

人工智能助力检验医学发展

周 泉¹, 齐素文², 肖 斌¹, 李乔亮², 孙朝晖¹, 李林海¹

¹中国人民解放军南部战区总医院检验科, 广东 广州 510010; ²深圳大学生物医学工程学院体外诊断系, 广东 深圳 518037

摘要:2017年以来,中国、美国和欧盟先后发布了国家级人工智能(AI)战略发展规划,人类历史即将迎来以“智能化”为主旋律的第四次工业革命。在医疗检验领域,人工智能的理论和技术的爆发也将为医学检验理论、方法与应用的发展提供新方向,本文首先从人工智能的理论与实践基础出发,介绍人工智能的发展历程以及算法、数据和算力三大要素的前沿进展,进而结合医学检验的关键应用维度,阐述“AI+检验”的结合创新,包括涉及检验标本处理的标本采集机器人、样本稀释机器人和样本传送机器人;检验项目挖掘如肿瘤标志物和药物基因组学;检验细胞形态学、检验数据处理、辅助诊断模型和互联网医学检验等。随着工业4.0时代的到来,人工智能技术必将推动医学检验从自动化发展到更高阶的智能化阶段。

关键词:人工智能;检验医学;智能化;大数据

Artificial intelligence empowers laboratory medicine in Industry 4.0

ZHOU Quan¹, QI Suwen², XIAO Bin¹, LI Qiaoliang², SUN Zhaohui¹, LI Linhai¹

¹Department of Medical Laboratory, General Hospital of Southern Theater of PLA, Guangzhou 51010, China; ²Department of In vitro Diagnostics, School of Biomedical Engineering, Shenzhen University, Shenzhen 518037, China

Abstract: Since 2017, China, the United States, and the European Union have successively issued national-level artificial intelligence (AI) strategic development plans, and the human history is about to witness the 4th industrial revolution with the theme of "intelligence". In the field of medical testing, the explosive growth of AI theories and technologies also provide a new direction for the development of medical testing theory, methods and applications. We review the evolution of AI and the recent progress in three major elements of AI, namely algorithms, data and computing power, and elaborate on the combined innovation of "AI+testing" in light of the key application dimensions of medical testing. The major applications include specimen collection robots, sample dilution robots and sample transfer robots involved in the processing of test specimens; test item mining such as tumor markers and pharmacogenomics; cytomorphology, laboratory medicine data processing, auxiliary diagnostic models, and internet-based medical tests. With the advent of the era of Industry 4.0, AI technology will promote the development of medical testing from automation to a highly intelligent stage.

Keywords: artificial intelligence; laboratory medicine; intellectualization; big data

《新一代人工智能发展规划》中提出,要发展智能医疗,我们的首要任务就是推广应用人工智能治疗的新模式新手段,建立快速精准的智能医疗体系^[1-2]。临床检验主要是运用物理学、化学和生物学等实验方法对各种标本进行定性或定量分析,为医生临床诊断提供可靠依据^[3]。近年来,人工智能(AI)在医疗检验领域大放异彩,人工智能在样本处理,形态学检验,检验结果审查等过程中发挥着重要作用,有效减少了检验工作人员因主观性导致的误差,提高了检验人员的工作效率,人工智能与检验的结合满足了临床对检验专业日益增高的质量要求^[4]。人工智能大潮汹涌而来,我们作为工作在临床一线的检

验人更应该顺应潮流,担当使命,积极接受新事物并将其转化为更先进的检验手段,更好地服务于医疗事业。

1 人工智能技术

1.1 人工智能的概念

人工智能是研究、开发用于模拟、延伸和扩展人的理论、方法、技术及应用系统的一门新的技术科学^[5]。人工智能由计算机科学衍生而来,尝试利用计算机模拟人脑做出智能反应,目前人工智能的研究涵盖机器人、语言识别、图像识别、自然语言处理和专家系统等多个领域。

1.2 人工智能发展历程

1945年世界上第1台通用电子计算机ENIAC出现,尽管该计算机计算速度有限,操作过程繁琐,但依然为人工智能的出现奠定了条件基础。1950年,图灵发出疑问“机器会思考吗?”,并且提出著名的图灵测试,该测试至今仍被视为人工智能水平测试标准之一,而图灵也被誉为人工智能之父。1956年,科学家们在达特茅

收稿日期:2020-02-17

基金项目:国家自然科学基金(31901035);广州市科技计划(201903010060);军队后勤科研项目(CWH17C017);军队保健专项(18BJZ14)

Supported by National Natural Science Foundation of China (31901035).

作者简介:周 泉,博士,主管技师,E-mail: springzhou@yeah.net

通信作者:李林海,主任技师,博士生导师,E-mail: mature303@126.com

斯会议中正式提出人工智能一词,并且将人工智能作为一个新概念划分为一个独立学科,故1956年被视为人工智能元年。由于当时计算机计算能力的局限性,人工智能黄金时期持续不久便遭遇寒冬,直到1980年卡耐基梅隆大学(CMU)研发的专家系统XCON正式投入使用,人工智能再一次复苏,随着机器学习的出现,人工智能进入繁盛时期^[6,7]。工业1.0是蒸汽机时代,工业2.0是电气化时代,工业3.0是信息化时代,而人工智能的出现标志着第四次工业革命的开始,我们迎来了工业4.0,即利用信息化技术促进产业变革的智能化时代。2016年,由谷歌(Google)旗下DeepMind公司开发的人工智能软件Alpha Go与国际围棋冠军李世石五番棋大战,李世石以1:4的比分惨败,向人们昭告人工智能技术已然成熟,人工智能已经走下神坛进入我们的生活。

1.3 智能化的理论基础

1.3.1 机器学习 机器学习是人工智能的核心研究领域,是实现人工智能的一种手段,其目的在于使计算机具有智能化学习能力。机器学习的定义为利用经验来改善计算机系统自身的性能,简而言之就是计算机利用已有的数据进行学习得出规律以便对未知的新数据进行预测^[8]。机器学习又分为监督学习和无监督学习,监督学习中所有数据都有对应的人为给定的特征标签,然后在这些特征中寻找数据之间的规律。而无监督学习则是要求计算机通过寻找数据间的规律主动确定数据特征。监督学习主要面向分类及回归问题,而无监督学习主要面向聚类分析以及数据降维等问题,机器学习概念如图1所示。

1.3.2 深度学习 深度学习是机器学习的一个分支,深

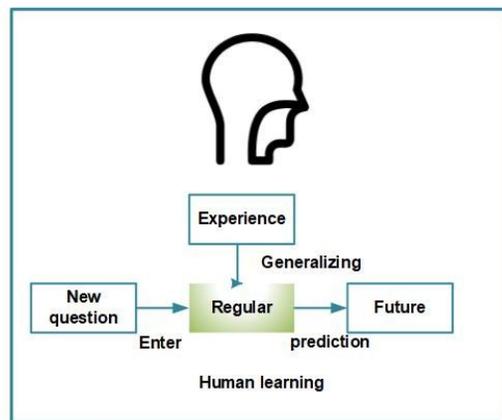
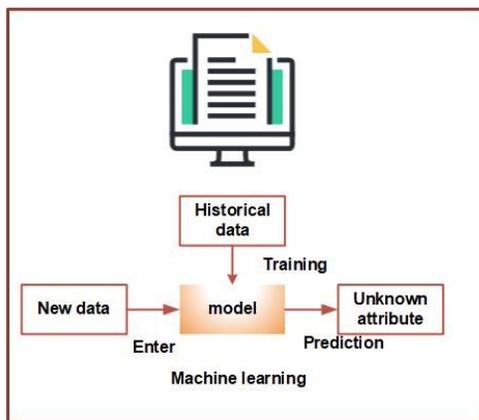


图1 机器学习概念图
Fig.1 Machine learning concept diagram.

度学习用无监督式或半监督式的特征学习和分层特征提取高效算法来替代传统特征提取算法^[9]。深度学习引入了感知器这一概念并且将神经元视作一个功能逻辑器,同时将神经元活动简化为了3个过程:输入信号线性加权,求和,非线性激活。神经元构成的神经网络由3部分组成:输入层,隐藏层,输出层,3者均由权值连接,输入层接收外界信息,隐藏层和输出层的神经元对信息进行函数处理,最后神经网络模拟人类大脑做出决策^[10]。深度学习依赖数据来提升算法准确率,数据集质量越高,深度学习对特征的学习效果就越好,深度学习概念如图2所示。

1.3.3 强化学习 强化学习也是机器学习的一个分支,强化学习有两个关键概念:策略和奖励。强化学习就是智能系统通过反复的执行某些策略并且得到反馈,然后朝奖励信号最大的方向学习,强化学习的自我学习通过实践来实现并且通过实践找到最好的结果。深度学习和强化学习首先都是自主学习系统。深度学习从训练集数据中学习是一种静态学习,但强化学习是通过不断

犯错再得到反馈来调整自身的策略以获得最优结果,是一种动态学习^[11-12]。事实上,深度学习与强化学习并不是相互排斥的概念,深度学习作为运作机制,强化学习进行算法优化,两者结合形成了一个更接近人类级别的通用智能。

1.3.4 迁移学习 深度学习和强化学习任务均从零知识开始,这意味着模型的结构和参数开始是随机猜测的。除此之外,实现机器学习需要大量标注好的高质量数据以及强大的计算资源和算力基础。因此学者们考虑对已有大数据集进行最大程度的开发,于是迁移学习诞生了。迁移学习也是实现机器学习的一种手段,迁移学习可以理解为将计算机在一个任务上学习到的知识迁移到另一个新任务上,作为其起点继续对新任务进行学习。迁移学习包括基于实例的迁移学习,其特点为源域实例类似目标域实例;基于特征的迁移学习,其特点为源域特征与目标域特征有重叠;以及在训练阶段共享一部分参数或者先验分布的基于共享参数的迁移学习^[13],迁移学习概念如图3所示。

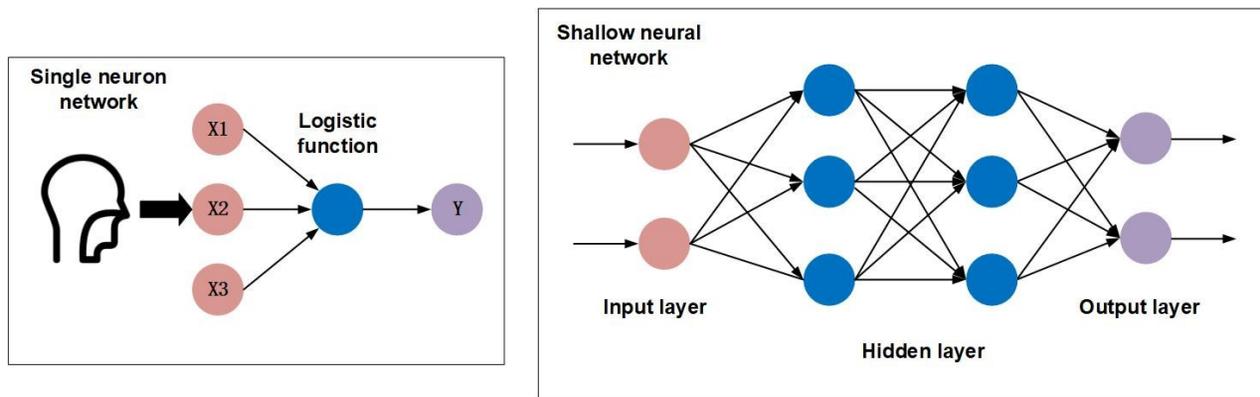


图2 深度学习概念图
Fig.2 Deep learning concept diagram.

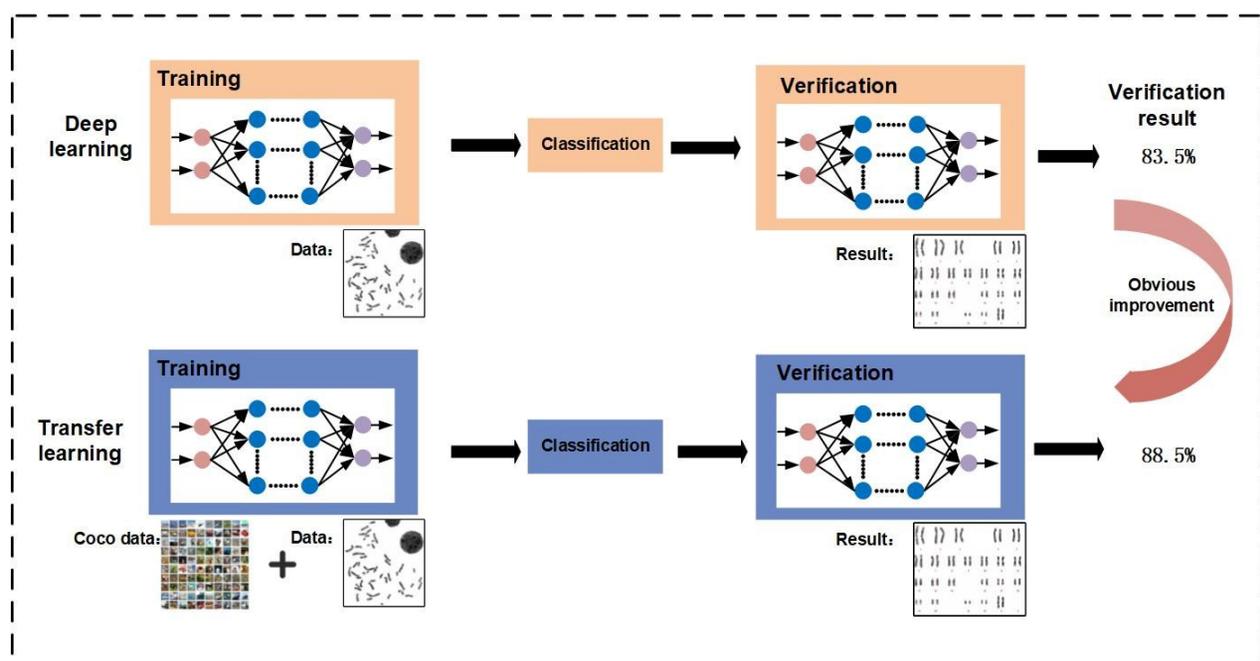


图3 迁移学习概念图
Fig.3 Transfer learning concept diagram.

1.4 智能化的实践基础

1.4.1 大数据 人工智能有3大要素:算法、数据、算力,数据是人工智能的基础之一。人工智能的发展需要通过特定数据进行反复的学习,总结经验与知识。数据的质量往往对人工智能算法的好坏有决定性作用,统一标准的数据集是对人工智能算法优劣进行准确评价的保证,大数据技术的发展迎来了人工智能发展的春天^[14]。“大数据”这个概念正式被提出是在2011年,并且在人工智能的催生下产生了大量优秀的带标注或反馈信息的数据集,如 ImageNet, COCO, 人脸数据库(AFLW)等。ImageNet是世界上最大的图像识别数据库,模拟人类的识别系统建立,目前 ImageNet 中共有 14,197,122 幅图像,涵盖了 21,841 个类别,每个类别中又有上百张图片。利用 ImageNet 训练的不同模型常在迁移学习中用做预训练,预训练在缩短计算机训练时间的同时

还可以提升模型训练效果。AFLW 人脸数据库是一个多姿态、多视角的大规模人脸数据库,该数据库大概包括 25,000 万张人脸图像,59%为女性,41%为男性,且每个人脸都被标注了 21 个特征点,这个数据集在人脸识别领域有极高的研究价值^[15]。

1.4.2 算力基础 1956 年,人工智能诞生却被当时时代的局限性限制,计算机计算能力无法达到人工智能的要求,人工智能一度面临寒冬。近年来与人工智能算法数据处理相关的芯片蓬勃发展,大致可以分为四种:CPU、GPU、NPU、TPU^[16]。CPU 擅长高效率处理指令,适合处理逻辑控制密集的计算任务。GPU 擅长图形类或非图形类无逻辑数据高度并行计算,更适合处理数据密集的计算任务,所以 GPU 比 CPU 更适合深度学习的大量矩阵、卷积运算^[17-18]。TPU 是为机器学习量身定做的一款计算芯片,以张量为处理单位可以加速人工智能算法框

架TensorFlow的运行,相比CPU和GPU,为机器学习量身定做的TPU执行每个操作所需的晶体管数量更少,效率更高^[19]。NPU是嵌入式神经网络处理器,专门为物联网人工智能而设计,面向神经网络算法进行加速,典型代表为中科院研制的寒武号系列^[20-21]。

1.4.3 软件平台 机器学习软件平台包括Caffe, TensorFlow, Keras, PyTorch等机器学习算法框架,这些框架的出现将机器学习常用的算法和模型封装成固定的组件,降低了机器学习的门槛。机器学习算法工作随着各类框架的出现变得通俗易懂,算法框架的出现不仅保证了运用的灵活也保证了运算的效率,目前机器学习领域各类优秀算法的百花齐放离不开软件平台的技术支持^[22-23]。

2 人工智能催化医学检验革命

2.1 检验标本处理

2.1.1 标本采集机器人 当前最典型,最成熟的标本采集机器人当属全自动采血机器人,全自动采血机器人通过红外线照射配合超声波可以得到与血管位置、直径相关的数据。机器视觉可将血管位置、直径等相关数据作为输入,选择合适的算法对数据进行计算然后预测出血管位置、直径以及针头采血的最佳位置及角度^[24]。王涛等人提出了一种新的基于计算机图像从血管造影图像中自动提取冠状动脉的图像分割方法。该方法基于多尺度自适应Hessian的增强方法对血管造影图像进行增强后,与统计区域合并技术结合从图像中分割出血管轮廓,提取中心线并且计算出血管直径^[25-26]。自动采血机器人利用了智能识别技术,用智能机器取代了人工,减少了人类的主观性失误,将工作人员从繁重且危险的工作中解脱出来。

2.1.2 样本稀释机器人 实验室及检验科所采集的样本

一般浓度较高,需要稀释到不超出检测方法学的线性范围。尿液,粪便及阴道分泌物有形成分显微镜检测时,样本均需要稀释到合适浓度。检测时,医务人员通常是直接将未进行稀释的样品进行检测,发现检测效果不佳后再做稀释。如阴道分泌物虽然呈液体状可以直接上镜检测,但是太浓稠的样本会导致镜检时有形成分黏连重叠,不利于医务人员观察。样本稀释机器人分为两类,一类利用光学原理,带有一个光电检测器,将透过样本检测到的透射光的大小作为输入,输出则为该样本是否需要稀释。另一类则需要具体的样本中某类成分的浓度数据。有学者利用基于多层感知器的人工神经网络对心肌肌钙蛋白I检测结果数据集进行建模得到可以用于判断心脏病患者体液样品是否需要稀释的算法模型,对不需要和需要进行稀释样品的正确推断率分别为100%和86.2%^[27]。样本稀释机器人的出现不仅解决了样品处理过程中耗时过长的的问题,也减少了样品浪费和多余的繁琐操作。

2.1.3 样本传送机器人 随着人工智能在商品物流上的应用趋向于成熟,人们开始把眼光转向人工智能在医学检测样品传送上的应用。样本的传送在检验科样品前处理过程中有重要地位,将直接影响检验的准确性。样本传送机器人可以通过机器视觉或者结合化学手段智能识别不同的样本和试剂,并且按照计算机的指令在固定区域装货卸货。样本传送机器人的出现让工作人员能够从医院各个病区、门诊、急诊中及时地获取标本并且送到指定的检验科室。如图4是ROBO公司的物流机器人RoboGo医用无人车,分别为密封型和开放型,样本传送机器人的无菌环境及其智能系统保证了运输的安全、准确、灵活、高效。样本传送机器人的出现减少了人工送检的误差,保证了样本的安全,避免了样本污染。



图4 RoboGo医用无人车
Fig.4 RoboGo medical drone.

2.2 检验项目挖掘

2.2.1 肿瘤标志物 肿瘤标志物又被称为肿瘤标记物,是指只存在于恶性肿瘤中,由恶性肿瘤细胞产生,或是宿主对肿瘤的刺激反应而产生并能反映肿瘤发生、发展情况,监测肿瘤对治疗反应的一类物质。检验人员一般是采集肿瘤患者的血清,组织,排泄物或体液,再通过一些生化分析方法如色谱-质谱结合方法、毛细管电泳-质谱结合法和核磁共振波谱等方法^[28]进行检测得到一些特异性物质的数据。特异性物质相关数据一般比较巨大,且不是所有数据都与肿瘤相关,这一问题的存在长期影响着检验医生对病情的分析,而数据挖掘技术的出现解决了这一问题,最常用的数据分析方法包括主成分

分析和人工神经网络^[29]。有学者在研究肝癌标志物时以肝癌患者的外周血清为样本,采用基于NMR的代谢组学分析技术,结合模式识别技术:主成分分析法(PCA)、偏最小二乘判别分析(PLS-DA)和正交偏最小二乘分析(OPLS-DA)、代谢物生物功能富集等方法,找出了肝癌的肿瘤标志物并进行分析,发现包括 α -葡萄糖、 β -葡萄糖苯、丙氨酸、琥珀酸、赖氨酸、酪氨酸、谷氨酰胺、肌酸、柠檬酸盐在内的九种标志物水平升高,包括低密度脂蛋白(LDL)、极低密度脂蛋白(VLDL)、N-乙酰基糖蛋白(NAG)、异亮氨酸、缬氨酸、胆碱酸盐和丙酮物质在内的7种标志物水平降低,并绘出了肝癌标志物通路图如图5^[30-31]。

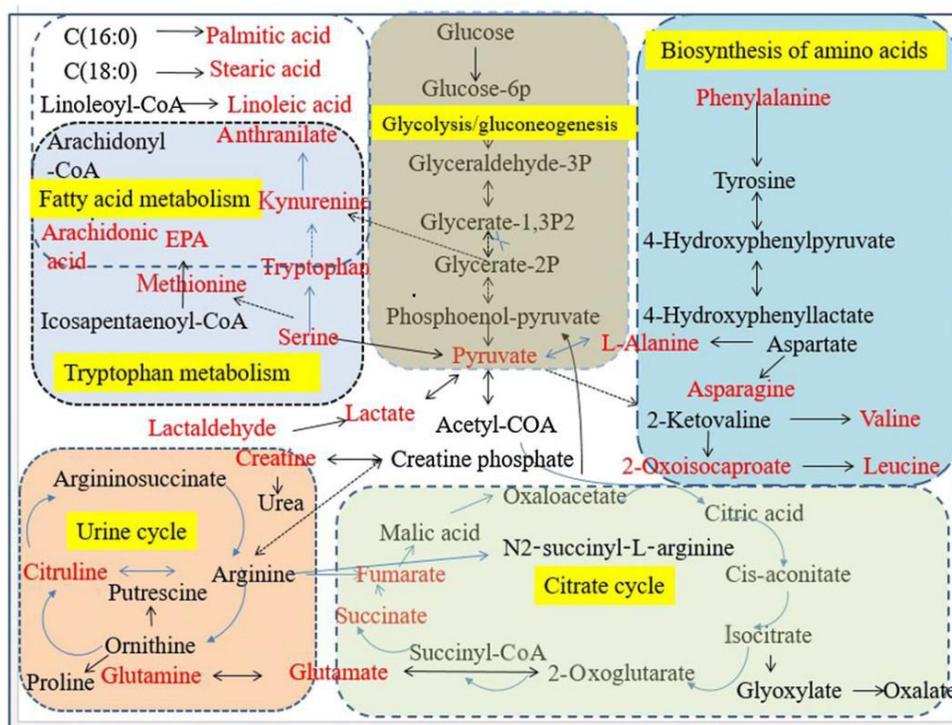


图5 肝癌标志物通路图^[30]
Fig.5 Liver cancer marker pathway map^[30].

2.2.2 核酸分子 核酸分子相关的数据挖掘常用方法有两类,一类是针对核酸序列进行预测分析,另一类是利用核酸与蛋白质相互比较进行预测分析。前者是通过分析核酸序列找出某种特定基因或者功能位点的位置,后者是通过比较核酸序列和蛋白质序列之间的相似性区域和保守性位点推测两者的进化关系^[32-33]。针对核酸序列进行预测分析的核心是核酸序列相似性研究,研究思路是给定一段核酸序列,在某个核酸序列库中查找与该段核酸序列相似性大于某个阈值的序列,典型的基于遗传算法的核酸相似性对比算法有SAGA(sequence alignment by geneticalgonthm)算法和MUSLE(multiple sequence alignmem)算法。Rybicki等人利用

SAGA算法基于肺和器官受累表型的协变量进行了多点连锁分析,证明了影响非裔美国人结节病临床表现的基因可能与疾病易感性的基因不同^[32]。郑霁等人以程序性细胞死亡相关基因PDCD5为靶分子,基于数据挖掘对其核酸与蛋白质序列关系进行分析,采用微阵列方法进行表达谱分析,得出PDCD5不仅仅参与细胞凋亡过程,还通过与核糖体的某些潜在作用影响特定蛋白质的翻译表达^[34]。

2.2.3 药物基因组学 药物基因组学结合了基因组学和分子药理学,通过计算机科学以及数学理论研究基因序列变异及其对应药物反应,高效地为特定人群寻找合适药物。检验人员一般通过服药后患者的体液或者排泄

物获取药物敏感性数据。Dong等^[35]使用癌细胞系百科全书(CCLE)中的基因表达特征和药物敏感性数据,基于支持向量机(Support Vector Machine, SVM)和递归特征选择工具来构建预测因子。通过在独立的数据集癌症基因组计划(Cancer Genome Project, CGP)进行交叉验证,验证了他们的模型在对GDSC(Genomics of Drug Sensitivity in Cancer, 抗癌药物敏感性基因组学)数据集常用的几种药物与其对应基因表达关系的预测具有较好的稳健性。Daemen等^[36]基于小二乘支持向量机和随机森林算法对70种乳腺癌细胞系和90种对应抗癌药物反应的相关分子特征进行分析,最终发现治疗反应是在基因组的多个水平上进行的,将患者与基于转录亚型的治疗相匹配将提高药物应答率。

2.3 形态学检验

目前检验科进行细胞形态学分析时均是以人工镜检为金标准,但是人工镜检主观因素影响大,需要专业的检验科人员进行操作,操作繁琐、耗费时间长,不适用于大规模筛查。人工智能的出现为形态学检验的困境带来了转机,人工智能辅助形态学检验的思路一般可以分为检测任务,分割任务和分类任务。检测任务一般是定位目标细胞并进行计数,分割任务一般是分割目标后进行面积或体积的量化分析,分类任务一般是面对种类较多较复杂的形态学分析时在检测或分割的基础上对目标细胞进行进一步的形态学分析,图6中列举了几种典型的细胞形态学分析的应用。

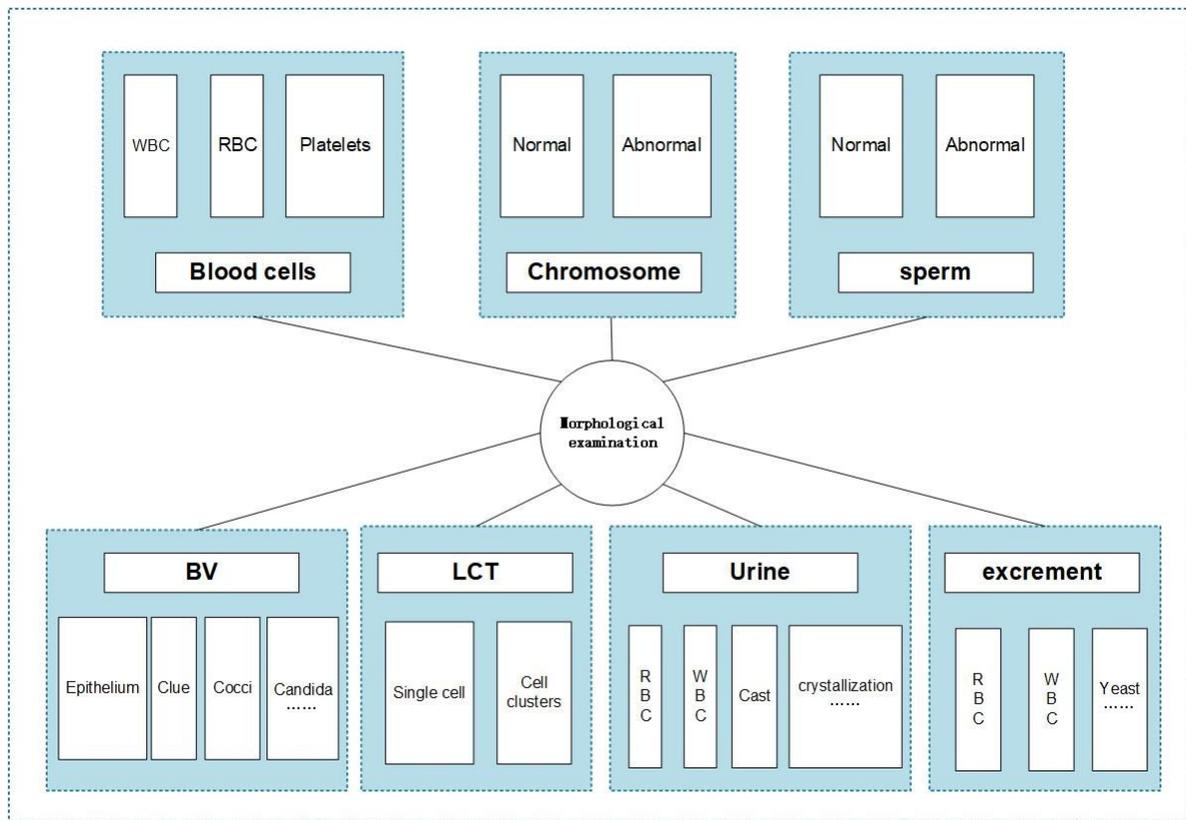


图6 人工智能辅助形态学检验
Fig.6 Artificial intelligence assisted morphology test.

2.3.1 血细胞形态学分析 血常规检查是诊断血液系统相关疾病的最基本检验项目,检验内容包括血细胞计数和血细胞形态学检查两个方面。由于血细胞形态学分析对检验人员水平要求较高且工作量大,故血细胞形态学分析是目前各级医院检验科的薄弱环节之一。血细胞形态学检查自20世纪50年代~60年代以来主要以显微镜为检测工具进行人工镜检^[37],人工智能的出现推动了自动化血细胞分析仪的发展,将检验医生们从繁重的血细胞形态学镜检工作中解放出来^[38]。陈畅等^[39]利用显微镜从外周血图片中采集分辨率为360像素×363像

素的五类白细胞单个图像作为训练数据,在深度学习框架Caffe上搭建由LetNet网络模型删减优化而得来的轻量高效新网络CCNet,该网络模型对979张5类细胞图像的最佳分类准确率达到99.69%。市面上常见的智能血细胞分析仪有希森美康XS-500i全自动血细胞分析仪和CellaVision的DM96全自动血细胞形态学分析仪等。

2.3.2 染色体核型分析 染色体核型分析是根据人类染色体数目和结构异常的类型,以推断染色体异常与疾病关系的一种诊断技术。检验医生在进行核型分析前一

般将细胞培养至有丝分裂中期并固定,再选取染色体分裂相完整独立,相对疏散,染色体形态适中无交联缠绕的部分进行分析。Abid等^[40]对近年来自动染色体分类系统领域的突破性研究进行了全面回顾,首先回顾了特征提取的方法,然后再进行基于神经网络的染色体分类器调查。特征提取主要是基于染色体形态、条带、着丝粒位置等特征进行分析,Poletti等^[41]使用从50个狭窄的骨中期细胞中提取的2287个染色体,开发了一款可以通过定位染色体着丝粒位置锁定染色体的算法。Oskouei等^[42]采用了具有两层概率的神经网络用于染色体的分类,将分类过程分为两个步骤,在第一步骤中,将染色体分为6组,在第二步中,将6组的染色体分为24个类。市面上常见的染色体核型分析系统有VideoTes公司的T-Karyo染色体核型分析系统,德适生物的AutoVision染色体智能分析工作站等。

2.3.3 精子质量分析 精子质量分析主要包括精子数量、精子体积、精子活跃程度,精子活跃程度又可以分为运动活跃型、非运动活跃型、完全不动型^[43]。精子形态学检验时,检验医生一般将精液稀释后制成样本在显微镜下肉眼观察精子形态、密度、活动率,并且进行计数。Ghasemian等^[44]开发了一套精子形态分析算法,检测和分析人体精子的不同部位,该算法在消除图像噪声的同时增强了图像的对比度,算法模型能够识别出精子的不同部分,分析每个部分的大小和形态,最后将精子分类为正常精子或者异常精子。Bijar等^[45]提出了一种分割精子顶体、核和中段的算法,该算法利用自适应混合方法(AMM)和马尔可夫随机场(MRF)模型来获取和升级分类条件概率密度函数(CCPDF)以及每个像素类别的先验概率,最后达到精子分割目的。

2.3.4 阴道分泌物有形成分分析 阴道分泌物有形成分分析一般是检验医生将稀释后的阴道分泌物制成样本在显微镜下肉眼观察上皮、线索、球菌、念珠菌、杆菌、滴虫、白细胞等有形成分的形态和数量,并且根据这些指标为阴道环境清洁度分级,是诊断细菌性阴道炎(BV)、滴虫性阴道炎、混合感染型阴道炎等多种妇科疾病的重要依据^[46]。胡静蓉等^[47]利用人工神经网络改进的模糊识别算法对白细胞进行识别,并且基于纹理特征和几何特征利用传统算法识别上皮细胞、杆菌和球菌。目前市面上已有的阴道分泌物形态学分析仪有丽拓生物的阴道分泌物分析仪LTS—V800等。

2.3.5 液基薄层细胞有形成分分析 液基薄层细胞有形成分分析主要是通过人工镜检检测宫颈是否存在细胞增生,从而判断宫颈是否病变,液基细胞检测对宫颈癌细胞的检出率为100%。郑欣等^[48]提出了一种基于YOLO(You Only Look Once)模型的算法对液基薄层细胞进行识别,该算法利用700张来自不同视野下的液

基薄层细胞进行训练。考虑到液基薄层细胞黏连严重,提取图像特征的基础网络过深会导致过拟合,故使用Resnet50作为基础网络,该算法模型对液基薄层细胞识别率为87%,对液基细胞团簇的识别准确率为75.9%。

2.3.6 尿液有形成分分析 尿液有形成分分析一般是医生将尿液制成样本直接在显微镜下用肉眼观察红细胞、白细胞、上皮细胞、各种复杂管型以及各种结晶。尿液显微镜检查是诊断泌尿系统疾病的重要手段^[49-50]。蔡立章等^[51]基于BP(back propagation)神经网络改进算法对白细胞进行识别,选用LeNet-5网络模型基于matlab平台实现网络对白细胞、红细胞、结晶的识别,然后将深度学习图像识别结果与传统算法进行比较,证明了深度学习算法的优越性。市面上常见的尿液有形成分分析仪有迈瑞的EH-2080全自动尿液有形成分分析系统等。

2.3.7 粪便有形成分分析 粪便有形成分分析一般是检验医生将稀释后的粪便制成样本在显微镜下用肉眼观察红细胞、白细胞、巨噬细胞、上皮细胞、酵母菌、脂肪滴、虫卵等有形成分,粪便有形成分分析对多种肠胃病诊断有重要参考意义,如果粪便中存在虫卵基本上可以断定患者感染寄生虫。Yang等^[52]提出了一种通过数字图像处理 and 人工神经网络从显微图像中提取粪便标本中人寄生虫卵的形态特征的算法,该算法利用了包含七个普通人类蠕虫卵的82个显微图像作为数据库。基于形状、外壳光滑度和大小的3个形态特征利用人工神经网络进行训练,算法模型正确分类率为90.3%。

2.3.8 骨髓细胞形态学分析 骨髓细胞形态学分析时,检验医生一般直接在显微镜下观察骨髓穿刺术所制得的骨髓图片,低倍镜下观察骨髓增生程度,有无异常细胞,并且对巨核细胞进行计数,油镜下对有核细胞分类计数并且观察细胞形态。骨髓形态学分析可以用于多种疾病的诊断,如急性白血病、自身免疫性疾病、感染性疾病、恶性肿瘤等。骨髓细胞形态学分析与血细胞形态学分析原理一致,只是检测细胞类别和样本获取位置有所不同。

2.4 检验数据处理

2.4.1 辅助诊断模型 医学专家系统是医学检验领域最常见的人工智能辅助诊断模型。医学专家系统集成大量医疗检验数据以及对应疾病症状和治疗预后,通过互联网与各个医院的不同科室紧密联系以保证医学信息的实时更新。人工智能利用这些医学信息学习各种检测数据与相关疾病的内在联系,最后得到一个可以诊断患者疾病或者给予可靠建议的算法模型,这样一个算法模型相当于一个人类医学专家。最早的医学专家系统是斯坦福大学的MYCIN系统,MYCIN可以辅助医生对血液感染患者进行筛查并且提出可靠性治疗建

议^[53]。John等^[54]基于贝叶斯网络,使用多层感知的人工神经网络,使用径向基函数的人工神经网络,支持向量机和专家系统分别利用入院时患者各项医学检测数据

开发可以预测是否需要重新住院的智能系统,并且通过临床试验证明除向量机外,所有模型均能提供良好的重新住院预测结果。医学专家系统概念如图7所示。

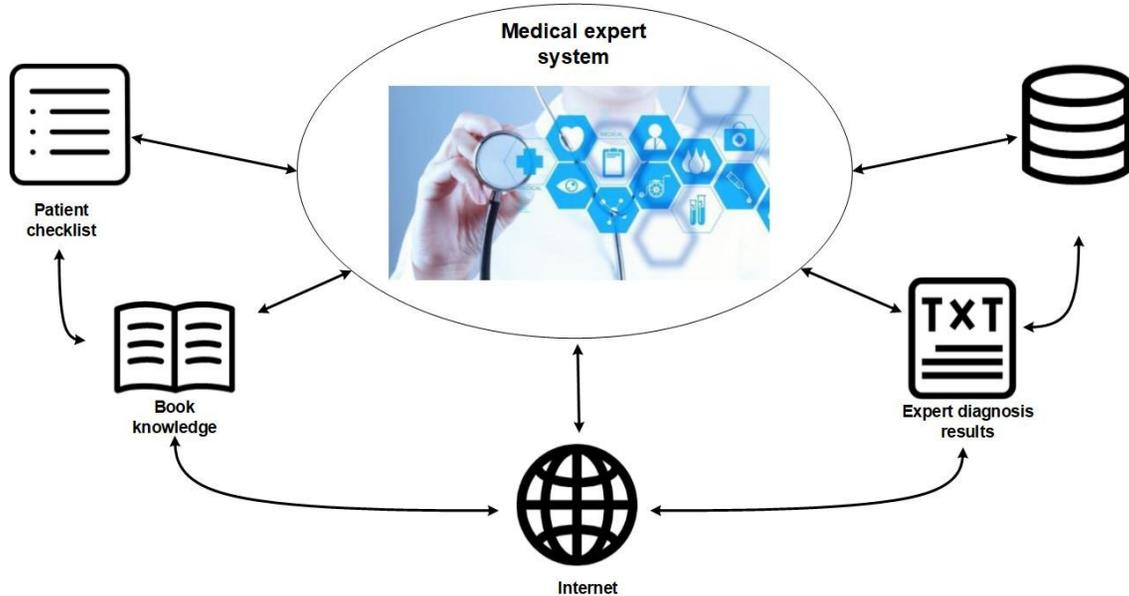


图7 医学专家系统
Fig.7 Medical expert system.

2.4.2 报告自动审核 检测报告审核是检验工作中必不可少的流程,报告审核的精准和快速保证了检测工作的高效。但是人工审核检验报告不仅耗时耗力,当检验科工作人员在高强度工作后还容易出现审核失误。为解决这些问题,自动化报告审核系统的搭建刻不容缓,而人工智能的出现为自动化报告审核系统带来了新的契机。自动化报告审核系统主要依靠HIS系统(医院信息系统)和LIS系统(实验室信息系统)实现。Shih等^[55]收集了569 001例患者检测信息作为数据库,通过LIS建立信息传递系统并且开发了一套自动验证算法,可输出极限检查,增量检查和一致性规则检查的范围。该算法消除了测试结果验证中的个体差异,缩短了周转时间(TAT),并减少了全职等效时间(FTE),从而使医疗技术人员能够投入更多的时间和精力来处理算法锁定的测试报告。

2.4.3 隐私数据保护 随着互联网医疗的兴起,医疗隐私数据的保护开始被各界关注。在医院普遍使用电子病历的当代,医学专家和计算机学者们联合起来对电子病历中的医嘱数据、诊断数据进行最大程度的开发,以补充医学知识的不足和医疗系统的缺陷。但是研究人员享受数据盛宴的同时,我们同样需要对隐私问题高度关注。胡荣磊等^[56]结合数据挖掘和数据储存技术提出了一种在大数据环境下对医学隐私数据进行脱敏的隐私保护方案,该方案先设定敏感级别再利用ETL(数据

仓库技术)对数据进行抽取,设计算法判别数据敏感级别,并且根据数据敏感级别的不同对数据进行不同程度的加密。

2.5 云计算

大数据、云计算、人工智能、物联网等一大批新兴信息技术的出现,在医疗检验领域产生了革命性的影响,推动了医疗检验事业的创新发展。云计算平台能对分布于不同地方,不同时间以及不同检验项目的数据进行收集、存储、分析、运算和质控。云平台的大部分数据处理都在云端完成,在保证数据安全、降低数据存储成本的同时为医疗诊断提供了可靠而快捷的诊断辅助,大大降低了基层医疗机构对操作人员的要求^[57]。

2.5.1 诊断字典 诊断字典是一个尚不成熟但是成长前景无限的概念,柯莉等^[58]以国际疾病分类(ICD)为基础,利用广州医科大学附属第三医院医学数据库建立了临床疾病诊断字典。该字典收集了广州医科大学附属第三医院各个临床科室疾病和诊断策略,统一规范疾病名称,将临床诊断疾病与ICD相对应。由于人力和医学资源的限制,柯莉等人所建立的字典内容是远远不够完整的,但是为我们医学检验人员更好地利用人工智能、大数据、云计算改革传统检验手段提供了一个好的思路,要想建立一个完善可靠的医学检验云计算中心,我们需要一个权威的医学标准。

2.5.2 互联网医学检验 互联网医学检验常见的技术手

段为第三方医学检验与即时检测(POCT),互联网医学检验以互联网为载体整合了云计算,大数据和线下医学检验数据,满足了医学检验系统对信息传送高效、诊断决策精准的要求。第三方检验实验室可以是实验室或者公司,服务对象为二级医院或者社区医院这些医学检验体系不够完善或者医学专家资源不足的级别较低医院,第三方实验室可以作为一个中转站,将低级别医院检验数据智能处理后个性化传送到特定的较高级别医院进行更专业的诊断。也可以是一个更高级的拥有丰富医学知识和医学诊断经验的医学专家系统,直接将诊断结果打印成报告传回较低级别医院^[59]。即时检测是医学检验领域的一个新概念,通俗的理解是一种试剂与仪器为便携式,在不固定场所均可进行的,检测报告结果即时化的检测手段。Sarwar等^[60]利用软电子技术开发了一款可以使用少量液体在柔性基板上执行电子或

光学监控,并将监控数据结合蛋白质组学进行数据挖掘分析的设备。该设备基于汗液的诊断程序可以为患者提供无痛的监测,患者在自己的家中舒适地跟踪各种医疗状况。

2.5.3 5G技术 2019年6月6日,工信部发放了5G商用牌照,标志着中国正式进入了5G商用元年,5G具备高速率、低时延和更大用户连接能力等特点,目前我国已经建成了全世界规模最大的5G网络^[61]。我国对于5G技术的应用并不仅限于通信,5G技术推动了智慧医疗的发展,让第三方实验室,即时检测等互联网检测手段变得更高效率,便捷,也为医学数据庞大数据的传送和医学数据复杂计算提供了实现的可能性。通过5G技术医护人员甚至患者能够在任意地点快速与医学云计算中心通讯,而专业检验医师也能第一时间获取患者信息并完成检测报告单。

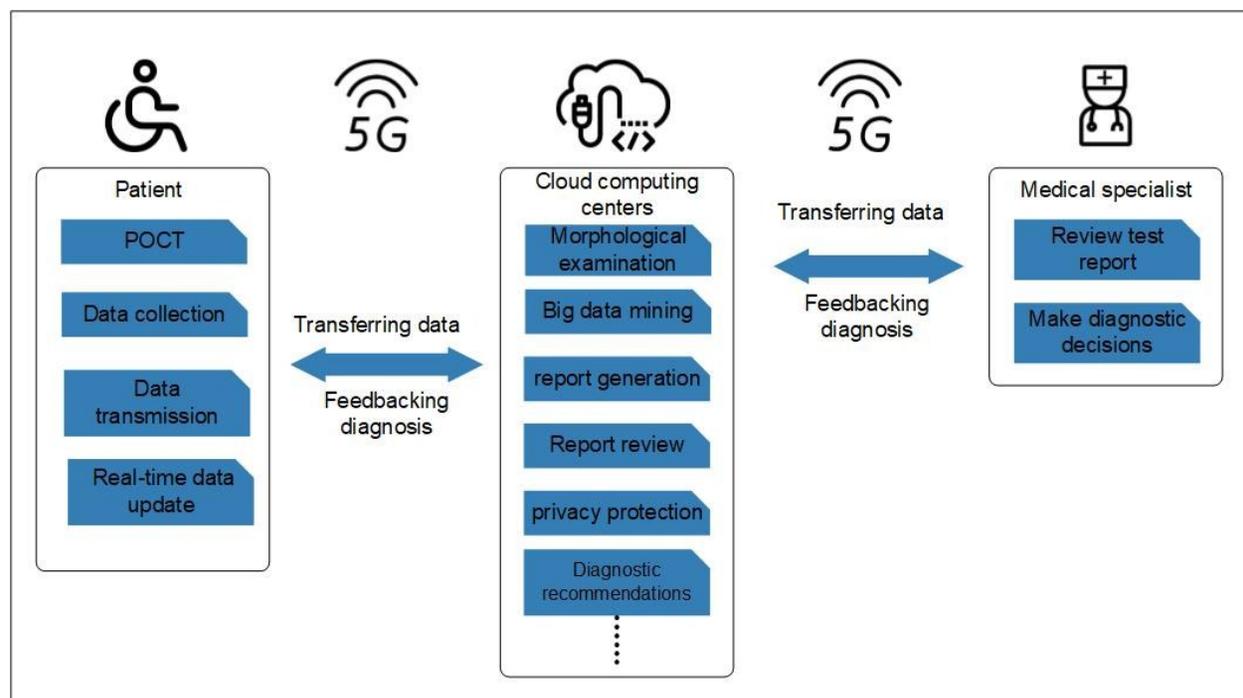


图8 5G技术在智能检测中应用
Fig.8 Application of 5G technology in intelligent detection.

3 智能化医学检验时代展望

我国人口的显著特征是人口基数大,人口老龄化严重,然而与我国对医疗服务需求大相对应的现状却是我国的医疗卫生资源分配极度不平均,经济落后地区和贫困山区的患者体验现代医学福音的机会少之又少。人工智能所携带的一系列先进检验技术如即时检测,第三方实验室,诊断字典等的出现在为医护人员带来便利,减少检测诊断误差的同时,也打破了我国医学资源分配不平均的僵局,让偏远地区的患者也能有机会享受最先进、专业的医学检测服务。

参考文献:

- [1] 张新新,刘华东. 出版+人工智能:未来出版的新模式与新形态--以《新一代人工智能发展规划》为视角[J]. 科技与出版, 2017, 12(12): 43-8.
- [2] 新一代人工智能发展规划[J]. 科技创新与生产力, 2017, 8(8): 52-66.
- [3] 田刚,何鹏,宋敏,等. 大数据背景下高校临床检验诊断学教学思考[J]. 现代医药卫生, 2019, 13(13): 2056-8.
- [4] 寿君妮,于观贞,余党会,等. 人工智能与医学--发展历程[J]. 第二军医大学学报, 39(8): 2.
- [5] 李德毅. 网络时代人工智能研究与发展[J]. 智能系统学报, 2009, 4(1): 1-6.
- [6] 陈超. 人工智能的三次发展高潮[J]. 机床与液压, 2018, 8(8): 131.
- [7] 韩韬. 计算机网络技术在人工智能中的运用[J]. 电子技术与软件工

- 程, 2017, 18(18): 29.
- [8] 何清, 李宁, 罗文娟, 等. 大数据下的机器学习算法综述[J]. 模式识别与人工智能, 2014, 27(4): 327-36.
- [9] 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. 计算机应用研究, 2012, 29(8): 2806-10.
- [10] 尹宝才, 王文通, 王立春. 深度学习研究综述[J]. 北京工业大学学报, 2015, 41(1): 48-59.
- [11] 陈学松, 杨宜民. 强化学习研究综述[J]. 计算机应用研究, 2010, 27(8): 2834-8, 2844.
- [12] 高阳, 陈世福, 陆鑫. 强化学习研究综述[J]. 自动化学报, 2004, 30(1): 86-100.
- [13] 庄福振, 罗平, 何清, 等. 迁移学习研究进展[J]. 软件学报, 2015, 26(1): 26-39.
- [14] 余滨, 李绍滋, 徐素霞, 等. 深度学习: 开启大数据时代的钥匙[J]. 工程研究-跨学科视野中的工程, 2014, 3(3): 233-43.
- [15] Hosang J, Benenson R, Piotr Dollár, et al. What Makes for Effective Detection Proposals[J]? IEEE T Pattern Anal, 2015, 38(4): 814-30.
- [16] 王珺. 人工智能, 大数据和算力的融合发展[J]. 科技经济导刊, 2018, 16(8): 132-3.
- [17] Rejimon T, Bhanja S. Scalable probabilistic computing models using Bayesian networks[C]. Symposium on Circuits & Systems. IEEE, 2005, 1: 712-5.
- [18] Sze V, Chen YH, Yang TJ, et al. Efficient processing of deep neural networks: a tutorial and survey[J]. P IEEE, 105(12): 2295-329.
- [19] Shin D, Lee J, Lee J, et al. DNPU: an Energy-Efficient Deep-Learning processor with heterogeneous Multi-Core architecture[J]. IEEE Micro, 38(5): 85-93.
- [20] Jiawei X, Yuxiang H, Kunlong Y, et al. Optimized Near-Zero Quantization Method for Flexible Memristor Based Neural Network[J]. IEEE Access, 2018, 6: 29320-31.
- [21] Ren L, Fletcher CW, Kwon A, et al. Design and implementation of the ascend secure processor[J]. IEEE T Depend Secure, 2017, 16(2): 204-16.
- [22] 余伟豪, 李忠, 安建琴, 等. 深度学习框架和加速技术探讨[J]. 软件, 2017, 6(6): 87-90.
- [23] Nikhil Ketkar. Introduction to Tensorflow[M]. Deep Learning with Python. 2017.
- [24] 陈楷丰. 基于改进Hessian矩阵的冠脉造影图像增强算法研究[J]. 网络安全技术与应用, 2018, 7(7): 52-4.
- [25] 李颖超, 刘越, 王涌天. 基于多尺度Hessian矩阵和Gabor滤波的造影图像冠脉中心线提取[J]. 中国医学影像技术, 2007, 23(1): 133-6.
- [26] Wan T, Shang X, Yang W, et al. Automated coronary artery tree segmentation in X-ray angiography using improved Hessian based enhancement and statistical region merging[J]. Comput Methods Programs Biomed, 2018, 157(1): 179-90.
- [27] Ialongo C, Pieri M, Bernardini S. Artificial neural network for total laboratory automation to improve the management of sample dilution: smart automation for clinical laboratory timeliness[J]. SLAS Technol, 2017, 22(1): 44-9.
- [28] Leung SM, Pitts RL. A novel approach using MALDI-TOF/TOF mass spectrometry and prestructured sample supports (AnchorChip Technology) for proteomic profiling and protein identification.[J]. Methods Mol Biol, 2008, 441: 57-70.
- [29] Qi S, Xu Y, Luo R, et al. Novel biochemical insights in the cerebrospinal fluid of patients with neurosyphilis based on a metabonomics study[J]. J Mol Neurosci, 2019, 69(1): 39-48.
- [30] Qi S, Huang S, Chen X, et al. Liver tissue metabolic profiling and pathways of non-alcoholic steatohepatitis in rats[J]. Hepatol Res, 2017, 47(13): 1484-93.
- [31] Qi SW, Tu ZG, Peng WJ, et al. ¹H NMR-based serum metabolic profiling in compensated and decompensated cirrhosis[J]. World J Gastroenterol, 2012, 18(3): 285-90.
- [32] Rybicki BA, Sinha R, Iyengar S, et al. Genetic linkage analysis of sarcoidosis phenotypes: the sarcoidosis genetic analysis (SAGA) study[J]. Genes Immun, 2007, 8(5): 379-86.
- [33] Zhang C, Wong AK. A genetic algorithm for multiple molecular sequence alignment[J]. Bioinformatics, 1997, 13(6): 565-81.
- [34] 郑冀, 马大龙. 人程序化死亡分子5(PDCD5)核酸和蛋白质序列的数据发掘[J]. 北京大学学报: 医学版, 2003, 35(4): 353-9.
- [35] Dong Z, Zhang N, Li C, et al. Anticancer drug sensitivity prediction in cell lines from baseline gene expression through recursive feature selection[J]. BMC Cancer, 2015, 15(1): 489.
- [36] Daemen A, Griffith OL, Heiser LM, et al. Erratum to: modeling precision treatment of breast Cancer[J]. Genome Biol, 2015, 16(1): 95.
- [37] 于红梅, 汤冬静, 刘晨. 外周血细胞形态学检查的重要性[J]. 医学检验与临床, 2012, 23(6): 88-90.
- [38] Liu Z, Xu D, Li T, et al. A stain normalization method for the white blood cell image [C]. International Conference on Automation, Computational and Technology Management (ICACTM), 2019: 19-23.
- [39] 陈畅, 程少杰, 李卫滨, 等. 基于卷积神经网络的外周血白细胞分类[J]. 中国生物医学工程学报, 2018, 37(1): 17-24.
- [40] Abid F, Hamami L. A survey of neural network based automated systems for human chromosome classification[J]. Artif Intell Rev, 2018, 49(1): 41-56.
- [41] Poletti E, Grisan E, Ruggeri A. A modular framework for the automatic classification of chromosomes in Q-band images [J]. Comput Methods Programs Biomed, 2011, 105(2): 120-30.
- [42] Oskouei BC, Shanbehzadeh J. Chromosome classification based on wavelet neural network [C] proceedings of the International Conference on Digital Image Computing: Techniques and Applications. Sydney, Australia, 2010: 605-10.
- [43] Qi S, Nie T, Li Q, et al. A sperm cell tracking recognition and classification method [C] proceedings of the 2019 International Conference on Systems, 2019: 163-7.
- [44] Ghasemian F, Mirroshandel SA, Monji-Azad S, et al. An efficient method for automatic morphological abnormality detection from human sperm images [J]. Comput Methods Programs Biomed, 2015, 122(3): 409-20.
- [45] Bijar A, Mikaeili M, Benavent AP, et al. Segmentation of sperm's Acrosome, nucleus and mid-piece in microscopic images of stained human semen smear [C] proceedings of the 2012 8th International Symposium on Communication Systems, 2012: 1-6.
- [46] Li Q, Wu Y, Mei J, et al. A method of automatically calculating vaginal cleanliness degree from leukorrhoea's microscope images [C] proceedings of the 2019 International Conference on Automation, 2019: 24-29.

(上接296页)

- [47] 胡静蓉. 白带显微图像中有型成分的智能识别技术研究[D]. 成都: 电子科技大学, 2017.
- [48] 郑欣, 田博, 李晶晶. 基于YOLO模型的宫颈细胞簇团智能识别方法[J]. 液晶与显示, 2018, 33(11): 965-71.
- [49] Xiang H, Chen Q, Wu Y, et al. Urine Calcium oxalate crystallization recognition method based on deep learning[C]proceedings of the 2019 International Conference on Automation, 2019: 30-3.
- [50] Li Q, Yu Z, Qi S, et al. A recognition method of urine cast based on deep learning[C/proceedings of the 2019 International Conference on Systems, 2019: 157-161.
- [51] 蔡立章. 基于卷积神经网络的体液细胞图像有形成分的特征识别方法研究[D]. 湘潭: 湘潭大学, 2017.
- [52] Yang YS, Park DK, Kim HC, et al. Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network[J]. IEEE Trans Biomed Eng, 2001, 48(6): 718-30.
- [53] Shortliffe EH, Davis R, Axline SG, et al. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system[J]. Comput Biomed Res, 1975, 8(4): 303-20.
- [54] Showalter JW, Subramanian GH. Predicting re-hospitalisations using intelligent systems: an exploratory study[J]. Int J Bus Inform Syst, 2013, 13(2): 151-72.
- [55] Shih MC, Chang HM, Tien N, et al. Building and validating an autoverification system in the clinical chemistry laboratory[J]. Lab Med, 2011, 42(11): 668-73.
- [56] 胡荣磊, 何艳琼, 曾萍, 等. 一种大数据环境下医疗隐私保护方案设计与实现[J]. 信息安全, 2018, 18(9): 48-54.
- [57] W BW, Rost P, Bartelt JS, et al. Benefits and impact of cloud computing on 5G signal processing[J]. IEEE Signal Process Mag, 2014, 31(6): 10.
- [58] 柯莉, 林伟研, 刘养奇, 等. 临床疾病诊断字典库的应用研究[J]. 中国数字医学, 2019, 6(6): 81-3.
- [59] 张婉, 曹永彤, 王云亭. 第三方医学检验机构的市场作用及发展前景分析[J]. 中国医院, 2015, 19(8): 27-8.
- [60] Sarwar M, Rodriguez P, Li CZ. Sweat-Based *in vitro* diagnostics (IVD): from sample collection to Point-of-Care testing(POCT) [J]. J Anal Test, 2019, 3(1): 80-8.
- [61] Chen SZ, Kang SL. A tutorial on 5G and the progress in China[J]. Front Inform Tech EI, 2018, 19(3): 309-21.

(编辑:余诗诗)