

医学图像分析深度学习研究方法研究与挑战

田娟秀^{1,2} 刘国才¹ 谷珊珊³ 鞠忠建³ 刘劲光^{1,2} 顾冬冬¹

摘要 深度学习 (Deep learning, DL), 特别是深度卷积神经网络 (Convolutional neural networks, CNNs), 能够从医学图像大数据中自动学习提取隐含的疾病诊断特征, 近几年已迅速成为医学图像分析研究热点. 本文首先简述医学图像分析特点; 其次, 论述深度学习基本原理, 总结深度 CNNs 在医学图像分析中的分类、分割框架; 然后, 分别论述深度学习在医学图像分类、检测、分割等各应用领域的国内外研究现状; 最后, 探讨归纳医学图像分析深度学习方法挑战及其主要应对策略和开放的研究方向.

关键词 深度学习, 医学图像分析, 卷积神经网络, 图像分类, 图像分割

引用格式 田娟秀, 刘国才, 谷珊珊, 鞠忠建, 刘劲光, 顾冬冬. 医学图像分析深度学习研究方法研究与挑战. 自动化学报, 2018, 44(3): 401–424

DOI 10.16383/j.aas.2018.c170153

Deep Learning in Medical Image Analysis and Its Challenges

TIAN Juan-Xiu^{1,2} LIU Guo-Cai¹ GU Shan-Shan³ JU Zhong-Jian³ LIU Jin-Guang^{1,2} GU Dong-Dong¹

Abstract Deep learning (DL) algorithms, such as convolutional neural networks (CNNs), can automatically extract hidden disease diagnosis features from medical image data, and are being used to analyze medical images now. We review most of the deep learning methods for medical image analysis. Firstly, we introduce the characteristics of medical image analysis briefly. Then, we analyze the principles of deep learning, highlight the popular CNNs and summarize the frameworks of image classification and segmentation. Thirdly, we describe the state-of-the-art of the medical image analysis methods based on deep learning. Finally, we discuss the challenges and practicable strategies in deep learning for medical image analysis, as well as open research.

Key words Deep learning (DL), medical image analysis, convolutional neural networks (CNNs), image classification, image segmentation

Citation Tian Juan-Xiu, Liu Guo-Cai, Gu Shan-Shan, Ju Zhong-Jian, Liu Jin-Guang, Gu Dong-Dong. Deep learning in medical image analysis and its challenges. *Acta Automatica Sinica*, 2018, 44(3): 401–424

生物医学影像已成为疾病诊断和治疗中不可或缺的组成部分, 且日益重要. 核磁共振成像 (Magnetic resonance image, MRI)、正电子发射断层扫描 (Positron emission tomography, PET)、计算机断层扫描 (Computer tomography, CT)、锥形束 CT、3D 超声成像等医学影像技术目前已广泛应用于临床检查、诊断、治疗与决策. 如何充分利用人工智能深度学习分析方法处理这些超大规模的医学

图像大数据, 为临床医学中各种重大疾病的筛查、诊断、治疗计划、治疗图像引导、疗效评估和随访提供科学方法和先进技术, 是当前医学图像分析领域急需解决的重大科学问题和前沿医学影像关键技术^[1].

医学图像分析最初主要采用边缘检测、纹理特征、形态学滤波以及构建形状模型和模板匹配等方法. 这类分析方法通常针对特定任务而设计, 被称为手工定制式设计方法. 机器学习以数据驱动方式分析任务, 能自动地从特定问题的大规模数据集中学习相关模型特征和数据特性. 与针对特定问题而显式地手工设计模型不同, 机器学习方法可直接从数据样本中隐式地自动学习医学图像特征, 其学习过程本质上是一个优化问题的求解过程. 通过学习, 模型从训练数据中选择正确的特征, 使分类器在测试新数据时做出正确决策. 因此, 机器学习在医学图像分析中起着至关重要的作用, 已经成为最有前途的研究领域^[2].

深度学习 (Deep learning, DL) 是一种机器学习方法, 源于人工神经网络的研究, 其动机在于建立模拟人脑分析理解数据的神经网络. 1959 年, 霍

收稿日期 2017-03-21 录用日期 2017-10-30
Manuscript received March 21, 2017; accepted October 30, 2017
国家自然科学基金 (61671204, 61271382, 61301254, 61471166), 湖南省科技计划重点研发专项基金 (2016WK2001) 资助

Supported by National Natural Science Foundation of China (61671204, 61271382, 61301254, 61471166), Key Research and Development Program of Hunan Province (2016WK2001)

本文责任编辑 桑农

Recommended by Associate Editor SANG Nong

1. 湖南大学电气与信息工程学院 长沙 410082 2. 湖南工程学院计算机与通信学院 湘潭 411104 3. 北京解放军总医院放疗科 北京 100853

1. College of Electrical and Information Engineering, Hunan University, Changsha 410082 2. College of Computer and Communication, Hunan Institute of Engineering, Xiangtan 411104 3. Departments of Radiation Oncology, Chinese PLA General Hospital, Beijing 100853

普金斯大学 Hubel 等通过观察猫的脑部视觉中枢对视网膜感知图像的处理方式发现, 视神经元对信息处理的方式是分工分层的, 不同神经元关注的对象特征不同, 每一层神经元抽象出对象的部分特征进行处理, 所有信息经过逐层激发, 在中枢最高层激发出整个对象认知. 这一发现给从事神经网络研究的计算机专家提供了重要的建模思路^[3]. 20 世纪 80 年代, 神经网络技术进一步发展, 通过在只有输入层和输出层的网络结构中添加中间隐层, 使神经网络可以解决更加复杂的分类问题^[4]. 但层数的增加为各层的神经节点之间的连接权重选取带来困难, 反向传播算法 (Back propagation, BP) 的出现一定程度上解决了权重选取问题^[5]. LeCun 等在 1989 年将 BP 算法应用到前馈多层神经网络学习过程, 用以识别手写的邮政编码^[6]. 随着层次的加深, 多层网络会出现梯度消失问题, 导致 BP 算法无法有效调整神经元连接之间的权重, 学习时间过长. 同时, 由于计算能力这一根本性限制, 神经网络研究一直未能在应用领域取得实质性突破. 直到 2006 年前后, Hinton 团队在深度前馈网络中采取无标注的数据集进行网络预训练, 解决了 BP 算法梯度消失问题. 他们先采用非监督贪心逐层训练方法, 有效降低了观察对象的维度, 然后用监督训练微调所有网络参数. 这一算法为解决深层结构相关优化难题带来了希望, 在图像目标识别等分类预测方面取得了突破性进展^[7-8]. LeCun 等提出的卷积神经网络 (Convolutional neural networks, CNNs) 利用空间相对关系减少参数数目以提高训练性能, 是第一个真正多层结构学习算法^[9]. Graves 等提出的长短时记忆神经网络 (Long short-term memory, LSTM) 也在图像手写字识别和语音识别方面取得了突破性进展^[10-12].

自 2006 年以来, 深度学习在多个领域取得了重要突破. 在语音识别方面, 微软研究人员通过与 Hinton 等合作, 首先将受限玻尔兹曼机 (Restricted boltzmann machine, RBM) 和深度信念网络 (Deep belief network, DBN) 引入到语音识别模型训练中, 在大词汇量语音识别系统中获得了巨大成功, 使语音识别错误率相对之前降低 30%. 2012 年 11 月, 微软展示了一个全自动的同声传译系统, 其支撑的关键技术就是深度学习^[13]. 在图像分类方面, 微软亚洲研究院 He 等^[14] 提出了残差学习框架, 其最重要的突破在于重构了学习过程, 重定向了深层神经网络信息流, 从而很好地解决了此前深层神经网络层数与准确度之间的矛盾. 在人脸识别领域, 香港中文大学 Sun 及其研究团队研发的 DeepID 在使用测试基准 LFW 数据库上获得 99.15% 的人脸识别率^[15]. 2015 年, DeepMind 团队在《Nature》杂志

上公布了自己的研究成果, 通过深度神经网络与强化学习等方法的结合, 计算机能够通过自学成为游戏高手, 在一系列复杂任务中的性能表现与人类相当^[16]. 2016 年 3 月, DeepMind 开发的 AlphaGo 程序以 4 : 1 击败韩国围棋冠军李世石, 成为近年来人工智能领域新的里程碑. 2017 年 5 月, AlphaGo 在中国以 3 : 0 击败世界围棋冠军柯洁, 再次证明了其强大的学习、分析、决策能力. DeepMind 团队自 2016 年起关注医疗领域, 试图将人工智能技术应用于医疗行业. DeepMind Health 开发了名为 Streams 的软件, 帮助临床医生更快地查看医疗结果, 只需几秒钟时间就能查看急性肾脏损伤风险病人的验血结果, 优化病人的治疗方案^[17].

近年来, 深度学习不断取得重大进展, 主要得益于不断提高的计算能力和持续增长的可用数据量, 以及深度学习模型及其算法的不断改进. 其实质是通过构建多隐层的机器学习模型, 利用海量的样本数据训练, 学习更精准的特征, 最终提高分类或预测的准确性^[18]. 深度学习这种从数据中学习层次特征的特点, 使得它非常适合发现高维数据中的复杂结构^[19], 已经应用到语音识别^[13]、自然语言处理^[20]、人脸识别^[15, 21]、目标检测^[22] 等领域和各种挑战赛中^[23-25], 取得了破纪录的好成绩. 有关深度学习的更详细发展历程和非医学图像分析领域的应用进展, 读者可参阅 2015 年由 LeCun 等联名撰写的综述^[19]、Schmidhuber 撰写的综述^[26]、Goodfellow 等撰写的专著^[27] 以及最近发表的中文综述^[28-30].

深度学习在计算机视觉领域的巨大成功, 激发了国内外许多学者将其应用于医疗图像分析. 哈佛大学医学院 Wells 教授在其综述中指出应用深度学习解决医学图像分析任务是本领域的发展趋势^[31]. 2016 年来, 已有多位专家对深度学习在医学图像分析中的研究现状及问题进行了总结、评述和讨论^[32-36]. 最近, Medical Image Analysis 上发表的综述对深度学习在医学图像分类、检测和分割、配准和检索等方面的研究进行了较全面的归纳总结^[37].

本文根据我们课题组近 3 年来收集、整理的文献资料和国家自然科学基金项目研究工作, 聚焦于综述深度学习在医学图像分析应用领域的研究现状和挑战. 首先, 简述医学图像分析特点. 其次, 论述深度学习方法自动提取多层次特征的基本原理; 然后, 重点论述计算机视觉和医学图像分析中深度 CNN 分类、分割框架; 系统梳理深度学习在医学图像分析各个应用领域的国内外研究现状; 最后, 总结深度学习方法应用于医学图像分析时面临的挑战与应对策略, 而且对开放的医学图像分析深度学习研究方向进行展望.

1 医学图像分析特点

医学图像分析已广泛应用于良恶性肿瘤、脑功能与精神障碍、心脑血管疾病等重大疾病的临床辅助筛查、诊断、分级、治疗决策与引导、疗效评估等方面。医学图像分类与识别、定位与检测、组织器官与病灶分割是当前医学图像分析深度学习方法研究主要应用领域。不同成像原理的医学图像分析和计算机视觉领域中的自然图像分析存在较大的差别。至今为止,国内外学者主要针对MRI、CT、X射线、超声、PET、病理光学显微镜等不同成像原理的医学图像分析任务开展了一系列的深度研究工作,因此,本节主要概述这几种医学图像及其主要分析任务。

1.1 常用医学图像特点

1) X射线图像:自德国物理学家伦琴于1895年发现X射线以来,X射线图像用于临床诊断已有100多年的历史。医学X射线图像是人体不同组织器官和病灶的电子密度度量影像。基于X射线的成像包括2D的计算机放射成像、数字化X射线摄影术、数字减影血管造影术和乳房X线摄影术,以及3D的螺旋计算机断层扫描术等,已广泛地应用于骨科^[38-39]、肺部、乳腺和心血管^[40]等临床疾病检测和辅助诊断,但2D X射线图像不能提供人体组织器官和病灶的三维立体信息,2D X射线图像中各组织器官和病灶信息重叠,自动识别比较困难。

2) CT图像:计算机断层扫描(CT)利用精确准直的X射线束对人体某部位一定厚度的断面进行照射扫描,并由与射线束一起旋转的探测器接收透射穿过该断面的X射线,最后,计算机根据探测器接收到的X射线信号数据重建相应人体断面的3D图像。CT图像具有亚毫米级的空间分辨率,能够提供清晰的人体骨性组织解剖结构和病灶影像,已广泛应用于多种临床疾病检查和辅助诊断。CT图像还能提供肿瘤放疗计划剂量计算所必需的组织和病灶电子密度信息,能够为肿瘤放疗提供更准确的靶区和危及器官定位和边界信息,是肿瘤临床放疗的基本图像^[41]。但CT图像不能提供清晰的软组织和病灶影像,因此,高精度的软组织器官和肿瘤CT图像分类识别、检测、定位和分割非常困难。

3) MRI图像:核磁共振图像(MRI)是人体组织器官和病灶中的氢原子核在外部强磁场作用下产生的磁共振信号大小的度量,并通过计算机对体外核磁共振信号探测器接收到的信息数据进行3D图像重建。MRI具有亚毫米级的空间分辨率,能够提供非常清晰的人体软组织解剖结构和病灶影像。功能核磁共振图像(Functional MRI, fMRI)和各种增强剂显像MRI可提供组织生理、

病理和生物化学信息。动态对比度增强(Dynamic contrast-enhanced, DCE)MRI具有非侵入方式评价整个肿瘤区域的能力,已应用于肿瘤学相关领域^[42]。特别是将DCE MRI与核磁共振弥散加权成像(Diffusion-weighted MRI, DW MRI)相结合,能够区分复杂的头颈部内不同区域的肿瘤^[43]。T1W、T2W、DCE、fMRI、DWI、磁共振血管成像、磁共振波谱成像等多种MRI影像增强技术使MRI比CT能更好地可视化、更精确地定位和区分肿瘤和正常软组织器官。但MRI一般不能提供骨性组织解剖结构影像,组织器官之间的空隙容易导致伪影,且难以避免和校正。外部磁场的变化也会产生难以校正的伪影。各种不同的成像序列使MRI图像的自动分析更加复杂和困难。

4) PET图像:正电子发射断层扫描(PET)利用 F^{18} 等放射性元素标记的示踪剂(如 F^{18} 脱氧葡萄糖)衰变时发射的正电子信息成像,因此,PET图像是相应示踪剂放射性活度的度量,能提供肿瘤生物学特性(如葡萄糖代谢、乏氧、增殖等)信息,其标准摄入量大小可用于临床辅助判别肿瘤良、恶性^[44]。PET能提供比CT、MRI更直观、更精确的可视化生物学与放射(抗辐射)生物学特性信息,目前已广泛用于恶性肿瘤临床诊断与转移检查、放疗靶区定位与勾画和肿瘤生物调强放疗计划设计评估。PET图像还常用于心、脑功能与疾病的辅助诊断与评估。然而,PET图像通常缺乏组织器官的解剖结构影像,而且目前临床PET系统的空间分辨率远低于CT和MRI系统,一般只有3~5毫米,部分容器效应大、噪声强,因此,PET图像自动分析更具挑战性。

5) 超声成像:利用超声束扫描人体,通过对反射信号的接收、处理,以获得体内器官的图像。近年来,超声成像技术不断发展,出现了3D彩超、超声全息摄影、体腔内超声成像、彩色多普勒成像及超声生物显微镜等新的超声成像技术。目前超声成像应用非常广泛,已成为临床妇产科、眼科及心血管等部位多种疾病诊断的首选方法,常用来鉴别胎儿发育是否正常,判断内脏器官的形态是否有异常,确定病灶的范围和物理性质^[45-46]。

6) 病理图像:是指切取一定大小的病变组织,采用苏木精和曙红(H & E)等染色方法将切片组织做成病理玻片,然后用显微镜成像技术对微观的细胞和腺体成像。通过对病理图像进行分析,可探讨病变产生的原因、发病机理、病变的发生发展过程,从而做出病理诊断。最近出现的数字全扫描仪能够获得细胞核方向、纹理、形状、结构等肿瘤空间信息,允许对切片组织形态进行量化分析。而识别这些量化特征的先决条件是需要检测和分割细胞核和腺体等组织学单元^[32]。

目前,临床医学图像分析深度学习研究对象并不限于前述 6 类医学图像,还包括裂隙灯成像^[47]、视网膜图像^[48]以及皮肤镜图像^[49]等。

1.2 主要医学图像分析任务

1) 医学图像分类与识别

临床医生常需要借助医学图像来辅助诊断人体内是否有病灶,并对病灶的轻重程度进行量化分级,因此自动识别图像中的病灶区域和正常组织器官是医学图像分析的基本任务。

2) 医学图像定位与检测

人体组织器官解剖结构和病灶区域的定位是临床治疗计划和干预流程中非常重要的预处理步骤,定位的精度直接影响治疗的效果^[37]。图像目标定位任务不仅需要识别图像中的特定目标,而且需要确定其具体的物理位置。图像目标检测任务则需要把图像中所有目标识别出来,且确定它们的物理位置和类别。

3) 医学图像分割任务

图像分割是识别图像中感兴趣的目标区域(如肿瘤)内部体素及其外轮廓,它是临床手术图像导航和图像引导肿瘤放疗的关键任务。

复杂的医学图像分析任务常常需要综合进行分类、检测与分割,如在文献[50]中,为了进行诊断乳房 X 射线图像中病灶的良/恶性,先后进行了病灶检测、病灶分割、病灶分类。由病理学图像分析判断癌症严重程度时,需要首先检测、分割细胞核,然后基于分割结果进行特征和统计分析,最后分类得到分级结果。

2 深度学习模型

本节论述常用的深度学习模型,包括栈式自编码器(SAE)、深度信念网络(DBN)、深度玻尔兹曼机(DBM)、卷积神经网络(CNNs)和循环神经网络(RNN)。重点论述如何通过各种模型从训练样本数据中学习多层次的图像特征。

2.1 无监督特征学习模型与网络微调

2.1.1 SAE

自动编码器(Autoencoder, AE)是特殊的两层神经网络^[51],如图 1(a)所示,若 AE 的输入 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, 隐层 $\mathbf{h} = [h_1, h_2, \dots, h_m]^T$, 输出 $\mathbf{x}' = [x'_1, x'_2, \dots, x'_n]^T$, 编码过程是从 \mathbf{x} 到 \mathbf{h} 的映射,一般采用非线性激活函数计算潜在特征表示:

$$\mathbf{h} = \sigma(W_{\mathbf{x},\mathbf{h}}\mathbf{x} + \mathbf{b}_{\mathbf{x},\mathbf{h}}) \quad (1)$$

解码过程是从隐层 \mathbf{h} 到输出 \mathbf{x}' 的映射:

$$\mathbf{x}' = W_{\mathbf{h},\mathbf{x}'}\mathbf{h} + \mathbf{b}_{\mathbf{h},\mathbf{x}'} \quad (2)$$

其中, $\sigma(\mathbf{x}) = 1/(1 + \exp(-\mathbf{x}))$, 为 sigmoid 函数。 $W_{\mathbf{x},\mathbf{h}}$ 是输入 \mathbf{x} 与隐层 \mathbf{h} 的连接权值矩阵,而 $\mathbf{b}_{\mathbf{x},\mathbf{h}}$ 为该连接的偏置量。 $W_{\mathbf{h},\mathbf{x}'}$ 是隐层 \mathbf{h} 与输出 \mathbf{x}' 的连接权值矩阵,而 $\mathbf{b}_{\mathbf{h},\mathbf{x}'}$ 为此连接的偏置量。权值矩阵 $W_{\mathbf{x},\mathbf{h}}$, $W_{\mathbf{h},\mathbf{x}'}$ 以及偏置量 $\mathbf{b}_{\mathbf{x},\mathbf{h}}$, $\mathbf{b}_{\mathbf{h},\mathbf{x}'}$ 通过求解最小化均方误差函数 J 得到最优值:

$$J = \|\mathbf{x} - \mathbf{x}'\|^2 = \|\mathbf{x} - (W_{\mathbf{h},\mathbf{x}'}(\sigma(W_{\mathbf{x},\mathbf{h}}\mathbf{x} + \mathbf{b}_{\mathbf{x},\mathbf{h}})) + \mathbf{b}_{\mathbf{h},\mathbf{x}'})\|^2 \quad (3)$$

$W_{\mathbf{h},\mathbf{x}'}$ 通常为 $W_{\mathbf{x},\mathbf{h}}^T$, AE 通过最小化网络的输入和输出值之间的重建误差学习输入数据的潜在特征或压缩表示。典型地, $m < n$, 这样,数据被投影到代表输入主要潜在结构的低维子空间中,达到维数约简的效果。

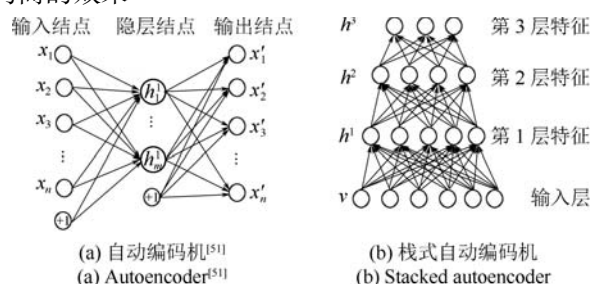


图 1 自动编码器及栈式自编码神经网络

Fig. 1 Autoencoder and stacked autoencoder

单层 AE 是简单的浅层结构,其表示能力非常有限。SAE 是由多层 AE 组成的神经网络,前一层 AE 的输出作为其后一层 AE 的输入,层与层之间采用全连接形式,如图 1(b)所示。SAE 通过自动编码-解码网络从像素级数据中抽取输入图像特征,从而提高模型的表示能力,已广泛应用于维数约简和特征学习。数据样本不仅是 SAE 的输入,同时还作为 SAE 的输出目标,因而可检测 SAE 中间层学到的特征是否符合要求。若对 SAE 中的各层加以稀疏性约束,可构成栈式稀疏自编码器(Stacked sparsely autoencoder, SSAE),使模型具有一定的抗噪能力,且模型泛化性更好^[52-53]。当输入向量用 SSAE 表示时,不同网络层表示不同层次的特征,即网络的较低层表示简单的模式,网络的较高层表示输入向量中更复杂抽象的本质模式。

2.1.2 基于 RBM 的深度模型

RBM 是具有一个可视层和一个隐层的无向图模型^[54]。它假设可视层和隐层之间是对称连接的,但层内结点之间不存在任何连接。给定输入向量,可以得到潜在特征表示,反之亦然。因此, RBM 是一个生成模型,可以从训练数据分布中采样生成新数据。图 2(a) 是一个 RBM 模型,可视层包含了 m 个可视结点 $\mathbf{v} = [v_1, v_2, \dots, v_m]^T$, 隐层包含 n 个隐藏结点 $\mathbf{h} = [h_1, h_2, \dots, h_n]^T$ 。模型参数 $\boldsymbol{\theta} = (W, \mathbf{b}, \mathbf{c})$,

其中, W 是一个 $m \times n$ 的矩阵, 表示可视结点与隐层结点之间的连接权值; $\mathbf{b} = [b_1, b_2, \dots, b_m]^T$, 是可视层对应的偏置值; $\mathbf{c} = [c_1, c_2, \dots, c_n]^T$, 是隐层对应的偏置值.

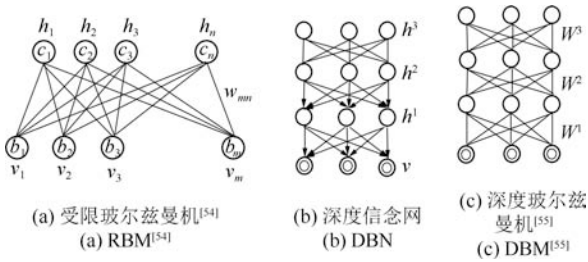


图2 受限玻尔兹曼机 RBM 及基于 RBM 的深度网络
Fig. 2 Restricted Boltzmann machine (RBM) and deep networks based RBM

若可视层和隐层结点服从伯努利分布, 状态 (\mathbf{v}, \mathbf{h}) 的能量函数可定义为

$$E(\mathbf{v}, \mathbf{h}) = \sum_{i=1}^n \sum_{j=1}^m W_{ij} h_i v_j - \sum_{j=1}^m b_j v_j - \sum_{i=1}^n c_i h_i \quad (4)$$

可视结点与隐层结点的联合分布 $p(\mathbf{v}, \mathbf{h})$ 可由能量函数 $E(\mathbf{v}, \mathbf{h})$ 给出:

$$p(\mathbf{v}, \mathbf{h}) = \frac{\exp(-E(\mathbf{v}, \mathbf{h}))}{Z} \quad (5)$$

式中, $Z = \sum_{\mathbf{v}} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h}))$ 为归一化因子. 可视结点的概率分布 $p(\mathbf{v}) = \sum_{\mathbf{h}} p(\mathbf{v}, \mathbf{h}) = \sum_{\mathbf{h}} \frac{\exp(-E(\mathbf{v}, \mathbf{h}))}{Z}$. Z 的计算包括了 2^{m+n} 项, 计算复杂度非常高, 通常需要进行简化.

模型的条件概率为

$$p(h_i = 1 | \mathbf{v}) = \sigma \left(\sum_{j=1}^m W_{ij} v_j + b_j \right) \quad (6)$$

$$p(v_j = 1 | \mathbf{h}) = \sigma \left(\sum_{i=1}^n W_{ij} h_i + c_i \right) \quad (7)$$

由于连接的对称性, 可从隐层表示生成输入观察, 因此, RBM 本质上是一个 AE. 模型参数通过最大化观察与输入的相似性进行优化, 通常采用对比散度算法 (Contrastive divergence, CD) 训练^[56]. RBM 有严谨数学框架, 可以输出概率, 用于解决分类问题.

将 RBM 视作搭建深度网络的基本模块, 可以构建出两种常用深度特征学习网络: DBN 和 DBM.

1) DBN 由 Hinton 等在 2006 年提出, 该模型将多个 RBM 堆叠起来, 构建出深度框架, 从而产生单个概率模型^[7-8]. DBN 包含一个可视层 \mathbf{v} 和一系

列隐层 $\mathbf{h}^1, \dots, \mathbf{h}^L$, 靠近可视层的部分隐层使用贝叶斯置信网络, 形成有向生成模型, 这些层的概率分布计算不依赖于高层, 如 \mathbf{h}^1 层仅依赖于可视层 \mathbf{v} 而不需考虑 \mathbf{h}^2 层, 从而加快了计算速度. 而最上面两层仍保持 RBM 无向生成模型的形式, 如图 2 (b)^[55].

2) DBM 也是以层次方式堆叠多个 RBM 的深度网络, 与上述 DBN 不同的是, DBM 中所有层保持无向生成模型的形式. 如图 2 (c), DBM 包含输入层和 L 个隐层^[55], 且只有相邻层结点之间才有连接. DBM 中间隐层的条件概率分布计算同时利用了其相邻两层信息, 如 DBM 第 l 隐层 (除第 1 隐层) 的条件概率依据第 $l+1$ 隐层和第 $l-1$ 隐层的联合信息来计算, 即 $p(\mathbf{h}^l | \mathbf{h}^{l+1}, \mathbf{h}^{l-1})$, 使 DBM 对有噪输入具有更加健壮的代表能力.

2.1.3 无监督学习与网络微调

从训练角度来看, 前述 SAE、DBN 和 DBM 三种深度模型中, SAE 通过重构误差来进行训练, 较后两者更容易学习, 因而常用来代替 RBM 构建深度结构. 三者形式上有所不同, 但存在相通之处. 首先, 由于 RBM 连接的对称性, 可从隐层特征生成输入观察, 因而, RBM 本质上是一个自动编码器. 其次, 这三种深度模型都是以基本模块 (RBM 或 AE) 逐层堆叠构成, 通常采用 Hinton 提出的贪婪逐层训练方法^[7-8, 55]. 贪婪层次学习的关键思想是每次只预训练一层网络, 即首先用训练样本数据输入训练第 1 隐层的参数, 然后用第 1 隐层的输出作为第 2 隐层的输入, 训练第 2 隐层的参数, 以此类推, 第 l 层的输出作为第 $l+1$ 层的输入以训练第 $l+1$ 层参数. 这种预训练技术以无监督的方式学习, 可利用未标记的样本作为训练数据集. 最后, 它们都是生成模型, 当它们从数据中学习特征表示时, 并未涉及到与分类任务对应的离散标签值或连续实数目标值, 因而学习到的特征不保证具有区分性的. 为了解决这个问题, 通常在非监督特征学习之后进行微调处理 (Fine-tuning).

对特定的分类或回归任务而言, 将 SAE、DBN 或 DBM 的最高隐藏层的输出 (即最高层特征) 作为分类器的输入, 预测输入的目标值, 构建特征学习深度神经网络预测模型^[33]. 为了微调深度神经网络的参数, 首先将预训练得到的各隐层参数作为有监督学习网络的初始值, 随机初始化最高隐层和输出层的连接权值, 然后利用标签数据, 以监督方式通过梯度下降和 BP 算法联合训练所有参数, 对网络参数进行调整. 经验证明, 通过预训练初始化参数有助于监督优化, 减少陷入局部最优的风险^[7, 57].

这三种深度模型可以学习潜在的本质特征, 从大量无标签的信息数据中学习高层次的特征表达,

结合标签数据对模型参数进行微调, 从而进一步达到期望的分类结果. 同时, 这种自底而上的生成结构能够自顶向下地反馈误差, 可减少网络前馈传递造成的误差, 更鲁棒地处理输入数据. 但是, 它们要求网络的输入通常为向量形式, 而对于医学图像, 像素或体素的邻域结构信息是一个重要的信息源, 向量化必然会破坏图像中的邻域结构关系.

2.2 监督学习模型

2.2.1 CNN

为更好利用空间结构信息, 1989 年 LeCun 提出了 CNN^[6]. CNN 可接受 2D 或 3D 图像块作为输入, 适于捕获视觉的局部信息. CNN 通常由多个交替出现的卷积层和池化 (Pooling) 层, 以及末端的全连接层构成多层次人工神经网络, 如图 3 所示. 其中, 卷积层具有固定的输入大小, 感受局部区域上下文信息, 具有权值共享和下采样的特点, 同时具有一定程度的位移、尺度和形变不变性, 大大减少了模型的自由度^[6, 9].

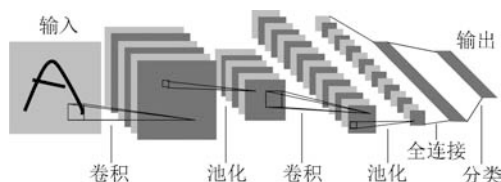


图 3 卷积神经网络框架^[9]

Fig. 3 Architecture of convolutional neural network^[9]

与传统固定卷积算子不同的是, 卷积层的卷积核通过学习得到. 卷积层通常采用多个卷积核, 用于提取多个特征映射. 设学习核 k_{ij}^l 表示第 $l-1$ 层的特征映射 i 与第 l 层特征映射 j 之间的连接权值, 卷积层的作用是利用核 k_{ij}^l 在输入不同位置检测局部特征.

$$A_j^l = f \left(\sum_{i=1}^{M^{l-1}} A_i^{l-1} * k_{ij}^l + b_j^{l-1} \right) \quad (8)$$

其中, $*$ 表示卷积操作, b_j^{l-1} 是偏置值, $f(\cdot)$ 是非线性激活函数. 具体来说, 卷积层 l 中第 j 个特征映射 A_j^l 依据其邻接较低 $l-1$ 层的特征映射 $A_i^{l-1}, i=1, 2, \dots, M^{l-1}$ 来计算, M^{l-1} 表示第 $l-1$ 层的特征映射数目.

池化层在卷积层之后, 对卷积层的特征映射进行子采样. 具体来说, 池化层特征映射中的每个结点特征值是基于对应卷积特征映射的局部感受野计算, 在感受野结点范围内找到一个代表值, 如最大值或平均值. 通常, 池化层中感受野的步长与子采样感受野的大小相等, 这样有助于 CNN 保持平移不变性^[33].

在卷积网络的末端, 通常是可作为分类的全连接层. 与 SAE 和 DBN 的逐层训练相反, CNN 通常以完全监督的方式端到端地训练. CNN 在本质上是一种输入到输出的映射, 通过学习使网络具有输入到输出的映射能力. 通常, CNN 网络参数的训练算法与传统 BP 算法相似, 通过前向传播计算输出值, 然后将输出值与理想标签值的误差, 通过梯度下降法对最小化误差问题寻优, 再利用反向传播梯度调整 CNN 的参数.

2.2.2 RNN

RNN 是一种具有反馈连接的循环神经网络, 其本质属性是网络的状态会随时间演化, 适用于提取数据的时序特征^[18]. RNN 在给定序列 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T$ 的情况下学习时序特征. 在某时间 t 隐层神经元状态 \mathbf{h}_t 由输入 \mathbf{x}_t 和前一状态 \mathbf{h}_{t-1} 确定:

$$\mathbf{h}_t = \sigma(W\mathbf{x}_t + R\mathbf{h}_{t-1} + \mathbf{b}) \quad (9)$$

W 表示隐层神经元与输入结点的连接权值矩阵, 而 R 表示隐层神经元之间的连接权值矩阵, $\boldsymbol{\theta} = (W, R, \mathbf{b})$. 对于分类任务, 通常在其后加全连接层和 softmax 分类层, 将序列映射到特定类别标签^[37].

$$P(Y = y | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T; \boldsymbol{\theta}) = \text{softmax}(\mathbf{h}_T; W_{\text{out}}; \mathbf{b}_{\text{out}}) \quad (10)$$

因为梯度需要从输出通过时间反向传播及 RNN 自身的深度特性, 因而会有学习中存在梯度消失或爆炸的问题^[58]. 1997 年由 Hochreiter 等提出的长短时记忆神经网络 LSTM 可解决这一问题^[59]. LSTM 模块由门控函数组成, 每个门受输入权值矩阵和前一个隐藏状态权值矩阵的控制. 模块的核心是一个存储单元 c , 可存储模块的输出并作为下一时刻的状态.

3 基于深度学习的医学图像分析

近几年, 深度学习方法不仅仅在以自然图像为分析处理对象的计算机视觉任务中展示了卓越的性能, 同时, 也在医学图像领域取得了突破性进展. 目前, 本领域学者大多利用深度 CNN 做分类研究, 主要用于医学图像病灶识别、检测和分割等医学图像分析任务. 在 2016 年 IEEE Trans on Medical Imaging 深度学习专刊出版的 18 篇论文中, 有 14 篇是深度 CNN 及其改进方法^[34]. 本节首先归纳计算机视觉和医学图像分析中深度 CNN 分类、分割框架, 然后, 总结深度学习在医学图像分类、检测、分割等应用领域的国内外研究现状.

3.1 计算机视觉中的深度 CNN

1998 年, LeCun 提出 LeNet, 并成功应用于手写数字识别^[9]. 随着 ReLU 和 Dropout 的提出, 以及 GPU 和大数据带来的历史机遇, CNN 在 2012 年迎来了历史突破, AlexNet 网络结构获得了当年的 ImageNet 图像分类竞赛冠军^[60]. Farabet 等率先引入多尺度 (Multi-scale) 和多处理流 (Multi-stream) 思想, 利用多个卷积神经网络流程分别处理输入图像金字塔的不同尺度信息, 提取多层次特征, 使模型具有强大捕获纹理、形状和上下文信息的能力^[61]. 2014 年提出的 VGG 框架与 AlexNet 相似, 但采用小卷积核实现更深的网络, 利用多尺度融合进一步提升准确率^[62]. 针对 CNN 经典结构要求固定尺寸的输入图像而可能导致损失信息的问题, He 等利用空间金字塔池化 (Spatial pyramid pooling, SPP) 深度网络模型, SPP 接受任意大小的输入图像, 输出等长的特征表示, 同时, SPP 结构增强了形变的鲁棒性^[63]. 2013 年底, Lin 等提出 Network in network (NIN) 结构, 采用微神经网络 MLP 代替卷积层中的线性滤波器, 在分类层使用全局均值池化取代全连接层减少过拟合风险^[64]. 在 NIN 的基础上, Google 提出了 Inception 结构, 将不同大小的卷积核输出与 3×3 的池化层输出级联起来, 增加了网络的宽度和尺度的适应性, 同时, 应用 1×1 的卷积来减少特征映射的维度. Google 将多个 Inception 结构串联起来, 搭建了一个 22 层的深度网络, 即 GoogLeNet, 获 ImageNet ILSVRC-2014 分类和检测任务的冠军^[65]. 2015 年, Inception 结构变得更深, 原有的 5×5 卷积核拆成两层 3×3 的卷积核, 并加入 $1 \times n$ 和 $n \times 1$ 的卷积核. 同时, He 等发现, 单纯增加深度会导致网络退化, 为此提出残差网络 ResNet, 允许模型中存在跳转连接 (Shortcut), 成功训练了深达 152 层的网络, 是 2015 年 ImageNet ILSVRC 不依赖外部数据的图像检测与图像分类竞赛的双料冠军.

近两年, 深度卷积网络朝着更深、更准确方向发展的同时, 也有不少工作朝着网络更简化、更灵活、更有效地训练方面努力. Google 进一步地将 Inception 结构与 Residual net 结合, 提出了 Inception-residual net 结构. 实验表明, 采用 Inception V4 结构的性能与 ResNet 相当, 但训练时间减半, 收敛更快^[66-67]. Long 等提出全卷积网络 (Fully convolutional network, FCN) 进行像素级分类, 避免了图像块之间的重叠而导致重复卷积计算的问题^[68]. Zagoruyko 等提出了宽度残差网络 (Wide residual networks, WRNs), 通过增加宽度来提高性能, 且在参数数量相当的情况下, WRNs 比深度 ResNet 好

训练. 实验表明 16 层 WRNs 网络在精度和效率上达到、甚至优于上千层的 ResNet 网络^[69]. Huang 等提出了密集连接卷积神经网络 DenseNet, 在该网络中, 任何两层之间都有直接连接. DenseNet 有效地缓解梯度消失问题, 强化了特征传播, 支持特征重用, 并大幅度降低网络参数数量, 可以更有效地训练^[70]. Chen 等提出了一种结合 ResNet 和 DenseNet 优势的新型卷积网络结构 DPN (Dual path networks), 该网络通过残差旁支通路再利用特征和密集连接通路探索新特征^[71]. 而 Iandola 等则致力于简化网络结构和减少网络参数, 提出了全新的网络结构 SqueezeNet, 该网络仅需 $1/50$ 的 AlexNet 参数量即可达到了 AlexNet 相同的精度^[72].

上述网络框架中卷积核的几何形状是固定的, 对模型几何变换的模拟能力受限. Dai 等提出了形变卷积网络 DCNN, 利用形变卷积 (Deformable convolution) 和形变感兴趣区池化 (Deformable ROI pooling) 来增强网络对于几何变换的建模能力. 该方法用带偏移的采样代替原来的固定位置采样, 且偏移量是可以通过学习得到, 并不需要额外的监督^[73].

2017 年 ILSVRC 图像分类任务冠军团队 WWW 提出的 SENet (Squeeze-and-excitation networks) 模型, 错误率已经低至 2.251%, 远远低于 5.1% 的人眼辨识错误率^[74].

近两年, 中国团队在计算机视觉挑战赛中表现出色. 商汤科技和香港中文团队、公安部三所、海康威视分别获得 2016 年 ILSVRC 物体检测、物体定位、场景分类冠军^[75]. 而南京信息工程大学和帝国理工学院组成的 BDAT 团队、新加坡国立大学与奇虎 360 合作团队 NUS-Qihoo-DPNs (DET) 分别拿下 2017 年 ILSVRC2017 物体检测 (识别)、物体定位冠军^[76]. 这些团队大都采用了多种网络整合的复合性网络, 如 NUS-Qihoo-DPNs (DET) 提出采用基于 DPN 和基本聚合的深度神经网络模型, 在两类分类定位任务中表现出色, 均取得了最低的定位错误率, 分别为 0.062263 和 0.061941, 刷新了世界纪录. 复旦大学 Li Heng-Duo (Henry) 使用具有软非最大抑制功能的可形变区域 FCN 和 ImageNet 预训练的 ResNet-50, 获得了 2017 年第 2 届 Nexar 挑战赛冠军^[77-78].

表 1 列出了计算机视觉分类任务的 CNN 经典框架. 从表 1 可看出, 基于 CNN 的深度学习模型展示了非常卓越的性能. 从 CNN 框架的演化过程来看, CNN 朝着更深、更准确方向发展的同时, 也有不少工作朝着网络更简化、更灵活、更有效地训练方面努力, 目标是得到更好的特征表示、减少计算复

表 1 基于 CNN 的计算机视觉分类任务经典框架
Table 1 Classical CNN frameworks for computer vision classification tasks

网络结构	特点	备注
LeNet ^[9]	多个卷积层和子采样层	美国手写数字识别
AlexNet ^[60]	提出了 ReLU 和 Dropout	刷新了 2012 年 ImageNet ILSVRC 物体分类竞赛的世界纪录
VGGNet ^[62]	提出采用小卷积核实现更深的网络以及多尺度融合	获 ILSVRC 2014 定位任务冠军、分类任务亚军
GoogleNet ^[65]	22 层网络, 多个 Inception 结构串联	获 ILSVRC 2014 分类和检测任务冠军
ResNet ^[14]	提出了残差网络, 引入跳转连接, 深达 152 层	2015 年 ILSVRC 物体检测与物体识别竞赛冠军
Inception ResNet ^[67]	Inception 结构与 Residual Net 结合	可获得与 ResNet 相当的性能, 但收敛速度加快
FCN ^[68]	密集性预测, 实现了像素级分类	避免了图像块之间的重叠而导致重复卷积计算的问题
DenseNet ^[70]	任何两层之间都有直接的连接	缓解梯度消失, 强化特征传播, 支持特征重用, 并降低网络参数数量
SqueezeNet ^[72]	简化网络结构和减少网络参数	仅需 1/50 的 AlexNet 参数量即可达到了 AlexNet 相同的精度
DCNN ^[73]	提出可变形深度卷积神经网络	增强了网络对于几何变换的建模能力
DPN ^[71]	结合了 ResNet 和 DenseNet 优势	基于 DPN 的团队取得 2017 年 ILSVRC 物体检测与物体识别竞赛冠军
SENet ^[74]	学习每个特征通道的重要程度, 强化有用特征	2017 年 ILSVRC 图像分类任务竞赛冠军

杂度和内存消耗. 这些高效网络拓扑结构不断融合新思想: 一方面, 卷积核趋于小尺度, 网络层次加深, 保证足够的感受野; 另一方面, 为了突破其固定感受野的局限性, 引入多尺度和多处理流思想, 采用多尺度卷积核形成 Inception 结构, 跳转连接提供了多尺度信息的融合, 而形变卷积可以增强网络对于几何变换的建模能力.

3.2 医学图像分析中的深度 CNN

医学图像分类、定位与检测和分割是医学图像分析的三大基本任务. CNN 在医学图像分析中的基本应用是基于图像块 (Patch) 的滑窗图像分类^[9, 79]. 而全卷积网络 (Fully convolutional network, FCN)^[68] 为实现图像语义分割提供了非常好的解决方案.

Sermanet 等采用对包围盒进行预测的方式工作, 实现了一个基于 CNN 的分类、定位和检测综合框架^[80]. 该研究表明, 目标定位与检测任务可以在图像分类的基础上实现. 医学图像分割问题是为图像中的像素或体素指定具有生物学意义的标记, 如解剖结构、组织类别或病灶, 因而, 可将分割问题转化为分类问题, 即对图像中的像素或体素进行特征提取, 然后基于特征将各个像 (体) 素分类. 因此, 定位、检测和分割任务都可以在图像分类的基础上实现.

1) 基于图像块的滑窗分类框架

CNN 可直接应用于分类图像中的每个像素或体素. 典型的像素级分类处理是以滑窗方式处理, 即以某个像素点为中心取一个图像块 (或立方体块) 作为样本训练深度学习模型, 使深度学习模型能够提取到深层特征, 然后基于特征训练分类器. 在测试阶段, 同样地, 以测试图片的每个像素点为中心采样

一个图像块作为输入, 通过模型进行预测, 输出结果作为该像素点的分类值. 基于深度学习的像素级分类方法早期主要采用这种方式. 如 Ciresan 等采用深度 CNN 网络作为像素分类器成功地在电子显微镜像中分割出生物神经膜^[79] 和在乳腺癌病理组织图像中检测有丝核分裂细胞^[81]. 然而, 这种逐像素取图像块进行分类的方法非常耗时, 且相邻图像块之间存在重叠、运算重复、效率低下, 另外一个不足是受图像块大小固定的限制, 无法建模较大的上下文信息, 从而影响分类性能. 为了对更多的上下文信息进行建模, 一类方法是加深网络层次, 另一类方法是引入多尺度和多处理流的思想. He 等提出的 SPP 与跳转连接为构建多尺度和多通道处理流程的信息融合提供了可行方案^[14]. 这种基于多尺度和多处理流的 CNN 框架, 可以处理不同大小的感受野, 有效地结合了局部和全局信息, 从而有效地提高分类准确率^[82–85].

2) FCN 语义分割框架

针对滑窗式分类中因相邻输入图像块之间的重叠而导致重复卷积计算的问题, Long 等提出了 FCN 进行像素级分类, 高效解决语义级别的图像分割问题^[68].

与经典 CNN 提取固定长度的特征向量进行分类不同, FCN 利用卷积和点积都是线性算子且可互换表示的特点, 将全连接层转换成卷积运算, 可以接受任意尺寸的输入图像. 图 4 是 FCN 的网络结构示意图.

具体来讲, FCN 模型的训练过程如下: 首先, 输入图像通过网络模型从左至右地利用卷积层、池化层和相应激活函数进行前馈计算, 逐层提取特征表示; 使用 1×1 卷积层代替全连接层, 并采用反卷积层对最后一个卷积层的特征图进行上采样, 使它恢

复到输入图像相同的尺寸, 从而可以对每个像素都产生一个预测; 然后, 利用分割金标准, 以监督方式通过反向传播误差, 训练调整网络参数. FCN 不需逐图像块计算过程, 保留了原始输入图像中的空间信息, 能够在上采样的特征图上进行逐像素分类, 即能够进行像素到像素的语义分割. 仅由最后特征图直接上采样得到的预测图边缘可能不光滑, 需要将浅层特征和高层抽象特征相融合, 然后再通过上采样得到输出. 这种方式兼顾局部和全局信息, 取得了非常好的分割效果.

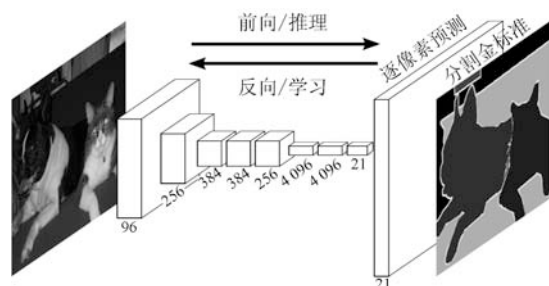


图 4 全卷积网络框架^[68]

Fig. 4 Architecture of fully convolutional network^[68]

最近提出的一些分割方法是在 FCN 的基础上实现的. Ronneberger 等在 FCN 思想基础上, 提出了 U-net 分割框架, U-net 采用对等的上采样和下采样层, 而且, 对应卷积和反卷积层之间存在着跳转连接, 可以结合不同层的特征. 与基于块的 CNN 相比, U-net 框架的优势在于可以考虑整个图像的上下文^[86]. 另一个相似方法将 U-net 扩展到 3D, 使其更适合 3D 医学图像体素处理^[87]. Dou 等提出了面向医学图像 3D 体数据分割框架 DSN, 该框架采用 FCN 进行密集性预测, 并利用逆卷积层将各层特征上采样到与输入相同大小, 有效地融合由粗而精的多尺度信息, 同时, 通过构建直接引导上下层训练的目标函数, 可进行端到端地 3D 深度监督学习, 最后用条件随机场方法 (Conditional random field, CRF) 精化分割结果^[88]. Milletari 等提出了一个专为 3D 医学图像分割而设计的改进型 U-net, 称之为 V-net, 该模型采用 3D 卷积和基于分割衡量指标 DICE 系数作为目标函数, 使之更适用于医学图像分割^[89].

3.3 医学图像分类

3.3.1 图像筛查

图像筛查是深度学习在医学图像分析领域中的最早应用之一, 是指将一个或多个检查图像作为输入, 通过训练好的模型对其预测, 输出一个表示是否患某种疾病或严重程度分级的诊断变量^[37].

图像筛查属于图像级分类, 用来解决此任务的

深度学习模型最初关注于 SAE、DBN 及 DBM 网络和非监督预训练方法. 研究主要集中在神经影像的分析上, 如通过神经影像诊断是否患有老年痴呆症 (Alzheimer's disease, AD) 或轻度认识功能障碍 (Mild cognitive impairment, MCI)^[90-92]. 这些算法通常利用多模态图像作为输入, 提取 MRI、PET 及 CSF 等模态中的互补特征信息. Suk 等分别采用 DBM 和 SAE 从 3D 神经影像图像块中发现其潜在层次特征表达, 构造 AD/MCI 诊断模型^[90-91]. 这两个方法都在 ADNI 数据集^[93] 上进行验证, 结果表明使用 SAE 比 DBM 的模型分类性能更好. 在其他方面也有少量基于非监督模型的图像分类研究, 如 Rahhal 等采用 SSAE 以弱监督方式学习特征, 对心电图信号进行分类^[94]; Abdel-Zaher 等先使用 DBN 无监督学习, 再利用反馈式监督学习调整网络, 对 Wisconsin 乳腺癌数据集分类^[95].

目前, CNN 正逐渐成为图像检查分类中的标准技术, 其应用非常广泛. 如 Arevalo 等提出了乳腺癌诊断的特征学习框架, 采用 CNN 自动学习区分性特征, 对乳房 X 线照片病变分类^[96]. Kooi 等比较了传统 CAD 中手动设计和自动 CNN 特征提取方法, 两者都在约 4.5 万乳房 X 线图像的大数据集上训练, 结果显示 CNN 在低灵敏度下优于传统 CAD 系统方法, 且在高灵敏度下两者相当^[97]. Spampinato 等应用深度 CNN 自动评估骨骼骨龄^[38]. 北京航空航天大学 Xu 等研究了在很少的人工标注情况下, 利用深度 CNN 自动提取特征, 结合多实例学习方法, 对组织病理学图像分类结肠癌^[98]. Gao 等探讨了影响 CT 脑图像分类任务的深度学习技术应用的重要性, 特别是利用 CNN, 旨在为 AD 早期诊断提供补充信息^[99]. Payan 等^[100] 和 Hosseiniasl 等^[101] 则采用 3D CNN 对神经影像进行 AD 的诊断. 另外, Abdi 等应用 CNN 对超声波心动图心尖四腔切面对心动回波质量进行自动分级^[102]. Gao 等通过融合两个 2D CNN 提取超声波心动图的时域和空域信息特征, 对超声波心动图进行视点分类, 从而辅助诊断心脏病^[103].

另外, 还有一些工作将 CNN 与 RNN 结合起来, 如 Gao 等利用 CNN 提取裂隙灯图像中的低层局部特征信息, 结合 RNN 进一步提取高层特征, 对核性白内障进行分级^[47].

3.3.2 目标或病灶分类

目标或病灶的检测与分类可以辅助医生对疾病进行诊断, 如对乳腺病灶进行良恶性分类. 其处理过程通常首先通过预处理方法识别或标记出的特定区域, 然后再对特定区域进行目标或病灶分类. 精确的分类不仅需要病灶外表的局部信息, 而且还需结合

其位置的全局上下文信息。

CNN 框架在病灶分类中应用广泛. Anthimopoulos 等利用 CNN 设计了一个区分毛玻璃状病变、蜂窝病变、钙化和肺小结节等间质性肺病模式的多分类框架, 准确率约 85.5%^[104]. Kawahara 等采用多处理流 CNN 对皮肤病灶分类, 其中每个流程处理不同分辨率的图像^[105]. Jiao 等利用 CNN 提取不同层次的深度特征, 提高了乳腺癌的分类准确率^[106]. Tajbakhsh 等就 CT 图像中检测肺结节且区分良性和恶性肺结节的任务, 比较了大规模训练人工神经网络 (Massive-training artificial neural networks, MTANNs) 与 CNN 这两种端到端训练的人工神经网络的性能, 其实验结果表明, 只有使用较少训练数据时, MTANN 的性能明显高于 CNN^[107].

另一方面, 一些学者把 CNN 与其他基本模型结合起来实现分类. 如 Kallenberg 结合 CNN 与 SAE 的特点, 采用非监督预训练卷积稀疏自动编码器 (Convolutional sparsely autoencoder, CSAE) 模型, 实现乳腺密度分割和乳腺的风险评估^[108]. 而 Van 等结合 CNN 的区分性与 RBM 的生成性特点, 利用 CRBM (Convolutional restricted boltzmann machines) 对肺部 CT 进行分析^[109]. 上海大学 Zhang 等构建点门限玻尔兹曼机和 RBM 的深度结构对剪切波弹性图像提取特征, 实现了更好的肿瘤良恶性分类性能, 准确率达到 93.4%^[110]. 而 Shi 等利用新的深度多项式网络对小样本超声数据集的肿瘤分类, 在胸部和前列腺数据集的分类准确度分别为 92.4% 和 90.28%, 优于基于 DBN 和 SAE 的方法^[45].

3.4 目标或病灶定位与检测

准确地在医学图像中定位特定生物标记或解剖结构在临床治疗中具有非常重要的意义, 直接关系到治疗效果的好坏. 医学图像定位常需要分析 3D 体素信息. 为了使用经典深度学习算法进行 3D 数据处理, 一些方法将 3D 空间看成 2D 正交面的组合, 这样可将定位任务转换成为分类任务, 利用通用深度学习框架进行处理. 如 Yang 等结合三个正交方向 CNN 的信息识别股骨末端的标记, 标记的 3D 位置定义为三个 2D 图块的交点^[111]. 香港中文大学 Chen 等将 FCN 扩展到 3D FCN, 提出的基于 3D FCN 的定位和分割方法, 在 2015 脊椎盘定位和分割挑战赛中取得了非常好的成绩^[112]. Vos 等通过将 3D CT 体积解析成 2D 形式, 识别目标 3D 矩形包围盒, 进而定位到感兴趣的心脏、主动脉弧和下降主动脉等解剖区域^[113]. 另外, LSTM 也用于处理包含在医学视频中的时间信息, 如 Kong 等结合 LSTM-RNN 和 CNN 检测心脏 MRI 视频中的舒张末和收缩末^[114]. 而 Cai 等采用深度 CRBM

以无监督方式提取和融合不同模态图像特征, 识别 MR/CT 图像中脊椎^[115-116].

图像的感兴趣目标区域 (Region of interest, ROI) 或病灶检测是诊断的关键部分, 在计算机辅助检测系统中已有较长的研究历史. 通常设计成自动检测病灶, 用以提高检测准确率或减少专家的读片时间^[37]. 实现过程分两个步骤: 在全图像空间中进行定位 ROI; 在感兴趣区进一步识别出小病灶区.

早在 1995 年, Lo 等就提出了第一个利用 RNN 的目标检测系统, 该系统使用 4 层的 CNN 在 X 光图像中检测结节^[117]. Ciresan 等采用深度 CNN 网络作为像素分类器成功地在乳腺癌病理组织图像中检测有丝核分裂细胞^[81]. Sirinukunwattana 等利用空间约束的 CNN, 结合邻域集成预测器, 有效地提高了在结肠癌病理图像中检测和分类细胞核的准确率^[118]. Li 等提出了基于深层 CNN 分类的青光眼检测方法^[119]. Roth 等利用深度卷积神经网络学习方法, 提高了 CT 图像目标检测的准确性. 他们的主要思想是用现有其他方法提取候选 ROI, 然后基于深度卷积神经网络学习目标隐含的高层次特征, 并根据这些特征进行目标对象检测和分割. 他们分别在腹部 CT 图像自动淋巴结检测^[120]、硬化转移和结肠息肉检测^[121-122] 等应用领域明显提高了目标检测精度. Wang 等利用 12 层 CNN 网络在乳房 X 光图像中检测乳房动脉钙化, 钙质量定量分析结果表明, 由此检测结果推断钙质量接近于金标准, 准确度达到了 96.24%^[123]. Quellec 等利用图像级标签监督训练基于 CNN 的病变检测器, 使用灵敏度分析检测在图像级预测中起作用的像素, 获得图像大小的热点图, 通过在训练 ConvNets 的同时增强热点图的稀疏性, 避免热点图受到伪影的影响. 该方案在 Kaggle 彩色眼底图像检测糖尿病性视网膜病变挑战赛数据集上获得了很好的成绩^[124].

图像的感兴趣目标或病灶检测的关键是对每个像素进行分类. 目前大多数基于深度学习目标检测系统采用 CNN 执行像 (体) 素分类任务, 之后采用某种形式的后处理方式得到目标. CNN 框架和方法类似于第 3.2 节中讨论的像 (体) 素级分类方法, 需要结合被分类像 (体) 素的邻域上下文或 3D 信息提高分类准确率, 如使用多处理流 CNN 整合不同视角剖面信息^[84] 或多模态图像数据^[125]. Albarqouni 等采用多尺度 CNN 方案在乳腺癌病理图像中检测有丝分裂^[126]. 香港中文大学 Chen 等利用多个 2D 深层特征近似表达 3D 医学图像的特征, 结合 SVM 分类器, 实现敏感性加权图像 (Susceptibility weighted imaging, SWI) 自动检测脑微出血 (Cerebral microbleeds, CMBs)^[127]. Dou 等在文献 [127] 基础上改进, 采用级联 3D CNN 框架,

充分利用 MR 中的空间上下文信息, 提取更能代表 CMBs 的高级特征, 该方法在具有 320 个 MR 图像的大数据集上进行了广泛的实验验证, 实现了 93.16% 高灵敏度^[128]. 该研究小组还用类似多级 3D CNN 框架在 CT 序列中检测肺结节, 在 ISBI 2016 举办的 LUNA16 挑战赛中得到验证, 该算法在减少假阳性指标上取得了 2016 年的最高成绩^[129]. van Grinsven 等利用 CNN 提取特征, 并采用正负样本均衡策略, 有效地实现了在眼底图像中检测出血^[130].

有少量工作采用其他深度学习方法来实现感兴趣目标或病灶检测, 如 Shin 等将 SAE 深度学习方法应用于 MRI 图像上检测腹部器官, 先以非监督方式学习空间特征, 然后基于“兴趣点”进行多器官检测^[52]. 南京信息工程大学 Xu 等利用 SSAE 网络学习深度特征, 识别组织病理学图像乳腺癌的细胞核, 确定乳腺癌的分期^[131]. Masood 等针对标记训练数据不足的情况, 提出基于结合 DBN 和 SVM 的半监督学习算法, 对皮肤镜黑色素瘤自动识别^[132]. 与传统 CNN 方法将原图像作为输入不同, Li 采用 Sobel 边缘轮廓特征和 Gabor 纹理特征作为输入, 利用 CNN 进行特征融合和深度特征提取, 提高了 C-arm X 射线图像自动检测腰椎准确率^[39].

最近有研究还将基于 CNN 的方法应用到检测和定位手术视频中的工具. 如 Girshick 等和 Sarikaya 等利用基于候选包围盒检测模型 Faster R-CNN 从机器人辅助手术视频中定位和检测自动化工具^[133-134]. Twinanda 等设计了一个新 CNN 框架 EndoNET 自动地从胆囊切除视频中学习视觉特征, 同时实现手术阶段 (Phase) 识别和术中多工具检测任务^[135]. 而 Chen 等提出了一个结合 CNN 和 LSTM 的方法, 自动检测超声影像中多个标准平面, 为实质性生物测量和诊断提供帮助^[136].

3.5 医学图像分割

3.5.1 器官与组织分割

医学图像中器官及其子结构的分割可用于定量分析体积和形状有关的临床参数, 如心脏的心室体积和收缩射出率. 另一方面, 在采用智能调强放疗技术对肿瘤进行治疗时, 危及器官勾画是制定放疗计划中非常重要的步骤之一. 深度学习在此任务中应用非常广泛, 主要应用于:

1) 组织病理学图像和显微镜图像分割

通过计算机提取来自手术和活检组织标本的图像特征可以帮助预测疾病侵袭性的程度, 从而进行疾病诊断和分级. 这些预测器的关键组成部分就是从组织病理图像挖掘的图像特征^[32].

最近, 深度学习方法已在组织病理学图像分割

中应用. 目前大多数组织病理学图像和显微镜图像分割方法都是基于 CNN 的. 许多学者利用图块训练网络取得了非常优秀的分割结果. Ciresan 等率先将深度 CNN 应用于医学图像分割, 他们以滑窗方式在电子显微镜图像中分割出生物神经膜^[79]. Kumar 等利用基于块的 CNN 对 H&E 染色的病理学图像进行细胞核分割^[137]. Xu 等则采用迁移学习方法, 利用 ImageNet 训练深度 CNN 作为特征提取, 在 MICCAI 2014 脑肿瘤数字病理挑战赛中取得了 97.5% 的分类准确度和 84% 的分割准确度^[138]. Qaiser 等利用 CNN 提取图像块特征, 构建基于拓扑特征的持续同调分布, 用于 H&E 染色的组织学图像的自动肿瘤分割^[139].

结合多分辨率和多处理流方法可以更好地结合上下文信息, 提高分割准确性, 如 Song 等采用自粗而精的多尺度 CNN 方法在宫颈癌病理图像中分割细胞核^[84, 140]. Fakhry 提出的残差反卷积网络由两个通路组成, 捕获全分辨率特征上下文信息, 用于电子显微图像 EM 的分割, 该方法采用端到端地训练, 提出的模型包括足够的上下文信息, 能够实现像素级预测, 是 EM 图像中分割 3D 神经突触开放挑战赛的最好方法之一^[141]. 香港中文大学的 Chen 等基于 FCN 的多级特征表示, 提出了一个高效深度轮廓感知网络 DCAN, 在 MICCAI' 2015 举行的结肠组织学图像腺体分割挑战赛中取得了冠军^[142]. Xu 等综合 FCN 前景分割、DCAN 边缘检测和基于包围盒的目标检测三个深度通道信息, 在 H&E 病理图像中进行腺体分割, 且该框架可扩展应用到其他分割任务中^[143]. 在 MICCAI' 2015 结肠组织学图像腺体分割挑战赛中表现出色方法大多是基于 CNN 或 FCN. 更多关于病理图像细胞核及腺体分割的文献, 可参考综述 [32, 144].

另外, Xie 等使用空间 RNN 去分割 H&E 组织病理学图像的肌束膜, 该网络考虑了当前图像块的行和列先验. 为结合左上角和右下角邻域的双向信息, RNN 在不同方向应用 4 次, 然后联合输入到全连接层, 为单个图块产生最终输出^[145].

为了得到好的分割结果, 有些工作将 CNN 分类的结果作为分割初始值, 利用水平集等形变模型或稀疏形状模型等精化细胞核分割结果^[140, 146].

2) 脑组织和结构分割

若中枢神经系统发生病变, 神经细胞的代谢紊乱, 将使脑脊液 (Cerebrospinal fluid, CSF) 的性状和成分发生改变, 因此, 脑脊液的检测成为重要的中枢神经系统受损辅助诊断手段之一. 白质 (White matter, WM) 异常的相关疾病很多, 包括多发性硬化症、精神分裂症、阿兹海默症等, 因此对中枢神经系统进行有效地分割出灰质 (Grey matter, GM),

WM 和 CSF 是研究热点之一. Zhang 等将 T1、T2 和 FA (Fractional anisotropy) 三种模态图像作为输入, 采用深度 CNN 网络解决婴儿 GM、WM 和 CSF 分割这一挑战性任务, 从而对婴儿脑发育优劣程度作评估^[147]. Stollenga 等为了利用 GPU 并列处理, 将传统的多维度 MD-LSTM 重新排列成金字塔的方式, 即 PyraMid-LSTM, 该模型在 MRBrainS13 数据集上的像素级脑部图像分割成 GM、WM 和 CSF, 取得了 2015 年比赛的最好效果^[148].

在采用智能调强放疗技术对脑肿瘤进行治疗时, 需要保护周围重要脑组织和结构, 因而正确地分割脑组织和结构有着重要的临床意义. 上海交通大学 Cui 等提出了一种基于图像块的方法使用 CNN 自动分割脑 MRI, 该方法从公共 MRI 数据集中抽取一部分图像块样本训练 CNN, 训练好的 CNN 用于脑 MRI 的白质、丘脑, 侧脑室的分割, 分割准确率达到了 90 %^[149]. 还有一些学者采用多尺度方法, 借助多 CNN 处理流程提高分割的准确率. 如 Choi 等采用两个 CNN 处理流程: 全局 CNN 确定纹状体的大致位置, 局部 CNN 预测所有体素的标签. 该方法应用于脑结构的纹状体分割, 实验表明其分割结果的平均 DICE 系数达到了 0.893^[83]. Moeskops 等采用多尺度 CNN 方法进行脑组织分割, 该方法在 8 个组织分类上取得了好结果, 在 5 个不同年龄段数据集 (从婴儿到中老年) 上进行验证, 分割结果的 DICE 相似系数分别为 0.87, 0.82, 0.84, 0.86 和 0.91^[150].

还有一些采用其他深度学习框架的分割方法, 如 Dolz 等提出了基于 SSAE 深度学习方法在罹患脑癌情况下从 MRI 图像中分割脑干^[151]. Mansoor 等利用 SAE 从脑 MRI 序列中提取有助于区分对比度差区域的深度特征, 如视束和病理部分, 并提出了基于局部形状和稀疏外观的前视觉通路自动分割方法, 该方法的鲁棒性与手动分割相当^[152].

最近的相关文献采用 FCN 方法减少基于滑窗分类的冗余计算, 如 Shakeri 等将 FCN 应用于脑皮层结构分割^[153]; Moeskops 等训练单个 FCN 分割脑 MRI、分割胸 MR 中胸部肌肉, 及在心脏 CT 血管造影图像中分割冠状动脉^[154].

3) 血管分割

深度学习还应用到非常有挑战性的血管分割任务中. Nasr-Esfahani 等采用基于图像块的 CNN 算法从 X 光血管造影图像中提取血管^[40]. Wu 提出识别血管结构的通用方法, 使用 CNN 学习目标的外观特征, 然后, 利用基于主成分分析的最近邻搜索估计局部结构分布, 进一步结合广义概率跟踪框架提取整个眼底血管连接树^[155]. Liskowski 提出一种基于监督的眼底血管分割技术, 使用深层 CNN 神经

网络, 采用预处理好的大样本 (高达 400 000) 训练, ROC 达到 0.99, 准确率达到 0.97, 网络性能大大超过了先前的算法^[156]. 山东大学 Wang 提出了一个综合 CNN 和随机森林的监督方法来解决视网膜血管病变分割的问题^[157]. 在这些方法中, 大多数是采用监督深度学习提取特征, 再结合其他已有技术和分类器保证分割的准确性. 与前述通过分类实现分割的研究不同的是, Li 等将视网膜血管分割问题转换为从视网膜图像到血管映射图的跨模态数据变换问题, 采用深度 AE 学习视网膜图像与血管映射图之间的关系^[158].

4) 心脏心室分割

从心脏 MRI 数据中分割出左心室是计算心室体积和收缩射出率等临床指标的重要步骤之一. Carneiro 等采用基于 DBN 学习特征对左心室外观建模, 利用监督学习模型在心脏超声波图像中自动分割出左心室^[46], 而 Avendi 等采用 SAE 学习深度特征初步推断左心室的形状, 再结合形变模型提高左心室分割的准确性和鲁棒性^[159]. Ngo 等结合 DBN 深度学习和水平集方法, 从心脏 MRI 图像自动分割出心脏左心室^[160]. Dou 等提出的基于 FCN 的 DSN 框架及 CRF 精化方法在心脏及主动脉的分割性能达到了领先水平^[88]. Tan 等通过 CNN 回归方法得到左心室中心, 利用两尺度 CNN 网络结合外轮廓之间的径向距离等形状先验参数来分割左心室^[161]. Zhen 等利用多尺度的 CRBM 进行无监督学习, 然后利用标签数据监督训练回归森林预测器, 从 MR 图像中可直接估计出双心室体积^[162].

深度学习还应用到其他方面, 如 Roth 等利用深度卷积神经网络学习方法, 实现了腹部 CT 图像胰腺全自动分割^[163–164]. Zilly 等提出基于集合学习的 CNN 架构方法从视网膜图像分割视盘^[48]. Guo 等联合 SSAE 模型深度特征学习和稀疏块匹配的前列腺分割方法^[165]. Korez 等使用 3D FCN 产生脊柱体相似度映射图, 驱动形变模型在 MR 图像中分割脊柱体^[166]. Yang 等利用 CNN 学习胸部 X 光照片的梯度与相应的骨骼图像之间的映射, 利用级联 CNN 逐渐细化预测的骨骼梯度, 最后将不同尺度的预测骨骼梯度融合到最大后验框架中, 以产生骨骼图像的最终估计. 该方法通过从原始 X 光照片中减去骨骼图像的估计, 从而起到抑制骨性结构影响的作用^[85].

从深度学习应用框架来看, 目前大多数图像分割方法都是基于 CNN 的. 许多学者利用基于图像块方式训练网络取得了很好的分割结果, 相关应用实例可参见文献 [40, 48, 79, 138–139, 147, 149, 163–164]. 结合多分辨率和多处理流方法可以更好地结合上下文信息, 提高分割准确性, 相关应用实

例可参见文献 [83, 141, 150]. 最近的文献有很多使用 FCN 方法减少基于滑窗分类的冗余计算, 且已扩展成 3D FCN 模型并应用到多目标分割中, 相关的例子可见文献 [88, 142, 153–154, 166]. 由 FCN 衍生出的 U-net^[86–87] 以及针对 3D 图像分割设计 DSN^[88] 和 V-net^[89] 在最近的图像分割任务中表现出色.

有少量工作采用其他深度学习方法来实现, 如采用 SAE 的方法有文献 [151, 159, 165], 采用 DBN 的实例有文献 [46, 160]. 另外, RNN 在分割任务中越来越流行. 例如, Xie 等使用空间 RNN 考虑当前图像块的行和列先验去分割 H&E 组织病理学图像的肌束膜^[145]. 多维递归神经网络 (MD-RNN), 特别是采用 LSTM 的 RNN, 可以通过几次扫描所有像素接收空-频信息, Stollenga 等利用 PyraMid-LSTM 将脑部图像分割成 GM、WM 和 CSF, 取得了 2015 年最好效果^[148]. Poudel 等结合 2D U-net 框架和 GRU-RNN, 提出了递归全卷积网络, 从 MRI 图像中分割心脏及左心室^[167].

基于体素分割方法挑战之一是存在假阳性, 不能确保分割的结果在空间上是连通的, 有可能存在孔洞或毛刺. 为了解决这个问题, 将深度学习方法与图割^[84]、MRF、CRF^[168–169] 和水平集方法^[160] 相结合可精化分割结果. 在大多数情况下, 图模型作为标记的正规化子应用在 CNN 或 FCN 产生的相似性映射图上, 用以生成连贯的精化区域. Gao 等提出了基于分割的标签传播方法, 利用深度 CNN 进行分类, 然后再结合 CRF 获得了高精度结果^[170].

3.5.2 病灶与肿瘤的分割

在对病灶或肿瘤进行任何治疗之前, 关键是准确地分割出病灶或肿瘤, 确保在治疗期间杀死肿瘤细胞的同时保护正常组织器官^[151]. 为了得到准确分割, 病灶与肿瘤分割通常需要结合多模态图像信息以及全局与局部上下文信息, 因此采用多模态图

像信息作为网络模型输入、不同尺度的多处理流网络^[82]、3D CNN^[88] 和非均匀采样图块的策略^[82, 171].

为方便比较, 我们梳理了近三年来采用脑肿瘤分割基准 (Brain tumor image segmentation benchmark, BRATS) 公共数据集^[172] 验证的几个具有代表性的优秀算法, 如表 2 所示. 这些算法都是基于 CNN 的方法, 值得注意的是, Kamnitsas 等采用全深度 3D CNN 网络, 利用多处理流网络结合全局和局部上下文信息, 在非常具有挑战性的包括从多模态数据分割创伤性脑损伤、脑肿瘤和中风等病变分割任务中展现了出色的性能, 特别是在脑肿瘤整体分割水平上已超过了专家的勾画水平^[168].

Yu 等结合残差网络和全卷积网的特点, 构造了深度残差全卷积网络 (Fully convolutional residual network, FCRN), 在皮肤镜图像中自动分割黑色素瘤, 在 ISBI2016 挑战赛中取得了第 2 名的成绩^[49].

在病灶分割中, 也看到了 U-net 的应用及利用全局和局部的相似框架. 如 Wang 等采用类似于 U-net 的结构, 由相同的下采样和上采样通路, 但没有使用跳转连接^[173]. 另一个类似 U-net 的框架应用于分割多发性硬化病灶, 但是使用了 3D 卷积, 且第一个卷积层和最后一个反卷积层之间存在单个跳转连接^[171].

值得注意的是, 图像中大多数像 (体) 素是属于正常组织, 病灶分割的一大挑战是类分布的非均衡性. 有关类分布非均衡性的应对策略将在第 4.2 节讨论.

4 挑战与对策

深度学习是以数据驱动方式学习各级抽象特征, 这些特征在很多应用领域表现了非常强大的特征表示能力和鲁棒性. 尽管深度学习在以自然图像为分析处理对象的计算机视觉任务中展示了卓越的性能, 但将深度学习方法成功应用到医学图像分析领域仍具有很大挑战性, 主要表现在以下几个方面:

表 2 脑瘤分割方法比较 (使用 BRATS 数据集验证)

Table 2 Comparison of methods for brain tumor segmentation (validation on BRATS database)

作者	方法	DICE		
		总肿瘤区	核心肿瘤区	活性肿瘤区
专家评定	医学训练和经验	0.88	0.93	0.74
Urban ^[174]	多模态输入, 训练 3D CNN	0.87	0.77	0.73
Zikic ^[175]	将 3D 立方体图像块转换成 2D 图像块, 训练 2D CNN 网络	0.837	0.736	0.69
Havaei ^[82]	2D 多模态输入, 双路径级联 CNN 架构, 综合了局部细节和更全局信息	0.88	0.79	0.73
Pereira ^[176]	3 × 3 的小的小卷积核, 更多的 CNN 层数和非线性运算, 更少的滤波器权重	0.88	0.83	0.77
Kamnitsas ^[168]	采用深度为 11 层的小滤波器 3D CNN 的双路径网络框架	0.898	0.75	0.721

1) 自然光学图像不仅空间分辨率、对比度比大多数医学图像高,而且还具有亮度、颜色、纹理等易于识别的许多自然视觉特征.而大多医学图像只有某种特殊信号的强度值,而且通常信噪比很低,因此,大多数医学图像中各组织器官解剖结构、病灶区域之间的分界线并不清晰,纹理差异也不大,同时还会因个体差异和成像原理不同而显著不同,因此,医学图像分析比自然图像分析更困难.

2) 医学图像各种成像方法存在局限性,不同模式的医学图像数据仅能反映人体内部的特定解剖、功能信息,各有优缺点,且不同成像设备和图像重建方法存在较大差异,临床上通常需要同时使用多种不同成像原理和成像方法的医学图像进行辅助分析诊断,导致医学图像自动分析处理比自然图像更复杂、更困难.

3) 目前计算机视觉中的很多分类任务是图像级的,而医学图像不仅仅用于图像级的疾病筛查,而且用于像素、体素级的治疗计划制定.如肿瘤调强放射治疗计划要求对肿瘤和危及的正常组织器官进行精确检测、识别、定位,并从 CT、MRI、PET 等医学图像中高精度地分割出肿瘤放疗靶区及其周围正常组织器官.而且异常病灶区域(肿瘤)非常复杂,其出现的位置、大小、形状个体差异很大,因而异常病灶区检测、识别和分割比正常组织器官更具挑战性,医学图像分析比自然图像分析计算复杂度更大,许多计算机视觉中的深度学习算法不能直接用于医学图像分析.

4) 自然图像分类任务中,业界已经建立了带有人工标注的超大规模的学习训练样本数据集,如 MNIST、CIFAR、ImageNet,而应用深度学习进行医学影像分析时,获取大规模的学习训练样本数据集非常困难,特别对病灶样本数据集,因为其变化很大,且需要临床专家标注,因此,可用的标注数据规模相对很小^[31, 35, 47].但是,异常的病灶检测、识别和分割具有重要的临床应用价值,比如自动筛查、自动诊断以及自动制定治疗计划.

5) 获得大规模医学数据分析标注样本数据困难. a) 它很难获得资金资助用于建设数据集; b) 需要稀缺和昂贵的医疗专家对高质量的医学图像数据进行标注; c) 医学图像大多位于医院封闭私有的数据库中,隐私保护条例可能会阻碍这些数据的访问,使得共享医疗数据比共享自然图像更难^[34]; d) 医学图像分析的应用很广,需要收集许多不同的数据集.

为了提高特征表示能力和分类准确率,需要足够的医学图像训练样本数据.当样本数据量不足时该怎么处理?如何使用小规模的学习训练样本数据最有效?如何利用各模态互补信息和图像空间上下文信息提高医学图像分类准确率?如何获取大规模的

医学数据分析标注样本数据?目前,主要的应对策略如下:

4.1 迁移学习与弱监督学习

CNN 的主要潜力在于能够在多层抽象中提取一系列区分性的特征.如前所述,深度 CNN 模型是监督学习模型,从零开始训练 CNN 是一大挑战.

1) CNN 要求大量标记训练样本数据,医学领域专家标注非常昂贵,有些病例(比如肿瘤)稀少;

2) 要求大量的内存和计算资源,否则训练过程将异常费时;

3) 训练深度 CNN 常会因过拟合和收敛问题而变得非常复杂,常需要重复调整框架或者网络中的学习参数^[34].

针对这一挑战,采用迁移学习方法,借助自然图像或来自不同医学领域的数据集以监督方式预训练 CNN 网络.迁移学习有两种典型使用策略:

1) 使用预训练网络作为特征提取器.研究发现,尽管医学图像与自然图像在成像原理和图像外观上有非常大的差异,用 ImageNet 训练好的 CNN 模型仍可用于医学图像识别应用^[177].如 Bar 等将预训练网络用作胸部病理学识别特征生成器^[178-179]. Ginneken 等将 CNN 特征与手工特征结合起来,提高了结节检测系统性能^[180].

2) 使用目标医疗数据微调预训练网络.将预训练 CNN 作为监督学习网络初始值,利用任务相关的有限标注数据集以监督方式对网络参数进行微调,调整网络中的几个或全部层^[181-182].Ciompi 等采用自然图像集 ImageNET 预先训练的卷积神经网络,用少量有标记 CT 数据序列进行有监督地学习精调网络,可自动检测肺裂周围结节^[183].Tajbakhsh 等分析表明深度微调比浅层精调性能有所提高,且训练集规模减少时使用精调网络的重要性提高了^[184].

两种策略都已得到广泛应用,前者的好处是不需要训练所有深度网络,方便将提取的特征替换插入到现有的医学图像分析流程中,但哪种策略效果最好目前尚未定论^[37].

除了采用迁移学习,另一种策略是弱监督学习,有效结合无监督预训练和监督学习的优点.虽然医学图像可用标注数据较少,但其本身的规模可能很大.可以充分利用未标注数据,采取预训练技术,以非监督的方式来提取该类图像的一般特征.然后将提取到的特征作为有监督学习的网络初始值,利用标注数据训练分类器.如 Kallenberg 等采用卷积稀疏自动编码器(Convolutional sparsely autoencoder, CSAE)进行乳腺密度分割和乳腺风险评估,CSAE 和经典 CNN 的主要区别在于 CSAE 卷积层以 SAE

的非监督方式逐层预训练^[108]。在获得目标标注数据非常昂贵的情况下, 利用多实例学习 (Multiple instance learning, MIL) 和深度学习相结合的方式也是一个可行替代方案, 如 Xu 等研究了使用结合监督和非监督特征学习的 MIL 框架, 结果表明基于 MIL 框架的性能优于手工特征, 与完全监督方法性能接近^[98]。另外, 有研究得到了一些有意思的结论, 即采用非专业医生完成标注工作, 其效果不比医学专家标注的差^[185]。Albarqouni 等提出由非专业的众包服务平台来完成标注工作, 或许是另一个可行解决方案^[126]。

4.2 正则化与均衡化

CNN 训练是多批次迭代优化网络参数的过程, 在每批次迭代中, 从训练数据随机选择样本子集, 输入到网络, 并通过反向传播更新其参数, 以最小化目标函数。使用医学图像数据需要考虑以下两个挑战:

1) 医学成像设备和图像重建方法差异。由于医学图像成像设备及图像重建方法存在差异, 可能会引起灰度不均匀和偏移场不一致的问题^[186]。虽然监督学习技术在使用标准化成像协议的有限实验中表现出很大的潜力, 但是可能在稍微不同的条件下获取的新图像上性能快速恶化。

针对以上问题, 可考虑进行以下处理来削弱其不良影响:

a) 批规范化: 对训练样本中每个小批量样本进行规范化, 通过将其均值标准化为 0, 方差归一化, 避免梯度消失和梯度溢出, 同时可以加速收敛, 可作为一种正则化技术来提高网络的泛化能力^[187]。

b) 正则化: 正则化是降低过拟合的有效策略, 通过在模型代价函数中加入正则化项 L_1 和 L_2 , 降低模型的复杂度, 从而减少过拟合^[82]。

c) Dropout: 在每次训练迭代中, 随机将比例为 p 的部分神经元输出设置为 0 (即解除部分结点的连接), Dropout 是避免网络过拟合的随机正则化策略, 同时也可看作隐式的模型集成^[188]。

2) 样本分布不均衡。医学领域的分类任务通常需辨别正常组织器官与病灶区域。每个病例中各种组织类别的数据分布是非均衡的, 而且正常组织器官类会被过度表示, 且大多数正常组织器官训练样本是高度相关的, 只有一部分正常组织器官样本的信息量大。例如, 脑肿瘤分割是一个高度数据不均衡问题, 其中健康组织体素约占总体素的 98%, 剩余约 2% 的病理体素中, 包括 0.18% 属于骨疽, 1.1% 的水肿, 0.12% 的非强化肿瘤, 0.38% 的强化肿瘤^[82]。在学习过程中均等地对待这些数据会导致许多训练迭代浪费在非信息样本上, 使得 CNN 训练过程耗费不必要的时间, 同时, 这种以健康组织样本主导的训

练, 会导致训练出的 CNN 模型存在问题。针对这个问题, van Grinsven 等通过识别信息量大的正常样本以及在训练过程中动态选择误分类的负样本, 提高 CNN 的学习效率并减少了训练时间^[130]。Havaei 等提出采用两次训练策略, 即初始时挑选所有类别等概率的图像块数据集进行训练, 然后, 保持其他各层的核固定, 仅在输出层使用更能代表真实分布样本进行重新训练^[82]。通过这种方式, 可以平衡应对所有类的多样性, 同时, 输出概率通过利用数据标签的真实分布再训练被校正。Brosch 等调整了损耗函数, 用加权敏感度和特异性的结合来定义, 特异性的权值越大, 使它对数据非均匀性越不敏感^[171]。还有一些方法通过在正样本上进行数据增强处理来平衡数据^[168, 176]。

4.3 融合多模态互补图像信息和图像空间上下文信息

由于医学图像成像的局限性, 不同模态 (成像原理) 的医学图像数据仅能反映人体特定的信息, 各有优缺点。如 CT 与 MRI 在骨组织与软组织成像上互补, CT、MRI 与 PET 在解剖成像与功能成像上互补。即使同一类成像, 利用不同对比度增强方法得到的图像也各有千秋。如不同的 MRI 模态产生不同类型的组织对比图像, 从而提供有价值的结构信息, 使肿瘤及其亚临床区域的诊断和分割成为可能。通常, T1 图像用于区分健康组织, 而 T2 图像对水肿区域敏感, 在图像上相应区域产生亮信号。在 T1-Gd 图像中, 肿瘤边界可以容易地通过在活性细胞区域中积累的造影剂 (钆离子) 的信号亮度来区分肿瘤组织。在 FLAIR 图像, 水分子的信号被抑制, 这有助于区分水肿区域脑脊髓液^[189]。联合多模态医学影像, 利用不同影像的互补信息, 可以提供清晰的功能和解剖结构信息, 从而提高分析的准确性, 因此大多分割算法是采用多模态医学图像作为输入^[147-149, 174-175]。

为了实现像素级的医学图像分类/分割, 仅有被分类像素点本身的信息是远远不够的, 还需要结合邻域像素, 提供更好的上下文信息。大多数医学图像是 3D 信息, 为了更好地考虑被分割体素上下文信息, 除了加深 CNN 网络层次、采用多尺度和多处理流 CNN 方法外, 还需要考虑如何对 3D 信息建模的问题。目前有以下两种不同的处理方式对 3D 信息建模:

1) 将 3D 体信息近似转换成 2D 图像块信息。以被分类体素为中心, 取多个视角的 2D 剖面近似表达其 3D 邻域上下文信息, 借助 2D 卷积简化计算, 提高计算效率^[120, 175]。将每个剖面视为一个信息源, 借助多通道或多处理流方式整合所有剖面信

息, 提高分类准确率. 如 Roth 等采用三个正交剖面输入到 CNN 三通道中, 检测结肠息肉或可疑淋巴结^[120]. Setio 等使用一个多处理流 CNN 框架对胸部 CT 感兴趣点是否为结节分类, 通过对感兴趣点提取了多达 9 个不同方向的剖面图像块作为多处理流 CNN 的输入, 最后在全连接层合并这些特征, 以获得最终分类结果^[190].

2) 利用 3D CNN 来处理. 以被分类体素为中心, 提取其 3D 立方体块, 借助 3D 卷积提取特征. 这种方法全面考虑 3D 邻域信息, 有助于提取更具区分性的特征, 一般来说, 分类准确率更高. 缺点是参与运算的 3D 体素多, 计算效率偏低^[168]. 随着计算速度的提高, 3D CNN 在近两年得到了广泛采用. 如 Nie 等利用 3D 本身信息训练 3D CNN, 用以评估重度胶质瘤病人生存率^[191].

另外, 通过与其他算法相结合, 可以进一步提高分类/分割的准确性和精化分割区域边界. 如先采用超像素分割方法生成候选区域 ROI, 再利用深度学习提取深度特征, 可以减少搜索空间以提高计算效率, 提高分类/分割准确性^[84, 120-121]; 结合深度学习和水平集方法, 从心脏磁共振 (MR) 数据中自动准确地分割出的心脏左心室^[160].

4.4 大规模的医学数据分析标注样本数据

一方面, 医学图像领域的挑战赛及大型公开数据集的出现对医学图像分析的发展有着极大的帮助. 自 2007 年以来, MICCAI、ISBI 和 SPIE 等医学成像研讨会组织挑战赛数据集已经成为惯例, 开放了大量用于基准研究的医学数据集^[34], 收录在网站 <http://www.grand-challenge.org/>. 使用这些公开基准数据集的明显优势在于其提供了任务和评估度量指标的精确定义, 便于公平、标准化地比较各种算法的性能. 在 Kaggle 近两年组织的彩色眼底图像糖尿病性视网膜病变检测及分期^[192] 和 MRI 图像分析测量心脏体积并导出射血量^[193] 的两次竞赛中, 顶级竞争者都使用了 CNN 网络. 同时, 国内也开始举办大型医学图像分析挑战赛, 如天池医疗 AI 大赛正在进行中, 该大赛旨在用深度学习等 AI 方法在肺部 CT 中检测肺小结节^[194]. 预计使用更大的数据集进行挑战的趋势将会继续, 而深度学习将会在最佳解决方案中占主导地位.

另一方面, 临床医学数据的收集为深度学习提供了另一个解决方案. 在大多数西方医院临床中 PACS (Picture archiving and communication systems) 系统使用至少有十年, 其中有上百万的医学图像, 且可数字化有效存取. 而我国人口基数大, 且县级以上的大多医院都陆续采用 PACS 系统, 潜在可用的医学图像规模非常大. 目前, 多地陆续成立

了医疗大数据中心及研究院. 2016 年, 我国已经启动国家重点研发计划“精准医学研究”专项项目, 如中南大学信息安全与大数据研究院承担的医学大数据安全与隐私保护课题, 可望在解决医学图像隐私问题的同时, 为深度学习研究提供可用的医疗大数据. 目前, 我们项目组承担的国家自然科学基金项目“肿瘤 PET/CT/MRI 纹理谱深度学习分析建模与智能放疗应用”进展顺利, 通过与北京 301 医院和湖南省肿瘤医院合作, 以临床医生勾画的肿瘤放疗靶区和危及器官信息作为参考金标准学习训练样本集, 搭建了结合多尺度信息的 3D CNN 靶区分割框架, 以 39 个鼻咽癌病例 PET/CT 图像数据进行训练, 用 8 例病例进行测试, 初步实验的 DICE 准确率为 0.68. 而以 PET/MRI 图像数据进行训练模型的准确率要比 PET/CT 好很多, 以 12 个鼻咽癌病例 PET/MRI 图像数据进行训练, 用 2 例病例进行测试, 实验的 DICE 准确率达到了 0.925, 目前正广泛收集更多病例进行验证, 有望获得更有说服力的结果.

5 开放的研究方向

综上所述, 深度学习具有自动地从数据中学习深层次、更具鉴别性特征的能力, 已应用于医学图像分析的多个研究领域, 并取得了突破性进展. 我们注意到, 在大多数文献中, 使用深度学习相关方法展示了其领先水平的性能, 这已由医学图像分析的若干计算挑战赛结果证明; 其次, 云计算和多 GPU 高性能并行计算技术的发展, 使得深度学习从海量的医学图像大数据中学习深层特征成为可能; 最后, 可公开访问的相关医学图像数据库的出现, 如脑瘤 MRI 图像 BRATS^[173] 数据集、老年痴呆症神经图像 ADNI 数据集^[93]、缺血性中风 ISLES 数据集^[195] 及多个医学图像分割挑战赛数据集, 使得基于深度学习的分割算法能够得到有效验证^[33, 36].

从现有文献梳理中可看出, 目前代表领先水平的深度学习方法大多是使用监督学习的方式, 特别是基于 CNN 的框架. 前期的研究侧重于预训练 CNN 和将 CNN 用作特征提取器, 可通过简单下载这些预训练网络并将其直接方便地使用到任何医学图像分析中. 而近两年端到端训练 CNN 成为医学图像分析的优先方法. 然而, 获得用于监督学习的标注数据是深度学习方法应用于医学数据分析的一大挑战^[31, 35]. 在目前已标记训练数据有限情况下, 充分利用医学图像中的非标注图像, 结合非监督和监督学习的优点, 进行弱监督和无监督学习将成为医学图像分析的一个重要发展方向.

另外, 电子病历和医生专家的文本报告中包含了丰富的以自然语言描述的临床诊疗信息, 这些信

息可作为图像标记数据的补充. 计算机视觉领域中的自然图像字幕生成方法, 结合 RNN 和 CNN 的文本和图像分析, 相信很快会应用于医学图像分析领域.

上述挑战为医学图像分析领域的研究人员提供了巨大的机会, 我们相信, 通过深度学习算法的不断改进, 借助高性能并行计算技术的发展和日益改善的医学图像质量与不断增长的医学图像标记样本集, 基于深度学习的医学图像分析将大有所为.

References

- 1 Bibault J E, Giraud P, Burgun A. Big data and machine learning in radiation oncology: state of the art and future prospects. *Cancer Letters*, 2016, **382**(1): 110–117
- 2 Suzuki K, Zhou L P, Wang Q. Machine learning in medical imaging. *Pattern Recognition*, 2017, **63**: 465–467
- 3 Hubel D H, Wiesel T N. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 1959, **148**(3): 574–591
- 4 Fukushima K, Miyake S, Ito T. Neocognitron: a neural network model for a mechanism of visual pattern recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 1983, **SMC-13**(5): 826–834
- 5 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, **323**(6088): 533–536
- 6 LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W, Jackel L D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, **1**(4): 541–551
- 7 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 8 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, **313**(5786): 504–507
- 9 Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, **86**(11): 2278–2324
- 10 Graves A, Schmidhuber J. Offline handwriting recognition with multidimensional recurrent neural networks. In: *Proceedings of the 21st International Conference on Neural Information Processing Systems*. British Columbia, Canada: Curran Associates Inc., 2012. 545–552
- 11 Cireşan D C, Meier U, Gambardella L M, Schmidhuber J. Deep, Big, Simple neural nets for handwritten digit recognition. *Neural Computation*, 2010, **22**(12): 3207–3220
- 12 Graves A, Jaitly N. Towards end-to-end speech recognition with recurrent neural networks. In: *Proceedings of the 31st International Conference on Machine Learning*. Beijing, China: PMLR, 2014. 1764–1772
- 13 Hinton G, Deng L, Yu D, Dahl G E, Mohamed A R, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T N, Kingsbury B. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine*, 2012, **29**(6): 82–97
- 14 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, 2016. 770–778
- 15 Sun Y, Wang X G, Tang X O. Deep learning face representation from predicting 10,000 classes. In: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014. 1891–1898
- 16 Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J, Bellemare M G, Graves A, Riedmiller M, Fidjeland A K, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. *Nature*, 2015, **518**(7540): 529–533
- 17 Helping clinicians get patients from test to treatment, faster [Online], available: <http://www.deepmind.com/health>, March 14, 2017.
- 18 Zhang Lei, Zhang Yi. Big data analysis by infinite deep neural networks. *Journal of Computer Research and Development*, 2016, **53**(1): 68–79
(张蕾, 章毅. 大数据分析的无限深度神经网络方法. *计算机研究与发展*, 2016, **53**(1): 68–79)
- 19 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, **521**(7553): 436–444
- 20 Collobert R, Weston J. A unified architecture for natural language processing: deep neural networks with multitask learning. In: *Proceedings of the 25th International Conference on Machine Learning (ICML)*. Helsinki, Finland: ACM, 2008. 160–167
- 21 Taigman Y, Yang M, Ranzato M, Wolf L. Deepface: closing the gap to human-level performance in face verification. In: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014. 1701–1708
- 22 Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection. In: *Proceedings of the 2013 Neural Information Processing Systems (NIPS)*. Lake Tahoe Nevada, USA: NIPS, 2013. 2553–2561
- 23 Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z H, Karpathy A, Khosla A, Bernstein M, Berg A C, Fei-Fei L. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, **115**(3): 211–252
- 24 VOC 2012 [Online], available: <https://coursys.sfu.ca/2016sp-cmpt-733-g2/pages/voc2012>, March 14, 2017.
- 25 Mitos-atypi-14 [Online], available: <http://mitos-atypia-14.grand-challenge.org/>, March 14, 2017.
- 26 Schmidhuber J. Deep Learning in neural networks: an overview. *Neural Networks*, 2015, **61**: 85–117
- 27 Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Massachusetts: MIT Press, 2016.
- 28 Duan Yan-Jie, Lv Yi-Sheng, Zhang Jie, Zhao Xue-Liang, Wang Fei-Yue. Deep learning for control: the state of the art and prospects. *Acta Automatica Sinica*, 2016, **42**(5): 643–654
(段艳杰, 吕宜生, 张杰, 赵学亮, 王飞跃. 深度学习在控制领域的研究现状与展望. *自动化学报*, 2016, **42**(5): 643–654)

- 29 Guan Hao, Xue Xiang-Yang, An Zhi-Yong. Advances on application of deep learning for video object tracking. *Acta Automatica Sinica*, 2016, **42**(6): 834–847
(管皓, 薛向阳, 安志勇. 深度学习在视频目标跟踪中的应用进展与展望. *自动化学报*, 2016, **42**(6): 834–847)
- 30 Luo Jian-Hao, Wu Jian-Xin. A Survey on fine-grained image categorization using deep convolutional features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318
(罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. *自动化学报*, 2017, **43**(8): 1306–1318)
- 31 Wells III W M. Medical image analysis-past, present, and future. *Medical Image Analysis*, 2016, **33**: 4–6
- 32 Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Medical Image Analysis*, 2016, **33**: 170–175
- 33 Shen D G, Wu G R, Suk H I. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 2017, **19**: 221–248
- 34 Greenspan H, van Ginneken B, Summers R M. Deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1153–1159
- 35 Weese J, Lorenz C. Four challenges in medical image analysis from an industrial perspective. *Medical Image Analysis*, 2016, **33**: 44–49
- 36 Rueckert D, Glocker B, Kainz B. Learning clinically useful information from images: past, present and future. *Medical Image Analysis*, 2016, **33**: 13–18
- 37 Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciompi F, Ghafoorian M, van der Laak J A W M, van Ginneken B, Sánchez C. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, **42**(9): 60–88
- 38 Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images. *Medical Image Analysis*, 2016, **36**: 41–51.
- 39 Li Y, Liang W, Zhang Y L, An H B, Tan J D. Automatic lumbar vertebrae detection based on feature fusion deep learning for partial occluded C-arm X-ray images. In: Proceedings of the 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA: IEEE, 2016. 647–650.
- 40 Nasr-Esfahani E, Samavi S, Karimi N, Soroushmehr S M R, Ward K, Jafari M H, Felfeliyan B, Nallamothu B, Najarian K. Vessel extraction in X-ray angiograms using deep learning. In: Proceedings of the 38th International Conference of the Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA: IEEE, 2016. 643–646.
- 41 Dobbs H J, Parker R P. The respective roles of the simulator and computed tomography in radiotherapy planning: a review. *Clinical Radiology*, 1984, **35**(6): 433–439
- 42 Niessen W J. MR brain image analysis in dementia: from quantitative imaging biomarkers to ageing brain models and imaging genetics. *Medical Image Analysis*, 2016, **33**: 107–113
- 43 Powell C, Schmidt M, Borri M, Koh D M, Partridge M, Riddell A, Cookd G, Bhide S A, Nutting C M, Harrington K J, Newbold K L. Changes in functional imaging parameters following induction chemotherapy have important implications for individualised patient-based treatment regimens for advanced head and neck cancer. *Radiotherapy and Oncology*, 2013, **106**(1): 112–117
- 44 Bagci U, Udupa J K, Mendhiratta N, Foster B, Xu Z Y, Yao J H, Chen X J, Mollura D J. Joint segmentation of anatomical and functional images: applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. *Medical Image Analysis*, 2013, **17**(8): 929–945
- 45 Shi J, Zhou S C, Liu X, Zhang Q, Lu M H, Wang T F. Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset. *Neurocomputing*, 2016, **194**: 87–94
- 46 Carneiro G, Nascimento J C, Freitas A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Transactions on Image Processing*, 2012, **21**(3): 968–982
- 47 Gao X T, Lin S, Wong T Y. Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, 2015, **62**(11): 2693–2701
- 48 Zilly J, Buhmann J M, Mahapatra D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging and Graphics*, 2017, **55**: 28–41
- 49 Yu L Q, Chen H, Dou Q, Qin J, Heng P A. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 2017, **36**(4): 994–1004
- 50 Dhungel N, Carneiro G, Bradley A P. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Medical Image Analysis*, 2017, **37**: 114–128
- 51 Bourlard H, Kamp Y. Auto-association by multilayer perceptrons and singular value decomposition. *Biological Cybernetics*, 1988, **59**(4–5): 291–294
- 52 Shin H C, Orton M R, Collins D J, Doran S J, Leach M O. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1930–1943
- 53 Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P A. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 2010, **11**(12): 3371–3408
- 54 Smolensky P. *Information Processing in Dynamical Systems: Foundations of Harmony Theory*. Massachusetts: MIT Press, 1986. 194–281
- 55 Salakhutdinov R, Hinton G. Deep Boltzmann machines. *Journal of Machine Learning Research*, 2009, **5**(2): 1967–2006

- 56 Carreira-Perpinan M A, Hinton G E. On contrastive divergence learning. In: Proceedings of the 2005 Artificial Intelligence and Statistics. Bridgetown, Barbados: AISTATS, 2005.
- 57 Larochelle H, Bengio Y, Louradour J, Lamblin P. Exploring strategies for training deep neural networks. *The Journal of Machine Learning Research*, 2009, **10**: 1–40
- 58 Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 1994, **5**(2): 157–166
- 59 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, **9**(8): 1735–1780
- 60 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 2012 International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates Inc., 2012. 1097–1105
- 61 Farabet C, Couprie C, Najman L, LeCun Y. Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1915–1929
- 62 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.
- 63 He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1904–1916
- 64 Lin M, Chen Q, Yan S C. Network in network. arXiv: 1312.4400, 2014.
- 65 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015. 1–9
- 66 Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception architecture for computer vision. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. 2818–2826
- 67 Szegedy C, Ioffe S, Vanhoucke V, Alemi A A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. arXiv: 1602.07261, 2016.
- 68 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, Massachusetts, USA: IEEE, 2015. 3431–3440
- 69 Zagoruyko S, Komodakis N. Wide residual networks. In: Proceedings of the 2016 British Machine Vision Conference. York, UK: BMVC, 2016. 87.1–87.12
- 70 Huang G, Liu Z, van der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2017. 2261–2269
- 71 Chen Y P, Li J A, Xiao H X, Jin X J, Yan S C, Feng J S. Dual path network. arXiv: 1707.01629, 2017.
- 72 Iandola F N, Han S, Moskewicz M W, Ashraf K, Dally W J, Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv: 1602.07360, 2016.
- 73 Dai J F, Qi H Z, Xiong Y W, Li Y, Zhang G D, Hu H, Wei Y C. Deformable convolutional networks. arXiv: 1703.06211, 2017.
- 74 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. arXiv: 1709.01507, 2017.
- 75 ILSVRC2016 [Online], available: <http://image-net.org/challenges/LSVRC/2016/results>, October 30, 2017.
- 76 LSVRC2017 [Online], available: <http://image-net.org/challenges/LSVRC/2017/results>, October 30, 2017.
- 77 Dai J F, Li Y, He K M, Sun J. R-FCN: object detection via region-based fully convolutional networks. arXiv: 1605.06409. 2016.
- 78 Nexar challenge II [Online], available: <https://www.getnexar.com/challenge-2/>, October 30, 2017.
- 79 Cireşan D C, Giusti A, Gambardella L M, Schmidhuber J. Deep neural networks segment neuronal membranes in electron microscopy images. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada: Curran Associates Inc., 2012. 2843–2851
- 80 Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: integrated recognition, localization and detection using convolutional networks. arXiv: 1312.6229, 2014.
- 81 Cireşan D C, Giusti A, Gambardella L M, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Medical Image Computing and Computer-Assisted Intervention*. Nagoya, Japan: Springer, 2013. 411–418
- 82 Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P M, Larochelle H. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 2017, **35**: 18–31
- 83 Choi H, Jin K H. Fast and robust segmentation of the striatum using deep convolutional neural networks. *Journal of Neuroscience Methods*, 2016, **274**: 146–153
- 84 Song Y Y, Zhang L, Chen S P, Ni D, Lei B Y, Wang T F. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Transactions on Biomedical Engineering*, 2015, **62**(10): 2421–2433
- 85 Yang W, Chen Y Y, Liu Y B, Zhong L M, Qin G G, Lu Z T, Feng Q J, Chen W F. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Medical Image Analysis*, 2017, **35**: 421–433
- 86 Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer, 2015, **9351**: 234–241
- 87 Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016, **9901**: 424–432

- 88 Dou Q, Yu L Q, Chen H, Jin Y M, Yang X, Qin J, Heng P A. 3D deeply supervised network for automated segmentation of volumetric medical images. *Medical Image Analysis*, 2017, **41**: 40–45
- 89 Milletari F, Navab N, Ahmadi S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: Proceedings of the 4th International Conference on 3D Vision. Stanford, CA, USA: IEEE, 2016. 565–571
- 90 Suk H I, Lee S W, Shen D G. The Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*, 2014, **101**: 569–582
- 91 Suk H I, Lee S W, Shen D G, The Alzheimer's Disease Neuroimaging Initiative. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Structure and Function*, 2015, **220**(2): 841–859
- 92 Liu S Q, Liu S D, Cai W D, Che H Y, Pujol S, Kikinis R, Feng D G, Fulham M J. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Biomedical Engineering*, 2015, **62**(4): 1132–1140
- 93 ANDI. Sharing Alzheimer's research data with the world [Online], available: <http://adni.loni.usc.edu/>, March 14, 2017.
- 94 Al Rahhal M M, Bazi Y, AlHichri H, Alajlan N, Melgani F, Yager R R. Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*, 2016, **345**: 340–354
- 95 Abdel-Zaher A M, Eldeib A M. Breast cancer classification using deep belief networks. *Expert Systems with Applications*, 2016, **46**: 139–144
- 96 Arevalo J, González F A, Ramos-Pollán R, Oliveira J L, Lopez M A G. Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 2016, **127**: 248–257
- 97 Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez C I, Mann R, den Heeten A, Karssemeijer N. Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis*, 2017, **35**: 303–312
- 98 Xu Y, Mo T, Feng Q W, Zhong P L, Lai M D, Chang E I C. Deep learning of feature representation with multiple instance learning for medical image analysis. In: Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Florence, Italy: IEEE, 2014. 1626–1630
- 99 Gao X W, Hui R, Tian Z M. Classification of CT brain images based on deep learning networks. *Computer Methods and Programs in Biomedicine*, 2017, **138**: 49–56
- 100 Payan A, Montana G. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arXiv: 1502.02506, 2015.
- 101 Hosseini-Asl E, Gimel'farb G, El-Baz A. Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. arXiv: 1607.00556, 2016.
- 102 Abdi A H, Luong C, Tsang T, Allan G, Nouranian S, Jue J, Hawley D, Fleming S, Gin K, Swift J, Rohling R, Abolmaesumi P. Automatic quality assessment of echocardiograms using convolutional neural networks: feasibility on the apical four-chamber view. *IEEE Transactions on Medical Imaging*, 2017, **36**(6): 1221–1230
- 103 Gao X H, Li W, Loomes M, Wang L Y. A fused deep learning architecture for viewpoint classification of echocardiography. *Information Fusion*, 2017, **36**: 103–113
- 104 Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1207–1216
- 105 Kawahara J, Hamarneh G. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. *International Workshop on Machine Learning in Medical Imaging*. Athens, Greece: Springer, 2016. 164–171
- 106 Jiao Z C, Gao X B, Wang Y, Li J. A deep feature based framework for breast masses classification. *Neurocomputing*, 2016, **197**: 221–231
- 107 Tajbakhsh N, Suzuki K. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition*, 2016, **63**: 476–486
- 108 Kallenberg M, Petersen K, Nielsen M, Ng A Y, Diao P F, Igel C, Vachon C M, Holland K, Winkel R R, Karssemeijer N, Lillholm M. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1322–1331
- 109 van Tulder G, de Bruijne M. Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted Boltzmann machines. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1262–1272
- 110 Zhang Q, Xiao Y, Dai W, Suo J F, Wang C Z, Shi J, Zheng H R. Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics*, 2016, **72**: 150–157
- 111 Yang D, Zhang S T, Yan Z N, Tan C W, Li K, Metaxas D. Automated anatomical landmark detection on distal femur surface using convolutional neural network. In: Proceedings of the 12th International Symposium on Biomedical Imaging. New York, NY, USA: IEEE, 2015. 17–21
- 112 Chen H, Dou Q, Wang X, Qin J, Cheng J C Y, Heng P A. 3D fully convolutional networks for intervertebral disc localization and segmentation. *Medical Imaging and Augmented Reality*. Cham: Springer, 2016, **9805**: 375–382
- 113 de Vos B D, Wolterink J M, de Jong P A, Viergever M A, Išgum I. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: Proceedings of the 9784, Medical Imaging 2016: Image Processing. San Diego, California, US: SPIE, 2016, **9784**: Article No. 97841Y
- 114 Kong B, Zhan Y Q, Shin M, Denny T, Zhang S T. Recognizing end-diastole and end-systole frames via deep temporal regression network. In: Medical Image Computing and Computer-Assisted Intervention-MICCAI. Athens, Greece: Springer, 2016. 264–272

- 115 Cai Y L, Landis M, Laidley D T, Kornecki A, Lum A, Li S. Multi-modal vertebrae recognition using Transformed Deep Convolution Network. *Computerized Medical Imaging and Graphics*, 2016, **51**: 11–19
- 116 Lei J, Li G H, Tu D, Guo Q. Convolutional restricted Boltzmann machines learning for robust visual tracking. *Neural Computing and Applications*, 2014, **25**(6): 1383–1391
- 117 Lo S C B, Lou S L A, Lin J S, Freedman M T, Chien M V, Mun S K. Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Transactions on Medical Imaging*, 1995, **14**(4): 711–718
- 118 Sirinukunwattana K, Raza S E A, Tsang Y W, Snead D R J, Cree I A, Rajpoot N M. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1196–1206
- 119 Li A N, Cheng J, Wong D W K, Liu J. Integrating holistic and local deep features for glaucoma classification. In: Proceedings of the 38th Annual International Conference of the Engineering in Medicine and Biology Society. Orlando, FL, USA: IEEE, 2016. 1328–1331
- 120 Roth H R, Lu L, Seff A, Cherry K M, Hoffman J, Wang S J, Liu J M, Turkbey E, Summers R M. A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: Proceedings of the 17th International Conference International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham, Germany: Springer, 2013, **8673**: 520–527
- 121 Roth H R, Lu L, Liu J M, Yao J H, Seff A, Cherry K, Kim L, Summers R M. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1170–1181
- 122 Roth H R, Yao J H, Lu L, Stieger J, Burns J E, Summers R M. Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications. *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Cham: Springer, 2015. 3–12
- 123 Wang J, Ding H J, Bidgoli F A, Zhou B, Iribarren C, Molloy S, Baldi P. Detecting cardiovascular disease from mammograms with deep learning. *IEEE Transactions on Medical Imaging*, 2017, **36**(5): 1172–1181
- 124 Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. *Medical Image Analysis*, 2017, **39**: 178–193
- 125 Teramoto A, Fujita H, Yamamuro O, Tamaki T. Automated detection of pulmonary nodules in PET/CT images: ensemble false-positive reduction using a convolutional neural network technique. *Medical Physics*, 2016, **43**: 2821–2827
- 126 Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci, Navab N. AggNet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1313–1321
- 127 Chen H, Yu L Q, Dou Q, Shi L, Mok V C T, Heng P A. Automatic detection of cerebral microbleeds via deep learning based 3D feature representation. In: Proceedings of the 12th International Symposium on Biomedical Imaging (ISBI). New York, NY, USA: IEEE, 2015. 764–767
- 128 Dou Q, Chen H, Yu L Q, Zhao L, Qin J, Wang D F, Mok V C T, Shi L, Heng P A. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1182–1195
- 129 Dou Q, Chen H, Yu L Q, Qin J, Heng P A. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, 2017, **64**(7): 1558–1567
- 130 van Grinsven M J J P, van Ginneken B, Hoyng C B, Theelen T, Sanchez C I. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1273–1284
- 131 Xu J, Xiang L, Liu Q S, Gilmore H, Wu J Z, Tang J H, Madabhushi A. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging*, 2016, **35**(1): 119–130
- 132 Masood A, Al-Jumaily A, Anam K. Self-supervised learning model for skin cancer diagnosis. In: Proceedings of the 7th International IEEE/EMBS Conference on Neural Engineering. Montpellier, USA: IEEE, 2015. 1012–1015
- 133 Girshick R. Fast R-CNN. In: Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 1440–1448
- 134 Sarikaya D, Corso J J, Guru K A. Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. *IEEE Transactions on Medical Imaging*, 2017, **36**(7): 1542–1549
- 135 Twinanda A P, Shehata S, Mutter D, Marescaux J, Mathelin M D, Padoy N. EndoNet: a deep architecture for recognition tasks on laparoscopic videos. *IEEE Transactions on Medical Imaging*, 2016, **36**(1): 86–97
- 136 Chen H, Wu L Y, Dou Q, Qin J, Li S, Cheng J Z, et al. Ultrasound standard plane detection using a composite neural network Framework. *IEEE Transactions on Cybernetics*, 2017, **47**(6): 1576–1586
- 137 Kumar N, Verma R, Sharma S, Bhargava S, Vahadane A, Sethi A. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 2017, **36**(7): 1550–1560
- 138 Xu Y, Jia Z P, Ai Y Q, Zhang F, Lai M D, Chang E I C. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In: Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). South Brisbane, QLD, Australia: IEEE, 2015. 947–951
- 139 Qaiser T, Sirinukunwattana K, Nakane K, Tsang Y W, Epstein D, Rajpoot N. Persistent homology for fast tumor segmentation in whole slide histology images. *Procedia Computer Science*, 2016, **90**: 119–124

- 140 Song Y Y, Tan E L, Jiang X D, Cheng J Z, Ni D, Chen S P, Lei B Y, Wang T F. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Transactions on Medical Imaging*, 2017, **36**(1): 288–300
- 141 Fakhry A, Zeng T, Ji S W. Residual deconvolutional networks for brain electron microscopy image segmentation. *IEEE Transactions on Medical Imaging*, 2017, **36**(2): 447–456
- 142 Chen H, Qi X J, Yu L Q, Heng P A. DCAN: deep contour-aware networks for accurate gland segmentation. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. 2487–2496
- 143 Xu Y, Li Y, Wang Y P, Liu M Y, Fan Y B, Lai M D, Chang E I C. Gland instance segmentation using deep multichannel neural networks. *IEEE Transactions on Biomedical Engineering*, 2017, **64**(12): 2901–2912
- 144 Sirinukunwattana K, Pluim J P W, Chen H, Qi X J, Heng P A, Guo Y B, Wang L Y, Matuszewski B J, Bruni E, Sanchez U, Böhm A, Ronneberger O, Cheikh B B, Racoceanu D, Kainz P, Pfeiffer M, Urschler M, Snead D R J, Snead N M. Gland segmentation in colon histology images: the glas challenge contest. *Medical Image Analysis*, 2016, **35**: 489–502
- 145 Xie Y P, Zhang Z Z, Sapkota M, Yang L. Spatial clockwork recurrent neural network for muscle perimysium segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016, **9901**: 185–193
- 146 Xing F Y, Xie Y P, Yang L. An automatic learning-based framework for robust nucleus segmentation. *IEEE Transactions on Medical Imaging*, 2016, **35**(2): 550–566
- 147 Zhang W L, Li R J, Deng H T, Wang L, Lin W L, Ji S W, Shen D G. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage*, 2015, **108**: 214–224
- 148 Stollenga M F, Byeon W, Liwicki M, Schmidhuber J. Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. arXiv: 1506.07452, 2015.
- 149 Cui Z P, Yang J, Qiao Y. Brain MRI segmentation with patch-based CNN approach. In: Proceedings of the 35th Chinese Control Conference. Chengdu, China: IEEE, 2016. 7026–7031
- 150 Moeskops P, Viergever M A, Mendrik A M, de Vries L S, Benders M J N L, Išgum I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1252–1262
- 151 Dolz J, Betrouni N, Quidet M, Kharroubi D, Leroy H A, Reyns N, Massotier L, Vermandel M. Stacking denoising auto-encoders in a deep network to segment the brainstem on MRI in brain cancer patients: a clinical study. *Computerized Medical Imaging and Graphics*, 2016, **52**: 8–18
- 152 Mansoor A, Cerrolaza J J, Idrees R, Biggs E, Alsharid M A, Avery R A, Linguraru M G. Deep learning guided partitioned shape model for anterior visual pathway segmentation. *IEEE Transactions on Medical Imaging*, 2016, **35**(8): 1856–1865
- 153 Shakeri M, Tsogkas S, Ferrante E, Lippe S, Kadoury S, Paragios N, Kokkinos I. Sub-cortical brain structure segmentation using F-CNN's. In: Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI). Prague, Czech: IEEE, 2016. 269–272
- 154 Moeskops P, Wolterink J M, van der Velden B H M, Gilhuijs K G A, Leiner T, Viergever M A, Išgum I. Deep learning for multi-task medical image segmentation in multiple modalities. *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016, **9901**: 478–486
- 155 Wu A, Xu Z Y, Gao M C, Buty M, Mollura D J. Deep vessel tracking: a generalized probabilistic approach via deep learning. In: Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI). Prague, Czech: IEEE, 2016. 1363–1367
- 156 Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 2016, **35**(11): 2369–2380
- 157 Wang S L, Yin Y L, Cao G B, Wei B Z, Zheng Y J, Yang G P. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 2015, **149**: 708–717
- 158 Li Q L, Feng B W, Xie L P, Liang P, Zhang H S, Wang T F. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 2016, **35**(1): 109–118
- 159 Avendi M R, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 2016, **30**: 108–119
- 160 Ngo T A, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic Resonance. *Medical Image Analysis*, 2017, **35**: 159–171
- 161 Tan L K, Liew Y M, Lim E, McLaughlin R A. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. *Medical Image Analysis*, 2017, **39**: 78–86
- 162 Zhen X, Wang Z, Islam A, Bhaduri M, Chan I, Li S. Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation. *Medical Image Analysis*, 2015, **30**: 120–129
- 163 Roth H R, Farag A, Lu L, Turkbey E B, Ronald M. Deep convolutional networks for pancreas segmentation in CT imaging. In: Proceedings of 9413, Medical Imaging 2015: Image Processing. Orlando, Florida, USA: SPIE, 2015, **9413**: Article No.94131G
- 164 Roth H R, Lu L, Farag A, Shin H C, Liu J M, Turkbey E B, Summers R M. DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In: Proceedings of the 18th International Conference Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Munich, Germany: Springer, 2015. 556–564
- 165 Guo Y R, Gao Y Z, Shen D G. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Transactions on Medical Imaging*, 2016, **35**(4): 1077–1089

- 166 Korez R, Likar B, Pernuš F, Vrtovec T. Model-based segmentation of vertebral bodies from MR images with 3D CNNs. *Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece: Springer, 2016. 433–441
- 167 Poudel R P K, Lamata P, Montana G. Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. *Reconstruction, Segmentation, and Analysis of Medical Images*. Cham: Springer, 2016. 83–94
- 168 Kamnitsas K, Ledig C, Newcombe V F J, Simpson J P, Kane A D, Menon D K, Rueckert D, Glocker B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 2017, **36**: 61–78
- 169 Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille. Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv: 1412.7062, 2014.
- 170 Gao M, Xu Z Y, Lu L, Wu A, Nogues I, Summers R M, Mollura D J. Segmentation label propagation using deep convolutional neural networks and dense conditional random field. In: Proceedings of the 13th International Symposium on Biomedical Imaging. Prague, Czech Republic: IEEE, 2016. 1265–1268
- 171 Brosch T, Tang L Y W, Yoo Y, Li D K B, Traboulsee A, Tam R. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1229–1239
- 172 BRATS2015 [Online], available: <https://www.smir.ch/BRATS/Start2015>, March 14, 2017.
- 173 Wang C H, Yan X C, Smith M, Kochhar K, Rubin M, Warren S M, Wrobel J, Lee H. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Milan, Italy: IEEE, 2015. 2415–2418
- 174 Urban G, Bendszus M, Hamprecht F A, Kleesiek J. Multimodal brain tumor segmentation using deep convolutional neural networks. In: Proceedings of the 2014 MICCAI Workshop on Multimodal Brain Tumor Segmentation Challenge. Boston, Massachusetts, USA: Springer, 2014. 31–35
- 175 Zikic D, Ioannou Y, Brown M, Criminisi A. Segmentation of brain tumor tissues with convolutional neural networks. In: Proceedings of the 2014 MICCAI Workshop on Multimodal Brain Tumor Segmentation Challenge. Boston, Massachusetts, USA: Springer, 2014. 36–39
- 176 Pereira S, Pinto A, Alves V, Silva C A. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1240–1251
- 177 Carneiro G, Nascimento J, Bradley A P. Unregistered multiview mammogram analysis with pre-trained deep learning models. *Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer, 2015. 652–660
- 178 Bar Y, Diamant I, Wolf L, Greenspan H. Deep learning with non-medical training used for chest pathology identification. In: Proceedings of 9414, Medical Imaging 2015: Computer-Aided Diagnosis. Orlando, Florida, USA: SPIE, 2015, **9414**: Article No. 94140V
- 179 Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, Greenspan H. Chest pathology detection using deep learning with non-medical training. In: Proceedings of the 12th International Symposium on Biomedical Imaging. New York, NY, USA: IEEE, 2015. 294–297
- 180 Ginneken B V, Setio A A A, Jacobs C, Ciompi F. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In: Proceedings of the 12th International Symposium on Biomedical Imaging. New York, NY, USA: IEEE, 2015. 286–289
- 181 Shin H C, Roth H R, Gao M C, Lu L, Xu Z Y, Nogues I, Yao J H, Mollura D, Summers R M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1285–1298
- 182 Chatfield K, Simonyan K, Vedaldi A, Zisserman A. Return of the devil in the details: delving deep into convolutional nets. In: Proceedings of the 2014 British Machine Vision Conference. Nottingham, UK: University of Nottingham, 2014.
- 183 Ciompi F, de Hoop B, van Riel S J, Chung K, Scholten E T, Oudkerk M, de Jong P A, Prokop M, van Ginneken B. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Medical Image Analysis*, 2015, **26**(1): 195–202
- 184 Tajbakhsh N, Shin J Y, Gurudu S R, Hurst R T, Kendall C B, Gotway M B, Liang J M. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1299–1312
- 185 Aroyo L, Welty C. Truth is a lie: crowd truth and the seven myths of human annotation. *AI Magazine*, 2015, **36**(1): 15–24
- 186 Zhang S T, Metaxas D. Large-scale medical image analytics: recent methodologies, applications and future directions. *Medical Image Analysis*, 2016, **33**: 98–101
- 187 Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd International Conference on Machine Learning. Lille, France: JMLR, 2015, **37**: 448–456
- 188 Hinton G E, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov R R. Improving neural networks by preventing co-adaptation of feature detectors. *Computer Science*, arXiv:1207.0580, 2012.
- 189 Işın A, Direkçioğlu C, Şah M. Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 2016, **102**: 317–324
- 190 Setio A A A, Ciompi F, Litjens G, Gerke P, Jacobs C, Riel V, Wille M M W, Naqibullah M, Sánchez C I, van Ginneken B. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1160–1169
- 191 Nie D, Zhang H, Adeli E, Liu L Y, Shen D G. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. In: Proceedings of the 19th International Conference Medical Image Computing and Computer-Assisted Intervention. Athens, Greece: Springer, 2016, **9901**: 212–220

- 192 Diabetic retinopathy detection [Online], available: <https://www.kaggle.com/c/diabetic-retinopathy-detection>, July 2, 2017.
- 193 Second annual data science bowl [Online], available: <https://www.kaggle.com/c/second-annual-data-science-bowl>, July 2, 2017.
- 194 Tianchi AI competition [Online], available: https://tianchi.aliyun.com/mini/medical2017.htm_spm=5176.100065.3333.1.FCejdd, July 2, 2017.
- 195 Maier O, Menze B H, Gablentz J V D, Hani L, Heinrich M P, Liebrand M, Winzeck S, Basit A, Bentley P, Chen L, Christiaens D, Dutil F, Egger K, Feng C, Glocker B, G'otz M, Haeck T, Halme H L, Havaei M, Iftekharruddin K M, Jodoin P M, Kamnitsas K, Kellner E, Korvenoja A, Larochelle H, Ledig C, Lee J H, Maes F, Mahmood Q, Maier-Hein K H, McKinley R, Muschelli J, Pal C, Pei L, Rangarajan J R, Reza S M, Robben D, Rueckert D, Sallie E, Suetens P, Wang C W, Wilms M, Kirschke J S, Kr Amer U M, M ünte T F, Schramm P, Wiest R, Handels H, Reyes M. ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical Image Analysis*, 2017, **35**: 250–269



田娟秀 湖南大学电气与信息工程学院博士研究生. 主要研究方向为医学图像分析, 模式识别, 深度学习. E-mail: juanxiutian@126.com (**TIAN Juan-Xiu** Ph.D. candidate at the College of Electrical and Information Engineering, Hunan University. Her research interest covers medical image analysis, pattern recognition, and deep learning.)



刘国才 湖南大学电气与信息工程学院教授. 主要研究方向为医学图像分析, 模式识别与智能系统. 本文通信作者.

E-mail: lgc630819@hnu.edu.cn (**LIU Guo-Cai** Professor at the College of Electrical and Information Engineering, Hunan University. His research interest covers medical image analysis, pattern recognition, and intelligent system. Corresponding author of this paper.)

His research interest covers medical image analysis, pattern recognition, and intelligent system. Corresponding author of this paper.)



谷珊珊 北京解放军总医院放疗科技师. 主要研究方向为医学图像分析与肿瘤放射治疗.

E-mail: guss1990@163.com

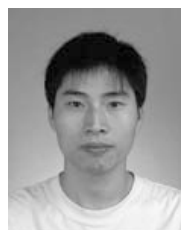
(**GU Shan-Shan** Physicist in the Department of Radiation Oncology, Chinese PLA General Hospital. Her research interest covers medical image analysis and tumor radiotherapy.)



鞠忠建 北京解放军总医院放疗科工程师. 主要研究方向为医学图像分析与肿瘤放射治疗.

E-mail: 15801234725@163.com

(**JU Zhong-Jian** Engineer in the Department of Radiation Oncology, Chinese PLA General Hospital. His research interest covers medical image analysis and tumor radiotherapy.)



刘劲光 湖南大学电气与信息工程学院博士研究生. 主要研究方向为医学图像分析, 放射治疗计划优化. E-mail: liujg201405@gmail.com

(**LIU Jin-Guang** Ph.D. candidate at the College of Electrical and Information Engineering, Hunan University. His research interest covers medical image analysis and optimization for radiotherapy plan.)



顾冬冬 湖南大学电气与信息工程学院博士研究生. 主要研究方向为医学图像分析, 模式识别. E-mail: gudongdongmia@163.com

(**GU Dong-Dong** Ph.D. candidate at the College of Electrical and Information Engineering, Hunan University. Her research interest covers medical image analysis and pattern recognition.)

Her research interest covers medical image analysis and pattern recognition.)