

融合注意力机制的U型对称网络的医疗辅助算法落地与运用

彭然, 王博, 邱子桐, 兰星辰, 彭跃, 陈先金, 高帅, 刘涛*

四川农业大学信息工程学院, 四川 雅安

收稿日期: 2024年2月20日; 录用日期: 2024年3月20日; 发布日期: 2024年3月27日

摘要

本文提出了一个融合注意力机制的U型对称网络的医疗辅助算法, 用于提高诊断的准确性和效率。利用深度学习技术, 对医学影像和病例数据进行自动化分析。本文使用了DRIVE、LiTS和SIIM-ACR等数据集进行训练, 在U-Net网络模型的基础上, 创新网络编码器和解码器之间引入了注意力模块, 以动态聚焦于图像中的关键特征。通过K折交叉验证来评估模型性能, 并使用L2权重衰减正则化防止模型过拟合。在系统落地方面, 设计了基于Django的医疗辅助小程序平台, 用户可以通过微信小程序进入用户界面, 实现医疗影像器官分割、在线问诊等功能。管理员则可以通过PC网页端对用户进行管理。该系统具有较高的准确性和效率, 可以有效改善医疗资源分配不均的问题。

关键词

人工智能, 医疗影像分割, U-Net网络, 注意力机制, 医疗辅助诊断

Landing and Use of Medical Assistance Algorithms for U-Shaped Symmetric Networks Incorporating Attention Mechanisms

Ran Peng, Bo Wang, Zitong Qiu, Xingchen Lan, Yue Peng, Xianjin Chen, Shuai Gao, Tao Liu*

College of Information Engineering, Sichuan Agricultural University, Ya'an Sichuan

Received: Feb. 20th, 2024; accepted: Mar. 20th, 2024; published: Mar. 27th, 2024

*通讯作者。

文章引用: 彭然, 王博, 邱子桐, 兰星辰, 彭跃, 陈先金, 高帅, 刘涛. 融合注意力机制的U型对称网络的医疗辅助算法落地与运用[J]. 计算机科学与应用, 2024, 14(3): 108-119. DOI: 10.12677/csa.2024.143062

Abstract

In this paper, we propose a medical assistance algorithm for U-shaped symmetric networks incorporating an attention mechanism for improving the accuracy and efficiency of diagnosis. Deep learning techniques are used to automate the analysis of medical images and case data. In this paper, datasets such as DRIVE, LiTS and SIIM-ACR are used for training, and an attention module is introduced between the innovative network encoder and decoder based on the U-Net network model to dynamically focus on key features in the image. The model performance is evaluated by K-fold cross-validation and L2 weight decay regularization is used to prevent model overfitting. In terms of system implementation, a Django-based medical assistance applet platform is designed, which allows users to access the user interface through a WeChat applet to realize functions such as medical image organ segmentation and online consultation. The administrator can manage the users through the PC web terminal. The system has high accuracy and efficiency, and can effectively improve the problem of uneven distribution of medical resources.

Keywords

Artificial Intelligence, Medical Image Segmentation, U-Net Network, Attention Mechanism, Medical Assisted Diagnosis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

医疗健康行业中人工智能技术的引入和发展应用十分重要，早期人工智能在医疗领域的探索出现在上世纪 70 年代，利兹大学开发的 AAPHelp 是记载人工智能系统在医疗领域最早的应用[1]。近年，随着机器学习与深度学习的发展，计算机视觉技术在医学图像分析中得到了广泛的应用。

伴随着现代人工智能相关技术的发展，基于大规模数据标注的深度学习方法在医学影像分析领域得到大规模的运用，深度学习技术已经成功地运用于临床医学的各个分支领域，例如检测胃癌[2]和乳腺癌[3]；肿块或正常乳腺组织分类[4]、肺结节分类[5]；肝脏[6]的分割等研究，且准确率达到 95% 以上，U 型对称网络(U-Net)凭借其强的非线性表达能力、利用来自较低分辨率图像的全局信息和来自较高分辨率图像的局部信息等优势在医学影像分割中被广泛运用。

尽管现代医疗系统在诊断和治疗疾病方面取得了显著进展，但仍面临诸多挑战。其中，诊断准确性、效率以及医疗资源分配不均[7]是尤为突出的问题。医生的主观判断、繁重的数据分析工作以及有限医疗资源的分配，都可能导致误诊、漏诊，甚至延误治疗。特别是在处理复杂疾病和分析大量医学影像数据时，这些问题更为显著。

为了解决这些问题，本研究提出了一种基于人工智能的自动识别病症系统。该系统利用深度学习技术，对医学影像和病例数据进行自动化分析，旨在提高诊断的准确性和效率。通过 skip connection 和注意力机制，我们的系统能够更好地融合上下文信息，捕捉疾病的细微特征。此外，系统的可扩展性和便携性使其能够在资源有限的地区提供高质量的医疗服务，从而改善医疗资源分配不均的问题。

2. 核心算法设计

2.1. 数据集来源

DRIVE 数据库的建立是为了能够对视网膜图像中血管的分割进行比较研究。视网膜血管分割和视网膜血管形态属性的描绘,如长度、宽度、曲折度、分支模式和角度,用于糖尿病、高血压、动脉硬化和脉络状新生血管等各种心血管和眼科疾病的诊断、筛查、治疗和评估,在 DRIVE 数据原有的 33 张图片的基础上,标注同类视网膜图像扩充数据样本共 200 张[8]。数据集展示如下图 1。

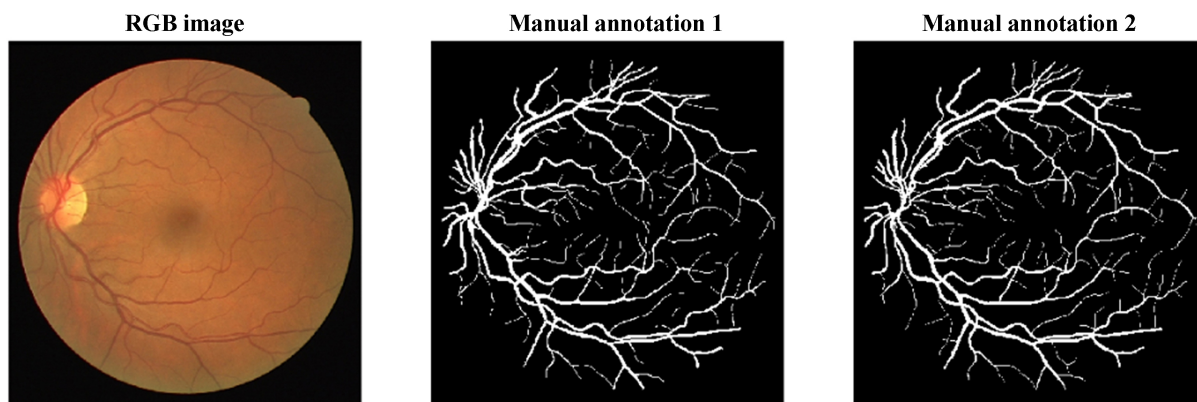


Figure 1. Presentation of the DRIVE dataset

图 1. DRIVE 数据集展示

LiTS 是 Liver Tumor Segmentation 数据集,是一个用于肝脏和肝癌分割的医学影像数据集。该数据集由纽约大学 Tandon 工程学院的研究人员开发,包含 131 个肝脏 CT 扫描图像和对应的分割掩模[9]。如下图 2 所示。

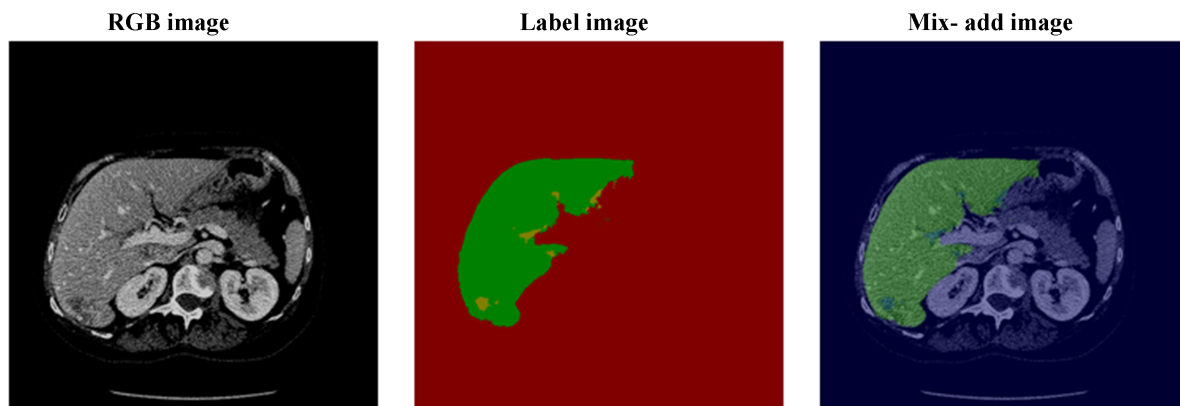


Figure 2. LiTS dataset presentation

图 2. LiTS 数据集展示

每个患者包含一个肿瘤,并且手动标注了盘状肝脏区域、肝动脉、门静脉、肝脏肿瘤等重要目标区域。其中,盘状肝脏区域是指肝脏整体的区域,肝动脉和门静脉是指这些血管的位置,肝脏肿瘤是指肝癌的位置。

胸部 X 光分割数据(肺部和心脏标签),来自 Kaggle 社区的竞赛数据由医学成像信息学会(SIIM)标注的 SIIM-ACR 气胸细分数据集,一共有 101 张胸片,101 个 nii 格式的标签数据,标签数据中 2 是肺部,

3 是心脏, 0 是背景[10]。展示数据如图 3 所示。

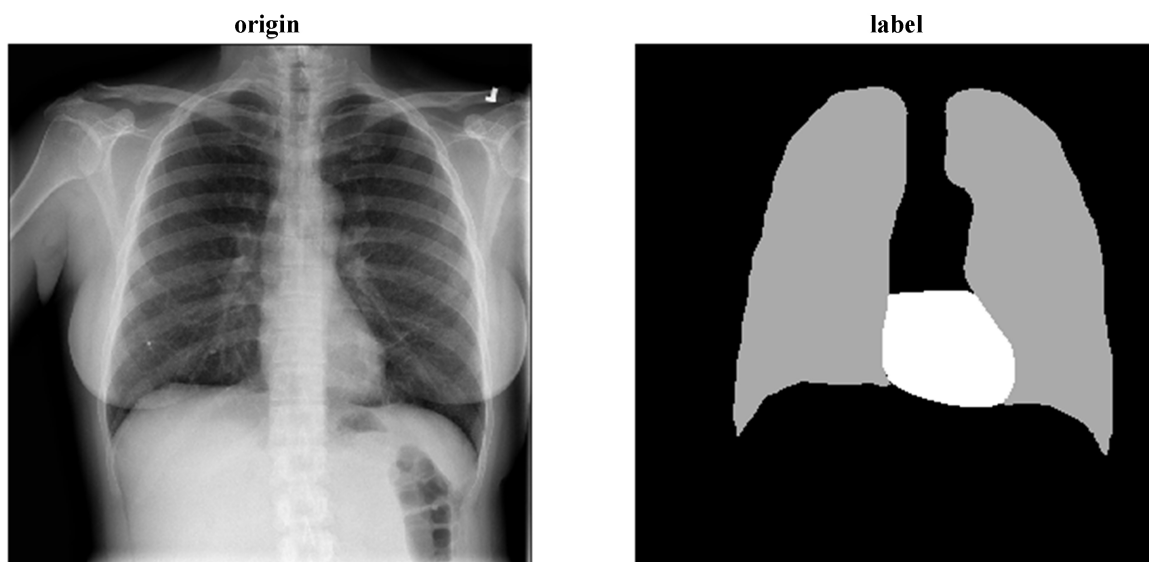


Figure 3. SIIM-ACR pneumothorax segmentation dataset
图 3. 胸部 X 光分割数据展示

2.2. 分割模型设计

U-Net [11]网络模型的特点是将图像输入和输出进行对称连接, 同时通过跳跃连接来引入低层次特征并进行细节保留。该模型被广泛用于医学图像分割, 如肺部 CT 图像、神经元、血管等的分割任务。

U-Net 网络模型主要组成部分包括编码器和解码器。编码器由一系列的卷积和池化层组成, 用于提取图像的高级特征。解码器则由一系列的反卷积和跳跃连接组成, 用于恢复图像的尺寸, 并通过跳跃连接引入低层次特征, 并进行细节保留, 网络模型如图 4 所示。

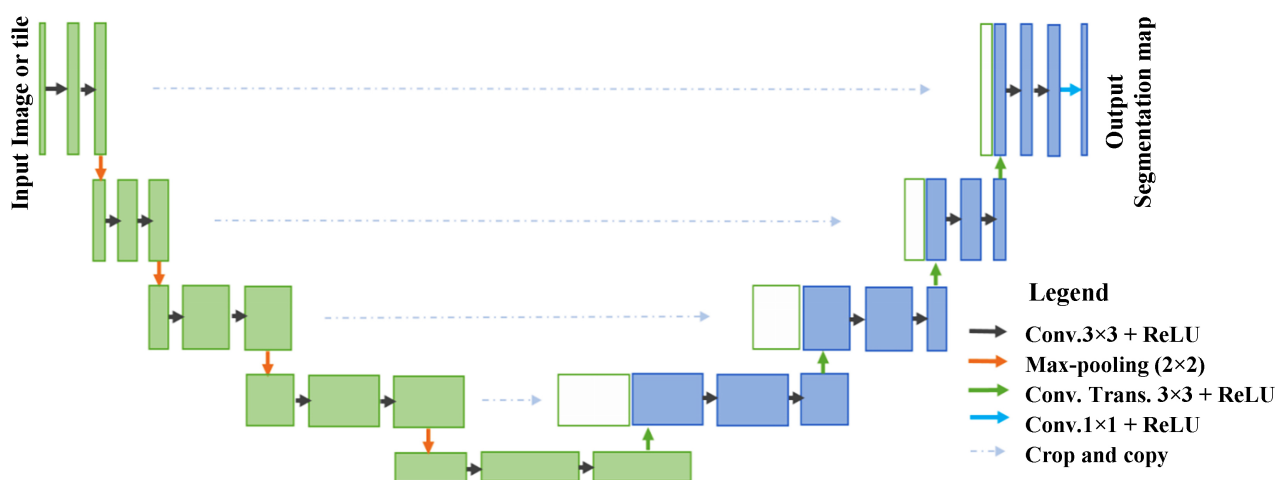


Figure 4. U-Net network model presentation
图 4. U-Net 网络模型展示

U-Net 网络模型的训练过程使用交叉熵损失函数[12], 并结合数据增强、Dropout [13]等技巧来增强模型的鲁棒性。在实际应用中, U-Net 网络模型已经被证明在医学图像分割领域具有优秀的性能, 特别是

对于小目标分割和细节保留方面优于其它模型。

为了进一步提升 Unet 网络在图像分割任务中的性能，我们对其原始结构进行了创新性设计。在 Unet 的基础上引入了一种新颖的注意力机制。该机制能够动态地聚焦于图像中的关键特征，从而提高网络对像素级细节的理解能力。具体操作上，我们在 Unet 的编码器和解码器之间加入了一个注意力模块，该模块实现不相关的区域被抑制，目标区域被关注。经过实验验证，我们提出的改进 Unet 结构在多个基准数据集上均取得了性能提升。改进后模型如图 5 所示。

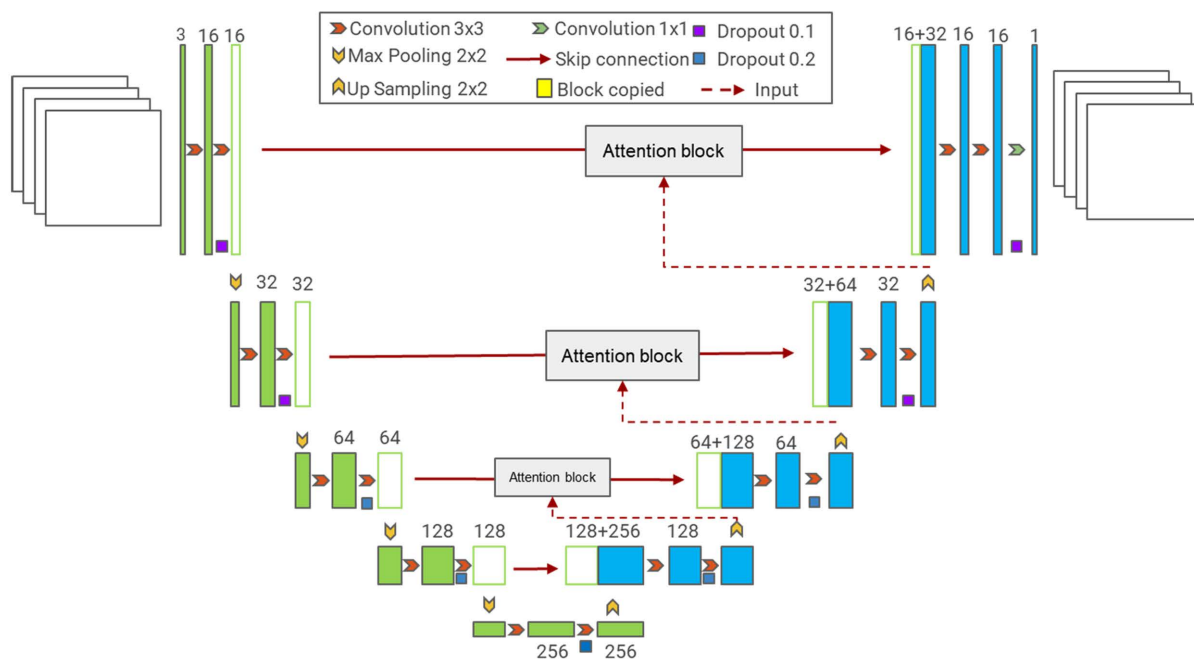


Figure 5. Improved U-net structure
图 5. 改进后的 U-net 结构

在我们提出的改进型 Unet 网络结构中，我们引入了一个新的注意力模块，该模块在传统的 skip connection 之前发挥作用。通过下采样操作获得的中间特征图与通过上采样操作恢复的相应特征图共同参与计算，以确定注意力权重。这些权重随后被应用于下采样特征图，以此强调对特定区域的关注，提升网络对细节的捕捉能力。处理完注意力权重后的特征图与未经过权重调整的特征图随后被拼接(concat)在一起，形成一个融合了上下文信息的特征序列，进一步提取特征并缩小空间维度。注意力机制模块结构图如下图 6 所示。

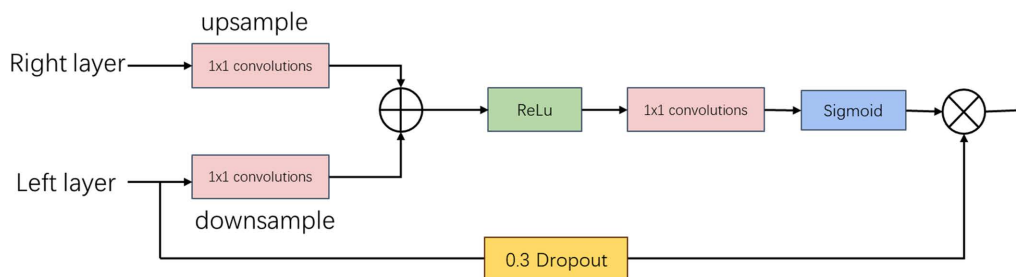


Figure 6. Attention mechanism module
图 6. 注意力机制模块

在评价指数上, 我们使用了 L2Decay 实现 L2 权重衰减正则化, 用于模型训练, 防止模型对训练数据过拟合。

$$\text{loss} = 0.5 * \text{coeff} * \text{reduce sum}(\text{square}(x))$$

在损失函数上, 使用了 DiceLoss [14]与 Focal Loss [15]共同作为损失函数。(X 表示真实分割图像的像素标签, Y 表示模型预测分割图像的像素类别, $|X \cap Y|$ 近似为预测图像的像素与真实标签图像的像素之间的点乘, 并将点乘结果相加, $|X|$ 和 $|Y|$ 分别近似为它们各自对应图像中的像素相加。)

$$\text{DiceLoss} = 1 - \text{Dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

DiceLoss 可以缓解样本中前景背景(面积)不平衡带来的消极影响。Dice Loss 训练更关注对前景区域的挖掘, 即保证有较低的 FN, 但会存在损失饱和问题, 为解决此问题我们同时使用 Focal Loss 解决计算机视觉任务中前景 - 背景不平衡的问题。

$$\text{loss}_{i,j} = \begin{cases} -\frac{1}{\text{fg_num}} * \alpha * (1 - \sigma(x_{i,j}))^\gamma * \log(\sigma(x_{i,j})) & (j+1) = \text{label}_{i,0} \\ -\frac{1}{\text{fg_num}} * (1 - \alpha) * \sigma(x_{i,j})^\gamma * \log(1 - \sigma(x_{i,j})) & (j+1) \neq \text{label}_{i,0} \end{cases}$$

2.3. 数据验证及指标评价

K 折交叉验证(K-fold cross-validation) [16]是一种评估模型性能的统计学方法, 通过将训练数据划分为多个子集(K 个), 并在每个子集上进行训练和测试, 来评估模型在不同数据集上的表现。

在数据验证时, 使用 K 折交叉验证用于模型调优, 找到使得模型泛化性能最优的超参值并在全部训练集上重新训练模型, 使用独立测试集对模型性能做出最终评价。评估模型在不同数据集上的泛化能力, 避免过拟合现象发生。原理如下图 7 所示。

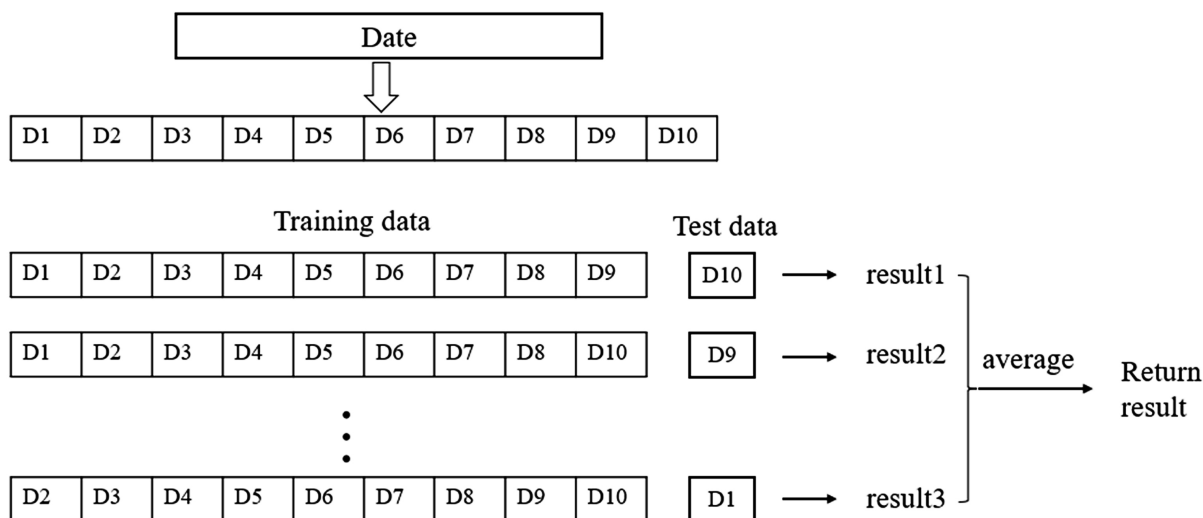


Figure 7. Schematic diagram of K-fold cross validation

图 7. K 折交叉验证示意图

我们在不同的数据集上验证模型的效果, 具体情况如下表 1 所示。

Table 1. Evaluation metrics of U-Net network model on different datasets

表 1. 在不同数据集上 U-Net 网络模型的评价指标

	mIoU	Acc	Kappa
	Unet/Our's	Unet/Our's	Unet/Our's
DRIVE 数据集	0.8540/0.8751	0.9763/0.9981	0.9178/0.9566
LiTS 数据集	0.8971/0.9223	0.9830/0.9964	0.9635/0.9789
SIIM-ACR 数据集	0.9319/0.9657	0.9585/0.9863	0.9600/0.9632

3. 系统架构设计

3.1. 总体功能模块设计

本平台使用者身份有用户和管理员。通过不同的使用平台来划分不同的用户功能，用户使用微信小程序进入用户界面，在小程序内实现医学影像器官分割、在线问诊等相关服务，管理者则通过 PC 网页端对用户进行管理，管理用户基本信息、分割图像情况、以及问诊情况等操作，系统功能模块图如图 8 所示。

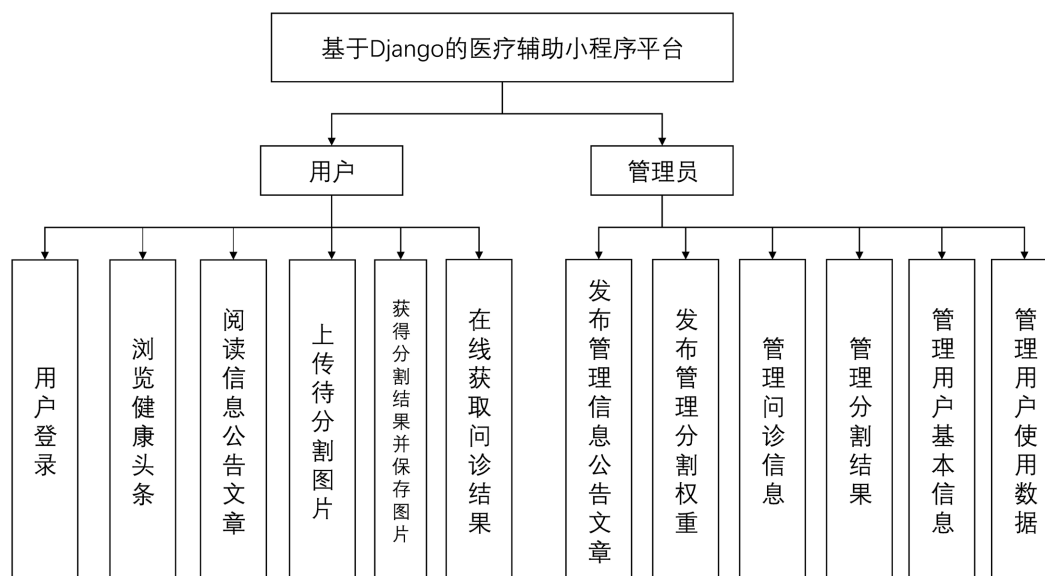


Figure 8. System functional module diagram

图 8. 系统功能模块图

用户功能详细阐述如表 2 所示。

Table 2. User function detail table

表 2. 用户功能详述表

具体功能	功能详情
医疗问答	支持用户对医疗问题提问，系统自动给出回答
新闻热点获取	展示网易新闻健康类新闻热点，点击跳转访问
图像分割	支持上传 CT 图像并且展示展示分割后的图像
结果保存	将分割后的图像保存到本地

续表

查看信息	进入主页查看管理员发布的信息，可以是文章也可以是公告
密码修改	对已有账号进行密码修改
信息留言	可在信息详情页留言

管理员功能详细阐述如表 3 所示。

Table 3. Detailed list of administrator functions

表 3. 管理员功能详述表

具体功能	功能详情
用户信息维护	对用户的基本信息进行管理，修改密码等
信息发布管理	进入后台发布页，对文章进行增删改查等操作
数据统计	统计后台用户的使用数据
分割情况管理	查看，修改，删除用户的分割数据
问诊信息管理	参看统计用户的问诊信息
权重管理	管理用户待使用的所需的权重文件

3.2. 页面结构设计

页面结构主要面向用户界面设计，在用户的视觉引导、使用观感中起重要的作用，是评判开发程序的重要指标之一。交互性较好、且页面美观、符合用户的使用逻辑的页面不仅可以提升用户体验，还可以增加用户黏度，防止用户流失。

本系统前端页面结构设计合计 10 个页面。其中一级界面 5 个，二级界面 5 个，下面对各个页面功能简单描述，如表 4、表 5 所示。

Table 4. Table of roles for level 1 pages

表 4. 一级页面作用表

页面名称	页面作用
首页、健康头条	展示健康头条的概览情况以及首页导航
登录	通过登录验证后进入系统
在线问诊	在线问诊的功能页，入口页面，在此编辑症状情况
个人中心	个人信息详情页面
消息发布	管理员发布的消息概览页

Table 5. Table of roles for secondary pages

表 5. 二级页面作用表

页面名称	页面作用
健康头条详情页	展示网易健康热头条的详情页
影像分割详情页	不同器官的分割详情页
影像分割结果页	展示分割结果，并允许在本页面下载分割图像
问诊回答页	显示诊断的详细结果
消息发布详情页	展示发布的消息的详情

3.3. 核心功能设计

1) 医疗影像分割功能。此功能须在用户登录后才能进行操作，在用户登录进入系统后，点击对应项目的入口模块进入分割页面，在微信小程序前端页面上上传图片，生成临时链接，并在页面展示临时上传的图像。用户点击提交后，以文件的形式上传到服务器指定路径，并调用预训练权重得到分割图片，透明化处理后再与原图叠加，保存图像路径到数据库，同时以 Json 的格式返回图像路径，在前端展示给用户，用户可以根据自己需要决定是否保存图片。医疗影像分割功能时序图如图 9 所示。

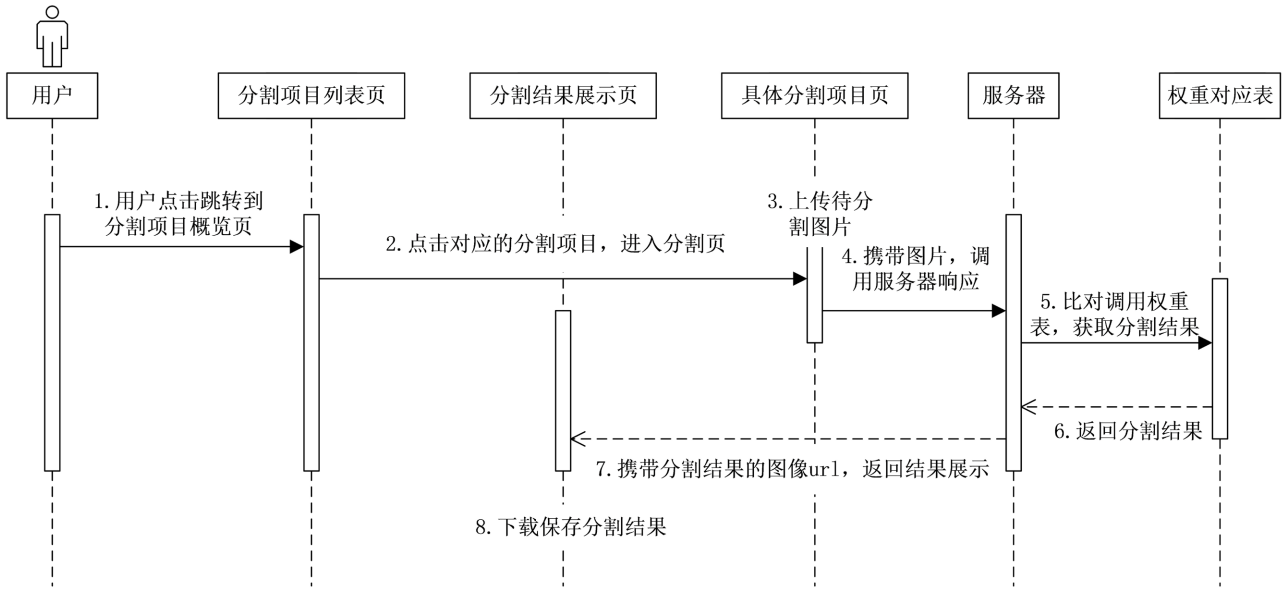


Figure 9. Timing diagram of medical image segmentation function
图 9. 医疗影像分割功能时序图

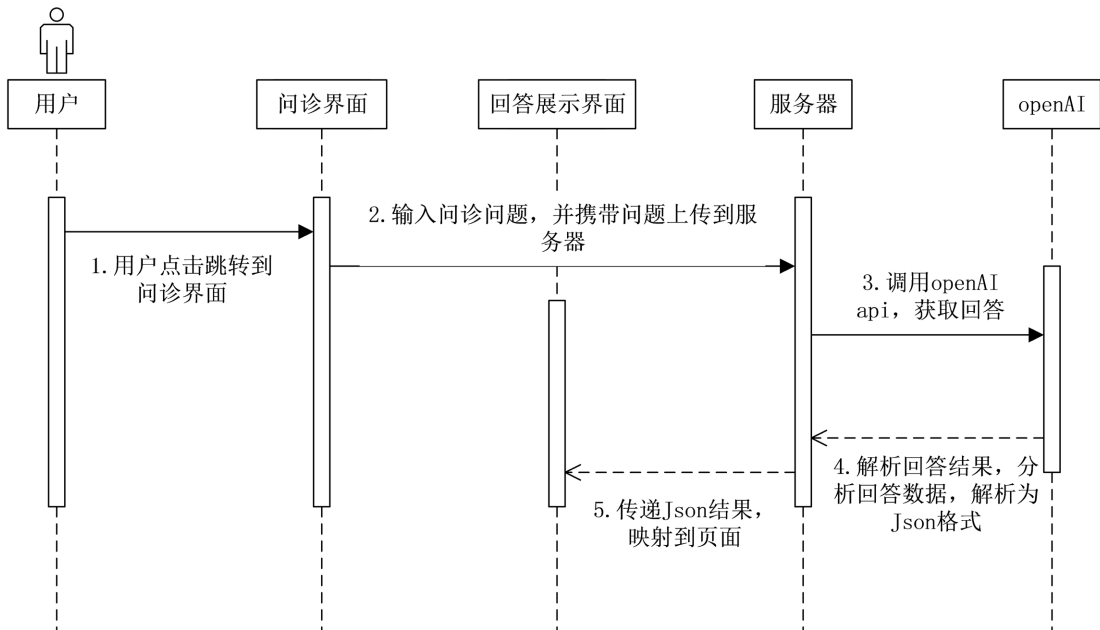


Figure 10. Timing diagram of the line inquiry function
图 10. 线问诊功能时序图

2) 在线问诊功能。在线问诊模块提供用户实时基本的疾病回答,用户在问诊界面输入问题症状或医疗建议,通过 Json 方式携带字符串向服务器发起请求,服务器解析携带信息,调用 openAI 获取问题回答,保存问题及回答并携带问题回答反馈给用户,在线问诊功能时序图如图 10 所示。

4. 核心功能展示

4.1. 医疗影像分割功能

医疗影像分割功能将医学影像中的结构、组织或病变区域分离出来,以便进行诊断、治疗和研究。在此方面功能模块中,用户可以上传医学 CT 图片,上传后服务器响应使用 U-Net 算法预训练权重通过对上传的图像进行分割,将分割后的结果图与原图叠加以文件路径的形式反馈给用户,并初步提供疑似病变区域以及初步的诊断结果。用户通过文件路径可以在线查看下载分割后的图片,如图 11 所示。

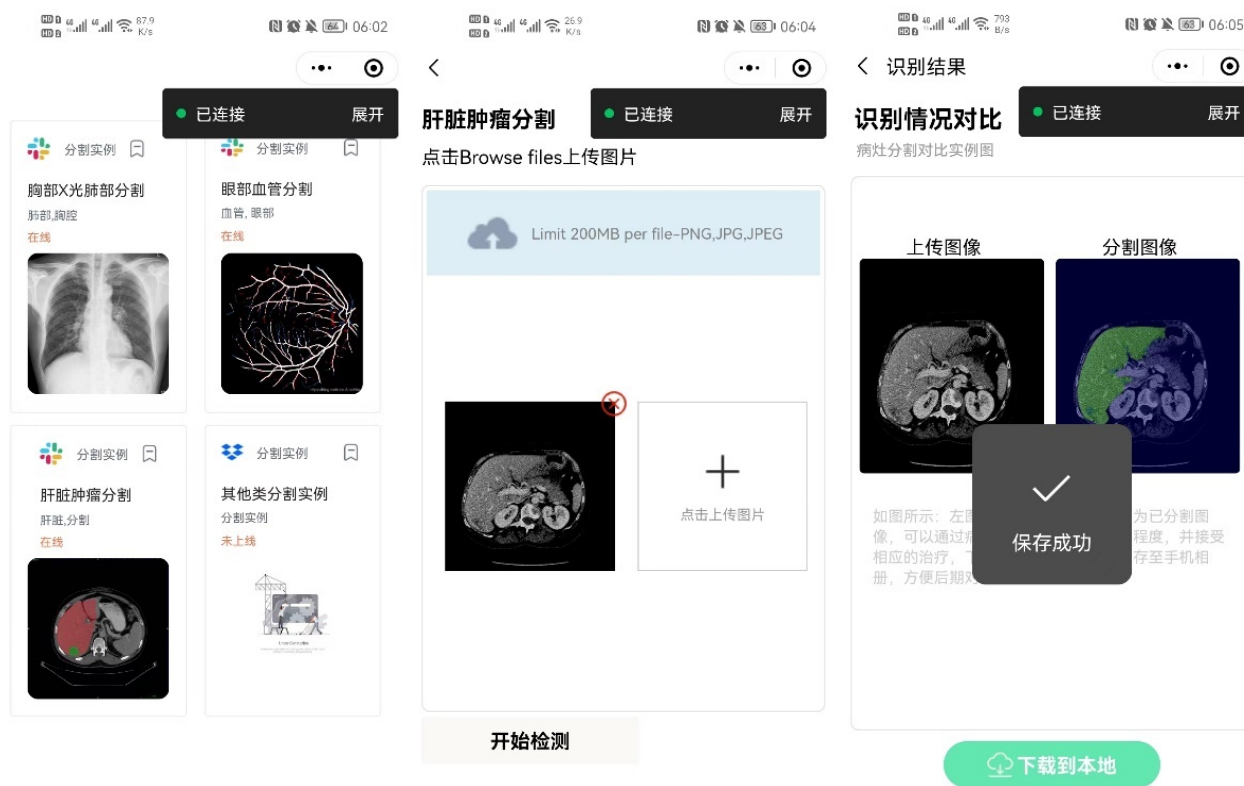


Figure 11. Effect of medical image segmentation interface

图 11. 医疗影像分割界面效果图

4.2. 在线问诊功能

在线问诊模块提供用户在线就医的平台,用户在问诊界面输入症状以及用药建议等相关问诊问题,以 Json 的格式上传到服务器,服务器解析 Json,使用 Davinci-003 模型,实现生成式文本问答,完成智能问诊,将参考结果返回给用户,如下图 12 所示。



Figure 12. Effect of online consultation function
图 12. 在线问诊功能效果图

5. 总论

本研究开发了一种基于人工智能的医疗辅助诊断系统，其核心目标在于提升医疗诊断过程的精确度和效率。该系统依托深度学习技术，融合了上下文信息处理和注意力机制，以捕捉疾病的微小细节。在数据集的选择上，我们使用了 DRIVE、LiTS 和 SIIM-ACR 等医学图像数据集来进行模型的训练和验证。在模型设计上，我们采用了 U-Net 网络结构，并通过引入注意力机制对其进行了优化，使其能够动态地关注图像中的关键区域。根据实验数据，改进后的模型在多个医学图像分割任务中表现出色。未来，我们计划对系统的鲁棒性进行进一步测试，并将其推广至乡镇医院和卫生所，以期减少由于医生主观判断、繁杂的数据分析工作以及医疗资源分配有限等因素所导致的误诊、漏诊和治疗延误等问题。

参考文献

- [1] 张戎, 刘洪臣. 人工智能技术在临床医疗中的应用概述[J]. 中华老年口腔医学杂志, 2021, 19(1): 40-44. <https://doi.org/10.19749/j.cn.cjgd.1672-2973.2021.01.010>
- [2] Hirasaw, T., Aoyama, K., Tanimoto, T., et al. (2018) Application of Artificial Intelligence Using a Convolutional Neural Network for Detecting Gastric Cancer in Endoscopic Images. *Gastric Cancer*, **21**, 653-660. <https://doi.org/10.1007/s10120-018-0793-2>
- [3] Hameed, Z., Zahia, S., Zafirain, B., et al. (2020) Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models. *Sensors*, **20**, Article 4373 <https://doi.org/10.3390/s20164373>
- [4] Rakhlin, A., Shvets, A., Iglovikov, V., et al. (2018) Deep Convolutional Neural Networks for Breast Cancer Histology

- Image Analysis. *Lecture Notes in Computer Science*, **10882**, 737-744. https://doi.org/10.1007/978-3-319-93000-8_83
- [5] Hua, K.L., Chen, Y.J., Hsu, C.H., *et al.* (2015) Computer-Aided Classification of Lung Nodules on Computed Tomography Images via Deep Learning Technique. *OncoTargets and Therapy*, **8**, 2015-2022. <https://doi.org/10.2147/OTT.S80733>
- [6] Che, H., Brown, L.G., Foran, D.J., Noshier, J.L., Hacihaliloglu, I. (2021) Liver Disease Classification from Ultrasound Using Multi-Scale CNN. *International Journal of Computer Assisted Radiology and Surgery*, **16**, 1537-1548. <https://doi.org/10.1007/s11548-021-02414-0>
- [7] 李芸. 智能医疗辅助诊断和问答系统关键技术研究[D]: [硕士学位论文]. 上海: 华东师范大学, 2020.
- [8] Wu, Y., Xia, Y., Song, Y., *et al.* (2018) Multiscale Network Followed Network Model for Retinal Vessel Segmentation. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*, Granada, 16-20 September 2018, 119-126. https://doi.org/10.1007/978-3-030-00934-2_14
- [9] Bilic, P., Christ, P.F., Vorontsov, E., *et al.* (2019) The Liver Tumor Segmentation Benchmark (LiTS). arXiv: 1901.04056.
- [10] Zawacki, A., Wu, C., Shih, G., Elliott, J., Fomitchev, M., Hussain, M., Lakhani, P., Culliton, P. and Bao, S.X. (2019) SIIM-ACR Pneumothorax Segmentation. <https://kaggle.com/competitions/siim-acr-pneumothorax-segmentation>
- [11] Kēniņš, R. (2020) Land Cover Classification Using Very High Spatial Resolution Remote Sensing Data and Deep Learning. *Latvian Journal of Physics and Technical Sciences*, **57**, 71-77. <https://doi.org/10.2478/lpts-2020-0009>
- [12] Wang, Y., Ma, X., Chen, Z., *et al.* (2019) Symmetric Cross Entropy for Robust Learning with Noisy Labels. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 322-330. <https://doi.org/10.1109/ICCV.2019.00041>
- [13] Srivastava, N., Hinton, G., Krizhevsky, A., *et al.* (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, **15**, 1929-1958.
- [14] Li, X.Y., *et al.* (2020) Dice Loss for Data-Imbalanced NLP Tasks. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5-10 July 2020, 465-476. <https://doi.org/10.18653/v1/2020.acl-main.45>
- [15] Lin, T.-Y., *et al.* (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [16] 胡局新, 张功杰. 基于 K 折交叉验证的选择性集成分类算法[J]. 科技通报, 2013, 29(12): 115-117. <https://doi.org/10.3969/j.issn.1001-7119.2013.12.039>