

人工智能技术在皮肤病辅助诊断的应用研究

傅裕^① 鲍迎秋^① 谢沂伯* 张晶^② 郝伟^② 李伟平^②

摘 要 目的: 皮肤病种类繁多, 采用机器学习、计算机视觉等人工智能技术可实现自动、快速分类及特征识别, 分析并研究国内外在皮肤病辅助诊断方面的人工智能技术应用。方法: 临床诊断使用的皮肤病图像主要包括皮肤镜图像以及皮肤组织病理图像等, 所使用的方法与样本数量及质量直接相关, 针对海量皮肤镜图像, 采用卷积神经网络迁移学习可实现疾病类型的分类; 针对小样本量皮肤镜图像, 结合皮肤学诊断原理和机器学习建模实现特定疾病类型诊断; 针对有限样本且复杂的皮肤组织病理图像, 运用图像分割和机器学习建模实现病理图像自动区域标注。结果: 在皮肤病辅助诊断方面, 结合人工智能技术, 国内外研究机构已经实现肤质特征识别、自动区域标注、疾病分类等应用。结论: 随着用于皮肤病诊断的人工智能技术发展, 有望大幅提高接诊效率, 为医生提供客观的辅助支撑, 减轻负担。

关键词 人工智能 皮肤病辅助诊断 机器学习 计算机视觉

Doi:10.3969/j.issn.1673-7571.2018.10.010

[中图分类号] R319

[文献标识码] A

Application Research of Artificial Intelligence in the Skin Diseases Assisted Diagnosis / FU Yu, BAO Ying-qiu, YIE Yi-bo, et al// China Digital Medicine. -2018 13(10): 29 to 31

Abstract Objective: There are many kinds of skin diseases. Artificial intelligence technology such as machine learning and computer vision can be used to realize automatic and rapid classification and feature recognition. This paper analyzes and studies the application of artificial intelligence technology in the skin diseases assisted diagnosis at home and abroad. Methods: The dermatological images used in clinical diagnosis mainly include dermoscopic images and skin tissue pathological images. The methods used are directly related to the number and quality of samples. For the images of massive dermoscopy, convolutional neural network transfer learning can be used to achieve disease types classification; for small sample dermoscopy images, combined with dermatological diagnostic principles and machine learning modeling to achieve disease type diagnosis; for limited sample and complex skin tissue pathology images, automatic segmentation of pathological images using image segmentation and machine learning modeling. Results: In terms of the skin diseases assisted diagnosis, combined with artificial intelligence technology, domestic and foreign research institutes have implemented applications such as skin feature recognition, automatic region labeling, and disease classification. Conclusion: With the development of artificial intelligence technology for the skin diseases assisted diagnosis, it is expected that the efficiency of admissions will be greatly increased, providing objective support for doctors and reducing the burden.

Keywords artificial intelligence, skin diseases assisted diagnosis, machine learning, computer vision

Corresponding author National Institute on Aging, Beijing Hospital, Beijing 100730, P.R.C.

1 引言

皮肤病是发生在皮肤和皮肤附属器官疾病的总称, 在人群中发生率非常高。皮肤是身体的第一道防线, 也是人体最大的器官, 人体机能的病症往往会在皮肤组织直接表现出来。因此, 皮肤病诊断具有至关重要的作用。当前我国皮肤病诊断面临着两大挑战, 一是皮肤病种类繁多, 外部环境、食物、微生物、遗传等都是引起皮肤病的致病因素, 疾病类型包括变态反应性、自身免疫性、感染性、遗传性、肿瘤性等呈多样化; 第二个挑战是皮肤科专家医生的短缺, 尤其是在基层医院, 缺少有经验的皮肤科临床医生。面对病人数量的日益增加和对疾病准确诊断的更高要求, 亟需利用人工智能技术^[1-2], 提高接

*通讯作者: 北京医院, 国家老年医学中心, 100730, 北京市东城区东单大华路1号

①北京医院, 国家老年医学中心, 100730, 北京市东城区东单大华路1号

②北京贝叶科技有限公司, 100083, 北京市海淀区成府路45号智道大街D座

诊效率，并提供客观准确的诊断结果^[3]。

2 用于皮肤病辅助诊断的人工智能技术

在医疗图像领域，将人工智能技术用于皮肤病辅助诊断，涉及计算机视觉与机器学习等技术，同时必须重点解决皮肤病类型多样化以及特征复杂难以提取等关键问题^[4]。

针对皮肤病类型多样化以及复杂特征难以提取等关键问题，需要根据不同规模、不同类型的皮肤病图像样本情况分析后进行建模。常用建模方法包括卷积神经网络（CNN）^[5]、支持向量机（SVM）^[6]等。CNN是一种常见的深度学习架构，受动物视觉皮层细胞检测光学信号的神经传导网络结构启发，具有多层神经网络。该网络在模式分类方面实现了端到端的图像目标识别，避免了复杂的图像前期预处理，可直接输入图像训练，获得了广泛应用。常用的CNN框架包括LeNet、AlexNet、ZF、GoogLeNet、VGG、ResNet（见图1）。

SVM具有完整的数学理论基础，既可以解决线性问题也可以解决非线性问题，适用于小样本、非线性及高维模式识别问题，基本思想是训练一个分类面，使特征空间中的点（即样本）到分类面的距离最大，SVM优化问题的数学模型为：

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + c \sum_{i=1}^N \xi_i \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) \geq 1 - \xi_i, i=1,2,\dots,N \\ & \xi_i \geq 0, i=1,2,\dots,N \end{aligned}$$

设该问题的解为 w^* 和 b^* ，则分类超平面为 $w^* \cdot x + b^* = 0$ ，分类决策函数为 $f(x) = \text{sign}(w^* \cdot x + b^*)$ 。这个模型用于训练线性分类模型，非线性模型的数学模型则可以表示为：

$$\min_{w,b,\xi} \quad \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i$$

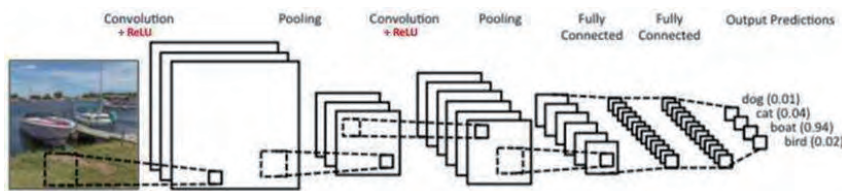


图1 卷积神经网络框架

$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i=1,2,\dots,N \end{aligned}$$

其中 $K(x_i, x_j)$ 是 x_i 和 x_j 的映射的内积， $K(x_i, x_j)$ 的定义如下：

$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ 表示由输入空间到新的特征空间的映射， $K(x_i, x_j)$ 是正定核函数，当映射函数是非线性函数时，在原特征空间中看来，训练得到的模型是一个非线性分类器。

CNN在样本量充足情况下可以自动提取特征并分类，SVM在计算机视觉算法提取图像特征后，再结合小样本进行准确分类，这两种方法在皮肤病诊断中均有广泛应用。

3 皮肤病辅助诊断的典型应用案例

临床诊断皮肤疾病类型的辅助手段包括皮肤镜、皮肤组织病理图像等，针对不同的样本规模和图像类型，采用不同算法进行处理。当皮肤镜图像样本量充足，并具有准确标注信息，适合采用深度学习方法，利用CNN进行分类，典型案例为斯坦福大学团队开展的皮肤癌检测，相关成果^[7]刊发在2017年1月底《自然》杂志；同时，当皮肤镜样本量有限，需要引入医生诊断原则，构建计算机视觉关键特征提取方法，使用SVM进行准确快速分类，典型案例为华中科技大学^[8]黑色素瘤识别；皮肤组织病理切片用于病症确诊，图像样本量有限，具有丰富图形学特征，适合采用边缘提取等计算机视觉方法，再结合SVM等机器学习

方法，进行皮肤组织细胞区域标注，典型案例为广东工业大学团队开展的病理图像自动标注，相关成果^[9]发表在2015年《计算机研究与发展》杂志。

斯坦福团队采用深度学习方法对皮肤镜和临床皮损图像进行自动分类，并取得了较好的结果。首先该团队构建了一个拥有129 450幅皮肤影像，包括2 032种不同皮肤病的数据库，为采用深度学习算法提供了充分的条件。该团队采用GoogLeNet Inception-V3^[10]分类框架，并利用迁移学习技术将模型进行微调得到了最终的CNN模型，算法的性能通过敏感性、特异性曲线进行衡量，分别代表正确识别恶性病变的能力和正确识别良性病变的能力。在区分细胞癌与脂溢性角化病、黑色素瘤与良性色素痣任务中，算法的表现与皮肤科医生鉴定能力相当，灵敏度曲线之下的区域达到整个曲线图区域的91%。同时，该算法的敏感性性能进行调整，研究者可以根据想要评估的东西从而调整其响应度。

华中科技大学团队结合皮肤病黑色素瘤症状识别ABCD-rule^[11]与SVM方法，提出一种基于模糊理论的SVM多分类算法，针对传统SVM“一对多”或“一对一”策略中出现的分类盲区问题，引入模糊隶属度函数，从而使该算法的分类能力更稳定。算法进行了图像数据库的预处理，包括图像采集、图像预处理（图像格式的转换、尺寸的统一、图像的增强、噪声的消除）、图像分割（在彩色空间转

换的基础上进行症状图像的分割)、特征计算(计算特征向量的17个指标参数)、数据存储。根据皮肤学诊断原理ABCD-rule,提出了基于颜色和边缘的特征提取算法,获得了症状的17个特征参数。同时,采用Canny^[12]算子在区域内进行边缘检测。在SVM分类系统的结构设计时,为了减少结构的复杂度,将高维特征空间压缩到低维进行特征选择,采用了PSL特征选择^[13]模块、SVM分类模块,最后用活检皮肤镜样本验证了所提方法的有效性。

广东工业大学团队识别的对象是病理图像,与皮肤镜图像不同,该团队面临两个挑战,一是医生是根据全局图像来判断,但是在文本诊断记录仅描述病理的特性,并未体现局部区域的关系;二是同一个标注词所对应的图像有很多种,存在很大差异。因此,该团队提出基于贝叶斯学习^[14]的多示例多标签稀疏集成算法,构建活检病理图像自动标注框架,利用带约束条件的分割算法把病理图像划分为若干不连续的区域,转化为多示例样本。在皮肤科活检样本数据集测试,达到了医学可接受的标注准确率,具有很高的实用价值。

4 我国皮肤病人工智能辅助诊断现状

我国在皮肤病人工智能辅助诊断方面,呈现了大型医院、研究所、科技公司通力合作的局面,并且已经取得一定进展。北京协和医院与北京航空航天大学使用皮肤镜图片自动识别疾病类型^[15],湘雅大学第二医院与丁香园、大拿科技实现了首个皮肤病人工智能诊断辅助系统^[16],中日友好医院发起了中国人群皮肤影像资料库项目^[17],建立可用于辅助诊断的中国

人群皮肤影像学习资源。人工智能技术在皮肤病诊断方面前景广阔,但目前依旧存在技术落地和推广的瓶颈。我国皮肤病特点是患病群体大,在临床诊断中炎症性皮肤病发病率高,多样性尤为明显,鉴别和诊断标准不统一,对于人工智能技术积累样本和分类准则制定都具有极大挑战;同时,可进行皮肤镜、皮肤组织病理切片诊断的医院主要为三甲医院,尤其在病理识别方面,高度依赖有经验的临床医生,大多数基层医院皮肤病医疗设备和医生储备方面均有欠缺,可用于训练的高质量标注图像样本量不足。因此,实现我国皮肤病的高效自动识别需要医院、研究院所和科技公司继续发挥医疗、技术优势,逐步提高人工智能技术在皮肤病医疗诊断方面的标准化,扩充样本规模,在精准识别方面取得突破。

5 应用前景与展望

随着我国经济建设的快速发展,进入人工智能时代,技术的发展趋势是不断提高效率,辅助医生临床诊断及治疗^[18]。将人工智能技术与皮肤病诊断深度融合,可以实现普及性强、稳定度高、不依赖人工经验的疾病类型识别方法,能够大大缓解患者与优质医疗资源紧张的矛盾,减轻医生负担,提高诊断效率和准确性。对于当前我国皮肤病人工智能辅助诊断应用,需要进一步聚焦关键问题,提高病种识别种类、模型精度等,增强人工智能技术对医院和医生的辅助作用。

参考文献

[1] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. The

bulletin of mathematical biophysics, 1943, 5(4): 115-133.

[2] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527-1554.

[3] 赵一鸣, 左秀然. PACS与人工智能辅助诊断的集成应用[J]. 中国数字医学, 2018, 13(4): 20-22.

[4] 张宇, 李姣. 电子健康档案数据挖掘与整合技术新进展[J]. 中国数字医学, 2017, 12(9): 41-44.

[5] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[6] Chen PH, Lin CJ, Scholkopf B. A tutorial on v-support vector machines[J]. Applied Stochastic Models in Business & Industry, 2005, 21(2): 111-136.

[7] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118.

[8] 艾武. 基于支持向量机的色素斑痣类皮肤病症状识别研究[D]. 武汉: 华中科技大学, 2007.

[9] 张钢, 钟灵, 黄永慧. 一种病理图像自动标注的机器学习方法[J]. 计算机研究与发展, 2015, 52(9): 2135-2144.

[10] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]// Computer Vision and Pattern Recognition. IEEE, 2016: 2818-2826.

[11] Nachbar F, Stolz W, Merkle T, et al. The ABCD rule of dermatoscopy[J]. Journal of the American Academy of Dermatology, 2014, 30(4): 551-559.

[12] Canny J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 679-698.

[13] 王凌. 智能优化算法(下转第38页)

年,这一时期中国的医疗人工智能研究已经达到美国第二阶段水平,研究热点主要是神经网络、分类识别、算法等人工智能技术以及疾病诊断等应用。第三阶段是2011—2014年,美国在这一时期因为技术瓶颈的限制而研究热度下降,中国虽然在这阶段同样没有进行人工智能技术的研究,但是却将力量集中研究医疗方面的一些应用以及管理方面,如智能卡、远程医疗信息系统、匿名、安全隐私等。第四阶段是2015—2018年,这一时期中国的研究热点集中在深度学习、特征提取、图像分类、支持向量机、大数据等人工智能技术以及电子皮肤、治疗等应用。

3 研究结论与建议

主要从文献计量学的视角,对1997—2018年间中美医疗人工智能期刊文献进行了比较,主要得到了以下结论。

第一,从研究年代分布来看,中美两国都是从2014年开始发表的文献数迅速上升,其中的原因:一是在2014年人工智能技术的新一轮爆发,二是两国都发布了政策文件促进医疗人工智能的应用。中国在2011年之前的文献远远少于美国,但在2014年的爆发后,到2018年已经赶上美国了。

第二,从研究力量分布来看,中美两国发文的核心期刊在载文量影响因子以及JCR分区上大致相同,说明中国的医疗人工智能研究水平当下与美国已经没有差距了,中国的医疗人工智能技术储备已经赶超美国,完全可以支持医疗人工智能的应用发展。两国的高产研究机构对比,美国的高产机构主要有高校研究机构和公共卫生机构,而中国更多集中于高校研究机构,建议中国的医疗卫生机构利用自

身的优势更多参与到研究。

第三,从研究的高频关键词看,中国比美国的起步较晚,两国的各个时期的研究热点都集中在当前最前沿的人工智能技术,以及相应的医疗应用,到2014—2018年这个阶段时,中美两国的人工智能技术已经处于同一水平,中国已经具备将人工智能的医疗应用的技术基础,但是实际应用进展和美国还有一定差距。

当前中国的人工智能技术发展已经达到和美国的同一水平,但我国目前医疗卫生机构和人工智能企业的合作比起美国远远不够,医疗人工智能在医疗领域的应用还不成熟^[11]。我国发展智能医疗行业具有良好的技术基础,在智能医疗的应用方面有广阔的发展前景,而且我国政府也很重视医疗人工智能行业的发展,出台一系列政策指导促进发展^[12]。建议高校研究机构、医疗卫生机构和企业进行三方合作,高校研究机构提供技术支持,医疗卫生机构拥有数据和应用场景,企业进行资本的提供以及医疗人工智能市场开发。

参考文献

- [1] 钟文艳.美国智能医疗产业发展现状分析[J].全球科技经济瞭望,2017,32(6):38-44.
- [2] 盖双双,刘雪立.我国图书馆学情报学研究的国际化现状——基于SSCI高影响力图书馆学情报学期刊引证分析[J].图书情报导刊,2015,25(11):156-160.
- [3] 马永双,蔡敏.中美STEM教育研究的文献计量学分析[J].比较教育研究,2018(2):104-112.
- [4] 吴向文,王志军.2001—2015年境内外教师教育研究文献计量分析及其启示[J].教师教育研究,2016,28(6):105-114.
- [5] 李俊平.人工智能技术的伦理问题及其对策研究[D].武汉:武汉理工大学,2013.

- [6] 林菲,杨舰.中国环境科学研究热点及其演化——基于文献计量学方法的量化分析[J].科学学研究,2016,34(9):1294-1300.
- [7] 朱彦君,程结晶.《中国图书馆学报》科学知识图谱分析——基于2005—2015年的文献数据[J].大学图书馆学报,2017,35(4):110-116.
- [8] 王素,邹俊伟.1970~2010年国际比较教育研究之演进——基于科学知识图谱方法的实证分析[J].外国教育研究,2011(9):37-44.
- [9] 王嘉雯.大数据背景下的物联网智能家居研究[J].信息与电脑(理论版),2017(24):144-145.
- [10] 王锡山.未来医学时代——人工智能诊疗[J].中华结直肠疾病电子杂志,2017,6(4):349-352.
- [11] 吴佳男.智能医疗:或成医疗机构新突破口[J].中国医院院长,2017(z1):89.
- [12] 周宣汝,赵丽亚,赵地,等.人工智能对科研信息化的推动作用[J].科研信息化技术与应用,2016,7(6).

【收稿日期:2018-08-27】

(责任编辑:肖婧婧)

(上接第31页)及其应用[M].施普林格出版社,2001.

- [14] Zhu J,Chen J,Hu W,et al.Big Learning with Bayesian methods[J].国家科学评论(英文版),2017.
- [15] 周航宁,谢凤英,姜志国,等.基于深度学习的皮肤影像分类[J].协和医学杂志,2018(1):15-18.
- [16] 丁香园.放大招!中国首个皮肤病人工智能辅助诊断系统发布[EB/OL].http://yyh.dxy.cn/article/519787,2017-05-19.
- [17] 黄硕,党超峰.人工智能助力诊断皮肤肿瘤[N/OL].http://cn.chinadaily.com.cn/2018-03/29/content_35942148.htm,2018-03-29.
- [18] 张静,常燕群.人工智能和虚拟现实技术在孤独症患者康复训练中的应用[J].中国数字医学,2013,8(7):83-86.

【收稿日期:2018-06-27】

【修回日期:2018-10-09】

(责任编辑:肖婧婧)