



## MorpheusAPI:基于大语言模型Agent的智能麻醉平台

王静, 申乐, 林飞, 张濛濛, 黄俊, 倪清桦, 田永林, 兰岭, 叶佩军, 吕宜生, 王飞跃

引用本文:

王静, 申乐, 林飞, 等. MorpheusAPI:基于大语言模型Agent的智能麻醉平台[J]. *智能系统学报*, 2026, 21(1): 156-166.

WANG Jing, SHEN Le, LIN Fei, et al. MorpheusAPI: an LLM Agent for intelligent anesthesia platform[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(1): 156-166.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202505004>

## 您可能感兴趣的其他文章

### 非结构化文档敏感数据识别与异常行为分析

Unstructured document sensitive data identification and abnormal behavior analysis  
*智能系统学报*. 2021, 16(5): 932-939 <https://dx.doi.org/10.11992/tis.202104028>

### 智能化医养融合服务平台关键技术及应用研究

Research on key technologies and applications of intelligent medical and care integration service platform  
*智能系统学报*. 2021, 16(5): 972-988 <https://dx.doi.org/10.11992/tis.202103022>

### 面向车规级芯片的对象检测模型优化方法

Object detection model optimization method for car-level chips  
*智能系统学报*. 2021, 16(5): 900-907 <https://dx.doi.org/10.11992/tis.202107057>

### 利用MISA多目标优化的置信规则库分类算法

Belief rule base classification algorithm using MISA multi-objective optimization  
*智能系统学报*. 2019, 14(5): 982-990 <https://dx.doi.org/10.11992/tis.201809022>

### 多移动机器人协同搬运技术综述

Technologies for cooperative transportation by multiple mobile robots  
*智能系统学报*. 2019, 14(1): 20-27 <https://dx.doi.org/10.11992/tis.201801038>

### 泛逻辑学理论——机制主义人工智能理论的逻辑基础

Universal logic theory: logical foundation of mechanism-based artificial intelligence theory  
*智能系统学报*. 2018, 13(1): 19-36 <https://dx.doi.org/10.11992/tis.201711033>

DOI: 10.11992/tis.202505004

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20251209.1031.002>

# MorpheusAPI: 基于大语言模型 Agent 的智能麻醉平台

王静<sup>1,2</sup>, 申乐<sup>3</sup>, 林飞<sup>1</sup>, 张濛濛<sup>2,4</sup>, 黄俊<sup>1</sup>, 倪清桦<sup>1</sup>, 田永林<sup>2</sup>, 兰岭<sup>3</sup>, 叶佩军<sup>2</sup>, 吕宜生<sup>2</sup>, 王飞跃<sup>1,5</sup>

(1. 澳门科技大学 创新工程学院, 澳门 999078; 2. 中国科学院自动化研究所 多模态人工智能系统全国重点实验室, 北京 100190; 3. 中国医学科学院北京协和医院 麻醉科, 北京 100730; 4. 中国科学院大学 人工智能学院, 北京 100049; 5. 中国科学院自动化研究所 复杂系统管理与控制国家重点实验室, 北京 100190)

**摘要:** 针对传统围术期麻醉管理模式主要依赖临床指南和麻醉医生临床判断, 使得麻醉医生在面对海量的实时生理数据、复杂的患者个体化情况以及瞬息万变的高风险场景时需要承担巨大的工作负担和决策压力的问题, 本文提出了 MorpheusAPI 框架, 一种基于大语言模型 (large language models, LLMs) 的多 Agents 智能麻醉平台。该平台包含执行大模型和影子大模型, 执行大模型整合了感知、预测、决策、验证和中央协调 5 个 Agents, 通过模型上下文协议 (model context protocol, MCP) 实现多模态数据高效整合, 思维链 (chain-of-thought, CoT) 提示增强风险推理, 检索增强生成 (retrieval-augmented generation, RAG) 确保临床决策可靠性, 麻醉影子大模型通过持续优化执行模型性能形成闭环系统。案例研究显示 MorpheusAPI 系统风险预测响应时间 0.4 s, 核心推理延迟 10~15 ms, 成功将丙泊酚诱导剂量优化至 2.0 mg/(kg·h), 维持平均动脉压不小于 65 mmHg, 结果验证了其在提升麻醉安全性和效率的巨大潜力, 可为智能麻醉系统设计与应用提供新思路。

**关键词:** 麻醉; 多代理系统; 大语言模型; 人工智能; 智能代理; 决策支持系统; 医疗计算; 风险评估

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2026)01-0156-11

中文引用格式: 王静, 申乐, 林飞, 等. MorpheusAPI: 基于大语言模型 Agent 的智能麻醉平台 [J]. 智能系统学报, 2026, 21(1): 156-166.

英文引用格式: WANG Jing, SHEN Le, LIN Fei, et al. MorpheusAPI: an LLM Agent for intelligent anesthesia platform[J]. CAAI transactions on intelligent systems, 2026, 21(1): 156-166.

## MorpheusAPI: an LLM Agent for intelligent anesthesia platform

WANG Jing<sup>1,2</sup>, SHEN Le<sup>3</sup>, LIN Fei<sup>1</sup>, ZHANG Mengmeng<sup>2,4</sup>, HUANG Jun<sup>1</sup>, NI Qinghua<sup>1</sup>,  
TIAN Yonglin<sup>2</sup>, LAN Ling<sup>3</sup>, YE Peijun<sup>2</sup>, LYU Yisheng<sup>2</sup>, WANG Feiyue<sup>1,5</sup>

(1. Faculty of Innovation Engineering, Macau University of Science and Technology, Macau 999078, China; 2. the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; 3. Department of Anesthesiology, Peking Union Medical College Hospital, Beijing 100730, China; 4. the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China; 5. the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** To address the reliance of traditional perioperative anesthesia management models on clinical guidelines and the clinical judgment of anesthesiologists, which causes anesthesiologists to bear a huge workload and decision-making pressure when faced with massive real-time physiological data, complex individualized patient conditions, and rapidly changing high-risk scenarios, MorpheusAPI, a multi-agent, intelligent anesthesia platform based on large language models (LLMs), was proposed. The platform includes an execution model and a shadow model. The execution model integrates five agents for perception, prediction, decision-making, verification and central coordination. It enables efficient integration of multi-modal data through the model context protocol, the chain-of-thought prompts to enhance risk reasoning, and retrieval-augmented generation (RAG) to ensure the reliability of clinical decisions. The anesthesia shadow model forms a closed loop that continuously optimizes the performance of the execution model. Case studies demonstrated that the risk prediction response time for the MorpheusAPI system is 0.4 s, the core reasoning delay is 10~15 ms, the propofol induction dose is successfully optimized to 2.0 mg/(kg·h), and the mean arterial pressure is maintained not less than 65 mmHg. The results verify the great potential of the model to improve anesthesia safety and efficiency and provide insights into the design and application of intelligent anesthesia systems.

**Keywords:** anesthetics; multi-agent systems; large language models; artificial intelligence; intelligent agents; decision support systems; medical computing; risk assessment

收稿日期: 2025-05-08. 网络出版日期: 2025-12-09.

基金项目: 澳门特别行政区科学与技术发展基金项目 (0093/2023/RIA2, 0145/2023/RIA3, 0157/2024/RIA2); 北京市自然科学基金海淀联合基金项目 (L222099); 首都卫生发展科研专项 (2024-2-4015); 四川科技厅重点研发计划项目 (2024YFHZ0011).

通信作者: 王飞跃. E-mail: [feiyue.wang@ia.ac.cn](mailto:feiyue.wang@ia.ac.cn).

围术期麻醉管理作为现代医疗体系中确保外科手术顺利进行和患者安全的关键环节, 涵盖了术前评估与优化、术中麻醉实施与精细调控、术后镇痛与早期康复支持的全过程。其目标在于为接受外科手术的患者提供系统性的医疗保障, 以

应对手术应激、维持生理稳定、减轻术后不适并促进顺利康复<sup>[1]</sup>。由于患者的生理状态、合并疾病和药物反应等个体差异巨大, 麻醉医师需进行细致的术前评估, 制定个体化的麻醉方案; 其次, 手术过程本身具有高度的不确定性, 术中出血、脏器功能变化等突发情况要求麻醉医师具备快速、准确的判断和应对能力; 此外, 患者的生理状态在术中会随着手术进程、药物作用等因素不断变化, 麻醉医师需要持续、精密地监测患者的心率、血压、氧饱和度、呼吸、体温和脑电活动等关键生理参数, 并根据实时数据动态调整麻醉深度、药物剂量和输液速度等, 以维持患者生理功能的稳定。这种多因素交互、实时变化的特性, 使得围术期麻醉管理成为一项极具挑战的专业领域。

传统麻醉管理主要依赖医生的临床经验知识储备和简单的基于规则的决策系统。然而, 这些方法存在异质性且缺乏灵活性, 往往不足以有效应对复杂且动态变化的临床场景<sup>[2-3]</sup>。基于规则的方法通常难以精确跟踪患者状态的实时变化并预测药物效应的动态变化<sup>[4]</sup>。此外, 单一预测模型(如传统线性回归或基础机器学习技术)在处理临床数据的复杂性和多模态性时面临重大挑战, 且在处理大规模数据和快速响应需求时常遇到计算瓶颈<sup>[5-7]</sup>。因此, 亟需引入智能化技术, 以提升麻醉管理的效率和准确性, 减轻麻醉医师的工作压力。

近年来, 大语言模型 (large language models, LLMs) 在医学领域的应用潜力逐渐显现。LLMs 凭借其强大的语言理解和生成能力, 已在医学知识问答、临床决策支持等方面展现出接近人类专家的水平<sup>[8]</sup>。与此同时, 大模型代理 (LLMs Agent) 系统通过多个 Agent 的协作, 能够有效处理复杂、动态的任务, 广泛应用于医疗资源优化和紧急医疗服务管理等领域<sup>[9]</sup>。赵毅飞等<sup>[10]</sup>提出的平行麻醉理论通过引入平行智能架构, 实现了从自动化到智慧型全周期麻醉管理的转变, 为多 Agent 协作提供了理论支持; 王惠珍等<sup>[11]</sup>提出的平行手术室概念进一步强调了围术期护理的系统性管理, 为智能化麻醉平台的设计提供了参考; Lin 等<sup>[12]</sup>通过边缘与云智能集成的平行医疗设备架构, 为实时性麻醉系统的分布式计算提供了支持。

尽管如此, 当前在麻醉领域, 鲜见将大模型和 Agent 技术进行有效整合的系统性研究。现有研究往往集中于单一技术应用, 如单纯使用深度学习预测麻醉深度, 或单一 Agent 实施麻醉剂量

优化, 缺乏全面的架构能够统筹处理数据采集、状态预测、剂量推荐、实时监控及术后跟踪等多个环节<sup>[13-16]</sup>。

基于此, 本文提出 MorpheusAPI 框架, 一种基于 Agent 协作的大语言模型架构, 麻醉执行大模型包括感知 Agent、预测 Agent、决策 Agent、验证 Agent 和中央协调 Agent 的自主协作, 通过模型上下文协议 (model context protocol, MCP)<sup>[17]</sup> 整合多模态数据, 思维链 (chain-of-thought, CoT)<sup>[18]</sup> 提示推理风险与干预, 检索增强生成 (retrieval-augmented generation, RAG)<sup>[19-20]</sup> 确保临床依据的准确性。框架采用 Agentic Workflow<sup>[21]</sup> 设计, 实现任务动态分配与高效通信, 满足术中实时性需求。验证 Agent 通过 MCP 和 RAG 进行质量控制, 中央协调 Agent 整合输出形成综合建议, 麻醉影子大模型分析运行数据优化性能, 构建从数据处理到临床决策的闭环。通过围术期低血压风险管理的案例, 验证框架的实际效果。本文主要创新点如下:

1) 模块化 Agent 架构: 设计五类 Agent 协同, 依托 MCP 与 CoT 提示, 整合多模态数据, 实现麻醉全流程精准支持。

2) 异步协作机制: 利用 Kafka 与注意力机制驱动 Agent 实时交互, 结合元学习适配动态临床场景。

3) 平行大模型与验证机制的系统优化: 麻醉执行大模型经微调和后蒸馏优化低延迟, 麻醉影子大模型通过云端预训练、微调和在线学习监控并优化协作策略。

## 1 国内外研究现状

麻醉管理作为一个复杂且高风险的医学领域, 其智能化技术的演进经历从传统规则系统到现代人工智能技术的显著变革。本节将综述麻醉智能化技术的发展历程、大模型在医学中的应用、Agent 系统的研究进展, 并分析当前研究的不足, 为 MorpheusAPI 的提出提供理论依据和研究背景。

### 1.1 麻醉智能化技术

近年来, 麻醉管理逐渐趋向于智能化与自动化, 广泛应用先进的人工智能技术, 以实现麻醉深度评估、药物剂量优化及患者生理状态实时监测等任务的精准控制。机器学习和深度学习技术的快速发展, 推动智能麻醉研究的显著进展。传统机器学习模型, 如决策树 (decision tree, DT)、支

持向量机 (support vector machine, SVM) 和随机森林 (random forest, RF), 在麻醉深度评估和剂量预测方面取得初步成果<sup>[22-24]</sup>。然而, 这些模型处理复杂的动态生理数据时难以有效捕捉时序依赖性和非线性关系, 存在明显局限性。

基于深度学习的时序模型, 如长短期记忆网络 (long short-term memory, LSTM) 和卷积神经网络 (convolutional neural network, CNN), 已逐渐成为麻醉智能研究的主流<sup>[25-26]</sup>。研究表明, 这些模型能更有效地捕捉生理信号的动态变化特征, 从而实现更高精度的麻醉深度预测和剂量优化。例如, 部分研究利用 LSTM 模型成功预测患者麻醉过程中的血压和心率变化趋势, 大幅提升临床决策的实时准确性。同时, 强化学习 (reinforcement learning, RL) 在麻醉剂量推荐领域展现出巨大潜力。通过实时互动和迭代优化, 强化学习能够为患者自动制定个性化麻醉策略。基于深度强化学习的临床模拟研究表明, 这种方法相比传统方法能够实现更加稳定和精准的麻醉剂量调整<sup>[27-28]</sup>。

然而, 这些技术尽管取得明显进展, 但它们通常仅专注单一任务, 缺乏整体任务协作与交互处理机制, 限制其在复杂麻醉场景中的广泛应用。因此, 更有效地整合多种人工智能技术、开发具备综合任务处理能力和协同智能的麻醉管理系统, 成为当前研究的重要方向之一。

## 1.2 大模型在医学中的应用

大语言模型, 如 GPT-4 (generative pre-trained transformer 4)<sup>[29]</sup>、Med-PaLM<sup>[30]</sup> 和 BioGPT<sup>[31]</sup>, 凭借其强大的预训练能力和迁移学习特性, 在医学领域展现出变革性潜力。这些模型通过在海量语料库上进行预训练, 捕获通用语言和领域知识, 随后通过微调 (fine-tuning) 适配特定任务, 显著降低对标注数据的需求。例如, Med-PaLM 在回答美国医学执照考试 (United States medical licensing examination, USMLE) 风格问题时, 首次以 67% 的得分超过及格线, 接近人类专家水平, 同时在生成临床笔记、辅助诊断和患者教育方面表现优异。此外, GPT-4 在生成临床对话摘要和回答医学考试题方面也表现出色, 显示出其在医学知识推理方面的潜力。张志成等<sup>[32]</sup> 提出了 OrthoGPT, 建立了骨科大模型体系; 肖建力等<sup>[33]</sup> 综述了 LLMs 在医疗领域的最新进展, 涵盖诊断、治疗建议和患者管理; 刘学博等<sup>[34]</sup> 分析了 LLMs 的关键技术 (如 CoT 和 RAG), 并展望其在精准医疗中的潜力; 马武仁等<sup>[35]</sup> 进一步指出 LLMs 在临床决策支持中的实时性和个性化优势, 为麻醉管理提供了借鉴。

近年来, 大型语言模型在麻醉学领域的应用迅速发展, 涵盖考试辅助、患者教育、术中决策支持和医疗记录管理等多个场景, 展现其在提升效率和精准性方面的潜力。在教育与考试辅助方面, LLMs 表现出色。例如, Angel 等<sup>[36]</sup> 比较多个 LLMs 在回答美国麻醉学会 (American society of anesthesiologists, ASA) 考试问题时的表现, 发现其在知识性问题上的正确率达 70%, 接近专家水平; Khan 等<sup>[37]</sup> 进一步验证 LLMs 在心胸麻醉考试中的适用性, 正确率约 65%, 但复杂推理仍需改进。类似地, Guthrie 等<sup>[16]</sup> 开发的 OARA (Operating and Anesthetic Reference Assistant) 通过微调 LLMs, 为住院医师教学生成高质量材料, 正确率达 90%, 显著提升教育效率。患者教育与沟通也是 LLMs 的重要应用场景。Ratnagandhi 等<sup>[38]</sup> 利用 LLMs 生成术前指导材料, 提高患者理解率至 85%, 而 Kuo 等<sup>[39]</sup> 发现 ChatGPT 在回答患者常见问题时信息全面性略优于单一医生, 表明 LLMs 可有效辅助医患交互。

在术中管理和数据处理方面, LLMs 的探索性应用展现出更大潜力。Wang 等<sup>[14]</sup> 开发的 Hypnos 模型通过微调优化术中剂量推荐, 模拟场景下误差率降至 10%。Feng 等<sup>[40]</sup> 提出的 AnesBench 评估框架显示, LLMs 在术前评估中 F1 分数达 0.85, 但在动态术中决策中表现较弱。此外, Wang 等<sup>[41]</sup> 通过 ATSCOT (anesthesia text structuring chain of thought) 模型利用 CoT 技术结构化麻醉记录, 错误率仅 5%, 为临床文档管理提供高效工具。针对非英语场景, Zhou 等<sup>[15]</sup> 和 Wang 等<sup>[42]</sup> 分别开发中文麻醉学数据集和 LLMs, 术前评估准确率达 88% 和 F1 分数 0.82, 填补语言多样性空白。然而, LLMs 在麻醉学中的应用仍面临挑战。当前模型多依赖静态知识, 缺乏实时数据整合和复杂推理能力, 限制其在术中动态场景的部署。此外, 数据隐私、模型可解释性和临床验证不足也阻碍其广泛应用。未来, 通过多模态数据融合和领域专用微调, LLMs 有望实现麻醉全流程的智能化支持, 为精准医疗注入新动力。

## 1.3 大模型 Agent 系统

大模型近年来显著推动医疗领域的智能化发展, 通过协作和自适应决策优化复杂任务处理。Agent 结合预训练语言能力与动态推理, 为医疗场景提供灵活支持<sup>[43-44]</sup>。在麻醉学中, Wang 等<sup>[41]</sup> 开发的 ATSCOT 模型利用 Agent 协作和 CoT 技术, 结构化麻醉医疗记录, 错误率降至 4.8%, 显著提高临床文档效率。类似地, Agent 在其他医疗

场景中展现广泛潜力。Sudarshan 等<sup>[45]</sup>提出 Agent workflows, 生成患者友好的医疗报告, 患者理解率达 85%。Yu 等<sup>[46]</sup>开发的 AIPatient 系统通过 Agent 协作, 基于电子病历 (electronic health record, EHR) 模拟患者交互, 真实度高达 90%, 为临床培训提供新工具。

在诊断与决策支持方面, Agent 的协作能力进一步凸显。Qiu 等<sup>[43]</sup>综述显示, Agent 系统通过任务分工优化诊断流程, 部分场景下准确率达 90%; Zou 等<sup>[47]</sup>指出, Agentic AI 作为医疗“队友”, 在模拟场景中诊断错误率降至 5%, 展现协同决策的潜力; Karunanayake 等<sup>[44]</sup>强调 Agent 系统在动态任务 (如急诊响应) 中的优势, 协作准确率达 88%; Yuan 等<sup>[48]</sup>则报告, Agent 优化患者管理工作流, 模拟效率提升 25%; Dolant 等<sup>[49]</sup>提出的自适应决策框架, 在术中决策支持模拟中准确率约 80%, 为麻醉等实时场景提供理论基础。然而, Agent 系统的计算复杂性和高资源需求限制实时部署, 数据隐私、可解释性及临床验证不足也成为瓶颈。此外, 部分研究如 BALROG (benchmarking agentic LLM and VLM reasoning on games) 虽提供协作推理启发, 但未直接应用于医疗, 迁移潜力待探索<sup>[50]</sup>。未来, 通过轻量化模型设计、多模态数据融合和标准化协议, Agent 有望实现医疗全流程的智能化协同, 为精准医疗注入新动力。

尽管麻醉智能化技术和 Agent 系统在各自领域取得显著进展, 但在围术期麻醉管理中仍面临诸多局限。现有方法多聚焦单一任务, 如剂量推荐或状态监控, 缺乏整合数据采集、风险预测、决策优化和结果验证的全流程解决方案, 难以适应患者个体差异和动态变化的临床环境。此外, 麻醉领域对 Agent 的系统性研究较为匮乏, 智能化技术的潜力尚未充分挖掘。

## 2 MorpheusAPI 框架

MorpheusAPI 框架基于平行大模型<sup>[51]</sup>, 通过麻醉执行大模型和麻醉影子大模型之间的协同进化以实现围术期麻醉智能管理。该架构包含感知 Agent、预测 Agent、决策 Agent、验证 Agent 和中央协调 Agent 共 5 个 Agents。MorpheusAPI 中的每个 Agent 都基于预训练的麻醉执行大模型如图 1 所示, 并通过微调、后蒸馏等技术进行优化, 以确保高效性和鲁棒性。同时, 部署在云端的麻醉影子大模型与各 Agent 运行的执行大模型进行持续互动, 以实现性能的闭环优化。

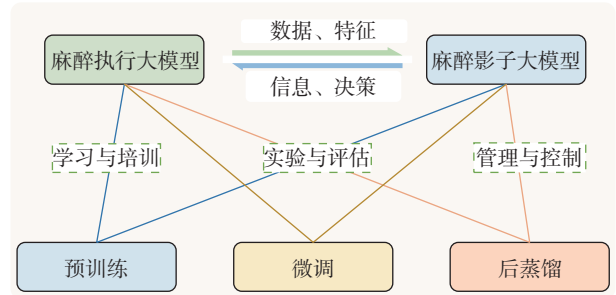


图 1 基于平行大模型的 MorpheusAPI 架构  
Fig. 1 MorpheusAPI architecture based on parallel deep foundation models

### 2.1 Agent 架构设计

麻醉执行大模型 ( $M_{core}$ ) 是一个在海量通用语料和医疗数据集 (如 MIMIC-IV) 上预训练的大型语言模型, 具备丰富的医学知识和语义推理能力。通过领域特定的麻醉数据进行微调,  $M_{core}$  被赋予理解和处理麻醉相关任务的能力, 并部署在各个 Agent 内部, 直接处理实时的临床数据和任务。为满足术中低延迟的要求, 执行大模型还会通过后蒸馏以减小模型体积和推理延迟。

麻醉影子大模型 ( $M_{shadow}$ ) 部署在云端, 与执行大模型结构相同或更大, 其主要职责是监控和分析各个执行大模型在实际运行过程中的数据, 并利用在线学习等技术持续优化模型参数和 Agent 的协作策略, 作为执行大模型的“教练”和“评估者”。

感知 Agent (perception Agent, PA) 的核心在于对多模态医疗数据  $D$  的高效解析和特征提取。其内部构建包括多模态数据接收与预处理模块、基于微调后执行大模型的特征工程与嵌入模块 (生成  $D_f, D_s$ ) 以及符合 MCP 协议的数据存储接口。 $D_f$  为结构化特征向量,  $D_s$  为对应的语义描述空间输入。为持续优化感知能力, 感知 Agent 会定期或在特定事件触发时, 将处理后的原始数据、提取的特征和嵌入信息发送至影子大模型。影子大模型分析这些数据, 评估感知 Agent 的特征提取效率和信息损失情况, 并为执行大模型提供微调建议或更新后的模型参数, 以提升其对复杂多模态数据的理解能力。

预测 Agent (prediction Agent, PDA) 的目标是基于感知 Agent 提供的特征和描述, 预测患者未来可能发生的风险和状态趋势。其内部构建包括数据接收与融合模块、基于微调后执行大模型和 CoT 的风险预测模型 (生成  $R$ )、基于微调后执行大模型的时间序列预测模型 (生成  $T$ ) 以及结果封装与存储接口。为提升预测的准确性, 预测 Agent 会将历史的输入数据 ( $D_f, D_s$ ) 和预测结果 ( $R, T$ ) 及

其对应的实际临床结果反馈给影子大模型。影子大模型通过对比预测结果和真实情况, 评估预测 Agent 的性能, 并优化预测模型的参数或 CoT 的提示策略。

决策 Agent (decision Agent, DA) 的职责是生成合理的麻醉管理建议。其内部构建包括数据接收与情境理解模块 (基于微调后执行大模型和 CoT)、基于强化学习模型或规则系统的决策模型 (以执行大模型作为策略或价值函数的一部分) 以及建议生成与存储接口 (生成  $Q_d$ )。为指导决策 Agent 学习更优的策略, 决策 Agent 会将生成的决策 ( $Q_d$ )、执行该决策后的患者状态变化以及最终的临床结果发送至影子大模型。影子大模型评估决策的有效性和安全性, 并更新决策模型的参数或强化学习的奖励函数。

验证 Agent (validation Agent, VA) 的目标是评估决策 Agent 生成结果的可靠性和合理性。其内部构建包括数据整合模块、利用 RAG 和微调后执行大模型进行知识检索与推理的模块 (生成  $V_r$ ) 以及结果报告模块。为提升验证的质量, 验证 Agent 会将其验证过程中的检索到的知识片段、推理过程和最终的验证结果 ( $V_r$ ) 发送给影子大模型。影子大模型分析验证 Agent 的推理过程和判断依据, 评估其验证的准确性和效率, 并优化 RAG 的检索策略或提供更全面的知识库。验证 Agent 基于指南和医学数据库定期更新, 确保实时性。为应对知识库不全, 结合影子大模型进行后处理, 分析运行数据, 降低不完整信息影响。

中央协调 Agent (center coordination Agent, CCA) 负责任务调度和结果整合。其内部构建包括临床优先级识别模块 CCA\_Prior、任务调度模块 (基于 MCP 和图神经网络) 以及结果整合模块 CCA\_Int。为提升整体系统的效率, CCA 会将任务调度的效率、Agent 之间的协作模式以及最终的综合建议和临床反馈发送给影子大模型。影子大模型分析这些信息, 评估当前的协作策略是否最优, 并为 CCA 提供优化的调度方案及 Agent 之间的通信策略。

MorpheusAPI 中的每个智能 Agent 都遵循通用的架构, 该架构的核心在于实现复杂动态决策和智能协作。图 2 为 Agent 内部结构。如图 2 所示, Agent 包含以下关键结构: 1) 输入接口, 负责接收和标准化来自系统内外部的各类数据, 如实时生理信号、历史病历、其他 Agent 输出及外部知识库; 2) 记忆模块, 用于支持上下文理解和经验学习, 通常包括存储近期信息的短期记忆 (如

上下文窗口) 和存储长期知识的长期记忆 (如向量数据库或知识图谱); 3) 核心处理单元, 作为 Agent 的智能核心, 依据输入和记忆执行认知任务, 其内部包含状态管理、存储知识与模型的模块 (如  $M_{core}$ 、知识图谱、强化学习策略、RAG)、执行推理与计算的模块 (利用 CoT 和 RAG) 以及进行行为选择与控制的模块; 4) 输出接口, 负责将 Agent 的处理结果以标准化格式传递给其他组件或用户。

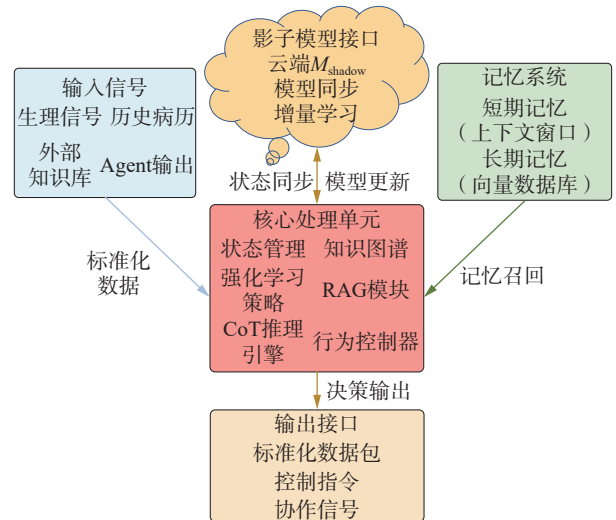


图 2 Agent 内部结构  
Fig. 2 Structure of Agent

影子模型交互接口则作为辅助机制, 支持 Agent 通过与云端  $M_{shadow}$  的双向通信实现持续优化。

输入接口的设计目标是确保异构数据能够被 Agent 的核心处理单元有效利用; 记忆模块通过短期和长期存储机制, 使 Agent 能够理解当前情境并从历史经验中学习; 核心处理单元是 Agent 执行智能任务的关键, 其内部的各个模块协同工作, 从状态维护到知识应用, 再到最终的行为决策; 输出接口保证 Agent 结果的有效传递。环境交互和影子模型交互接口则分别扩展 Agent 的物理交互能力和持续学习优化的途径。

通过以上结构, MorpheusAPI 中的每个 Agent 能够有效地执行其特定任务, 并通过记忆机制和与影子模型的交互实现持续的性能提升和知识进化, 最终支持整个系统的智能协作。

## 2.2 Agent 动态协作机制

MorpheusAPI 采用 Agent 间的动态协作机制, 以实现围术期麻醉状态的实时智能管理, 具体如图 3 所示。该机制的核心是一个持续监测、智能诊断与敏捷响应的闭环流程, 其算法框架如算法 1 所示。

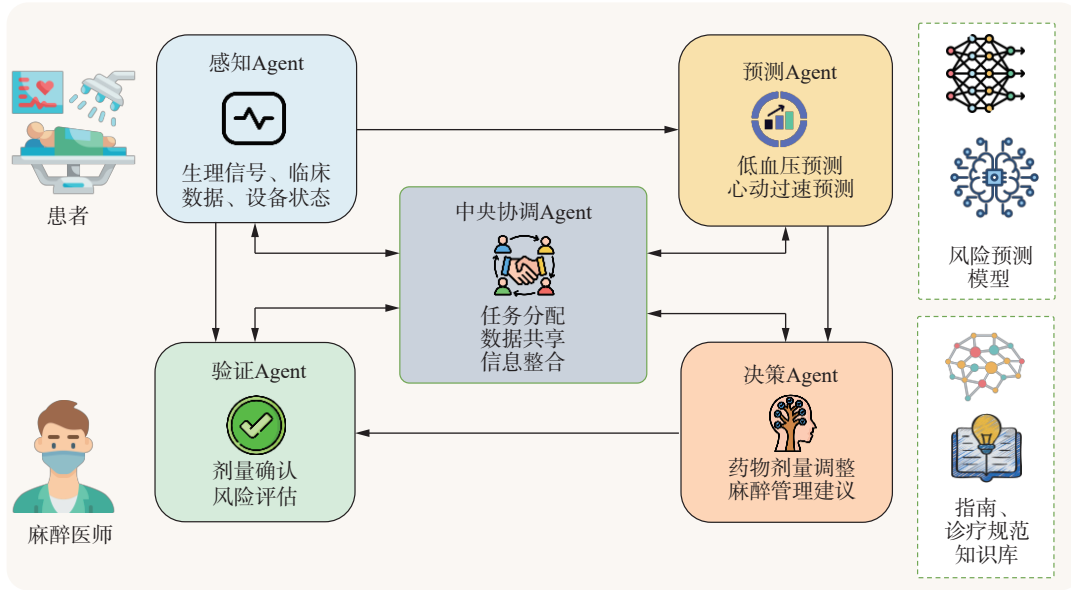


图 3 MorpheusAPI 多 Agent 协作框架

Fig. 3 Multi-agent collaboration framework of MorpheusAPI

算法 1 MorpheusAPI Agent 动态协作机制

输入: 多模态数据  $D$

输出: 实施综合建议  $A$

- 1)  $M_{core} \leftarrow \text{PreTrain}()$ ;  $V_a \leftarrow \text{“假”}$
- 2) while 实时麻醉管理中 do
- 3)  $D_f, D_s \leftarrow \text{PA}(D) \xrightarrow{\text{MCP}} \text{KafkaPool}$
- 4)  $P_p \leftarrow \text{CCA\_Prior}(\text{CCA\_Sub}(\text{KafkaPool}))$
- 5)  $R, T \leftarrow \text{PDA\_Pred}(D_f, D_s) \xrightarrow{\text{CoT}} \text{KafkaPool}$
- 6) if  $\text{CheckAnomaly}(R, T, D_f, D_s) == \text{“真”}$  then
- 7)  $\text{Trigger}(\text{DA}, R, T)$ ;  $V_a \leftarrow \text{真}$
- 8) while  $V_a == \text{“真”}$  and 实时麻醉管理中 do
- 9)  $Q_d \leftarrow \text{DA\_Dec}(D_f, R, T) \xrightarrow{\text{CoT}} \text{KafkaPool}$
- 10)  $V_r \leftarrow \text{VA\_Val}(D_f, R, Q_d) \xrightarrow{\text{RAG}} V_r$
- 11) if  $V_r == \text{“通过”}$  then
- 12)  $A \leftarrow \text{CCA\_Int}(D_f, R, Q_d, V_r) \xrightarrow{\text{Push}}$  医生界面
- 13)  $V_a \leftarrow \neg \text{CheckAnomaly}(R, T, D_f, D_s)$
- 14) else
- 15)  $\text{CCA\_Adjust}(P_p, \text{PDA}, \text{DA})$
- 16) end if
- 17)  $D_f, D_s, R, T \leftarrow \text{SensePredict}()$   $\Delta$  合并感知与预测
- 18)  $V_a \leftarrow \text{CheckAnomaly}(R, T, D_f, D_s)$
- 19) end while
- 20) end if
- 21) end while
- 22)  $M_{core} \leftarrow \text{MetaOpt}(M_{core})$ ;  $\text{VerEnh}(V_a)$
- 23) return 流程结束

系统首先通过预训练的麻醉执行大模型  $M_{core}$  完成初始化, 并将表征异常状态的布尔变量  $V_a$  初

始化为“假”。

在实时麻醉管理过程中, PA 持续采集包括生理信号、患者病史等多模态医疗数据  $D$ 。PA 对  $D$  进行处理, 提取关键特征向量  $D_f$  并生成相应的语义描述  $D_s$ 。处理后的数据遵循 MCP, 存储至 Kafka 数据池 (KafkaPool)。

CCA 通过订阅 KafkaPool, 获取相关数据,  $\text{CCA\_Sub}(\text{KafkaPool})$  利用优先级识别模块  $\text{CCA\_Prior}$  分析数据, 确定当前的临床优先级  $P_p$ 。随后, PDA 接收特征向量  $D_f$  和语义描述  $D_s$ , 并借助 CoT 推理生成风险评分  $R$  和未来状态趋势  $T$ , 预测结果同样被存储至 KafkaPool。

系统通过异常检测模块 (CheckAnomaly) 对风险评分  $R$ 、趋势预测  $T$  以及原始特征  $D_f$  和语义描述  $D_s$  进行综合评估, 判断当前麻醉状态是否存在异常。若检测到异常, 系统将触发 DA, 并将异常状态  $V_a$  置为“真”。在异常状态持续且实时麻醉管理进行中时, DA 基于  $D_f$ 、 $R$ 、 $T$ , 利用 CoT 推理生成具体的剂量调整或其他干预建议  $Q_d$ , 并将建议存储至 KafkaPool。VA 随后整合  $D_f$ 、 $R$ 、 $Q_d$ , 并通过 RAG 检索外部医学知识库进行验证, 产生验证结果  $V_r$ 。若验证结果  $V_r$  为“通过”, 则 CCA 通过整合各 Agent 的输出, 生成最终的综合建议  $A$ , 并推送至医生界面。同时, 异常状态  $V_a$  根据最新的感知和预测结果进行更新。若验证结果  $V_r$  为“未通过”, 则 CCA 调用调整模块  $\text{CCA\_Adjust}$  调整任务调度策略, 并重新触发预测和决策过程。系统通过 SensePredict 持续进行感知和预测, 并再次评估是否存在异常  $V_a$ , 直至异常解除。

最终,在整个实时麻醉管理流程结束后,系统通过元学习方法 MetaOpt( $M_{core}$ )优化核心模型  $M_{core}$  的性能,并通过验证增强模块 VerEnh( $V_a$ )提升验证过程的准确性。整个流程旨在实现对麻醉状态的实时监测、异常情况的快速响应以及对医生的智能决策支持。

### 3 围术期低血压风险管理案例研究

本章通过 Agentic Workflow 形式,描述 MorpheusAPI 框架中感知 Agent、预测 Agent、决策 Agent、验证 Agent 和中央协调 Agent 这 5 个 Agents 在围术期低血压风险管理中的自主协作。本案例基

于 MIMIC-IV 模拟数据,针对 ASA III 级患者,以一名 60 岁男性患者接受腹部手术为背景,旨在验证 MorpheusAPI 框架的工作流程。为探讨个体差异影响,未来将纳入不同 ASA 分级 (I-IV 级) 患者数据进行测试。Agent 依托麻醉执行大模型的语义推理能力,通过 MCP、CoT 和 RAG,整合多模态数据,预测低血压风险,生成干预建议,验证可靠性,并通过麻醉影子大模型优化性能。本章阐释 Agent 协作流程,图 4 以流程图形式可视化 Agentic Workflow,突出验证 Agent 的质量控制和影子大模型的闭环优化,为围术期麻醉管理提供高效支持。

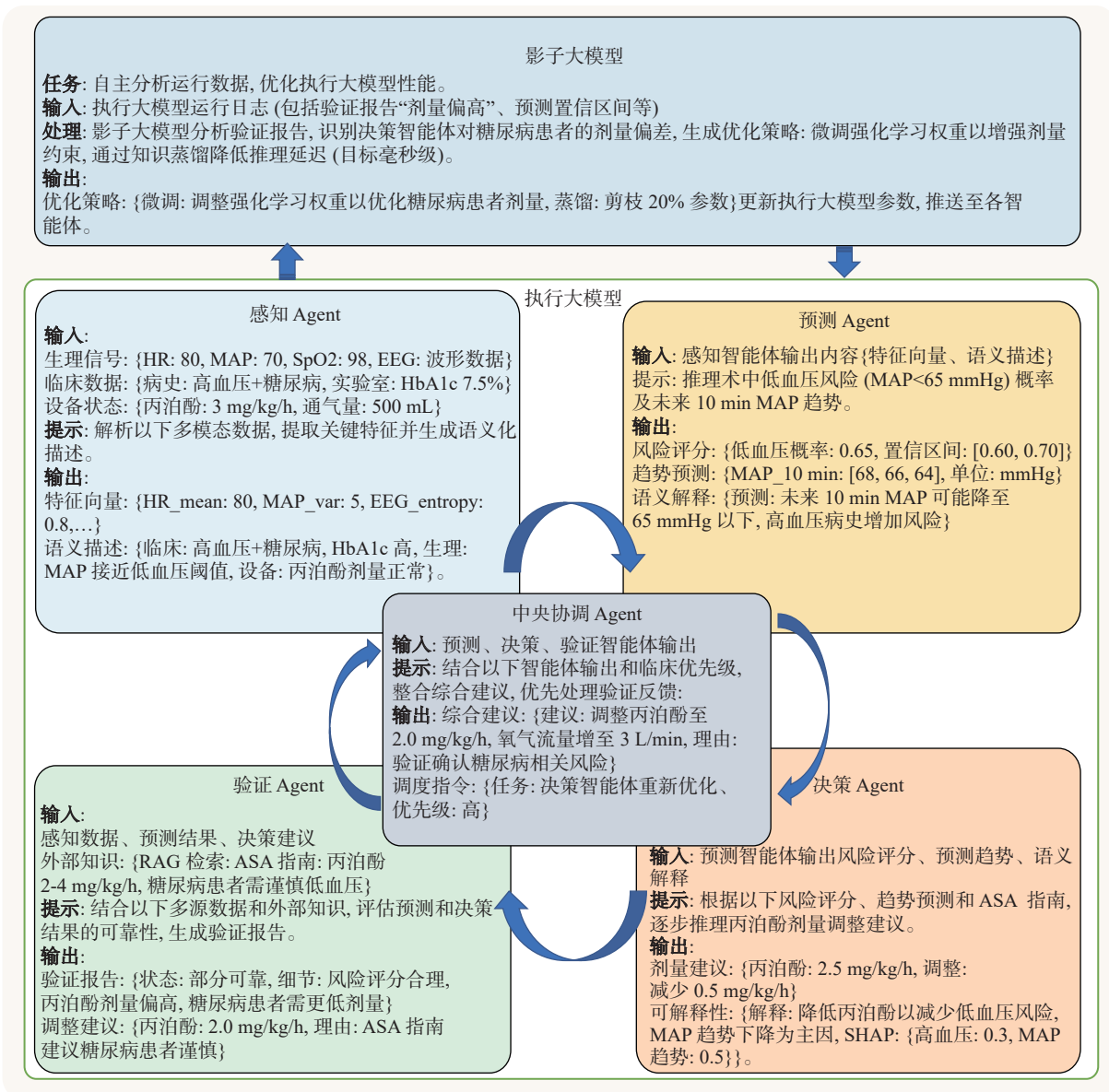


图 4 MorpheusAPI 围术期低血压风险管理案例研究

Fig. 4 Case study of perioperative hypotension risk management with MorpheusAPI

#### 3.1 案例背景

患者为 60 岁男性, 体重 70 kg, 美国麻醉医师

协会分级 III 级, 病史包括高血压和 2 型糖尿病, 接受 3 h 腹部手术。手术中, 监测设备包括麻醉

机、心电监护仪, 采用丙泊酚输注全静脉麻醉, 数据采集符合行业标准, 电子病历提供术前病史和实验室结果。临床目标为维持平均动脉压 (mean arterial pressure, MAP) 在 65~100 mmHg, 调整丙泊酚剂量 (初始 8 mg/(kg·h)) 以降低低血压风险。MorpheusAPI 框架通过 Agentic Workflow 协调 5 个 Agents 的协作, 依托 Kafka 数据池实现异步交互, 麻醉执行大模型驱动语义推理, 影子大模型分析运行数据优化性能。

### 3.2 Agentic Workflow 描述

在术中低血压风险管理中, 感知 Agent 首先采集和标准化多模态数据, 为后续 Agent 提供语义化输入。其处理的数据包括生理信号 (如心率 80 次/分钟、MAP 70 mmHg、血氧饱和度 98% 和脑电图)、电子病历中的高血压与糖尿病病史以及设备状态 (如丙泊酚输液 3 mg/(kg·h) 和通气量 500 mL)。感知 Agent 通过麻醉执行大模型的 MCP 整合这些数据, 解析非结构化病历生成语义描述, 同时对生理信号进行去噪, 提取时域、频域和非线性特征, 形成结构化特征向量和语义描述。这些输出存储至 Kafka 数据池, 延迟约 1 ms, 触发预测 Agent 和验证 Agent 的处理。

预测 Agent 订阅感知 Agent 的输出, 自主评估低血压风险并预测 MAP 趋势。其利用执行大模型的 CoT 提示, 基于患者病史和当前 MAP 值, 结合 MIMIC-IV 训练的医学知识, 通过 Transformer 模块预测未来 10 min MAP 趋势。预测结果显示低血压概率为 65%, MAP 下降至 64 mmHg, 并附带语义解释, 说明高血压病史是风险增高的主要因素。预测 Agent 将风险评分、趋势预测和解释存储至 Kafka 数据池, 延迟约 10 ms, 推送至决策 Agent 和验证 Agent, 继续协作流程。

决策 Agent 接收预测 Agent 的低血压风险评分, 结合感知 Agent 的麻醉深度指数 (bispectral index, BIS) 与 MAP 数据, 自主生成丙泊酚剂量建议。其通过 LLM 的 CoT 提示, 参考 ASA 指南, 融合强化学习优化剂量, 从 3.0 mg/(kg·h) 降至 2.5 mg/(kg·h), 平衡麻醉深度 (BIS 50~60)、MAP 稳定性和副作用。SHAP 值解释显示, BIS 下降与 MAP 趋势为主要依据。建议与解释存储至 Kafka 数据池, 延迟约 10 ms, 供验证 Agent 评估。

验证 Agent 负责评估预测和决策结果的可靠性, 订阅感知 Agent、预测 Agent 和决策 Agent 的输出。其通过执行大模型的 MCP 整合多源数据, 包括特征向量、风险评分和剂量建议, 并结合 RAG 提示检索 ASA 指南, 确认风险评分合理, 但

发现丙泊酚剂量 2.5 mg/(kg·h) 对糖尿病患者偏高。验证 Agent 生成验证报告, 指出需进一步降低剂量至 2.0 mg/(kg·h), 以符合指南对糖尿病患者的谨慎要求。报告和调整建议存储至 Kafka 数据池, 延迟约 15 ms, 因 RAG 检索增加开销。验证报告为协调 Agent 提供指导, 并为影子大模型的优化提供偏差反馈, 形成关键反馈循环。

中央协调 Agent 整合预测 Agent、决策 Agent 和验证 Agent 的输出, 生成综合临床建议并动态调度任务。其通过执行大模型的 MCP 优先采纳验证 Agent 的调整建议, 生成综合建议: 将丙泊酚调整至 2.0 mg/kg/h, 同时增加氧气流量至 3 L/min, 以应对糖尿病相关的低血压风险。协调 Agent 还触发决策 Agent 重新优化剂量, 调度指令和建议存储至 Kafka 数据池, 延迟约 5 ms, 并通过 Web-Socket 推送至医生界面, 为临床干预提供支持。

影子大模型分析 Kafka 运行日志, 优化执行大模型性能。其关注验证报告中的剂量偏差, 识别决策 Agent 对糖尿病患者的剂量偏高问题, 生成优化策略: 微调强化学习权重以增强剂量约束, 并通过后蒸馏降低推理延迟。优化后的执行大模型参数推送至各 Agent, 完成闭环改进, 处理时间约 1 s, 采用云端异步方式。这一反馈循环确保系统适应复杂病例, 提升长期性能。

Agentic Workflow 的协作机制通过 Agent 的自主性和分布式交互得以实现。感知 Agent 通过 MCP 标准化数据, 预测 Agent 和决策 Agent 利用 CoT 生成前瞻性输出, 验证 Agent 通过 MCP 和 RAG 确保可靠性, 协调 Agent 通过 MCP 整合建议, 影子大模型驱动优化。Kafka 数据池支持毫秒级交互, MCP 整合上下文, CoT 和 RAG 增强推理和知识检索, 推理延迟约 10~15 ms, 共同实现高效、低延迟的临床支持。

### 3.3 结果与讨论

MorpheusAPI 的 Agentic Workflow 在约 0.4 s 内完成低血压风险管理, 满足术中实时性要求, Kafka 交互延迟约 1 ms, 推理延迟约 10~15 ms。预测 Agent 生成 65% 低血压风险评分, 准确反映患者高血压和糖尿病病史, MAP 趋势预测为后续决策提供依据。决策 Agent 建议将丙泊酚从 3.0 mg/(kg·h) 降至 2.5 mg/(kg·h), 验证 Agent 通过 MCP 整合数据和 RAG 检索 ASA 指南, 进一步调整至 2.0 mg/(kg·h), 符合糖尿病患者的需求。协调 Agent 整合验证反馈, 生成综合建议, 助力医生快速干预, 假设后续 MAP 维持在 65 mmHg 以上。影子大模型利用验证报告优化决策 Agent 的强化学习模块, 提升未来剂量精度。

本案例凸显 Agentic Workflow 在围术期麻醉管理中的应用优势。感知 Agent 通过 MCP 标准化多模态数据,为预测和决策 Agent 提供可靠输入;预测和决策 Agent 通过 CoT 生成风险评估和剂量建议;验证 Agent 通过 MCP 和 RAG 确保结果可靠性;协调 Agent 通过 MCP 整合输出并动态调度任务,形成高效闭环。执行大模型的语义推理能力通过 MCP 整合上下文,CoT 和 RAG 增强推理,确保低延迟,适合术中应用。影子大模型通过分析验证报告优化系统,提升对复杂病例的适应性。验证 Agent 的质量控制和可解释性增强医生信任,确保患者安全。然而,RAG 检索依赖知识库更新,受限于最新指南的可用性。

MorpheusAPI 通过多 Agent 协作(感知、预测、决策、验证、中央协调)和 MCP 协议实现全流程麻醉管理,单一低血压案例展示了工作流的初步可行性。然而,作为案例研究,单一案例无法代表多场景(如不同手术类型、ASA 分级)下的鲁棒性与普适性,缺乏真实临床数据验证。案例未涉及量化指标(如预测准确率),仅用于 workflow 演示。未来通过多中心真实临床数据试验,验证框架在复杂场景下的适用性与可靠性。

### 3.4 局限性与挑战

本研究通过单一低血压案例验证了 MorpheusAPI 的多 Agent 协作 workflow,展示了感知、预测、决策、验证和中央协调 Agent 的设计可行性。然而,因仅基于单一场景(低血压管理),未涵盖复杂临床场景,且缺乏真实临床数据集和量化指标,框架的鲁棒性、普适性和临床有效性尚待验证。

在可解释性方面,尽管 CoT 提示增强了低血压案例的决策透明性,但大模型复杂推理链缺乏直观语义解释,难以获得临床医生的充分信任<sup>[52-53]</sup>。此外,训练数据偏差可能导致 AI 幻觉(如错误剂量建议),引发伦理风险,威胁患者安全<sup>[54-55]</sup>。验证 Agent 通过 RAG 缓解偏见,但因本研究未使用真实临床数据,伦理影响的评估尚不充分。

为应对上述局限,未来将通过多中心临床合作获取真实数据,开发推理可视化工具以增强决策透明性,引入多源知识库和联邦学习优化 RAG(通过分布式数据训练保护患者隐私并提升模型泛化能力),优化 MCP 和影子大模型的轻量化设计,并融入价值敏感设计以降低伦理风险,从而提升框架的可靠性和临床适用性。

## 4 结束语

MorpheusAPI 通过多 Agent 协作(感知、预

测、决策、验证、中央协调)与 MCP,实现了围术期麻醉管理的智能化 workflow,在低血压案例中展示了实时性和适用性。核心优势在于分布式 Agent 协作、闭环优化和 Kafka 数据池支持的毫秒级交互,显著提升临床效率。影子大模型通过验证反馈优化主模型性能,增强复杂病例的适应性。

未来,框架需解决单一场景和数据不足的局限性,通过多中心临床试验获取真实数据以验证鲁棒性和普适性;开发推理可视化工具以提升决策透明性,解决大模型推理封闭性问题;引入联邦学习保护数据隐私并优化 RAG,结合价值敏感设计降低伦理风险,从而进一步提升框架的可靠性和临床适用性,推动智能化麻醉管理的发展。

## 参考文献:

- [1] ZENG Shiyue, QING Qi, XU Wei, et al. Personalized anesthesia and precision medicine: a comprehensive review of genetic factors, artificial intelligence, and patient-specific factors[J]. *Frontiers in medicine*, 2024, 11: 1365524.
- [2] CANNESON M, ANI F, MYTHEN M M, et al. Anaesthesiology and perioperative medicine around the world: different names, same goals[J]. *British journal of anaesthesia*, 2015, 114(1): 8-9.
- [3] SESSLER D I, MEYHOFF C S, ZIMMERMAN N M, et al. Period-dependent associations between hypotension during and for four days after noncardiac surgery and a composite of myocardial infarction and death: a substudy of the POISE-2 trial[J]. *Anesthesiology*, 2018, 128(2): 317-327.
- [4] FUTIER E, LEFRANT J Y, GUINOT P G, et al. Effect of individualized vs standard blood pressure management strategies on postoperative organ dysfunction among high-risk patients undergoing major surgery: a randomized clinical trial[J]. *Journal of the American medical association*, 2017, 318(14): 1346-1357.
- [5] NAIR B G, GABEL E, HOFER I, et al. Intraoperative clinical decision support for anesthesia: a narrative review of available systems[J]. *Anesthesia & analgesia*, 2017, 124(2): 603-617.
- [6] SHICKEL B, TIGHE P J, BIHORAC A, et al. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record(EHR) analysis[J]. *IEEE journal of biomedical and health informatics*, 2018, 22(5): 1589-1604.
- [7] SHORTLIFFE E H, SEPÚLVEDA M J. Clinical decision support in the era of artificial intelligence[J]. *Journal of the American medical association*, 2018, 320(21): 2199.
- [8] THIRUNAVUKARASU A J, TING D S J, ELANGOVAN K, et al. Large language models in medicine[J]. *Nature medicine*, 2023, 29(8): 1930-1940.
- [9] SAFDARI R, MALAK J S, MOHAMMADZADEH N, et al. A multi agent based approach for prehospital emergency management[J]. *Bulletin of emergency and trauma*, 2017, 5(3): 171-178.

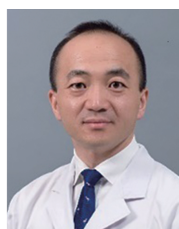
- [10] 赵毅飞, 申乐, 叶佩军, 等. 平行麻醉: 从麻醉自动化走向智慧型全周期麻醉平台[J]. 智能科学与技术学报, 2024, 5(2): 234–246.  
ZHAO Yifei, SHEN Le, YE Peijun, et al. Parallel anesthesia: from anesthesia automation to intelligent full-cycle anesthesia platform[J]. Chinese journal of intelligent science and technology, 2023, 5(2): 234–246.
- [11] 王惠珍, 张捷, 俞怡, 等. 平行手术室: 围术期护理流程与智慧手术平台管理的新模式[J]. 模式识别与人工智能, 2023, 36(10): 867–876.  
WANG Huizhen, ZHANG Jie, YU Yi, et al. Parallel operating rooms: a new model of perioperative nursing process and smart surgical platform management[J]. Pattern recognition and artificial intelligence, 2023, 36(10): 867–876.
- [12] LIN Fei, GAO T, SUN Dali, et al. Parallel medical devices and instruments: integrating edge and cloud intelligence for smart treatment and health systems[J]. IEEE/CAA journal of automatica sinica, 2025, 12(4): 651–654.
- [13] PATNAIK S. S, HOFFMANN U. Comparison of ChatGPT vs. Bard to anesthesia-related queries[EB/OL]. (2023–06–30)[2025–04–20]. <https://www.medrxiv.org/content/10.1101/2023.06.29.23292057v1>.
- [14] WANG Zhonghai, JIANG Jie, ZHAN Yibing, et al. Hypnos: a domain-specific large language model for anesthesiology[J]. Neurocomputing, 2025, 624: 129389.
- [15] ZHOU Bohao, ZHAN Yibing, WANG Zhonghai, et al. Benchmarking medical LLMs on anesthesiology: a comprehensive dataset in Chinese[J]. IEEE transactions on emerging topics in computational intelligence, 2025, 9(4): 3057–3071.
- [16] GUTHRIE E, LEVY D, DEL CARMEN G. The operating and anesthetic reference assistant(OARA): a fine-tuned large language model for resident teaching[J]. The American journal of surgery, 2024, 234: 28–34.
- [17] Anthropic. 2024. Introducing the Model Context Protocol[EB/OL]. (2024–11–25)[2025–04–06]. <https://www.anthropic.com/news/model-context-protocol>.
- [18] WEI J, WANG Xuezhi, SCHUURMANS D, et al. Chain-of-thought prompting elicits reasoning in large language models[EB/OL]. (2023–01–10)[2025–04–06]. <https://arxiv.org/abs/2201.11903>.
- [19] LEWIS P, ETHAN P, ALEKSANDRA P, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks[EB/OL]. (2020–05–22)[2025–03–20]. <https://arxiv.org/abs/2005.11401>.
- [20] YANG Jing, TIAN Yonglin, LIN Fei, et al. From RAG to ARG: agent reinforced generation for agentic intelligence[J]. The international journal of intelligent control and systems, 2025, 30(1): 76–82.
- [21] ZHANG Jiayi, XIANG Jinyu, YU Zhaoyang, et al. AFlow: automating agentic workflow generation[EB/OL]. (2025–04–15)[2025–06–15]. <https://arxiv.org/abs/2410.10762>.
- [22] CONNOR C W. Artificial intelligence and machine learning in anesthesiology[J]. Anesthesiology, 2019, 131(6): 1346–1359.
- [23] JOHN DOYLE D, DAHABA A A, LEMANACH Y. Advances in anesthesia technology are improving patient care, but many challenges remain[J]. BMC anesthesiology, 2018, 18(1): 39.
- [24] HASHIMOTO D A, WITKOWSKI E, GAO Lei, et al. Artificial intelligence in anesthesiology[J]. Anesthesiology, 2020, 132(2): 379–394.
- [25] SCHMIERER T, LI Tianning, LI Yan. Harnessing machine learning for EEG signal analysis: innovations in depth of anaesthesia assessment[J]. Artificial intelligence in medicine, 2024, 151: 102869.
- [26] HASHEMI S, YOUSEFZADEH Z, ALI ABIN A, et al. Machine learning-guided anesthesiology: a review of recent advances and clinical applications[J]. Journal of cellular & molecular anesthesia, 2024, 9(1): 1–29.
- [27] SCHAMBERG G, BADGELEY M, MESCHÉDE-KRASA B, et al. Continuous action deep reinforcement learning for propofol dosing during general anesthesia[J]. Artificial intelligence in medicine, 2022, 123: 102227.
- [28] LEE H, YOON H K, KIM J, et al. Development and validation of a reinforcement learning model for ventilation control during emergence from general anesthesia[J]. NPJ digital medicine, 2023, 6(1): 145.
- [29] OpenAI, ACHIAM J, ADLER S, et al. GPT-4 Technical Report[EB/OL]. (2024–03–04)[2025–04–20]. <https://arxiv.org/abs/2303.08774>
- [30] SINGHAL K, AZIZI S, TU Tao, et al. Large language models encode clinical knowledge[J]. Nature, 2023, 620(7972): 172–180.
- [31] LUO Renqian, SUN Liai, XIA Yingce, et al. BioGPT: generative pre-trained transformer for biomedical text generation and mining[J]. Briefings in bioinformatics, 2022, 23(6): bbac409.
- [32] 张志成, 王静, 张阳, 等. OrthoGPT: 面向精准诊疗的多模态骨科大模型[J]. 智能科学与技术学报, 2024, 6(3): 338–346.  
ZHANG Zhicheng, WANG Jing, ZHANG Yang, et al. OrthoGPT: multimodal generative pre-trained transformer models for precise diagnosis and treatment of orthopedics[J]. Chinese journal of intelligent science and technology, 2024, 6(3): 338–346.
- [33] 肖建力, 许东舟, 王浩, 等. 医疗领域的大型语言模型综述[J]. 智能系统学报, 2025, 20(3): 530–547.  
XIAO Jianli, XU Dongzhou, WANG Hao, et al. Survey of large language models in healthcare[J]. CAAI transactions on intelligent systems, 2025, 20(3): 530–547.
- [34] 刘学博, 户保田, 陈科海, 等. 大模型关键技术与未来发展方向——从 ChatGPT 谈起[J]. 中国科学基金, 2023, 37(5): 758–766.  
LIU Xuebo, HU Baotian, CHEN Kehai, et al. Key technologies and future development directions of large models — starting from ChatGPT[J]. Science foundation in China, 2023, 37(5): 758–766.
- [35] 马武仁, 弓孟春, 戴辉, 等. 以 ChatGPT 为代表的大语言模型在临床医学中的应用综述[J]. 医学信息学杂志, 2023, 44(7): 9–17.  
MA Wuren, GONG Mengchun, DAI Hui, et al. A comprehensive review of the applications of large language models in clinical medicine with ChatGPT as a representative [J]. Journal of medical informatics, 2023, 44(7): 9–17.
- [36] ANGEL M C, RINEHART J B, CANNESON M P, et al.

- Clinical knowledge and reasoning abilities of AI large language models in anesthesiology: a comparative study on the American board of anesthesiology examination[J]. *Anesthesia & analgesia*, 2024, 139(2): 349–356.
- [37] KHAN A A, YUNUS R, SOHAIL M, et al. Artificial intelligence for anesthesiology board-style examination questions: role of large language models[J]. *Journal of cardiothoracic and vascular anesthesia*, 2024, 38(5): 1251–1259.
- [38] RATNAGANDHI J A, GODAVARTHY P, GNANESWARAN M, et al. Enhancing anesthetic patient education through the utilization of large language models for improved communication and understanding[J]. *Anesthesia research*, 2025, 2(1): 4.
- [39] KUO F H, FIERSTEIN J L, TUDOR B H, et al. Comparing ChatGPT and a single anesthesiologist's responses to common patient questions: an exploratory cross-sectional survey of a panel of anesthesiologists[J]. *Journal of medical systems*, 2024, 48(1): 77.
- [40] FENG Xiang, JIANG Wentao, WANG Zhenmao, et al. AnesBench: multi-dimensional evaluation of LLM reasoning in anesthesiology[EB/OL]. (2025-04-03)[2025-04-20]. <https://arxiv.org/abs/2504.02404>.
- [41] WANG Zhonghai, ZHOU Bohao, LIU Weifeng. ATSCOT: chain of thought for structuring anesthesia medical records[C]//2024 International Conference on New Trends in Computational Intelligence. Qingdao: IEEE, 2024: 145–149.
- [42] WANG Zhonghai, JIANG Jie, ZHAN Yibin, et al. Towards training a Chinese large language model for anesthesiology[EB/OL]. (2024-03-05)[2025-02-16]. <https://arxiv.org/abs/2403.02742>
- [43] QIU Jianing, LAM K, LI Guohao, et al. LLM-based agentic systems in medicine and healthcare[J]. *Nature machine intelligence*, 2024, 6(12): 1418–1420.
- [44] KARUNANAYAKE N. Next-generation agentic AI for transforming healthcare[J]. *Informatics and health*, 2025, 2(2): 73–83.
- [45] SUDARSHAN M, SHIH S, YEE E, et al. Agentic LLM workflows for generating patient-friendly medical reports[EB/OL]. (2024-08-05)[2025-04-20]. <https://arxiv.org/abs/2408.01112>.
- [46] YU HuiZhi, ZHOU Jiayan, LI Lingyao, et al. AIPatient: simulating patients with EHRs and LLM powered agentic workflow[EB/OL]. (2024-09-27)[2025-02-15]. <https://arxiv.org/abs/2409.18924>.
- [47] ZOU J, TOPOL E J. The rise of agentic AI teammates in medicine[J]. *The lancet*, 2025, 405(10477): 457.
- [48] YUAN Han. Agentic large language models for healthcare: current progress and future opportunities[J]. *Medicine advances*, 2025, 3(1): 37–41.
- [49] DOLANT A, KUMAR P. Agentic LLM framework for adaptive decision discourse[EB/OL]. (2025-02-16)[2025-04-20]. <https://arxiv.org/abs/2502.10978>.
- [50] PAGLIERI D, CUPIAL B, COWARD S, et al. BALROG: benchmarking agentic LLM and VLM reasoning on games[EB/OL]. (2025-04-14)[2025-05-05]. <https://arxiv.org/abs/2411.13543v2>.
- [51] TIAN Yonglin, LIN Fei, WANG Cong, et al. Parallel deep foundation model: a co-evolution framework for analogical imagination and embodied cognition of parallel intelligence[J]. *The international journal of intelligent control and systems*, 2025, 30(1): 83–90.
- [52] 陈小平. 大模型关联度预测的形式化和语义解释研究[J]. *智能系统学报*, 2023, 18(4): 894–900.
- CHEN Xiaoping. Research on formalization and semantic interpretations of correlation degree prediction in large language models[J]. *CAAI transactions on intelligent systems*, 2023, 18(4): 894–900.
- [53] 陈小平. 人工智能中的封闭性和强封闭性: 现有成果的能力边界、应用条件和伦理风险[J]. *智能系统学报*, 2020, 15(1): 114–120.
- CHEN Xiaoping. Criteria of closeness and strong closeness in artificial intelligence—limits, application conditions and ethical risks of existing technologies[J]. *CAAI transactions on intelligent systems*, 2020, 15(1): 114–120.
- [54] 古天龙, 马露, 李龙, 等. 符合伦理的人工智能应用的价值敏感设计: 现状与展望[J]. *智能系统学报*, 2022, 17(1): 2–15.
- GU Tianlong, MA Lu, LI Long, et al. Value sensitive design of ethical-aligned AI applications: current situation and prospect[J]. *CAAI transactions on intelligent systems*, 2022, 17(1): 2–15.
- [55] 黄河燕, 李思霖, 兰天伟, 等. 大语言模型安全性: 分类、评估、归因、缓解、展望[J]. *智能系统学报*, 2025, 20(1): 2–32.
- HUANG Heyan, LI Silin, LAN Tianwei et al. A survey on the safety of large language model: classification, evaluation, attribution, mitigation and prospect[J]. *CAAI transactions on intelligent systems*, 2025, 20(1): 2–32.

### 作者简介:



王静, 博士研究生, 主要研究方向为平行医疗理论与方法、智慧医疗系统的构建与应用。E-mail: [wangjing2014@ia.ac.cn](mailto:wangjing2014@ia.ac.cn)。



申乐, 教授, 博士, 中国医学科学院北京协和医院麻醉科主任, 主要研究方向为基于麻醉大数据和人工智能的医疗质量提升与患者安全, 疼痛与瘙痒的机制与生物干预研究。主持国家级省部级课题 10 项, 发表学术论文 60 余篇。E-mail: [pumchshenle@163.com](mailto:pumchshenle@163.com)。



王飞跃, 教授, 博士, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室主任, 主要研究方向为平行系统的方法与应用、社会计算、平行智能以及知识自动化。先后当选 IEEE、IFAC、AAAS、ASME 等多项国际会士, 获得国家自然科学二等奖 1 项。E-mail: [feiyue.wang@ia.ac.cn](mailto:feiyue.wang@ia.ac.cn)。