

深度学习在医学影像中的研究进展及发展趋势

王丽会^{1,2}, 秦永彬^{1,2}

1. 贵州省智能医学影像分析与精准诊断重点实验室, 贵州 贵阳 550025;
2. 贵州大学计算机科学与技术学院, 贵州 贵阳 550025

摘要

医学影像是临床诊断的重要辅助工具, 医学影像数据占临床数据的90%, 因此, 充分挖掘医学影像信息将对临床智能诊断、智能决策以及预后起到重要的作用。随着深度学习的出现, 利用深度神经网络分析医学影像已成为目前研究的主流。根据医学影像分析的流程, 从医学影像数据的产生、医学影像的预处理, 到医学影像的分类预测, 充分阐述了深度学习在每一环节的应用研究现状, 并根据其面临的问题, 对未来的发展趋势进行了展望。

关键词

深度学习; 医学影像; 图像处理; 人工智能; 卷积神经网络

中图分类号: TP183

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2020056

State of the art and future perspectives of the applications of deep learning in the medical image analysis

WANG Lihui^{1,2}, QIN Yongbin^{1,2}

1. Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, Guiyang 550025, China
2. School of Computer Science and Technology, Guizhou University, Guiyang 550025, China

Abstract

Medical imaging is an important auxiliary tool for clinical diagnosis. Medical images occupy almost 90% of clinical data. Therefore, mining medical image information will be beneficial for intelligent diagnosis, decision-making and prognosis prediction. With the emergence of deep learning, using deep neural networks to analyze medical images has become a hot research topic. Following the process of medical image analysis, from the image acquisition, the image pre-processing to the classification and prediction, the state-of-the-art applications of deep learning in each step of medical image analysis was elaborated, and according to the existed issues and the challenges, the future perspectives were finally discussed.

Key words

deep learning, medical imaging, image processing, artificial intelligence, convolutional neural network

1 引言

医学成像已成为临床诊断的重要辅助手段,其包括计算机断层扫描(computed tomography, CT)成像、磁共振成像(magnetic resonance imaging, MRI)、正电子发射断层扫描(positron emission tomography, PET)成像、超声(ultrasound, US)成像、X射线(X-ray)成像等。如何借助大数据和人工智能技术,深入挖掘海量的医学图像信息,实现基于影像数据的智能诊断、智能临床决策以及治疗预后,已成为目前的研究热点。

深度学习属于机器学习的分支,是目前实现人工智能技术的重要手段。随着深度学习技术在图像处理和计算机视觉领域的广泛应用,利用深度学习技术辅助临床诊断和决策已成为医学图像分析领域的研究重点。医学影像智能诊断的流程可大致分为3个步骤,首先获取大量高质量的图像数据,然后对图像进行预处理,最后挖掘图像信息,进行分析预测。其具体环节如图1所示。其中海量、高质量的图像数据是深度学

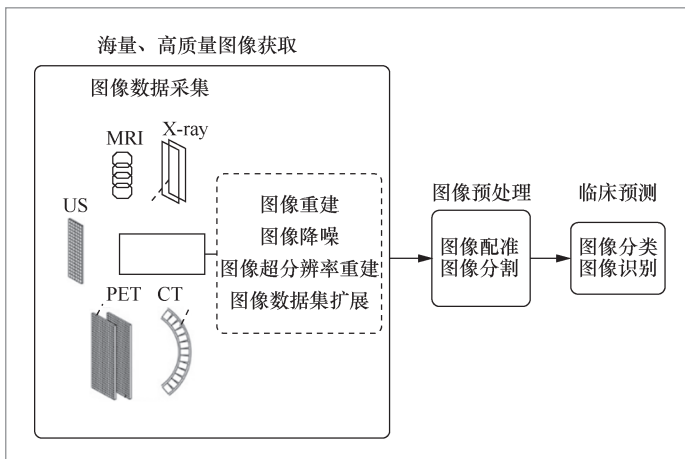


图1 医学图像处理分析过程

习训练的基础,图像预处理(如配准、感兴趣区域提取)是后续分析准确度的基本保障,挖掘信息、建立预测模型是临床智能决策的关键。因此,本文将分别围绕这3个方面,阐述深度学习在医学图像处理分析流程中每个环节的主要应用现状,最后总结深度学习在医学影像研究中的发展趋势。

2 医学图像复原、重建与合成

2.1 医学图像复原与重建

海量、高质量的医学图像数据是利用深度学习技术实现影像精准诊断的基础。然而,由于成像设备和采集时间等因素的限制,在医学成像的过程中不可避免地会受到噪声、伪影等因素的影响。同时,针对某些成像方式,需要在成像分辨率和采集时间上进行折中,例如在CT成像中,为了降低辐射的影响,需要减少投影采集数目;在磁共振成像中,为了减少患者运动或者器官自身运动引起的伪影,需要降低K空间的采样率以减少采集时间,然而低采样率会严重影响图像的重建质量。为了获得高质量的采集图像,经常需要进行图像降噪、图像超分辨率重建、图像去伪影等复原与重建工作。下面将分别阐述深度学习在这几方面的研究现状。

2.1.1 医学图像降噪

基于深度学习的医学图像降噪主要应用在低剂量CT图像中。卷积降噪自动编码器(convolutional neural network-denoise auto-encoder, CNN-DAE)是早期用于医学图像降噪的深度学习模型^[1]。该模型通过一些堆叠的卷积层,以编码和解码的方式从噪声图像中学习无噪图像,其鲁棒性较差,对噪声类型变化较为敏感。随后,Chen H等人^[2]提出RED-CNN

降噪模型,将残差网络与卷积自动编码器相结合,通过跳跃连接形成深度网络,实现低剂量CT图像的降噪。同年,Kang E等人^[3]首先对低剂量CT图像进行方向小波变换,然后将深度卷积神经网络模型应用于小波系数图像,实现降噪,并使用残差学习架构加快网络训练速度,提高性能。

虽然这些网络结构的降噪性能相较于传统方法得到了显著的提升,但是其网络训练均以复原CT图像与相应正常剂量CT图像之间的均方误差最小为优化目标,使得降噪图像存在细节模糊和纹理缺失等问题。为了解决这一问题,研究者提出改进损失函数和模型结构的方法来优化低剂量CT图像的降噪效果。WGAN-VGG模型通过引入感知损失,采用WGAN(Wasserstein generative adversarial network)模型进行降噪,利用Wasserstein距离和感知损失提高降噪图像与真实图像的相似性^[4]。基于WGAN-GP(gradient penalty)的SMGAN(structurally-sensitive multi-scale generative adversarial net)模型将多尺度结构损失和 L_1 范数损失结合到目标函数中,并利用相邻切片之间的信息降噪,其结果优于WGAN-VGG模型^[5]。但是梯度惩罚的使用削弱了生成式对抗网络(generative adversarial network, GAN)的表示能力。为了解决这个问题,Ma Y J等人^[6]提出基于最小二乘生成对抗网络(least-square GAN, LS-GAN)的残差生成器结构,通过引入结构相似度和 L_1 范数损失来提高降噪能力,生成器负责学习噪声,降噪图像为生成器的网络输入与网络输出的相减结果。除了生成模型,为了提高降噪效果,Yin X R等人^[7]同时在投影域和图像域采用3D残差网络进行降噪,并利用滤波反投影重建算法,实现投影域和图像域的相互转化,通过迭代的思想实现图像降噪。Wu D F等人^[8]提出一致性神经网络模型,实现了无监督的图像

降噪方法,其不需要无噪图像标签,仅利用有噪图像对模型进行训练,从而获得降噪图像。

可以看出,在利用深度学习进行降噪时,常需要利用有噪图像和无噪图像来训练模型,学习噪声类型,或者学习无噪图像与有噪图像之间的对应关系,进而实现图像降噪。这种方式具有一定的局限性,在临床的某些应用上,很难获得真实的无噪图像。因此,如何采用无监督或者自监督模型,仅利用有噪图像实现医学图像降噪将是未来研究的主要方向。

2.1.2 医学图像超分辨率重建

高分辨率的医学图像可以提供更多的临床诊断细节,然而由于采集设备的限制,临床上高分辨率图像较难获取。因此,如何利用深度学习技术从一幅或者多幅低分辨率医学图像中获得高分辨率图像成为当前主要研究热点之一。随着深度学习模型在自然图像超分辨率重建中的成功应用,采用深度学习模型进行医学图像超分辨率重建的研究逐渐开展起来。然而,医学图像与自然图像有本质的区别,其超分辨率重建不仅需要在图像切片平面上进行,还需要在切片之间进行,如图2所示。

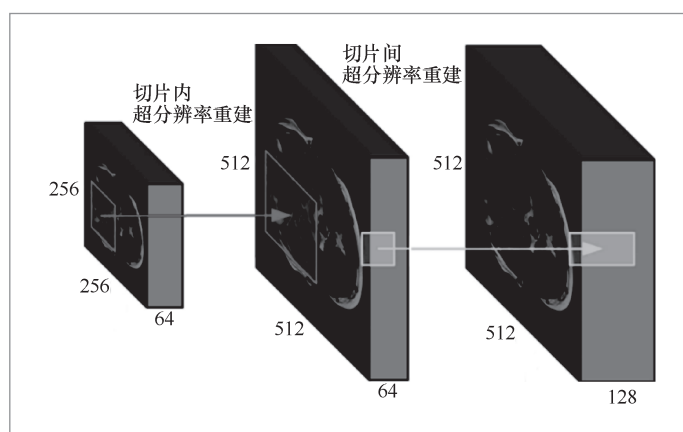


图2 医学图像超分辨率图像示意图(此图部分来自参考文献[9])

除了将自然图像中的超分辨率重建模型直接应用到医学图像, Oktay O等人^[10]采用深度残差卷积网络从多个2D心脏磁共振(magnetic resonance, MR)图像中重建出3D高分辨率MR图像, 提高了层间分辨率。Pham C H等人^[11]将SRCNN模型拓展到3D, 以实现脑部MR图像的超分辨率重建。McDonagh S等人^[12]提出对上下文敏感的残差网络结构, 可以得到边界和纹理清晰的高分辨率MR图像。Zheng Y等人^[13]提出多个Dense模块和多路分支组合的MR高分辨重建模型, 该模型具有较好的重建结果和泛化能力。Zhao X L等人^[14]提出通道可分离的脑部MR图像高分辨重建模型, 一个通道采用残差结构, 一个通道采用密集连接结构, 实现了特征的有效利用, 从而提高高分辨率图像的重建质量。Tanno R等人^[15]结合3DSubpixelCNN和变分推论实现了磁共振扩散张量图像的超分辨率重建。Peng C等人^[16]提出空间感知插值网络(spatially aware interpolation network, SAINT), 充分利用不同切面的空间信息提高超分辨率图像的重建质量, 该模型在对CT图像进行2倍、4倍和6倍分辨率重建时, 均取得了较好的结果。Shi J等人^[17]提出一种多尺度全局和局部相结合的残网络(multi-scale global local residual learning, MGLRL)模型, 实现了MR图像的超分辨重建, 该模型可以增强图像重建细节。Lyu Q等人^[18]采用GAN实现了多对比度MR图像的超分辨率重建。

与医学图像降噪相似, 基于深度学习的超分辨率图像重建需要低分辨率图像样本和高分辨率图像样本对对网络进行训练。通常采用下采样的方式进行高/低分辨率图像样本对的构造。然而针对不同模态的医学成像, 其成像原理大不相同, 高分辨率和低分辨率之间的对应关系也不尽相同。因此, 采用人工下采样的方式获得训

练数据, 学习低分辨率图像与高分辨率图像的对应关系, 很可能与实际采集中低分辨率图像与高分辨率图像的对应关系不相符, 进而导致重建的高分辨图像无意义, 因此如何构建符合实际的高/低分辨率图像样本对是利用深度学习进行超分辨重建的难点。

2.1.3 医学图像重建

医学图像重建是指将采集的原始数据重建为临床上可视图像的过程, 如CT采集的原始数据为投影图像, MR采集的原始数据为K空间数据, 需要重建算法才能获得临床上用于诊断的图像。在实际应用中, 由于一些采集条件的限制(如在CT中尽量减少投影数目, 缩短采集时间, 以降低辐射影响; 在MR成像中, 减少K空间填充数目, 缩短采集时间, 以避免患者的不适或者由患者运动带来的图像伪影), 需要降低原始数据的采集率。然而, 降低原始数据的采集率必然会影响图像的重建质量。因此, 研究合适的重建算法, 保证在原始数据低采样率下仍能获得高质量的重建图像, 成为医学图像重建中的研究重点。

目前采用深度学习模型进行医学图像重建的方法主要分为两类: 一类是从原始数据直接到图像的重建, 另一类是基于后处理的方式提高重建图像的质量。第一类方法的代表模型有: ADMM-Net^[19], 其用深度迭代的方式学习传统交替方向乘子(alternating direction method of multipliers, ADMM)优化算法中的超参数, 可以直接从欠采样的K空间数据中重构出MR图像; Adler J等人^[20]提出对偶学习模型, 用其代替CT重建中的滤波反投影方法, 实现了投影数据到CT图像的准确重建; Cheng J等人^[21]在此基础上提出原始-对偶网络(primal-dual network, PD-Net), 实现了MR图像的快速重建;

Zhang H M等人^[22]提出JSR-Net (joint spatial-Radon domain reconstruction net), 利用深度卷积神经网络模型, 同时重建CT图像及其对应的Radon投影变换图像, 得到了比PD-Net更好的重建结果。第二类方法是目前主要的重建方式, 即采用图像去伪影的后处理模型进行重建。用于图像降噪、超分辨率重建的模型都可以用于该类型的图像重建, 如Lee D等人^[23]提出带有残差模块的U-Net模型结构来学习重建图像与原始欠采样图像之间的伪影; 随后, 他们又提出利用双路U-Net模型对相位图像和幅度图像进行重建, 进而提高了MR图像的重建质量^[24]; Schlemper J等人^[25]采用深度级联的卷积神经网络 (convolutional neural network, CNN) 模型, 学习动态MR图像采集的时序关系, 进而在快速采集下提高动态MR图像的重建质量; Han Y等人^[26]采用域适应微调方法, 将CT图像重建的网络应用到MR图像重建上, 可以实现高采样率下的准确重建; Eo T等人^[27]提出KIKI-Net, 同时在K空间和图像空间域上使用深度学习网络进行重建, 提高了MR图像重建的性能; Bao L J等人^[28]采用一个增强递归残差网络, 结合残差块和密集块的连接, 用复数图像进行训练, 得到了较好的MR图像重建结果; Dai Y X等人^[29]基于多尺度空洞卷积设计深度残差卷积网络, 以较少的网络参数提高了MR图像的重建精度; 受到GAN在视觉领域成功应用的启发, Yang G等人^[30]提出一种深度去混叠生成对抗网络 (DAGAN), 以消除MRI重建过程中的混叠伪影; Quan T M等人^[31]提出一种具有周期性损失的RefinGAN模型, 以极低的采样率提高了MR图像的重建精度; Mardani M等人^[32]基于LS-GAN损失, 采用ResNet的生成器和鉴别器来重建MR图像, 获得了较好的可视化结果。

图像降噪、图像超分辨率重建、图像重建等均属于反问题求解。因此, 其模型可互相通用, 本文不对其进行一一阐述。

2.2 医学图像合成

2.2.1 医学图像数据扩展

目前, 临床上医学图像合成主要有两个目的。其一, 扩展数据集, 以获得大量医学影像样本来训练深度学习模型, 从而提高临床诊断和预测的准确度。尽管已有很多数据扩展方法, 如平移、旋转、剪切、加噪声等, 但是其数据扩展方式无法满足数据多样性的需求, 在提升深度学习模型的预测精度以及泛化能力上仍有待提高。其二, 模拟成像。由于不同模态的医学图像可以提供不同的信息, 融合不同模态的医学影像信息可以提高临床诊断精度。然而同一个病人的多模态影像信息很难获取, 此时图像合成便提供了一种有效的手段。此外, 某些新兴的成像技术对成像设备具有较高的要求, 仅少数的医院及科研机构可以满足要求, 因此图像合成为获取稀缺的影像数据提供了可能。

随着GAN模型在自然图像合成上的成功应用, 应用GAN的衍生模型进行医学图像合成已成为近几年的研究热点。在医学图像数据集扩展方面, 主要采用无条件的GAN模型进行合成, 即主要从噪声数据中生成医学图像。常用的方法是以深度卷积生成对抗网络 (deep convolutional GAN, DCGAN) 为基线模型进行改进。如Kitchen A等人^[33]基于DCGAN模型成功地合成了前列腺的病灶图像; Schlegl T等人^[34]基于DCGAN提出一种AnoGAN模型, 用来生成多样的视网膜图像, 以辅助视网膜疾病的检测; Chuquicusma M J M等人^[35]采用DCGAN模型生成肺结节数据, 其结

果可达到临床放射科医生无法辨别的程度；Frid-Adar M等人^[36]使用DCGAN生成了3类肝损伤（即囊肿、转移酶、血管瘤）的合成样本，以提高肝病分类的准确性；Bermudez C等人^[37]采用DCGAN的原有训练策略，生成了高质量的人脑T1加权MR图像。

尽管DCGAN在医学图像合成上取得了众多有价值的成果，但其仅能合成分辨率较低的图像。为了提高医学图像合成的质量，一些改进的GAN模型被提出，如Baur C等人^[38]采用LAPGAN，基于拉普拉斯金字塔的思想，利用尺度逐渐变化来生成高分辨率的皮肤病变图像，该方法生成的图像可以有效地提高皮肤疾病分类的准确性。此外，基于渐进生长生成对抗网络（progressive grow GAN, PGGAN）在高分辨率图像合成方面的优势，Korkinof D等人^[39]利用PGGAN合成了分辨率为1 280×1 024的乳腺钼靶X光图像。

2.2.2 医学图像模态转换

医学图像的模态转换合成可以分成两类。一类是单模态的转换，如低剂量CT到普通计量CT图像的转换^[40-41]、3T磁共振仪器采集的MR图像到7T磁共振仪器采集的MR图像的生成^[42-43]，其目的是提高图像质量。另一类是跨模态的一对一转换，如为了解决CT在软组织中对比度不高且会带来辐射的问题，Nie D等人^[44-45]提出上下文感知生成模型，通过级联3D全卷积网络，利用重建损失、对抗损失、梯度损失，采用配对图像进行训练，实现了MR图像到CT图像的合成，提高了合成CT图像的真实性。除了级联模型，在多模态图像转换任务中，常采用的深度模型网络架构为编码-解码结构，典型代表为Pix2Pix^[46]以及CycleGAN^[47]模型。如Maspero M等人^[48]采用Pix2Pix

的网络结构，实现了MR图像到CT图像的转换，进而实现放疗过程中辐射剂量的计算；Choi H等人^[49]基于Pix2Pix模型，从PET图像生成了结构信息更加清晰的脑部MR图像。尽管Pix2Pix模型可以较好地实现多模态图像的转换，但是其要求源图像与目标图像必须空间位置对齐。这种训练数据在临床上是很难获取的。针对源图像和目标图像不匹配的问题，通常采用CycleGAN模型进行图像生成。Wolterink J M等人^[50]使用不配对数据，利用CycleGAN从头部MRI图像合成了其对应的CT图像，合成图像更真实。目前，CycleGAN已成为多模态医学图像转换中广泛采用的手段，如心脏MR图像到CT图像的合成^[51]、腹部MR图像到CT图像的合成^[52]、脑部CT图像到MR图像的合成^[53]等。然而CycleGAN有时无法保留图像的结构边界。Hiasa Y等人^[54]引入梯度一致性损失，对CycleGAN模型进行了改进，该损失通过评估原始图像与合成图像之间每个像素梯度的一致性来保留合成图像的结构边界，进而提高了合成图像的质量。

3 医学图像配准与分割

在很多医学图像分析任务中，获得高质量的图像数据后，经常需要对图像进行配准，并对感兴趣区域进行分割，之后才能进行图像分析和识别。本节分别对深度学习在医学图像配准以及分割领域的应用进行详细的阐述。

3.1 医学图像配准

图像配准是对不同时刻、不同机器采集的图像进行空间位置匹配的过程，是医

学图像处理领域非常重要的预处理步骤之一,在多模态图像融合分析、图谱建立、手术指导、肿瘤区域生长检测以及治疗疗效评价中有广泛的应用。目前,深度学习在医学图像配准领域的研究可以分成3类,第一类是采用深度迭代的方法进行配准,第二类是采用有监督的深度学习模型进行配准,第三类是基于无监督模型的深度学习配准。第一类方法主要采用深度学习模型学习相似性度量,然后利用传统优化方法进行配准的形变^[55-57]。该类方法配准速度慢,没有充分发挥深度学习的优势,因此近几年鲜见报道。本文主要集中介绍有监督学习和无监督学习的医学图像配准。

基于有监督学习的配准在进行网络训练时,需要提供与配准对相对应的真实变形场,其配准框架如图3所示。

网络模型的训练目标是缩小真实变形场与网络输出变形场的差距,最后将变形场应用到待配准的图像上,从而得到配准结果。在有监督学习的医学图像配准中,变形场的标签可以通过以下两种方式获得:一种是将经典配准算法获得的变形场作为标签;另一种是对目标图像进行模拟形变,将形变参数作为真实标签,将形变图像作为待配准图像。

在基于有监督学习的刚性配准方面,

Miao S等人^[58-59]首先结合CNN,采用回归的思想将3D X射线衰减映射图与术中实时的2D X射线图进行刚体配准;Salehi S S M等人^[60]结合深度残差回归网络和修正网络,采用“先粗配准,再细配准”的策略,基于测地线距离损失实现了3D胎儿大脑T1和T2加权磁共振图像的刚体配准,建立了胎儿大脑图谱;随后,Zheng J N等人^[61]采用域自适应的思想,利用预训练网络实现了2D和3D射线图像配准,其设计了成对域适应模块,用来调整模拟训练数据与真实测试数据之间的差异,以提高配准的鲁棒性。

在非线性配准方面,模拟非线性变形场比模拟刚性变形场困难很多,因此在基于有监督学习的非线性配准中,大多采用经典方法获得变形场,并以其为标签,对模型进行训练。Yang X等人^[62]首先以U-Net网络模型为基线结构,利用微分同胚算法获得变形场,并将其作为标签,实现2D和3D脑部MR图像的端到端配准。因为非线性变形场较难模拟,所以在监督学习中引入弱监督配准和双监督配准的概念。弱监督配准指利用解剖结构标签做配准的标记,学习变形场。Hu Y P等人^[63]使用前列腺超声图像和MR图像的结构标记训练CNN模型,学习变形场,然后将变形

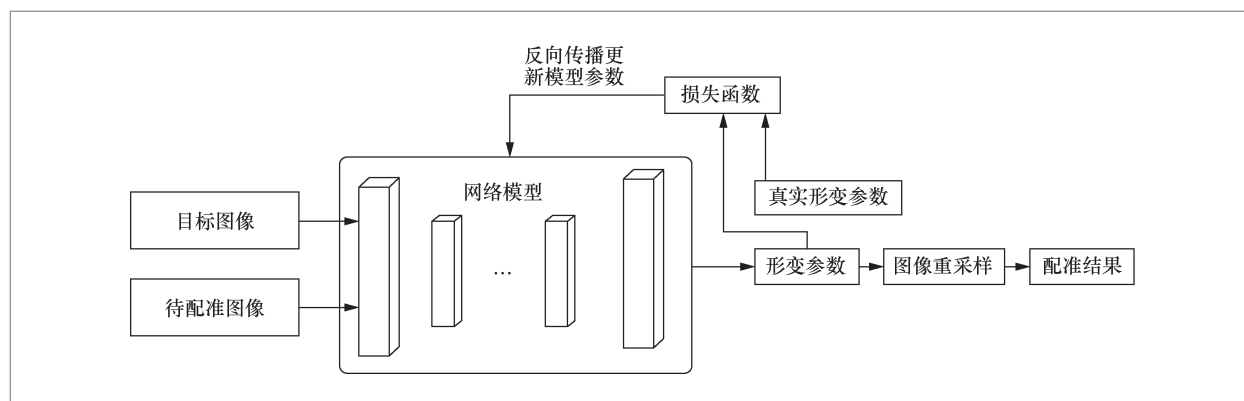


图3 有监督深度学习医学图像配准框架

场施加在灰度图像上,从而实现MR图像和超声图像的配准。Hering A等人^[64]采用相似度测量和组织结构分割标签,同时训练配准网络,提高了心脏MR图像的配准精度。双监督配准是指模型采用两种监督形式的损失函数进行训练,如Cao X H等人^[65]在进行MR图像和CT图像配准时,先利用生成网络将MR图像转换为其对应的CT图像,将CT图像转换为其对应的MR图像,在配准的过程中,同时计算原始MR图像与生成MR图像之间的相似性损失以及原始CT图像与生成CT图像之间的相似性损失,通过两种损失的优化,提高配准的精度; Fan J F等人^[66]结合有监督模型损失和无监督模型损失,实现了脑部MR图像的准确配准。有监督学习的医学图像配准的精度取决于标签的可靠性,因此,如何生成可靠的标签并设计合适的损失函数,是有监督学习的医学图像配准中待解决的难点。

随着空间变换网络(spatial transformer network, STN)^[67]的问世,利用无监督深度学习模型进行医学图像配准成为研究热点。其配准网络框架如图4所示。

Yoo I等人^[68]结合卷积自动编码器(convolutional auto-encoder, CAE)和STN模型,实现了神经组织显微镜图像的

配准,其中CAE负责提取待配准图像与目标图像的特征,基于该特征计算相似性损失,结果表明,该种损失能取得较好的配准结果。2018年, Balakrishnan G等人^[69]提出VoxelMorph网络结构,以U-Net为基线模型,结合STN模块,实现了MR图像的非线性配准;随后,其对模型进行了改进,引入分割标记辅助损失,进一步提高了配准的Dice分数^[70]。Kuang D等人^[71]提出空间变换模块,用于替代U-Net网络结构,在降低模型参数的前提下,实现了脑部MR图像的准确配准。Zhang J^[72]为了进一步提高无监督配准的准确度,除了相似度损失,还引入了变换平滑损失、反向一致性损失以及防折叠损失。其中,变化平滑损失和防折叠损失是为了保证变形场的平滑性。反向一致性损失在互换待配准图像与目标图像时,可保证变形场满足可逆关系。Tang K等人^[73]利用无监督网络实现了脑部MR图像的端到端配准,即网络模型同时学习了仿射变换参数和非线性变换参数。

除了基于CNN模型的无监督配准,采用GAN模型进行配准也已成为一种研究趋势,即采用条件生成对抗网络进行医学图像配准。其中,生成器用来生成变换参

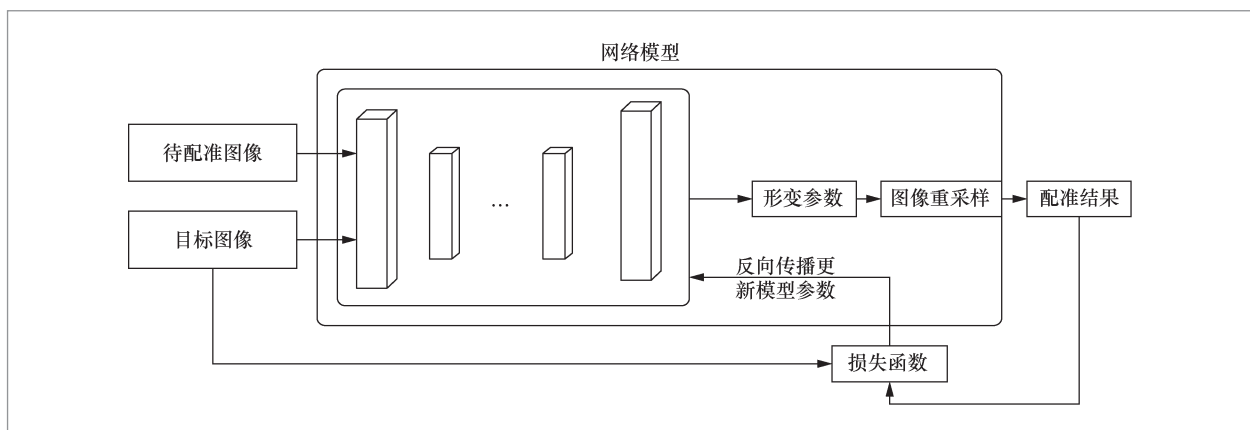


图4 无监督深度学习图像配准网络框架

数或者配准后的图像, 判别器用于对配准图像进行鉴别。通常在生成器与判别器之间插入STN模块, 以进行端到端训练。目前, 基于GAN模型的医学图像配准有较多的应用, 如前列腺MR图像与超声图像配准^[74], 以CycleGAN为基线模型的多模态视网膜图像、单模态MR图像配准^[75], CT图像和MR图像配准^[76]等。在基于GAN的医学图像配准中, GAN模型或者起到正则化的作用, 用来调节变形场及配准图像, 或者用来进行图像转换, 利用交叉域配准提高配准的性能。**表1**总结了典型的无监督配准模型

和有监督配准模型。

3.2 医学图像分割

医学图像分割是计算机辅助诊断的关键步骤, 是进行感兴趣区域定量分析的前提。随着深度学习在语义分割中的快速发展, 将自然图像分割模型扩展到医学图像已成为主要趋势。在医学图像分割中, 采用的主流网络框架有CNN、全卷积网络(full convolutional network, FCN)、U-Net、循环神经网络(recurrent neural

表1 深度学习配准的代表性模型总结

模型类型	配准类型	数据集	变形场来源	器官	模型	评价指标	文献
有监督配准模型	刚体配准	内部数据集	合形成变场	骨骼X-ray	CNN	mTREproj: 0.282 mm	[69-70]
	刚体配准	内部数据集	合形成变场	脊柱CT和X-ray	PDA	TRE: 5.65 mm	[61]
	非线性配准	LBPA40、ISBR18、CUMC12、MGH10	真实形变场	脑部MR	3D-CNN		[62]
	非线性配准	内部数据集	真实形变场	前列腺超声和MR	CNN	TRE=8.5 mm Dice=0.86	[63]
	非线性配准	ACDC dataset	分割监督	心脏MR	CNN	Dice=0.865	[64]
	非线性配准	内部数据集	标签监督	前列腺CT和MR	CNN	ASD=1.58 mm Dice=0.873	[65]
	非线性配准	IBSR18、CUMC12、MGH10、IXI30	双监督学习形变场	脑部MR	BIRNet	Dice>0.728 (分脑区比较)	[66]
无监督配准模型	非线性配准	ADNI、OASIS、ABIDE、ADHD200、MCIC、PPMI、HABS、Harvard GSP		脑部MR	VoxelMorph	Dice=0.78	[69]
	非线性配准	Mindboggle101		脑部MR	FAIM	Dice=0.533 (左上顶叶)	[71]
	非线性配准	ADNI		脑部MR	ICNet	Dice=0.88 ASD=0.71 mm HD=12.71 mm (白质)	[72]
	非线性配准	内部数据集		脑部MR	ADMIR	Dice=0.91 HD=2.68 ASD=0.59	[73]
	非线性配准	内部数据集		前列腺MR和超声	AirNet	TRE=3.48 mm	[74]
	非线性配准	内部数据集、Sunnybrook		视网膜图像、心脏MR	GAN	Dice=0.887/0.79 HD=8.0/5.12 (视网膜/心脏)	[75]
	非线性配准	VISCERAL Anatomy3 benchmark		全身MR、CT	GAN	Dice=0.757 (胸部) Dice=0.783 (腹部)	[76]

network, RNN)和GAN模型。目前常用的医学图像分割模型包括2.5D CNN,即分别在横断面、矢状面、冠状面上使用2D卷积进行分割,在节约计算成本的前提下,充分利用三维空间的邻域信息提高分割的准确度^[77]。FCN是深度学习语义分割的初始模型,通过全卷积神经网络和上采样操作,可以粗略地获得语义分割结果。为了提高分割细节,采用跳跃连接将低层的空间信息和高层的语义信息相结合,以提高图像分割的细腻度。FCN^[78]及其变体(如并行FCN^[79]、焦点FCN^[80]、多分支FCN^[81]、循环FCN^[82]等)已被广泛应用到各种医学图像分割任务中,且表现良好。

U-Net是由一系列卷积和反卷积组成的编码和解码结构,通过跳跃连接实现高级语义特征和低级空间信息的融合,进而保证分割的准确度。U-Net及其变体(如Nested U-Net^[83]、V-Net^[84]、循环残差U-Net^[85])医学图像分割上取得了较好的分割结果,是目前医学图像分割的主流基线模型。

RNN类分割模型主要考虑医学图像分割中切片和切片之间的上下文联系,进而将切片作为序列信息输入RNN及其变体中,从而实现准确分割。典型的模型有CW-RNN(clockwork RNN)^[86]和上下文LSTM模型^[87],其通过抓取相邻切片的相互关系,锐化分割边缘。在此基础上,Chen J X等人^[88]提出双向上下文LSTM模型——BDC-LSTM,即在横断面双向、矢状面双向和冠状面双向上学习上下文关系,其结果比采用多尺度分割的金字塔LSTM模型要好。

基于GAN的分割的主要思想是生成器被用来生成初始分割结果,判别器被用来细化分割结果。一般在分割网络中,生成器常采用FCN或者U-Net网络框架,判别器为常见的分类网络结构,如ResNet、

VGG等。基于GAN的医学图像分割已经被应用到多个器官和组织的医学图像分割任务中^[89-92]。表2为常见医学图像分割模型所用的数据集以及其分割性能对比。

4 医学图像分类及识别

4.1 医学图像分类

医学图像分类和识别是计算机辅助诊断(computer-aided diagnosis, CAD)的最终目标。在深度学习出现前,常采用人工定义的图像特征(如图像的纹理、形状、图像的灰度直方图等),经过特征选择后,再基于机器学习模型(如支持向量机、逻辑回归、随机森林等)进行分类。典型代表为影像组学方法,其在肿瘤的分型分期、治疗的预后预测方面取得了很多重要的成果^[93]。然而,人工定义特征以及特征选择方式很大程度上影响了分类的可靠性和鲁棒性。

近年来,深度学习模型的飞速发展,尤其是CNN的广泛应用,使得利用神经网络模型自动提取和选择特征并进行分类成为主流趋势。CNN模型的不同变体已经在基于医学影像的临床疾病诊断中得到了广泛的应用,例如基于Kaggle公司的眼底图像公开数据集,Shanthi T等人^[94]使用改进的AlexNet进行糖尿病视网膜病变的分类,其精度可以达到96.6%左右;基于VGG,利用胸片进行肺结节的良恶性分类,其精度可高达99%^[95]。目前,在常见的CNN变体中,ResNet和VGG在医学影像分类中的表现最好,因此大多数的肿瘤检测、脑神经系统疾病分类、心血管疾病检测等将这两种模型作为基线模型进行研究。

与自然图像数据相比,医学图像数据

中满足模型训练需求的数据较少。因此,为了提高临床影像智能诊断的准确性,通过知识迁移来训练医学图像分类模型已成为主流。常见的知识迁移包含自然图像到医学图像的迁移、基于临床知识的指导迁移^[96-98]。在自然图像到医学图像的迁移中,主要有两种方式:一种是固定利用自然图像训练的网络模型的卷积层参数,利用该参数提取医学影像特征,然后利用该特征结合传统的机器学习方法进行分类;另一种是将自然图像训练的网络模型参数作为医学图像训练模型的初始化参数,通过微调来实现医学图像分类。除了自然图像到医学图像的迁移,还可以利用其他医学图像数据集,采用多任务学习的方式进行数据信息共享,弥补数据不足带来的分类缺陷^[99]。

基于临床知识的指导迁移将临床医生诊断的经验(如医生的经验学习方式、影

像诊断方式以及诊断关注的图像区域和特征等)融入模型,根据临床医生诊断的经验,即先掌握简单的疾病影像诊断,再进行复杂疾病诊断,研究者们提出了“课程学习”模型,将图像分类任务从易到难进行划分,模型训练先学习简单的图像分类任务,再学习较难的分类任务^[100-101]。基于该方式的学习可以提高分类的准确度。基于医生诊断的方式(如迅速浏览全部医学图像,再选择某些切片进行诊断),研究者提出基于全局和局部的分类模型,其在胸片^[102]和皮肤疾病^[103]的诊断上取得了较好的效果。基于诊断时关注的影像区域,带有注意力机制的分类模型被提出,典型的代表有AG-CNN(attention-based CNN for glaucoma detection)^[104]、LACNN(lesion aware CNN)^[105]和ABN(attention branch network)^[106],通过引入注意力,网络可以关注某些区域,从而提高分类的精度。此

表2 典型的深度学习医学图像分割方法

数据集	模型	器官	损失	精度	文献
内部数据集	2.5D CNN	脑部、乳腺MR,心脏血管造影图像	交叉熵损失		[77]
2016 MICCAI IVDs挑战赛数据集	3D FCN	椎间盘MR	加权交叉熵损失	Dice=0.912	[78]
2017 MICCAI grand challenge on infant brain MRI	Multi-stream 3D FCN	脑部MR	似然损失	Dice=0.954 ASD=0.127 MHD=9.62 (脑脊液)	[81]
MICCAI 2009 LV Segmentation Challenge	Recurrent FCN	心脏MR	交叉熵损失	Dice=0.90 APD=2.05	[82]
TCIA(ProstateX, QINHEADNECK)	U-Net	多器官	Combo损失	Dice=0.92	[92]
DRIVE、STARE、CHASH_DB1	R2 U-Net	多器官	二值交叉熵损失	Dice=0.86	[85]
PROMISE 2012 challenge	V-Net	前列腺MR	Dice损失	Dice=0.87 HD=5.71 mm	[84]
DRIVE Dataset、ISIC 2018	Bi-LSTM	多器官	二值交叉熵损失	F1-Score>0.99	[87]
MICCAI BRATS 2013、2015	SegGAN	头部MR	多尺度L ₁ 范数损失	Dice=0.84/0.85 (BRATS 2013数据集/BRATS 2015数据集)	[100]
INbreast dataset、DDSM dataset	cGAN	乳腺钼靶图像	Dice以及对抗损失	Dice=0.94 (INbreast)	[90]

外,根据医生诊断用到的经验特征,如肿瘤的形状、大小、边界等信息,将人工定义的特征与深度模型提取的特征进行融合,提高医学图像的分类精度,也是一种趋势。如 Majtner T等人^[107]将人工特征分类结果与深度学习分类结果进行融合,提高了皮肤癌分类的准确度;Chai Y D等人^[108]将人工特征和深度学习特征进行融合并训练分类器,从而实现青光眼图像的分类;Xie Y T等人^[109]将人工提取的特征图像块与深度

学习图像块同时作为ResNet模型的输入,实现肺结节的准确分类。如何将深度学习特征与传统人工特征进行有效的融合,是该类模型设计的难点。

4.2 医学图像目标识别

医学图像目标识别也属于临床诊断的一种,即在一幅图像中标记出可能病变的区域,并对其进行分类,如图5所示。

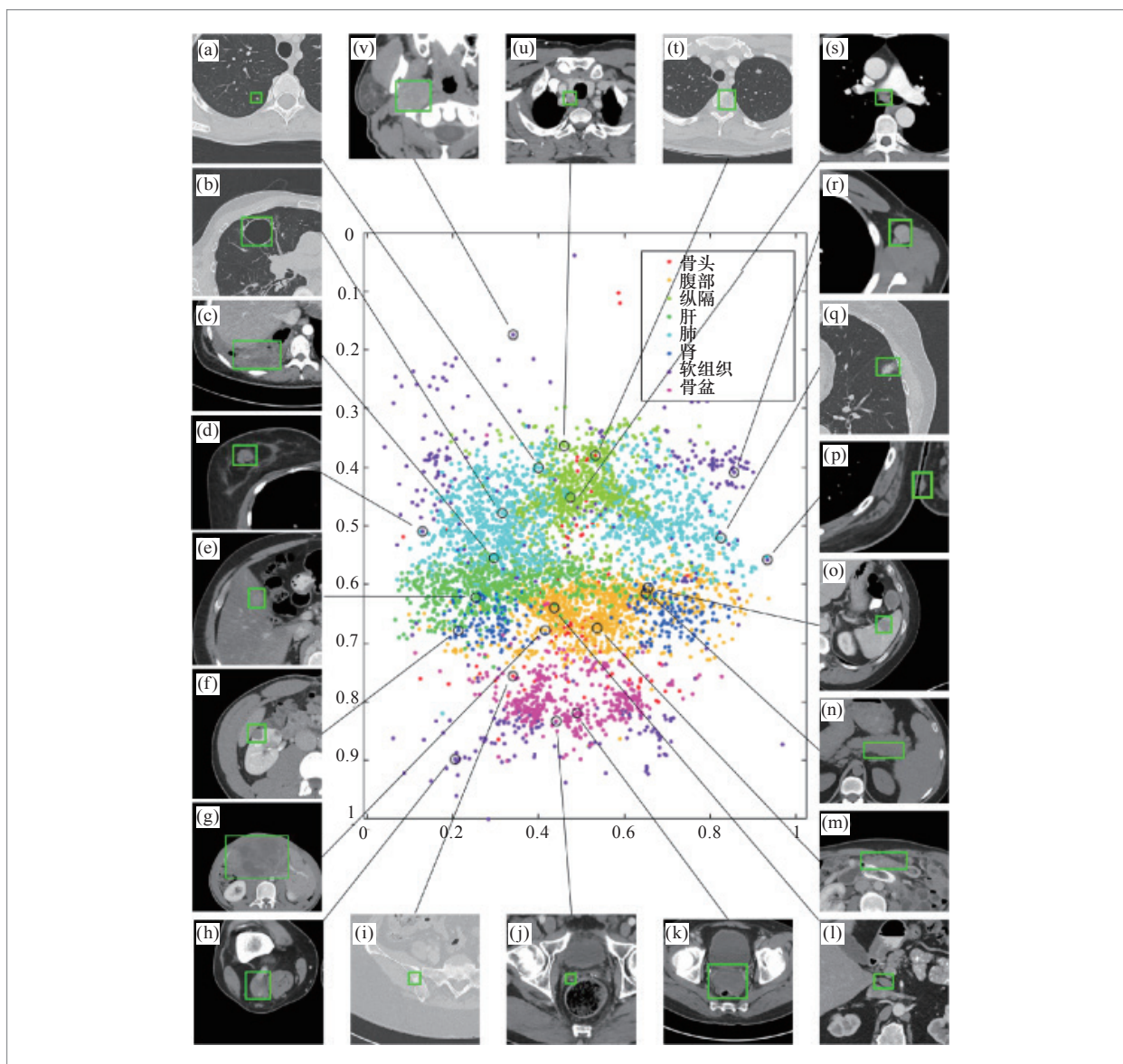


图5 医学图像目标识别示意图^[110]

传统的人工标记识别费时费力。最初将深度学习模型应用于目标识别时,主要是将图像分成小块,逐块输入由CNN等组成的二分类模型中,判断其是否属于目标区域^[111-112]。随着深度学习模型在目标检测领域的快速发展,尤其是Fast R-CNN模型和Mask R-CNN模型的出现,将整幅医学图像输入模型,即可一次找到所有可能的目标区域。但是在这两类模型中均存在一个区域建议模块和一个分类模块,二者需要进行迭代更新,模型的速度并不能满足临床的实时性要求。YOLO (you only look once)和SSD (single shot multibox detector)模型的问世解决了目标检测的实时性问题。基于此类模型, Lin T Y等人^[113]提出RetinaNet模型,并将其扩展应用到病理图像和钼靶图像乳腺肿瘤识别^[114-115]、CT图像的肺结节检测^[98,116]中。上述模型均针对2D图像进行目标检测,忽略了3D图像中切片和切片之间的空间信息。为了提高识别的准确度,基于RNN和LSTM的识别模型被应用到医学图像中^[117-119]。

此外,在医学图像目标识别中,同样存在数据不充足的问题。为了解决这个问题,基于迁移学习的医学图像识别逐渐开展起来,如基于ImageNet数据进行模型迁移,实现肺结节^[120]、乳腺癌^[99]和结直肠息肉的检测^[121]。同时,基于临床经验知识指导的迁移学习也被应用到医学图像的目标检测中。典型代表有AGCL模型,其基于注意力的课程学习,实现胸片中的肿瘤检测^[121]; CASED (curriculum adaptive sampling for extreme data imbalance)模型,其可检测CT图像中的肺结节^[122]; 特征金字塔模型 (feature pyramid network, FPN), 其采用不同对比度的图像,利用多尺度注意力模型实现肿瘤检测^[123]。

图像分类和图像目标识别是医学影像临床诊断的最终目标,是目前人工智能技术与临床紧密结合的研究方向。笔者仅对分类识别的几种情况进行了阐述,以便掌握其发展方向。**表3**给出了肿瘤分类中常用的医学图像数据集以及深度学习模型,并对比了其分类性能。

表3 基于深度学习的医学图像分类总结

器官	数据集	方法	目标	精度	文献
肺	LIDC-IDRI、ANODE09 challenge、DLCST	卷积神经网络	结节分类	Sensitivity=0.854	[95]
	LIDC-IDRI	知识学习	结节分类	AUC=0.957	[109]
	Chest X-ray 14	注意力卷积神经网络	肺部疾病分类	AUC=0.871	[102]
	HUG database	迁移学习	间质性肺疾病分类	F1-score=0.88	[97]
	LIDC以及内部数据集	迁移学习	肺结节检测	AUC=0.812	[98]
	LIDC-IDRI	知识学习	肺结节分类	AUC=0.957	[109]
	LIDC-IDRI	流形学习	肺结节分类	ACC=0.90	[110]
乳腺	DDSM、MIAS、BCDR	迁移学习	乳腺癌分类	AUC=0.997 (MIAS) AUC=0.956 (BCDR)	[99]
	皮肤	迁移学习	皮肤病分类	AUC=0.914	[96]
皮肤	ISIC 2016, 2017	迁移学习	皮肤病分类	AUC=0.914	[96]
	2017 ISBI Challenge on Skin Lesion Analysis Towards Melanoma Detection、EDRA、ISIC	知识学习	皮肤病分类	AUC=0.917	[103]
	ISIC	人工与学习特征结合	皮肤病分类	AUC=0.780	[107]
眼底	Messidor	AlexNet	糖尿病视网膜病变分级	ACC=0.966	[94]
	内部数据集	注意力卷积神经网络	青光眼分类	AUC=0.975	[104]
	UCSD、NEH	病变感知卷积神经网络	眼底疾病分类	AUC>0.96	[103]
	内部数据集	知识学习	青光眼分类	ACC=0.915	[108]

5 结束语

本文从医学图像数据产生、医学图像预处理,以及医学图像识别和分类等方面,阐述了深度学习模型在医学图像分析领域的应用现状。尽管深度学习模型(如CNN、LSTM、GAN、注意力机制、图模型、迁移学习等)在医学图像分析中已取得众多突破,然而将深度学习应用于临床,辅助临床进行精准诊断和个性化治疗仍受到以下几方面的限制。

首先,现有的深度学习模型对影像数目和质量有较高的要求,而临床上带有标记的医学影像数据难以获取,且目前临床诊断预测常使用的方法是有监督学习,数据的不充足势必会影响预测的准确性和稳定性。因此,如何在只有少量有标签数据的情况下,采用弱监督、迁移学习以及多任务学习的思想,提高分类预测的准确度,将是持续的研究热点。

其次,临床应用对可解释性要求较高,而目前深度学习模型所学习的特征无法进行有效的解释。尽管现阶段已有研究学者提出采用可视化以及一些参数分析来对模型和结果进行解释,但是与临床需求中要求的形成可解释的影像学标记还有一定的距离。因此,研究深度学习模型的可解释方法将是医学图像领域的研究热点。

最后,如何提高模型预测的鲁棒性是待解决的难点。现有深度学习模型多数仅针对单一数据集效果较好,无法在不训练的情况下,较好地预测其他数据集。而医学影像由于采集参数、采集设备、采集时间等因素的不同,相同疾病的图像表现可能大不相同,这导致现有模型的鲁棒性和泛化性较差。如何结合脑认知思想改进模型结构以及训练方式,提高深度学习模型

的泛化能力,也是医学图像应用领域中待研究的关键问题。

参考文献:

- [1] GONDARA L. Medical image denoising using convolutional denoising autoencoders[C]// 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). [S.l.:s.n.], 2016: 241-246.
- [2] CHEN H, ZHANG Y, KALRA M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network[J]. IEEE Transactions on Medical Imaging, 2017, 36(12): 2524-2535.
- [3] KANG E, MIN J H, YE J C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction[J]. Medical Physics, 2017, 44(10): 360-375.
- [4] YANG Q S, YAN P K, ZHANG Y B, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1348-1357.
- [5] YOU C Y, YANG Q S, SHAN H M, et al. Structurally-sensitive multi-scale deep neural network for low-dose CT denoising[J]. IEEE Access, 2018, 6: 41839-41855.
- [6] MA Y J, WEI B, FENG P, et al. Low-dose CT image denoising using a generative adversarial network with a hybrid loss function for noise learning[J]. IEEE Access, 2020, 8: 67519-67529.
- [7] YIN X R, ZHAO Q L, LIU J, et al. Domain progressive 3D residual convolution network to improve low-dose CT imaging[J]. IEEE Transactions on Medical Imaging, 2019, 38(12): 2903-2913.
- [8] WU D F, GONG K, KIM K, et al. Consensus neural network for medical

- imaging denoising with only noisy training samples[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. [S.l.:s.n.], 2019: 741–749.
- [9] GEORGESCU M I, IONESCU R T, VERGA N. Convolutional neural networks with intermediate loss for 3D super-resolution of CT and MRI scans[J]. IEEE Access, 2020, 8: 49112–49124.
- [10] OKTAY O, BAI W J, LEE M, et al. Multi-input cardiac image super-resolution using convolutional neural networks[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2016: 246–254.
- [11] PHAM C H, DÍEZ C T, MEUNIER H, et al. Multiscale brain MRI super-resolution using deep 3D convolutional networks[J]. Computerized Medical Imaging and Graphics, 2019: 77.
- [12] MCDONAGH S, HOU B, ALANSARYET A, et al. Context-sensitive super-resolution for fast fetal magnetic resonance imaging[C]// Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Cham: Springer, 2017: 116–126.
- [13] ZHENG Y, ZHEN B, CHEN A, et al. A hybrid convolutional neural network for super-resolution reconstruction of MR images[J]. Medical Physics, 2020, 47(7): 3013–3022.
- [14] ZHAO X L, ZHANG Y L, ZHANG T, et al. Channel splitting network for single MR image super-resolution[J]. IEEE Transactions on Image Processing, 2019, 28(11): 5649–5662.
- [15] TANNO R, WORRALL D E, GHOSH A, et al. Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2017: 611–619.
- [16] PENG C, LIN W A, LIAO H F, et al. SAINT: spatially aware interpolation network for medical slice synthesis[C]// The IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.:s.n.], 2020: 7750–7759.
- [17] SHI J, LIU Q P, WANG C F, et al. Super-resolution reconstruction of MR image with a novel residual learning network algorithm[J]. Physics in Medicine & Biology, 2018, 63(8): 85011.
- [18] LYU Q, SHAN H M, STEBER C, et al. Multi-contrast super-resolution MRI through a progressive network[J]. IEEE Transactions on Medical Imaging, 2020, 39(9): 2738–2749.
- [19] YANG Y, SUN J, LI H B, et al. ADMM-CSNet: a deep learning approach for image compressive sensing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(3): 521–538.
- [20] ADLER J, ÖKTEM O. Learned primal-dual reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1322–1332.
- [21] CHENG J, WANG H F, YING L, et al. Model learning: primal dual networks for fast MR imaging[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. [S.l.:s.n.], 2019: 21–29.
- [22] ZHANG H M, DONG B, LIU B D. JSR-Net: a deep network for joint spatial-radon domain CT reconstruction from incomplete data[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. [S.l.:s.n.], 2019: 3657–3661.
- [23] LEE D, YOO J, YE J C. Deep residual learning for compressed sensing MRI[C]// 2017 IEEE 14th International Symposium on Biomedical Imaging. Piscataway: IEEE Press, 2017: 15–18.
- [24] LEE D, YOO J, TAK S, et al. Deep residual learning for accelerated MRI using magnitude and phase networks[J]. IEEE Transactions on Biomedical Engineering, 2018, 65(9): 1985–1995.

- [25] SCHLEMPER J, CABALLERO J, HAJNAL J V, et al. A deep cascade of convolutional neural networks for dynamic MR image reconstruction[J]. *IEEE Transactions on Medical Imaging*, 2017, 37(2): 491–503.
- [26] HAN Y, YOO J, KIM H H, et al. Deep learning with domain adaptation for accelerated projection-reconstruction MR[J]. *Magnetic Resonance in Medicine*, 2018, 80(3): 1189–1205.
- [27] EO T, JUN Y, KIM T, et al. KIKI-Net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images[J]. *Magnetic Resonance in Medicine*, 2018, 80(5): 2188–2201.
- [28] BAO L J, YE F Z, CAI C B, et al. Undersampled MR image reconstruction using an enhanced recursive residual network[J]. *Journal of Magnetic Resonance*, 2019, 305: 232–246.
- [29] DAI Y X, ZHUANG P X. Compressed sensing MRI via a multi-scale dilated residual convolution network[J]. *Journal of Magnetic Resonance Imaging*, 2019, 63: 93–104.
- [30] YANG G, YU S M, DONG H, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction[J]. *IEEE Transactions on Medical Imaging*, 2017, 37(6): 1310–1321.
- [31] QUAN T M, NGUYEN-DUC T, JEONG W K. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1488–1497.
- [32] MARDANI M, GONG E H, CHENG J Y, et al. Deep generative adversarial neural networks for compressive sensing MRI[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(1): 167–179.
- [33] KITCHEN A, SEAH J. Deep generative adversarial neural networks for realistic prostate lesion MRI synthesis[J]. *arXiv preprint*, 2017, arXiv: 1708.00129.
- [34] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery[C]// *International Conference on Information Processing in Medical Imaging*. Cham: Springer, 2017: 146–157.
- [35] CHUQUICUSMA M J M, HUSSEIN S, BURT J, et al. How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis[C]// *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. [S.l.:s.n.], 2018: 240–244.
- [36] FRID-ADAR M, DIAMANT I, KLANG E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification[J]. *Neurocomputing*, 2018, 321: 321–331.
- [37] BERMUDEZ C, PLASSARD A J, DAVIS L T, et al. Learning implicit brain MRI manifolds with deep learning[C]// *SPIE Medical Imaging Conference*. [S.l.:s.n.], 2018.
- [38] BAUR C, ALBARQOUNI S, NAVAB N. MelanoGANs: high resolution skin lesion synthesis with GANs[J]. *arXiv preprint*, 2018, arXiv: 1804.04338.
- [39] KORKINOF D, RIJKEN T, O'NEILL M, et al. High-resolution mammogram synthesis using progressive generative adversarial networks[J]. *arXiv preprint*, 2018, arXiv: 1807.03401.
- [40] KANG E, CHANG W, YOO J, et al. Deep convolutional framelet denoising for low-dose CT via wavelet residual network[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1358–1369.
- [41] WOLTERINK J M, LEINER T, VIERGEVER M A, et al. Generative adversarial networks for noise reduction in low-dose CT[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2536–2545.
- [42] BAHRAMI K, SHI F, ZONG X P, et al.

- Reconstruction of 7T-like images from 3T MRI[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(9): 2085–2097.
- [43] BAHRAMI K, REKIK I, SHI F, et al. Joint reconstruction and segmentation of 7T-like MR images from 3T MRI based on cascaded convolutional neural networks[C]// *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.:s.n.], 2017: 764–772.
- [44] NIE D, CAO X H, GAO Y Z, et al. Estimating CT image from MRI data using 3D fully convolutional networks[C]// *Deep Learning and Data Labeling for Medical Applications*. Cham: Springer, 2016: 170–178.
- [45] NIE D, TRULLO R, PETITJEAN C, et al. Medical image synthesis with context-aware generative adversarial networks[C]// *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.:s.n.], 2017: 417–425.
- [46] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks[C]// *The IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.:s.n.], 2017: 1125–1134.
- [47] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// *The IEEE International Conference on Computer Vision*. Piscataway: IEEE Press, 2017: 2223–2232.
- [48] MASPERO M, SAVENIJE M, DINKLA A M, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy[J]. *Physics in Medicine & Biology*, 2018, 63(18): 185001.
- [49] CHOI H, LEE D S. Generation of structural MR images from amyloid PET: application to MR-less quantification[J]. *Journal of Nuclear Medicine*, 2018, 59(7): 1111–1117.
- [50] WOLTERINK J M, DINKLA A M, SAVENIJE M H F, et al. Deep MR to CT synthesis using unpaired data[C]// *International Workshop on Simulation and Synthesis in Medical Imaging*. Cham: Springer, 2017: 14–23.
- [51] CHARTSIAS A, JOYCE T, DHARMAKUMAR R, et al. Adversarial image synthesis for unpaired multi-modal cardiac data[C]// *International Workshop on Simulation and Synthesis in Medical Imaging*. [S.l.:s.n.], 2017: 3–13.
- [52] ZHENG J M, CAO J W, WANG Z X, et al. Semi-automatic synthetic computed tomography generation for abdomens using transfer learning and semi-supervised classification[J]. *Journal of Medical Imaging and Health Informatics*, 2019, 9(9): 1878–1886.
- [53] JIN C B, JUNG W, JOO S, et al. Deep CT to MR synthesis using paired and unpaired data[J]. *Sensors*, 2019, 19(10): 2361.
- [54] HIASA Y, OTAKE Y, TAKAO M, et al. Cross-modality image synthesis from unpaired data using CycleGAN[C]// *International Workshop on Simulation and Synthesis in Medical Imaging*. [S.l.:s.n.], 2018: 31–41.
- [55] CHENG X, ZHANG L, ZHENG Y F. Deep similarity learning for multimodal medical images[J]. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2018, 6(3): 248–252.
- [56] SEDGHI A, LUO J, MEHRTASH A, et al. Semi-supervised deep metrics for image registration[J]. *arXiv preprint*, 2018, arXiv: 1804.01565.
- [57] SIMONOVSKY M, GUTIÉRREZ-BECKER B, MATEUS D, et al. A deep metric for multimodal registration[C]// *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.:s.n.], 2016: 10–18.
- [58] MIAO S, WANG Z J, ZHENG Y F, et al. Real-time 2D/3D registration via

- CNN regression[C]// 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Piscataway: IEEE Press, 2016: 1430–1434.
- [59] MIAO S, WANG Z J, LIAO R. A CNN regression approach for real-time 2D/3D registration[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1352–1363.
- [60] SALEHI S S M, KHAN S, ERDOGMUS D, et al. Real-time deep pose estimation with geodesic loss for image-to-template rigid registration[J]. *IEEE Transactions on Medical Imaging*, 2018, 38(2): 470–481.
- [61] ZHENG J N, MIAO S, WANG Z J, et al. Pairwise domain adaptation module for CNN-based 2-D/3-D registration[J]. *Journal of Medical Imaging*, 2018, 5(2): 21204.
- [62] YANG X, KWITT R, STYNER M, et al. Quicksilver: fast predictive image registration—a deep learning approach[J]. *Neuroimage*, 2017, 158: 378–396.
- [63] HU Y P, GIBSON E, GHAVAMI N, et al. Label-driven weakly-supervised learning for multimodal deformable image registration[C]// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). [S.l.:s.n.], 2018: 1070–1074.
- [64] HERING A, KUCKERTZ S, HELDMANN S, et al. Enhancing label-driven deep deformable image registration with local distance metrics for state-of-the-art cardiac motion tracking[M]// *Bildverarbeitung für die Medizin 2019*. Wiesbaden: Springer Vieweg, 2019: 309–314.
- [65] CAO X H, YANG J H, WANG L, et al. Deep learning based inter-modality image registration supervised by intra-modality similarity[C]// *International Workshop on Machine Learning in Medical Imaging*. Cham: Springer, 2018: 55–63.
- [66] FAN J F, CAO X H, YAP P T, et al. BIRNet: brain image registration using dual-supervised fully convolutional networks[J]. *Medical Image Analysis*, 2019, 54: 193–206.
- [67] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[C]// *The 28th International Conference on Neural Information Processing Systems*. New York: ACM Press, 2015: 2017–2025.
- [68] YOO I, HILDEBRAND D G C, TOBIN W F, et al. ssEMnet: serial-section electron microscopy image registration using a spatial transformer network with learned features[C]// *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer, 2017: 249–257.
- [69] BALAKRISHNAN G, ZHAO A, SABUNCU M R, et al. VoxelMorph: a learning framework for deformable medical image registration[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1788–1800.
- [70] ZHAO A, BALAKRISHNAN G, DURAND F, et al. Data augmentation using learned transformations for one-shot medical image segmentation[C]// *The IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2019: 8543–8553.
- [71] KUANG D, SCHMAH T. FAIM—a ConvNet method for unsupervised 3D medical image registration[M]// *Machine Learning in Medical Imaging*. [S.l.:s.n.], 2019: 646–654.
- [72] ZHANG J. Inverse-consistent deep networks for unsupervised deformable image registration[J]. *arXiv preprint*, 2018, arXiv1809.03443.
- [73] TANG K, LI Z, TIAN L L, et al. ADMIR—affine and deformable medical image registration for drug-addicted brain images[J]. *IEEE Access*, 2020, 8: 70960–70968.
- [74] YAN P K, XU S, RASTINEHAD A R, et al.

- Adversarial image registration with application for MR and TRUS image fusion[C]// International Workshop on Machine Learning in Medical Imaging. Cham: Springer, 2018: 197–204.
- [75] MAHAPATRA D, ANTONY B, SEDAI S, et al. Deformable medical image registration using generative adversarial networks[C]// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Piscataway: IEEE Press, 2018: 1449–1453.
- [76] TANNER C, OZDEMIR F, PROFANTER R, et al. Generative adversarial networks for MR–CT deformable image registration[J]. arXiv preprint, 2018, arXiv: 1807.07349.
- [77] MOESKOPS P, WOLTERINK J M, VAN DER VELDEN B H M, et al. Deep learning for multi–task medical image segmentation in multiple modalities[C]// International Conference on Medical Image Computing and Computer–Assisted Intervention. Cham: Springer, 2016: 478–486.
- [78] LI X M, DOU Q, CHEN H, et al. 3D multi–scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi–modality MR images[J]. Medical Image Analysis, 2018, 45: 41–54.
- [79] BI L, KIM J, KUMAR A, et al. Stacked fully convolutional networks with multi–channel learning: application to medical image segmentation[J]. The Visual Computer: International Journal of Computer Graphics, 2017, 33(6–8): 1061–1071.
- [80] ZHOU X Y, SHEN M, RIGA C, et al. Focal FCN: towards small object segmentation with limited training data[J]. arXiv preprint, 2017, arXiv: 1711.01506.
- [81] ZENG G D, ZHENG G Y. Multi–stream 3D FCN with multi–scale deep supervision for multi–modality isointense infant brain MR image segmentation[C]// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Piscataway: IEEE Press, 2018: 136–140.
- [82] POUDEL R P K, LAMATA P, MONTANA G. Recurrent fully convolutional neural networks for multi–slice MRI cardiac segmentation[C]// Reconstruction, Segmentation, and Analysis of Medical Images. Cham: Springer, 2016: 83–94.
- [83] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHS N, et al. Unet++: a nested U–Net architecture for medical image segmentation[C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3–11.
- [84] MILLETARI F, NAVAB N, AHMADI S A. V–Net: fully convolutional neural networks for volumetric medical image segmentation[C]// 2016 4th International Conference on 3D Vision (3DV). Piscataway: IEEE Press, 2016: 565–571.
- [85] ALOM M Z, HASAN M, YAKOPCIC C, et al. Recurrent residual convolutional neural network based on U–Net (R2U–Net) for medical image segmentation[J]. arXiv preprint, 2018, arXiv: 1802.06955.
- [86] XIE Y P, ZHANG Z Z, SAPKOTA M, et al. Spatial clockwork recurrent neural network for muscle perimysium segmentation[C]// International Conference on Medical Image Computing and Computer–Assisted Intervention. Cham: Springer, 2016: 185–193.
- [87] AZAD R, ASADI–AGHBOLAGHI M, FATHY M, et al. Bi–directional ConvLSTM U–net with densely connected convolutions[C]// 2019 IEEE/CVF International Conference on Computer Vision Workshop. Piscataway: IEEE Press, 2019: 406–415.
- [88] CHEN J X, YANG L, ZHANG Y Z, et al. Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation[C]// Advances in Neural Information Processing Systems. [S.l.:s.n.], 2016: 3036–3044.
- [89] XUE Y, XU T, ZHANG H, et al. SegAN:

- adversarial network with multi-scale L_1 loss for medical image segmentation[J]. *Neuroinformatics*, 2018, 16(3-4): 383-392.
- [90] SINGH V K, RASHWAN H A, ROMANI S, et al. Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network[J]. *Expert Systems with Applications*, 2020, 139: 112855.
- [91] ZHANG C Y, SONG Y, LIU S D, et al. MS-GAN: GAN-based semantic segmentation of multiple sclerosis lesions in brain magnetic resonance imaging[C]// 2018 Digital Image Computing: Techniques and Applications (DICTA). Piscataway: IEEE Press, 2018: 1-8.
- [92] TAGHANAKI S A, ZHENG Y F, ZHOU S K, et al. Combo loss: handling input and output imbalance in multi-organ segmentation[J]. *Computerized Medical Imaging and Graphics*, 2019, 75: 24-33.
- [93] 韩冬, 李其花, 蔡巍, 等. 人工智能在医学影像中的研究与应用[J]. *大数据*, 2019, 5(1): 39-67.
- HAN D, LI Q H, CAI W, et al. Research and application of artificial intelligence in medical imaging[J]. *Big Data Research*, 2019, 5(1): 39-67.
- [94] SHANTHI T, SABEENIAN R S. Modified AlexNet architecture for classification of diabetic retinopathy images[J]. *Computers & Electrical Engineering*, 2019, 76: 56-64.
- [95] LI X C, SHEN L L, XIE X P, et al. Multi-resolution convolutional networks for chest X-ray radiograph based lung nodule detection[J]. *Artificial Intelligence in Medicine*, 2020, 103: 101744.
- [96] MAHBOD A, SCHAEFER G, ELLINGER I, et al. Fusing fine-tuned deep features for skin lesion classification[J]. *Computerized Medical Imaging and Graphics*, 2019, 71: 19-29.
- [97] CHRISTODOULIDIS S, ANTHIMOPOULOS M, EBNER L, et al. Multisource transfer learning with convolutional neural networks for lung pattern analysis[J]. *IEEE Journal of Biomedical and Health Informatics*, 2016, 21(1): 76-84.
- [98] HARSONO I W, LIAWATIMENA S, CENGGORO T W. Lung nodule detection and classification from thorax CT-scan using RetinaNet with transfer learning[J]. *Journal of King Saud University-Computer and Information Sciences*, 2020: Accepted.
- [99] ALKHALEEF M, MA S C, CHANG Y L, et al. Double-shot transfer learning for breast cancer classification from X-ray images[J]. *Applied Sciences*, 2020, 10(11): 3999.
- [100] ABBAS A, ABDELSAMEA M M, GABER M M. Detrac: transfer learning of class decomposed medical images in convolutional neural networks[J]. *IEEE Access*, 2020, 8: 74901-74913.
- [101] OKSUZ I, RUIJSINK B, PUYOL-ANTÓN E, et al. Automatic CNN-based detection of cardiac MR motion artefacts using k-space data augmentation and curriculum learning[J]. *Medical Image Analysis*, 2019, 55: 136-147.
- [102] GUAN Q J, HUANG Y P, ZHONG Z, et al. Diagnose like a radiologist: attention guided convolutional neural network for thorax disease classification[J]. *arXiv preprint*, 2018, arXiv: 1801.09927.
- [103] GONZÁLEZ-DÍAZ I. Dermaknet: incorporating the knowledge of dermatologists to convolutional neural networks for skin lesion diagnosis[J]. *IEEE Journal of Biomedical and Health Informatics*, 2018, 23(2): 547-559.
- [104] LI L, XU M, WANG X F, et al. Attention based glaucoma detection: a large-scale database and CNN model[C]// The IEEE Conference on Computer Vision and Pattern Recognition. [S.l.:s.n.], 2019: 10571-10580.
- [105] FANG L Y, WANG C, LI S T, et al. Attention to lesion: Lesion-aware convolutional neural network for retinal optical coherence tomography image classification[J]. *IEEE Transactions on*

- Medical Imaging, 2019, 38(8): 1959–1970.
- [106] MITSUHARA M, FUKUI H, SAKASHITA Y, et al. Embedding human knowledge in deep neural network via attention map[J]. arXiv preprint, 2019, arXiv: 1905.03540.
- [107] MAJTNER T, YILDIRIM-YAYILGAN S, HARDEBERG J Y. Combining deep learning and hand-crafted features for skin lesion classification[C]// 2016 6th International Conference on Image Processing Theory, Tools and Applications. Piscataway: IEEE Press, 2016: 1–6.
- [108] CHAI Y D, LIU H Y, XU J. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models[J]. Knowledge-Based Systems, 2018, 161: 147–156.
- [109] XIE Y T, XIA Y, ZHANG J P, et al. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT[J]. IEEE Transactions on Medical Imaging, 2018, 38(4): 991–1004.
- [110] YAN K, WANG X S, LU L, et al. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning[J]. Journal of Medical Imaging, 2018, 5(3): 36501.
- [111] XUE Z, JAEGER S, ANTANI S, et al. Localizing tuberculosis in chest radiographs with deep learning[C]// Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications. [S.l.:s.n.], 2018.
- [112] DING J, LI A X, HU Z Q, et al. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2017: 559–567.
- [113] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(2): 318–327.
- [114] LOTTER W, DIAB A R, HASLAM B, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using annotation-efficient deep learning approach[J]. arXiv preprint, 2019, arXiv: 1912.11027.
- [115] MERCAN C, BALKENHOL M, VAN DER LAAK J, et al. From point annotations to epithelial cell detection in breast cancer histopathology using RetinaNet[C]// International Conference on Medical Imaging with Deep Learning-Extended Abstract Track. [S.l.:s.n.], 2019.
- [116] ZLOCHA M, DOU Q, GLOCKER B. Improving RetinaNet for CT lesion detection with dense masks from weak RECIST labels[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. [S.l.:s.n.], 2019: 402–410.
- [117] SANTERAMO R, WITHEY S, MONTANA G. Longitudinal detection of radiological abnormalities with time-modulated LSTM[C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 326–333.
- [118] SU Y T, LU Y, CHEN M, et al. Spatiotemporal joint mitosis detection using CNN-LSTM network in time-lapse phase contrast microscopy images[J]. IEEE Access, 2017, 5: 18033–18041.
- [119] GAO R Q, HUO Y K, BAO S X, et al. Distanced LSTM: time-distanced gates in long short-term memory models for lung cancer detection[C]// International Workshop on Machine Learning in Medical Imaging. [S.l.:s.n.], 2019: 310–318.
- [120] SHIN H C, ROTH H R, GAO M C, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning[J]. IEEE Transactions on Medical Imaging, 2016,

- 35(5): 1285–1298.
- [121] ZHANG R K, ZHENG Y L, MAK T W C, et al. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain[J]. IEEE Journal of Biomedical and Health Informatics, 2016, 21(1): 41–47.
- [122] JESSON A, GUIZARD N, GHALEHJEGH S H, et al. CASED: curriculum adaptive sampling for extreme data imbalance[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2017: 639–646.
- [123] LIU J Y, CAO L L, AKIN O, et al. 3DFPN-HS 2: 3D feature pyramid network based high sensitivity and specificity pulmonary nodule detection[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Cham: Springer, 2019: 513–521.

作者简介



王丽会 (1982–), 女, 博士, 贵州大学计算机科学与技术学院、贵州省智能医学影像分析与精准诊断重点实验室副教授, 主要研究方向为医学成像、机器学习与深度学习、医学图像处理、计算机视觉。



秦永彬 (1980–), 男, 博士, 贵州大学计算机科学与技术学院、贵州省智能医学影像分析与精准诊断重点实验室教授, 主要研究方向为大数据治理与应用、文本计算与认知智能。

收稿日期: 2020-07-06

通信作者: 秦永彬, ybqin@foxmail.com

基金项目: 国家自然科学基金资助项目 (No.61661010, No.62066008); 中法“蔡元培”交流合作项目 (No.2018N*41400TC)

Foundation Items: The National Natural Science Foundation of China (No.61661010, No.62066008), The Program PHC-Cai Yuanpei (No. 2018N*41400TC)