

# 一种基于改进损失函数的 UNet 网络医学图像分割方法研究——以 COVID-19 和 ACDC 数据集为例

## 摘要

医学图像分割是医学影像分析中的关键任务，对于疾病诊断和治疗具有重要意义。本文针对经典 UNet 网络模型，提出了一种改进的损失函数，以提高分割精度。在传统交叉熵损失函数的基础上，引入 Dice 损失函数，从而兼顾类别不平衡问题和细节分割精度。实验采用 COVID-19 和 ACDC 医学图像数据集，涵盖肺部感染和心脏解剖结构的分割任务。实验结果表明，改进后的损失函数有效提升了分割性能，尤其在边缘轮廓和小目标区域表现突出。本文的研究为医学图像分割提供了一种更为稳健的解决方案。

**关键词：**医学图像分割；UNet；DiceLoss

## 1 引言

### 1.1 选题背景

医学图像分割技术在现代医学中发挥着至关重要的作用。随着计算机视觉和深度学习的发展，基于卷积神经网络的分割方法取得了显著进展。其中，UNet 网络因其对小样本数据的适应性和对边缘信息的高效捕捉能力，广泛应用于医学图像分析任务。然而，医学图像中通常存在类别不平衡和模糊边界问题，导致传统交叉熵损失函数难以满足高精度分割需求。因此，设计更适合医学图像分割的损失函数具有重要研究价值。

### 1.2 选题依据

在医学图像分割任务中，传统交叉熵损失函数侧重于像素级分类，但忽略了目标区域的整体形态信息，尤其是对于边缘区域和小目标的分割效果较差。而 Dice 损失函数能够直接优化分割区域的重叠程度，弥补交叉熵的不足。因此，将 Dice 损失与交叉熵损失相结合，有望进一步提高模型对复杂医学图像的分割性能。

### 1.3 选题意义

本文旨在通过改进损失函数优化 UNet 网络的分割性能，以应对医学图像数据集中的挑战。COVID-19 数据集涉及肺部感染区域的检测，而 ACDC 数据集涵盖心脏解剖结构的分割任务。这两类医学应用均要求模型具备高鲁棒性和高精度分割能力。本文的研究不仅有助于提升医学图像分割技术水平，也为医学诊断和治疗提供更为可靠的数据支持。

## 2 相关工作

### 2.1 UNet 网络结构

UNet [2] 最早由 Ronneberger 等人提出，是一种专门针对医学图像分割设计的卷积神经网络。如图1所示，UNet 采用编码-解码结构，包括下采样和上采样两个阶段。编码阶段通过卷积和池化层提取特征，并逐步降低分辨率，而解码阶段通过反卷积逐步恢复分辨率，并利用跳跃连接将编码阶段的特征直接传递给解码阶段，从而保留高分辨率信息。这种设计特别适合小样本训练，同时保持了较高的分割精度。

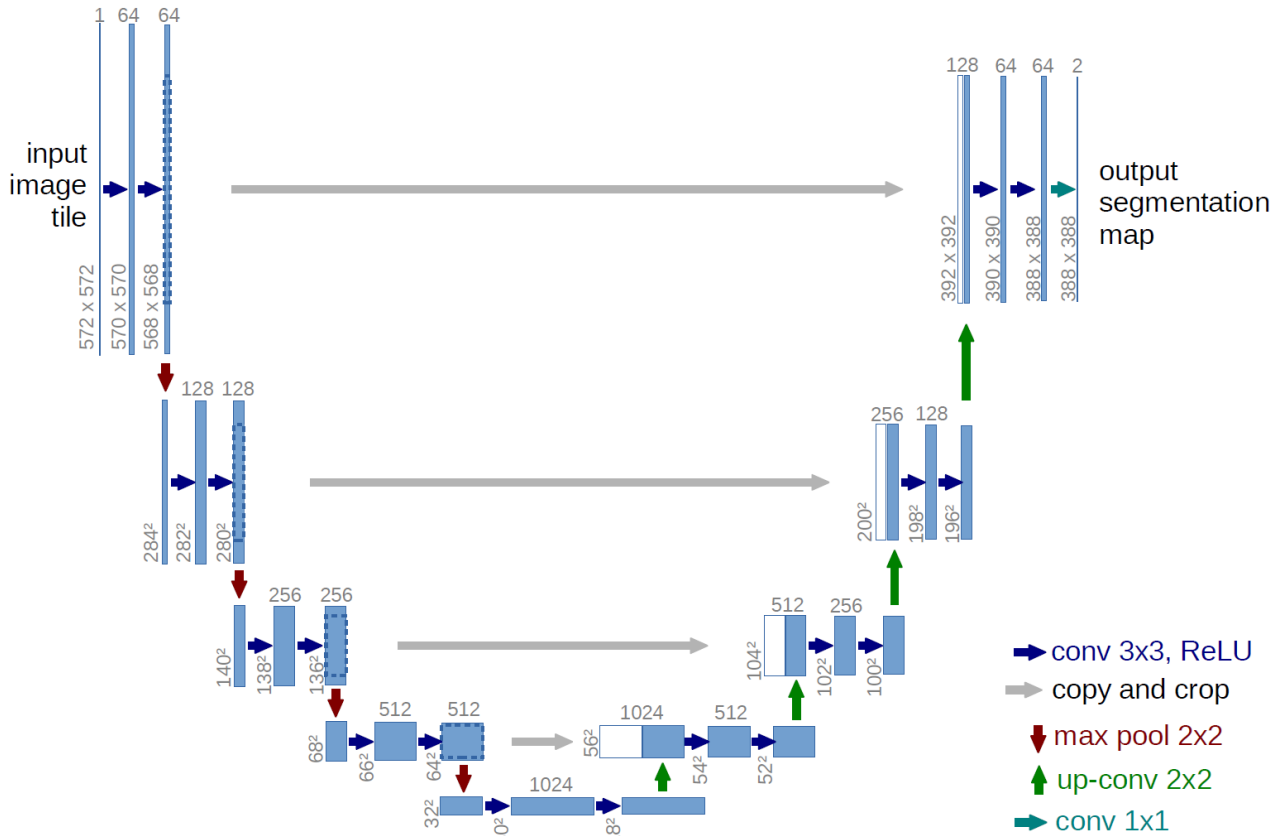


图 1. U-net 架构（最低分辨率下 32x32 像素的示例）。每个蓝色框对应一个多通道特征图。通道数量在框的顶部表示。x-y 尺寸在框的左下角提供。白色框表示复制的特征图。箭头表示不同的操作。

UNet 与全卷积网络（FCN）关系密切。FCN [1] 是最早应用于图像语义分割的深度学习网络之一，采用卷积层替代全连接层，实现端到端的密集预测。UNet 在 FCN 的基础上进行了改进，增加了对称的解码路径和跳跃连接，显著增强了特征融合能力，特别适用于需要高精度边界恢复的医学图像分割任务。

U-Net 网络架构作为一种经典分割模型为后续图像分割技术的出现奠定了基础，如 3D-Unet、Attention U-Net、U-Net++ 等模型都是以 U-Net 作为基础架构。基于 U-net 的架构在医学图像分析中具有相当潜力和价值。

## 2.2 数据集

### 2.2.1 COVID-19

COVID-19 数据集包含来自新冠肺炎患者的 CT 扫描图像，主要用于检测和分割肺部感染区域。该数据集的挑战在于感染区域的形态多变、边界模糊以及与正常组织的对比度较低。模型需要在感染区域与正常组织之间进行精确区分，从而辅助临床诊断和治疗。

### 2.2.2 ACDC

ACDC 数据集是心脏磁共振 (MRI) 影像的标准数据集，包含左心室、右心室和心肌的分割任务。该数据集用于评估心脏解剖结构分析中的分割性能，特别适合心脏疾病的诊断和随访。由于心脏结构复杂且解剖部位边界相邻密切，该数据集对模型的边界识别能力提出了较高要求。

## 2.3 评价指标

本文中主要使用以下三个评价指标：Dice 系数、交并比 (IoU) 和像素错误率 (Pixel error)。

### 2.3.1 Dice 系数

Dice 系数是一种衡量两个样本集合相似度的指标，广泛应用于医学图像分割任务中，取值范围  $[0, 1]$  (0 表示完全不重叠即预测失败，1 表示完全重叠即完美预测)。它强调预测结果与真实标签的重叠部分，用于评估分割模型的性能。计算公式如下：

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

其中， $A$  和  $B$  分别表示预测结果和真实标签的像素集合。Dice 系数对小目标更加敏感，适合医学图像分割等领域。强调重叠区域，能够很好地衡量预测与真实值之间的相似度。

### 2.3.2 交并比 (IoU)

交并比是评价分割精度的经典指标，表示预测结果和真实标签重叠区域占二者并集的比例，取值范围  $[0, 1]$  (0 表示完全不重叠，1 表示完全重叠)。计算公式如下：

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

其中， $A$  和  $B$  分别表示预测结果和真实标签的像素集合。交并比对于预测与真实标签部分重叠但未完全重合的情况比较敏感。

### 2.3.3 Pixel Error

像素错误率衡量的是预测结果与真实标签之间分类错误的像素比例，取值范围  $[0, 1]$  (0 表示完全正确的预测，1 表示完全错误的预测)。它反映模型在像素级别的分类准确性。计算公式如下：

$$Pixel Error = \frac{FP + FN}{Total Pixels} \quad (3)$$

其中,  $FP$  为假正例数量,  $FN$  为假负例数量,  $Total Pixels$  为图像中的总像素数量。像素错误率注重全局像素错误率, 更适合整体评估而非细节匹配。对少量错误预测较不敏感, 不适合边界区域或小目标的精细评估。

### 3 本文方法

#### 3.1 本文方法概述

本文基于 UNet 网络结构, 结合交叉熵损失函数与 Dice 损失函数, 以解决医学图像分割任务中的类别不平衡问题, 并提升边缘轮廓与小目标区域的分割精度。模型采用 PyTorch 框架实现, 充分利用其灵活性与强大的计算能力。

#### 3.2 网络结构

本文沿用了 UNet 的编码-解码结构, 输入图像通过一系列卷积和池化操作提取特征, 并在解码阶段进行上采样恢复分辨率。同时, 通过跳跃连接保留高分辨率特征, 有助于边界信息的恢复。

其中, DoubleConv 模块由两个  $3 \times 3$  卷积 (无填充卷积) 的重复应用组成, 每个卷积后接一个修正线性单元 (ReLU) 和一个步长为 2 的  $2 \times 2$  最大池化操作以进行下采样。上采样通过反卷积操作实现。反卷积, 顾名思义, 就是反着卷积。卷积是让 feature map 越来越小, 反卷积就是让 feature map 越来越大。

#### 3.3 损失函数定义

##### 3.3.1 交叉熵损失

交叉熵是信息论中的一个概念, 用于量化两个概率分布之间的差异。在机器学习中, 交叉熵损失函数用于评估模型输出的概率分布与真实标签分布之间的差距。其计算公式如下:

一般形式, 对于两个概率分布  $P$  和  $Q$ , 交叉熵的计算公式为:

$$H(P, Q) = - \sum_x P(x) \log(Q(x)) \quad (4)$$

其中,  $P(x)$  是真实分布的概率,  $Q(x)$  是模型预测的概率。

在二分类问题 (如 COVID19 数据集) 中, 交叉熵损失函数可以简化为:

$$L = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})] \quad (5)$$

其中,  $y$  是真实标签 (0 或 1),  $\hat{y}$  是模型预测的概率值。

对于多分类问题 (如 ACDC 数据集), 假设有  $C$  个类别, 交叉熵损失函数为:

$$L = - \sum_{i=1}^C p_i \log(q_i) \quad (6)$$

其中,  $p_i$  是真实标签的概率 (通常为独热编码形式),  $q_i$  是模型预测每个类别的概率。

### 3.3.2 Dice 损失

Dice 损失函数是一种广泛应用于图像分割任务的损失函数, 特别是在医学图像分割中表现突出。它的核心思想是通过计算预测结果与真实标签之间的重叠部分来优化模型性能。其计算公式如下:

$$\text{DiceLoss} = 1 - \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2} \quad (7)$$

其中,  $N$  表示样本中的总像素数量,  $p_i$  表示模型预测的第  $i$  个像素的值, 通常为  $[0, 1]$  之间的概率值。  $g_i$  表示真实标签 (ground truth) 的第  $i$  个像素的值, 取 0 或 1 (0 表示背景, 1 表示目标)。

Dice Loss 衡量的是模型预测区域与真实标签区域, 之间的相似度, 取值范围在  $[0, 1]$  之间。值越小表示模型的预测与真实标签越接近, 损失越小, 分割效果越好。

### 3.3.3 联合损失函数

本文采用交叉熵损失 (CE) 与 Dice 损失 (DL) 的加权组合, 以充分利用二者的优点, 既保证整体分类精度, 又提高边缘轮廓和小目标区域的分割性能。具体计算公式如下:

$$\text{Loss} = \text{CE} + \alpha \cdot \text{DL} \quad (8)$$

其中,  $\alpha$  是权重超参数用于调节 Dice 损失的权重。

使用交叉熵损失和 Dice 损失的加权组合兼顾像素级和区域级的分割性能, 提高预测结果的整体准确性和边界细节质量。并且适合类别不平衡问题和目标区域较小的医学图像分割任务, 例如心脏分割、肿瘤检测等。对不同目标形状和大小的适应性更强, 更加鲁棒。

## 4 复现细节

### 4.1 与已有开源代码对比

本文没有参考任何相关源代码。本文使用 PyTorch 深度学习框架实现了 UNet 的模型框架, 使用交叉熵和 Dice 损失结合的损失函数分别在 COVID19 和 ACDC 数据集上训练和测试, 并使用评价指标对模型的训练效果进行评估。训练过程中通过调节权重超参数  $\alpha$  来调整损失函数 ( $\alpha$  为 0 时表示仅使用交叉熵损失)。

### 4.2 实验

实验使用 COVID19 和 ACDC 数据集并将输入图片尺寸统一调整为 256x256 的单通道灰度图像, 损失函数的权重超参数  $\alpha$  分别设置为 0 和 1 进行对比实验。 $\alpha=0$  表示仅使用论文中的交叉熵损失函数,  $\alpha=1$  表示我们的联合损失函数并且 Dice 损失和交叉熵损失同等重要。实验训练 100 轮, 学习率设置为 0.01, 使用 Adam 优化算法, 结果如下:

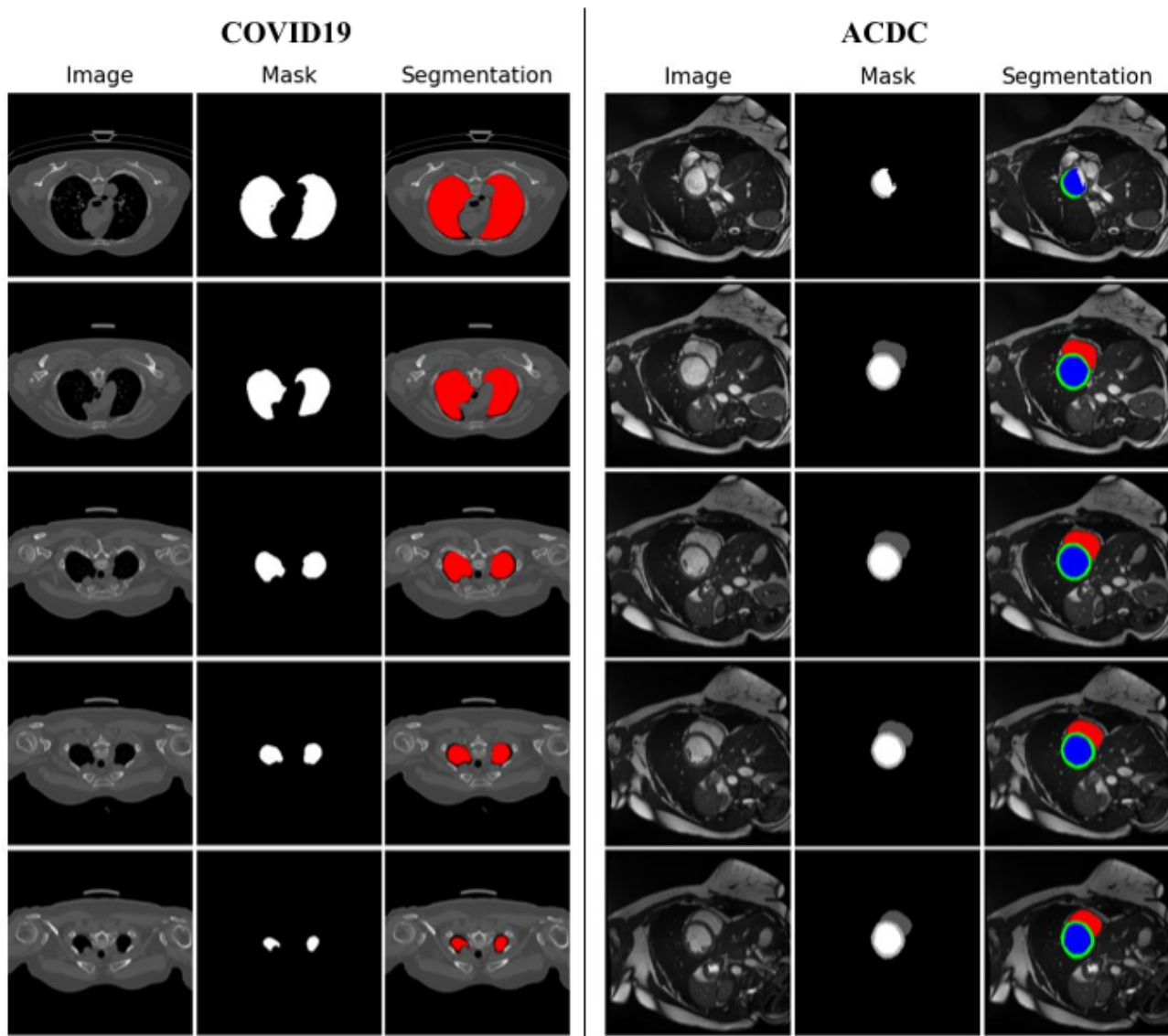


图 2.  $\alpha=0$ , 取验证集中 5 个样本的预测结果

表 1. 使用 CE 和 CE + DL 损失函数对 COVID19 和 ACDC 数据集的性能指标进行比较

Loss	COVID19			ACDC		
	Dice	IoU	Pixel Error	Dice	IoU	Pixel Error
CE	0.9497	0.9111	0.0176	0.9090	0.8420	0.0069
<b>CE + DL</b>	<b>0.9527</b> $\uparrow$	<b>0.9169</b> $\uparrow$	<b>0.0163</b> $\downarrow$	<b>0.9213</b> $\uparrow$	<b>0.8608</b> $\uparrow$	<b>0.0060</b> $\downarrow$

### 4.3 创新点

- 提出将交叉熵损失与 Dice 损失相结合的损失函数，以解决类别不平衡和边界模糊问题。
- 在 COVID-19 和 ACDC 两个医学图像数据集上进行验证，证明改进损失函数的通用性和鲁棒性。
- 基于 PyTorch 框架实现改进的 UNet 模型，提高实验复现性与灵活性。

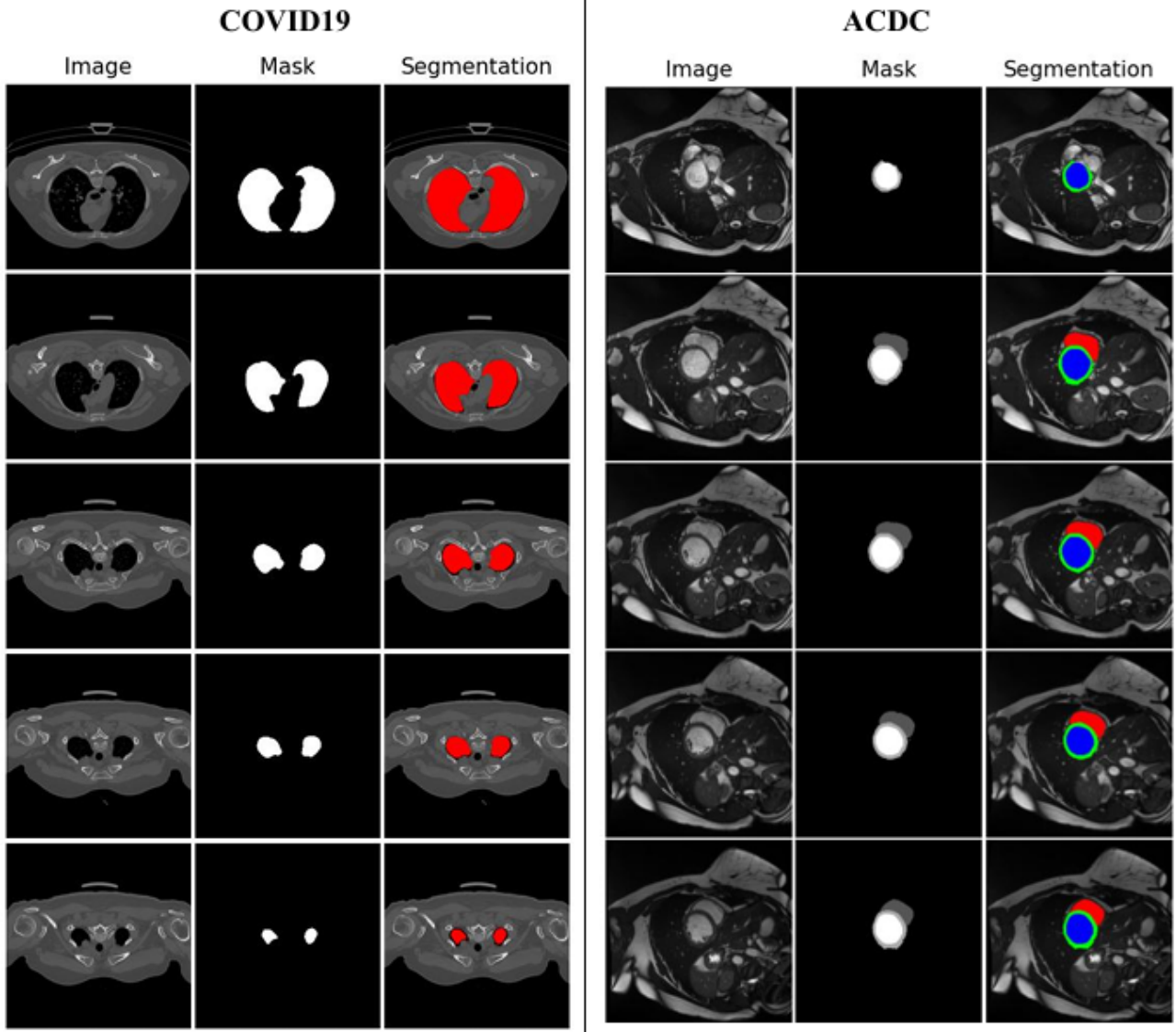


图 3.  $\alpha=1$ ，取验证集中 5 个样本的预测结果

## 5 实验结果分析

实验结果如表1所示，结果表明引入 Dice 损失的改进模型在 COVID-19 和 ACDC 数据集上的分割性能显著优于传统交叉熵损失模型。具体分析如下：

图4和图5表明，在 Dice 系数和 IoU 指标上，改进后的模型均有提升。从图2和图3可以看出，在像素错误率上，改进模型减少了误分率，尤其在边缘和小目标区域表现更优。可视化结果显示，改进模型更准确地捕捉边界轮廓和细节信息，降低了过分割和欠分割问题。

## 6 总结与展望

本文提出了一种基于改进损失函数的 UNet 网络医学图像分割方法，通过将交叉熵损失函数与 Dice 损失函数相结合，提升了模型在类别不平衡和边缘模糊问题上的分割性能。实验结果验证了本文方法在 COVID-19 和 ACDC 数据集上的有效性和鲁棒性，为医学图像分割任务提供了更优的解决方案。

然而，本文仍存在一些不足之处：首先，模型对复杂形态和重叠区域的识别精度仍有提



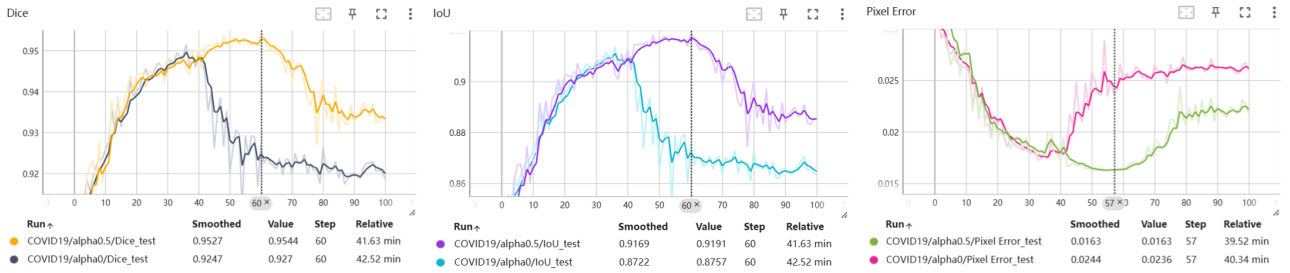


图 4. 对比  $\alpha=0$  和  $\alpha=1$  训练 COVID19 数据集时评价指标的变化情况

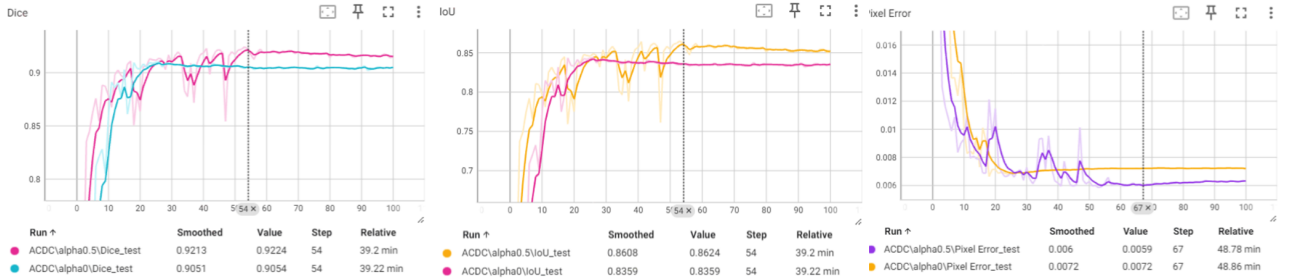


图 5. 对比  $\alpha=0$  和  $\alpha=1$  训练 ACDC 数据集时评价指标的变化情况

升空间；其次，训练过程的计算复杂度较高，影响了应用的实时性；而且训练是基于有监督的，在医学图像数据集匮乏的背景下有监督学习并不是一个好的解决方案。未来研究将侧重于以下几个方面：

- 引入注意力机制或 Transformer 架构进一步提升模型的特征提取能力。
- 探索轻量化模型结构，以满足实时分割需求。
- 在更多医学图像数据集上进行验证，提升模型的泛化能力。
- 探索基于高斯混合模型的无监督分割方法。

## 参考文献

- [1] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.