

# 基于条件扩散模型的结直肠图像染色归一化<sup>\*</sup>

李子成<sup>1</sup> 贾伟<sup>1,2</sup> 赵雪芬<sup>1</sup> 高宏娟<sup>1,2</sup>

(1. 宁夏大学信息工程学院 银川 750021; 2. 宁夏“东数西算”人工智能与信息安全重点实验室 银川 750021)

**摘要:** 由于现有染色归一化方法无法准确提取结直肠病理图像的复杂结构特征,导致丢失部分结直肠病理图像的结构信息,无法生成高质量的染色归一化结直肠病理图像。为解决该问题,提出一种基于条件扩散模型的结直肠病理图像染色归一化方法,该方法包括条件扩散模型和图像特征重建。在条件扩散模型中,首先,使用马尔科夫链前向过程对结直肠病理原图像进行加噪声。然后,将噪声图像和条件图像输入到增强去噪网络中进行去噪,在这过程中利用增强激活模块,学习结直肠病理图像的深层特征,捕获更多的图像上下文信息。在编码器和解码器之间引入跳跃连接空间注意力模块,准确提取结直肠病理图像的位置空间信息。在图像特征重建中,设计金字塔特征提取模块,提取多尺度条件图像与生成图像的特征,并构建重建损失函数,优化整个网络的性能。实验结果表明,与现有方法相比,所提染色归一化方法在公共数据集 GlaS 和 CRAG 上的能生成质量更高的染色归一化结直肠病理图像。

**关键词:** 条件扩散模型;结直肠病理图像;染色归一化;去噪网络

中图分类号: TP391; TN29 文献标识码: A 国家标准学科分类代码: 520.6

## Diffusion model-based staining normalization for colorectal image

Li Zicheng<sup>1</sup> Jia Wei<sup>1,2</sup> Zhao Xuefen<sup>1</sup> Gao Hongjuan<sup>1,2</sup>

(1. School of Information Engineering, Ningxia University, Yinchuan 750021, China;

2. Ningxia Key Laboratory of Artificial Intelligence and Information Security for Channeling Computing  
Resources from the East to the West, Yinchuan 750021, China)

**Abstract:** Existing staining normalization methods are unable to accurately extract the complex structural features of colorectal pathological images, resulting in the loss of partial structural information and the inability to generate high-quality staining-normalized colorectal pathological images. To address this issue, a staining normalization method for colorectal pathological images based on a conditional diffusion model is proposed. The proposed method includes conditional diffusion model and image feature reconstruction. In conditional diffusion model, firstly, the Markov chain forward process is employed to add noise to the original colorectal pathological images. Then, the noisy images and conditional images are input into an enhanced denoising network for denoising. During this process, an enhanced activation module is utilized to learn the deep features of the colorectal pathological images and capture more contextual information. A skip-connection spatial attention module is introduced between the encoder and decoder to accurately extract the positional spatial information of the colorectal pathological images. Finally, a pyramid feature extraction module is designed to extract the features of the multi-scale conditional images and generated images, and a reconstruction loss function is constructed to optimize the performance of the entire network. Experimental results demonstrate that compared with existing methods, the proposed staining normalization method can generate higher-quality staining-normalized colorectal pathological images on public datasets GlaS and CRAG.

**Keywords:** conditional diffusion model; colorectal pathological images; staining normalization; denoising network

## 0 引言

癌症研究机构估计,2020年全球结直肠癌新发病例约为193万例,并且结直肠癌位居癌症发病排行榜第三位<sup>[1]</sup>。预计在2022年,中国结直肠癌新发病例将达到59万例,超

过世界其他国家<sup>[2-3]</sup>。对结直肠癌病理图像的准确分析可以用于判断病因,排除或确认恶性肿瘤,从而制定合理的诊疗方案,提高结直肠癌患者生存率<sup>[4-5]</sup>。但是在结直肠病理图像的制作过程中,由于受到不同制备载玻片的设备和制作技术的影响,导致从不同设备和不同制作技术获得结直肠病理图像之间的颜色存在差异,影响了后续对结直肠病理图像的结构和纹理等信息的分析,尤其是难以对结直肠病理图像中腺体形态进行准确分析。

为解决不同病理图像之间的颜色存在差异的问题,一些学者对病理图像的染色归一化进行研究,尝试通过染色归一化,消除染色差异对病理图像特征的影响,使得病理图像中的结构、纹理和腺体形态等信息中更加清晰、易于识别。传统的病理图像染色归一化方法包括基于统一转换的方法<sup>[6]</sup>和基于非负矩阵分解的方法<sup>[7]</sup>。随着人工智能的发展,深度学习已成功应用在医学图像分割<sup>[8-10]</sup>、医学图像生成<sup>[11-13]</sup>等任务中。一些学者提出了基于深度学习的病理图像的染色归一化方法,例如,由于生成对抗网络(generative adversarial networks, GAN)具有良好的图像生成能力,Cong 等<sup>[14]</sup>提出一种基于自适应 GAN 的病理图像颜色归一化方法(pathological image color normalization method based on adaptive GAN, CAGN),利用双解码器的相互学习,对颜色进行重新染色。Vahadane 等<sup>[15]</sup>提出一种使用 Cycle-GAN 保留结构信息的染色归一化框架(stain normalization framework using Cycle-GAN to preserve structural information, SPCN),该方法将 Cycle-GAN 作为主干网,在预训练的语义网络和染色归一化网络之间集成不同网络层中的语义信息。Salehi 等<sup>[16]</sup>提出一种基于像素到像素的染色归一化方法(pix2pix-based stain-to-stain translation method, STST),该方法在成对的组织病理学图像及其对应的灰度图像上进行训练,以实现染色转移。Zhao 等<sup>[17]</sup>提出一种基于自监督机制的网络 RestainNet,利用单个图像的内部信息来学习染色标准化的映射,通过模拟染色过程的物理特性,实现了对图像的重新染色。Barua 等<sup>[18]</sup>提出一种采用多域颜色转换技术的方法(multi-domain stain color translation, MDST),该方法在条件生成对抗网络中引入注意力机制,能够更加聚焦于图像的重要区域,并通过门控机制控制信息流,从而提高染色颜色转换的准确性。这些基于深度学习的病理图像的染色归一化方法可以在染色特异性、结果保存和颜色一致性之间提供良好的平衡。

虽然现有方法在病理图像的染色归一化研究中取得了一定成果,但是,由于结直肠病理图像的组织结构复杂,且腺体形状具有多样性特点,导致现有病理图像的染色归一化方法在提取结直肠病理图像特征时,丢失部分结直肠病理图像的结构特征信息,并在结直肠病理图像生成过程中产生少量噪点,导致生成的染色归一化图像质量欠佳。

条件扩散模型是一种能够产生新样本的生成模型,其

核心机制是通过掌握数据的条件概率分布来实现这一过程,由于该模型具有保留数据语义结构的能力,使得其在图像分析中取得了较好的应用效果,尤其是在生成高质量图像方面,条件扩散模型已经超越了 GAN。例如,在生成数字图像方面,Wang 等<sup>[19]</sup>提出一种将条件扩散模型和多尺度空间相结合的散斑图像生成方法,利用条件扩散模型解决了难以生成高度复杂的背景纹理和随机的散斑图像问题。在利用文本生成图像方法,Zhu 等<sup>[20]</sup>提出一种基于条件扩散模型的文本图像生成方法,通过手写脚本和场景文本生成图像,利用条件扩散模型提升生成照片级真实和多样化图像样本的能力。在医学图像的分割和分类方面,利用条件扩散模型生成高质量图像可以进一步提高分割和分类的性能。Chen 等<sup>[21]</sup>提出一种基于条件伯努利扩散模型的用于医学图像分割方法,该方法将伯努利噪声作为扩散核,以增强扩散模型对二元分割任务的能力。为提升条件扩散模型的区域注意力,Wu 等<sup>[22]</sup>提出一种采用动态条件编码和特征频率解析器的条件扩散模型,用于眼底图像、脑肿瘤图像和甲状腺结节的分割。Yang 等<sup>[23]</sup>提出一种基于双条件引导策略的条件扩散模型,利用条件扩散模型消除图像中的噪声和干扰因素,对超声图像、皮肤镜图像和眼底图像进行分类。

虽然条件扩散模型已经应用到了图像分析中,并利用其良好的图像生成能力,在不同的应用中体现出了优于 GAN 等技术的性能,如果能将条件扩散模型有效应用到结直肠病理图像的染色归一化研究中,将有利于进一步提升染色归一化性能。但是条件扩散模型尚未被应用在结直肠病理图像的染色归一化研究中,由于结直肠病理图像具有复杂的纹理和不规则的腺体结构,现有基于条件扩散模型的方法无法直接应用到结直肠病理图像的染色归一化研究中,如何将其有效应用到该领域是一项具有挑战性的任务。受到上述研究的启发,为进一步提升染色归一化方法的性能,本文首次将条件扩散模型应用到结直肠病理图像的颜色归一化方法中,提出一种基于条件扩散模型的结直肠病理图像染色归一化方法(diffusion model-based staining normalization method for colorectal pathological image, DMSC),将目标染色风格图像作为 DMSC 方法输入的条件图像,并且与加噪后的对应原图像作为共同输入,进行去噪网络训练,产生颜色亮度对比较为明显,图像质量较高的结直肠病理图像。

本文主要贡献包括以下 3 点:

1) 提出一种基于结直肠病理图像染色归一化方法 DMSC,该方法包括条件扩散模型和图像特征重建两部分。为提升条件扩散模型的性能,重新设计条件扩散模型的去噪网络。在图像特征重建中,重新设计了特征金子塔网络(feature pyramid network, FPN),从而提升特征提取能力,进而能够生成结构信息完整、颜色对比明显的结直肠病理图像。

2) 现有基于卷积神经网络的条件扩散模型

(convolutional neural network-based conditional diffusion model, CNCDM)<sup>[24]</sup> 的去噪网络以卷积神经网络 (convolutional neural network, CNN) 与自注意力的结合的形式实现去噪, 缺少全局颜色信息的捕获能力, 不易学习到结直肠病理图像的整体颜色信息。为解决该问题, 在改进的条件扩散模型 (improved conditional diffusion model, ICDM) 中, 设计了增强去噪网络 (enhanced denoising network, EDNet), 在去噪过程中, 受到 Peng 等<sup>[25]</sup> 提出的 Transformer 块的启发, 为捕获结直肠病理图像的深层结构特征信息, 设计了增强激活模块 (enhance activation module, EAM), 使用 Transformer 网络代替 CNN, 将 Transformer 网络作为骨干网, 增强全局颜色信息的捕获能力, 并在编码器和解码器之间重新设计了跳跃连接空间注意力模块 (skip connected spatial attention module, CSAM), 用于获取结直肠病理图像的位置空间信息。

3) 在图像特征重建部分, 受到 Zhu 等<sup>[26]</sup> 提出的 FPN 思想启发, 在此思想的基础上, 根据结直肠病理图像纹理复

杂和不规则的腺体结构, 为提取不同细粒度的特征, 加入了多尺度特征信息的提取, 设计了金字塔特征提取模块 (pyramid feature extraction module, EPFM), 用于增强提取多尺度条件图像与生成图像特征的能力, 为特征重建提供更准确的特征信息。

## 1 网络架构

DMSC 的结构如图 1 所示, 其包括条件扩散模型和图像特征重建。在条件扩散模型中, 首先对结直肠病理原图像进行扩散过程, 即逐步加入噪声, 直至获得一个完全噪声化的图像。随后, 本文将目标图像作为条件信息, 连同噪声图像一起输入到 EDNet 中。经过训练, EDNet 能够估计出噪声, 这一估计值将用于计算扩散损失。最后, 通过 EDNet 得到的噪声图像, 本文进行采样以恢复为生成图像。为了进一步提升训练效果, 本文利用 EPFM 对目标图像和生成图像进行特征比对, 构建特征重建损失。这里的条件图像是与原图像对应, 并且经过扫描仪处理后的病理图像。

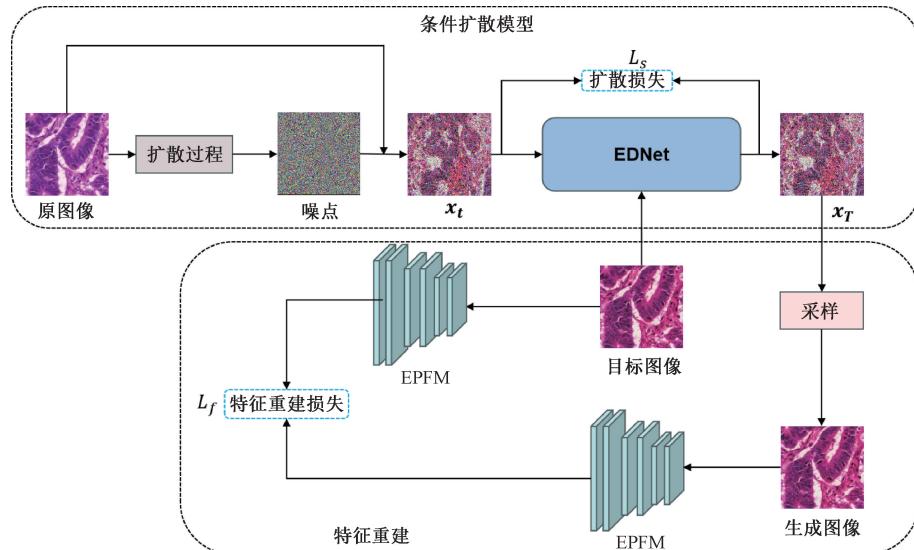


图 1 DCGW 整体结构

Fig. 1 Overall structure diagram of DCGW

### 1.1 扩散与去噪过程

将结直肠病理原图像进行扩散过程, 添加噪声的每一步尽量保持相同的噪声扩散幅度,  $t \in (0, T)$  时刻下的概率密度是  $x_0$  时刻至  $x_{t-1}$  时刻条件概率密度的连乘, 其中条件概率密度服从均值为  $\sqrt{1-\beta_t}x_{t-1}$  方差为  $\beta_t I$  的高斯分布。过程如式(1)和(2)所示。

$$q(x_{1:T} | x_0) = \prod_{t=1}^T q(x_t | x_{t-1}) \quad (1)$$

$$q(x_t | x_{t-1}) = N(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I) \quad (2)$$

扩散模型在去噪过程中不能满足规定颜色转化后的图像样式, 存在生成图像不确定的情况。为解决此问题, 设计了一种特定染色规则的条件扩散模型。给定输入

$(x_t, c^*, t)$ , 其中,  $x_t$  是噪声图像,  $c^*$  表示条件图像, 也指的是目标图像,  $t$  表示  $0 \sim T$  的某一时刻。此网络用于预测噪声分布  $\epsilon_\theta$ 。在训练过程中, 使用  $L_1$  损失来优化网络。

$$L_s = \| \epsilon_t - \epsilon_\theta(x_t, c^*, t) \| \quad (3)$$

条件扩散模型中的反向过程是一个逆向马尔可夫过程。本文将开始用纯高斯噪声  $x_U$  估计  $q(x_{t-1} | x_t)$ 。

$$p_\theta(x_{0:U} | c^*) = p(x_U) \prod_{t=1}^U p_\theta(x_{t-1} | x_t, c^*) \quad (4)$$

$$p_\theta(x_{t-1} | x_t, c^*) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, c^*, t), \Sigma_\theta(x_t, c^*, t)) \quad (5)$$

当  $\beta_t$  的值非常小时, 每个逆向过程通过设计一个网络来估计均值  $\mu_\theta$  和方差  $\Sigma_\theta(x_t, c^*, t)$ , 均值计算过程如下:

$$\mu_\theta(x_t, c^*, t) = \frac{1}{\sqrt{1-\beta_t}} \left( x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} f_\theta(x_t, c^*, t) \right) \quad (6)$$

然后, 使用重参数化技巧对高斯分布进行采样。给定一个高斯分布, 可以通过使用公式  $z = \mu\sigma \odot \epsilon$  来采样一个  $x$ , 其中,  $\epsilon$  是正态高斯分布,  $\odot$  表示逐元素乘积。最后, 使用以下方程迭代生成样本图像, 直到获得最终的生成图像  $x_0$ 。

$$x_{t-1} = \frac{1}{\sqrt{1-\beta_t}} \left( x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} \int_\theta (x_t, c^*, t) \right) + \sigma_t z \quad (7)$$

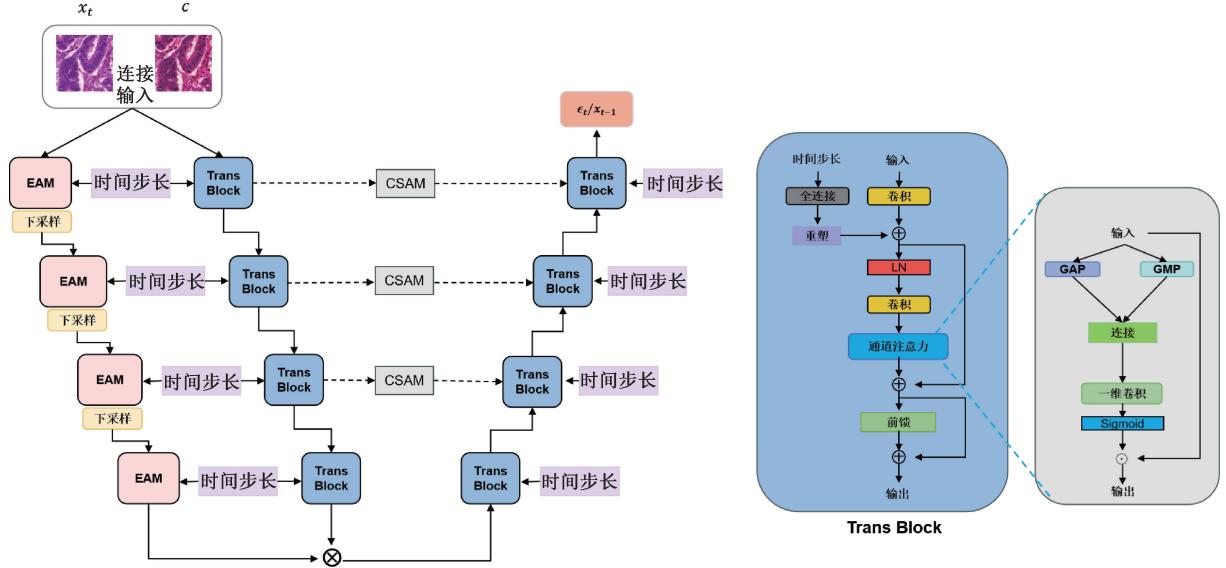


图 2 EDNet 结构

Fig. 2 Structure diagram of EDNet

给定输入噪声图像  $x_t$ , 条件图像  $c^*$  和时刻  $t \in (0, T)$ , 其中  $T$  为时间步长。首先, 将两个图像连接成 6 通道图像  $B_t$ , 然后将其输入到两个分支网络中。由于现有 Transformer 中自注意力模块捕获的信息缺少颜色信息, 受到 Peng 等<sup>[25]</sup>提出的 Transformer 块的启发, 本文重新设计的 Transformer 块的结构如图 2 所示。设计包含通道注意力的 Transformer 块。在 Transformer 块中, 使用卷积层调整通道数量, 并且将时间步长输入到全连接层进行编码。然后, 特征向量重塑为特征图  $R_t$ , 并将其添加到 6 通道图像的特征图中。过程如式(8)~(10)所示:

$$B_t = \Phi_\theta(Con(x_t, c^*)) \quad (8)$$

$$R_t = R(D_\theta(t)) \quad (9)$$

$$F = B_t + R_t \quad (10)$$

其中,  $\Phi(\cdot)$ ,  $Con(\cdot)$ ,  $D(\cdot)$  和  $R(\cdot)$  分别是卷积、连接、全连接、重塑操作。 $F$  表示输入特征。

接着, 输入特征  $F$  进行层归一化的操作, 再经过  $1 \times 1$  的卷积, 调整通道数量。随后, 经过通道注意力模块, 并与输入特征  $F$  残差相加, 送入前馈层, 得出结果。过程如式(11)和(12)所示。

## 1.2 增强去噪网络

现有去噪预测网络, 采用 CNN 与自注意力的结合, CNN 擅长处理图像的局部特征, 且自注意力机制无法有效地捕获颜色信息, 不容易学习到相关的颜色信息。为解决该问题, 本文设计了 EDNet, 该网络将 Transformer 网络作为骨干网。EDNet 结构如图 2 所示, 编码器阶段包含 Transformer 分支以及 EAM 分支, 解码器采用 Transformer 解码器。并且在解码器和编码器之间设计 CSAM。

$$\hat{F} = Att(Con_{1 \times 1}(Ln(F))) + F \quad (11)$$

$$\tilde{F} = FF(Ln(\hat{F})) + \hat{F} \quad (12)$$

其中,  $F, \hat{F}, \tilde{F} \in R^{H \times W \times C}$  分别是输入特征、中间特征和输出特征图。 $Ln(\cdot)$  是层归一化,  $Con_{1 \times 1}(\cdot)$  是  $1 \times 1$  卷积。 $Att(\cdot)$  表示通道注意力, 其包含一个平均池化层和一个最大池化层, 以及只有一个可学习参数的卷积层, 参数量更少。 $Att(\cdot)$  过程如式(13)所示:

$$\hat{F} = Sigmoid(\phi_\theta(C_t(G(F) + M(F)))) \times F \quad (13)$$

其中,  $G(\cdot)$  和  $M(\cdot)$  表示全局平均池化和全局最大池化,  $C_t$  表示连接操作, 将全局平均池化和全局最大池化的结果连接在一起。 $\phi_\theta$  为一维卷积操作。

### 1) 增强激活模块

虽然 Transformer 块可以捕获长距离依赖关系, 但在处理图像的局部特征和空间关系时效果不好, 因此, 本文设计 EAM, EAM 结构如图 3 所示, 主要由深度卷积、通道注意力、门操作和空洞卷积组成。深度卷积能够学习图像的深层特征, 空洞卷积可以在不增加参数数量的前提下,

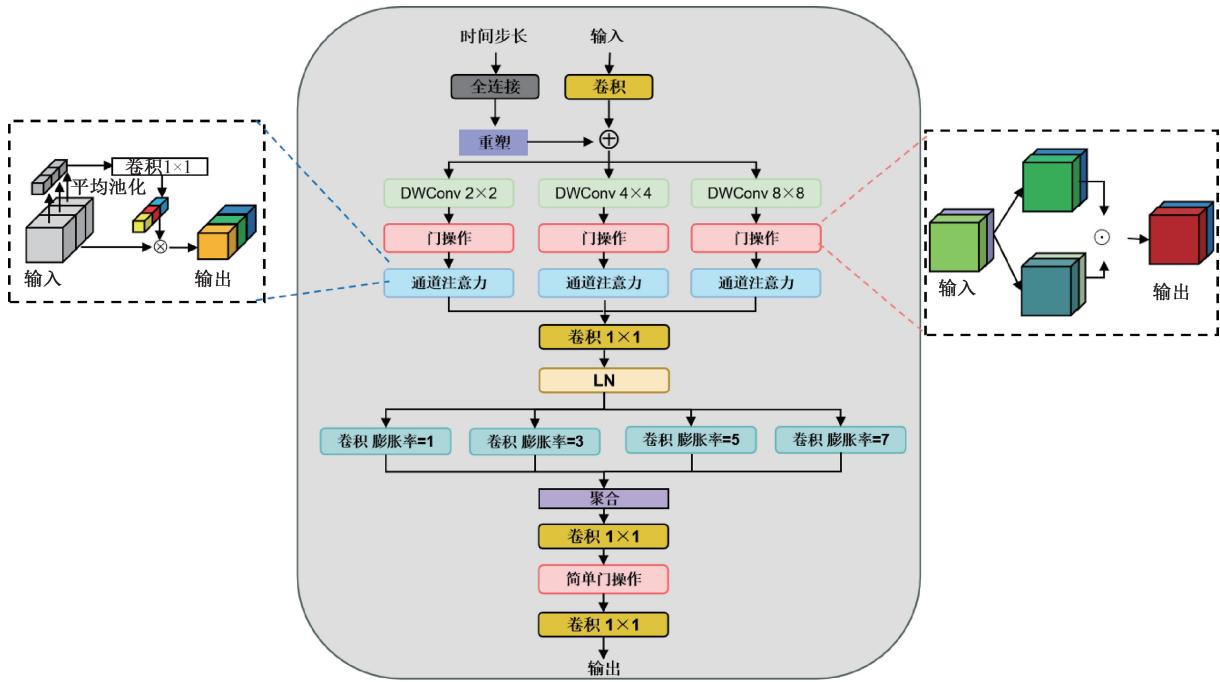


图 3 EAM 结构

Fig. 3 Structure diagram of EAM

增大卷积核的感受野,捕获更多的上下文信息。

给定输入特征  $F$ ,首先,按照利用多尺度深度卷积(depthwise convolution, DWConv)获取多尺度信息。在每个尺度内,通过通道注意力和门操作,来整合额外的非线性表示。门操作本质上是对特征图进行逐元素乘积运算。多尺度简单激活过程可以表示为:

$$\begin{cases} F^2 = SCA(SimpleGate(f_{2 \times 2}(F))) \\ F^4 = SCA(SimpleGate(f_{4 \times 4}(F))) \\ F^8 = SCA(SimpleGate(f_{8 \times 8}(F))) \end{cases} \quad (14)$$

其中,  $SimpleGate$  是门操作,其将经过多尺度深度卷积提取的特征,依据通道数均分为两部分,并分别作为输入进行后续处理,  $SCA$  是通道注意力。然后,设置  $1 \times 1$  卷积聚合多尺度表示  $F' = f_{1 \times 1}(Concat(F^2, F^4, F^8))$ , 在层归一化(layer normalization, LN)后,利用不同膨胀率空洞卷积获取上下文信息,其过程如下:

$$\begin{cases} F_1^p = f_{1 \times 1}^p(F') \\ F_3^p = f_{3 \times 3}^p(F') \\ F_5^p = f_{5 \times 5}^p(F') \\ F_7^p = f_{7 \times 7}^p(F') \end{cases} \quad (15)$$

再聚合空洞卷积信息表示为  $\bar{F} = Concat(F_1^p, F_3^p, F_5^p, F_7^p)$ ,最后,输出  $Y$  可以通过以下公式得到:

$$Y = f_{1 \times 1}(SimpleGate(f_{1 \times 1}(\bar{F}))) \quad (16)$$

其中,  $F_c^* \in R^{\sqrt{N} \times \sqrt{N}}$  是 patch 级的局部和全局特征,  $I^{N \times 1}$  重塑算子,用来将  $M_c$  变平匹配与  $A^*$  相乘的大小,而  $I^{\sqrt{N} \times \sqrt{N}}$  是将其重塑到  $M_c$  大小的算子。

## 2) 跳跃连接空间注意力模块

特征经过编码器和解码器部分的最大池化和上采样操作会丢失部分位置空间等信息,导致去噪神经网络不能准确提取结直肠病理图像的位置空间信息。为解决这一问题,在跳跃连接中加入 CSAM。空间注意力模块通过直接级联编码层和对应反卷积层的特征,融合两者互补的特征,抑制上采样产生的噪声。CSAM 结构如图 4 所示。首先,输入的特征图经过卷积调整通道数,随后分别进入全局最大池化层(global maximum pooling, GMP)和全局平均池化层(global average pooling, GAP),并通过池化后的结果相加,保留特征中的位置信息。然后,通过 ReLU 激活函数处理生成的特征映射,并利用卷积减少特征映射通道的维度,再使用 Sigmoid 函数产生注意力权重。最后,将输入的特征图与注意力权重逐元素相乘,并作为输出结果传递到解码器块,与上采样的特征图进行级联。

## 3) 金字塔特征提取模块

在设计特征重建损失时,使用 ResNet50 进行特征提取,这种方式不能提取结直肠病理图像中的多尺度空间特征信息,从而影响整个网络训练的效果。为了解决此问题,设计 EPFM 来提取多尺度空间特征信息。EPFM 的结构如图 5 所示。输入特征,根据通道数量分成两个分支。然后,第一个分支分别进行深度卷积操作,全局平均池化层。计算过程如式(17)和(18)所示。

$$Q_1 = P_{avg}(C_{dw}(I_1)) \quad (17)$$

$$Q_2 = P_{avg}(C_{dw}(I_2)) \quad (18)$$

其中,  $C_{dw}$  和  $P_{avg}$  是深度卷积  $1 \times 1$  和全局平均池化。

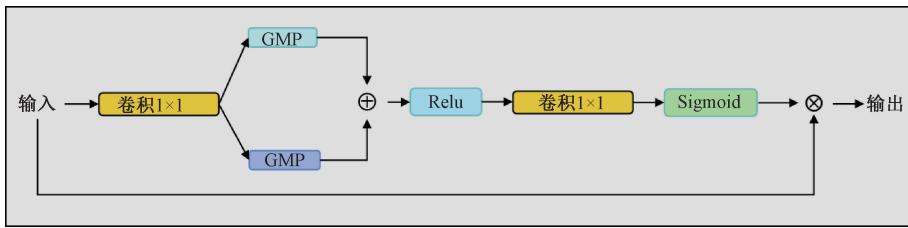


图 4 CSAM 结构

Fig. 4 Structure diagram of CSAM

将得到的  $Q_1^a$  和  $Q_2^a$  进行卷积操作, 并将它们连接起来以生成注意力特征图, 再经过 Softmax 进行归一化。

$$Q = \text{Softmax}(Q_1 + Q_2) \quad (19)$$

分支的输出结果  $Q$  通过残差连接与  $I$  进行逐元素相乘再相加。

$$I^* = (Q \cdot I) \oplus I \quad (20)$$

得出的结果  $I^*$  与第 2 个分支的输出  $I^c$  进行逐元素相加, 得出  $I'$ 。这里  $I^c$  和  $I'$  计算过程如下:

$$I' = I^* \oplus I^c \quad (21)$$

$$I^c = C_{dw}^{3 \times 3}(I) \quad (22)$$

其中,  $C_{dw}^{3 \times 3}$  是深度卷积  $3 \times 3$ 。 $I'$  再经过全局最大池化层和卷积后, 与经过通道注意力后的  $I^*$  连接起来, 作为最后输出。

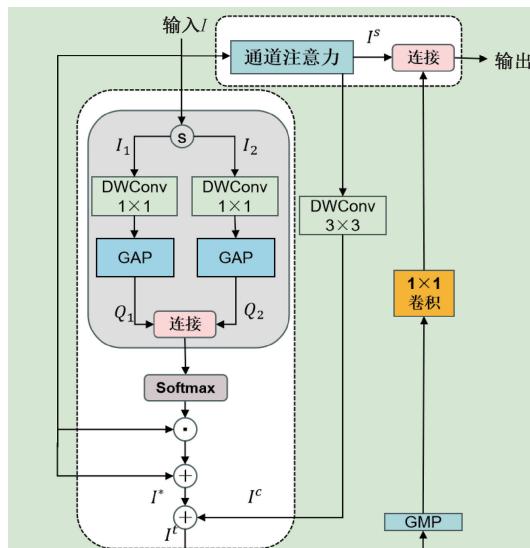


图 5 EPFM 结构

Fig. 5 Structure diagram of EPFM

### 1.3 损失函数的定义

为了每个时间步  $t$  恢复的图像应尽可能接近目标图像, 本文提出如下损失函数:

$$L = L_s + \gamma L_f \quad (23)$$

$$L_f = D_{\cos}(G(x_0), G(\tilde{x}_0)) \quad (24)$$

其中,  $L_s$  是扩散损失,  $G(\cdot)$  是输入图像的特征提取过程,  $D_{\cos}$  是两个输入之间的余弦距离。 $\gamma$  是这些项的平衡

权重。在本文的所有实验中, 因为生成网络中主要损失项是  $L_s$ , 为了避免  $L_f$  损失项对总损失的影响较大, 但是对整个网络可以进行一部分的优化,  $\gamma$  值的选定, 将通过实验分析确定。

## 2 实验设计及分析

### 2.1 数据集

为了验证所提方法对结直肠病理图像染色归一化的性能, 在两个公共数据集上进行了实验, 第 1 个数据集是 2015 年 MICCAI Gland Challenge 数据集 Glas<sup>[27]</sup>, 该数据包含 165 张结直肠腺体病理图像, 其中训练集 85 张, 测试集 80 张。第 2 个数据集是结直肠癌腺体<sup>[28]</sup> (ColoRectal adenocarcinoma gland, CRAG), 包含 173 张训练图像和 40 张测试图像。

### 2.2 实验设置

所有实验过程采用 PyTorch1.7.1 框架, Cuda 版本为 12.2, 显卡内存为 24 GB。网络训练使用 AdamW 优化器来最小化目标函数, 学习率设置为 0.005。将两个公共数据集进行大小为  $128 \times 128$  的随机裁剪, 为了平衡批大小和图像大小, 将批大小和图像大小分别设置为 8 和  $128 \times 128$ 。条件扩散模型的训练时间步长设置为  $T=1500$ 。在  $\beta$  范围为  $[10^{-6}, 10^{-2}]$  进行抽样。在测试阶段, 图像输入尺寸为  $256 \times 256$ 。用于训练和测试的 GPU 是 NVIDIA GeForce RTX 3090。

### 2.3 评价指标

为了评价染色归一化后的图像质量, 本文使用四个评价指标: 结构相似性<sup>[29]</sup> (structural similarity index, SSIM), 峰值信噪比 (peak signal-to-noise ratio, PSNR), 空间相关系数 (spatial correlation coefficient, SCC)<sup>[30]</sup>, 相对平均光谱误差 (relative average spectral error, RASE)<sup>[31]</sup>。第一个评价指标 SSIM 用于衡量两幅图像之间结构相似度的指标, 它由  $l(a, b)$ ,  $s(a, b)$  和  $c(a, b)$  组合而成。

$l(a, b)$ ,  $s(a, b)$  和  $c(a, b)$  分别表示包括亮度、结构和对比度在内的 3 个因素。亮度是图像表面照明与反射率的乘积, 但结构与照明无关。SSIM 方法满足对称性、有界性和唯一最大值的参数要求。人类视觉对视觉场景中的结构和亮度信息高度敏感, 而 SSIM 比较原图像和生成图像之间的亮度、结构和对比度的 3 个函数。SSIM 方程

由式(25)组合而成,为测量原图像和生成图像之间结构、亮度和对比度的差异提供了一个参考度量。生成图像的质量是相对于初始图像或原图像来评估的,SSIM 数值的范围从 0~1,数值越接近 1 表示染色归一化效果越好。

$$\begin{cases} l(a, b) = \frac{\mu_a \mu_b + t_1}{\mu_a^2 + \mu_b^2 + t_1} \\ c(a, b) = \frac{2\sigma_{ab} + t_2}{\sigma_a^2 + \sigma_b^2 + t_2} \\ s(a, b) = \frac{\sigma_{ab} + t_3}{\sigma_a \sigma_b + t_3} \end{cases} \quad (25)$$

其中,  $a$  是原图像,  $b$  是生成图像, 变量  $\mu_a$  和  $\mu_b$  分别代表原图像和生成图像的样本均值, 而  $\sigma_a$  和  $\sigma_b$  则分别表示它们的标准差。原图像和生成图像之间的相关系数用  $\sigma_{ab}$  表示, 而常数  $t_1$ 、 $t_2$  和  $t_3$  用于在 SSIM 接近零时稳定其值。SSIM 的数学公式具体如式(26)所示。

$$SSIM(a, b) = \left\{ \left( \frac{\mu_a \mu_b + t_1}{\mu_a^2 + \mu_b^2 + t_1} \right) \left( \frac{2\sigma_{ab} + t_2}{\sigma_a^2 + \sigma_b^2 + t_2} \right) \right\} \quad (26)$$

第二评价指标 PSNR 通过比较原始图像和处理后图像的像素差异来评估图像质量, PSNR 的值越高, 表示图像失真越小, 质量越好。计算过程如式(27)和(28)所示。

$$MSE(f, g) = \frac{1}{M^* N^*} \sum_{s=1}^{M^*} \sum_{e=1}^{N^*} (f_{se} - g_{se})^2 \quad (27)$$

$$PSNR(f, g) = 10 \log_{10} (255^2 / MSE(f, g)) \quad (28)$$

其中,  $M^*$  和  $N^*$  是图像大小,  $f$  是原图像,  $g$  是生成图像,  $MSE$  是均方误差。

第三评价指标 SCC 考虑了图像中相邻像素之间的相似性和关联程度, SCC 的值较高, 表示图像的空间结构保

持得较好, 图像质量较高。计算过程如式(29)所示。

$$SCC = \frac{\text{cov}(X, Y)}{(\sigma_a \times \sigma_b)} \quad (29)$$

其中,  $X$  和  $Y$  分别是处理后的图像和原始图像中的像素值,  $\text{cov}(X, Y)$  是它们之间的协方差,  $\sigma_a$  和  $\sigma_b$  分别是它们的标准差。

第四评价指标 RASE 用于评估重构或处理后的图像与原始图像在光谱信息上的差异。RASE 值越小, 说明处理后的图像在光谱信息上与原始图像越接近, 图像质量越高。计算过程如式(30)和(31)所示。

$$RMMSE = \sqrt{\frac{\sum_{i=1}^n (f(x^i) - g(x^i))^2}{n}} \quad (30)$$

$$RASE = \frac{1}{M} \sqrt{\frac{\sum_{i=1}^N RMSE^2(x^i)}{N}} \quad (31)$$

其中,  $f(x^i)$  和  $g(x^i)$  分别表示处理后图像和原始图像,  $M$  表示原始多光谱图像  $N$  个光谱波段的平均辐射值,  $x^i (i = 1, 2, 3)$  分别对应  $R, G, B$  波段。

### 3 实验结果与分析

#### 3.1 损失函数权重选择实验

本文的损失函数由扩散损失  $L_s$  和重建损失  $L_f$  组成, 由于生成网络中主要损失项是  $L_s$ , 因此设置  $L_s$  损失权重为 1, 将通过实验选择  $L_f$  损失权重  $\gamma$  的值。权重选择实验如表 1 所示, 通过不断降低  $\gamma$  的值, 观察方法的性能变化, 当  $\gamma = 0.001$  时, DMSC 方法生成图像质量最好, 指标最高。

表 1 权重参数  $\gamma$  的选择实验

Table 1 Selection experiment of weight parameter  $\gamma$

$\gamma$	SSIM		PSNR		SCC		RASE	
	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG
0.01	0.831	0.836	24.02	23.21	0.262	0.264	1 320	1 309
0.009	0.856	0.843	24.14	23.18	0.273	0.272	1 229	1 216
0.008	0.861	0.855	24.26	23.23	0.282	0.283	1 181	1 156
0.007	0.863	0.853	24.13	23.11	0.272	0.264	1 056	1 089
0.006	0.869	0.863	24.21	23.18	0.276	0.271	1 146	1 123
0.005	0.873	0.854	24.03	23.21	0.281	0.276	1 121	1 096
0.004	0.872	0.867	24.06	23.33	0.288	0.281	1 063	1 056
0.003	0.867	0.862	24.21	23.38	0.295	0.286	1 011	1 106
0.002	0.875	0.867	24.46	23.42	0.289	0.294	995	1 006
0.001	<b>0.878</b>	<b>0.869</b>	<b>24.56</b>	<b>23.45</b>	<b>0.301</b>	<b>0.297</b>	<b>987</b>	<b>1 021</b>
0.0009	0.872	0.868	24.23	23.21	0.285	0.274	1 002	1 210

### 3.2 消融实验

本文提出了一种条件扩散模型方法, 在去噪过程中使

用 EDNet, 并且在其中引入了 EAM, 以及在编码器与解码器之间设计 CSAM。在重建特征损失过程中, 设计 EPFM

进行特征提取。为验证各模块的有效性,设计了消融实验。在消融实验中,设置 EPFM 和 ResNet50 分别与条件扩散模型其主干为 Transformer 网络、EAM、CSAM 进行组合,并且在 GlaS 和 CRAG 数据集上进行实验,以全面评估所提方法的性能。

各模块的消融实验如表 2 所示,从表 2 中可观察到,Transformer+EPFM 和 Transformer+ResNet50 的方案

在性能上比单独使用 Transformer 实现了明显的提升,分别 SSIM 指标上提高了 7.3% 和 2.4%,说明在增加重建损失项后,对整个网络的性能具有优化的作用。在 Transformer+EPFM 和 Transformer+ResNet50 方案上进行其他模块的组合,可以发现,各个组合性能都高于 Transformer+EPFM 和 Transformer+ResNet50 方案,因此,EAM、CSAM 在提高性能过程中都发挥着重要作用。

表 2 各模块消融实验

Table 2 Ablation experiment of various modules

方法	SSIM		PSNR		SCC		RASE	
	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG
Transformer	0.728	0.712	16.08	15.96	0.256	0.242	1 427	1 356
Transformer+ResNet50	0.752	0.756	16.56	16.03	0.259	0.252	1 359	1 345
Transformer+EAM+ResNet50	0.818	0.810	16.85	16.26	0.264	0.252	1 325	1 421
Transformer+CSAM+ResNet50	0.812	0.799	16.72	16.96	0.263	0.251	1 409	1 339
Transformer+CSAM+EAM+ResNet50	0.862	0.852	23.12	22.65	0.296	0.295	1 012	1 213
Transformer+EPFM	0.801	0.798	16.55	16.09	0.260	0.247	1 328	1 342
Transformer+EAM+EPFM	0.821	0.815	17.23	16.36	0.264	0.241	1 321	1 426
Transformer+CSAM+EPFM	0.815	0.802	16.23	17.02	0.261	0.254	1 416	1 348
Transformer+CSAM+EAM+EPFM	<b>0.878</b>	<b>0.869</b>	<b>24.56</b>	<b>23.45</b>	<b>0.301</b>	<b>0.297</b>	<b>987</b>	<b>1 021</b>

最后,在 Transformer+CSAM+EAM+EPFM 和 Transformer+CSAM+EAM+ResNet50 方案中,重建损失过程中使用 EPFM 在 SSIM 评价指标上,2 个数据集分别比使用 ResNet50 提高了 1.6% 和 1.7%。这有效证明了 EPFM 与其他模块全组合时,在整个结直肠病理图像网络训练中,起到了优化网络的性能。并且其 4 个评价指标达到所有组合方案最佳,各个模块发挥着重要的作用,学习到结直肠病理图像的深层特征,并且捕获到图像上下文信息,所生成的图像质量较好,且病理图像结构信息较为完整。

时间步长在 2 个公共数据集的选择实验如表 3 所示,从表中可以看出,在反向恢复图像时,一般采用均匀分布对时间步长进行采样。当训练时间步长设置为  $T=4000$ ,反向过程中的采样步长设置为 40 时,时间间隔为 100,因此,采样的时间序列是  $\{4000, 3600, 3200, 400, 0\}$ 。在本文实验中,找到一个合适的方案,本文尝试使用不同数量的步长在反向过程中恢复图像。选择  $S_t = \{40, 20, 10\}$ ,结果如表 3 所示。对于所有不同步长,4 个评价指标值都保持较高。当步长  $S_t = 10$ ,其运行时间最小,评价指标性能较好,因此,在反向过程中设置采样步长为 10。

表 3 时间步长在 2 个公共数据集选择实验

Table 3 Experiments on the selection of time steps in two public datasets

时间	运行时间/ s	SSIM		PSNR		SCC		RASE	
		GlaS	CRAG	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG
40	1.78	0.877	0.868	24.3	23.46	0.302	0.294	982	1 029
20	1.35	0.872	0.861	24.6	23.39	0.299	0.291	988	1 031
10	1.21	0.878	0.869	24.5	23.45	0.301	0.297	987	1 021

### 3.3 实验结果对比

本文所提方法 DMSC 是在对现有条件扩散模型 CNCDM 和特征提取网络 FPN 进行改进的基础上提出的方法,该方法将 ICDM 和 EPFM 组合在一起,ICDM 对 CNCDM 中的去噪网络进行改进,提出在去噪过程中使用 EDNet。EPFM 是对 FPN 进行改进,提出使用 EPFM 提

取特征。为说明改进后的效果,将 DMSC 的 ICDM 和 EPFM 分别与 CNCDM 和 FPN 及其现有改进模型基于噪声估计网络的条件扩散模型 (noise estimation network-based conditioned diffusion model, NECDM)<sup>[32]</sup> 和采用 CNN 和 Transformer 的金字塔卷积网络 (CNN and transformer-based pyramid convolution network,

PYCN)<sup>[33]</sup> 进行实验比较, 具体包括 CNCDM + FPN、CNCDM + PYCN、CNCDM + EPFM、NECDM + FPN、NECDM + PYCN、NECDM + EPFM、ICDM + FPN、ICDM + PYCN 和 ICDM + EPFM 九组组合形式, 9 组比较结果如表 4 所示, 从表中可以看出, CNCDM + FPN 组合在一起的 4 项评价指标的性能表现明显低于其他组合。与 CNCDM + FPN 组合相比, 对 CNCDM 和 FPN 改进后的各组合的指标性能都有所提升, 说明改进 CNCDM 和 FPN 的组合都能够提高染色归一化方法的性能。本文提出的

方法 DMSC 将 ICDM 和 EPFM 同时应用到染色归一化方法中, 4 项指标的性能均高于其他组合形式, 说明与 NECDM 和 PYCN 相比较, 同时使用 ICDM 和 EPFM 能生成更高质量的图像, 保留的结直肠病理图像的纹理和腺体结构信息更加完整。这是因为通过改进 CNCDM 和 FPN, ICDM 中的去噪网络 EDNet 在准确定位结直肠的腺体位置空间信息的同时, 能够获得结直肠病理图像的整体颜色信息和深层结构特征信息。而 EPFM 能够有效提取结直肠病理图像中的多尺度空间特征信息。

表 4 条件扩散模型和特征提取网络的改进效果实验

Table 4 Experiments on the improved effects of conditional diffusion models and feature extraction network

方法	SSIM		PSNR		SCC		RASE	
	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG	GlaS	CRAG
CNCDM + FPN	0.803	0.793	19.21	19.02	0.251	0.242	1.278	1.264
CNCDM + EPFM	0.822	0.815	21.18	20.04	0.258	0.256	1.159	1.257
CNCDM + PYCN	0.836	0.829	21.82	20.39	0.262	0.259	1.132	1.235
NECDM + FPN	0.840	0.831	21.93	20.82	0.271	0.263	1.107	1.203
NECDM + PYCN	0.847	0.839	22.06	21.15	0.275	0.272	1.056	1.189
NECDM + EPFM	0.852	0.841	22.56	21.77	0.281	0.279	1.041	1.148
ICDM + FPN	0.859	0.851	22.97	22.65	0.284	0.282	1.035	1.121
ICDM + PYCN	0.862	0.855	23.83	22.93	0.295	0.287	9.91	1.098
ICDM + EPFM	<b>0.878</b>	<b>0.869</b>	<b>24.56</b>	<b>23.45</b>	<b>0.301</b>	<b>0.297</b>	<b>9.87</b>	<b>1.021</b>

DMSC 使用了条件扩散模型进行染色归一化研究, CAGN<sup>[14]</sup>、SPCN<sup>[15]</sup>、STST<sup>[16]</sup> 和 MDST<sup>[18]</sup> 都使用了 GAN 进行染色归一化研究, 条件扩散模型和 GAN 都属于生成模型的范畴, 能够通过学习数据的生成分布, 实现对图像内在结构的学习, 条件扩散模型是基于扩散过程生成图像, 而 GAN 是通过生成器和判别器的对抗生成图像。RestainNet<sup>[17]</sup> 从自监督机制的角度研究病理图像的染色归一化。目前, 生成对抗网络和自监督机制都在染色归一化研究中取得了较好的效果, 为验证本文所提方法 DMSC 的性能, 将 DMSC 与现有染色归一化方法 CAGN, SPCN, STST, RestainNet 和 MDST 进行比较。本文将 DMSC 方法与现有几种组织学染色归一化方法在 GlaS 测试集上的 4 个评价指标对比结果如表 5 所示。

表 5 不同方法在 GlaS 测试集上染色归一化

Table 5 Different methods for stain normalization on the GlaS test set

方法	SSIM	PSNR	SCC	RASE
CAGN <sup>[14]</sup>	0.835	17.24	0.266	1.262
SPCN <sup>[15]</sup>	0.841	17.53	0.263	1.427
STST <sup>[16]</sup>	0.845	20.28	0.282	1.012
RestainNet <sup>[17]</sup>	0.874	18.65	0.278	1.003
MDST <sup>[18]</sup>	0.869	22.47	0.297	9.96
DMSC	<b>0.878</b>	<b>24.56</b>	<b>0.301</b>	<b>9.87</b>

从表 5 可以看出, 本文方法与 CAGN, SPCN, STST, RestainNet 和 MDST 对比, 4 个评价指标性能都达到最佳水平。尤其在 SSIM 评价指标上, 它对结直肠病理图像的亮度、对比度和结构进行比较, DMSC 方法相较于 SPCN 方法, 在 SSIM 指标上显著提高了 3.7%。本文方法在 4 个评价指标上与 MDST 方法对比时, SSIM 指标提高了 0.9%, 本文方法采用条件扩散模型的方式, 并且引入 EAM、CSAM 和 EPFM, 其较好的保持了图像的结构信息, 具有保留数据语义结构的能力, 其性能好于基于 GAN 和基于自监督机制的染色归一化方法。

本文同时也在 CRAG 数据集上进行了方法的对比实验, 结果如表 6 所示。从表中可以明显看出, 本文方法在 4 个评价指标上都取得了较好的性能, DMSC 的性能优于

表 6 不同方法在 CRAG 测试集上染色归一化

Table 6 Different methods for stain normalization on the CRAG test set

方法	SSIM	PSNR	SCC	RASE
CAGN <sup>[14]</sup>	0.831	17.96	0.255	1.473
SPCN <sup>[15]</sup>	0.833	18.04	0.260	1.521
STST <sup>[16]</sup>	0.840	19.56	0.251	1.221
RestainNet <sup>[17]</sup>	0.864	22.31	0.245	1.038
MDST <sup>[18]</sup>	0.856	20.98	0.281	1.123
DMSC	<b>0.869</b>	<b>23.45</b>	<b>0.297</b>	<b>1.021</b>

基于 GAN 和基于自监督机制的方法。本文所提方法在图像失真方面较小, 图像质量的损失程度较小, 保持较好的结构信息, 染色归一化后的图像与条件图像在光谱信息上接近, 图像质量较高。

为了直观展现 DMSC 方法在 2 个公共测试集上染色归一化的效果, 本文绘制了如图 6 所示的可视化图。此图

展示了 DMSC 方法的加噪与去噪流程, 以及测试集的原图像与经方法处理后的生成图像。特别地, 在图 6 的第二列中, 展示了一个结直肠癌病理图像的示例。该图像中的腺体形状发生了明显的变形与扭曲, 针对此类复杂图像, 本文提出的 DMSC 方法能够生成结构清晰、信息完整且染色区域统一的病理图像。

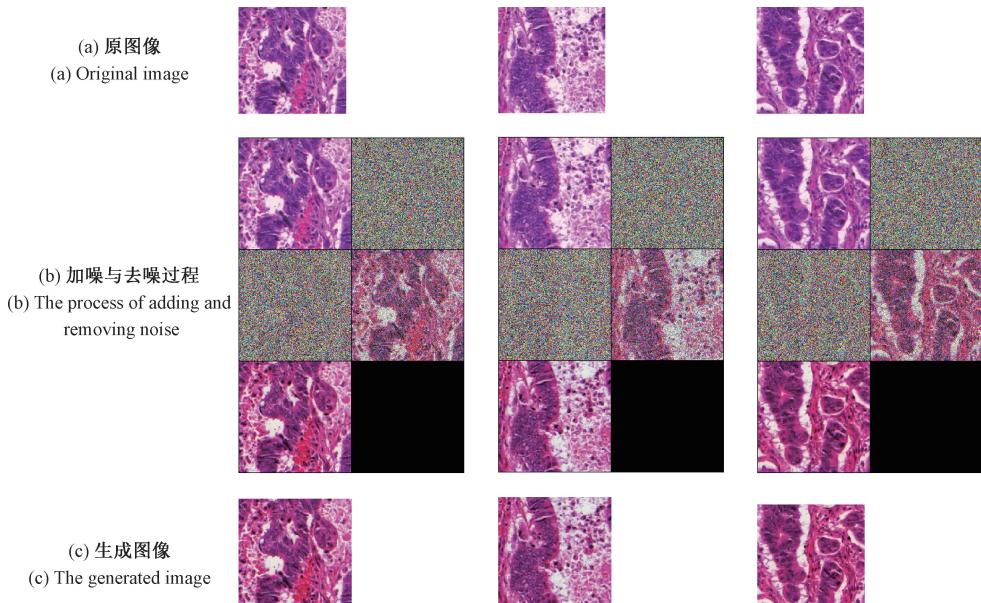


图 6 DMSC 方法的可视化图

Fig. 6 Visualization results of DMSC method

在两个公共数据集的可视化结果如图 7 所示, 各方法在结直肠病理图像的生成效果呈现出差异, 本文提出的

DMSC 方法展现出了最佳的生成性能。以图 7 第 1 行蓝色方框内的腺体区域为例, CAGN、SPCN 和 STST 方法在生

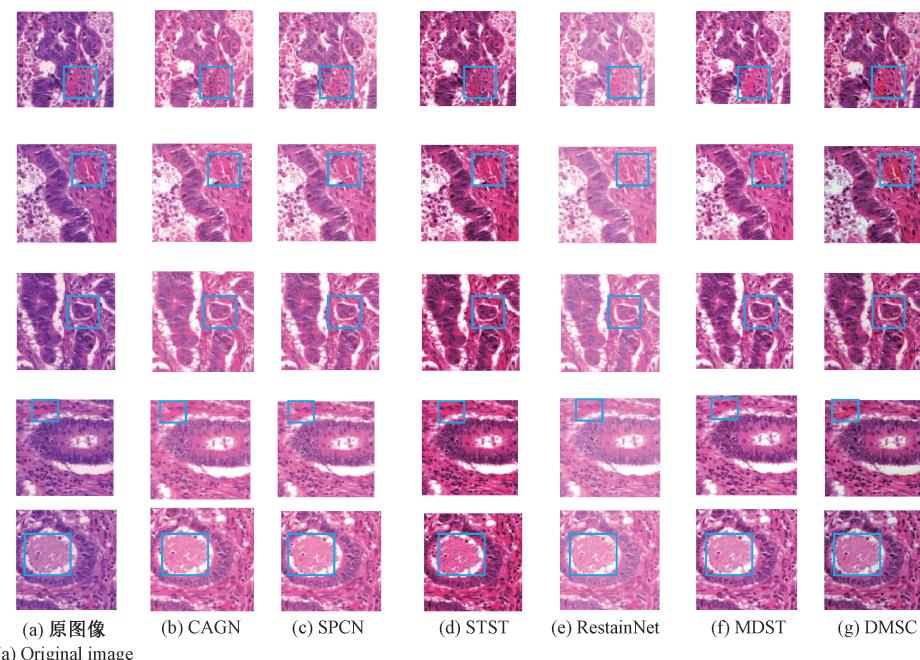


图 7 不同方法在公共数据集上的可视化结果

Fig. 7 Visualization results of different methods on public datasets

成图像时出现了明显的噪点。虽然 STST 方法采用变体 GAN 结构和染色转换相结合的方法,其生成的腺体结构相较于 CAGN 和 SPCN 有所改善,但是效果仍有限。RestainNet 和 MDST 方法分别通过自监督方式和变体 GAN 中注意力模块,较好地处理了图像噪点问题,对图像细节的处理也较好,但仍存在结构特征信息缺失情况。第 2 行和第 4 行的蓝色方框中,CAGN、SPCN、RestainNet 和 MDST 方法在此区域虽然细节部分产生的纹理清晰,但是产生的颜色亮度对比较为不明显。在第 3 行和第 5 行的蓝色方框中,原图中的黑色组织区域亮度对比度明显,而 SPCN、RestainNet 和 MDST 方法对于黑色组织亮度对比度较不明显。SPCN 和 STST 方法甚至会产生少量噪点。

本文提出的 DMSC 方法在生成效果上表现最佳。其利用 EAM 的深度卷积学习能力,捕捉了图像中更精细的结构特征;同时,CSAM 有效弥补了编码器和解码器在上采样和下采样过程中丢失的空间特征信息。此外,重建特征损失分支中的 EPFM 提取了多尺度空间特征信息,进一步提升了方法的整体性能。因此,DMSC 方法在生成结直肠病理图像时,能够保持较高的图像亮度和清晰的腺体结构细节,且未出现生成噪点的情况。

## 4 结 论

本文提出一种结直肠病理图像染色归一化方法 DMSC,该方法包括条件扩散模型和图像特征重建。在条件扩散模型中设计了 EDNet,利用 EDNet 中的 EAM 和 CSAM 增强语义结构的感知能力,获得更加准确的结直肠病理图像的空间结构特征。在图像特征重建过程中设计了 EPFM,通过获得的结直肠病理图像的多粒度特征信息生成高质量的结直肠病理图像。在 2 个公共数据集上的实验结果表明,DMSC 能够从结直肠病理图像的复杂纹理和不规则腺体结构中提取出有效的特征信息,进一步提升了结直肠病理图像的生成质量。本文提出的 DMSC 存在模型参数多且规模大的问题,影响了 DMSC 的生成图像的效率,在未来的研究工作中,将进一步研究如何优化网络结构和降低参数量,提升 DMSC 生成图像的效率。

## 参考文献

- [1] 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]. CA: A Cancer Journal for Clinicians, 2021, 71(3): 209-249.
- [2] XIA CH F, DONG X S, LI H, et al. Cancer statistics in China and United States, 2022: Profiles, trends, and determinants[J]. Chinese Medical Journal, 2022, 135(5): 584-590.
- [3] BUCCAFUSCA G, PROSERPIO I, TRALONGO A C, et al. Early colorectal cancer: Diagnosis, treatment and survivorship care [J]. Critical Reviews in Oncology/Hematology, 2019, 136: 20-30.
- [4] CIARDIELLO F, CIARDIELLO D, MARTINI G, et al. Clinical management of metastatic colorectal cancer in the era of precision medicine[J]. CA: A Cancer Journal for Clinicians, 2022, 72(4): 372-401.
- [5] 闫雯, 汤烨, 张益肇, 等. 深度学习在数字病理中的应用[J]. 中国生物医学工程学报, 2018, 37(1): 95-105.
- [6] YAN W, TANG Y, ZHANG Y ZH, et al. Deep learning in digital pathology analysis [J]. Chinese Journal of Biomedical Engineering, 2018, 37 (1): 95-105.
- [7] ASAFA M Z, RAO B, AKRAM M U, et al. Dual contrastive learning based image-to-image translation of unstained skin tissue into virtually stained H&E images [J]. Scientific Reports, 2024, 14 (1): 2335-2342.
- [8] VAHADANE A, PENG T, SETHI A, et al. Structure-preserving color normalization and sparse stain separation for histological images [J]. IEEE Transactions on Medical Imaging, 2016, 35(8): 1962-1971.
- [9] YANG K, YANG Y T, LU Y F, et al. Automatic segmentation of arteries and veins in multispectral retinal imaging [J]. Electronic Measurement Technology, 2024, 46(10): 84-91.
- [10] ZHANG ZH X, TIAN CH N, BAI H X, et al. Discriminative error prediction network for semi-supervised colon gland segmentation [J]. Medical Image Analysis, 2022, 79: 102458.
- [11] 刘肇隆, 范馨月. 基于全尺度跳跃连接的 TransUNet 医学图像分割网络[J]. 国外电子测量技术, 2023, 42(11): 42-48.
- [12] LIU ZH L, FAN X Y. TransUNet medical image segmentation network based on full-scale jump connectivity [J]. Foreign Electronic Measurement Technology, 2023, 42(11): 42-48.
- [13] ZINGMAN I, FRAYLE S, TANKOYEU I, et al. A comparative evaluation of image-to-image translation methods for stain transfer in histopathology [C]. Paris: Medical Imaging with Deep Learning, 2024: 1509-1525.
- [14] ASAFA M Z, RAO B, AKRAM M U, et al. Dual contrastive learning based image-to-image translation of unstained skin tissue into virtually stained H&E images [J]. Scientific Reports, 2024, 14 (1): 2335-2342.
- [15] 史韶杰, 韩建宁, 李大威等. 基于多尺度曝光融合的医学 CT 图像增强方法[J]. 电子测量技术, 2022, 45(6): 106-111.

- SHI SH J, HAN J N, LI D W, et al. Medical CT image enhancement method based on multi-scale exposure fusion [J]. *Electronic Measurement Technology*, 2022, 45(6): 106-111.
- [14] CONG C, LIU S D, IEVA A D, et al. Colour adaptive generative networks for stain normalisation of histopathology images[J]. *Medical Image Analysis*, 2022, 82: 102580.
- [15] VAHADANE A, PENG T, SETHI A, et al. Structure-preserving color normalization and sparse stain separation for histological images [J]. *IEEE transactions on medical imaging*, 2016, 35 (8): 1962-1971.
- [16] SALEHI P, CHALECHALE A. Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis[C]. *Qom:2020 International Conference on Machine Vision and Image Processing*, 2020: 1-7.
- [17] ZHAO B CH, HAN CH, PAN X P, et al. RestainNet: A self-supervised digital re-stainer for stain normalization [J]. *Computers and Electrical Engineering*, 2022, 103: 1-10.
- [18] BARUA B, BORA K, DAS A K, et al. Stain color translation of multi-domain OSCC histopathology images using attention gated cGAN[J]. *Computerized Medical Imaging and Graphics*, 2023, 106: 1-12.
- [19] WANG X, YUE Q R, LIU X G. Conditional diffusion model-based generation of speckle patterns for digital image correlation [J]. *Optics and Lasers in Engineering*, 2024, 175:1-14.
- [20] ZHU Y ZH, LI ZH H, WANG T W. Conditional text image generation with diffusion models [C]. *Vancouver:Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 14235-14245.
- [21] CHEN T, WANG C, SHAN H. BerDiff: Conditional bernoulli diffusion model for medical image segmentation [C]. *Vancouver: Medical Image Computing and Computer Assisted Intervention*, 2023: 491-501.
- [22] WU J, FU R, FANG H, et al. MedSegDiff: Medical image segmentation with diffusion probabilistic model[C]. *Machine Learning Research*, 2024: 1623-1639.
- [23] YANG Y, FU H, RIVERO A A, et al. DiffMIC: dual-guidance diffusion network for medical image classification [C]. *Vancouver: Medical Image Computing and Computer Assisted Intervention*, 2023: 95-105.
- [24] PRABHAT K C, MYERS K J, FARHANGI M M, et al. On use of augmentation for the DNN-based CT image denoising[C]. *San Francisco: 7th International Conference on Image Formation in X-Ray Computed Tomography*, SPIE, 2022:1-18.
- [25] PENG L T, ZHU CH L, BIAN L H. U-shape transformer for underwater image enhancement[J]. *IEEE Transactions on Image Processing*, 2023, 32: 3066-3079.
- [26] ZHU L X, LEE F, CAI J W, et al. An improved feature pyramid network for object detection [J]. *Neurocomputing*, 2022,483: 127-139.
- [27] RASTOGI P, KHANNA K, SINGH V. Gland segmentation in colorectal cancer histopathological images using U-net inspired convolutional network[J]. *Neural Computing and Applications*, 2022, 34 (7): 5383-5395.
- [28] AWAN R, SIRINUKUNWATTANA K, EPSTEIN D, et al. Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images[J]. *Scientific Reports*, 2017, 7(1): 1-13.
- [29] SAIFULLAH S, Dreżewski R. Modified histogram equalization for improved cnn medical image segmentation[J]. *Procedia Computer Science*, 2023, 225: 3021-3030.
- [30] MULAY S, RAM K, SIVAPRAKASAM M. Attention adaptive instance normalization style transfer for vascular segmentation using deep learning [J]. *Applied Intelligence*, 2023, 53: 29638-29655.
- [31] ELMOASRY A, KHAN L S, KHAN M, et al. A dual layer security scheme for medical images using Hessenberg and singular value decompositions [J]. *Multimedia Tools and Applications*, 2022, 81 (4): 14001-14022.
- [32] GAO Q, SHAN H. CoCoDiff: A contextual conditional diffusion model for low-dose CT image denoising [C]. *California: Developments in X-Ray Tomography XIV*, SPIE, 2022: 92-98.
- [33] LIU X W, HU Y K, CHEN J G. Hybrid CNN-transformer model for medical image segmentation with pyramid convolution and multi-layer perceptron [J]. *Biomedical Signal Processing and Control*, 2023, 86: 105331.

### 作者简介

李子成,硕士研究生,主要研究方向为医学图像处理与分析。

E-mail:18895010216@163.com

贾伟(通信作者),博士,副教授,主要研究方向为医学图像处理与分析。

E-mail:jiawnx@163.com

赵雪芬,博士,副教授,主要研究方向为图像处理与分析。

E-mail:snownfen@163.com

高宏娟,博士,副教授,主要研究方向为图像处理与分析。

E-mail:ghj\_y2004@nxu.edu.cn