

Li 等学者提出了自适应扰动步长的策略。实验表明,该策略在 CIFAR-10 数据集和 ResNet-50 模型的对抗训练中,可在保持模型鲁棒性不变的情况下,将对抗样本生成时间缩短 35%,有效缓解了 PGD 的计算成本压力。

#### 2.4.2 跨攻击泛化性差

在未接触过的攻击方式(如 DeepFool、JSMA)下,模型鲁棒准确率会显著下降。2024 年, Baniecki 等提出了可解释性引导对抗训练方案。该方案利用类激活映射(CAM)图定位脆弱特征区域,在训练时优先针对该区域生成对抗样本,迫使模型学习更稳健的特征表示。实验表明,该方案在 AutoAttack 测试下,能使模型鲁棒准确率从 74.3% 提升至 86.3%,提升幅度达 12%。

#### 2.4.3 边缘设备适配难

在边缘设备场景中,对抗训练因硬件算力限制难以落地。例如,自动驾驶车载芯片无法承载多步迭代生成对抗样本的计算,导致鲁棒性模型难以部署。2025 年,Zhang 等提出量化感知对抗训练方案,在训练阶段模拟 8 位量化误差,使模型提前适配低精度计算环境。实验表明,该方法能够使模型在边缘设备上的推理速度提升 50%,在 CIFAR-10/ResNet-50 基准测试中,PGD 攻击下的鲁棒准确率仅下降 3%,有效平衡了算力约束与模型鲁棒性需求。

### 3 结语

对抗训练作为对抗机器学习的核心防御技术,在高安

全需求领域(如医疗影像、自动驾驶)和数据隐私敏感场景(如医疗、金融)中展现出显著优势。然而,对抗训练在低资源场景下容易引发过拟合,和在多模态场景中难以同步优化两种模态的鲁棒性的弊端也在一定程度上限制了其工程落地。

作为对抗机器学习的核心防御技术,未来,对抗训练研究应聚焦于小样本和多模态场景,强化理论保障,优化“鲁棒性、泛化性、计算成本”的三维平衡,为 AI 安全落地提供坚实的技术支撑。

#### 参考文献

- [1] Madry A, Makelov A, Schmidt L. Towards Deep Learning Models Resistant to Adversarial Attacks[J]. JMLR, 2018,19(1):2847-2880.
- [2] Zhang H, Li M, Wang Y. Gradient Reuse for Efficient Adversarial Training[C]//ICML, 2021:12345-12354.
- [3] 周纯毅,王磊.联邦对抗蒸馏在医疗 AI 中的应用[J].软件学报, 2024,35(2):789-805.
- [4] Baniecki H, Biecek P. Explainable Guided Adversarial Training[J]. TPAMI, 2024,47(3):1201-1215.
- [5] Li Q, Zhang Q, Xu H. Adaptive Step-Size for Fast Adversarial Training[C]//ICML, 2025:9876-9885.
- [6] Zhang L, Chen J. Quantization-Aware Adversarial Training for Edge Devices[C]//ICASSP, 2025:4567-4571.
- [7] Carlini N, Wagner D. Towards Evaluating the Robustness of Neural

# AI Large Models Reshaping the Future: Applications, Challenges, and Cutting-Edge Prospects in the Healthcare and Health Management Field

Yonggang Weng

Oriental Huaxia Cardiovascular Health Research Institute, Suzhou Industrial Park, Suzhou, Jiangsu, 215000, China

## Abstract

Artificial intelligence (AI) and big data, especially the new generation of AI technologies represented by large language models (LLMs), are emerging as the core engine of new quality productive forces, unleashing unprecedented potential for integration and innovation in the healthcare sector. From assisting in diagnosis, personalized treatment to drug development and health management, AI applications are profoundly transforming every aspect of medical services. By integrating massive data and advanced algorithms, AI not only improves medical quality and management efficiency, enhances patient satisfaction, but also has the potential to address core challenges such as uneven distribution of medical resources and high costs, driving the field of medical health management towards a more intelligent, inclusive new era of sustainable development.

## Keywords

artificial intelligence application; big data; large language model; medical health management

## AI 大模型重塑未来：医疗健康管理领域的应用、挑战与前沿展望

翁勇刚

苏州工业园区东方华夏心血管健康研究院，中国·江苏 苏州 215000

## 摘要

人工智能（AI）和大数据，特别是以大语言模型（LLM）为代表的新一代AI技术，正作为新质生产力的核心引擎，在医疗健康领域释放出前所未有的融合与创新潜力。从辅助诊断、个性化治疗到药物研发和健康管理，AI的应用正深刻变革着医疗服务的各个环节。通过整合海量数据和先进算法，AI不仅能提升医疗质量与管理效率，改善患者满意度，更有望解决医疗资源不均、成本高昂等核心挑战，推动医疗健康管理领域迈向一个更智能、更普惠的可持续发展新纪元。

## 关键词

人工智能应用；大数据；大语言模型；医疗健康管理

## 1 AI 与大数据：医疗健康领域的变革之力

随着科学技术的飞速发展，人工智能和大数据已不再是遥远的概念，而是深度渗透医疗健康领域的现实力量。它们通过强大的数据分析和模式识别能力，为医生提供更精准的诊断依据，优化患者的诊疗体验。特别是近年来，以 GPT-4、Claude 3.5 等为代表的大语言模型展现出惊人的能力，能够理解和生成复杂的医学文本，支持智能问诊、定制化治疗方案，为构建全新的智慧医疗生态奠定了坚实基础。

### 1.1 提高医疗服务效率与运营水平

AI 与大数据能够显著优化医院的运营管理模式。通过对门诊量、住院率、手术安排等历史和实时数据进行分析，AI 可以预测未来的医疗需求，实现床位、人员和设备的动态优化配置，从而在有限的资源下服务更多患者，显著降本增效。这推动了医疗机构从传统的“诊疗中心”向覆盖预防、诊断、治疗、康复全流程的“全生命周期健康管理中心”转型。

### 1.2 提升患者体验与健康结果

AI 技术正将医疗的重心从“治疗”向“预防”前移。借助可穿戴设备和健康大数据分析，AI 能够及早预警个人健康风险并进行恰当干预，显著提升患者的生活质量。更重要的是，AI 使真正的个性化医疗成为可能。通过分析个体的基因、生活方式、电子病历（EHR）等海量数据，AI

【作者简介】翁勇刚（1977-），男，中国浙江衢州人，在读博士，工程师，从事人工智能开发和应用研究。

可以为每位患者量身定制治疗方案，避免“一刀切”的治疗模式，从而大幅提升疗效、减少副作用，实现精准医疗。

### 1.3 降低医疗成本与研发门槛

高昂的成本是制约医疗普惠性的关键因素。AI 和大数据的应用从多个环节入手，有效降低了医疗成本。在资源利用方面，大数据分析能减少医疗资源的闲置与浪费。在药物研发方面，AI 能够大幅缩短新药靶点发现和临床试验的周期，降低研发成本。这使得更多创新疗法能够以更可及的价格进入市场，惠及广大患者。

## 2 核心应用与前沿突破

人工智能和大数据在医疗健康领域的应用已从理论走向实践，并在多个细分领域取得了令人瞩目的突破，特别是大语言模型的加入，进一步拓宽了应用边界。

### 2.1 智能诊断：超越人类视觉的洞察力

医学影像与病理分析

AI 在医学影像领域的应用最为成熟，其表现已在某些方面超越人类专家。通过深度学习模型训练，AI 能够精准识别 X 光、CT、MRI 中的微小病变。例如，在肺结节、乳腺癌、脑肿瘤等疾病的筛查中，AI 展现了极高的准确性。在病理诊断领域，AI 能够辅助医生细致地分析病理切片，自动识别和计数癌细胞，显著提升诊断效率和一致性。

案例聚焦：AI 在放射学诊断中的卓越表现

一项由麻省总医院（MGH）与麻省理工学院（MIT）合作的研究显示，其开发的 AI 系统在通过 CT 扫描检测肺结节方面的准确率达到 94%，显著高于人类放射科医生 65% 的平均水平。这不仅减轻了放射科医生的工作负担，也为患者赢得了宝贵的早期治疗时间。

此外，像 PathAI 和 Paige 这样的公司，正致力于利用 AI 技术推动数字病理学的革命，为癌症诊断提供更精准的工具。



图 1 人工智能在医疗领域中的应用框架，展示了 AI 在辅助诊疗、医学影像、健康管理等多个层面的应用。

大语言模型赋能复杂病例诊断

新一代的大语言模型（LLM）正在成为医生的“超级诊断助手”。它们能够快速阅读和理解海量的医学文献、病历和检查报告，从中提炼关键信息，并形成鉴别诊断列表。

一项研究表明，GPT-4 在处理《新英格兰医学杂志》发布的复杂临床病例挑战时，诊断准确率达到 57%，表现优于绝大多数人类答题者。这预示着 LLM 在辅助医生处理罕见病和复杂病例方面具有巨大潜力。

### 2.2 精准治疗与手术辅助

#### 2.2.1 个性化治疗方案

AI 正在推动肿瘤治疗进入“千人千面”的精准时代。通过整合患者的基因测序数据、影像学特征和临床信息，AI 可以预测不同治疗方案（如化疗、靶向药、免疫疗法）的有效性，帮助医生制定最佳个性化策略。然而，这一领域也曾经历波折。

案例反思：IBM Watson for Oncology 的经验与教训

IBM 的 Watson for Oncology 曾是 AI 医疗的明星项目，旨在为癌症患者提供个性化治疗建议。然而，由于其训练数据主要来自美国纪念斯隆-凯特琳癌症中心，导致其建议在其他国家和地区存在“水土不服”的问题，且未能充分融入临床工作流。这一案例深刻揭示了 AI 医疗应用必须重视数据多样性、本地化适应以及与临床实践的深度融合。

#### 2.2.2 AI 驱动的手术机器人

以“达·芬奇手术系统”为代表的手术机器人，在 AI 的加持下正变得越来越智能。AI 不仅能帮助医生进行精确的手术路径规划和风险评估，还能在术中通过增强现实技术叠加关键解剖结构，提高手术的精准度和安全性。未来的手术机器人将具备更高的自主性，能够在医生的监督下完成部分标准化操作，从而提升手术效率。

### 2.3 主动健康管理与虚拟助手

#### 2.3.1 从被动监测到主动干预

智能手环、手表等可穿戴设备正在持续收集用户的生命体征数据。AI 通过分析这些数据，可以识别出潜在的健康风险，如心律不齐、睡眠呼吸暂停等，并及时发出预警。例如，AI 算法可以分析远程心电图（ECG）记录，快速检测房颤，使临床医生能够更快地提供干预。这种模式将健康管理从医院延伸到了日常生活中。

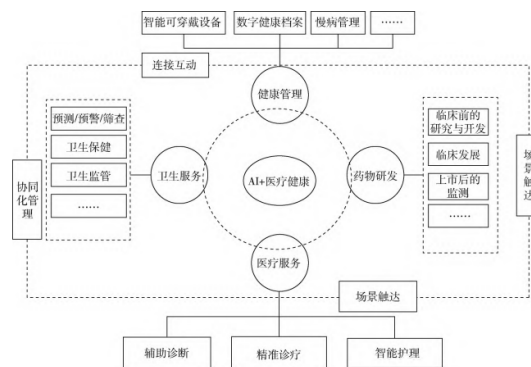


图 2 人工智能在医疗健康中的应用生态，涵盖健康管理、药物研发、医疗服务和卫生服务四大领域。

### 2.3.2 AI 赋能心理与精神健康

心理健康是 AI 应用的新兴热点领域。基于 LLM 的聊天机器人，如 Woebot、Wysa 等，可以为用户提供 24/7 的心理支持，通过认知行为疗法（CBT）等循证方法帮助用户管理焦虑、抑郁等情绪问题。这些虚拟助手拓宽了心理治疗的渠道，为更多人提供了便捷、私密的支持。

## 2.4 革新药物研发与临床试验

### 2.4.1 AI 加速新药发现

传统药物研发周期长、成本高、失败率高。AI 正在从根本上改变这一现状。AI 模型可以快速筛选数以亿计的分子，预测其成药性，从而发现全新的药物靶点和候选分子。这一领域的成功案例正不断涌现。

### 2.4.2 案例聚焦：AI 发现的药物进入临床试验

由 Insilico Medicine 公司利用其生成式 AI 平台发现的一款治疗特发性肺纤维化（IPF）的候选药物，已进入 II 期临床试验阶段。这是全球首批完全由 AI 设计并进入临床阶段的药物之一，标志着 AI 在药物发现领域的里程碑式突破。

### 2.4.3 AI 赋能药物重定位与临床试验

除了从头设计新药，AI 在“老药新用”（药物重定位）方面也大有可为。例如，哈佛大学开发的 TxGNN 模型，通过分析庞大的医学知识图谱，能够为缺乏有效疗法的罕见病等发现潜在的治疗药物。在临床试验阶段，AI 可以更精准地招募合适的受试者，实时监控试验数据，预测并发现潜在的安全性问题，从而提高试验效率和成功率。

## 3 挑战、伦理与治理框架

尽管 AI 在医疗健康领域的前景广阔，但其发展和应用仍面临着技术、伦理和监管等多重挑战。审慎应对这些挑战是确保技术能够安全、公平、有效地服务于人类健康的关键。

### 3.1 技术与数据挑战

AI 模型的性能高度依赖于高质量、大规模、多样化的数据。然而，医疗数据往往存在于孤立的系统中，格式不一，质量参差不齐，形成了“数据孤岛”。此外，许多先进的 AI 模型，特别是深度学习模型，其决策过程如同一个“黑箱”，缺乏透明度和可解释性，这在要求高可靠性和高问责性的医疗领域是一个重大障碍。

### 3.2 算法偏见与健康公平

AI 模型是从数据中学习的，如果训练数据本身存在偏见（例如，某些人群的数据代表性不足），模型就可能复制甚至放大这些偏见。这可能导致 AI 系统在不同性别、种族或社会经济地位的患者群体中表现出性能差异，从而加剧现有的健康不平等。确保算法的公平性是 AI 医疗伦理的核心议题。

### 3.3 伦理与隐私安全

医疗数据是极其敏感的个人隐私。在利用这些数据训

练 AI 模型时，如何确保数据安全、保护患者隐私、获得合法授权，是必须严格遵守的红线。此外，当 AI 系统出错导致医疗事故时，责任应如何界定——是开发者、医院还是医生？这些复杂的伦理和法律问题亟待解决。

## 3.4 监管框架与可解释 AI (XAI)

全球各地的监管机构正在积极探索适应 AI 医疗发展的监管新范式。例如，美国食品药品监督管理局（FDA）提出了针对“软件即医疗设备”（SaMD）的监管框架，并引入了“预定变更控制计划”（PCCP），允许 AI 模型在上市后持续学习和迭代。同时，为了解决“黑箱”问题，可解释人工智能（XAI）的研究和应用变得至关重要。XAI 技术（如 LIME、SHAP）旨在揭示模型做出特定决策的原因，增强临床医生对 AI 的信任和采纳率。

## 4 未来展望：迈向 AI 原生智慧医疗

展望未来，AI 技术将更深层次地融入医疗健康的全流程，塑造一个以“AI 原生”为特征的智慧医疗新生态。

**多模态 AI 的融合应用：**未来的医疗 AI 将不再局限于单一类型的数据。通过融合影像、基因、病理、临床文本和实时生理信号等多模态数据，AI 将能够构建对疾病和患者更全面、更立体的理解，从而做出更精准判断。

**生成式 AI 的普及：**以 LLM 为代表的生成式 AI 将在临床文档自动化、医患沟通、医学教育和虚拟诊疗等场景中发挥更大作用，将医生从繁重的行政工作中解放出来，回归到对患者的直接关怀。

**从辅助到协同：**AI 的角色将从一个被动的“辅助工具”转变为与医生协同工作的“智能伙伴”。AI 将无缝集成到临床工作流程中，在医生需要时主动提供决策支持和风险预警。

**超个性化与预测性维护：**基于持续的健康监测和深度数据分析，AI 将为每个人提供动态调整的超个性化健康管理方案，实现对疾病的精准预测和前瞻性干预。

## 5 结语

综上所述，人工智能和大数据正在以前所未有的广度和深度重塑医疗健康领域。从智能诊断、精准治疗到药物研发和健康管理，其应用价值已得到广泛验证。然而，技术的飞速发展也伴随着严峻的挑战。未来，我们必须在推动技术创新的同时，建立健全的伦理规范、法律法规和治理体系，解决数据、偏见和安全等核心问题。只有这样，我们才能充分释放 AI 的潜力，打造一个更加高效、公平、人性化的智慧医疗生态，最终实现全人类健康福祉的提升。

### 参考文献

- [1] Mishra V, Lurie Y, Mark S. Ethical Considerations in AI-Enabled Healthcare[M]//Singh T P, Kumar C J, Abraham A, et al. Revolutionizing Healthcare: Impact of Artificial Intelligence on Diagnosis, Treatment, and Patient Care. Cham: Springer, 2025: