

基于文献计量分析的动物遗传育种与AI交叉领域研究态势

王宁¹, 张博², 张浩^{2*}

(1. 中国农业大学图书馆, 北京 100193; 2. 中国农业大学动物科学技术学院, 北京 100193)

摘要: 本文探究了动物遗传育种与人工智能(Artificial Intelligence, AI)交叉领域的研究进展, 旨在揭示该领域学术研究与技术应用的研究热点与前沿态势, 为后续研究提供参考。研究通过文献计量分析方法, 整合SCIE论文和incoPat专利数据, 从年度分布、机构合作网络及主题演化等维度系统分析趋势, 并利用VOSviewer、CiteSpace和ITGInsight构建知识图谱识别热点布局。结果表明, 2019年为动物遗传育种与AI领域发展的转折点, 论文与专利数量均进入快速增长阶段。中国在论文和专利总量上居全球首位, 但合作模式呈现差异: 论文领域形成以中国、美国、巴西为主的国际合作网络, 而专利合作以本土机构为主, 跨国协作薄弱。研究主题上, 论文集中于机器学习与深度学习等模型在性状预测、遗传机制解析与基因组选育等方向的优化与创新; 专利则侧重于运用AI技术开发面向畜禽全生命周期的智能化管理系统。尽管AI技术显著提升了动物遗传育种的精准性与效率, 但其应用仍受限于模型泛化能力、数据标准化与硬件兼容性等瓶颈。未来需加强多源数据整合、轻量化算法开发及跨场景技术融合, 推动智能育种系统向全流程自动化发展, 为畜牧业可持续发展提供技术支撑。

关键词: 文献计量; 动物遗传育种; 人工智能; 研究热点与前沿

中图分类号: S126

文献标识码: A

DOI编号: 10.19556/j.0258-7033.20250603-02

动物遗传育种技术的持续创新对于应对2050年全球动物蛋白需求增长70%的挑战至关重要^[1]。人工智能(Artificial Intelligence, AI)凭借其在高维基因组数据解析与非加性遗传效应捕捉方面的优势, 已成为新一代育种体系的关键支撑。研究表明, 机器学习可提升基因组预测准确率19%以上^[2]; 随机森林与支持向量机可提升关键基因标记筛选效率^[3]; 多组学融合AI技术可增强基因组-环境互作表型预测模型的精度^[4]; AI亦在疾病识别与监测方面展现出潜力^[5]。总体而言, AI已突破性能测定、基因挖掘、育种值评估与配种优化等瓶颈, 可推动“生物技术+AI+大数据+信息技术”驱动的“育种4.0”时代的到来。

随着AI应用的不断深化, 动物遗传育种领域的研

收稿日期: 2025-06-03; 修回日期: 2025-08-25

资助项目: 国家重点研发计划(2023YFD1300401)和国家现代农业产业技术体系(CARS-40)

作者简介: 王宁(1990-), 女, 山东菏泽人, 硕士, 主要从事文献计量、知识产权服务相关研究, E-mail: zoe@cau.edu.cn

*通讯作者: 张浩(1968-), 男, 安徽萧县人, 博士, 教授, 主要从事动物遗传育种相关研究, E-mail: hzhang@cau.edu.cn

究成果与视角日益多元。为助力科研工作者把握学科前沿与发展趋势, 为理论与实践提供决策依据, 亟需对该领域的研究热点与前沿进行梳理。文献计量学通过量化论文主题分布、合作网络及演化轨迹^[6-9], 结合专利技术属性与时空分布分析^[10-11], 能客观描绘学术研究与技术布局的全景。基于此, 本文整合SCIE论文与incoPat专利数据, 通过知识图谱构建, 从时间、机构及主题演化等多方面剖析交叉领域的发展态势, 旨在明确研究前沿, 优化项目布局, 并为该领域的持续深化提供阶段性总结与决策参考。

1 数据来源与研究方法

1.1 数据来源 本研究数据来源于Web of Science核心合集SCIE(Science Citation Index Expanded)的论文及incoPat全球专利数据库, 聚焦动物遗传育种与AI交叉领域, 通过专业检索策略与专家筛选, 截至2025年5月1日, 共获取文献408篇、专利115件。

1.2 研究方法 本研究采用文献计量学与科学知识图谱分析方法, 系统构建科技论文与专利数据的分析框

架。运用美国德雷塞尔大学陈超美教授团队研发的 CiteSpace，其基于引文空间算法可有效探测学科演进路径与前沿趋势^[12]；采用荷兰莱顿大学 Eck 教授团队开发的 VOSviewer，该工具在共现网络构建与可视化方面具有显著优势^[13]；同时选用 ITGInsight 专业科技文本挖掘系统进行专利情报分析^[14]。

2 结 果

本文首先对论文与专利的年度产出趋势及主要国家和机构的分布进行量化分析，以勾勒该交叉领域的基本特征；随后，以论文被引频次与专利价值度为指标，评估文献的关注度与影响力，揭示研究热点；接着，构建了国家和机构合作网络，客观呈现各参与主体的学术地位与协作模式；最后，绘制出主题演化与前沿态势图谱，深入剖析核心学术议题与技术布局，归纳总结未来研究方向。

2.1 总体趋势 根据 SCIE 和 incoPat 检索结果，自 2019 年起，论文与专利申请量快速增长，标志着动物遗传育种与 AI 技术的交叉已进入快速发展阶段。从图 1A 可以看出，发文量年度趋势可分为 2 个阶段：2015—2018 年为学科交叉的萌芽期，年发表论文 13~20 篇；自 2019 年起进入技术融合的爆发期，论文量迅速攀升，年均增速达 7.67%。2018 年，美国科学院院士、玉米遗传学家 Edward Buckler 教授首次提出“育种 4.0”——以分子设计育种为基础，融合 AI 等技术，实现育种数据的高效采集、存储与分析，为作物育种决策提供智能支持并推动精准高效发展^[15]。该概念发布后，动植物智能育种研究热度大幅提升，2019 年发文量显著激增也充分验证了这一趋势（2025 年数据因完整性未纳入分析）。

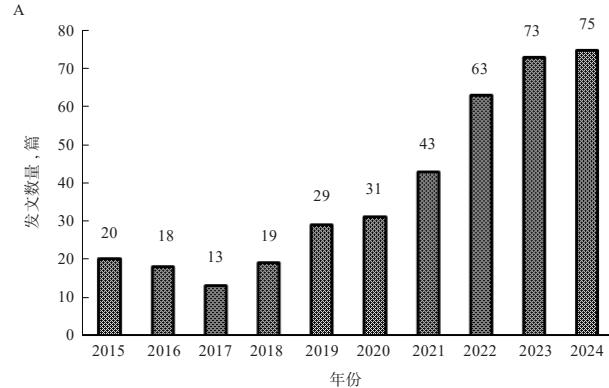
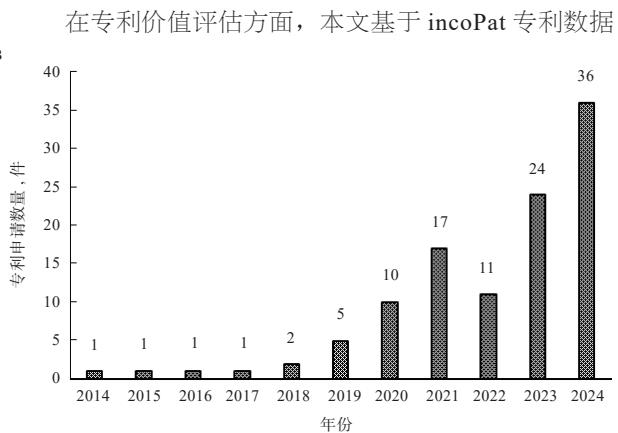


图 1 论文发表趋势 (A) 与专利申请趋势 (B)

专利申请趋势亦呈现类似态势（图 1B）：2014—2018 年年均仅有 1~2 项申请，技术储备处于起步阶段；自 2019 年起申请量迅速攀升，截至 2024 年达 36 件，比 2019 年增长近 7 倍，年均增幅 6.20%。尽管 2022 年出现短期波动，但整体专利增长势头持续，表明 AI 技术在动物遗传育种中的应用效能正在逐步显现。

2.2 主要国家与机构 中国已在动物遗传育种与 AI 交叉领域实现科研与专利“双领先”，尤其以中国农业科学院与中国农业大学的产出更为突出。论文方面（图 2A），中国以 104 篇位居全球首位，美国 71 篇位列第二，其余国家均不足 40 篇；机构层面（图 2B），中国农业科学院（24 篇）、中国农业大学（22 篇）和中国农业农村部（20 篇）位列全球发文量前三，且为 TOP10 机构中发文量突破 20 篇的单位。美国则有威斯康星大学麦迪逊分校（18 篇）、乔治亚大学（10 篇）和美国农业部（10 篇）入围前十。专利申请方面（图 2C、2D），中国贡献 87 件，占比 73.04%；前 9 位申请主体悉数来自中国，其中中国农业大学（9 件）与中国农业科学院北京畜牧兽医研究所（5 件）位居前列，且 88.89% 的机构为高校及科研院所。相比之下，美国虽在论文发表上排名第 2 位，但专利布局相对薄弱。

2.3 论文被引频次与专利价值度 被引频次作为衡量学术论文质量与影响力的重要指标，不仅在学术共同体中具有广泛的认同，也是评估学者、研究机构及学术期刊影响力的重要工具^[16]。表 1 统计显示，被引频次 TOP10 文献集中于机器学习与深度学习模型在基因组育种预测中的应用，同时涉及遗传关联解析、动物行为识别及体重预测模型构建等主题，反映出该领域研究热点的集中指向。



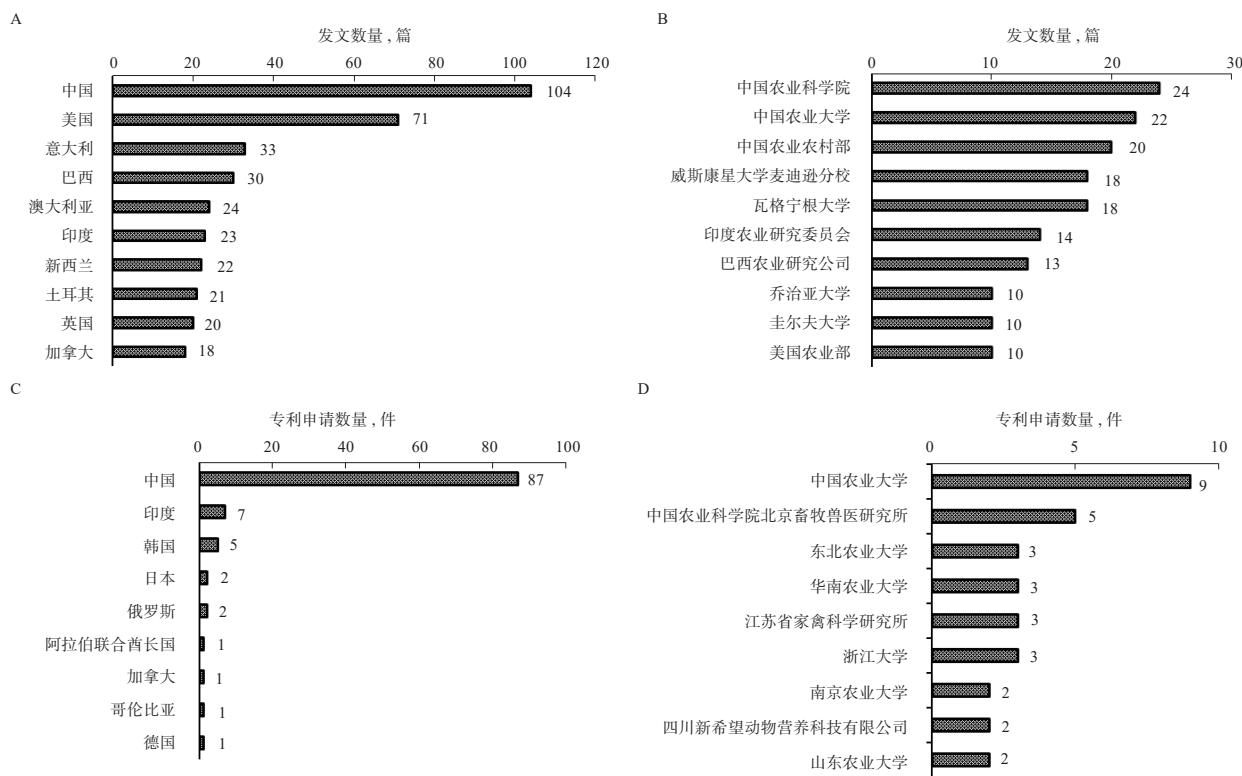


图 2 论文发文量 TOP10 国家 (A)、机构 (B) 与专利申请 TOP9 国家 (C)、机构 (D)

表 1 论文被引频次 TOP10 情况

序号	论文标题	年份	被引频次, 次
1	Genomic prediction of breeding values using a subset of snps identified by three machine learning methods 基于三种机器学习方法筛选的 SNP 子集进行基因组育种值预测	2018 年	139
2	Benchmarking parametric and machine learning models for genomic prediction of complex traits 复杂性状基因组预测中参数模型与机器学习模型的性能对比	2019 年	113
3	Ancestral relationships using metafounders: finite ancestral populations and across population relationships 基于元祖先的遗传关系解析：有限祖先群体与跨群体亲缘关系	2015 年	109
4	Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes 深度学习与参数模型及集成方法在复杂表型基因组预测中的对比	2020 年	108
5	Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning 猪牛行为识别：从计算机视觉到深度学习的演进	2021 年	105
6	Weight prediction of broiler chickens using 3D computer vision 基于三维计算机视觉的肉鸡体重预测	2016 年	99
7	Development of an early warning algorithm to detect sick broilers 肉鸡疾病预警检测算法的开发	2018 年	93
8	A review of deep learning algorithms for computer vision systems in livestock 畜禽计算机视觉系统中的深度学习算法综述	2021 年	73
9	A 100-Year Review: Methods and impact of genetic selection in dairy cattle-From daughter-dam comparisons to deep learning algorithms 百年综述：奶牛遗传选择方法与影响——从母女对比到深度学习算法	2017 年	70
10	Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle 反向传播神经网络在荷斯坦牛和德系西门塔尔牛复杂性状基因组预测中的应用	2015 年	64

构建的多维度评价体系，依托专利类型、同族数量、引证频次、权利要求范围等 20 项技术指标，通过加权算法模型对专利技术价值进行量化，价值度 ≥ 7 的专利

界定为高价值专利^[17]。图 3 所示价值度 ≥ 9 的专利聚焦于结合全基因组信息与机器学习、深度视觉与行为监测、自适应优化算法，构建集基因组选配、分子标记鉴别、

表型监控及饲料与近交优化于一体的智能化动物育种系统。从申请主体看，高校与科研机构为主要发明人，企业专利多以合作方式申请；从时间分布看，2023—2024年高价值度专利达8项，2023年以前仅3项，表明近2年内新申请专利的技术价值显著提升。

2.4 合作情况

2.4.1 论文国际及机构合作情况 基于 VOSviewer 构建了科研合作网络图谱（图 3），其中节点代表发文量，连边宽度代表合作强度，节点颜色代表聚类得到的作者所属群^[18]。国家合作的网络结构呈现 3 大集群（图 3A）：集群 1 以中国（发文量最多）和美国为核心，包括澳大利亚、加拿大及巴西；集群 2 以意大利为主导，涵盖德国、法国、英国等欧盟国家；集群 3 以荷兰与印度为主体，包括西班牙、挪威等国。其中，中国、美国

表 2 价值度达到 9 分及以上的专利情况

序号	标题	申请人	申请日
1	一种基于全基因组 SNP 信息的华西牛基因组选配方法及应用	中国农业科学院北京畜牧兽医研究所	2024 年 7 月 9 日
2	一种基于 GCS-YOLO 算法的牛场多目标检测方法及系统	山东农业大学	2024 年 4 月 3 日
3	基于混合遗传算法优化牛饲料配比的方法	华南农业大学	2024 年 3 月 21 日
4	基于运动传感器的家禽活动行为监测方法与系统	江苏省家禽科学研究所；和盛食品集团有限公司	2024 年 3 月 20 日
5	基于 XGBoost 模型的 20 个藏羊品种分子鉴别的 SNP 位点组合及应用	中国农业科学院北京畜牧兽医研究所	2023 年 11 月 20 日
6	一种生猪养殖场多源数据异常检测与不良数据修正方法	山东农业大学	2023 年 10 月 20 日
7	一种基于卷积神经网络预测哺乳动物家畜增强子序列的装置、方法及计算机可读存储介质	中国农业大学	2023 年 9 月 8 日
8	基于叠层饲养模式下鸡异常粪便的疾病早期预警溯源系统	中国农业大学	2023 年 4 月 27 日
9	一种基于计算机视觉的母猪哺乳行为细粒度识别方法	南京农业大学	2022 年 9 月 15 日
10	一种基于指派问题解法控制家禽近交的方法	中国农业大学	2021 年 4 月 25 日
11	提高黄羽肉种鸡合格种蛋数的选育方法、模型及选育系统	江苏兴牧农业科技有限公司；南京农业大学；江苏立华牧业股份有限公司	2021 年 2 月 25 日

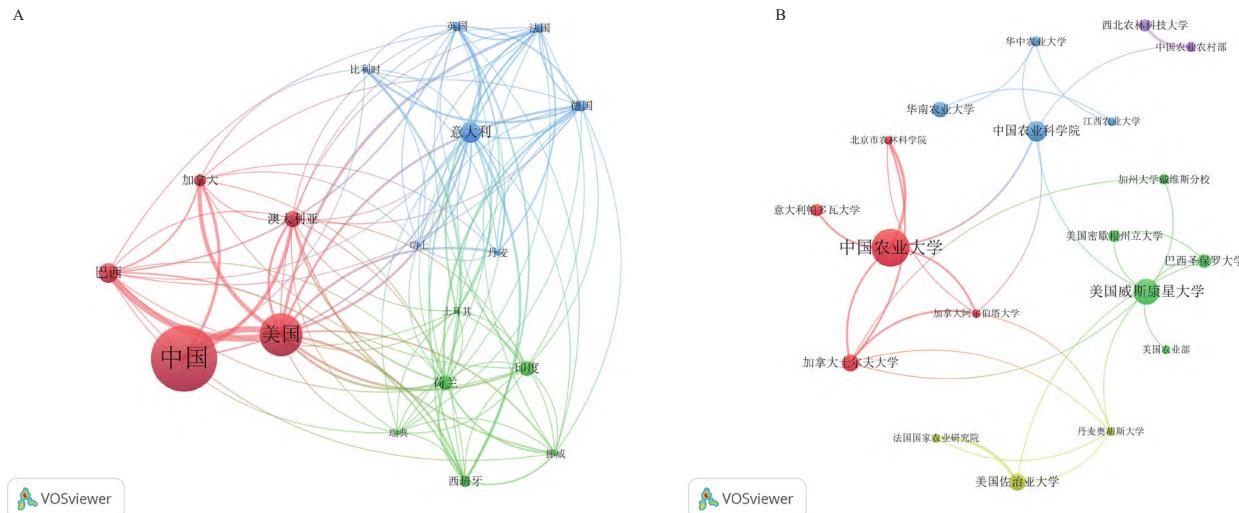


图 3 论文国家合作（A）和机构合作（B）情况

和巴西 3 个国家相互间合作最为紧密。

在机构合作层面（图 3B），中国农业大学、中国农业科学院及美国威斯康星大学构成发文量前 3 位的研究主体。中国农业大学除与北京市农林科学院保持本土合作外，还与意大利帕多瓦大学、加拿大圭尔夫大学及阿尔伯塔大学形成跨国合作体系。数据分析表明，上述 3 所机构在学术产出规模及国际合作广度方面均处于领先地位。

使用 ITGInsight 软件对专利数据进行聚类分析，构建发明人与专利权人合作图谱（图 4），其中节点代表发明人专利产出规模，连边宽度代表合作频次，不同类团代表合作群体。图 4A 显示，发明人合作网络形成 7 类群体。其中，包含丁露雨等 6 位发明人的类团，隶属于北京市农林科学院信息技术研究中心；涵盖李俊雅

等5位发明人的类团，主要来自中国农业科学院北京畜牧兽医研究所。图4B专利权人合作网络分析表明，国内机构间合作呈现显著地域封闭性。中国农业大学与中国农业科学院北京畜牧兽医研究所专利申请量虽居前，但未形成合作关系。国内有效合作仅见于新希望集团内部，尚未延伸至外部企业或高校。而跨国合作群体由Anita Chaturvedi等6位国际专利权人构成，显示出海外机构在技术研发中的协同创新特征。

2.5 研究主题与前沿

2.5.1 论文研究主题 关键词共现分析作为文献计量学重要方法，通过共现频次构建语义网络、高频词凸显研究焦点、共现强度映射学科主题关联性^[19]，对学术概念间的关联强度进行解析，进而揭示学科知识结构演化规律与研究热点分布特征^[10-21]。使用CiteSpace进行关键词聚类分析，一般认为聚类图谱中节点面积越大，热度越高；节点外圈颜色越深，中心性越高；热度与中心性越高，则研究主题越重要。

图5关键词聚类图谱显示，网络模块化指数Q值为0.496 (Q>0.3)，检引度指数S值达0.766 2 (S>0.7)，表明聚类结构显著、类团合理。研究识别出9个主题聚类，包括机器学习（Machine Learning）、肉质（Meat Quality）、基因组选择（Genomic Selection）、预测（Prediction）、AI、识别模式（Spotting Pattern）、基因组选配（Genomic Mating）、深度学习（Deep Learning）、遗传力（Heritability）。其中，机器学习、基因选择、基因预测、深度学习等节点兼具高热度与高中心性，构成领域重要的研究主题。

根据以上9个主题，收集相应文献并按引用次数

降序排列，采用人工研读标题与摘要、结合专家评议标注研究方向并参照正文内容，提炼各主题下的主要研究方向。

机器学习：该主题聚焦算法驱动的育种范式转型。采用集成学习（XGBoost、SVM）与深度学习（CNN）模型突破传统线性方法对非线性性状预测的限制^[22-23]，融合近红外光谱（NIRS）与傅里叶红外光谱（FTIR）技术实现基因型无创鉴别与生理指标解析^[24-25]，并通过超低密度SNP芯片（20~50个标记）同时满足遗传多样性保护与精准个体鉴定需求^[26-27]。

肉质：该主题聚焦多模态数据融合的性状评价体系。结合近红外光谱与化学计量学对肉质成分进行无损检测^[25,28-29]，基因组选择模型优化遗传育种值估算^[30]，应用结构方程模型解析脂肪酸代谢对大理石花纹的遗传调控^[31]，并利用CHAID决策树筛选关键生产变量构建分类模型^[32]，实现客观化、标准化的肉质育种评价框架。

基因组选择：该主题聚焦复杂性状解析的技术革新。通过深度残差网络（ResNet）捕捉非加性遗传效应，显著提升预测精度^[33-34]，结合群体分层与表型方差控制的训练集优化策略增强模型的普适性^[35-36]，利用Elastic-Net算法在低密度芯片上精准筛选标记并填补基因型数据^[37]，进而在多物种基因组交配设计中于控制近交率的前提下最大化遗传增益^[38-39]。

预测：该主题聚焦多维数据赋能的遗传评估。该主题下建立基因组-表型-环境多元预测体系，例如线性模型（GBLUP）与非线性算法（随机森林/神经网络）协同优化基因组关系矩阵^[40]；径向基函数网络整合繁殖记录构建表型预测模型^[41]；基因型填补技术

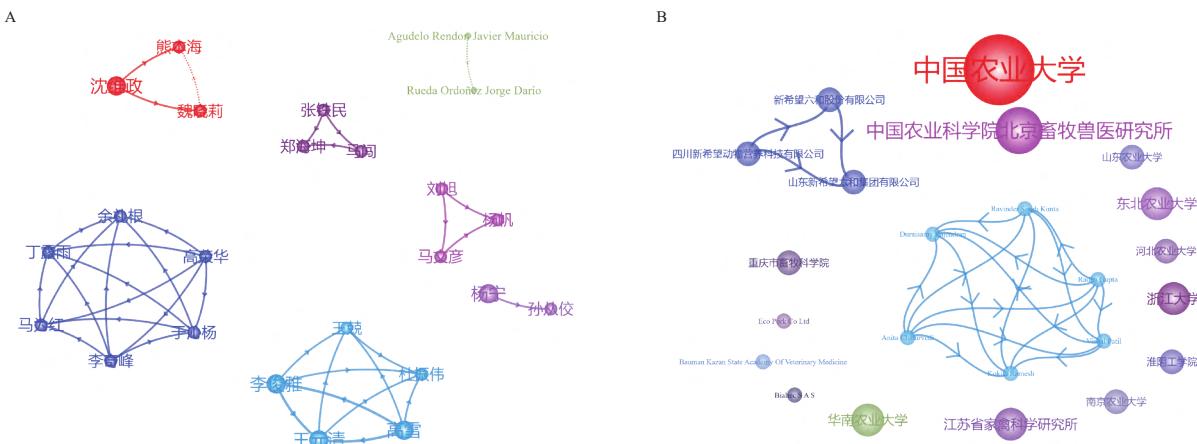


图4 发明人合作（A）和专利权人合作（B）情况

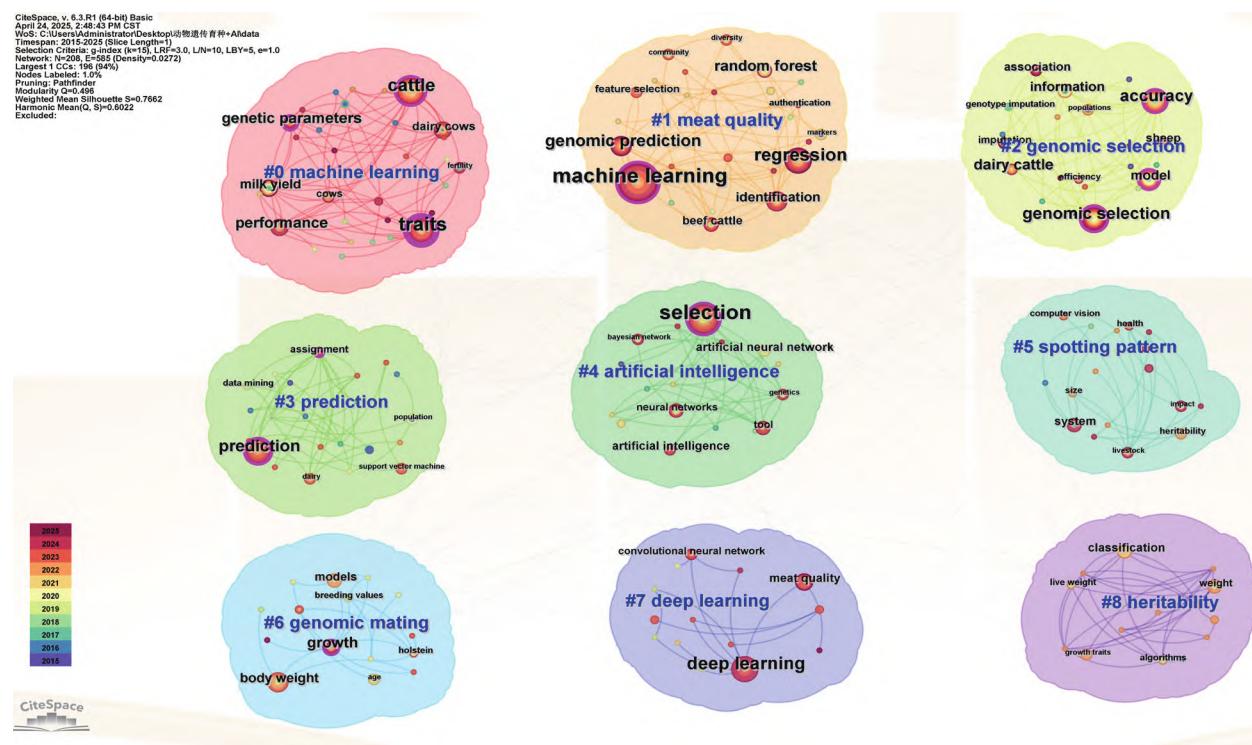


图5 论文主题聚类

(fastPHASE)降低全基因组分型成本。该主题的研究通过高维基因组数据与智能算法的深度融合，有效增强了复杂性状的预测可靠性，为精准育种提供了方法论支持^[42]。

AI: 该主题聚焦场景化技术集成创新。通过AI技术构建覆盖育种全流程的技术矩阵，如随机森林/XGBoost优化关键性状预测模型^[43-45]；YOLOv5架构实现动物表型实时量化；物联网传感数据驱动繁殖管理智能决策^[46-47]；声纹识别突破性别鉴定技术瓶颈^[48]。其创新价值在于打通“数据采集-算法解析-场域应用”技术闭环，但硬件适配性与小样本建模仍是产业化落地关键障碍^[48-49]。

识别模式：该研究主题聚焦遗传互作网络的量化解析。通过遗传信息识别，揭示遗传-表型-环境三元互作规律，例如遗传算法量化代谢性状与繁殖性能的遗传相关性^[50]；结构方程模型解析初产年龄的遗传拮抗机制^[51]；传感器网络支持发情动态监测与预警^[48]等。

基因组选配：即以SNP/ROH矩阵为基础的最优贡献选择(GOCS)算法，实现遗传增益(ΔG)与近交率(ΔF)的协同控制^[52]，采用显性效应模型(HIBLUP)提升生产性能^[53]，并通过多世代模拟验证交配策略的长期稳健性^[40]，构建智能化遗传资源管控新范式。

深度学习：该主题聚焦非线性遗传机制的破译工具。通过深度学习方法实现复杂遗传效应的深度解析，例如改进YOLO架构，结合三维医学影像/光谱技术实现活体动物体重^[54]、生殖表型^[55]等非接触式精准预测；deepGBLUP模型提升非加性效应预测精度^[56-57]；LSTM网络构建热应激风险预警系统^[58]；染色体异常诊断模型实现SNP数据驱动决策^[59]。其技术优势在于捕捉显性/上位效应等非线性遗传规律，但受限于高质量标注数据获取成本及模型生物学可解释性等问题^[60-61]，技术普适性受到制约。

遗传力：该主题聚焦智能赋能的遗传参数体系。系列研究致力于改进复杂性状遗传力量化方法，比如贝叶斯-神经网络混合模型解析阈性状动态特征^[62]、全基因组标记整合提升遗传参数估计精度^[63]、环境互作分析建立甲烷排放遗传选择指标^[64]等。

2.5.2 论文研究前沿 作为学科趋势探测方法，高频主题的演化分析能够有效揭示领域主题的生命周期动态^[65]。通过高频主题词聚类分析，不仅能识别研究热点，更能为学科发展提供理论框架与方法论支撑。本文利用CiteSpace构建关键词突现图谱与时间序列图谱，系统呈现交叉领域研究热点的演进轨迹及其前沿态势。

一些关键词如“信息”“准确性”“育种值”“基

因参数”“性状”“随机森林”等长期保持热度，反映了领域内的核心议题。近年来，“人工智能”和“深度学习”等关键词迅速崛起，成为新的研究热点，表明动物遗传育种正向着智能化、精准化转变。

突现词时序分析表明，“计算机视觉”“支持向量机”“人工智能”“深度学习”等技术已形成技术簇群，其应用聚焦于肉质性状量化评估、多性状协同选择模型构建及基因组选择精度提升等关键领域。

2.5.3 专利研究主题 本文基于 ITGInsight 软件提取高频关键词（Top50）构建共现网络，通过模块化聚类算法生成主题聚类图谱（图 7）。专利领域的研究热点主题词主要聚集于神经网络、储存介质、遗传算法、计算机、AI、基因组等。其中以“存储介质”“神经网络”等为代表词的类团体现遗传算法与 AI 方法的融合应用，涵盖神经网络架构优化、图像特征提取及分布式存储介质设计等跨领域技术整合；以“人工智能”“畜牧业”代表词的类团映射畜牧业生产场景下的 AI 技术融合。针对这 2 个类群内的专利，通过人工研读标题和摘要并结合专家评议，对每项专利方向进行标注，最终提炼出各聚类的核心研究内容。

以“存储介质”“神经网络”等为代表词的类团（G06N3/08：学习方法）：该主题聚焦 AI 算法与跨学科技术融合，突破遗传评估与养殖监测技术瓶颈。例如遗传

算法优化深度视觉感知系统实现畜禽行为精准识别^[66-67]，时空序列-图神经网络模型解析多维遗传互作机制^[66,68]，卷积神经网络驱动基因组调控元件因果推断^[69]，构建“算法 - 数据 - 硬件”协同的智能化技术基座。

以“人工智能”“畜牧业”代表词的类团（G06Q50/02：农业）：该主题聚焦通过育种产业化构建全链条智能决策体系的相关内容，例如全基因组 SNP 深度挖掘支撑表型 - 基因型 - 环境多源联动选配优化^[70-72]，机器学习融合生物信息学实现疫病风险动态预测与防控策略自适应调整^[73]，智能算法与畜牧学原理深度耦合形成“基因筛选 - 群体管理”技术闭环，推动经验育种向数据驱动模式转型。

2.5.4 专利研究前沿趋势 基于 ITGInsight 软件构建动物遗传育种领域主题词演化网络，通过提取国内外高频主题词（Top50）揭示技术发展轨迹。网络节点代表主题词频，连线代表共现演化关系，主题词的变化路径反映同一主题的时序演进特征。图 8 显示，2016—2021 年技术内核聚焦神经网络、遗传算法与 AI 基础算法层；2022—2025 年研究边界向应用场景延伸，形成“计算机系统 - 储存介质 - 基因组技术 - 养殖场域 - 测定方法”的复合技术生态。

基于主题词频次统计构建的突破性识别主题（图 9）显示：神经网络、管理系统与遗传算法稳居持续性热

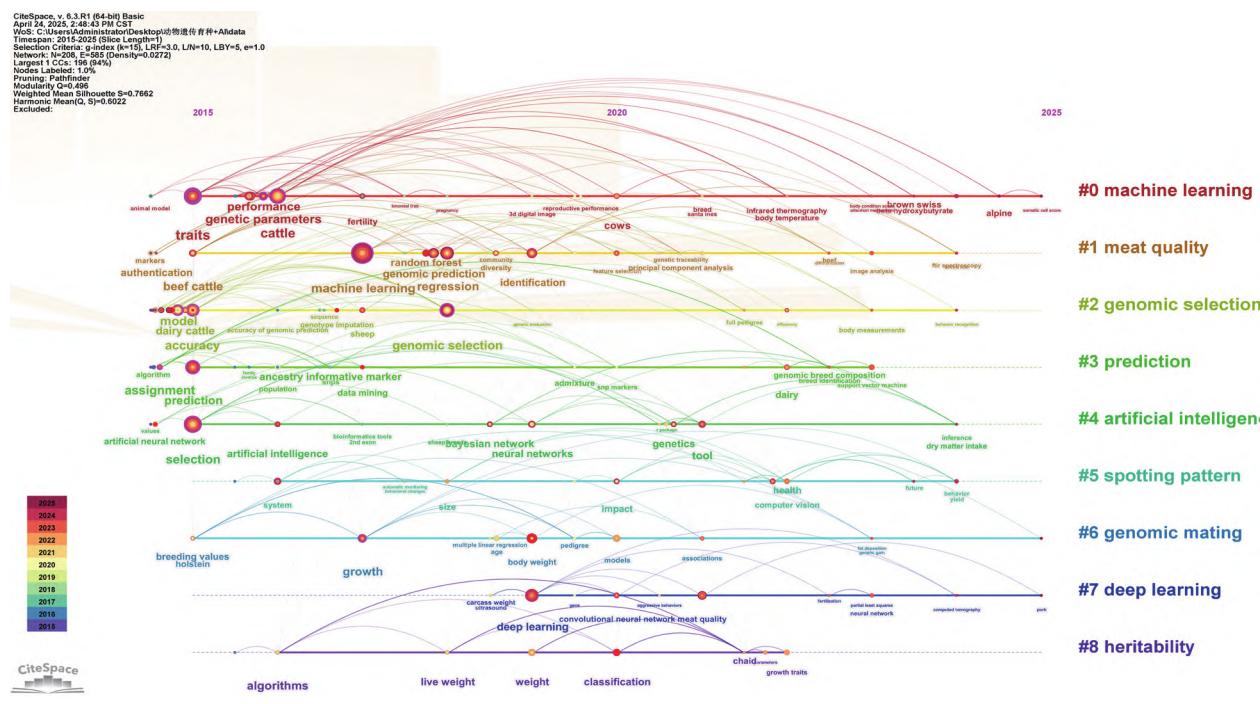


图 6 论文主题演进

表3 TOP21关键词突现情况

关键词	年份	突现强度	突现开始年份	突现结束年份	2015—2025年
information (信息)	2015年	3.09	2015年	2020年	—————
association (组合)	2015年	2.70	2015年	2017年	—————
accuracy (准确性)	2015年	2.65	2015年	2018年	—————
dairy cattle (奶牛)	2015年	2.19	2015年	2017年	—————
breeding value (育种值)	2015年	2.06	2015年	2020年	—————
algorithm (算法)	2015年	1.81	2015年	2016年	—————
growth (生长性状)	2017年	3.06	2017年	2019年	—————
genotype imputation (基因型估算)	2017年	3.04	2017年	2018年	—————
artificial neural network (人工神经网络)	2015年	2.64	2017年	2019年	—————
data mining (数据挖掘)	2017年	2.62	2017年	2019年	—————
model (模型)	2015年	2.11	2017年	2020年	—————
random forest (随机森林)	2018年	3.75	2018年	2021年	—————
cattle (牛)	2016年	2.31	2021年	2021年	—————
heritability (遗传性)	2022年	2.71	2022年	2022年	—————
body weight (体重)	2019年	2.57	2022年	2023年	—————
computer vision (计算机视觉)	2022年	1.90	2022年	2023年	—————
meat quality (肉质)	2021年	2.97	2023年	2023年	—————
feature selection (特征选择)	2020年	2.36	2023年	2023年	—————
support vector machine (支持向量机)	2023年	2.14	2023年	2023年	—————
artificial intelligence (人工智能)	2016年	1.99	2023年	2025年	—————
deep learning (深度学习)	2019年	3.08	2024年	2025年	—————

注：最后一列加粗黑色表示关键词在对应年份突现持续时间。

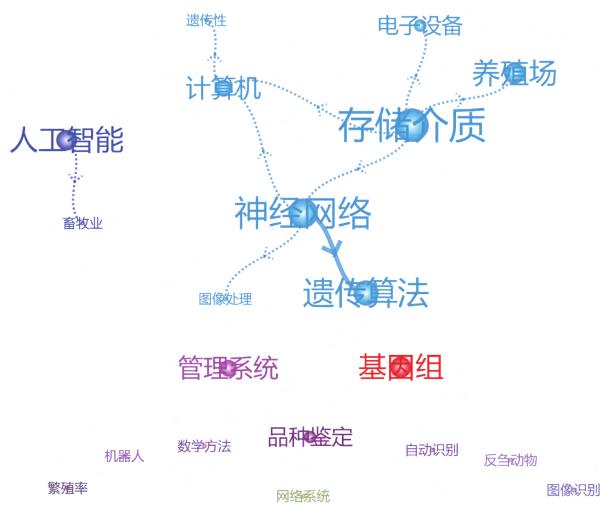


图7 专利主题词聚类

点；养殖场、AI与存储设备构成近期热点，反映应用场景的不断扩展；机器人、图像处理、预警系统及图像识别则成为新兴主题，标志着技术集成度的显著提升。

3 结论与建议

通过对动物遗传育种与AI交叉领域论文与专利的

多维度分析，本文得出以下结论：

1) 发文数量与专利申请方面：论文自2019年由萌芽期跃迁至爆发期，且中国机构贡献了全球主要论文产出；专利则表现为“厚积薄发”，2019年前处于技术储备阶段，随后因AI革新进入快速增长，中国申请量稳居全球首位；

2) 合作网络方面：论文合作方面，中国、美国、巴西形成紧密合作，但中国机构跨国合作范围有限，本土合作强度显著高于国际。专利合作情况，国内合作以机构内部或地域邻近为主，跨国合作仅由6位国际专利权人维系，缺乏系统性技术联盟；

3) 研究主题及前沿态势方面：论文侧重算法、模型的优化与创新，机器学习与深度学习在基因组选择与遗传参数预测等基础领域占据主导，并向计算机视觉和多模态数据融合等场景延伸；而专利开发则以技术集成为导向，早期聚焦神经网络与遗传算法等底层技术，近年已拓展至智能装备（如机器人、图像识别）与场景化系统（养殖场管理与预警平台）。

尽管AI技术在基因组选择、多性状预测和精准育

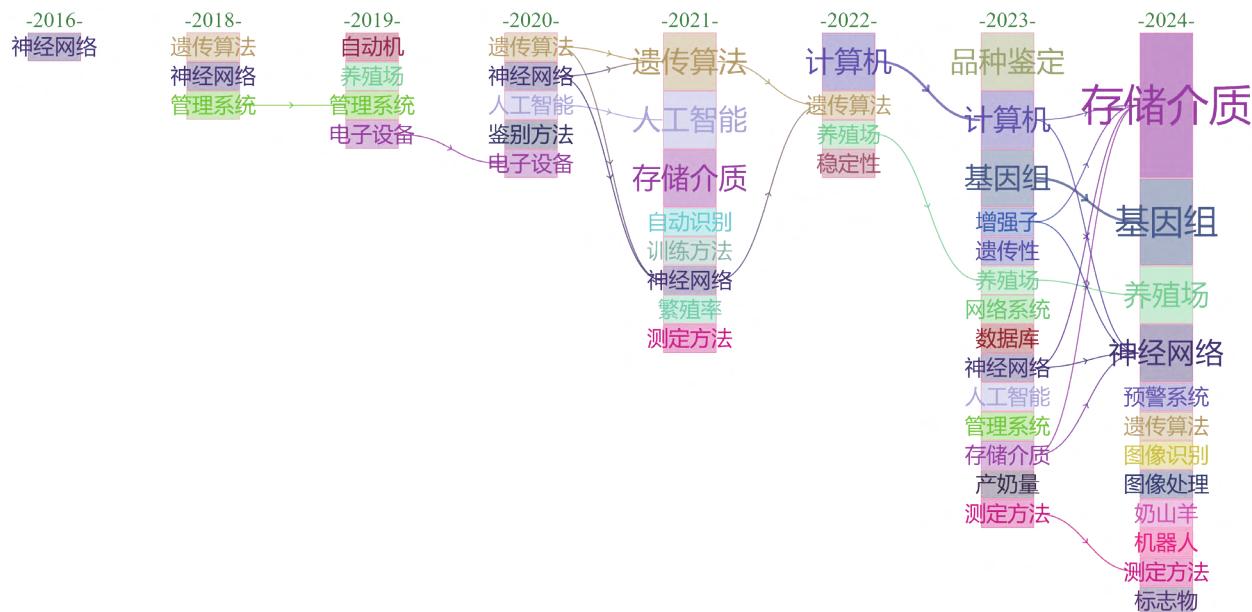


图8 专利主题词演进

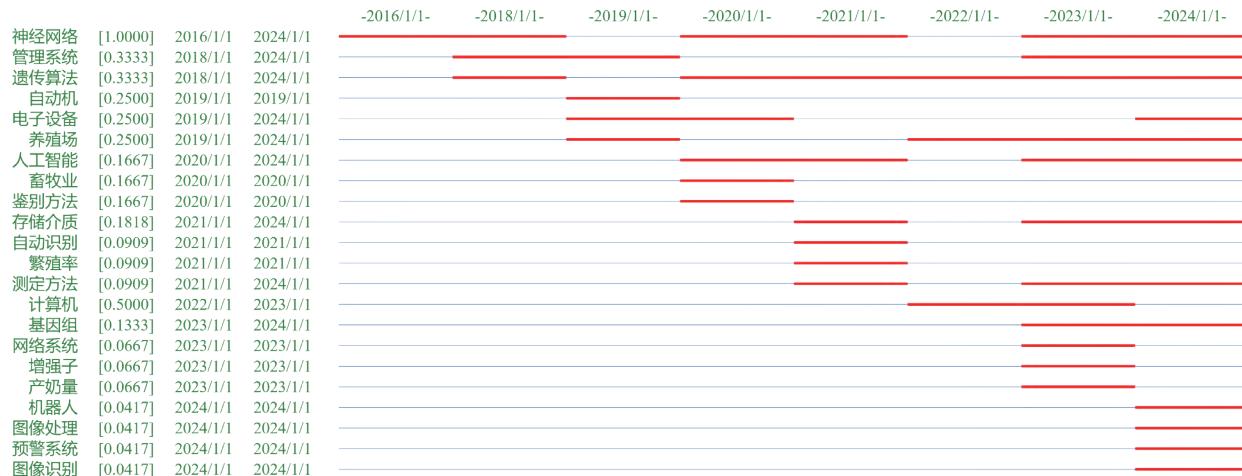


图9 专利主题词突现图谱

种设计等领域取得了显著成果，但在模型可解释性、跨群体泛化能力、小样本偏差以及数据标准化体系等方面仍存在不足^[74]。目前已有一些基于机器学习、深度学习和非接触式传感技术的多组学数据融合框架，但全流程自动化的智能育种系统尚未完善，硬件兼容性和成本效益也亟待提升。因此，未来研究应聚焦解决关键技术瓶颈，构建理论突破与产业化落地协同发展的创新生态。在方法层面，应重点深化模型可解释性和因果推断理论研究，通过揭示基因与表型关联的生物学内在逻辑，提高育种决策的科学可信度。同时，对于跨群体适用性不足和小样本偏差等问题，需要探索领域自适应学习、迁移学习及数据增强等策略，提升模型在不同遗传背景和

环境条件下的泛化能力。此外，为突破多源异构数据的整合瓶颈，亟需建立统一的表型本体框架和数据标准，构建开放共享的基准测试平台以促进技术迭代和方法验证。在技术转化层面，需要推动多组学数据与视觉、时空等信息的深度融合，研发兼具高预测精度和低计算负载的轻量化算法，以平衡模型性能和实际部署约束。同时应强化软硬件系统的兼容性设计，完善模块化测定与验证流程，并通过典型示范场景的试点，推动辅助工具升级为智能决策核心引擎，最终实现从算法创新到产业化应用的衔接。

总之，该领域的发展在于实现跨学科知识与工程化实践的深度结合，将基因组选择、精准育种设计与智能

化生产管理有机结合，为智慧畜牧业提供可持续的技术支撑和创新动能。

参考文献：

- [1] Mrode R, Tarekegn G M, Mwacharo J M, et al. Invited review: Genomic selection for small ruminants in developed countries: how applicable for the rest of the world? [J]. Animal, 2018, 12(7): 1333-1340.
- [2] Wang X, Shi S, Wang G, et al. Using machine learning to improve the accuracy of genomic prediction of reproduction traits in pigs[J]. J Anim Sci Biotechno, 2022, 13(1): 60-72.
- [3] Yousefzadeh-najafabadi M, Eskandari M, Torabi S, et al. Machine-learning-based genome-wide association studies for uncovering QTL underlying soybean yield and its components [J]. Int J Mol Sci, 2022, 23(10): 5538.
- [4] Shi Q, Chen W, Huang S, et al. Deep learning for mining protein data[J]. Brief Bioinform, 2021, 22(1): 194-218.
- [5] Uddin S, Khan A, Hossain M E, et al. Comparing different supervised machine learning algorithms for disease prediction [J]. BMC Med Inform Decis Mak, 2019, 19(1): 281.
- [6] 赵阳.面向企业技术创新的专利技术信息挖掘与分析研究[J].现代情报, 2020, 40(12): 84-89.
- [7] 冯仁涛.基于专利文献的专利维持时间影响因素分析[J].情报杂志, 2020, 39(7): 202-207.
- [8] 邱均平,余波,杨思洛.大数据背景下一门交叉学科的兴起——论数据计量学的构建[J].中国图书馆学报, 2021, 47(5): 48-58.
- [9] 廖政贸.国际科技期刊研究的特征与演进——基于WoS的文献计量分析(1980-2022)[J].中国科技期刊研究, 2024, 35(4): 547-558.
- [10] Hammarfelt B. Linking science to technology: The “patent paper citation” and the rise of patentometrics in the 1980s[J]. J Doc, 2021, 77(6): 1413-1429.
- [11] Xuan Y, Zhang S, Li X, et al. Identify cross-country knowledge flow and innovation trajectory: Insights from patent citation network analysis of 5G technology[J]. Technol Anal Strateg Manage, 2024, 36(10): 2997-3019.
- [12] Chen C. Searching for intellectual turning points: Progressive knowledge domain visualization[J]. Proc Natl Acad Sci USA, 2004, 101: 5303-5310.
- [13] van Eck N J, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping[J]. Scientometrics, 2010, 84(2): 523-538.
- [14] 刘玉琴,刘晶,张勇斌.中国图书情报领域专利研究的计量分析[J].情报工程, 2018, 4(6): 87-97.
- [15] 齐学礼,陈艳艳,王永霞,等.中国作物育种先进技术的研究现状与发展建议[J].分子植物育种, 2024: 1-11.
- [16] Waltman L. A review of the literature on citation impact indicators[J]. J Informetr, 2016, 10(2): 365-391.
- [17] 谢智敏,范晓波,郭倩玲.专利价值评估工具的有效性比较研究[J].现代情报, 2018, 38(4): 124-129.
- [18] 盖赢钊,邵子涵,周群.《中国农机化学报》2012—2021年文献计量与研究热点分析[J].中国农机化学报, 2023, 44(3): 214-222.
- [19] 闫伟娜.我国科普期刊研究的进展、热点与趋势——基于CiteSpace知识图谱的可视化分析[J].中国科技期刊研究, 2024, 35(2): 163-170.
- [20] 范少萍,李迎迎,张志强.国内外共词分析研究的文献计量分析[J].情报杂志, 2013, 32(9): 104-109.
- [21] 魏玉梅,滕广青.网络视域下领域重要关键词提取方法的比较研究[J].情报资料工作, 2020, 41(3): 97-104.
- [22] Pedrosa V B, Chen S Y, Gloria L S, et al. Machine learning methods for genomic prediction of cow behavioral traits measured by automatic milking systems in north American holstein cattle[J]. J Dairy Sci, 2024, 107(7): 4758-4771.
- [23] Xiang T, Li T, Li J, et al. Using machine learning to realize genetic site screening and genomic prediction of productive traits in pigs[J]. Faseb J, 2023, 37(6): e22961.
- [24] Reis W, Franca T, Calvani C, et al. Enhancing early identification of high-fertile cattle females using infrared blood serum spectra and machine learning[J]. Sci Rep, 2024, 14(1): 19446.
- [25] Serva L, Marchesini G, Cullere M, et al. Testing two NIRs instruments to predict chicken breast meat quality and exploiting machine learning approaches to discriminate among genotypes and presence of myopathies[J]. Food Control, 2023, 144: 109391.
- [26] Xu Z, Diao S, Teng J, et al. Breed identification of meat using machine learning and breed tag SNPs[J]. Food Control, 2021, 125: 107971.
- [27] Liu R, Xu Z, Teng J, et al. Evaluation of six machine learning classification algorithms in pig breed identification using SNPs array data[J]. Anim Genet, 2023, 54(2): 113-122.
- [28] Vasconcelos L, Dias L G, Leite A, et al. Contribution to characterizing the meat quality of protected designation of origin serrana and preta de montesinho kids using the near-infrared reflectance methodology[J]. Foods, 2024, 13(10): 1581.
- [29] Kamruzzaman M, Elmasry G, Sun D W, et al. Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis[J]. Anal Chim Acta, 2012, 714: 57-67.
- [30] Tang X, Xie L, Yan M, et al. Genomic selection for meat quality traits based on VIS/NIR spectral information[J]. J Integr Agr, 2025, 24(1): 235-245.
- [31] Inoue K, Valente B D, Shoji N, et al. Inferring phenotypic causal structures among meat quality traits and the application of a structural equation model in Japanese black cattle[J]. J Anim Sci, 2016, 94(10): 4133-4142.
- [32] Ariza A G, González F J N, Jurado J M L, et al. Data mining as a tool to infer chicken carcass and meat cut quality from autochthonous genotypes[J]. Animals, 2022, 12(19): 2702.
- [33] Su R, Lv J, Xue Y, et al. Genomic selection in pig breeding: Comparative analysis of machine learning algorithms[J]. Genet Sel Evol, 2025, 57(1): 13.
- [34] Liang M, Miao J, Wang X, et al. Application of ensemble

- learning to genomic selection in Chinese simmental beef cattle[J]. *J Anim Breed Genet*, 2021, 138(3): 291-299.
- [35] Isidro J, Jannink J L, Akdemir D, et al. Training set optimization under population structure in genomic selection[J]. *Theor Appl Genet*, 2015, 128(1): 145-158.
- [36] Spindel J, Begum H, Akdemir D, et al. Genomic selection and association mapping in rice (*oryza sativa*): Effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines[J]. *PLoS Genet*, 2015, 11(2): e1004982.
- [37] Mao R, Zhou L, Wang Z, et al. A comprehensive strategy combining feature selection and local optimization algorithm to optimize the design of low-density chip for genomic selection [J]. *Agriculture-Basel*, 2023, 13(3): 614.
- [38] Seyum E G, Bille N H, Abtew W G, et al. Genomic selection in tropical perennial crops and plantation trees: A review[J]. *Mol Breeding*, 2022, 42(10): 58.
- [39] Wang Y, Zhu B, Wang J, et al. Evaluation of genomic mating approach based on genetic algorithms for long-term selection in Huaxi cattle[J]. *BMC Genomics*, 2024, 25(1): 1140.
- [40] Cuyabano B C D, Su G, Lund M S. Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population[J]. *BMC Genomics*, 2014, 15: 1171.
- [41] Eyduran E, Zaborski D, Waheed A, et al. Comparison of the predictive capabilities of several data mining algorithms and multiple linear regression in the prediction of body weight by means of body measurements in the indigenous Beetal goat of Pakistan[J]. *Pak J Zool*, 2017, 49(1): 257-265.
- [42] Azodi C B, Bolger E, McCarren A, et al. Benchmarking parametric and machine learning models for genomic prediction of complex traits[J]. *G3-Genes Genom Genet*, 2019, 9(11): 3691-3702.
- [43] Hamadani A, Ganai N A. Artificial intelligence algorithm comparison and ranking for weight prediction in sheep[J]. *Sci Rep*, 2023, 13(1): 13242.
- [44] Hamadani A, Ganai N A, Mudasir S, et al. Comparison of artificial intelligence algorithms and their ranking for the prediction of genetic merit in sheep[J]. *Sci Rep*, 2022, 12(1): 18726.
- [45] Kannan A, Basu J, Roy R, et al. Gender identification of chicks using vocalisation signals, artificial intelligence and machine learning techniques: Current status and future prospects[J]. *Worlds Poult Sci J*, 2025, 81(1): 87-102.
- [46] Campanholi S P, Garcia neto S, Pinheiro G M, et al. Can in vitro embryo production be estimated from semen variables in Senepol breed by using artificial intelligence?[J]. *Front Vet Sci*, 2023, 10: 1254940.
- [47] Verhoeven S, Chantziras I, Bernaerdt E, et al. The evaluation of an artificial intelligence system for estrus detection in sows [J]. *Porcine Health Manag*, 2023, 9(1): 9.
- [48] Arshad M F, Burrai G P, Varcasia A, et al. The groundbreaking impact of digitalization and artificial intelligence in sheep farming[J]. *Res Vet Sci*, 2024, 170: 105197.
- [49] Soloshenkov A D, Soloshenkova E A, Semina M T, et al. Artificial intelligence and classical methods in animal genetics and breeding[J]. *Russ J Genet*, 2024, 60(7): 843-856.
- [50] Buaban S, Duangjinda M, Suzuki M, et al. Genetic relationships of fertility traits with test-day milk yield and fat-to-protein ratio in tropical smallholder dairy farms[J]. *Anim Sci J*, 2016, 87(5): 627-637.
- [51] Jafaroghi M, Ghafouri-kesbi F, Khorami S J, et al. Application of structural equation models for genetic evaluation of lifetime reproductive traits and age at first lambing in Moghani sheep[J]. *Small Ruminant Res*, 2022, 214: 106761.
- [52] Zhao F, Zhang P, Wang X, et al. Genetic gain and inbreeding from simulation of different genomic mating schemes for pig improvement[J]. *J Anim Sci Biotechnol*, 2023, 14(1): 87.
- [53] Tang Z, Yin L, Yin D, et al. Development and application of an efficient genomic mating method to maximize the production performances of three-way crossbred pigs[J]. *Brief Bioinform*, 2023, 24(1): 1-10.
- [54] Liu Z, Du H, Lao F D, et al. PIMFP: An accurate tool for the prediction of intramuscular fat percentage in live pigs using ultrasound images based on deep learning[J]. *Comput Electron Agr*, 2024, 217: 108552.
- [55] Kou X, Yang Y, Xue H, et al. Revealing in vivo broiler chicken growth state: Integrating CT imaging and deep learning for non-invasive reproductive phenotypic measurement[J]. *Comput Electron Agr*, 2024, 227: 109477.
- [56] Lee H J, Lee J H, Gondro C, et al. deepGBLUP: Joint deep learning networks and GBLUP framework for accurate genomic prediction of complex traits in korean native cattle[J]. *Genet Sel Evol*, 2023, 55(1): 56.
- [57] Han J, Gondro C, Reid K, et al. Heuristic hyperparameter optimization of deep learning models for genomic prediction[J]. *G3-Genes Genom Genet*, 2021, 11(7): jkab032.
- [58] Chapman N H, Chlingaryan A, Thomson P C, et al. A deep learning model to forecast cattle heat stress[J]. *Comput Electron Agr*, 2023, 211: 107932.
- [59] Bouwman A C, Hulsegge I, Hawken R J, et al. Classifying aneuploidy in genotype intensity data using deep learning[J]. *J Anim Breed Genet*, 2023, 140(3): 304-315.
- [60] Sanchez T, Cury J, Charpiat G, et al. Deep learning for population size history inference: Design, comparison and combination with approximate bayesian computation[J]. *Mol Ecol Resour*, 2021, 21(8): 2645-2660.
- [61] Manzoori S, Farahani A H K, Moradi M H, et al. Detecting SNP markers discriminating horse breeds by deep learning[J]. *Sci Rep*, 2023, 13(1): 11592.
- [62] Mair C, Stear M, Johnson P, et al. A bayesian generalized random regression model for estimating heritability using overdispersed count data[J]. *Genet Sel Evol*, 2015, 47: 51.
- [63] Peters S O, Sinecen M, Kizilkaya K, et al. Genomic prediction with different heritability, QTL, and SNP panel scenarios using artificial neural network[J]. *IEEE Access*, 2020, 8: 147995-148006.

- [64] Souza L L, Dominguez-castaño P, Gianvecchio S B, et al. Heritability estimates and genome-wide association study of methane emission traits in Nellore cattle[J]. J Anim Sci, 2024, 102: skae182.
- [65] 刘晶.中国高校科技成果研究的文献计量学分析[J].情报工程,2019,5(6): 98-108.
- [66] 张玉良,黄煜,尤园.一种猪群扎堆检测方法、装置、设备及可读存储介质:202110485028.9[P].2021-07-09.
- [67] 郑吉星,胡清华,马锐.一种大型养殖场体征异常禽类检测系统及检测方法:201910506337.2[P].2020-06-30.
- [68] 李东锋,张进,张康宁,等.提高黄羽肉种鸡合格种蛋数的选育方法、模型及选育系统:202110212210.7[P].2024-04-09.
- [69] 吴克亮,王冀,赵兴波,等.一种基于卷积神经网络预测哺乳动物家畜增强子序列的装置、方法及计算机可读存储介质:202311156784.2[P].2023-12-01.
- [70] 赵书红,杨翔,李新云,等.生猪全域数字智能育种系统:202410311051.X[P].2024-06-07.
- [71] 吴筠彤,杨齐心,申李,等.一种生猪全域数字智能育种系统:202510054980.1[P].2025-02-18.
- [72] 高雪,李俊雅,王元清,等.一种基于全基因组SNP信息的华西牛基因组选配方法及应用:202410915694.5[P].2024-11-01.
- [73] 羊扬,王娟,宋瑞龙,等.一种用于畜牧养殖的智能化防疫方法及系统:202411053324.1[P].2024-11-19.
- [74] Novakovsky G, Dexter N, Libbrecht M W, et al. Obtaining genetics insights from deep learning via explainable artificial intelligence[J]. Nat Rev Genet, 2023, 24(2): 125-137.

Research Trends in the Intersection of Animal Genetic Breeding and Artificial Intelligence: A Bibliometric Analysis

WANG Ning¹, ZHANG Bo², ZHANG Hao^{2*}

(1. Library of China Agricultural University, Beijing 100193, China;
2. College of Animal Science and Technology, China Agricultural University, Beijing 100193, China)

Abstract: This study explores the research landscape at the intersection of artificial intelligence (AI) and animal genetic breeding, aiming to identify key academic and technological trends to provide a reference for future research. Utilizing bibliometric analysis, the study integrates data from SCIE-indexed journal articles and incoPat patents to examine annual output, institutional collaboration, and thematic evolution. VOSviewer, CiteSpace, and ITGInsight were employed to construct knowledge maps and identify thematic hotspots. Results indicate 2019 as a critical turning point with accelerated growth in publications and patents. China ranks first globally in both the number of publications and patent applications. However, collaboration patterns differ between literature and patents: academic research features international networks—primarily among China, the U.S., and Brazil—while patent activity is dominated by domestic institutions, with limited cross-border cooperation. Thematic analysis reveals that academic research focuses on the optimization of machine learning and deep learning models for trait prediction, genetic analysis, and genomic selection. Patents emphasize the integration of AI into intelligent management systems covering the full lifecycle of livestock production. Although AI has improved the accuracy and efficiency of animal breeding, challenges remain—namely, limited model generalizability, lack of standardized data, and hardware compatibility issues. Future efforts should prioritize multi-source data integration, lightweight algorithm development, and cross-scenario application to advance full-process automation and support the sustainable transformation of animal husbandry.

Keywords: Bibliometric analysis; Animal genetics breeding; Artificial intelligence; Research highlights and fronts

(责任编辑:赵楠)