

中山大学硕士学位论文

基于GAN的带病灶标签的配准多模态医学影像的 合成

Synthesis of Registered Multimodal Medical Images with
Lesion Label Based on GAN

学位申请人： 瞿毅力

指导老师： 卢宇彤教授

专业名称： 软件工程

答辩委员会（签名）：

主席：

委员：

二零二零年五月十六日

论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

学位论文使用授权声明

本人完全了解中山大学有关保留、使用学位论文的规定，即：学校有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版；有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆、院系资料室被查阅；有权将学位论文的内容编入有关数据库进行检索；可以采用复印、缩印或其他方法保存学位论文；可以为建立了馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

保密论文保密期满后，适用本声明。

学位论文作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

基于GAN的带病灶标签的配准多模态医学影像的合成

专业：软件工程

硕士生：瞿毅力

指导老师：卢宇彤教授

摘要

医学影像数据的采集和标注是非常困难的，尤其是配准的多模态数据。合成的医学影像数据可以很好地缓解此问题，但医学影像中具有复杂的生理结构信息，直接合成很容易生成不合理的生理结构。此外，当前医学影像合成的研究还存在模态数量少、需要配准多模态训练数据、不能合成指定病灶并检验、无法从随机矩阵合成、需要额外的数据产生生理结构信息、合成质量评价不客观等各项未能很好解决的问题。

因此，我们设计了一种无监督的基于GAN的配准多模态医学影像合成方法，无需配准训练数据，先从随机矩阵合成具有生理结构信息的结构特征图，进而生成一组有病灶标签的多模态配准医学影像。我们在多个数据集上验证了我们的合成数据中病灶信息的有效程度和合成数据在病灶处理任务中的可用程度。本文的主要工作如下：

1. 我们提出了一种基于Sobel边缘检测算子的结构特征图的提取与随机生成方法，无需额外的解剖结构分割标签或标签提取训练，可直接从任意模态的真实影像提取得到结构特征，用以辅助GAN学习生成更合理的合成影像。

2. 我们实现了带标签多模态配准影像的合成。我们将随机生成的结构特征图随机选取的病灶标签融合，通过生成网络合成多模态医学影像。我们通过实现生成的不同模态影像间的互相转换确保了它们的互相配准，通过对生成的影像的病灶信息的再处理还原出病灶标签确保了生成影像根据输入的标签生成了对应的病灶信息。

3. 我们对合成数据的可用性进行了客观的验证。我们使用不同数据量的合成数据和真实数据构建的数据集来训练病灶处理网络，验证了合成数据可以在医学影像智能处理任务作为预训练数据和增强数据来提高模型的能力。

关键词：医学影像、生成对抗网络、图像合成、多模态配准、边缘检测

Synthesis of Registered Multimodal Medical Images with Lesion Label Based on GAN

Major: Software Engineering

Name: Yili Qu

Supervisor: Prof. Yutong Lu

Abstract

The acquisition and labeling of medical image data is very difficult, especially for registered multi-modal data. Synthesized medical image data can alleviate this problem well, but medical images have complex physiological structure information, and direct synthesis can easily generate unreasonable physiological structures. In addition, the current research on medical image synthesis also includes a small number of modalities, the need to register multi-modal training data, the inability to synthesize designated lesions and test them, the inability to synthesize from random matrices, the need for additional data to generate physiological structure information, and the evaluation of synthesis quality. Objectives and other issues that are not well resolved.

Therefore, we designed an unsupervised GAN-based registration multi-modal medical image synthesis method. Without the need to register training data, we first synthesized a structural feature map with physiological structure information from a random matrix, and then generated a set of lesion labels. Multimodal registration medical image. We verified the validity of the lesion information in our synthetic data and the availability of the synthetic data in the lesion processing task on multiple data sets. The main work of this article is as follows:

1. We propose a method for extracting and randomly generating structural feature maps based on the Sobel edge detection operator. No additional anatomical structure segmentation labels or label extraction training are required, and structural features can be directly extracted from real images of arbitrary modes To assist GAN learning to generate more reasonable synthetic images.

2. We have realized the synthesis of labeled multi-modal registration images. We randomly fuse the randomly generated lesion feature maps with selected lesion labels and synthesize multi-modal medical images by generating a network. We ensure the mutual registration by implementing the mutual conversion between the different modal images that are generated. By reprocessing the lesion information of the generated images, we restore the lesion labels to ensure that the generated images generate corresponding lesion information based on the input labels. .

3. We have objectively verified the usability of the synthesized data. We used data sets constructed from synthetic data of different data amounts and real data to train the lesion processing network, and verified that the synthetic data can be used as pre-trained data and enhanced data in medical image intelligent processing tasks to improve the model's ability.

Keywords: Medical images, generative adversarial networks, image synthesis, multi-modal registration, edge detection

目录

摘要	I
Abstract	II
第一章 绪论	1
1.1 研究的背景	1
1.2 研究的现状	2
1.3 论文的结构	3
第二章 基础方法和数据集	5
2.1 数据集	5
2.2 生成对抗网络	6
2.3 变分自编码器	6
2.4 图像分割	6
2.5 物体检测	6
第三章 带病灶标签的配准多模态医学影像的合成	7
3.1 整体架构	7
3.2 结构特征图的提取和生成	7
3.3 多模态影像的合成	11
3.4 多模态影像的配准	12
3.5 病灶信息的添加和合成	13
3.6 合成数据集的构建	14
第四章 合成医学影像的评估与展示	17
4.1 评估指标	17
4.2 训练设置	17
4.3 合成方法在脑肿瘤MRI数据集上的消融实验	17
4.4 合成方法在不同数据集上的量化评估	17
4.5 合成病灶的有效性验证实验	19

4.6 合成数据在智能医学影像处理任务中的可用性验证实验	21
4.7 合成数据效果展示	22
第五章 结语	24
参考文献	25
致谢	32

第一章 绪论

1.1 研究的背景

- **医学影像及其模态** 医学影像是指为了医疗或医学研究，对人体或人体某部分，以非侵入方式取得的内部组织影像，其中不同的成像方式得到的医学影像我们称之为不同的模态，常见的医学影像模态有核磁共振成像（MRI）、CT成像、PET成像、B超成像、X射线成像等。有的模态在成像时设置不同的参数将得到具有明显视觉差异的不同的子模态，例如CT分低剂量和高剂量、MRI包括T1、T2、T1c等子模态。不同的模态对医生具有不同的参考价值，医生往往需要多个模态的影像互相对照才能做出准确的判断。在医学影像的智能处理任务的训练和学习中，我们往往也期望获得更多模态的影像，例如采用卷积神经网络（CNN）[1]或生成对抗网络（GAN）[2]进行的医学图像处理任务。
- **多模态图像的配准** 当同一个病人的同一个部位通过不同的成像技术得到不同的模态时，如果成像位置和视角是一致的，那么得到的不同模态的影像就是对齐的，我们称之为这些模态之间是配准的。相较于单模态数据，配准的多模态影像数据能提供更多的信息，可以支撑更多和更复杂的应用场景，满足深度神经网络对训练数据的需求，有助于提供更加高效可靠的智能诊断服务。对于医生来说，获取不同模态的影像需要花费更长的时间并且需要患者的耐心配合。
- **病灶的重要意义** 病灶是指人体器官的病变区域，例如肺结节、肿瘤、结石等。医学影像中的病灶信息对医生来说至关重要，它们是医生进行诊断的重要依据。在医学影像中，病灶区域与正常区域不一定具有明显的视觉差异，研究者们尝试使用人工神经网络来学习潜在的差异，以帮助医生进行诊断。对于医学影像智能处理任务的研究者来说，多模态的医学影像数据集十分稀缺，收集难度非常大，尤其是罕见病数据，而配准的数据则更加稀少，这使得很多的训练任务无法实现。因此，通过应用图像合成技术扩展数据集，从已有的单模态图像转换为配准的多模态图像、从随机噪声生成配准的多模态医学影像，有着广泛的用途和深远的意义。

- **图像合成的原理和发展** 图像合成分为从随机噪声合成图像和从带有指导信息的草图合成目标图像，后者进一步发展为像素到像素的图像转换。图像合成本质上是一个数据分布到另一个数据分布的转换。随着生成对抗网络（GAN）GAN [2]的出现，从随机噪声合成图像发展迅速。图像转换包括风格迁移、人像转换等诸多应用，我们最关注的其中的医学影像的模态转换，即以以一种医学影像模态为输入合成另一种模态的医学影像。
- **深度学习在医学影像上的应用** 除了GAN在医学影像的模态转换上的应用，深度学习还广泛应用于多种医学影像处理任务，如图像分割、病灶检测、图像分类等。 [3], [4], [5]等研究中详细阐述了该方向的广阔前景。

1.2 研究的现状

- **图像合成的研究现状** 图像到图像转换任务在最近的研究中被公式化为使用编码器和解码器组成的CNN实现对像素到像素的映射 [6], [7], [8], [9], [10], [11]。GAN可采用无监督或自监督训练，也可进行有监督训练，还能使用合成数据进行辅助训练 [12], [13], [14], [15]，因此在数据集样本不足的场景下，GAN更为适用。

在人像合成领域，多域转换的发展最近已经有了进展 [16], [17], [18], [6]。从这些研究中可以看出，采用基本的生成对抗网络进行处理主要有这些缺点：(i) n 个域需要训练 $n(n-1)$ 数量个转换模型；(ii) 在学习特定的转换模型时，不可利用其他域的数据。对此，诸如CycleGAN、IcGAN [19]、StarGAN [18]、ContrastGAN [17]等可实现图像到图像的多域转换方案或模型被陆续提出，最近的ModularGAN [16]、ComboGAN [20]和XGAN [21]将网络模块化为多个部件开启了另一种思路。

- **医学影像合成的研究现状** 一些研究通过不同模态数据之间的转换 [22], [23], [24], [25]来减少诊断和治疗中给医生和病人带来不必要的代价（例如病人辐射剂量的减少），甚至提高治疗的可行性 [24]，同时不同模态数据之间的转换可以很好的缓解数据样本稀少的难题 [12]。在GAN之前，一些研究使用图字典映射

[24]、稀疏编码 [26], [27], CNN [28]等探索了医学影像的跨模态转换。在GAN展现了强大的生成能力之后, 出现了许多基于GAN的医学影像转换研究 [22], [23], [29], [28], [25]。许多研究使用GAN实现了更高质量的转换结果 [16], [17], [30], [18]。最近一些研究实现了基于未配准的多模态数据的转换 [22], [31]。GAN逐渐被应用到各个部位的器官, 诸如脑部MRI到CT图像的转换 [23], [25]、视网膜血管注释到图像的转换 [32]、采用CycleGAN [30]的无监督心脏MRI到CT图像的相互转换与分割 [23]等。当前, 在医学影像处理任务中, GAN被广泛应用于医学影像重建 [33], [34]、合成 [12], [32], [14], [15]、转换 [22], [23], [29], [28], [25]、超分辨率 [35], [36]等各类研究。在CycleGAN [30]实现了不成对的图像到图像转换之后, 许多基于CycleGAN的研究从各个角度对GAN进行了发展。

但在当前的医学影像合成的研究中, 两模态之间的转换合成很多 [22], [23], [24], [27], [29], [28], [25], 对多模态的研究依然稀少 [37], [31], [12]。对于多模态影像的合成, [37]实现多输入多输出的MRI合成, 但对输入的多模态数据要求配准, [31]进一步实现了未配准的多输入合成模型, 能够从其输入的任何子集执行MRI图像合成, 但限制了输出为单一模态。[38]针对医学图像配准进行了深入研究。[12]应用GAN合成脑肿瘤图像实现数据增强和数据匿名化, 但需要额外训练解剖结构分割网络。

此外, [32]研究了基于变分自编码器(VAE) [39], [40]的思想实现血管注释图的随机生成, 进而合成彩色视网膜图像。

1.3 论文的结构

本文一共有五章, 每一章的内容安排如下:

第一章: 绪论。介绍本文的研究内容及其背景, 对国内外的研究现状进行简单的介绍与总结。

第二章: 基础方法和数据集。介绍本文后续章节中的方法和实验将会运用的基础方法和数据集的简介, 包括基础的生成对抗网络的介绍、变分自动编码器的介绍、图像分割任务的介绍、物体检测任务的介绍。

第三章: 带病灶标签的配准多模态医学影像的合成。本章是完整的方法描述, 先

介绍整体架构，再依次介绍结构特征图的提取和生成方法、多模态影像的合成和配准方法、辅助的模态转换训练过程、病灶信息的添加和合成方法、合成数据集的构建过程。

第四章：合成医学影像的评估与展示。本章节介绍对合成的医学影像的评估方法和实验结果，包括通用评估指标、消融实验、通用量化评估结果、合成的病灶的有效性评估、合成数据集在智能医学影像处理任务上的可用性评估。

第五章：结语。总结本文的研究成果，罗列本文的不足之处，分析本文方法落地运用的困难，并对进一步的研究方向进行展望。

第二章 基础方法和数据集

2.1 数据集

在我们的实验中使用了六个公开数据集：

- **BRATS2015数据集 [41]**公开数据集BRATS2015包含T1 / T2 / T1c / Flair的四个配准模态脑部3D MRI。训练集每个模态有274个大小为 $155 \times 240 \times 240$ 张图及对应的274张肿瘤分割标签图。我们将样本按9:1重新划分训练集和测试集，然后取每个3D MRI的55-105之间的50个切片构建2D数据集。在数据预处理时，我们已对每个图像都进行标准化。
- **Kaggle Chest X-Ray数据集¹**该数据集包括5863个病毒性肺炎、细菌性肺炎和正常肺部的正面2D X射线灰度图，图片尺寸从 384×127 到 2772×2304 不等。在数据预处理时，我们已对每个图像都进行标准化，尺寸缩放为 512×512 。
- **Kaggle Lung CT数据集²**该数据集包含267个胸部至腹部横向截面的2D CT灰度图，尺寸为 512×512 。在数据预处理时，我们已对每个图像都进行标准化。
- **DRIVE视网膜图像数据集³**该数据集的训练集和测试集各包含20张2D彩色眼底视网膜照片，尺寸均为 565×584 ，训练集还有20张对应的视网膜血管主视图。在数据预处理时，我们已对每个图像都进行标准化，尺寸统一插值为 512×512 。
- **FIRE视网膜图像数据集⁴**该数据集包含268张 2912×2912 的2D彩色眼底视网膜照片。在数据预处理时，我们已对每个图像都进行标准化，尺寸缩放为 512×512 。
- **天池全球数据智能大赛(2019)数据集⁵**该数据集包含肺部3D CT扫描共1837张，训练集1470张，测试集145张。训练集提供的标注信息为：中心坐标+直径（单

¹<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

²<https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data/>

³<http://www.isi.uu.nl/Research/Databases/DRIVE/>

⁴<https://projects.ics.forth.gr/cvrl/fire/>

⁵<https://tianchi.aliyun.com/competition/entrance/231724/information>

位为mm)+类别(1-结节, 2-肺密度增高影, 3-肺气肿或肺大泡, 5-索条, 31-动脉硬化或钙化, 32-淋巴结钙化, 33-胸膜增厚)。在数据预处理时, 我们根据标注信息切取有标注的slice及其前后slice组成的3通道图, 再对每个图像都进行标准化, 尺寸缩放为 512×512 。这样我们得到一个包含17156张 $512 \times 512 \times 3$ 的CT图, 新的标签为中间通道的原始标签。

2.2 生成对抗网络

2.3 变分自编码器

2.4 图像分割

2.5 物体检测

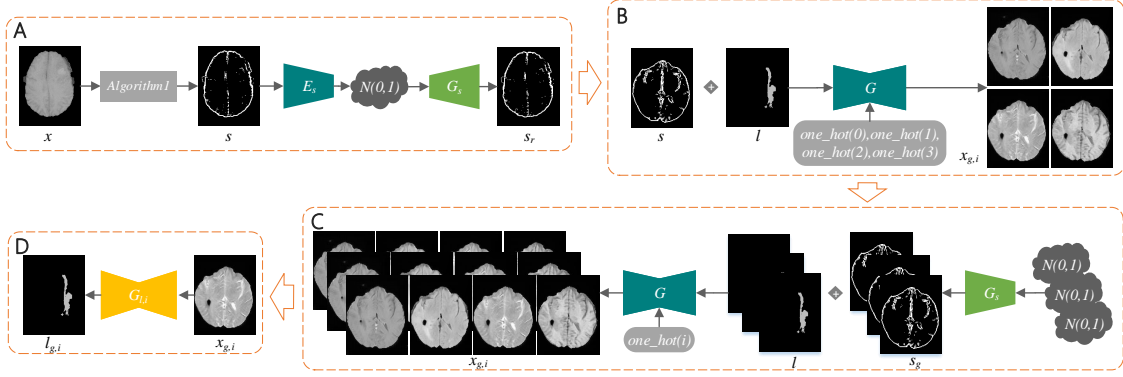


图 1: Architecture.

第三章 带病灶标签的配准多模态医学影像的合成

3.1 整体架构

图 1 所示包括四个主要阶段。在结构特征图提取和生成阶段，我们将获得一个结构特征图生成器，该生成器可以从随机正态分布矩阵生成结构特征图。在多模态影像合成阶段，我们以结构特征图为输入训练一个条件生成器，该条件生成器可以根据不同条件合成不同模态的影像。若需要在合成影像中添加指定的病灶信息，则可将病灶标签与结构特征图融合后作为输入，同时添加病灶生成指导损失。若对合成多模态的影像有高配准度要求，则可以对合成的多模态影像进行模态转换训练保证像素级配准。在合成数据集构建阶段，我们使用前面两个阶段生成的模型从随机正态分布矩阵合成配准的多模态的MRI。最后是对合成数据使用阶段，合成数据可在智能医学影像处理任务中作为增强样本进行运用。

3.2 结构特征图的提取和生成

GAN直接从随机噪声中生成的医学图像很难生成现实的结构信息。我们将提供基本轮廓和结构信息的图像称为结构特征图。例如，视网膜血管分布图可以看作是视网膜图像的结构特征图 [32]。结构特征图可以为合成医学图像提供必要的基本指导。在合成医学影像时，一些研究从组织分割标签 [12] 获取了基本的结构信息。但是，诸如视网膜血管图和脑组织分割标签之类的一般结构特征，在从原始图像中提取之前需

要额外的数据来训练一个提取模型。为此，我们首先设计了一种直接从医学影像直接提取结构特征图的方法，该方法具有操作快速，无需训练，无需附加数据的优点。

3.2.1 结构特征图提取方法

在传统的数字图像处理方法中，Roberts算子，Prewitt算子，Sobel算子等是出色的边缘检测算子。Sobel运算符通常用于处理医学图像。如算法 1 中所示，我们探索了一种从Sobel算子生成的边缘检测图中进一步提取结构特征图的方法。在算法 1 中，

Algorithm 1 Structural feature map extraction

- 1: Input a real image x and pixel threshold β
 - 2: $f1 = \text{reduce_min}(\text{sobel}(x))$
 - 3: $f2 = \text{reduce_max}(\text{sobel}(x))$
 - 4: $f1 = \text{mean}(f1) - f1$
 - 5: $f2 = f2 - \text{mean}(f2)$
 - 6: $f1 = \text{ones} \times (f1 > \beta)$
 - 7: $f2 = \text{ones} \times (f2 > \beta)$
 - 8: $f = \text{ones} \times ((f1 + f2) > 0.)$
-

我们使用Sobel运算符从真实图像中提取水平和垂直边缘检测图。每个边缘检测图执行最大缩小和最小缩小以获得两个边缘检测融合图，然后每个融合图计算与平均像素值的差。根据像素阈值对两个差异图进行二值化，然后对两个二进制图像求和，然后完全二值化。最终结果是我们需要的结构特征图。

3.2.2 结构特征图生成训练

生成结构特征图时，[12]仍需要输入真实图像以获取生成的结构特征图，这大大减少了合成数据的多样性。[32]实现了一种基于VAE的方法，用于根据多维正态分布生成视网膜血管分布图。以此为基础，我们设计了一种混合网络，结合了VAE和GAN的特征，可以从随机正态分布矩阵生成结构特征图，该图具有更好的多样性并且没有其他训练数据。此外，我们训练了一个生成器 MG ，该生成器从大脑结构特征图中获取了目标器官区域的掩膜，以供以后用于匹配病变标签。生成器与结

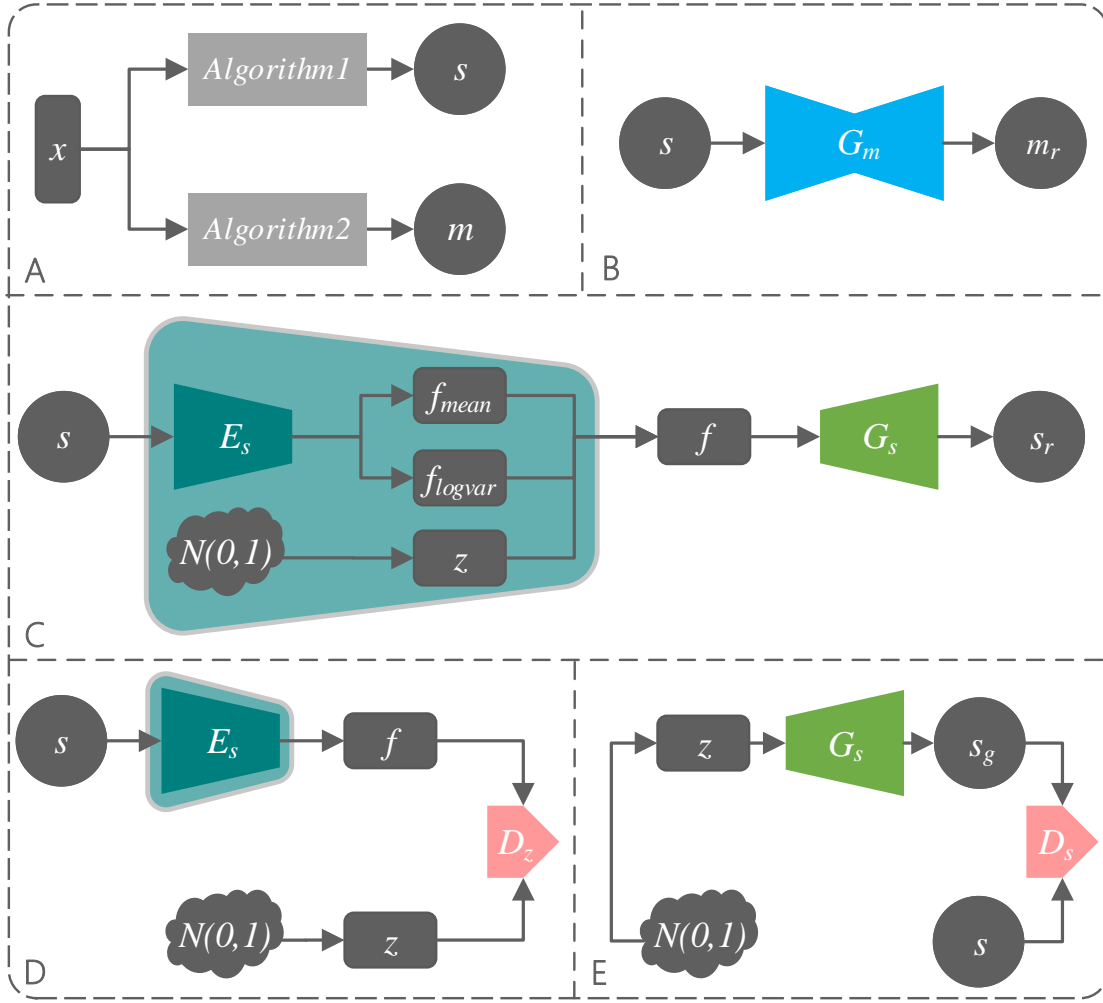


图 2: Structural feature map generation training.

Algorithm 2 Mask Extraction

- 1: Input a real image x and the expanded pixel value p
 - 2: $mask = 1.0 - ones \times (x > 0.)$
 - 3: $new_size = [x.width() + p, x.length() + p]$
 - 4: $mask = resize(mask, new_size)$
 - 5: $mask = crop_padding(mask, p)$
-

构特征图生成的训练同步。在MG训练期间，算法 2提取的掩膜用作标签数据。如图 2所示，具体的训练过程如下：

- 结构特征图 f 使用算法 1从 x 获得，掩膜 $mask$ 通过算法 2从 x 获得；
- 使用VAE编码器 EC_f 编码 f 以获得 $code_{f,mean}$ 和 $code_{f,logvar}$ ，然后从多维正态分布 $\mathcal{N}(0, 1^2)$ 中获得随机噪声 $code_n$ ，因此近似正态分布矩阵 $code_f = code_{f,mean} + exp(0.5 \times code_{f,logvar}) \times code_n$ ；
- 用VAE解码器 DC_f 解码 $code_f$ 以获得重建的结构特征图 f_r ；
- 使用掩码生成器MG从 f 中提取掩码 $mask_r$ ；
- 随机生成一个矩阵 $code_{f,g}$ 服从正态分布 $\mathcal{N}(0, 1^2)$ ；
- 用VAE解码器 DC_f 解码 $code_{f,g}$ ，以获得生成的随机结构特征图 f_g ；
- 使用遮罩生成器MG从 f 中提取遮罩 $mask_g$ ；
- 结构特征图鉴别器 D_f 分别标识 f 和 f_g ，前者是正样本，后者是负样本；
- 特征鉴别器 FD_f 标识 $code_{f,g}$ 和 $code_f$ ，前者是正样本，后者是负样本。

对于近似正态分布矩阵的约束，我们在原始的VAE编码器损失基础上添加一个由特征鉴别器 FD_f 为VAE编码器提供的对抗性损失。同时，我们还一个使用L2损失来约束均值矩阵的均值趋近于0和方差矩阵的均值趋近于1。完整的损失项如下，其中 ω_{ij} 是每个损失项的权重， $mean()$ 是均值函数：

- Part A: Structural feature map reconstruction loss

$$\mathcal{L}_r(E_s, G_s) = \mathbb{E}_{s,f} [\|s - s_r\|_2^2], \quad (1)$$

where $s_r = G_s(f)$.

- Part B: Structural feature map discriminator and generator loss

$$\mathcal{L}_{d1}(D_s) = \mathbb{E}_{s,z} [\|D_s(s) - 1\|_2^2 + \|D_s(s_g)\|_2^2], \quad (2)$$

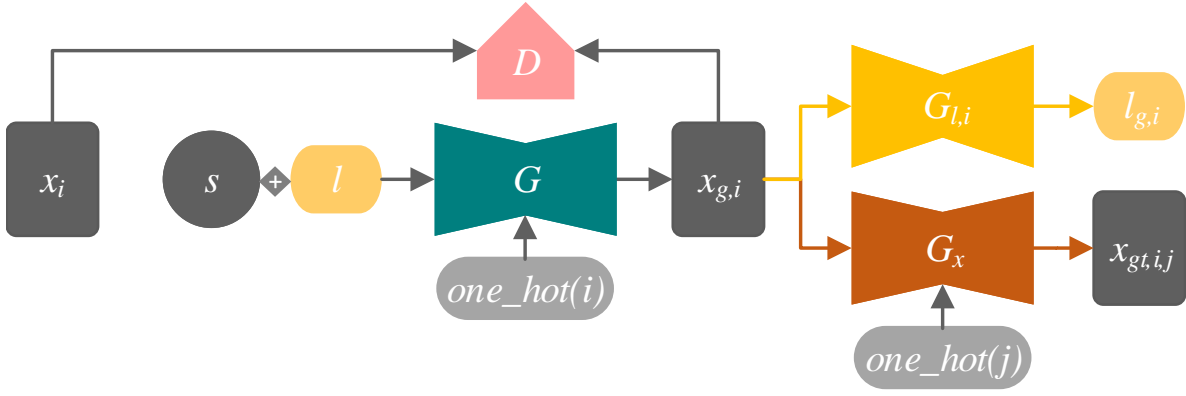


图 3: Synthesis of multimodal MRI.

$$\mathcal{L}_{g1}(E_s, G_s) = \mathbb{E}_z[\|D_s(s_g) - 1\|_2^2], \quad (3)$$

where $s_g = G_s(z)$.

- Part C: Mask reconstruction loss

$$\mathcal{L}_m(G_m) = \mathbb{E}_{m,s}[\|m - m_r\|_2^2], \quad (4)$$

where $m_r = G_m(s)$.

3.3 多模态影像的合成

多模态影像合成过程如图 3。首先，我们将结构特征图通过结构特征图编码器 EC_r 编码获得语义特征图，再将语义特征图连接到不同的独热条件矩阵作为影像解码器的输入，然后获得不同模态的合成图像。我们使用影像鉴别器提供的对抗性损失和类别指导损失来约束合成图像以逼近真实影像。

多模态影像合成过程的损失项如下，其中 d_i 和 c_i 是影像鉴别器 $D(x_i)$ ， $d_{g,i}$ ， $c_{g,i}$ 是 $D(x_{g,i})$ 的输出。 $x_{g,i}$ 是模态 i 的合成图像， $x_{gt,j,i}$ 是模态 i 的合成图像，由模态 j 的合成图像翻译而成。 $code_g$ 是由融合图编码器 EC_R 获得的语义特征图， $code_{g,i}$ 是从 x_i 编码的语义特征图。 l 是输入标签， $l_{g,i}$ 是病变标签生成组件从 $x_{g,i}$ 获得的标签。 f 是输入结构特征图， $f_{g,i}$ 是算法 1 从 $x_{g,i}$ 中提取的结构特征图：

- Discriminator loss

$$\mathcal{L}_{d2}(D) = \mathbb{E}_{x,s,l} \left[\sum_{i=0} (\|d(x_i) - 1\|_2^2 + \|d(x_{g,i})\|_2^2 + \|c(x_i) - i\|_2^2 + \|c(x_{g,i}) - i\|_2^2) \right], \quad (5)$$

where $x_{g,i} = G(f_g, i)$, $f_g = E(s, l)$.

- Adversarial and category guidance loss

$$\mathcal{L}_{g2}(E, G) = \mathbb{E}_{s,l} \left[\sum_{i=0} (\|d(x_{g,i}) - 1\|_2^2 + \|c(x_{g,i}) - i\|_2^2) \right]. \quad (6)$$

- Lesion generation loss

$$\mathcal{L}_{les}(E, G) = \mathbb{E}_{s,l} \left[\sum_{i=0} (\|l - l_{g,i}\|_2^2) \right], \quad (7)$$

where $l_{g,i} = G_{l,i}(x_{g,i})$.

- MRI registration loss

$$\mathcal{L}_{reg}(E_x, G) = \mathbb{E}_{s,l} \left[\sum_{j=0} \sum_{i=0, i \neq j} (\|x_{g,i} - x_{gt,j,i}\|_2^2) \right], \quad (8)$$

where $x_{gt,i,j} = G(f_{g,i}, j)$, $f_{g,i} = E_x(x_{g,i})$.

- Semantic consistency loss

$$\mathcal{L}_s(G, E_x) = \mathbb{E}_{s,l} \left[\sum_{i=0} (\|f_g - f_{g,i}\|_2^2) \right]. \quad (9)$$

3.4 多模态影像的配准

为了确保合成的不同模态的影像精确配准，我们将合成图像通过模态转换器在继续进行模态转换，并通过损失来约束所有语义特征图的一致性。模态转换器通过真实的多模态数据预先训练完成，如图 4所示。模态转换器训练时，编码器 EC 对模态 i 的真实影像 x_i 进行编码以获得语义特征图 $code_i$ ，则连接到不同的条件矩阵，并通过解码器 DC 获得所有模态。在循环重建中，我们使用编码器对所有获得的转换图像进行

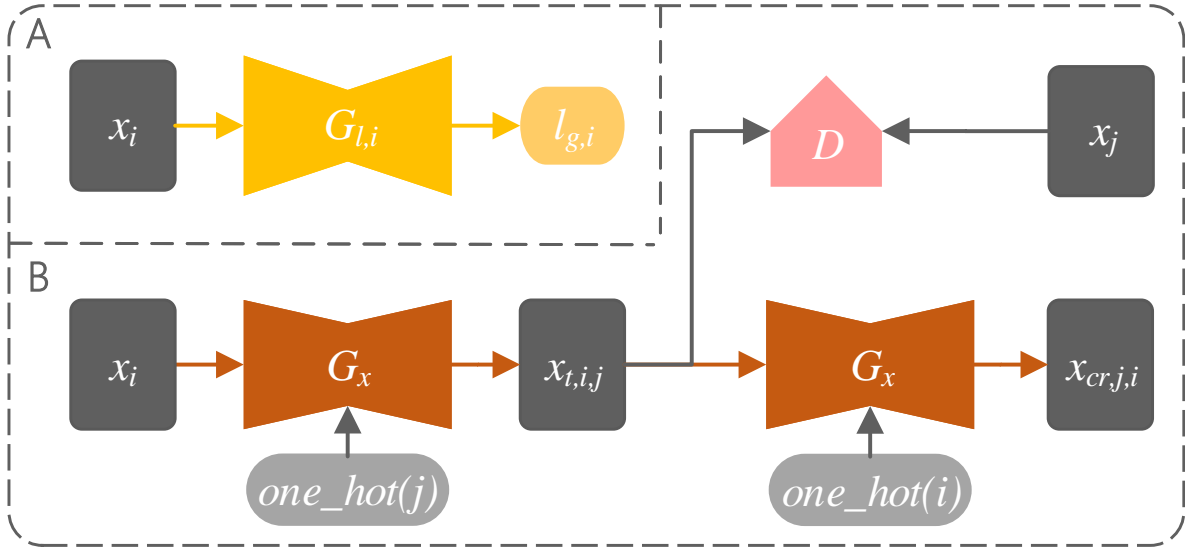


图 4: Reconstruction and translation training.

重新编码，将所有重新编码的语义特征图与模态条件向量 i 连接起来，最后由解码器对其进行解码，以获得循环重建图像 x_{rcji} 。鉴别器组件均独立更新，鉴别器训练过程如图 ?? 所示。

除了CycleGAN中的损失设置外，我们还为模态转换器的训练设计了以下损失：

- Feature Discriminator loss

$$loss_{fd} = \sum_{i=0} (\|FD(code_i) - 1\|_2^2) \quad (10)$$

- Lesion segmentation loss

$$loss_{lesion,1} = \sum_{i=0} \|label_i - label_{r,i}\|_2^2 \quad (11)$$

3.5 病灶信息的添加和合成

当我们需要在合成的多模态影像中添加指定的病灶信息时，我们需要在生成结构特征图 f_g 之后随机选择病灶信息标签 $label$ 融合再输入给多模态影像生成器。具体来说， $label$ 有 n 类病变则转换为 n 通道的独热矩阵 $onehot_l$ ，每个通道对应一个病灶类别，每个通道中的像素值为0或1，像素为1的区域为病灶区域，然后，我们计算每个通

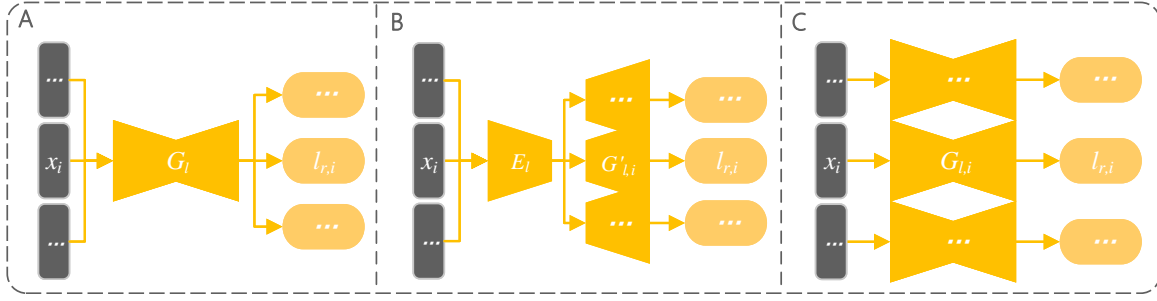


图 5: Lesion segmentation.

道 $onehot_l$ 与 f_g 的加权总和（）加权系数默认为0.2和0.8，得到一个融合了 f_g 和 $label$ 信息的新矩阵，即是融合后的输入。

值得注意的是，由于随机选择标签中指定的病灶的生成位置可能会出现在结构特征图的目标器官轮廓之外，因此我们需要使用掩膜生成器 MG 为结构特征图生成掩膜 $mask$ 。如果 $mask$ 与选定的 $label$ 的乘积为0，则病灶将位于 $mask$ 的目标器官轮廓内，可以采用该轮廓，否则需要重新选择 $label$ 。

为确保合成的多模态图像已根据输入的病变标签合成了相应的病变内容，我们使用一个病灶生成指导组件对每个合成影像的赋早进行提取，还原出输入的病灶标签。病灶生成指导组件用真实的影像和标签数据预先训练完成。

如图5中所示，每个模态均由有独立的病灶生成指导组件 $G_{l,i}$ ，训练过程损失为：

$$\mathcal{L}_{seg}(G_{l,i}) = \mathbb{E}_{l,x}[\|l_i - l_{r,i}\|_2^2], \quad (12)$$

where $l_{r,i} = G_l(x_i)$.

3.6 合成数据集的构建

如图6所示，我们可以通过经过训练的结构特征图解码器从随机正态分布矩阵生成任意数量的结构特征图。然后，我们随机缩放，旋转，平移，翻转原始标签集，以获得随机病变标签集。然后，我们将生成的结构特征图与从随机病变标签集中随机选择的标签融合。像训练阶段一样，我们可以通过掩膜生成器 MG 从结构特征图中获取蒙版，从而选择合适的标签。

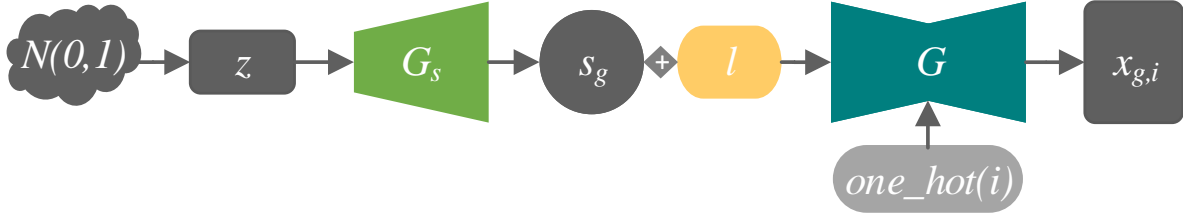


图 6: Construction of synthetic datasets.

由于存在一些结构特征图质量较差、目标器官轮廓未闭合的情况，我们为此设计了一种结构特征图过滤算法。首先，我们使用生成器 MG 生成结构特征图的掩膜。然后，我们在结构特征图上执行高斯模糊 [42]，并使用OpenCV⁶提供的轮廓搜索算法和填充算法来获得高斯的所有闭合轮廓模糊图像并填充它们。因此，我们通过传统算法获得了另一个掩膜。最后，我们计算两个掩膜的平均绝对误差（MAE）。如果MAE低于我们设置的阈值，则意味着结构特征图的目标器官轮廓相对完整，可以使用；否则，意味着通过传统算法得到的遮罩内部是空心的，与掩膜生成器产生的掩膜有很大的不同，因此需要重新生成。该算法表示为算法 3。在多模态合成影像数据集中，

Algorithm 3 Structural Feature Map Filtering

```

1: Input a MAE threshold  $mae$ 
2: function  $get\_mask(f)$ 
3:    $contours = OpenCV.findContours(f)$ 
4:    $mask = OpenCV.drawContours(f, contours)$ 
5:   return  $mask$ 
6: do
7:    $f = DC_f()$ 
8:    $m = MG(f)$ 
9:    $m' = get\_mask(f)$ 
10: while  $MAE(m', m) \leq mae$ 
    
```

仍存在病灶生成质量较差的样本。在这一点上，我们通过在真实数据上进行训练过的病变处理器对病灶进行处理，根据处理结果即可进行过滤（骰子得分的默认阈值

⁶<https://opencv.org/>

为0.95)。经过多级过滤后，我们获得了由结构特征图，掩膜，病灶标签和多模态合成影像组成的最终合成数据集。

第四章 合成医学影像的评估与展示

4.1 评估指标

我们以 [43]作为我们的基准。这项工作中，我们使用以下三个指标来评估合成医学图像的性能，包括多尺度结构相似性（MS-SSIM），Sliced Wasserstein距离（SWD），以及Fresnet Inception距离（FID）。MS-SSIM 是一种广泛使用的指标，用于测量配对图像的相似性，其中 MS-SSIM 越高，性能越好。SWD 是计算随机近似于earth mover's distance的有效指标，也用于测量 GAN 性能，其中 SWD 越低，性能越好。FID 在像素级别计算真实图像和假图像之间的距离，其中 FID 越低，性能越好。我们使用Dice Score [44]和均方误差（MSE）[45]来评估分割结果。

4.2 训练设置

每个实验的迭代次数等于训练数据集的100个epoch。学习率为 $1e-4$ ，无权重衰减。我们使用beta为1为0.5的Adam优化器并在输入层上执行0.1的Dropout，批处理大小为1。评估结果是2D图像上多模态结果的平均值，每个实验都经过了4次训练保留最佳结果。

4.3 合成方法在脑肿瘤MRI数据集上的消融实验

我们执行简单的消融实验，以改变不同组件的影响。我们进行了以随机噪声代替结构特征图作为MRI生成训练输入的实验。我们通过选择没有掩码限制的输入标签进行实验。我们进行了删除合成MRI注册的翻译培训的实验。我们进行了删除真实MRI的翻译训练的实验。我们在MRI综合训练中进行了消除病变导引丢失的实验。将以上5个实验与我们的完整方法进行比较。图7显示了在消融实验中生成的合成图像的示例。

4.4 合成方法在不同数据集上的量化评估

如表 1所示，我们的方法在各个数据集上进行了量化评估，并与当前最先进的结

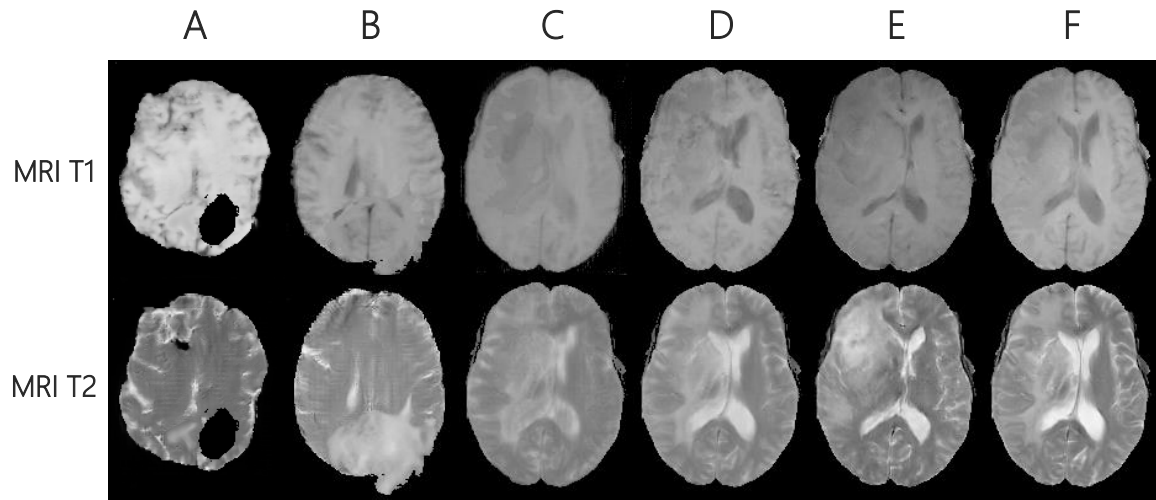


图 7: 消融实验的合成图像。模型A: 用随机噪声替换结构特征图, 而在脑轮廓上没有结构特征图的约束, 模型A生成的图像符合MRI的特征, 但不符合大脑的结构特征。B型: 取下面罩, 很容易看到B型产生的肿瘤超出了大脑的轮廓。模式C: 取消翻译培训。模式D: 取消重建训练。由模型C, D生成的图像的配准效果不是很好, 并且合成的质量也不高。模型E: 在多模式MRI综合训练过程中去除病变损失, 可以看出生成的图像中的病变与输入的病变标签不太吻合。F型: 我们完整的模型。

果进行对比。

表 1: 不同数据集上合成图像质量的量化评估（一）

Dataset	Metric	Ours	SkrGAN [43]	DCGAN [46]	ACGAN [47]	WGAN [48]	PGGAN [49]
Color Fundus	SWD ↓	xxx	0.025	0.160	0.149	0.078	0.036
	MS-SSIM ↑	xxx	0.614	0.418	0.490	0.584	0.537
	FID ↓	xxx	27.59	64.83	96.72	240.7	110.8
Chest X-ray	SWD ↓	xxx	0.026	0.118	0.139	0.196	0.031
	MS-SSIM ↑	xxx	0.506	0.269	0.301	0.401	0.493
	FID ↓	xxx	114.6	260.3	235.2	300.7	124.2
Lung CT	SWD ↓	xxx	0.020	0.333	0.317	0.236	0.057
	MS-SSIM ↑	xxx	0.359	0.199	0.235	0.277	0.328
	FID ↓	xxx	79.97	285.0	222.5	349.1	91.89

表 2: 不同数据集上合成图像质量的量化评估（二）

Dataset	Metric	Ours	SkrGAN [43]	DCGAN [46]	ACGAN [47]	WGAN [48]	PGGAN [49]
BRATS MRI	SWD ↓	xxx	0.025	0.160	0.149	0.078	0.036
	MS-SSIM ↑	xxx	0.614	0.418	0.490	0.584	0.537
	FID ↓	xxx	27.59	64.83	96.72	240.7	110.8

4.5 合成病灶的有效性验证实验

4.5.1 脑肿瘤MRI合成

我们以肿瘤分割标签为病灶标签与结构特征图融合作为输入，合成4个模态的MRI时，每个模态采用一个训练好的肿瘤分割器作为病灶生成指导组件，由肿瘤分割器提供分割结果与输入标签的自监督损失来确保病灶信息的合成。

4.5.2 对合成脑肿瘤MRI的肿瘤区域分割检验

如表 3所示，我们在BRATS2015训练数据集上训练了4个模态的肿瘤病变分割器，

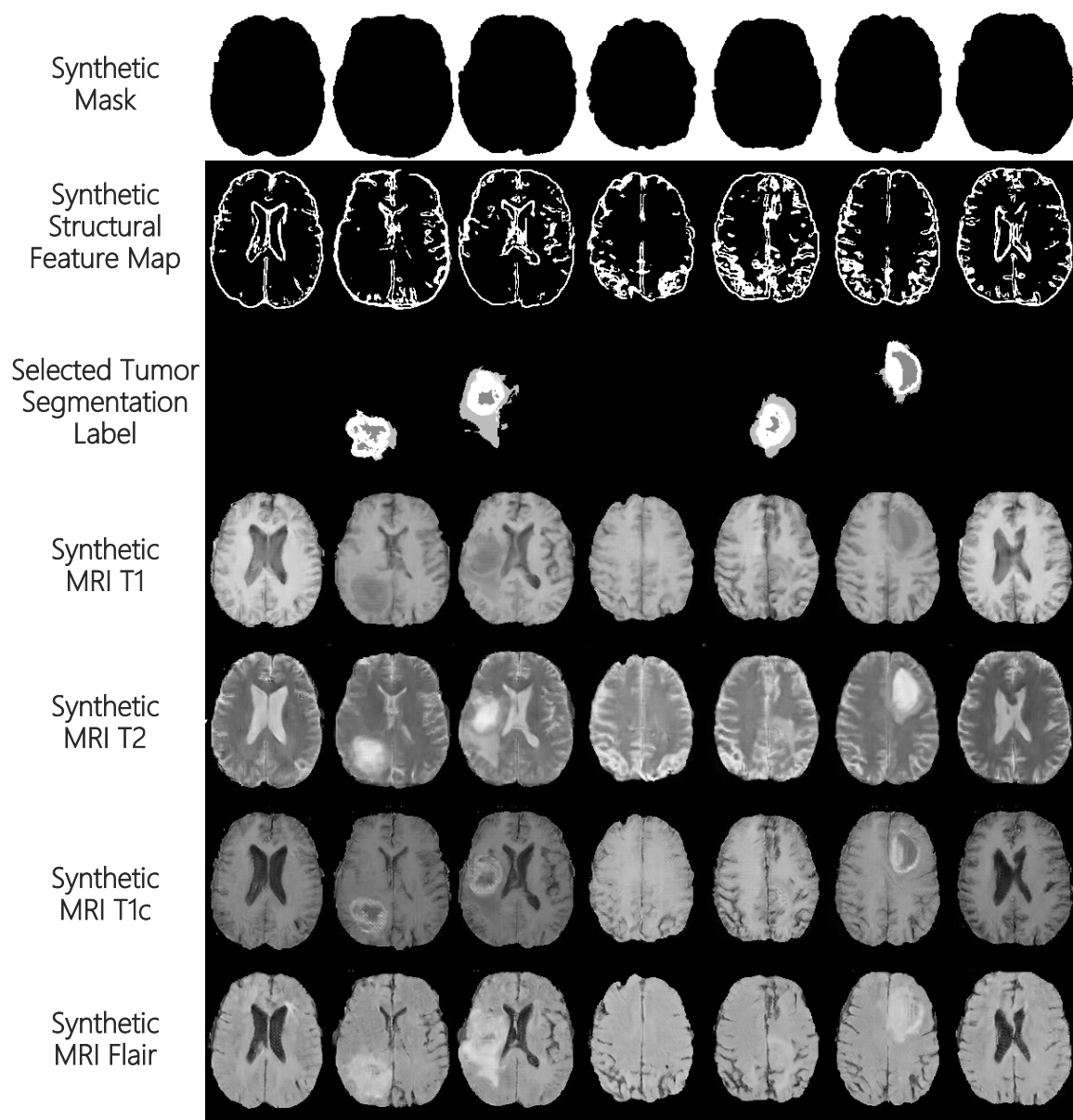


图 8: Multimodal MRIs synthesized from structural feature maps and lesion labels.

表 3: Lesion generation methods experiments.

testing dataset	MSE	Dice Score
real	0.026	0.915
synthetic	0.043	0.838

并在BRATS2015测试数据集上对其进行了测试。然后，我们使用训练有素的分割器对未过滤合成数据进行分割。

4.6 合成数据在智能医学影像处理任务中的可用性验证实验

4.6.1 合成脑肿瘤MRI数据用于肿瘤区域分割任务

如表 4所示，我们将不同量的BRATS 2015训练数据与真实BRATS合成数据混合，然后使用混合数据集进行多分割器方法进行分割训练，并在真实BRATS 2015上评估模型的分割能力测试数据集。所有实验均以相同的迭代次数进行了完全训练，该迭代次数等于BRATS2015训练数据集上的100个epoch。同时，我们设置了三种数据混合模式：随机混合，首先进行实数据训练和首先进行综合数据训练。在实优先实验和综合优先实验中，来自不同来源的数据的训练迭代次数与数量成正比数据的。除表格中的条件外，其他条件相同。如表 4所示，实验NO.3-NO.5的结果表明，合成数据不能完全替代训练中的真实数据。NO.6-NO.8的结果表明，大量合成数据的预训练和少量真实数据的微调性能与完整真实数据的训练相似。在NO.9-NO.11中，不同混合比的结果也完全不同。当两个比率接近时，分割结果与NO.1相似。当合成数据所占比例较高时，合成数据将干扰实际数据的学习，结果低于NO.1。当合成数据所占比例较低时，可以通过合成数据提高模型的泛化能力，结果高于NO.1。在NO.12-NO.17中，我们进一步尝试将不同数量的合成数据添加到真实数据中，这表明添加少量合成数据可以增强学习效果，并且合成数据越多，增强效果越好，但是当综合数据达到一定百分比然后继续增加时，它会达到相反的效果。在NO.18-NO.23中，我们比较了合成数据和通过常规数据增强方法生成的增强数据的增强效果。我们发现两者在增强效果和数据量增加之间的趋势上相似但不相等。总的来说，增强模型在模型对增强数据量的

敏感性方面更健壮，但增强效果较高。合成数据的限制远高于增强数据的限制。我们将NO.24-NO.25与NO.15进行了比较，发现合成数据用作预训练数据时性能最佳，而用作补充训练数据时性能较差。当用作增强数据与真实数据混合时，合成数据也可以实现某些增强。

通常，如果存在大量真实数据，则可以将少量合成数据用作增强数据，或者可以将大量合成数据用于预训练，然后对真实数据进行训练。如果真实数据较少，则可以使用大量的综合数据进行预训练，然后对少量真实数据进行微调，其结果可以与完整真实数据的结果相抗衡，因此得出的结论是与 [12]一致。我们不建议将合成数据完全用于训练，也不建议将合成数据用于补充训练。

4.6.2 合成眼底视网膜数据用于视网膜血管注释任务

4.6.3 合成胸部X光线数据用于肺炎分类任务

4.6.4 合成肺部CT数据用于肺结节检测任务

4.7 合成数据效果展示

4.7.1 正态分布效果展示

4.7.2 多样性效果展示

4.7.3 指定病灶生成效果展示

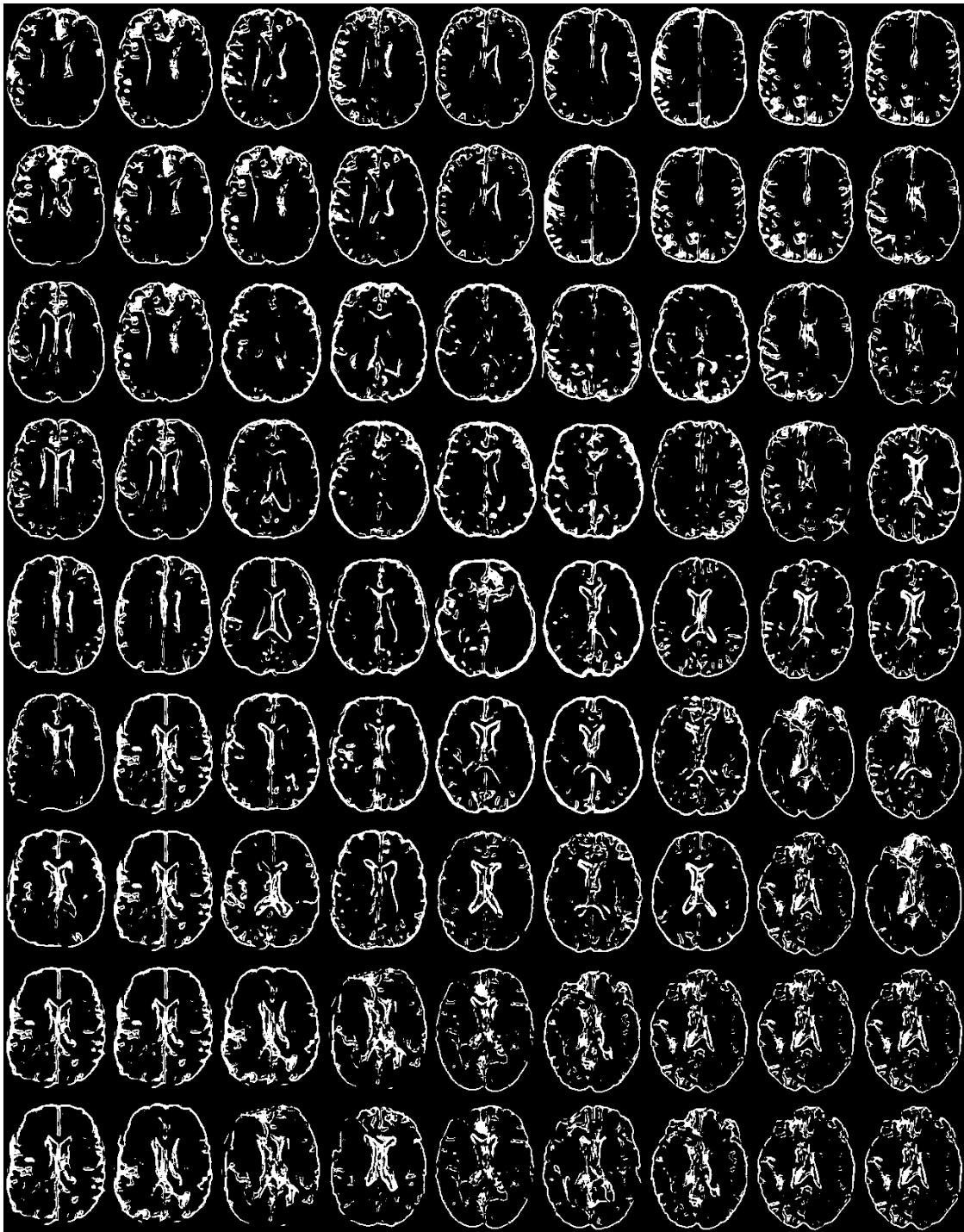


图 9: Synthetic structural feature maps from continuous sampling.

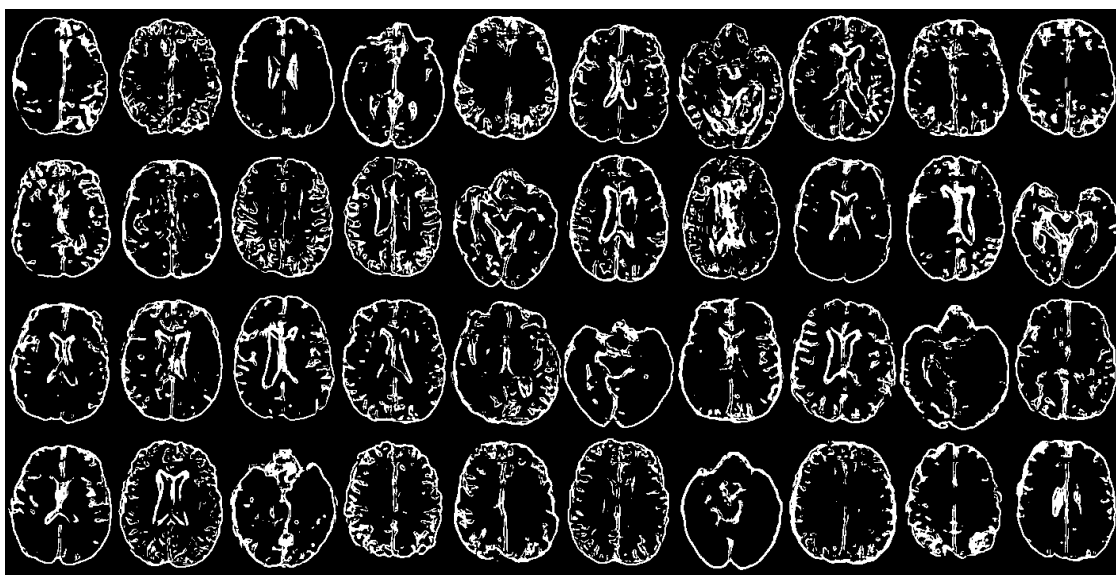


图 10: Synthetic structural feature maps from random sampling.

第五章 结语

简而言之，我们提出了一种结构特征图提取方法，无需训练或附加标签数据即可直接从医学图像中提取解剖结构信息。我们提出一种结构特征图生成方法，以从多维正态分布矩阵生成结构特征图。我们通过无监督训练实现了从随机正态分布矩阵合成配准的多模式MRI图像，从而可以自由添加病变信息。我们验证了合成MRI图像可用作智能医学图像处理任务的预训练数据或增强数据，并且可以通过病变分割实验显著提高模型的泛化能力。在本文中，我们的贡献如下：

- 我们提出一种结构特征图提取方法，以直接从医学图像中提取解剖结构信息，而无需训练或附加标签数据。
- 我们提出一种结构特征图生成方法，以从多维正态分布矩阵生成结构特征图。
- 我们实现了从结构特征图和随机选择的病变标签中合成具有相应病变信息的配准多模态医学图像。
- 我们通过许多数据可用性测试，验证合成数据可以用作智能医学图像处理任务的预训练数据或增强数据。

参考文献

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *neural information processing systems*, 141(5):1097–1105, 2012.
- [2] Ian J Goodfellow, Jean Pougetabadie, Mehdi Mirza, Bing Xu, David Wardefarley, Sherjil Ozair, Aaron C Courville, and Yoshua Bengio. Generative adversarial nets. *neural information processing systems*, pages 2672–2680, 2014.
- [3] Geert J S Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A W M Van Der Laak, Bram Van Ginneken, and Clara I Sanchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.
- [4] Junegoo Lee, Sanghoon Jun, Youngwon Cho, H S Lee, Guk Bae Kim, Joon Beom Seo, and Namkug Kim. Deep learning in medical imaging: General overview. *Korean Journal of Radiology*, 18(4):570–584, 2017.
- [5] Dinggang Shen, Guorong Wu, and Heung Il Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19(1):221–248, 2017.
- [6] Phillip Isola, Junyan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *computer vision and pattern recognition*, pages 5967–5976, 2017.
- [7] Mingyu Liu, Thomas M Breuel, and Jan Kautz. Unsupervised image-to-image translation networks. *neural information processing systems*, pages 700–708, 2017.
- [8] Taeksoo Kim, Moon-su Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. *international conference on machine learning*, pages 1857–1865, 2017.

- [9] Yizhe Zhu, Mohamed Elhoseiny, Bingchen Liu, and Ahmed M Elgammal. Imagine it for me: Generative adversarial approach for zero-shot learning from noisy texts. *arXiv: Computer Vision and Pattern Recognition*, 2017.
- [10] He Zhang and Vishal M Patel. Densely connected pyramid dehazing network. *computer vision and pattern recognition*, pages 3194–3203, 2018.
- [11] Yunye Gong, Srikrishna Karanam, Ziyang Wu, Kuanchuan Peng, Jan Ernst, and Peter C Doerschuk. Learning compositional visual concepts with mutual consistency. *computer vision and pattern recognition*, pages 8659–8668, 2018.
- [12] Hoochang Shin, Neil A Tenenholtz, Jameson K Rogers, Christopher G Schwarz, Matthew L Senjem, Jeffrey L Gunter, Katherine P Andriole, and Mark Michalski. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. *arXiv: Computer Vision and Pattern Recognition*, pages 1–11, 2018.
- [13] Yuankai Huo, Zhoubing Xu, Shunxing Bao, Albert Assad, Richard G Abramson, and Bennett A Landman. Adversarial synthesis learning enables segmentation without target modality ground truth. *international symposium on biomedical imaging*, pages 1217–1220, 2018.
- [14] Juan Eugenio Iglesias, Ender Konukoglu, Darko Zikic, Ben Glocker, Koen Van Leemput, and Bruce Fischl. Is synthesizing mri contrast useful for inter-modality analysis? *medical image computing and computer assisted intervention*, pages 631–638, 2013.
- [15] Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Joshua Susskind, Wenda Wang, and Russell Webb. Learning from simulated and unsupervised images through adversarial training. *computer vision and pattern recognition*, pages 2242–2251, 2017.
- [16] Bo Zhao, Bo Chang, Zequn Jie, and Leonid Sigal. Modular generative adversarial networks. *european conference on computer vision*, pages 157–173, 2018.

- [17] Xiaodan Liang, Hao Zhang, Liang Lin, and Eric P Xing. Generative semantic manipulation with mask-contrasting gan. *eupean conference on computer vision*, pages 574–590, 2018.
- [18] Yunje Choi, Minje Choi, Munyoung Kim, Jungwoo Ha, Sunghun Kim, and Jaegul Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. *computer vision and pattern recognition*, pages 8789–8797, 2018.
- [19] Guim Perarnau, Joost Van De Weijer, Bogdan Raducanu, and Jose M Alvarez. Invertible conditional gans for image editing. *arXiv: Computer Vision and Pattern Recognition*, 2016.
- [20] Asha Anoosheh, Eirikur Agustsson, Radu Timofte, and Luc Van Gool. Combogan: Unrestrained scalability for image domain translation. *computer vision and pattern recognition*, pages 783–790, 2018.
- [21] Amelie Royer, Konstantinos Bousmalis, Stephan Gouws, Fred Bertsch, Inbar Mosseri, Forrester Cole, and Kevin P Murphy. Xgan: Unsupervised image-to-image translation for many-to-many mappings. *arXiv: Computer Vision and Pattern Recognition*, 2018.
- [22] Zizhao Zhang, Lin Yang, and Yefeng Zheng. Translating and segmenting multimodal medical volumes with cycle- and shape-consistency generative adversarial network. *computer vision and pattern recognition*, pages 9242–9251, 2018.
- [23] Dong Nie, Roger Trullo, J Lian, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. Medical image synthesis with context-aware generative adversarial networks. *medical image computing and computer assisted intervention*, pages 417–425, 2017.
- [24] Ninon Burgos, M J Cardoso, Filipa Guerreiro, Catarina Veiga, Marc Modat, J McClelland, Antjechristin Knopf, Shonit Punwani, D Atkinson, Simon R Arridge, et al. Robust ct synthesis for radiotherapy planning: Application to the head and neck region. *medical image computing and computer assisted intervention*, pages 476–484, 2015.

- [25] Konstantinos Kamnitsas, Christian F Baumgartner, Christian Ledig, Virginia Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Aditya V Nori, Antonio Criminisi, Daniel Rueckert, et al. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. *information processing in medical imaging*, pages 597–609, 2017.
- [26] Yawen Huang, Ling Shao, and Alejandro F Frangi. Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding. *computer vision and pattern recognition*, pages 5787–5796, 2017.
- [27] Raviteja Vemulapalli, Hien Van Nguyen, and Shaohua Kevin Zhou. Unsupervised cross-modal synthesis of subject-specific scans. *international conference on computer vision*, pages 630–638, 2015.
- [28] Hien Van Nguyen, Kevin S Zhou, and Raviteja Vemulapalli. Cross-domain synthesis of medical images using efficient location-sensitive deep network. *medical image computing and computer assisted intervention*, pages 677–684, 2015.
- [29] Anton Osokin, Anatole Chessel, Rafael E Carazo Salas, and Federico Vaggi. Gans for biological image synthesis. *international conference on computer vision*, pages 2252–2261, 2017.
- [30] Junyan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *international conference on computer vision*, pages 2242–2251, 2017.
- [31] Thomas Joyce, Agisilaos Chatsias, and Sotirios A Tsaftaris. Robust multi-modal mr image synthesis. *medical image computing and computer assisted intervention*, pages 347–355, 2017.

- [32] Pedro Costa, Adrian Galdran, Maria Ines Meyer, Michael D Abramoff, Meindert Niemeijer, Ana Maria Mendonca, and Aurelio Campilho. Towards adversarial retinal image synthesis. *arXiv: Computer Vision and Pattern Recognition*, 2017.
- [33] Zhiwen Fan, Liyan Sun, Xinghao Ding, Yue Huang, Congbo Cai, and John Paisley. A segmentation-aware deep fusion network for compressed sensing mri. *european conference on computer vision*, pages 55–70, 2018.
- [34] Rushil Anirudh, Hyojin Kim, Jayaraman J Thiagarajan, K Aditya Mohan, Kyle M Champley, and Timo Bremer. Lose the views: Limited angle ct reconstruction via implicit sinogram completion. *computer vision and pattern recognition*, pages 6343–6352, 2018.
- [35] Chenyu You, Guang Li, Yi Zhang, Xiaoliu Zhang, Hongming Shan, Shenghong Ju, Zhen Zhao, Zhuiyang Zhang, Wenxiang Cong, and Michael W. Vannier. Ct super-resolution gan constrained by the identical, residual, and cycle learning ensemble(gan-circle). *arXiv preprint arXiv:1808.04256*, 2018.
- [36] Qing Lyu, Chenyu You, Hongming Shan, and Ge Wang. Super-resolution mri through deep learning. *arXiv preprint arXiv:1810.06776*, 2018.
- [37] Agisilaos Chartsias, Thomas Joyce, Mario Valerio Giuffrida, and Sotirios A Tsaftaris. Multimodal mr synthesis via modality-invariant latent representation. *IEEE Transactions on Medical Imaging*, 37(3):803–814, 2018.
- [38] Shun Miao, Sebastien Piat, Peter Walter Fischer, Ahmet Tuysuzoglu, Philip Mewes, Tommaso Mansi, and Rui Liao. Dilated fcn for multi-agent 2d/3d medical image registration. *national conference on artificial intelligence*, pages 4694–4701, 2018.
- [39] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *international conference on learning representations*, 2014.

- [40] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *international conference on learning representations*, pages 1278–1286, 2014.
- [41] Bjoern Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lenczi, Elisabeth Gerstner, Marc-Andre Weber, Tal Arbel, Brian Avants, Nicholas Ayache, Patricia Buendia, Louis Collins, Nicolas Cordier, Jason Corso, Antonio Criminisi, Tilak Das, Hervé Delingette, Cagatay Demiralp, Christopher Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan Iftekharuddin, Raj Jena, Nigel John, Ender Konukoglu, Danial Lashkari, Jose Antonio Mariz, Raphael Meier, Sergio Pereira, Doina Precup, S. J. Price, Tammy Riklin-Raviv, Syed Reza, Michael Ryan, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos Silva, Nuno Sousa, Nagesh Subbanna, Gabor Szekely, Thomas Taylor, Owen Thomas, Nicholas Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, page 33, 2014.
- [42] Alle Meije Wink and Jos B T M Roerdink. Denoising functional mr images: a comparison of wavelet denoising and gaussian smoothing. *IEEE Transactions on Medical Imaging*, 23(3):374–387, 2004.
- [43] Tianyang Zhang, Huazhu Fu, Yitian Zhao, Jun Cheng, Mengjie Guo, Zaiwang Gu, Bing Yang, Yuting Xiao, Shenghua Gao, and Jiang Liu. Skrgan: Sketching-rendering unconditional generative adversarial networks for medical image synthesis. *arXiv: Computer Vision and Pattern Recognition*, 2019.
- [44] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

- [45] N G N Prasad and J N K Rao. The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85(409):163–171, 1990.
- [46] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv: Learning*, 2015.
- [47] Augustus Odena, Christopher Olah, and Jonathon Shlens. Conditional image synthesis with auxiliary classifier gans. *arXiv: Machine Learning*, 2016.
- [48] Martin Arjovsky, Soumith Chintala, and Leon Bottou. Wasserstein generative adversarial networks. pages 214–223, 2017.
- [49] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv: Neural and Evolutionary Computing*, 2017.

致谢

致谢

瞿毅力

二零二零年四月

表 4: Synthetic data availability verification experiments.

NO.	real data	synthetic data	enhanced data	mixing modes	MSE	Dice Score
1	$\times 1$	0	0	-	0.026	0.915
2	$\times 50\%$	0	0	-	0.032	0.902
3	0	$\times 1$	0	-	0.205	0.708
4	0	$\times 2$	0	random mixing	0.206	0.736
5	0	$\times 3$	0	random mixing	0.205	0.754
6	$\times 10\%$	$\times 1$	0	synthetic first	0.031	0.908
7	$\times 10\%$	$\times 2$	0	synthetic first	0.028	0.907
8	$\times 10\%$	$\times 3$	0	synthetic first	0.030	0.907
9	$\times 20\%$	$\times 80\%$	0	random mixing	0.041	0.850
10	$\times 50\%$	$\times 50\%$	0	random mixing	0.031	0.904
11	$\times 80\%$	$\times 20\%$	0	random mixing	0.024	0.935
12	$\times 1$	$\times 20\%$	0	random mixing	0.025	0.921
13	$\times 1$	$\times 50\%$	0	random mixing	0.023	0.939
14	$\times 1$	$\times 80\%$	0	random mixing	0.026	0.916
15	$\times 1$	$\times 1$	0	random mixing	0.027	0.913
16	$\times 1$	$\times 2$	0	random mixing	0.033	0.901
17	$\times 1$	$\times 3$	0	random mixing	0.034	0.897
18	$\times 1$	0	$\times 20\%$	random mixing	0.027	0.911
19	$\times 1$	0	$\times 50\%$	random mixing	0.025	0.927
20	$\times 1$	0	$\times 80\%$	random mixing	0.026	0.920
21	$\times 1$	0	$\times 1$	random mixing	0.026	0.915
22	$\times 1$	0	$\times 2$	random mixing	0.032	0.898
23	$\times 1$	0	$\times 3$	random mixing	0.036	0.885
24	$\times 1$	$\times 1$	0	real first	0.195	0.795
25	$\times 1$	$\times 1$	0	synthetic first	0.021	0.940