心脏T1 mapping图像分割中融合策略的自动搜索

王陈星宇^{1,3} 沈文枫¹ 高济英彰² 清野真也² 朱欣³

¹上海大学计算机工程与科学学院 上海 200444 ²福岛医科大学附属医院 日本福岛

³会津大学生物医学信息工程实验室 日本福岛

star_0912@i.shu.edu.cn

摘要:

多种医学图像的融合可以提高图像质量,便于特征提取为疾病诊断和治疗提供帮助。在心脏 磁共振(MR)图像T1 mapping中,多张MR图像被用来估计T1 mapping;因此,我们需要从图像中分 割得到心肌,以便对T1值进行进一步分析。传统上,每幅图像都各自进行分割,并且将多幅图像 的分割结果复杂地融合起来,可以得到更好的分割结果。传统的融合方法要求所有图像的分割结 果都满足一定的标准,因此经常由于图像质量低而失效。然而,设计一种用于整合多个图像的信 息的融合策略是非常困难的。在本研究中,我们设计了基于卷积神经网络的Segfuse网络,以获得即 使在某些图像质量低的情况下也能提高T1 mapping图像的分割性能的融合策略。在来自42名受试者 的112张T1 mapping图像验证集验证后,基于Segfuse的分割精度平均Dice为80.2%,与用于训练的传 统方法几乎一致。

关键词:心肌纤维化, T1 mapping, 融合策略, 图像分割, 图像融合

1.介绍

有时间序列的医学图像,如超声波图像、内 镜图像、MR图像等,能够为疾病的诊断和治疗提 供时间信息。也有通过序列图像来估计组织特征 的可能。例如,通过连续的MR图像计算得到的T1 mapping图像可以用来检查心肌纤维化的程度。通过 观察T1 mapping图像中的心肌区域,可以确定心肌纤 维化的程度和范围。图像分割技术可以帮助医生解 释图像,大大减轻医生的负担。

经典的图像分割方法包括阈值分割、区域增长 方法、分水岭算法。近年来,基于神经网络的图像 分割算法,如FCN、Unet和Segnet^[1-3],在医学图像处 理^[4]中得到了广泛的应用。对于序列图像,如MR图 像,需要多对一的分割方法综合来自多个图像^[6]的信 息。然而,通过一张图像获得一张分割结果^[6]是很常 见的。因此,我们需要在不改变分割模型本身的情 况下找到一种用于融合分割结果的融合策略。

图像融合在生物特征识别^[7,8]、医学图像^[5,9]、 遥感^[10]等领域,通过多图像之间的信息互补性增强成 像质量。不同模式的医学图像,如CT和MR,通常用 图像融合来结合其特征。Zong等人^[11]提出了一种基于 稀疏表示的新的医学图像融合方法。利用方向梯度 特征的直方图对图像斑块进行了分类,并学习了几 种子字典。他们采用L1范数和选择最大策略重建融 合图像。深度学习方法在红外和视觉图像融合或多 模态图像融合中得到了广泛的研究。Li等人^[12]提出了 一种基于残差网络(ResNet)和零相分量分析(ZCA)的图 像融合方法。他们使用ResNet从源图像中提取深度特 征,并使用ZCA和soft-max操作标准化深度特征,以 获得权重图。最后,他们利用加权平均策略重建了 具有源图像和权值图的融合图像。Haithem Hermessi 等人^[13]设计了一种基于卷积神经网络(CNN)的CT和 MR图像的融合方法。实验表明了CNN在多模态图像 融合中的潜力。由于图像融合方法可以用于融合不 同模式的图像,因此可以将分割结果用图像融合方 法进行融合。

本文研究了一种基于图像融合的自动化融合 策略获取方法。我们设计了基于卷积神经网络的 Segfuse网络来获取策略。为了证明融合策略是可学 习的,我们手动筛选了一个有效的策略,然后尝试 使用Segfuse来拟合它。我们利用单个T1 mapping图像 的分割作为基准参考,比较分割结果的平均Dice系 数,并通过Mann-Whitney-Wilcoxon测试(MWW测试) 来判断它们是否有显著差异。实验结果表明,该网 络可以有效地从训练集中学习融合策略。

2. 数据和方法

我们研究了从MR图像中分割左心室,用于分析 T1 mapping图像。我们使用FCN、Segnet和Unet作为基 础分割模型来测试融合策略。这些模型包括特征提 取网络的改进模型,如VGG和ResNet50。我们用来自 福岛医科大学附属医院的20名健康志愿者和209名患 者的图像数据训练了这些模型。这些图像是使用1.5 Tesla的MRI扫描仪(MRT200PP3TM,日本佳能医疗系 统公司)获得的。

这项研究已获得福岛医科大学医院机构审查委 员会的批准。所有来自20名健康志愿者和43名患者 的图像都由放射技术人员进行了标注。

分割模型的训练集由来自10名健康志愿者和11 名患者的450张图像组成。原始图像已选择感兴趣区 域,我们将其尺寸调整为256像素×256像素。 我们使用了Matlab提供的图像标签器工具来标记 标签。左心室和左心室腔部分被独立标注。

Segfuse网络基于Densefuse模型^[14]进行了改 进。Densefuse模型被设计用来融合RGB和红外图像, 我们用Segfuse来融合分割结果。图1显示了Segfuse网 络的结构。它的结构可以分为三个部分。在编码器 部分中,输入的二维图像通过几个卷积层被编码为 16个大小相等的特征图。每个卷积层都接受前两个 卷积层输出的特征图,因此编码器的深度可以很容 易地增加。在解码器部分中,32个特征图被二个卷 积层重建为二维图像输出。在训练阶段,融合层只 连接编码器和解码器。融合图像时,它将编码器输 出的两个输入图像的特征图相加,以达到融合的目 的。

在本文中,我们直接使用了通过Unet-ResNet50 分割模型得到的28组融合后的分割结果作为Segfuse的 训练集。这些标注是根据公式(1)融合得到。



图1 Segfuse网络的网络结构。(a)显示了Segfuse网络的编码 - 解码结构。(b)展示如何使用图像标签训练集来训练网 络。融合层只连接编码器和解码器。(c)显示如何融合图像1和图像2。融合层将图像1和图像2的特征图相加,解码器将 它们恢复到分割结果。

Segfuse网络的网络结构。(a)显示了Segfuse网络的编码-解码结构。(b)展示如何使用图像标签训练集 来训练网络。融合层只连接编码器和解码器。(c)显 示如何融合图像1和图像2。融合层将图像1和图像2 的特征图相加,解码器将它们恢复到分割结果。

T1 mapping是使用佳能提供的极性修正(PC)T1 准备工具^[15]获得的。由于该方法中的每个T1 mapping 图像通过5个输入的MR图像得到,因此每组图像将进 行5次融合。

$$\begin{cases}
M_{myocardium} = M_{ventricle} - M_{cavity} \\
M_{ventricle} = \bigcup_{k=1}^{n} M_{ventricle}^{k} \\
M_{cavity} = \bigcap_{k=1}^{n} M_{cavity}^{k}
\end{cases}$$
(1)

其中, k是用于计算T1 mapping的MR图像的数字 索引, M是这组MR图像的分割结果。

验证集中包括112组来自志愿者和患者的T1 mapping图像。福岛医科大学附属医院的医生和技术 人员已经确认了验证集中图像的质量。 图2显示了我们在本文中使用的分割过程。在 实验中,我们将基准参考和用不同的融合策略情 况下最终分割结果之间的Dice系数进行了比较。在 本文中,我们将来自公式(1)的策略定义为predesigned。由Segfuse得到的策略被称为fuse。baseline 是一个单一图像的分割结果,没有使用任何组合策 略。他们使用了来自相同的分割模型得到的分割结 果。我们分别计算了baseline, pre-designed 和fuse与 基准参考之间的Dice系数。这些结果被绘制在在T1 mapping图像上。我们采用MWW检验来确定它们在验 证集中的Dice系数是否存在统计学差异。实验结果见 第3节。

3. 实验结果

3.1 Segfuse network的效果

(1) Dice系数

表1表明,该融合策略几乎在大多数分割模型上都有效。除了FCN-32和FCN-32-VGG之外,Segfuse网络的融合策略在验证集上取得的效果与pre-designed的组合策略非常相似。

(2) 在验证集上的分割结果

图3展示了分割模型在单个图像上的分割结果, 其中包含了baseline策略。图4比较了pre-designed、 fuse的分割结果和T1 mapping上的基准参考。

	Segmentati	on Model				
	FCN-32	FCN-32-VGG	Segnet - VGG	Segnet - Resnet50	Unet - Resnet50	
Baseline	0.617	0.636	0.690	0.729	0.752	
Pre-designed	0.670	0.671	0.757	0.803	0.813	
Fuse	0.645	0.651	0.758	0.793	0.802	

表1 在验证集上的各种分割模型的平均Dice系数。



图2 实验中的分割过程。分割结果的融合部分是主要变量。



图3 由FCN-32、FCN-32-VGG、Segnet-VGG、Segnet-ResNet50和Unet-ResNet50取得的一些分割结果样例。



(a) FCN-32

(b) FCN-32-VGG





(d) Segnet-ResNet50 (e) Unet-ResNet50 图4 基准参考(红色线)、pre-designed 策略(蓝色线)与fuse 策略(黑色线)之间的比较

表2	在MWW测试中	,不同策略之间的Dice系数的p值。	加粗值小干0.05	,这意味着它们遵循不同的分	六 布
184				,这志怀自口们这咱们吗的。	,

	Segmentation Model				
	FCN-32	FCN-32-VGG	Segnet - VGG	Segnet - Resnet50	Unet-Resnet50
Baseline to Pre-designed	0.004	0.007	0.000	0.000	0.000
Baseline to Fuse	0.195	0.196	0.000	0.000	0.000
Pre-designed to Fuse	0.105	0.122	0.812	0.479	0.269

表3 在验证集中具有不同数量卷积层的不同Segfuse网络的平均Dice系数。加粗值是各个分割模型的最大值。

Amount of conv layers	Segmentation Model						
	FCN-32	FCN-32-VGG	Segnet-VGG	Segnet - Resnet50	Unet-Resnet50		
2	0.643	0.650	0.757	0.788	0.804		
3	0.645	0.651	0.758	0.793	0.802		
4	0.643	0.649	0.756	0.791	0.801		
5	0.642	0.650	0.756	0.788	0.798		
10	0.638	0.646	0.757	0.788	0.801		

3.2 该融合策略的有效性证明

表2列出了MWW测试的结果,证明了采用多图 像的融合与否的分割模型的分割效果是不同的,而 fuse 和pre-designed两种组合策略几乎是等效的。fuse 策略使用了具有3个卷积层的编码器。此外,基 于FCN-32和FCN-32-VGG的网络的p值比较大表 明,Segfuse不能很好地处理弱分割的融合。

3.3 编码器的深度对融合策略的影响

表3比较了我们研究的不同深度编码器的Segfuse 网络对各个分割模型的融合效果。具有3个卷积层的 编码器的Segfuse在大多数分割模型中表现良好。此 外值得注意的是,加大编码器深度并不是这个提高 融合效果的最佳选择。

4. 讨论

表2中FCN-32和FCN-32-VGG的检测结果和图3 表明,Dice系数取决于分割模型的性能和融合策略的 质量,如公式(2)所示。

 $Loss = l_{mod\ell} \implies Loss = l_{mod\ell} + l_{strategy}$ (2)

对于单个图像分割,分割模型的性能直接影 响分割效果。对于多图像分割,融合策略应该优化 模型,但也可能负优化模型。在这种情况下,对分 割模型和融合策略的改进都将得到很好的结果。在 过去,我们总是根据经验来选择这个策略。实验表 明,它可以从训练集中学习。这意味着在多输入分 割问题中,我们只需通过提供一组原始图像和想要 得到的图像输出,就可以得到一个很好的融合策 略。该方法还通过融合多输入所提供的信息提高了

分割效果。

表3表明,增加编码器的深度根本并不能提高学 习得到的融合策略的性能。原因可能是分割结果之 间的融合不涉及特定的图像信息。它只是在二进制 图像之间选择。层数过多的网络可能会使融合策略 包含太多的非必要因素。

5. 结论

本文设计了Segfuse网络用于获得融合多输入图 像分割的融合策略。MWW试验结果表明,baseline 策略与fuse存在统计差异,但pre-designed策略与fuse 之间无统计差异。pre-designed的分割结果和fuse的 分割结果在各种分割模型上是相似的。证明了通过 深度学习可以学习多输入图像的分割结果的融合策 略。这意味着我们可以通过一个简单的端到端训练 集得到一个根据需要融合多个图像的方法。这将大 大降低劳动力成本。

致谢

这项研究部分得到了会津大学的竞争研究 基金(2020-P-3)、中国国家重点研究和开发项目 (NO.2017YFB0701600),上海市科技委员会(STCSM) (NO.19511121002),以及由上海市经济与信息化委 员会主办的上海人工智能创新开发专项基金(2019-RGZN-01080)的支持。

参考文献:

[1] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[2] Ronneberger, Olaf, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." International Conference on Medical Image Computing and Computer - Assisted Intervention Springer International Publishing, 2015.

[3] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder - decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39.12 (2017): 2481 - 2495.

[4] Chang, Yakun, et al. "Automatic segmentation and cardiopathy classification in cardiac mri images based on deep neural networks." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018.

[5] Zhang, Nan, et al. "Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation." Computer Vision and Image Understanding 115.2 (2011): 256-269.

[6] Zhang, Zhenli, et al. " Exfuse: Enhancing feature fusion for semantic segmentation." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

[7] Abaza, Ayman, and Thirimachos Bourlai. "On ear - based human identification in the mid - wave infrared spectrum." Image and Vision Computing 31.9 (2013): 640 - 648.

[8] Nair, S. Anu H., P. Aruna, and M. Vadivukarassi. "PCA based image fusion of face and iris biometric features." International Journal on Advanced Computer Theory and Engineering (IJACTE) 1.2 (2013): 106-112.

[9] Xia, Kai-jian, Hong-sheng Yin, and Jiang-qiang Wang. "A novel improved deep convolutional neural network model for medical image fusion." Cluster Computing 22.1 (2019): 1515-1527.

[10] Zhang, Mengliang, et al. "Image fusion employing adaptive spectral-spatial gradient sparse regularization in UAV remote sensing." Signal Processing 170 (2020): 107434.

[11] Zong, Jing - jing, and Tian - shuang Qiu. "Medical image fusion based on sparse representation of classified image patches." Biomedical Signal Processing and Control 34 (2017): 195 - 205.

[12] Li, Hui, Xiao-jun Wu, and Tariq S. Durrani. "Infrared and visible image fusion with ResNet and zero-phase component analysis." Infrared Physics & Technology 102 (2019): 103039.

[13] Hermessi, Haithem, Olfa Mourali, and Ezzeddine Zagrouba. "Convolutional neural network - based multimodal image fusion via similarity learning in the shearlet domain." Neural Computing and Applications 30.7 (2018): 2029 - 2045.

[14] Li, Hui, and Xiao-Jun Wu. "Densefuse: A fusion approach to infrared and visible images." IEEE Transactions on Image Processing 28.5 (2018): 2614-2623.

[15] Kuhara, Shigehide, et al. "Polarity - corrected TI prep tool for delayed - enhancement MR imaging and T1 mapping." Proc. Intl. Soc. Mag. Reson. Med. Vol. 22. 2014.