**DiffuseMorph: Unsupervised Deformable Image Registration Along Continuous Trajectory Using Diffusion Models**

一、摘要：

可变形图像配准是医学成像和计算机视觉的基本任务之一。经典的配准算法通常依赖于迭代优化方法来提供精确的变形，这需要很高的计算代价。虽然已经开发了许多基于深度学习的方法来实现快速图像配准，但在估计形变场的同时减少拓扑折叠问题仍然是具有挑战性的。此外，这些方法仅能够配准到单个固定图像，并且不可能在运动图像和固定图像之间获得连续变化的配准结果。针对这一问题，本文提出了一种新的基于扩散模型的概率图像配准方法，称为DiffuseMorph。具体地说，我们的模型学习了运动图像和固定图像之间变形的得分函数。与现有的扩散模型类似，DiffuseMorph不仅可以通过反向扩散过程提供合成变形图像，还可以使运动图像与潜在空间一起进行不同程度的变形。在2D人脸表情图像和3D脑图像配准任务上的实验结果表明，该方法能够提供灵活、准确的变形，并具有拓扑保持能力。

可变形图像配准是估计运动图像和固定图像对之间的体素对应关系。这对于诸如疾病诊断和治疗监测的医学图像分析尤其重要，因为医学图像的解剖结构或形状根据对象、扫描时间、成像方式等而不同。因此，在过去的几十年中已经研究了各种图像配准方法，并且这些经典方法通常试图通过求解计算代价高昂的优化问题来对齐图像。为了解决这一计算问题，基于深度学习的图像配准方法最近被广泛研究，训练神经网络将运动图像和固定图像作为网络输入来估计变形场。这些方法在保持配准精度的同时提供快速变形。然而，它们通常需要用于监督训练的地面真实变形场[31，33]，复杂的附加微分同形积分层[10，25]，或者用于拓扑保持的循环一致性[22]。最近，去噪扩散概率模型(DDPM)[18，36]在计算机视觉的许多领域都显示出令人印象深刻的性能[8，15，21，35，37]。DDPM学习从高斯噪声分布到数据分布的马尔可夫变换，并通过随机扩散过程提供不同的样本。为了生成具有所需语义的图像，还提出了条件去噪扩散模型[9，34]。然而，将DDPM应用到图像配准任务中是具有挑战性的，因为现有的方法生成的样本可能会丢失图像身份，并且正确的配准应该由运动图像的变形场来执行，而不是图像生成。

本文提出了一种新的无监督图像配准方法DiffuseMorph，该方法采用DDPM来生成变形图像，并允许沿着运动图像和固定图像之间的连续轨迹进行基于变形场的配准。该模型由扩散网络和变形网络组成：扩散网络学习运动图像和固定图像之间变形的得分函数，变形网络利用该得分估计形变场，并提供变形图像。这两个模型以端到端的学习方式进行联合训练，使得DiffuseMorph不仅可以测量运动图像变形方向的马尔可夫变换，还可以估计运动图像扭曲成固定图像的配准区域。由于这种端到端的学习，所提出的方法在推理阶段执行有效的图像配准。具体地说，如图1所示，一旦网络被训练，我们的模型可以通过简单地内插用作变形网络的输入的分值来提供从运动图像到固定图像的中间变形。此外，提出的DiffuseMorph算法可以快速生成类似于固定图像的合成变形运动图像。具体地说，我们不是从随机的高斯噪声开始，而是从运动图像执行向前一步的扩散，然后使用DDPM在扩散网络中的反向扩散进行迭代细化。这允许样本保留原始运动图像的身份，并显著减少扩散步骤的数量。

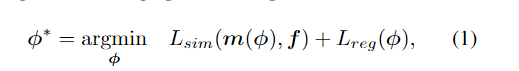
为了验证该方法的性能，我们将该模型应用于2D人脸图像配准和3D脑MR图像配准任务。实验结果表明，该方法具有较高的精度。证明了我们的模型在变形精度上达到了很高的性能。此外，即使我们的模型没有对微分同胚配准的约束，结果也显示出保持拓扑结构的光滑形变场。此外，我们的方法实现了运动图像和固定图像之间沿连续轨迹的可变形配准，这是任何现有的图像配准方法都不可能实现的。我们的贡献概括为：

* 我们提出了DiffuseMorph方法，这是第一种基于运动图像和固定图像的去噪扩散模型的图像配准方法。
* 由于扩散网络和变形网络的端到端学习，我们的模型不仅通过快速反向扩散生成变形图像，而且能够沿着运动图像和固定图像之间的连续轨迹进行图像配准。
* 验证了该方法可以应用于2D和3D图像配准任务，并提供了精确的变形和拓扑保持。

二、背景与相关工作：

**2.1变形图像配准**

给定运动图像m和固定图像f，经典的可变形图像配准方法通过求解以下优化问题来执行：



(1)其中φ\*是将运动图像变形为固定图像的最优配准域。LSIM是计算变形图像和固定图像之间相似度的相异函数，LREG是形变场的正则化惩罚。通过最小化能量函数，通过对运动图像进行扭曲来估计变形图像m(φ)。特别地，当对场φ施加附加约束时，可以实现微分同胚配准，使得变形映射是可微的和可逆的，并且提供拓扑保持[2，5，41]。

基于学习的方法由于传统的配准方法计算量大、运行时间长，近年来深度学习方法得到了广泛的研究，一旦训练好神经网络，深度学习方法就可以实时估计变形。例如，监督学习方法用地面真实注册场[6，7，33，43]训练网络，但它们并不实用，因为训练需要高质量的变形场标签。相比之下，非监督学习方法通过计算变形图像和固定参考图像[4，12，22，28，29] 之间的相似度来训练网络。为了保证拓扑保持，提出了具有变形概率参数的微分同胚配准方法[10，11，25]。然而，微分同胚约束也是推理所必需的，这增加了额外的复杂性。

**2.2扩散概率模型**

最近，去噪扩散概率模型(DDPM)[18，36]被提出作为产生式模型，它学习了一个马尔可夫链过程，将简单的高斯分布转换为数据分布。在前向扩散过程中，使用马尔可夫链将噪声逐渐添加到数据0，其中采样潜在变量xt，t∈[0，T]的每一步都表示为高斯变换：

****

其中0<βt<1是噪声的方差。然后，在给定X0的情况下，XT的结果分布表示为：

****

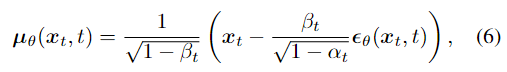
其中。因此，给定£∼N(0，i)，XT可以被采样为：

****

对于执行反向扩散的生成过程，DDPM学习参数化高斯过程p(xt−1|xt)，其表示为：θ：

****

其中，σt是固定方差，而μθ(xt，t)是定义为：

****

其中在训练阶段期间估计参数化模型θ。事实上，模型θ(xt，t)只是得分函数的θ(xt，t)[38]的缩放版本，它是logpθ(Xt)的梯度。一旦训练了模型，通过以下随机生成步骤对数据进行采样：xt−1=√(xt，t)+∼θμθ，其中zβN(0，i)。

**条件生成方法：**

由于我们对给定运动图像和参考图像执行的图像配准感兴趣，这里我们回顾了现有的基于条件扩散的图像生成模型[9，19，30，34，37，39]。为了通过反向过程控制图像的生成，ddim[37]提出了一个从初始条件开始的确定性的非马尔可夫生成过程。SR3[34]是通过用条件图像训练DDPM来实现超分辨率的。ILVR[9]提出了使用无条件模型的条件性迭代生成过程。然而，现有的方法不会使条件图像的形状变形，或者可能会改变图像的身份，这在图像配准任务中是不应该发生的。相反，在一对运动图像和固定图像的情况下，我们的扩散模型会产生运动图像与固定图像对齐的变形，而不会丢失其一致性。

三、提出方法

3.1DiffuseMorph

通过利用DDPM的能力，我们旨在开发一种新的基于扩散模型的无监督变形方法。由于图像配准的任务是利用变形场对图像进行变形，因此我们设计了两个网络，如图2所示：一个是扩散模型Gθ，用于估计运动图像和固定图像条件下变形的得分函数；另一个是变形模型Mψ，它实际使用得分函数输出配准场。具体地，对于运动源图像m和固定参考图像f，训练扩散网络Gθ以学习在条件c=(m，f)下运动图像和固定图像之间的变形的得分函数。为此，我们通过Eq对目标的潜变量x进行采样。(4)以固定图像为目标，即x0=f。此外，为了使网络Gθ知道噪声的级别，我们直接给出网络噪声α的方差的适当统计，类似于[8，34]。另一方面，变形网络Mψ取扩散网络ˆ以及运动源图像m的分数输出，并且估计变形场φ。然后，通过使用空间变换层[20]对运动图像m进行扭曲来生成变形图像m(φ)。在实验中，为了对2D/3D图像进行变形，我们采用了双线性/三线性插值法的变换函数。

损失函数回顾扩散网络Gθ和变形网络Mψ是通过端到端学习联合训练的。因此，对于我们的模型的训练，我们设计了如下目标函数：



其中Ldif和Lregist分别是扩散损失和配准损失，λ是一个超参数。具体地说，扩散损失由下式给出：



其中∼N(0，i)和训练∼U(αt∈1，αt)，其中tα[0，α]是均匀采样的时间步长。此外，配准损失是为了估计变形场，使得变形后的图像具有与固定图像相似的形状。基于方程中传统的能量函数，提出了一种新的能量函数--能量函数。(1)、登记遗失设计如下：



其中φ=Mψ(m，£)，其中ˆ指的是扩散网络输出，并且λR是超参数。第一个(9)是相异度量的局部归一化互相关[4]，第二项是变形场的光滑性惩罚。我们设置λR=1。值得注意的是，两个损失函数的净效果是训练Gθ学习变形的得分函数。因此，当与反向扩散相结合时，分数引导反向扩散以从运动图像初始化生成固定图像。此外，该分数有助于变形网络向固定图像生成变形场

3.2使用DiffuseMorph的图像配准

一旦训练了模型的网络，在推理阶段，它们通过估计要扭曲成固定图像的运动图像的变形场来提供图像配准。由于我们的模型进行了端到端的训练，扩散模型引导变形网络生成规则的变形场。具体地说，使用学习到的G∗∗和Mθψ的参数，通过以下方式估计变形场φ：



其中，x0被设置为固定目标图像f。变形图像m(φ)通过空间变换使用估计的φ来计算。因此，我们的模型一步完成图像配准，配准区域平滑，折叠问题较少。

**沿连续轨迹的图像配准**，

因为扩散网络估计变形场的分数函数，我们的模型还提供了运动图像沿轨迹朝向固定图像的连续变形，这在现有的配准方法中是不可能的。具体地，如果输入到变形网络的分数被设置为零，则网络产生几乎不会使运动图像变形的变形场。因此，对于给定的网络输出∗(c，f，ˆ0)f=Gθα(c，f，ˆ0)，可以通过简单地内插得分值来生成使运动图像沿连续轨迹朝向固定图像变形的配准区域φη：



For 0≤≤η1。这一有趣的现象是由于学习了变形的分数函数而发生的，这将在后面的实验中得到证实。

**基于生成过程的图像变形**

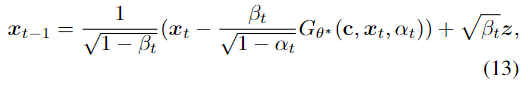
除了连续的图像配准外，学习的变形分数函数可以帮助通过反向扩散过程生成合成变形图像。这里，我们描述了一种仅使用扩散网络Gθ来生成变形图像的方法。

具体地说，当得分函数学习运动图像向固定图像的变形时，我们的图像生成从运动图像开始，而不是现有的DDPM[9，34]的条件生成过程，该条件生成过程从纯高斯噪声XT∼N(0，i)开始。当我们用原始运动图像m设置初始状态，通过以下方式执行一步前向扩散：



其中，∼N(0，I)和训练T是在时间步长T≤T处的噪声电平。这里，时间步长T被设置为不丢失图像的身份的值。这种前向采样允许运动图像分布接近固定图像分布，如图3所示，这减少了反向扩散步骤的数量和生成时间。

然后，通过从Xt开始，通过从t=T到t=1的以下反向扩散处理来执行适合于固定图像f的变形图像X0的生成：



其中z∼N(0，i)。这里，在选择总采样步长T时，我们采用了比DDPM更有效的推断方法[8]。由于我们的去噪模型直接以噪声水平α为条件，可以灵活地设置采样步数。在我们的实验中，我们设置了最大200个采样步长。算法1中描述了DiffuseMorph的这个生成过程的伪代码。

四、实验

为了证明DiffuseMorph使用高质量的配准区域从运动图像到固定图像生成变形图像，我们将我们的方法应用于两个应用程序。本文首先对二维人脸表情图像进行了变形实验。此外，我们将所提出的方法应用于3D脑MR配准，在该方法中，单个脑图像被变形到共同的图谱。数据集和培训详细信息如下，更多详细信息见附录

数据集

**面部表情图像** 用于2D面部表情图像配准，Radboud Faces数据库(Rafd)[27]。它包含了来自67个受试者的8种不同的面部表情：中性、愤怒、轻蔑、厌恶、恐惧、高兴、悲伤和惊讶。对于每个面部表情，提供了三个不同的凝视方向，因此总共有1608张图像。

**在脑MRI中**，我们使用OASIS-3数据集[26]进行了3D脑MR配准实验，该数据集提供了来自Freesurfer的脑MR图像和相应的体积分割图[16]。我们使用了1156次T1加权扫描，通过图像重采样到256×256×256网格和1mm3各向同性体素，仿射空间归一化和脑提取进行了预处理。图像被裁剪了160×192×224。我们分别使用了1027次和129次扫描进行训练和测试。

实验细节

该模型是利用DDPM[18]和VoxelMorph[4]的网络结构实现的。我们将DDPM中设计的网络用于扩散网络G训练，通过线性调度将噪声水平从10−6设置到10−2，其中Tθ=2 0 0 0。对于形变网络Mψ，我们使用了体素变形网络。这里，我们根据图像的维度配置网络的层，例如用于2D图像配准的2D卷积层。使用单个NVIDIA Quadro RTX 6000图形处理器，我们使用ADAM优化算法[23]训练我们的模型。具体地说，对于人脸图像配准，我们将超参数设置为−=2，并以5×10λ6的学习率训练模型40个历元。另一方面，对于脑MR图像配准，我们将超参数设置为−=10，并以1×10λ4的学习率训练网络60个时代。所提出的方法是使用PyTorch库在Python语言中实现的。

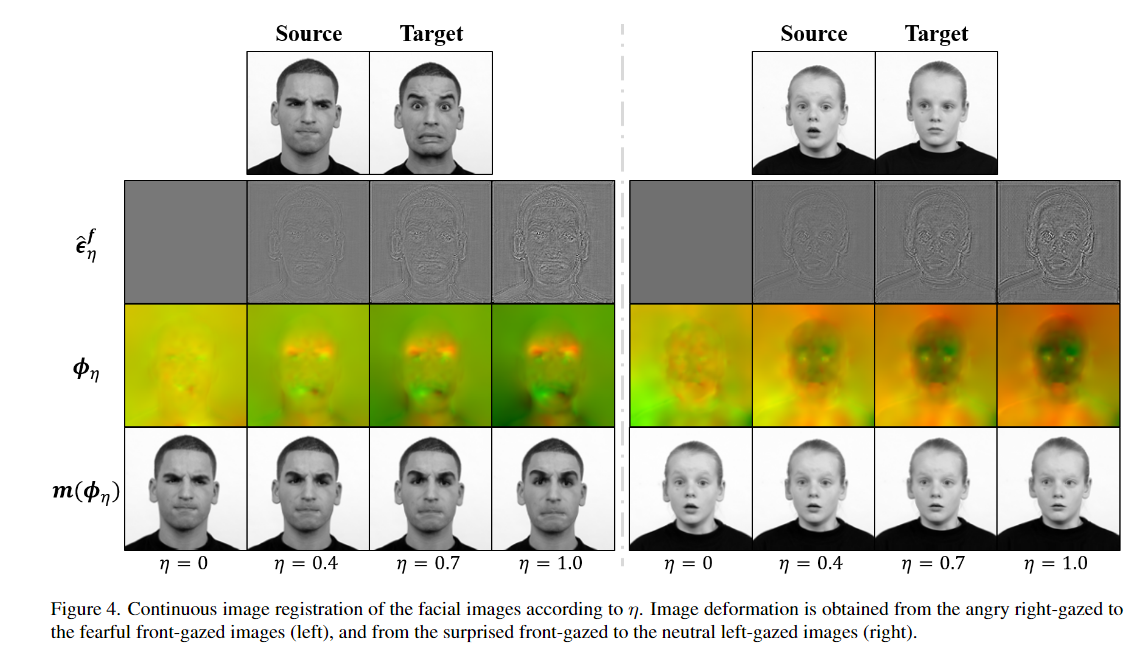
评估

为了评估配准性能，我们计算了形变场(|J≤|φ0)上雅可比行列式的非正值的百分比，这表明配准的一对一映射已经丢失。此外，对于所有变形图像都有标签的人脸图像，我们测量了变形图像和固定目标图像之间的NMSE和SSIM。对于脑MR图像，我们将形变场应用到原始运动图像的分割图中，并计算了30个解剖结构的Dice分数。当与现有的基于学习的模型进行比较时，我们使用相同的变形网络体系结构进行公平的比较。

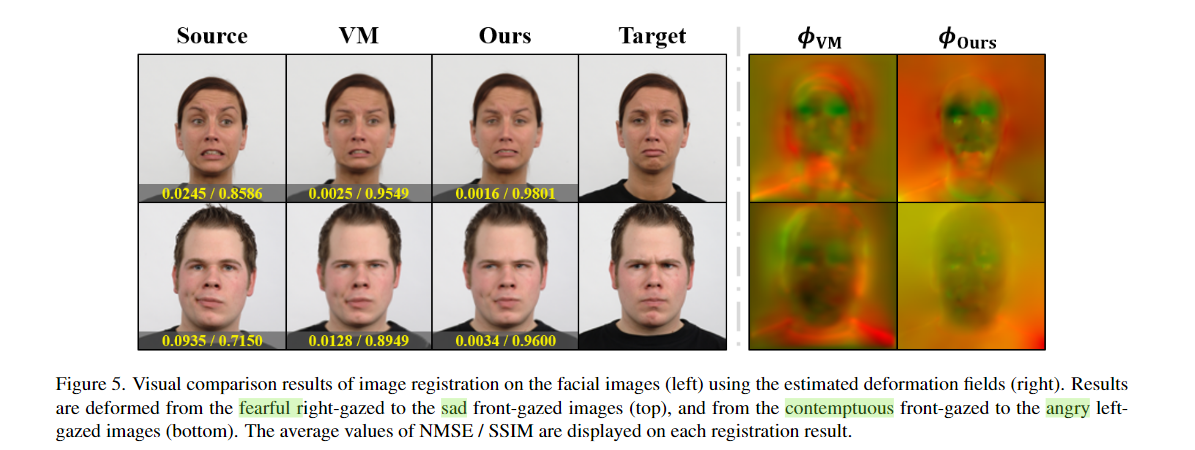
五、结果

5.1 2D脸部表情图像配准

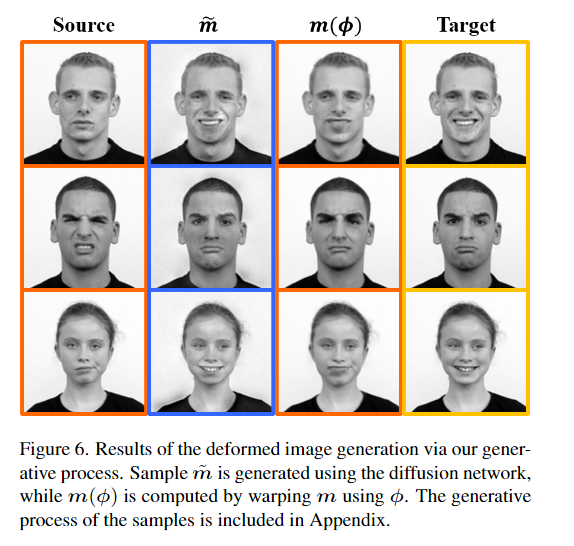
图4示出了根据变量η对面部表情图像配准的DiffuseMorph的结果。我们可以看到，图像正在慢慢地随着η的增加向固定图像移动。此外，生成的变形场不是最终变形场的简单缩放版本。相反，注册场的变化不是均匀的，这取决于中间形变水平的变化的重要性。这验证了我们的模型提供了运动源图像和固定目标图像之间的连续变形。此外，这也证实了分数作为变形的潜在变量的重要性。



此外，我们还比较了我们的模型与VoxelMorph(Vm)[4]和Vm-Diff[10]对固定图像的变形性能。在这里，我们通过将相同的灰度图像的变形场应用于每个RGB通道来获得具有RGB通道的变形图像。图5示出了配准结果的可视比较。与VM相比，我们的模型对源图像进行变形，使其与目标图像更精确地对齐。还有，Tab。1给出了量化评价结果。与初始源图像和其他比较方法相比，我们的模型获得了更低的NMSE和更高的SSIM。此外，该方法还减少了形变场的雅可比行列式的度量。这些结果表明，所提出的DiffuseMorph算法能够在较少的折叠问题下提供高质量的图像配准。更多的面部表情图像结果显示在附录中



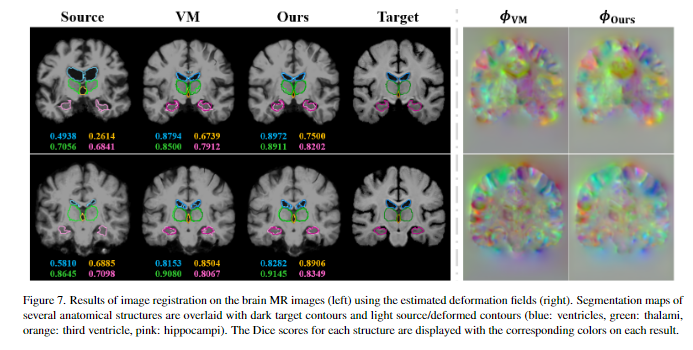
合成变形图像的生成为了验证DiffuseMorph生成图像的能力，我们利用人脸图像进行了生成过程的研究。因为面部数据被分成了训练集与测试集，使用看不见的图像进行图像生成。图6示出了给定动动源和固定目标图像时生成的样本̃m。样本是从噪声水平为α200的噪声运动图像中获得的，用于前向扩散。我们将反向扩散步数设置为80。结果表明，对于不同的人脸表情对，我们的模型提供了与目标图像相似的合成变形图像。此外，当与使用变形场的扭曲图像m(φ)相比较时，我们可以观察到，如果运动图像没有显示在固定图像中的牙齿，则所提出的生成过程能够有效地提供图像变形。

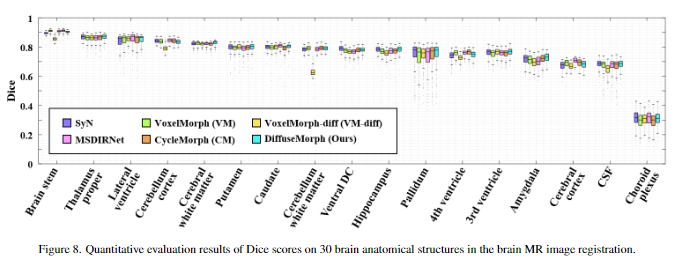


5.2 3D脑部MR图像配准

我们使用比较方法：SYN[2]、VM[4]、VM-DIFF[10]、MSDIRNet[28]和CM[22]来评估脑MRI配准结果。图7示出了视觉配准结果。我们在附录中显示了所有方法的结果。结果表明，我们的模型估计了平滑的形变场，并提供了与固定图像对齐的更准确的变形运动图像，这也可以通过几个大脑解剖结构的分割图的轮廓观察到。图8和表2.2还报道了与现有的基于学习的方法相比，该方法获得了更高的Dice分数和更少的非正雅可比行列式。

配准损失的影响研究:





为了分析模型训练中的注册丢失对模型性能的影响，我们在EQ中改变λ的值进行了对比研究。(7)使用脑部磁共振数据。如选项卡中所述。3、当λ较低时，Dice评分降低，但变形场的规律性较好。这表明配准损失迫使模型输出更平滑的变形场，我们需要在权衡两者之间进行权衡。

六、限制

在本文中，我们只验证了DiffuseMorph仅在二维图像上生成图像的能力，因为现有的扩散模型通常是在二维图像生成上进行测试的。未来的工作应该为3D图像生成建模适当的架构。然而，从图像配准的角度来看，我们的模型通过联合训练扩散网络和变形网络来学习变形的分数函数，为可变形图像的配准和生成提供了非常灵活的工具。

**潜在负面影响：**

虽然配准任务在医学图像领域有重要的应用，但我们的方法的应用可以用来转换图像的身份。它可以扩展为恶意使用，如deepfake。滥用图像配准方法会引起诸如侵犯版权和隐私问题等社会问题。为了减轻此类风险，必须执行社会和隐私法规。

七、结论

采用与变形网络联合训练的扩散概率模型，提出了一种新的无监督图像配准的DiffuseMorph结构。由于具有学习形变分数函数的能力，该方法不仅可以生成合成形变图像，而且可以通过估计形变场沿固定图像的连续轨迹提供高质量的配准。我们相信DiffuseMorph是一个很有前途的算法。

附录：

1. 数据集细节

对于本文中图像配准任务的实验，我们使用了以下两个公共数据集。

A.1 数据集

Radboud面部数据库[27]包含67个受试者的照片，包括白人成人和儿童。在受控环境下，每个受试者有8种不同的情绪表情和3种凝视方向。所有的情绪表达都是基于面部动作编码系统[13]的原型。根据其官方网站的许可协议，该数据集可公开用于非营利性科学研究。此外，数据集是奈梅亨大学行为科学研究所的财产。虽然目前尚不清楚所提供的数据是否经过IRB的批准，但RaFD数据集广泛用于非商业科学研究，也可用于面部表情注册任务。该数据集可从以下网址下载:www.rafd.nl。

OASIS-3数据集[26]由影像研究开放获取系列(OASIS)提供，其中包含42至95岁人类受试者的脑磁共振成像数据扫描。该数据集还包含脑解剖结构的体积分割图。它可以在OASIS数据使用条款(DUT)下使用，该条款具有知识共享署名4.0许可。参与受试者同意Knight adr创建的项目遵循华盛顿大学医学院机构审查委员会(IRB)批准的程序，因此所有数据的受保护健康信息(PHI)已被排除在外。数据可从https://www.oasis-brains.org下载。

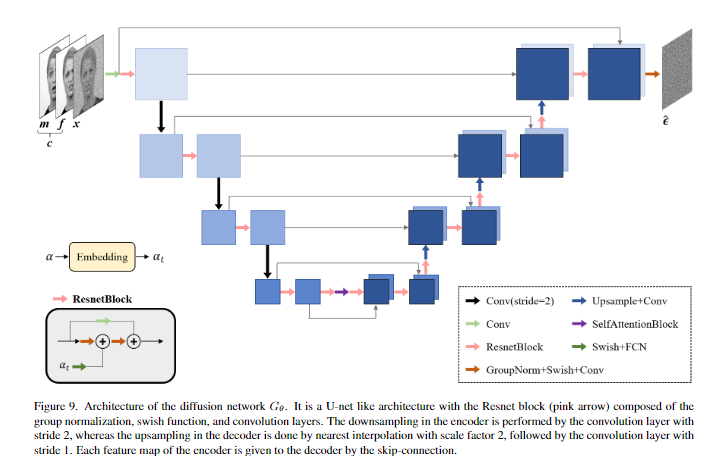
**B.实施细节**

**B.1代码**

我们基于SR3[34]实现了我们提出的DiffuseMorph，提供了条件扩散去噪模型的公共代码。此外，利用voxelmorph1[4]的网络架构，我们训练和测试了所提出的图像配准方法，并使用Pytorch 1.4.0库[32]在Python中实现。复制模型的源代码可从https://github.com/DiffuseMorph/DiffuseMorph获得。

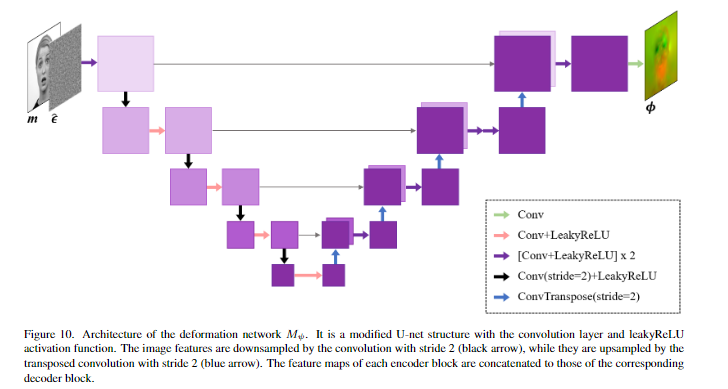
**B.2网络结构**

扩散网络如本文所述，我们采用了DDPM[18]的网络体系结构提出了融合网络Gθ的方法。扩散网络的结构如图9所示。具体来说，它由四个编码器和四个解码器块组成。每个块都有由组归一化[42]、swish函数[14]和卷积层组成的Resnet块，它接受噪声级α的嵌入输入以及前一层的输出。在最后一个编码器块，特征映射由自关注模块[40]参与。此外，每个解码器块还通过跳过连接获取编码器输出的特征映射，该跳过连接使解码器能够使用输入的编码信息。因此，当将运动图像m、固定图像ef和扰动目标图像x给定给噪声等级为α的扩散网络时，网络估计变形的分数函数。每个块的卷积通道数等细节如表4所示。在这里，我们根据输入图像的尺寸来配置卷积层的核维。此外，我们将3D扩散网络的通道数量设置为尽可能深的GPU内存。



**变形网络：**

对于变形网络Mψ，我们实现了voxelmorph1[4]，它呈现了图像配准的网络架构。如图10所示，变形网络与扩散网络类似，是一个带有编码器和解码器块的u形网络。但是，与Resnet块不同，每个块都有convolution和leakyReLU[17]层，称为CL单元，使网络能够专注于图像处理和学习特征提取。同样，由于卷积的下采样是用步幅2进行的，上采样是用步幅2的转置卷积进行的。最终的网络输出由步长为1的卷积生成。因此，给定运动图像m和扩散网络的输出，变形网络估计将运动图像扭曲成固定图像的配准场φ。变形网的详细结构如图5所示。与扩散网络类似，卷积层的配置取决于输入图像的尺寸。



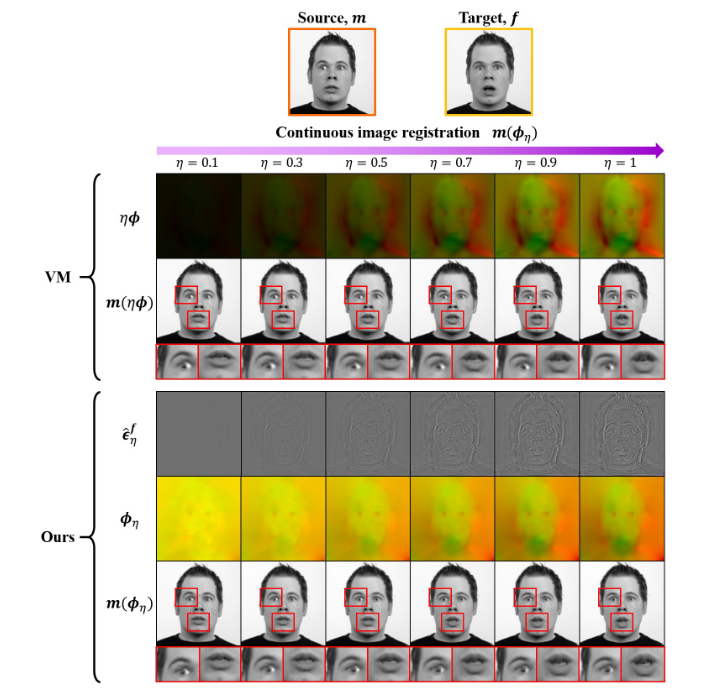
**B.3数据处理**

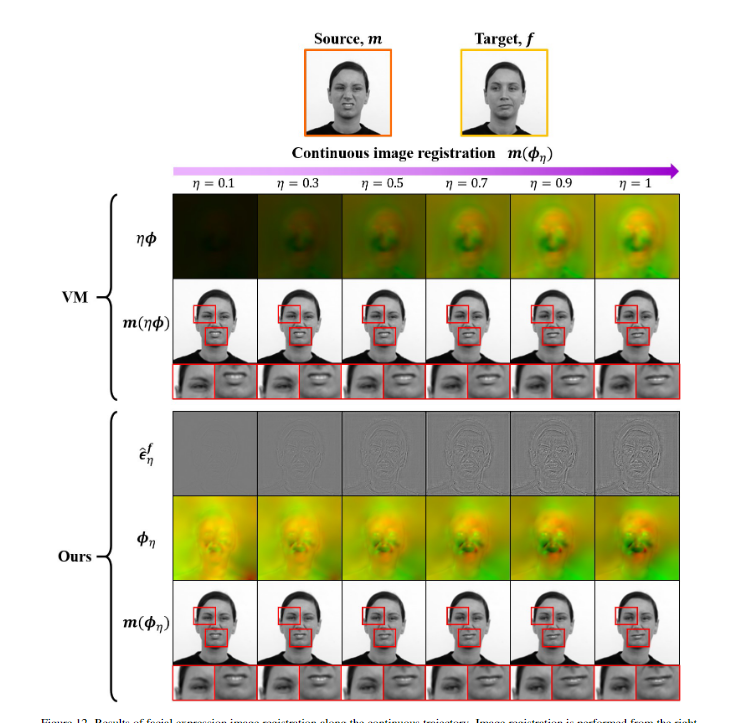
主论文实验中使用的灰度人脸数据和预处理后的脑MR数据的强度范围为[0,1]。我们将数据的强度范围缩放为[- 1,1]。然后，使用缩放后的图像对噪声目标进行采样，并将运动和固定图像的条件作为输入输入到模型中。在这里，由于运动图像是通过线性插值的空间变换对变形场进行变形的，所以我们在对运动图像进行翘曲之前将运动图像重新缩放为[0,1]。对于面部数据的数据增强，我们使用了随机水平翻转。另外，我们使用随机水平/垂直翻转和随机旋转90度来获取大脑数据。

**C.额外的实验结果**

**C.1沿着持续策略的配准**

该方法可以实现运动图像沿轨迹向固定图像的连续配准，这是本文的主要贡献之一。在这里，我们通过变化η乘以扩散网络输出的分数来显示连续变形的附加结果。图11和图12显示了面部表情图像配准任务的结果。对于VM[4]的比较方法，在配准域中进行线性插值，而对于我们的方法，通过调整η在分数函数的潜在空间中进行线性插值。从结果可以看出，虚拟机的估计变形场仅在尺度上发生变化，而配准场的相对空间分布没有变化。因此，眼睛等特定面部运动的变化并不清晰可见。另一方面，在我们提出的方法中，从缩放分数函数值估计的变形场不仅仅是虚拟机中矢量场的缩放版本，而是根据位置显示非常动态变化的运动场(例如，更具体地说，与其他背景相比，沿着眼睛和嘴巴的一致运动)。因此，由此产生的中间变形图像与运动和固定图像有明显的变化。

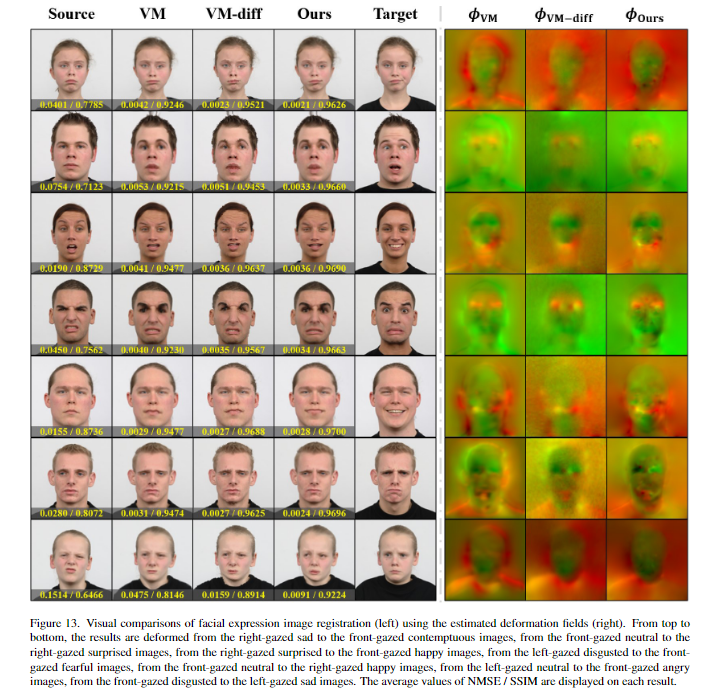




**C.2配准对比结果**

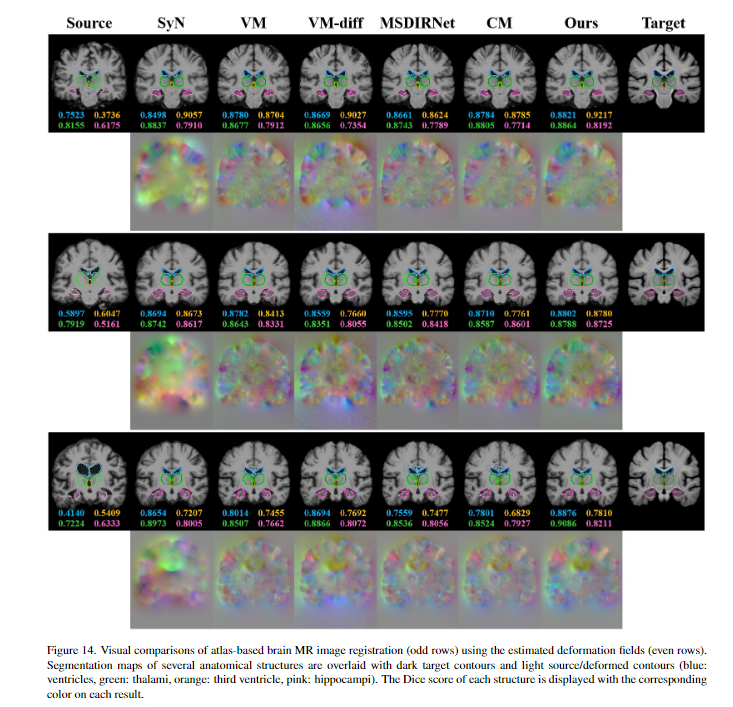
**二维面部表情图像配准**

对于面部表情图像配准任务，我们将所提出的方法与VM-diff[10]和VM-diff[10]进行了比较。我们使用我们模型的变形网络Dψ的相同架构来实现这些方法，并训练网络直到损失收敛以进行公平比较。图13是对男性、女性和儿童的各种面部表情图像进行视觉对比的结果。论文主体部分报告了定量评价结果。与其他方法相比，我们可以看到，我们的模型通过平滑变形场将运动源图像更精确地变形为固定目标图像。这也可以通过显示在每个结果上的NMSE和SSIM的定量评估来观察



3D脑MR图像配准

为了验证三维脑图像配准的性能，我们采用了几种比较方法:SyN [2] by Advanced Normalization Tools (ANTs)[3]、VM[4]、VM-diff[10]、MSDIRNet[28]和CM[22]。对于基于学习的方法，我们使用变形网络Dψ的三维模型作为基线网络，并设置相同的参数进行公平比较。图14为基于图谱的脑图像配准结果。每个结果都显示了几个结构的Dice分数，总体定量评估结果可以在主论文中找到。这表明，与其他方法相比，所提出的DiffuseMorph方法使运动源图像的变形更接近于固定目标图像。此外，从变形结果上多个脑结构的重叠分割边界可以观察到，我们模型的配准可以使运动源图像更准确地对齐到固定图像中，不仅在整个形状，而且在细节结构上。



**C.3合成变形图像的生成**

除了图像配准之外，由于扩散和变形网络的联合训练，我们的DiffuseMorph通过反向扩散过程提供图像生成。如本文所述，在给定一对运动图像和固定图像的条件下，生成过程从运动图像上向前扩散一步开始。然后，通过反向扩散步骤迭代细化具有一定噪声水平的运动图像。面部表情图像的生成过程如图15所示。从噪声等级为α200的运动图像开始，经过80步扩散得到采样结果。随着我们的方法学习了各种对面部表情图像的变形分数函数，我们可以看到，从运动图像中生成的样本变得与固定图像相似。这清楚地表明我们的模型具有控制图像生成和图像配准的能力。

