**学号：2021030509**

**北京化工大学**

**毕业设计（论文）开题报告**

**论文题目：基于GAN的医疗机器人图像生成与修复方法研究**

**学院名称：机电工程学院**

**专 业：机器人工程**

**学生姓名：廖文萱**

**导师姓名：郄龙飞（讲师）**

**开题日期：2024.12.9**

|  |  |  |
| --- | --- | --- |
| **考核**  **成绩** | **合格** | |
| **审核小组成员以及职称** | **姓 名** | **职 称** |
|  |  |
|  |  |
|  |  |
|  |  |

# 开题报告

# 一、选题背景和意义

随着人工智能技术的飞速发展，医疗诊断机器人在临床应用中正逐渐成为一种重要的辅助工具。视觉判定作为医疗诊断机器人的核心技术之一，其算法的强化对于提升诊断的准确性和效率具有至关重要的作用。然而，当前的视觉判定算法在实际应用中仍存在诸多不足，如噪声去除不彻底、图像增强效果不佳、病灶识别准确率低等问题，这些不足严重影响了医疗诊断机器人的实际操作性和可靠性[1]。

为了解决上述问题，本研究聚焦于“基于生成对抗网络（GAN）的医疗机器人图像生成与修复方法研究”。生成对抗网络作为一种先进的深度学习技术，已经在图像生成、图像修复、图像超分辨率等领域展现出强大的性能。通过引入生成对抗网络技术，本研究旨在提升医疗诊断机器人的图像处理能力，从而增强其在临床环境中的应用效果。

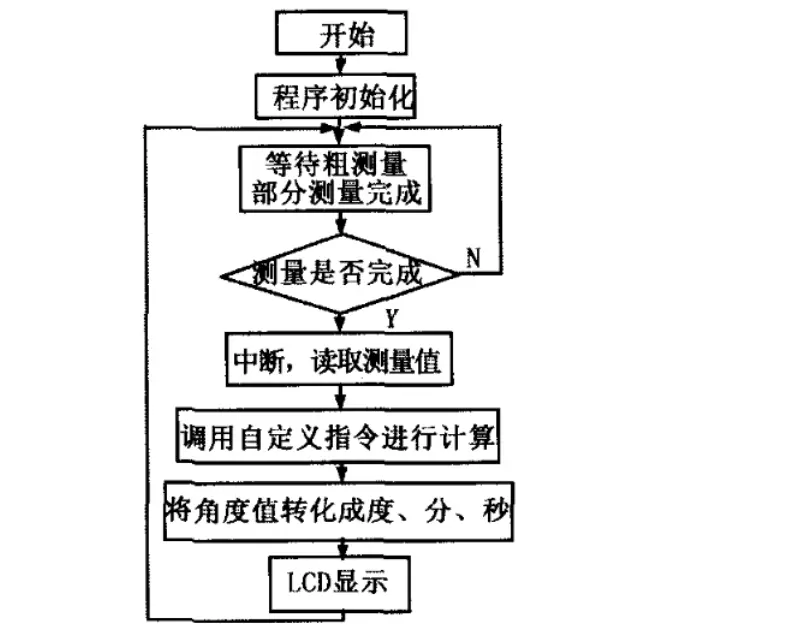
本研究具有一定学术意义和实际应用价值。首先，通过引入生成对抗网络技术，可以有效解决现有医疗图像处理算法在噪声去除、图像增强和病灶识别等方面的局限性，提高图像处理的准确性和效率。其次，本研究将为医疗诊断机器人的视觉判定算法提供新的思路和方法，推动该领域的技术进步。此外，本研究还将探讨生成对抗网络在医疗图像生成与修复中的应用潜力，为未来的研究提供理论和实践基础。

## 二、国内外研究现状

## **2.1 医疗机器人图像处理研究**

医疗机器人技术是近年来迅速发展的研究领域，它结合了医学、机器人工程、自动化和先进制造等多个学科的技术，以改善医疗保健工作质量和效率。医疗机器人在外科手术、诊断、治疗等许多方面都发挥着重要作用，如胶囊机器人、超声机器人、微创穿刺机器人等。

在医疗机器人图像处理领域，国内外学者已经开展了大量的研究工作。例如，2014年，李杰等提出了针对胶囊机器人小体积限制其携带能量的问题，提出了微视觉技术的思想，并利用基于Bayer型彩色过滤阵列的图像传感器采集到的Raw RGB数据进行传输、存储操作。该论文还提供了一种基于最优化彩色空间转换的多梯度彩色插值算法，实验结果显示，此方法减小了误差，提高了峰值信噪比，改善了图像的边缘模糊现象[1]。



**图2-1：微视觉技术处理流程**

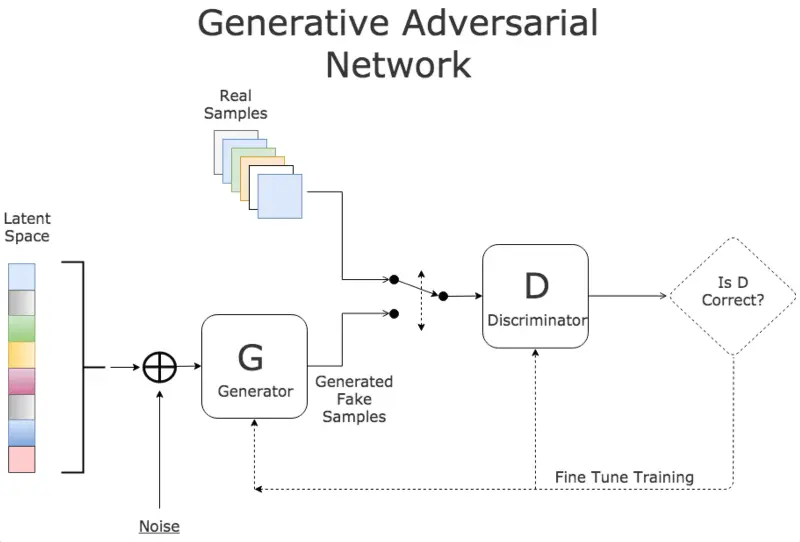
早期，王田苗等对应用在医疗外科领域中的机器人进行了定义和发展历史概述，分析了医疗机器人的分类和结构特点，阐述了医用机器人的应用优缺点，并对其未来发展进行了展望[2]。2021年，赵新刚等则强调了医疗机器人的高速发展来源于需求与技术的共同促进，其中医学、计算机科学、自动化、先进制造等学科的发展和高度融合催生出大量革新技术[3]。

2022年，李等人介绍了医疗领域超声机器人的主要应用方向，包括超声机器人的优点以及在手术、放射治疗、远程诊断中的应用[4]。2023年，Shuai等人回顾了医疗影像导航技术在微创穿刺机器人中的应用研究现状，突出了其未来的发展趋势和挑战[5]。

## **2.2 生成对抗网络的引入**

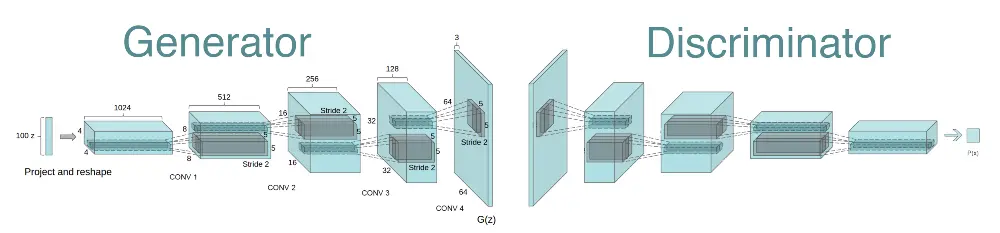
生成对抗网络（GAN）是深度学习中的一种重要技术，它通过两个神经网络的相互竞争和学习，可以生成逼真的数据。GAN在计算机视觉、自然语言处理等领域有着广泛的应用，如图像生成、文本合成图片等。近年来，随着研究的深入，GAN模型也在不断地改进和优化，以适应更多的应用场景。

在医疗机器人图像处理领域，生成对抗网络的应用也逐渐受到关注。例如，2020年的一篇论文对生成对抗网络的发展动向和理论研究的最新进展进行了阐述，详细介绍了GAN的基本架构、原理及其优势和劣势，并对几种常见的GAN改进和变体进行了比较[6]。



**图2-2：生成对抗网络基本逻辑图**

2021年，一篇论文从零和博弈的角度出发，对GAN模型进行了深入的研究，认为GAN可以通过无监督学习获得数据的分布，并能生成较逼真的数据。该论文对GAN的改进和扩展的研究成果进行了广泛的研究，并从图像超分辨率重建、文本合成图片等多个实际应用领域展开讨论，系统地梳理、总结出GAN的优势与不足[7]。

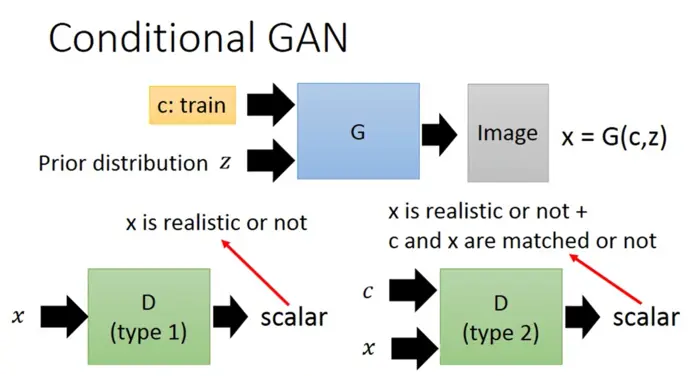


**图2-3：深度卷积生成对抗网络(DCGAN)**

2023年，一篇论文对GAN的研究进展和基本思想进行了介绍，并对一些经典的GAN，如深度卷积生成对抗网络（DCGAN）、条件生成对抗网络（CGAN）、WGAN和超分辨率生成对抗网络（SRGAN）等进行了综述。最后，该论文对GAN的相关工作进行了总结与展望[8]。

## **2.3 GAN在医疗机器人图像处理中的应用**

生成对抗网络在医疗机器人图像处理中的应用主要集中在医学图像生成、图像去噪、图像修复等方面。例如，2019年，Arjun等人提出了使用条件生成对抗网络（cGAN）来生成解剖学上准确的全尺寸CT图像。他们的方法是基于最近发现的风格迁移概念，并提议将两个单独的CT图像的风格和内容混合起来生成新图像。他们认为，通过在基于风格迁移的架构中使用这些损失以及cGAN，可以多倍地增加临床准确、带注释的数据集的大小。他们的框架可以为所有器官生成具有新颖解剖结构的高分辨率全尺寸图像，并且只需要有限数量的患者输入数据[9]。

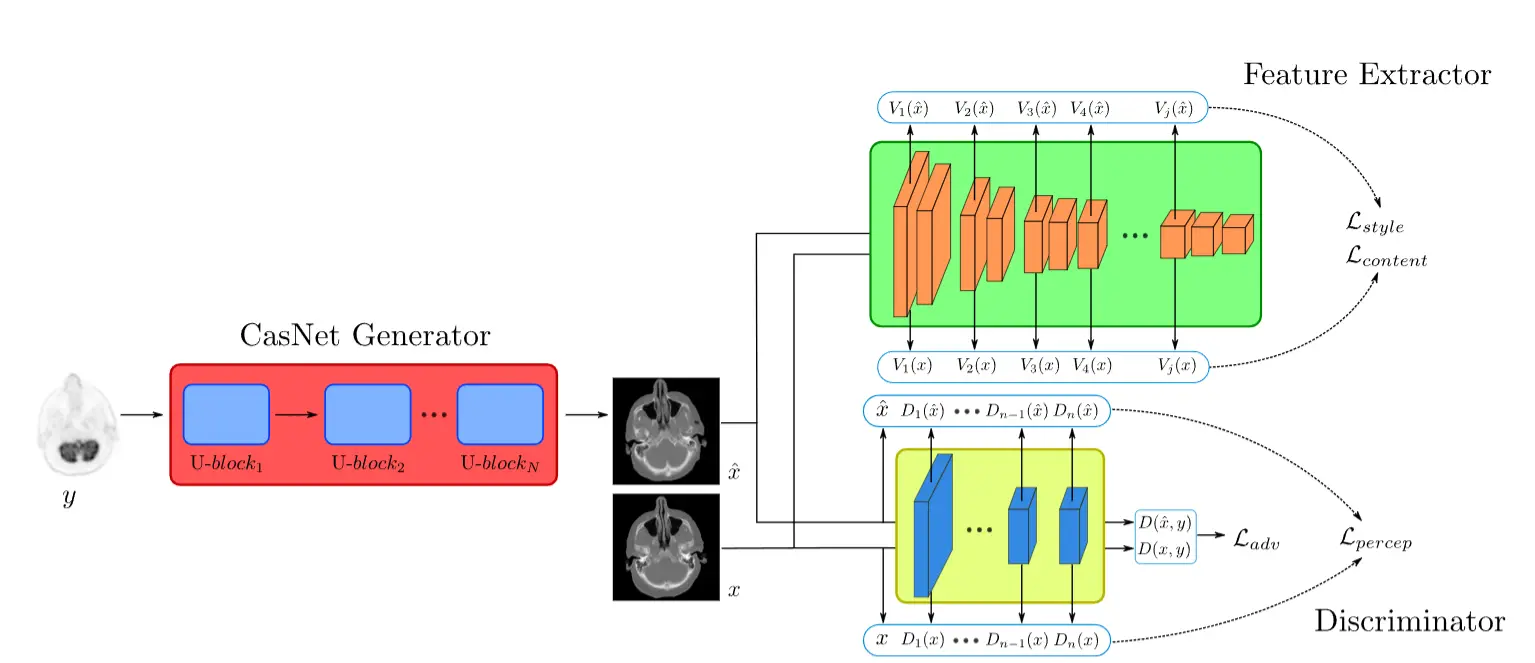


**图2-4：CGAN基本逻辑**

2021年，Euijin等人针对自然图像设计的现有模型难以生成高质量的3D医学图像的问题，提出了一种新的cGAN方法。这种方法首先使用基于注意力的2D生成器生成一系列2D切片，然后通过一组2D和3D鉴别器来确保3D空间中的一致性。此外，他们还提出了一种基于注意力分数的自适应身份损失，以适当地转换与目标条件相关的特征。实验结果表明，该方法可以生成在不同阿尔茨海默症阶段的平滑且逼真的3D图像[10]。

2022年，Bo等人指出，在医学图像获取过程中，未知的混合噪声会影响图像质量。然而，现有的去噪方法通常只关注已知的噪声分布。他们提出了一种基于噪声生成网络的盲去噪方法，该方法能够有效地去除未知的混合噪声，提高图像质量[11]。

此外，生成对抗网络还被应用于高分辨率医学图像生成。例如，2019年，一篇论文提出了使用生成对抗网络（GANs）来合成高质量的视网膜图像及其对应的语义标签图。与其他方法不同，该方法采用了两步策略：首先，通过逐步增长的GAN生成描述血管结构的语义标签图；然后，使用图像到图像的转换方法从生成的血管结构中获得真实的视网膜图像。这种方法只需要少量的训练样本就可以生成逼真的高分辨率图像，从而有效地扩大了小型数据集[12]。



**图2-5：MedGAN框架**

在2021年，另一篇论文针对医学图像领域中数据稀缺的问题，提出了一种基于生成对抗网络的数据增强协议。该协议在像素级（分割掩码）和全局级信息（采集环境或病变类型）上对网络进行条件化，从而控制合成图像的全局类特定外观。为了刺激与分割任务相关的特征的合成，还在对抗游戏中引入了一个额外的被动玩家[13]。

到了2023年，考虑到医学图像与典型的RGB图像在复杂性和维度上的差异，一篇论文提出了一种自适应的生成对抗网络，称为MedGAN。该方法首先使用Wasserstein损失作为收敛度量来衡量生成器和鉴别器的收敛程度，然后基于此度量自适应地训练MedGAN。实验结果验证了MedGAN在模型收敛、训练速度和生成样本的视觉质量方面的优势[14]。

综上所述，生成对抗网络在医疗机器人图像处理中的应用已经取得了显著的进展，但仍存在一些挑战，如模式崩溃、梯度消失等问题。未来的研究需要进一步优化GAN模型，以提高其在医疗图像生成与修复中的应用效果。

## 三、拟研究的主要内容和思路

## **3.1 主要研究内容**

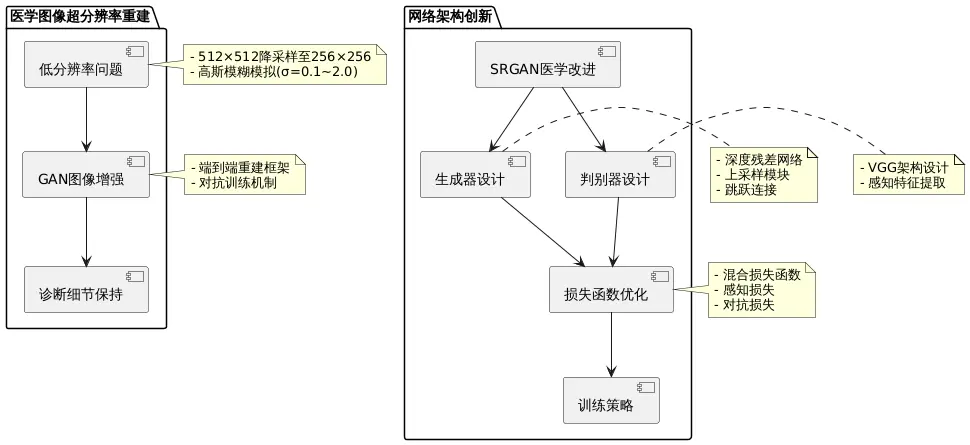
### **3.1.1 医学图像超分辨率重建**

在医学图像超分辨率重建方面，将重点关注三个核心问题： 首先，针对医学图像的低分辨率问题，深入研究MRI医学图像在采集和传输过程中的分辨率损失现象，重点分析512×512高分辨率MRI图像的特征分布和细节表现，探索256×256低分辨率图像到512×512高分辨率图像的映射关系，并研究高斯模糊(σ=0.1~2.0)对图像质量降级的模拟效果。其次，通过生成对抗网络(GAN)提高图像质量，设计基于GAN的端到端超分辨率重建框架，研究生成对抗训练在医学图像重建中的应用，分析对抗学习对图像纹理和细节重建的影响，探索GAN在保持医学图像诊断价值方面的优势。最后，在保持医学诊断细节方面，研究图像归一化处理对医学特征保持的影响，分析RGB三通道信息对医学图像表达的重要性，探索不同数据增强方法对诊断细节的保护机制，并评估重建图像在医学诊断应用中的实用性。

### 3.1.2 网络架构创新

在网络架构创新方面，将基于SRGAN进行深入改进: 首先，针对医学图像的特点，设计专门的数据预处理流程，研究医学图像数据集的构建和批处理策略，探索特定的数据增强技术，并分析自定义transforms对网络性能的影响。其次，在生成器网络结构设计上，将采用深度残差网络结构进行特征提取，深入研究上采样模块在医学图像重建中的作用，分析skip connection对细节保持的贡献，并探索注意力机制在医学特征增强中的应用价值。在判别器设计方面，将基于VGG架构设计判别器，研究判别器深度对重建质量的影响，分析感知特征提取的层次选择，并探索适用于医学图像的特定判别策略。针对损失函数的优化，本研究将设计针对医学图像的混合损失函数，研究像素级损失与感知损失的平衡关系，分析对抗损失对图像真实度的提升效果，并探索医学特征保持损失的引入机制。最后，在训练策略创新上，将设计渐进式训练方案，研究学习率调整对网络收敛的影响，分析批大小选择对训练稳定性的作用，并探索模型微调和迁移学习策略的应用。

### 3.1.3 研究内容架构图



**图3-1：研究内容架构**

## **3.2 主要研究思路**

### 3.2.1数据处理与准备

首先着手构建高质量的医学影像数据集，通过系统性收集和筛选MRI医学影像数据，建立以512x512分辨率为标准的基准数据集，并重点确保数据在临床应用场景中具有充分的代表性和多样性。

在数据预处理环节，设计并实现一套标准化的图像预处理流程，包括建立高低分辨率图像的精确配对机制，同时引入多样化的数据增强技术，如随机旋转、翻转和对比度调整等，以提升模型在实际应用中的鲁棒性和泛化能力。通过这些严谨的数据处理步骤，为后续的模型训练和验证奠定坚实的数据基础。

### 3.2.2 改进型超分辨率模型设计

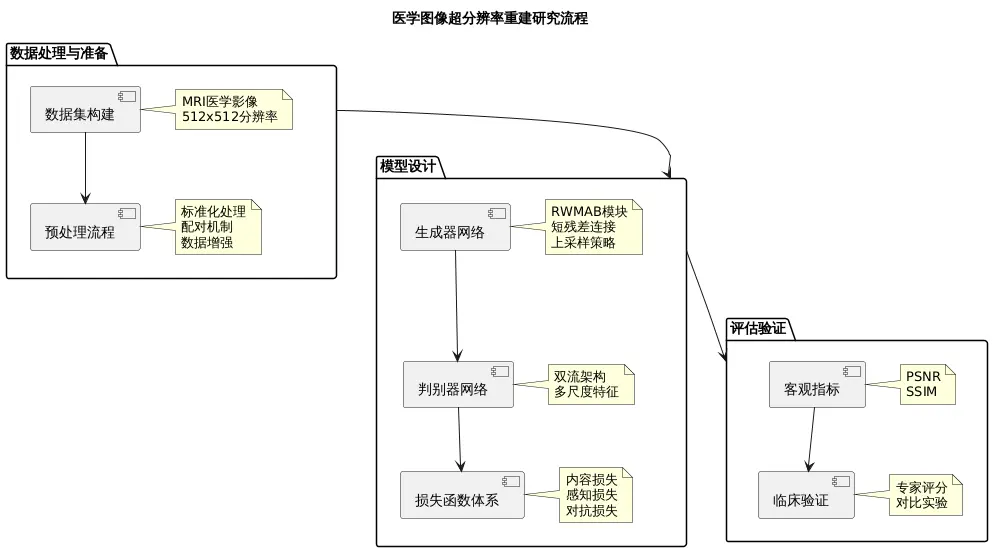
就模型设计而言，从生成器网络、判别器网络和损失函数体系三个关键方向展开深入优化。针对生成器网络，创新性地设计残差宽多注意力模块(RWMAB)以显著增强特征提取能力，同时引入短残差连接机制来改善网络的梯度传播效果，并对上采样策略进行优化以更好地保持医学图像的细节信息。

进行判别器网络创新，构建双流判别架构以提升对图像真实度的判别能力，设计多尺度特征提取模块以捕获不同层级的图像特征，并对损失函数的权重分配进行精细化调整。

而对于损失函数体系的构建，将采用多目标优化策略，通过结合内容损失来保证重建图像的结构准确性，整合感知损失以提升整体视觉质量，引入对抗损失来增强图像真实感，同时创新性地设计针对医学图像特定的质量约束项，以确保重建结果满足临床应用需求。

### 3.2.3 评估与验证体系

建立一套全面的评价体系，从客观评价指标和临床应用验证两个维度展开。在客观评价方面，采用峰值信噪比(PSNR)作为基础指标评估重建图像的准确度，运用结构相似性指数(SSIM)衡量重建图像与原始图像在结构特征上的保持程度，同时引入专门针对医学图像特征的质量度量指标，以全面评估重建效果。在临床应用验证层面，将系统性地开展对比实验研究，通过与现有主流超分辨率方法进行性能对比，并建立由专业医生参与的评分机制，确保重建结果满足实际临床诊断需求，为模型的实用性提供可靠的验证依据。



**图3-2：医学图像超分辨率重建研究流程**

# 四、研究的难点及创新点

## **4.1 研究难点:**

首先，医学图像具有显著的特殊性，不仅对图像细节的保真度有着极其严格的要求，而且不同成像模式(如CT、MRI等)之间存在明显的特征差异，同时还需要确保对诊断至关重要的关键结构信息得到完整保留。

其次，在模型设计层面存在诸多挑战，需要在图像质量提升和超分辨率倍数之间寻找最佳平衡点，同时要有效避免重建过程中伪影和噪声的引入，此外模型的训练过程往往需要消耗大量计算资源并且耗时较长。

最后，在评估标准的制定方面也存在一定困难，由于传统的图像评价指标可能并不完全适用于医学图像的特殊需求，而目前医学图像质量评估尚缺乏统一的标准体系，这给研究工作带来了额外的挑战。

## **4.2 创新点：**

在网络架构方面，通过设计专门的医学图像特征提取模块来适应医学影像的独特性质，同时引入多尺度特征融合机制以捕获不同层级的图像信息，并结合注意力机制来增强对关键区域的特征学习能力。

在损失函数设计上，创新性地结合医学图像特征构建感知损失，引入结构相似性保持损失以确保重建图像的结构准确性，并对传统对抗损失进行改进以提升生成效果。针对数据处理环节，建立特定的医学图像预处理流程，优化数据增强策略以扩充训练样本，并开发多模态图像处理能力以适应不同类型的医学影像数据。

在应用创新层面，着重提升系统的实用价值，支持CT、MRI等多种医学成像模式的处理，具备实时处理能力，并充分考虑临床实际应用需求，确保研究成果能够切实服务于医疗实践。

## **4.3 技术亮点：**

(1)端到端的超分辨率重建架构，实现从低分辨率到高分辨率图像的直接转换；

(2)通过引入自适应特征学习机制，使模型能够根据不同类型医学图像的特征自动调整学习策略；

(3)设计了高效的训练策略，在保证模型性能的同时有效降低了计算资源消耗；

(4)良好的泛化能力，能够适应不同成像设备和不同病例的图像处理需求。

## 五、进度计划

|  |  |  |
| --- | --- | --- |
|  | 设计（论文）各阶段名称 | 起止日期 |
| 1 | 文献调研与理论准备 | 2024.12 - 2025.1 |
| 2 | 数据集构建与预处理 | 2025.1 - 2025.2 |
| 3 | 基础模型实现 | 2025.2 - 2025.3 |
| 4 | 模型优化与改进 | 2025.3 - 2025.4 |
| 5 | 实验评估与性能验证 | 2025.4 - 2025.5 |
| 6 | 论文撰写与项目总结 | 2025.5 - 2025.6 |

## 六、参考文献

1. 李杰,程磊,徐建省,吴怀宇,陈洋.基于彩色图像插值算法的胶囊机器人微视觉[J].计算机测量与控制,2014,22(2):503-506509
2. 王田苗,宗光华,张启先.新应用领域的机器人——医疗外科机器人[J].机器人技术与应用,1997(2):7-9
3. 赵新刚, 段星光, 王启宁, 夏泽洋. 医疗机器人技术研究展望[J]. 机器人, 2021, 43(4): 385-385.
4. 李奇轩,张钒,奚谦逸,焦竹青,倪昕晔.医疗超声机器人的研究进展[J].中国医疗设备,2022,37(8):21-2461
5. Hu S, Lu R, Zhu Y, Zhu W, Jiang H, Bi S. Application of Medical Image Navigation Technology in Minimally Invasive Puncture Robot. Sensors. 2023; 23(16):7196. https://doi.org/10.3390/s23167196
6. 彭泊词,邵一峰.生成对抗网络研究及应用[J].现代计算机,2020,26(27):42-48
7. 张恩琪,顾广华,赵晨,赵志明.生成对抗网络GAN的研究进展[J].计算机应用研究,2021,38(4):968-974
8. 于文家,樊国政,左昱昊,陈怡丹.生成对抗网络研究综述[J].电脑编程技巧与维护,2023(5):174-176
9. Krishna, A., & Mueller, K. (2019). Medical (CT) image generation with style. 15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine.
10. Euijin Jung, Miguel Luna, and Sang Hyun Park. 2021. Conditional GAN with an Attention-Based Generator and a 3D Discriminator for 3D Medical Image Generation. In Medical Image Computing and Computer Assisted Intervention – MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VI. Springer-Verlag, Berlin, Heidelberg, 318–328. https://doi.org/10.1007/978-3-030-87231-1\_31
11. Fu, B., Zhang, X., Wang, L., Ren, Y., & Thanh, D. N. H. (2022). A blind medical image denoising method with noise generation network. Journal of X-ray science and technology, 30(3), 531–547. https://doi.org/10.3233/XST-211098
12. Andreini, P., Ciano, G., Bonechi, S., Graziani, C., Lachi, V., Mecocci, A., Sodi, A., Scarselli, F., & Bianchini, M. (2022). A Two-Stage GAN for High-Resolution Retinal Image Generation and Segmentation. Electronics, 11(1), 60. https://doi.org/10.3390/electronics11010060
13. Qasim, A.B., Ezhov, I., Shit, S., Schoppe, O., Paetzold, J.C., Sekuboyina, A., Kofler, F., Lipkova, J., Li, H. & Menze, B.. (2020). Red-GAN: Attacking class imbalance via conditioned generation. Yet another medical imaging perspective.. Proceedings of the Third Conference on Medical Imaging with Deep Learning, in Proceedings of Machine Learning Research 121:655-668. https://proceedings.mlr.press/v121/qasim20a.html.
14. Guo, Kehua & Chen, Jie & Qiu, Tian & Guo, Shaojun & Luo, Tao & Chen, Tianyu & Ren, Sheng. (2023). MedGAN: An adaptive GAN approach for medical image generation. Computers in Biology and Medicine. 163. 107119. 10.1016/j.compbiomed.2023.107119.

|  |
| --- |
| **指导教师意见：**  指导教师签名：  20xx年xx月xx日 |
| **审核小组意见：**  审核小组组长签字：  20xx年xx月xx日 |
| **学院意见：**    20xx年xx月xx日 |
| **备注：** |