

人工智能在医学影像中的 研究与应用

韩冬¹, 李其花², 蔡巍³, 夏雨薇², 宁佳¹, 黄峰¹

1. 沈阳东软医疗系统有限公司, 辽宁 沈阳 110167; 2. 慧影医疗科技(北京)有限公司, 北京 100192
3. 东软集团股份有限公司, 辽宁 沈阳 110179

摘要

近年来,人工智能成为学术界和工业界的研究热点,并已经成功应用于医疗健康等领域。着重介绍了人工智能在医学影像领域最新的研究与应用进展,包括智能成像设备、智能图像处理与分析、影像组学、医学影像与自然语言处理的结合等前沿方向。分析了研究和发展从源头入手的全链条人工智能技术的重要性和可行性,阐述了学术界和工业界在这一重要方向上的创新性工作。同时指出,人工智能在医学影像领域中的研究尚处于起步阶段,人工智能与医学影像的结合将成为国际上长期的研究热点。

关键词

人工智能;医学影像;成像方法;图像处理与分析;自然语言处理

中图分类号:TP18

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2019004

Research and application of artificial intelligence in medical imaging

HAN Dong¹, LI Qihua², CAI Wei³, XIA Yuwei², NING Jia¹, HUANG Feng¹

1. Neusoft Medical Systems, Shenyang 110167, China
2. Huiying Medical Technology (Beijing), Beijing 100192, China
3. Neusoft Corporation, Shenyang 110179, China

Abstract

In recent years, artificial intelligence (AI) has become the research hotspot in both academic and industrial societies, and has been applied to biometrics, natural language processing (NLP), finance and healthcare, et al., with huge success. The most recent progress of research and application of AI in medical imaging was introduced, including intelligent imaging system, intelligent image processing and analysis, radiomics and the combination of medical image and NLP, et al. The importance and feasibility of developing full-pipeline AI techniques were elaborated, and the related innovative work from both academic and industrial societies in recent years was introduced. Research of AI in medical imaging is still in its early stage and much more work needs to be done in the future, which will make this research area a long-term international hotspot.

Key words

artificial intelligence, medical imaging, imaging method, image processing and analysis, natural language processing

1 引言

人工智能 (artificial intelligence, AI) 是当下学术界和产业界的一个热点。经过近几年的高速发展,深度学习已经实现了在传统的图像、视频、语音识别等领域的落地,并迅速地向文本处理、自然语言理解、人机对话、情感计算等方面渗透,并在安防、物流、无人驾驶等行业发挥了重要作用。

人口老龄化问题的显现以及人们对健康与日俱增的要求,对目前有限的医疗资源和医疗技术提出了更大的挑战。医疗领域亟需新的技术满足这些需求。与此同时,国内外与医疗相关的人工智能技术也在飞速地发展,科研和创业项目如雨后春笋,为解决医疗领域的挑战提供了新的机遇。目前已经出现了计算机辅助诊断、智能专家系统、手术机器人、智能药物研发以及健康管理等多种产品。

在众多的医疗信息中,医学影像是疾病筛查和诊断、治疗决策的最主要的信息来源。基于医学影像的诊断和治疗是一个典型的长链条、专业化的领域,涵盖了医学影像成像、图像处理与分析、图像可视化、疾病早期筛查、风险预测、疾病辅助检测与诊断、手术计划制定、术中辅助导航、随访跟踪与分析、康复计划制定等一系列方向。目前,医院存储的信息超过90%是影像信息,影像信息已经形成了巨大的数据积累。为此,基于医学影像大数据的人工智能技术与应用就成为医疗机构、科研、产业和政府共同关注的焦点。

医学影像链可以分为成像和图像挖掘两部分。首先,作为信息源头的医学成像设备,其成像质量会对后续疾病的检测、诊断与治疗起到至关重要的作用。利用AI技术

可以实现医学影像成像质量的提升,AI优化的扫描 workflow 可以显著提高扫描效率,并使成像质量趋于标准化,从而给整个医疗健康链条带来深远的影响,具有重要的临床与科研价值。

其次,理解医学图像、提取其中具有诊断和治疗决策价值的关键信息是诊疗过程中非常重要的环节。AI辅助诊断可以承担繁琐的病灶筛查工作,迅速地从海量数据中提取出与诊断相关的有价值的信息,同时避免人工阅片带来的主观性差异。AI辅助图像处理算法还可以迅速地完成分割配准等复杂功能,为用以治疗的医疗设备(例如手术导航和手术机器人)提供精准的病灶结构信息。

此外,目前AI在医学影像方面的应用还存在着诸多挑战,比如扫描成本和患者隐私问题使得医疗大数据难以被获得;医学图像的标注者需要具有一定的医学背景,获得高质量的医学图像标注甚至比获得医学图像代价更高;不同病变类型与正常的医学图像在数据量上的差距巨大。一些基于AI的非结构医学信息数据分析方法,可以结合影像和医生撰写的诊断报告,自动或者半自动地生成标注数据,扩充训练数据集。

医学影像大数据和人工智能涉及影像链中的成像、图像挖掘、利用文本和影像的关联解决部分图像标注这3方面的问题。本文主要从智能医学成像系统、医学影像大数据与智能图像分析、医学影像与自然语言文本处理的结合分析这3个方面详细介绍国内外的研究现状与发展趋势。

2 智能医学成像系统

随着各种大型医学成像设备在各层级

医院的普及,医学影像检查已经成为各种疾病诊断过程中最为重要的检查手段和诊断依据。医学影像相比于生化等其他检查,可以提供病变的位置、结构和功能等重要信息,为疾病的诊断和治疗提供直观的信息和参考。目前,各种医疗成像系统还面临着成像速度较慢、图像质量受患者配合度影响、成像 workflow 繁琐等各种挑战。本节将介绍人工智能对医学成像系统各方面的改进。

2.1 医学成像系统简介

临床中常用的医学影像模态有多种,包括电子计算机断层扫描(computed tomography, CT)、磁共振成像(magnetic resonance imaging, MRI)、正电子发射计算机断层显像(PET-CT)、X射线、超声等。不同模态的医学影像适合观察不同类型的生理病理信息。医学影像设备采集人体中与特定生理病理信息相关的物理信号,并依据信号传播的数学物理模型重建生理病理信息精确的二维、三维空间分布(即医学图像)。CT图像反映的是人体内不同组织对特定谱段X射线的衰减系数分布情况,由于衰减系数与物质密度直接相关,因此CT图像间接地反映出人体组织的三维密度分布。MRI可以反映人体中质子的密度、纵向弛豫时间(T_1)、横向弛豫时间(T_2)、质子扩散等多种对比度。PET图像可以反映出人体内氟代脱氧葡萄糖(^{18}F -FDG)的代谢分布情况。

三维医学图像成像过程属于逆问题(inverse problem)范畴,即通过在体外采集到的物理信号,重建能够产生该观测信号的人体内生理病理信息的分布情况。这一过程通常是不稳定的,其原因可能有多种,例如,由于扫描时间和剂量的限制而造成的采集信号量不足、原始采集信号中

存在着固有噪声、原始信号中掺杂着其他干扰信号、成像设备硬件的不完美性、患者在成像过程中不自主的随机运动等。这些因素使得基于理想数学物理模型的解析重建方法通常无法得到满足临床要求的图像。为了解决这一问题,传统上需要在重建过程中加入针对图像的特定先验信息(prior information)作为约束条件,以使重建过程更加稳定。常用的约束包括图像 L_1 范数、 L_2 范数、全变分(total variation, TV)约束等。这些简单的约束条件具有良好的通用性,但却无法准确地反映数据的本质特征,对于特定的影像模态与成像模式无法得到最优的重建结果,因此如何设计更好的先验约束,一直是医学影像成像领域的难题之一。

近年来,AI技术的快速发展,特别是其在计算机视觉(computer vision, CV)、图像处理与分析等领域的重要突破,使得国际上的研究人员逐渐认识到将AI技术应用于医学影像成像领域的可能性。AI技术(特别是深度学习技术)摒弃了传统的人工设计的图像先验信息,采用一种完全数据驱动(data-driven)的方式,学习图像固有的深层次先验信息。这些学习到的先验信息具有更加专业化的特点,将其应用于特定领域中,效果远优于传统的浅层次先验信息。目前,AI技术与医学影像成像方法的结合已经成为领域内的研究热点,相关的研究成果呈爆发式增长。

2.2 快速医学影像成像方法

在临床医学影像扫描过程中,成像速度始终是倍受关注的重要因素之一,长扫描时间会降低影像科室的日均流量,还会给患者带来不适。扫描过程中患者的不自主运动也会对成像质量产生不良的影

响。在快速成像方面,国际上相关研究主要集中在磁共振成像加速方面,目前已经发表了大量研究工作,是AI与成像相结合的热点方向之一。

Mardani M等人^[11]提出了一种基于生成对抗网络(generative adversarial network, GAN)的磁共振(magnetic resonance, MR)压缩感知(compressed sensing, CS)快速成像方法,该方法利用GAN对高质量MR图像的低维流形(manifold)进行建模。GAN由生成器(generator)和判别器(discriminator)组成,生成器的作用是将低质量的MR图像映射到高质量图像的流形上,判别器的作用是对映射后的图像质量进行评判。生成器网络的损失函数(loss function)由图像域 L_1/L_2 范数和GAN损失函数组成,其中 L_1/L_2 范数用于抑制图像中的噪声,而GAN损失函数用于保留图像的细节信息。为了保证生成的图像真实可靠,笔者将k空间(k-space)数据保真(data fidelity)项引入网络。实验结果表明,该方法可以实现至少5倍的扫描加速,同时成像结果明显优于传统的压缩感知算法。

Schlemper J等人^[12]提出了一种基于级联深度神经网络(cascaded DNN)的MR快速成像方法。级联深度神经网络由若干个网络单位级联而成,每个网络单元包含卷积神经网络(convolutional neural network, CNN)和数据保真项两个部分,其中CNN以残差网络(residual network, ResNet)的形式构建。因此,CNN学习到的是降采(under-sampled)图像与满采(fully-sampled)图像之间的差异。通过CNN与数据保真项的多次交替处理,可以将原始复杂的MR图像重建问题转化为一系列子过程的顺序执行,而每一个子过程仅需对前一子过程的结果进行进一步的优化即可。相比于整个重建问题,其难度显著

地降低了,从而使重建过程变得更加稳定。实验结果表明,级联深度神经网络的重建图像质量相比于传统的压缩感知方法以及基于字典学习的图像重建方法有了明显的提升,同时其重建一幅二维心脏图像的时间仅为23 ms,基本达到了准实时的效果。

为了将传统迭代重建方法与深度学习各自的优势结合起来,Yang Y等人^[13]提出了一种基于交替方向乘子算法(alternating direction method of multipliers, ADMM)的MR图像重建方法——ADMM-Net,该方法将经典的ADMM迭代重建方法利用神经网络进行重新实现。ADMM-Net对特定迭代次数的ADMM方法进行建模,在每次迭代中,利用CNN解决ADMM算法中的3个子优化问题,整个网络以端到端(end-to-end)的方式进行训练。ADMM-Net的优势在于各种参数可以通过完全数据驱动的、端到端的学习方式得到。实验结果证明,ADMM-Net的重建结果明显优于传统方法。此外,ADMM-Net的构建参照了经典的ADMM,因此,网络的重建结果具有更好的可解释性。

深度神经网络自身的复杂性以及端到端的学习特性,使其通常被看作一个黑盒(black box)方法。为了进一步证明将深度学习应用于成像方向的理论上的合理性,Ye J C等人^[14]提出利用卷积框架(convolution framelets)方法从理论上加以解释。卷积框架最初用来拓展低秩Hankel矩阵(low-rank Hankel matrix)理论在逆问题中的应用。Ye J C等人提出了一种深度卷积框架神经网络(deep convolutional framelet neural network, DCFNN),并证明了在采用修正线性单元(rectified linear unit, ReLU)非线性激活函数的情况下,该网络可以实现完美重建,同时也证明了常用的网络组件(如

residual blocks、concatenated ReLU等)确实可以促进完美重建的实现。此外,基于文章中的理论分析,作者指出了现有的基于深度学习的成像方法的局限性,并通过实验验证了DCFNN方法优于现有的基于深度学习的方法。

通常,一次MR扫描可以生成多种不同对比度的图像。现有的基于深度学习的单一对比度快速成像方法没有充分利用不同对比度图像之间的结构相似性,因此限制了其可以达到的加速比。为了进一步提升重建图像的质量,Chen M等人^[5]提出了一种Multi-echo图像联合重建方法,该方法采用U-Net^[6]实现图像重建,通过将6-echo的图像作为不同的通道输入网络中,使得在卷积过程中能够充分利用不同echo图像间的结构相似性,从而为网络的训练加入更多的约束条件,让训练过程变得更加稳定。实验结果表明,该方法可以实现4.2倍的MR成像加速,重建图像在均方根误差(root mean square error, RMSE)和结构相似性指数(structural similarity index, SSIM)等方面均优于单一对比度重建方法。

由于CNN卷积操作的空间局部特性,目前绝大部分基于深度学习的快速成像方法选择在图像域进行处理。然而,一些因k空间数据不完备性造成的图像伪影却很难在图像域完美地解决。为了解决这一问题,Eo T等人^[7]提出了一种基于双域深度学习的MR快速成像方法,在图像域和频率域均设计了对应的深度CNN,试图从两个不同空间分别对未采集的数据进行恢复,同时图像域与频率域通过数据保真项被关联起来,从而保证重建得到的图像的真实可靠性。实验结果表明,图像域CNN和频率域CNN在图像重建过程中的作用是不同的,相比于仅采用图像域CNN的成像方法,将二者结合起来可以获得质量更高

的图像重建结果。

在工业界,目前关于AI技术应用于快速成像的相关报道较少,尚处于研究探索阶段。其中,国产医疗设备厂商沈阳东软医疗系统有限公司(以下简称东软医疗)在这一领域进行了一些工作,例如东软医疗研发的BrainQuant技术(头部一站式多对比度定量成像技术),可以实现同时获得至少10种高分辨率的、包括定性和定量值在内的三维全脑图像。近年来,定量成像(例如T1 mapping、T2* mapping、定量磁敏感成像等)在临床上的价值受到了越来越多的关注。传统的磁共振扫描技术在获得这些定性和定量性质图像时,是基于多个独立的扫描的,要获得全部图像需要几十分钟甚至更长时间,因此定量成像在临床中的应用受到了长扫描时间的严重制约。BrainQuant通过创新性的数据采集技术,同时结合了AI技术在快速成像方面的优异性能,可在1.5 T设备上5 min内获取至少10种定性和定量的三维高分辨图像,具有重要的临床应用价值。图1为BrainQuant技术同时获取到的10种不同对比度的图像。

2.3 医学图像质量增强方法

在成像过程中,采集数据量的不足、信号中的固有噪声、患者不自主的运动等原

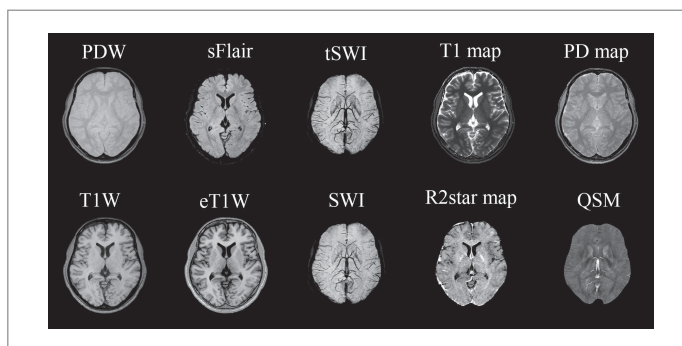


图1 BrainQuant技术在5 min内同时获取到的10种对比度图像

因造成了重建图像中容易出现伪影和噪声等影响医学临床诊断的问题。传统方法在处理这类问题时存在很多局限性。近年来,国际上的学者开始将AI技术应用于医学图像质量增强领域,并取得了长足的进展。

2.3.1 CT图像质量增强

有关CT图像质量的增强研究主要集中在如何利用AI技术处理由于降低放射剂量而带来的噪声和由于减少投影数量而带来的条状伪影(streak artifact)。在低剂量图像去噪方面,Chen H等人^[8]提出了一种基于残差自编码器(residual auto-encoder)的CT图像去噪方法,该方法利用神经网络构建一个自编码器(auto-encoder, AE),不同之处在于网络的编码器和解码器部分采用残差的方式连接。这样做的好处是可以将不同层次的图像特征结合起来,提升网络的建模能力;还可以使训练过程中误差的反向传播更加有效,提升网络的训练效果。此外,网络采用残差的方式连接,使网络实际上学习到的是噪声图像到噪声的映射,这比直接学习从噪声图像到高质量图像的映射更加容易。参考文献[8]给出了仿真实验结果和临床图像实验结果,相比于传统的图像去噪方法(如BM3D),Chen H等人提出的残差自编码器方法在峰值信噪比(peak signal to noise ratio, PSNR)、SSIM等指标上均有明显的优势,同时去噪速度也更快。

通常而言,基于深度学习的图像去噪方法容易产生一定的过平滑(over-smooth)现象,这是因为网络的损失函数通常采用整幅图像的 L_1/L_2 范数,并没有对细节区域进行特别处理,而细节区域的误差在整体误差中的比例很小,所以容易在网络的训练过程中被“淹没”掉,最终

导致图像细节丢失。为了解决这一问题,Wolterink J M等人^[9]设计了一种基于GAN的CT图像去噪方法,GAN用于学习从低剂量图像到正常剂量图像的映射,判别网络用于判别生成的去噪后的图像是否处于正常剂量图像所在的流形中,即是否和真实的正常剂量图像在视觉上相似。本质上,判别网络可以看作一个计算机自己学到的损失函数,相比于人工设计的损失函数,其能够学到更加高层次和细节化的图像特征,因此可以得到更加准确的训练结果。实验结果表明,Wolterink J M等人提出的基于GAN的图像去噪方法能够有效地去除低剂量CT图像中的噪声,同时能够很好地保护图像的细节信息,使去噪后的图像在视觉上更加自然可信。

此外,关于神经网络在CT图像条状伪影抑制方面也有一些相关的研究。低剂量图像中的噪声通常是局部的,但稀疏投影采样造成的条状伪影是全局的,因此在构建网络时需要采用更大的感受野(receptive field)。Han Y等人^[10]提出了一种基于U-Net的去条状伪影算法,和其他的基于U-Net的去伪影算法不同,作者基于Hankel矩阵理论从原理上证明了经典U-Net方法在处理条状伪影时的不足,并给出了具体的改进策略,提出了dual-frame U-Net和tight-frame U-Net。实验结果显示,作者提出的两种改进网络的伪影抑制效果明显优于经典的U-Net网络,解剖结构细节保留更加完整。

在某些情况下,由于物理、机械等条件的限制,只能获取到一定角度范围内的CT投影数据。传统的解析重建方法和迭代重建方法在处理这类数据不全的问题时,重建得到的图像通常包含严重的伪影和模糊。为了解决这一问题,Anirudh R等人^[11]提出了一种基于深度学习的有限角度CT图像去伪影算法(CT-Net)。其基本思想是

在CT-Net的训练过程中直接学习从不完整弦图(sinogram)到CT图像的映射,损失函数结合了图像域 L_2 范数和GAN,保证了增强后的图像具有较高的信噪比和丰富的细节信息。在应用过程中,首先利用CT-Net得到增强后的CT图像,然后利用该图像补全缺失的弦图,最后采用解析或迭代重建方法利用补全后的弦图重建出最终的图像。在实验中,作者仅采集了 90° 的弦图数据,利用CT-Net依然可以重建出质量较高的图像,而直接用传统的解析或迭代重建方法无法获得清晰的重建结果。

2.3.2 PET图像质量增强

由于PET成像需要事先向患者体内注射放射性示踪剂(如 ^{18}F -FDG),为了降低患者接受的辐射剂量,临床上对低剂量PET成像有很高的需求,然而剂量的降低会造成图像噪声的增加和对比度的下降,影响疾病的临床诊断。针对这一问题,Xu J等人^[12]提出了一种基于残差编解码器(residual encoder-decoder)的PET图像增强方法。和传统的非局部均值(non-local means, NLM)、块匹配三维滤波(block-matching and 3D filtering, BM3D)等方法相比,作者提出的方法可以在0.5%正常剂量的情况下得到高质量的PET增强结果,同时,处理一张2D PET图像的时间仅为19 ms,远少于传统方法所需的处理时间。

2.3.3 MR图像质量增强

为了实现成像加速,通常会在k空间进行数据截断和填零,这会导致重建图像中存在Gibbs伪影。传统的MR图像去伪影方法通常基于k空间滤波,然而k空间滤波无法很好地区分伪影信号和有用信号,使得增强后的图像往往存在着过平滑、细节丢失等问题。为了解决这一难题,东软医疗提出了一种基于多任务学习(multi-task learning, MTL)的MR图像增强方法^[13],该方法基于U-Net和ResNet网络结构,可以实现Gibbs伪影抑制。图2为该方法与传统的k空间滤波方法的对比实验结果,其中图2(a)为经过不同强度的k空间滤波得到的增强图像,可以看出,基于k空间滤波的伪影抑制是以图像分辨率为代价的,图2(b)为深度网络对Gibbs伪影去除的结果,图2(c)为满采结果。从这些结果中可以看出,基于MTL的MR图像增强方法能够在保护图像分辨率的情况下有效抑制Gibbs伪影。

2.4 医学成像智能化 workflow

进行临床医学影像(如CT、MRI等)扫描需要繁琐的准备工作:扫描医师首先需要确认患者的身份信息,并陪同患者进入扫描间;然后指导患者进行扫描前的准备工作(如摘掉随身携带的金属物等),



图2 对比实验结果

并进行手动摆位。在正式开始扫描之前还有一系列定位的流程：扫描技师首先需要采集一组患者的定位图像，然后在定位图像上手动设置成像参数和确定扫描视野（field of view, FOV）。上述这些繁琐的、重复性的工作会带来如下几个问题：对于大型医院而言，每天就诊的患者数量非常大，扫描医师一直处于高强度的工作状态，容易产生各种误操作，从而影响扫描图像的质量以及后续疾病诊断的效果；对于基层医院而言，由于扫描医师的经验相对不足、技能水平参差不齐，难以保证获取到患者高质量的医学图像，从而影响疾病的诊断；由于不同医院、不同医师的经验与习惯存在着明显的个体差异，因此扫描得到的医学图像很难具有良好的一致性，为远程会诊、分级诊疗带来了难以解决的困难，同时也给后续基于AI的疾病辅助检测与诊断带来了更大的挑战。

近年来，AI技术的快速发展使得智能化影像扫描 workflow 逐渐成为可能。智能化扫描 workflow 涵盖了患者身份智能认证、智能语音交互、智能患者摆位、智能化扫描参数设定等功能，贯穿影像扫描的整个流程，其目的在于显著地降低扫描医师的重复工作，提高医院患者的流通量，并提升患者的就医体验，同时使不同医师采集到的影像数据具有更好的一致性。

目前，学术界在智能化 workflow 领域的研究工作较少，现有工作主要集中在智能化扫描定位方面，其中快速精准的人体解剖结构全自动定位是实现其功能的核心所在。Kelm B M等人^[14]提出了一种称为边缘空间学习（marginal space learning, MSL）的人体解剖结构自动定位方法，将解剖结构定位建模为在医学图像中对特定解剖结构的搜索过程。其搜索空间（包括位置、尺寸、角度等维度）巨大，导致穷举搜索方法带来的时间消耗是不可接受的。

而MSL的原理是在搜索过程中对不可能的情况进行提前剪枝，从而避免了大量无用的搜索，其有效搜索空间仅是完整搜索空间的很小部分，因此称为边缘空间学习。MSL的应用范围很广，可以实现对不同人体解剖结构的快速定位。该参考文献介绍了MSL用于MR图像脊柱自动定位的实验，结果表明，CPU版本的MSL算法可以在平均11.5 s的时间内检测到所有的腰椎间盘突出，灵敏度达到98.64%，每个个体数据的平均假阳性率仅为0.073 1，具有良好的临床应用价值。

除了组织器官的自动定位外，关键点（landmark）的自动定位在智能化扫描 workflow 中也十分重要。现有的大部分方法首先学习一个结构与纹理的特征模型，然后基于该模型在图像中搜索感兴趣的关键点，通常这些特征模型是基于图像局部信息计算的，容易陷入局部极值中。为了解决上述问题，Ghesu F C等人^[15]提出了一种新颖的关键点定位方法，该方法将关键点的特征建模过程和搜索过程看作一个统一的过程来处理。具体来说，该方法利用深度学习实现多层次的图像特征提取，并利用增强学习（reinforcement learning, RL）方法实现高效的搜索，同时使用深层神经网络将二者结合在一起，实现了端到端的学习过程，有效地提升了算法的整体检测效果。该参考文献分别在二维MR图像、二维超声图像和三维CT图像上进行了算法测试，实验结果表明，该算法在精度和速度上远优于现有的关键点检测算法，平均误差为1~2个像素，当关键点不存在时，该算法也能够自动地给出相应的提示，具有较广的应用范围与良好的实用价值。

针对三维CT和MR图像，Zhang P等人^[16]提出了一种细粒度人体区域自动识别方法。相比于计算机视觉领域，医学影像

领域的有标签数据是相对较少的,为了解决网络训练过拟合的问题,通常可以采用迁移学习(transfer learning)的方法。然而自然图像和医学图像存在着较大的差异,因此基于自然图像的迁移学习在很多情况下无法获得最优的效果。该参考文献提出的方法的创新之处在于,设计了一种无标签自监督(self-supervised)的网络迁移学习方法,这样就可以利用CT或MR图像本身进行自学习,从而避免了不同领域图像差异较大带来的问题。实验结果表明,相比于从自然图像到医学图像的跨领域迁移学习,该参考文献提出的领域内无标签自监督迁移学习能够获得明显更优的识别效果。

在工业界,目前已经有与智能化扫描 workflow 相关的工作被报道了。德国西门子股份公司研发了智能化辅助扫描 workflow (fully assisting scanner technologies, FAST) 系统,该智能化 workflow 利用高精度 3D 相机实现了精准的患者自动摆位。具体来说,利用红外光技术,3D相机可以获取到患者身体的三维轮廓,从而计算出患者的体型等有用信息,基于这些信息实现等中心点自动定位、扫描范围自动设定等功能,从而有效地降低不必要的辐射剂量,并提高影像扫描的一致性。目前,FAST系

统已经应用于西门子商业化的CT设备中。

在MR智能化扫描方面,东软医疗研发了基于2D定位片的头部扫描FOV自动设定方法^[17]。该方法基于深度学习技术,可以实现针对冠状面、矢状面、横断面3个不同方向的4种自动定位模式。相比于现有的基于3D定位片的方法,2D定位片的获取速度更快、成本更低,同时也更符合医生的临床操作习惯。实验结果表明,该算法可以在0.6 s内完成特定方向的FOV自动设定,定位结果具有良好的准确性和一致性。图3为基于2D定位片的头部MR扫描FOV自动设定方法的示意。

总体来说,医学影像智能化扫描 workflow 领域的研发尚处于起步阶段,仅在一些点上实现了突破与创新,整个影像扫描链条尚未完全打通,未来需要继续开展大量具有临床价值的创新性研究,从而提升患者的诊疗效果与就医体验,减轻扫描医师繁重的重复性工作负担。

3 医学影像大数据与智能图像分析

医学影像设备的发展和技术进步为医生提供了更多的影像信息,这些信息被用于对疾病进行诊断和鉴别。但是对

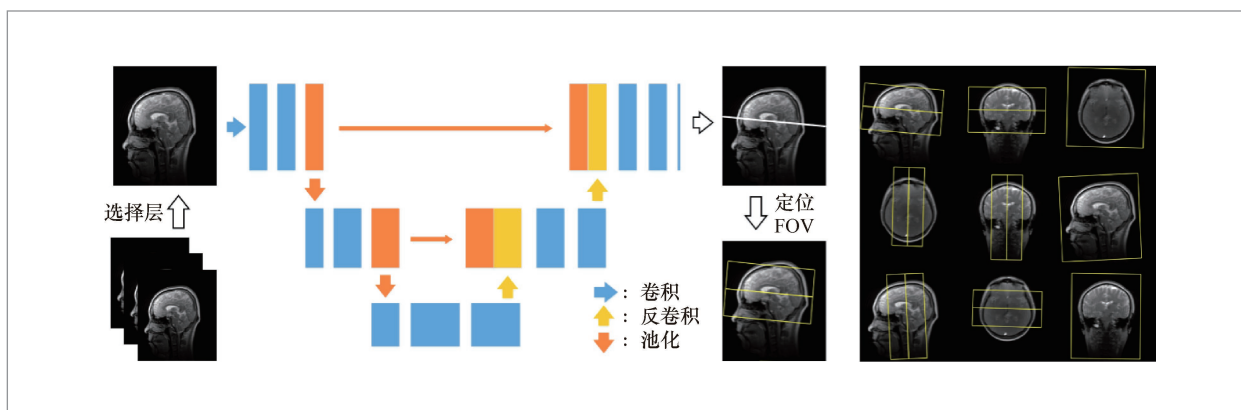


图3 基于2D定位片的头部MR扫描FOV自动设定方法的示意

大量图像信息的整合和解读也提高了对医生医学影像知识水平的要求,并且也增加了医生的阅片时间。为此,使用计算机辅助诊断技术对医学影像信息进行进一步的智能化分析挖掘,以辅助医生解读医学影像,成为现代医学影像技术发展的重要需求。

近年来,人工智能成为计算机辅助诊断的研究热点。在过去10年时间里,关于人工智能计算机辅助诊断的研究增加了近10倍。其中,影像组学和深度学习算法在现阶段更是被广泛研究和应用,被应用在医学图像的分类、检测、分割和配准等任务中。相关研究在X射线、CT、PET和MRI等领域取得了重大突破。

3.1 影像组学的概念及其在医学影像中的应用

影像组学起源于肿瘤学领域,最早是由荷兰学者Lambin P等人^[18]于2012年正式提出的,即高通量地提取大量描述肿瘤特性的影像特征。同年,Kumar V等人^[19]进一步对概念进行了完善,即影像组学是高通量地从MRI、PET及CT影像中提取大量高维的定量影像特征,并进行分析。影像组学将传统的医学影像转化为可挖掘的高通量影像特征,用于定量描述影像中的空间时间异质性,揭示出肉眼无法识别的图像特征,有效地将医学影像转换为高维的可识别的特征空间,并对生成的特征空间进行统计学分析,从而建立具有诊断、预后或预测价值的模型,为个性化诊疗提供有价值的信息。近年来,该领域成为研究热点^[20-26],以Radiomics为关键词从Web of Science数据库中检索近10年的相关SCI论文发表情况,发现自2012年正式提出影像组学概念以来,2013年关于影像组学的论文只有7篇。而2018年1月至11日,

影像组学论文发表量已经高达600余篇(如图4所示)。

3.1.1 影像组学实现过程

如图5所示,影像组学的主要流程包括图像获取和标注、感兴趣区图像分割、影像组学特征的提取、特征值选择和降维、预测模型的训练和性能评估5个步骤。

第一步,图像获取和标注。影像组学大数据要求病人数据临床问题明确、格式规范、信息完整,而目前常规临床使用的影像,由于采集时的成像参数、病人体位、重建算法以及扫描仪器的不同,具有很大的差异性,因此识别和整理大量具有相似临床参数的图像数据示例(如疾病分期)也是一项挑战。同样,对于前瞻性研究而言,制定统一的数据规范也是实现影像组学研究的关键^[27]。

第二步,感兴趣区图像分割。将感兴趣区域(如肿瘤等)在影像上分割出来是实现后续特征提取和信息分析的基础。通常以专家的手动分割结果作为标准,但是由于评价者自身和评价者之间的差异性^[28-29]以及分割工作本身的繁杂性,使手动分割不可能满足影像组学的要求。而全自动或半自动的分割方法会降低由于人的参与造成的差异性,使分割结果的可重复性更高。传统应用于医学影像的自动和半自动分割方法(如阈值分割法、边缘检测法、区域生长法、水平集法、模糊集法、活动轮廓模型法、图谱引导法)在实际中分割的精度和速度很难满足要求,近年来使用深度卷积神经网络的分割方法成为研究的热点。

第三步,影像组学特征的提取。一旦确定了感兴趣区域,就可以提取高通量的影像组学特征。影像组学特征可以分为两大类:一类是使用数学公式定量计算的感兴趣区特征,包括形状特征、灰度特征和

纹理特征以及小波变换、高斯变换后的特征；另一类是使用神经网络提取的图像的深度特征，但对于深度特征的具体物理意义没有直接的对照。

第四步，特征值选择和降维。初步提取的图像特征一般是数以千计甚至万计的数据。为了选出可重复性好、信息量大、无冗余的特征用于最终模型的建立，一般需要对高通量的特征进行降维处理^[30]。常见的特征降维方法如下。

- 方差分析。将数据列变化非常小(即列包含的信息量少)的特征直接滤除。
- 相关性度量。数值列之间的相似性通过计算相关系数来表示，名词类列的相关系数通过计算皮尔逊卡方值来表示，相关系数大于某个阈值的两列只保留一列。
- 组合决策树方法。对目标属性产生许多巨大的树，然后根据对每个属性的统计结果找到信息量最大的特征子集。
- 主成分分析法。将原始的 n 维数据集通过正交变换转换成不相关的被称作主成分的数据集，变换后方差最大的特

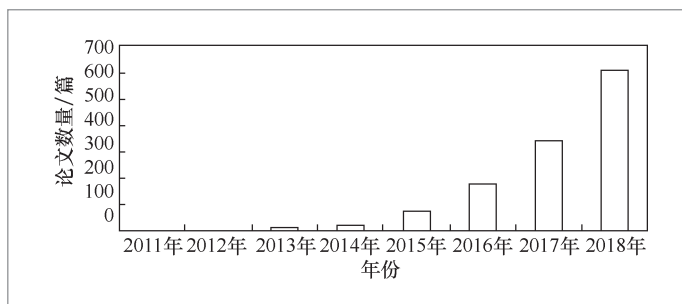


图 4 2011年至2018年11月 Web of Science 数据库关于影像组学的论文发表量统计

征即第一个主成分，其后的成分在与前述主成分正交条件限制下具有最大方差，保存前 m ($m < n$) 个主成分就能保存最大的信息量。

- 特征一致性度量。由手动分割结果计算得来的特征值，需要对其进行再测信度(test-retest)检验，计算一致性相关系数，将一致性相关系数小于某个值的顽固性低的特征滤除等。

第五步，预测模型的训练和性能评估。依据分析的类型(即有无预测标签)可以将分析方法分为监督学习和无监督学习。

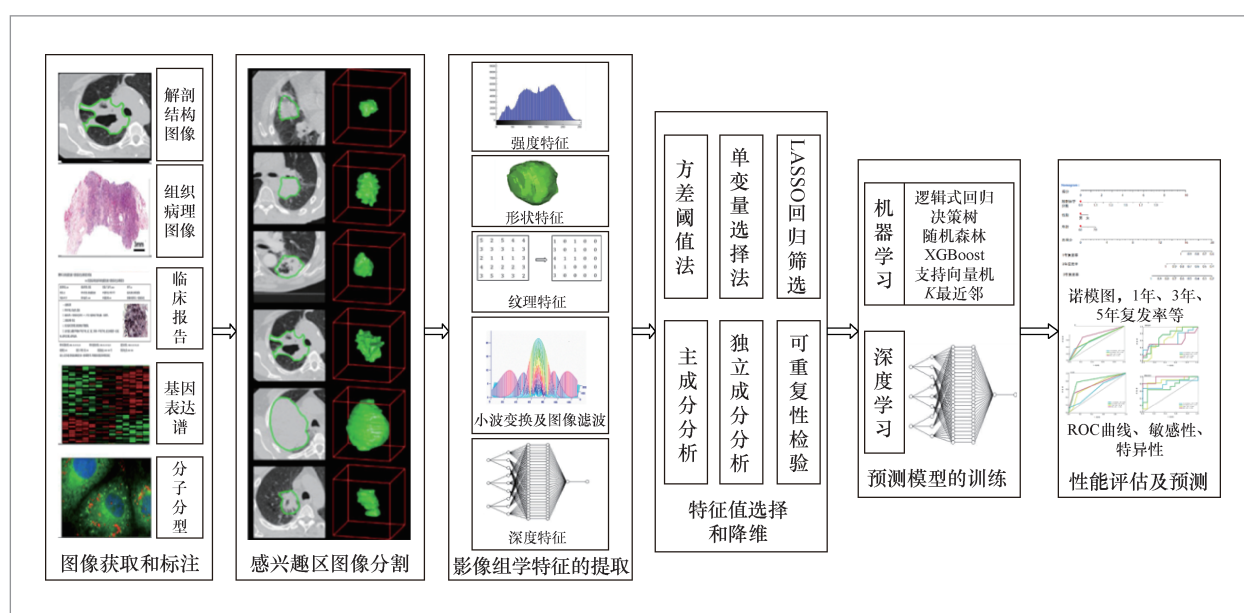


图 5 影像组学分析流程

聚类是最常用的无监督学习方法^[31], 分层聚类是最常用的聚类算法。监督学习常用的方法有感知机、 K 最近邻法、决策树、线性回归、支持向量机、随机森林、神经网络和深度学习等。影像组学分析的最终目的是使建立的模型不仅对现有的数据有很好的预测能力, 而且对未知的数据也有很好的预测能力, 这就要求选择适当的模型使测试误差最小。常用的模型选择方法有正则化和交叉验证。依据分析目的(即标签类型), 也可以将监督学习分为分类预测和回归预测。相应的模型性能评价指标一般采用的是分类准确率和损失函数。

影像组学的核心是将低维视觉特征、高维复杂特征和临床经验特征结合, 全面分析感兴趣区异质性, 寻找影像组学特征与感兴趣区的表现特征、分子标志物之间的联系。这就需要在影像组学研究中收集病人的临床信息、基因分子表达信息等。并且为了使最终建立的模型可以推广应用, 就必须考虑模型的泛化能力, 最好的解决方法就是进行多中心验证。实现这些过程需要建立可以规范化管理和高效利用的数据平台, 以用于影像、临床信息和基因表达状态等数据的存储、检索和分析, 同时将影像组学特征提取、选择和模型训练等方法有效地整合到数据平台中, 一体化地实现影像组学的应用^[32-34]。

3.1.2 影像组学在医学影像分析中的应用

作为医学领域一种新兴的研究方法, 影像组学利用信息挖掘等信息技术, 通过从不同模态的影像中提取定量的高通量影像特征, 在一定程度上实现了感兴趣区域异质性的诊疗和预后评估^[22]。例如Aerts H J等人^[20]于2014年回顾分析了1 019例肺癌和头颈癌患者的CT影像, 利用影像组学分析方法, 非侵入式地分析了影像特征与

临床分型、基因表达图谱的关联性, 揭示了影像特征与基因表达的潜在关系, 提出了一种可以量化和监控治疗期间肿瘤表型改变的方法, 引发了国内外影像组学研究的热潮。Cui Y等人^[35]回顾分析了两个研究中心79例胶质母细胞瘤患者的资料信息, 从T1增强序列和磁共振成像液体衰减反转恢复序列(FLAIR)两个模态的MR影像中提取了多区域的影像特征, 将影像特征与患者总生存期进行了关联性预测分析, 采用多参数的LASSO回归, 构建了患者总生存期预测模型, 提供了一种肿瘤内部子区域分割方法, 验证了影像组学可以为患者提供具有生存期预测价值的信息。Huang Y Q等人^[36]回顾分析了500余例进行结直肠癌手术的患者资料, 利用影像组学方法, 对影像特征和临床病理特征(血清标记物和临床指标)进行关联性分析, 构建了结直肠癌淋巴结转移术前预测模型, 与传统CT影像学评估相比, 影像组学预测模型术前淋巴结预测准确率提高了14.8%, 为医生进行结直肠癌的术前决策提供了重要参考。目前, 影像组学被用于多种疾病的良恶性判定、生存期预测、生物分子标志物状态及淋巴结转移风险等^[37-50], 为医生的诊断、治疗决策、预后管理等提供了具有参考价值的预测模型, 具有重要的临床价值和应用前景。

3.2 深度学习介绍及其在医学影像中的应用

近年来, 随着算法、算力和数据量的不断提升, 在计算机视觉中, 深度卷积神经网络已成为首选技术。深度卷积神经网络的优势在于它能够自动学习重要的低级特征(如线条和边缘), 并且能够从低级特征中迭代地提取更复杂和更高级的特征(如形状等), 其端到端的设计给模型提供了更

多可以根据数据自动调节的空间,增加了模型的整体契合度。基于卷积神经网络的深度学习方法在2012年的全球计算机视觉竞赛ImageNet Classification中取得了压倒性的胜利,引起了广泛关注。医学图像分析也是深度学习的一个活跃的研究领域,笔者回顾了近年来深度学习在医学领域中的应用,统计了自2012年起相关的会议和期刊论文发表情况。笔者发现,2012—2014年结合深度学习的医学图像分析论文总数不足30篇,而相关研究在2015年后迅速增长,截至2018年11月,相关论文发表量已经高达350余篇(如图6所示)。

深度学习一般包括监督学习模式和无监督学习模式。

监督学习模式中的卷积神经网络是目前医学图像分析中研究最多的机器学习算法^[51]。其主要原因是CNN在过滤输入图像时保留了空间关系。空间关系在放射学中至关重要,例如,骨骼边缘与肌肉、正常组织与癌组织的连接等。循环神经网络(recurrent neural network, RNN)可以利用内部记忆处理任意时序的输入序列^[52-53],常被用于时间性分布数据。长短期记忆(long short-term memory,

LSTM)网络是RNN的一种特殊类型,它的出现解决了RNN在实际训练过程中出现的梯度消失问题。在医学图像分析领域,RNN主要用于分割任务。

无监督学习模式中的自动编码器以无监督的方式学习编码,无需训练标签,减少了模型的维度和复杂性,同时通过重建输出,可以生成与输入数据类似的新数据,用以解决医学图像分析中标记数据稀缺的问题^[54]。受限玻尔兹曼机(restricted boltzmann machine, RBM)^[54-56]是由可见层和隐藏层组成的浅层神经网络,层间彼此连接,但层内无连接。RBM使用输入数据的反向传递生成重建,并估计原始输入的概率分布。RBM被用于降维、分类、回归、协同过滤、特征学习等,是组成深度置信网络(deep belief network, DBN)的基础部件。GAN^[57]通过生成模型和判别模型的互相博弈,学习并产生更好的输出,是近年来复杂分布上无监督学习颇具前景的方法之一。GAN及其扩展已被用来解决很多具有挑战性的医学图像分析问题,如医学图像去噪、重建、分割、检测或分类。此外,GAN合成图像的能力也被用于解决医学领域中标记数据的长

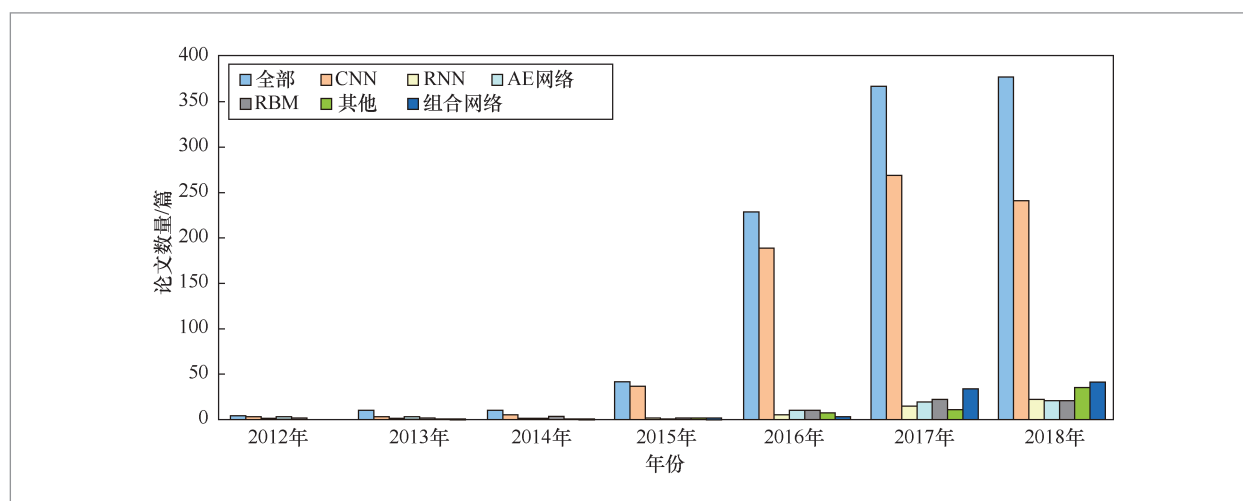


图6 2012年至2018年11月结合深度学习的医学图像分析论文发表数量统计

期稀缺性的问题。深度学习在医学影像领域的应用主要包括分类、检测、分割和配准。

3.2.1 分类

分类主要涉及病变良恶性分类和多种疾病的鉴别。Lo S C B等人^[58]早在1995年就通过CNN检测到了胸片中的肺结节，他们使用了55个胸片和包含两层隐藏层的CNN判断区域是否有肺结节。随后，Rajkomar A等人^[59]将胸片训练样本增加到150 000个，输入GoogLeNet^[60]中微调(fine-tune)，他们将图像的方向分为正面视图和侧面视图，准确率接近100%。肺炎是世界范围内常见的健康问题。Rajpurkar P等人^[61]采用改良的DenseNet^[62]，使用来自ChestXray14的112 000张图像^[63]数据集，对胸片上看到的14种不同疾病进行分类。要实现该任务并且得到高精度的预测模型，通常需要大量带标签同时标注异常位置的图像，然而这样的数据获取起来十分昂贵，尤其是具有位置注释的数据。Li Z等人^[64]通过弱监督方法设计了统一模型，

对只含少量位置注释的图像进行训练。在训练阶段，使用多实例学习的方法学习两种类型的图像；在测试阶段，模型同时预测出疾病的类别和定位，如图7所示。Shen W等人^[65]基于LIDC-IDRI数据集中1 010位患者的带标记的CT肺部图像，将CNN与支持向量机(support vector machine, SVM)和随机森林(random forest, RF)分类器结合使用，将肺结节分类为良性或恶性，该方法对肺结节分类的准确率为86%，同时，还发现该模型对不同级别的噪声输入具有很强的顽健性。除肺部病变外，还有许多其他应用。Esteva A等人^[66]使用皮肤病照片和皮肤镜图像训练的皮肤病诊断模型准确率已优于人类皮肤科医生。Pratt H等人^[67]在约90 000个眼底图像上训练CNN模型，对糖尿病性视网膜病变(DR)严重程度进行分类，准确率为75%。

无监督学习方法也是一个活跃的研究领域。Suk H I等人^[68]将fMRI图像分类，得到健康或轻度认知障碍的诊断，使用RBM的堆叠结构学习不同大脑区域之间的分层功能关系。

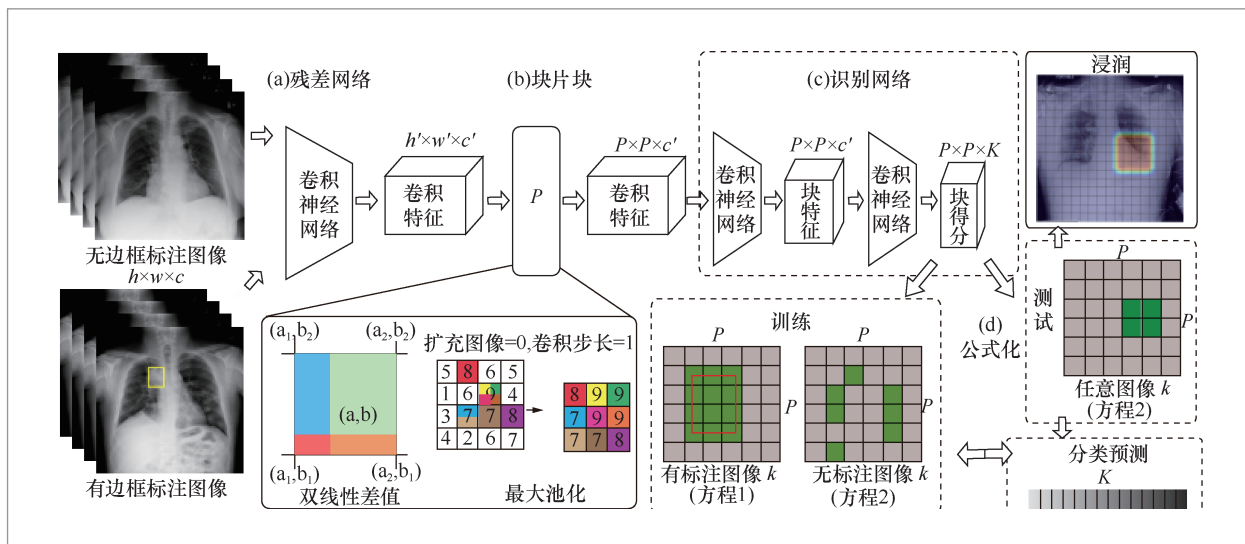


图7 弱监督胸部疾病定位与分类流程

3.2.2 检测

计算机辅助检测 (computer-aided detection, CADe) 是一个重要的研究领域,因为在检查中遗漏病变会对患者和临床医生产生严重后果。CADe的目标是在图像中定位异常或可疑区域,从而提醒临床医生。CADe旨在提高患病区域的检出率,同时降低假阴性率。Shin H C等人^[69]评估了5种CNN结构,用于CT扫描检测胸腹部淋巴结和间质性肺病。检测淋巴结很重要,因为它们可能是感染或癌症的标志物。他们使用GoogLeNet获得了纵隔淋巴结检测模型,AUC (area under curve) 评分为0.95,灵敏度为85%。Becker A S等人^[70]通过训练模型从乳房钼靶图像中检测乳腺癌,AUC评分为0.82,与经验丰富的放射科医师相当。Wang X等人^[71]采用T2加权MR图像训练CNN模型,用以检测前列腺癌,AUC评分为0.84,明显高于传统的机器学习方法(如基于尺度不变特征变换特征的支持向量机模型,AUC为0.70)。

DeepLesion^[72]是迄今全球规模最大的多类别、病灶级别标注的开放获取临床

医疗图像数据集,含有32 735个带标记的病灶实例,包括来自全身各个部位的关键影像学发现,比如肺结节、肝肿瘤、淋巴结肿大等。Yan K等人^[72]基于DeepLesion数据集,开发了一种通用的病变检测器,为帮助放射科医生找到患者身上所有类型的病灶提供了技术可能,如图8所示。通用病灶检测的难度远高于特定病灶检测,DeepLesion中包含肺、肝、肾、淋巴、胰腺、骨骼、软组织等各种病灶,病灶类内差异大,类间差异小(肺、肝的病灶相对容易检测一些,而一些腹腔中的病灶与周围正常组织差异较小)。为了改进病灶检测的精度,Yan K等人^[73]又提出了一种利用3D信息的检测算法,将病灶识别准确率提高到了84.37%。

组织病理学图像目前也越来越数字化,Ciresan D C等人^[74]使用11~13层CNN识别来自MITOS数据集的50个乳房组织学图像中的有丝分裂图。他们的方法分别达到了88%的精确度和70%的召回率。Yang X L等人使用5~7层CNN将肾癌组织病理学图像分类为肿瘤或非肿瘤,达到97%~98%的准确度。Sirinukunwattana K等人^[75]使用CNN检测100个结肠直肠腺癌

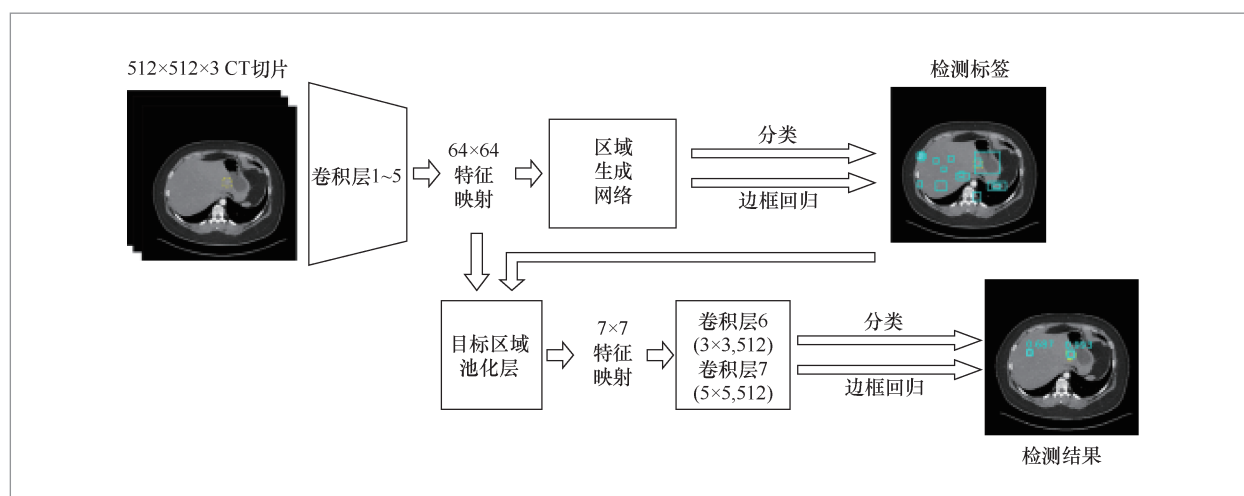


图8 基于 DeepLesion 构建的通用病灶检测流程

组织学图像中的细胞核。

3.2.3 分割

CT和MRI的图像分割研究涵盖了肝脏、前列腺和膝关节软骨等多种器官,但大量工作主要集中在脑部图像分割,如肿瘤分割。肿瘤分割在外科手术计划中尤其重要,可确定肿瘤的确切边界,指导手术切除。Moeskops P等人^[76]通过集成3个CNN模型,将22个早产儿和35个成人的MRI脑图像分类和分割成不同的组织类别,如白质、灰质和脑脊液,该算法Dice系数在0.82和0.87之间。大多分割研究是关于二维图像切片的,但Milleterai F等人^[77]应用三维CNN分割了来自PROMISE2012挑战数据集的MRI前列腺图像。受到参考文献[6]中U-Net架构的启发^[78],他们提出了V-Net,并在MRI前列腺扫描中进行了训练,Dice系数为0.869。Stollenga M F等人^[79]使用3D LSTM-RNN在6个方向上对脑部MR图像进行了分割,用金字塔方式重新排列了MD-LSTM中传统的长方体计算顺序,

在2015年MRBrainS挑战赛中取得了很好的分割结果。Andermatt S等人^[80]使用带有门控单元的3D RNN分割脑MR图像中的灰质和白质,结合数据预处理和后处理操作,进一步提高了分割准确率。Singh V K等人^[81]提出了一种基于条件生成对抗网络(conditional generative adversarial network, cGAN)的乳腺肿块分割方法,生成网络不断学习肿瘤的内在特征,对抗网络不断进行强制分割,该方法在乳腺钼靶数据库(digital database for screening mammography, DDSM)公开数据集和内部数据集中提取的数十个恶性肿瘤上进行了验证,获得了0.94的Dice系数和0.89的Jaccard指数,如图9所示。

3.2.4 配准

图像配准用于神经外科手术或脊柱外科手术,以定位肿瘤或脊柱骨界标,便于手术切除肿瘤或植入脊柱螺钉。Yang X等人^[82]使用来自开放获取系列影像研究(open access series of imaging studies,

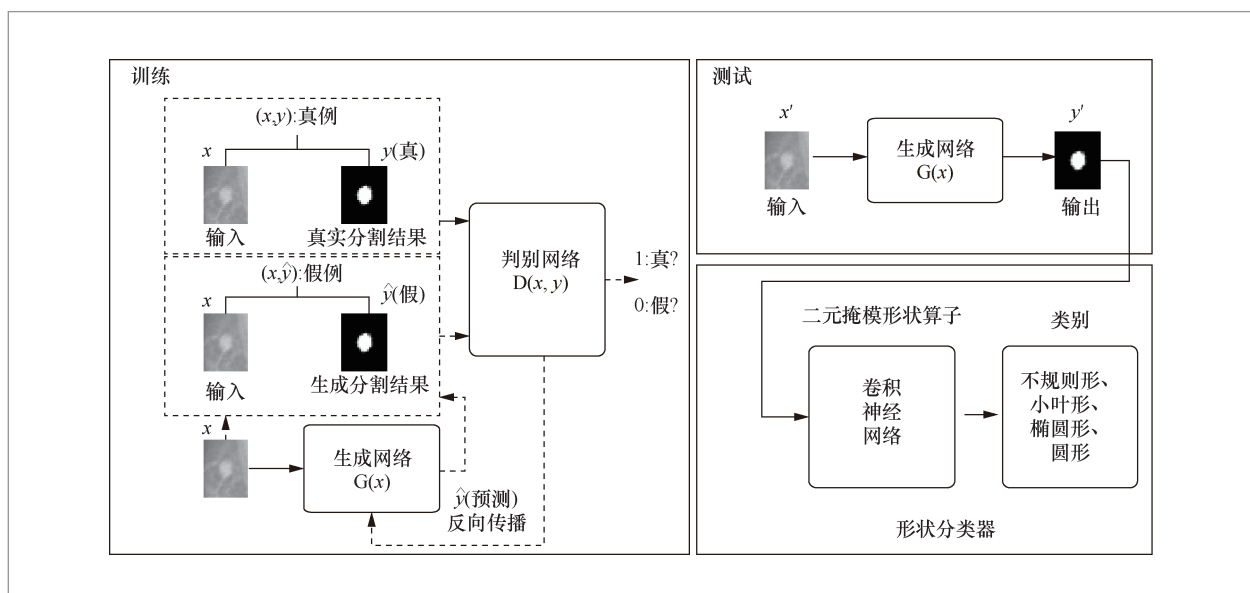


图9 基于cGAN的乳腺肿块分割和形态分类流程

OASIS) 数据集的MRI脑部扫描,以编码器-解码器方式堆叠卷积层,以预测输入像素将如何变形为最终像素。他们引用了高度形变微分同胚度量映射 (large deformation diffeomorphic metric mapping, LDDMM) 模型进行配准,同时在计算时间方面也取得了显著的进步。Miao S等人^[83]在合成X射线图像上训练5层CNN,以便将膝关节植入物、手部植入物和经食道探针的三维模型配准到二维X射线图像上,以便估计他们的姿势。Yan P等人^[84]提出了对抗图像配准网络 (adversarial image registration, AIR-net) 配准框架,应用于MR和经直肠超声 (TRUS) 图像融合配准,训练生成器和判别器两个深度神经网络,不仅可以获得用于图像配准的网络,还获得可以帮助评估图像配准质量的度量网络,如图10所示。

人工智能在医学影像应用中也有着一定的挑战。在医学图像分析中,长期缺乏可公开的训练数据,高质量的标记数据更加稀缺。本文中提供的大多数数据集涉及的患者不到100人。然而,尽管训练数据不足,但本文中的参考文献在各项任务中的表现相对令人满意。训练集中的数据或类不平衡也是医学图像分析中的重要问题^[85]。尽管存在过度拟合的风险,但通过使用数据增广

来生成更多稀有或异常数据的训练图像,可以改善这种数据不平衡效应。

同时,还有一项非技术性的挑战是人类对人工智能分析结果的接受程度。机器学习算法在图像识别任务中已超过人类表现,并且它们在医学图像分析中也可能比人类表现更好。然而,当患者被误诊、因AI或AI辅助医疗管理而患病时,就会出现有关法律和道德责任的问题。由于深度学习现在还不具有可解释性,人们无法完全解释黑盒算法的内部工作机制,因此,更加突出了这方面问题的严重性。

4 医学影像与自然语言文本处理的结合分析

4.1 自然语言文本处理在医学影像分析中的作用

神经网络模型在图像视觉识别中表现卓越。主流的、应用效果较好的深度神经网络计算机视觉模型大多基于有监督的训练过程,而有监督的训练过程依赖于大规模高质量的有标注数据集。现有的大规模图像数据集 (如ImageNet) 通过AMT (Amazon mechanical turk) 等大量人工

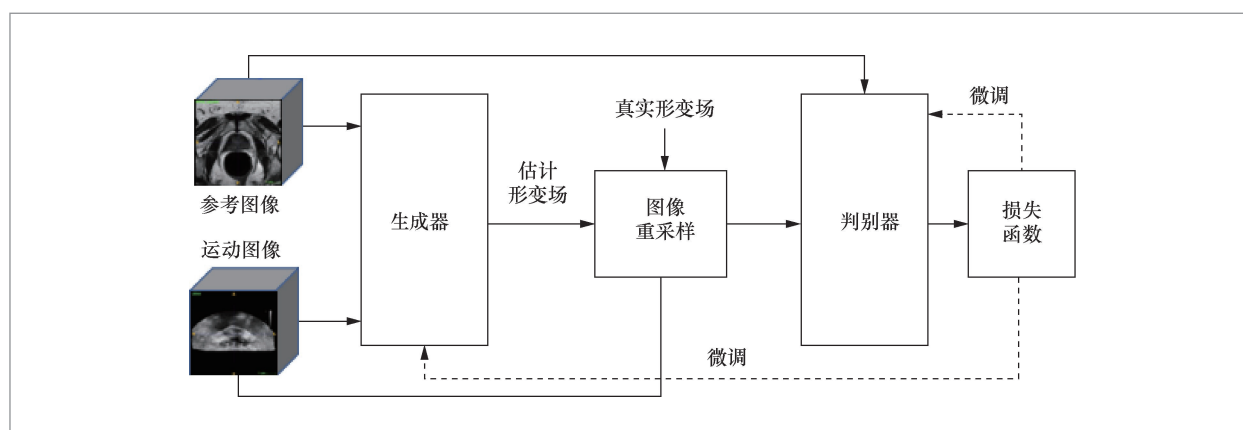


图10 基于AIR-net的MR和TRUS图像配准流程

操作完成图像标注。在医疗影像的辅助诊疗等应用场景下,众多的研究机构和公司投入了大量的资源用于构建更大规模的有标注的医疗影像数据集。

然而,在医疗影像研究方面,通过人工方式对医疗影像进行标注构建训练数据集具有非常大的挑战性,标注者需要具备相当程度的医学专业知识,不能像普通的图像标注任务一样进行众包。另外,医疗影像的标注工作要求更加细致,病灶的人工识别过程往往需要仔细阅片,并尽可能地发现和准确标注微小病灶,标注速度慢,标注成本高,因此很多情况下标注的准确度往往不令人满意。

现代医院已经逐步完成了信息化改造,并逐步向更加标准化、数字化的方向发展,医疗影像归档与通信系统(picture archiving and communication system, PACS)中不仅包含大量的医疗影像,也包含与其相关的各种文本报告,这些文本报告中包含相当多的专业诊断信息。使用自然语言处理技术从这些文本中提取有价值的标签,对相关的影像进行标注,是自动化生成有标注影像数据集的一种有效途径。

4.2 医学影像领域结合自然语言文本处理的研究进展

Schleg T等人^[86]最先发表了“用文本报告代替人工标注医疗影像”的研究。在研究中收集使用了157例频域光学相干断层扫描技术(spectral-domain optical coherence tomography, SD-OCT)的视网膜图像和相关的文本报告,通过自然语言处理技术挖掘文本报告中与病理相关的物体信息,对图像进行“某(病理相关的)物体是否出现”的标注,在此基础上训练CNN模型,训练得到的模型可根据视网膜影像预测“视网膜内囊样液体

(intraretinal cystoid fluid, IRC)”“视网膜下液体(subretinal fluid, SRF)”“视网膜正常”3种结果。在具体实验中, Schleg T等人使用SVM对报告文本进行解析,挖掘出“与病理相关的物体”和“位置”的对结构([obj, loc]),在此基础上,构建([物体1, 有/无], [物体2, 有/无], ...)的向量,作为标签对影像数据进行二分(出现或未出现)标注,并训练CNN模型。

Shin H C等人^[87]构建了包括78万例来自PACS的CT/MRI影像(含头、胸等多个身体部位)和文本报告的数据库,使用基于隐含狄利克雷分布(latent dirichlet allocation, LDA)主题模型的文档主题学习(document topic learning),分层次挖掘文本报告中的词汇,对影像进行标注,并训练深度CNN模型,训练的模型可以根据CT/MRI影像生成来自3个不同层级的词汇。Shin H C等人对文本报告的LDA挖掘是分层次的,包括顶级词汇(如“MRI”“胸”)、次级词汇(如“肿块”“增强”),目的是挖掘文本报告中尽可能多的“潜在主题”。在此基础上,使用ImageNet的预训练模型(pre-trained model),以主题词作为标签,对影像进行迁移学习,通过预训练模型进行微调,得到初始CNN模型。为了得到更好的“图像到文本”的输出,Shin H C等人使用了一系列自然语言处理技术:通过文本向量化(word to vector)建模和Skip-Gram建模去除词汇级别的歧义;利用疾病本体知识(disease-ontology)构建二元语法语言模型(bi-grams),进行影像与“病理文字描述”之间的关系挖掘匹配。主题词是分层次的,训练得到的CNN模型预测结果也是分层次的,最终得到的CNN模型可以根据影像生成3个不同层级的输出词汇,如图11所示。

Wang X S等人^[88]提出了参考文献[87]

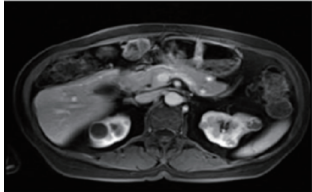
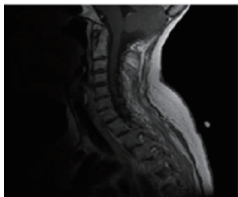
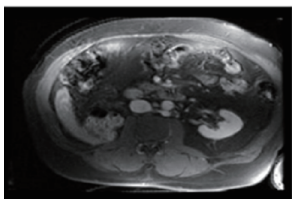

输入图像	输出文字	原始文字
	diameter(直径) mass(肿块) kidney(肾脏) 平均距离 0.33	"... and solid lobulated mass arises from the anterior lower pole of right kidney and measures 1.6 cm in diameter ..."
	spine(脊柱) chest(胸部) scoliosis(脊柱侧屈) 平均距离 0.42	"... it measures a few mm in diameter and is best appreciated on series 3 image 6 in the thoracic spine no definite enhancing lesions are present. in the lumbosacral spine ..."
	diameter(直径) lesion(病变) kidney(肾脏) 平均距离 0.33	"... 2 apparently cystic lesion in the retroperitoneum adjacent to the crus 3 liver lesions for example series 17 image 22 series 16 image 172 and image 137 the lateral lesion ..."
	adenopathy(腺病) masses(肿块) lung(肺部) 平均距离 0.21	"... dozens of masses of various sizes in or near right pleura and or peripheral lung without definite change ... by areas of right lung consolidation atelectasis and or confluent masses ..."

图 11 分层次的主题词输出

中的两个问题：类别极度不平衡（最多的类别包含113 037个图像，某些类别仅包含几十个图像）；类别不是“视觉关联”的，导致Shin2015训练出的CNN模型不像ImageNet的CNN模型那样适合用来迁移学习。

为了解决以上问题，Wang X S等人设计了LDPO框架（looped deep pseudo-task optimization framework），如图12所示，该框架的核心思想如下：

- 训练CNN模型的过程中，当聚合出未知的类别时，使用“伪标签”进行标注（而不是先从文本报告中挖掘标签标注）；
- 使用通用的ImageNet预训练CNN

模型和通过文本挖掘出主题标签（topic label）的CNN模型进行特征提取与编码，不断迭代优化；

- 当图像聚类相对稳定时，对每个聚类对应的文本报告分别进行语义标签挖掘。

Wang X S等人使用和Shin2015相同的数据集进行实验，得到了更优的语料标签输出和更适合迁移学习的模型。

Shin H C^[89]随后基于医学主题词表（medical subject heading, MeSH）标签进行了X光胸片的研究，此研究的关注点是通过CNN-RNN联合学习（joint learning）训练可以生成X光胸片影像描述（主题）的模型。Shin团队使用的数据

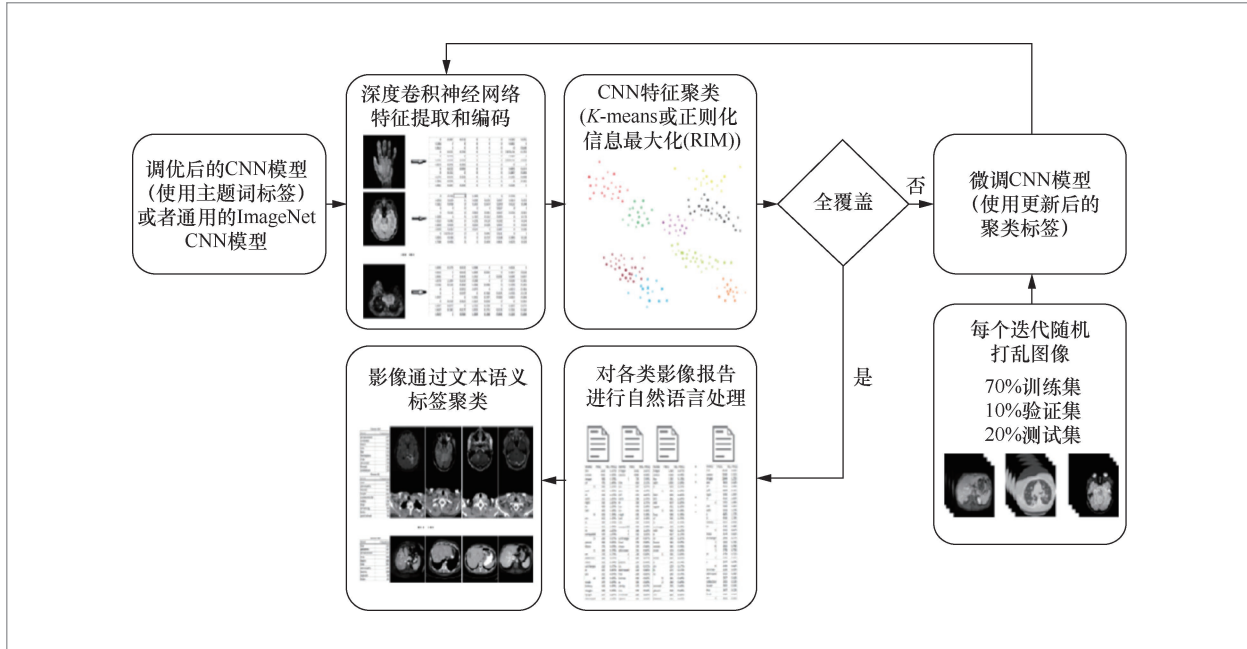


图 12 LDPO 框架

集是OpenI中的部分X光胸片影像(已有MeSH标注)和文本报告。MeSH是一种对影像数据的标注标准,数据集中的所有影像已经打好MeSH标签(如图13所示),包含17种疾病的标注模式(disease annotation pattern)。Shin首先对这些打好MeSH标签的影像数据进行迁移学

习,完成图像的初始编码,得到初始的CNN模型。

在此基础上,Shin团队使用迭代瀑布模型(recurrent cascade model)进行影像/文本上下文的关联,生成关联向量,并不断迭代优化,得到词汇序列标签。具体使用的RNN实现包括LSTM和门控递归单元(gated recurrent unit, GRU)。

最终得到的CNN模型可对图像进行预测,并将预测标签作为RNN模型的输入,生成5个词汇的描述,如图14所示。

Wang X S等人^[90]提出要建设大规模多标签胸片X光影像数据集ChestX-ray8,以解决医疗影像领域缺乏已标注数据库的难题。Wang X S的团队从PACS中采集了32 717个病人的脱敏数据,包含108 948张X光正片,并使用人机联合的方式(基于少量人工标注的弱监督学习方式)对疾病类型和位置进行了多标签标注,如图15所示。

使用自然语言处理技术挖掘文本报告中的8种常见病。具体步骤包括:使用

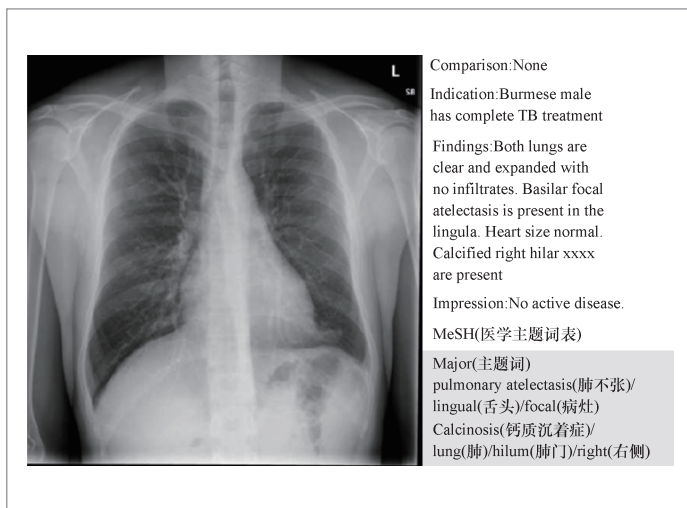


图 13 标注 MeSH 标签的 X 光胸片

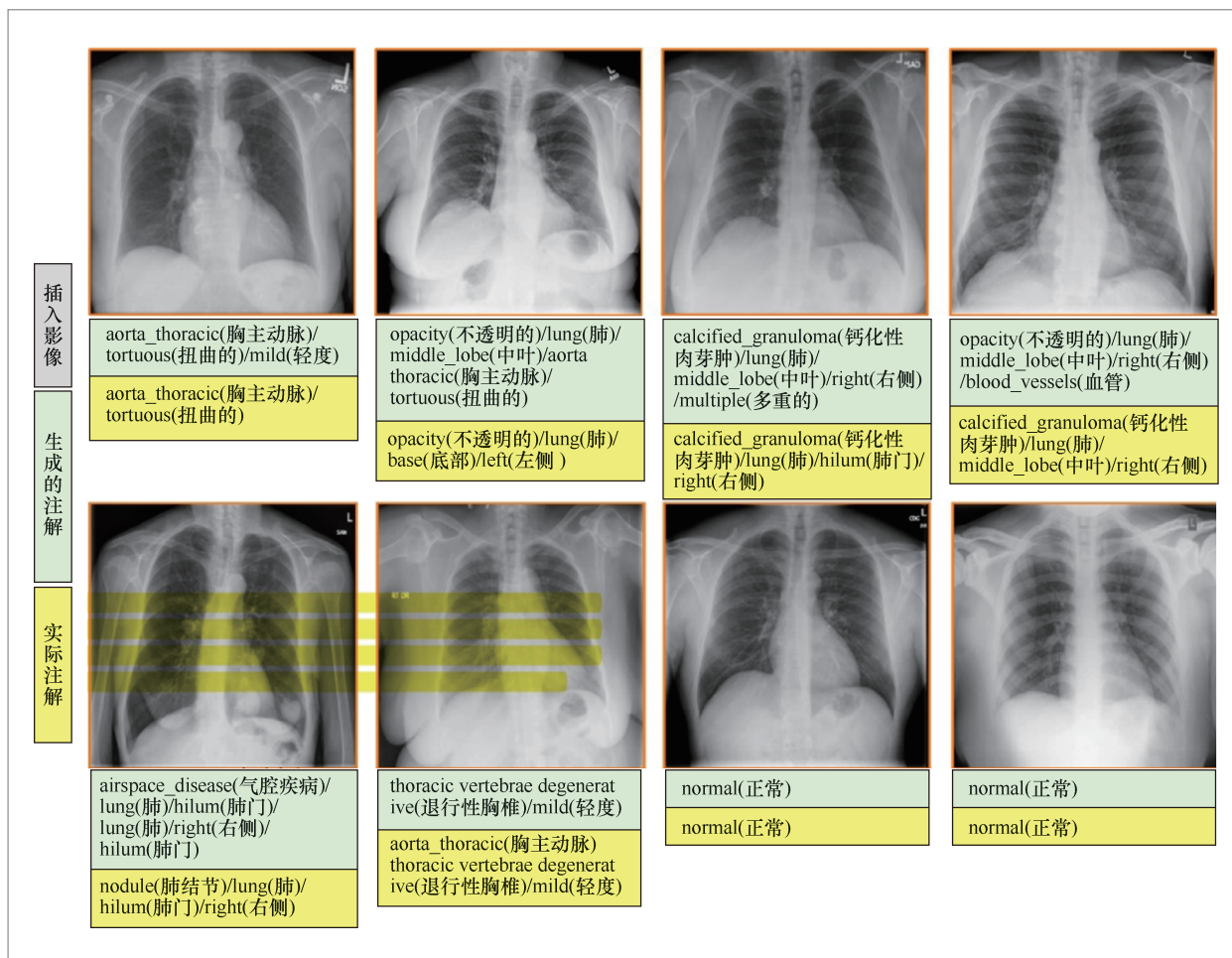


图 14 输出 5 个词汇的标签描述

DNorm和MetaMap挖掘文本中的病理关键词（疾病名称和疾病相关的实体），并使用Python开源自然语言处理工具包（natural language toolkit, NLTK）和布朗实验室的语言信息处理分析器（David McClosky模型^[91]）处理文本中的模糊词和否定词（如“疑似气胸”“并非气胸”等）。在此基础上，使用8种疾病名称作为标签对图像进行第一轮标注。利用第一轮标注结果（8种疾病标签），在每种疾病中选出200个实例（共1 600个实例），并由一名专业医师通过边界矩形框（bounding box, B-Box）方式标注出每种疾病的位置，保

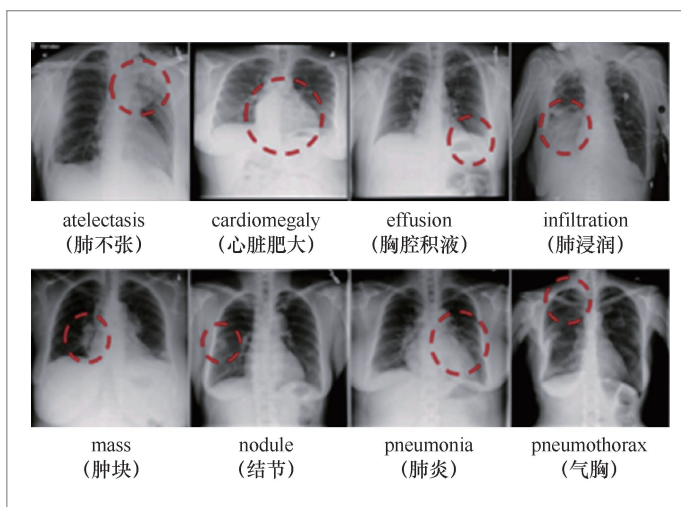


图 15 参考文献 [89] 中的 8 种常见胸部疾病

存在XML中,作为后续CNN训练的基础。深度学习训练过程中用到了ImageNet的预训练模型(即AlexNet、GooLeNet)以及VGGNet-16和ResNet-50等模型。最终训练的模型可用于标注胸片X光影像中的疾病和位置。

在后续的研究中,Wang X等人^[92]进一步通过扩展论文的arxiv版本建设ChestX-ray8库,扩充影像实例,加入更多的疾病类型的标注,形成了Chest X-ray14数据集,并进一步提出了RNN-CNN结合方法基于该数据集的影像分类模型^[93-94],该数据集是目前有标注的规模最大的医疗影像开放数据集,其中包含14种疾病、30 805名患者的112 120张前胸X光图片。

ChestX-ray14的超大规模迅速吸引了诸多研究者在该数据集上进行研究,Yao L等人^[95]在数据集上训练了基于LSTM的分类模型,取得了平均80%以上的准确率;吴恩达团队的Rajpurkar P等人^[96]基于ChestX-ray14数据集训练了一个121层的卷积神经网络CheXNet,并取得了平均84%以上的准确率,并在数个疾病的分类准确率超过了90%,研究人员在文献中称“新技术已经在识别胸透照片中肺炎等疾病的准确率上超越了人类专业医师”。

另外一种自动化的数据集构建趋势也值得关注,参考文献[97]首先在12 600张无标注影像数据中人工标注了100张影像数据,通过学习100张少量的人工标注数据,构建对无标注数据的粗糙标注后由人工进行审核,并将人工审核通过的数据作为标注数据进行下一轮迭代,整体进行6轮迭代后,得到了较高质量的半自动化手部骨骼影像数据集。

随着医疗人工智能场景落地的需求加剧,工业界也注意到了利用文本挖掘技术辅助构建医疗影像有标注数据集的应用前景。在一些工作中,应用自然语言处理结合图像

认知模型赋能真实世界临床数据,通过人在环路(human-in-the-loop)的方法,迭代构建基于临床数据的有标注、标准化、大规模数据,通过医工结合和医学研究大数据的方法,有望夯实医疗影像分析的数据基础,实现医疗影像辅助诊断、影像智能设备、影像自诊等智能场景的大范围突破。

5 结束语

综上所述,AI在医学影像中有非常广泛的应用,本文列举的只是很小的一部分。从这些研究成果中不难看出,AI已经在推动医学成像设备智能化、数据采集规范化和标准化、数据分析自动化等方面取得了重要的进展。随着数据的积累和技术的进一步成熟,AI与医疗的结合将产生巨大的社会效益和经济效益,改变医疗资源紧缺的现状,在技术的驱动下迅速提高基层医生的诊疗水平,切实解决看病难的问题。

参考文献:

- [1] MARDANI M, GONG E, CHENG J Y, et al. Deep generative adversarial neural networks for compressive sensing (GANCS) MRI[J]. IEEE Transactions on Medical Imaging, 2019, 38(1): Accepted.
- [2] SCHLEMPER J, CABALLERO J, HAJNAL J V, et al. A deep cascade of convolutional neural networks for MR image reconstruction[C]//International Conference on Information Processing in Medical Imaging, June 25-30, 2017, North Carolina, USA. Heidelberg: Springer Press, 2017: 647-658.
- [3] YANG Y, SUN J, LI H, et al. Deep ADMM-Net for compressive sensing MRI[C]//The 30th Conference on Neural

- Information Processing Systems, December 5–10, 2016, Barcelona, Spain. Cambridge: MIT Press, 2016.
- [4] YE J C, HAN Y, CHA E. Deep convolutional framelets: a general deep learning framework for inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2018, 11(2): 991–1048.
- [5] CHEN M, WANG Y, SUN A, et al. Advantages of joint multi-echo MRI reconstruction via deep learning[C]// *ISMRM Workshop on Machine Learning Part II*, October 25–28, 2018, Washington, USA. Hoboken: Wiley Press, 2018.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, October 5–9, 2015, Munich, Germany. Heidelberg: Springer, 2015: 234–241.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] HAN Y, YE J C. Framing u-net via deep convolutional framelets: application to sparse-view CT[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1418–1429.
- [11] ANIRUDH R, KIM H, THIAGARAJAN J J, et al. Lose the views: limited angle CT reconstruction via implicit sinogram completion[C]// *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 18–22, 2018, Salt Lake City, USA. Piscataway: IEEE Press, 2018: 6343–6352.
- [12] XU J, GONG E, PAULY J, et al. 200x low-dose PET reconstruction using deep learning[C]// *NIPS Workshop on Machine Learning for Health*, December 8, 2017, Long Beach, USA. Cambridge: MIT Press, 2017.
- [13] HAN L, JU G, MEI L, et al. Multitask DNN for liver imaging enhancement[C]// *ISMRM Workshop on Machine Learning Part II*, October 25–28, 2018, Washington, DC, USA. Hoboken: Wiley Press, 2018.
- [14] KELM B M, WELSM, ZHOU S K, et al. Spine detection in CT and MR using iterated marginal space learning[J]. *Medical Image Analysis*, 2013, 17(8): 1283–1292.
- [15] GHESU F C, GEORGESCU B, MANSI T, et al. An artificial agent for anatomical landmark detection in medical images[C]// *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, October 17–21, 2016, Athens, Greece. Heidelberg: Springer, 2016: 229–237.
- [16] ZHANG P, WANG F, ZHENG Y. Self supervised deep representation learning for fine-grained body part recognition[C]// *IEEE 14th International Symposium on Biomedical Imaging*, April 18–21, 2017, Melbourne, Australia. Piscataway: IEEE Press, 2017: 578–582.
- [17] ZOU D, GUO N, WANG J, et al. Automatic FOV positioning of brain MRI using 2D scout images[C]// *ISMRM Workshop on Machine Learning Part II*, October 25–28, 2018, Washington, DC, USA. Hoboken: Wiley Press, 2018.
- [18] LAMBIN P, RIOS-VELAZQUEZ E, LEIJENAAR R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *European Journal of Cancer*, 2012, 48(4): 441–446.
- [19] KUMAR V, GU Y H, BASU S, et al. Radiomics: the process and the challenges[J]. *Magnetic Resonance Imaging*, 2012, 30(9): 1234–1248.

- [20] AERTS H J, VELAZQUEZ E, LEIJENAAR R T H, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J]. *Nature Communications*, 2014, 5(1): 4006.
- [21] VALLIERES M, FREEMAN C R, SKAMENE S R, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities[J]. *Physics in Medicine and Biology*, 2015, 60(14): 5471-5496.
- [22] GUTMAN D A, COOPER L A, HWANG S N, et al. MRI predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set[J]. *Radiology*, 2013, 267(2): 560-569.
- [23] HUANGS Y, FRANCO B L, HARNISH R J, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis[J]. *NPJ Breast Cancer*, 2018(4): 24.
- [24] LI K, SUN H Z, LU Z M, et al. Value of [(18)F]FDG PET radiomic features and VEGF expression in predicting pelvic lymphatic metastasis and their potential relationship in early-stage cervical squamous cell carcinoma[J]. *European Journal of Radiology*, 2018, 106: 160-166.
- [25] SAH B R, OWCZARCZYK K, SIDDIQUE M, et al. Radiomics in esophageal and gastric cancer[J]. *Abdom Radiol (NY)*, 2018.
- [26] KAMILA N, REITER J G, BRACHTEL E, et al. Origins of lymphatic and distant metastases in human colorectal cancer[J]. *Science*, 2017, 357(6346): 55-60.
- [27] ZHAO B S, TAN Y Q, TSAI W Y, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging[J]. *Scientific Reports*, 2016(6): 23428.
- [28] SHROUT P, FLEISS J L. Intraclass correlations: uses in assessing rater reliability[J]. *Psychological Bulletin*, 1979, 86(2): 420-428.
- [29] LIN H S, CHEN Y J, LU H L, et al. Test-retest reliability of mandibular morphology measurements on cone-beam computed tomography-synthesized cephalograms with random head positioning errors[J]. *BiomedEng Online*, 2017.
- [30] YANG J Z, ZHANG L F, FAVE X J, et al. Uncertainty analysis of quantitative imaging features extracted from contrast-enhanced CT in lung tumors[J]. *Computerized Medical Imaging and Graphics*, 2016, 48(1): 1-8.
- [31] PARMAR C, LEIJENAAR R T H, GROSSMANN P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck Cancer[J]. *Scientific Reports*, 2015, 5: 11044.
- [32] SHU J, TANG Y Q, CUI J J, et al. Clear cell renal cell carcinoma CT-based radiomics features for the prediction of fuhrman grade[J]. *European Journal of Radiology*, 2018(109): 8-12.
- [33] ZHANG Y, ZHU Y, SHI X, et al. Soft tissue sarcomas: preoperative predictive histopathological grading based on radiomics of MRI[J]. *Academic Radiology*, 2018.
- [34] FENG D Y, ZHOU Y Q, XING Y F, et al. Selection of glucocorticoid-sensitive patients in interstitial lung disease secondary to connective tissue diseases population by radiomics[J]. *Therapeutics and Clinical Risk Management*, 2018, 14: 1975-1986.
- [35] CUI Y, THAKK K, TERASAKA S, et al. Prognostic imaging biomarkers in GBM: development and validation on the basis of multiregion and quantitative analysis of MRI[J]. *Radiology*, 2016, 278(2): 546-553.
- [36] HUANG Y Q, LIANG C H, HE L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer[J]. *Journal of Clinical Oncology*, 2016, 34(18): 2157-2164.
- [37] LAMBIN P, LEIJENAAR R T H, DEIST T M, et al. Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nature Reviews Clinical*

- Oncology, 2017, 14(12): 749–762.
- [38] ZHANG B, TIAN J, DONG D, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma[J]. *Clinical Cancer Research*, 2017, 23(15): 4259–4269.
- [39] HUANG Y Q, LIU Z, HE L, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in Early-Stage (I or II) NSCLC[J]. *Radiology*, 2016, 281(3): 947–957.
- [40] WANG J, LIU X, DONG D, et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method[C]//The 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), August 16–20, 2016, Orlando, USA. Piscataway: IEEE Press, 2016: 1272–1275.
- [41] WEI K, SU H, ZHOU G, et al. Potential application of radiomics for differentiating solitary pulmonary nodules[J]. *OMICS J Radiol*, 2016, 5(2).
- [42] ZHOU Y, HE L, HUANG Y Q, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma[J]. *Abdominal Radiology*, 2017, 42(6): 1695–1704.
- [43] LIANG C S, HUANG Y Q, HE L, et al. The development and validation of a CT-based radiomics signature for the preoperative[J]. *Oncotarget*, 2016, 7(21): 31401–31412.
- [44] GIESEL F L, SCHNEIDER F, KRATOCHWIL C, et al. Correlation between SUVmax and CT radiomic analysis using lymph node density in PET/CT-based lymph node staging[J]. *Journal of Nuclear Medicine*, 2017, 58(2): 282–287.
- [45] HEBERT A V, HARINI V, MAURA M, et al. A novel representation of inter-site tumour heterogeneity from pre-treatment computed tomography textures classifies ovarian cancers by clinical outcome[J]. *European Radiology*, 2017, 27(9): 3991–4001.
- [46] YAMASHITA K, HIWATASHI A, TOGAO O, et al. MR imaging-based analysis of glioblastoma multiforme: estimation of IDH1 mutation status[J]. *American Journal of Neuroradiology*, 2016, 37(1): 58–65.
- [47] STOYANOVA R, TAKHAR M, TSCHUDI Y, et al. Prostate cancer radiomics and the promise of radiogenomics[J]. *Translational Cancer Research*, 2016, 5(4): 432–447.
- [48] CAMERON A, KHALVATI F, HAIDER MA, et al. MAPS a quantitative radiomics approach for prostate cancer detection[J]. *IEEE Transactions on Biomedical Engineering*, 2016, 63(6): 1145–1156.
- [49] BRAMAN N M, ETESAMI M, PRASANNAP, et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI[J]. *Breast Cancer Research*, 2017, 19(1): 57.
- [50] FARBER N J, WU Y, ZOU L, et al. Challenges in RCC imaging renal insufficiency, post-operative surveillance, and the role of radiomics[J]. *Journal of Kidney Cancer*, 2015, 13(4): 84–90.
- [51] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. *Medical Image Analysis*, 2017, 42(9): 60–88.
- [52] SUTSKEVER I, MARTENS J, HINTON G E. Generating text with recurrent neural networks[C]//The 28th International Conference on Machine Learning, June 28–July 2, 2011, Bellevue, USA. [S.l. : s.n.], 2011: 1017–1024.
- [53] KARPATHY A, LI F F. Deep visual-semantic alignments for generating image descriptions[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7–12, 2015, Boston, USA. Piscataway: IEEE Press, 2015: 3128–3137.
- [54] KALLENBERG M, PETERSEN K, NIELSEN M, et al. Unsupervised deep learning applied to breast density segmentation and mammographic risk

- scoring[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1322–1331.
- [55] ACKLEY D H, HINTON G E, SEJNOWSKI T J. A learning algorithm for boltzmann machines[M]// *Connectionist models and their implications: readings from cognitive science*. New York: Ablex Publishing Corp, 1988.
- [56] MCCLELLAND J. *Information processing in dynamical systems: foundations of harmony theory[M]//Parallel distributed processing: explorations in the microstructure of cognition: foundations*. Cambridge: MIT Press, 1986.
- [57] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// *The 27th International Conference on Neural Information Processing Systems*, December 8–13, 2014, Montreal, Canada. Cambridge: MIT Press, 2014: 2672–2680.
- [58] LO S C B, LOU S L A, LIN J S, et al. Artificial convolution neural network techniques and applications for lung nodule detection[J]. *IEEE Transactions on Medical Imaging*, 1995, 14(4): 711–718.
- [59] RAJKOMAR A, LINGAM S, TAYLOR A G, et al. High-throughput classification of radiographs using deep convolutional neural networks[J]. *Journal of Digital Imaging*, 2017, 30(1): 95–101.
- [60] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7–12, 2015, Boston, USA. Piscataway: IEEE Press, 2015.
- [61] RAJPURKAR P, IRVIN J, ZHU K, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning[J]. *ScienceOpen*, 2017.
- [62] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21–26, 2017, Honolulu, USA. Piscataway: IEEE Press, 2017.
- [63] WANG X, PENG Y, LU L, et al. ChestX-Ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21–26, 2017, Honolulu, USA. Piscataway: IEEE Press, 2017.
- [64] LI Z, WANG C, HAN M, et al. Thoracic disease identification and localization with limited supervision[C]// *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 18–22, 2018, Salt Lake City, USA. Piscataway: IEEE Press, 2018: 8290–8299.
- [65] SHEN W, ZHOU M, YANG F, et al. Multi-scale convolutional neural networks for lung nodule classification[J]. *Information Processing in Medical Imaging*, 2015, 24: 588–599.
- [66] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542(7639): 115–118.
- [67] PRATT H, COENEN F, BROADBENT D M, et al. Convolutional neural networks for diabetic retinopathy[J]. *Procedia Computer Science*, 2016, 90: 200–205.
- [68] SUK H I, WEE C Y, LEE S W, et al. State-space model with deep learning for functional dynamics estimation in resting-state fMRI[J]. *Neuroimage*, 2016, 129: 292–307.
- [69] SHIN H C, ROTH H R, GAO M, 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.
- [70] BECKER A S, MARCON M, GHAFOR S, et al. Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer[J]. *Investigative Radiology*, 2017, 52(7): 434.
- [71] WANG X, YANG W, WEINREB J, et al. Searching for prostate cancer by fully

- automated magnetic resonance imaging classification: deep learning versus non-deep learning[J]. *Scientific Reports*, 2017(7): 15415.
- [72] YAN K, WANG X, LU L, et al. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning[J]. *Medical Image Analysis*, 2018, 5(3).
- [73] YAN K, BAGHERI M, SUMMERS R M. 3D context enhanced region-based convolutional neural network for end-to-end lesion detection[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, September 16-20, 2018, Granada, Spain. Heidelberg: Springer, 2018: 511-519.
- [74] CIRESDAN D C, GIUSTI A, GAMBARDELLA L M, et al. Mitosis detection in breast cancer histology images with deep neural networks[J]. *Medical Image Computing and Computer Assisted Intervention*, 2013, 16(Pt 2): 411-418.
- [75] SIRINUKUNWATTANA K, RAZA S E A, TSANG Y W, et al. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1196-1206.
- [76] MOESKOP S P, VIERGEVER M A, MENDRIK A M, et al. Automatic segmentation of MR brain images with a convolutional neural network[J]. *IEEE Transactions on Medical Imaging*, 2016(1): 1.
- [77] MILLETARI F, NAVAB N, AHMADI S A. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//The 4th International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, USA. Piscataway: IEEE Press, 2016.
- [78] 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, October 5-9, 2015, Munich, Germany. Heidelberg: Springer, 2015: 234-241.
- [79] STOLLENGA M F, BYEON W, LIWICKI M, et al. Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation[J]. *Computer Science*, 2015.
- [80] ANDERMATT S, PEZOLD S, CATTIN P. Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data[M]// *Deep Learning and Data Labeling for Medical Applications*, October 17-21, 2016, Athens, Greece. Heidelberg: Springer, 2016: 142-151.
- [81] SINGH V K, ROMANI S, RASHWAN H A, et al. Conditional generative adversarial and convolutional networks for X-ray breast mass segmentation and shape classification[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, September 16-20, 2018, Granada, Spain. Heidelberg: Springer, 2018: 833-840.
- [82] YANG X, KWITT R, STYNER M, et al. Quicksilver: fast predictive image registration—a deep learning approach[J]. *Neuroimage*, 2017, 158: 378-396.
- [83] 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.
- [84] YAN P, XU S, RASTINEHAD A R, et al. Adversarial image registration with application for MR and TRUS image fusion[J]. [C]//International Workshop on Machine Learning in Medical Imaging, September 16, 2018, Granada, Spain. Heidelberg: Springer, 2018: 197-204.
- [85] MAZUROWSKI M A, HABAS P A, ZURADA J M, et al. Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance[J]. *Neural Networks*, 2008, 21(2-3): 427-436.
- [86] SCHLEGL T, WALDSTEIN S M, VOGL W D, et al. Predicting semantic descriptions from medical images with convolutional

- neural networks[M]//Information processing in medical imaging. Heidelberg: Springer, 2015: 437-448.
- [87] SHIN H C, LU L, KIM L, et al. Interleaved text/image deep mining on a large-scale radiology database[J]. Journal of Machine Learning Research, 2016, 17(1): 3729-3759.
- [88] WANG X S, LU L, SHIN H C, et al. Unsupervised category discovery via looped deep pseudo-task optimization using a large scale radiology image database[J]. Computer Science, 2016, arXiv: 1603. 07965.
- [89] SHIN H C, ROBERTS K, LU L, et al. Learning to read chest x-rays: recurrent neural cascade model for automated image annotation[C]//The IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, USA. Piscataway: IEEE Press, 2016: 2497-2506.
- [90] WANG X S, PENG Y F, LU L, et al. Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, USA. Piscataway: IEEE Press, 2017: 3462-3471.
- [91] MCCLOSKEY D. Any domain parsing: automatic domain adaptation for natural language parsing[D]. Providence: Brown University, 2009.
- [92] WANG X, PENG Y, LU L, et al. ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, USA. Piscataway: IEEE Press, 2017: 3462-3471.
- [93] PENG Y, WANG X, LU L, et al. NegBio: a high-performance tool for negation and uncertainty detection in radiology reports[J]. Summits on Translational Science Proceedings, 2018: 188-196.
- [94] WANG X, PENG Y, LU L, et al. TieNet: text-image embedding network for common thorax disease classification and reporting in chest X-rays[C]// IEEE CVPR 2018, June 18-22, 2018, Salt Lake City, USA. Piscataway: IEEE Press, 2018.
- [95] YAO L, POBLENZ E, DAGUNTS D, et al. Learning to diagnose from scratch by exploiting dependencies among labels[J]. Computer Science, 2017, arXiv: 1710. 10501.
- [96] RAJPURKAR P, IRVIN J, ZHU K, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning[J]. Computer Science, 2017, arXiv: 1711. 05225.
- [97] IGLOVIKOV V, RAKHLIN A, KALININ A, et al. Pediatric bone age assessment using deep convolutional neural networks[J]. Computer Science, 2017, arXiv: 1712. 05053.

作者简介



韩冬(1983-),男,博士,沈阳东软医疗系统有限公司人工智能中心算法研发负责人,主要研究方向为人工智能、图像处理、计算机视觉等。

作者简介



李其花(1987-),女,慧影医疗科技(北京)有限公司合作科学家,主要研究方向为医学人工智能。



蔡巍(1982-),男,东软集团股份有限公司先行产品研发事业部人工智能专家,东软智能医疗研究院研究员,东软智能医疗研究云平台研发负责人,主要研究方向为人工智能、复杂系统与复杂网络、类脑计算、自然语言处理。



夏雨薇(1993-),女,慧影医疗科技(北京)有限公司合作科学家,主要研究方向为计算机视觉。



宁佳(1991-),女,博士,沈阳东软医疗系统有限公司磁共振研发中心算法工程师,主要研究方向为磁共振快速和定量成像、非笛卡儿成像等。



黄峰(1973-),男,博士,曾任飞利浦医疗主任研究员和临床应用总监,现任沈阳东软医疗系统有限公司人工智能中心首席科学家兼中心主任、磁共振首席科学家兼卓越临床总监。中国医学装备协会人工智能联盟理事、影像装备人工智能联盟企业副主任委员。发表国际核心期刊文章43篇,发表会议论文超过200篇;申请的国内国际专利已公布40余项。担任13种期刊的审稿人,曾任 IEEE transaction on Biomedical Engineering 的副主编和《中国医学装备》编委,作为主要负责人获得国家“十三五”和上海市基金累计超过2 000万元。数十次在国际会议上口头汇报或受邀演讲。多项研究成果已经广泛应用于工业界。

收稿日期: 2018-12-04

基金项目: 2016北京市科技计划专项课题基金资助项目(No.Z16010101038); 2017北京市自然科学基金资助项目(No.7181010, No.7182040); 2018年度医药协同科技创新研究基金资助项目(No.Z181100001918010); 2018年沈阳市科学技术计划基金资助项目(No.18-400-2-01)

Foundation Items: 2016 Beijing Science and Technology Program(No.Z16010101038), 2017 Beijing Municipal Natural Science Foundation(No.7181010, No.7182040), 2018 Pharmaceutical Collaborative Science and Technology Innovation Research Project(No.Z181100001918010), 2018 Shenyang Science and Technology Program(No.18-400-2-01)