

机器学习在蚊虫及蚊媒传染病研究中的应用进展

孙燕群^{1,2}, 张守刚¹, 赵姗姗¹, 陆墨原¹, 张艳¹, 王冲¹, 李成国¹

1. 南京市疾病预防控制中心消毒与病媒生物防制科, 南京医科大学附属南京疾病预防控制中心, 江苏 南京 210003; 2. 军事科学院军事医学研究院微生物流行病学研究所, 病原微生物生物安全国家重点实验室, 北京 100071

摘要: 该文主要介绍机器学习在全球蚊虫及蚊媒传染病研究中的应用进展, 系统搜索国内外数据库进行文献调研, 简单回顾了机器学习的主要方法, 对机器学习在蚊虫及蚊媒传染病研究中的几大应用进行了系统总结, 主要聚集在蚊虫和蚊媒传染病预测预警、蚊虫图像和声音识别、蚊虫生物学等研究领域, 为国内蚊虫及蚊媒传染病防制研究提供新的视角。

关键词: 蚊媒; 蚊媒传染病; 机器学习; 研究进展

中图分类号: R181.2³ **文献标识码:** A **文章编号:** 1003-8280(2021)04-0503-06

DOI: 10.11853/j.issn.1003.8280.2021.04.024

Application progress of machine learning in mosquito and mosquito-borne disease research

SUN Yan-qun^{1,2}, ZHANG Shou-gang¹, ZHAO Shan-shan¹, LU Mo-yuan¹, ZHANG Yan¹, WANG Chong¹, LI Cheng-guo¹

1. Department of Disinfection and Vector Control, Affiliated Nanjing Center for Disease Control and Prevention, Nanjing Center for Disease Control and Prevention, Nanjing Medical University, Nanjing, Jiangsu 210003, China; 2. State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China
Supported by the Nanjing Medical Science and Technology Development Fund (No. YKK17200, YKK18178) and the Tenth Cycle of Nanjing's Key Medical Specialty (Prevention and Control of Infectious Diseases)

Abstract: This article mainly introduced the application progress of machine learning in global mosquito and mosquito-borne disease research, systematically searched domestic and foreign databases for literature research, briefly reviewed the main methods of machine learning, and systematically summarized several major applications of machine learning in mosquito and mosquito-borne disease research. It is mainly concentrated in the research fields of mosquito and mosquito-borne disease prediction and early warning, mosquito image and sound recognition, and mosquito biology, providing a new perspective for domestic mosquito and mosquito-borne disease prevention and control.

Key words: Mosquito; Mosquito-borne disease; Machine learning; Research progress

近年来人工智能和机器学习呈现暴发式发展势头, 在诸多领域取得重要研究成果, 在医学和公共卫生领域的应用也逐渐增多。媒介蚊虫可引起重要的公共卫生问题, 它们不仅叮咬骚扰人群, 而且可以传播疟疾、登革热、流行性乙型脑炎(乙脑)、丝虫病、西尼罗河病、基孔肯雅热等蚊媒传染病, 近年来新发和重现的蚊媒传染病成为影响全球公共健康重大威胁, 蚊虫及蚊媒传染病研究成为目前公共卫生领域较热门的研究内容之一^[1-7], 涉及昆虫生态学、传染病流行病学、病原学、数理统计学、地理信息科学、气象学、仿生学等多领域, 这种学科之间的交叉融合在发展和应用机器学习技术上具有一定的优势。本文总结了机器学习的定义、算法等基础概念以及在蚊虫及蚊媒传染病研究中的应用, 概括了机器学习在蚊虫及蚊媒传染病研究领域的应用现状, 针对机器学习提出了在蚊虫及蚊媒传染病研究领域深入挖掘的设想。

1 机器学习模型方法概述

1.1 机器学习(machine learning)简介 机器学习是人工智能领域重要的组成, 也是实现人工智能的一个重要途径, 它是计算机利用已有的大规模数据, 分析推出某种模型, 并用于预测的一种方法^[8]。

1.2 机器学习分类 机器学习的数据处理主要分为无监督学习和监督学习 2 种^[9]。无监督学习: 处理不具有标签的数据, 利用计算机本身的自动化功能, 发现数据中的隐藏模式和规律, 完成学习过程。无监督学习不需要提前进行模型训练, 机器有选择地自行学习探索, 直接对样本数据进行建模分析。无监督学习多用于聚类问题的解决。监督学习: 借助有标签的数据进行学习辅助, 从而达到学习目标。具体学习过程是使用有标签的数据向算法提供示例进行模型训练产生期望输出, 并将训练好的模型用来测试新的数据产生的

基金项目: 南京市卫生科技发展专项资金(YKK17200, YKK18178); 南京市第十周期医学重点专科(传染病预防控制)

作者简介: 孙燕群, 男, 主管医师, 主要从事蚊虫及蚊媒病监测与防制工作, E-mail: sunyq@njcdc.cn

结果。在该学习方式下,初始标签数据是不可或缺的要素,进行模型训练前必须搜集。监督学习可充分发挥机器学习的泛化能力,对规则以及规律数据等进行预测。监督学习多用于处理分类问题和回归问题。

2 机器学习经典研究方法

在机器学习使用的算法种类较多,其中典型的算法有以下几种:人工神经网络、决策树、随机森林、支持向量机、深度学习等^[10-12]。

2.1 人工神经网络(artificial neural network, ANN) ANN 是一种能够识别输入数据与输出数据集之间复杂非线性关系灵活的机器学习模型,ANN 具有结构简单、能够大规模运行、容易用硬件实现,并且具有学习和记忆能力、自适应和多样性等特性,优良特性使得其显示出强有力的生命力。ANN 的概念由来已久,起源于对生物神经网络的模拟研究,随着 20 世纪 40 年代以来人们对人体神经系统的了解越深入,研究者们基于 ANN 的工作原理,综合数学、物理以及信息处理等科学方法对人脑神经网络进行抽象,并建立简单的模型,称之为 ANN。目前,ANN 作为知名度最广的机器学习方法,已经在多个领域得到相当广泛的应用,包括流行病学、疾病诊断、疾病预后等^[10,12]。

2.2 决策树(decision tree) 决策树是根据属性值来进行排序并且进行分组的树类型,主要用于分类。每棵树都由节点和分支组成,每个节点表示要分类的组的属性,每个分支表示节点可以采用的值。决策树运算中分 2 个阶段:构造决策树以及以递归方式修剪树枝。当无法继续进行数据分割时,递归便会随之结束。决策树常用来解决分类和回归问题^[10,12]。

2.3 随机森林(random forest, RF) RF 是一种典型的 Bagging 模型,RF 算法是基于决策树的算法,将若干数量的决策树作为弱学习器,随机重复抽取样本来训练每一棵树,并且用剩余的样本数据对所训练的树进行评估,从而获得若干数量的决策树形成的森林,当有需要预测的新样本进入随机森林分类器时,所有决策树便会通过投票的方式得出最终的结果预测。RF 目前也已经在多个领域得到较为成熟的应用,有多个研究利用其进行空间分布的预测^[10,12]。

2.4 支持向量机(support vector machine, SVM) SVM 是最近广泛使用的一种机器学习技术,按照边距计算的原理,在 2 个类别之间创建一个决策边界,使边距与类别之间的距离最大,从而使分类的误差最小。它在统计分类和回归分析中应用较广泛,可以解决文本分类和图像识别等问题^[10,12]。

2.5 深度学习 深度学习是机器学习领域中一个新的研究方向,与传统的浅层机器学习不同,深度学习的神经网络含有很多隐藏层,机器能够自动学习数据各个层次的特征,充分挖掘数据的信息并对其进行分析处理,包括深度神经网络、卷积神经网络(CNN)、递归神经网络和新兴架构(深度时空神经网络 DST-NNs、多维递归神经网络 MD-RNNs 和卷积自动编码器 CAEs)4 类^[10-11]。

3 机器学习在蚊虫研究中的应用

蚊虫研究一般涉及蚊虫的生态学、昆虫学、生物学等领域,机器学习可利用传统的蚊虫研究数据进行深入挖掘,进行模型的预测和拟合。

3.1 蚊虫分布与密度预测预警 传统的蚊虫分布与密度研究主要通过人工或自动的蚊虫监测点开展蚊虫数据收集工作,如诱蚊灯法(普通诱蚊灯、二氧化碳诱蚊灯)、布雷图指数法、BG-trap 法、路径法等^[13]。此类方法主要观察蚊虫是否存在以及蚊虫分类的量化分级,用以评估蚊虫侵害的风险。蚊虫的分布与否以及密度受到多种外在因素的影响,如气象、土地利用方式、植被、水体类型等线性或其他非线性因素,用传统的统计分析方法存在共线性问题以及忽略了某些因素(如气象因素)对蚊虫的孳生繁殖的滞后作用,因此可以运用机器学习方法对蚊虫的分布与密度进行分析预测。

3.1.1 伊蚊分布与密度预测 伊蚊是众多蚊媒传染病的媒介,如登革热、寨卡病毒病、黄热病等,其中白纹伊蚊(*Aedes albopictus*)是半家栖蚊种,除了能孳生在室内外,多数孳生在人类居住区或活动区及其周围;埃及伊蚊(*Ae. aegypti*)是典型的家栖蚊种,主要孳生在室内外及其周围。伊蚊对健康的影响较大,对其分布及密度进行探索预测,有助于进一步开展伊蚊防治。

肖冰^[14]利用决策树预测城市水绿复合系统幼蚊分布及密度,其中分类树预测不同变量下幼蚊是否孳生,回归树预测幼蚊在不同分类下的密度分布;黄建华等^[15]利用伊蚊诱捕器监测法得出指标与气象资料建立多变量灰色预测模型[multivariable grey model, MGM(1, n)],对伊蚊密度进行短期预测;李卫红等^[16]基于遗传算法-后向传播(GA-BP)神经网络模型的登革热时空扩散模拟;周毅彬等^[17]利用多元逐步回归和神经网络进行白纹伊蚊密度与气象因素关系的研究;于德宪等^[18]运用 ANN 模型用于分析气候因素对白纹伊蚊密度的影响;Früh 等^[19]利用 4 种机器学习方法,分别是决策树、Logistic 回归、RF、SVM 对德国侵入伊蚊进行研究,结果显示包含 2 个或 3 个模型的拟合效果要优于单个模型;Kerkow 等^[20]基于 SVM 模拟日本伊蚊(*Ae. japonicus*)在德国入侵的潜在分布;Demertzis 等^[21]提出了一种集成的机器学习(ML)模型在温度和降雨条件下来预测白纹伊蚊和日本伊蚊栖息地的适宜性;Zheng 等^[22]运用回归树模型根据现有白纹伊蚊的监测数据、气候和环境数据预测白纹伊蚊潜在的季节分布范围,平均准确率为 98%,平均受试者工作曲线下面积(AUC)为 99%;Scavuzzo 等^[23]基于多种机器学习技术(SVM、ANN、K 近邻、决策树回归)运用遥感数据对阿根廷北部地区埃及伊蚊的产卵获得进行时间建模,结果发现机器优于线性模型,尤其是最近邻回归(KNNR)性能最好;Ding 等^[24]利用 3 种机器学习模型(SVM、梯度提升机、RF)对全球埃及伊蚊和白纹伊蚊进行模拟,RF 的拟合效果最好。

3.1.2 其他蚊种分布与密度预测 裘炯良等^[25]利用 ANN 在入境航空器外来医学媒介生物输入开展风险评估;Wieland

等^[26]利用 SVM 对蚊虫分布与气候数据进行建模,并对蚊虫分布进行预测。

3.2 蚊虫图像声音识别 机器学习在图像声音文本识别领域具有天然的优势,包括蚊虫在内的各种昆虫具有丰富多彩的图像数据和声音数据,因此可以通过各种机器学习方法对蚊虫进行分类鉴定或者蚊媒感染图像进行识别鉴定,主要通过分类进化树、神经网络、SVM 等分类机器学习进行研究。

3.2.1 蚊虫图像识别 我国台湾省洪铭鸿^[27]基于边缘计算和深度学习开展病媒蚊虫的分类鉴定研究; Sanchez-Ortiz 等^[28]开展了基于卷积神经网络的幼蚊分类方法,其中使用幼蚊数据集进行训练,以便机器学习 2 种类型的蚊虫,即伊蚊属和“其他”属; Amarasinghe 等^[29]用定向梯度直方图(HOG)算法通过无人机图像识别蚊虫繁殖区域的明显着色来预测蚊虫繁殖区域; Park 等^[30]基于深度卷积神经网络(DCNNs)对蚊虫的视觉特征进行分类,准确率达 97% 以上; Motta 等^[31]基于 CNN 中的 3 种网络(LeNet、AlexNet、GoogleNet)对蚊虫进行分类,其中 GoogleNet 获得了最佳结果,对伊蚊和库蚊的分类准确率分别达到了 100% 和 90%; Case 等^[32]利用 CNN 训练图像对美国东北部地区无人机图像开展白纹伊蚊孳生地识别预测,准确率高达 67%; Genoud 等^[33]比较了有监督的机器学习算法(如线性判别分析、决策树、SVM、K 近邻和朴素贝叶斯)通过蚊虫的翅拍频率和光学截面来识别美国新泽西州存在的蚊虫种类、性别等。

3.2.2 蚊虫声音识别 李振宇^[34]开展了基于翅振频率的 CNN 方法在蚊虫分类中的应用研究,以翅振波形的频率分量作为特征向量训练的后向传播算法(BP)CNN 有较好的识别率。

3.3 蚊虫生物学 罗嘉鹏^[35]利用 SVM 技术研究组学数据检测昆虫的抗药性和入侵性; 宋帅葆等^[36]利用 RF 进行冈比亚按蚊(*Anopheles gambiae*)犬尿氨酸甲酰胺酶抑制剂的虚拟筛选; 王婷婷等^[37]用最大似然法、贝叶斯法和最大简约法探讨 4 种重要蚊虫离子受体基因的氨基酸序列的系统发育关系; 毛启萌等^[38]用 ML 法和贝叶斯法构建林氏按蚊(*An. lindesayi*)和按蚊属其他 32 种蚊虫的系统发育树; Smith 等^[39]运用贝叶斯区间映射和 RF 研究了 6 个与盐水耐受性有关的基因组; Saxena 和 Mishra^[40]用 CNN 对基孔肯雅热的表位进行预测和分析,用于疫苗的筛选和设计。

4 机器学习在蚊媒传染病研究的应用

蚊媒传染病预测预警主要涉及重要的蚊媒传染病,例如登革热、疟疾以及其他蚊媒传染病(基孔肯雅热、丝虫病、西尼罗河病等)。与蚊虫预测类似的是,对于伊蚊所传播的各种蚊媒传染病预测较多,可能与伊蚊传播的新发及再发传染病较多,其次伊蚊传播的虫媒传染病波及范围广,影响较大,伊蚊生存的环境也与人居环境高度相关。因此机器学习领域在伊蚊传播蚊媒传染病研究中着墨较多。

蚊媒传染病的发生发展与流行主要包括自然因素和社会因素影响,应用机器学习主要研究自然因素对蚊媒传染病的影响,与拟合蚊虫不同的是,对于蚊媒传染病的研究中机器学习同时也要考虑人的影响,因此有些研究将人口社会学

因素纳入影响研究范围。

4.1 登革热预测预警 黄宇琳等^[41]利用 RF 回归模型构建小空间尺度的登革热风险评估工具,模型拟合度较好;任红艳等^[42]利用反向传播神经网络模型对广东省登革热疫情进行预测;赵永谦^[43]运用 RF 回归模型建立珠江三角地区精细空间尺度的登革热风险评估模型;陈斌^[44]在评估广东省登革热防控能力及社区干预实验中利用 BP-ANN 模型筛选登革热易感者的特征;宋晓晴^[45]基于深圳市登革热本地染病病例数据,运用 RF 算法构建本地染病风险模型;陈业滨和李卫红^[46]进行基于向量机模型的登革热时空扩散预测研究;Zhao 等^[47]分别运用 RF 和 ANN 模型对哥伦比亚登革热病例进行预测,RF 模型预测误差小于 ANN,预测中环境和气象因子对短期预测范围相对重要,而人口社会学预测因子对长期预测范围相对重要;Kesorn 等^[48]基于 SVM 利用埃及伊蚊的登革热感染率对登革热发病率进行预测,预测精度为 88%;Ho 等^[49]基于 3 种预测模型(决策树、深层神经网络、逻辑回归)用 4 个输入变量(年龄、体温、白细胞计数和血小板)对登革热进行诊断预测,预测模型效果较好,AUC 分别为 85%、86% 和 84%。

4.2 疟疾预测预警 温亮^[50]分析气象因素与海南省万宁市疟疾发病率的相关性时发现,利用气象因素构建的 BP 神经网络模型较逐步回归模型具有更好的发病率拟合效果;郭华^[51]将机器学习方法、经济学模型以及生物统计学模型相结合,提出了面向主动监控的疟疾感染风险分布建模与挖掘方法;高春玉^[52]开展了疟疾流行现状及气象因素对疟疾发病影响的神经网络模型研究,其预测符合率达 80% 以上。

4.3 其他蚊媒传染病预测预警 高文等^[53]利用贝叶斯判别分析方法研究乙脑发生与蚊密度时间序列预测模型的关系;Vidal 等^[54]使用潜在类聚类分析(LCCA)对哥伦比亚加勒比地区基孔肯雅热病例症候学和表型反应之间的关系,总体准确率高达 92%;Eneanya 等^[55]使用泛化增强模型和 RF 对尼日利亚淋巴丝虫病进行了传播风险预测。

4.4 蚊媒传染病病原体预测 朱远林^[56]利用深度学习技术,在薄片图像中的疟原虫感染细胞检测运用视觉几何神经网络和 SVM 结合(VGG+SVM),在厚血片图像中的疟原虫感染细胞检测运用基于 Faster-RCNN 深度学习模型,2 种学习模型准确率均较高(分别为 92% 和 90%);Zhao 等^[57]使用卷积神经网络模型对薄血片的疟原虫感染进行预测,准确率为 97%。

5 机器学习可能存在的问题

机器学习算法的分析能力极高,优于经典统计和建模方法,但其主要缺点依然是所谓的黑箱问题:机器学习算法进行数据分析是基于机器阐述的数学函数、算法或模型,理解机器学习算法中获得所研究问题的机械答案不容易。机器学习虽可以从变量中预测结果,但是无法给出结果发生的原因,人们无法理解得出结果的机制,即机器学习无法回答结果是如何产生的(有相关性不等于有因果关系)。因此如果在数据拟合过程中出现数据不完整、不稳定、偏差甚至是错误的情况,相关性都是在不完全真实的数据上产生的拟合结果,可能对拟合效果产生错误的判断。因此基于机器学习的

蚊媒研究,一定要在真实可靠的数据基础上,采用平稳可信的模型进行模拟分析。

6 机器学习在蚊媒研究中的总结和展望

基于机器学习的蚊媒研究目前绝大多数是对蚊虫和蚊媒传染病的预测预警分析,基于蚊虫研究新技术的发展,基于宏基因组、二代测序技术、结构基因组学、功能基因组学、蛋白质组学、生物芯片等应用到微观蚊媒研究中,机器学习也会相应地从宏观的蚊媒生态学研究走向微观研究中去;同时深度学习技术的应用也会带来蚊媒研究的新体验,以往的蚊媒机器学习研究一般集中在决策树、RF、SVM 等,但随着深度学习技术的成熟,包括 ANN 在内的深度学习将不断开拓蚊媒研究的新领域;另外学科交叉也会催生机器学习应用于蚊媒研究的新领域,遥感、定位、地理信息系统代表的“3S”技术应用在蚊媒研究领域、传染病动力学在蚊媒传染病中的作用、社会因素对于蚊媒传染病的传播驱动等新方向都会再一次使得机器学习与蚊媒研究更加紧密结合,必将对蚊媒研究带来更多更好的丰硕成果。

利益冲突 无

参考文献

- [1] 徐承龙,姜志宽. 蚊虫防制(一):蚊虫的危害与形态分类[J]. 中华卫生杀虫药械, 2006, 12(4): 289-293. DOI: 10.3969/j.issn.1671-2781.2006.04.027.
Xu CL, Jiang ZK. Mosquito control (1): The damage and morphological classification of mosquitoes[J]. Chin J Hyg Insect Equip, 2006, 12(4): 289-293. DOI: 10.3969/j.issn.1671-2781.2006.04.027.
- [2] 瞿逢伊. 我国蚊虫种质资源现状及其共享利用[J]. 中国寄生虫学与寄生虫病杂志, 2006, 24 增刊: 13-16. DOI: 10.3969/j.issn.1000-7423.2006.z1.003.
Qu FY. Current status, utilization and sharing of mosquito germplasm resources in China[J]. Chin J Parasit Parasit Dis, 2006, 24 Suppl: S13-16. DOI: 10.3969/j.issn.1000-7423.2006.z1.003.
- [3] 张菊仙,龚正达. 中国蚊类研究概况[J]. 中国媒介生物学及控制杂志, 2008, 19(6): 595-599. DOI: 10.3969/j.issn.1003-4692.2008.06.047.
Zhang JX, Gong ZD. Survey of mosquito research in China[J]. Chin J Vector Biol Control, 2008, 19(6): 595-599. DOI: 10.3969/j.issn.1003-4692.2008.06.047.
- [4] 张仪. 新发媒传疾病及其防控[J]. 中国血吸虫病防治杂志, 2012, 24(5): 501-504. DOI: 10.3969/j.issn.1005-6661.2012.05.001.
Zhang Y. Emerging vector-borne diseases and control[J]. Chin J Schisto Control, 2012, 24(5): 501-504. DOI: 10.3969/j.issn.1005-6661.2012.05.001.
- [5] 边长玲,龚正达. 我国蚊类及其与蚊媒病关系的研究概况[J]. 中国病原生物学杂志, 2009, 4(7): 545-551. DOI: 10.13350/j.cjpb.2009.07.010.
Bian CL, Gong ZD. Mosquitoes and mosquito-borne diseases in China[J]. J Parasit Biol, 2009, 4(7): 545-551. DOI: 10.13350/j.cjpb.2009.07.010.
- [6] 郑学礼. 我国蚊媒研究概况[J]. 中国病原生物学杂志, 2014, 9(2): 183-187. DOI: 10.13350/j.cjpb.140222.
Zheng XL. Advances in research on mosquitoes in China[J]. J Parasit Biol, 2014, 9(2): 183-187. DOI: 10.13350/j.cjpb.140222.
- [7] 王梦蕾,苏昊,吴焜,等. 中国蚊媒病流行现状及防治进展[J]. 热带医学杂志, 2012, 12(10): 1280-1285.
Wang ML, Su H, Wu K, et al. Current status of mosquito-borne diseases in China[J]. J Trop Med, 2012, 12(10): 1280-1285.
- [8] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016: 1-14.
Zhou ZH. Machine learning[M]. Beijing: Tsinghua University Press, 2016: 1-14.
- [9] Pirracchio R, Cohen MJ, Malenica I, et al. Big data and targeted machine learning in action to assist medical decision in the ICU[J]. Anaesth Crit Care Pain Med, 2019, 38(4): 377-384. DOI: 10.1016/j.accpm.2018.09.008.
- [10] 汤胜男,辛学刚. 机器学习在生物信息学领域的应用与研究进展[J]. 人工智能, 2020(1): 84-93. DOI: 10.16453/j.cnki.issn2096-5036.2020.01.009.
Tang SN, Xin XG. Application and research progress of machine learning in the field of bioinformatics[J]. AI-View, 2020(1): 84-93. DOI: 10.16453/j.cnki.issn2096-5036.2020.01.009.
- [11] 吴亚飞,方亚. 机器学习方法在慢性病研究中的应用进展[J]. 中国卫生统计, 2020, 37(4): 624-628.
Wu YF, Fang Y. Progress in the application of machine learning methods in the research of chronic diseases[J]. Chin J Health Stat, 2020, 37(4): 624-628.
- [12] 杜唯佳,徐振东,刘志强. 人工智能在麻醉学领域的应用进展[J]. 国际麻醉学与复苏杂志, 2020, 41(8): 800-803. DOI: 10.3760/cma.j.cn321761-20190620-00099.
Du WJ, Xu ZD, Liu ZQ. Advances of artificial intelligence in anesthesiology[J]. Int J Anesthesiol Resusc, 2020, 41(8): 800-803. DOI: 10.3760/cma.j.cn321761-20190620-00099.
- [13] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会. GB/T 23797-2009 病媒生物密度监测方法 蚊虫[S]. 北京:中国标准出版社, 2009.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Administration Standardization. GB/T 23797-2009 Surveillance methods for vector density-Mosquito[S]. Beijing: Standards Press of China, 2009.
- [14] 肖冰. 城市水绿复合系统的蚊虫孳生现状及成因分析:以上海和池州为例[D]. 上海:华东师范大学, 2019.
Xiao B. Mosquito breeding status and cause analysis of the urban water-green combined system: A case study of Shanghai and Chizhou[D]. Shanghai: East China Normal University, 2019.
- [15] 黄建华,石挺丽,陈远源,等. 多变量灰色模型 MGM(1, n)在白纹伊蚊密度预测中的应用[J]. 中华疾病控制杂志, 2016, 20(1): 87-90. DOI: 10.16462/j.cnki.zhjbkz.2016.01.022.
Huang JH, Shi TL, Chen YY, et al. Application of multivariable grey model (1, n) in prediction of *Aedes albopictus* density[J]. Chin J Dis Control Prev, 2016, 20(1): 87-90. DOI: 10.16462/j.cnki.zhjbkz.2016.01.022.
- [16] 李卫红,陈业滨,闻磊. 基于 GA-BP 神经网络模型的登革热时空扩散模拟[J]. 中国图象图形学报, 2015, 20(7): 981-991.

- DOI:10.11834/jig.20150715.
- Li WH, Chen YB, Wen L. Simulation of spatio-temporal diffusion of dengue fever based on the GA-BP neural network model[J]. J Image Graph, 2015, 20(7):981-991. DOI:10.11834/jig.20150715.
- [17] 周毅彬, 冷培恩, 顾君忠, 等. 上海市白纹伊蚊密度与气象因素关系的研究[J]. 中国媒介生物学及控制杂志, 2014, 25(5):405-407. DOI:10.11853/j.issn.1003.4692.2014.05.005.
- Zhou YB, Leng PE, Gu JZ, et al. Study on relationship between population density of *Aedes albopictus* and meteorological factors in Shanghai, China[J]. Chin J Vector Biol Control, 2014, 25(5):405-407. DOI:10.11853/j.issn.1003.4692.2014.05.005.
- [18] 于德宪, 林立丰, 罗雷, 等. 人工神经网络模型用于分析气候因素对白纹伊蚊密度影响的初步探讨[J]. 南方医科大学学报, 2010, 30(7):1604-1605, 1609. DOI:10.12122/j.issn.1673-4254.2010.07.030.
- Yu DX, Lin LF, Luo L, et al. Establishment of an artificial neural network model for analysis of the influence of climate factors on the density of *Aedes albopictus* [J]. J South Med Univ, 2010, 30(7):1604-1605, 1609. DOI:10.12122/j.issn.1673-4254.2010.07.030.
- [19] Früh L, Kampen H, Kerkow A, et al. Modelling the potential distribution of an invasive mosquito species: comparative evaluation of four machine learning methods and their combinations [J]. Ecol Modell, 2018, 388: 136-144. DOI:10.1016/j.ecolmodel.2018.08.011.
- [20] Kerkow A, Wieland R, Koban MB, et al. What makes the Asian bush mosquito *Aedes japonicus japonicus* feel comfortable in Germany? A fuzzy modelling approach[J]. Parasit Vectors, 2019, 12(1):106. DOI:10.1186/s13071-019-3368-0.
- [21] Demertzis K, Iliadis L, Anezakis VD. Commentary: *Aedes albopictus* and *Ae. japonicus*-two invasive mosquito species with different temperature niches in Europe [J]. Front Environ Sci, 2017, 5:85. DOI:10.3389/fevns.2017.00085.
- [22] Zheng XL, Zhong DB, He YL, et al. Seasonality modeling of the distribution of *Aedes albopictus* in China based on climatic and environmental suitability[J]. Infect Dis Poverty, 2019, 8(1):98. DOI:10.1186/s40249-019-0612-y.
- [23] Scavuzzo JM, Trucco F, Espinosa M, et al. Modeling dengue vector population using remotely sensed data and machine learning [J]. Acta Trop, 2018, 185: 167-175. DOI:10.1016/j.actatropica.2018.05.003.
- [24] Ding FY, Fu JY, Jiang D, et al. Mapping the spatial distribution of *Aedes aegypti* and *Ae. albopictus* [J]. Acta Trop, 2018, 178: 155-162. DOI:10.1016/j.actatropica.2017.11.020.
- [25] 裘炯良, 孙志, 王军, 等. 人工神经网络在外来医学媒介生物输入风险评估中的应用研究[J]. 中华卫生杀虫药械, 2016, 22(5):456-460. DOI:10.19821/j.1671-2781.2016.05.013.
- Qiu JL, Sun Z, Wang J, et al. Application of back propagation neural network on the risk assessment of exotic medical-vector [J]. Chin J Hyg Insect Equip, 2016, 22(5):456-460. DOI:10.19821/j.1671-2781.2016.05.013.
- [26] Wieland R, Kerkow A, Früh L, et al. Automated feature selection for a machine learning approach toward modeling a mosquito distribution[J]. Ecol Modell, 2017, 352:108-112. DOI:10.1016/j.ecolmodel.2017.02.029.
- [27] 洪铭鸿. 基于边缘计算和深度学习之病媒蚊分类系统[D]. 中国台北:台湾师范大学, 2019.
- Hong MH. A vector mosquitoes classification system based on edge computing and deep learning [D]. Taipei, China: National Taiwan Normal University, 2019.
- [28] Sanchez-Ortiz A, Fierro-Radilla A, Arista-Jalife A, et al. Mosquito larva classification method based on convolutional neural networks [C]//Proceedings of 2017 International Conference on Electronics, Communications and Computers. Cholula, Mexico: IEEE, 2017. DOI:10.1109/CONIELECOMP.2017.7891835.
- [29] Amarasinghe A, Suduwella C, Elvitigala C, et al. A machine learning approach for identifying mosquito breeding sites via drone images [C]//Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems. Delft, The Netherlands: ACM, 2017. DOI:10.1145/3131672.3136986.
- [30] Park J, Kim DI, Choi B, et al. Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks [J]. Sci Rep, 2020, 10(1):1012. DOI:10.1038/s41598-020-57875-1.
- [31] Motta D, AÁBSantos, Winkler I, et al. Application of convolutional neural networks for classification of adult mosquitoes in the field [J]. PLoS One, 2019, 14(1):e0210829. DOI:10.1371/journal.pone.0210829.
- [32] Case E, Shragai T, Harrington L, et al. Evaluation of unmanned aerial vehicles and neural networks for integrated mosquito management of *Aedes albopictus* (Diptera: Culicidae) [J]. J Med Entomol, 2020, 57(5):1588-1595. DOI:10.1093/jme/tjaa078.
- [33] Genoud AP, Gao YP, Williams GM, et al. A comparison of supervised machine learning algorithms for mosquito identification from backscattered optical signals [J]. Ecol Inform, 2020, 58:101090. DOI:10.1016/j.ecoinf.2020.101090.
- [34] 李振宇. 基于翅振频率的人工神经网络方法在蚊虫分类中的应用研究[D]. 雅安:四川农业大学, 2005.
- Li ZY. Application of artificial neural network to identify species of mosquitoes (Diptera: Culicidae) based on wingbeat frequency [D]. Ya'an: Sichuan Agricultural University, 2005.
- [35] 罗嘉鹏. 利用组学数据检测昆虫的抗药性和入侵性[D]. 南京:南京师范大学, 2018.
- Luo JP. Detect the insecticide resistance and invasiveness of insects with omics data [D]. Nanjing: Nanjing Normal University, 2018.
- [36] 宋帅葆, 艾上杰, 关怀, 等. 冈比亚按蚊犬尿氨酸甲酰胺酶抑制剂的虚拟筛选[J]. 昆虫学报, 2018, 61(1):68-78. DOI:10.16380/j.kcxb.2018.01.008.
- Song SB, Ai SJ, Guan H, et al. Virtual screening of inhibitors for kynurenine formamidase of *Anopheles gambiae* (Diptera: Culicidae) [J]. Acta Entomol Sin, 2018, 61(1):68-78. DOI:10.16380/j.kcxb.2018.01.008.
- [37] 王婷婷, 郝友进, 何正波, 等. 4种重要医学媒介蚊虫离子受体基因 *IR8a* 和 *IR25a* 的特征及分类地位[J]. 昆虫学报, 2017, 60(4):379-388. DOI:10.16380/j.kcxb.2017.04.003.
- Wang TT, Hao YJ, He ZB, et al. Characteristics and classification position of the ionotropic receptor genes *IR8a* and *IR25a* in four vector mosquito species of medical importance [J]. Acta Entomol

- Sin, 2017, 60(4):379-388. DOI: 10.16380/j.kcxb.2017.04.003.
- [38] 毛启萌, 李廷景, 付文博, 等. 林氏按蚊线粒体全基因组序列的测定及基于线粒体基因组的按蚊属系统发育分析(英文)[J]. 昆虫学报, 2019, 62(1): 101-116. DOI: 10.16380/j.kcxb.2019.01.011.
- Mao QM, Li TJ, Fu WB, et al. Sequencing of the complete mitochondrial genome of *Anopheles lindesayi* and a phylogenetic analysis of the genus *Anopheles* (Diptera: Culicidae) based on mitochondrial genomes[J]. Acta Entomol Sin, 2019, 62(1): 101-116. DOI: 10.16380/j.kcxb.2019.01.011.
- [39] Smith HA, White BJ, Kundert P, et al. Genome-wide QTL mapping of saltwater tolerance in sibling species of *Anopheles* (malaria vector) mosquitoes [J]. Heredity (Edinb), 2015, 115(5):471-479. DOI: 10.1038/hdy.2015.39.
- [40] Saxena P, Mishra S. Study of the binding pattern of HLA class I alleles of Indian frequency and cTAP binding peptide for Chikungunya vaccine development [J]. Int J Pept Res Ther, 2020, 26(2): 2437-2448. DOI: 10.1007/s10989-020-10038-2.
- [41] 黄宇琳, 赵永谦, 曹峥, 等. 基于随机森林回归模型的登革热风险评估研究[J]. 华南预防医学, 2019, 45(1): 26-31. DOI: 10.13217/j.scjpm.2019.0026.
- Huang YL, Zhao YQ, Cao Z, et al. Risk assessment of dengue fever based on random forest model [J]. South China J Prev Med, 2019, 45(1): 26-31. DOI: 10.13217/j.scjpm.2019.0026.
- [42] 任红艳, 吴伟, 李乔玄, 等. 基于反向传播神经网络模型的广东省登革热疫情预测研究[J]. 中国媒介生物学及控制杂志, 2018, 29(3): 221-225. DOI: 10.11853/j.issn.1003.8280.2018.03.001.
- Ren HY, Wu W, Li QX, et al. Prediction of dengue fever based on back propagation neural network model in Guangdong, China [J]. Chin J Vector Biol Control, 2018, 29(3): 221-225. DOI: 10.11853/j.issn.1003.8280.2018.03.001.
- [43] 赵永谦. 珠三角地区精细空间尺度的登革热风险评估模型构建研究[D]. 广州:暨南大学, 2018.
- Zhao YQ. Building models for risk assessment of dengue fever in the Pearl River Delta, China: a study based on the fine spatial scale[D]. Guangzhou: Jinan University, 2018.
- [44] 陈斌. 广东登革热防控能力评估及社区干预实验研究[D]. 北京:中国疾病预防控制中心, 2017.
- Chen B. Capacity assessment and community intervention study on dengue control in Guangdong, China [D]. Beijing: Chinese Center for Disease Control and Prevention, 2017.
- [45] 宋晓晴. 基于大规模手机定位数据的城市内部登革热防控模拟研究[D]. 武汉:武汉大学, 2017.
- Song XQ. Simulating interventions on intra-urban dengue outbreaks using large-scale mobile phone tracking data [D]. Wuhan: Wuhan University, 2017.
- [46] 陈业滨, 李卫红. 支持向量机模型的登革热时空扩散预测[J]. 测绘科学, 2017, 42(2): 65-70. DOI: 10.16251/j.cnki.1009-2307.2017.02.013.
- Chen YB, Li WH. Simulation of spatio-temporal diffusion trend of dengue fever based on the SVM model [J]. Sci Surv Mapp, 2017, 42(2): 65-70. DOI: 10.16251/j.cnki.1009-2307.2017.02.013.
- [47] Zhao NZ, Charland K, Carabali M, et al. Machine learning and dengue forecasting: comparing random forests and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia [J]. PLoS Negl Trop Dis, 2020, 14(9): e0008056. DOI: 10.1371/journal.pntd.0008056.
- [48] Kesorn K, Ongkruk P, Chompoonsri J, et al. Morbidity rate prediction of dengue hemorrhagic fever (DHF) using the support vector machine and the *Aedes aegypti* infection rate in similar climates and geographical areas [J]. PLoS One, 2015, 10(5): e0125049. DOI: 10.1371/journal.pone.0125049.
- [49] Ho TS, Weng TC, Wang JD, et al. Comparing machine learning with case-control models to identify confirmed dengue cases [J]. PLoS Negl Trop Dis, 2020, 14(11): e0008843. DOI: 10.1371/journal.pntd.0008843.
- [50] 明亮. 海南省疟疾流行预测方法及基于 GIS 的疟疾监测预警系统的初步构建[D]. 西安:第四军医大学, 2004.
- Wen L. Research on prediction of malaria epidemic and construction GIS-based malaria surveillance and early warning system in Hainan province, China [D]. Xi'an: Fourth Military Medical University, 2004.
- [51] 郭华. 面向主动监控的感染风险时空模式建模与挖掘[D]. 长春:吉林大学, 2014.
- Guo H. Active surveillance oriented spatiotemporal patterns modeling and mining of infection risk [D]. Changchun: Jilin University, 2014.
- [52] 高春玉. 我国疟疾流行现状及气象因素对疟疾发病影响的神经网络模型初步研究[D]. 重庆:第三军医大学, 2002.
- Gao CY. Study on epidemic situation of malaria in China and neural network model of disease influenced by meteorological factors [D]. Chongqing: Third Military Medical University, 2002.
- [53] 高文, 黄钢, 韩晓莉. 基于蚊密度差分自回归移动平均模型预测流行性乙型脑炎的贝叶斯判别分析研究[J]. 中国媒介生物学及控制杂志, 2018, 29(6): 557-563. DOI: 10.11853/j.issn.1003.8280.2018.06.003.
- Gao W, Huang G, Han XL. Application of Bayes analysis in Japanese encephalitis prediction based on multiple seasonal autoregressive integrated moving average model [J]. Chin J Vector Biol Control, 2018, 29(6): 557-563. DOI: 10.11853/j.issn.1003.8280.2018.06.003.
- [54] Vidal OM, Acosta-Reyes J, Padilla J, et al. Chikungunya outbreak (2015) in the Colombian Caribbean: latent classes and gender differences in virus infection [J]. PLoS Negl Trop Dis, 2020, 14(6): e0008281. DOI: 10.1371/journal.pntd.0008281.
- [55] Eneanya OA, Cano J, Dorigatti I, et al. Environmental suitability for lymphatic filariasis in Nigeria [J]. Parasit Vector, 2018, 11(1): 513. DOI: 10.1186/s13071-018-3097-9.
- [56] 朱远林. 基于深度学习的疟疾自检测与分类算法研究[D]. 北京:北京交通大学, 2019.
- Zhu YL. Research on self-detection and classification algorithm of malaria based on deep learning methods [D]. Beijing: Beijing Jiaotong University, 2019.
- [57] Zhao OS, Kolluri N, Anand A, et al. Convolutional neural networks to automate the screening of malaria in low-resource countries [J]. PeerJ, 2020, 8: e9674. DOI: 10.7717/peerj.9674.