

基于深度学习的跨模态医学图像转换

董国亚^{1,2}, 宋立明^{1,2,3}, 李雅芬³, 李文³, 谢耀钦³

1. 省部共建电工装备可靠性与智能化国家重点实验室(河北工业大学), 天津 300132; 2. 河北省电磁场与电器可靠性重点实验室(河北工业大学), 天津 300132; 3. 中国科学院深圳先进技术研究院, 广东 深圳 440305

【摘要】运用深度学习的方法基于脑部CT扫描图像合成相应的MRI。将28例患者进行颅脑CT和MRI扫描得到的CT和MRI的断层图像进行刚性配准,随机选取20例患者的图像输入U-Net卷积神经网络进行训练,利用训练好的网络对未参与训练的8例患者的CT图像进行预测,得到合成的MRI。研究结果显示:通过对合成的MRI进行定量分析,利用基于L2损失函数构建的U-Net网络合成MRI效果良好,平均绝对平均误差(MAE)为47.81,平均结构相似性指数(SSIM)为0.91。本研究表明可以利用深度学习对CT图像进行转换,获得合成MRI,现阶段可以达到扩充MRI医学图像数据库的目的,随着合成图像精度的提高,可以用于帮助诊断等临床应用。

【关键词】深度学习;CT;MRI;U-Net;卷积神经网络;图像模态转换;合成MRI

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2020)10-1335-05

Cross-modality medical image synthesis based on deep learning

DONG Guoya^{1,2}, SONG Liming^{1,2,3}, LI Yafen³, LI Wen³, XIE Yaoqin³

1. State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300132, China; 2. Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability of Hebei Province, Hebei University of Technology, Tianjin 300132, China; 3. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 440305, China

Abstract: The purpose of this research is to synthesize the corresponding magnetic resonance imaging (MRI) based on brain computed tomography (CT) images by deep learning. The tomographic images of CT and MRI obtained by brain CT and MRI scanning are rigidly registered in 28 patients, and the images of 20 patients are randomly input into U-Net convolutional neural network for training. The CT images of 8 patients who do not participate in the training are predicted by the trained network, thereby obtaining the synthetic MRI. The results reveal that through the quantitative analysis on synthetic MRI, the U-Net network constructed based on L2 loss function has a good performance in synthesizing MRI, with a mean absolute error of 47.81 and an average structural similarity index of 0.91. This study shows that deep learning method can be used to obtain synthetic MRI by converting CT images, thus achieving the purpose of expanding MRI medical database. With the improvement of the accuracy of image synthesis, it can be used in diagnosis and other clinical applications.

Keywords: deep learning; computed tomography; magnetic resonance imaging; U-Net; convolutional neural network; cross-modality image synthesis; synthetic magnetic resonance imaging

前言

【收稿日期】2020-05-12

【基金项目】深圳市配套项目(GJHS20170314155751703);国家重点研发计划(2016YFC0105102);国家自然科学基金(61871374);广东省特支计划领军人才(2016TX03R139);深圳市基础研究计划(JCYJ20170413162458312);广东省自然科学基金(2017B20170413162458312, 2015B020233011, 2014A03-0312006)

【作者简介】董国亚,博士,副教授,主要研究方向:医学信号和图像处理、神经工程, E-mail: dong_guoya@126.com

【通信作者】谢耀钦,博士,研究员, E-mail: yq.xie@sia.ac.cn

在医学辅助诊断中,磁共振成像(Magnetic Resonance Imaging, MRI)由于对软组织优秀的成像能力以及无创的特性而占有重要地位,但是MRI扫描也存在不可避免的缺点。首先,MRI扫描比CT花费更多时间,一次完整的磁共振扫描通常需要20 min以上;其次,MRI在长时间狭小环境扫描的同时伴随巨大噪音,患者需佩戴专用耳罩配合检查,这使患有幽闭恐惧症的患者和小孩等很难完成检查;此外,由于MRI机器产生的强磁场伴随磁场效应,机房内禁止放置金属监护和抢救设备,急诊和危重患者以及体内含有普通心脏起搏器、金属冠脉支架、人工关节等金属植入物的患者

无法进行该检查;最后,在机器价格及治疗成本方面MRI扫描比CT高出许多,限制了MRI的普及和应用。这也造成了CT图像与MRI数量的差距,MRI的医学图像数据集相对较少。

近年来,计算机技术和大规模数据集处理技术的迅速发展使得深度学习不再停留在理论阶段,而是越来越多地应用于图像识别、自然语言处理等领域。基于深度学习的算法与传统机器学习算法不同,这种算法能够从单一模态医学图像中自动学习有用的特征^[1],具有图像处理和分类的优势,并且可以将学习到的特征运用到多模态图像处理中。模态是指事物发生或存在的方式,多模态是指两个或者两个以上模态的各种形式的组合^[2]。多模态机器学习指通过建立模型使机器从多个模态中学习各个模态的信息,从而实现各个模态信息的交流和转换。Baltrušaitis等^[3]总结了多模态机器学习的研究情况,提出围绕多模态机器学习发展中的5个挑战:模态表示、模态转换、模态对齐、模态融合和合作学习。深度学习网络中用于图像处理的卷积神经网络(Convolutional Neural Network, CNN),已成为医学影像分析的常用方法^[4]。CNN能够处理多维多通道数据,捕获输入和输出之间复杂的非线性映射^[5],网络通常包含卷积层、池化层和全连接层^[6]。Long等^[7]首次提出了一种端到端(end-to-end)用于语义分割的全卷积神经网络(Fully Convolutional Networks, FCN)。FCN将全连接层替换为卷积层,不再包含全连接层,可适应任意尺寸输入。Ronneberger等^[8]根据FCN设计的U-Net网络采用了“拼接”这一种完全不同的特征融合方式,将特征在不同维度拼接在一起,形成更厚的特征。在利用卷积网络对医学图像处理的使用中,Ben-Cohen等^[9]使用FCN与生成对抗网络(Generative Adversarial Network, GAN)^[10]网络相结合的方式通过深度学习将CT图像进行模态转换得到PET图像,优点在于主要针对没有人工标记数据的恶性病变。Nie等^[11]使用3D FCN将骨盆MRI合成CT图像,优点在于直接对三维图像进行处理。同样是运用深度学习将MRI合成CT图像。Han^[12]使用自己设计的深度卷积网络(Deep Convolutional Neural Network, DCNN)在脑肿瘤患者图像上测试,并与基于Atlas的方法进行比较,结果显示DCNN可以在不牺牲测试速度的情况下轻松处理大量训练数据。在Jin等^[13]研究中,第一次试图运用深度学习的方法将CT图像转化成MRI。通过GAN网络,将已配准数据与未配准数据结合训练数据的一致性损失,同时使用基于配对数据的像素损失来合成MRI。本研究运用U-Net网络对20例患者进行颅脑CT和MRI扫描得到的医学图像进行学习,运用训练好的网络将CT

图像合成MRI。本次研究针对MRI扫描的局限性以及MRI数据集较少的情况,通过深度学习的方法基于CT图像合成MRI,达到扩充医学图像数据集的目的并借此探讨不同模态间进行相互转化的可能性。

1 方法

本文实验的流程图如图1所示,包含以下两个模块:训练模块(Training)和测试模块(Testing)。在训练模块中,首先将采集到的CT和MRI分别传入Dataset模块进行数据预处理,将处理好的数据传入卷积神经网络(U-Net)进行学习,得到训练好的网络(Trained U-Net)。在测试模块中,分别选取不同的CT图像输入训练好的网络,并对输出数据进行一定处理得到合成的MRI。

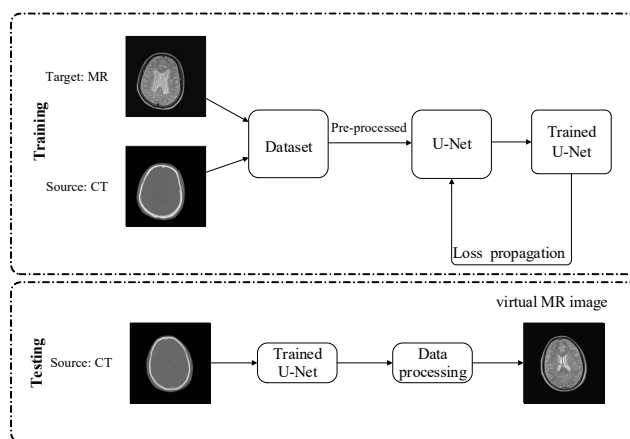


图1 合成MRI流程图

Fig.1 Flowchart of synthesizing MRI

1.1 数据

采集28例患者CT和MRI头部数据,使用SOMATOM Definition Flash CT机进行CT扫描,扫描参数为120 kV、330 mA,图像分辨率 512×512 ,层厚1 mm。使用西门子MAGNETOM Avanto 1.5T磁共振系统进行MRI扫描,图像分辨率为 256×256 ,层厚为1 mm。

1.2 预处理

将采集得到的28例患者的CT图像和MRI进行刚性配准^[14],对CT图像分辨率进行降采样,使得CT图像和MRI保持相同分辨率为 256×256 。将数据分为训练集与测试集,训练集中包含随机选取20例患者数据,剩余8例患者数据作为测试集。每个患者CT或MRI平均包含超过100个二维轴向图像切片。对训练数据进行预处理,基于式(1)将CT和MRI像素值,重新采样至在256灰度范围内均匀分布。

$$\text{Image} = \left(\frac{I - \min(I)}{\max(I) - \min(I)} \right) \times 255 \quad (1)$$

1.3 网络结构

U-Net的设计结构如图2所示。网络总共包含9部分,其中前5部分用于特征提取,后4部分设计为上采样。为了增加模型的非线性,提取数据更高维度的特征,在特征提取和上采样过程中,每一部分U-Net结构中卷积设计均采用两个卷积计算连用的形式。最后一部分设计略有不同,增加了一步卷

积核长度为1卷积计算。具体卷积计算设计如下:在特征提取部分采用3×3卷积核,卷积步长为1,池化层步长为2进行特征提取;上采样部分为了恢复图像尺寸采用2×2卷积核,卷积步长为1进行逆卷积计算。第一部分是两个卷积连用提取特征,其他部分是步长为2的池化层后面连续进行两个卷积计算。

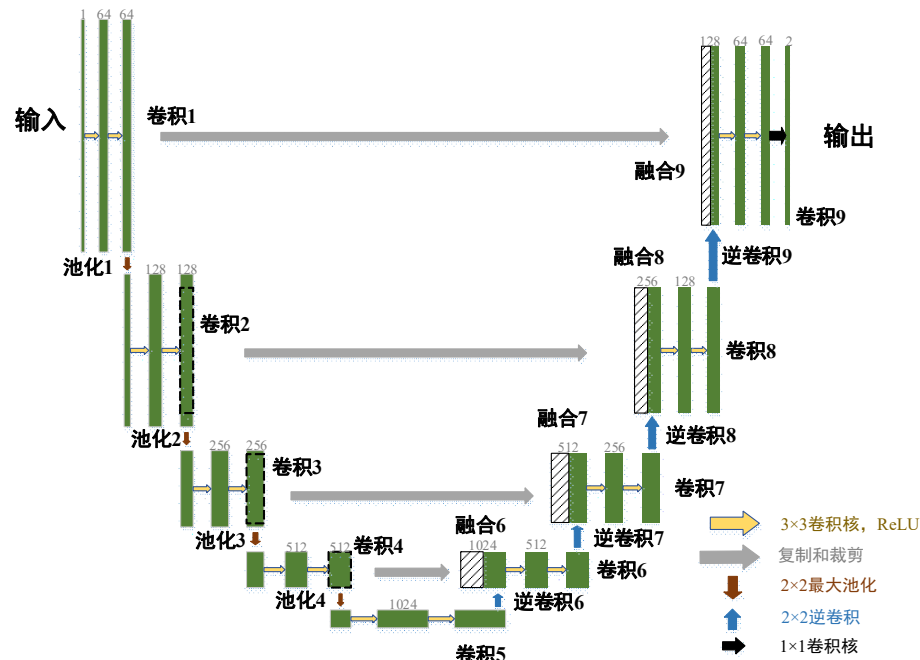


图2 U-Net结构图
Fig.2 U-Net architecture

网络均基于PyTorch^[15]库设计,采用内存为12 GB的NVIDIA TITAN X进行训练。在U-Net训练过程中使用采用均方损失(MSE Loss)函数验证准确率,使用ADAM梯度下降算法^[16],随着训练周期降低学习率。为了加快训练速度,通过4进程来导入数据。训练批处理参数设置为16,测试批处理参数设置为1。

1.4 结果评估

为了避免预测图像产生误差相互抵消,准确显示实际预测误差的大小,选择绝对平均误差(Mean Absolute Error, MAE)作为定量评价指标。结构相似性指数(Structural Similarity Index, SSIM)^[17]将结构信息定义为独立于亮度、对比度的,反映场景中物体结构的属性,并将失真建模为亮度、对比度和结构3个不同因素的组合。通过用均值作为亮度的估计,标准差作为对比度的估计,协方差作为结构相似程度的度量来衡量两幅图像的相似度^[18-19]。本实验选取SSIM评估预测图像与真实图像的结构相似性。MAE与SSIM定义如式(2)、式(3)所示:

MAE = \frac{1}{N} \sum_{i=0}^{N-1} |I_{MR}(i) - I_{sMR}(i)| \tag{2}

SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3}

其中,N为除去像素值为零的背景像素点以外的像素点数量,I_{MR}(i)和I_{sMR}(i)分别为标准MRI和合成MRI中在i处的像素值。μ_x是x(标准MRI)的平均值,μ_y是y(合成MRI)的平均值,σ_x²是x的方差,σ_y²是y的方差,σ_{xy}是x和y的协方差。c₁=(k₁L)²,c₂=(k₂L)²是用来维持稳定的常数。L是像素值的动态范围。k₁=0.01,k₂=0.03^[15]。

2 结果

为了评估从CT图像合成MRI的质量,使用不同的图像质量指标将合成的MRI与标准MRI进行比较,对比较结果分别进行成像误差分析和像素误差分析。

2.1 定性结果

对测试结果与标准图像成像分析如图3所示,从左至右为输入的标准CT图像、标准MRI、合成MRI和标准MRI与合成MRI之间的差值。将合成MRI与标准CT图像对比,观察颅内软组织区域可以看出所设计的U-Net网络能够做到在CT图像中区分具有相似像素

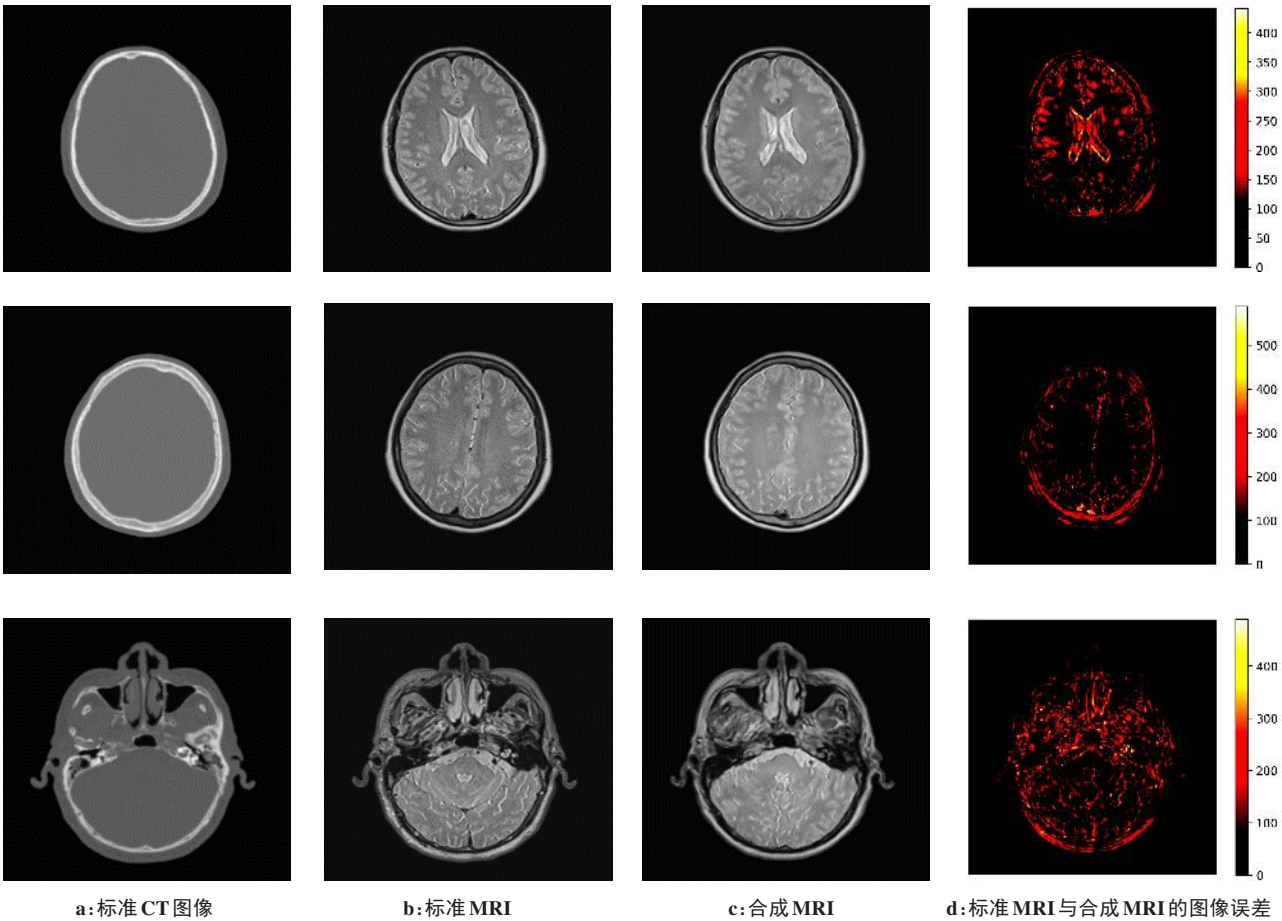


图 3 合成图像与标准图像的差异
Fig.3 Differences between synthetic images and standard images

强度的不同解剖结构,如骨骼、脑回和脑部软组织。

2.2 定量结果

对测试结果进行定量分析,测得 8 例测试数据标准 MRI 与合成 MRI 的 MAE 和 SSIM,分别计算平均值和标准差并记录每一例数据中单张图像的最大值和最小值,结果如表 1 所示。

从表 1 可以看出,全部 8 例测试集中单张图像最小 MAE 为 47.81,最大 SSIM 为 0.95,全部测试数据平均 MAE 为 73.68,平均 SSIM 为 0.91。使用 U-Net 网络合成 MRI 能够达到和标准图像 91% 左右相似度。

3 讨论

本研究通过总结在医学图像模态转换方面部分已经完成的研究工作^[20-21],提出运用 U-Net 网络将标准颅脑 CT 图像合成 MRI 的想法,并通过实验结果验证。对实验结果进行成像误差分析,结果显示标准 MRI 与合成 MRI 间在骨结构区域差异较大,在软组织区域差异较小。综合考虑成像结果与 CT、MRI 各自成像原理,造成的差异可能是由于 CT 图像提供了更多关于骨结构的细节来弥补合成 MRI 的不足,而标准 MRI 主要针对的是软组织成像。

将合成 MRI 与标准 MRI 逐像素点进行统计学分析,得到误差分析结果表明合成的 MRI 与标准图像平均可以达到 91% 的相似度。具体到单个测试集的合成 MRI 与标准 MRI 分析,发现在软组织占比例高的图像如扫描中段得到的断层图像中 MAE 值较低而 SSIM 值则较大,最小为 47.81,最大为 0.95。MAE 越小则说明两幅图像的实际预测误差越小。SSIM 的取值范围为 0 到 1,数值越接近于 1 代表两幅图像越相似。基于 U-Net 图像模态转换对软组织的转换效果较好,对骨组织的转换精度有待进一步提高。

在医疗领域基于机器学习的医疗辅助研究已经取得初步成果,其中多数研究成果是针对医学图像进行的配准或是分割,二者均需要大量图像数据训练网络,而且应用于图像处理方面的深度学习模型大多基于自然图像,针对临床应用的医学图像模型特别是医学肿瘤图像模型相对较少。医学图像与自然图像之间存在十分明显的不同,医学图像多为灰度图像,上下层图像之间关系紧密且连续性强。自然图像多为彩色图像,数据之间关联性不强。相比于自然图像,医学图像数据量小且涉及患者隐私,加之不同医疗中心之间数据不通,图像数据更加难以获得。针对数据不足的情况,

表1 合成MRI与标准MRI间的MAE和SSIM

Tab.1 Mean absolute error and structural similarity index between synthetic MRI and standard MRI

病例	MAE			SSIM		
	均值(标准差)	最小值	最大值	均值(标准差)	最小值	最大值
1	81.06(8.70)	57.44	97.42	0.91(0.02)	0.87	0.93
2	58.16(4.60)	47.81	65.01	0.92(0.02)	0.89	0.94
3	96.88(7.87)	80.86	110.16	0.87(0.03)	0.75	0.91
4	67.12(13.58)	52.64	97.94	0.92(0.05)	0.79	0.95
5	74.77(14.98)	58.57	122.08	0.91(0.06)	0.74	0.95
6	75.84(4.54)	66.91	83.44	0.91(0.02)	0.87	0.94
7	70.41(8.38)	60.68	96.14	0.92(0.03)	0.82	0.95
8	65.18(8.36)	55.35	88.87	0.92(0.05)	0.75	0.95
总计	73.68(10.99)	47.81	122.08	0.91(0.02)	0.74	0.95

计算机可以通过深度学习合成与输入图像相匹配的输出图像,缓解图像数量较少的问题。使得医学图像合成具有广大应用前景。但是不管是图像配准、图像分割还是图像的生成,研究结果大多停留在理论阶段,距离投入临床使用尚缺少大量数据对模型进行训练以及多中心数据对模型进行检验。

本次研究显示利用深度学习方法基于CT图像的合成MRI,用于扩充医学图像数据以及不同模态的医学图像间进行相互转化具有可行性。这种处理图像的方法还可以进行扩展,对MRI、CT和PET等多模态间进行相互转换。后续研究将会从此方面入手,在此基础上探究不同网络、不同算法对不同模态医学图像生成的影响。

致谢:中科院人机智能协同系统重点实验室和深圳市介入式诊疗一体化关键技术工程实验室。

【参考文献】

[1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[2] 刘建伟, 丁熙浩, 罗雄麟. 多模态深度学习综述[J]. 计算机应用研究, 2020, 37(6): 1601-1614.

LIU J W, DING X H, LUO X L. Survey of multimodal deep learning [J]. Application Research of Computers, 2020, 37(6): 1601-1614.

[3] BALTRUŠAITIS T, AHUJA C, MORENCY L P. Multimodal machine learning: a survey and taxonomy[J]. IEEE Trans Pattern Anal, 2017, 41(2): 423-443.

[4] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Med Image Anal, 2017, 42(9): 60-88.

[5] LU F, WU F, HU P, et al. Automatic 3D liver location and segmentation via convolutional neural network and graph cut[J]. Int J Comput Ass Rad, 2017, 12(2): 171-182.

[6] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks[J]. Pattern Recogn, 2018, 77(C): 354-377.

[7] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal, 2015, 39(4): 640-651.

[8] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional

networks for biomedical image segmentation [C]//International Conference on Medical image computing and computer-assisted intervention. Springer, 2015: 234-241.

[9] BEN-COHEN A, KLANG E, RASKIN S P, et al. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection[J]. Eng Appl Artif Intel, 2019, 78: 186-194.

[10] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Advances in Neural Information Processing Systems. 2014: 2672-2680.

[11] NIE D, CAO X, GAO Y, et al. Estimating CT image from MRI data using 3D fully convolutional networks[M]// Deep Learning and Data Labeling for Medical Applications. Springer, 2016: 170-178.

[12] HAN X. MR-based synthetic CT generation using a deep convolutional neural network method[J]. Med Phys, 2017, 44(4): 1408-1419.

[13] JIN C B, KIM H, LIU M, et al. Deep CT to MR synthesis using paired and unpaired data[J]. Sensors, 2019, 19(10): 2361.

[14] KLEIN S, STARING M, MURPHY K, et al. Elastix: a toolbox for intensity-based medical image registration [J]. IEEE Trans Med Imaging, 2009, 29: 196-205.

[15] PASZKE A, GROSS S, MASSA F, et al. Pytorch: an imperative style, high-performance deep learning library [C]//Advances in neural information processing systems. 2019: 8026-8037.

[16] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. [2017-1-30]. <https://arxiv.org/abs/1412.6980>.

[17] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition. IEEE, 2010: 2366-2369.

[18] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment [C]//The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers. IEEE, 2003, 2: 1398-1402.

[19] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Trans Image Process, 2004, 13(4): 600-612.

[20] DU J, LI W, LU K, et al. An overview of multi-modal medical image fusion[J]. Neurocomputing, 2016, 215(26): 3-20.

[21] 周沛, 陈后金, 于泽宽, 等. 跨模态医学图像预测综述[J]. 电子学报, 2019, 47(1): 220-226.

ZHOU P, CHEN H J, YU Z K, et al. Review of cross-modality medical image prediction[J]. Acta Electronica Sinica, 2019, 47(1): 220-226.

(编辑:陈丽霞)