

深度学习技术在医学影像领域的应用

夏黎明¹, 沈 坚², 张荣国², 王少康², 陈 宽²

¹ 华中科技大学同济医学院附属同济医院放射科, 武汉 430030

² 北京推想科技有限公司, 北京 100025

通信作者: 沈 坚 电话: 010-85795622, E-mail: contact@infervision.com

【摘要】深度学习技术, 作为最近几年人工智能最热门的研究领域, 已成为全世界关注的焦点。深度学习在很多行业中展现出强大的应用能力, 在某些视听识别任务中的表现甚至超越了人类。在医学领域, 深度学习也逐渐成为研究者们分析大数据, 尤其是医学影像的首选方法。本文简要介绍深度学习的历史与概况, 结合国内外最新和最有影响力的研究成果, 阐述深度学习在医学影像领域的科学研究进展, 同时介绍深度学习在医学影像领域产品化应用及其未来的机遇与挑战。

【关键词】深度学习; 人工智能; 医学影像; 临床研究

【中图分类号】R445 **【文献标志码】**A **【文章编号】**1674-9081(2018)01-0010-05

DOI: 10.3969/j.issn.1674-9081.2018.01.003

Application of Deep Learning in Medical Imaging Research

XIA Li-ming¹, SHEN Jian², ZHANG Rong-guo², WANG Shao-kang², CHEN Kuan²

¹Department of Radiology, Tongji Hospital, Tongji Medical College of HUST, Wuhan 430040, China

²Infervision Inc., Beijing 100025, China

Corresponding author: SHEN Jian Tel: 010-85795622, E-mail: contact@infervision.com

【Abstract】Deep learning, as the most popular research field in artificial intelligence, has been developing rapidly in recent years and become the focus of global attention. Deep learning has demonstrated a powerful role in many application areas. In some visual and auditory recognition tasks, deep learning even shows better performance than human beings. In medical domain, deep learning has become the top choice for researchers to analyze big data, especially medical imaging. This review briefly introduces the history and development of deep learning, and elaborates on the progress of research on deep learning in medical imaging by reviewing the latest and most influential research results. In addition, this paper briefly discusses application of deep learning in medical imaging analysis, as well as the future prospect and challenges of deep learning.

【Key words】deep learning; artificial intelligence; medical imaging; clinical research

Med J PUMCH, 2018,9(1):10-14

深度学习技术, 作为人工智能的热门研究领域, 自2006年由多伦多大学教授Geoffrey Hinton^[1]提出后, 已在各行各业呈现爆炸式增长, 并涌现出像AlphaGo、自动驾驶汽车这样的重量级应用, 继而引发了全球范围内的广泛讨论及深度学习发展的潮流。

深度学习是一种对数据进行表征学习的方法, 其使用包含复杂结构或由多重非线性变换构成的多个处理层对数据进行高层抽象^[2]。深度学习网络可理解为传统神经网络的延伸, 是通过深层非线性网络结构无限逼近复杂函数, 以获取数据的原始特征, 进而取代传统的由

人工提取数据特征的方法, 实现让机器自主学习并提取数据特征。同时, 由于网络模型具有层次深、运算量大、表达能力强等特点, 深度学习非常适用于大数据, 尤其是图像、语音等直观意义不明显的的数据。

2016年是深度学习高速发展的一年, 这主要得益于大数据的支撑, 图形处理器强大的运算能力, 以及新型算法的不断更新。工业界巨头如国外的 Google、Facebook 以及国内的百度、阿里巴巴、腾讯等一系列创新型大公司纷纷对外公开表示人工智能尤其是深度学习将是其下一个战略重心。同时, 许多创业公司在各行各业纷纷成立, 均瞄准了深度学习的广阔前景。

1 应用于医学影像的深度学习方法及应用分类

自计算机技术得到广泛应用起, 医学影像研究者们就尝试建立影像自动分析系统。20世纪70年代至90年代, 研究者们通过对边、线、圆等特殊图形的数学建模进行医学影像分析; 90年代末开始, 以训练数据为基础的监督学习系统逐渐占据主流, 但其中的关键步骤, 提取影像特征仍由人工完成; 为了实现计算机的自动特征提取, 卷积神经网络 (convolutional neural network, CNN) 应运而生, 虽然最初 CNN 并不受重视, 但 Krizhevsky 等^[3]在2012年 ImageNet 大赛中应用深度学习网络 AlexNet 将准确率足足提高了一个档次, 使得深度学习神经网络成为计算机视觉以及医学影像分析领域的热门技术。自此, 深度学习在医学影像领域的研究逐年增加, 尤其在2015年之后更是呈井喷式增长^[4]。

1.1 深度学习方法

1.1.1 监督学习

监督学习需要通过有标注的数据进行训练得到模型参数, 即先输入计算机一些问题的正确答案, 然后在这些经验的基础上判断和分析新案例。在医学影像领域监督学习的应用较多, 其中常见模型包括 CNN, 以及近来颇受关注的的循环神经网络 (recurrent neural networks, RNN)。

1.1.2 无监督学习

无监督学习则是在未标记数据的基础上直接从数据中寻找规律, 生成模型参数, 典型案例是在一些无标签的肿瘤图片中分出其良恶性。

1.2 应用分类

1.2.1 分类问题

分类问题又可称为诊断问题, 是深度学习最先

涉及的问题, 即输入完整或局部的医学影像, 计算机输出单一的诊断结果, 如是否存在异常。CNN 是最常用的方法, 此外, 由于医学影像训练数据的量级 (成百上千) 往往远小于计算机视觉模型所需的量级 (数百万数千万), 迁移学习也很常见。迁移学习可利用由其他图片 (通常是自然图片) 训练生成的网络, 弥补医学影像训练数据量不足的问题。迁移学习在实际操作中有两种, 一是直接用网络进行医学影像特征提取; 二是用医学影像对网络进行进一步微调。这两种方法均很流行, 至于哪种方法更好, 学术界各执一词, 目前尚无定论。总体上, CNN 是目前影像分类问题最尖端的技术, 在某些疾病的分类判断上可达到人类专家级别的准确度。

1.2.2 检测问题

检测问题又可分为器官、区域和界标定位以及目标或病灶位置检测。其中目标或病灶位置检测是诊断过程中的关键步骤, 也是医生最费时费力的工作之一。检测问题相当于对每一个像素进行分类, CNN 依然是主要的解决方法。但与分类问题相比, 检测问题的一个显著且关键的区别在于目标或病灶区域相对于整幅图小很多, 所以对每个像素分类之后, 非目标类的数量远远超过目标类, 这就意味着训练数据的分类极度不平衡。对此, van Grinsven 等^[5]提出将分类错误的样本重新送回模型进行训练, 以增加 CNN 对这些区域的敏感性。Facebook 人工智能研究所的何凯明团队最近也提出将 focal loss 引入输入样本的权重来解决数据不平衡的问题^[6]。

1.2.3 分割问题

分割问题是医学领域的一个重要问题, 也是学术成果产出最多的研究课题。对器官或其他子结构的精确分割有利于影像上目标物体积、形状等临床指标的定量分析。应用深度学习对医学图像进行分割对将来利用计算机生成结构化报告也具有重要意义。在处理分割问题的过程中, 研究者们提出了很多以 CNN 为基础的新模型和新框架, 其中最具影响力的是 Ronneberger 等^[7-8]提出的 U-net, 其特点是拥有同样数目的上采样及下采样层数, 可对整个扫描的所有影像直接训练, 可更有效地利用全局信息。此外, RNN 在分割问题上的应用亦越来越多。

除了以上提到的3大类问题, 深度学习在医学影像的应用还包括图像配准、影像复原、影像生成与增强等, 其在医学影像领域的应用仍在不断渗透。

2 深度学习在医学影像领域应用案例

2.1 脑部磁共振图像用于儿童自闭症早期诊断

北卡罗来纳大学教堂山分校的精神病学专家 Hazlett 团队利用深度学习算法预测 2 岁前自闭症高危儿童（有家族史）是否会在 2 岁后被诊断为自闭症，其准确率高达 88%，远远超过准确率只有 50% 的传统行为问卷调查法，且可更早诊断^[9]。

Hazelett 团队发现患有自闭症儿童的大脑会在早期过度增长，往往比正常儿童大，但增大的时间点一直无法确定，且儿童自闭症在实际诊断中与许多其他因素有关，所以即使测得相关数据，医生也很难下定论。于是 Hazelett 团队使用深度学习算法，首先对 106 名有家族史的高风险儿童和 42 名低风险儿童在 6 个月和 12 个月的时候分别进行磁共振扫描，并输入每个儿童的性别、大脑表层面积、大脑颅内体积和皮质厚度等数据（左 39 区至右 39 区皮层半球，约 312 个测量数据），一共 315 个输入数据。整个算法结构采用三步带权重的深度降维网络，分别从输入 315 个数据降至 100 个，再降至 10 个，最后一层降至 2 个输出数据判断是否会患上自闭症。这一算法在 143 个测试数据中取得了 88% 的准确率，同时仅有 4 个假阳性。

其他的后续研究支持了这一结论，但该理论的鲁棒性依然有待验证，且距离临床应用仍具有一定距离。主要问题在于磁共振扫描的差异和提取数据方法的不同为检测脑结构的微小变化增加了困难，因此是否有其他更加稳定且精确的脑成像技术亦是问题。有学者对该研究是否适用于普通婴儿的预测尚存疑问。如果这套算法被证实可应用于临床，将意味着可提前判断儿童是否会患上自闭症，从而进行早期干预以降低幼儿患自闭症的风险。

2.2 视网膜眼底图像用于糖尿病性视网膜病变检测

糖尿病性视网膜病变是目前越来越被人们重视的致盲病因，全球约有 4.5 亿糖尿病患者存在该病风险，若能及早发现，可被治愈。

Google 公司研究者 Gulshan^[10] 团队在 2016 年采用深度学习网络^[11-12] 对 128 175 张已由 54 名美国专家标注过的视网膜眼底图像进行训练，然后在单独的测试数据集上进行了验证，准确率达到曲线下面积 91%，检测效能可与人类专家相当。

这项研究证明了糖尿病性视网膜病变人工智能自

动检测的应用前景，可大大提高效率和可复制性，扩大筛查范围，减小患者就诊难度，实现早诊早治。

2.3 乳腺钼靶影像用于乳腺癌筛查

乳腺癌是对女性生命威胁最大的疾病之一，目前拍摄乳腺钼靶图像是乳腺癌最主要的筛查手段。但由于乳腺肿块差异性较大，图像信噪比较低，易受影像医生诊断经验影响，乳腺癌筛查存在较高的误诊率。

来自荷兰的研究者 Kooi 等^[14] 应用深度学习网络检测乳腺肿块和钙化，对 45 000 张有标签的图像进行训练，采用和 OxfordNet^[13] 类似的网络结构，准确率高达 90%，超过传统的计算机辅助设计系统。此外，哈佛大学团队用课程学习的方法来训练多尺度卷积神经网络，并结合 ResNet^[15] 结构，在一个公开的乳腺钼靶影像数据集（Digital Database for Screening Mammography, DDSM）上训练并达到了 92% 的准确率^[16]。

然而，整体数据量级相对于深度学习所需要的数据量级来说仍太小，且现有模型所训练的数据样本不平衡问题突出，即正样本（有病灶的）数量远远超过了负样本（无病灶的）数量，导致预测结果偏向正样本，这亦是今后研究所需解决的问题。

2.4 皮肤表面图像用于皮肤癌检测

皮肤癌是人类最常见的恶性肿瘤之一，但因存在于皮肤表面，常易被忽视。皮肤癌的筛查主要通过视觉判断，一般首先进行临床筛查，之后进行皮肤镜检查、活检和皮肤组织病理学分析。

2017 年初，斯坦福大学工程学院和医学院合作团队在 *Nature* 上发布了在皮肤癌诊断领域的最新突破^[17]。其利用深度学习网络 Google Inception-v3 结合迁移学习技术，以 ImageNet 上基于 128 万张图 1000 种类数据集得到的预训练参数，训练了 129 450 张皮肤表面图片，其中包含 2032 种不同种类疾病。最后，此模型和 21 位皮肤科医生进行了两轮较量：第一轮为最常见的皮肤癌识别，第二轮为最致命的皮肤癌识别，人工智能系统在这两轮较量中均表现出与人类专家不相上下的水平，人工智能的准确率达到 69.4%，而人类专家的准确率在 66% 左右。除此之外，该团队还展望未来在手机智能移动端使用人工智能系统检测皮肤癌，以大幅度降低皮肤癌筛查费用。

3 深度学习产品化应用

基于深度学习的医学影像处理技术目前已被证明

在某些领域达到甚至超过了人类专家水平，这项技术也逐渐走向临床，开始辅助医生的日常工作。

2017年初，美国食品药品监督管理局（Food and Drug Administration, FDA）批准了全球首个深度学习影像临床应用平台 Arterys Cardio DL，意味着深度学习逐渐获得认可。该平台可用于诊断与检测多种心血管疾病，自动采集心室内外轮廓数据，具有耗时短、精度高的特点。其主打的产品 ViosWorks 革新了传统的心脏磁共振成像方法，将扫描时间从 1 h 缩短至 6~10 min，患者无需再屏住呼吸。此外，这款产品可做到从 7 个维度呈现心脏，包括 3 个空间维度、1 个时间维度和 3 个速度维度。

中国医学人工智能的发展丝毫不亚于发达国家，并且正以难以置信的速度增长。最重要的原因是中国医疗数据的量级远远超过其他国家，而数据规模对于深度学习算法起到了非常关键的作用。2015 年以来，中国陆续出现一批基于深度学习算法的医学影像人工智能公司。与此同时，许多顶级三甲医院积极配合并且推进深度学习在医学影像领域的临床化。目前问世的基于深度学习技术的产品主要包括 X 线辅助筛查系统和 CT 肺结节辅助筛查系统等。其中 X 线筛查产品可迅速发现胸部 X 线片中 20 多种异常影像学表现，帮助医生降低 X 线的漏诊率。CT 肺结节筛查产品不仅可快速识别不同大小的实性、亚实性、磨玻璃结节，还可提供前后图像对比功能，极大减轻了影像医生的工作负担。当然，基于深度学习的医疗影像产品还不够成熟，主要体现在诊断准确率不足、假阳性率高和鲁棒性较差 3 个方面。

4 趋势与展望

4.1 优质的海量数据

深度学习需要量大且质优的数据，尤其是监督学习需要人类专家标记数据并教给机器正确答案，才能产生更加准确的结果，所以更多的数据库共享、医学影像存档与通信系统的普及、以及 2D 到 3D 图像甚至是更高维数据的同步训练有利于影像数据的积累与利用，也是未来发展的趋势。

4.2 模型的鲁棒性

目前已经出现大量优质的算法用以实现各种需求，以后将会涌现出许多功能更强大、效率更高、稳定性更好的算法与框架。算法的优化是一方面，数据的人口代表性是另一方面。要保证模型的鲁棒性还有

赖于收集更全面、更具有统计学意义的训练数据集。

4.3 非监督学习比重提高

目前监督学习占主导地位，用神经网络的方法来训练医学影像的分类、检测、分割任务极度依赖医生的准确标记，耗时耗力。但从人类和动物拥有能够在自然环境中学习的本能来看，非监督学习比重提高必然是未来兴起的方向^[2]，可显著降低影像标记工作的时间和人力成本。

4.4 人工智能地位进一步提高

目前人工智能的发展仍然处于快速上升阶段，可能面临一些来自政策和伦理等方面的挑战。“人工智能是否会取代影像科医生”的话题在多个学术会议上引发激烈讨论，人机协作的具体落实还有待临床流程的设计与优化。在 2017 年 3 月，人工智能首次被列入政府工作报告，显示了国家对这项高科技的重视，其未来发展不可估量。

参 考 文 献

- [1] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks [J]. *Science*, 2016, 313: 504-507.
- [2] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521: 436-444.
- [3] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks [C]. *International Conference on Neural Information Processing Systems*, 2012: 1097-1105.
- [4] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis [J]. *Med Image Anal*, 2017, 42: 60-88.
- [5] van Grinsven MJ, van Ginneken B, Hoyng CB, et al. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images [J]. *IEEE Trans Med Imaging*, 2016, 35: 1273-1284.
- [6] Lin Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [C]. In *ICCV*, 2017.
- [7] Ronneberger O, Philipp F, Thomas B. U-net: Convolutional networks for biomedical image segmentation [C]. *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2015.
- [8] Özgün C, Abdulkadir A, Lienkamp SS, et al. 3D U-net: learning dense volumetric segmentation from sparse annotation [C]. *International Conference on Medical Image Computing and Computer Assisted Intervention*, 2016.

- [9] Hazlett HC, Gu H, Munsell BC, et al. Early brain development in infants at high risk for autism spectrum disorder [J]. *Nature*, 2017, 542: 348-351.
- [10] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs [J]. *JAMA*, 2016 316: 2402-2410.
- [11] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [J]. *Computer Science*, 2015: 2818-2826.
- [12] Christian S, Sergey I, Vincent V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning [C]. *AAAI*, 2017.
- [13] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. *Computer Science*, 2014, 1556: 1409.
- [14] Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions [J]. *Med Image Anal*, 2017, 35: 303-312.
- [15] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [16] Lotter, William, Greg S, et al. A multi-scale CNN and curriculum learning strategy for mammogram classification [C]. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2017: 169-177.
- [17] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. *Nature*, 2017, 542: 115.

(收稿日期: 2017-10-11)

· 医学新闻 ·

北京协和医院研究首次揭示 V δ 2 T 细胞参与类风湿关节炎的发病机制

北京协和医院风湿免疫科张焯教授团队最新研究成果表明,天然免疫细胞 V δ 2 T 细胞在类风湿关节炎 (rheumatoid arthritis, RA) 发病中具有重要作用,为监测 RA 病情活动提供了新型生物标志物,并为 RA 治疗提供了新靶点。该研究成果近期发表于国际风湿病学顶级期刊《风湿病学年鉴》(*Ann Rheum Dis*, 影响因子 12.811),是本年度国内风湿病首个发表于该杂志的研究成果,北京协和医院为第一作

者单位。

RA 是最常见的慢性侵蚀性关节炎,以多发关节侵蚀破坏为特征,病程迁延反复,给患者、家庭和社会造成了严重的经济负担。RA 发病机制尚未完全明确,因此寻找发病过程中的关键免疫细胞,阻断其作用机制,具有潜在的临床意义。

(北京协和医院宣传处 陈华 陈明雁)