

# 基于条件生成对抗网络的配准多模态脑MRI合成

Yili Qu, Wanqi Su, Chufu Deng,  
Ying Wang, Yutong Lu, Nong Xiao, Zhiguang Chen

School of Data and Computer Science

Sun Yet-Sen University

quyli@mail2.sysu.edu.cn

## Abstract

在基于大量数据驱动的医学影像智能处理任务中，医学影像数据的收集和采集是非常困难的，尤其是配准的多模态医学影像数据。合成的医学影像数据可以很好地缓解数据不足的问题。我们基于无监督的条件生成对抗网络模型实现了完全从随机噪声生成配准的多模态医学影像，并且可以根据自由选定的病灶标签有效地生成对应的病灶信息。我们在BRATS2015上进行了多项验证实验，验证了我们的合成MRI可以在医学影像智能处理任务中作为预训练数据或增强数据使用，并能大幅度提高模型的泛化能力。

## 1 Introduction

核磁共振成像（MRI）是一种常见的医学影像，根据成像参数的不同可以有多种模态，例如T1、T2、T1c等。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）(Krizhevsky, Sutskever, and Hinton 2012)或生成对抗网络（GAN）(Goodfellow et al. 2014)进行的医学图像处理任务。

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。对于医学影像智能处理任务的研究者来说，多模态的MRI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。

在GAN流行之前，研究者们关注于将已有模态的医学图像转换为其他模态的医学图像(Burgos et al. 2015),(Huang, Shao, and Frangi 2017),(Vemulapalli, Van Nguyen, and Zhou 2015),(Van Nguyen, Zhou, and Vemulapalli 2015)。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(Zhang, Yang, and Zheng 2018),(Nie et al. 2017),(Osokin et al. 2017),(Van Nguyen, Zhou, and Vemulapalli 2015),(Kamnitsas et al. 2017)。许多研究使用GAN实现了更高质量的转换结果(Zhao et al. 2018),(Liang et al. 2018),(Zhu et al. 2017),(Choi et al. 2018)。最近一些研究实现了基于未配准的多模态数据的转换(Zhang, Yang, and Zheng 2018),(Joyce, Chartsias, and Tsaftaris 2017)。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(Nie et al. 2017),(Kamnitsas et al. 2017)、视网膜血管注释到图像的转换(Costa et al. 2017)、采用CycleGAN(Zhu et al. 2017)的无监

督心脏MRI到CT图像的相互转换与分割(Nie et al. 2017)等。对于多模态影像的合成, (Chartsias et al. 2018)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, (Joyce, Chartsias, and Tsaf-taris 2017)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。(Miao et al. 2018)针对医学图像配准进行了深入研究。(Shin et al. 2018)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, (Costa et al. 2017)研究了基于变分自编码器(VAE)(Kingma and Welling 2014),(Rezende, Mohamed, and Wierstra 2014)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。在医学影像处理领域之外, 多域转换的研究十分踊跃(Zhao et al. 2018),(Liang et al. 2018),(Choi et al. 2018),(Isola et al. 2017)。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(Zhang, Yang, and Zheng 2018),(Nie et al. 2017),(Bur-gos et al. 2015),(Vemulapalli, Van Nguyen, and Zhou 2015),(Osokin et al. 2017),(Van Nguyen, Zhou, and Vem-ulapalli 2015),(Kamnitsas et al. 2017), 对多模态的研究依然稀少(Chartsias et al. 2018),(Joyce, Chartsias, and Tsaf-taris 2017),(Shin et al. 2018)。

目前针对医学影像合成的研究存在模态数量难以扩展、需要配准训练数据、依赖于复杂的大型网络、无法添加或保留病灶、无法从随机矩阵开始生成、需要额外的训练数据等各项问题, 且大多数研究对合成数据的评价依赖于经验医师的人工视觉效果评估, 没有进行客观的量化检验。因此, 我们设计了一种基于条件生成对抗网络(CGAN)(Mirza and Osindero 2014)的配准多模态MRI生成的方法, 采用无监督学习方法, 训练数据无需配准, 输入随机正态分布矩阵进而生成一组有病灶标签的多模态配准MRI。我们在BRATS2015数据集上进行了带肿瘤分割标签的脑多模态MRI生成实验, 并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肿瘤病灶分割实验中的可用程度。我们将开源我们的代码。总的来说, 我们的工作主要体现在以下三个方面:

- **结构特征图的提取与随机生成** 我们针对脑MRI提出了一种解剖结构特征的提取方法, 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。我们训练了一个结构特征图生成器实现了从多维正态分布生成结构特征图。我们的提取方法可以直接获取真实影像的解剖结构特征, 在提升合成影像质量同时不带来额外参数、计算开销小, 而随机生成方法可以无限地生成丰富多样的结构特征图。
- **带标签多模态配准影像的合成** 我们使用随机生成的结构特征图, 融合随机的病灶标签, 通过生成器合成配准的多模态MRI。我们探讨了多种病灶生成指导方法, 并通过病灶生成指导方法实现了在多模态MRI的合成过程中根据输入病灶标签有效地生成对应的病灶信息。在训练时我们无需配准的数据, 除病灶标签外无需额外的标签数据, 而合成的数据是配准的, 输入的随机病灶标签即为合成数据的病灶标签。我们的方法能够便捷快速地构建带标签的配准多模态MRI数据集。
- **合成数据可用性的客观验证方法** 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分割网络, 验证了合成数据可以在医学影像智能处理任务可以作为预训练数据和增强数据来提高模型的泛化能力从而提高分割精度。对比传统合成影像质量的主观评价方法, 我们更加客观地呈现了合成数据在智能病灶处理任务中的可用性。

## 2 方法

我们在多模态脑MRI的合成任务上展示我们的方法。在本文的实例中, 我们合成的病灶为肿瘤, 病灶标签为肿瘤分割标签, 病灶处理任务为病灶分割。我们的方法对合成部位、病灶类型、病灶处理任务、具体的模态和模态数量等不做限制, 通过下述实例的展示, 我们能很容易的将该方法推广应用到其他的类似任务。

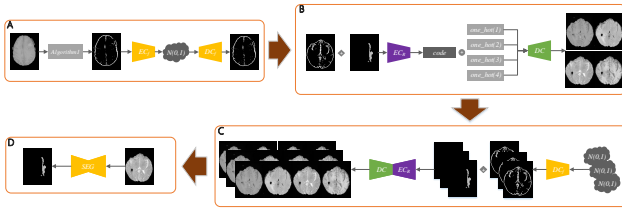


图 1: 整体架构图.Overall Architecture.

## 2.1 整体架构

如图 1所示，我们的方案包括结构特征图提取和生成、多模态MRI生成、构建合成数据集、合成数据可用性验证四个主要阶段。

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，能从随机的正态分布矩阵生成结构特征图。该阶段我们训练的模型模包括一个结构特征图编码器、一个结构特征图解码器、一个结构特征图鉴别器、一个编码分布鉴别器和一个结构特征图掩膜生成器。

多模态MRI生成阶段我们产出一个条件生成器，其以结构特征图为输入，能根据不同的独热条件向量生成不同模态的MRI，并且可在结构特征图上添加病灶标签使得生成的MRI具有对应的病灶信息。该阶段我们训练的模型模块包括一个结构特征图与病灶标签的融合图编码器、不同模态的病灶分割器、一个MRI编码器、一个MRI解码器、一个MRI鉴别器和一个MRI编码鉴别器。

在构建合成数据集阶段，我们使用前两个阶段产出的模型先从随机正态分布矩阵生成足量的结构特征图，再与随机病灶标签进行信息融合，最后通过条件生成器生成配准的多模态MRI，从而构建出一个合成数据集。

在合成数据可用性验证阶段，我们首先根据真实的数据为每个MRI模态单独训练一个病灶分割器，并在真实数据集中进行分割能力测试，再用该分割器对采用不同的病灶生成组件来指导病灶生成的合成数据进行分割测试。然后，我们使用由不同数据量的合成数据和真实数据构建的数据集来对病灶分割网络进行训练，训练充分后再在真实测试数据集上进行分割测试，对比各项测试结果，以验证合成数据在肿瘤病灶分割训练中的可用性。

## 2.2 结构特征图提取方法

直接从随机噪声通过生成对抗训练生成的医学影像通常训练困难且难以生成真实的结构信息。我们将医学影像中提供基本轮廓和结构信息的图像称为其结构特征图，例如视网膜血管分布图可视为视网膜图像的结构特征图(Costa et al. 2017)。结构特征图可以为医学影像的合成提供必要的基础指导信息，例如合成脑部MRI图像时一些研究从脑分割标签图获取基本的结构信息(Shin et al. 2018)。然而，视网膜血管分布图和脑分割标签图等常用的结构特征图都需要额外的数据和训练才能实现从原图提取出结构特征图。为此，我们首先设计了下述直接从脑MRI提取结构特征图的方法，该方法具有运算快、无需训练、无需额外数据等优点。

在传统的数字图像处理方法中，Roberts算子(Roberts 1965)、Prewitt算子(Prewitt 1970)、Sobel算子(Sobel 1970)等是十分优秀的边缘检测算子，其中Sobel算子常用于脑部医学图像的处理，其卷积核参数和计算公式如图所示。我们探索出了从Sobel算子生成的边缘检测图中进一步提取结构特征的方法，如算法 1所示。

### Algorithm 1 Structural Feature Extraction

- 1: Input a real image  $x$ ,  $\beta$  is pixel threshold
- 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:  $f1 = \text{ones} * (f1 > \beta)$
- 7:  $f2 = \text{ones} * (f2 > \beta)$
- 8:  $f = f1 + f2$
- 9:  $f = \text{ones} * (f > 0.)$

在算法 1中，我们对一张真实图像用Sobel算子提取得到其横向和纵向的边缘检测图，对两张边缘检测图进行最大值规约和最小值规约得到两张新的边缘检测融合图，然后两张边缘检测融合图分别与各自的平均像素值求差，再对两张差值图根据设定像素阈值进行二值化处理，两张二值图求和后再进行完全的二值化，最后得到的就是我们需要的结构

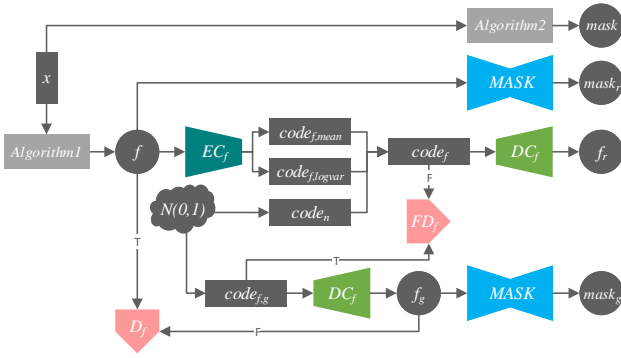


图 2: 随机结构特征图的生成训练.Training of Random Structural Feature Map Generation.

特征图。

### 2.3 随机结构特征图的生成训练

在生成结构特征图时, (Shin et al. 2018)仍然需要真实的MRI作为输入来得到生成的结构特征图, 这大大降低了生成数据的多样性, (Costa et al. 2017)实现了一种基于VAE的从多维正态分布生成视网膜血管分布图的方法。我们结合了VAE与GAN的特点设计了一种混合网络, 同样从随机正态分布矩阵生成脑部结构特征图, 无需额外数据且具有更好的多样性。此外, 为了防止通过结构特征图生成的脑MRI像素区域超出结构特征图的脑轮廓线之外, 我们还训练了一个从脑结构特征图获取脑部区域掩膜的生成器MASK, 该生成器与结构特征图的生成训练进行同步训练。训练MASK时, 将真实脑MRI $x$ 通过掩膜提取算法提取得到的掩膜作为训练标签数据, 掩膜的提取算法如算法2所示。如图 2所示, 从正态分

#### Algorithm 2 Mask Extraction

- 1: Input a real image  $x$ ,  $p$  is expanded pixel value
- 2:  $mask = 1.0 - ones * (x > 0.)$
- 3:  $shape = get\_shape(x)$
- 4:  $mask = resize(mask, size = [shape[1] + p, shape[2] + p])$
- 5:  $mask = crop\_padding(mask, crop\_length = p, crop\_width = p)$

布矩阵解码得到随机结构特征图的具体处理过程如下:

- 从真实MRI $x$ 中用算法 1得到结构特征图 $f$ , 用算法 2生成掩模 $mask$ ;
- 用VAE编码器 $EC_f$ 对结构特征图 $f$ 进行编码获得 $code_{f,mean}$ 及 $code_{f,logvar}$ , 从正态分布 $\mathcal{N}(0,1^2)$ 的获取随机矩阵 $code_n$ , 由三个编码求得近似正态分布矩阵 $code_f = code_{f,mean} + exp(0.5 * code_{f,logvar}) * code_n$ ;
- 用VAE解码器 $DC_f$ 对 $code_f$ 解码得到重建的结构特征图 $f_r$ ;
- 用掩模生成器MASK从 $f$ 提取得到掩模 $mask_r$ ;
- 随机生成符合正态分布 $\mathcal{N}(0,1^2)$ 的矩阵 $code_{f,g}$ ;
- 用VAE解码器 $DC_f$ 对 $code_{f,g}$ 解码得到生成的随机结构特征图 $f_g$ ;
- 用掩模生成器MASK对 $f_g$ 提取得到掩模 $mask_g$ ;
- 结构特征图鉴别器 $D_f$ 分别对 $f$ 和 $f_g$ 进行鉴别学习, 前者为正样本, 后者为负样本;
- 编码矩阵分布鉴别器 $FD_f$ 分别对 $code_f$ 和 $code_{f,g}$ 进行鉴别, 前者为负样本, 后者为正样本。

其中, 对于近似正态分布矩阵的损失约束, 我们没有采用VAE原本的编码器损失, 而是通过一个编码矩阵分布鉴别器为编码器提供了对抗性损失。此外, 我们还通过L2正则损失指导均值矩阵的均值逼近0值, 标准差矩阵的均值逼近1值。训练过程中的各项损失函数如下, 其中,  $\omega_{i,j}$ 为各损失项的权重:

- 编码鉴别器损失Discriminator Loss of Code Distribution

$$loss_{FD_f} = \|FD_f(code_{f,g}) - 1\|_2^2 + \|FD_f(code_f)\|_2^2$$

- 结构特征图鉴别器损失Discriminator Loss of Structural Feature Map

$$loss_{D_f} = \|D_f(f) - 1\|_2^2 + \|D_f(f_g)\|_2^2$$

- 对抗性损失Adversarial Loss

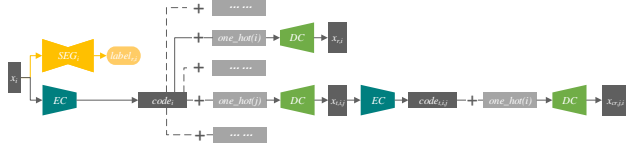


图 3: 辅助的模态重建和模态转换训练.Auxiliary modality reconstruction and modality translation training.

$$loss_{G_f} = \|FD_f(code_f) - 1\|_2^2 + \|D_f(f_g) - 1\|_2^2$$

- 结构特征编码的分布监督损失Supervised Loss of Structural Feature Code Distribution

$$loss_{normal} = \|\text{mean}(code_{f,mean})\|_2^2 + \|\text{mean}(\exp(0.5 * code_{f,logvar})) - 1\|_2^2$$

其中,  $\text{mean}()$  函数为均值函数。 where  $\text{mean}()$  is a mean function.

- 结构特征图及掩模的自监督损失Self-supervised Loss of Structural Feature Map and Mask

$$loss_{sv} = \|f - f_r\|_2^2 + \|f_r * mask\|_2^2$$

- 掩膜生成器损失Mask Generator Loss

$$loss_{mask} = \|mask - mask_r\|_2^2 + \|f * mask_r\|_2^2 + \|f_r * mask_r\|_2^2 + \|f_g * mask_g\|_2^2$$

## 2.4 真实MRI重建和转换训练

我们在真实数据上进行重建和转换训练, 并与后续章节的多模态MRI生成训练同步, 以此来约束各个组件在多模态MRI生成的过程中完成我们指定的任务。此外我们还通过一个编码鉴别器来对两个训练过程进行一致性指导。鉴别器训练过程如图 5 所示。

如图 3 所示, MRI重建和转换时, 编码器将模态  $i$  的真实MRI  $x_i$  编码得到语义特征图  $code_i$ , 然后我们将其与不同的条件向量堆叠, 通过解码器解码出全部的模态。循环重建时, 我们对所得到的转换图采用编码器全部进行再编码, 将全部再编码得到的

语义特征图均与模态  $i$  的条件向量进行连接, 最后再用解码器全部解码得到循环重建的  $x_{rc,j,i}$ 。MRI重建的循环重建都是自监督训练。在上述过程中, 我们以原始输入模态  $x_i$  对应的病灶标签  $l_i$  作为病灶生成训练的监督标签, 对  $x_i$  用病灶标签生成组件得到  $l_{r,i}$ 。

我们的鉴别器组件独立更新, 其他组件通过一个优化器更新训练, 损失项包括鉴别器提供的对抗性损失、MRI重建自监督损失、MRI循环重建自监督损失、MRI循环重建一致性损失、语义一致性损失、病灶生成监督损失。详细损失如下, 其中  $x_{r,i}$  表示模态  $i$  重建得到的MRI,  $x_{t,j,i}$  指由模态  $j$  转换生成的模态  $i$  的MRI,  $d_{t,j,i}$  和  $c_{t,j,i}$  分别为鉴别器对  $x_{t,j,i}$  的真假鉴别和类别鉴别结果,  $x_{cr,j,i}$  表示模态  $i$  转换为模态  $j$  再转换回模态  $i$  的MRI;  $code_i$  表示编  $x_i$  编码器编码后得到的语义特征图,  $code_{t,i,j}$  表示  $x_i$  转换生成的模态  $j$  的MRI再经过编码器编码后得到的语义特征图;  $l_i$  表示  $x_i$  的真实病灶标签,  $l_{r,i}$  表示  $x_i$  经过病灶标签生成组件生成的病灶标签:

- 鉴别器损失discriminator loss

$$loss_{D,assist} = \sum_{j=0, j \neq i} \sum_{i=0} (\|d_{t,j,i}\|_2^2 + \|c_{t,j,i} - i\|_2^2)$$

- 对抗性损失和类别指导损失adversarial loss and category guidance loss

$$loss_{G,assist} = \sum_{j=0, j \neq i} \sum_{i=0} (\|d_{t,j,i} - 1\|_2^2 + \|c_{t,j,i} - i\|_2^2)$$

- MRI重建自监督损失self-supervised loss of MRI reconstruction

$$loss_{sv} = \sum_{i=0} (\|x_i - x_{r,i}\|_2^2)$$

- MRI循环重建自监督损失self-supervised loss of MRI cycle-reconstruction

$$loss_{cycle} = \sum_{j=0, j \neq i} \sum_{i=0} (\|x_i - x_{cr,j,i}\|_2^2)$$

- MRI循环重建一致性损失consistency loss of MRI cycle-reconstruction

$$loss_{cycle,consistency} = \sum_{k=0, k \neq j, k \neq i} \sum_{j=0, j \neq i} \sum_{i=0} (\|x_{cr,j,i} - x_{cr,k,i}\|_2^2)$$



- 语义一致性损失 **semantic consistency loss**

$$loss_{code,consistency} = \sum_{j=0, j \neq i} \sum_{i=0} (\|code_i - code_{t,i,j}\|_2^2)$$

- 病灶生成监督损失 **supervised loss of lesion generation**

$$loss_{sv,l} = \sum_{i=0} \|label_i - label_{r,i}\|_2^2$$

## 2.5 结构特征图与病灶分割标签的融合

结构特征图与病灶分割标签融合时，我们先生成结构特征图  $f_g$ ，再随机选择合适的病灶分割标签  $label$ ，然后再将包含  $n$  个类别的病灶分割标签转为  $n$  个通道的独热矩阵  $onehot_l$ ，每个通道对应一个分割类别，每个通道内的像素值为0或1，各个1像素区域与原始分割标签图中的各个类别的分割区域是配准的。最后，我们将  $onehot_l$  的每个通道与  $f_g$  按位求取加权求和，就得到一个新的融合了  $f_g$  和  $label$  信息的矩阵。

如果结构特征图  $f'$  是从随机MRI  $x$  中提取的，那么提取出的结构特征有可能包含肿瘤结构信息，会对随机标签  $l$  中的肿瘤信息产生干扰而影响融合后生成的MRI，所以  $f'$  需要在与随机标签  $label$  融合前消除肿瘤信息，得到无肿瘤信息的结构特征图  $f$ ，使生成图像的肿瘤信息只来源于标签  $label$ 。我们对  $x$  的分割标签  $label_x$  通过算法 2 生成无边界扩充的分割掩膜  $mask_{l,x}$ ，则  $f = mask_{l,x} \times f'$ 。

由于随机选择的病灶的位置可能出现在结构特征图的脑部轮廓之外，因此我们选取标签图时需要用算法 2 获取结构特征图的脑部区域掩膜  $mask$ ，若  $mask$  与选取的  $label$  求积后为0，则说明肿瘤标签像素在  $mask$  的脑轮廓内部，可以采用，否则需要重新选取  $label$ 。

## 2.6 多模态MRI生成训练Multimodal MRI generation training

我们通过算法 1 从真实MRI提取结构特征图  $f$ ，并随机选择真实的病灶标签  $label$  与之融合。得到的融合图包含了目标部位的基本解剖结构信息和病灶信

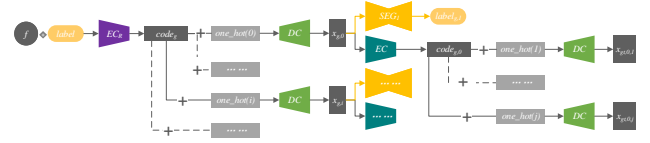


图 4: 多模态MRI生成, generation of Multimodal MRI.

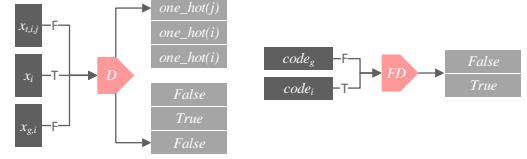


图 5: 真实MRI重建和转换训练和多模态MRI生成训练过程中的鉴别器训练. reconstruction and translation training of real MRI and discriminator training during multimodal MRI generation training.

息。多模态MRI生成过程如图 4 所示，首先，我们使用一个单独的融合图编码器对信息融合图进行编码得到语义特征图，语义特征图与不同的条件向量堆叠后通过一个MRI解码器解码，得到不同模态的合成图。这些合成图再通过MRI编码器和MRI解码器进行模态间的互相转换。我们通过一个MRI鉴别器提供的对抗性损失和类别指导损失来使得生成的各个模态的合成图逼近于真实的MRI。我们通过损失对所有语义特征图和转换图进行一致性约束，以此保证了生成的多模态MRI的互相配准。此外，我们使用一组病灶标签生成组件从各合成MRI中分割还原出肿瘤病灶分割标签，确保生成的多模态影像根据输入病灶标签生成了对应的病灶内容。

- 鉴别器真假鉴别损失 **true/false Discrimination loss of Discriminator**

$$loss_D = \sum_{i=0} (\|d_i - 1\|_2^2 + \|d_{g,i}\|_2^2)$$

- 鉴别器模态鉴别损失 **modality Discrimination loss of Discriminator**

$$loss_{D,class} = \sum_{i=0} (\|c_i - i\|_2^2 + \|c_{g,i} - i\|_2^2)$$

- 对抗性损失 **adversarial loss**

$$loss_G = \sum_{i=0} (\|d_{g,i} - 1\|_2^2)$$

- 模态类别指导损失 **Modality category guidance loss**

$$loss_{G,class} = \sum_{i=0} (\|c_{g,i} - i\|_2^2)$$

- 输入的结构特征图的重建自监督损失 **Reconstruction self-supervised loss of input structural feature map**

$$loss_{sv,f} = \sum_{i=0} (\|f - f_{g,i}\|_2^2)$$

- 病灶标签生成监督损失 **supervised loss of Lesion label generation**

$$loss_{sv,l} = \sum_{i=0} (\|label - label_{g,i}\|_2^2)$$

- MRI配准监督损失 **supervision loss of MRI registration**

$$loss_{trans} = \sum_{j=0, j \neq i} \sum_{i=0} (\|x_{g,i} - x_{gt,j,i}\|_2^2)$$

- 语义一致性损失 **semantic consistency loss**

$$loss_{trans,code} = \|code_g - code_{g,i}\|_2^2 + \sum_{j=0, j \neq i} \sum_{i=0} (\|code_{g,i} - code_{g,j}\|_2^2)$$

## 2.7 病灶标签生成指导方案 **Lesion label generation guidance method**

我们设计了如下三种病灶标签生成组件来提供多模态MRI生成训练中病灶生成的指导损失：

- 单分割器方案 **Single segmentor** 每个模态由一个共同的完整的分割器从合成的MRI还原得到各自的病灶标签。
- 单病灶编码器+多病灶解码器方案 **Single lesion encoder + multiple lesion decoders** 不同模态的分割器由一个共同的病灶编码器与不同的病灶解码器组合得到。
- 多分割器方案 **multiple segmentors** 每个模态由一个独立的完整的分割器从合成的MRI还原得到各自的病灶标签。

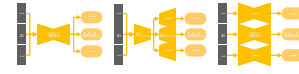


图 6: 病灶分割.lesion segmentation.



图 7: 构建合成数据集.construction of synthetic datasets.

上述三种方案的损失函数与前文所述损失函数一致，三种方案中的各个组件均只使用真实MRI重建和转换训练中的病灶生成监督损失进行训练。另外，我们还将单独的对这三种方案进行独立的训练。在不同的实验中，我们将采用不同量的合成数据或真实数据对各个方案中的分割器进行训练。训练的损失函数如下，其中 $label_i$ 是监督标签， $label_{r,i}$ 是分割结果：

$$loss_l = \sum_{i=0} \|label_i - label_{r,i}\|_2^2$$

## 2.8 构建合成数据集 **construction of synthetic datasets**

如图所示，我们通过训练好的结构特征图解码器即可从随机生成的正态分布矩阵生成任意数量的结构特征图。然后，我们对原始的标签集进行了随机的缩放、旋转、平移、翻转等改变得到随机病灶标签集。我们再将随机生成的结构特征图和从随机病灶标签集中随机选择的病灶标签融合，同样地，我们通过掩膜生成器 *MASK* 生成结构特征图的掩膜，可以筛选得到合适的随机病灶标签。最后，我们通过多模态MRI生成组件即可从融合信息图中生成配准的多模态MRI，选取的病灶标签就是生成的多模态MRI的病灶标签。由此，我们可以从随机正态分布矩阵构建带有病灶标签的多模态配准MRI数据集。

由于训练时去除肿瘤信息操作的影响，合成的结构特征图中，存在一些质量较差的脑轮廓未闭合结构特征图，我们对此设计了一个结构特征图过滤算法。首先，我们使用生成器生成一张结构特征图的掩膜，我们先对结构特征图进行高斯模糊(Wink and Roerdink 2004)，再采用OpenCV 【】提供的轮廓

查找算法和填充算法获取高斯模糊图所有的闭合轮廓并进行填充，这样我们得到一个采用传统算法的得到的掩膜，最后我们计算两张掩膜的均差(MAE)。若MAE低于我们设定的阈值则说明该结构特征图主要的脑部轮廓较为完整，该特征图可以使用；否则说明该结构特征图主要的脑部轮廓有残缺，采用传统算法得到的掩膜内部是空心的，与生成器生成的掩膜差异较大，因此，需要重新生成。算法表示如下：

---

**Algorithm 3** Structural feature map filtering

---

```

1: function GetMaskFromF(img)
2:   contours = OpenCV.findContours(img)
3:   img = OpenCV.drawContours(img, contours)
4:   return img
5: end function
6:
7: mae = 0.05
8: do
9:   f, m = Generator()
10:  m' = GetMaskFromF(f)
11: while MAE(m', m) <= mae

```

---

从筛选出来的结构特征图和匹配的病灶标签得到的多模态MRI中，同样存在病灶信息生成情况较差的样本。此时，我们通过预先训练好的病灶分割网络对我们的合成MRI数据进行分割，然后将分割结果与输入的病灶标签进行骰子评分评估，可以过滤得到评分高于设定阈值（默认0.95）的样本。

经过多重筛选，我们得到最终的由随机结构特征图、配对的掩膜、随机病灶分割标签、多模态MRI组成的合成数据集。我们要求经过分割过滤后的数据集在使用真实数据训练得到的分割器上能取得0.98以上的骰子分数，然后才能将其用于数据可用性验证实验。

## 3 实验Experiments

### 3.1 BRATS2015数据集BRATS2015 dataset

我们采用了公开的BRATS2015(Menze et al. 2014)数据集进行实验，该数据集包含已配准

的T1、T2、T1c、Flair四个模态，训练集每个模态有274张3D MRI，大小为 $155 \times 240 \times 240$ ，同时配有274张相同尺寸的肿瘤分割标签。我们将样本按9:1划分训练集和测试集，取每张3D MRI第55-105间的50个slice构建2D的数据集。在数据预处理阶段，我们将每张图进行了标准化。

### 3.2 BRATS合成数据集BRATS synthetic dataset

我们采用 2.8中的方法构建了一个配准的包含T1、T2、T1c、Flair四个模态的具有肿瘤标签的合成数据集。合成数据集样本的尺寸与BRATS2015数据集一致，但样本的多少可以根据实验需要进行任意数量的合成。

### 3.3 BRATS增强数据集BRATS Enhanced dataset

我们对原始的BRATS2015数据集进行了随机的缩放、旋转、平移、翻转等改变，得到增强数据。增强数据集样本的尺寸与BRATS2015数据集一致，但样本的多少可以根据实验需要进行任意数量的生产。

### 3.4 训练设置Training settings

每项实验的迭代次数与BRATS2015训练数据集的100个epoch相等；基础学习率分别为 $1e-4$ ，无权重衰减；采用优亚当优化器，beta1取0.5；在输入层进行0.1的Dropout；Batch size为1；在生成器中使用均值滤波器的参数进行参数初始化，具体来说，对于一个卷积核尺寸为 $[k, k, f]$ 的卷积层，我们初始化该层的 $k \times k \times f$ 个卷积核参数均为 $1/(k \times k \times f)$ ，偏置量为0。我们采用骰子分数(Dice 1945)和均方差(MSE) (Prasad and Rao 1990)进行分割结果的评估，评估结果为2D图像的评估结果的均值，每项实验训练四次保留最佳结果。

### 3.5 病灶生成组件各方案对比实验Contrast experiment of each method of lesion generation component

我们在BRATS2015训练集上采用4SEG方案训练



表 1: 病灶检测实验结果.Lesion segmentation result

| synthetic method | test data type | MSE   | Dice Score |
|------------------|----------------|-------|------------|
| 4SEG             | real           | 0.026 | 0.915      |
| 1SEG             | synthetic      | 0.053 | 0.741      |
| 1ECL+4DCL        | synthetic      | 0.055 | 0.808      |
| 4SEG             | synthetic      | 0.043 | 0.838      |

了肿瘤病灶的分割器，然后我们在BRATS2015测试数据集上使用该分割器对通过不同病灶生成指导方案得到的未经分割器筛选的合成数据进行分割测试。除了测试数据来源的不同外，测试数据的样本量等其他条件完全相同。其中，我们在真实数据上采用的分割方案为章节 2.7中的多分割器方案，即每个模态各训练一个独立的分割器。

### 3.6 合成数据可用性验证实验 verification experiment of Synthetic data availability

如表所示，我们将真实的BRATS2015训练数据与BRATS合成数据进行了不同数量的混合，再用构建的混合数据集进行分割训练，最后再在真实的BRATS2015测试数据上进行模型的分割能力评估，所有实验都进行相同迭代步数的充分训练，除了训练数据源的不同外其他条件完全相同。同时，我们还进行了单独的合成数据的训练、真实数据与通常的数据增强数据的混合训练作为对比。我们设定了随机混合、先真后假、先假后真三种数据混合方式。我们选用实验 3.5中，在合成数据集上表现最好的分割方案来作为验证实验的分割方案。

## 4 结果Results

### 4.1 病灶检测实验结果Lesion segmentation result

如表 1所示，我们在BRATS2015训练数据经过相同的迭代步数的充分训练后，在真实的测试数据集上，分割测试结果达到了0.026的MSE和0.915的Dice Score。之后我们使用这个表现优秀的分割网络对我们未过滤的合成数据进行分割测试，在与真实测试数据集相同数据量的合成数据集上，不同病灶标

表 2: 合成数据可用性验证实验结果.Verification experiment results for synthetic data availability

| Num   | real data  | synthetic data | Enhanced data | mixing modes    | MSE   | Dice Score |
|-------|------------|----------------|---------------|-----------------|-------|------------|
| 1-1   | 15070      | 0              | 0             | -               | 0.026 | 0.915      |
| 2-31  | 15070× 0.5 | 0              | 0             | -               | 0.032 | 0.902      |
| 3-2   | 0          | 15070          | 0             | -               | 0.205 | 0.708      |
| 4-3   | 0          | 15070× 2       | 0             | random mixing   | 0.206 | 0.736      |
| 5-4   | 0          | 15070× 3       | 0             | random mixing   | 0.205 | 0.754      |
| 6-7   | 15070× 0.1 | 15070          | 0             | synthetic first | 0.031 | 0.908      |
| 7-8   | 15070× 0.1 | 15070× 2       | 0             | synthetic first | 0.028 | 0.907      |
| 8-9   | 15070× 0.1 | 15070× 3       | 0             | synthetic first | 0.030 | 0.907      |
| 9-13  | 15070× 0.2 | 15070× 0.8     | 0             | random mixing   | 0.041 | 0.850      |
| 10-12 | 15070× 0.5 | 15070× 0.5     | 0             | random mixing   | 0.031 | 0.904      |
| 11-14 | 15070× 0.8 | 15070× 0.2     | 0             | random mixing   | 0.024 | 0.935      |
| 12-32 | 15070      | 15070× 0.2     | 0             | random mixing   | 0.025 | 0.921      |
| 13-33 | 15070      | 15070× 0.5     | 0             | random mixing   | 0.023 | 0.939      |
| 14-34 | 15070      | 15070× 0.8     | 0             | random mixing   | 0.026 | 0.916      |
| 15-15 | 15070      | 15070          | 0             | random mixing   | 0.027 | 0.913      |
| 16-18 | 15070      | 15070× 2       | 0             | random mixing   | 0.033 | 0.901      |
| 17-19 | 15070      | 15070× 3       | 0             | random mixing   | 0.034 | 0.897      |
| 18-35 | 15070      | 0              | 15070× 0.2    | random mixing   | 0.027 | 0.911      |
| 19-36 | 15070      | 0              | 15070× 0.5    | random mixing   | 0.025 | 0.927      |
| 20-37 | 15070      | 0              | 15070× 0.8    | random mixing   | 0.026 | 0.920      |
| 21-22 | 15070      | 0              | 15070         | random mixing   | 0.026 | 0.915      |
| 22-23 | 15070      | 0              | 15070× 2      | random mixing   | 0.032 | 0.898      |
| 23-24 | 15070      | 0              | 15070× 3      | random mixing   | 0.036 | 0.885      |
| 24-16 | 15070      | 15070          | 0             | real first      | 0.195 | 0.795      |
| 25-17 | 15070      | 15070          | 0             | synthetic first | 0.021 | 0.940      |

签生成组件设计方案都取得了较好的分割结果，其中每个模态训练一个独立的分割器的方案取得了最好的结果，Dice Score也达到了0.838。

### 4.2 合成数据可用性验证实验结果Verification experiment results for synthetic data availability

如表 2所示，实验NO.3-NO.5的结果表明，在训练中合成数据还不能完全替代真实数据。实验NO.6-NO.8结果表明使用大量合成数据进行预训练再在少量真实数据上微调能取得和完全使用真实数据训练十分接近的结果。实验NO.9-NO.11中，不同的混合比例取得的结果差异也很大。当两者比例相近时，分割结果与实验NO.1相差不大。合成数据比例偏高时，合成数据会干扰模型对真实数据的学习，结果比实验NO.1低。合成数据比例偏低时，合成数据能提高模型的泛化能力，取得高于实验NO.1的结果。实验NO.12-NO.17中我们进一步尝试在全量的真实数据中添加不同数据量的合成数据，我们发现添加少量的合成数据能起到数据增强的作用，并且合成数据越多增强效果越好，但当合成数据到达

一定比例后再继续增加将取得反效果。实验NO.18-NO.23中，我们使用通常的数据增强方法产生的增强数据来与前述合成数据的增强效果进行对比，我们发现在增强数据量与增强效果的变化趋势上，两者是一致的，但两者在具体的增强效果随着增强数据量变化的变化曲线上是有一些差异的。整体来说，在模型对增强数据量的敏感程度上，增强数据更加鲁棒，但合成数据能取得的增强效果的上限比增强数据要高很多。在实验NO.24-NO.25中，我们对比实验NO.15发现，合成数据作为预训练数据先于真实数据训练时表现最好，作为补充训练数据集在真实数据集之后训练则表现很差，作为增强数据与真实数据混合训练时可以取得一定的增强效果。

总的来说，当真实数据量较多时，可以使用少量的合成数据作为增强数据混合使用，也可以使用大量的合成数据先进行预训练，再在真实数据集上训练；当真实数据较少时，可以使用大量的合成数据进行预训练，再在少量真实数据上微调，最后能得到和完全真实数据训练相竞争的结果，这点我们与(Shin et al. 2018)的结论一致；我们不建议完全使用合成数据进行训练，与(Shin et al. 2018)的结论不同的是我们也不建议使用合成数据进行补充训练。

### 4.3 合成图像展示

图8展示了我们从随即正态分布矩阵生成的结构特征图示例。图9中展示了我们从随即正态分布矩阵生成的结构特征图、从结构特征图生成的对应的掩膜、随机选择的病灶分割标签、从结构特征图生成和的病灶分割标签生成的多模态MRI的几个示例。

## 5 结论与未来工作

我们基于条件生成对抗网络，通过无监督的训练，实现了从随机正态分布矩阵生成配准的多模态MRI并能自由添加病灶信息。我们通过病灶分割实验验证了合成的MRI可以作为智能医学影像处理任务的预训练数据或增强数据来使用并能显著提高模型的泛化能力。具体来说，我们的贡献包括以下几点：

- 我们提出了一种从医学影像直接提取解剖结构信

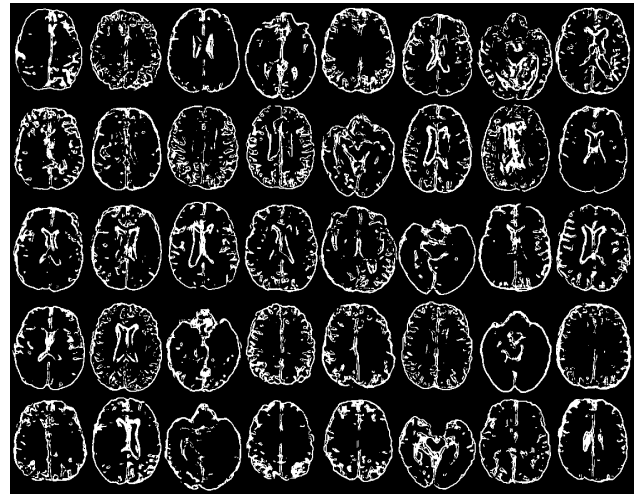


图 8: 合成的结构特征图.

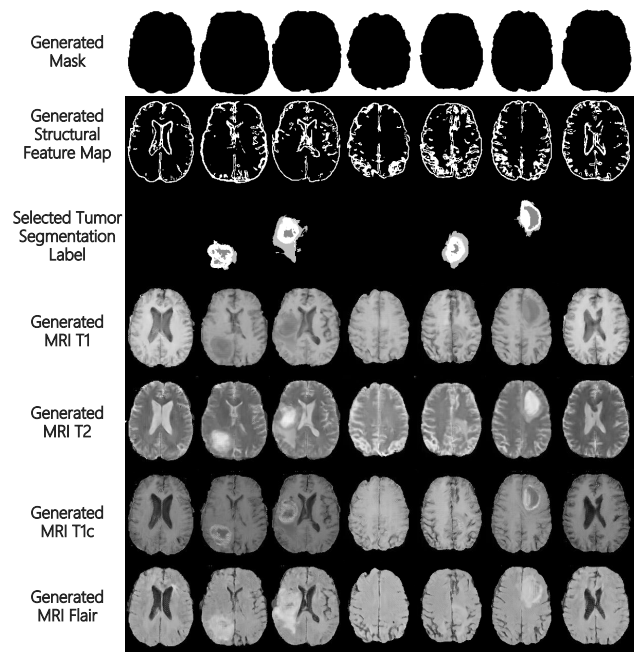


图 9: 从随机结构特征图和病灶标签生成多模态MRI.

息的结构特征图提取方法，无须训练，无需额外数据；

- 我们提出了一种从多维正态分布采样生成结构特征图的随机结构特征图生成方法，使结构特征图能便捷地大量生成并具有良好的多样性；
- 我们实现了从结构特征图及随机选择的病灶标签生成具有对应病灶信息的配准的多模态MRI；
- 我们通过多项在合成数据上进行的数据可用性测试，验证了合成数据可以作为智能医学影像处理任务的预训练数据或增强数据来使用。

未来，我们将进一步在CT、PET等不同模态，心脏、肺等其他部位，检测、分类等其他病灶处理任务中对我们的方法进行改进。我们将致力于进一步简化训练过程和合成更高质量的医学影像。

## 6 Acknowledgments

感谢NSCCGZ提供的计算支持。

## References

- Burgos, N.; Cardoso, M. J.; Guerreiro, F.; Veiga, C.; Modat, M.; McClelland, J.; Knopf, A.; Punwani, S.; Atkinson, D.; Arridge, S. R.; et al. 2015. Robust ct synthesis for radiotherapy planning: Application to the head and neck region. *medical image computing and computer assisted intervention* 476–484.
- Chartsias, A.; Joyce, T.; Giuffrida, M. V.; and Tsaftaris, S. A. 2018. Multimodal mr synthesis via modality-invariant latent representation. *IEEE Transactions on Medical Imaging* 37(3):803–814.
- Choi, Y.; Choi, M.; Kim, M.; Ha, J.; Kim, S.; and Choo, J. 2018. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. *computer vision and pattern recognition* 8789–8797.
- Costa, P.; Galdran, A.; Meyer, M. I.; Abramoff, M. D.; Niemeijer, M.; Mendonca, A. M.; and Campilho, A. 2017. Towards adversarial retinal image synthesis. *arXiv: Computer Vision and Pattern Recognition*.
- Dice, L. R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26(3):297–302.
- G, R. L. 1965. Machine perception of three-dimensional solids. *Optical and Electro-Optical Information Processing*.
- Goodfellow, I. J.; Pougetabadie, J.; Mirza, M.; Xu, B.; Wardefarley, D.; Ozair, S.; Courville, A. C.; and Bengio, Y. 2014. Generative adversarial nets. *neural information processing systems* 2672–2680.
- Huang, Y.; Shao, L.; and Frangi, A. F. 2017. Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding. *computer vision and pattern recognition* 5787–5796.
- Isola, P.; Zhu, J.; Zhou, T.; and Efros, A. A. 2017. Image-to-image translation with conditional adversarial networks. *computer vision and pattern recognition* 5967–5976.
- Joyce, T.; Chartsias, A.; and Tsaftaris, S. A. 2017. Robust multi-modal mr image synthesis. *medical image computing and computer assisted intervention* 347–355.
- Kamnitsas, K.; Baumgartner, C. F.; Ledig, C.; Newcombe, V.; Simpson, J. P.; Kane, A. D.; Menon, D. K.; Nori, A. V.; Criminisi, A.; Rueckert, D.; et al. 2017. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. *information processing in medical imaging* 597–609.
- Kingma, D. P., and Welling, M. 2014. Auto-encoding variational bayes. *international conference on learning representations*.
- Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *neural information processing systems* 141(5):1097–1105.
- Liang, X.; Zhang, H.; Lin, L.; and Xing, E. P. 2018. Generative semantic manipulation with mask-contrasting gan. *european conference on computer vision* 574–590.

- Menze, B.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; Lanczi, L.; Gerstner, E.; Weber, M.-A.; Arbel, T.; Avants, B.; Ayache, N.; Buendia, P.; Collins, L.; Cordier, N.; Corso, J.; Criminisi, A.; Das, T.; Delingette, H.; Demiralp, C.; Durst, C.; Dojat, M.; Doyle, S.; Festa, J.; Forbes, F.; Geremia, E.; Glocker, B.; Golland, P.; Guo, X.; Hamamci, A.; Iftekharuddin, K.; Jena, R.; John, N.; Konukoglu, E.; Lashkari, D.; Antonio Mariz, J.; Meier, R.; Pereira, S.; Precup, D.; Price, S. J.; Riklin-Raviv, T.; Reza, S.; Ryan, M.; Schwartz, L.; Shin, H.-C.; Shotton, J.; Silva, C.; Sousa, N.; Subbanna, N.; Szekely, G.; Taylor, T.; Thomas, O.; Tustison, N.; Unal, G.; Vasseur, F.; Wintermark, M.; Hye Ye, D.; Zhao, L.; Zhao, B.; Zikic, D.; Prastawa, M.; Reyes, M.; and Van Leemput, K. 2014. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging* 33.
- Miao, S.; Piat, S.; Fischer, P. W.; Tuysuzoglu, A.; Mewes, P.; Mansi, T.; and Liao, R. 2018. Dilated fcn for multi-agent 2d/3d medical image registration. *national conference on artificial intelligence* 4694–4701.
- Mirza, M., and Osindero, S. 2014. Conditional generative adversarial nets. *Computer Science* 2672–2680.
- Nie, D.; Trullo, R.; Lian, J.; Petitjean, C.; Ruan, S.; Wang, Q.; and Shen, D. 2017. Medical image synthesis with context-aware generative adversarial networks. *medical image computing and computer assisted intervention* 417–425.
- Osokin, A.; Chessel, A.; Salas, R. E. C.; and Vaggi, F. 2017. Gans for biological image synthesis. *international conference on computer vision* 2252–2261.
- Prasad, N. G. N., and Rao, J. N. K. 1990. The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association* 85(409):163–171.
- Prewitt, J. 1970. Object enhancement and extraction. *Picture processing and Psychopictorics*.
- Rezende, D. J.; Mohamed, S.; and Wierstra, D. 2014. Stochastic backpropagation and approximate inference in deep generative models. *international conference on learning representations* 1278–1286.
- Shin, H.; Tenenholz, N. A.; Rogers, J. K.; Schwarz, C. G.; Senjem, M. L.; Gunter, J. L.; Andriole, K. P.; and Michalski, M. 2018. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. *arXiv: Computer Vision and Pattern Recognition* 1–11.
- Sobel, I. 1970. Camera models and machine perception. *Ph.D.dissertation*.
- Van Nguyen, H.; Zhou, K. S.; and Vemulapalli, R. 2015. Cross-domain synthesis of medical images using efficient location-sensitive deep network. *medical image computing and computer assisted intervention* 677–684.
- Vemulapalli, R.; Van Nguyen, H.; and Zhou, S. K. 2015. Unsupervised cross-modal synthesis of subject-specific scans. *international conference on computer vision* 630–638.
- Wink, A. M., and Roerdink, J. B. T. M. 2004. Denoising functional mr images: a comparison of wavelet denoising and gaussian smoothing. *IEEE Transactions on Medical Imaging* 23(3):374–387.
- Zhang, Z.; Yang, L.; and Zheng, Y. 2018. Translating and segmenting multimodal medical volumes with cycle- and shape-consistency generative adversarial network. *computer vision and pattern recognition* 9242–9251.
- Zhao, B.; Chang, B.; Jie, Z.; and Sigal, L. 2018. Modular generative adversarial networks. *european conference on computer vision* 157–173.
- Zhu, J.; Park, T.; Isola, P.; and Efros, A. A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *international conference on computer vision* 2242–2251.