 文件夹中包含多个 dcm 格式的文件和一个 xml 格式的文件

通常一个病例的文件夹 LIDC-IDRI-XXXX，包含多个以病例的检查实例号命名的文件夹，每个文件夹代表该病人的一种检查实例。在每种类型的检查目录下，还包含多个以检查的序列实例号命名的文件夹，代表该可能存在多个检查影像。每个序列文件夹下包含一个医疗影像的多个 dicom 文件和 xml 文件，用来存储影像和描述影像信息，如图 4-2所示。因此序列实例号和影像是一一对应的。虽然每个序列文件夹下，只包含一个影像且都是 dicom 类型的文件，但是这些影像对应的检查种类各不相同。在 LIDC-IDRI 中包含四种检查类型，分别是：CT（计算机断层扫描）、SEG（DICOM 分割）、CR（计算机 X 线成像）和 DX（数字化 X 线成像）。本文选用的是 CT 类型的图像，因此在进行信息筛选时，需要筛选出图像类型是 CT 类型的数据。

除了病例、检查、影像的一对多关系之外，每个影像中可能包含多个结



图 4-2: 某检查序列的文件夹示例图

节，它们同样是一个一对多的关系，如图 4-3 所示。在样本获取过程中，每个正样本的原始数据都与一个实际的结节对应的，是一对一的关系。因此，信息提取获得的每条数据也同样是以结节作为单位的，一条数据将会对应一个结节，包含多项结节有关数据。

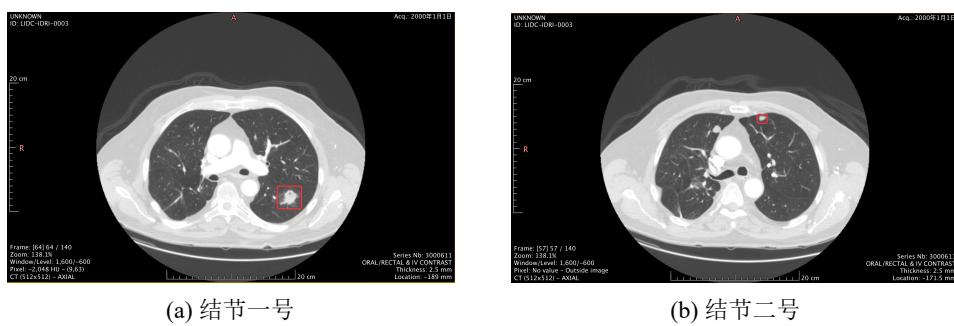


图 4-3: 3000611 检查中的多个结节

第三章介绍了 dicom 文件中包含的多种影像的元数据，这些数据能够辅助进行病灶信息处理，比如有：病例的检查实例号 (Study Instance UI)、切片厚度 (Slice Thickness)、像素间隔 (Pixel Spacing)、用于转换为 HU 值的数据 (Rescale Intercept、Rescale Slope) 等。在信息获取时，可以通过

SimpleITK^①得到 dicom 中的这些关键信息。同理，在下一步的样本获取流程中，也可以通过此方法进行信息获取，实现读取、重建、展示以及转化，详细内容将在下一小节中介绍。

在 LIDC-IDRI 数据集中，dicom 文件着重于对影像本身的描述，而 xml 文件则是着重记录该影像中结节的相关信息。XML 是一种标记语言 [59]，主要用于存储、传输数据。在 xml 文件中，标记成对出现，形成一个一个节点，文件整体类似树形结构。在 LIDC-IDRI 的 xml 文件中，通过节点的组合，记录了结节的关键数据。比如：检查的序列实例号可以用以定位是哪一个影像；readingSession 包含一个医生对该影像中结节的所有标注信息；readingSession 下又包含多个肺结节节点（包含假阳性），每个肺结节节点下又有具体的特征及轮廓描述，如图 4-4 所示。xml 文件中的信息非常详细，进行一定的加工、计算便可以获得结节数据。

```

<?xml version="1.0" encoding="UTF-8"?>
<LidcReadMessage uid="1.3.6.1.4.1.14519.5.2.1.6279.6001.1307390687803.0"
xmlns="http://www.nih.gov" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.nih.gov
http://troll.rad.med.umich.edu/lidc/LidcReadMessage.xsd">
<ResponseHeader>
<Version>1.8.1</Version>
<MessageId>-421198203</MessageId>
<DateRequest>2007-11-01</DateRequest>
<TimeRequest>12:30:44</TimeRequest>
<TaskDescription>Second unblinded read</TaskDescription>
<SeriesInstanceUid>1.3.6.1.4.1.14519.5.2.1.6279.6001.1790493736364387050597206031
92</SeriesInstanceUid>
<DateService>2008-08-18</DateService>
<TimeService>02:05:51</TimeService>
<ResponseDescription>1 - Reading complete</ResponseDescription>
<StudyInstanceUID>1.3.6.1.4.1.14519.5.2.1.6279.6001.2988061372886334532469756301
78</StudyInstanceUID>
</ResponseHeader>
<readingSession>
  <annotationVersion>3.12</annotationVersion>
  <servicingRadiologistID>540461523</servicingRadiologistID>
  <unblindedReadNodele>
    <noduleID>Nodule 001</noduleID>
    <characteristics...>
      <roi...>
      <roi...>
      <roi...>
    </unblindedReadNodele>
    <unblindedReadNodele...>
    <unblindedReadNodele...>
    <nonNodule...>
  </readingSession>
  <readingSession...>
  <readingSession...>
  <readingSession...>
</LidcReadMessage>

```

图 4-4: xml 文件详情

^①<https://simpleitk.org/>

具体的信息提取算法如下：

1. 根据数据库元数据文件遍历所有影像
2. 根据 xml 文件，筛选 CT 类型影像
3. 根据 xml 文件，获取信息。
 - (a) 获取序列实例号。
 - (b) 获取结节轮廓位置。遍历 readingSession 节点，获取每名医生的诊断信息。在此节点下，有多个肺结节节点 <unblindedReadNodule>，每个节点代表一个肺结节诊断信息，肺结节轮廓的标记是按照切片来记录的，该节点下每个 <roi> 节点代表本结节在 z 轴切片上的信息。每个切片上，轮廓是以 (x,y) 坐标形式记录的。遍历 <roi> 节点，可以获得结节在 z 轴上的最大坐标及最小坐标，通过计算可以获得结节在 x 轴、y 轴上的最大坐标及最小坐标，形成标记框。最值进行计算又可以获得长度差、宽度差和高度差。
 - (c) 不同医师会有重复的结节标记，也会有不同的结节标记，记录结节被医师标记的次数，可以选择更加具有真实性的结节。
4. 根据 dicom 文件，使用 SimpleITK 获得切片厚度、像素间隔等数据
5. 数据保存

通过以上算法，能够将 xml 和 dicom 文件中的结节信息提取出来。将得到的信息进行规范化处理并保存，用 xml/csv 文件进行存储。获取到的结节信息如表 4-1 所示。

表 4-1: 结节信息

编号	结节信息 NoduleInfo
1	结节号 (NoduleId)
2	检查实例号 (StudyId)
3	序列实例号 (SeriesId)
4	结节最小 X 坐标 (minX)
5	结节最大 X 坐标 (maxX)
6	结节最小 Y 坐标 (minY)
7	结节最大 Y 坐标 (maxY)
8	结节最小 Z 坐标 (minZ)
9	结节最大 Z 坐标 (maxZ)
10	层间距 (distance)
.....	

4-5是 xml 文件解析的部分代码。代码通过查找 xml 文件中的重要节点来获取数据，能够得到检测序列号以及对应的多个结节信息。

```
def xml_info(xml_path):
    with open(xml_path, 'r') as xml_file:
        markup = xml_file.read()
        xml = BeautifulSoup(markup, features="xml")
        # 提取患者号
        patient_id = xml.LidcReadMessage.ResponseHeader.SeriesInstanceUid.text
        # 提取专家对该病例做出的诊断
        reading_sessions = xml.LidcReadMessage.find_all("readingSession")

        dict_info = {}
        for reading_session in reading_sessions:
            nodules = reading_session.find_all("unblindedReadNodule")
            for nodule in nodules:
                nodule_id = nodule.noduleID.text
                rois = nodule.find_all("roi")

                minX = sys.maxsize
                maxX = -sys.maxsize - 1
                minY = sys.maxsize
                maxY = -sys.maxsize - 1
                for roi in rois:
                    imageZposition = roi.imageZposition.text
                    edgeMaps = roi.find_all("edgeMap")
                    for edgeMap in edgeMaps:
                        x = int(edgeMap.xCoord.text)
                        y = int(edgeMap.yCoord.text)
                        minX = min(x, minX)
                        maxX = max(x, maxX)
                        minY = min(y, minY)
                        maxY = max(y, maxY)
                        # 其他信息...
                dict_info[nodule_id] = [minX, maxX, minY, maxY]
    return dict_info
```

图 4-5: 结节信息读取代码

4.2 样本获取算法

深度学习模型训练的效果很大程度上取决于训练样本的质量，所以在样本获取这一步格外重要。在进行上一流程获取到结节信息之后，便可以将其与原数据集结合进行样本获取。医疗数据集不同于通用图像，其图像格式、所含信息都有其独特性。并且从二维图像上升到三维图像，其涉及的图像信息处理又

变得更加复杂，因此在将医疗数据集影像转换为训练样本时，要首先了解医疗图像的特点。

医疗成像是通过非侵入方式获得人体或人体某部位的内部组织影像。根据观测影像的信号能反推展现人体组织的特性，反映在图像上展示某人体部位的结构和特征。本文使用的计算机断层扫描，又称计算机体层成像，是使用 X 射线对人体部位进行扫描，根据不同方向得到的数值，形成该部位的一系列横向切片，这些切片图像也称作断层。因此，正如 LIDC-IDRI 中数据存储的形式一样，它是通过一层一层的横切面，即二维图像形式来保存的。但是在医疗影像的观测当中，往往需要结合冠状位、矢状位、轴状位三个角度来观察来分析，如图 4-6 所示，因此才出现了众多以三维医疗图像为研究对象的实验。

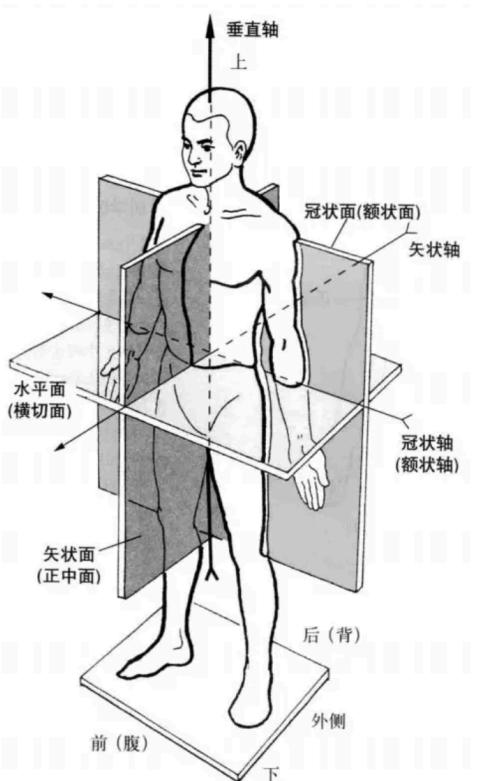


图 4-6: 人体基本面

Dicom 文件实际上也满足三维图像重建的需要。dicom 文件根据人体确定了坐标系，建立以指向病人左侧的方向为 x 轴正方向、以指向病人背部的方向为 y 轴正方向、以指向病人头部的方向为 z 轴正方向的坐标系，如 4-7 中所示。一次检测中的多个 dicom 文件其实就是沿着 z 轴的一组横切面。dicom 文件中提供了层间距信息，为三维图像的重建提供了更多可能性。

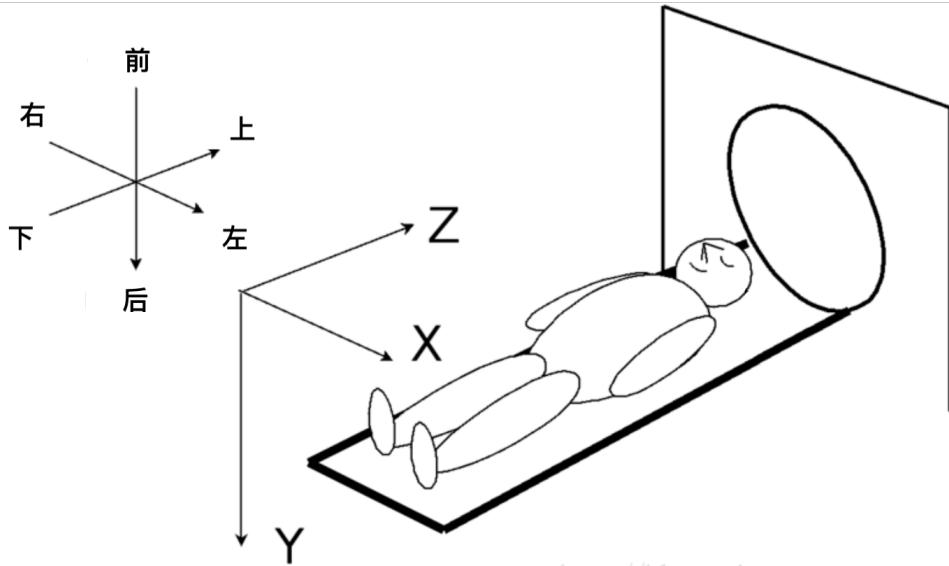


图 4-7: Dicom 坐标系

事实上，dicom 文件中还存储记录了图像位置、图像方向、像素间距等属性。图像位置代表了影像最左上角的像素的的 x、y、z 坐标，图像方向代表了图像的行向量和列向量三维方向上的余弦。根据图像位置，可以建立以此为原点的坐标系，再结合图像方向便可以找到每一个图像像素对应的三维坐标。而在肺结节的提取中，因为样本的提取只关注样本周围的立方体，并且关注其形态大小而非方向角度，所以在本项目的图像转换中并不需要对于原图像进行坐标轴的改动。因此，在进行样本生成时，项目选择统一将多个 dicom 文件按照 z 坐标进行排序，并垂直 z 轴进行横切平面的叠加，从而实现简单的三维 CT 影像重建。

三维图像中体素的数值能代表不同的特性。在 CT 扫描中，这些值其实就是亨利单位（Hounsfield unit，简称 HU），又被称为 CT 值，表示身体对于 X 光的吸收程度，用以测定人体某组织器官的密度大小，范围是-1024-3071。亨氏单位为其所测量之介质的线性衰减系数的线性映射，其定义将蒸馏水与空气在标准温度与标准压力下的阻射率分别设为 0 HU 和-1,000 HU。因此，在线性衰减系数介于平均值间的体素中，其亨氏单位 [60] 为：

$$HU = 1000 \times \frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}} \quad (4-1)$$

其中 μ_{water} 与 μ_{air} 分别为水与空气的线性衰减系数。

HU 值通常用于肿瘤等组织的辨识 [61]，它也是医学图像处理和使用中最

常用的数据形式。但由于不同仪器或机构的标准不同，不同 dicom 文件提供的值的类型不同，一些是 CT 值，而一些是灰度值（CT 影像是灰度图像）。一般来说，通过按对数关系在白色与黑色之间分成若干级可以得到灰度等级。灰度值的数值范围一般从 0（黑色）到 255（白色）。本文使用标准的 CT 图像，它是 12 位灰度图像。因为灰度图像是单通道的，灰度图像的像素值就等于其灰度值，所以一些 dicom 中提供的灰度值也可以称之为像素值。为了满足医学图像处理的需要，必须要对这些灰度值（像素值）进行转化，变为 HU 值（CT 值）。

已知，HU 值的计算公式：

$$HU = pixel * slope + intercept \quad (4-2)$$

pixel: 灰度值（像素值）

intercept: dicom 文件中的 *RescaleIntercept*

slope: dicom 文件中的 *RescaleSlope*

事实上，在利用公式进行转化之前，还需要判断该 dicom 文件像素数组是否需要进行转化。根据 dicom 文件规则，如果 *slope* = 1, *intercept* = 0 则不需要进行转化。

经过上述的操作，可以得到统一化的由二维图像重建生成的三维 CT 影像。至此，结合之前筛选出的结节信息可以实现样本获取。

图像本质其实就是矩阵，三维图像对应的便是三维矩阵。三维矩阵可以由二维矩阵叠加生成，相应的，三维图像同样可以由一系列二维图像生成。所以，三维肺部图像重建的关键就是处理好二维肺部图像切片的次序，之后进行组合。因此，样本获取过程中的关键环节是读取 dicom 文件，并获取每个文件的层坐标进行升序排列，之后进行二维图像的叠加，形成一个三维的图像。这时，再依照结节信息提取缓解生成的结节信息文件，能够方便地找到本次的检测影像中包含的所有结节，以及他们的轮廓信息。根据每个结节的 X 轴最大值和最小值、Y 轴的最大值和最小值以及 Z 轴的最大值和最小值就能够获得一个三维的立方体结构。每个这样的三维立方体都包含了一个阳性肺结节。本文实验需要尺寸大小为 32*32*32 的三维肺结节图像作为训练样本输入到三维卷积生成对抗网络中进行训练。根据生成的结节信息文档可知，数据集中所有肺结节的大小都不超出 32*32*32。所以，结合结节信息文档中的轮廓信息与重建得到的一系列三维胸部 CT 影像，能够获得训练所需的样本。

图 4-8 是三维重建的关键代码。它实现了将一次检测中的一组二维肺部 CT 图像重建成三维肺部 CT 图像。

```

for directorypath in directorypaths:
    filebroken=False
    series_ID = sitk.ImageSeriesReader.GetGDCMSeriesIDs(directorypath)[0]
    if os.path.exists(savePath + series_ID):
        print(series_ID,"file already exist")
        exist=exist+1
        continue

    # 所有切片路径
    #选取其中一个序列ID,获得该序列的若干文件名
    dicom_names = reader.GetGDCMSeriesFileNames(directorypath,series_ID)

    # 读取单个文件
    imageFileReader = sitk.ImageFileReader()
    imageFileReader.GetLoadPrivateTags()

    dict_img = {}
    for dicomFileName in dicom_names:
        imageFileReader.SetFileName(dicomFileName)
        image = imageFileReader.Execute()

        if '0020|1041' in imageFileReader.GetMetaDataKeys():
            location = float(imageFileReader.GetMetaData('0020|1041'))
            imageArray = sitk.GetArrayFromImage(image)[0]
            dict_img[location] = imageArray
        else:
            print(dicomFileName+"not exist location")
            fail=fail+1
            filebroken=True
            break

    if not filebroken:
        with open(savePath + series_ID, 'wb') as handle:
            pickle.dump(dict_img, handle, protocol=pickle.HIGHEST_PROTOCOL)
        success=success+1

    print("success",success)
    print("fail",fail)
    print("already exist",exist)

```

图 4-8: 三维重建实现代码

样本获取阶段主要进行了三方面的工作：首先使用二维胸部 CT 切片进行三维肺部图像的重建，之后根据 dicom 文件中的属性值进行像素值与 CT 值的转化，最后结合结节信息文档和三维肺部图像得到用于卷积生成对抗网络训练的大小为 32*32*32 的结节样本。这里展示一些实验中获取到的三维肺结节样

本，如 4-9 所示。

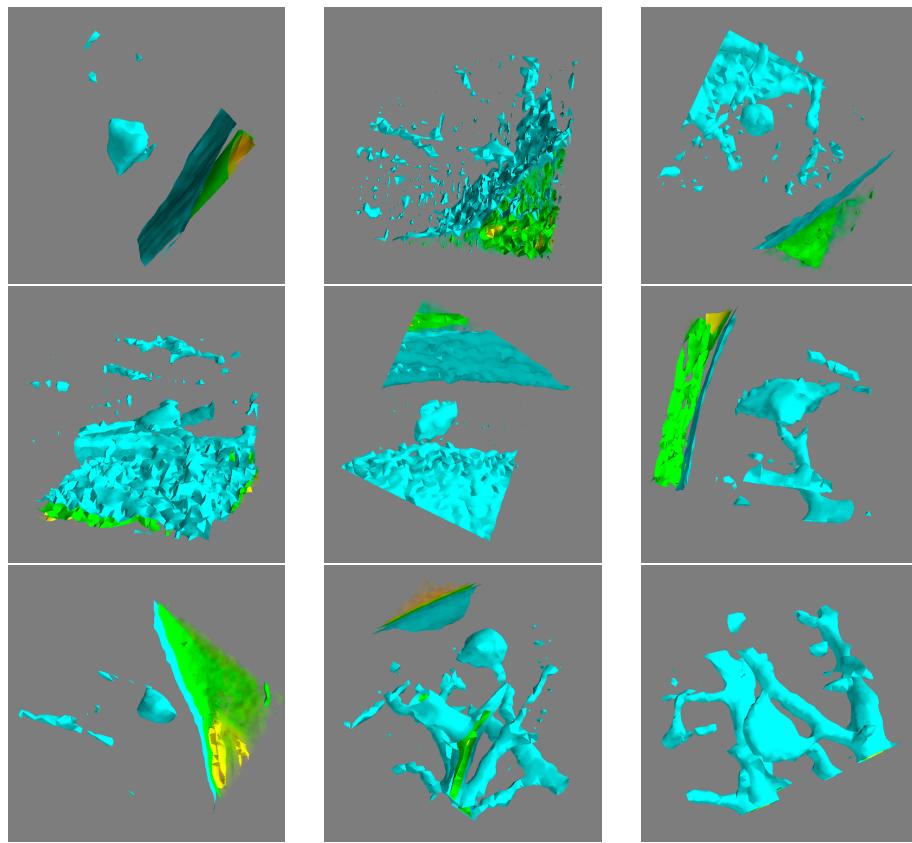


图 4-9: 获取的原始样本

4.3 通用图像扩增算法

前文介绍了通用图像扩增方法的多种分类方式：根据一次扩增变换中使用到的样本数量为标准，图像扩增可以分为单样本扩增和多样本扩增；根据进行扩增的图像变异范围不同，图像扩增可以分为针对图像本身的整体扩增和针对病灶（或非病灶）部分的局部扩增；根据对像素点的不同变异方式，可以粗略将图像扩增分为仿射扩增和弹性扩增。结合胸部 CT 图像和肺结节影像的特点，能够大致确定需要使用的通用扩增手段，即针对单图像样本进行的整体扩增，其中将使用仿射扩增和弹性扩增中的图像处理方法。仅针对扩增手段进行分析还不够，在选用具体的仿射扩增和弹性扩增方法时，必须结合肺部 CT 图

像特点进行。以下是几个经过综合分析最终选用的方法。

首先，数字图像是由一个个像素点组成的，它可以用矩阵来表示，图像的每个像素点都与一个矩阵元素对应。对灰度图像而言，图像矩阵中元素的值其实就是对应图像该位置像素的灰度值。二维医疗灰度图像可以用一个二维矩阵来表示。图像的高度与矩阵的行长相等，图像的宽度与矩阵的列长相等。本文使用肺部 CT 形成三维灰度图像，同理这个三维灰度图像可以使用三维的矩阵来表示。不论二维还是三维图像，因为它们均可以用矩阵来表示，所以想要对图像进行变换，其实都可以采用矩阵变换来实现。

由于结节的角度变化并不会对其的性质产生本质变化，所以本文选用了旋转和反转这两种方法。它们既不会对结节大小产生变动（缩放等方法可能会引起），又不会导致样本中结节关键部分缺失（平移、裁剪等方法可能会引起）。因为三维图像的仿射变换是相对于三维坐标轴原点和坐标轴进行的变换。齐次坐标对于仿射变化非常方便，三维空间中点的变换可以表示为点的齐次坐标与四阶的三维变换矩阵相乘，

4.3.1 旋转

针对结节图像，本文采用旋转的仿射变换，原因已在上文中分析。这里的旋转特指将样本沿 x 轴、 y 轴、 z 轴分别进行旋转，改变位置但整体形状不变。在三维图像旋转中有右手定则：当使用右手拇指指向围绕旋转的坐标轴的正向，其他四根手指弯曲的方向代表了这次旋转角的正向。

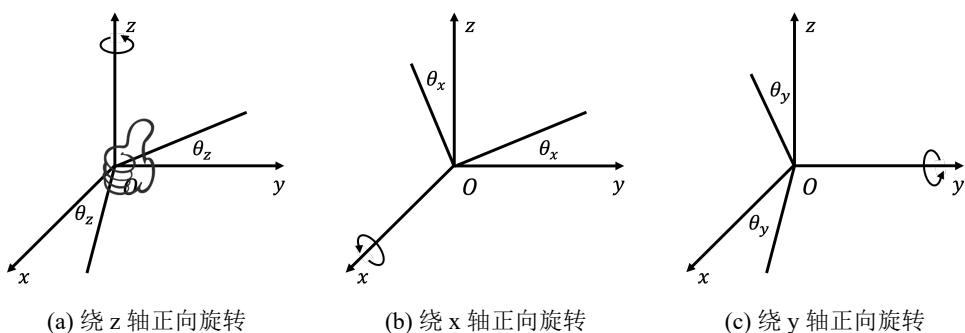


图 4-10: 绕坐标轴旋转

这里以绕 x 轴为例来分析变换式子。如果有 $P'(x,y,z)$ 是 $P(x,y,z)$ 绕 x 轴旋转 θ_x 度得到的点，经过这个变化 x 坐标未发生变化，在 yoz 平面上发生了绕原点 o 的旋转，角度为 θ_x ，因此有转换式子：

$$\begin{aligned}x' &= x \\y' &= y \cos \theta_x - z \sin \theta_x \\z' &= y \sin \theta_x + z \cos \theta_x\end{aligned}\tag{4-3}$$

有矩阵变换式子：

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}\tag{4-4}$$

旋转矩阵为

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix}\tag{4-5}$$

绕 y 轴, z 轴同理, 只需与变换矩阵进行乘法便可获得新的旋转图像, 实现代码如 4-11。

```
# 旋转
def img_rot(image, rnum, axes): # img rnum 旋转几次 axes 轴
    if axes == 0:
        # 绕 z 轴
        image = np.rot90(image, rnum, axes=(0, 1))
    elif axes == 1:
        # y
        image = np.rot90(image, rnum, axes=(0, 2))
    elif axes == 2:
        # x
        image = np.rot90(image, rnum, axes=(1, 2))
    return image
```

图 4-11: 旋转实现代码

据此, 本文使用 python 语言对原样本矩阵进行矩阵变换, 得到旋转后的图像。为了直观地展示, 项目使用 Mayavi^①和 VTK^②实现了三维图像可视化, 下列三维图像展示均是基于该技术。

^①<https://docs.enthought.com/mayavi/mayavi/>

^②<https://docs.enthought.com/mayavi/vtk/>

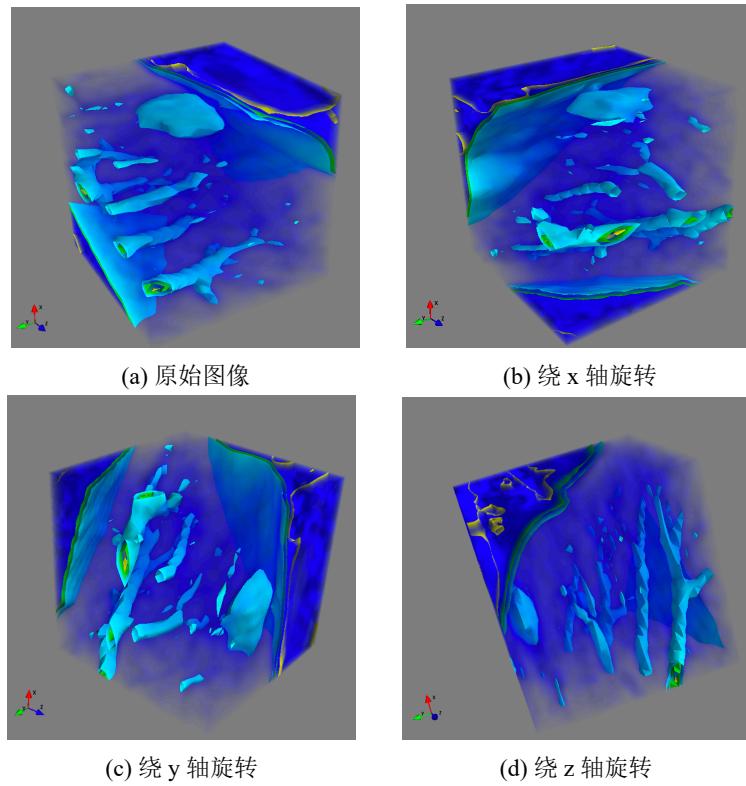


图 4-12: 肺结节样本旋转变换

4.3.2 翻转

与旋转同理，结节对称的变换不会改变其内部特征和性质。本文采用的翻转是关于坐标平面 xoy 和 xoz 的翻转，即上下翻转和左右翻转。对称变换的坐标变换简单，例如在进行上下翻转时有：

$$[x', y', z', 1] = [x, y, -z, 1] \quad (4-6)$$

有变换矩阵

$$T_{XOY} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4-7)$$

实现代码如 4-13 所示。

```
# 翻转
def img_flip(image, dir): # img dir方向 0上下 1左右
    if dir == 0:
        image = np.flipud(image)
    elif dir == 1:
        image = np.fliplr(image)
    return image
```

图 4-13: 翻转实现代码

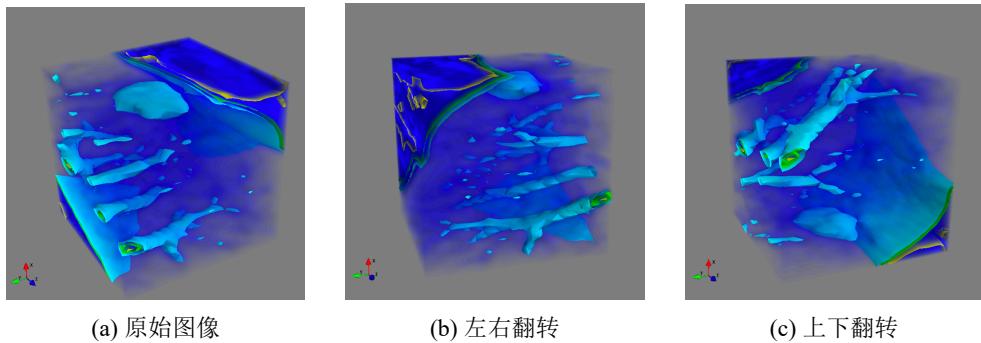


图 4-14: 肺结节样本翻转变换

4.3.3 亮度

亮度变换是常用的图像扩增手段，适用于常见的各个领域。具体到医疗图像领域，通过改变影像的亮度，不仅可以模拟不同仪器的质量差异，还可以模拟环境因素导致的影像采样差异。亮度度量了眼睛对光的强度的感受。人眼不仅接受发光体发射的光线，同时也感受着物体反射光。图像扩增方法中常用的就有对图像进行随机的亮度调整。这种方法属于弹性变换。但如果在进行随机亮度调整时变化幅度过大，那么将可能影响到图片的其他属性，比如饱和度。根据亮度定义，在更改一个像素点的亮度时，可以采取以下公式，进行简单的线性像素变换：

$$f'(x, y, z) = a * f(x, y, z) + b \quad (4-8)$$

其中 $f'(x, y, z)$ 代表原像素值，改变 a 可以改变图像对比度，改变 b 改变图片亮度。

实现代码如 4-15 所示。

```
# 亮度
def img_brightness(image, brightness):
    contrast = 1 # 对比度
    pic_turn = cv2.addWeighted(image, contrast, image, 0, brightness)
    # cv2.addWeighted(图像1,图像1权重,图像2,图像2权重, 亮度) 实现的是图像透明度
    # 的改变与图像的叠加

    return pic_turn
```

图 4-15: 亮度变换实现代码

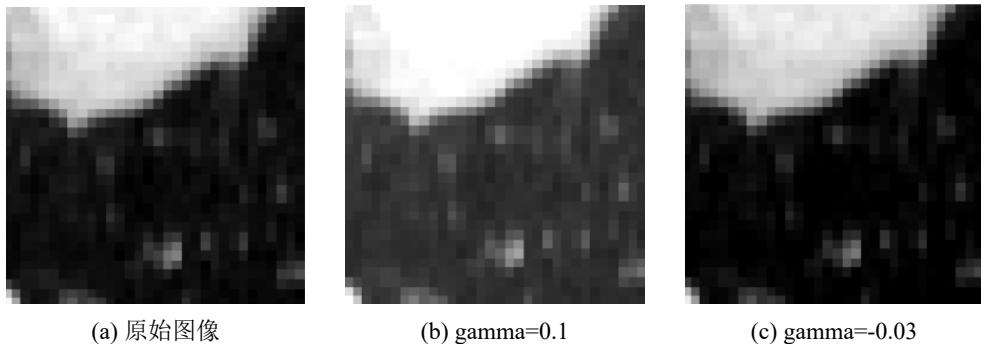


图 4-16: 肺结节样本亮度变换

4.3.4 噪声

医疗图像一般都会受到一定的噪声影响，通过为图像添加噪声能够模拟更多合理样本。添加噪声的操作非常简单非常方便，噪声干扰的随机性使得它可以生成大量有效数据。同时，实验也已证明了随机添加噪声方法对于提升模型的泛化能力非常有效。噪声的种类非常多样，本文根据医疗领域影响特征，结合实际情况，选择高斯噪声、椒盐噪声和泊松噪声对图像进行随机干扰。

高斯噪声的概率密度函数遵循正态分布。假设随机变量 X 呈位置参数为 μ 、尺度参数为 σ 的正态分布，分布可以记为：

$$X \sim N(\mu, \sigma^2) \quad (4-9)$$

其概率密度函数为：

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4-10)$$

高斯噪声是分布比较密集的噪声类型，其噪点深度具有随机性，使用高斯噪声干扰图像能够降低一些无用的高频特征对模型产生的负面影响，从而促进

模型对图像中真正核心的关键特征进行有效识别。所以，本文在图像上增加均值为 0 的高斯噪声，提升模型泛化能力。

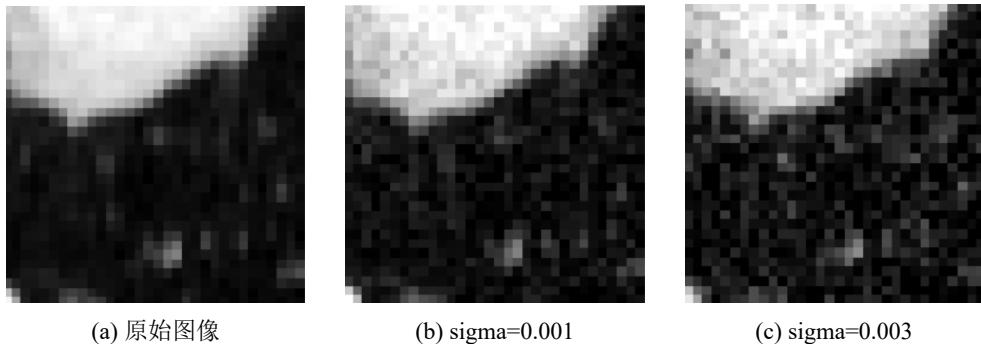


图 4-17: 肺结节样本添加高斯噪声

椒盐噪声，又称为脉冲噪声，同它的名字一样，它包含两种噪声类型，一种叫椒噪声，一种叫盐噪声。其中，椒噪声是黑色像素点，盐噪声是白色像素点。椒噪声是低灰度的，盐噪声是高灰度。这种噪声的出现往往是由于采集时影像传输受到了一些干扰。椒噪声和盐噪声都是在图像中随机出现的，像素深度基本不变。

在添加椒盐噪声干扰时，有一个非常重要的变量叫做信噪比（Signal Noises Rate），他是信号功率与噪声功率之比。一般情况下，为了方便计算，会令信号均值为 0，直接使用信号方差比上噪声方差来衡量噪声干扰程度。如下公式所示：

$$SNR = \frac{\sigma_{\text{signal}}^2}{\sigma_{\text{noise}}^2} \quad (4-11)$$

其中 σ_{signal}^2 代表信号方差， σ_{noise}^2 代表噪声方差，SNR 就是信噪比的简称 $([0, 1])$ 。

对于灰度图像而言，其信噪比可以表示为更加容易计算的形式：

$$SNR = \frac{\text{无噪声图片像素点的灰度值之和}}{|\text{噪声图片的灰度值之和} - \text{无噪声图片的灰度值之和}|} \quad (4-12)$$

$$NP = SP * (1 - SNR) \quad (4-13)$$

其中，NP 代表加噪的像素数量，SP 代表总像素数，SNR 代表信噪比

([0, 1])。

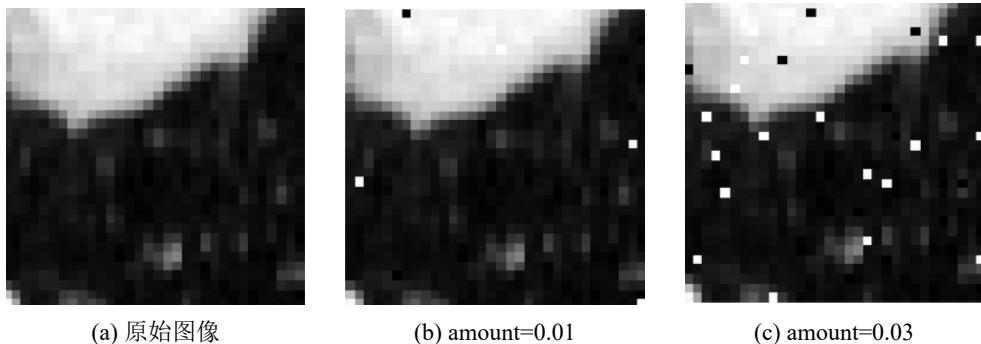


图 4-18: 肺结节样本添加椒盐噪声

在进行实验观测中，当电路中的电子或光学仪器中的光子等粒子的数量较少时，进行数据读取会读出散粒噪声，又称之为泊松噪声。该噪声其实是由于光的粒子性造成的。当光源发出光打在 CMOS（一种可以作为感光元件的半导体）上会形成光点，CMOS 接收到的光子越多，像素灰度值越大。但是从发射到接收这个过程中会出现一些干扰，使得接收到的光子数不稳定，这种噪声其实是数据读取中统计涨落的可视化的表现。这种噪声符合泊松分布所以又称为泊松噪声。当粒子数非常多时，分布就会趋向于正态分布。此时，泊松噪声的标准差就等于平均粒子数的平方根，所以有信噪比公式：

$$SNR = \frac{N}{\sqrt{N}} = \sqrt{N} \quad (4-14)$$

其中，N 代表接收的平均粒子数，SNR 代表信噪比。

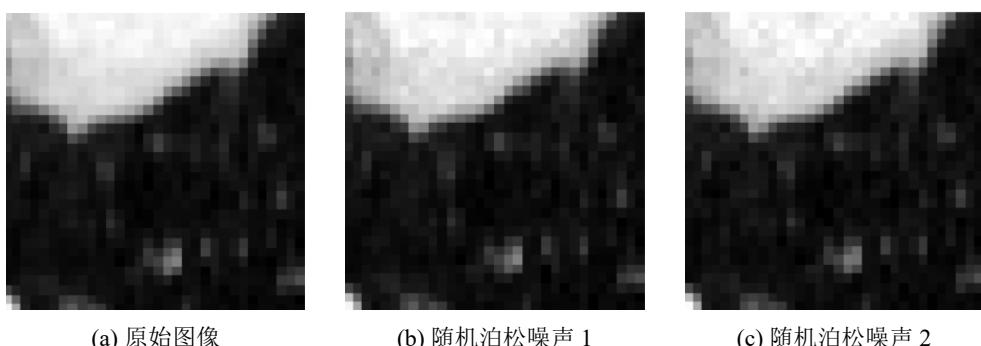


图 4-19: 肺结节样本添加泊松噪声

实现代码如 4-20 所示。

```
# 噪声
# 1.高斯噪声
# 高斯噪声是指概率密度函数服从高斯分布的一类噪声。如果一个噪声的幅度分布服从高斯分布，而它的功率谱密度又是均匀分布的，则称它为高斯白噪声。
# input:均值为means,标准差为sigma
def img_noise_gaussian(image, mean, sigma):
    image_modified = util.random_noise(image, mode='gaussian', seed=None,
clip=False, mean=mean, var=sigma)
    return image_modified

# 2.椒盐噪声
# 椒盐噪声是由图像传感器、传输信道、解码处理等产生的黑白相间的亮暗点噪声。椒盐噪声是指两种噪声，一种是盐噪声，另一种是胡椒噪声。前者是高灰度噪声，后者属于低灰度噪声。一般两种噪声同时出现，呈现在图像上就是黑白杂点。
# input:amount 噪声占比
def img_noise_saltpepper(image, amount):
    image_modified = util.random_noise(image, mode='salt&pepper', seed=None, clip=False,
amount=amount)
    return image_modified

# 3.泊松噪声
# 泊松噪声就是符合泊松分布的噪声模型，泊松分布适合于描述单位时间内随机事件发生的次数的概率分布。
def img_noise_poisson(image):
    try:
        image_modified = util.random_noise(image, mode='poisson', seed=None,
clip=False)
    except ValueError:
        print('ValueError: lam < 0 or lam contains NaNs')
        return image
    return image_modified
```

图 4-20: 噪声实现代码

至此，本文介绍了可以用于进行三维医疗图像扩增的一般性方法。这些手段生成的图像将会应用在验证实验中，用来作为对照技术与本文的对抗生成网络技术进行对比，从而验证本文技术的有效性。

4.4 3D-DCGAN 生成结节算法

4.4.1 卷积生成对抗网络技术

Alec Radford 等人在原始 GAN 的基础上提出了 DCGAN，即在生成器和判别器特征提取层用卷积神经网络代替了原始 GAN 中的多层感知机，使得生成器能够产生质量更优秀的图像，学习到影像中的深度特征。由于医疗图像生成对质量要求高，更需要网络能够学习到深层图像特征，生成较高质量的医学影像，保证医学领域对于影像的高精度要求，所以选择 DCGAN[62] 来进行病灶的扩增生成更加合适。DCGAN 的公式原理与 GAN 近似，所以在此不做赘述。仅针对 DCGAN 与 GAN 的不同之处进行描述。

卷积神经网络 [63]（简称 CNN）通常用于图像识别，它是依据生物视觉处理过程为启发提出的。它由卷积层、池化层、全连接层组成，能够解决全连接神经网络效率低下、过拟合等问题。在卷积神经网络中，卷积可以看作是对特征的一种提取，将图像中蕴含的大维度的信息变为小维度的信息。神经元不使用全连接方式，而是与前一层中的局部部分进行连接，只感知局部的图像信息。之后只要将局部的图像信息结合起来，就能实现对全局信息的把握。这里使用到的思想就是局部视野感知。卷积神经网络中还有一个重要概念，就是权值共享。其含义就是将部分区域提取出的特征应用在其他区域。在卷积提取特征之后，网络将会使用学习到的特征进行分类。为了降低数据空间大小和参数数量，降低复杂度，避免过拟合，卷积神经网络使用池化的思想，对局部区域的特征进行聚合。CNN 中在卷积层中加入了池化层，不仅能降低计算量，也能降低卷积层的边缘敏感性。池化是非线性形式的降维采样，具体手段取决于使用的池化函数。全连接层中的神经元与上一层中的所有激活都有联系，通过矩阵乘法加偏差量实现激活，与全连接神经网络相同。

CNN 在图像识别中有着优秀的表现，并且生成对抗网络中的判别器实际上的功能就是进行识别，这点与 CNN 功能一致。因此 DCGAN 结合了 GAN 和 CNN，将 GAN 中生成器和判别器换为卷积网络，将卷积网络的优势融入对抗生成网络中。在 DCGAN 中以 GAN 为基础进行了一些调整：

1. 生成对抗网络中的生成器 (G) 网络和判别器 (D) 网络使用卷积神经网络实现
2. 将 G、D 网络中的池化层改为卷积层，判别器使用步长卷积，生成器使用小

数步长卷积

3. 取消完全连接层
4. G 网络中输出层使用 Tanh 作为激活函数，其他层使用 ReLU 作为激活函数
5. D 中所有层都使用 ReLU 作为激活函数

上采样和下采样是 DCGAN 中的重要概念。下采样又称为降采样，是通过对原图像变换缩小图像，生成符合显示区域大小的图像或生成缩略图。在卷积网络中，卷积层通过下采样进行特征提取，而池化层则通过下采样进行降维。DCGAN 中，为了使判别器能够自己学习下采样，将判别器网络中的池化层去掉，改为卷积层实现。DCGAN 判别器使用步长卷积进行采样的过程就是卷积层进行正向卷积的过程，可以使用矩阵运算来表示：

$$Y = CX \quad (4-15)$$

其中，X 代表输入的元素矩阵，Y 代表输出的元素矩阵，C 代表卷积矩阵，对卷积核 kernel 进行重新排列可以得到。通过定义卷积核、步长和填充就能够对输入图像进行卷积。卷积输入输出尺寸关系如下：

$$o = (i + 2p - k)/s + 1 \quad (4-16)$$

其中 o 代表输出矩阵尺寸，i 代表输入矩阵尺寸，p 代表填充，s 代表步长。

与下采样对应的概念是上采样，它主要是通过对原图像变换来放大图像，生成更高分辨率的图像，常见方法有内插值、反卷积和反池化。在 DCGAN 中，生成器就是通过上采样来使用低维度特征生成图像的。DCGAN 同样使用卷积层代替池化层，让网络学习上采样能力。这里使用反卷积来实现上采样，从而生成完整尺寸的图片。DCGAN 使用小数步长卷积方法 [64] 进行上采样。在进行正向卷积时，如果有步长 $s > 1$ ，那么它对应的反卷积步长 $s' < 1$ 。所以在进行反卷积时，可以通过以下方式实现。即在输入矩阵的值之间插入 $s-1$ 个 0，将其当作新的输入，此时步长 $s' = 1$ ，有反卷积层输入与输出得尺寸关系：

$$o' = s(i' - 1) + k - 2p \quad (4-17)$$

其中 o' 代表输出矩阵尺寸， i' 代表输入矩阵尺寸，p 代表填充，s 代表步长。

在神经网络中，输入分布经过每层的变化后都会发生一定程度的改变。当

经过网络层数较多时，其产生的影响将会进一步扩大，其输入分布会产生明显偏移。在训练过程中，深层网络中的隐层的输入分布会由于参数变化发生偏移，这就叫做内部协变量转移。而这种偏移的产生往往会使网络训练收敛困难。所以本文在网络中使用了批量标准化的操作（Batch Normalization），将各层的输入值的分布进行规范化矫正，把数据分成小批量进行随机梯度下降，使得输入值的分布更加标准，令输入值落在对输入敏感区，从而增大梯度、加快收敛。网络中在激活函数之前进行批标准化，使得输入的均值为0，方差为1，批量标准化不会损坏网络已学到的特征。具体的批量标准算法可以描述为：

有批处理输入 $x : B = \{x_1, \dots, m\}$ ，能够计算其均值为 $\mu_B = \frac{1}{m} \sum_{i=1}^m x_i$ ，计算其方差 $\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$ ，根据公式 $\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ 进行规范化处理，得到标准化后的数据 $y_i = \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i)$ 。

4.4.2 判别器实现算法

判别器的任务是根据输入的病灶图像进行特征分析后，判断该样本是否为真实的样本图像，输出一个(0,1)之间的数字用于表示该输入是真实病灶图像的概率。本文期望能够在判别器和生成器的依次优化训练间获得能够生成逼真样本的生成器。根据上述理论依据，本文使用卷积神经网络来实现判别器的任务，判别器将接收32*32*32尺寸的样本，经过多层卷积层并经sigmoid激活函数处理后得到输出结果，即判别器对样本真伪的判断。同时，本文在判别器层中插入了批量标准化处理进行数据规范化，使用Leaky ReLU作为激活函数，将函数 α 设为0.2。实验中使用卷积核4*4*4，设置卷积步长为2*2*2，设置batch尺寸为32。具体层设计如下：

1. 输入层 InputLayer
2. 卷积层 h0
 - (a) 使用 Conv3dLayer 进行卷积运算
 - (b) 使用 Leaky ReLU 作为激活函数， $\alpha=0.2$
 - (c) 输入 size: batchsize * 32 * 32 * 32 * 1(批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
 - (d) 输出 size: batchsize * 16 * 16 * 16 * 32
3. 卷积层 h1
 - (a) 使用 Conv3dLayer 进行卷积运算

- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 Leaky ReLU 作为激活函数, $\alpha=0.2$
- (d) 输入 size: $batchsize * 16 * 16 * 16 * 32$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 8 * 8 * 8 * 64$

4. 卷积层 h2

- (a) 使用 Conv3dLayer 进行卷积运算
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 Leaky ReLU 作为激活函数, $\alpha=0.2$
- (d) 输入 size: $batchsize * 8 * 8 * 8 * 64$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 4 * 4 * 4 * 128$

5. 卷积层 h3

- (a) 使用 Conv3dLayer 进行卷积运算
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 Leaky ReLU 作为激活函数, $\alpha=0.2$
- (d) 输入 size: $batchsize * 4 * 4 * 4 * 128$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 2 * 2 * 2 * 256$

6. 使用 FlattenLayer、DenseLayer 进行全连接, 使用 sigmoid 作为激活函数

4.4.3 生成器实现算法

生成器的任务是根据肺结节病灶的特征, 经过上采样生成三维肺结节图像。在一开始, 生成器的输入是随机生成的噪音, 经过上采样——这里使用卷积层反卷积来实现, 生成较为逼真的样本图片。将这些生成的图片输入判别器中, 期望能够迷惑判别器对其真伪性的识别。通过生成器与判别器的博弈, 优化两者, 最终得到优秀的生成器模型。生成器接收随机的高斯噪声, 经过多层卷积层处理, 得到输出结果。在层中同样插入批量标准化处理进行数据规范化, 使用 ReLU 作为激活函数。实验中使用反卷积核 $4*4*4$, 设置反卷积步长为 $2*2*2$, 设置 batch 尺寸为 32。具体层设计如下:

1. 输入层 InputLayer

2. 使用 DenseLayer 进行全连接

- (a) 使用 ReshapeLayer 进行 reshape
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 ReLU 作为激活函数
- (d) 输出 size: $batchsize * 4 * 4 * 4 * 512$

3. 卷积层 h1

- (a) 使用 DeConv3dLayer 进行反卷积
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 ReLU 作为激活函数
- (d) 输入 size: $batchsize * 4 * 4 * 4 * 512$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 8 * 8 * 8 * 256$

4. 卷积层 h2

- (a) 使用 DeConv3dLayer 进行反卷积
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 ReLU 作为激活函数
- (d) 输入 size: $batchsize * 8 * 8 * 8 * 256$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 16 * 16 * 16 * 128$

5. 卷积层 h3

- (a) 使用 DeConv3dLayer 进行反卷积
- (b) 使用 BatchNormLayer 进行批量标准化
- (c) 使用 tanh 作为激活函数
- (d) 输入 size: $batchsize * 16 * 16 * 16 * 128$ (批量数 * 图像长 * 图像宽 * 图像高 * 卷积核数量)
- (e) 输出 size: $batchsize * 32 * 32 * 32 * 1$

4.4.4 判别器、生成器的博弈

图 4-21 和图 4-22 是本文设计的判别器和生成器网络的简单示意图。

生成对抗网络的原理就是利用判别器与生成器之间的互相优化，逐步提升判别器辨别样本真伪的能力和生成器生成逼真的样本的能力。本文使用的

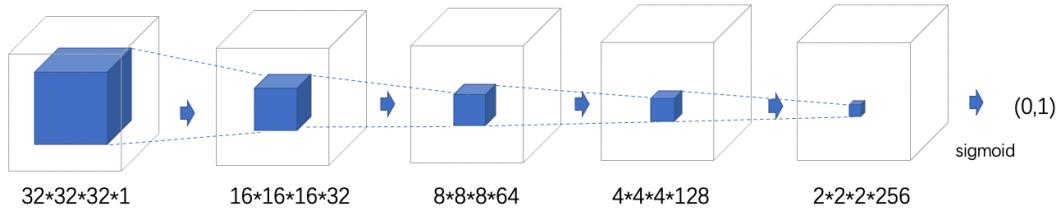


图 4-21: 针对肺结节生成网络的判别器

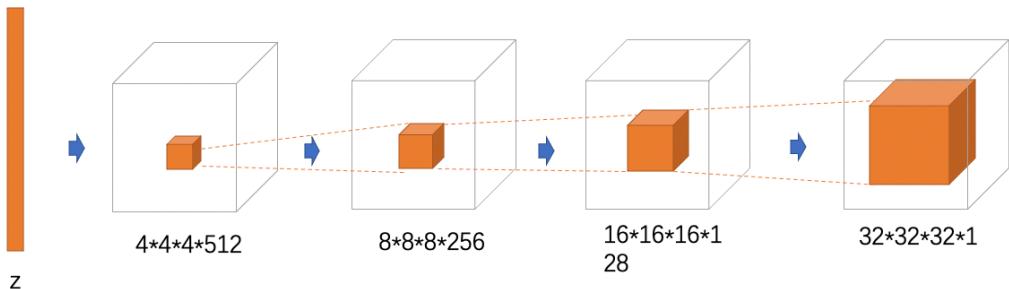


图 4-22: 针对肺结节生成网络的生成器

3D-DCGAN 也不例外，同样也是循环训练判别器和生成器，直到生成器生成样本分布接近真实数据分布。具体算法如下：

1. 固定生成器，优化判别器得到 D^*
2. 固定判别器，优化生成器，令生成的数据能够迷惑 D^* 输出错误的判断，得到优化后的 G^*
3. 重复前两过程进行博弈，直到生成器分布接近真实数据分布，判别器难以判断样本真伪。

4.4.5 损失函数

第二章介绍了相对熵（KL 散度）和交叉熵及二者的关系：

$$D_{KL}(A||B) = -S(A) + H(A, B) \quad (4-18)$$

其中， $D_{KL}(A||B)$ 代表 A 与 B 的相对熵， $S(A)$ 代表 A 的熵， $H(A, B)$ 代表 A 与 B 的交叉熵。当 $S(A)$ 为常量， $D_{KL}(A||B)$ 与 $H(A, B)$ 等价。

因为训练数据是确定的，所以其熵是确定值，所以使用交叉熵和相对熵作为损失函数的作用一致。在实现中，为了减少计算量，本文模型选用交叉熵作为损失函数。

在生成器中，期望判别器将生成器生成的图片识别为真，因此需要使用 $D(G(z))$ 与 1 的交叉熵作为损失函数，以期令判别器将生成器生成的三维肺结节判断为真实的肺结节样本。在判别器中，期望判别器将真实的三维肺结节识别为真，将生成器生成的三位肺结节识别为假，因此，对于真实图片，判别器的输出期望为 1（真），所以计算与 1 的交叉熵；对于生成器生成的图片，判别器输出期望为 0（假），所以计算与 0 的交叉熵即可。判别器的损失函数等于真实样本和生成样本的损失相加：

$$d_loss = \arg \min_G \max_D \mathbb{E}_{x \sim p_r(x)}[\log D(x)] + \mathbb{E}_{z \sim p(z)}[\log(1 - D(G(z)))] \quad (4-19)$$

项目使用 tensorflow 实现了 3D-DCGAN 生成病灶的技术，其网络结构如图 4-23 所示。在代码实现本章所述所有关键算法后，便需要对实现的网络进行训练，从而实现技术应用，这就是下一章所要介绍的内容。

4.5 本章小结

本章是文章的核心章节。首先，本章针对技术的实现做了全流程、分模块的介绍，阐述了各模块的主要功能及在流程中的作用。之后，又针对技术所涉及到的四个关键算法进行详细解释。其中，3D-DCGAN 是本文的核心算法，通过生成器与判别器的博弈获得了优秀的生成器，该生成器能够生成逼真的三维病灶图像。结节信息提取和样本获取算法则是针对医疗图像数据集特点提出的数据预处理算法，解决了医疗图像文件较大、信息存储特殊的问题。虽然本文是以肺结节作为目标进行算法设计，但样本处理算法仍然能够应用到其他的医疗图像的处理中。同时，本章介绍了验证实验中使用的通用图像扩增算法：旋转、翻转、亮度调整、噪声添加，阐述了其算法实现，为质量评估实验的顺利进行打下了坚实的基础。

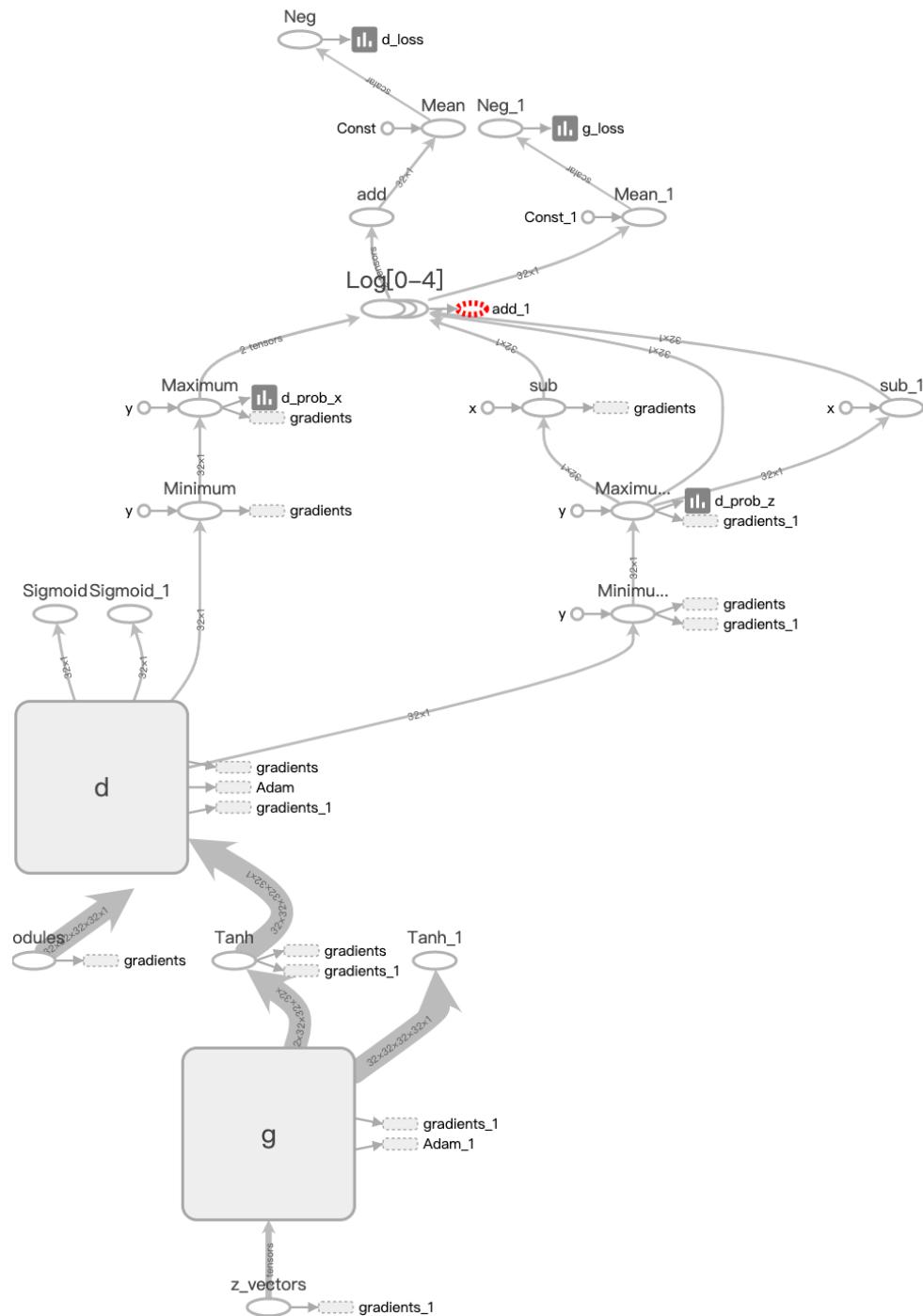


图 4-23: 网络结构

第五章 生成病灶及质量评估实验

5.1 3D-DCGAN 网络训练

5.1.1 实验环境

由于 3D-DCGAN 网络训练对硬件 CPU、GPU、内存等具有较高要求，所以本文实验使用远程服务器进行。使用远程服务器 15GB 内存的 CPU 和一块 RTX 2080 Ti GPU 进行深度学习网络的训练。

医疗影像数据集通常包含文件内容多，信息丰富。本实验使用的原始肺部影像数据集 LIDC-IDRI 已包含了 125G 的原始数据。在经过样本获取、图像变换、图像生成等环节后，相关的图片数据能够达到 400G。因此，在进行本实验时，不仅使用远程服务器硬盘进行数据存储，同时结合网盘进行数据的保存。

本文中所有的算法都使用 python 实现，并借助 tensorflow-gpu、tensorflowlayer 等库完成了 3D-DCGAN 网络的构建。

表 5-1: 实验环境

条目	信息
服务器	AutoDL 硬盘：300GB SSD 网盘：200GB
CPU	4 核 Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz 内存：15GB
GPU	RTX 2080 Ti * 1 显存：11GB
操作系统	Ubuntu 20.04.3 LTS
其他配置	Python: 3.7 Tensorflow-GPU: 1.15.0 Tensorlayer: 1.11.1 Cuda: 10.0 cuDNN: 7.6.0

5.1.2 模型训练与病灶生成

本文使用 tensorflow 实现第四章中的算法，从 LICD-IDRI 数据集中提取出肺结节样本并进行简易扩增，得到 5058 个训练样本，用他们对 3D-DCGAN 网络进行训练。

表 5-2: DCGAN 模型相关参数

条目	信息
D-卷积核尺寸	4*4*4
D-卷积步长	2*2*2
G-反卷积核尺寸	4*4*4
G-反卷积步长	2*2*2
优化器	Adam beta1=0.5 D 学习速率 alpha_d=0.00004 G 学习速率 alpha_g=0.0001 D 学习速率 alpha_d_2=0.000015 G 学习速率 alpha_g_2=0.00005 (4500 次迭代后)

实验尝试了数十组 Adam[65] 学习率，最终选用 $\alpha_g = 0.0001$ ， $\alpha_d = 0.00004$ 进行训练。在每个迭代中，从样本集中随机选取 32 个样本用于训练，在训练到 4500 次迭代后，重新设定 $\alpha_g = 0.00005$ ， $\alpha_d = 0.000015$ ，可以观察到经过 20000 次迭代，生成器和判别器的损失函数趋于平稳，如图 5-1、5-2 所示。

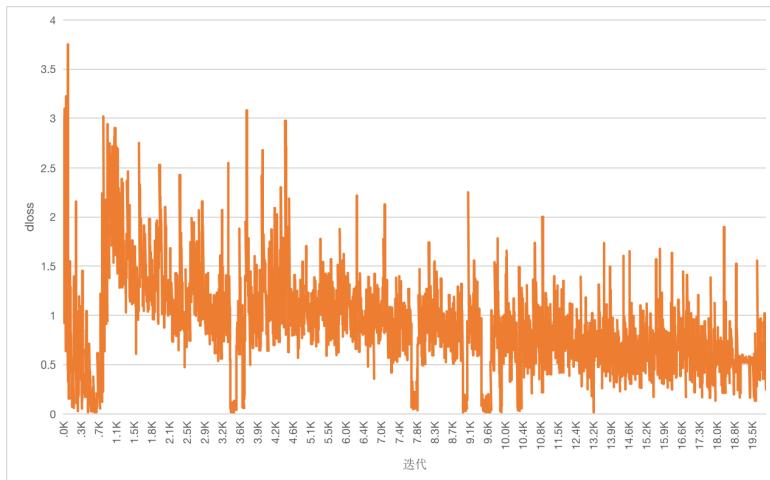


图 5-1: DCGAN 网络训练-判别器损失函数

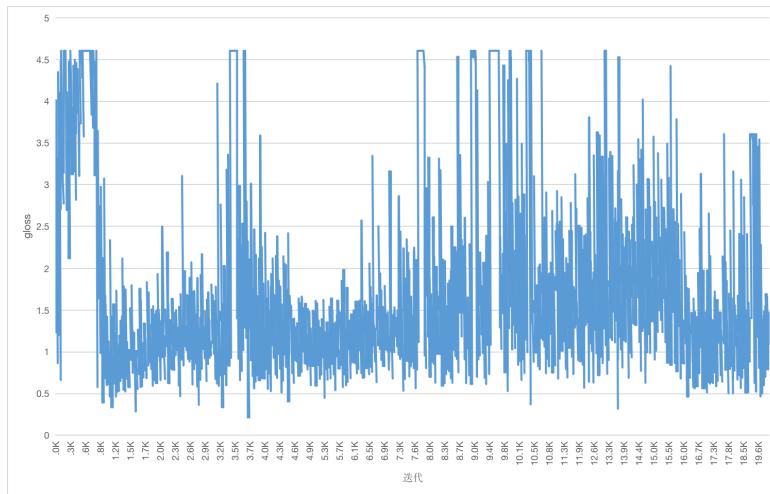


图 5-2: DCGAN 网络训练-生成器损失函数

这时获得了优化后的生成器网络，它可以根据随机输入的高斯分布噪声，生成一系列的肺结节病灶样本。使用训练模型生成结节的代码如图 5-3 所示。

```
# 使用DCGAN生成肺结节

tf.reset_default_graph()
z_size = 100
z_vector = tf.placeholder(shape=[None,z_size],dtype=tf.float32, name='z_vectors')
net_g_test = generator(z_vector, is_train=False, reuse=False)
saver = tf.train.Saver()

with tf.Session() as sess:
    ckpt = tf.train.get_checkpoint_state(model_directory)

    # 选择加载的模型
    # print(ckpt.all_model_checkpoint_paths)
    model = ckpt.all_model_checkpoint_paths[0]
    saver.restore(sess, save_path=model)

    # 噪声
    noise = np.random.normal(0, 0.33, size=[batch_size,z_size]).astype(np.float32)

    # 生成
    nodule_samples = sess.run(net_g_test.outputs,feed_dict={z_vector:noise})

    # 保存
    nodule_samples.dump('samples')
```

图 5-3: 使用模型生成肺结节

如图 5-4、5-5、5-6，这是由生成器生成的肺结节图像的三维立体展示。

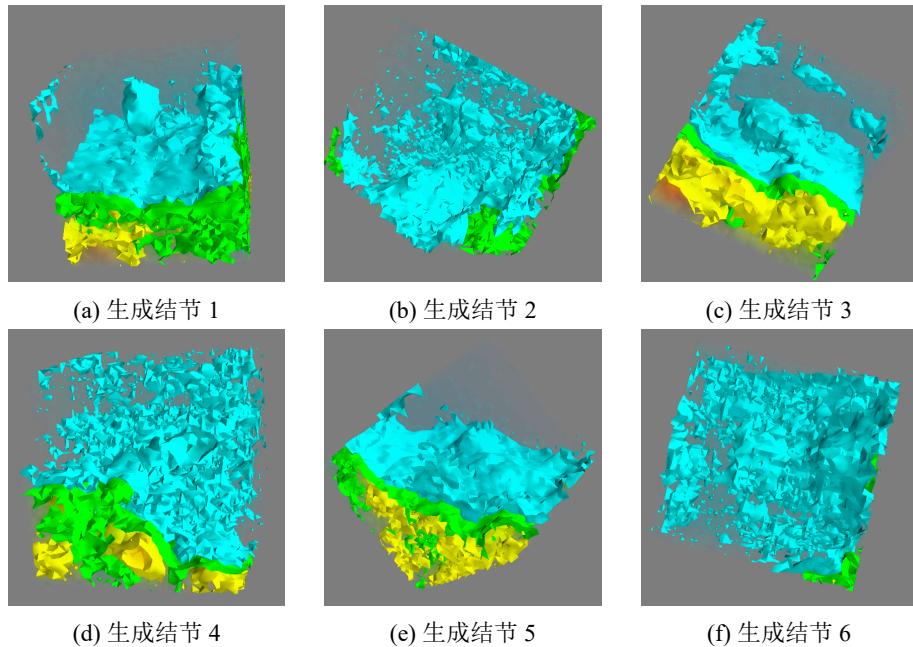


图 5-4: 生成结节-7500 次迭代

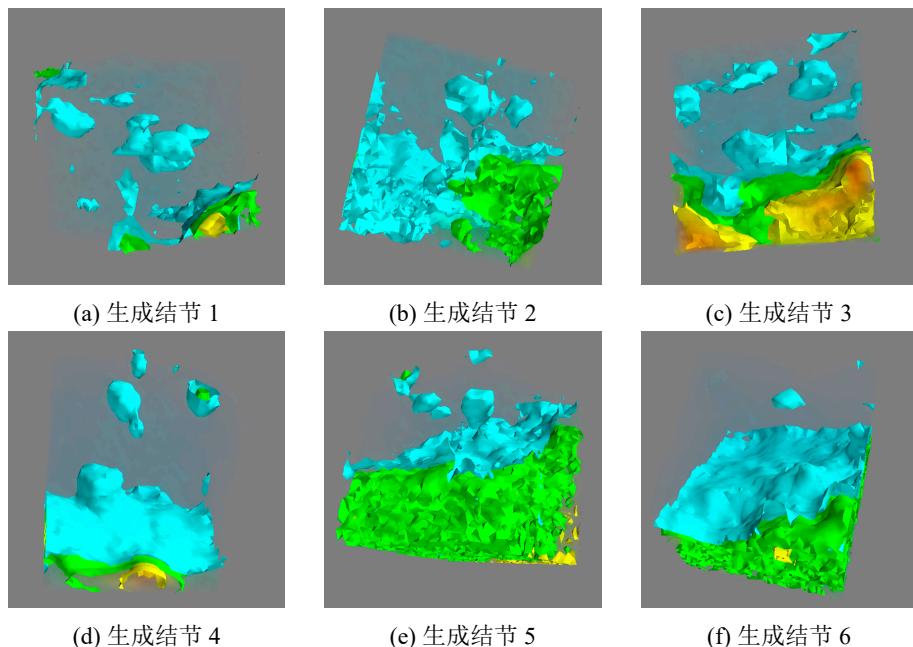


图 5-5: 生成结节-15000 次迭代

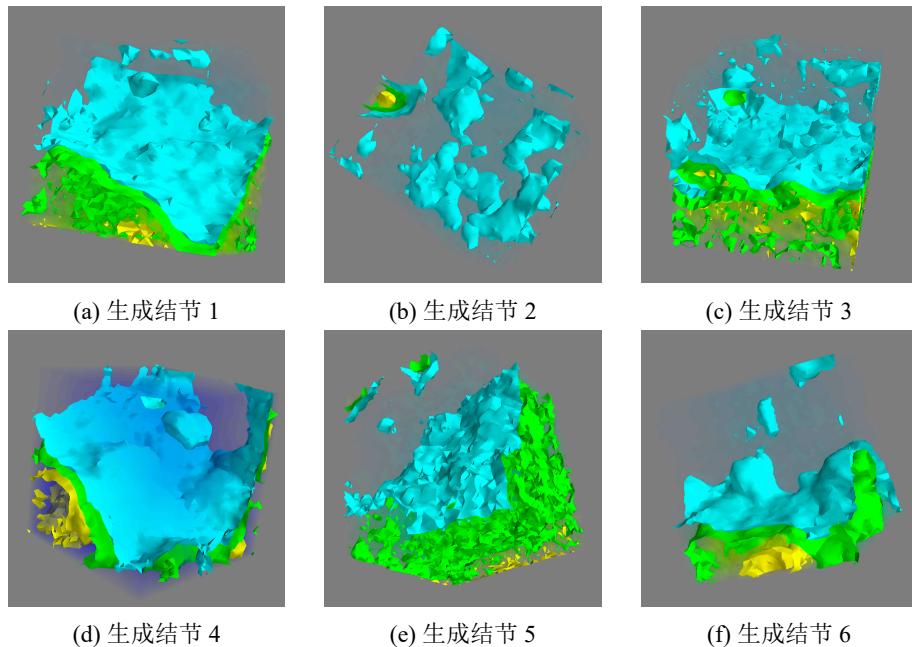


图 5-6: 生成结节-20000 次迭代

5.2 质量检测手段

为了验证本文技术的可靠性，本文需要对生成的肺结节图像进行质量检测与评估。在图像质量评估领域，一般对于扩增得到的图像的评估会使用基于扩增数据集与原数据集差异对比的手段进行评估。比如，使用峰值信噪比（PSNR）计算扩增得到的图像与原图像之间像素误差的全局大小来进行质量的检验，也有使用结构相似度（SSIM）进行度量，它是通过对比图像间的亮度相似性、对比度相似性和结构相似性来进行评估的。常见的质量度量指标详见表 5-3。但对于本文使用的生成对抗网络技术来说，难以确定进行比较的原图像且过程繁琐，所以本文没有选用这种基于扩增图像与原图像对比的质量评估方案。

本文采用根据模型分类结果进行评估的手段，对模型生成的数据集进行质量评估。具体来说，就是将扩增生成的图像用于病灶判别模型的训练，通过分析训练后模型准确度等指标的变化来评估生成图像的质量，分析生成图像是否能够有助于病灶判别模型的训练，从而验证本文技术的有效性。

表 5-3: 质量度量指标

指标	简介
峰值信噪比 (PSNR)	PSNR 用于量化压缩得到的图像或视频的重建质量，通过计算图像间的全局像素误差衡量图像质量。 PSNR 越大，代表失真越小，MSE 越大，代表失真越大。
均方误差 (MSE)	PSNR 可以用 MSE 来定义： $MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$ 其中，I 是无噪声的 $m \times n$ 单色图像，K 是其噪声近似值 $PSNR = 20 \cdot \log_{10} (MAX_I) - 10 \cdot \log_{10} (MSE).$ 其中， MAX_I 是图像的最大可能像素值
结构相似度 (SSIM)	从三个角度（亮度相似度、对比度相似度、结构相似度）衡量两幅图像相似度。SSIM 的取值范围是-1 到 1，当两张图片相同时，SSIM 取 1。
多层级结构相似性 (M-SSIM) [66]	对于图像 x, y 有： $SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$ 其中， μ_x 是 x 的平均值， μ_y 是 y 的平均值， σ_x^2 是 x 的方差， σ_y^2 是 y 的方差， σ_{xy} 是 x 和 y 的协方差。 $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ 是用来维持稳定的常数。 L 是像素值的动态范围。 $k_1 = 0.01, k_2 = 0.03$ 。 当图像有 M 个 patch 时有： $MSSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j)$
平均绝对差值 (MAD) [67]	$D = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H S(x, y) - T(x, y) $ 其中， $S(x, y)$ 表示原始图像在坐标 (x, y) 的像素值， $T(x, y)$ 为待评价图像在坐标 (x, y) 的像素值。 W 代表图像宽度， H 代表图像高度。 D 代表待评估图像与原始图像的失真度，其值越小越好。
信息保真度准则 (IFC) 视觉信息保真度 (VIF、VIFp)[68, 69]	拓展了图像与人眼之间的联系，也是通过比较待评价的图像与原始图像之间的信息进行质量度量。

5.3 质量验证实验

5.3.1 研究问题

通用的扩增方法已被证实能够生成对于模型训练有一定帮助的图像，而本文使用的生成对抗网络生成样本技术期望能够在通用扩增方法的基础上进行进一步的补充加强。因此，本文技术质量评估实验的研究问题着眼于 GAN 生成的三维医疗病灶图像是否能够提升病灶判别模型的训练效果，以及 GAN 生成的三维病灶是否能补充通用扩增获得的样本集，进一步提升模型训练的质量。针对以上研究问题，本文进行了实验的设计。

5.3.2 实验设计

常用的扩增方法（如第四章中描述的旋转等仿射变换以及添加噪声等弹性变换）已被验证能够生成对于模型训练有一定帮助的图像。本文技术（生成对抗网络生成样本）期望能够在常用扩增方法的基础上进行进一步的补充加强。基于此目的，本次质量验证实验选用基于模型分类结果的评估手段，选择用于肺结节判别的 3D-CNN 模型作为训练对象，设计了 4 组实验，通过对比来验证生成样本的质量：

实验组 1：使用未经扩增的原数据集（记训练集大小为 N1）训练肺结节判别模型；

实验组 2：使用通用扩增图像（随机经过弹性变换：亮度调整、噪声添加或仿射变换：旋转、翻转）训练模型（记训练集大小为 N2， $N2 > N1$ ）；

实验组 3：使用更多样本的通用扩增图像（记训练集大小为 N3， $N3 > N2$ ）训练模型；

实验组 4：使用通用扩增图像与生成样本图像（训练集大小为 N3，其中通用扩增图像个数为 N2）训练模型。

通过对不同数据集下的模型训练效果能够有效检验生成肺结节的质量，检验本文提出的病灶生成技术的有效性。本文使用的相关评估指标包括 Accuracy（模型识别准确率）、ROC 曲线和 AUC 数值。

在实验的最初，固定了验证集（340 个正样本和 340 个负样本）和测试集（681 个正样本和 681 个负样本）用于对训练效果的验证和测试。实验中通过对样本进行一系列扩增变化，包括随机进行一般扩增或增加由卷积生成对抗网

络生成的样本，来进行对比实验。具体实验数据集设置如表 5-4 所示。为了验证数据扩增对训练集的扩充作用，验证数据扩增技术在解决样本缺乏问题上的有效性，实验设计了实验组 1 作为基准对照组。实验组 1 中的数据集是未经过任何扩增变换的。为了验证通用的整体变换扩增方法能够一定程度上应用于三维医疗病灶图像扩增，实验设计了实验组 2。实验组 2 中的数据是经过通用扩增方法扩增得到的，也包含一部分未经变换的原始样本。通用扩增方法在特殊领域的效果有限，实验设计了实验组 3，期望验证这一结论。实验组 4 是验证本文技术有效性的关键实验组。实验组 4 包含了与实验组 3 等量的样本。不同的是，实验组 3 中的数据除了一些是未经变换的，其余的都是经过通用扩增方法扩增得到的，而实验组 4 中的数据集除了包含了与实验组 2 等量的通用扩增数据外，还包含了 3D-DCGAN 生成的肺结节样本。

表 5-4: 实验数据集设计

实验编号	1	2	3	4
训练集	正样本	1248	3744	4867
	负样本	1248	3744	4867
	总	2496	7488	9734
验证集	正样本	340	340	340
	负样本	340	340	340
	总	680	680	680
测试集	正样本	681	681	681
	负样本	681	681	681
	总	1362	1362	1362
正样本总数		2269	4765	5888
负样本总数		2269	4765	5888

随后，按照设计中的 4 组实验对一个 3D-CNN 的肺结节判别模型进行训练，具体实验参数如表 5-5 所示。

表 5-5: 验证实验参数

条目	值
优化器	Adam (学习率 =0.002)
损失函数	交叉熵 categorical_crossentropy
标签	One-Hot 编码
批大小	50
迭代次数	50

5.3.3 评估指标

在深度学习领域，经常使用混淆矩阵来实现预测分析的数据可视化。它具有两个维度（实际和预测），它的列代表预测类别，它的行代表数据的真实类别。通过建立混淆矩阵，能够直观地展示模型预测结果，看出模型是否预测错误（即，将某类别的数据预测为其他类别），能够计算模型的准确度或其他评估指标。具体到二分类任务中，混淆矩阵可以表示为图 5-7。

		预测类别	
		真	假
真实类别	真	真正类 True Positive (TP)	假负类 False Negative (FN)
	假	假正类 False Positive (FP)	真负类 True Negative (TN)

图 5-7: 混淆矩阵

图 5-7 中，真正 (TP) 代表样本真实的类别为正，模型判断结果为正，判断正确。假负 (FN) 代表样本真实的类别为负，模型判断的结果为负，判断错误。假正 (FP) 代表样本真实的类别为正，模型判断的结果为负，判断错误。真负 (TN) 代表样本真实的类别为负，模型判断的结果为负，判断正确。表格中的每一格代表该类别包含的样本数，通过结合 TP、TN、FN、FP 进行计算，能够得到多种模型评估指标。这里详细介绍本项目实验使用的三个指标：Accuracy、ROC、AUC。

Accuracy (准确率) 是机器学习模型评估最常用的指标。将模型判断正确的样本数比上样本的总数就能够计算模型的准确度。结合混淆矩阵，它的计算

公式可表示为：

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (5-1)$$

显然，Accuracy 越大（越接近 1）说明模型的识别准确度越高，说明模型的质量越高。

同时，本文实验使用了 ROC 曲线（受试者工作特征曲线，Receiver Operating Characteristic Curve）辅助判断模型质量。AUC（曲线下面积，Area Under Curve）则是以 ROC 曲线为基础得来的更加直观的数值。在二分类的任务中，ROC 曲线的横纵坐标分别为 False Positive Rate（FPR）和 True Positive Rate（TPR），其中 FPR 代表了模型对负样本（FP+TN）判断错误的比率，TPR 代表了模型对正样本（TP+FN）判断正确的比率，所以有：

$$FPR = \frac{FP}{FP + TN} \quad (5-2)$$

$$TPR = \frac{TP}{TP + FN} \quad (5-3)$$

ROC 曲线能够避免样本不平衡导致的质量评估不准的问题。曲线越接近左上角，代表该模型性能越好。

AUC（曲线下面积，Area Under Curve）是 ROC 曲线下的面积大小，它代表了模型对随机选择的一正一负的样本进行判断时，正样本的排名在负样本之前的概率 [70]。AUC 越大（越接近 1）代表模型质量越高。相较于 ROC，AUC 作为一个确切数值更加直观。具体实现中，使用 sklearn 的 metrics 模块即可得到该值。

5.3.4 实验结果

通过以上实验，得到了一系列数据结果，如表 5-6 所示。

四组训练实验中的 Acc 曲线、Loss 曲线及 ROC 曲线如图 5-8、5-9、5-10 所示。

通过观察曲线、分析实验结果数值可以知道，在使用原数据集进行训练时，准确率仅能达到 0.916。而通用扩增方法确实能够解决样本数量缺少的问题，在一定程度上将识别准确度提高至 0.9567，AUC 相应也有增长。

表 5-6: 验证实验结果

实验	AUC	ACC
1	0.977599	0.9162
2	0.990609	0.9567
3	0.99142	0.9569
4	0.991976	0.964

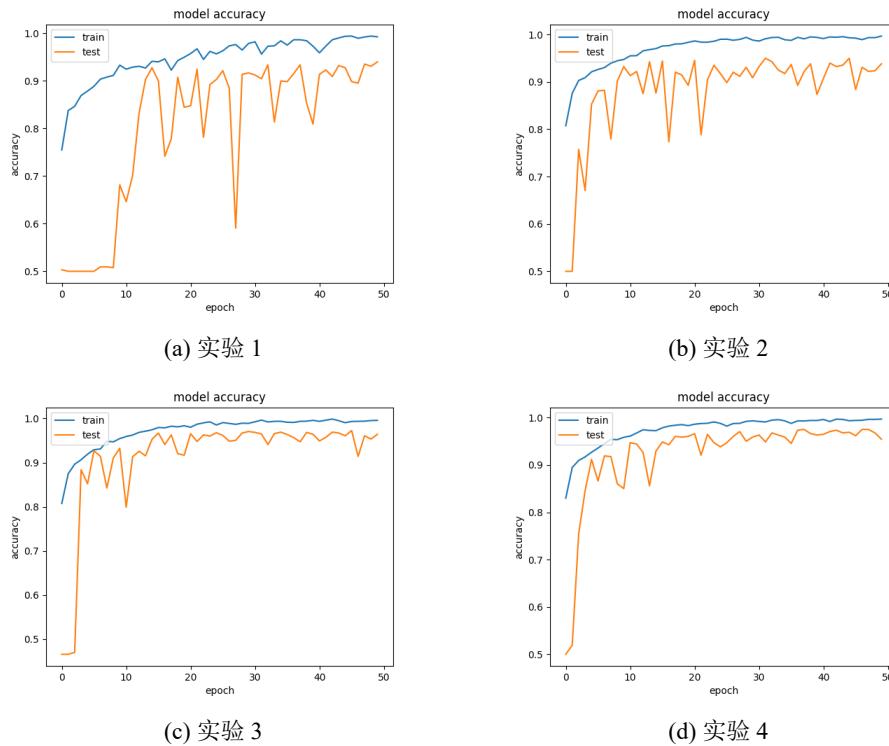


图 5-8: 训练 3D-CNN-Acc 曲线对比

但是，如实验 3 结果所示，继续增加由通用扩增方法生成的样本并不能更显著的提升模型识别准确度，AUC 亦然。

实验 4 使用与实验 3 同样大小的数据集，但其中的样本由 3744 个（同实验 2）通用扩增样本和 1123 个对抗生成样本组成，训练模型后发现，该数据集能够实现对模型准确度的进一步提升达到 0.964，AUG 也有微小提升。

由此能够证明，使用对抗网络生成的样本不仅能够对 3D-CNN 肺结节判别模型的训练有帮助，并且能够对通用扩增手段（仿射扩增、弹性扩增）起到补充作用，在此基础上进一步的提高模型的质量，这符合实验预期。进而能够证明，使用本文技术生成的病灶样本质量较高，本文提出的对抗生成网络生成三维病灶的技术确实有效，能够有效的提升模型识别的准确度。

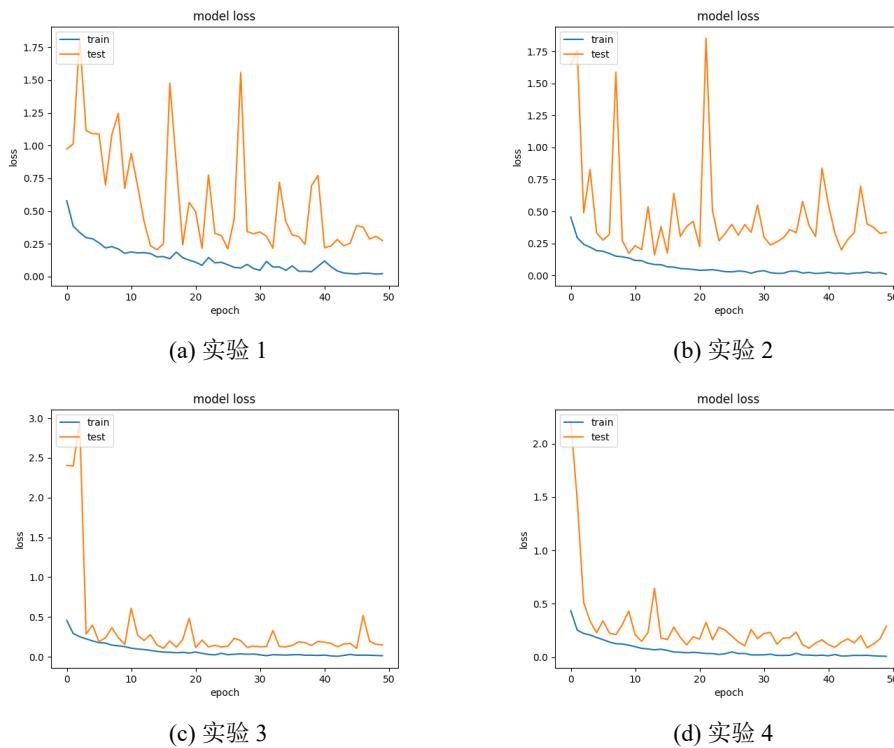


图 5-9: 训练 3D-CNN-Loss 曲线对比

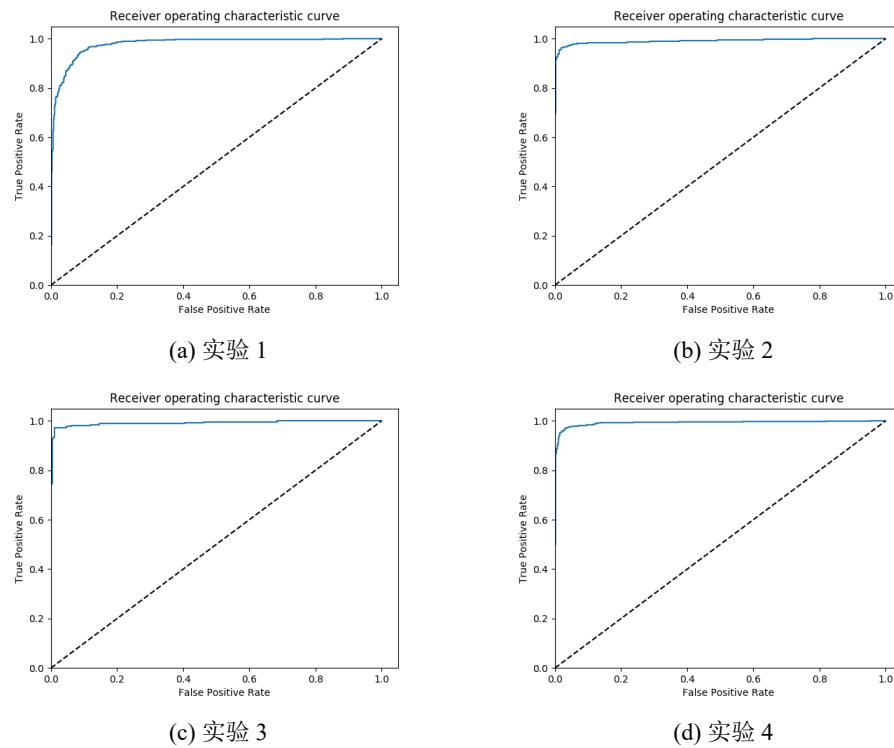


图 5-10: 训练 3D-CNN-ROC 曲线对比

5.4 本章小结

本章介绍了本文所涉及到的两个实验：3D-DCGAN 的训练实验以及使用 3D-CNN 进行技术质量评估的验证实验。使用从 LIDC-IDRI 中规范化提取得到的肺结节样本训练 3D-DCGAN。经过 20000 次迭代，最终获得了能够生成逼真三维肺结节的生成器。向生成器输入随机的高斯噪声能够生成一系列的三维肺结节样本。为了验证生成图像的质量以及本文技术的可行性，本文设计了基于模型训练准确度的对比验证实验。通过对比四组实验结果，能够证明使用本技术生成的肺结节病灶样本不仅能够提升 CNN 病灶判别模型的准确度，也能够在通用扩增手段的基础上进一步扩充数据集，提升训练模型的质量。

第六章 总结与展望

本文使用卷积生成对抗网络实现了三维肺结节病灶样本生成技术，实现了对三维医疗病灶图像的数据扩增，解决了医疗领域训练样本缺乏的问题。与通用图像扩增技术不同，本文不仅引入了 DCGAN 技术，还将扩增对象从原来的二维图像升至三维图像，满足了目前先进的医疗检测诊断模型的样本需求。本文技术实现由数据预处理、生成病灶、质量评估三个大模块组成。在数据预处理部分，本项目针对医疗图像的特殊信息存储规则和读取规则，实现了规范化的数据预处理流程，解决了医疗影像集因数据存储庞大导致的操作困难的问题。该流程不仅适用于本文研究对象——肺结节（LIDC-IDRI 数据集，胸部 CT 影像），本质上也可以用于大多数类型的医疗图像集处理。病灶生成部分选用卷积生成对抗网络实现，利用生成器和判别器的博弈，使得生成器生成与真实样本分布接近的病灶数据，而判别器难以分辨真实数据和生成器生成的数据。在训练完成后，将随机的高斯噪声输入训练好的生成器，就能够方便地获得逼真的病灶样本，从而实现数据扩增。同时，为了验证本文技术的可靠性，项目通过基于模型训练的评估手段，使用多组数据对肺结节检测模型（3D-CNN）进行训练。实验表明，本技术能够生成质量较高的三维病灶样本，有效提升判别模型的质量。

本文中，为了验证生成病灶质量，选用 3D-CNN 肺结节检测模型进行训练，通过对不同数据集训练后模型的准确度进行验证。在之后的研究过程中，可以选用多种类型的模型进行更全方位的质量评估。另外，文中选用高斯噪声作为生成器输入进行三维医疗图像生成。相关研究表明，不同的噪声类型会对生成器生成的结果产生影响，因此，在本文技术基础上，可以继续探究其他噪音类型对生成对抗网络质量的影响。

致 谢

一转眼我已经在南京大学生活学习了六年，时光匆匆而过，无数喜悦、失落、满足、焦虑的情绪化作当下的一声轻叹。这篇毕业论文也是为这六年，特别是近两年研究生学习生涯做一个总结。在日常科研学习以及毕业论文的写作过程中，我得到了很多人的帮助，在此向他们表示由衷的感谢。

首先，我要感谢我的导师陈振宇教授。三年前，陈老师引领我初次接触学术研究，并结合我的想法为我规划科研方向。在实验室中，我接触到了软件领域前沿的研究成果，得到了实验室老师和同学的大力支持，增长了自己的见闻，提升了自己的能力。在刘佳玮学姐和王黛薇学姐的指导下，我取得了一定的研究成果。她们不仅仅是师姐，更像是朋友，在学习和生活中帮助我前行。

其次，我要感谢辅导员黄蕾老师和各位任课老师。在软件学院我不仅学习到了专业知识，更学习到了严谨、负责、包容、担当的态度。

感谢我的室友和朋友。感谢何云同学、贺璐同学、姜福艳同学，3223 的小伙伴为我两年研究生生活带来了无数的快乐与欢笑。感谢王锐同学，我所有的欢笑都能在你那里得到回应，我所有的烦闷都能在你那里得到安慰。聊天窗口永远排在前几个，大概就是我们每天互相鼓励、抱团取暖的证明。

感谢我的父母。善良、博爱、平等、尊重，我从你们那里学到了太多。不论在任何时候，我都能从家里得到支持与信任，重拾动力。感谢方渤鎔同学，谢谢你永远敞开怀抱拥抱失意的我，让我能不顾及身后，勇敢地向前飞。希望未来也能继续一起走过。感谢薛雨晴同学，我所有不安的小情绪都讲给了你听，谢谢一直都有你。

最后，感谢所有在论文写作过程中为我提供支持的朋友们。感谢范天逸、汪启俊、王维家、马大庆、郑乃馨对本文作者精神上的大力支持。

参考文献

- [1] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述 [J]. 计算机工程与应用, 2019, 55(12): 17.
- [2] 田娟秀, 刘国才, 谷珊珊, et al. 医学图像分析深度学习方法研究与挑战 [J]. 自动化学报, 2018, 44(3): 24.
- [3] RAZZAK M I, NAZ S, ZAIB A. Deep Learning for Medical Image Processing: Overview, Challenges and the Future[M/OL] // DEY N, ASHOUR A S, BORRA S. Classification in BioApps: Automation of Decision Making. Cham : Springer International Publishing, 2018 : 323–350.
https://doi.org/10.1007/978-3-319-65981-7_12.
- [4] SHEN D, WU G, SUK H-I. Deep Learning in Medical Image Analysis[J/OL]. Annual Review of Biomedical Engineering, 2017, 19(1): 221–248.
<https://doi.org/10.1146/annurev-bioeng-071516-044442>.
- [5] 马婷, 陈清财. 基于开放医疗大数据的人工智能研究 [J]. 医学与哲学, 2022, 43(1): 1–4.
- [6] SHORTEN C, KHOSHGOFTAAR T. A survey on Image Data Augmentation for Deep Learning[J/OL]. Journal of Big Data, 2019, 6.
<http://dx.doi.org/10.1186/s40537-019-0197-0>.
- [7] ANABY-TAVOR A, CARMELI B, GOLDBRAICH E, et al. Do Not Have Enough Data? Deep Learning to the Rescue![J/OL]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(05): 7383–7390.
<https://ojs.aaai.org/index.php/AAAI/article/view/6233>.
- [8] LI R, LI X, HENG P-A, et al. Pointaugment: an auto-augmentation framework for point cloud classification[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020 : 6378–6387.

- [9] SINGH S P, WANG L, GUPTA S, et al. 3D Deep Learning on Medical Images: A Review[J/OL]. Sensors, 2020, 20(18).
<https://www.mdpi.com/1424-8220/20/18/5097>.
- [10] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J/OL]. Medical Image Analysis, 2017, 42 : 60 – 88.
<https://www.sciencedirect.com/science/article/pii/S1361841517301135>.
- [11] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation Applied to Handwritten Zip Code Recognition[J/OL]. Neural Computation, 1989, 1(4) : 541 – 551.
<http://dx.doi.org/10.1162/neco.1989.1.4.541>.
- [12] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J/OL]. Proceedings of the IEEE, 1998, 86(11) : 2278 – 2324.
<http://dx.doi.org/10.1109/5.726791>.
- [13] CERNAZANU C, STEFAN H. Segmentation of Bone Structure in X-ray Images using Convolutional Neural Network[J/OL]. Advances in Electrical and Computer Engineering, 2013, 13 : 87 – 94.
<http://dx.doi.org/10.4316/aece.2013.01015>.
- [14] SU H, LIU F, XIE Y, et al. Region segmentation in histopathological breast cancer images using deep convolutional neural network[C/OL] // 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). 2015 : 55 – 58.
<http://dx.doi.org/10.1109/ISBI.2015.7163815>.
- [15] BENZAMIN A, CHAKRABORTY C. Detection of Hard Exudates in Retinal Fundus Images Using Deep Learning[C/OL] // 2018 Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR). 2018 : 465 – 469.
<http://dx.doi.org/10.1109/ICIEV.2018.8641016>.
- [16] SETIO A A A, CIOMPI F, LITJENS G, et al. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Net-

- works[J/OL]. IEEE Transactions on Medical Imaging, 2016, 35(5) : 1160–1169.
<http://dx.doi.org/10.1109/TMI.2016.2536809>.
- [17] SETIO A A A, TRAVERSO A, de BEL T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge[J]. Med Image Anal, 2017, 42 : 1–13.
- [18] HAMIDIAN S, SAHINER B, PETRICK N, et al. 3D Convolutional Neural Network for Automatic Detection of Lung Nodules in Chest CT[J]. Proc SPIE Int Soc Opt Eng, 2017, 10134.
- [19] HUANG X, SHAN J, VAIDYA V. Lung nodule detection in CT using 3D convolutional neural networks[C/OL] // 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). 2017 : 379–383.
<http://dx.doi.org/10.1109/ISBI.2017.7950542>.
- [20] ORNEK A H, CEYLAN M. Comparison of Traditional Transformations for Data Augmentation in Deep Learning of Medical Thermography[C/OL] // HERENCSAR N. 42nd International Conference on Telecommunications and Signal Processing, TSP 2019, Budapest, Hungary, July 1-3, 2019. [S.l.] : IEEE, 2019 : 191–194.
<https://doi.org/10.1109/TSP.2019.8769068>.
- [21] HOU Y, LIU J, WANG D, et al. TauMed: Test Augmentation of Deep Learning in Medical Diagnosis[M/OL] // Proceedings of the 30th ACM SIGSOFT International Symposium on Software Testing and Analysis. New York, NY, USA : Association for Computing Machinery, 2021 : 674 – 677.
<https://doi.org/10.1145/3460319.3469080>.
- [22] PEZESHK A, SAHINER B, ZENG R, et al. Seamless Insertion of Pulmonary Nodules in Chest CT Images[J/OL]. IEEE Trans. Biomed. Engineering, 2015, 62(12) : 2812–2827.
<https://doi.org/10.1109/TBME.2015.2445054>.

- [23] PEZESHK A, PETRICK N, CHEN W, et al. Seamless Lesion Insertion for Data Augmentation in CAD Training[J/OL]. IEEE Trans. Med. Imaging, 2017, 36(4): 1005 – 1015.
<https://doi.org/10.1109/TMI.2016.2640180>.
- [24] BERMUDEZ C, PLASSARD A J, DAVIS T L, et al. Learning Implicit Brain MRI Manifolds with Deep Learning[J]. Proc SPIE Int Soc Opt Eng, 2018, 10574.
- [25] YAN Q, WANG B, GONG D, et al. COVID-19 Chest CT Image Segmentation - A Deep Convolutional Neural Network Solution[J/OL]. CoRR, 2020, abs/2004.10987.
<https://arxiv.org/abs/2004.10987>.
- [26] OZTURK T, TALO M, YILDIRIM E A, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images[J/OL]. Computers in Biology and Medicine, 2020, 121 : 103792.
<http://www.sciencedirect.com/science/article/pii/S0010482520301621>.
- [27] 王坤峰, 苟超, 段艳杰, et al. 生成式对抗网络 GAN 的研究进展与展望 [J]. 自动化学报, 2017, 43(3) : 321 – 332.
- [28] KER J, WANG L, RAO J, et al. Deep Learning Applications in Medical Image Analysis[J/OL]. IEEE Access, 2018, 6 : 9375 – 9389.
<http://dx.doi.org/10.1109/ACCESS.2017.2788044>.
- [29] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks[J], 2014.
- [30] HUANG H, YU P S, WANG C. An Introduction to Image Synthesis with Generative Adversarial Nets[J], 2018.
- [31] 梁俊杰, 韦舰晶, 蒋正锋. 生成对抗网络 GAN 综述 [J]. 计算机科学与探索, 2020, 14(1) : 1 – 17.
- [32] ISOLA P, ZHU J-Y, ZHOU T, et al. Image-to-Image Translation with Conditional Adversarial Networks[J], 2018.

- [33] ZHU J-Y, PARK T, ISOLA P, et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks[J], 2020.
- [34] 陈佛计, 朱枫, 吴清潇, et al. 生成对抗网络及其在图像生成中的应用研究综述 [J]. 计算机学报, 2021.
- [35] 张杰, 赵惠军, 李贤威, et al. 基于生成式对抗网络技术合成医学图像的研究进展与展望 [J]. 医疗卫生装备, 2019, 40(10): 6.
- [36] 唐贤伦, 杜一铭, 刘雨微, et al. 基于条件深度卷积生成对抗网络的图像识别方法 [J]. 自动化学报, 2018, 44(5): 10.
- [37] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述 [J]. 计算机科学, 2019, 46(3): 63 – 73.
- [38] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C] // Icml. 2010.
- [39] MAAS A L, HANNUN A Y, NG A Y, et al. Rectifier nonlinearities improve neural network acoustic models[C] // Proc. icml : Vol 30. 2013 : 3.
- [40] NIELSEN M A. Neural networks and deep learning : Vol 25[M]. [S.l.] : Determination press San Francisco, CA, USA, 2015.
- [41] INOUE H. Data Augmentation by Pairing Samples for Images Classification[J/OL], 2018.
<https://arxiv.org/abs/1801.02929>.
- [42] ZHANG H, CISSE M, DAUPHIN Y N, et al. mixup: Beyond Empirical Risk Minimization[C/OL] // International Conference on Learning Representations. 2018.
<https://openreview.net/forum?id=r1Ddp1-Rb>.
- [43] 张巧丽, 赵地, 迟学斌. 基于深度学习的医学影像诊断综述 [J]. 计算机科学, 2017, 44(B11): 7.
- [44] 李坤成, 卢洁. 医学影像学新进展 [J]. 首都医科大学学报, 2013, 34(3): 364 – 367.

- [45] SUNG H, FERLAY J, SIEGEL R L, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries[J/OL]. CA: A Cancer Journal for Clinicians, 2021, 71(3): 209–249.
<https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.3322/caac.21660>.
- [46] VALENTE I R S, CORTEZ P C, NETO E C, et al. Automatic 3D pulmonary nodule detection in CT images: A survey[J/OL]. Comput. Methods Programs Biomed., 2016, 124: 91–107.
<https://doi.org/10.1016/j.cmpb.2015.10.006>.
- [47] ZHOU Q, FAN Y, WANG Y, et al. [China National Guideline of Classification, Diagnosis and Treatment for Lung Nodules (2016 Version)][J]. Zhongguo Fei Ai Za Zhi, 2016, 19(12): 793–798.
- [48] ABERLE D R, ADAMS A M, BERG C D, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening[J]. N Engl J Med, 2011, 365(5): 395–409.
- [49] TEAM N L S T R, OTHERS. Lung cancer incidence and mortality with extended follow-up in the National Lung Screening Trial[J]. Journal of Thoracic Oncology, 2019, 14(10): 1732–1742.
- [50] PASTORINO U, SILVA M, SESTINI S, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy[J]. Ann Oncol, 2019, 30(7): 1162–1169.
- [51] de KONING H J, van der AALST C M, de JONG P A, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial[J]. N Engl J Med, 2020, 382(6): 503–513.
- [52] 胡荣慧, 闻颂苏, 王贵美, et al. 肺部低剂量螺旋CT放射剂量的研究 [J]. 医学影像学杂志, 2004, 14(12): 4.
- [53] American Cancer Society. Cancer Facts Figures 2022[EB]. Atlanta : American Cancer Society, 2022.

- [54] 李惠民, 肖湘生. 肺结节 CT 影像评价 [J]. 中国医学计算机成像杂志, 2001(01): 30–41.
- [55] 王琪, 陈志强, 邬小平, et al. X 射线安全检查技术综述 [J]. CT 理论与应用研究, 2004, 13(1): 32–37.
- [56] 任法云, 付克广. 低剂量 CT 在肺结节筛查中的临床应用 [J]. 中国医学计算机成像杂志, 2014, 20(5): 5.
- [57] ARMATO III S, MCLENNAN G, BIDAUT L, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans[J/OL]. Medical Physics, 2011, 38 : 915–931.
<http://dx.doi.org/10.1118/1.3528204>.
- [58] BOQIANG L, ZHU M, ZHANG Z, et al. Medical Image Conversion with DICOM[C/OL] // . 2007 : 36 – 39.
<http://dx.doi.org/10.1109/CCECE.2007.16>.
- [59] 瞿裕忠, 张剑锋, 陈峥, et al. XML 语言及相关技术综述 [J]. 计算机工程, 2000, 26(12): 4–6,30.
- [60] LEV M, GONZALEZ R. 17 - CT Angiography and CT Perfusion Imaging[G/OL] // TOGA A W, MAZZIOTTA J C. Brain Mapping: The Methods (Second Edition). Second Edition. San Diego : Academic Press, 2002 : 427 – 484.
<https://www.sciencedirect.com/science/article/pii/B9780126930191500198>.
- [61] 张鹏, 李佳翔, 刘新国, et al. 基于深度学习的计算机辅助肺癌诊断方法 [J]. 计算机辅助设计与图形学学报, 2018, 30(1): 10.
- [62] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J], 2016.
- [63] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks[J/OL]. Pattern Recognition, 2018, 77 : 354–377.
<https://www.sciencedirect.com/science/article/pii/S0031320317304120>.

-
- [64] SHI W, CABALLERO J, THEIS L, et al. Is the deconvolution layer the same as a convolutional layer?[J], 2016.
 - [65] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
 - [66] WANG Z, BOVIK A, SHEIKH H, et al. Image quality assessment: from error visibility to structural similarity[J/OL]. IEEE Transactions on Image Processing, 2004, 13(4) : 600 – 612.
<http://dx.doi.org/10.1109/TIP.2003.819861>.
 - [67] LARSON E, CHANDLER D. Most apparent distortion: Full-reference image quality assessment and the role of strategy[J/OL]. J. Electronic Imaging, 2010, 19 : 011006.
<http://dx.doi.org/10.1117/1.3267105>.
 - [68] SHEIKH H, BOVIK A, de VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J/OL]. IEEE Transactions on Image Processing, 2005, 14(12) : 2117 – 2128.
<http://dx.doi.org/10.1109/TIP.2005.859389>.
 - [69] SHEIKH H, BOVIK A. Image information and visual quality[J/OL]. IEEE Transactions on Image Processing, 2006, 15(2) : 430 – 444.
<http://dx.doi.org/10.1109/TIP.2005.859378>.
 - [70] FAWCETT T. An introduction to ROC analysis[J/OL]. Pattern Recognition Letters, 2006, 27(8) : 861 – 874.
<https://www.sciencedirect.com/science/article/pii/S016786550500303X>.

简历与科研成果

基本信息

侯韵晗，女，汉族，1998年2月出生，河南省新乡人。

教育背景

2020年9月—2022年6月	南京大学软件学院	硕士
2016年9月—2020年6月	南京大学软件学院	本科

攻读工程硕士学位期间完成的学术成果

1. Yunhan Hou, Jiawei Liu, Daiwei Wang, Jiawei He, Chunrong Fang, and Zhenyu Chen, “TauMed: Test Augmentation of Deep Learning in Medical Diagnosis,” in *Proc. 30th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA) 2021*, July. 2021.