



计算行为学研究进展

任炜^{1,3,4}, 余山^{2,3,5*}, 张永清^{1,3,4*}

1. 中国科学院遗传与发育生物学研究所, 分子发育生物学国家重点实验室, 北京 100101;

2. 中国科学院自动化研究所, 北京 100190;

3. 中国科学院脑科学与智能技术卓越创新中心, 上海 200031;

4. 中国科学院大学生命科学学院, 北京 100049;

5. 中国科学院大学未来技术学院, 北京 100049

* 联系人, E-mail: shan.yu@nlpr.ia.ac.cn; yqzhang@genetics.ac.cn

2021-02-25 收稿, 2021-05-05 修回, 2021-05-06 接受, 2021-05-07 网络版发表

国家重点研发计划(2017YFA0105203, 2019YFA070710)、国家自然科学基金(31921002, 31830036)、北京市科学技术委员会(Z1811000015180010)和中国科学院战略性先导科技专项(XDBS1020100)资助

摘要 20世纪的行为学研究通过在自然状态下和受控的实验室条件下的观察揭示了动物行为的众多本质特点. 但这两种途径都存在各自的局限, 不能完全满足动物行为学研究发展的需要. 近年来, 计算技术的发展与突破给行为学研究带来了新的机遇, “计算行为学”这一新兴学科正在形成. 计算行为学的发展、完善和应用使得高效量化分析动物行为成为可能. 本文对计算行为学的发展历史和前沿进展进行了系统梳理和总结, 介绍了人工智能技术用于动物行为分析的基本概念、方法和应用场景. 此外, 还对计算行为学在行为的神经机制解析和神经精神疾病诊断方面的应用进行了概述, 分析了目前该领域所面临的挑战与机遇, 希望能为行为科学、脑科学, 以及神经精神疾病的临床转化研究提供参考.

关键词 动物行为, 人工智能, 计算行为学, 姿态估计, 机器学习

基因和环境共同塑造了动物的神经系统. 神经系统直接控制行为, 而行为往往随着外界环境的变化而变化. 对高等动物来说, 脑是动物行为的指挥中心, 而行为则是脑功能的表现, 所以系统的行为学分析是揭示脑功能的重要研究手段.

达尔文认为, 动物行为在物种演化过程中受到环境选择的压力, 所以行为有其生物学基础, 并且类似的行为能在同类后代和相近的物种中表现出来^[1]. 自然状态下和实验场景下的行为研究是20世纪行为学研究的两种主要范式.

通过观察自然状态下的动物行为, 行为学家想要探索动物在自然环境中本能行为的基本规律. 早期动物行为学研究领域的代表性学者Tinbergen、Lorenz和

von Frisch因对动物个体与群体行为模式的开创性研究而获得1973年的诺贝尔生理学或医学奖. Tinbergen^[2]在本能行为的研究中提出了信号刺激(signal stimulus)的理论. 他发现海鸥雏鸟一生下来就表现啄母鸟鸟喙的行为, 并证明母鸟鸟喙上的红点为一种信号刺激, 引发了雏鸟“啄”的行为. 他也发现, 雏鸟张大的菱形嘴巴是母鸟喂食行为的信号刺激. von Frisch^[3]发现, 蜜蜂间存在一种简单的社交行为, 用以传达花蜜的位置和距离信息. 当蜜蜂发现一处蜜源时, 会根据蜜源地的远近作出两种行为: 圆舞(round dance)与摇摆舞(waggle dance). 圆舞是表达蜂蜜就在附近, 摇摆舞则是传递蜜源与蜂巢距离的信息, 蜜源距离越远, 蜜蜂跳摇摆舞的时间越久^[3]. 行为印记(imprinting)是Lorenz^[4]发现的一

引用格式: 任炜, 余山, 张永清. 计算行为学研究进展. 科学通报, 2021, 66: 3799–3810

Ren W, Yu S, Zhang Y Q. Progress on Computational Ethology (in Chinese). Chin Sci Bull, 2021, 66: 3799–3810, doi: 10.1360/TB-2021-0213

种动物学习行为. 印记发生在个体发育的早期阶段, 有一个明显的学习敏感期. 印记作用一旦发生就很难改变, 所以被称为行为印记. 同时Lorenz^[4]也提出了先天诱发机制(innate releasing mechanism): 动物具有某些先天性的反应能力, 一旦遇到特定的刺激, 就会产生相应的本能反应. Tinbergen^[5]与Simmons和Young^[6]认为动物运动具有相对规律的模式. 这种模式表现在行为的不同层次上, 所以行为是简单运动模块的不同组合方式所形成的复杂表现. 行为模式的启动、相互协调和终止受环境的控制, 并与环境产生交互作用, 而不是孤立存在的^[5,6]. 行为谱(ethogram)用专业术语描述动物个体的一系列行为^[7], 常用于解析动物的特定行为是如何由简单模块所组成的.

动物行为学的另一条研究途径是依托于生理学与心理学框架下的比较心理学. 这一领域开发了一系列的行为学实验范式, 试图在一个更可控的实验场景下开展对动物行为的探索和研究. 通过精心设计的实验范式得以理解动物的习得性行为, 包括经典条件反射和操作性条件反射^[8,9]. 通过该研究途径发现了学习和动机的基本原理, 发展出一系列有关行为组织规则及其机制的理论, 比如驱动力满足理论(drive satisfaction)^[10]和认知地图理论(cognitive maps)^[11]等. 驱动力满足理论认为, 行为的产生与选择是动物受本能欲望等驱动力(或称为动机)的影响而满足自身欲求的结果. 认知地图理论认为, 动物的大脑通过对环境信息的加工, 借由视觉、本体感觉、嗅觉及听觉等感官的输入来构建和积累与空间相关的信息^[11~13]. 此外, 还建立了不同的实验范式来研究各种细分的行为模式及其背后的影响因素, 比如用高架十字迷宫^[14]和旷场实验^[15]来研究动物的抑郁、焦虑和恐惧等多种情绪.

虽然早期动物行为学研究通过自然状态和实验范式中的观察获得了许多有价值的发现, 但两种方法都有各自的局限性. 自然状态下的行为观察难以控制各种环境变量, 因而影响观察结果的可重复性. 动物的行为在特征空间(feature space)常常具有高维度的属性, 包含的信息丰富^[16]. 视频录像技术出现之前的早期行为谱研究依靠简单观察与书面描述, 信息量较少, 量化分析较简单.

为了解决自然状态下观察的局限性, 实验范式中的观察有可控制的实验变量, 实验结果更有可重复性. Colgan^[17]主张通过使用近交系和标准饲养条件来减少动物个体之间的行为差异. Skinner^[18]建议进行更严格

的实验控制, 通过训练动物执行特定任务使动物个体之间的行为只存在微小的差异. 但这类解决方案的问题是以牺牲实验结果的普适性为代价来换取实验结果的一致性^[19]. 动物被限制在特定的空间范围内, 只能开展受限的行为, 一些实验范式还需要对动物进行长期训练. 这些限制和人为的训练同样会给实验结果带来偏差, 同时这样的范式所记录的实验数据维度也相对较低, 不能全面地反映动物的行为与心理状态.

计算行为学(Computational Ethology)经过20年的发展, 与完善逐渐得到学界的重视(图1). 计算行为学是数学、工程学、计算机科学与动物行为学深度交叉融合的结果, 是一门运用人工智能方法对各种条件, 包括自由活动条件下的动物行为进行测量、描述和分析的学科^[20]. 人工智能领域的发展为动物行为学研究提供了强大的分析工具. 人工智能在这里指的是能模仿人类认知功能的软硬件综合体^[21], 其研究领域包括机器人、语言识别、图像识别、自然语言处理和专家系统等. 越来越多的研究利用人工智能技术对动物行为进行自动化的定量分析, 并挖掘未知的行为特征^[20].

计算行为学涉及的主要技术包括动物追踪技术与自动行为识别技术. 动物追踪技术的发展经历了从粗略的运动信息提取到精细的姿态信息提取. 自动行为识别技术的发展经历了从早期个体行为的分析到群体行为的分析, 以及从一种动物的行为识别逐渐拓展到多种动物的行为识别. 自动行为识别涉及机器学习技术, 包括有监督分类技术与无监督聚类技术. 从最初常用的有监督分类技术到最近发展的无监督聚类技术,

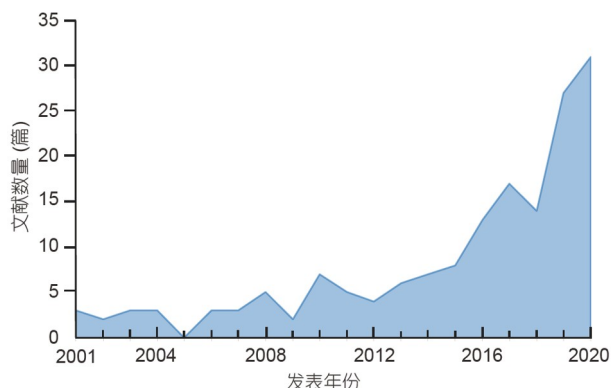


图1 (网络版彩色)PubMed数据库内以“计算行为学”为关键词查询所得到的最近20年文献发表数量. 该统计以2001年为计算行为学领域形成的起始点^[20]

Figure 1 (Color online) Number of publications over the last 20 years recorded in PubMed with the keyword: Computational Ethology. The outset of Computational Ethology is the year of 2001^[20]

使得动物行为标识的人工依赖逐渐降低。结合高精度的行为学数据与多模态数据进行综合分析成为最近计算行为学发展的一个新方向^[22~24]。我们对计算行为学目前存在的一些问题进行了总结,并梳理了可能的解决方案。希望未来的计算行为学能帮助我们解析行为控制的神经机制,加深对脑疾病的理解。

1 计算行为学前沿研究进展

1.1 动物追踪技术发展

视频录像技术出现之后,在人工智能技术广泛应用于行为学分析之前,行为学定量分析主要依赖数字图像处理技术(digital image processing)。数字图像处理可以提取图像中的频率、颜色、边界、区域、纹理、形状等特征。数字图像处理技术常常通过计算动物所占面积的质心位置,将所观察的动物简化为一个点来分析^[25]。但是这种方法只能得到动物的位置信息。为了获取动物躯干的朝向信息,动物图像被近似为椭圆形,椭圆的长轴朝向代表了动物躯干的朝向^[25]。早期的追踪技术只能对单只动物进行追踪,随后针对多只动物的追踪技术逐渐成熟。多动物追踪技术利用视频中的时间信息对不同动物个体进行区分^[26~29]。首先选定一帧图像并指定好不同个体,再假定不同个体每帧图像之间的移动速度相对恒定且移动的距离较短,从而计算其中某个个体在上一帧与下一帧之间的距离,相互距离最小的两帧图像被认定来自同一个体。以此类推,重复该过程可以自动区分不同的个体^[26~29]。

诺达思信息技术公司(Noldus Information Technology)在2001年发布了小鼠行为分析软件Ethovision^[25]。Ethovision采集小鼠旷场实验和水迷宫实验视频数据后,利用数字图像处理技术提取图像特征,计算出每帧图像中小鼠的移动轨迹、移动速度、身体转弯角度等运动参数。2009年,加州理工学院(California Institute of Technology)的de Chaumont等人^[26]开发了果蝇行为分析软件CTrax。该软件能对单只果蝇的个体行为与多只果蝇的社交行为进行分析评判,在此基础上还能准确区分果蝇性别和不同表型对应的基因型。随后动物追踪技术开始从二维平面向三维空间发展。2011年,Straw等人^[30]开发了基于多视角三维重建的动物飞行轨迹捕捉系统Flydra,该系统能够追踪果蝇和鸟类等动物在飞行过程中的身体方向和三维空间位置。2014年,Pérez-Escudero等人^[31]开发了idTracker多动物追踪算法

平台,能对3种实验室常用的模式生物,包括小鼠、果蝇和斑马鱼进行运动追踪。

在更强大的人工智能技术辅助下,动物行为定量分析从原来粗略的个体运动追踪发展为更加精细的姿态分析。

人工神经网络(artificial neural network)是一种模仿生物神经网络的结构和功能的数学模型。深度学习(deep learning)又称深度神经网络(deep neural network),是一种以人工神经网络为架构,对输入数据进行表征学习的算法。深度神经网络由简单的神经元组成,这些神经元按层组织,然后依次堆叠形成网络。深度神经网络能学会从原始数据中提取有价值的信息^[32,33]。

姿态估计(pose estimation)是图像识别领域一个具有挑战性的技术问题。动物行为学中的姿态估计指的是对于视频中的动物肢体关键点进行定位与标记。姿态估计是动物行为学研究的一个重要技术手段,使得动物肢体关键点的自动追踪成为可能。心理物理学家Johansson^[34]认为,人类和许多其他动物的运动是由肢体相对于关节点摆动形成的几何结构所决定的。他的研究表明,若干个关键点的连贯移动足以被感知为运动。这为利用动物运动过程中的肢体关键点来分析动物行为提供了理论依据。

由于深度学习的发展,基于深度学习的姿态估计技术在过去5年中取得了巨大的进步。目前计算行为学已经广泛采用了基于深度学习的姿态估计技术,我们将近期发表的相关工具软件归纳总结如表1。基于深度学习的动物姿态估计已经在果蝇^[35~38,40,42]、啮齿类^[35~37,39,42,44]、狗^[41]、马^[47]、猕猴^[41,43,45,48]和狨猴^[49]上实现。但是其中大部分是一些方法学的研究,还缺乏与脑中行为控制机制的关联分析。从2015年起,计算机视觉领域开始关注利用深度学习解决从二维物体图像估计三维形状的问题^[50],并在鸟类^[51]、犬^[52]以及其他多种动物^[53~57]中进行了尝试。

相比于传统的有标记法姿态估计,基于深度学习的无标记姿态估计精确度较低。但传统研究中在动物身上佩戴的标记容易脱落,而且由于标记附着在松弛的皮肤上,会随之移动,不能准确地标记关节点。最近的研究表明,通过手术附着在大鼠体表的方法可较好地固定标记位置,从而提高标记姿态估计的精确度^[58]。

1.2 基于动物姿态的自动行为识别

目前基于姿态估计的行为学数据分析的基本流程

表 1 基于深度学习的动物姿态估计工具总览^{a)}

Table 1 Overview of deep learning tools for animal pose estimation

	动物种类	三维重构	多动物追踪	图形用户界面	引用量(截至2021-01-20)
Deeplabcut ^[35]	种类广泛	有	有	有	581
LEAP ^[36]	种类广泛	无	有	有	143
DeepPoseKit ^[37]	种类广泛	无	无	部分	90
DeepFly3D ^[38]	有限种类	有	无	部分	33
DeepBehavior ^[39]	有限种类	有	有	无	22
Anipose ^[40]	种类广泛	有	无	无	9
MONET ^[41]	有限种类	有	无	无	8
Optiflex ^[42]	种类广泛	无	无	部分	5
OpenMonkeyStudio ^[43]	有限种类	有	有	无	5
FreiPose ^[44]	有限种类	有	无	无	3
MacaquePose ^[45]	有限种类	无	有	无	2
AlphaTracker ^[46]	有限种类	无	有	有	0

a) 更多的工具软件请参考网站: https://github.com/anl13/animal_papers

为图像采集和处理以及数据加工和分析(图2)。图像采集和处理的基本流程包括: (a) 数据收集. 动物的运动是在三维空间中进行的, 姿态的精确估计需要进行三维重构. 多个相机同步拍摄收集动物运动时不同角度的信息, 用于姿态的三维重构. 除此之外, 如果已经建立了基于多相机重建三维姿态的标注数据库, 可以通过深度学习利用单相机视频估计动物的三维姿态^[61]. (b) 姿态估计. 姿态估计在动物身上佩戴标记点的有标记法与基于深度学习的无标记法. 基于深度学习的无标记姿态估计自动识别并标记动物的肢体关键点. (c) 相机标定及姿态三维重建. 多个拍摄角度的视频信息经过关键点自动识别追踪得到不同角度的关键点图像坐标, 进一步根据标定相机得到的相机参数计算出关键点空间三维坐标.

姿态数据加工和分析的基本流程包括: (d) 数据降维. 姿态三维重建得到高维的时间序列数据, 分析起来有较大难度, 计算成本高, 一般通过数据降维以方便计算和可视化^[59]. 数据降维有主成分分析(principal component analysis)^[62]、 t 分布随机邻近嵌入法(t -distributed stochastic neighbor embedding)^[58]、UMAP(uniform manifold approximation and projection for dimension reduction)^[43]等方法. (e) 运动模块划分. 基于降维后的数据进行运动模块划分, 针对数据特征进行有监督的分类和无监督的聚类分析. (f) 运动模块可视化. 在特征空间中, 高维度的姿态信息降维后可发现某个特定

运动模块反复出现, 特征空间中的不同轨迹代表不同的运动模块^[60]. (g) 运动模块组合. 动物行为往往是由反复出现的、不同时序的运动模块叠加组合而成^[6]. (h) 计算行为模块转换规律. 根据不同行为模块在时间序列上的变化计算得到行为模块的转换规律.

计算行为学利用机器学习技术来自动识别行为(图3). 有监督的行为模块分类主要分为两个步骤. 首先是从视频中将动物每一帧的行为特征抽提出来, 如动物的运动速度、朝向角度及动物和喂食器的距离等人为定义的特征. 然后将这些特征作为分类器(分类器是一种通过学习已有数据建立的数学模型)的输入, 由分类器进行自动分类. 2000年, Rousseau等人^[63]利用前馈神经网络模型作为分类器对大鼠行为进行了分类. 但是该模型的泛化能力较差, 面对训练集之外的大鼠数据时, 正确率只有63.74%. 2010年, Jhuang等人^[64]采用隐马尔可夫模型和支持向量机混合模型(hidden Markov model support vector machine)所构成的分类器对小鼠的居家笼行为进行分析. 虽然分类器的准确度和手工方法接近, 但是仍然存在一些局限. 该模型的输入包括小鼠的位置信息, 如果观察记录的实验环境发生变化, 则同一行为的位置信息会发生较大的改变, 这将导致分类器准确性下降. 另外, 由于该研究采用单视角拍摄, 如果改变拍摄角度也会使分类器准确性下降^[64]. 2012年, Burgos-Artizzu等人^[65]利用Adaboost分类器分析了小鼠的社交行为. 2013年, Kabra等人^[66]开发了动

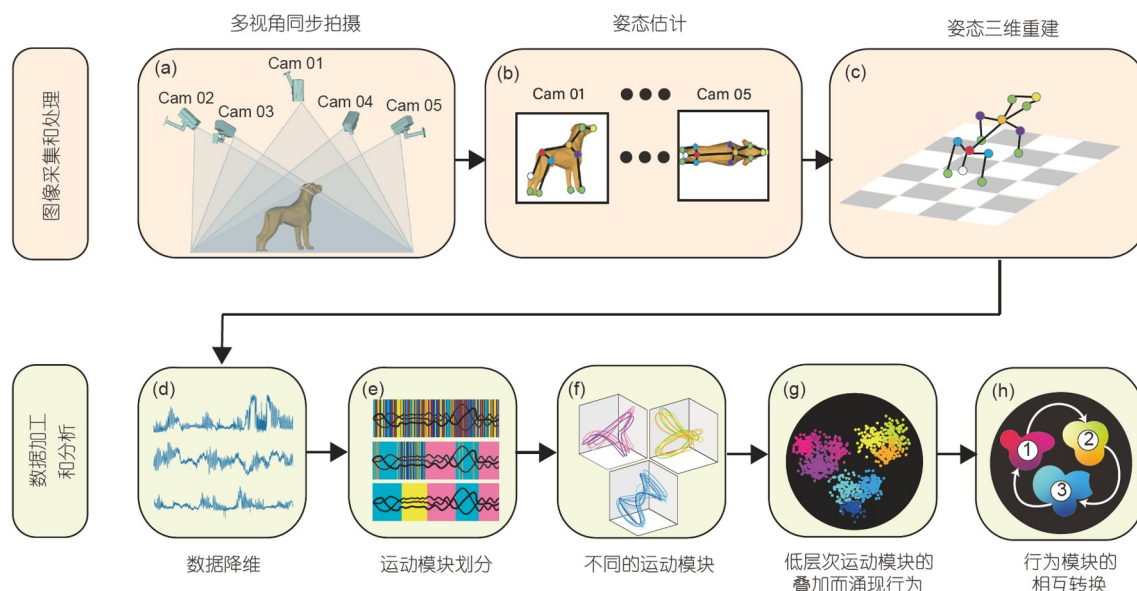


图 2 (网络版彩色)行为数据收集和分析流程。(a) 多个相机在不同角度的同步拍摄动物运动, 获得的视频用于姿态的三维重建; (b) 姿态估计采取在动物身上佩戴标记点的有标记法与基于深度学习的无标记法。基于深度学习的无标记姿态估计自动识别标记动物的肢体关键点; (c) 在标定相机以后, 通过对不同视角的肢体关键点二维坐标进行三角测量得到肢体关键点的三维坐标; (d) 具有高维数据属性的动物姿态需要数据降维, 方便计算与可视化^[59]; (e) 降维后的数据进行运动模块划分, 针对数据特征进行分类或聚类分析(不同的颜色代表不同的运动模块); (f) 降维聚类产生的动作模块在行为空间上的轨迹^[60]; (g) 动物行为是由反复出现的、不同时序的运动模块叠加组合而成^[6]; (h) 根据不同行为模块在时间序列上的变化计算得到行为模块的转换规律

Figure 2 (Color online) A pipeline for the collection and analysis of behavioral data. (a) Multi-view synchronized cameras capture animal motion, and acquired videos are used for 3D pose reconstruction; (b) marker-based tracking and markerless tracking are both applied in animal pose estimation. Markerless tracking based on deep learning is able to automatically track body parts; (c) triangulation of 2D position of body parts from multi-view synchronized videos produce 3D position of body parts after camera calibration; (d) dimensionality reduction which facilitates visualization and calculation^[59]; (e) clustering analysis or classification analysis on data features after module segmentation (motion modules are identified by different colors); (f) trajectories of animal motion modules produced by dimensionality reduction and clustering analysis^[60]; (g) behavior is a stereotyped sequence of motion modules along the time series^[6]; (h) transition between behavior modules according to time series of behavior modules

物行为分析平台JAABA。JAABA采用了Gentleboost分类器, 能对小鼠、果蝇成虫、果蝇幼虫的个体行为与群体行为进行分类。近年来, 利用无监督聚类分析动物行为的研究开始出现: 2015年, Wiltschko等人^[67]开发了小鼠自然行为分析算法平台MoSeq, 对小鼠在旷场实验中的自然行为进行分析。MoSeq的行为聚类算法是一种概率图模型——自回归隐马尔可夫模型(auto-regressive hidden Markov model, AR-HMM), 该模型能对降维后的小鼠姿态信息进行聚类。结果显示, 小鼠行为由亚秒级的运动模块组成。MoSeq还成功检测到小鼠嗅闻2,5-二氢-2,4,5-三甲基噻唑啉(狐狸粪便的成分, 能诱发恐惧反应)与*Ror1β*基因突变导致的行为改变, 说明MoSeq能检测到环境刺激与基因突变对行为的影响^[67]。2020年, Mearns等人^[68]利用亲和力传播算法(affinity propagation clustering)对斑马鱼捕食过程中的游动姿态进行聚类分析发现, 斑马鱼在捕食过程中会出现不

同的运动模式, 能够根据猎物的情况灵活选择。

从以上介绍可以看出, 动物行为识别由个体行为分析逐渐发展到群体行为的分析, 同时从一种动物的行为识别逐渐拓展到对多种动物的行为识别。相比于有监督的行为分类, 无监督聚类技术分析动物行为还处在起步阶段。无监督聚类技术更多地用于挖掘之前未发现的动物行为特征, 作为有监督分类技术的补充。

1.3 计算行为学在人类精神疾病诊断中的应用

计算行为学在人类神经精神疾病的诊断中也发挥了重要的作用。伊利诺伊大学芝加哥分校(University of Illinois at Chicago)的Cao等人^[69]开发了名为DeepMood的预测神经网络。该神经网络能对双相情感障碍患者在使用智能手机过程中的打字速度与智能手机内部自带的加速度传感器数据进行分析。通过对少于1 min的

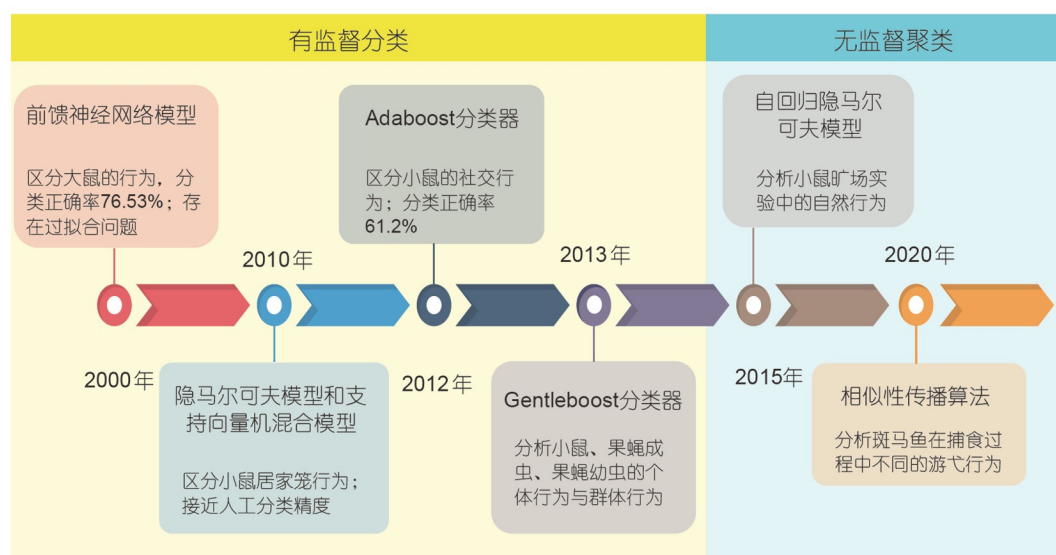


图3 (网络版彩色)过去20年动物行为自动识别分析技术重要进展时间线。计算行为学中基于机器学习的行为分析主要包括有监督分类与无监督聚类分析。有监督分类的应用进展: 2000年, Rousseau等人^[63]采用前馈神经网络模型作为分类器对大鼠行为进行分类。2010年, Jhuang等人^[64]采用隐马尔可夫模型和支持向量机混合模型分类器对小鼠的居家笼行为进行分类。2012年, Burgos-Artizzu等人^[65]采用Adaboost分类器对小鼠社交行为进行分类。2013年, Kabra等人^[66]采用Gentleboost分类器对小鼠、果蝇成虫、果蝇幼虫的个体行为与群体行为进行分类。无监督聚类的应用进展: 2015年, Wiltchko等人^[67]采用自回归隐马尔可夫模型分析旷场实验中小鼠的自然行为。2020年, Mearns等人^[68]采用相似性传播算法分析斑马鱼的捕食行为

Figure 3 (Color online) Advances in the analysis of automated behavior recognition over the past 20 years. Supervised classification and unsupervised clustering in behavior recognition. Advances in the application of supervised classification: In 2000, feed-forward neural network was used by Rousseau et al.^[63] in classification of rat behavior. In 2010, hidden Markov model support vector machine was used by Jhuang et al.^[64] in home-cage behavioral phenotyping of mice. In 2012, Adaboost classifier was used by Burgos-Artizzu et al.^[65] in mouse social behavior recognition. In 2013, Gentleboost classifier was used by Kabra et al.^[66] in classification of individual and social behavior for different organisms, including mice and adult and larval *Drosophila*. Advances in the application of unsupervised clustering: In 2015, auto-regressive hidden Markov model was used by Wiltchko et al.^[67] in mouse naturalistic behavior analysis in open field. In 2020, affinity propagation clustering was used by Mearns et al.^[68] to deconstruct hunting behavior of zebrafish

打字速度与手机移动加速度的分析, DeepMood就可以对使用者的抑郁情况进行评分, 准确率达到90.31%。纽卡斯尔大学(Newcastle University)的Hammerla等人^[70]通过手腕式运动传感器收集现实环境中帕金森患者的运动数据, 利用神经网络对疾病的严重程度进行区分, 展现了良好的临床应用前景。罗德岛大学(University of Rhode Island)的Goodwin等人^[71]利用可穿戴式的加速度传感器自动检测自闭症儿童的身体摆动和手部拍打等刻板行为, 能正确识别大约90%的刻板行为。乔治佐亚理工学院(Georgia Institute of Technology)的Westeyn等人^[72]使用一种称为隐马尔可夫模型(hidden Markov model)的统计模型可区分自闭症儿童不同类型的刻板行为。精神分裂症患者通常在情绪和语言方面表现出障碍, 这些障碍常常伴随面部表情的异常特征。这启发了伦敦玛丽女王大学(Queen Mary University of London)的Bishay等人^[73]开发出一种新的神经网络工具SchiNet, 它能自动检测与精神分裂症相关的面部表情

异常。这些研究表明, 人工智能结合行为特征的分析在神经精神疾病的诊断和治疗效果评估中具有广阔的应用前景。

1.4 多模态数据的联合分析

为了更深入地理解控制行为的神经网络机制, 需要将计算行为学方法与其他多模态的数据进行有效的整合, 例如基因、脑网络结构, 以及脑活动等数据。目前大规模神经网络记录技术(电生理、微型双光子、微内窥)的发展为计算行为学结合脑活动数据的分析提供了可能。Lin等人^[22]以斑马鱼为研究对象, 利用全脑钙成像技术在决策行为的控制机制研究上获得了突破。该研究利用热刺激作为惩罚使斑马鱼学会将尾部朝向不被惩罚的一边。他们发现, 小脑在行为决策上发挥了重要作用, 小脑的脑活动能提前10 s以上预测斑马鱼的行为。但该研究为了观察斑马鱼的全脑钙成像, 需要将斑马鱼的头部固定, 斑马鱼的活动受限, 所研究的

行为比较简单^[22]。Cong等人^[23]开发了一种新的斑马鱼脑成像技术与三维运动跟踪技术,可以监测自由活动中斑马鱼幼体的全脑神经活动。利用在斑马鱼自由活动条件下的全脑记录有望为研究更加自然、丰富的行为提供条件。另外,Nourizonoz等人^[24]开发了名为“EthoLoop”的动物行为分析平台。该平台能通过无线光遗传操纵技术和无线设备刺激诱发特定行为,并利用无线技术记录动物脑电活动。但该工作只是验证了已知的行为控制机制,尚没有深入研究未知的行为控制机制^[24]。

2 计算行为学面临的挑战与机遇

2.1 数据标注

目前基于深度学习姿态估计的精确度较大程度上取决于训练样本的质量和数量,高质量的海量训练样本需要投入大量的人力、物力和时间成本。解决该问题目前有几条可行的途径。一是小样本学习。小样本学习旨在通过少量样本学习得到解决问题的模型。目标检测(object detection)领域出现了利用小样本来训练神经网络的尝试^[74]。姿态识别领域尚需要结合新的神经网络架构和训练方式从而降低对训练样本的依赖。另一种解决途径是多任务学习。与单任务学习只使用单个任务的训练样本不同,多任务学习通过共同训练和优化建立不同任务之间的联系^[75]。这种训练模式充分促进任务之间的信息交换并达到相互学习的目的,尤其是在各自任务训练样本有限的条件下,各个任务之间可以相互启发,间接利用其他任务的数据,从而缓解了对大量标注数据的依赖,达到提升各自学习模型准确度的目的^[75]。

动物的姿态估计可以通过多任务学习提升整体效果。例如,未来可以尝试将只有骨架标注和只有身体部位标注的数据联合训练,研究是否可以在不增加标注成本的情况下增加精度。自监督学习(self-supervised learning)是一种不需要人工标注训练样本,直接利用训练样本数据本身作为监督信息,减少数据标注成本。对比学习(contrastive learning)是自监督学习的一个分支。对比学习是自监督学习的一个重要部分,已被广泛运用在计算机视觉和自然语言处理等领域^[75]。前置任务(pretext tasks)是对比学习中非常重要的一个环节。在对比学习的前置任务之中,同一样本经过数据增强的不同版本被当作正样本(positive sample),其余的样本

被当作负样本(negative sample),对比学习就是训练一个模型用于区分正负样本^[76],这一过程无需任何人工标注。完成前置任务训练的网络再通过较少量的标注数据进行监督学习。目前对比学习已经在目标检测领域进行了尝试^[77-80]。我们相信,在不久的将来,对比学习也会在行为分析领域发挥重要作用。还有一种解决数据量不足的方法是利用合成数据(synthetic data)作为训练集。合成数据通过使用计算机模拟的场景来获取特定的图像、视频和其标注信息。目前已有利用合成数据进行动物姿态估计的尝试^[81],预计未来的研究将进一步减小合成数据和真实数据之间的差异。

2.2 动物追踪

如何同时提高追踪的分析速度和准确性是动物追踪算法中的一个难题。2020年,Kane等人^[82]提出了两个有效的解决方案:使用图形处理器(graphics processing unit)与动态裁剪(dynamic cropping)。图形处理器拥有比中央处理器更多的计算单元,能有效提升图形计算的效率。动态裁剪指计算机在检测到动物时将图片中包含动物的部分进行裁剪,这样可以缩减需要计算的图片文件体积,同时缩小姿态估计的范围,达到同时提高分析速度和准确性的目的。变化的环境背景会影响动物追踪的稳定性。以表1中提到的LEAP为例,虽然其神经网络模型简洁,对特定的动物行为的检测快速准确,但难以应用到变化的环境中^[36]。该问题可以从两个方面入手解决:如果在不同的实验室条件下进行拍摄,可以事先拍下无动物的不同条件的静止背景画面,在进行数据分析前通过帧差法进行背景减除^[83]。如果是在动态变化的自然环境中进行拍摄,则可以尝试通过增加训练样本数的办法来提升动物追踪的稳定性。

计算行为学研究的另一个挑战是动物追踪算法面对动物出现遮挡时的表现较差,如果是多只动物同时追踪则容易出现身份混淆。穿戴式传感器技术可以获得动物姿态信息,它有效解决了佩戴标记点的有标记法与基于深度学习的无标记法存在的视线遮挡问题,以及多只动物追踪识别混淆的问题。但是目前常用的传感器需要与电源搭配使用,穿戴式的传感器体积偏大,增加了传感器的重量,佩戴在动物身上易引起不适和脱落,导致行为测量的误差,所以限制了传感器在动物姿态估计中的应用。传感器技术的发展为该问题提供了新的解决思路。WISP(Wireless Identification and Sensing Platform)是具有实时传输数据功能的无线识

别传感平台,它不需电池或连接电源线便能运作^[84]. WISP能够接收超高频射频识别读写器(UHF RFID reader)发出的无线电波作为能源,通过微处理器驱动控制传感平台的运行,进而采集所需数据,最后再通过超高频射频识别读写器,将采集到的数据实时传输回计算机上.在自由活动的实验场景下,理论上在动物肢体皮下植入此类传感器可以在没有电源的情况下长时间记录和发送动物的姿态信息.预期该技术将在动物行为研究中得到广泛的应用.

2.3 行为序列分割

计算行为除了在姿态估计中存在以上问题,在后续的数据分析处理中也存在不小的挑战.动物在执行某种行为时姿态随着时间的变化而改变,通过姿态估计能得到不同肢体位点随时间变化的曲线,这种曲线是一种时间序列.时间序列的分割是一大挑战.动物可能在进行一种行为同时存在另一种行为,一种行为也能划分成不同层次的次级行为.动物行为的这些特征给行为序列的分割带来困难.时间序列符号表示(symbolic representations of time series)是一种有潜力的解决方案,其在数据挖掘领域得到越来越多的重视^[85].该算法指的是将时间序列表示成具有可读性的符号序列.转化为符号序列后,可以尝试利用文本处理等领域的算法进行进一步分析.目前时间序列符号表示在人体运动识别领域中已有了初步尝试^[86].

2.4 计算行为学向临床的转化

计算行为学在脑科学的基础研究与重大神经精神

疾病的临床转化研究有着巨大的应用潜力.计算行为学为建立新的疾病动物模型提供了有效的分析方法.精神疾病的研究面临特殊的挑战,因为诊断这些疾病主要依靠对病人的行为评估,而这些人类异常行为在常用的啮齿类动物模型中缺乏明确的对应^[87].与小鼠相比,猴的行为与人类的更为接近.但是猴的繁殖周期长、成本高,因此难以大规模用于疾病模型研究^[88].目前针对这些复杂精神疾病缺乏理想的动物模型,阻碍了对病理机制的理解以及药物靶点和生物标记物的探索 and 发现,同时也给临床检测和治疗手段的开发带来了挑战.借助计算行为学的研究方法,辅以基因编辑技术,有机会对自然界中和实验条件下相关动物的行为有更加深刻的理解,对不同基因突变造成的行为差异有更清晰的认识,并为脑科学的基础研究与神经精神疾病研究寻找新的动物模型.

3 小结

本文回顾了行为学研究的两种主要研究途径:自然状态下的行为观察与实验范式中的行为观察.这两种研究途径对早期的行为科学发展都作出了重要贡献,但两种方法都有各自的局限.计算机科学与技术的进步推动了计算行为学的形成与发展,进而促进了动物行为学的研究,为脑科学和神经精神疾病的研究提供了有力的工具.随着人工智能技术的快速发展,以及与脑科学、行为科学、临床医学等学科的不断深入交叉融合,我们相信,计算行为学将在未来发挥出更大的作用,为理解动物行为的组织结构及其神经机制带来前所未有的新机遇.

致谢 衷心感谢中国科学院遗传与发育生物学研究所吴梁和中国科学院自动化研究所胡古月与华娇娇参与讨论并提出宝贵建议.

参考文献

- 1 Darwin C. The Expression of the Emotions in Man and Animals. New York: Oxford University Press, 1998. 27–49
- 2 Tinbergen N. The herring gull's world: A study of the social behaviour of birds. *Bird-Banding*, 1953, 25: 81–82
- 3 von Frisch K. The Dancing Bees. New York: Springer Press, 1954. 100–136
- 4 Lorenz K. Über die Bildung des Instinkt Begriffes. *Naturwissenschaften*, 1937, 25: 289–300
- 5 Tinbergen N. The study of instinct. *J Nerv Ment Dis*, 1954, 120: 141
- 6 Simmons P, Young D. Nerve Cells and Animal Behaviour. Cambridge: Cambridge University Press, 1989. 7
- 7 Lehner P N. Design and execution of animal behavior research: An overview. *J Anim Sci*, 1987, 65: 1213–1219
- 8 Konorski J. Conditioned reflexes and neuron organization. *Phys Ther*, 1948, 29: 237
- 9 Pavlov P I. Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Ann Neurosci*, 2010, 17: 136
- 10 Hull C L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943. 422

- 11 Tolman E C. Cognitive maps in rats and men. *Psychol Rev*, 1948, 55: 189–208
- 12 Colwill R M, Rescorla R A. Postconditioning devaluation of a reinforcer affects instrumental responding. *J Exp Psychol-Anim Behav Process*, 1985, 11: 120–132
- 13 Dickinson A, Nicholas D J, Adams C D. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol Sect B*, 1983, 35: 35–51
- 14 Pellow S, Chopin P, File S E, et al. Validation of open: Closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J Neurosci Methods*, 1985, 14: 149–167
- 15 Hall C S. Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J Comp Psychol*, 1934, 18: 385–403
- 16 Gomez-Marín A, Paton J J, Kampff A R, et al. Big behavioral data: Psychology, ethology and the foundations of neuroscience. *Nat Neurosci*, 2014, 17: 1455–1462
- 17 Colgan P W. Quantitative Ethology. New Jersey: Wiley Interscience, 1978. 7–24
- 18 Skinner B F. Cumulative Record. New York: Appleton-Century-Crofts, 1959. 100–110
- 19 Altmann J. Observational study of behavior: Sampling methods. *Behaviour*, 1974, 49: 227–266
- 20 Anderson D J, Perona P. Toward a science of Computational Ethology. *Neuron*, 2014, 84: 18–31
- 21 Russell S, Norvig P. Artificial Intelligence: A Modern Approach. New Jersey: Prentice Hall, 2002. 1–35
- 22 Lin Q, Manley J, Helmreich M, et al. Cerebellar neurodynamics predict decision timing and outcome on the single-trial level. *Cell*, 2020, 180: 536–551.e17
- 23 Cong L, Wang Z, Chai Y, et al. Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (*Danio rerio*). *eLife*, 2017, 6: e28158
- 24 Nourizonoz A, Zimmermann R, Ho C L A, et al. EthoLoop: Automated closed-loop neuroethology in naturalistic environments. *Nat Methods*, 2020, 17: 1052–1059
- 25 Spink A J, Tegelenbosch R A J, Buma M O S, et al. The EthoVision video tracking system—A tool for behavioral phenotyping of transgenic mice. *Physiol Behav*, 2001, 73: 731–744
- 26 Branson K, Robie A A, Bender J, et al. High-throughput ethomics in large groups of *Drosophila*. *Nat Methods*, 2009, 6: 451–457
- 27 de Chaumont F, Coura R D S, Serreau P, et al. Computerized video analysis of social interactions in mice. *Nat Methods*, 2012, 9: 410–417
- 28 Gershow M, Berck M, Mathew D, et al. Controlling airborne cues to study small animal navigation. *Nat Methods*, 2012, 9: 290–296
- 29 Swierczek N A, Giles A C, Rankin C H, et al. High-throughput behavioral analysis in *C. elegans*. *Nat Methods*, 2011, 8: 592–598
- 30 Straw A D, Branson K, Neumann T R, et al. Multi-camera real-time three-dimensional tracking of multiple flying animals. *J R Soc Interface*, 2011, 8: 395–409
- 31 Pérez-Escudero A, Vicente-Page J, Hinz R C, et al. idTracker: Tracking individuals in a group by automatic identification of unmarked animals. *Nat methods*, 2014, 11: 743–748
- 32 Schmidhuber J. Deep learning in neural networks: An overview. *Neural Netw*, 2015, 61: 85–117
- 33 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521: 436–444
- 34 Johansson G. Visual perception of biological motion and a model for its analysis. *Percept Psychophys*, 1973, 14: 201–211
- 35 Mathis A, Mamidanna P, Cury K M, et al. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci*, 2018, 21: 1281–1289
- 36 Pereira T D, Aldarondo D E, Willmore L, et al. Fast animal pose estimation using deep neural networks. *Nat Methods*, 2019, 16: 117–125
- 37 Graving J M, Chae D, Naik H, et al. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife*, 2019, 8: e47994
- 38 Günel S, Rhodin H, Morales D, et al. DeepFly3D, a deep learning-based approach for 3D limb and appendage tracking in tethered, adult *Drosophila*. *eLife*, 2019, 8: e48571
- 39 Arac A, Zhao P, Dobkin B H, et al. DeepBehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data. *Front Syst Neurosci*, 2019, 13: 20
- 40 Karashchuk P, Rupp K L, Dickinson E S, et al. Anipose: A toolkit for robust markerless 3D pose estimation. *bioRxiv*, 2020, doi: <https://doi.org/10.1101/2020.05.26.117325>
- 41 Yao Y, Jafarian Y, Park H S. Monet: Multiview semi-supervised keypoint detection via epipolar divergence. In: Curless B, Levoy M, Ikeuchi K, et al., eds. In: Proceedings of the IEEE International Conference on Computer Vision, Seoul, 2019. 753–762
- 42 Liu X, Yu S Y, Flierman N, et al. OptiFlex: Video-based animal pose estimation using deep learning enhanced by optical flow. *bioRxiv*, 2020, doi: <https://doi.org/10.1101/2020.04.04.025494>
- 43 Bala P C, Eisenreich B R, Yoo S B M, et al. Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. *Nat*

- Commun, 2020, 11: 1–12
- 44 Zimmermann C, Schneider A, Alyahyay M, et al. FreiPose: A deep learning framework for precise animal motion capture in 3D spaces. bioRxiv, 2020, doi: <https://doi.org/10.1101/2020.02.27.967620>
- 45 Labuguen R, Matsumoto J, Negrete S, et al. MacaquePose: A novel ‘in the wild’ macaque monkey pose dataset for markerless motion capture. bioRxiv, 2020, doi: <https://doi.org/10.1101/2020.07.30.229989>
- 46 Chen Z, Zhang R, Zhang Y E, et al. AlphaTracker: A multi-animal tracking and behavioral analysis tool. bioRxiv, 2020, doi: <https://doi.org/10.1101/2020.12.04.405159>
- 47 Mathis A, Biasi T, Schneider S, et al. Pretraining boosts out-of-domain robustness for pose estimation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2021. New York: IEEE, 2021. 1859–1868
- 48 Berger M, Agha N S, Gail A. Wireless recording from unrestrained monkeys reveals motor goal encoding beyond immediate reach in frontoparietal cortex. *eLife*, 2020, 9: e51322
- 49 Ebina T, Obara K, Watakabe A, et al. Arm movements induced by noninvasive optogenetic stimulation of the motor cortex in the common marmoset. *Proc Natl Acad Sci USA*, 2019, 116: 22844–22850
- 50 Han X, Laga H, Bennamoun M. Image-based 3D object reconstruction: State-of-the-art and trends in the deep learning era. In: Proceedings of the 2019 IEEE Transactions on Pattern Analysis and Machine Intelligence. New York: IEEE, 2019. 1578–1604
- 51 Badger M, Wang Y, Modh A, et al. 3D bird reconstruction: A dataset, model, and shape recovery from a single view. In: Proceedings of the 2020 European Conference on Computer Vision. Berlin: Springer, 2020. 1–17
- 52 Biggs B, Boyne O, Charles J, et al. Who left the dogs out? 3D animal reconstruction with expectation maximization in the loop. In: Proceedings of the 2020 European Conference on Computer Vision. Berlin: Springer, 2020. 195–211
- 53 Kulkarni N, Gupta A, Fouhey D F, et al. Articulation-aware canonical surface mapping. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2020. 449–458
- 54 Zuffi S, Kanazawa A, Black M J. Lions and tigers and bears: Capturing non-rigid, 3D, articulated shape from images. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018. 3955–3963
- 55 Biggs B, Roddick T, Fitzgibbon A, et al. Creatures great and small: Recovering the shape and motion of animals from video. In: Proceedings of the 2018 Asian Conference on Computer Vision, Australia. Berlin: Springer, 2018. 3–19
- 56 Kim A. Single-view 3D reconstruction of animals. Doctor Dissertation. Maryland: University of Maryland, 2017
- 57 Kanazawa A, Kovalsky S, Basri R, et al. Learning 3D deformation of animals from 2D images. *Comput Graph Forum*, 2016, 35: 365–374
- 58 Marshall J D, Aldarondo D E, Dunn T W, et al. Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire. *Neuron*, 2021, 109: 420–437.e8
- 59 Pareek P, Thakkar A. A survey on video-based human action recognition: Recent updates, datasets, challenges, and applications. *Artif Intell Rev*, 2021, 54: 2259–2322
- 60 Brown A E X, de Bivort B. Ethology as a physical science. *Nat Phys*, 2018, 14: 653–657
- 61 Gosztolai A, Günel S, Abrate M P, et al. LiftPose3D, a deep learning-based approach for transforming 2D to 3D pose in laboratory animals. bioRxiv, 2020, doi: <https://doi.org/10.1101/2020.09.18.292680>
- 62 Stephens G J, Johnson-Kerner B, Bialek W, et al. Dimensionality and dynamics in the behavior of *C. elegans*. *PLoS Comput Biol*, 2008, 4: e1000028
- 63 Rousseau J B I, Van Lochem P B A, Gispén W H, et al. Classification of rat behavior with an image-processing method and a neural network. *Behav Res Methods Instr Comput*, 2000, 32: 63–71
- 64 Jhuang H, Garrote E, Mutch J, et al. Automated home-cage behavioural phenotyping of mice. *Nat Commun*, 2010, 1: 1
- 65 Burgos-Artizzu X P, Dollár P, Lin D, et al. Social behavior recognition in continuous video. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2012. 1322–1329
- 66 Kabra M, Robie A A, Rivera-Alba M, et al. JAABA: Interactive machine learning for automatic annotation of animal behavior. *Nat Methods*, 2013, 10: 64–67
- 67 Wiltshko A B, Johnson M J, Iurilli G, et al. Mapping sub-second structure in mouse behavior. *Neuron*, 2015, 88: 1121–1135
- 68 Mearns D S, Donovan J C, Fernandes A M, et al. Deconstructing hunting behavior reveals a tightly coupled stimulus-response loop. *Curr Biol*, 2020, 30: 54–69.e9
- 69 Cao B, Zheng L, Zhang C, et al. Deepmood: Modeling mobile phone typing dynamics for mood detection. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017. New York: Association for Computing Machinery, 2017. 747–755
- 70 Hammerla N, Fisher J, Andras P, et al. PD disease state assessment in naturalistic environments using deep learning. In: Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2015, 29. 1742–1748

- 71 Goodwin M S, Intille S S, Albinali F, et al. Automated detection of stereotypical motor movements. *J Autism Dev Disord*, 2011, 41: 770–782
- 72 Westeyn T, Vadas K, Bian X, et al. Recognizing mimicked autistic self-stimulatory behaviors using HMMs. In: Ninth IEEE International Symposium on Wearable Computers (ISWC'05). Osaka: IEEE, 2005. 164–167
- 73 Bishay M, Palasek P, Priebe S, et al. Schinet: Automatic estimation of symptoms of schizophrenia from facial behaviour analysis. *IEEE Trans Affect Comput*, 2019, 99: 1
- 74 Xian Y, Lampert C H, Schiele B, et al. Zero-shot learning—A comprehensive evaluation of the good, the bad and the ugly. *IEEE Trans Pattern Anal Mach Intell*, 2018, 41: 2251–2265
- 75 Caruana R. Learning many related tasks at the same time with backpropagation. In: *Proceedings of the 7th International Conference on Neural Information Processing Systems*. Colorado: MIT Press, 1995. 657–664
- 76 Jaiswal A, Babu A R, Zadeh M Z, et al. A survey on contrastive self-supervised learning. *Technologies*, 2021, 9: 2
- 77 Wang X, Zhang R, Shen C, et al. Dense contrastive learning for self-supervised visual pre-training. In: *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New York: IEEE, 2021
- 78 Xie E, Ding J, Wang W, et al. DetCo: Unsupervised contrastive learning for object detection. 2021, arXiv: 2102.04803
- 79 Yang C, Wu Z, Zhou B, et al. Instance localization for self-supervised detection pretraining. 2021, arXiv: 2102.08318
- 80 Liu S, Li Z, Sun J. Self-EMD: Self-supervised object detection without ImageNet. 2021, arXiv: 2011.13677
- 81 Fangbemi A S, Lu Y F, Xu M Y, et al. ZooBuilder: 2D and 3D pose estimation for quadrupeds using synthetic data. 2021, arXiv: 2009.05389
- 82 Kane G A, Lopes G, Saunders J L, et al. Real-time, low-latency closed-loop feedback using markerless posture tracking. *eLife*, 2020, 9: e61909
- 83 Alex D S, Wahi A. BSFD: Background subtraction frame difference algorithm for moving object detection and extraction. *J Theor Appl Inf Technol*, 2014, 60: 623–628
- 84 Sample A P, Smith J R. The wireless identification and sensing platform. In: Sample A P, Smith J R, eds. *Wirelessly Powered Sensor Networks and Computational RFID*. New York: Springer, 2013. 33–56
- 85 Elsworth S, Güttel S. Time series forecasting using LSTM networks: A symbolic approach. 2021, arXiv: 2003.05672
- 86 Terzi M, Cenedese A, Susto G A. A multivariate symbolic approach to activity recognition for wearable applications. *IFAC-PapersOnLine*, 2017, 50: 15865–15870
- 87 Jennings C G, Landman R, Zhou Y, et al. Opportunities and challenges in modeling human brain disorders in transgenic primates. *Nat Neurosci*, 2016, 19: 1123–1130
- 88 Zhao H, Jiang Y H, Zhang Y Q. Modeling autism in non-human primates: Opportunities and challenges. *Autism Res*, 2018, 11: 686–694

Summary for “计算行为学研究进展”

Progress on Computational Ethology

Wei Ren^{1,3,4}, Shan Yu^{2,3,5*} & Yong Q. Zhang^{1,3,4*}

¹ State Key Laboratory for Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China;

² Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;

³ Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China;

⁴ College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;

⁵ School of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China

* Corresponding authors, E-mail: shan.yu@nlpr.ia.ac.cn; yqzhang@genetics.ac.cn

Animal behavior, mostly controlled by the central brain, has been studied in natural environments and controlled laboratory settings. In the early 20th century, researchers studied behavior in natural environments to reveal how it is built from components and organized over time in response to stimuli. In the laboratory settings, researchers study the ability of brain to generate behaviors in response to rewards and punishments. However, there are limitations in quantifying animal behaviors in these two approaches. Recent advances and breakthroughs in computer science provide an important opportunity for overcoming these limitations of behavioral studies.

In this review, we focus on an emerging new discipline called “Computational Ethology”, which uses a wide variety of techniques, including computer vision and machine learning, to measure and analyze the patterns of animal behaviors. Computational Ethology allows quantitative analyses of animal behaviors with high efficiency, and has made significant progress in recent years towards a better understanding of behaviors as well as their underlying neural mechanism.

Over the last decade, with the application of artificial intelligence in animal behavior study, numerous methods have been developed to automatically quantify animal behavior, including automatic tracking of movements. Classical computer vision methods estimate centroids and ellipses of animals, which reflect the locomotion and orientation, respectively. This estimation was then extended to multiple animals. Nevertheless, non-locomotion information of animal pose cannot be reliably captured. To solve this problem, deep learning-based pose estimation of behaviors in the laboratory setting has been developed for various animal species.

In this review, we present a pipeline of collecting and analyzing behavioral data. High-dimensional raw behavioral data are first subjected to dimension reduction. The low-dimensional data are then segmented and analyzed by supervised classification or unsupervised clustering in order to produce behavioral modules. After this, transition probabilities between behavioral modules over time are calculated to elucidate the pattern of behaviors. We also review the application of artificial intelligence to analyze behaviors for diagnosing and evaluating neuropsychiatric diseases and discuss the opportunities and challenges in Computational Ethology.

Computational Ethology, a result of substantial interdisciplinary research of neuroscience, psychology, physics, computer science, and ethology has great potential towards a deeper understanding of the nature of animal behaviors. Nevertheless, it is still in its infancy and many questions remain to be explored. This review provides a summary as well as a reference resource for this new yet rapidly advancing discipline. We hope that this review will be informative and useful for a wide interdisciplinary scientific community studying animal behaviors and artificial intelligence.

animal behavior, artificial intelligence, Computational Ethology, pose estimation, machine learning

doi: [10.1360/TB-2021-0213](https://doi.org/10.1360/TB-2021-0213)