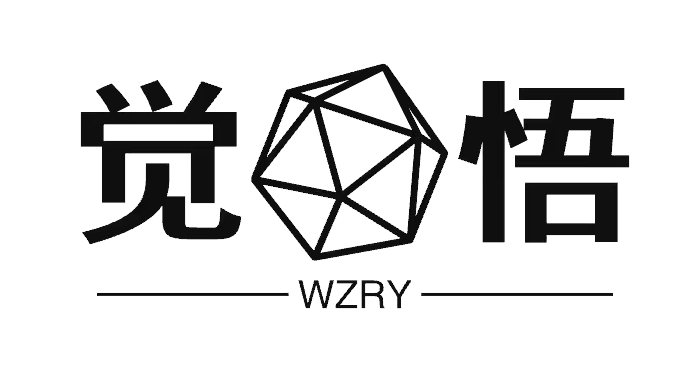
|  |  |
| --- | --- |
|  | **第十八届全国大学生软件创新大赛-软件设计创新赛** |
| **文档编号：SWC2025-第一滴血** |

****

**觉悟-王者荣耀智能体**

**Awakening: Honor of Kings Intelligent Agent**

**项目计划书**

**Version: 1.0**



**第一滴血**

**2024-12-29**

**All Rights Reserved**

目录

[1 项目概述 1](#_Toc186311163)

[1.1 项目背景 1](#_Toc186311164)

[1.2 项目内容 3](#_Toc186311165)

[1.3 创新点 4](#_Toc186311166)

[1.3.1 技术创新点 4](#_Toc186311167)

[1.3.2 功能创新点 6](#_Toc186311168)

[1.3.3 其他创新点 7](#_Toc186311169)

[2 技术方案 8](#_Toc186311170)

[2.1 总体架构 8](#_Toc186311171)

[2.2 功能详述 10](#_Toc186311172)

[2.2.1 系统流程图 10](#_Toc186311173)

[2.2.2 核心功能 12](#_Toc186311174)

[2.3 关键技术 13](#_Toc186311175)

[2.4 其他相关技术 14](#_Toc186311176)

[2.4.1 相关技术 14](#_Toc186311177)

[2.4.2 环境部署 14](#_Toc186311178)

[项目计划 15](#_Toc186311179)

[2.5 可行性分析 15](#_Toc186311180)

[2.5.1 技术可行性分析 15](#_Toc186311181)

[2.5.2 资源可行性分析 16](#_Toc186311182)

[2.5.3 市场可行性分析 17](#_Toc186311183)

[2.6 排期规划 17](#_Toc186311184)

[3 总结与展望 18](#_Toc186311185)

[3.1 可行性分析 18](#_Toc186311186)

[参考资料 19](#_Toc186311187)

[附录 20](#_Toc186311188)

文档修订历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **修订原因** | **版本号** | **作者** | **修订日期** | **备注** |
| **1.0** | **编写文档** | **1.0** | **赵元川** | **2024、12、29** | **无** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 项目概述

## 项目背景

随着人工智能技术的迅猛发展，AI在各行业的深度应用不断拓展了人类对未来的想象空间。从几年前难以想象的对话式AI ChatGPT的出现，到如今针对特定问题领域的智能解决方案层出不穷，AI正在以其强大的适应性和创新能力推动各领域的变革。在游戏领域，AI更是取得了非凡的成就。从1952年艾伦·图灵开发算法进行国际象棋对弈，到AlphaGo在围棋领域击败世界冠军，再到OpenAI Five和AlphaStar在复杂战略游戏中的突破，AI展现了其在处理完美信息和非完美信息游戏中的卓越能力。这些技术成果反映了AI在游戏中的潜力，以及未来AI代理游戏的可能性。

图1.1 星际争霸 图1.2 围棋

不仅是人工智能行业，电子竞技行业也拥有庞大的市场以及无限的潜力。尤其是移动端MOBA类游戏《王者荣耀》已吸引了数亿用户。根据全球游戏数据统计平台 Newzoo的数据所示，当前游戏市场规模非常庞大，2024年总产出收益达到了1877亿美元。相比去年总共增长了2.1%。其中移动端游戏更是占据了接近一半的游戏市场，2024年总产出收益达到了926亿美元。相比去年增长了3.0%。移动端游戏市场已然成为全球最大的游戏细分市场。王者荣耀作为其中翘楚，拥有数亿注册用户和庞大的活跃玩家群体，这为AI游戏服务提供了巨大的潜在市场。

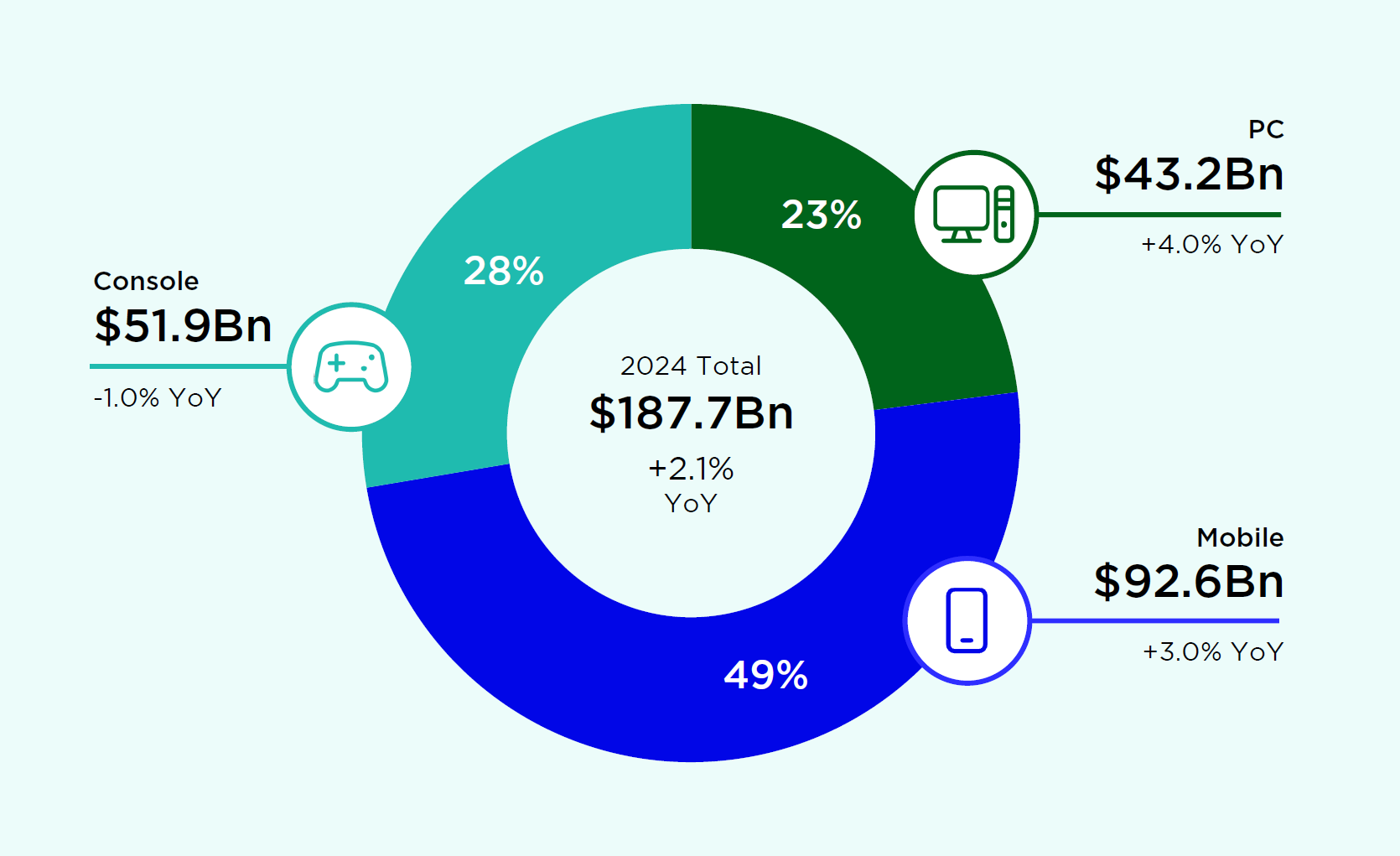


图1.3 2024游戏市场收益总和

然而，玩家在竞技体验中仍面临多重挑战。新手玩家由于对地图、英雄技能和经济发育等关键要素缺乏理解，往往难以获得良好的游戏体验；进阶玩家在提升操作与策略水平的过程中缺少专业指导，难以进一步突破自身能力；休闲玩家则希望拥有一个智能可靠的队友来共同享受游戏乐趣。这种需求可以体现在游戏各个行业的市场规模。如陪玩市场，从博思数据发布的《2024-2030年中国游戏陪玩市场分析与行业调查报告》可以看出游戏陪玩市场正处于快速发展的时期。

表1.1 游戏陪玩市场规模以及增速

又如游戏直播市场。2024年3月中国互联网信息中心(CNNIC)发布《中国互联网络发展状况统计报告》显示，2023年中国网络直播用户首次突破8亿人次，规模增长至8.16亿人，同比上涨8.66%，网络直播用户使用率74.7%，较2022年上涨1.4个百分点。中国演出行业协会2024年6月发布《中国网络表演(直播与短视频)行业发展报告(2023-2024)》，报告显示，2023年中国网络表演(直播)市场规模达2095亿元，较2022年增长5.15%。这些数据表明，中国网络直播行业市场规模已经相当庞大，且依然保持增长态势。

表1.2 游戏直播市场规模

此外，整个游戏生态对智能化技术的需求日益增加，例如自动化赛事分析和优化玩家决策路径等，成为提升整体体验和效率的重要方向。

本项目旨在构建一个能够自动游玩《王者荣耀》的AI Agent，服务于移动游戏市场，特别是在竞技类游戏领域。该AI代理将通过模拟真实玩家的操作和决策，实现自动化的游戏操作，也可作为陪玩与玩家一同游戏，为用户提供新的游戏思路和训练工具。《王者荣耀》爱好者和电子竞技选手可以通过与AI的对战提高自己的反应速度、战术理解和操作技巧。游戏开发者也可以利用AI代理进行游戏测试，优化游戏设计，并提供更加个性化的游戏体验。

## 项目内容

本项目以《王者荣耀》这一现象级MOBA手游为核心应用场景，通过构建智能AI Agent，旨在服务普通玩家、高阶玩家以及游戏开发者和赛事团队，为提升玩家技能、优化游戏生态以及推动智能化升级提供全方位支持。

《王者荣耀》凭借其庞大的玩家基数和复杂的竞技系统，为用户提供了丰富的游戏体验。然而，玩家在学习曲线、竞技水平提升和生态平衡性维护方面依然面临挑战。本项目通过构建智能AI Agent，模拟真实玩家行为，针对不同用户群体提供精准解决方案。

对于普通玩家，AI Agent 提供高质量的操作示范和技能指导，帮助新手快速上手并提升游戏体验，同时延长游戏生命周期。高阶玩家可以通过AI Agent 模拟高水平对手或队友，从中学习复杂的战术策略和操作技巧，提升竞技水平并获得清晰的训练思路。对于游戏开发者和赛事团队，AI Agent 的模拟能力和数据支持为游戏测试、平衡性调整和赛事分析提供了强有力的技术支持，助力游戏设计优化和赛事生态发展。

AI Agent 的核心功能包括操作示范、策略模拟和智能分析。在操作示范方面，AI Agent 复现高水平玩家的决策与操作流程，为普通玩家提供明确的学习路径。在策略模拟中，AI Agent 可作为智能对手或队友，帮助玩家深入理解战术要点并提升决策能力。同时，通过智能分析，AI Agent 能够在复杂场景下进行多局对战模拟，为开发者提供可靠的平衡性测试数据和优化参考。

本项目的AI Agent依托实时获取的《王者荣耀》对局数据动态生成训练数据集，并通过深度学习与强化学习相结合的机制实现自我优化。数据集涵盖游戏实时画面、英雄技能使用情况、地图环境、当前对局时间、实时经济数据等多维信息，全面反映游戏过程中的操作细节与局势变化。基于这些数据，AI Agent利用深度学习进行初始模型训练，掌握基础操作与决策逻辑，并通过强化学习框架在模拟对局中逐步优化策略。

自定义奖励函数的设计综合考虑了击杀得分、团队配合和经济效率等关键指标，使AI Agent能够在动态复杂的环境中持续提升决策能力。同时，实时在线学习机制确保了模型能够快速适应最新的玩家行为和版本更新，保持高水平的智能性与适配性。通过这种动态更新与自我优化的训练模式，AI Agent展现了强大的战术深度和操作能力，不仅能复现真实玩家的行为，还能在高水平对局中提供超越人类的策略支持。

当前，AI Agent 的核心功能集中在操作示范和策略模拟，但其未来发展潜力同样令人期待。AI Agent 可拓展为智能陪玩助手，为玩家提供实时配合与互动体验，例如模拟不同水平的战术风格、提供策略建议以及缓解单排模式下的孤独感。这一功能不仅丰富了游戏的趣味性，也拓展了AI在游戏领域的应用边界。

## 创新点

### 技术创新点

#### 左右手单独网络设计

本项目基于Double DQN构建了左右手单独网络结构，模仿玩家的实际操作方式，将左手和右手操作分配给两个独立的网络。左手网络专注于角色的移动控制，右手网络负责攻击、技能释放等行为。

在训练过程中，左右手网络通过共享状态字典进行初始化，确保同步学习。每个网络根据当前游戏状态分别预测最优动作，并通过各自的Q值计算损失。这样的协作机制，不仅能够提高AI在复杂场景下的操作效率，还能更接近人类玩家的操作习惯。

#### 长短期记忆网络（LSTM）的引入

由于游戏状态是动态变化的，本项目引入了长短期记忆网络（LSTM）来捕捉时间序列数据中的长期依赖关系。通过帧堆叠技术，将连续帧信息整合为输入，LSTM能够更好地理解游戏中的时序性变化，从而预测未来状态并指定相应的策略。

这一技术使得AI能够在游戏对抗中捕捉关键状态变化，做出更加精准和连贯的操作，显著提升了整体策略的可行性和稳定性。

#### 动态奖励函数的设计

本项目采用动态奖励机制，通过多维度奖励设计，引导AI在《王者荣耀》中逐步掌握复杂操作和战略决策。该机制基于实时游戏状态，结合关键因素计算奖励，确保AI在不同游戏阶段做出合理选择并优化策略。

移动奖励根据AI英雄与目标位置间的距离变化评估。当AI靠近目标位置时，给予正向奖励；远离目标则惩罚。在游戏前期，若AI血量较低，系统引导其返回泉水恢复，并动态调整奖励值，确保资源获取与生存策略的合理性，使AI在不同情境下做出最佳移动决策。

技能使用奖励通过技能冷却状态与使用时机评估。AI在攻击状态下释放技能会获得奖励；非攻击状态下使用技能则被惩罚。针对特定技能（如治疗技能），系统根据AI血量状态进行差异化奖励，确保技能使用的有效性与效率，使AI能够根据实时状态灵活调整技能策略。

攻击奖励根据AI的攻击行为及游戏阶段动态调整。攻击野怪或敌方英雄时，系统根据目标范围和游戏阶段奖励不同分值，确保AI在资源争夺与团队作战中制定合理攻击策略。

胜负奖励在比赛结束时根据结果给予高额奖励或惩罚。团队胜利时奖励分数，失败则施加惩罚，引导AI优化全局策略以追求胜利，通过长期训练掌握团队协作与战略决策。

经济奖励通过监测AI经济增长和击杀数动态调整。当AI经济或击杀数增加时，给予正向奖励，鼓励其在资源获取与击杀敌方英雄方面的表现，确保其经济发育能力，从而在资源争夺中占据优势。

动态奖励机制还针对游戏阶段设计。在前期，奖励侧重生存与资源获取；后期则更强调团队作战与战略目标达成。通过阶段化奖励，AI能够灵活调整策略，使其行为始终与全局目标一致。

#### 采样权重优先策略

在强化学习中，本项目采用了优先级经验回放（Priority Replay Buffer）。这一机制根据经验的重要性动态调整其被采样的概率，使得高优先级经验更频繁地用于训练。例如：

* 高权重经验：更可能改善策略的经验被赋予更高的权重。
* 探索与利用平衡：新加入的经验被赋予较高优先级，鼓励AI探索未充分了解的状态。

这一策略能够加速AI的学习效率，同时保证学习过程的多样性和全面性。

### 功能创新点

#### 基于图像识别的动态环境感知

本项目通过结合多种图像处理技术与深度学习模型，实现了对复杂游戏环境的实时感知与分析。YOLO模型在系统中负责检测并识别敌我双方的位置。OCR技术则用于提取游戏画面中的数字信息，包括当前经济数据、击杀数量等关键指标。OpenCV算法处理了其他关键任务，例如检测血量、自身角色位置、判断攻击是否有效、攻击对象、对局状态（如是否开始、死亡或重生）以及技能冷却状态等。

OpenCV因采用高效算法，识别速度较快，并且对其他模型的输入数据进行了预处理，从而显著提升了整体识别的速度与准确性。该图像识别框架通过高效协作，在动态复杂的游戏环境中能够快速、精准地获取关键信息，为系统的后续决策提供可靠支持。这种设计显著优化了游戏环境感知能力。

#### 强化学习驱动的策略优化与自动化操作

系统通过强化学习技术完成了AI的策略学习与优化，模拟人类玩家的操作行为。

AI采用强化学习框架，通过视觉输入学习如何操控角色，包括移动、攻击与技能释放。在训练过程中，AI模拟真实用户的触屏点击与滑动操作，而非直接调用游戏API，从而完全符合人类操作习惯。

通过动态奖励机制，AI能够根据操作表现实时调整策略，例如在技能命中率较低时优化释放时机，或在面对敌方英雄时优先选择撤退或追击。这种训练方式使得AI能够无缝衔接游戏的各个阶段，从开局到结束实现完全自动化游玩。

#### 大语言模型驱动的局势分析与战术建议

系统整合了大语言模型，用于分析游戏局势并提供战术建议。大语言模型基于当前画面图像，敌我双方英雄位置，自身状态生成自然语言描述，对游戏当前局势进行分析。例如，经过分析当前局势，判断目前是否需要抢夺公共资源并给出建议。此外，大语言模型还为玩家提供操作建议和局面总结，包括分析阵容强弱点、团队配合优劣势等。这一功能不仅提升了AI的策略性，还为玩家提供了可参考的战术方向，帮助其提升游戏水平。

### 其他创新点

#### 多线程数据采集与训练

为了提高训练效率，本项目在训练模块中实现了多线程并行处理。通过创建两个线程池，分别执行数据采集任务和训练任务，使得数据采集和模型训练能够同时进行，显著减少了训练过程中的等待时间。多线程设计提高了整体训练效率，并且增强了模型的稳定性。

#### 深度融合的操作模拟能力

本项目通过强化学习框架与多模态图像识别技术的深度融合，极大地提升了AI在游戏中的操作模拟能力和动态环境感知水平。在训练过程中，AI充分整合YOLO、OCR和OpenCV的功能，对游戏环境中的多维信息进行全面感知和解析。YOLO快速检测敌我双方的实时位置，并能够精准区分敌方英雄与野怪，为AI的全局策略提供支持。OCR负责提取游戏界面中的关键数值信息，如经济状况、击杀数等，为AI提供动态的量化指标。OpenCV则在细粒度分析上发挥了重要作用，包括血量检测、自身角色定位、攻击效果判定、对局状态分析（如开局、死亡或重生）以及技能冷却时间的动态监控。同时，OpenCV还对图像数据进行了预处理，使YOLO和OCR的识别速度与准确率进一步提升。

在此基础上，AI与ChatGPT推荐的坐标数据相结合，构建了一套动态的移动规划机制。ChatGPT提供大致的目标坐标，而AI通过强化学习不断优化行动路径。当英雄偏离推荐坐标时，系统通过奖励机制调整行动策略，引导AI重新规划路径，从而实现更高效的移动和资源利用。通过多模态信息的深度融合，AI能够实时感知野怪分布并优化资源获取路径，同时结合环境动态调整团战切入时机，选择最佳攻击和防守策略。此外，AI能够根据多维数据快速判断敌方状态和战场形势，在复杂对局中展现出流畅且精准的操作模拟能力。

这一融合框架不仅赋予了AI对动态环境的高效适应能力，还使其在操作层面达到了接近高水平玩家的表现。例如，AI能够在战斗中快速调整技能释放节奏，避免冗余操作，同时在团战中根据实时数据灵活选择站位和策略。这种深度结合的操作模拟能力，不仅显著提升了AI在游戏中的表现，还为复杂动态环境下的智能行为建模提供了全新的思路和技术支持。

#### 多模态数据分析能力的应用

本项目充分利用多模态数据分析能力，将图像识别技术与多种分析手段相结合，实现了对游戏场景的深度感知与实时决策支持。YOLO模型与OpenCV的结合在其中发挥了核心作用，通过快速精准地检测敌方英雄、野怪、防御塔等游戏关键元素，为AI提供了实时的位置信息和状态数据。同时，OCR技术负责从游戏界面中提取经济、击杀等数值信息，而OpenCV则用于更细粒度的图像处理，如血量检测、自身角色定位、判断攻击是否有效、游戏是否开始或结束、技能冷却时间以及角色死亡或重生状态。

这种多模态数据的融合应用使得AI不仅能够感知游戏中离散的图像信息，还能综合多种来源的数据进行智能分析。例如，YOLO与OpenCV结合检测全局目标和环境动态，OCR提取关键数值信息，而其他书写的OpenCV的算法识别进一步提升了对细节变化的感知速度与精度。通过这种多模态数据的协同处理，AI能够快速捕捉游戏场景的全貌和细节变化，形成全面、精准的动态环境感知能力。

在此基础上，AI实现了基于场景感知的动态决策调整。在游戏初期，AI通过对野怪、资源点等的实时识别和分析，优化了资源获取路径。而在游戏中后期，AI则利用多模态数据分析能力识别团战时机、监控关键目标状态，并动态规划最佳策略。这一能力大大提升了AI的决策效率和适应性。

# 技术方案

## 总体架构

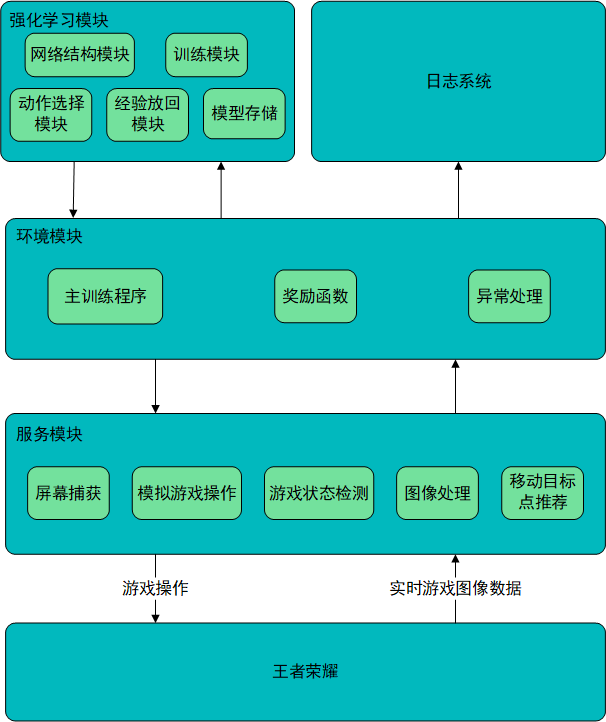


图2.1 模块架构图

本项目的架构设计采用模块化结构。如图2.1所示，整体框架分为强化学习模块、环境模块、日志系统模块以及服务模块。各模块在功能上独立分工，在系统运行中通过接口进行协作，以确保训练效率和实时性。以下将对各模块的功能和作用进行详细描述。

强化学习模块是系统的核心部分，负责实现智能体策略的优化与推理。该模块包含了网络结构模块、训练模块、动作选择模块、经验回放模块和模型存储模块等核心功能。网络结构模块为强化学习提供了可扩展的神经网络框架，能够根据游戏环境的复杂性调整模型的深度和结构。训练模块实现了强化学习的核心算法逻辑，通过不断优化模型参数提升策略的准确性和稳定性。动作选择模块用于根据当前环境状态输出智能体的动作决策，同时动态平衡探索与利用，以确保模型训练的多样性和全面性。经验回放模块存储并管理智能体与环境交互的数据记录，通过批量采样增强训练的稳定性和效率。模型存储模块提供了高效的模型存储与加载功能，支持训练中断恢复以及推理模型的快速部署。

环境模块负责模拟游戏中的动态环境，是强化学习模块训练和推理的重要支持组件。该模块通过主训练程序实现强化学习模型与虚拟游戏环境的交互，包括状态更新、动作执行和奖励反馈等功能。奖励函数模块根据游戏目标和状态变化计算智能体的奖励值，为策略优化提供明确的目标方向。例如，当英雄完成关键任务时奖励分数会增加，而未达成预期目标或远离移动目标时则会受到惩罚。异常处理模块检测并处理训练过程中可能出现的数据异常或逻辑冲突，保证系统运行的稳定性和训练过程的可靠性。路径规划模块结合游戏状态与目标，为智能体提供合理的移动方向，距离的偏离程度也会直接影响奖励函数的计算结果。

日志系统模块贯穿于系统的各个流程，用于记录系统运行中的关键数据，支持调试、性能分析和结果复现。数据记录子模块以结构化的形式捕获强化学习模块与环境模块的交互数据，包括状态、动作、奖励和策略变化等信息。这些记录为系统的优化和问题的定位提供了详实的数据支持。异常日志功能主要用于捕捉运行过程中可能出现的错误或异常事件，例如数据输入格式错误、网络训练超时等，并将这些信息存储为诊断依据。日志系统模块还提供性能分析功能，能够对强化学习模型的收敛速度、策略表现和胜率变化进行定量分析，从而为调整训练参数和优化系统提供数据支持。

服务模块是系统与《王者荣耀》游戏界面的交互桥梁，负责图像采集、信息提取和指令执行。屏幕捕获子模块实时获取游戏图像，为环境模块提供小地图、英雄与敌方单位位置、野怪状态等动态信息。图像处理模块结合YOLO、OCR和OpenCV等技术，提取关键信息：YOLO识别敌我单位和防御塔，OCR提取经济值和击杀数，OpenCV分析血量、技能冷却时间及游戏阶段，全面提升数据解析的准确性。模拟游戏操作模块根据强化学习生成的指令，完成技能释放、移动和攻击操作，精准响应游戏环境变化。游戏状态检测功能实时分析游戏画面，判断进程状态（如加载、进行或结束），触发策略调整。路径推荐子模块结合强化学习输出与环境反馈，借助ChatGPT优化移动策略，提升智能体的灵活性与适应性。

模块之间通过数据和接口进行高效的协作。服务模块通过屏幕捕获功能实时获取游戏图像数据并进行预处理，将提取的关键信息传递至环境模块。环境模块根据当前状态调用强化学习模块的接口，生成针对当前状态的最优动作指令。强化学习模块的动作选择直接影响服务模块的模拟操作，最终通过与游戏界面的交互完成智能体的动作执行。日志系统模块在这一流程中记录关键数据，为系统的优化和问题诊断提供全面的数据支持。

整体架构设计以模块化为核心，强化学习模块作为策略生成的核心引擎，环境模块提供动态反馈与模拟，服务模块处理数据输入输出，日志系统模块保障系统运行的透明性和可追溯性。各模块间的协作确保了系统的功能完整性与运行稳定性，为智能体在《王者荣耀》中的高效表现提供了技术保障。

## 功能详述

### 系统流程图

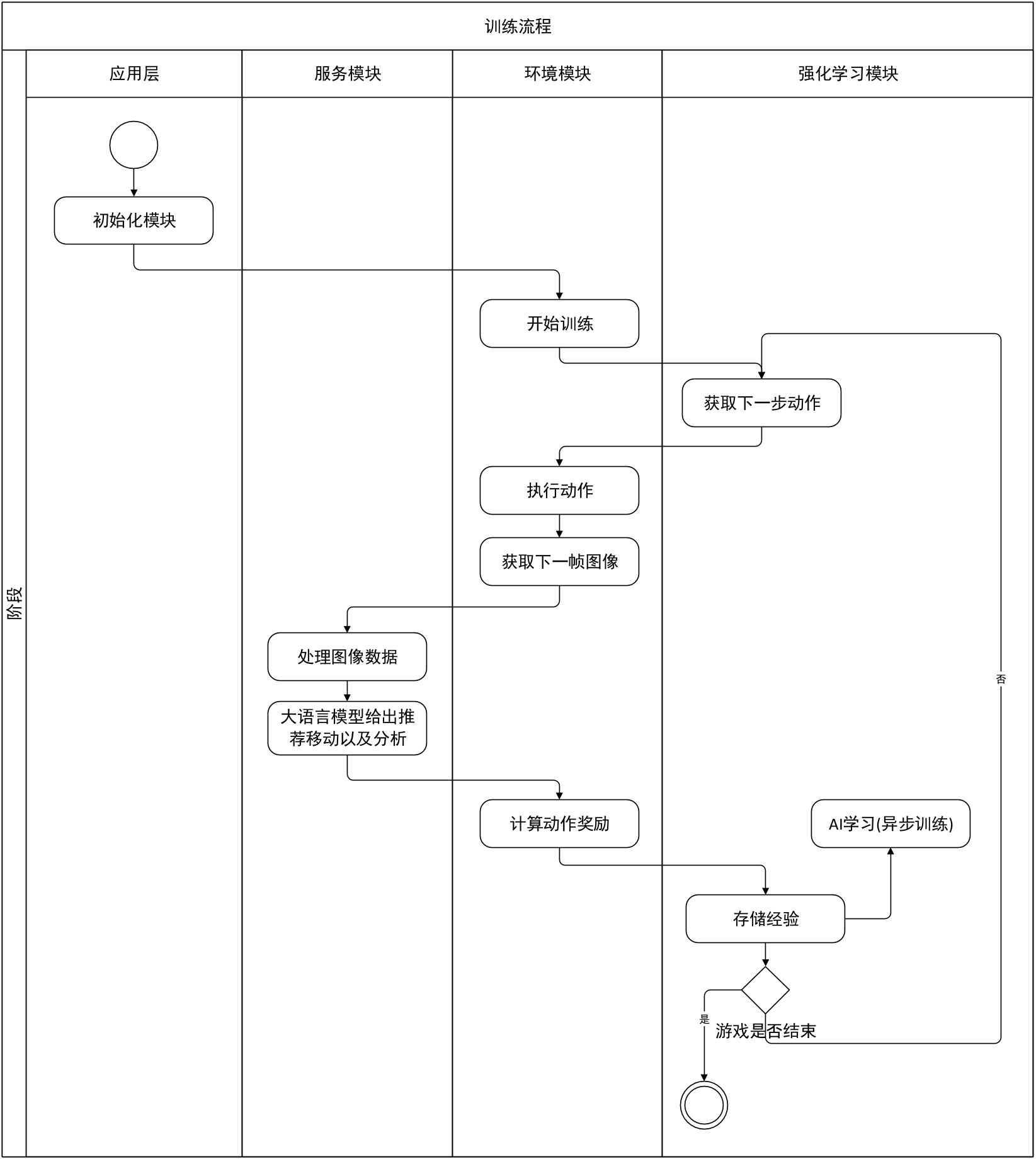


图2.2 功能结构

如图2.2所示，系统的运行流程由应用层、服务模块、环境模块和强化学习模块共同构成，形成一个循环优化的闭环系统。以下是对流程的详细说明。

系统启动时，应用层首先完成初始化操作。该步骤包括加载训练模型、配置参数和基础环境资源的准备，为系统后续运行提供必要的支撑和稳定性。在初始化完成后，控制权交由环境模块，系统正式进入训练流程。

环境模块的第一步是调用强化学习模块的API，以获取当前状态下的下一步动作。强化学习模块通过分析当前环境状态，结合优化的策略模型，返回一个可执行的动作指令。环境模块接收该指令后，将会执行接收到的具体的指令。随后，环境模块与服务模块进行交互，从游戏界面中捕获当前帧的图像数据，以获取最新的环境反馈信息。

服务模块的任务是对图像数据进行处理。首先，服务模块利用图像识别和数据分析技术，从捕获的图像中提取环境中的关键信息，例如角色位置、敌人动态、防御塔分布等。这些信息不仅为后续的动作评估提供了数据基础，还会通过服务模块调用大语言模型生成推荐的移动方向和策略建议。这会优化前期未成熟AI的移动策略以及一定程度上加速AI的学习收敛。

服务模块处理完成后，环境模块接收图像分析结果和大语言模型的建议，结合奖励函数计算当前执行动作的得分。奖励函数根据预设规则评估动作的效果，例如角色是否成功击杀敌人、是否规避了攻击或是否朝目标方向移动，从而生成正向或负向的反馈值。计算结果被传递回强化学习模块，作为训练数据的一部分。

强化学习模块接收来自环境模块的数据，包括当前状态、执行动作和奖励值等信息。这些数据被存储到经验回放池，为模型优化提供训练样本。随后，强化学习模块利用这些样本使用异步的方式对模型进行优化，以提升其决策能力。

如果游戏尚未结束，系统将再次进入环境模块，重复上述流程。环境模块会基于强化学习模块的最新策略指导角色执行新的操作，并通过服务模块分析最新的环境状态，不断调整决策模型。如此循环，系统在持续运行中不断优化智能体的行为策略，直至满足训练终止条件。我们可以发现整个流程是自动化进行的，完全不需要人为干预。

### 核心功能

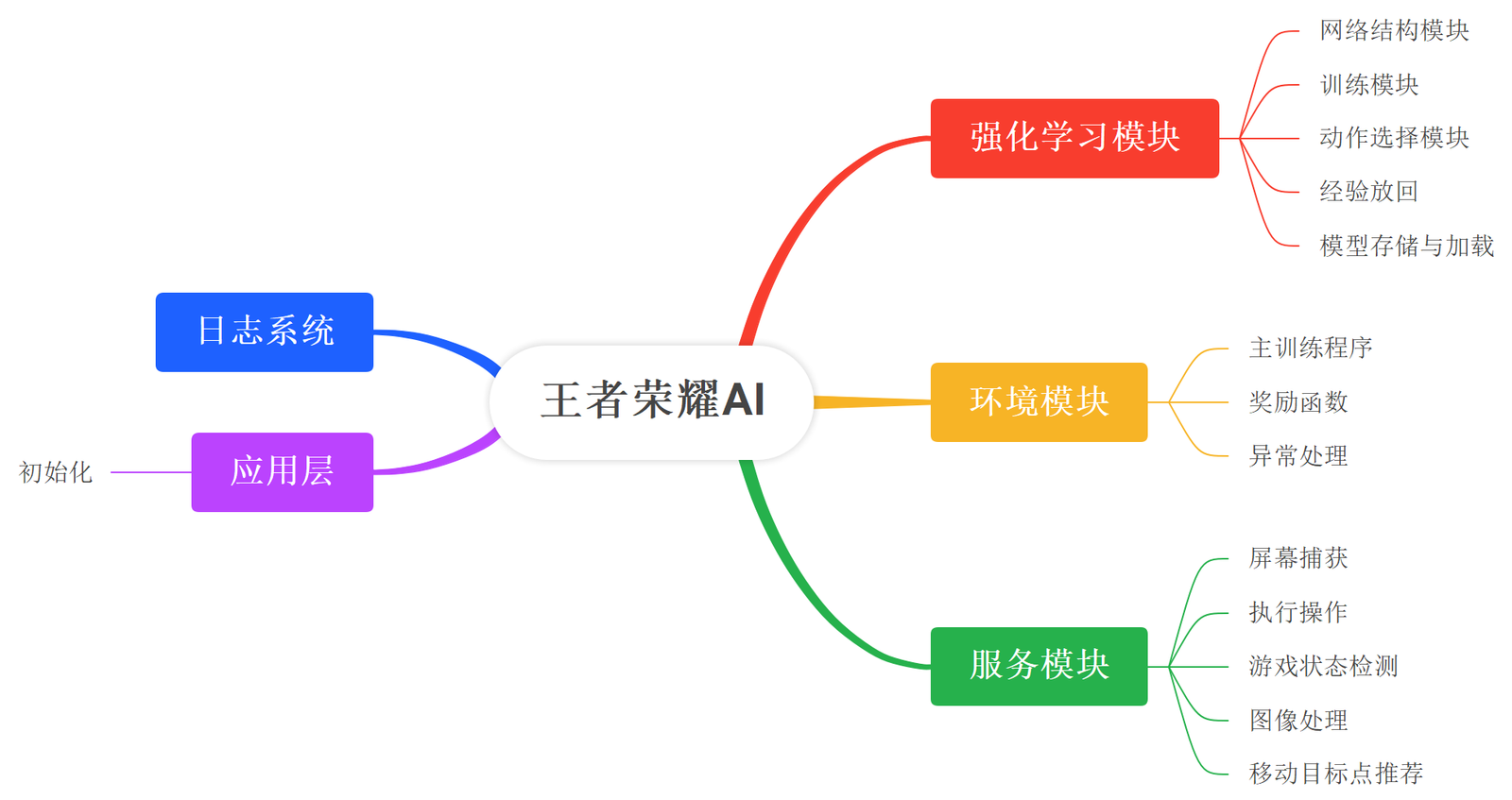


图2.3 系统模块思维导图

如图2.3所示，本项目的核心功能在于构建由多个模块协作组成的高效智能体训练系统，能够实现全自动化的训练流程。该系统以强化学习技术为核心，结合环境模拟和服务支持，提供了一个高度集成和优化的解决方案，使智能体能够在复杂的任务场景中通过自主学习不断优化其决策能力。

系统的设计注重模块间的有机协作，以确保高效和稳定的运行。强化学习模块是系统的核心部分，负责智能体的策略学习和优化。通过网络结构模块，系统能够建立适应特定任务的深度学习模型；动作选择模块结合模型预测结果和当前环境状态，输出智能体的最优行动；经验回放模块记录智能体与环境交互的历史数据，用于优化模型的训练过程；模型存储模块确保每个训练阶段的模型状态可追溯，实现断点续训和长期优化。

环境模块在系统中发挥了实时交互的作用，承担了训练中模拟任务场景的职责。当系统进入训练阶段时，环境模块通过主训练程序初始化任务场景，并调用强化学习模块获取智能体的下一步动作。随后，环境模块执行该动作并获取新的环境状态，例如下一帧的图像或任务目标的变化。同时，该模块负责通过奖励函数对智能体的动作进行评估，计算当前行为的得分，并将结果反馈给强化学习模块。这种环境交互机制确保了训练的实时性和针对性，使智能体的行为能够快速适应不同的任务变化。

服务模块在整个系统中充当了信息处理和外部交互的桥梁。首先，服务模块通过屏幕捕获和图像处理功能，实时提取训练环境中的关键信息，例如目标位置和当前状态。其次，服务模块能够调用大语言模型（如ChatGPT）进行策略分析和下一个移动目标的推荐，从而为智能体提供辅助决策支持。此外，服务模块还负责对外部游戏界面进行操作模拟，确保智能体的训练在真实任务环境中高效开展。

系统的训练流程完全实现了自动化，从环境初始化到强化学习策略的更新均由系统自主完成。在训练过程中，强化学习模块与环境模块形成闭环，通过实时的交互和反馈不断优化模型性能。如果环境模块未检测到任务结束信号，系统会自动进入下一次训练循环，直至智能体达到预期的任务目标。这种自动化的闭环设计不仅减少了人工干预的需求，还大大提高了智能体的训练效率。

本项目的核心创新在于通过模块化设计和技术集成，实现了强化学习算法、实时环境交互和辅助策略分析的有机结合。系统在每一步决策中都充分利用了环境提供的实时反馈和大语言模型的分析能力，从而确保训练过程的高效性和智能体的快速成长。这种设计为复杂场景中的智能体训练提供了一种全面而有效的解决方案。

## 关键技术

本项目的关键技术涵盖强化学习算法、大语言模型辅助决策、图像处理技术和环境建模技术，通过模块化协作实现了智能体全自动训练的完整技术体系。强化学习算法作为系统核心，通过策略优化与动态交互，提升智能体在复杂环境中的决策能力和适应性能。大语言模型提供辅助分析与策略建议，增强了智能体的多维度决策能力和策略灵活性。图像处理技术实时感知环境，提供精确的数据输入，环境建模技术保障了训练场景的真实性和稳定性，使系统能够构建一个闭环、高效的训练流程。

强化学习模块通过采用前沿的深度强化学习算法，尤其是深度Q网络（DQN），实现了高效的策略优化和智能体性能提升。

深度Q网络（DQN）通过结合深度神经网络与Q学习方法，解决了高维状态空间下的不稳定性问题。DQN设计中引入经验回放和目标网络，优化了离散动作空间的决策性能。在本项目中，DQN用于特定策略路径的评估与选择，显著提高了探索效率，减少了随机性对决策的影响，加速了智能体的收敛。

强化学习模块中还包含经验回放机制，通过存储智能体与环境交互的状态、动作与奖励等数据，随机抽取批量数据进行训练，减少数据间的时间相关性干扰。多维度奖励机制对智能体的行为设计了多样化的正向奖励与负向惩罚策略，确保其行为不仅追求任务完成效率，还能兼顾稳定性与多样性。例如，奖励机制会对接近目标的行为赋予正向激励，而对偏离目标或非最优行为施加惩罚，从而引导智能体逐步优化决策。

服务模块利用大语言模型（如ChatGPT）辅助决策，提供智能体在复杂环境中的实时策略建议和目标点推荐。通过结合图像处理技术和环境感知能力，模块能够获取当前游戏状态，包括玩家位置、队友和敌人分布、游戏时间等，并将这些信息传递给大语言模型进行分析。大语言模型不仅理解文本，还能够结合图像数据，为智能体提供综合决策支持。基于当前环境，模型实时生成策略建议，推荐最佳的移动目标、行动路线，并预测敌人位置或避开危险区域。此外，模块能够根据智能体的反馈信息动态调整决策，确保系统能够适应变化的环境并持续优化策略。通过这一模块，智能体不仅能快速做出决策，还能灵活应对各种环境变化，提升决策的全面性、实时性和适应性，进一步增强系统的自主性和多样性。

## 其他相关技术

### 相关技术

本项目使用了Python作为主要编程语言，并采用多项关键技术构建高效智能体训练系统。CUDA用于调用GPU资源，加速模型训练与推理过程。自主训练的YOLO模型用于图像对象的位置识别，Tesseract-OCR实现数字信息提取，如金币数量等关键数据。NumPy支持强化学习中的数值计算与状态表示。

OpenCV负责图像数据的处理（使用算法精准识别某些特性场景下的数据），PyTorch作为深度学习框架支持神经网络训练。Gym提供标准化环境接口以实现智能体交互，OpenAI API通过大语言模型增强策略生成与分析。PyQt5用于构建直观的可视化界面，Pillow、Pytesseract和Ultralytics分别支持图像处理、OCR识别与YOLOv8模型开发。

### 环境部署

推荐使用anaconda创建python版本为3.12的虚拟环境，便于之后资源管理。

本项目使用pip作为环境部署工具，在本项目的根目录下输入如下指令。

|  |
| --- |
| pip install -r requirements.txt |

这会下载本项目在代码层面上的所有第三方依赖。随后需要下载PyTorch和CUDA，输入如下指令。

|  |
| --- |
| pip3 install torch torchvision torchaudio  --index-url <https://download.pytorch.org/whl/cu118> |

接下来需要安装用于GPU推理的onnxruntime-gpu。

如果CUDA的版本是11则输入如下指令进行安装。

|  |
| --- |
| pip install onnxruntime-gpu |

如果CUDA的版本是12则输入如下指令进行安装。

|  |
| --- |
| pip install onnxruntime-gpu --extra-index-url  https://aiinfra.pkgs.visualstudio.com/PublicPackages/\_packaging/onnxruntime-cuda-12/pypi/simple/ |

接下来我们需要配置模拟器用于运行王者荣耀（模拟器推荐使用 mumu 模拟器，安装包放在该项目的 wzry/new\_wzry\_ai/app\_tools/mumu 目录下，可以直接使用），接下来我们需要启动模拟器，并将窗口命名为 wzry\_ai，分辨率设置为 1600\*720，同时确保屏幕缩放率为 100%。项目启动后程序会自动搜寻名字为 wzry\_ai 的窗口并从中获取数据开始训练。

在模拟器中安装王者荣耀程序，并进行键位配置，若使用 mumu 模拟器，可采用该方案（分享《王者荣耀》按键操作方案，复制此分享码：mumu0364377751。打开 MuMu 模拟器，在方案管理中导入使用！方案适用分辨率：1600\*720）

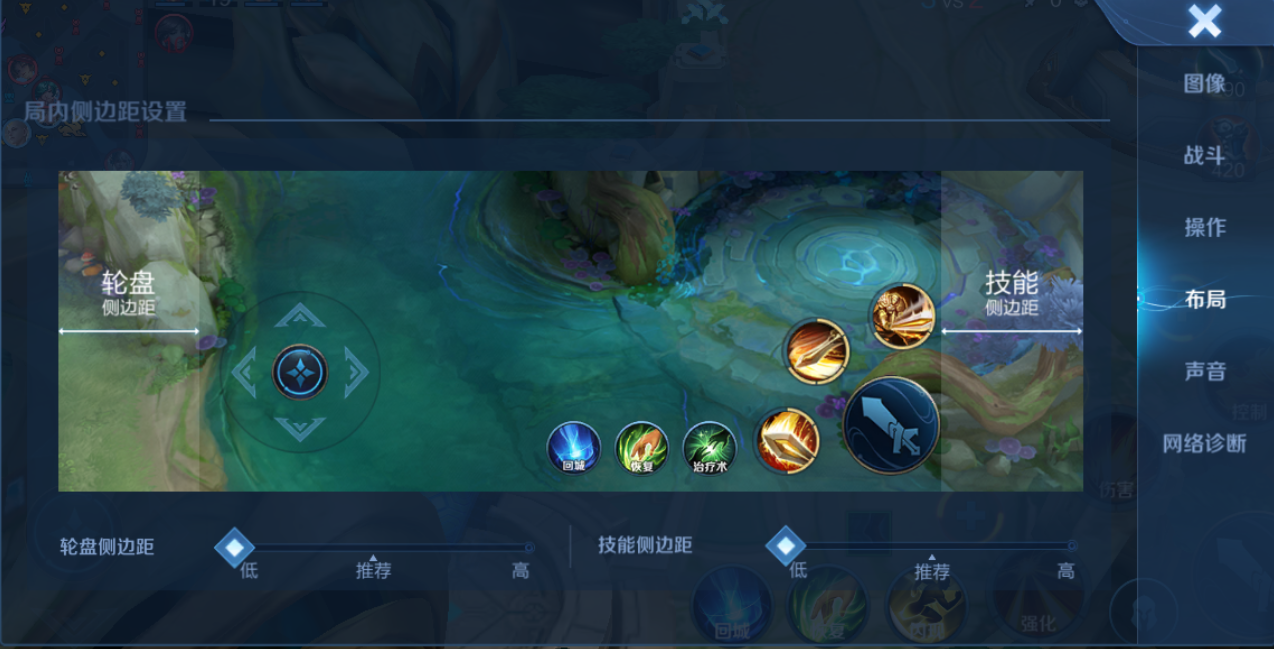
若为其他模拟器，可参考下图：



同时，我们要修改王者荣耀内部布局，总布局为右手模式，局内出装位置为右侧



轮盘侧边距与技能侧边距均调整为最低



接下来我们需要配置第三方工具。因为项目中使用了 OCR 来识别图片中的相关字符数字，因此，要在本机配置 OCR 工具。

工具已存放在该项目下的 wzry/new\_wzry\_ai/app\_tools/Tesseract-OCR 目录下，需要将该目录的绝对路径添加到系统的环境变量中。

最后请将 wzry/new\_wzry\_ai/config/default\_config.py 文件中第 180 行变量 openai\_api\_key\_4\_0\_2 设置为自己的 chatgpt api\_key。

接下来我们需要启动模拟器，并将窗口命名为wzry\_ai。项目启动后程序会自动搜寻名字为wzry\_ai的窗口并从中获取数据开始训练。

# 项目计划

## 可行性分析

### 技术可行性分析

本项目通过多种先进技术的集成，实现了可自动游玩《王者荣耀》的AI。包括了图像识别、大语言模型分析以及强化学习训练。这些技术模块在各自领域的实践中都已表现出了非凡的效果。

本项目采用OpenCV技术、OCR技术以及YOLO模型进行游戏界面和目标物体的实时识别。

OpenCV处理图像细节，如血量检测、角色位置、攻击对象、游戏状态等。由于其优化算法，OpenCV能够在实时游戏画面中高效工作，确保系统能处理多维信息，提供精准的决策支持。

OCR用于提取游戏画面中的数字信息，如金币、击杀数量等。通过快速且准确的数字识别，OCR为系统提供了必要的经济数据和战斗信息，支持智能体做出合理决策。

YOLO以其快速、高效的特性，能够在保证识别精度的同时实时处理游戏画面。通过定制训练数据集，模型可以精准识别英雄位置。这一模块的技术成熟度高，能够稳定地为后续决策提供精准输入。

本项目使用大语言模型（如ChatGPT）分析游戏局面，为AI提供策略建议。模型通过自然语言描述对当前局势进行评估，例如判断当前最优移动方向或技能释放时机。这种分析可以结合强化学习模块的决策结果，提升操作的智能化与人性化水平。该技术基于现有成熟的自然语言处理能力，能够很好地与游戏场景需求相适配。

强化学习部分使用了基于Double DQN的架构，结合LSTM和优先级经验回放机制，提升了策略学习的效率和连贯性。Double DQN通过分离动作选择和价值评估，避免了策略过估计问题，使得AI能够在复杂游戏场景中学习到稳定可靠的操作策略。LSTM的引入增强了AI的时间序列建模能力，使其能够结合历史信息进行动态决策。优先级经验回放则通过提高重要经验的采样概率，加速了关键行为模式的学习。

奖励函数设计针对游戏的多样化目标进行了优化，涵盖局部即时反馈和全局胜负影响，从而引导AI逐步掌握复杂的游戏策略。多阶段学习策略的应用，使得AI能够从基础操作学习逐步过渡到高级策略优化，表现出较高的适应性和学习能力。

项目采用的技术架构具有高度的模块化和扩展性。图像识别、大语言模型与强化学习之间的协作使得AI能够在感知、分析和决策三个层面实现高效联动。各模块的技术选型基于成熟的方法和工具，确保了实现的可行性和稳定性。同时，针对游戏的特殊需求（如实时性和对抗性），项目通过并行计算和增量训练等手段优化了性能。

### 资源可行性分析

本项目在现有条件下能够充分满足开发和实施的需求，资源的配备和获取渠道完全保障了项目的可行性。

硬件设备方面，项目的开发和运行对计算能力有一定要求，尤其是在深度学习模型的训练和推理阶段。现有的高性能GPU设备，以及云计算平台提供的分布式计算服务，能够高效支持YOLO模型的训练和强化学习的模拟迭代。这些设备可以在训练阶段完成大规模数据处理，同时在推理阶段提供实时响应能力，确保AI在游戏中的操作流畅。消费级硬件设备也能满足部署需求，适合实际游戏环境下的长时间运行。

人力资源方面，项目团队由多名经验丰富的AI开发者组成，具备深度学习、强化学习和游戏开发的相关技能。团队成员熟练使用主流深度学习框架和强化学习工具库，能够快速构建和优化模型。对于游戏中的目标检测和交互逻辑设计，团队成员的实际开发经验确保了AI行为与游戏规则的完全契合。此外，团队还掌握高效的数据标注和自动化测试工具，能够在短时间内完成高质量的数据准备和模型验证工作。

数据资源方面，项目所需的训练数据具备多种获取渠道。游戏画面数据可以通过录制实际游戏操作获得，并结合数据增强技术进一步扩充数据集规模。目标检测所需的标注信息可以利用现有工具高效生成，无需额外资源投入。强化学习部分的交互数据来源于模拟器的自动生成，避免了对人工操作的依赖。通过设计合理的模拟环境和奖励函数，AI能够自主生成大量高质量的训练数据。此外，项目实现提供的训练状态存储模块能够保证高效存储和管理这些数据，为训练和推理提供稳定的支持。

软件支持方面，项目所依赖的开发工具和框架已经非常成熟，能够显著降低开发难度。深度学习框架如PyTorch和TensorFlow，以及强化学习库如Stable-Baselines3，提供了丰富的预训练模型和工具函数，可直接用于项目的核心算法实现。游戏模拟器可以准确还原游戏环境，使AI行为能够在高度还原的场景中进行测试和优化。开源开发工具和版本管理平台为代码开发和团队协作提供了稳定的技术保障，完善的调试和测试工具则进一步提升了开发效率。部署阶段可以依赖边缘设备或云平台，支持实时推理和高并发请求处理。

### 市场可行性分析

本项目在市场上具有显著的潜力和竞争优势，其市场可行性体现在潜在用户需求的广泛性、市场竞争环境的空白点、技术与应用的前瞻性，以及商业化路径的清晰性。

本项目拥有坚实的市场基础。游戏玩家对人工智能的兴趣和需求正在快速增长，尤其是在MOBA类游戏中，《王者荣耀》的用户群体庞大且活跃。本项目提供的智能化AI系统，不仅能够辅助用户理解游戏对局形势，还能在训练和娱乐中为玩家提供高质量的互动体验。AI模拟人类玩家的行为模式，能够帮助新手玩家迅速提升技能，也可以为高端玩家提供策略参考和对抗训练，显著扩大目标用户覆盖范围。此外，AI的自动化能力还能吸引游戏内容创作者和主播，通过展示AI对局生成独特内容，进一步提高市场吸引力。

本项目相比同行拥有显著的优势。目前市面上的游戏辅助工具和AI产品，大多停留在简单的数据分析和规则引擎层面，缺乏对复杂游戏场景的深度理解和动态适应能力。传统外挂通常依赖内存数据修改，不仅存在法律和道德风险，还对用户设备造成潜在威胁。相比之下，本项目的AI通过图像识别和模拟人类行为实现游戏操作，完全避开了内存篡改的风险，同时在技术和合规性上更具优势。作为一款创新型游戏AI，本项目在这一市场空白中占据了独特的竞争位置，能够快速吸引用户并建立品牌认知。

本项目具有技术与应用的前瞻性。随着人工智能在游戏行业的应用日益广泛，AI的智能化表现和人机交互能力已成为行业发展的重要趋势。本项目不仅展示了YOLO和强化学习等前沿技术在实际场景中的应用，还创新性地结合了大语言模型，为玩家提供实时对局分析与建议，显著提升了游戏体验的个性化和智能化水平。这种技术与应用的结合不仅满足了当前市场需求，还为未来技术升级和功能扩展提供了充分的空间。

本项目拥有清晰的可商业化路径，以及可持续发展的动力。本项目可以通过多种方式实现商业变现，包括面向个人用户的订阅服务、为电竞机构提供对抗训练的AI系统，以及为游戏公司开发专属AI对手或游戏模式。此外，基于项目技术能力的内容创作和AI直播功能，还可以通过广告分成和观众打赏获得收益。与行业内的相关平台和社区合作，通过推广和赞助进一步扩大市场影响力，也为商业化提供了多样化的选择。

## 排期规划

目前初赛阶段期间，项目已经初步完成，目前系统能够正常运行和训练。接下来的排期规划主要围绕功能优化、模块完善、系统集成和文档整理展开，以确保项目的整体性能和质量进一步提升。

未来的工作中，强化学习模块将集中优化核心算法，通过调整参数配置和细化多维度奖励机制来提升策略收敛速度和训练效率。同时，ChatGPT推荐模块将在现有基础上强化策略推荐的精准性与智能化，完善服务层逻辑以保障数据交互的流畅性。图像处理模块将聚焦提升YOLO目标检测的识别精度和速度，同时对OCR模块进行优化以确保数字信息的高效识别。系统的服务层组件也将在测试过程中进一步完善，确保各模块间的高效协作。

系统集成和性能测试将在排期中占据重要位置。针对强化学习、推荐策略和图像处理模块的整合，系统将通过联合调试验证整体运行的稳定性与高效性。同时优化性能瓶颈并完善资源分配策略，确保硬件资源得到充分利用。针对集成测试中的问题，相关模块将进行逐步优化，以提升整体系统的稳定性和鲁棒性。

在完成系统功能优化后，团队还将专注于文档整理与演示材料的准备。所有技术文档将涵盖系统架构设计、算法原理、模块说明和测试结果，为后续的开发和交付提供全面支持。同时，基于典型应用场景生成的实验数据将用于系统演示，突出系统的核心能力和技术亮点。

在团队分工上，强化学习负责人将专注于算法优化和测试。ChatGPT模块开发者负责策略优化以及服务层组件的完善。图像处理负责人将继续改进YOLO和OCR模块，同时参与系统联调和性能优化。文档整理和技术说明由专人负责，以确保文档的完整性和系统的易复用性。

# 总结与展望

## 可行性分析

本项目的开发和实施充分展现了人工智能技术在游戏领域的强大潜力，同时为AI技术的多元化应用提供了新的方向和可能性。

本项目的价值主要体现在技术创新和实际应用的双重突破。在技术层面，项目通过整合YOLO图像识别、强化学习以及大语言模型（ChatGPT）等先进技术，构建了一个高度智能化、模块化且具备良好扩展性的决策框架。该框架的核心特性在于其能够模拟人类玩家的操作方式，从而实现高度接近真实玩家行为的游戏体验。这种创新有效解决了现有工具智能化程度不高、交互性不足的痛点，为玩家提供了更逼真且兼具策略性的游戏AI。此外，项目通过合规的方式避开了传统外挂工具存在的法律和道德风险，不仅提高了用户体验，也为行业树立了人工智能健康应用的标杆。在实际应用层面，项目具备广泛的适用性，无论是普通玩家、职业选手，还是游戏内容创作者，都能够从中受益。同时，项目为游戏公司提供了开发AI对手或训练工具的技术范例，进一步提升了其商业价值和社会影响力。

目前本项目的核心功能已经基本实现并达到了较高的成熟度。基于YOLO的游戏场景实时识别模块能够精准捕捉游戏中的动态变化，为AI决策提供可靠的视觉输入。强化学习算法驱动的操作模拟系统展现了卓越的环境适应能力和任务完成效率，能够在动态对局中快速调整策略，实现高水平的操作表现。由ChatGPT驱动的战术分析功能则为AI增添了实时建议生成的能力，进一步提高了游戏对局中的智能化水平。初步测试结果表明，项目在核心功能的完整性和实用性方面表现良好，具备较高的部署价值。然而，针对复杂场景的策略深度以及AI在长时间运行下的稳定性等方面，仍然存在进一步优化的空间，这些改进方向也为未来的技术提升提供了明确的目标。

未来，项目具备在功能、技术和商业化方向上的多元化拓展潜力。在功能升级方面，优化AI的策略深度和灵活性是首要目标，使其能够应对更复杂的战术场景，例如动态资源分配与多角色协同的团战策略。进一步提升用户体验是另一个关键方向，通过引入实时语音或文本接口，AI可以以更加直观的方式与玩家交互，提供实时操作指导或战术建议。在技术拓展方面，项目可以将现有框架应用于其他类型的游戏，包括策略类、冒险类甚至开放世界类游戏，从而验证AI模型的通用性与适应性，并为更多领域提供智能解决方案。在商业化方向上，项目可以与电竞赛事平台、直播平台以及游戏开发商合作，开发基于AI的全新娱乐内容，例如AI对战直播或智能陪练功能，为用户创造更加丰富的游戏与观赛体验。

# 参考资料

1. 新浪财经, "预见2024：《2024年中国网络直播行业全景图谱》

(附市场规模、竞争格局和发展前景等)" 新浪财经, 9月 11日,

2024, [https://finance.sina.com.cn/roll/2024-09-11/doc-incntvru8723832.shtml](https://www.google.com/url?sa=E&q=https%3A%2F%2Ffinance.sina.com.cn%2Froll%2F2024-09-11%2Fdoc-incntvru8723832.shtml).

1. 博思bosi, "游戏陪玩市场火爆，专业陪玩师成新职业宠儿" 搜狐, 4月 4日, 2024, <https://www.sohu.com/a/769227360_100233772>.
2. Newzoo (2024) *Global Games Market Report 2024*.
3. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), 779–88
4. Ray Smith, "An Overview of the Tesseract OCR Engine," in *Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)* (Curitiba, Brazil: IEEE, 2007), 1045–52, doi:10.1109/ICDAR.2007.4376991.

# 附录