



计算机科学

COMPUTER SCIENCE

基于子空间特征相互学习的MRI与PET/SPECT图像融合

张瑛, 聂仁灿, 马朝振, 余仕双

引用本文

张瑛, 聂仁灿, 马朝振, 余仕双. 基于子空间特征相互学习的MRI与PET/SPECT图像融合[J]. 计算机科学, 2022, 49(11A): 211000171-6.

ZHANG Ying, NIE Ren-can, MA Chao-zhen, YU Shi-shuang. MRI and PET/SPECT Image Fusion Based on Subspace Feature Mutual Learning [J]. Computer Science, 2022, 49(11A): 211000171-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于机器学习的剩余使用寿命预测实证研究](#)

Empirical Research on Remaining Useful Life Prediction Based on Machine Learning

计算机科学, 2022, 49(11A): 211100285-9. <https://doi.org/10.11896/jsjcx.211100285>

[基于深度神经网络与联邦学习的污染物浓度预测二次建模](#)

Secondary Modeling of Pollutant Concentration Prediction Based on Deep Neural Networks with Federal Learning

计算机科学, 2022, 49(11A): 211200084-5. <https://doi.org/10.11896/jsjcx.211200084>

[深度神经网络的对抗攻击及防御方法综述](#)

Survey of Adversarial Attacks and Defense Methods for Deep Neural Networks

计算机科学, 2022, 49(11A): 210900163-11. <https://doi.org/10.11896/jsjcx.210900163>

[融合多层次视觉信息的人物交互动作识别](#)

Human-Object Interaction Recognition Integrating Multi-level Visual Features

计算机科学, 2022, 49(11A): 220700012-8. <https://doi.org/10.11896/jsjcx.220700012>

[R-YOLOv5:自动切割的旋转的文本检测模型](#)

R-YOLOv5:Auto-cutting,Rotated Text Detection Model

计算机科学, 2022, 49(11A): 210900185-6. <https://doi.org/10.11896/jsjcx.210900185>

基于子空间特征相互学习的 MRI 与 PET/SPECT 图像融合

张 瑛 聂仁灿 马朝振 余仕双

云南大学信息学院 昆明 650500

(zhangying@mail.ynu.edu.cn)

摘要 在医学图像中, MRI 图像提供包含细节的纹理结构信息和较好的分辨率, 而 PET/SPECT 图像保留了分子活性信息以及颜色功能信息, 因此, 将它们进行融合是一项重要的任务。大部分现有的方法在融合过程中存在颜色失真、模糊和噪声等问题。为此, 提出了一种新的基于子空间注意力孪生自编码网络 (Subspace Attention-Siamese Auto-encoding Network, SSA-SAEN) 来融合 MRI 和 PET/SPECT 图像中有意义的信息。在图像融合网络中提出 SSA-SAEN, 引入了子空间特征相互学习概念, 利用子空间注意力模块, 使 MRI 和 PET/SPECT 图像能够在自己学习自己特征的同时互相学习彼此的特征, 同时减少信息冗余, 保证高效、完整的特征提取。此外, 通过条件概率模型对所提取的特征进行互补融合, 同时将加权保真项、梯度损失项加入到训练网络中, 以达到网络优化的目的。在公共数据集上进行的大量定性和定量实验表明, 该方法能够得到一幅清晰的融合图像, 表明了该方法与其他先进方法相比的优越性和有效性。

关键词: 子空间注意力; 互补学习; 神经网络; MRI 与 PET/SPECT 图像融合

中图分类号 TP911.73

MRI and PET/SPECT Image Fusion Based on Subspace Feature Mutual Learning

ZHANG Ying, NIE Ren-can, MA Chao-zhen and YU Shi-shuang

School of Information Science and Engineering, Yunnan University, Kunming 650500, China

Abstract In medical imaging, MRI images provide detailed texture information and better resolution, while PET/SPECT images retain molecular activity information and color function information, so fusing them is an important task. Most of the existing methods have some problems in the fusion process, such as color distortion, blur and noise. Therefore, a new subspace attention-siamese auto-encoding network (SSA-SAEN) is proposed to fully fuse meaningful information from MRI and PET/SPECT images. SSA-SAEN is proposed in image fusion network, and the subspace feature mutual learning concept is introduced. By using subspace attention module, MRI and PET/SPECT images can learn each other's features, while reducing information redundancy and ensuring efficient and complete feature extraction. In addition, the conditional probability model is used to complement and fuse the extracted features, and the weighted fidelity gradient loss term is added into the training network to achieve the goal of network optimization. A large number of qualitative and quantitative experiments on public datasets show that the proposed method can obtain a clear fused image, which demonstrates the superiority and effectiveness of the proposed method compared with other advanced methods.

Keywords Subspace attention, Mutual learning, Neural network, MRI and PET/SPECT image fusion

1 引言

随着医学成像技术的不断发展, 医学图像成为了诊断疾病的有效工具^[1]。由于技术的限制, 每种成像方式只能提供有限的信息。例如, 磁共振图像 (MRI) 显示优越的软组织信息和较高的分辨率, 而正电子发射断层显像 (PET) 和单光子发射计算机断层显像 (SPECT) 可以描述血流和显著代谢变化等功能信息, 分辨率却很低。图像融合是将多模态成像设备在同一场景中获得的多源图像中的显著或互补信息进行融合, 是增强人类感知和机器视觉处理的有效技术^[2]。通过医学图像融合技术, 融合结果既保留了 MRI 图像中的位置信息

(如颅骨轮廓), 又保留了 PET 和 SPECT 图像中的分子活性信息 (如肿瘤和器官), 展示了更丰富的细节信息, 从而提高医生临床诊断的可靠性。本文主要研究 MRI 和 PET/SPECT 多模态医学图像融合任务。

目前已经提出了多种医学图像融合技术, 常见的图像融合技术大致可以分为两类, 一类是基于传统方法的图像融合技术, 另一类是基于深度学习方法的图像融合技术。多尺度变换 (Multi-Scale Transform, MST) 方法是最常用的传统方法。基于 MST 的方法, 首先利用特定的规则对源图像进行不同尺度的分解, 然后采用合适的融合策略对分解结果进行融合, 最后通过逆变换得到融合结果。近年来提出了很多

基金项目: 国家自然科学基金 (61966037, 61463052); 中国博士后科学基金 (2017M621586); 云南大学研究生科学基金资助项目 (2020314)

This work was supported by the National Natural Science Foundation of China (61966037, 61463052), China Postdoctoral Science Foundation (2017M621586) and Postgraduate Science Foundation of Yunnan University (2020314).

通信作者: 聂仁灿 (rnie@ynu.edu.cn)

基于 MST 的方法,如基于离散余弦变换 (Discrete Cosine Transform, DCT)^[3]、基于离散小波变换 (Discrete Wavelet Transform, DWT)^[4]、基于拉普拉斯重分解 (Laplacian Re-Decomposition, LRD) 框架^[2]、基于非下采样轮廓波变换 (Non-Subsampled Contourlet Transform, NSCT)^[5]、基于双树复小波变换 (Dual-Tree Complex Wavelet Transform, DT-CWT)^[6]。这些基于 MST 的方法能很好地保留来自源图像的细节,但对噪声的敏感性不高,容易造成融合结果模糊重复,存在出现吉布斯伪影等问题。此外,还有基于混合技术的图像融合方法被提出,例如,基于边界测量脉冲耦合神经网络 (Pulse-Coupled Neural Network, PCNN)^[7] 融合策略和非下采样轮廓变换 (Non-Subsampled Shearlet Transform, NSST)^[8] 能量属性融合策略的多模态医学图像融合算法。上述传统的图像融合方法分解时间长,人工设计的融合规则复杂,因此难以在实际中应用。

近年来,深度学习被广泛用于多模态医学图像融合,与传统的融合方法相比,深度学习具有融合效果稳定、速度快的优点。在过去的几年里,有很多基于深度学习的方法被提出,如基于卷积神经网络 (Convolutional Neural Networks, CNN)^[9]、基于无监督图像融合网络 (Unified Unsupervised Image Fusion Network, U2Fusion)^[10]、基于梯度强度比例维护 (Proportional Maintenance of Gradient and Intensity, PM-GI)^[11] 的快速统一图像融合网络,以及基于双鉴别器条件生成对抗网络 (Dual Discriminator Conditional Generative Adversarial Network, DDCGAN)^[12] 的融合方法。然而,利用深度学习进行医学图像融合的主要障碍是缺乏用于监督学习的标签图像。上述基于 CNN 的方法中都只使用最后一层的结果作为图像特征,这一操作会丢失网络中间层获得的有用信息。在现有的基于传统生成式对抗网络 (GAN)^[10-13] 的融合方法中,融合后的图像训练后仅与一幅源图像相似,会丢失另一幅源图像中包含的一些信息。

针对传统图像融合方法和大多数深度学习技术的不足,本文提出了一种基于子空间注意力孪生自编码网络 (Subspace Attention-Siamese Auto-encoding Network, SSA-SAEN)。SSA-SAEN 包含特征学习和特征融合两部分,通过基于子空间注意力机制的孪生自编码网络相互学习源图像特征,再运用条件概率公式对源图像进行特征互补融合。特征学习部分由基于子空间注意力 (Subspace Attention, SSA)^[14] 的孪生自编码网络构成。源图像通过 SSA 模块相互学习彼此的深层特征,SSA 模块中通过投影和变换图像特征到信号子空间,使细节特征提取更充分,并且能够减少噪声信息的干扰。特征融合部分通过条件概率融合规则将提取到的特征归一化处理作为条件概率,与源图像结合生成互补的融合图像。

2 相关工作

2.1 子空间注意力 (SSA)

Cheng 等^[14]提出了一种子空间注意力机制 (SSA) 用于图像去噪,通过 SSA 去除深度网络中冗余的信息,提取更加准确和细节的图像特征。如图 1 所示,SSA 模块包含两个关键步骤,根据图像特征生成子空间基向量,再将图像特征变换到信号子空间。我们假设 $\mathbf{X}_1, \mathbf{X}_2 \in R^{H \times W \times C}$ 表示两个特征,它们是 CNN 的中间激活特征,具有相同的尺寸。与文献[14]中的

SSA 不同的是,我们的 Conv-block 只用了简单两个卷积层,没有采用不同网络层跳连,降低了网络复杂度。首先,我们基于 $\mathbf{X}_1, \mathbf{X}_2$ 估计 K 个基向量 $[v_1, v_2, \dots, v_K]$, 每个 $v_i \in R^N$ ($N = HW$) 为信号子空间的基向量;然后将 \mathbf{X}_1 投影到上述子空间。

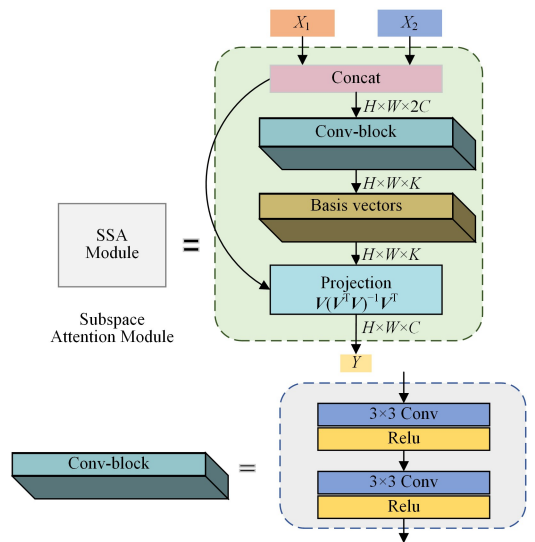


图 1 子空间注意力模块

Fig. 1 Subspace attention module

2.1.1 空间基向量生成

假设 $f_\theta: (R^{H \times W \times C}, R^{H \times W \times C}) \rightarrow R^{N \times K}$ 表示空间基生成函数,空间基生成过程描述如下:

$$\mathbf{V} = f_\theta(\mathbf{X}_1, \mathbf{X}_2) \quad (1)$$

其中, $\mathbf{X}_1, \mathbf{X}_2$ 表示图像特征, \mathbf{V} 表示由基向量构成的矩阵。

2.1.2 子空间投影

给定了上述矩阵 $\mathbf{V} \in R^{N \times K}$, 我们可以将图像特征 \mathbf{X}_1 通过正交线性投影投影到上述空间。

假设 $P: R^N \rightarrow V$, 表示信号子空间的正交投影, 而 P 可以通过式(2)计算得到:

$$\mathbf{P} = \mathbf{V} (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \quad (2)$$

最后图像特征 \mathbf{X}_1 可以在信号子空间重建为 \mathbf{Y} , 表示如下:

$$\mathbf{Y} = \mathbf{P} \mathbf{X}_1 \quad (3)$$

2.2 融合规则

假设给定一对医学图像 \mathbf{A} 和 \mathbf{B} , 多模态医学图像融合任务是生成具有多种模式医学图像信息互补的融合图像。通常,融合任务可以表示为 \mathbf{A}, \mathbf{B} 及其权重分别赋给 \mathbf{A} 和 \mathbf{B} 的加权模型, 如式(4)所示:

$$\mathbf{F} = \mathbf{P}(\mathbf{A}/\mathbf{B}) \odot \mathbf{A} + \mathbf{P}(\mathbf{B}/\mathbf{A}) \odot \mathbf{B} \quad (4)$$

其中, \odot 表示两个矩阵的点乘。 \mathbf{F} 为融合结果, $\mathbf{P}(\mathbf{A}/\mathbf{B})$ 和 $\mathbf{P}(\mathbf{B}/\mathbf{A})$ 分别表示赋给 \mathbf{A} 和 \mathbf{B} 的归一化条件概率, 如式 $\mathbf{P}(\mathbf{A}/\mathbf{B}) + \mathbf{P}(\mathbf{B}/\mathbf{A}) = 1$ 。

3 子空间注意力孪生自编码网络 (SSA-SAEN)

3.1 网络结构

本文的网络是基于子空间注意力孪生自编码网络 (SSA-SAEN), 如图 2 所示, 通过孪生自编码网络学习源图像的特征。同时在编码解码过程中, 子空间注意力模块 (SSA) 被叠加在每一层卷积层之后, 将提取到的两幅源图像的特征输入 SSA, 使两幅图像相互学习彼此的子空间特征, 以提取更完整、有效的信息。

3.2 权值计算

3.2.1 显著性和亮度

图像中显著性高的像素通常是人们最感兴趣的特征。因此,我们可以利用医学图像的显著图像来构造损失函数中保真项的权值,并将较高的权值赋给显著值较高的像素,以促进源图像中显著信息的融合。显著性检测只依赖于像素之间的绝对灰度差,与像素的亮度无关。然而,图像中的高亮度像素并不总是显著的,仅用图像的显著性来做权值估计不能保证源图像的亮度信息完全融合,也不能从根本上避免融合图像的亮度退化。为了有效地融合源图像的显著特征,并且充分融合源图像的亮度信息,我们采用图像的显著性和亮度作为一部分保真项的权值图,其中显著性的计算采用的是基于 LC 的显著性计算方法^[15]。

$$\mathbf{U} = \mathbf{S} \odot \mathbf{L} \quad (5)$$

其中, \mathbf{U} 为保真度项的权值图的显著特征计算, \mathbf{S} 和 \mathbf{L} 分别为图像的显著性特征和亮度特征。

3.2.2 相位一致性和梯度

近年来,神经生物学的一些研究使用功能磁共振成像(Functional Magnetic Resonance Imaging, fMRI)^[16]进一步证实,我们可以提取相位一致性(Phase Congruency, PC)^[17]高

的像素点的强信息特征。考虑到相位一致性是对比度不变的,而图像局部对比度对人类视觉系统(Human Visual System, HVS)对图像质量的感知有影响,我们计算图像梯度大小(Gradient Magnitude, GM)作为次要特征来保留对比度信息。PC和GM是互补的,因为它们反映了HVS在评价图像局部质量方面的差异。我们在构建损失函数的保真项时,将图像的PC和GM相乘作为保真项权重的一部分。

$$\mathbf{V} = \mathbf{PC} \odot \mathbf{GM} \quad (6)$$

其中, \mathbf{V} 表示保真度项的权值图的局部对比度特征计算。

3.2.3 保真项的权值

通过计算图像的显著特征和局部对比度特征的混合方法来度量每个像素的重要性。 \mathbf{W}_A 和 \mathbf{W}_B 分别表示为:

$$\mathbf{W}_A = \frac{\mathbf{U}_A^\alpha + \mathbf{V}_A^\beta}{(\mathbf{U}_A^\alpha + \mathbf{V}_A^\beta) + (\mathbf{U}_B^\alpha + \mathbf{V}_B^\beta)} \quad (7)$$

$$\mathbf{W}_B = \frac{\mathbf{U}_B^\alpha + \mathbf{V}_B^\beta}{(\mathbf{U}_A^\alpha + \mathbf{V}_A^\beta) + (\mathbf{U}_B^\alpha + \mathbf{V}_B^\beta)} \quad (9)$$

其中, $\mathbf{U}_A, \mathbf{V}_A$ 和 $\mathbf{U}_B, \mathbf{V}_B$ 分别是 \mathbf{A} 和 \mathbf{B} 的显著特征和局部对比度特征, α 和 β 是两个可调节常数,用于平衡我们混合测量中显著特征和局部对比特征的重要性。在本文的实验案例中,实证经验性分别取为1和10。

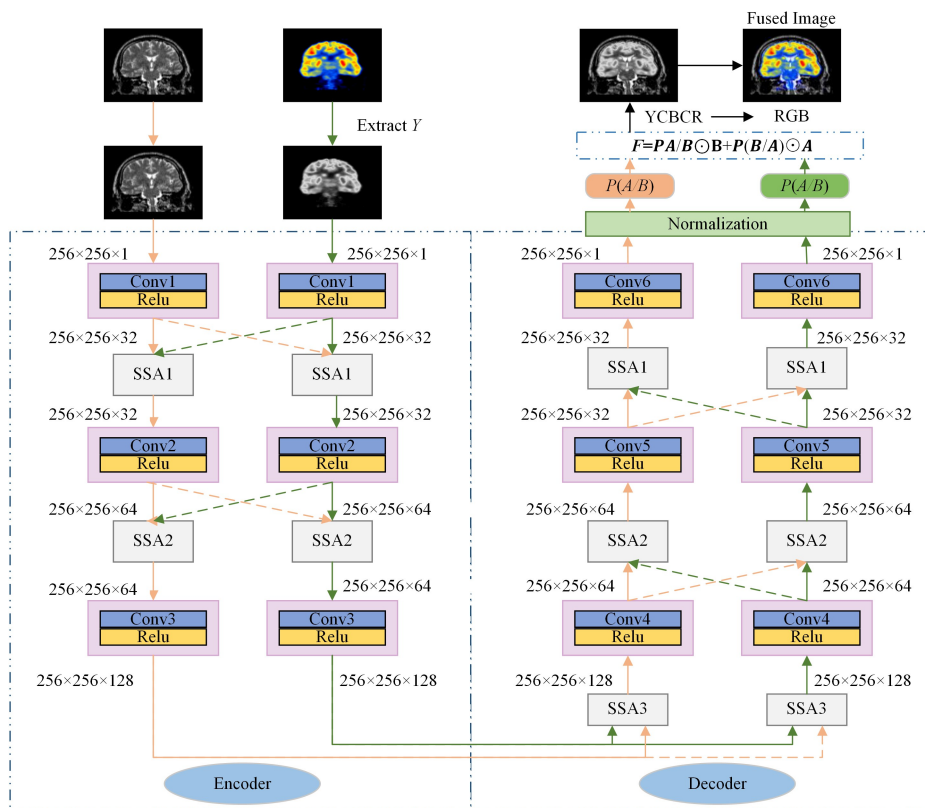


图2 子空间注意力孪生自编码网络

Fig. 2 Subspace attention siamese auto-encoding network

3.3 损失函数

损失函数是在无监督的情况下训练我们的网络模型,本文的损失函数主要由以下两部分构成:

$$Loss_{SSA-SAEN} = L_{w_f} + \gamma L_G \quad (10)$$

其中, L_{w_f} 为加权保真项损失, L_G 为结果与源图像之间的梯度损失, γ 是为了平衡上述两种损失重要性的可调节参数。梯度大小测量用来衡量重建图像与源图像的局部对比质量。

$$L_{w_f} = \|\mathbf{W}_A \odot (\mathbf{F} - \mathbf{A})\|_2^2 + \|\mathbf{W}_B \odot (\mathbf{F} - \mathbf{B})\|_2^2 \quad (11)$$

其中, \mathbf{A} 和 \mathbf{B} 是一对源图像, \mathbf{W}_A 和 \mathbf{W}_B 代表图像 \mathbf{A} 和图像 \mathbf{B} 的权值, \mathbf{F} 是重建的结果图像, $\mathbf{W}_A + \mathbf{W}_B = 1$ 。

L_G 梯度损失:我们对梯度损耗进行了新颖的设计,无论融合图像和源图像之间的 x 方向还是 y 方向,都可以保证每个方向的像素是相似的。因此,融合结果可以进一步增强,如式(12)所示:

$$L_G = \|(F-A)_x\|_2^2 + \|(F-A)_y\|_2^2 + \|(F-B)_x\|_2^2 + \|(F-B)_y\|_2^2 \quad (12)$$

4 实验

在本节中,为了验证所提方法的有效性,我们首先将其与7种方法(NSCT,LRD,PCNN,ZL,DDCGAN,U2FUSION,PMGI)针对4种不同类型的多模式医学图像进行定性比较。对于定量比较,我们利用4个指标来评估融合结果。对比方法均使用网上开源的代码及训练好的模型进行测试得到的结果。另外还进行了参数分析以及消融研究。

4.1 实验细节

为了训练网络,我们选择了74对不同模式的医学图像,包括MR-T1和PET、MR-T2和PET、MR-T2和SPECT_Tc、MR-T2和SPECT_TI,这些图像源自哈佛大学医学图书馆^[18],所有图像大小均为 256×256 像素。

对于彩色PET/SPECT图像,我们从RGB颜色空间转换到YCBCR颜色空间来进行融合。首先,利用所提出的图像融合方案将MRI灰度图像与PET或SPECT图像的亮度分量Y进行融合。将网络得到的融合结果作为最终融合图像的亮度分量Y,替换原彩色图像在YCBCR空间中的Y分量,得到最终的彩色融合图像,整体流程如图3所示。

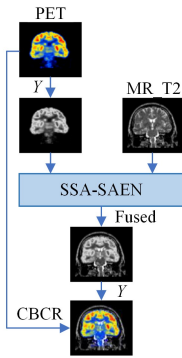


图3 颜色空间转换

Fig. 3 Color space transformation

4.2 评价指标

由于仅凭视觉效果难以准确评价融合图像的性能,因此有必要通过测量客观指标来进行客观评价。我们选择4个客观指标来评估融合结果。标准化互信息(Normalized Mutual Information,NMI)^[19-20]用来衡量源图像中的信息是如何传输到融合图像中的。杨氏结构相似度指标 Q_Y ^[21]用于评价源图像和融合图像之间的结构信息。利用边缘信息传递因子 $Q^{AB/F}$ ^[22]来评价源图像向融合图像的梯度转移的程度。可以利用一个通用图像质量指数(Universal Image Quality Index,UIQI)^[23]来评价不同灰度图像融合方案的性能。

4.3 结果分析

为了证明本文方法的优越性,我们测试了4种不同的MRI/PET,MRI/SEPCT医学图像数据集,从中选取每种类型的10对医学图像。为节省空间,每种医学图像类型只选择一对视觉结果进行展示,如图2、图3所示,其中Set-1代表MR-T1/PET,Set-2代表MR-T1/PET,Set-3代表MR-T2/SPECT_Tc,Set-4代表MR-T2/SPECT_TI,以及本文方法的融合结果,其评价指标测量结果如表1所列。每种类型的10对图像的指标平均值如表2所列。

从图4可以看出,本文方法在MRI/PET,MRI/SEPCT 4种不同类型的图像融合问题上都获得满意的结果。相比之下,NSCT和PCNN的融合结果具有明显的结构信息,但降低了PET/SPECT图像的颜色强度(绿色方框放大部分)。相反,LRD可以保留PET/SPECT图像的色彩功能信息,但会导致MRI图像纹理结构信息的丢失(红色方框放大部分)。ZL方法学习MRI细节部分出现噪声(红色方框显示),DDCGAN和PMGI均出现了细节模糊的问题,同时DDCGAN和U2FUSION的融合结果还出现了颜色畸变。从图4可以看出,我们的方法不仅对光谱信息有很好的处理性能,而且含有丰富的纹理结构。表1、表2列出了不同图像融合方法的定量评价,从表1可以看到,所提方法在图3中的每个图像集的所有指标中得分最高。

表1 不同方法融合图像客观评价指标

Table 1 Evaluation metrics of fused images for different methods

Inputs	评价指标	Nset	Lrd	Pcnn	Zl	Ddcgan	U2fusion	Pmgi	Ours
Set-1 Mr-T1/Pet	NMI	0.6295	0.5634	0.5222	0.8914	0.3832	0.5921	0.5624	0.9938
	Q_Y	0.8754	0.7817	0.6957	0.8656	0.1459	0.6247	0.2890	0.9514
	$Q^{AB/F}$	0.7528	0.5919	0.6042	0.7800	0.3543	0.6400	0.3623	0.8092
	UQI	0.5888	0.5341	0.4558	0.5819	0.0422	0.3319	0.1707	0.6281
Set-2 Mr-T2/Pet	NMI	0.6510	0.6136	0.5824	0.8709	0.4427	0.6493	0.6214	1.0918
	Q_Y	0.8745	0.8083	0.6588	0.8671	0.1390	0.5725	0.3206	0.9722
	$Q^{AB/F}$	0.7560	0.6566	0.6028	0.7752	0.4241	0.6482	0.4334	0.8451
	UQI	0.6114	0.5737	0.4317	0.6117	0.0265	0.2967	0.1773	0.6643
Set-3 Mr-T2/Spect_Tc	NMI	0.7389	0.7785	0.7207	0.9570	0.4988	0.7481	0.6907	1.0807
	Q_Y	0.8728	0.8736	0.8215	0.8655	0.1017	0.4483	0.2407	0.9358
	$Q^{AB/F}$	0.7602	0.7314	0.7357	0.7276	0.3591	0.6630	0.4620	0.7891
	UQI	0.5471	0.5594	0.4995	0.5512	0.0207	0.2310	0.1361	0.5925
Set-4 Mr-T2/Spect_Ti	NMI	0.7450	0.7772	0.6527	0.8605	0.3521	0.6322	0.4987	1.1370
	Q_Y	0.9277	0.9091	0.8401	0.8376	0.0516	0.4170	0.0946	0.9828
	$Q^{AB/F}$	0.8294	0.7923	0.7756	0.7031	0.2990	0.7102	0.3936	0.8939
	UQI	0.5254	0.5179	0.4734	0.4730	0.0052	0.2187	0.0613	0.5548

表 2 不同方法融合图像客观评价指标均值

Table 2 Average evaluation metrics of fused images for different methods

Methods	Metrics			
	NMI	Q_Y	$Q^{AB/F}$	UQI
NSCT	0.7092	0.8878	0.7787	0.5702
LRD	0.6915	0.8446	0.6956	0.5493
PCNN	0.6297	0.7533	0.6812	0.4666
ZL	0.9067	0.8641	0.7541	0.5595
DDCGAN	0.4201	0.1074	0.3506	0.0217
U2fusion	0.6547	0.5084	0.6462	0.2655
PMGI	0.5999	0.2328	0.4137	0.1342
Ours	1.0776	0.9602	0.8355	0.6121

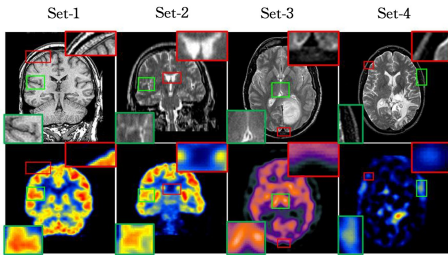


图 4 测试源图像(电子版为彩图)

Fig. 4 Test source images

从表 2 和图 5 可以看出,在所有测试集上的结果表明,我们的方法在每个指标中平均得分最高。因此,结果表明,该方法具有较好的性能,优于其他对比方法。

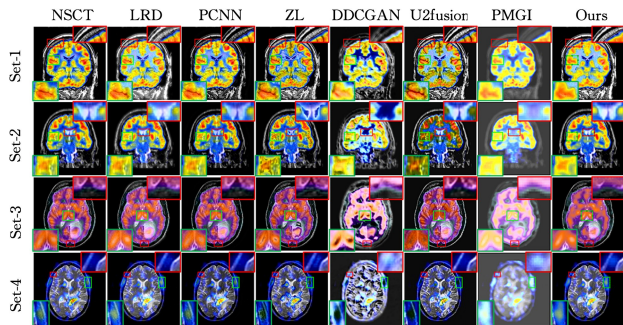


图 5 不同方法的融合图像

Fig. 5 Fused images of different methods

4.4 参数分析

为了进一步证明 γ 在损失函数中的影响对有效训练 SSA-SAENs 非常重要,我们还在各种设置下进行了实验。选择不同的 γ 来训练 SSA-SAENs,以最小化损失 $Loss_{SSA-SAENs}$,产生不同的融合结果,并计算所有指标的平均评价指标。

最后,表 3 列出了采用不同值 γ 时我们方法的融合性能。可以看出,除了 UQI 之外,在 $\gamma=0.05$ 时,我们的方法在各指标下均能获得最佳的融合结果。因此,在我们的例子中,最终选择 γ 为 0.05。

表 3 参数分析客观评价指标均值

Table 3 Average evaluation metrics of fused images for different parameters

Parameters	Metrics			
	NMI	Q_Y	$Q^{AB/F}$	UQI
0.05(Ours)	1.0776	0.9602	0.8355	0.6121
0.10	1.0699	0.9599	0.8354	0.6123
0.15	1.0562	0.9601	0.8355	0.6128
0.20	1.0538	0.9596	0.8351	0.6128
0.25	1.0480	0.9598	0.8340	0.6130

4.5 消融实验

为了验证本文方法中损失函数的优越性,我们将提出的损失函数降级,让 $Loss_{SSA-SAENs}$ 退化为一个去除梯度损失 (L_G) 的版本,以演示梯度损失对所提方法的影响。从表 4 可以看出,与我们的方法相比,降级版本除了 NMI,在其他每个指标上都表现出更差的性能。可见,梯度损失均有利于提高融合质量。

表 4 损失函数消融实验

Table 4 Loss function ablation experiment

Parameters	Metrics			
	NMI	Q_Y	$Q^{AB/F}$	UQI
No- L_G	1.1007	0.9589	0.8354	0.6113
Ours	1.0776	0.9602	0.8355	0.6121

结束语 本文提出了一种基于子空间注意力的孪生自编码网络框架,称为 SSA-SAEN。我们利用子空间注意力机制来相互学习一对源图像的特征信息,并将它们转换为条件概率决策图,通过条件概率模型来进行互补特征融合。此外,我们在特征提取网络中构建了加权保真项和梯度损失项,前者恢复了融合图像的结构,而后者保证了融合图像的细节信息。我们对保真项进行加权,加权构造结合了显著性特征、亮度特征以及局部对比度特征,以达到加权保真度损失的目的。在公共数据集上进行了本文方法与最新方法的大量定性和定量实验,结果验证了本文方法在视觉和定量分析方面的优越性。

参考文献

- [1] WANG L F, WANG R F, LIN S Z, et al. Multimodal Medical Image Fusion Based on Dual Residual Hyper Densely Networks [J]. Computer Science, 2021, 48(2): 160-166.
- [2] LI X, GUO X, HAN P, et al. Laplacian Re-Decomposition for Multimodal Medical Image Fusion [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69 (9 Pt. 2): 6880-6890.
- [3] WANG M, SHANG X. A Fast Image Fusion with Discrete Cosine Transform [J]. IEEE Signal Processing Letters, 2020, 27: 990-994.
- [4] RAJARSHI K, HIMABINDU C. DWT based medical image fusion with maximum local extrema [C]// Proceedings of the International Conference on Computer Communication & Informatics, 2016.
- [5] ZHU Z, ZHENG M, QI G, et al. A Phase Congruency and Local Laplacian Energy Based Multi-Modality Medical Image Fusion Method in NSCT Domain [J]. IEEE Access, 2019, 7: 20811-20824.
- [6] LEWIS J J, O'CALLAGHAN R, NIKOLOV S G, et al. Pixel- and region-based image fusion with complex wavelets [J]. Information Fusion, 2007, 8(2): 119-30.
- [7] WEI T, TIWARI P, PANDEY H M, et al. Multimodal medical image fusion algorithm in the era of big data [J]. Neural Computing and Applications, 2020(3): 1-21.
- [8] LIU Y, LIU S, WANG Z. A general framework for image fusion based on multi-scale transform and sparse representation [J]. Information Fusion, 2015, 24: 147-164.
- [9] XU H, MA J, LE Z, et al. FusionDN: A Unified Densely Con-

- nected Network for Image Fusion [C] // Proceedings of the AAAI. 2020, 44(1):502-518.
- [10] XU H, MA J, JIANG J, et al. U2Fusion: A Unified Unsupervised Image Fusion Network [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.
- [11] ZHANG H, XU H, XIAO Y, et al. Rethinking the Image Fusion: A Fast Unified Image Fusion Network based on Proportional Maintenance of Gradient and Intensity [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020:12797-804.
- [12] MA J, XU H, JIANG J, et al. DDcGAN: A Dual-Discriminator Conditional Generative Adversarial Network for Multi-Resolution Image Fusion [J]. IEEE Transactions on Image Processing, 2020, 29:4980-4995.
- [13] JM A, PL A, WEI Y A, et al. Infrared and visible image fusion via detail preserving adversarial learning [J]. Information Fusion, 2020, 54:85-98.
- [14] CHENG S, WANG Y, HUANG H, et al. NBNet: Noise Basis Learning for Image Denoising with Subspace Projection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:4896-4906.
- [15] SHI Y, YI Y, YAN H, et al. Region contrast and supervised locality-preserving projection-based saliency detection [J]. The Visual Computer, 2015, 31(9):1191-205.
- [16] HENRIKSSON L, HYVÄRINEN A, VANNI S. Representation of cross-frequency spatial phase relationships in human visual cortex [J]. Journal of Neuroscience, 2009, 29 (45): 14342-14351.
- [17] ZHANG L, ZHANG L, MOU X, et al. FSIM: A feature similarity index for image quality assessment [J]. IEEE transactions on Image Processing, 2011, 20(8):2378-2386.
- [18] JOHNSON K A, BECKER J A. The Whole Brain Atlas of Harvard Medical School [EB/OL]. <http://www.med.harvard.edu/AANLIB/>.
- [19] GUO X, NIE R, CAO J, et al. FuseGAN: Learning to fuse multi-focus image via conditional generative adversarial network [J]. IEEE Transactions on Multimedia, 2019, 21(8):1982-96.
- [20] HOSSNY M, NAHAVANDI S, CREIGHTON D. Comments on Information measure for performance of image fusion [J]. Electronics letters, 2008, 44(18):1066-1067.
- [21] YANG C, ZHANG J Q, WANG X R, et al. A novel similarity based quality metric for image fusion [J]. Information Fusion, 2008, 9(2):156-160.
- [22] XYDEAS C A, PETROVIC V. Objective image fusion performance measure [J]. Electronics Letters, 2000, 36:308-309.
- [23] TOET A, HOGERVORST M A. Performance comparison of different gray-level image fusion schemes through a universal image quality index [C] // Proceedings of SPIE. 2003:552-561.



ZHANG Ying, born in 1997, postgraduate. Her main research interests include information fusion, image processing and deep neural network.



NIE Ren-can, born in 1982, Ph.D, associate professor, master supervisor. His main research interests include neural networks, image processing and machine learning.