arXiv:2304.01992v2 [eess.IV] 4 Jul 2023

用于结直肠组织分类的跨调制小样本图像生成

AmandeepKumar[1], Ankan kumar Bhunia[1], Sanath Narayan[3], Hisham Cholakkal[1], Rao Muhammad Anwer[1,2], Jorma Laaksonen[2], 和 Fahad Shahbaz Khan[1]

[1]MBZUAI，阿联酋； [2]阿尔托大学，芬兰；[3] 技术创新研究院，阿联酋
amandeep.kumar@mbzuai.ac.ae

摘要。 在本研究中，我们提出了一种小样本结直肠组织图像生成方法，用于解决罕见癌症组织组织病理学训练数据稀缺的问题。我们的小样本生成方法名为 XM-GAN，以一张基准图像和一对参考组织图像作为输入，生成高质量且多样化的图像。在 XM-GAN 中，我们提出了一种新颖的可控融合模块，根据参考图像局部区域与基准图像中对应区域的相似性，对参考图像的局部区域进行密集聚合，从而获得局部一致特征。据我们所知，我们是首个研究结直肠组织图像小样本生成的团队。我们通过广泛的定性、定量以及基于专业人员（病理学家）的评估，对小样本结直肠组织图像生成进行了评估。具体而言，在基于专业人员的评估中，病理学家仅在 55% 的情况下能够区分 XM-GAN 生成的组织图像与真实图像。此外，我们利用这些生成图像进行数据增强，以解决小样本组织图像分类任务，相较于基础小样本分类器，平均准确率提升了 4.4%。代码：https://github.com/VIROBO-15/ XM-GAN。

关键词： 少样本图像生成 · 交叉调制 .

1 引言

组织病理学图像分析是癌症诊断的重要环节。然而，全球病理学家短缺，加上组织病理学数据的复杂性，使这项任务既耗时又充满挑战。因此，近年来，利用深度学习的最新进展开发自动且准确的组织病理学图像分析方法受到了广泛关注。在本研究中，我们探讨结直肠癌的诊断问题。结直肠癌是全球最常见的癌症死亡原因之一，在欧洲和美洲尤为如此 [26]。

现有基于深度学习的结直肠组织分类方法 [24,19,25]通常需要大量标注的组织病理学训练数据，才能对

所有组织类型进行分类。然而，获取大量训练数据颇具挑战，尤其是对于罕见癌症组织。为此，我们希望开发一种结直肠组织少样本分类方法，它能够从拥有充足训练数据的已见组织类别中学习，并将这些知识迁移到仅有未见 （新颖）组织类别上，这些类别只有少量 示例训练图像。

虽然生成对抗网络（GAN） [6]已被用于合成图像，但通常需要使用大量相应类别的真实图像进行训练，这在上述少样本设置中并不可行。因此，我们提出一种少样本图像生成方法，利用有限的样本生成高质量且多样化的新类别结直肠组织图像。此外，我们展示了这些生成图像在具有挑战性的少样本结直肠组织分类任务中的适用性。

贡献：我们提出了一种少样本结直肠组织图像生成框架，称为 XM-GAN，旨在同时生成高质量且多样化的图像。在该组织图像生成框架中，我们引入了一种新颖的可控融合模块（CFB），根据参考组织图像的局部区域与基础组织图像中对应区域的一致性，对这些局部区域进行密集聚合。CFB 采用交叉注意力机制，在基础组织图像特征（查询）与参考组织图像特征（键、值）之间进行特征聚合。这种交叉注意力机制能够从全局感受野聚合参考特征，从而得到局部一致的特征。因此，生成的结直肠组织图像伪影更少。

为进一步提升生成组织图像的多样性和质量，我们在 CFB 中引入了映射网络和可控交叉调制层归一化（cLN）。映射网络根据参考组织图像的全局特征和控制参数生成“元权重”。随后，利用这些元权重计算 cLN 中用于特征重新加权的调制权重。这样，便可依据参考组织图像特征及相关控制参数，以可控方式对交叉注意力处理后的组织图像特征重新加权并加以丰富。因此，我们基于 Transformer 的框架生成的组织图像多样性得以提升（见图 3）。

我们通过广泛的定性、定量评估以及基于专业人士（病理学家）的评估，验证了 XM-GAN 在结直肠组织少样本图像生成任务中的表现。XM-GAN 能生成逼真且多样的结直肠组织图像（见图 3）。在基于专业人士（病理学家）的评估中，病理学家仅在 55% 的情况下能够区分 XM-GAN 生成的结直肠组织图像与真实图像。此外，我们在训练结直肠组织少样本图像分类器时，将生成的组织图像用作数据增强，以评估其有效性；与未经改进的少样本分类器相比，平均分类准确率绝对提高了 4.4%。

2 相关工作

生成模型 [6,15]和 Transformer [21,20] 拟合各种数据分布的能力，推动了图像生成 [12,13,3,22,16],等任务取得长足进展。尽管取得了成功，这些生成模型通常仍需要大量数据进行训练，以避免过拟合。相比之下，少样本图像生成方法 [4,17,2,9,7]致力于从与训练数据同属一个领域、但类别互不重叠的新类别中生成自然图像。现有的自然图像少样本生成方法，根据变换 [1], 优化 [4,17] 和融合 [9,10,7], 大致可分为三类。基于变换的方法学习执行通用的数据增强，从单张条件图像生成类内图像。另一方面，基于优化的方法通常利用元学习技术，通过在新领域的少量参考图像上进行优化，适应不同的图像生成任务。与这两种更适用于简单图像生成任务的范式不同，基于融合的方法首先聚合参考图像的潜在特征，然后利用解码器根据这些聚合特征生成同一类别的图像。

我们的方法： 虽然上述研究探讨了自然图像的少样本生成，但据我们所知，我们是首个研究结直肠组织图像少样本生成的工作。在本研究中，我们探讨多类别结直肠组织分析问题，其中类别包含低级别和高级别肿瘤。本研究使用的相应数据集 [14]广泛用于结直肠癌组织学中的多类别纹理分类，包含八种组织类型：肿瘤上皮、简单间质、复杂间质、免疫细胞、碎屑、正常黏膜腺体、脂肪组织和背景（无组织）。生成这些不同类别的结直肠组织图像是一项具有挑战性的任务，尤其是在少样本设置下。生成逼真且多样化的组织图像需要确保全局与局部纹理一致性（模式）。我们的 XM-GAN 在全局感受野范围内，密集聚合来自参考组织图像所有相关局部区域的特征 [23,5], 并利用根据输入参考组织图像特征计算得到的元权重，采用可控机制调制组织图像特征。因此，即使在少样本设置下，也能生成高质量且多样化的结直肠组织图像。

3 方法

问题定义： 在我们的小样本结直肠组织图像生成框架中，目标是根据 K 个输入样本 X 生成一组多样化图像，这些图像来自未见（新颖）的组织类别。设 $\mathcal{D}^s$ 和 $\mathcal{D}^u$ 分别为已见类别和未见类别的集合，其中 $\mathcal{D}^s \cap \mathcal{D}^u = \emptyset$。在训练阶段，我们从 $\mathcal{D}^s$ 中采样图像，并训练模型学习可迁移的生成能力，以便为未见类别生成新的组织图像。在推理阶段，给定来自 $\mathcal{D}^u$中某个未见类别的 K 张图像，训练好的模型无需进一步微调，即可努力为该未见类别生成多样且合理的图像。

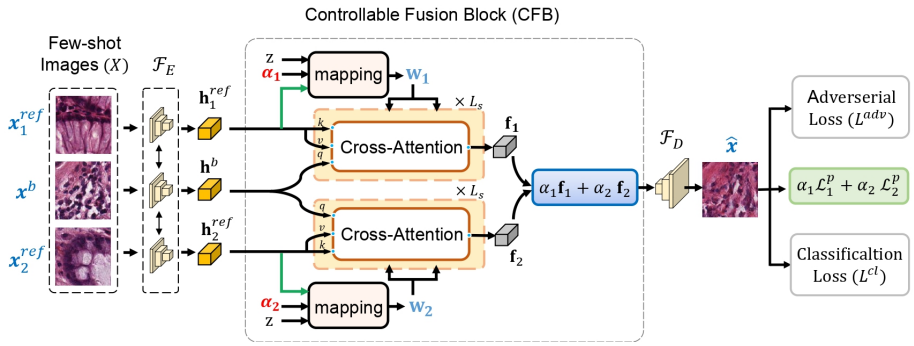


图1：我们的 XM-GAN 由一个 CNN 编码器、一个基于 Transformer 的可控融合模块（CFB）和一个用于组织图像生成的 CNN 解码器组成。对于K-shot 设置，共享编码器 $\mathcal{F}_E$ 接收一张基础组织图像 x^b 和 $K-1$ 张参考组织图像 $\{x_i^{ref}\}_{i=1}^{K-1}$，并分别输出视觉特征 $\mathbf{h}^b$ 和 $\{\mathbf{h}_i^{ref}\}_{i=1}^{K-1}$。在 CFB 中，映射网络计算元权重 $\mathbf{w}_i$，用于生成调制权重，以便在交叉注意力过程中对特征进行重新加权。交叉注意力特征 $\mathbf{f}_i$ 经融合后输入解码器 $\mathcal{F}_D$，生成图像 $\hat{x}$。

整体架构： 图 1 展示了我们提出的 XM-GAN 框架的整体架构。在这里，我们从 X 中随机选择一张组织图像作为基准图像 x^b，并将其余 $K-1$ 张组织图像记为参考图像 $\{x_i^{ref}\}_{i=1}^{K-1}$。给定输入图像 X，我们将其输入共享编码器 $\mathcal{F}_E$，从而获得基准组织图像和每张参考组织图像的特征表示。接下来，将编码后的特征表示$\mathbf{h}$ 输入可控融合块（CFB），在基准特征$\mathbf{h}^b$ 与参考特征$\mathbf{h}_i^{ref}$ 之间执行交叉注意力 [23]。在 CFB 中，我们引入映射网络以及可控交叉调制层归一化（cLN），以计算元权重$\mathbf{w}_i$，随后利用这些元权重生成调制权重，用于在 cLN 中进行重新加权。得到的融合表示$\mathbf{f}$ 被输入解码器 $\mathcal{F}_D$，以生成组织图像 $\hat{x}$。整个框架遵循生成对抗网络（GAN）范式进行训练。除了$\mathcal{L}^{adv}$ 和 $\mathcal{L}^{cl}$之外，我们还提出使用引导感知损失项 $\mathcal{L}^p$，并利用控制参数 α_i。接下来，我们将详细介绍 CFB。

3.1 可控融合模块

图 2 展示了我们提出的 CFB 架构，它由一个共享的交叉变换器和一个特征融合机制组成。这里，交叉变换器基于多头交叉注意力机制，根据基础组织图像中每个位置与参考组织图像中每个区域之间的成对注意力分数，密集聚合相关的输入图像特征。查询嵌入 $\boldsymbol{q}^m \in \mathbb{R}^{n\times d}$ 由基础特征$\mathbf{h}^b \in \mathbb{R}^{n\times D}$计算得出，而键 $\boldsymbol{k}_i^m \in \mathbb{R}^{n\times d}$ 和值 $\boldsymbol{v}_i^m \in \mathbb{R}^{n\times d}$ 则来自

参考特征$\mathbf{h}_i^{ref} \in \mathbb{R}^{n\times D}$，其中 $d = {}^{D}/_{M}$， M 表示注意力头的数量。接下来，交叉注意力函数使用键值对将查询映射为输出$\boldsymbol{r}_i^m$。最后，将所有 M 个注意力头的输出 $\boldsymbol{r}_i^m$拼接起来，并通过可学习权重矩阵$\mathbf{W} \in \mathbb{R}^{D\times D}$ 进行处理，从而生成交叉注意力特征 $\boldsymbol{c}_i \in \mathbb{R}^{n\times D}$，表示为

$$\boldsymbol{c}_i = [\boldsymbol{r}_i^1; \cdots; \boldsymbol{r}_i^M]\mathbf{W} + \mathbf{h}^b, \qquad \text{where} \quad \boldsymbol{r}_i^m = \text{softmax}(\frac{\boldsymbol{q}^m \boldsymbol{k}_i^{m\top}}{\sqrt{d}})\boldsymbol{v}_i^m. \tag{1}$$

接下来，我们在交叉变换器中引入一种可控特征调制机制，以进一步增强生成图像的多样性和质量。

可控特征调制：上文所述的标准交叉注意力机制能够计算出局部一致的特征，从而生成伪影较少的图像。然而，鉴于交叉注意力的确定性，以及参考图像数量有限，在少样本设置下同时生成多样且高质量的图像仍然是一项挑战。为此，我们在 CFB 中引入了一种可控特征调制机制，旨在提升生成图像的多样性和质量。所提出的调制机制在特征聚合与细化步骤中融入了随机性，并增强了控制能力。具体而言，我们利用映射网络的输出，对交叉变换器中层归一化模块内的视觉特征进行调制。

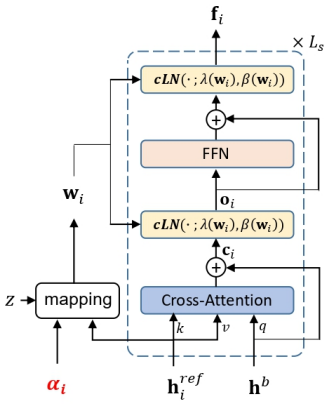


图 2：在我们的 CFB 中，使用可控交叉调制层归一化（cLN）对基础组织图像特征和参考组织图像特征进行交叉注意力处理。此处，参考特征 $\mathbf{h}_i^{ref}$ 、噪声 $\boldsymbol{z}$ 和控制参数 α_i 被输入映射网络，以生成元权重 $\mathbf{w}_i$。所得$\mathbf{w}_i$ 通过 $\lambda(\mathbf{w}_i)$ 和 $\beta(\mathbf{w}_i)$ 对我们的 cLN 中的特征进行调制。通过这种可控特征调制，输出特征 $\mathbf{f}_i$ 能够生成多样且与输入组织图像语义一致的组织图像。

映射网络： 元权重 $\mathbf{w}_i \in \mathbb{R}^D$由映射网络生成，如下所示：

$$\mathbf{w}_i = \mathbf{g}_i^{ref} \odot \psi_\alpha(\alpha_i) + \psi_z(\boldsymbol{z}), \tag{2}$$

其中， $\psi_\alpha(\cdot)$ 和 $\psi_z(\cdot)$ 是线性变换， $\boldsymbol{z} \sim \mathcal{N}(0, 1)$ 是高斯噪声向量， α_i 是控制参数。$\mathbf{g}_i^{ref}$ 是通过线性变换和全局平均池化操作，从参考特征 $\mathbf{h}_i^{ref}$ 计算得到的全局级特征

随后，元权重$\mathbf{w}_i$ 用于调制特征，这一过程通过我们的交叉调制层归一化实现，具体描述如下。

可控交叉调制层归一化（*cLN*）：我们的 cLN 学习依赖样例的调制权重，以对特征进行归一化，因为我们希望

生成与少样本样例相似的图像。对特征进行这种动态调制，使我们的框架能够生成高质量且多样化的图像。为此，我们使用元权重 $\mathbf{w}_i$ 来计算层归一化模块中的调制参数 λ 和 β。给定作为输入的交叉注意力特征 $\mathbf{c}_i$，我们的 cLN 对输入进行调制，生成输出特征 $\mathbf{o}_i \in \mathbb{R}^{n \times D}$，表示为

$$\mathbf{o}_i = \mathrm{cLN}(\mathbf{c}_i, \mathbf{w}_i) = \lambda(\mathbf{w}_i) \odot \frac{\mathbf{c}_i - \mu}{\sigma} + \beta(\mathbf{w}_i), \tag{3}$$

其中，μ 和 σ^2 分别是输入 $\mathbf{c}_i$的估计均值和方差。此处，$\lambda(\mathbf{w}_i)$ 是元权重 $\mathbf{w}_i$ 与样本无关的可学习权重 $\boldsymbol{\lambda} \in \mathbb{R}^D$逐元素相乘得到的，即 $\boldsymbol{\lambda} \odot \mathbf{w}_i$。对 $\beta(\mathbf{w}_i)$ 也执行类似计算。因此，本文提出的归一化机制能够基于参考图像输入对输入特征进行可控调制，从而提升生成图像的多样性和质量。随后，将得到的特征 $\mathbf{o}_i$ 输入前馈网络（FFN），再经过另一层条件层归一化（cLN）进行逐点特征细化，如图 2 所示。之后，使用控制参数 α_i 聚合交叉注意力特征 $\mathbf{f}_i$，得到融合特征表示 $\mathbf{f} = \sum_i \alpha_i \mathbf{f}_i$，其中$i \in [1, \cdots, K-1]$。最后，解码器 $\mathcal{F}_D$ 生成最终图像 $\hat{}x$。

3.2 训练与推理

训练： 整个框架按照铰链版本的生成对抗网络 [18] 公式进行端到端训练。生成器 $\mathcal{F}_G$ 由我们的编码器、CFB 和解码器共同构成，判别器为 $\mathcal{F}_{Dis}$，对抗损失 $\mathcal{L}^{adv}$ 定义如下：

$$\begin{gathered}\mathcal{L}^{adv}_{\mathcal{F}_{Dis}} = \mathop{\mathbb{E}}_{x \sim real}[\max(0, 1 - \mathcal{F}_{Dis}(x))] + \mathop{\mathbb{E}}_{\hat{x} \sim fake}[\max(0, 1 + \mathcal{F}_{Dis}(\hat{x}))] \\ \text{and} \quad \mathcal{L}^{adv}_{\mathcal{F}_G} = - \mathop{\mathbb{E}}_{\hat{x} \sim fake}[\mathcal{F}_{Dis}(\hat{x})].\end{gathered} \tag{4}$$

此外，为了鼓励生成图像 $\hat{}x$ 在感知上与基于指定控制参数 $\boldsymbol{\alpha}$的参考图像相似，我们采用参数化形式的标准感知损失 [11], ，定义如下：

$$\mathcal{L}^p = \sum_i \alpha_i \mathcal{L}^p_i, \qquad \text{where} \quad \mathcal{L}^p_i = \mathbb{E}[\|\phi(\hat{x}) - \phi(x^{ref}_i)\|_2]. \tag{5}$$

此外，分类损失 $\mathcal{L}^{cl}$ 确保解码器生成的图像被分类为输入少样本样本所对应的类别。随后，我们使用以下公式训练 XM-GAN：$\mathcal{L} = \mathcal{L}^{adv} + \eta_p \mathcal{L}^p + \eta_{cl} \mathcal{L}^{cl}$，其中 η_p 和 η_{cl} 是用于对损失项加权的超参数。

推理： 在推理过程中，通过改变控制参数 α_i，可为一组固定的 K样本生成多张高质量且多样化的图像$\hat{}x$。基准图像 x^b 和 α_i 可随机选取，而我们的框架允许用户根据 α_i 值的选择来控制生成过程。

4 实验

我们在人类结直肠癌数据集 [14]上进行实验。该数据集包含 8 类结直肠组织：肿瘤、间质、淋巴组织、复合组织、碎屑、黏膜、脂肪组织和空白，每类有 625 张图像。为实现少样本设置，我们将这 8 类划分为 5 个已见类别（用于训练）和 3 个未见类别（用于评估），每类包含 40 张图像。我们使用两项指标评估所提方法：Fréchet Inception 距离（FID） [8] 和学习感知图像块相似度（LPIPS） [27]。我们的编码器 $\mathcal{F}_E$ 和解码器 $\mathcal{F}_D$ 均包含五个带批量归一化和 Leaky-ReLU 激活的卷积块，与 [7]中的设置相同。输入图像和生成图像的尺寸为 128×128。线性变换 $\psi(\cdot)$ 通过 1×1 卷积实现，输入和输出通道数均设为 D。权重 η_p 和 η_{cl} 分别设为 50 和 1。除非另有说明，所有实验中均将 $K=3$ 设为指定值。我们的 XM-GAN 使用批量大小为 8 的设置进行训练，采用 Adam 优化器，固定学习率为 10^{-4}。

4.1 最先进方法比较

少样本组织图像生成：在表 1 中，我们将用于少样本组织图像生成的 XM-GAN 方法与该数据集上的最先进方法 LoFGAN [7]进行 [14]比较。我们提出的 XM-GAN 利用全局感受野内相关局部信息的密集聚合，并结合可控特征调制，性能显著优于 LoFGAN，FID 分数提升了 30。1，达到 55。8。此外，XM-GAN 的 LPIPS 分数也更优。图 3 展示了 XM-GAN 与 LoFGAN [7]的定性比较。

低数据分类：在此，我们评估了 XM-GAN 生成的组织图像作为数据增强来源，对于未见类别的低数据结直肠组织分类这一下游任务的适用性。未见数据集被划分为 D_{tr}，D_{val}，D_{test}。未见类别的图像按 10:15:15 划分。按照 [7],，已见类别用于初始化 ResNet18 骨干网络，并使用 D_{tr}训练一个新的分类器。我们将此方法称为标准方法。随后，我们使用 XM-GAN 生成的 30 张组织图像对 D_{tr} 进行增强

表 1：我们的 XM-GAN 在 FID 和 LPIPS 分数上均取得稳定提升，优于 [14]数据集上的 LoFGAN。

方法	FID(↓)	
LoFGAN [7]	85.9	0.44
本文方法：XM-GAN	**55.8**	**0.48**

图像，并针对每个未见类别使用相同的 D_{tr} 作为少样本样例。表 2 展示了分类性能对比。与 LoFGAN [7], 相比，我们的 XM-GAN 取得了 2.8% 的绝对提升。

4.2 消融研究

在此，我们开展消融研究，以验证所提出贡献的优势。表3展示了在 [14] 数据集上的基线比较。我们的基线包含一个编码器、一个具有标准

表2：低数据图像分类。与近期提出的 LoFGAN 相比，所提出的 XM- GAN 在分类性能方面更优。

方法	准确率（%）
标准	68.1
LoFGAN [7]	69.7
我们的方法：XM-GAN	**72.5**

表3：将参数化感知损失（PPL）和条件层归一化（cLN）整合到基线中的影响。更多详情请参阅第4.2节。

方法	FID(↓)	
基线	73.6	0.451
基线 +PL	69.2	0.467
基线 +PPL	66.5	0.471
基线 +PPL + 条件层归一化†	62.1	0.475
我们的方法：基线 + PPL +条件层归一化	**55.8**	**0.482**

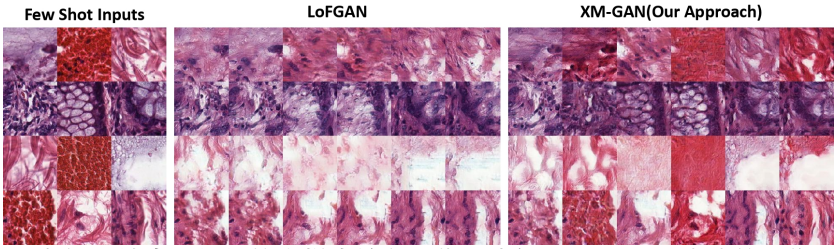


图 3：左侧：结直肠组织的少样本输入图像。中间：LoFGAN 生成的图像。右侧：我们的 XM-GAN 生成的图像。与 LoFGAN 相比，我们的 XM-GAN 生成的图像质量更高且更多样。建议放大查看。更多结果见补充材料。

层归一化（LN）层的标准交叉变换器以及一个解码器。该模型记为基线。基线+PL表示在基线的基础上进一步集成标准感知损失。我们使用α_i 使得$\sum_i \alpha_i{=}1$ 的随机值进行额外实验，以计算融合特征$\mathbf{f}$ 和参数化感知损失（公式5）。我们将其称为基线+PPL。此处所称的最终XM-GAN，即基线+PPL+cLN，包含新颖的CFB。在我们的CFB中，我们还通过仅使用公式2中的高斯噪声 $\boldsymbol{z}$ ，计算元权重$\mathbf{w}_i$ ，来验证参考特征对特征调制的影响。该模型在此记为基线+PPL+cLN†。基于新颖CFB的方法在所有基线方法中取得了最佳性能。

4.3 人工评估研究

我们对一组由十名病理学家组成的团队开展了一项研究，他们的平均从业年限为 8.5 年。向每位病理学家随机展示 20 张图像（10 张真实图像和 10 张由 XM-GAN 生成的图像），并要求他们判断图像是真实的还是生成的。研究表明，病理学家仅有 55% 的时间能够区分 AI 生成图像与真实图像仅，这一结果与二元分类问题中的随机预测相当，说明我们提出的生成框架能够生成逼真的结直肠图像。

5 结论

我们提出了一种小样本结直肠组织图像生成方法，其中包含一个可控融合模块（CFB），该模块根据局部区域与基础组织图像中区域的相似性，对参考组织图像中的局部区域进行密集聚合，从而生成局部一致的特征。我们在 CFB 中引入了映射网络和跨调制层归一化，以提升生成图像的质量和多样性。我们通过定量、定性和基于人工评估的方式对 XM-GAN 进行了全面验证，并取得了当前最先进的结果。

致谢：我们衷心感谢为本项目作出重要贡献的病理学家。我们尤其感激来自印度科泽科德政府医学院的Hima Abdurahiman博士；来自印度科泽科德MVR癌症中心与研究所的Sajna PV博士；来自英国北德文郡地区医院的Binit Kumar Khandelia博士；来自印度科泽科德Aster母亲医院的Nishath PV博士；来自印度特里凡得琅Girija博士诊断实验室与扫描中心的Mithila Mohan博士；来自印度科泽科德Aster MIMS的Kavitha博士；以及其他几位未具名的病理学家。他们在我们研究工作的各个阶段提供了专业建议、宝贵意见和深刻反馈。

参考文献

1. Antoniou, A., Storkey, A., Edwards, H.: 数据增强生成对抗网络。载于：arXiv预印本 arXiv:1711.04340（2017） 2. Bartunov, S., Vetrov, D.: 使用生成式匹配网络进行少样本生成建模。载于：ICAIS（2018） 3. Brock, A., Donahue, J., Simonyan, K.: 用于高保真自然图像合成的大规模GAN训练。载于：arXiv预印本 arXiv:1809.11096（2018） 4. Clouâtre, L., Demers, M.: Figr：使用Reptile进行少样本图像生成。载于：arXiv预印本 arXiv:1901.02199（2019） 5. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., 等：一张图像胜过16×16个词：面向大规模图像识别的Transformer。载于：arXiv预印本 arXiv:2010.11929（2020） 6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: 生成对抗网络。载于：NeurIPS（2014） 7. Gu, Z., Li, W., Huo, J., Wang, L., Gao, Y.: LoFGAN：融合局部表征的少样本图像生成。载于：ICCV（2021） 8. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: 采用双时间尺度更新规则训练的GAN收敛至局部纳什均衡。载于：NeurIPS（2017） 9. Hong, Y., Niu, L., Zhang, J., Zhang, L.: MatchingGAN：基于匹配的少样本图像生成。载于：ICME（2020） 10. Hong, Y., Niu, L., Zhang, J., Zhao, W., Fu, C., Zhang, L.: F2gan：用于少样本图像生成的融合与填充GAN。载于：ACM MM（2020）

11. Johnson, J., Alahi, A., Fei-Fei, L.: 用于实时风格迁移和超分辨率的感知损失。载于：欧洲计算机视觉国际会议（2016） 12. Karras, T., Aila, T., Laine, S., Lehtinen, J.: GAN渐进式增长：提升质量、稳定性和多样性。载于：arXiv预印本 arXiv:1710.10196（2017） 13. Karras, T., Laine, S., Aila, T.: 用于生成对抗网络的风格化生成器架构。载于：计算机视觉与模式识别会议（2019） 14. Kather, J.N., Zöllner, F.G., Bianconi, F., Melchers, S.M., Schad, L.R., Gaiser, T., Marx, A., Weis, C.A.: 结直肠癌组织学纹理集（2016年5月）。https://doi.org/10.5281/zenodo.53169，https://doi.org/10.5281/zenodo.53169 15. Kingma, D.P., Welling, M.: 自编码变分贝叶斯。载于：arXiv预印本 arXiv:1312.6114（2013） 16. Kumar, A., Ghose, S., Chowdhury, P.N., Roy, P.P., Pal, U.: Udbnet：通过对抗博弈实现的无监督文档二值化网络。载于：2020年第25届国际模式识别会议（ICPR）。页码：7817–7824。电气与电子工程师协会（2021） 17. Liang, W., Liu, Z., Liu, C.: Dawson：一种领域自适应少样本生成框架。载于：arXiv预印本 arXiv:2001.00576（2020） 18. Lim, J.H., Ye, J.C.: 几何生成对抗网络。载于：arXiv预印本 arXiv:1705.02894（2017） 19. Ohata, E.F., Chagas, J.V.S.d., Bezerra, G.M., Hassan, M.M., de Albuquerque, V.H.C., Filho, P.P.R.: 一种用于结直肠癌组织学图像分类的新型迁移学习方法。《超级计算杂志》页码：1–26（2021） 20. Shamshad, F., Khan, S., Zamir, S.W., Khan, M.H., Hayat, M., Khan, F.S., Fu, H.: 医学影像中的Transformer：综述。arXiv预印本 arXiv:2201.09873（2022） 21. Thawakar, O., Anwer, R.M., Laaksonen, J., Reiner, O., Shah, M., Khan, F.S.: 使用时空Transformer进行三维线粒体实例分割。arXiv预印本 arXiv:2303.12073（2023） 22. Vahdat, A., Kautz, J.: NVAE：一种深度分层变分自编码器。载于：NeurIPS（2020） 23. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.:《Attention is All You Need》。载于：NeurIPS（2017） 24. Wang, C., Shi, J., Zhang, Q., Ying, S.: 使用双线性卷积神经网络进行组织病理图像分类。载于：2017年第39届电气与电子工程师协会医学与生物学工程学会国际年会（EMBC）。页码：4050–4053。电气与电子工程师协会（2017） 25. Wang, K.S., Yu, G., Xu, C., Meng, X.H., Zhou, J., Zheng, C., Deng, Z., Shang, L., Liu, R., Su, S., 等：基于组织病理图像和人工智能对结直肠癌进行准确诊断。《BMC医学》**19**(1), 1–12（2021） 26. Yu, G., Sun, K., Xu, C., Shi, X.H., Wu, C., Xie, T., Meng, R.Q., Meng, X.H., Wang, K.S., Xiao, H.M., 等：通过在病理图像上进行半监督深度学习准确识别结直肠癌。《自然·通讯》**12**(1), 6311（2021） 27. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: 深度特征作为感知度量的非凡有效性。载于：计算机视觉与模式识别会议（2018）