

人工智能在微流控系统中的应用

王宇 方群*

(浙江大学化学系微分析系统研究所, 杭州 310058)

摘 要 微流控系统可以对微流体进行精准操控,因此被广泛应用于化学、生物、医学、药学等诸多领域。近年来取得跨越式发展的人工智能技术,在应对微流控系统中海量数据的分析与挖掘方面具有很大的优势,被应用于很多微流控系统中,并在生物研究、医学诊断、药物研发等多个领域展现了突出的潜力。本文综述了数种典型的人工智能模型及其在微流控系统中的应用,重点介绍了人工智能在微流控系统的目标检测、关联预测与结果分类方面的研究和应用进展,并对其未来发展趋势进行了展望。

关键词 微流控; 人工智能; 数据挖掘; 大数据; 评述

1 引言

微流控(Microfluidics)系统是指在微米级结构中操纵微量流体(体积为 $10^{-9} \sim 10^{-12}$ L)的系统,由于其对微流体的精准操控能力,被广泛应用于化学、生物、医学、药学等领域,具有试样和试剂消耗量小、反应效率高、分析通量高、体积小、集成化和自动化程度高等优点,也常被称为芯片实验室(Lab on a chip)。近 30 年来,微流控系统在生命科学基础研究^[1~4]、化学分析^[5,6]、临床诊断^[7,8]、药物研发^[9,10]、食品分析^[11]、环境监测^[12]等多个应用领域展现出巨大的应用潜力,改变了许多传统的实验方法。微流控系统凭借强大的微量样品操纵能力,配合多种检测手段和完备的数据采集系统,可以高通量、自动化、低成本地产生海量高质量数据,包括实验图像、流体参数、化学和生物检测信号等。但是,这些数据所包含的信息通常是不可直接获取的。以实验图像为例,一张单细胞图像中包含了细胞大小、细胞形态、表面形貌、运动方向、荧光强度、与周围环境的联系等信息,但却很难直接从图像中提取。由于微流控系统在高通量、自动化等方面的长足进展,在展示其强大的获取海量实验数据能力的同时,也使其在数据分析与挖掘方面面临着严峻的挑战。

人工智能(Artificial intelligence, AI)作为研究和开发用于模拟、延伸和扩展人的智能的理论、方法、技术、系统及应用的技术和科学,具有强大的分析与运算能力,能够处理复杂环境下的决策问题,在大数据的分析与挖掘方面具有天然的优势。尽管在 20 世纪 50 年代左右人工智能技术就已经诞生^[13],并引领过多次研究热潮,但是受限于数据量与计算能力,一直未能展现出足够的数据处理能力。近几年,得益于神经网络模型的发展、快速图形处理器(GPU)^[14]的运用、Tensorflow^[15]等框架的开发和分布式计算的普及,人工智能技术已经在图像识别^[16]、自然语言翻译^[17]、语音处理^[18]、个性化推荐^[19]等多个领域大放异彩,成为数据分析与挖掘的首选利器。不同于传统的计算机算法,基于人工智能开发的机器学习技术通常无需直接编程即可学习,最终数据处理性能的优劣主要取决于其算法训练过程。因此,在一些复杂度较高、特征不明显的场景(如图像识别或自然语言翻译等)中,人工智能算法相比于传统计算机算法具有无可比拟的优势。

近年来,随着人工智能在多个领域取得应用突破,越来越多的微流控领域学者开始关注其强大的分析与分类能力,并针对微流控系统发展了许多人工智能数据分析与挖掘方法,不仅大大提高了系统的数据处理能力,还构建了许多状态评估与结果预测模型,使得微流控系统逐渐从自动化向智能化发展。其中,人工智能所擅长的回归、分类和聚类等问题尤其受到了微流控领域学者的青睐,并逐渐将其应用于目标检测、关联预测和结果分类等多个方面,显示了突出的应用潜力和发展前景。

2019-11-18 收稿; 2020-01-16 接受

本文系国家自然科学基金项目(Nos. 21827806, 21435004)资助

* E-mail: fangqun@zju.edu.cn

本文对三类典型的人工智能模型以及其在微流控系统中的应用进行综述,并对其未来发展趋势进行展望。

2 人工智能模型简介

人工智能是一个相对宽泛但又比较通俗直观的概念,旨在模拟、延伸和扩展人的智能。目前,人工智能主要用于解决 4 个方面的问题,即回归、分类、聚类与强化问题。其中,回归问题指的是根据数据样本中的特征参数,预测连续值结果,例如预测癌细胞侵袭概率;分类问题指的是根据数据中的特征,判定其所属类别,如编码液滴的识别;聚类问题指的是将样本划分为若干个类别,如疾病亚型的划分;强化问题指的是如何基于环境而行动,以取得最大收益,如临床给药剂量的优化^[20]。此外,除了强化学习自成体系外,根据训练数据集是否明确输入与输出的映射关系,又可以分为有监督学习(包括半监督学习)和无监督学习。其中,回归问题与分类问题一般属于有监督学习,而聚类问题通常属于无监督学习。

线性回归是一种基本的回归算法,可以用一条直线较为精确地描述数据之间的关系。如果数据间的关系并非线性,也可以使用多项式回归、岭回归等其它更加复杂的回归算法。因此,回归算法可以较好地描述连续数据间复杂的关系,是一种发现多种参数间关联,并定量描述其对结果影响的方法,常用于函数拟合、模型开发等。

相比于输出结果为连续且确定数值的回归问题,分类问题的输出结果却是定性的离散值。

分类模型有多种类型,如用回归算法预测出事件概率,并设定判断阈值,就是一个简单的二元分类算法,即逻辑回归。此外,朴素贝叶斯分类(Naïve Bayesian classification, NB)是一种以贝叶斯理论为基础,依据概率原则进行分类的算法。其基本原理为,根据已知的先验概率,利用贝叶斯公式求出样本属于某一类的后验概率,然后选择后验概率最大的类作为该样本所属的类。 k 近邻(k -Nearest neighbor, k NN)算法则是一种基于距离度量的分类算法,将整个数据集作为训练集,先确定待分类样本与每个训练样本之间的距离,再找出与待分类样本距离最近的 k 个样本作为其 k 个近邻,最终选择占比最大的类别作为待分类样本的类别。决策树(Decision tree)算法是根据数据属性采用树状结构建立决策模型的算法,常用于解决分类和回归问题。若将多个决策树模型作为基分类器,投票决定最终类别,即成为随机森林(Random forest, RF)模型。支持向量机(Support vector machine, SVM)是一种基于统计学 VC 维理论与结构风险最小原理的有监督二分类器,其决策边界是对学习样本求解的最大边距超平面。

除了以上几种典型的分类算法,人工神经网络(Artificial neural network, ANN)作为人工智能近年来的研究热点,更是体现出了极强的实用性能。ANN 是一种模拟生物神经网络进行信息处理的数学模型,由大量神经元节点按一定的体系架构连接成网状结构,一般包含输入层、隐含层和输出层。根据给定的训练样本,调整人工神经网络参数,即可使网络输出接近于已知样本所在类的标记。若神经网络的深度很深、参数很多,模型的复杂度通常很高,也意味着可完成比较复杂的学习任务,即成了典型的深度学习(Deep learning)模型。ANN 有很多分支,如卷积神经网络(Convolutional neural network, CNN)、循环神经网络(Recurrent neural network, RNN)等,都应用广泛。其中, CNN 是一类包含卷积计算且具有深度结构的前馈神经网络(Feedforward neural network, FNN),是深度学习的代表算法之一,由于可以将图像作为输入数据,现已成为计算机视觉领域的主流研究方法。RNN 则是一类以序列数据为输入,在序列的演进方向进行递归,且所有节点(循环单元)按链式连接形成闭合回路的递归神经网络。该网络的特点是对之前的信息具有“记忆”功能,并将其应用于当前的输出计算中,因而在对时间序列进行分析时具有独特的优势。

聚类问题也有很多算法模型可以支持。例如,通过将样本划分 k 个集群的 k -均值(k -Means)算法、基于质心计算的均值漂移聚类、基于密度的 DBSCAN 聚类、基于余弦距离的余弦相似度聚类、在不同层次对数据进行划分的层次聚类等。此外,通过将高维数据降维,也可以实现类似聚类的效果,比如主成分分析(Principal component analysis, PCA)。

强化问题的对应模型主要通过与环境交互提升性能,但目前在微流控领域几乎没有应用。

表 1 是对几种典型的人工智能模型的简单介绍。在实际应用中,对典型模型进行优化与改进,或将

多种模型联合使用,都是解决复杂问题时常用的方法。

表 1 经典人工智能模型及其应用

Table 1 Classic artificial intelligence models and their applications

目标问题 Target problems	经典模型 Classic models	代表算法 Representative algorithms	应用 Applications
回归问题 Regression problem	线性模型 Linear model	线性回归、多项式回归、岭回归 Linear regression, polynomial regression, ridge regression	定量描述参数关联、函数拟合、模型开发、概率预测、状态评估等 Quantitative description of parameter correlation, function fitting, model development, probability prediction, state assessment, etc.
分类问题 Classification problem	线性模型 Linear model	逻辑回归 Logistics regression	图像识别、目标检测、已知类别分类、目标动态跟踪、场景综合评估等 Image recognition, target detection, classification of known categories, target dynamic tracking, comprehensive scene assessment, etc.
	贝叶斯模型 Bayesian model	朴素贝叶斯分类 (NB)	
	度量学习 Metric learning	k 近邻 k-Nearest neighbor (kNN)	
	决策树 Decision tree	随机森林 Random forest (RF)	
	统计学习 Statistical learning	支持向量机 Support vector machine (SVM)	
聚类问题 Clustering problem	人工神经网络 Artificial neural network (ANN)	卷积神经网络、循环神经网络 (CNN), Recurrent neural network (RNN)	潜在关联发现、未知类别聚类、现象发现、关联效果预测等 Potential association discovery, unknown category clustering, phenomenon discovery, association effect prediction, etc.
	原型聚类 Prototype-based clustering	k 均值 k-means	
	质心聚类 Centroid-based clustering	均值漂移 Mean shift	
	密度聚类 Density-based clustering	DBSCAN Density-based spatial clustering of applications with noise (DBSCAN)	
	距离度量 Distance measure	余弦相似度聚类 Cosine similarity clustering	
	层次聚类 Hierarchical clustering	AGNES AGglomerative NESTing (AGNES)	
强化问题 Reinforcement problem	降维 Dimension reduction	主成分分析 (PCA)	通过与环境的交互提升性能,目前在微流控系统中几乎无应用 Improving performance through interaction with the environment, no application in microfluidic systems currently
	时序差分学习等 Temporal difference (TD), etc.	Q-学习 Q-learning	

综上,人工智能包含了多种算法原理,包括基于数学与统计学的算法和基于神经网络的深度学习算法。因此,如何选择合适的的人工智能模型,灵活运用,并最终取得理想的效果,对研究者在人工智能领域的知识积累与实践能力有较高的要求。以微流控系统中的应用为例,在图像识别、目标检测方面,可以使用 CNN 模型等分类算法与均值漂移等聚类算法;在关键参数发现、关联效果预测方面,可以使用回归算法与 ANN 模型;在结果分类、现象发现方面,可以使用分类算法、聚类算法以及其它 ANN 模型。

3 人工智能在微流控系统目标检测中的应用

目前,人工智能在微流控系统中一个十分重要且广泛的应用,就是对微流控装置内海量的成像图片进行快速精准识别,以获得更加强大的目标检测能力。其中,CNN 是人工智能技术在图像识别领域最主要的研究模型。

虽然,人工智能相关的研究已经有了数十年的历史,但近十年才发展出了包含三层及以上隐含层的

深度神经网络模型,并展现了其在性能上的显著提升。2012 年,Krizhevsky 等^[14]提出了 AlexNet 模型,并利用 GPU 参与相关计算,使得模型训练速度得到数量级的提升。此后,Tensorflow^[21]和 Caffe^[22]等深度学习框架逐渐推广,更大规模的数据得以用于训练更为复杂的神经网络模型,最终使得深度学习性能实现了质的突破。因此,目前所认知的人工智能技术是在近几年模型复杂度、运算能力、数据量都得到大幅提升后,才展现出其巨大潜能的。

在微流控系统中,对于微小目标的检测一直是一个较大的挑战,尤其是当检测图像的分辨率较低时,检测更为困难。Huang 等^[23]利用微流控通道与互补金属氧化物半导体(CMOS)图像传感器构建了一个微型化的无透镜细胞计数系统,并使用极限学习机(Extreme learning machine, ELM)与 CNN 模型进行超分辨率(Super-resolution, SR)处理(图 1A),将细胞图像的分辨率提高了 4 倍,成功实现了血细胞计数。此后,又通过改进硬件结构,对最小直径仅 2 μm 的血小板进行了识别与计数^[24]。人工智能超分辨率技术的运用,使得微流控系统有望突破光学等检测手段在分辨率上的限制,实现更高精度的目标检测。

流式细胞术(Flow cytometry, FCM)是对悬液中的单细胞或其它生物粒子,通过检测标记的荧光信号,实现高速、逐一的细胞定量分析和分选的技术。近年来,针对传统流式细胞术缺少细胞形态和内在结构等信息的缺点,发展了成像流式细胞术,许多都使用了人工智能图像识别技术。Göröcs 等^[25]先通过无透镜全息显微镜获得微通道内微小物体的衍射图像,再基于深度学习的相位恢复与图像重建方法实时重建全息衍射图像,从而在无需外部标记的条件下成功识别了目标,通量可达 100 mL/h。通过这种方法测量得到了洛杉矶 6 个公共海滩中潜在有毒藻类浓度,与公共卫生部门的测量结果一致。该设备质量仅为 1 kg,可通过笔记本电脑操作,体现了人工智能技术在降低硬件要求、缩小设备体积、节约制造成本方面的优势,为实现大型专用仪器的微型化提供了更多途径。与此同时,Nitta 等^[16]也以 CNN 为基础,开发了一种基于图像的实时智能细胞分类技术(图 1B),能够以每秒 100 个细胞的速度进行细胞分选,成功对克氏衣藻、薄足页藻、人血小板团簇、红细胞、白血球、H1975 腺癌细胞等进行了分选,纯度高达 99.1%。相比于已经发展数十年的传统流式细胞术,基于该方法的成像流式细胞术在分选速度、分选准确性等方面的性能达到了相当的水平,体现了人工智能在图像识别方面强大的分析能力,以及其革新传统微流控系统的潜力。

此外,动态目标的识别与实时监控也是人工智能目标检测技术的一个重要应用方向。Wang 等^[26]在利用 3D 微流控装置模拟血管生成网络中的细胞迁移时,使用 CNN 模型来检测候选区域,再将后向卡尔曼滤波与多个假设跟踪相结合,实现了多细胞迁移跟踪,准确度高达 86.4%。该工作体现了人工智能在实时目标检测方面的优势,为微小物体的监控与追踪提供了更多技术途径。Rizkin 等^[27]使用非接触红外热成像装置监控微流控系统的热扩散,并利用计算机视觉人工智能神经网络快速分析其热响应,提出了一种新的连续流化学反应器热控制方法(图 1C)。相比于使用传统的比例积分导数控制法,该方法的温度稳定时间从 144.00 s 减少到 11.43 s,最大偏差也降低到了 0.2℃左右。而且,此类监控装置结构简单,成本低廉,有望在化学工程实时监控与智能调控等领域广泛应用。

除了较为精确的图像识别与计数,人工智能在相对模糊的密度估算、状态评估、模式识别等方面也有较好的应用。Kim 等^[28]使用快速傅里叶变换(FFT)技术检测微流控系统中细菌生长显微图像频率的变化,再基于深度学习进行数据回归,成功对微流控系统内的细菌密度进行了估算(图 1D),并将其应用于微流控抗生素敏感性测试。该方法没有试图对模糊的细胞图形进行精准识别,而是根据图像中的微小变化进行粗略的密度估算,为解决微流控系统内的数据估算问题提出了一种新的思路。与之相似,Moore 等^[29]搭建了一种可构建肿瘤微环境的肿瘤活检微流控装置,并利用基于机器学习的自定义图像分析算法成功量化了肿瘤浸润淋巴细胞的死亡时间(图 1E)。该工作充分发挥了人工智能在状态评估方面的优势,为定量表征微流控系统内的待测物提供了更多可能。

总之,由于人工智能在提升数据质量、跟踪动态目标、评估模糊场景等方面的独特性能,使其在微流控目标检测中展现出了明显的应用潜力。

4 人工智能在微流控系统关联预测中的应用

在传统的科研范式中,寻找物质运动及其相互作用的规律,一直是需要科研人员依赖其科研技能去

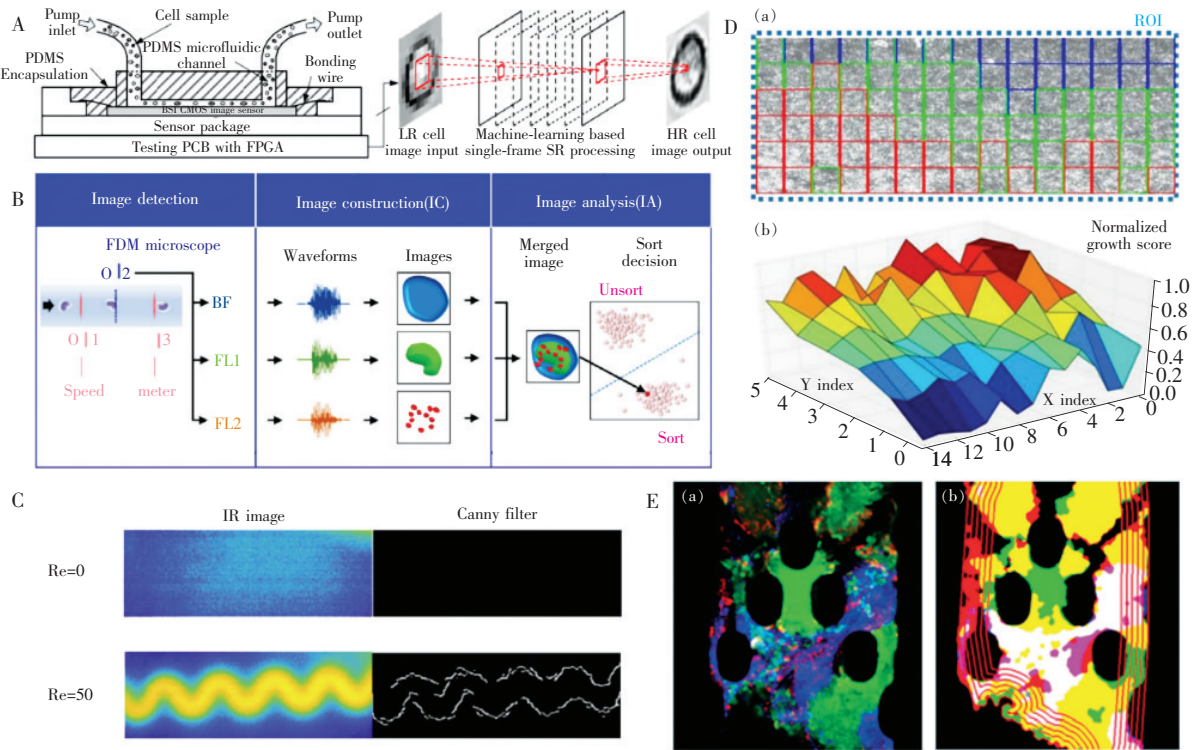


图 1 人工智能在微流控系统目标检测中的应用

(A) 微小图像超分辨率处理^[23]；(B) 成像流式细胞术^[16]；(C) 热成像动态监测^[27]；(D) 细菌密度估算^[28]；(E) 肿瘤渗透过程识别^[29]。

Fig. 1 Applications of artificial intelligence in target detection of microfluidic systems

(A) Super-resolution processing of tiny sample images^[23]；(B) Imaging flow cytometry^[16]；(C) Dynamic monitoring with thermal imaging^[27]；(D) Bacterial density estimation^[28]；(E) Identification of tumor infiltration process^[29]。

解决的问题。然而,随着人工智能技术的快速发展,发现重要变量、寻找不同变量之间的潜在关联、预测实验结果,也可以通过人工智能实现。

在微流控系统的数值模拟方面,人工智能已经有了许多应用。早在 2006 年, Magargle 等^[30]就利用 ANN 模型对芯片实验室系统中的注射器装置进行建模,通过基础质量传递偏微分方程的有限元模拟训练,将注射器行为映射到由系统物理变量参数化的一组性能分析函数中(图 2A),不仅计算速度比数值模拟快 4 个数量级,均方误差也可精确至 10^{-4} 。该工作首次将人工智能用于微流控装置建模,即展现出了远超传统数值模拟方法的运算效率,为此后人工智能在微流控系统深度应用提供了有益的指导。此外, Han 等^[31]利用递归神经网络模型对微流控柔性传感器进行表征(图 2B),在压力响应滞后的情况下,依然模拟出了柔性传感器的非线性特征,并估算了其所受压力刺激的大小与位置,其平均定位精度可达 85.42%,标准偏差为 5.81%。该研究展现了人工智能在寻找非线性关联特征方面的优势,有可能为微流控柔性传感器提供更多的表征与模拟工具,并简化机械与控制系统的结构。

除了数值模拟,人工智能与微流控系统结合,还能快速分析不同参数对实验结果的影响,并构建预测模型。Nyberg 等^[32]以微流控单细胞实验为基础,利用 k-近邻算法对 6 种细胞物理表型进行分析,发现了一组预测癌细胞入侵的 4 种物理表型(图 2C(a)),其组合后的最高准确度可达 96%,最后通过训练多元线性回归模型成功生成了癌细胞侵袭预测模型(图 2C(b))。该研究表明,人工智能有助于寻找影响事件发生的关键因素,并发现不同因素间的潜在联系,为医学研究模型、疾病诊断模型的开发提供了更多手段。此外, Yasemi 等^[33]在使用微流控系统进行鞣酸的提取实验时,利用 ANN 模型分析了多种实验参数对提取效率的影响,构建了一种鞣酸提取效率模型。根据该模型得到了鞣酸最佳提取条件,实际提取率为 $95.01\% \pm 0.63\%$,与预测的最大提取率 96.4% 基本一致。因此,人工智能与微流控系统的

结合,有效提高了实验参数优化的进程。另外,Huang 等^[34]使用 ANN 训练出了一种可以预测微通道特征系数的模型,用于研究微通道中的流体阻尼,特征系数预测值与实验测量值的平均偏差仅为 4.7%。该研究表明,人工智能可以预测微流控系统某些关键参数,对于更加精确的数值模拟与流体操控也有很大的帮助。

甚至对于很多无法利用数字精确描述的直观经验,人工智能技术也能进行描述。例如,Khor 等^[35]使用卷积自动编码器模型对微流体管道中的乳液液滴形状进行描述,精度可达 91.7%,再对其中包含的多维参数进行解释,最终找到了影响液滴破裂的 3 个主要因素,总结出了微流体管道中乳液形状与其后续变化之间的关系,成功预测了微流体通道中不同形状乳液液滴的稳定性(图 2D)。因此,人工智能在描述特征方面也有一定的优势,相关技术的应用对于归纳、解释微流控系统实验现象,加深对微流控系统的理解,具有重要的意义。

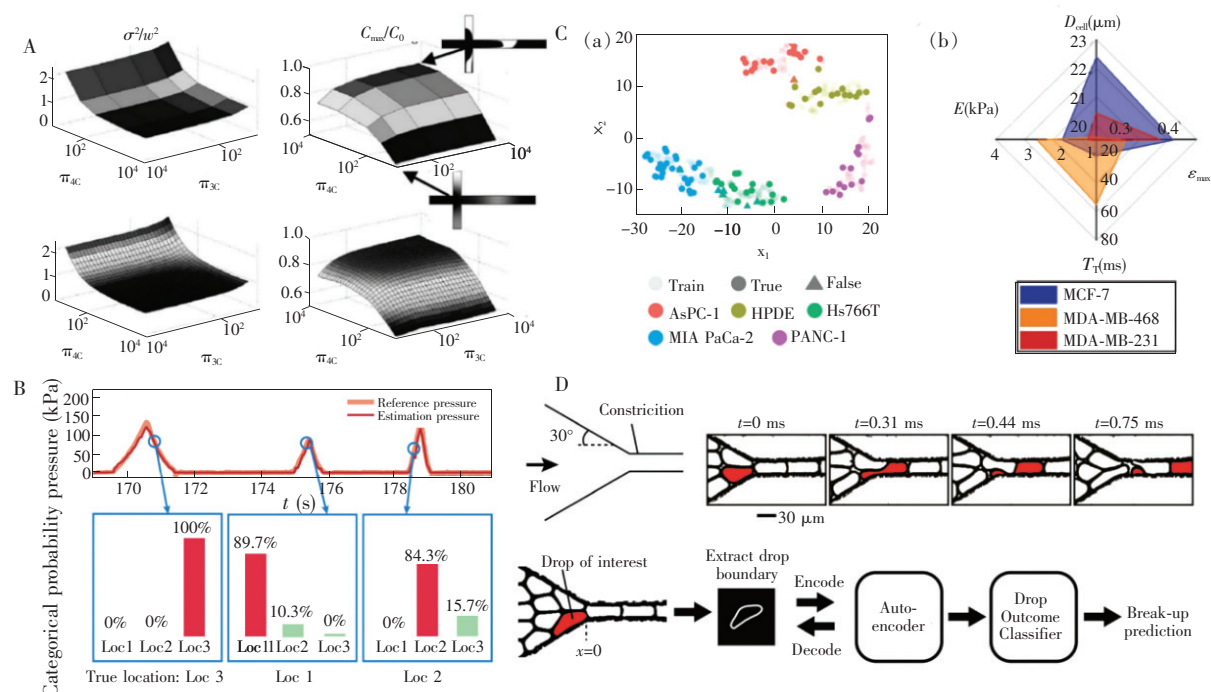


图 2 人工智能在微流控系统关联预测中的应用

(A) 高分辨率数值模拟^[30]; (B) 柔性传感器表征与定位^[31]; (C) 物理表型分析与癌细胞侵袭预测模型构建^[32]; (D) 乳液液滴稳定性参数发现^[35]。

Fig. 2 Applications of artificial intelligence in association prediction of microfluidic systems

(A) High-resolution numerical simulation^[30]; (B) Characterization and localization of soft sensors^[31]; (C) Physical phenotype analysis and cancer cell invasion prediction model construction^[32]; (D) Discovery of emulsion droplet stability parameters^[35].

所以,人工智能在发现关键参数、预测多因素关联、描述特征、总结归纳变化规律等方面具有优势,其在微流控系统中的应用将推动整体系统的智能化发展。

5 人工智能在微流控系统结果分类中的应用

微流控系统由于其自动化、高通量的特点,在快速完成大量相似实验的同时,也会得到海量的实验结果。对这些结果进行分析与分类,并加以差异化处理,已经成为对微流控数据分析与挖掘的基本要求。得益于人工智能优秀的分类与聚类能力,许多实用的微流控数据处理技术得以发展。

首先,恰当的分类可以有效区分不同的被测物状态,如 Ellett 等^[8]使用微流控系统从一滴稀释血液测量中性粒细胞的自发运动性,并通过基于机器学习的评分系统成功区分了败血症患者与其他患者的血样。对 42 名患者中的败血症患者识别敏感率为 97%,特异性为 98%。这种可以区分疾病患者的

分类模型已经起到了类似疾病诊断模型的作用。类似的,Manak 等^[36]将前列腺癌和乳腺癌患者的组织样本在微流控装置上成像,再利用随机森林模型进行分析,对癌症病人的风险情况进行分层,实现了疾病的分级(图 3A)。此外,Khan 等^[37]使用基于线性判别分析的分类器,为实时诊断微流控装置开发了一种设备健康状态诊断与预测系统,用于监控设备状态,准确度>98%。因此,利用人工智能对微流控系统实验结果进行恰当分类的意义,不亚于设计一个专门的研究模型,这也体现了人工智能在数据挖掘方面的能力。

其次,高效的分类算法有助于开发便携式智能检测设备。Potluri 等^[7]开发了一种基于手机的低成本排卵检测设备(图 3B),重量仅 208 g,可以利用微流控芯片对唾液进行处理并产生特定纹路,再由手机上基于 MobileNet 模型(一种 CNN 模型)开发的 APP 进行分析,即可判断妇女的排卵能力(准确率>99%),帮助调控生育计划。该方法将微流控芯片与手机相结合,体现了微流控技术和人工智能技术在家用检测方面的潜力。此外,在可穿戴设备开发方面,Kim 等^[38]使用两个微流控柔性传感器检测人类运动过程中的步态,再利用半监督深度学习模型完成不同运动状态下的步态分类(图 3C),从而以较小的数据量实现了设备校准。该研究说明人工智能有助于简化可穿戴设备结构,降低设备成本,对于相关设备的开发与推广,具有重要意义。

再次,依据结果进行的分类也有助于识别某些稀有或难以标注的目标。如 Guo 等^[39]使用基于固态微孔的微流控阻抗细胞计数器,对 600 个红细胞与癌细胞进行检测,再通过 SVM 模型分析信号脉冲中的两个重要参数,测得癌细胞含量为(38.2%±2.0%),与商品化流式细胞仪测得的结果(36.4%±0.5%)基本一致,为癌细胞的识别提供了一种相对简单的方法。Wang 等^[40]利用微量移液器吸取肿瘤细胞的时间表征单个细胞的细胞质粘度,又根据细胞质粘度,将基于神经网络的模式识别用于不同细胞的分类,最高取得了 76.7% 的分类成功率。该工作体现了人工智能在发掘单个样品特征的应用价值方面,具有突出的能力。Jagannadh 等^[41]利用 PCA、SVM 模型,对显微镜观察的微流控系统 中的癌细胞进

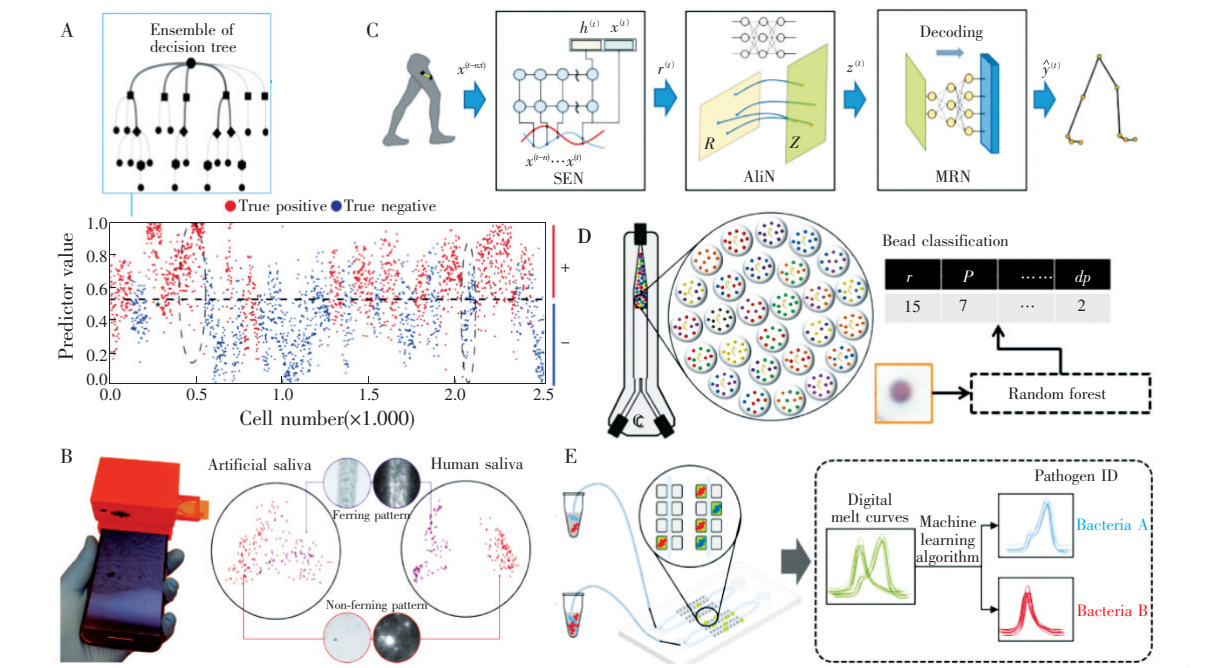


图 3 人工智能在微流控系统结果分类中的应用

(A) 癌症病人风险分层^[36]; (B) 基于手机的排卵检测设备^[7]; (C) 步态分类^[38]; (D) 大规模液滴编码^[42]; (E) 细菌鉴定与药物敏感性测试^[43]。

Fig. 3 Applications of artificial intelligence in results classification of microfluidic systems

(A) Risk stratification of cancer patients^[36]; (B) Ovulation detection equipment based on mobile phones^[7]; (C) Gait classification^[38]; (D) Large-scale droplet coding^[42]; (E) Bacterial identification and drug sensitivity testing^[43]

行了无标记分类,降低了癌细胞识别难度,准确率>97%。这些方法都为稀有或特征不明显目标的识别与标注提供了更多的解决方案。

最后,强大的目标分类技术也为样品的大规模分析提供了一种新的解决办法。Svensson 等^[42]使用彩色微球对微流控液滴进行编码后,利用随机森林分类器和贝叶斯分类器进行解码,成功对 20 种实验条件、每种条件超过 100000 个液滴的大规模液滴群进行了编码与解码(图 3D),解码准确率高达 99.6%,展现了大规模微流控液滴作为细胞反应器的应用潜力。Athamanolap 等^[43]在微流控阵列平台中使用数字 PCR 测定方法,在 4 h 内完成了通常需要数天才能实现的快速细菌鉴定与抗菌药物敏感性测试(图 3E),采用机器学习算法,依据基因熔解曲线自动识别了细菌种类,大大加快了识别速度。因此,人工智能在海量数据处理方面的突出性能,有助于突破微流控大规模样品分析方面的技术瓶颈,推动更多先进设备的开发。

由此可知,人工智能对微流控系统结果的高效分类,在医学模型构建、设备状态评估、便携式智能检测设备开发、特殊目标识别、大规模样品分析等方面,都具有十分重要的意义。

6 结论与展望

微流控系统与人工智能分别在大数据生成、采集与数据处理方面的优势,使得二者的结合成为了一种必然的趋势。首先,人工智能在大规模数据处理方面的优势,使其可对微流控系统自动化、高通量的分析流程进行快速实时的分析与判断。其次,人工智能高通用性的信号检测与目标识别能力,可以减少对特定检测设备的依赖,降低硬件要求,缩小设备体积,节约检测成本,推动微流控系统在微型化、集成化方面的发展。再次,人工智能可以对微流控系统产生的海量数据进行深入挖掘,辅助研究者发现数据间的潜在关联,归纳总结物性规律,构建系统模拟与预测模型。最后,人工智能可以实时监控系统运行状态,综合评估实验结果,并智能调节其中关键步骤,充分发挥微流控系统的流体控制能力,推动其由自动化向智能化发展,以应对更加复杂的实验环境与更加广泛的应用需求。

基于当前的应用趋势,微流控领域的研究人员对人工智能技术的理解正在逐渐加深,将理论运用于实际情境的工程能力也越来越强。一方面,从简单的回归模型到相对复杂的统计模型,再到复杂度较高的深度学习模型,人工智能在微流控系统中所应用的模型复杂度逐渐提升;另一方面,从单一人工智能模型到多种模型联用,从标准模型到各种优化模型,所用模型的灵活性也在逐步提高。这既得益于人工智能技术的逐步普及,也体现了微流控研究者对人工智能的关注与迫切需要。随着 Tensorflow 等学习框架、分布式计算,人工智能基础教育的推广,学习并使用人工智能技术的门槛将进一步降低,人工智能将成为微流控系统中如同数学、化学、物理学和计算机一般的基础技术。

虽然,人工智能在微流控系统中已取得了许多重要的应用成果,但仍然面临诸多挑战。例如,CNN 等常用的监督学习算法通常需要使用大量标记数据进行训练,而标记数据比较费时费力,因此无标签、半监督的高效模型可能会成为未来的研究热点。此外,在面对复杂的研究体系时,采用降维学习等相对简单的模型,可能尚不足以深入探究这些体系,而高度复杂的模型又带来模型构建与学习效率等方面的困难。因此,为了获得更好的学习性能,还需进一步发展深度学习、强化学习等高复杂度模型,并开发更多针对复杂体系的研究策略。另外,由于开发难度的限制,目前人工智能主要用于对微流控系统的数据处理,但是其在状态评估、系统决策、智能优化等方面仍然具有很大的应用潜力。

总之,人工智能在微流控系统中的应用展现出蓬勃的发展态势,在目标检测、关联预测、结果分类等方面展现了明显的优势与良好的实用效果。随着技术的快速进步,人工智能技术必将逐步覆盖微流控系统中数据处理、状态评估、智能决策、自动优化等方面,大大提高微流控系统的智能化水平,并最终在化学、生物学、医学、药学、材料科学等领域取得更加广泛且重要的应用,甚至有望在当前依赖专家经验、人工尝试的科研模式之外,发展出利用人工智能参与实验设计、操作和结果处理的新的科研模式。

References

- 1 Qasaimeh M A, Ricoult S G, Juncker D. *Lab Chip*, **2013**, 13(1): 40–50

- 2 Riordon J, Sovilj D, Sanner S, Sinton D, Young E W K. *Trends Biotechnol.* , **2019**, 37(3): 310–324
- 3 Tang J, Ibrahim M, Chakrabarty K, Karri R. *IEEE Trans. Comput-Aided Des. Integr. Circuits Syst.* , **2019**, 38(4): 589–603
- 4 Wu Q, Kumar N, Velagala V, Zartman J J. *J. Biol. Eng.* , **2019**, 13(1): 16–33
- 5 Lee W, Gonzalez A, Arguelles P, Guevara R, Gonzalez-Guerrero M J, Gomez F A. *Electrophoresis* , **2018**, 39(12): 1443–1451
- 6 Hassoun M, Rüger J, Kirchberger-Tolstik T, Schie I W, Henkel T, Weber K, Cialla-May D, Krafft C, Popp J. *Anal. Bioanal. Chem.* , **2018**, 410(3): 999–1006
- 7 Potluri V, Kathiresan P S, Kandula H, Thirumalaraju P, Kanakasabapathy M K, Kota Sai Pavan S, Yarravarapu D, Soundararajan A, Baskar K, Gupta R, Gudipati N, C. Petrozza J, Shafiee H. *Lab Chip* , **2019**, 19(1): 59–67
- 8 Ellett F, Jorgensen J, Marand A L, Liu Y M, Martinez M M, Sein V, Butler K L, Lee J, Irimia D. *Nat. Biomed. Eng.* , **2018**, 2(4): 207–214
- 9 Schneider G. *Nat. Rev. Drug Discov.* , **2017**, 17(2): 97
- 10 Thomford N, Senthebane D, Rowe A, Munro D, Seele P, Maroyi A, Dzobo K. *Int. J. Mol. Sci.* , **2018**, 19(6): 1578
- 11 Neethirajan S, Kobayashi I, Nakajima M, Wu D, Nandagopal S, Lin F. *Lab Chip* , **2011**, 11(9): 1574–1586
- 12 Marle L, Greenway G M. *TrAC-Trends Anal. Chem.* , **2005**, 24(9): 795–802
- 13 Samuel A L. *IBM J. Res. Dev.* , **1959**, 3(3): 211
- 14 Krizhevsky A, Sutskever I, Hinton G E. *Adv. Neural Inform. Process. Syst.* , **2012**: 1097–1105
- 15 Rampasek L, Goldenberg A. *Cell Syst.* , **2016**, 2(1): 12–14
- 16 Nitta N, Sugimura T, Isozaki A, Mikami H, Hiraki K, Sakuma S, Iino T, Arai F, Endo T, Fujiwaki Y, Fukuzawa H, Hase M, Hayakawa T, Hiramatsu K, Hoshino Y, Inaba M, Ito T, Karakawa H, Kasai Y, Koizumi K, Lee S, Lei C, Li M, Maeno T, Matsusaka S, Murakami D, Nakagawa A, Oguchi Y, Oikawa M, Ota T, Shiba K, Shintaku H, Shirasaki Y, Suga K, Suzuki Y, Suzuki N, Tanaka Y, Tezuka H, Toyokawa C, Yalikun Y, Yamada M, Yamagishi M, Yamano T, Yasumoto A, Yatomi Y, Yazawa M, Di Carlo D, Hosokawa Y, Uemura S, Ozeki Y, Goda K. *Cell* , **2018**, 175(1): 266–276
- 17 Sutskever I, Vinyals O, Le Q V. *Adv. Neural Inform. Process. Syst.* , **2012**: 3104–3112
- 18 Bahdanau D, Chorowski J, Serdyuk D, Brakel P, Bengio Y. *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* , **2016**: 4945–4949
- 19 Wu S, Ren W, Yu C, Chen G, Zhang D, Zhu J. *Proc. IEEE Int. Conf. Data Eng.* , **2016**: 1218–1229
- 20 Nemati S, Ghassemi M M, Clifford G D. *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.* , **2016**: 2978–2981
- 21 Abadi M. *ACM Sigplan Not.* , **2016**, 51(9): 1
- 22 Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. *Proc. 22nd ACM Int. Conf. Inf. Multimedia* , **2014**: 675–678
- 23 Huang X, Jiang Y, Liu X, Xu H, Han Z, Rong H, Yang H, Yan M, Yu H. *Sensors* , **2016**, 16(11): 1836
- 24 Liu X, Huang X, Jiang Y, Xu H, Guo J, Hou H W, Yan M, Yu H. *IEEE Trans. Biomed. Circuits Syst.* , **2017**, 11(4): 794–803
- 25 Göröcs Z, Tamamitsu M, Bianco V, Wolf P, Roy S, Shindo K, Yanny K, Wu Y, Koydemir H C, Rivenson Y, Ozcan A. *Light Sci. Appl.* , **2018**, 7(1): 66
- 26 Wang M, Ong L S, Dauwels J, Asada H H. *J. Med. Imag.* , **2018**, 5(2): 024005
- 27 Rizkin B A, Popovich K, Hartman R L. *Comput. Chem. Eng.* , **2019**, 121: 584–593
- 28 Kim K, Kim S, Jeon J. *Sensors* , **2018**, 18(2): 447
- 29 Moore N, Doty D, Zielstorff M, Kariv I, Moy L Y, Gimbel A, Chevillet J R, Lowry N, Santos J, Mott V, Kratchman L, Lau T, Addona G, Chen H, Borenstein J T. *Lab Chip* , **2018**, 18(13): 1844–1858
- 30 Magargle R, Hoburg J F, Mukherjee T. *IEEE Trans. Comput-Aided Des. Integr. Circuits Syst.* , **2006**, 25(2): 378–385
- 31 Han S, Kim T, Kim D, Park Y, Jo S. *IEEE Robot. Autom. Lett.* , **2018**, 3(2): 873–880
- 32 Nyberg K D, Bruce S L, Nguyen A V, Chan C K, Gill N K, Kim T, Sloan E K, Rowat A C. *Integr. Biol.* , **2018**, 10(4): 218–231
- 33 Yasemi M, Rahimi M, Heydarinasab A, Ardjmand M. *Chem. Prod. Process Model.* , **2017**, 12(1): 20160053

- 34 Huang L, Nie W, Wang X, Shen T. *Microsyst. Technol.*, **2017**, 23(6): 2297–2305
- 35 Khor J W, Jean N, Luxenberg E S, Ermon S, Tang S K Y. *Soft Mat.*, **2019**, 15(6): 1361–1372
- 36 Manak M S, Varsanik J S, Hogan B J, Whitfield M J, Su W R, Joshi N, Steinke N, Min A, Berger D, Saphirstein R J, Dixit G, Meyyappan T, Chu H, Knopf K B, Albala D M, Sant G R, Chander A C. *Nat. Biomed. Eng.*, **2018**, 2(10): 761–772
- 37 Khan H, Al-Gayem Q, Richardson A M. *IEEE Trans. Device Mater. Rel.*, **2017**, 17(2): 438–449
- 38 Kim D, Kim M, Kwon J, Park Y, Jo S. *IEEE Robot. Autom. Lett.*, **2019**, 4(3): 2501–2507
- 39 Guo J, Chen Z, Ban Y, Kang Y. *IEEE Trans. Emerg. Topics Comput.*, **2017**, 5(4): 518–525
- 40 Wang K, Sun X H, Zhang Y, Zhang T, Zheng Y, Wei Y C, Zhao P, Chen D Y, Wu H A, Wang W H, Long R, Wang J B, Chen J. *Rev. Soc. Open Sci.*, **2019**, 6(3): 181707
- 41 Jagannadh V K, Gopakumar G, Subrahmanyam G R K S, Gorthi S S. *Med. Biol. Eng. Comput.*, **2017**, 55(5): 711–718
- 42 Svensson C M, Shvydkiv O, Dietrich S, Mahler L, Weber T, Choudhary M, Tovar M, Figge M T, Roth M. *Small*, **2018**, 15(4): 1802384
- 43 Athamanolap P, Hsieh K, Wang A T. *40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, **2018**: 5346

Application of Artificial Intelligence in Microfluidic Systems

WANG Yu, FANG Qun*

(*Institute of Microanalytical Systems, Department of Chemistry, Zhejiang University, Hangzhou 310058, China*)

Abstract Microfluidic systems are widely applied in many fields including chemistry, biology, medicine, and pharmacy, because of their precise control ability to microfluids. In recent years, artificial intelligence technology has achieved leap-forward development with great advantages in dealing with the analysis and mining of massive data. The application of artificial intelligence technology in microfluidic systems has shown great potential in many fields such as biological research, medical diagnosis, and drug discovery, and so on. This paper reviews several typical artificial intelligence models and their applications in microfluidic systems. It focuses on the progress of artificial intelligence in target detection, correlation prediction and result classification of microfluidic systems, and forecasts the future development trend based on its application status.

Keywords Microfluidics; Artificial intelligence; Data mining; Big data; Review

(Received 18 November 2019; accepted 16 January 2020)

This work was supported by the National Natural Science Foundation of China (Nos. 21827806, 21435004).