

## РАЗРАБОТКА АВТОНОМНОГО АГЕНТА ДЛЯ ИГРЫ В VIZDOOM

**Кахикало К.Р.<sup>1</sup>**

**Научный руководитель – Сукачев П. П.<sup>1</sup>**

<sup>1</sup>Университет ИТМО

kkakhikalo@yandex.ru

### **Введение**

Уровни в видеоигре Doom предоставляют среду с большим пространством возможных действий. Уровни спроектированы как лабиринты и для их успешного прохождения нужно искать ключи, двери, которые они открывают, и одновременно вести бой с противниками. Для того чтобы проходить подобные уровни, агенту нужно одновременно использовать инструменты для навигации, осуществлять тактическое планирование, стрелять во врагов и уклоняться от их выстрелов. В работах, посвящённых созданию агентов для VizDoom решается конкретная задача в специально сконструированной для этого среде [1-2]. Цель этой работы состоит в том, чтобы создать модульного агента, который способен проходить стандартные, не адаптированные под агента уровни без предварительного обучения под каждый из них так, как это делал бы человек — используя только изображение с экрана. Модульность агента позволит менять его компоненты на другие, для того чтобы было возможно быстро проверять гипотезы о влиянии эффективности конкретного компонента на решение общей задачи прохождения уровня, а не на конкретный, специально созданный бенчмарк.

### **Основная часть**

Существующие модели уже сейчас способны определять своё положение в игровом мире, составлять карту и с помощью неё достигать своей цели [1-3]. Подход с обучением с подкреплением позволяет создать агента, который способен играть в достаточно сложные игры, вроде Doom, используя те же данные, которые видит игрок, без доступа к каким-то скрытым техническим переменным [4]. Объединив существующие подходы к навигации с моделью для выполнения простых игровых действий, возможно получить агента, который будет способен проходить простые уровни в игре, без отдельного предварительного обучения на них. Для анализа игрового поведения возможно использовать большие языковые модели [5], у которых будет доступ к базе знаний, к снимкам экрана, метрикам моделей более низкого уровня и к управлению моделями более низкого уровня. Это позволит планировать действия, использовать все знания, которыми обладает выбранная языковая модель и быстро реагировать на изменения в игровой среде с помощью модели низкого уровня.

### **Выводы**

Создана низкоуровневая модель, способная выполнять базовые игровые задачи в VizDoom, вроде стрельбы в противников. Обучена модель для визуального распознавания образов для игры VizDoom. Спроектирована общая архитектура для автономного агента.

### **Литература**

1. Bhatti S., Desmaison A., Miksik O., Nardelli N., Siddharth N., Torr P. H. S. Playing Doom with SLAM-Augmented Deep Reinforcement Learning [Электронный ресурс]. arXiv:1612.00380 [cs.AI]. 2016. URL: <https://arxiv.org/abs/1612.00380> (дата обращения: 17.12.2025).
2. Savinov N., Dosovitskiy A., Koltun V. Semi-parametric Topological Memory for Navigation [Электронный ресурс]. arXiv:1803.00653 [cs.LG]. 2018. URL: <https://arxiv.org/abs/1803.00653> (дата обращения: 17.12.2025).
3. Bakulin S., Akhtyamov T., Fatykhov D., Devchich G., Ferrer G. PixelNav: Towards Model-based Vision-Only Navigation with Topological Graphs [Электронный ресурс]. arXiv:2507.20892 [cs.RO]. 2025. URL: <https://arxiv.org/abs/2507.20892> (дата обращения: 17.12.2025).
4. Wydmuch M., Kempka M., Jaskowski W. ViZDoom Competitions: Playing Doom From Pixels // IEEE Transactions on Games. 2019. Vol. 11, No. 3. P. 248–259. DOI: 10.1109/TG.2018.2877047.
5. Wang C., Tang L., Yuan M., Yu J., Xie X., Bu J. Leveraging LLM Agents for Automated Video Game Testing [Электронный ресурс]. arXiv:2509.22170 [cs.SE]. 2025. URL: <https://arxiv.org/abs/2509.22170> (дата обращения: 17.12.2025).