

# 多尺度特征学习和边缘增强的医学图像融合

肖万新<sup>1,2</sup>, 李华锋<sup>1,2</sup>, 张亚飞<sup>1,2\*</sup>, 谢明鸿<sup>1</sup>, 李凡<sup>1,2</sup>

<sup>1</sup>昆明理工大学信息工程与自动化学院, 云南 昆明 650500;

<sup>2</sup>云南省人工智能重点实验室, 云南 昆明 650500

**摘要** 医学图像融合作为一种对多种不同形式的医学图像所包含的信息进行整合的有效方法, 被广泛用于疾病诊断和治疗规划等各种临床应用中。然而, 现有的医学图像融合方法没有很好地解决不同器官之间的边界模糊问题, 使得融合后的图像较难理解。为了解决这个问题, 提出了一种基于多尺度特征学习和边缘增强的医学图像融合模型。具体地, 首先使用多个不同空洞率的空洞卷积扩大感受野, 使模型学习到源图像更具判别性的多尺度特征。然后, 根据最大值融合策略融合源图像特征, 得到融合特征, 并通过卷积层对其进行重建, 得到融合图像。同时, 引入边缘增强模块, 增强融合图像中的边缘信息, 从而较好地解决了医学图像融合中不同器官之间的边界模糊问题。实验结果证明, 所提方法得到的结果无论从主观视觉效果还是客观定量评价, 均优于对比方法。

**关键词** 医用光学; 医学图像融合; 多尺度特征学习; 深度学习; 空洞卷积; 边缘增强

中图分类号 TP391

文献标志码

doi: 10.3788/LOP202259.0617029

## Medical Image Fusion Based on Multi-Scale Feature Learning and Edge Enhancement

Xiao Wanxin<sup>1,2</sup>, Li Huafeng<sup>1,2</sup>, Zhang Yafei<sup>1,2\*</sup>, Xie Minghong<sup>1</sup>, Li Fan<sup>1,2</sup>

<sup>1</sup>Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;

<sup>2</sup>Yunnan Key Laboratory of Artificial Intelligence, Kunming, Yunnan 650500, China

**Abstract** As an effective method for integrating the information in different forms of medical images, medical image fusion has been commonly used in various clinical applications, such as disease diagnosis and treatment planning. However, the existing medical image fusion methods do not effectively solve the problem of blurred boundaries between different organs, making the fused images more difficult to understand. Therefore, to solve this problem, this paper proposes a medical image fusion model based on multi-scale feature learning and edge enhancement. First, the receptive field is expanded using multiple dilated convolutions with different dilate rates to enable the model to learn more discriminative multi-scale features of the source images. Then, according to the maximum fusion strategy, the source image features are fused to obtain the fused feature. The convolutional layer is used to reconstruct it to obtain the fused image. Further, the edge enhancement module is introduced to enhance the edge information in the fused image to better solve the problem of blurred boundaries between different organs in medical image fusion. The experimental results show that the results obtained by using the proposed method are superior to comparison methods in terms of subjective visual effects and objective quantitative evaluation.

**Key words** medical optics; medical image fusion; multi-scale feature learning; deep learning; dilated convolution; edge enhancement

收稿日期: 2021-10-25; 修回日期: 2021-11-16; 录用日期: 2021-11-29

基金项目: 国家自然科学基金(62161015, 61966021)

通信作者: \*zyfeimail@163.com

## 1 引言

不同模态的医学图像提供不同类型的信息,它们在临床诊断中发挥着越来越重要的作用。由于成像机制不同,不同模态的医学图像关注不同类型的器官/组织信息。例如,计算机断层扫描(CT)图像显示密集结构的信息,如骨骼和植入物,而T2加权磁共振成像(MR-T2)图像显示高分辨率的解剖信息,如软组织<sup>[1]</sup>。为了获得足够的信息进行准确的诊断,医生经常需要对用不同模式捕获的医学图像进行顺序分析,但这种分离式的诊断方式在许多情况下会带来不便。解决这一问题的一种有效方法是医学图像融合,它的目的是生成一幅复合图像,对多种不同形式的医学图像中所包含的互补信息进行整合<sup>[2-5]</sup>。

现有的医学图像融合方法大致可以分为三类,即基于多尺度变换的方法<sup>[6-8]</sup>、基于字典学习的方法<sup>[9-10]</sup>和基于深度学习的方法<sup>[11-13]</sup>。在基于多尺度变换的方法中,常用的多尺度变换包括小波变换<sup>[4]</sup>、双树复小波变换<sup>[14]</sup>和非下采样轮廓波变换<sup>[15]</sup>等。与多尺度变换方法不同的是,基于字典学习的稀疏表示方法可以有效地避免基于多尺度变换方法的缺陷,表现出良好的融合性能<sup>[16-17]</sup>。由于深度学习技术可以有效地提取源图像的显著特征,其在图像融合领域得到了广泛应用。但是,医学图像空间分辨率较低,细节内容较少,深度神经网络结构过深容易造成信息丢失。同时,不同源图像的亮度语义是完全不同的,例如,CT图像的亮度代表组织的密度,MR-T2图像的亮度代表组织的流动性和磁性。现有图像融合方法忽略了不同源图像的语义冲突,导致一些融合图像中的脑组织边界模糊,影响临床医生的诊断。

为了解决以上问题,本文提出一种基于多尺度特征学习和边缘增强的医学图像融合模型。具体地,该模型由4个模块组成:特征提取模块、融合模块、重建模块和边缘增强模块。其中,特征提取模块主要由多个不同空洞率的空洞卷积构成,作用是扩大感受野,使网络学习到源图像更具判别性的多尺度特征。融合模块采用最大值策略融合源图像特征。重建模块含有数个不同尺寸的卷积层,作用是将融合特征重建成融合图像。医学图像的语义信息大部分由图像中的边缘纹理表征。如果要使融合图像中的边缘更加清晰,可以让不同器官之间

的边界更加明显,从而更好地帮助临床医生进行诊断。基于这一思路,在模型中引入了边缘增强模块,它由两个边缘增强支路组成,作用是得到源图像的边缘图,并将它们与重建出的融合图像结合,得到最终的融合图像。

## 2 基于深度学习的图像融合

近年来,深度学习由于可以很好地从源图像中提取关键信息,被成功应用于图像融合领域。在多聚焦图像融合中,Liu等<sup>[18]</sup>提出了一种基于卷积神经网络(CNN)的多聚焦图像融合方法。在多模态图像融合中,Zhong等<sup>[19]</sup>提出了一种基于CNN的能同时进行图像融合和超分辨率重建的方法。Liu等<sup>[20]</sup>提出了一种用于图像融合的卷积稀疏表示,其中反卷积网络旨在构建层级结构,每层由一个编码器和一个解码器组成。在遥感图像融合中,Masi等<sup>[21]</sup>提出了一种有效的三层架构来解决全色锐化问题,通过添加几个非线性辐射指数图来增强输入,以提升融合性能。在红外和可见光图像融合中,Lahoud等<sup>[22]</sup>将源图像分解为基础层和细节层,基础层在视觉显著图的指导下进行融合,而细节层则根据CNN模型生成的权重图进行合并。Zhang等<sup>[23]</sup>提出了一种端到端的图像融合架构,包括特征提取、融合和重建阶段。Jung等<sup>[24]</sup>提出了一种无监督的深度图像融合方法,通过应用结构张量表示来设计损失函数,该表示可以衡量融合图像和源图像在梯度域中的差异。Ma等<sup>[12]</sup>首次将生成对抗网络引入图像融合领域,用于融合红外和可见光图像。在他们的方法中,生成器网络用于从源图像生成融合图像,而鉴别器网络用于进一步从可见光图像中提取空间细节。在医学图像融合中,Hou等<sup>[25]</sup>将卷积神经网络和非下采样剪切波变换方法结合,用于CT和MRI图像融合。Fan等<sup>[26]</sup>以U-net网络为基础,构建了一个医学图像融合模型,并使用重构损失来衡量重构图像和源图像之间的语义差别。Liang等<sup>[27]</sup>提出一种端到端的深度融合网络,该网络使用多层级联技术,避免了浅层特征的丢失。Zhang等<sup>[28]</sup>提出了一种基于DenseNet的医学图像融合模型,该模型在使用更少参数数量的同时取得了较好的融合效果。

尽管上述图像融合方法有着相对良好的表现,但它们都忽略了不同源图像中的边缘语义信息,导致医学图像融合结果中不同器官之间的边界模糊,

影响临床医生的诊断。因此,在本文提出的医学图像融合模型中加入了边缘增强模块,通过将源图像中的边缘信息与融合结果相结合,强化了融合图像中的边缘信息,从而使不同器官之间的边界更清晰。

### 3 方法介绍

#### 3.1 多尺度特征学习与融合网络

针对医学图像融合问题,提出了多尺度特征学习和边缘增强融合网络(MFEnet)。该网络的框架如图 1 所示,主要由特征提取模块、融合模块、重建模块和边缘增强模块 4 部分组成。源图像信息特征的有效表示对图像融合任务非常重要,通过扩大感受野和学习多尺度特征来获取源图像更具判别性

的特征。在相同的计算成本下,空洞卷积可以显著扩大卷积核的感受野,因此以多个不同空洞率的空洞卷积为基础,构建了特征提取模块。具体地,先经过一个  $3 \times 3$  的卷积层提取源图像的浅层特征后,使用一个多尺度的特征提取层提取特征,该特征提取层包含 3 个空洞卷积。空洞卷积的卷积核尺寸都为 3,空洞率分别为 1,3,5。不同空洞率的空洞卷积可以提取不同尺度的特征。将三个不同尺度的特征拼接,得到包含多尺度的特征信息,从而在提高源图像特征判别性的同时使融合图像具有更好的纹理细节。为了补偿空洞卷积造成的信息丢失,将浅层特征跳跃连接到多尺度特征。最后使用两个  $3 \times 3$  的卷积层进一步提取深层特征。

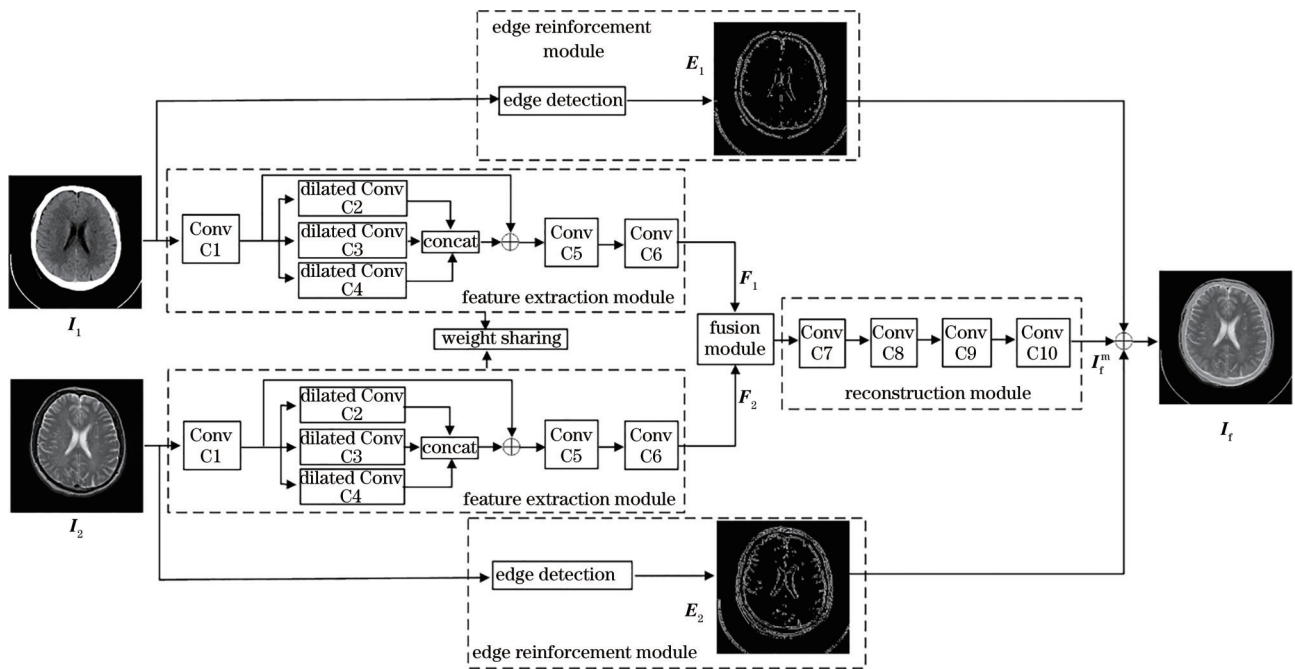


图 1 多尺度特征学习和边缘增强融合网络的基本框架

Fig. 1 Basic framework of multi-scale feature learning and edge enhancement fusion network

融合模块利用最大值融合策略对编码器提取到的源图像特征进行融合,可描述为

$$F_i(i, j) = \begin{cases} F_1(i, j), & |F_1(i, j)| \geq |F_2(i, j)| \\ F_2(i, j), & |F_1(i, j)| < |F_2(i, j)| \end{cases}, (1)$$

式中:  $F_1$  和  $F_2$  分别表示待融合图像的特征;  $(i, j)$  表示特征的坐标。

重建模块由 4 个  $3 \times 3$  的卷积层组成,作用是将融合特征重建为融合图像。整个网络所用卷积层的参数如表 1 所示。特征提取模块和重建模块作为一个整体,在训练阶段一起训练。训练好后,分别通过特征提取模块提取源图像特征,然后使用融合

模块融合这两类特征,最后通过重建模块重建融合图像,同时与边缘增强模块得到的结果相结合,得到最终的融合图像。

#### 3.2 边缘增强模块

边缘是医学图像的重要特征之一,它有助于确定器官的正确形状、结构和边界,分辨重叠的器官。因此,在医学融合图像中,保留源图像中的边缘信息非常重要,它能够帮助临床医生分辨不同的器官,从而更好地诊断病情。但是,现有的方法忽视了医学融合图像中的语义冲突问题,导致融合结果中边缘不明确,加大了临床医生的诊断难度,如图 2 所示。图 2(a) 中 CT 图像的亮度表示组织的密度,



表 1 MFEnet 结构

Table 1 Structure of MFEnet

Module	Convolutional layer	Size	Number of input channels	Number of output channels	Activation layer
Feature extraction module	C1	3	1	64	ReLU
	C2	1	64	64	ReLU
	C3	3	64	64	ReLU
	C4	5	64	64	ReLU
	C5	1	192	64	ReLU
	C6	3	64	64	ReLU
Reconstruction module	C7	3	64	64	ReLU
	C8	3	64	32	ReLU
	C9	3	32	16	ReLU
	C10	3	16	1	ReLU

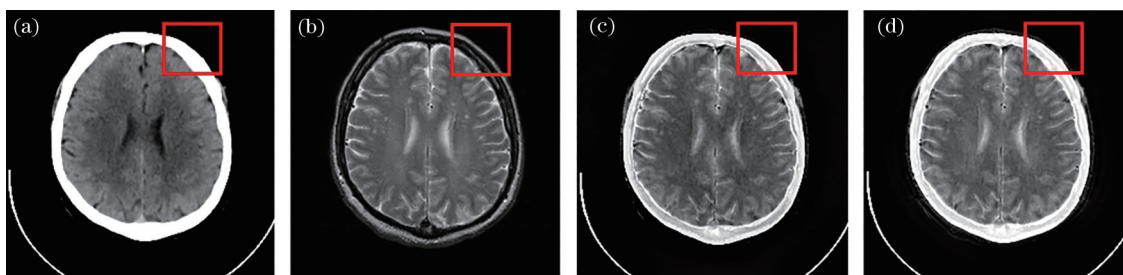


图 2 CNN 和 NSCT-PRPCNN 方法的融合结果。(a) CT 源图像; (b) MR-T2 源图像; (c) CNN; (d) NSCT-PRPCNN

Fig. 2 Fusion results obtained by the CNN and NSCT-PAPCNN methods. (a) Source CT image; (b) source MR-T2 image; (c) CNN; (d) NSCT-PRPCNN

高亮度部分为高密度低流动性的颅骨,图 2(b)中 MR-T2 图像的亮度表示组织的流动性,高亮度部分为低密度高流动性的脑脊液。也就是说,不同源图像的亮度语义信息完全不同。现有 CNN 和 NSCT-PRPCNN 的融合结果,分别如图 2(c)、(d)所示。融合图像中的脑脊液和颅骨的边界比较模糊,这加大了临床医生的诊断难度。

为了解决这个问题,将源图像的边缘信息跟融合图像结合,从而使融合图像中不同器官的边界变得清晰。具体地,使用 Canny 边缘检测算法分别得到待融合图像的边缘图,然后将它们叠加到解码器重建出来的融合图像中,得到最终的边缘增强后的融合图像。边缘增强过程可表示为

$$I_f = I_r^m + \text{Canny}(I_1) + \text{Canny}(I_2), \quad (2)$$

式中: $I_1$ 和 $I_2$ 分别表示待融合的源图像; $I_f$ 代表最终的融合图像; $I_r^m$ 代表重建的融合图像; $\text{Canny}(\cdot)$ 是 Canny 边缘检测算子。

### 3.3 损失函数

损失函数在深度学习优化中扮演着重要角色,选取合适的损失函数对深度学习模型的结果有重要影响,它用于衡量网络的输出结果和标签数据之

间的差异,网络训练的最终目的是使损失函数最小化。由于医学图像中的亮度信息由像素强度表征,而纹理细节信息可以部分由其梯度表征,MFEnet 网络的总损失函数强制使网络的输出图像和输入图像拥有相似的像素强度和梯度,它由两项组成:

$$L = \frac{1}{WH} \left( \|I_o - I_i\|_F^2 + \alpha \|\nabla I_o - \nabla I_i\|_F^2 \right), \quad (3)$$

式中: $W$ 和 $H$ 分别表示输入图像的宽度和高度; $I_i$ 是 MFEnet 编码器的输入; $I_o$ 是解码器的输出图像; $\|\cdot\|_F$ 是矩阵的 Frobenius 范数; $\nabla$ 代表梯度算子; $\alpha$ 为经过实验选取的平衡参数,在实验中设置为 6。为了清晰地表达所构建融合网络 MFEnet 的训练和测试过程,训练和测试的算法如表 2 所示。

## 4 实 验

### 4.1 数据集和训练细节

实验使用的数据集为从 <http://www.med.harvard.edu/AANLIB/home.html> 上获得的 CT 和 MR-T2 医学图像。两种图像各有 360 张,图像编号一一对应组成 360 个图像对,其中选取 240 个图像对作为训练集,80 个图像对作为验证集,并从剩下的图像对中选取 20 个作为实验的测试集,部分样本图像如

表 2 MFEnet 训练和测试的算法  
Table 2 Algorithms of MFEnet training and testing

**Algorithm** MFEnet training and testing algorithms

**Training**

**Input:** Training set source image  $I_i$

**Output:** Reconstructed image  $I_o$

- 1) Randomly select  $m$  source images from the training set  $\{I^{(1)}, \dots, I^{(m)}\}$
- 2) Input  $m$  source images into the feature extraction module to generate source image features  $F$
- 3) Input  $F$  into the reconstruction module to generate a reconstructed image  $I_o$ .
- 4) Use Adam optimizer to update the parameters of the feature extraction module and reconstruction module:

$$\nabla_{\theta} \frac{1}{WH} \left( \|I_o - I_i\|_F^2 + \alpha \|\nabla I_o - \nabla I_i\|_F^2 \right)$$

- 5) If the number of iterations is equal to epoch, the training ends, otherwise repeat steps 1)-4)

**Testing**

**Input:** Testing set source images  $I_1$  and  $I_2$

**Output:** Fused image  $I_f$

- 1) Input  $I_1$  and  $I_2$  into the feature extraction module to get the source image features  $F_1$  and  $F_2$
- 2) Input  $I_1$  and  $I_2$  into the edge enhancement module to get the source image edge maps  $E_1$  and  $E_2$
- 3) Input  $F_1$  and  $F_2$  into the fusion module and the reconstruction module to obtain the intermediate fusion image  $I_f^m$
- 4) Combine  $I_f^m$ ,  $E_1$ , and  $E_2$  to get the final fusion image  $I_f$

图 3 所示,第一行为 CT 图像,第二行为 MR-T2 图像。

在训练阶段,将特征提取模块和重建模块作为一个整体进行训练,摒弃融合模块,输入为从训练集随机选取的单张医学图像,输出为重建图像。训练过程的伪代码如表 2 所示。训练共进行 60 代,使用 Adam 优化器优化损失函数,学习率设为  $1 \times$

$10^{-4}$ ,批次大小设为 16,实验环境为 Pytorch1.7,使用 NVIDIA RXT3090 GPU 训练网络。训练完编码器和解码器后,将待融合图像分别输入特征提取模块,得到图像对应的特征,并通过融合模块和重建模块得到中间融合图像,最后与边缘增强模块的结果相叠加,得到最终的融合图像。

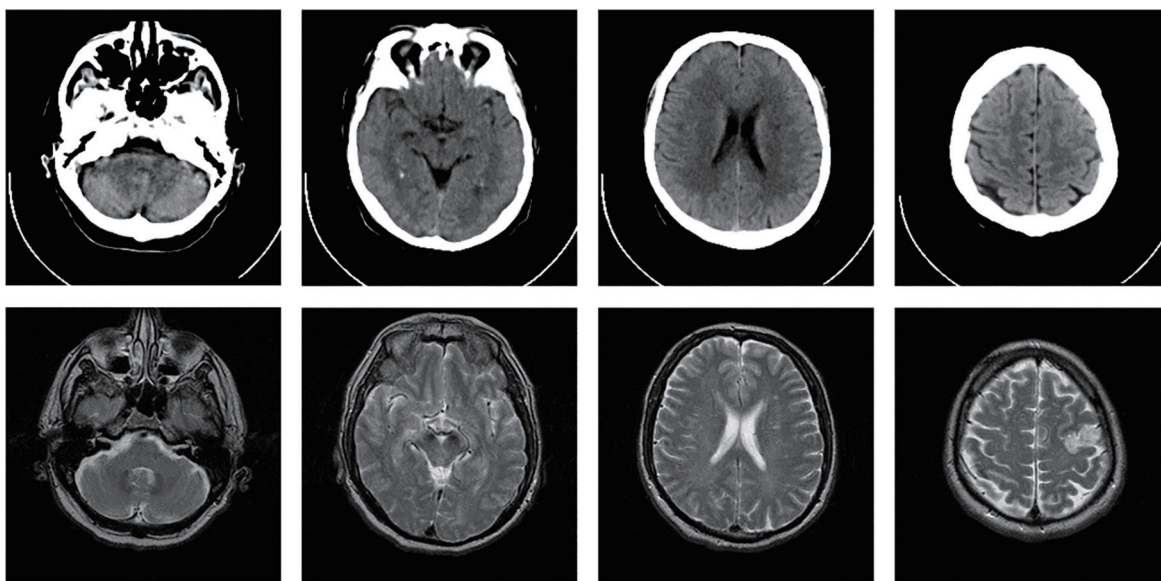


图 3 测试集源图像样本

Fig. 3 Source image samples in testing dataset

**4.2 融合图像质量客观评价指标**

选取了 5 个常用的客观评价指标来评价所提方

法和对比方法的融合性能。它们分别是像素特征互信息(FMI<sub>pixel</sub>)<sup>[29]</sup>、差异相关性的总和(SCD)<sup>[30]</sup>、结

构相似度(SSIM)<sup>[31]</sup>、语义损失(SL)<sup>[27]</sup>和基于结构相似度的度量( $Q_Y$ )<sup>[32]</sup>。 $FMI_{\text{pixel}}$ 是一种基于互信息的无参考客观图像融合评价指标,计算从源图像到融合图像的像素特征信息量,可以测量融合图像中含有多少源图像信息。SCD通过计算源图像和融合图像之间的差异相关性之和来评估图像融合的效果。SSIM衡量融合图像与源图像之间的结构相似度。SL度量融合图像的重建误差,表示融合图像中的语义冲突。 $Q_Y$ 度量融合图像中保留的结构信息量。需要指出的是,除了SL,其他4个指标都是越大越好。

### 4.3 融合方法对比分析

为了验证所提方法的有效性,分别从主观视觉效果和客观质量评价两个角度同5种医学图像融合方法作对比,其中包括3种传统方法,分别为GF<sup>[33]</sup>、NSCT<sup>[7]</sup>、NSCT-PRPCNN<sup>[1]</sup>,2种基于深度学习的

方法,分别为CNN<sup>[18]</sup>、FW-Net<sup>[27]</sup>。5种对比方法和所提方法得到的融合结果如图4所示。FW-Net方法的融合结果的方框中,虽然CT源图像中的亮度信息保留了下来,但可以看到含有较多模糊区域,且颅骨和脑脊液的边界模糊,这说明FW-Net没有很好地保留MR-T2源图像中的纹理边缘细节信息。GF方法的融合结果中,CT源图像的亮度信息丢失严重。在CNN、NSCT、NSCT-PRPCNN方法的融合图像中,丢失了MR-T2源图像中的边缘信息,最外层的边缘几乎看不见。所提方法使用了多尺度特征学习,使得从源图像中提取的特征更丰富全面,且考虑到融合图像中不同器官之间边界模糊问题,通过使用边缘增强模块,保留了源图像中的边缘信息。从图4(h)可以看出,所提方法的融合图像中的各种器官之间的边界清晰,源图像中的亮度和纹理细节也较好保留了下来。

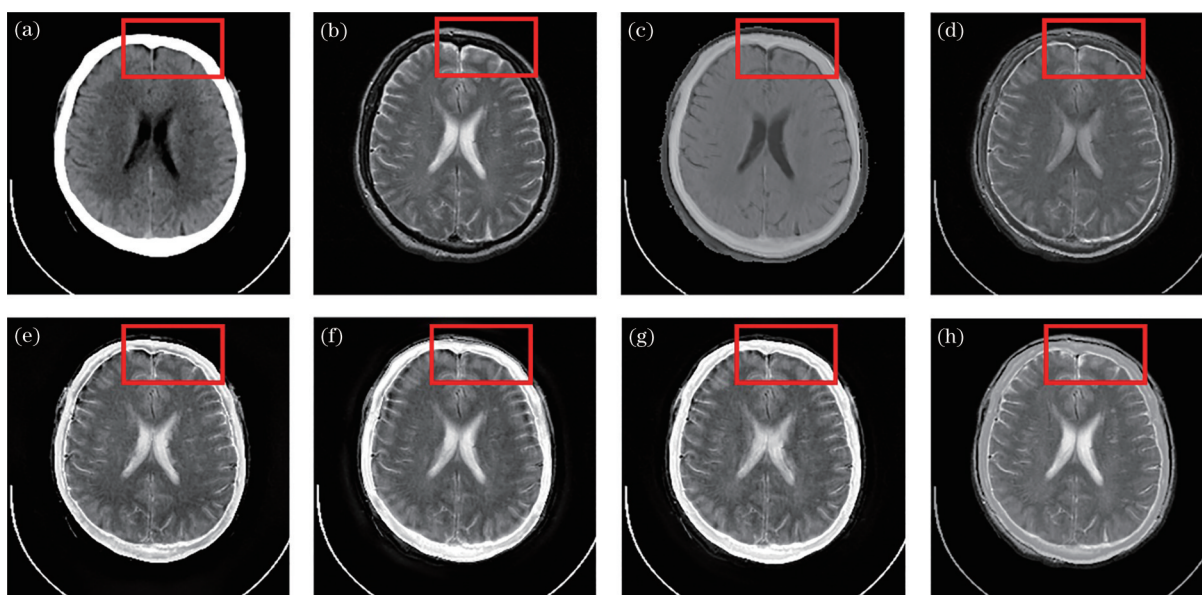


图4 不同方法在视觉上的对比。(a) CT源图像;(b) MR-T2源图像;(c) FW-Net;(d) GF;(e) CNN;(f) NSCT;(g) NSCT-PRPCNN;(h) 所提方法

Fig. 4 Visual comparison of the different methods. (a) Source CT image; (b) source MR-T2 image; (c) FW-Net; (d) GF; (e) CNN; (f) NSCT; (g) NSCT-PRPCNN; (h) proposed method

表3展示了各种对比方法在测试集上的评价指标的平均值,最高的质量评价数值由黑体字标识。从表3可以看出,所提方法的评价指标取得了最优值,这证明了所提方法的有效性。从时间效率上看,所提方法耗时最短,说明融合效率更高。因此,所提使用多尺度特征学习和边缘增强的医学图像融合模型在进行融合时能够获得更高质量的图像。

### 4.4 消融实验分析

#### 4.4.1 边缘增强模块分析

为了分析所提方法中边缘增强模块对融合效果的影响,图5分别展示了无边缘增强模块和有边缘增强模块的融合效果图。从图5的方框中明显可以看出,引入边缘增强模块后,融合结果的边缘细节增多,不同器官之间的边界变得清晰。这说明所提边缘增强模块有助于提升图像的融合质量。



表 3 20 幅融合图像的质量评价指标平均值

Table 3 Average value of quality evaluation index of 20 fused images

Method	FMI <sub>pixel</sub>	SCD	SSIM	SL	Q <sub>v</sub>	Time /s
FW-Net	0.8182	0.7776	0.5108	0.0072	0.6600	0.0199
GF	0.8595	0.8910	0.6612	0.0167	0.8407	0.0912
CNN	0.8637	1.1123	0.5717	0.0134	0.7165	9.601
NSCT	0.873	0.969	0.6160	0.0155	0.7311	3.43
NSCT-PRPCNN	0.8448	1.2690	0.6350	0.0122	0.7575	9.601
Proposed method	<b>0.8741</b>	<b>1.2879</b>	<b>0.7412</b>	<b>0.0069</b>	<b>0.8422</b>	<b>0.0189</b>

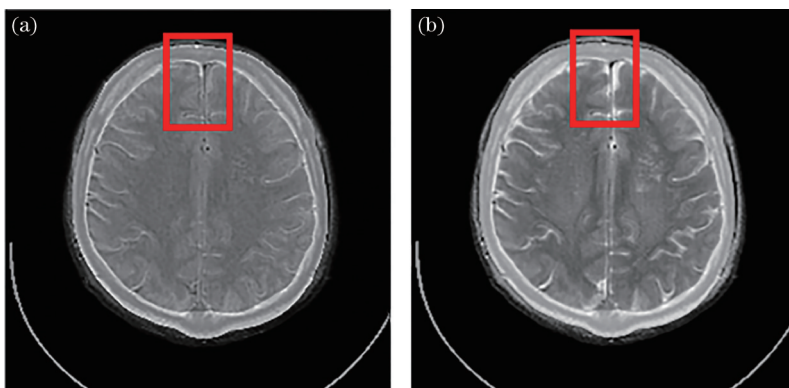


图 5 有无边缘增强支路的融合结果。(a)无边缘增强分支;(b)有边缘增强分支

Fig. 5 Fusion results without and with edge reinforcement branches. (a) Without edge reinforcement branch; (b) with edge reinforcement branch

#### 4.4.2 融合规则分析

融合规则对图像融合任务的融合性能起着重要作用,为了分析所使用的融合规则的合理性,选取 4 种主流的融合规则作为对比:最大值策略<sup>[34]</sup>、相加<sup>[35]</sup>、平均<sup>[36]</sup>和加权平均<sup>[37]</sup>。其中,最大值策略在融合时选取源图像对应位置上的最大特征值作为融合特征;相加策略在融合时直接求和源图像特征值;平均策略将源图像特征平均值作为融合特征;加权平均策略先分别对源图像特征使用 Softmax 函数,得到两张权重图,然后在融合时取对应位置源

图像特征的加权平均作为融合特征。4 种融合规则分别用于所提模型,得到的融合结果质量评价如表 4 所示。从表 4 可以看出,最大值策略的评价指标大部分取得了最优值,并且评价指标 SCD 和 Q<sub>v</sub> 优势明显,这是因为最大值融合规则可以增强融合图像中目标区域和非目标区域的对比度,这有助于增加医学融合图像中不同器官的区分度,帮助网络更好地保留源图像的结构信息,从而提高网络的融合性能。基于上述分析,选用最大值策略作为融合规则。

表 4 4 种不同融合策略融合结果的质量评价指标平均值

Table 4 Average value of the quality evaluation index of fusion results with four different fusion rules

Fusion rule	FMI <sub>pixel</sub>	SCD	SSIM	SL	Q <sub>v</sub>
Addition <sup>[35]</sup>	0.8728	1.1951	<b>0.7422</b>	0.0071	0.8116
Average <sup>[36]</sup>	0.8689	1.1703	0.7270	<b>0.0068</b>	0.8095
Weighted average <sup>[37]</sup>	0.8708	1.2456	0.7337	0.0073	0.8119
Max value <sup>[34]</sup>	<b>0.8741</b>	<b>1.2879</b>	0.7412	0.0069	<b>0.8422</b>

#### 4.5 在其他医学图像融合任务上的表现

为了验证所提方法的泛化能力,本部分展示了所提方法在 T1 加权磁共振成像(MR-T1)图像和 MR-T2 医学图像融合任务上的表现。MR-T1 和 MR-T2 图像在临床诊断中被广泛使用,其中 MR-

T1 图像通常用于观察解剖结构,脂肪、黑色素、蛋白质等在图像中呈现明亮的高信号,脑脊液呈现为低信号。而 MR-T2 图像显示组织病变情况较好,脑脊液呈现为明亮的高信号。图 6 展示了所提方法得到的 MR-T1 和 MR-T2 图像融合结果。其中

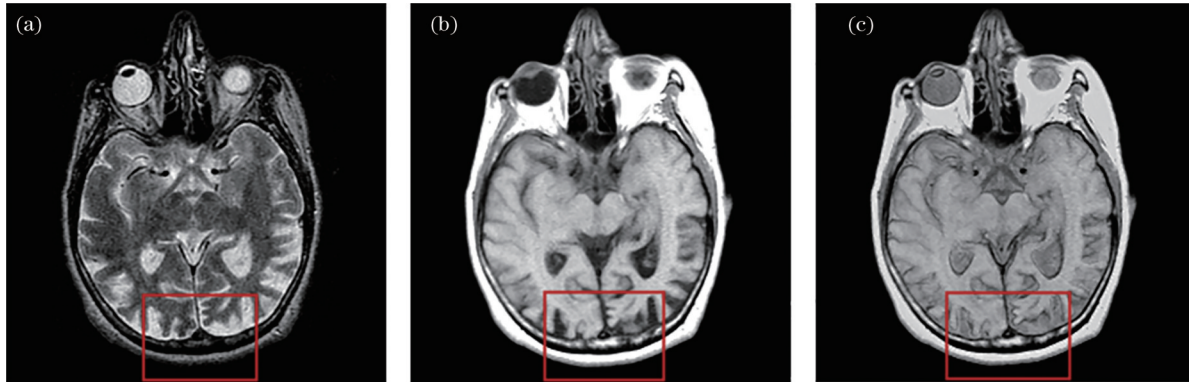


图6 所提方法在MR-T1和MR-T2图像融合任务上得到的融合结果。(a) MR-T1源图像;(b) MR-T2源图像;(c)所提方法  
Fig. 6 Fusion results obtained by the proposed method in MR-T1 and MR-T2 image fusion task. (a) Source MR-T1 image;  
(b) source MR-T2 image; (c) proposed method

图6(a)是MR-T1源图像,图6(b)是MR-T2源图像。从融合结果中可以看出,源图像中的边缘细节信息被保留了下来,不同器官之间边界清晰(如方框中的脑脊液和颅骨边界)。这表明所提方法有较好的泛化能力。

## 5 结 论

提出了一种基于多尺度特征学习和边缘增强的医学图像融合方法。该方法首先使用训练集训练特征提取模块和重建模块,然后用最大值融合策略融合特征提取模块得到的源图像特征,最后利用重建模块重建出中间图像,并与源图像的边缘图结合,得到最终的融合图像。所构建的融合模型能有效解决医学融合图像中器官边界模糊的问题,融合后的图像中不同器官之间的边界清晰,且具有丰富的细节信息。实验表明,不论是主观视觉效果上,还是客观评价指标上,所提方法的融合结果皆优于对比方法。

## 参 考 文 献

- [1] Yin M, Liu X N, Liu Y, et al. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsamped shearlet transform domain [J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, 68(1): 49-64.
- [2] Wang J X, Chen S, Xie M H. Multi-source image fusion based on low-rank decomposition and convolutional sparse coding[J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 2210009.  
王加新, 陈升, 谢明鸿. 基于低秩分解和卷积稀疏编码的多源图像融合[J]. *激光与光电子学进展*, 2021, 58(22): 2210009.
- [3] Li W, Li Z M. NSST-based perception fusion method for infrared and visible images[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 2010014.  
李威, 李忠民. NSST域红外和可见光图像感知融合[J]. *激光与光电子学进展*, 2021, 58(20): 2010014.
- [4] Yang Y. Multimodal medical image fusion through a new DWT based technique[C]//2010 4th International Conference on Bioinformatics and Biomedical Engineering, June 18-20, 2010, Chengdu, China. New York: IEEE Press, 2010: 11495757.
- [5] Du J, Li W S, Lu K, et al. An overview of multimodal medical image fusion[J]. *Neurocomputing*, 2016, 215: 3-20.
- [6] Zhao H, Zhang J X, Zhang Z G. PCNN medical image fusion based on NSCT and DWT[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 2017002.  
赵贺, 张金秀, 张正刚. 基于NSCT与DWT的PCNN医学图像融合[J]. *激光与光电子学进展*, 2021, 58(20): 2017002.
- [7] Zhu Z Q, Zheng M Y, Qi G Q, 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.
- [8] Li H F, Liu X K, Yu Z T, et al. Performance improvement scheme of multifocus image fusion derived by difference images[J]. *Signal Processing*, 2016, 128: 474-493.
- [9] Li H F, Wang Y T, Yang Z, et al. Discriminative dictionary learning-based multiple component decomposition for detail-preserving noisy image fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(4): 1082-1102.
- [10] Zhu Z Q, Yin H P, Chai Y, et al. A novel multi-modality image fusion method based on image



- decomposition and sparse representation[J]. *Information Sciences*, 2018, 432: 516-529.
- [11] Liu Y, Chen X, Wang Z F, et al. Deep learning for pixel-level image fusion: recent advances and future prospects[J]. *Information Fusion*, 2018, 42: 158-173.
- [12] Ma J Y, Yu W, Liang P W, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, 48: 11-26.
- [13] Jung H, Kim Y, Jang H, et al. Unsupervised deep image fusion with structure tensor representations[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3845-3858.
- [14] Yuan X C, Pun C M, Chen C L P. Robust Mel-frequency cepstral coefficients feature detection and dual-tree complex wavelet transform for digital audio watermarking[J]. *Information Sciences*, 2015, 298: 159-179.
- [15] da Cunha A L, Zhou J, Do M N. The nonsubsampling contourlet transform: theory, design, and applications[J]. *IEEE Transactions on Image Processing*, 2006, 15(10): 3089-3101.
- [16] Liu Y, Liu S P, Wang Z F. A general framework for image fusion based on multi-scale transform and sparse representation[J]. *Information Fusion*, 2015, 24: 147-164.
- [17] Zhang Y F, Yang M Y, Li N, et al. Analysis-synthesis dictionary pair learning and patch saliency measure for image fusion[J]. *Signal Processing*, 2020, 167: 107327.
- [18] Liu Y, Chen X, Peng H, et al. Multi-focus image fusion with a deep convolutional neural network[J]. *Information Fusion*, 2017, 36: 191-207.
- [19] Zhong J Y, Yang B, Li Y H, et al. Image fusion and super-resolution with convolutional neural network [M]//Tan T, Li X L, Chen X L, et al. *Pattern recognition. Communications in computer and information science*. Singapore: Springer, 2016, 663: 78-88.
- [20] Liu Y, Chen X, Ward R K, et al. Image fusion with convolutional sparse representation[J]. *IEEE Signal Processing Letters*, 2016, 23(12): 1882-1886.
- [21] Masi G, Cozzolino D, Verdoliva L, et al. Pansharpening by convolutional neural networks[J]. *Remote Sensing*, 2016, 8(7): 594.
- [22] Lahoud F, Süssstrunk S. Fast and efficient zero-learning image fusion[EB/OL]. (2019-05-09)[2021-05-05]. <https://arxiv.org/abs/1905.03590>.
- [23] Zhang Y, Liu Y, Sun P, et al. IFCNN: a general image fusion framework based on convolutional neural network[J]. *Information Fusion*, 2020, 54: 99-118.
- [24] Jung H, Kim Y, Jang H, et al. Unsupervised deep image fusion with structure tensor representations[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3845-3858.
- [25] Hou R C, Zhou D M, Nie R C, et al. Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model [J]. *Medical & Biological Engineering & Computing*, 2019, 57(4): 887-900.
- [26] Fan F D, Huang Y Y, Wang L, et al. A semantic-based medical image fusion approach[EB/OL]. (2019-06-01)[2021-05-05]. <https://arxiv.org/abs/1906.00225>.
- [27] Liang X C, Hu P Y, Zhang L G, et al. MCFNet: multi-layer concatenation fusion network for medical images fusion[J]. *IEEE Sensors Journal*, 2019, 19(16): 7107-7119.
- [28] Zhang B, Jiang C, Hu Y X, et al. Medical image fusion based a densely connected convolutional networks[C]//2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), March 12-14, 2021, Chongqing, China. New York: IEEE Press, 2021: 2164-2170.
- [29] Haghghat M, Razian M A. Fast-FMI: non-reference image fusion metric[C]//2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), October 15-17, 2014, Astana, Kazakhstan. New York: IEEE Press, 2014: 14916890.
- [30] Aslantas V, Bendes E. A new image quality metric for image fusion: the sum of the correlations of differences[J]. *AEU-International Journal of Electronics and Communications*, 2015, 69(12): 1890-1896.
- [31] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [32] 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.
- [33] Li S T, Kang X D, Hu J W. Image fusion with guided filtering[J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2864-2875.
- [34] Jian L H, Yang X M, Liu Z, et al. SEDRFuse: a symmetric encoder-decoder with residual block network for infrared and visible image fusion[J]. *IEEE Transactions on Instrumentation and*

- Measurement, 2021, 70: 1-15.
- [35] Prabhakar K R, Srikar V S, Babu R V. DeepFuse: a deep unsupervised approach for exposure fusion with extreme exposure image pairs[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4724-4732.
- [36] Liu H X, Zhu T H, Zhao J J. Infrared and visible image fusion based on region of interest detection and nonsubsampling contourlet transform[J]. Journal of Shanghai Jiaotong University (Science), 2013, 18(5): 526-534.
- [37] Yang B, Yang C, Huang G Y. Efficient image fusion with approximate sparse representation[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2016, 14(4): 1650024.