

流形学习流模型在重构心脏图像的应用

赵嘉俊

摘要

心脏图像分割对于心脏相关功能参数计算以及相关疾病的诊断起到至关重要的作用。对于无监督自动分割模型,推断数据概率密度分布是必不可少的。虽然许多深度生成模型在概率密度评估方面取得较好效果,但是它们并不能很好兼顾概率密度易处理性和低维数据流形的计算高效性。为此,复现论文提出了一种新的生成模型,称为流形学习流模型 (\mathcal{M} -flow)。而且, \mathcal{M} -flow 结合了归一化流、生成对抗网络、自动编码器和基于能量的模型,有可能更可靠地用流形结构表示数据集,可提供降维、去噪和分布外检测的处理。因此,本文将 \mathcal{M} -flows 应用到心脏图像生成,研究潜在空间维度与 \mathcal{M} -flow 模型性能的关系。实验结果表明, \mathcal{M} -flow 模型的表示数据能力强,而且增加潜在变量空间的维度能稍微提高图片表示质量。

关键词: 心脏图像分割; 深度生成模型; \mathcal{M} -flow

1 引言

近年来,心血管疾病发病频率逐年升高,每年因心血管疾病死亡的人数已成为非传染性疾病死亡人数的首位。心脏核磁共振 (Magnetic Resonance, MR) 图像对于心脏相关功能参数计算以及相关疾病的诊断起到至关重要的作用。临床中,一般对心脏图像各部位的分割主要靠专家手动标注。由于心脏各部位在各个阶段变化明显,导致分割更为困难。分割过程不但耗时,并且各专家分割时会产生主观差异,因此需要研究一种心脏图像自动分割系统。在无监督自动分割模型训练中,要求模型能从样本数据推断数据的概率分布。

目前,越来越多地使用深度生成模型来解决这个问题。生成对抗网络 (Generative Adversarial Networks, GANs)^[1]和变分自编码器 (Variational Autoencoders, VAEs)^[2]都基于低维潜在空间和到数据空间的可学习映射,本质上是描述嵌入数据空间中低维数据流形。虽然它们允许有效的采样,但它们的概率密度是难以处理的,限制了它们在推理任务中的使用性。另一方面,归一化流^[3]是基于与数据空间维数相同的潜在空间和微分同胚映射;它们的易处理概率密度贯穿整个数据空间,而不局限于低维曲面。

流方法可能不适合那些没有填充完整数据空间且被限制到较低维流形的数据^[4]。构造归一化流不能准确地表示这样的数据结构,反而会学习一个模糊的、不支持数据流形的版本。图 1 说明了这一点。此外,对潜在空间要求与数据空间具有相同的维数,这增加了模型的内存占用和计算成本。因此,如何将归一化流的概率密度易处理性与低维数据流形的计算高效性合理结合起来是非常重要的。Brehmer 等人^[5]提出了一种新的生成模型,称为流形学习流 (\mathcal{M} -flow)。它是基于从低维潜在空间到数据空间的单射且可逆的归一化流,可同时学习数据流形以及该流形上易求解的概率密度。它们结合了归一化流、GANs、自动编码器等方面,从而更可能可靠地用流形结构表示数据集。因此,本文将 \mathcal{M} -flows 应用到心脏图像生成。

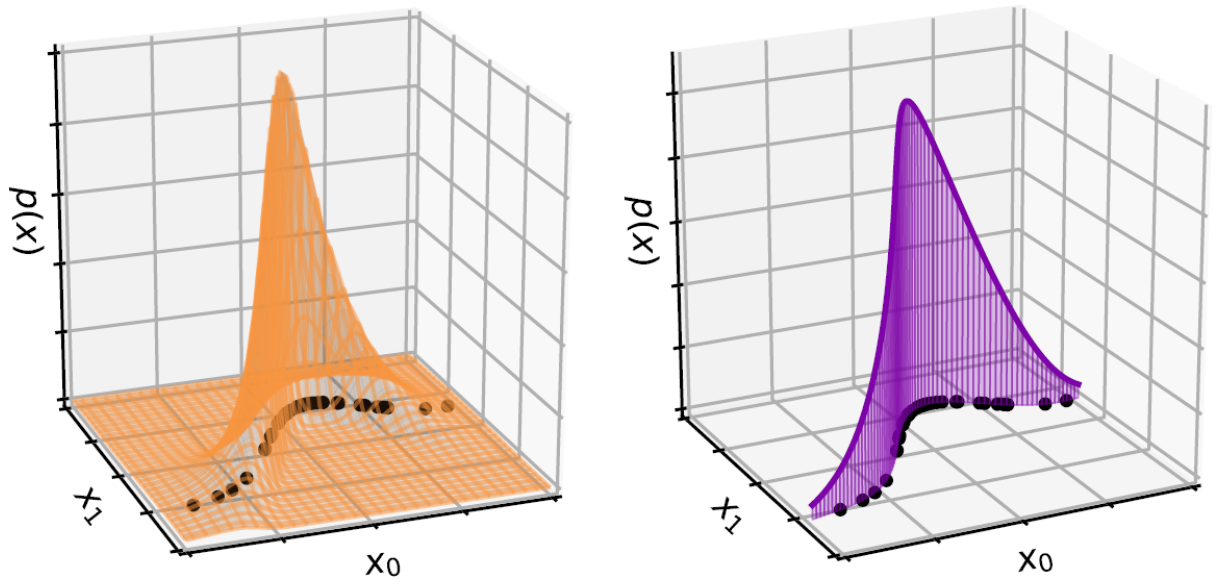


图 1: 描述数据空间中的标准流 (左) 和 \mathcal{M} 流 (右) 是如何在流形上建模数据

2 相关工作

在介绍新的流形学习流 (\mathcal{M} -flow) 之前, 我们将首先回顾生成模型的现有类是如何解决数据概率密度分布问题。为了简单起见, 我们将流形在拓扑上等价于 \mathbb{R}^n , 并假设其维数 n 是已知的。为了便于直接比较, 我们将用潜在变量 $u \in U$ 和 $v \in V$ 的两个向量来描述所有生成模型, 其中 $u = \mathbb{R}^n$ 是映射到已知流形的潜空间 \mathcal{M} , 也就是流形的坐标。 $v = \mathbb{R}^{d-n}$ 参数化任何剩余的潜在变量, 表示“离开流形”的方向。接下来, 我们简单描述这些方法, 这些方法总结在图 2 中。

2.1 环境流场

环境流场 (Ambient flow, AF) 是数据空间中的标准的欧几里得归一化流, 是微分同胚映射函数 $f: U \times V \mapsto X$ 加可解决的基密度 $p_{uv}(u, v)$ 。在 X 中的密度表达式为:

$$p_x(x) = p_{uv}(f^{-1}(x)) |\det J_f(f^{-1}(x))|^{-1}. \quad (1)$$

在生成模式中, 流从它们的基密度中采样 u 和 v , 并应用变换 $x = f(u, v)$ 产生 $x \sim p_x(x)$ 。值得注意的是, u 和 v 之间没有区别, 该模型也没有数据流形的概念。

2.2 在给定流形上的流

在给定流形上的流 (Flow on a prescribed manifold, FOM) 指的是, 当一个函数 $g^*: U \mapsto \mathcal{M}^* \subset X$ 是一个已知先验, 那么就可以在该流 \mathcal{M}^* 中构造一个流。因此, 密度定义为:

$$p_{\mathcal{M}^*}(x) = p_u(g^{*-1}(x)) |\det [J_g^T(g^{*-1}(x)) J_g(g^{*-1}(x))]|^{-\frac{1}{2}}. \quad (2)$$

密度 $p_u(u)$ 用 n 维的标准归一化流建模, 即用可学习的微分同胚 h 映射到另一组潜在变量 \tilde{u} 和相应的基密度 $p_{\tilde{u}}(\tilde{u})$ 。

2.3 生成对抗网络

生成对抗网络 (Generative adversarial networks, GANs) 将一个 n 维潜在空间映射到数据空间, 主要通过可学习函数 $g: U \mapsto \mathcal{M} \subset X$ 。 g 既不被限制为可逆的、也不被限制为单射的, 意味着可以有多个 u 对应同一个数据点 x 。严格地说, g 不是一个‘chart’, 也不一定是黎曼流形的变换。

2.4 伪逆编码器

伪逆编码器（Pseudo-invertible encoder, PIE）将环境流的潜在变量分成两个具有不同基密度的向量 u 和 v 。 u 的分布指定表示流形上的坐标，是用 n 维欧氏流建模的，即变换函数 h 。该函数将 u 映射到另一个与基密度相关的潜在变量 \tilde{u} 。对于 v ，它应该学习潜在空间的非流形方向，其基密度 $p_v(v)$ ，这个密度函数在 0 附近快速达到峰值。

2.5 概率自编码器

概率自编码器（Probabilistic autoencoder, PAE）用一个编码器 $e : X \mapsto U$ 编码数据，而且使用一个解码器 $g : U \mapsto X$ 将潜在变量映射到数据的无噪声版本。值得注意的是， g 不要求具有可逆性。虽然放宽可逆性的要求可能会使 PAE 更有表现力，但是它同时失去了模型的可处理密度分布特性。

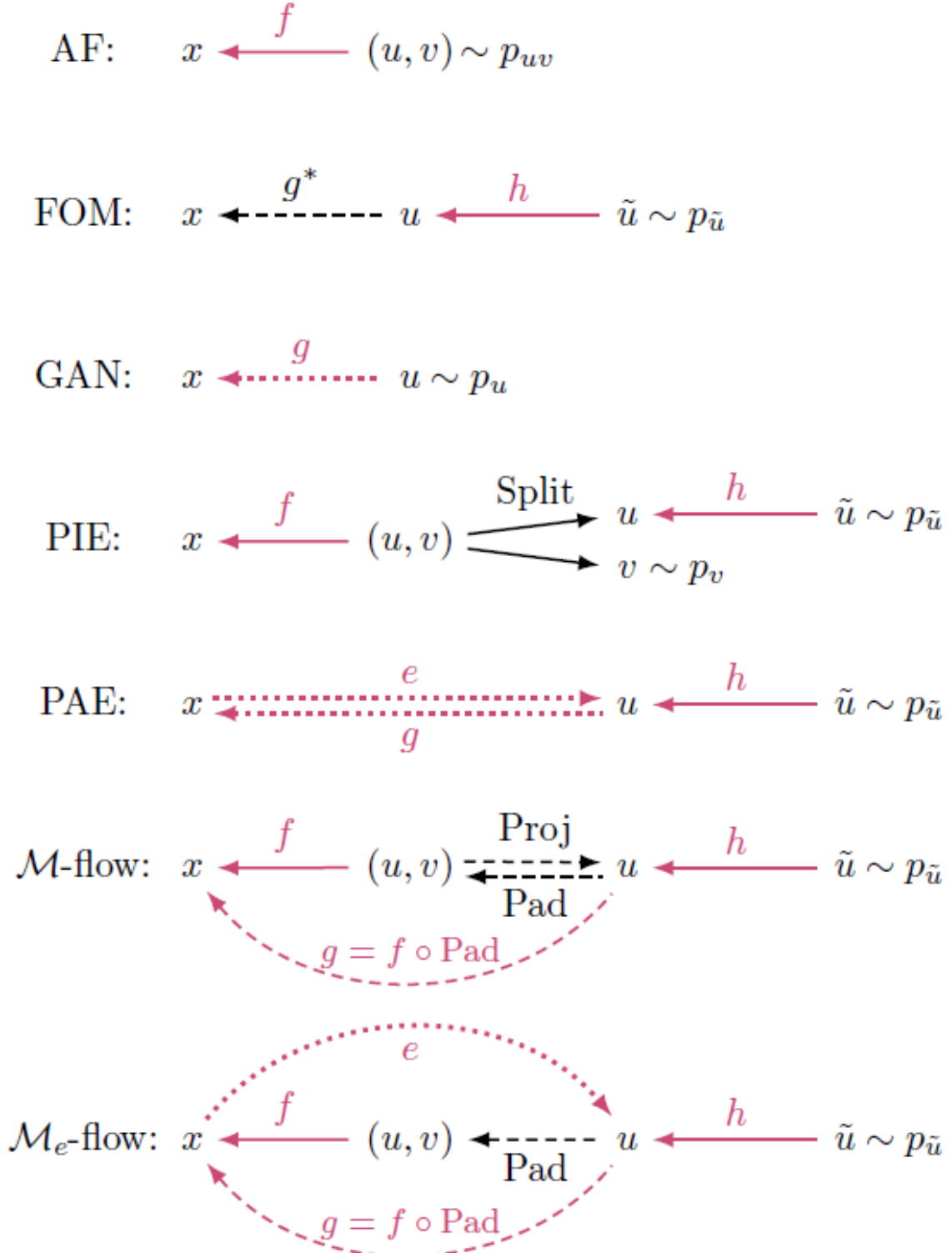


图 2: 生成模型的方法总结

3 本文方法

3.1 流形学习的流模型概述

流形学习的流模型（Manifold-learning flow, \mathcal{M} -flow）是复现工作提出的模型，它将 GANs 的可学习流形与 FOMs 的可处理密度相结合，而不引入生成模式与可处理似然之间的不一致。在 \mathcal{M} -flow 中，首先用微分同胚函数 $f : U \times V \mapsto X$ 来建模潜在空间和数据空间之间的关系，就像 AF 或 PIE 一样。此外，通过水平集定义模型流形 \mathcal{M} ，表示如下：

$$g : U \mapsto \mathcal{M} \subset X \quad \text{with} \quad u \mapsto g(u) = f(u, 0). \quad (3)$$

实际上，这个转换实现为一个零填充，然后是一系列可逆变换，即 $g = f_k \circ \dots \circ f_1 \circ \text{Pad}$ ，其中 Pad 表示用 $d - n$ 个零填充到一个 n 维向量中。与 FOM 和 PIE 一样，我们用 n 维流 h 来模拟基密度 $p_u(u)$ ，该流将 u 映射到另一个潜变量 \tilde{u} ，其相关的易处理基密度是 $p_{\tilde{u}}(\tilde{u})$ 。因此，流形上的概率密度可推导为：

$$p_{\mathcal{M}}(x) = p_{\tilde{u}}(h^{-1}(g^{-1}(x))) |\det J_h(h^{-1}(g^{-1}(x)))|^{-1} |\det [J_g^T(g^{-1}(x)) J_g(g^{-1}(x))]|^{-\frac{1}{2}}. \quad (4)$$

除了有一个可学习的转换 g 外，这与 FOM 模型相同。

从一个 \mathcal{M} -flow 中采样是很简单的：先取潜在变量 $\tilde{u} \sim p_{\tilde{u}}(\tilde{u})$ 并将潜在变量映射到数据空间中，即 $u = h(\tilde{u})$ ，然后使用 $x = g(u) = f(u, 0)$ 引导在流形的数据点，这是遵循 $x \sim p_{\mathcal{M}}(x)$ 的。

作为 \mathcal{M} -flow 的最后一个组成部分，需为任意且可能不在流形上的数据点添加了一个评估函数。 g 从低维潜在空间映射到数据空间，因此是一个解码器。我们定义一个对应的编码器 g^{-1} ，等于 f^{-1} 加 u 的投影函数，即 $g^{-1} : X \mapsto U$ ，意味着 $x \mapsto g^{-1}(x) = \text{Proj}(f^{-1}(x))$ ，其中 $\text{Proj}(u, v) = u$ 。这便将 g 的可逆性扩展到了整个数据空间 X 。而且，与自编码器一样，结合 g 和 g^{-1} 就可以计算重构误差 $\|x - x'\| = \|x - g(g^{-1}(x))\|$ 。当且仅当 $x \in \mathcal{M}$ 时重构误差才为 0。与标准的自动编码器不同，对于流形上的点，编码器和解码器是确切互相可逆的。

对任意点 $x \in X$ ，一个 \mathcal{M} -flow 允许计算三个量（见图 3）：在流形中的投射点 $x' = g(g^{-1}(x))$ 可能是输出的无噪声版本；重构误差 $\|x - x'\|$ 不仅对流形的学习很重要，可能也对异常检测有用；当在流形上的似然表示为 $p_{\mathcal{M}}(x')$ 时， \mathcal{M} -flow 将分离了数据流形和数据流的密度——在 AF 中这两个概念很容易混在一起。因此， \mathcal{M} -flow 可利用基于能量模型的想法来处理“off-the-manifold”问题，但在学习到的数据流形上仍然存在可处理的、准确的似然。

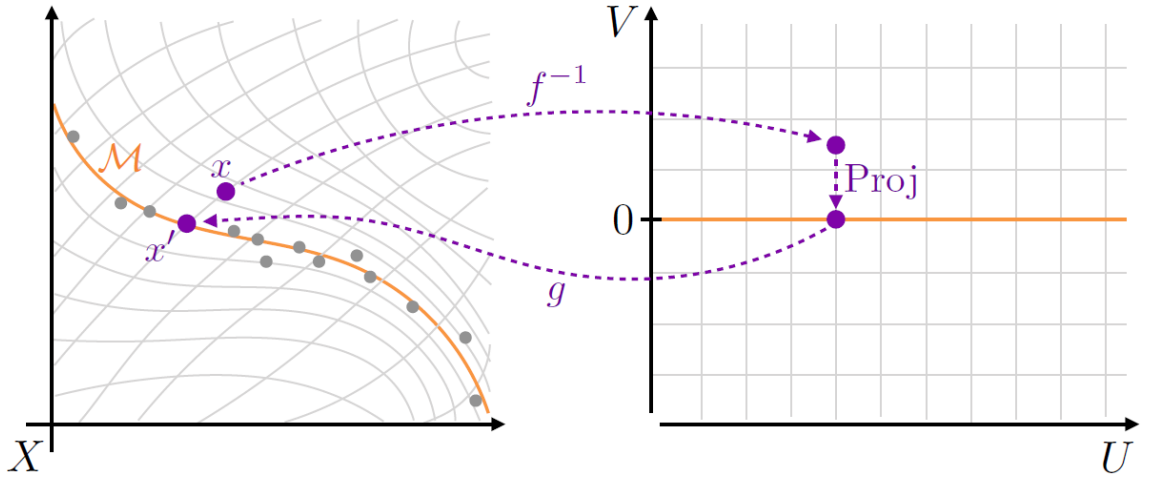


图 3: 生成模型的方法总结

3.2 具有编码器的 \mathcal{M} -flow

最后，我们介绍一个 \mathcal{M} -flow 模型的变体 (\mathcal{M}_e -flow)。 \mathcal{M}_e -flow 不是使用逆函数 f^{-1} 加 u 的投影函数作为编码器，而是使用单独函数 $e: X \mapsto U$ 编码数据。这个编码器不要求可逆的或者有一个易于处理的雅可比矩阵，使得可能增加了网络的表达能力。就像 \mathcal{M} -flow 方法一样，对于给定的数据点 x \mathcal{M}_e -flow 模型返回一个在已学习流形的投影点，然后评估重构误差和似然。这种编码器增加的表达能力是以编码器和解码器之间潜在不一致为代价的，因此要像标准自编码器一样在训练过程中添加不一致惩罚项。

3.3 M/D 训练法

复现工作将训练分为两个阶段。在流形学习阶段中只更新 f 的参数，它通过一个水平集定义了流形 \mathcal{M} 和映射函数 g 。类似于自动编码器，我们最小化投影到流形上的重构误差 $\|x - g(g^{-1}(x))\|$ 。对于 \mathcal{M}_e -flow 模型，编码器 e 的参数也在此阶段进行更新。在密度分布学习阶段，只根据最大化似然更新 h 的参数，固定流形参数，因为 h 只影响密度 p_u 。

4 复现细节

4.1 与已有开源代码对比及其创新点

在本次复现中，主要代码是使用复现论文^[5]提供的源码。特别地，本文工作包括三项，下面将详细描述。

第一，将 \mathcal{M}_e -flow 模型用于 MICCAI-2020 多中心、多供应商和多疾病心脏图像分割挑战的公共数据集（简称心脏图像数据）^[6]。由于此数据集不可直接使用，需要进行预处理。在数据集中只获取每个个体数据的中间层图片，因为其他层的图片的心脏占比过小；此外，对数据进行中心对齐和数据增强等操作。

第二，为了让 \mathcal{M}_e -flow 模型能顺利运行心脏图像数据，本文实现相应的数据获取和利用代码，例如缩放图片大小以满足模型要求和图片归一化等。

第三，潜在空间的变量维度与重构图像的好坏存在何种关系是未知的。因此，本文重点研究和分析适合心脏图像数据的最佳变量维度。

4.2 实验环境搭建与使用说明

在实验中，本文使用了一张 NVIDIA GeForce RTX 3090 的显卡和 pytorch 1.0 框架。其中 pytorch 需要一些额外包：matplotlib \geq 3.0.0, numpy \geq 1.16.0, python \geq 3.6.0, scikit-learn \geq 0.21.0, scipy \geq 1.2.0, torchvision \geq 0.4.0, palettable。

在本文中，模型训练的运行命令是：python train.py --algorithm emf --sequential --dataset cardiac --actnorm --batchnorm --modellatentdim xxx。其中，“emf”代表使用 \mathcal{M}_e -flow 算法；“--sequential”代表先训练流形再训练似然；“cardiac”代表使用心脏图像数据进行训练；“--actnorm --batchnorm”是神经网络的归一化方式；“--modellatentdim”用来设置潜在变量的维度。

5 实验结果分析

在实验设置方面，我们使用文献 [5] 中降采样到 64×64 分辨率的部分心脏图像数据集，其中包含 2475 张训练图像和 946 张测试图像。 \mathcal{M}_e -flow 模型为转换器 f 使用具有四层的多尺度设置。每一层有 five 步，每一步由一个“actnorm”层，一个可逆的 1×1 卷积和一个有理-二次耦合变换器组成。编码器 e 是一个残差卷积网络，具有跨越三维空间分辨率的六个残差块，紧接着进行线性转换到 n 维潜在空间。 h 由另外 8 个可逆、线性有理-二次耦合变换器组成。

所有模型都使用 AdamW 优化器进行训练，初始学习率为 $3 \cdot 10^{-4}$ ，采用余弦退火策略，权值衰减为 10^{-5} 。采用顺序 M/D 算法训练 \mathcal{M}_e -flow 模型。在 M/D 训练的流形阶段，我们使用的是平均重构误差；而在密度分布学习阶段，损失函数是负的平均对数似然。我们训练 500 个周期，样本批量大小为 25。模型权值在每一代之后被检查，最后恢复到导致最小验证损失的模型。为了研究潜在空间维度 n 与 \mathcal{M}_e -flow 模型性能的关系，我们考虑 $n = 128$ 的 \mathcal{M}_e -flow (Flow128)、 $n = 256$ 的 \mathcal{M}_e -flow (Flow256) 和 $n = 512$ 的 \mathcal{M}_e -flow (Flow512)。

表 1: 在心脏测试集中各模型的生成质量结果

Model	Flow128	Flow256	Flow512
Reconstruction error	2816.5315 \pm 467.3036	2397.1912 \pm 479.8530	2312.6079\pm443.2921

表 1 展示使用均方误差 (Mean Square Error, MSE) 作为重构误差得到的流模型生成质量对比。从表可知，Flow512 模型达到最好的性能，Flow128 模型的性能排名最后。这暗示了，增加潜在变量空间的维度能提高生成质量。值得注意的是，与 Flow512 相比，Flow128 的性能只是稍差，可是它节省了一定的模型计算能力。图 4 展示了在心脏测试集中各模型的生成质量结果。从图可知，各模型生成的图片质量相差不大。总的来说， \mathcal{M}_e -flow 模型的生成能力强，而且增加潜在变量空间的维度能稍微提高图片生成质量。

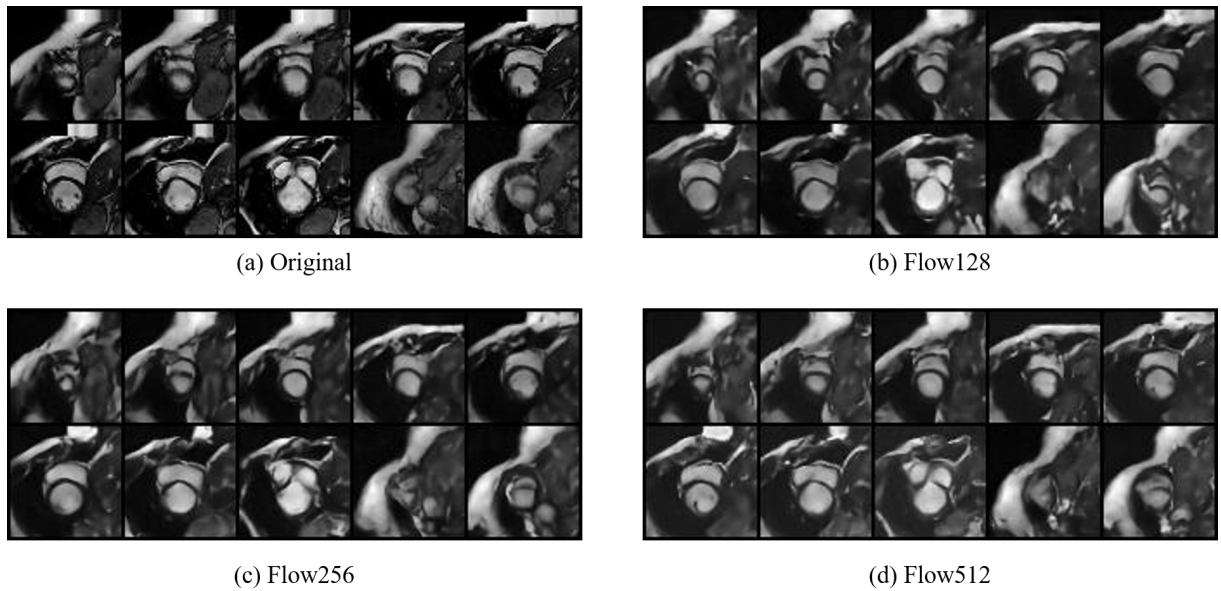


图 4: 心脏图像流形样本。左上部分显示了测试样本，其他部分显示了从各种流模型生成的样本。

6 总结与展望

本文复现的工作引入了流形学习流 (\mathcal{M}_e -flow)，是一种新的生成模型，它结合了归一化流、自动编码器和 GAN 的各个方面。 \mathcal{M}_e -flow 暗示数据在嵌入数据空间的低维流形上具有可处理的概率密度。流形和密度都是从数据中学到的。复现的工作使用稳定、高效的 M/D 训练策略进行模型的训练。本文将心脏数据集进行了预处理后，应用到 \mathcal{M}_e -flow 模型中，研究潜在空间维度与 \mathcal{M}_e -flow 模型性能的关系。实验结果发现， \mathcal{M}_e -flow 模型的生成能力强，而且增加潜在变量空间的维度能稍微提高图片生成质量。但是，由于未对潜在空间变量的表示做约束，可能导致相似图片的潜在表示相差很大，这是不合理的。因此，未来的研究工作将是加入一些约束条件或者函数来解决这个问题。

参考文献

- [1] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [2] KINGMA D P, WELLING M. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013.
- [3] DINH L, KRUEGER D, BENGIO Y. Nice: Non-linear independent components estimation[J]. arXiv preprint arXiv:1410.8516, 2014.
- [4] FEINMAN R, PARTHASARATHY N. A Linear Systems Theory of Normalizing Flows[J]. arXiv preprint arXiv:1907.06496, 2019.
- [5] BREHMER J, CRANMER K. Flows for simultaneous manifold learning and density estimation[J]. Advances in Neural Information Processing Systems, 2020, 33: 442-453.
- [6] CAMPELLO V M, GKONTRA P, IZQUIERDO C, et al. Multi-centre, multi-vendor and multi-disease cardiac segmentation: the M&Ms challenge[J]. IEEE Transactions on Medical Imaging, 2021, 40(12): 3543-3554.