基于改进nnU-Net的3D腹部多器官图像智能分割解析平台

组长姓名，

成员姓名，

成员姓名，

成员姓名，

**摘要**：本大作业基于nnUNet模型进行腹部多器官分割预测。通过介绍nnUNet的基本原理和工作流程，包括数据预处理、网络结构和训练策略等，使用该模型对腹部CT图像中的多个器官进行分割预测。项目中应用高斯模糊、伽马变换等技术优化数据，并调整训练参数和引入注意力机制以提高模型性能。

**关键词:**腹部多器官分割，nnUNet，深度学习，医学图像处理，数据预处理，注意力机制

**引言**

计算机视觉是人工智能领域的一个重要分支，致力于让计算机理解和处理视觉信息。近年来，计算机视觉技术在各个领域取得了显著进展，尤其在医疗影像分析中展现出了巨大潜力。医学影像作为临床疾病诊断的重要工具，通过计算机视觉技术可以实现高效、准确的图像分析，从而辅助医生进行诊断和治疗规划。

腹部多器官分割预测是医学图像处理中的一个关键任务，其目标是从腹部CT图像中自动识别并分割出多个器官（如肝脏、胆囊、胰腺等）。传统的图像分割方法通常依赖于人工特征提取和经典的机器学习算法，这不仅耗时费力，还容易受到主观因素的影响。随着深度学习技术的发展，基于深度神经网络的方法在图像分割任务中表现出色，其中nnUNet模型作为一种通用的分割框架，凭借其自适应性和高精度成为医学图像分割领域的热门选择。

本项目旨在利用nnUNet模型实现腹部多器官的自动分割预测。通过深入理解nnUNet的基本原理和工作流程，并结合数据预处理、训练策略和注意力机制等技术手段，优化模型性能，提升分割精度。本研究不仅为腹部多器官分割提供了一种有效的方法，还为计算机视觉在医学影像中的应用探索了新的方向。

在本次设计中，我们将系统地介绍nnUNet模型的架构和实现过程，详细阐述数据预处理方法和模型优化策略，并通过实验评估模型的性能，展示其在实际应用中的效果。希望通过该研究，进一步推动计算机视觉技术在医学图像处理中的应用发展，为医疗行业带来更多创新和变革。

**算法原理及实现**

1. **nnUNet模型**

nnUNet是一种基于U-Net架构的深度学习模型，它通过端到端的学习过程，能够对各种医疗图像进行自动化的、高精度的分割。nnUNet采用了编码器-解码器结构，其中编码器用于捕获图像的高级特征，而解码器则用于生成分割结果。

nnUNet的主要特点包括：

（1）自动化配置：根据输入数据的特性自动调整网络结构、预处理参数和训练策略。

（2）多尺度特征提取：通过不同分辨率的特征图进行融合，增强模型对不同尺度目标的分割能力。

（3）数据预处理：包括归一化、数据增强（如旋转、翻转、缩放）等，提高模型的泛化能力。

（4）损失函数：采用联合损失函数（如Dice系数和交叉熵），更好地处理类别不平衡问题。

1. **数据准备与预处理**

为了使模型更好地理解和处理输入数据，我们需要对原始的腹部CT图像进行一些预处理操作，如高斯模糊、伽马变换、高斯滤波器等。这些操作可以增强模型的泛化能力，并加快训练速度。

**2.1下载实验数据**

本次课程设计抽取AMOS2022数据集中的部分数据作为训练和评测数据。

多器官数据集共有15个器官的标注，包括脾脏、右肾、左肾、胆囊、食道、肝、胃、主动脉、下腔静脉、胰腺、右肾上腺、左肾上腺、十二指肠、膀胱、前列腺/子宫，数据集展示如图2-1。

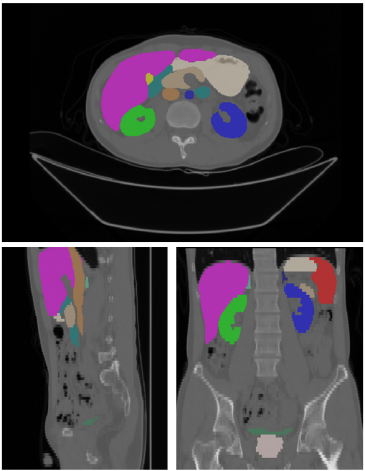


图 2-1 数据集展示

使用下面的代码可以读取nii.gz格式的医疗数据文件:

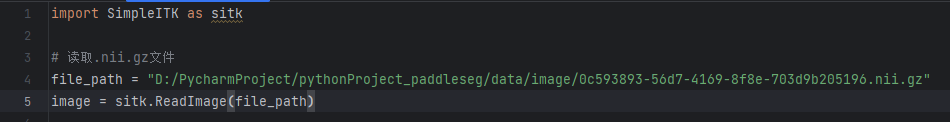


图 2-2 医疗数据读取

**2.2初步分析数据**

使用以下代码获取数据的基本信息：



图 2-3 分析数据代码片段

数据集的基本信息：

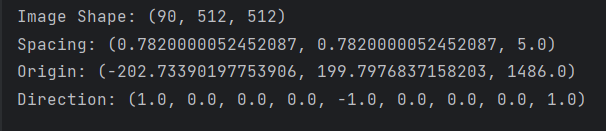


图 2-4 数据集信息

该条数据的大小为（90,512,512），体素间距为（0.78,0.78,5.0），数据集包含以下内容： -imagesTr 训练集目录，160个训练数据 -labelsTr 训练集标签 -imagesTs 测试集目录，40个测试数据 -dataset.json 数据集说明文件。

**数据可视化：**

为了更深入地理解数据，我们不仅需要对其进行基本的统计分析，还需要进行更高级的数据三视图分析和3D可视化。数据三视图分析可以从不同的角度对数据进行全面的审视，帮助我们发现数据中的潜在规律和趋势。而3D可视化则可以将数据以更直观、更生动的方式呈现出来，使我们能够更轻松地理解和解释数据。通过这些高级的分析技术，我们可以更好地把握数据，为决策提供更准确、更全面的支持。

**数据三视图分析：**

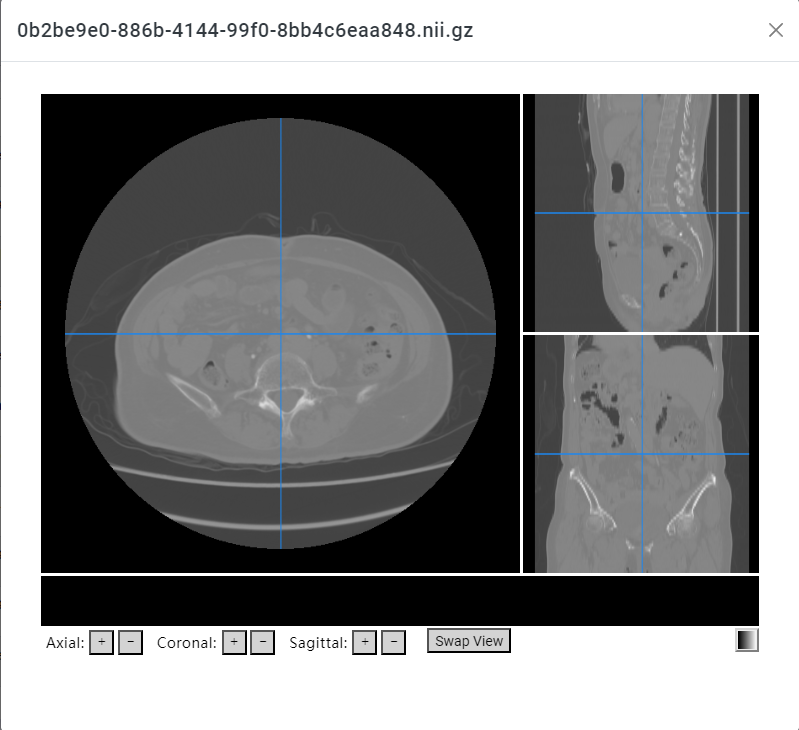


图 2-5 三视图分析

**数据3D可视化：**

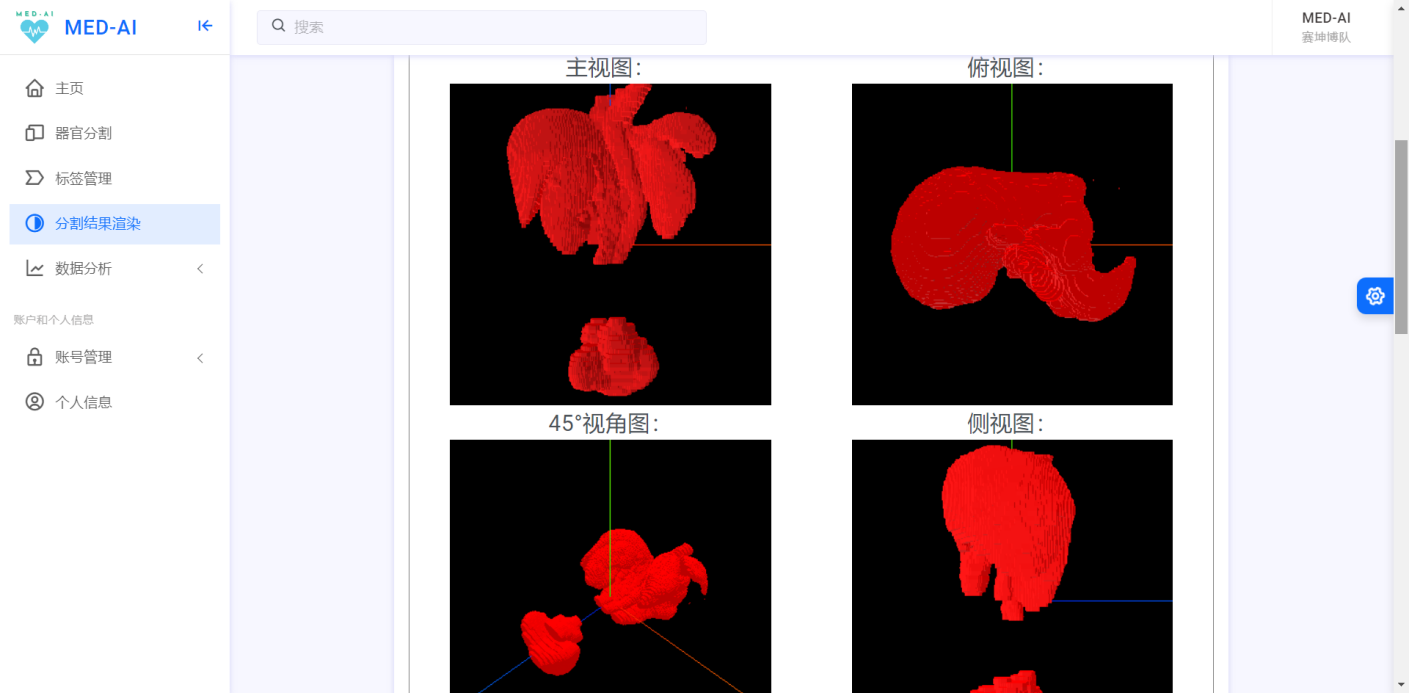


图 2-6 3D可视化

**2.3样本数据预处理**

**高斯模糊**

高斯模糊是一种常用的图像平滑技术，它通过一个高斯核对图像进行卷积操作，从而减小图像中的噪声和细节。高斯核是一个二维的矩阵，其中的元素按照高斯分布进行排列。

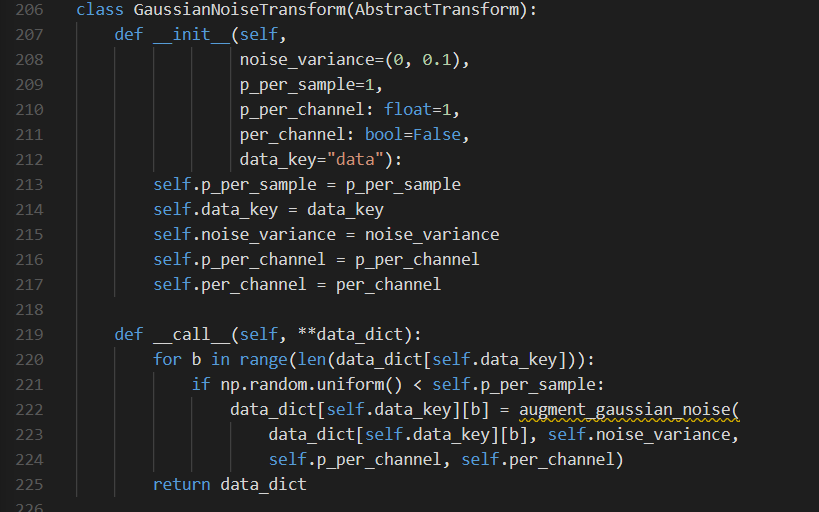


图 2-7 高斯模糊

**伽马变换**

伽马变换是一种将图像从线性空间转换到非线性空间的方法，它可以用来调整图像的亮度和对比度。其公式为I′=I^(1/γ)，其中I是原始图像的像素值，I′是变换后的像素值，γ是伽马系数。当γ>1时，图像的暗部会变得更亮，而亮部则会变得更暗；当0<γ<1时，图像的亮部会变得更亮，而暗部则会变得更暗。在实施伽马变换时，我们需要首先计算每个像素的伽马系数，并将其应用于原始图像。然后，我们可以通过对变换结果进行反向变换，将图像恢复到原始的空间域。

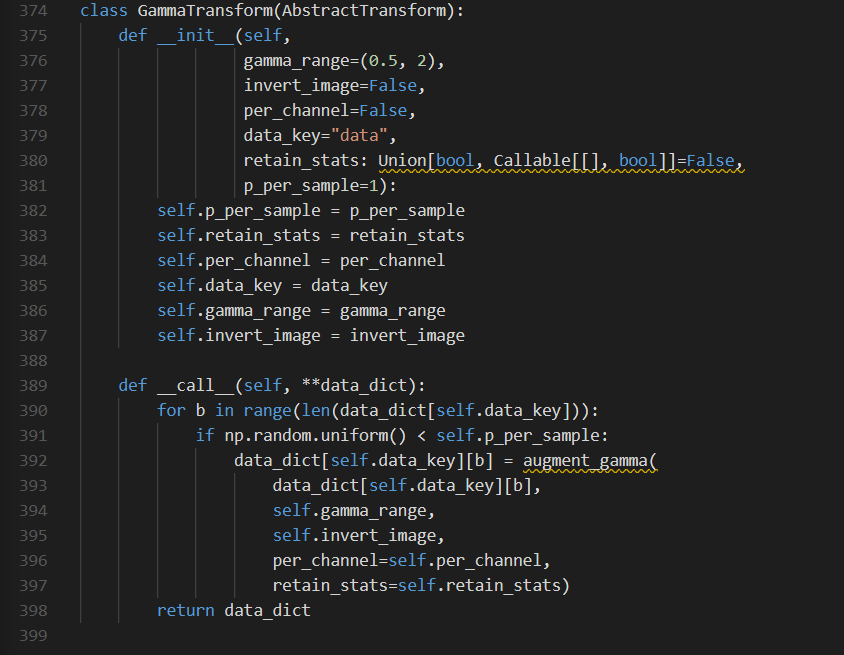


图 2-8 伽马变换

**高斯滤波器**

高斯滤波器是一种基于高斯核的滤波器，它可以用来去除图像中的噪声和不相关特征。在高斯滤波器中，首先定义一个高斯核，然后将其与图像进行卷积操作。通过这种方式，保留图像中的主要结构，同时去除噪声和边缘效应。

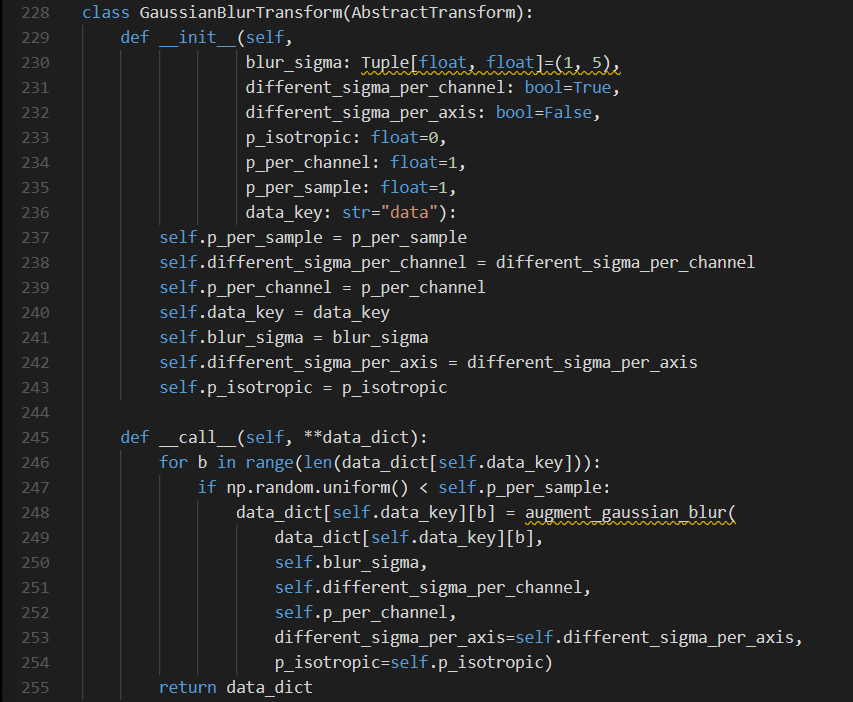


图 2-9 高斯滤波器

1. **模型优化**

**3.1增加训练轮次**

**增加训练轮次的目的**

让模型有更多机会从训练数据中学习和理解模式，随着训练轮次的增加，模型能够逐渐捕获更多的信息，并通过这些信息进行更准确的预测。

**实施过程**

在实际操作中，将训练周期设置为50个epoch，然后观察每个epoch后的模型性能变化。如果发现模型的性能在某个epoch后开始下降，意味着模型已经开始过拟合，此时停止训练以防止过拟合。

**检测指标**

为了判断是否需要增加训练轮次，我们需要监控一些关键的性能指标，如验证集上的精度、召回率或IoU等。这些指标可以帮助我们了解模型的泛化能力以及是否存在过拟合现象。

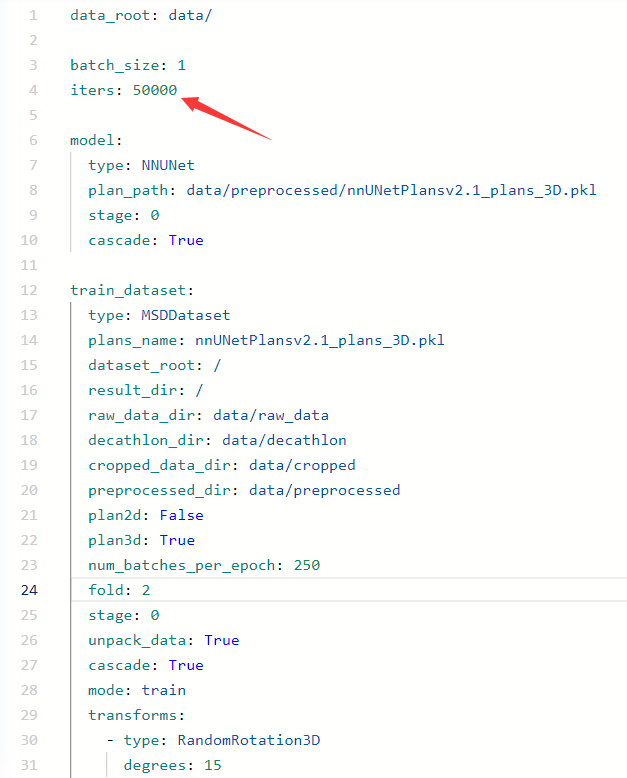


图 3-1 训练轮次设定

**3.2修改Batch Size**

**修改Batch Size的目的**

提高训练效率：通过增加Batch Size，可以减少每次迭代时需要处理的数据量，从而加快训练速度。这对于大规模数据集的训练尤其有利，因为大型数据集需要更多的时间来处理每个Batch。

增加模型泛化能力：Batch Size的增加可以减少模型对训练数据的过拟合，从而增强模型的泛化能力。这是因为更大的Batch Size意味着模型在更新权重时能够看到更多的数据样本，从而更好地泛化到未见过的数据。

**定义实验参数**

我们需要定义一组Batch Size值进行实验。这可以包括一个或多个值，如8、16、32等。

**实施实验**

我们使用这些Batch Size值分别训练模型，并记录每个Batch Size下的模型性能，分析实验结果，以确定最佳的Batch Size值。通过比较不同Batch Size下的验证集精度、召回率或IoU等指标，选择性能最好的那个作为最佳Batch Size。

**验证最佳Batch Size**

最后，我们使用最佳Batch Size重新训练模型，并在测试集上评估其性能。如果测试集上的性能与验证集上的性能一致，那么我们可以认为这个Batch Size是最适合当前任务的。

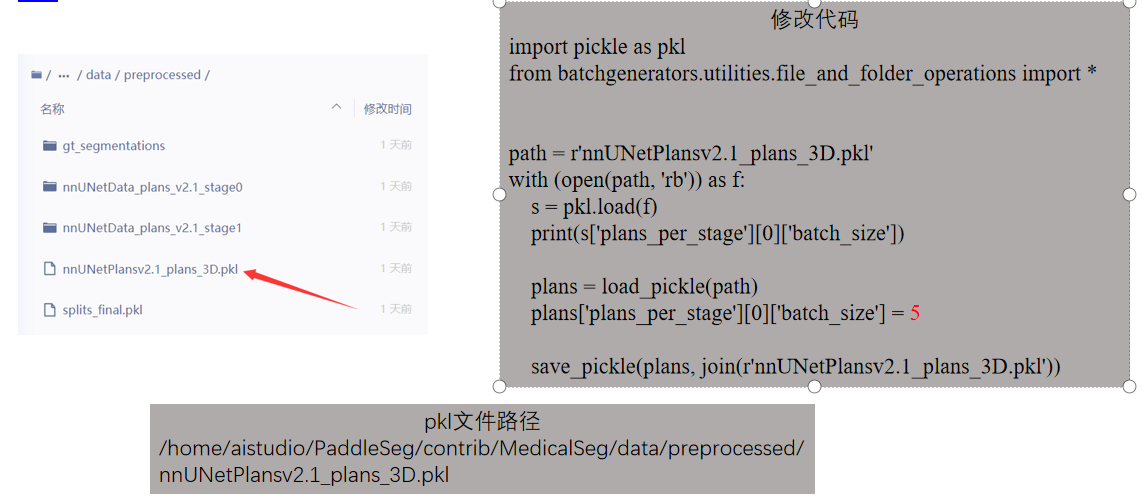


图 3-2 Batch Size修改

**3.3更改优化器**

**优化器的选择**

在深度学习中，优化器是一个非常重要的组成部分，它负责更新模型的权重和偏差。常见的优化器包括梯度下降、随机梯度下降、Adam等。

**梯度下降**

梯度下降是最基本的优化算法，它通过沿着损失函数的梯度方向最小化损失来更新模型的参数。然而，梯度下降可能会遇到局部最优的问题，并且对初始点的选择敏感。

**随机梯度下降**

随机梯度下降是一种改进的梯度下降算法，它通过每次只使用一个样本或一小批样本进行梯度计算，从而加快了训练速度。但是，随机梯度下降可能会导致模型的收敛不稳定。

**Adam**

Adam（Adaptive Moment Estimation）是一种自适应的学习率优化算法，它结合了动量和RMSprop的优点，可以自动调整学习率，并保持模型的稳定收敛。

**更改优化器的影响**

不同的优化器会对模型的训练过程和最终性能产生影响。通过对比不同优化器下的模型性能，我们可以找到最适合当前任务的优化器。



图 3-3 更改优化器

### **3.4五折交叉训练**

**五折交叉训练的概念**

五折交叉训练是一种将数据集分为五个子集的方法，其中四个子集用于训练，一个子集用于测试。然后，我们轮流使用不同的子集进行测试，从而得到五次不同的测试结果。

**五折交叉训练的优势**

五折交叉训练可以有效减少过拟合和欠拟合的问题，因为它可以让模型接触到更多的数据。此外，通过比较五次测试结果，我们可以得到更可靠和稳定的性能指标。

**实施过程**

在实施五折交叉训练时，我们需要首先将数据集随机划分为五个子集。然后，我们轮流使用四个子集进行训练，并用剩下的一个子集进行测试。最后，我们将五次测试结果进行平均，得到最终的性能指标。

**结果分析**

通过对五折交叉训练的结果进行分析，我们可以了解到模型的优点和不足，并提出改进的建议。此外，我们还可以通过与其他方法的比较，来进一步验证模型的有效性和优越性。

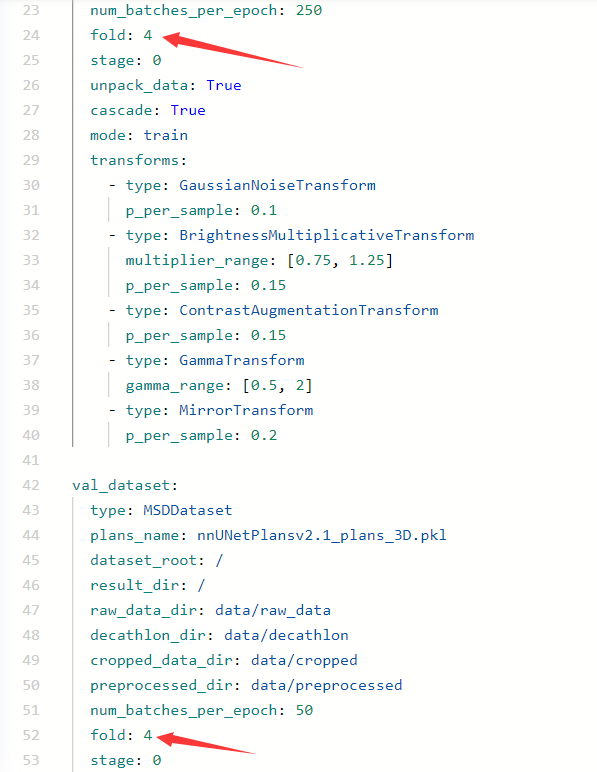


图 3-4 五折交叉训练

### **3.5注意力机制**

我们针对腹部多器官分割任务对医学图像分割领域的SOTA模型nnU-Net进行了改进，引入了LKA（大核卷积注意力）模块以增强模型对于全局信息的理解，捕捉到不同位置的详细信息，以实现长距离依赖性和空间自适应性的同时保留，以提高模型对空间信息的捕捉能力和对特征的表达能力，提高了模型的分割精度。

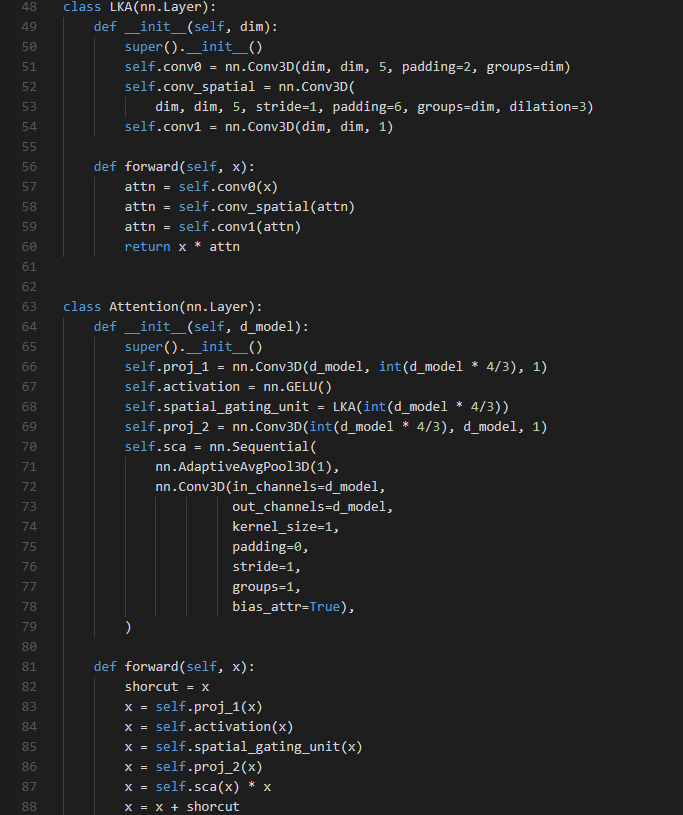


图 3-5 注意力机制

1. **实验结果**

### **4.1实验环境**

1) 系统测试软件配置参考

操作系统：Windows10

Python版本: Python 3.7

框架：PaddleSeg 2.6

项目开发环境：PyCharm2019、IntelliJ IDEA 2022.3.2、Unreal Engine 4

2) 系统硬件配置参考

CPU：8核

GPU: RTX 3050 Ti

RAM（内存）：16GB

显存：16GB

硬盘容量：100GB

### **4.2实验过程**

**4.2.1 系统实现**

平台基于Spring Boot框架进行开发，采用分层架构设计，将应用划分为表现层、应用层、领域层和基础设施层。

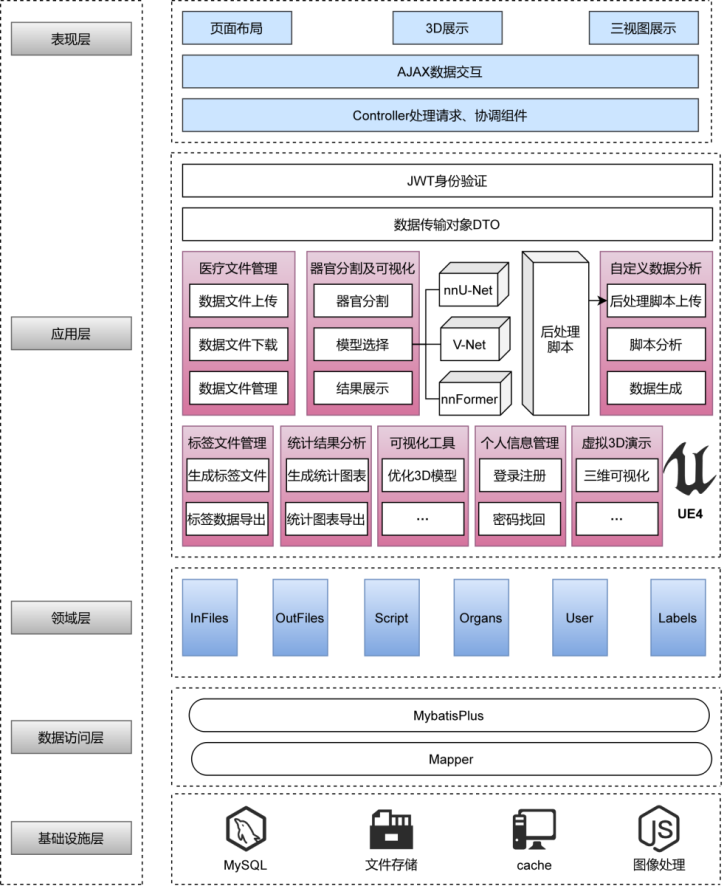


图 4-1 系统总体架构设计

在这个架构中，用户通过与表现层进行交互来获取所需的功能。用户的数据请求由AJAX发送，然后通过控制器（Controller）层进行处理。控制器层负责接收用户请求，并调用其他层次的组件来处理请求。领域层封装了与业务相关的数据实体，是整个应用程序的核心。

应用层负责协调表现层和领域层之间的交互。应用层接收来自表现层的用户请求，并将请求转发给领域层进行处理，也可以对领域层返回的结果进行处理，并将处理结果返回给表现层。在请求数据经过应用层的JWT身份校验成功后，Spring的服务组件（Service）处理业务规则和数据逻辑。服务层通过数据访问层的MyBatis Plus和Mapper与MySQL数据库进行数据交互，以满足相关数据处理的业务需求。

应用层通过将用户上传的医疗数据文件经由训练模型nnU-Net或V-Net进行医学分割，模型的分割结果作为转化为nii文件和stl文件存储相关信息到数据库中，应用层可读取数据库中的相应信息并通过基础设施层的图像处理API进行展示。同时，应用层利用虚幻引擎4来进行医疗图像的图形渲染、动画和场景管理等操作，以实现对医疗图像的优化和增强。

**4.2.2系统功能模块设计**

腹部多器官智能分割解析平台分成了个人信息管理模块、医疗文件管理模块、器官分割及可视化模块、标签文件管理模块、数据分析与统计模块、虚拟3D演示模块、分割结果渲染模块和诊断报告生成模块等八个功能模块。

**个人信息管理模块**主要包括用户的登录注册功能、允许用户管理其个人身份和相关信息；

**医疗文件管理模块**的主要功能为用户上传所需要的医疗数据文件，并对自己上传的医疗数据文件进行管理和三视图查看操作；

**器官分割和可视化模块**的主要功能为进行医疗文件的初步预览并将上传的医疗文件通过算法进行模型生成以及展示模型分割结果；

**标签文件管理模块**主要功能为用户对选择的医疗文件进行标注，并结合其他丰富的工具进行医疗文件的分析以及标签文件的导出；

**数据分析与统计模块**的功能为对选择的医疗文件进行直径、面积、体积等相关数据的统计和展示，并支持用户自行上传数据分析脚本文件对医疗文件进行数据分析；

**虚拟3D演示模块**使用虚幻引擎对分割结果进一步渲染，提高了可视化效果与可交互性，给用户提供最直观的感受；

**分割结果渲染模块**的主要功能为将分割后的医疗模型进行整体的三维展示，根据需求对整体的医疗模型进行分离以便观察单个器官；

**诊断报告生成模块**支持医生填写相关信息和建议后基于算法生成word形式的诊断报告，并可以在web端快速浏览报告信息。

系统功能模块设计如图4-2所示。

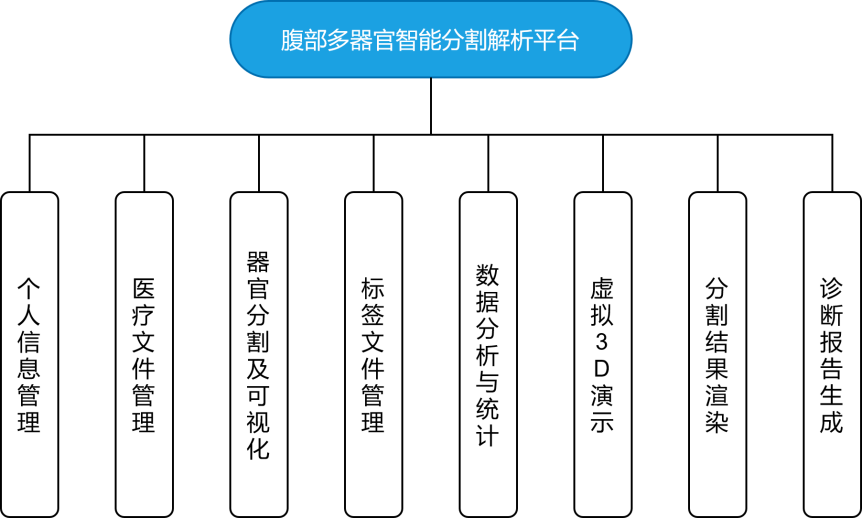
****

图4-2 系统功能模块设计

4.2.3**系统服务端设计与实现**

当客户端发送请求时，首先由拦截器（interceptor）对于除用户登录注册外的请求设置请求拦截。对于初次注册的账号由JWT进行token令牌签发，转发到客户端，当客户端后续访问服务端时需要携带token令牌，由JWT进行身份校验成功后方可通过请求。请求通过后，经由controller层进行请求转发并协调service层处理业务逻辑并返回数据到客户端，当请求需要使用py脚本或分割模型时，service层将调用模型或脚本完成业务逻辑。Service层处理业务逻辑时，需要与repository的数据持久化层交互以从数据库中存储或取出数据，从而完成整个业务逻辑的流程。

系统服务端总体架构设计，如图4-3所示。

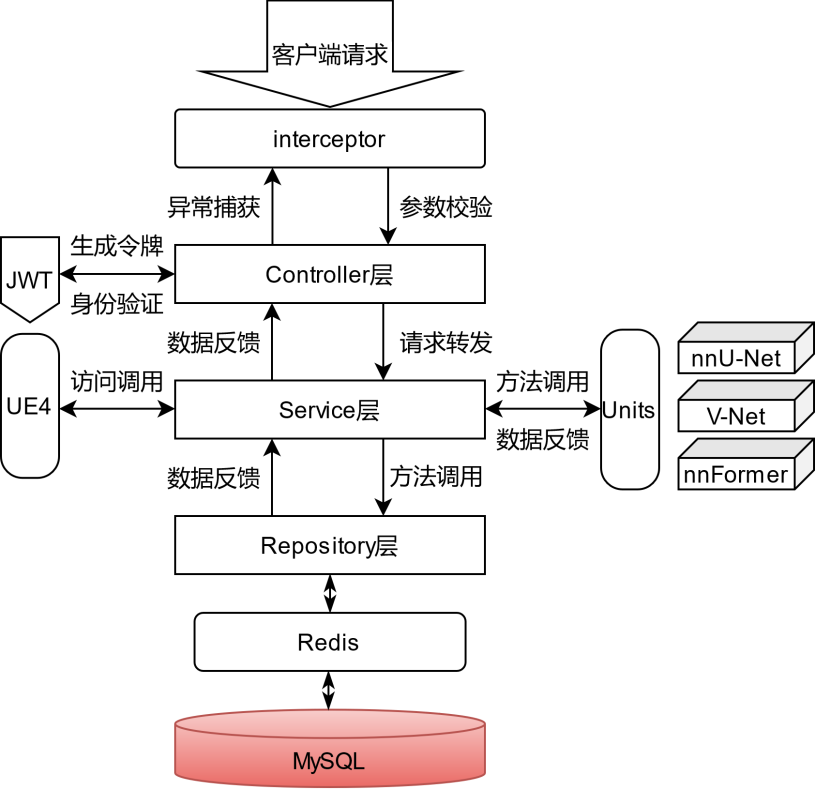


图4-3 系统服务端总体架构

4.2.4实验结果分析

**1) 用户登录界面**：用户访问腹部多器官智能分割解析平台前需要进行登录，已注册的用户可通过输入账号和密码，然后点击登录按钮，若用户信息填写错误会有相关提示信息，信息填写无误即可成功访问智能分割解析平台的系统主页，从而获取所需功能。

用户登录界面如图4-4所示。



图4-4 用户登录界面

**2)医疗数据文件导入与管理界面:**在医疗数据文件导入与管理界面中，用户可以进行医疗数据文件的上传、管理等操作。用户点击选择文件按钮或直接将文件拖拽到选择文件区域即可完成文件上传，文件上传功能支持文件的批量导入与展示。上传文件成功后，界面将以表格的形式展示上传的文件名称、文件类型、文件大小等医疗文件信息，当用户点击表格的删除按钮后，即可从数据库删除相关医疗数据。

医疗数据文件导入与管理界面如图4-5和图4-6所示。



图4-5 医疗数据文件导入界面



图4-6 医疗数据文件管理界面

**3)器官分割及可视化界面:**在器官分割及可视化界面，用户可以通过下拉框根据不同的分割需求选择不同的分割模型，比如选择快速分割模型V-Net、高精度模型nnU-Net和nnFormer模型等适用于不同的分割需求。用户可点击下拉框选择需要分割的医疗数据文件，下拉框选择完毕后，点击开始分割按钮即可根据分割需求对医疗数据文件进行分割操作，分割结束后可以在在数据文件下拉框的下方选择分割结果展示其三视图。

三视图支持在不同视图下进行鼠标操作带动其他视图。在三视图下方提供了多种工具按钮，可以进行三视图的逐帧展示、切换视图以及图像渲染等操作，为用户提供直观的分割结果展示。

器官分割界面与分割结果三视图展示界面如图4-7和图4-8所示。



图4-7 器官分割界面

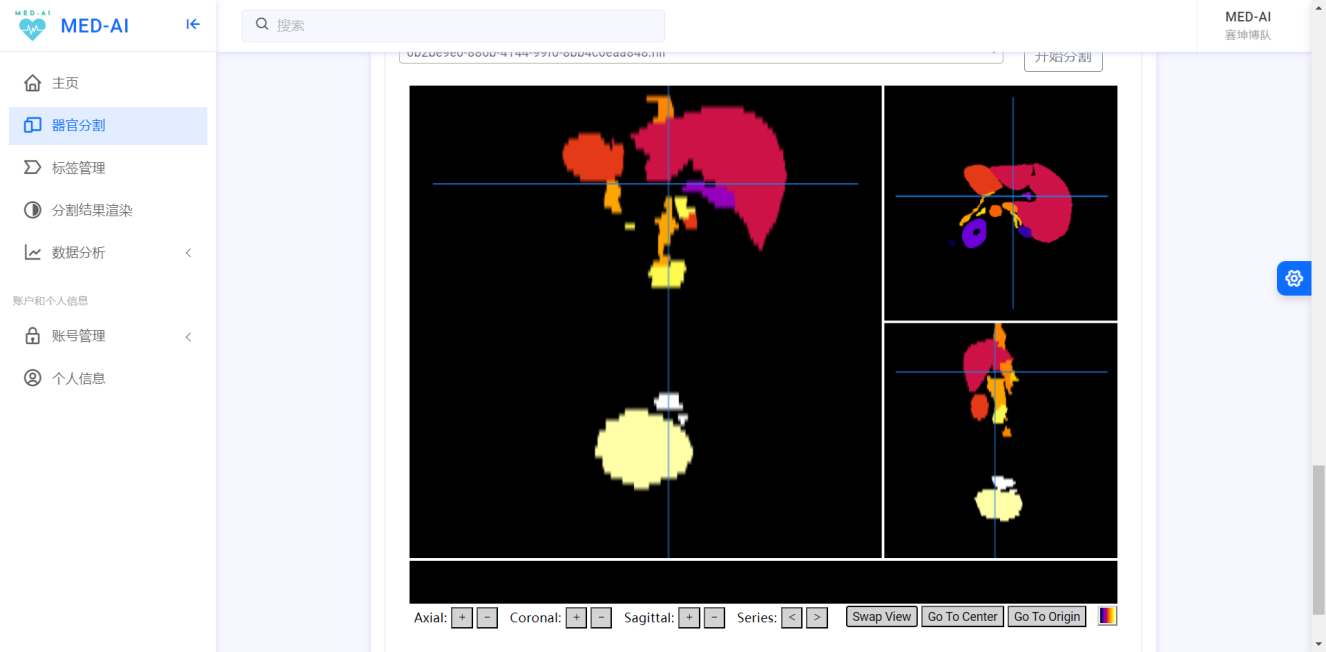


图4-8 分割结果三视图展示界面

**4)分割结果渲染界面:**用户可以在文件渲染界面通过下拉框选择分割结果文件，并通过点击确认按钮来渲染出分割结果的多个视图。每个视图支持鼠标拖动以改变查看方位，同时鼠标拖动还会影响3D原点的变化。文件渲染界面还支持缩放操作，通过滑动鼠标滚轮可以放大或缩小，以便详细查看3D分割结果。

器官分割结果三视图展示界面如图4-9所示。

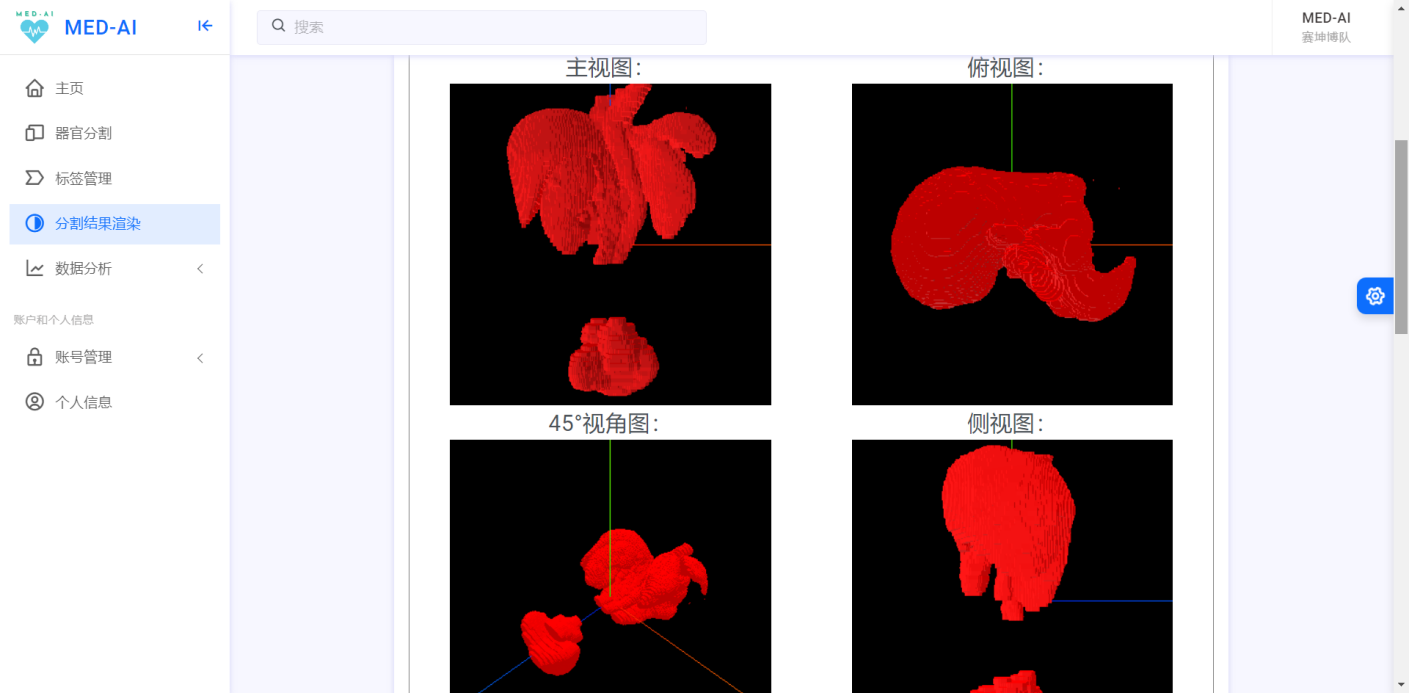


图4-9 器官分割结果三视图展示界面

**4.结论**

本项目成功开发了一款腹部多器官智能分割解析平台，基于先进的深度学习技术和优化算法，能够高效、准确地对腹部CT影像进行多器官自动分割。实验结果表明，我们采用的nnU-Net和V-Net模型在肝脏、脾脏、肾脏等多个器官的分割任务中均表现出色，尤其是对肝脏的分割准确率达到了0.97652。通过引入Rician噪声的数据增强技术，我们进一步提升了模型的泛化能力和分割精度。

平台的设计与实现不仅在准确性上取得了突破，还在系统架构和用户体验上进行了优化。基于Spring Boot框架的分层架构设计，使得系统具备良好的扩展性和维护性；结合虚幻引擎4的三维可视化展示，医生能够更直观地观察到病变区域和器官结构，提高了诊断效率和准确性。

通过对模型的反复训练与测试，我们总结出了一套有效的模型优化策略和系统实现方案，验证了平台在实际应用中的可行性和实用性。平台的应用不仅为医生提供了高效的辅助诊断工具，也为患者提供了快捷的医疗服务，具有重要的临床应用价值和社会意义。未来，我们将继续优化模型和系统功能，进一步提升平台的性能和用户体验，为智能医疗领域的发展贡献力量。

**参考文献**

[1] Isensee F, Petersen J, Klein A, et al. nnu-net: Self-adapting framework for u-net-based medical image segmentation[J]. arXiv preprint arXiv:1809.10486, 2018.

[2] Xie Y, Zhang J, Shen C, et al. Cotr: Efficiently bridging CNN and transformer for 3D medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part III 24. Springer International Publishing, 2021: 171-180.

[3] Luo X, Hu M, Song T, et al. Semi-supervised medical image segmentation via cross teaching between CNN and transformer[C]//International Conference on Medical Imaging with Deep Learning. PMLR, 2022: 820-833.

[4] Azad R, Arimond R, Aghdam E K, et al. DAE-Former: Dual attention-guided efficient transformer for medical image segmentation[J]. arXiv preprint arXiv:2212.13504, 2022.

[5] Isensee F, Ulrich C, Wald T, et al. Extending nnU-Net is all you need[J]. arXiv preprint arXiv:2208.10791, 2022.

[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.

[7] Isensee F, Jäger P F, Kohl S A A, et al. Automated design of deep learning methods for biomedical image segmentation[J]. arXiv preprint arXiv:1904.08128, 2019.