

# 面向光学相干层析指纹的条件扩散生成方法

戚佳锦<sup>1</sup>,刘义鹏<sup>1</sup>,李静<sup>2</sup>

<sup>1</sup>(浙江工业大学 计算机科学与技术学院,杭州 310023)

<sup>2</sup>(浙江省立同德医院(浙江省中医药研究院),杭州 310012)

E-mail:liuyipeng@zjut.edu.cn

**摘要:**光学相干断层扫描技术以其高分辨率和捕捉指尖皮肤三维结构的能力而闻名,能够增强指纹识别系统的防伪能力.然而,与其它生物识别技术相比,数据集的稀缺严重阻碍其广泛应用.由于采集困难且出于隐私考虑不便公开,研究数据生成是应对该挑战的解决方案之一.本文提出一种基于扩散模型的条件生成方法,利用层分割掩码作为先验知识引导生成过程,通过逐步去噪直接在像素空间建模,避免潜在扩散模型的精度损失,从而生成高保真 OCT 指纹图像.实验表明,该方法生成的样本具有逼真的皮肤结构特征,通过对 30 位领域专家进行的主观评价实验证明其生成结果解剖结构清晰、像素分布真实.进一步实验证明,使用生成数据扩充训练集可显著提升多种防伪模型的性能.

**关键词:**光学相干断层扫描;指纹;条件生成;伪造攻击检测

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)05-1205-07

## Conditional Diffusion Generation for Optical Coherence Tomography Fingerprints

QI Jiajin<sup>1</sup>, LIU Yipeng<sup>1</sup>, LI Jing<sup>2</sup>

<sup>1</sup>(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China)

<sup>2</sup>(Tongde Hospital of Zhejiang Province, Zhejiang Academy of Traditional Chinese Medicine, Hangzhou 310012, China)

**Abstract:** Optical Coherence Tomography (OCT) is renowned for its high resolution and ability to capture the three-dimensional structure of fingertip skin, significantly enhancing the anti-spoofing capability of fingerprint recognition systems. However, compared to other biometric technologies, the scarcity of datasets severely hinders its widespread application. Due to the difficulty of data collection and privacy concerns that restrict public sharing, synthetic data generation presents a more practical solution to this challenge. This paper proposes a conditional generation method based on a diffusion model, which leverages layer segmentation masks as prior knowledge to guide the generation process. By employing iterative denoising to directly model the pixel space, our approach avoids the precision loss associated with latent diffusion models, thereby producing high-fidelity OCT fingerprint images. Experiments demonstrate that the generated samples exhibit realistic skin structural features, with subjective evaluations by 30 domain experts confirming their anatomically accurate structures and authentic pixel distributions. Further validation shows that augmenting training datasets with synthetic data significantly improves the performance of various anti-spoofing models.

**Keywords:** optical coherence tomography; fingerprint; conditional generation; presentation attack detection

## 0 引言

指纹作为独特的生物特征<sup>[1]</sup>,其个体差异性体现在指尖嵴线构成的独特纹路模式.作为生物识别的关键技术,指纹识别通过分析纹路特征实现身份认证,应用领域持续扩展以满足日益多元化的安全性与便捷性需求<sup>[2,4]</sup>.但现有技术仍存在局限:表层指纹的采集易受汗渍、磨损及环境因素干扰,导致特征提取时出现纹路模糊或缺失等问题.

指纹识别系统面临的重要安全威胁是表现攻击(PA)<sup>[5]</sup>,其被定义为“通过干扰生物特征数据采集子系统来影响生物识别系统运行的行为”<sup>[6]</sup>.传统攻击方法<sup>[7,8]</sup>通常使用蜡、硅胶等材料制作手掌等实体模型,进而采集指纹以欺骗指纹检测系统.除传统的实体伪造手段外,近年来利用深度学习

算法伪造指纹图像的攻击方式也呈现逐渐增多的趋势<sup>[9]</sup>.

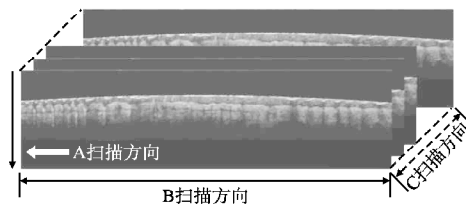


图1 OCT 指尖皮肤体数据

Fig.1 Volume data of OCT fingertip skin

光学相干断层扫描(OCT)技术凭借其非侵入性、高速运行和高分辨率等优异特性,已在医学影像领域获得广泛应用. OCT 采集到的指尖皮肤体数据如图 1 所示.与传统仅采集表

皮指纹的2D接触式指纹技术不同,OCT技术可深入指尖内部,捕获皮下多层组织的特征性结构<sup>[10]</sup>,从而具备更强的表现攻击检测(PAD)能力。

当前制约OCT指纹领域发展的关键因素之一在于缺乏公开的大规模数据集。由于生物特征数据集涉及隐私问题,许多指纹数据集已撤出公开访问渠道,这使得可公开获取的数据集更为稀缺。这一现象在OCT指纹领域尤为突出,目前可用的数据集数量极其有限<sup>[11,12]</sup>。此外,数据标注在OCT领域也是一个重大挑战:尽管其精细程度要求可能不及OCT眼底视网膜图像,但OCT指纹语义分割图的标注工作仍需要高度的专业水准。

已有若干研究采用生成对抗网络(GAN)来合成2D指纹图像<sup>[13-15]</sup>。然而,由于数据集图像风格高度集中,基于GAN的方法往往面临模式坍塌问题,导致生成指纹的多样性不足。此外,传统GAN在训练过程中存在不稳定性,可能学习到有偏的分布特征。指纹的曲率、脊线间距等细节特征直接影响其生物识别特性,但传统GAN难以精确控制这些参数,因而无法准确捕捉指纹图像中复杂的纹理特征与皮肤组织结构。

扩散模型<sup>[16,17]</sup>作为生成任务中的新一代前沿方法,在多个领域展现出卓越性能。然而,多模态大模型严重依赖海量训练数据。在OCT指纹等细分领域,有限的数据集会显著制约模型的泛化能力,使得基于ImageNet等通用数据集预训练的模型难以直接适用。相较于成本高昂的微调过程,针对OCT指纹生成任务设计专用领域模型更具实践价值。

通过在2D U-Net的输入中添加条件变量,可以实现条件生成,当前常见的文生图和图生图模型均采用了类似的嵌入或拼接方法。本研究采用指纹分割掩码作为先验条件变量,通过无分类器引导的方式,将其与含噪图像拼接形成双通道输入张量,使2D U-Net能够预测当前时间步对应的噪声。经过充分训练后,当模型能准确预测各时间步的噪声时,即可获得一个能将分割掩码映射为OCT指纹的生成器。

本研究的主要贡献总结如下:

- 1) 提出一种高度逼真的条件生成方法,能够根据分割掩码生成包含指尖皮肤层结构的OCT指纹图像。
- 2) 采用逐步去噪的扩散模型将生成任务分解为多个连续时间步的较小子任务,通过在每一步最小化条件分布误差更精确地捕捉目标分布的微观细节特征。通过主观评价实验实施的定性分析验证了该方法的有效性。
- 3) 直接在像素空间模拟扩散过程从而规避编解码过程中的精度损失,并以分割掩码作为指导学习指尖皮肤结构的丰富先验知识,有效学习真实OCT指纹的像素分布特征。将生成的指纹样本加入PAD模型的训练集后,其性能指标的变化证实:本方法生成的指纹在像素分布上与真实指纹具有高度一致性。

## 1 相关工作

图像生成是人工智能领域的重要研究方向,近年来取得了显著进展。本节将概述当前主流的生成模型方法及其最新发展,特别关注其在生物特征合成领域的应用潜力。

### 1.1 图像生成方法

根据任务目标的不同,神经网络模型可以分为两类<sup>[18]</sup>:判别模型和生成模型。判别模型专注于学习分布之间的决策边界,并构建映射以将输入数据划分为不同的类别或标签。生成模型则侧重于学习输入数据的具体分布,并从该分布中生成与训练数据相似但不完全相同的新的样本。

生成模型主要包括变分自编码器(VAE)<sup>[19]</sup>,其通过学习潜在变量分布来生成数据;生成对抗网络(GAN)<sup>[20]</sup>,其通过判别器和生成器的对抗训练来优化生成质量;基于流的模型(FBM)<sup>[21,22]</sup>,其通过可逆变换直接建模数据的概率分布;以及最先进的扩散模型(DDPM)<sup>[17]</sup>,其通过逐步去噪的过程生成高质量数据。与GAN相比,扩散模型生成高质量输出,并且训练过程通常更加稳定<sup>[23]</sup>,其逐步去噪的方法使其特别适合生成复杂结构,例如医学图像<sup>[24]</sup>。

GAN已成为深度学习的重要研究方向,其通过生成器和判别器的对抗训练生成逼真数据,尽管早期存在训练不稳定等问题,但已发展出多种改进方法,并在图像生成、风格转换和超分辨率重建等领域取得了显著成果。在医学图像领域,GAN被广泛应用于数据增强、图像去噪和超分辨率重建等任务(例如生成合成MRI图像<sup>[25]</sup>、去噪低剂量CT图像<sup>[26]</sup>),显著提高了医学图像分析的性能和诊断准确性。

去噪扩散概率模型(DDPM)<sup>[17]</sup>通过逐步加噪和去噪的过程生成高质量样本。在此基础上,去噪扩散隐式模型(DDIM)<sup>[27]</sup>引入了确定性逆向过程替代逐步采样方式,以牺牲部分生成质量为代价,显著提升了生成效率。基于分数的生成模型<sup>[28]</sup>通过优化数据分布的梯度(即分数函数)进行迭代去噪,进一步扩展了高质量生成的尺寸范围。潜在扩散模型(LDM)<sup>[29]</sup>则借助变分自编码器将图像映射到低维潜在空间中进行扩散处理,再解码回图像空间,显著降低了计算成本。LDM的出现开启了多模态大模型的发展篇章,为多模态生成任务提供了更高效的解决方案。

除了生物特征数据集常见的隐私问题外,OCT指纹研究还面临独特的挑战,将扩散模型应用于OCT指纹生成领域的研究仍处于初步阶段。OCT指纹的复杂细节,例如汗腺和汗孔,需要高分辨率才能准确合成。此外,指尖皮肤的表皮和真皮分层结构需要生成模型具备更强的细节捕捉能力。通过在像素空间中直接进行扩散并结合先验的指尖皮肤结构知识来应对这些挑战,确保高分辨率OCT指纹B扫描图像的生成。

### 1.2 指纹合成方法

SfinGe<sup>[30]</sup>是第一个合成指纹生成方法,后续的研究工作分别采用了手工方法<sup>[31-33]</sup>、基于学习的方法<sup>[34,35]</sup>以及这两种方法的结合<sup>[36]</sup>。尽管这些方法在合成逼真的指纹数据集方面取得了开创性的贡献,但它们在许多方面仍存在局限性,主要在于未能充分利用和反映指纹中包含的全部信息,生成的指纹虽然可控,但不够逼真。基于学习的方法虽然可以生成逼真的指纹,但通常难以生成同一手指的多个不同印迹。

只有结合指纹先验知识的混合方法,才能在真实性和生成数量方面取得更好的结果。首先,基于手工方法,引入先验知识指导指纹生成过程以涵盖足够的细节点。然后,通过深度学习对指纹进行渲染,以确保指纹图像具有足够的逼真度。比如PrintsGAN<sup>[14]</sup>将2D指纹的生成过程分为两个步骤,并分别训练两个具有不同参数的BigGAN。而SpoofGAN<sup>[15]</sup>则提

高了生成不同按压风格指纹的能力,这可以有效扩展指纹类别的丰富性.然而,目前在 OCT 指纹领域,尚未开展类似的数据扩充研究.现有的研究主要集中在指纹提取和 PAD 方法上.

## 2 基于扩散模型的 OCT 指纹条件生成方法

图 2 展示了本文提出的基于 DDPM 的 OCT 指纹 B 扫描图像生成器的全局流程.直接在  $512 \times 512$  的像素空间中模拟逐步扩散的正向和反向过程,以确保高质量的生成效果,避免采样器扩散模型逆过程侧重于速度而做出的权衡.引入分割掩码作为先验信息,并将其与噪声图像拼接,形成双通道输入,从而建立从分割掩码到噪声的条件映射.

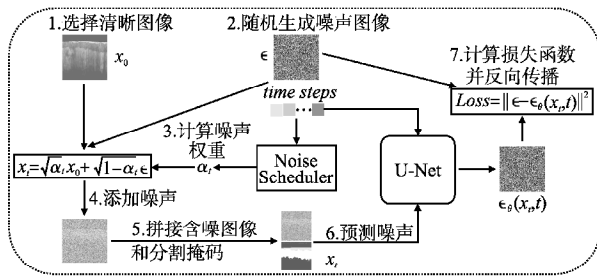


图 2 方法流程图

Fig. 2 Method flow diagram

### 2.1 标准扩散过程

如图 3 所示,扩散模型的核心是通过马尔可夫链模拟正向扩散和逆向去噪过程,其数学表示由公式(1)表示.正向扩散过程  $q$  逐渐向给定的干净图像添加噪声,而逆向去噪过程  $p$  通过神经网络逐步去除噪声,恢复原始图像分布.

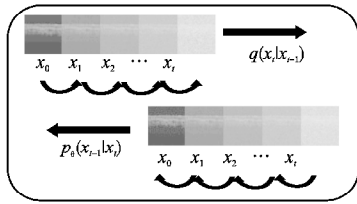


图 3 扩散过程图

Fig. 3 Diffusion process

$$P(x_t | x_{t-1}, x_{t-2}, \dots, x_0) = P(x_t | x_{t-1}) \quad (1)$$

在扩散模型中,噪声水平随时间步变化,公式(2)提供了每个步骤中噪声图像的计算公式,

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon \quad (2)$$

其中,  $\epsilon$  是符合标准高斯分布  $N(0,1)$  的一个噪声采样.

采用 U-Net 作为去噪网络,其多尺度特性能够有效捕捉局部细节与全局上下文信息,适用于渐进式去噪.为了学习所有时间步的去噪能力,在训练过程中会随机选择一个时间步来预测相应噪声.算法 1 详细描述了无条件 DDPM 的训练流程.

#### 算法 1. 无条件 DDPM 的训练过程

重复

$$\begin{aligned} x_0 &\sim q(x_0) \\ t &\sim \text{Uniform}(\{1, \dots, T\}) \end{aligned}$$

$$\epsilon \sim N(0,1)$$

执行梯度下降

$$\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(\sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon, t)\|^2$$

直至收敛

DDPM 采用均方误差 (MSE) 作为损失函数来测量预测噪声与真实噪声之间的误差. MSE 简单有效,能够直接优化噪声预测的准确性,并且在训练过程中表现出良好的稳定性.此外, MSE 损失函数与扩散模型的数学推导一致,确保模型在反向去噪过程中逐渐逼近真实数据分布.

### 2.2 引入分割掩码引导生成

在 OCT 指纹图像生成任务中,由于皮肤组织结构复杂,生成高质量且解剖学上合理的 OCT 指纹是一个挑战. ControlNet 引入了额外的条件输入(如边缘图和姿态图),并结合了一个独立的约束网络分支来精确控制生成过程.受 ControlNet 的启发,本文提出将分割掩码作为条件信息整合到扩散模型中.这些分割掩码在横向方向上提供了重要的空间结构信息,而对解剖结构施加的显式约束有效地指导了生成过程.这使得模型能够生成与目标解剖结构高度一致的 OCT 图像,确保每层的厚度、形态和边界清晰且逼真.

将给定的 OCT 指纹图像对应的分割掩码与噪声图像沿着通道维度进行拼接,形成一个 shape 为  $(\text{batch\_size}, 2, \text{img\_size}, \text{img\_size})$  的双通道张量.在去噪过程中同时处理噪声图像和分割掩码,从而生成更符合指尖皮肤真实结构的 OCT 指纹.通过这种方法,模型有效地学习到从分割掩码到噪声的条件映射,从而能够生成具有一致解剖结构的高质量 OCT 指纹切片.

### 2.3 方法实施细节

在扩散模型中,无论是训练过程还是采样过程,噪声都是一个关键要素.整个过程的第一步是生成一个符合标准高斯分布的噪声图像.随后,基于预先设定的总步数,在一个预定义的时间间隔内随机选择一个时间步.然后,利用公式(2)计算出当前时间步对应的噪声图像.为了实现更高质量的生成效果,DDPM 通常会使用相对较多的时间步,通常从几百到几千不等.随机选择时间步的训练策略确保了模型能够在所有时间步上有效地执行去噪操作.

噪声强度的过快增长会使训练过程过于具有挑战性,而噪声强度的缓慢增长会导致效率降低.这就需要不同的调度器来计算每个步骤  $t$  的噪声强度.常用的调度器包括线性调度器,它在每一步都有固定的增量,以及余弦调度器,它最初提供缓慢的增长,然后在接近结束时迅速增加.

本文方法采用线性调度器.相比余弦调度器,能够提供一个非常稳定的训练过程.将步数设置为 1000,不会出现收敛问题,并且能够学到足够的分布信息.按照算法 2 中概述的步骤训练 100 个 epoch.

#### 算法 2. 条件生成方法训练过程

重复

$$\begin{aligned} x_0 &\sim q(x_0) \\ t &\sim \text{Uniform}(\{1, \dots, T\}) \\ \epsilon &\sim N(0,1) \\ x_t &= \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon \\ m &\sim q(m | x_0) \\ \text{input} &= [x_t, m] \end{aligned}$$

执行梯度下降  $\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(input, t)\|^2$

直至收敛

与 OCT 指纹图像对应的分割掩码随后与噪声图像沿着通道维度进行拼接, 形成一个 shape 为 (batch\_size, 2, img\_size, img\_size) 的双通道张量. 然后, 将这个张量以及与噪声图像对应的时间步  $t$  输入到一个 2D U-Net 中进行预测. 不对分割掩码添加噪声的原因如下: 分割掩码图像由手动标注软件导出, 是灰度图格式, 且仅包含 3 种像素值—0、1 和 2. 大部分黑色背景区域的像素值为 0. 因此, 添加噪声并不会引起显著差异.

### 3 实验与分析

基于 OCT 指纹的实际应用需求, 当前研究主要聚焦于 PAD 和层分割任务, 二者共同致力于提升对伪造指纹攻击的防御性能. 首先进行主观评价测试, 然后在 PAD 和层分割下游任务上进行实验, 使用代表性指标评价生成的 OCT 指纹质量.

在 PAD 任务中, AUC 通过计算接收者工作特征曲线下面积来评估模型区分阳性和阴性样本的能力, AUC 越大表示性能越好. EER 表示错误接受率等于错误拒绝率的点, EER 越低性能越好. 这两个指标均从模型的二进制输出 (真或假) 和基于决策阈值计算的置信度分数中派生出来.

为了验证方法的有效性, 将其与来自不同生成框架的代表性模型进行比较:

1) 基于 VAE 的方法, 包括标准 VAE<sup>[19]</sup>、VQVAE<sup>[37]</sup> 和 VQVAE2<sup>[38]</sup>, 以潜空间编码和重建能力而闻名;

2) 基于 GAN 的方法, 包括 pix2pix<sup>[39]</sup> 和 StyleGAN3<sup>[40]</sup>, 擅长于高质量的图像合成场景, 其中 pix2pix 是专门为成对数据生成任务设计的;

3) 基于扩散模型的方法, 包括 DDPM<sup>[17]</sup> 和 BBDM<sup>[41]</sup>, 利用迭代噪声预测来产生高保真结果, 其中 BBDM 针对成对数据生成任务专门设计了扩散逻辑. 通过比较这些方法, 全面展示本文方法在生成具有准确像素分布和层结构的逼真 OCT 指纹方面的优势.

#### 3.1 主观评价实验

通过观察图 4 中列出的各种方法生成的 OCT 指纹, 可以明显看出, 与本文方法相比, 其它生成方法很难完全复制真实的数据分布, 具有一定的局限性. 相比之下, 本文方法可以生成高度真实的清晰指纹.

基于 AE 的方法基本不能生成清晰的图像, 相反, 它们生成的基本上是水平的模糊像素带. 基于 GAN 的方法则不能看出明显的层边界, 比如 StyleGAN3 表现出了明显的团状模糊. 这是 StyleGAN 系列模型的经典问题, 即使在 ImageNet 这样色彩丰富的数据集上训练也无法避免.

基于扩散模型的方法能够生成较为清晰的 OCT 指纹, 层与层之间也具有明显的分界线. 然而, DDPM 存在重复和断裂的经典问题, 代表样例如图 5 所示. 进行了 5 次独立训练, 并统计了 10000 个生成样本内的异常情况. 统计结果表明, 异常率在 10% ~ 20% 之间波动. 实验表明, 增加训练轮数可以将异常率降低到 10% 以下. 然而, 即使将 epoch 轮数增加到 1000 轮, 导致明显的过拟合, 仍然无法彻底避免. 几种 pair-to-

pair 的条件生成方法有效缓解这一问题. 然而, 与其它方法生成的图像相比, 本文方法生成的指纹具有更多的层内细节, 甚至能精确渲染汗腺这类微观皮肤结构.

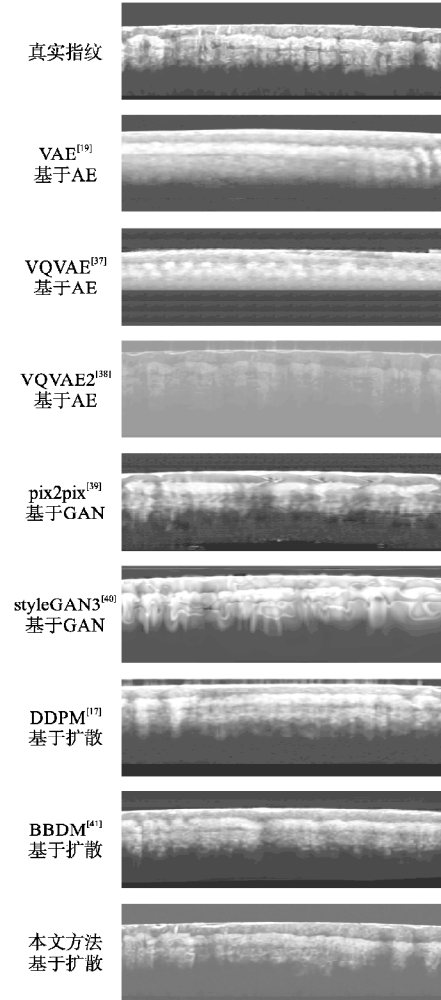


图 4 不同方法的生成结果

Fig. 4 Results generated by different methods

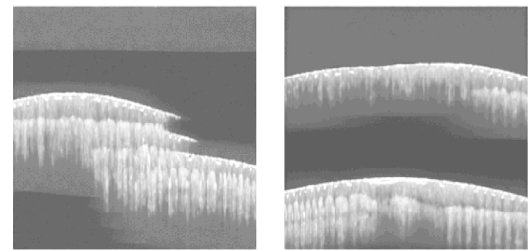


图 5 DDPM 生成的异常情况

Fig. 5 Anomalies in DDPM generated results

为了验证上述关于生成的 OCT 指纹图像清晰度的观点, 邀请了 30 位领域专家对 1 幅真实图像和 8 种方法生成的图像进行 5 轮主观评估 (每轮对 5 组图像按真实程度 1~8 分排序). 表 1 中统计的 150 组评分结果 (均值  $\pm$  标准差) 显示本文方法生成的图像在真实感评分上最接近真实样本. ANOVA 检验和 Kruskal-Wallis H 检验的 p 值远低于严格的置信水平 0.01, 可以证明这些数据是真实可信的.

研究人员指出,在评估 OCT 指纹时,首要关注的是像素分布.本文方法生成的 OCT 指纹具有清晰而明确的边界,这代表了更精确的皮肤结构表示.考虑的第 2 个方面是单个层内的像素分布细节.本文方法生成的 OCT 指纹在这些区域具有更丰富的细节,强度变化与真实指纹的分布即为接近.此外,汗腺等精细结构的存在,也为生成的指纹提升了可信度.这些优点使得受试人员倾向于给本文方法生成的 OCT 指纹更高的真实度分数.

表 1 生成结果真实度打分对比

Table 1 Realism scores comparison for generated results

生成方法	评分均值	标准差
VAE <sup>[19]</sup>	4.53	1.34
VQVAE <sup>[37]</sup>	3.17	1.35
VQVAE2 <sup>[38]</sup>	1.20	0.40
pix2pix <sup>[39]</sup>	4.43	0.85
styleGAN3 <sup>[40]</sup>	3.40	2.81
DDPM <sup>[17]</sup>	5.76	1.56
BBDM <sup>[41]</sup>	6.23	0.92
本文方法	<b>7.20</b>	0.63

### 3.2 表示攻击检测

在本研究中使用的 PAD 方法均为仅依赖真实样本的无监督模型,包括 AE<sup>[42]</sup>, VAE<sup>[19]</sup>, MAE<sup>[43]</sup>, OCPAD<sup>[44]</sup>, PPAD<sup>[45]</sup> 和 UOCFR<sup>[46]</sup>. 在训练完成之后,在相同的测试集上计算 PAD 任务的关键评价指标并进行防伪模型之间的横向比较.由于带分割标签的真实样本数据集数量仅有 1000 张左右,以 1000 个真实样本作为基线.在此基础上,依次将生成的 500、1000、1500、2000 张图像添加到训练集中.

表 2 PAD 模型指标对比

Table 2 Metrics comparison of PAD models

PAD 方法	r1	r1g0.5	r1g1	r1g1.5	r1g2
<b>AUC</b>					
AE <sup>[42]</sup>	0.834	0.857	0.875	0.883	0.892
VAE <sup>[19]</sup>	0.870	0.907	0.916	0.921	0.931
MAE <sup>[43]</sup>	0.893	0.916	0.925	0.930	0.936
OCPAD <sup>[44]</sup>	0.865	0.886	0.893	0.898	0.906
PPAD <sup>[45]</sup>	0.916	0.944	0.959	0.969	0.970
UOCFR <sup>[46]</sup>	0.901	0.923	0.943	0.945	0.949
<b>EER</b>					
AE <sup>[42]</sup>	0.239	0.201	0.184	0.183	0.181
VAE <sup>[19]</sup>	0.201	0.155	0.136	0.095	0.049
MAE <sup>[43]</sup>	0.156	0.143	0.140	0.130	0.119
OCPAD <sup>[44]</sup>	0.218	0.187	0.179	0.171	0.162
PPAD <sup>[45]</sup>	0.166	0.113	0.089	0.082	0.071
UOCFR <sup>[46]</sup>	0.200	0.195	0.173	0.163	0.143

表 2 展示了在不同训练集配置下 PAD 模型的性能表现,其中"r1"代表 1000 个真实 OCT B-scan 切片,"r1g0.5"~"r1g2"表示在此基础上分别添加 500~2000 个生成指纹.经过 5 次独立实验验证,随着生成图像数量从 500 增至 2000,所有 PAD 模型的 AUC 值平均提升达到 5.79%,EER 值至少降低 23.72%.这一稳定上升趋势充分证明了本文生成方法能有效增强数据集质量,所生成指纹与真实数据分布高度一致,显著提升了模型的泛化能力和鲁棒性.

表 3 数据表明,在增强各种 PAD 模型的性能方面,本文方法始终优于其它生成方法,能够使所有 PAD 模型在增加不同生成样本数量的情况下,AUC 显著改善.

表 3 不同生成模型对 PAD 性能的提升效果

Table 3 Performance improvement of PAD models with different generative methods

生成方法	r1	r1g0.5	r1g1	r1g1.5	r1g2
<b>PAD method:AE</b>					
VAE <sup>[19]</sup>	0.834	0.821	0.820	0.818	0.812
VQVAE <sup>[37]</sup>	0.834	0.828	0.819	0.813	0.808
VQVAE2 <sup>[38]</sup>	0.834	0.836	0.844	0.845	0.847
pix2pix <sup>[39]</sup>	0.834	0.848	0.859	0.866	0.871
styleGAN3 <sup>[40]</sup>	0.834	0.841	0.852	0.858	0.863
DDPM <sup>[17]</sup>	0.834	0.846	0.860	0.868	0.881
BBDM <sup>[41]</sup>	0.834	0.848	0.857	0.869	0.883
本文方法	0.834	0.857	0.875	0.883	0.892
<b>PAD method:VAE</b>					
VAE <sup>[19]</sup>	0.870	0.868	0.868	0.849	0.851
VQVAE <sup>[37]</sup>	0.870	0.861	0.858	0.855	0.849
VQVAE2 <sup>[38]</sup>	0.870	0.870	0.867	0.866	0.863
pix2pix <sup>[39]</sup>	0.870	0.880	0.886	0.894	0.894
styleGAN3 <sup>[40]</sup>	0.870	0.872	0.888	0.885	0.888
DDPM <sup>[17]</sup>	0.870	0.887	0.892	0.893	0.903
BBDM <sup>[41]</sup>	0.870	0.890	0.893	0.898	0.914
本文方法	0.870	0.907	0.916	0.921	0.931
<b>PAD method:MAE</b>					
VAE <sup>[19]</sup>	0.893	0.889	0.889	0.882	0.888
VQVAE <sup>[37]</sup>	0.893	0.897	0.889	0.875	0.873
VQVAE2 <sup>[38]</sup>	0.893	0.893	0.898	0.898	0.895
pix2pix <sup>[39]</sup>	0.893	0.910	0.910	0.913	0.910
styleGAN3 <sup>[40]</sup>	0.893	0.892	0.891	0.904	0.904
DDPM <sup>[17]</sup>	0.893	0.902	0.914	0.916	0.922
BBDM <sup>[41]</sup>	0.893	0.913	0.919	0.923	0.925
本文方法	0.893	0.916	0.925	0.930	0.936
<b>PAD method:OCPAD</b>					
VAE <sup>[19]</sup>	0.865	0.812	0.806	0.807	0.800
VQVAE <sup>[37]</sup>	0.865	0.833	0.822	0.823	0.813
VQVAE2 <sup>[38]</sup>	0.865	0.847	0.839	0.838	0.825
pix2pix <sup>[39]</sup>	0.865	0.866	0.875	0.885	0.890
styleGAN3 <sup>[40]</sup>	0.865	0.854	0.864	0.871	0.874
DDPM <sup>[17]</sup>	0.865	0.872	0.884	0.886	0.897
BBDM <sup>[41]</sup>	0.865	0.878	0.888	0.889	0.895
本文方法	0.865	0.886	0.893	0.898	0.906
<b>PAD method:PPAD</b>					
VAE <sup>[19]</sup>	0.916	0.909	0.907	0.904	0.904
VQVAE <sup>[37]</sup>	0.916	0.911	0.906	0.902	0.903
VQVAE2 <sup>[38]</sup>	0.916	0.922	0.920	0.907	0.906
pix2pix <sup>[39]</sup>	0.916	0.916	0.919	0.928	0.937
styleGAN3 <sup>[40]</sup>	0.916	0.913	0.912	0.924	0.926
DDPM <sup>[17]</sup>	0.916	0.920	0.926	0.935	0.946
BBDM <sup>[41]</sup>	0.916	0.932	0.937	0.948	0.953
本文方法	0.916	0.944	0.959	0.969	0.970

实验数据充分验证了不同生成方法在 OCT 指纹任务中的性能差异:在训练集中增加 2000 张生成图像,基于 VAE 的方法因模型容量有限导致生成图像模糊,使 PAD 性能平均下降 2.84%;基于 GAN 的方法虽能部分保留皮肤结构,但训练不稳定性限制了其性能提升(如 styleGAN3 在 PPAD 任务中 AUC 仅能达到 0.926,而基于扩散的方法中效果最低的 DDPM 也能提升到 0.946,本文方法更是能达到 0.970)。

相比之下,传统扩散方法(DDPM/BBDM)在最终扩充到2000张时能将AUC提升3.2~5.8%。而本文提出的条件扩散方法通过引入分割掩码先验知识,在关键层状结构区域(角质层-活表皮连接处)实现了最优生成质量,将性能优势进一步扩大,提升幅度为4.7~7.0%。特别是在结构敏感的OCPAD任务中,本文方法使得AUC(0.906)较扩散模型和GAN方法中表现最好的DDPM(0.897)和pix2pix(0.890)分别提升1%和1.8%;在PPAD任务中AUC更达到0.970的卓越性能。值得注意的是,本文方法在MAE任务中对层间明暗变化的建模能力使5种PAD模型上的平均AUC提升幅度为5.88%,达到传统扩散模型DDPM(3.92%)的1.5倍和BBDM(4.38%)的1.34倍;5项任务最高AUC提升7.0%,较传统扩散模型DDPM(5.6%)和BBDM(5.8%)分别额外获得1.4%和1.2%的性能增益,充分验证了结构条件引导对提升OCT指纹生成质量的有效性。

本节实验证明,本文方法将分割掩码作为条件信息,通过引入分割掩码,本文方法可以更准确地学习真实数据的分布,特别是在层与层之间像素值明暗变化极为强烈的关键区域,从而进一步提高生成图像的质量和一致性。

### 3.3 指纹层分割

OCT体数据对所有B-scan图像进行层分割可重建出完整的二维指纹图像,其中内指纹因深嵌皮下具有更强的防伪性和活体检测价值。为评估不同生成方法对分割性能的影响,基于OCT指纹层分割模型GPGT-Net<sup>[47]</sup>,对比pix2pix、BBDM和本文方法在相同训练集规模下的表现。

在OCT指纹图像层分割任务中,长达1800像素的不规则边界对生成方法的细粒度控制能力提出了极其严苛的要求。A层对应指纹采集的玻璃层,曲率相对规则。由于玻璃层不反射光,呈现为黑色背景,该边界很容易被检测出来。B层是表皮层的上边界。当手指按压时,它没有完全粘附在玻璃层上,导致它们之间出现一个空腔区域。皮肤组织的差异导致两侧像素值差异较大,局部对比度水平明显提高。C层是角质层和活性表皮层之间的边界,起伏程度大,难以精确定位。使用平均绝对误差(MAE)评价分割结果。

表4定量分析表明,本文提出的条件扩散方法在3个关键边界的分割任务中均展现出显著优势。当使用相同数量级的生成数据扩充训练集时,本节方法在同一测试集上取得的MAE指标显著低于pix2pix和BBDM。特别是随着生成数据量的逐渐增加,由于生成质量的限制,pix2pix和BBDM在提高分割模型性能方面面临瓶颈甚至倒退。相比之下,本节方法确保了持续的性能改进,稳定地将每层的MAE值降低到与基线组相比相对较低的水平。

对于易检测的A层,3种方法性能相近,但本文方法在数据扩充至2000张时取得最优MAE(0.689),较pix2pix(0.697)和BBDM(0.713)分别降低1.14%和3.37%。在B层分割中,本文方法MAE(1.123)相比pix2pix(1.163)和BBDM(1.151)分别提升3.44%和2.43%,这得益于对空腔区域对比度的精确建模。对于最具挑战性的C层分割结果,本文方法MAE(1.775)显著优于BBDM(1.783)和pix2pix(1.896)。本文方法通过分割掩码条件引导,能准确地学习表

表4 分割模型MAE对比

Table 4 MAE comparison of segmentation models

生成方法	数据集组成	A层	B层	C层
pix2pix <sup>[39]</sup>	r1	0.767	1.198	1.948
	r1g0.5	0.763	1.197	1.941
	r1g1	0.741	1.191	1.939
	r1g1.5	0.708	1.178	1.907
	r1g2	0.697	1.163	1.896
BBDM <sup>[41]</sup>	r1	0.767	1.198	1.948
	r1g0.5	0.761	1.196	1.938
	r1g1	0.738	1.188	1.916
	r1g1.5	0.719	1.179	1.799
	r1g2	0.713	1.151	1.783
本文方法	r1	0.767	1.198	1.948
	r1g0.5	0.757	1.193	1.945
	r1g1	0.729	1.178	1.898
	r1g1.5	0.691	1.129	1.785
	r1g2	0.689	1.123	1.775

皮层中复杂的层间结构变化,特别是在C层这类曲率变化剧烈的区域。实验数据证明,随着生成数据量的增加,本文方法在所有边界的分割精度提升幅度均持续领先,验证了其在OCT指纹层分割任务中的显著优势。实验结果证明了本文方法在捕捉指尖皮肤结构信息方面表现优异,极高的整图生成质量和精确的像素级控制能力有效提升了分割模型的性能。

## 4 总结与展望

针对OCT指纹研究中的核心挑战—数据集规模受限以及分割标注缺失,本文提出了基于扩散模型的OCT指纹条件生成框架,训练后可以得到一个高度逼真的条件生成模型。该框架将层分割掩码作为条件变量,通过双通道U-Net架构实现以解剖结构为引导的精细生成,充分利用指尖皮肤结构的大量先验知识。该生成器能够以分割掩码为指导条件,在逐步扩散和去噪过程中生成包含真实指尖皮肤层次结构的高分辨率OCT指纹。在分割掩码的指导下,本文方法有效地学习到足够的指尖皮肤结构先验知识,生成的数据能够增强PAD和层分割下游任务中相应模型的性能。

## References:

- [1] Cappelli R, Ferrara M, Maltoni D. Minutia cylinder-code: a new representation and matching technique for fingerprint recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(12): 2128-2141.
- [2] Das A, Galdi C, Han H, et al. Recent advances in biometric technology for mobile devices [C]//Proceedings of IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2018: 1-11.
- [3] Khan T M, Bailey D G, Khan M A U, et al. Efficient hardware implementation for fingerprint image enhancement using anisotropic gaussian filter [J]. IEEE Transactions on Image Processing, 2017, 26(5): 2116-2126.
- [4] Chugh T, Jain A K. Oct fingerprints: resilience to presentation attacks [EB/OL]. https://arxiv.org/abs/1908.00102, 2019.
- [5] Goicoechea Telleria I, Fernandez Saavedra B, Liu Jimenez J, et al. An evaluation of presentation attack detection of fingerprint biometric systems applying iso/iec 30107-3 [C]//Proceedings of 4th International Biometric Performance Testing Conference (IBPC), Gaithersburg, 2016.

- [6] Information technology—biometric presentation attack detection—part 1; framework [M]. Geneva: International Standards Organization, 2016.
- [7] Sousedik C, Busch C. Presentation attack detection methods for fingerprint recognition systems: a survey [J]. *IET Biometrics*, 2014, 3(4): 219-233.
- [8] Liu F, Liu G, Wang X. High-accurate and robust fingerprint anti-spoofing system using optical coherence tomography [J]. *Expert Systems with Applications*, 2019, 130(9): 31-44.
- [9] Bontrager P, Roy A, Togelius J, et al. Deepmasterprints: generating masterprints for dictionary attacks via latent variable evolution [C]//Proceedings of IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2018: 1-9.
- [10] Liu F, Zeng W, Li Y, et al. A lightweight and noise-robust method for internal OCT fingerprint reconstruction [J]. *IEEE Transactions on Information Forensics and Security*, 2024, 19: 5492-5505.
- [11] Sun H, Wang H, Zhang Y, et al. Zjut-eifd: a synchronously collected external and internal fingerprint database [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 46(4): 2267-2284.
- [12] Liu F, Shen C, Liu H, et al. A flexible touch-based fingerprint acquisition device and a benchmark database using optical coherence tomography [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(9): 6518-6529.
- [13] Zhu Y, Yin X, Hu J. Fingergan: a constrained fingerprint generation scheme for latent fingerprint enhancement [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(7): 8358-8371.
- [14] Engelsma J J, Grosz S A, Jain A K. Printsgan: synthetic fingerprint generator [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(5): 6111-6124.
- [15] Grosz S A, Jain A. Spoofgan: synthetic fingerprint spoof images [J]. *IEEE Transactions on Information Forensics and Security*, 2022, 18: 730-743.
- [16] Sohl Dickstein J N, Weiss E A, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics [EB/OL]. <https://arxiv.org/abs/1503.03585>, 2015.
- [17] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models [EB/OL]. <https://arxiv.org/abs/2006.11239>, 2020.
- [18] Wu Y N, Gao R, Han T, et al. A tale of three probabilistic families: discriminative, descriptive and generative models [EB/OL]. <https://arxiv.org/abs/1810.04261>, 2018.
- [19] Kingma D P, Welling M. Auto-encoding variational bayes [EB/OL]. <https://arxiv.org/abs/1312.6114>, 2022.
- [20] Goodfellow I J, Pouget Abadie J, Mirza M, et al. Generative adversarial networks [J]. *Communications of the ACM*, 2014, 63(11): 139-144.
- [21] Dinh L, Sohl Dickstein J N, Bengio S. Density estimation using real nvp [EB/OL]. <https://arxiv.org/abs/1605.08803>, 2016.
- [22] Kingma D P, Dhariwal P. Glow: generative flow with invertible 1x1 convolutions [EB/OL]. <https://arxiv.org/abs/1807.03039>, 2018.
- [23] Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis [EB/OL]. <https://arxiv.org/abs/2105.05233>, 2021.
- [24] Kazerouni A, Aghdam E K, Heidari M, et al. Diffusion models in medical imaging: comprehensive survey [J]. *Medical Image Analysis*, 2023, 88: 102846.
- [25] Han C, Hayashi H, Rundo L, et al. Gan-based synthetic brain mr image generatio [C]//Proceedings of 15th IEEE International Symposium on Biomedical Imaging (ISBI), 2018: 734-738.
- [26] Yang Q, Yan P, Zhang Y, et al. Low dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss [J]. *IEEE Transactions on Medical Imaging*, 2017, 37(6): 1348-1357.
- [27] Song J, Meng C, Ermon S. Denoising diffusion implicit models [EB/OL]. <https://arxiv.org/abs/2010.02502>, 2020.
- [28] Song Y, Sohl Dickstein J N, Kingma D P, et al. Score-based generative modeling through stochastic differential equations [EB/OL]. <https://arxiv.org/abs/2011.13456>, 2020.
- [29] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 10674-10685.
- [30] Cappelli R, Maio D, Maltoni D, et al. Synthetic fingerprint-image generation [C]//Proceedings of 15th International Conference on Pattern Recognition (ICPR), 2000: 471-474.
- [31] Zhao Q, Jain A K, Paultier N G, et al. Fingerprint image synthesis based on statistical feature models [C]//Proceedings of IEEE 5th International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2012: 23-30.
- [32] Johnson P A, Hua F, Schuckers S. Texture modeling for synthetic fingerprint generation [C]//IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2013: 154-159.
- [33] Kirchgasser S, Kauba C, Uhl A. Assessment of synthetically generated mated samples from single fingerprint samples instances [C]//Proceedings of IEEE International Workshop on Information Forensics and Security (WIFS), 2021: 1-6.
- [34] Bouzaglo R, Keller Y. Synthesis and reconstruction of fingerprints using generative adversarial networks [EB/OL]. <https://arxiv.org/abs/2201.06164>, 2022.
- [35] Bahmani K, Plesh R, Johnson P A, et al. High fidelity fingerprint generation: quality, uniqueness, and privacy [C]//Proceedings of IEEE International Conference on Image Processing (ICIP), 2021: 3018-3022.
- [36] Wyzkowski A B V, Segundo M P, de P Lemes R. Level three synthetic fingerprint generation [C]//Proceedings of 25th International Conference on Pattern Recognition (ICPR), 2020: 9250-9257.
- [37] van den Oord A, Vinyals O, Kavukcuoglu K. Neural discrete representation learning [C]//Proceedings of 31st International Conference on Neural Information Processing Systems (NIPS), 2017: 6309-6318.
- [38] Razavi A, Van den Oord A, Vinyals O. Generating diverse high-fidelity images with vq-vae-2 [C]//Proceedings of 33rd International Conference on Neural Information Processing Systems (NeurIPS), 2019: 14866-14876.
- [39] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5967-5976.
- [40] Karras T, Aittala M, Laine S, et al. Alias-free generative adversarial networks [C]//Proceedings of the 35th International Conference on Neural Information Processing Systems (NeurIPS), 2021: 852-863.
- [41] Li B, Xue K, Liu B, et al. BBDM: image-to-image translation with brownian bridge diffusion models [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 1952-1961.
- [42] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323: 533-536.
- [43] Gong D, Liu L, Le V, et al. Memorizing normality to detect anomaly: memory-augmented deep autoencoder for unsupervised anomaly detection [C]//Proceedings of IEEE International Conference on Computer Vision (ICCV), 2019: 1705-1714.
- [44] Liu F, Liu H, Zhang W, et al. One-class fingerprint presentation attack detection using auto-encoder network [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2394-2407.
- [45] Liu Y P, Zuo W, Liang R, et al. Prototype-guided autoencoder for OCT-based fingerprint presentation attack detection [J]. *IEEE Transactions on Information Forensics and Security*, 2023, 18: 3461-3475.
- [46] Zhang W, Liu H, Liu F, et al. A uniform representation model for oct-based fingerprint presentation attack detection and reconstruction [J]. *Pattern Recognition*, 2024, 145: 109981.
- [47] Liu Y P, Zhong Q, Liang R, et al. Layer segmentation of oct fingerprints with an adaptive Gaussian prior guided transformer [J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-15.