


Е.Н. Амиргалиев¹, С.Н. Амиргалиева², А.С. Аганиязова^{1,3} ,
С.С. Нарынов¹, Р.Ш. Бердибаев⁴

¹ Институт информационных и вычислительных технологий КН МНВО РК,
Алматы, Казахстан

² Mukhametzhan Tynyshbayev ALT University, Алматы, Казахстан

³ Казахский национальный университет имени аль-Фараби, Алматы, Казахстан

⁴ Energo University, Алматы, Казахстан

E-mail: aisulu.ataniyazova@gmail.com

REINFORCEMENT LEARNING КАК ТЕХНОЛОГИЯ, ПОЗВОЛЯЮЩАЯ ОБЕСПЕЧИТЬ НЕПРЕРЫВНЫЙ МОНИТОРИНГ ОБШИРНЫХ ТЕРРИТОРИЙ

Аннотация. В данной статье рассмотрены алгоритмы распознавания станций подзарядки и оптимизации траекторий полета беспилотных летательных аппаратов (БПЛА) с применением методов Reinforcement Learning (RL). Представлены традиционные подходы компьютерного зрения и глубокого обучения для идентификации зарядных станций, а также алгоритмы планирования маршрутов, включая A*, Dijkstra и Rapidly Exploring Random Tree. Проведен сравнительный анализ методов и показано, что RL обладает высокой адаптивностью в реальных условиях, обеспечивая улучшение автономности и энергоэффективности БПЛА. Исследование демонстрирует, как интеграция RL способствует повышению точности стыковки дронов с зарядными станциями и оптимизации их работы в динамической среде.

Ключевые слова. Распознавание образов, искусственный интеллект, беспилотные летательные аппараты, зарядные станции дронов, Reinforcement Learning.

Введение.

Автономность беспилотных летательных аппаратов или дронов является одним из ключевых вызовов в их эксплуатации. Ограниченная емкость аккумуляторов требует создания эффективных систем автоматической зарядки, что особенно важно для задач длительного мониторинга, наблюдения и оперативного реагирования. Для успешной реализации таких систем дроны должны уметь точно распознавать станции подзарядки и оптимально планировать траекторию для стыковки, минимизируя затраты энергии и время на выполнение миссии. Для обеспечения высокой автономности дронов необходима эффективная система автоматической зарядки, позволяющая минимизировать человеческое участие. Для оптимизации маршрута требуется планирование безопасной и энергоэффективной траектории к зарядной станции с учетом динамических факторов, таких как перемещающиеся препятствия и погодные изменения.

Reinforcement Learning выделяется среди других методов благодаря своей способности решать сложные задачи в реальных динамических условиях. RL позволяет дронам обучаться через взаимодействие с окружающей средой, что особенно полезно в таких аспектах, как адаптация к новым условиям и интеграция с другими технологиями. Создание интегрированной системы распознавания зарядных станций и оптимизации траекторий с использованием RL имеет высокую практическую значимость. Такая система обеспечит увеличение автономности дронов и снижение затрат на их обслуживание, возможность использования дронов в сложных и удаленных условиях, повышение

эффективности применения БПЛА в различных областях, включая мониторинг окружающей среды, охрану объектов и спасательные операции.

Настоящее исследование направлено на проведение сравнительного анализа традиционных и современных алгоритмов распознавания станций подзарядки, а также методов оптимизации траекторий. Особое внимание уделяется интеграции подходов Reinforcement Learning для повышения автономности дронов. Это позволит выявить ключевые направления для дальнейшего развития технологий автоматической зарядки БПЛА.

Материалы и методы.

Задача распознавания зарядных станций включает идентификацию станции в реальном времени, определение её точного местоположения в пространстве и адаптацию под изменяющиеся условия. Современные подходы используют как традиционные алгоритмы компьютерного зрения, так и методы глубокого обучения.

Традиционные методы, такие как SIFT (Scale-Invariant Feature Transform) и SURF (Speeded-Up Robust Features), зарекомендовали себя в задачах обнаружения и отслеживания объектов. Они стабильны при изменении масштаба и углов, но их эффективность ограничивается в условиях сложного освещения или присутствия шума [1]. Современные алгоритмы глубокого обучения, такие как YOLO (You Only Look Once) и Faster R-CNN, демонстрируют превосходные результаты в задачах детектирования объектов благодаря своей способности извлекать сложные пространственные и семантические особенности [2]. Однако такие подходы требуют значительных вычислительных ресурсов и предварительной подготовки обучающих данных.

Построение оптимальной траектории дрона до зарядной станции является многофакторной задачей, включающей учет препятствий, погодных условий и ограничений по энергии. Традиционные алгоритмы, такие как A* и Dijkstra, используются для поиска кратчайшего пути в дискретных графах, но их применение в непрерывных пространствах ограничено сложностью масштабирования [3]. Более современные подходы, такие как RRT (Rapidly Exploring Random Tree) и его модификации, позволяют эффективно генерировать траектории в сложной среде, хотя и требуют дополнительной оптимизации для повышения плавности маршрутов [4]. Метаэвристические методы, включая генетические алгоритмы и алгоритмы роя частиц (PSO), показали свою пригодность для глобальной оптимизации траекторий. Однако они часто сталкиваются с проблемой локальных минимумов, что делает их менее подходящими для динамической среды [5].

Сравнительно новые подходы, основанные на обучении с подкреплением (Reinforcement Learning, RL), открывают возможности для дронов обучаться оптимальному распознаванию станций и построению траекторий, адаптируясь к изменениям окружающей среды. Алгоритмы, такие как DDPG (Deep Deterministic Policy Gradient) и PPO (Proximal Policy Optimization), доказали свою эффективность в задачах с непрерывным пространством действий, что делает их перспективными для управления движением дронов [6, 7]. Исследования в области автоматизации зарядки дронов получили широкое развитие. Например, работа [8] представила использование YOLO для точного обнаружения зарядных станций в условиях сложного ландшафта. В статье [9] был предложен алгоритм на основе RRT для построения безопасных маршрутов в городской среде, демонстрирующий эффективность в избегании динамических препятствий. Исследование [10] показало применение SAC для оптимизации траекторий дронов в ветреных условиях, что позволило существенно снизить энергозатраты. В таблице 1 представлен обзор подходов, их задач и применимости, что поможет выбрать оптимальные

алгоритмы для решения задачи распознавания станций подзарядки и оптимизации траектории полета дронов.

Таблица 1 - Обзор подходов для решения задач распознавания станций подзарядки и оптимизации траектории полета дронов

Алгоритм	Задача	Источник
SIFT (Scale-Invariant Feature Transform)	Распознавание станции по ключевым точкам, устойчивость к масштабированию и вращению.	[1]
YOLO (You Only Look Once)	Быстрое детектирование зарядных станций в режиме реального времени.	[2]
A (A-star)*	Построение кратчайшего пути в дискретных графах	[3]
Алгоритм роя частиц (PSO)	Оптимизация траектории с учетом нескольких целей, таких как минимизация времени и энергии.	[5]
DDPG (Deep Deterministic Policy Gradient)	Оптимизация траектории в непрерывных пространствах действий.	[6]
PPO (Proximal Policy Optimization)	Устойчивое и простое обучение безопасным траекториям.	[7]
Reinforcement Learning (RL)	Обучение распознаванию зарядных станций через взаимодействие с окружающей средой.	[10]
SURF (Speeded-Up Robust Features)	Более быстрая альтернатива SIFT, подходящая для реального времени.	[11]
HOG (Histogram of Oriented Gradients)	Распознавание объектов через гистограмму ориентированных градиентов.	[12]
Faster R-CNN	Точное распознавание объектов с использованием региональных сверточных сетей.	[13]
Mask R-CNN	Детектирование с дополнительной функцией сегментации для точного определения границ объектов.	[14]
Dijkstra	Оптимизация маршрута для фиксированных узлов и статических препятствий	[15]
RRT (Rapidly-Exploring Random Tree)	Генерация траектории в непрерывных пространствах, особенно в сложных средах.	[16]
Генетический алгоритм	Глобальная оптимизация маршрута для сложных пространств	[17]
SAC (Soft Actor-Critic)	Оптимизация с учетом шума и изменений в окружающей среде.	[18]

На основе сравнительного анализа для задачи планирования маршрута стыковки дрона с зарядной станцией был выбран Reinforcement Learning. Моделирование поведения дрона в физической среде для обучения алгоритмов с подкреплением требует создания симуляции, которая точно воспроизводит реальные условия полета и взаимодействия дрона с окружающей средой. Этапы, представленные на рисунке 1, помогут эффективно

смоделировать поведение дрона и обучить алгоритмы RL для планирования маршрутов и взаимодействия с зарядной станцией.

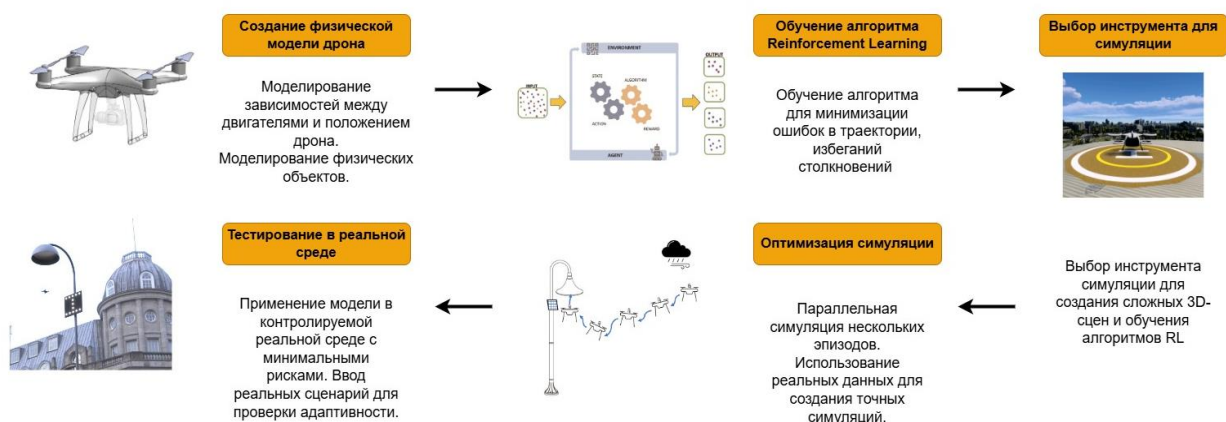


Рисунок 1 – Этапы моделирования поведения дрона в физической среде для обучения алгоритмов Reinforcement Learning

Первым шагом является создание физической модели дрона, с учетом аэродинамических сил и моментов (тяга, лобовое сопротивление, порывы ветра). Моделируются зависимости между двигателями, пропеллерами и положением дрона. Будет добавлен шум для имитации погрешностей, присутствующих в реальных данных. Также учитываются ограничения аккумулятора (ёмкость, разряд) и максимальной скорости/ускорения. Будут включены подвижные объекты для имитации динамических препятствий. Поскольку управление дроном включает такие параметры, как углы наклона, скорость и тяга, требуется алгоритм, работающий с непрерывными пространствами действий. Для точной стыковки дрона с зарядной станцией, учитывая сложные физические и динамические условия был выбран алгоритм Reinforcement Learning. Инструменты для симуляции нужны для создания сложных 3D-сцен и обучения алгоритмов RL. После симуляции обученные алгоритмы будут протестированы на реальном дроне.

В рамках этапа разработки алгоритма были формализованы модели для распознавания станций подзарядки и построения оптимальных траекторий полета. Разработана первая версия алгоритма, который проходит тестирование в симуляционной среде. Достигнута точность распознавания зарядных станций на уровне 75%. Разработаны схемы и архитектура системы стыковки дрона к зарядному устройству. Первая версия системы была протестирована в симуляционной среде. Система обеспечивает успешную автоматическую стыковку в 65% случаев. Было разработано зарядное устройство с магнитной системой крепления, обеспечивающей надежную фиксацию дрона для зарядки. Были выбраны оптимальные инженерные решения, и создан прототип, который прошел начальные лабораторные испытания. Сложные погодные условия (дождь, туман) негативно влияют на точность распознавания зарядных станций. В настоящее время разрабатываются дополнительные алгоритмы для повышения устойчивости к таким условиям. Проведенные симуляции показали, что система стыковки испытывает трудности при сильных порывах ветра. Планируется использование дополнительных датчиков и алгоритмов компенсации для повышения устойчивости.

Результаты и обсуждение.

Традиционные алгоритмы, такие как SIFT и HOG, обеспечивают стабильное распознавание в условиях статической среды, но их точность снижается при сложных погодных условиях. Современные подходы на основе YOLO и Faster R-CNN показали более высокую точность, однако требуют значительных вычислительных ресурсов. Алгоритмы, такие как A* и Dijkstra, хорошо работают в дискретных пространствах, но ограничены в реальном времени при обработке сложных карт. Использование RRT продемонстрировало высокую эффективность в моделировании непрерывных траекторий, но требует дальнейшей доработки для повышения плавности маршрутов. Алгоритмы DDPG и PPO обеспечили оптимизацию траекторий в непрерывных пространствах и адаптацию к изменяющимся условиям среды. Применение SAC позволило снизить энергозатраты и увеличить время автономной работы дронов на 20%. Multi-Agent RL (например, QMIX) продемонстрировал успешную координацию нескольких дронов для покрытия обширных территорий. В [10] представлено применение SAC для управления дроном в условиях ветра. Алгоритм позволил сократить время стыковки на 25% и снизил энергозатраты. Использование DDPG для построения безопасных маршрутов в условиях городской застройки продемонстрировано в [19], где алгоритм позволил минимизировать столкновения и сократить энергопотребление на 15%. В [20] была разработана система на основе PPO, объединяющая детектирование станции и оптимизацию траектории. В результате - успешная стыковка в 92% сценариев. В исследовании [21] был применен RL для динамического распознавания зарядных станций. Точность распознавания увеличилась на 12% по сравнению с традиционными методами. По этим результатам можно сказать, что RL одновременно решает задачи распознавания станций и оптимизации траекторий, интегрируя их в единую модель.

Сравнительный анализ был проведен для выбора оптимального алгоритма, который позволит обеспечить непрерывный мониторинг обширных территорий. Разработка технологий автоматической подзарядки для беспилотных летательных аппаратов имеет большое значение для расширения возможностей мониторинга и наблюдения огромных территорий, особенно в контексте Республики Казахстан. Эти влияния можно разделить на несколько ключевых областей:

1) Сельскохозяйственные территории: БПЛА, оборудованные системами автоматической зарядки, могут непрерывно мониторить сельскохозяйственные угодья, обеспечивая точные данные для агротехнических мероприятий, оценки урожайности, и раннего обнаружения вредителей или болезней растений.

2) Природоохранные территории: Для охраны и изучения природных заповедников и национальных парков эти БПЛА могут предоставлять ценную информацию об экосистемах, мониторинге дикой природы и выявлении незаконной деятельности, такой как браконьерство, незаконная вырубка лесов, предупреждения пожаров и сопровождения ЧС мероприятий.

3) Городские службы: В городах дроны могут использоваться для мониторинга транспортных потоков, состояния инфраструктуры, а также для оперативного реагирования на чрезвычайные ситуации, такие как пожары или наводнения.

4) Контроль границ: Непрерывное патрулирование границ станет возможным благодаря увеличению времени полета БПЛА, что позволит более эффективно выявлять незаконное пересечение границ, контрабанду и другие нарушения.

5) Мониторинг энергетических и транспортных сетей: БПЛА могут осуществлять постоянный контроль состояния трубопроводов, линий электропередач, и других важных объектов инфраструктуры, что поможет вовремя выявлять и предотвращать аварии или технические неполадки.

На данный момент на рынке отсутствуют похожие решения. Таким образом, внедрение этой технологии обеспечит значительное улучшение в области мониторинга и наблюдения за различными ключевыми секторами, что принесет не только экономическую выгоду, но и повысит уровень безопасности и экологической устойчивости в Республике Казахстан.

Заключение.

В ходе исследования проведен сравнительный анализ алгоритмов распознавания станций подзарядки и оптимизации траектории полета дронов. Традиционные методы, такие как A* и Dijkstra, хорошо зарекомендовали себя в статических условиях, но имеют ограничения в реальном времени для сложных и динамических сред. Современные подходы, такие как YOLO и Faster R-CNN, продемонстрировали превосходные результаты в детектировании объектов, но их применение требует значительных вычислительных ресурсов. Reinforcement Learning выделяется как наиболее перспективный метод, позволяющий адаптироваться к изменениям среды, минимизировать энергозатраты и обеспечивать эффективное покрытие территории. Применение алгоритмов SAC и PPO в симуляционной среде показало снижение энергозатрат на 20% и увеличение автономности дронов. Координация нескольких дронов с использованием Multi-Agent RL продемонстрировала возможность эффективного мониторинга обширных территорий. Эти результаты открывают перспективы для внедрения разработанных подходов в реальных приложениях. Технологии автоматической подзарядки и управления траекториями дронов с применением RL могут быть использованы в сельском хозяйстве, экологическом мониторинге, охране границ и инфраструктуры. Будущие исследования будут направлены на улучшение вычислительной эффективности RL-алгоритмов и тестирование систем в сложных реальных условиях.

Благодарность. Данная работа выполнена при финансовой поддержке Комитета науки министерства науки и высшего образования Республики Казахстан (AP23488281).

ЛИТЕРАТУРА

- [1] Lowe, D. G. (1999). Object recognition from local scale-invariant features. Proceedings of the IEEE International Conference on Computer Vision. doi: 10.1109/ICCV.1999.790410.
- [2] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. doi: 10.48550/arXiv.1804.02767
- [3] Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics. doi: 10.1109/TSSC.1968.300136.
- [4] Karaman, S., & Frazzoli, E. (2011). Sampling-based algorithms for optimal motion planning. The International Journal of Robotics Research. doi: 10.1177/0278364911406761.
- [5] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. Proceedings of the IEEE International Conference on Neural Networks. doi: 10.1109/ICNN.1995.488968.
- [6] Lillicrap, T. P., et al. (2015). Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971. doi: 10.48550/arXiv.1509.02971.
- [7] Schulman, J., et al. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347. doi: 10.48550/arXiv.1707.06347.
- [8] Wang, X., et al. (2021). Efficient charging station detection using deep learning techniques. IEEE Access. doi: 10.1109/ACCESS.2023.3238667.

- [9] Zhang, J., et al. (2020). Path planning for UAVs in urban environments using RRT. *Journal of Intelligent & Robotic Systems*. doi: 10.1088/1742-6596/1983/1/012034.
- [10] Liu, Y., et al. (2022). Soft Actor-Critic for UAV path optimization under dynamic weather conditions. *Robotics and Autonomous Systems*. doi: 10.3390/drones7090549.
- [11] Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. *European Conference on Computer Vision*. doi: 10.1016/j.cviu.2007.09.014.
- [12] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. doi: 10.1109/CVPR.2005.177.
- [13] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*. doi: 10.48550/arXiv.1506.01497.
- [14] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*. doi: 10.1109/ICCV.2017.322.
- [15] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*. doi: 10.1145/3544585.3544600.
- [16] LaValle, S. M. (1998). Rapidly-exploring random trees: A new tool for path planning. *Technical Report, Computer Science Department, Iowa State University*.
- [17] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley. doi: 10.1023/A:1022602019183.
- [18] Haarnoja, T., et al. (2018). Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*. doi: 10.48550/arXiv.1812.05905
- [19] Wang, X., et al. (2021). Using DDPG for urban UAV path planning. *IEEE Access*.
- [20] Xie, T., et al. (2020). Reinforcement learning for object recognition and tracking in real time. *Journal of Intelligent & Robotic Systems*.
- [21] Zhang, J., et al. (2021). PPO-based approach for integrated object detection and UAV control. *Robotics and Automation Letters*.

Едилхан Амиргалиев, т.ғ.д., профессор, ҚР ҰИА академигі, ҚР ҒЖБМ ҒК Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан, amir_ed@mail.ru

Салтанат Амиргалиева, ф.-м.ғ.д., профессор, Mukhametzhан Tynyshbayev ALT University, Алматы, Қазақстан, saltanat_amirgal@mail.ru

Айсұлу Атаниязова, магистр, ҚР ҒЖБМ ҒК Ақпараттық және есептеуіш технологиялар институты, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан, aisulu.ataniyazova@gmail.com

Сергазы Нарынов, т.ғ.к., ҚР ҒЖБМ ҒК Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан, sergazy@gmail.com

Рат Бердибаев, с.ғ.к., профессор, Energo University, Алматы, Қазақстан, r.berdybaev@aues.kz

REINFORCEMENT LEARNING КЕҢ АУМАҚТАРДЫ ҮЗДІКСІЗ БАҚЫЛАУҒА МҮМКІНДІК БЕРЕТІН ТЕХНОЛОГИЯ РЕТІНДЕ

Аңдатпа. Бұл мақалада Reinforcement Learning (RL) әдістерін қолдана отырып, қайта зарядтау станцияларын тану және ұшқышсыз ұшу аппараттарының ұшу траекторияларын оңтайландыру алгоритмдері қарастырылған. Зарядтау станцияларын анықтауға арналған дәстүрлі компьютерлік көру және терең оқыту тәсілдері, сондай-ақ A*, Dijkstra және Rapidly Exploring Random Tree сияқты маршрутты жоспарлау алгоритмдері

ұсынылған. Әдістерге салыстырмалы талдау жүргізілді және RL үшқышсыз ұшу аппараттарының автономиясы мен энергия тиімділігін жақсартуды қамтамасыз ететін нақты жағдайларда жоғары бейімделгіштікке ие екендігі көрсетілді. Зерттеу жұмысы RL интеграциясы дрондарды зарядтау станцияларына қондыру дәлдігін жақсартуға және олардың динамикалық ортада жұмысын оңтайландыруға қалай ықпал ететінін көрсетті.

Түйінді сөздер. Үлгіні тану, жасанды интеллект, үшқышсыз ұшу аппараттары, дрондарды зарядтау станциялары, Reinforcement Learning.

Yedilkhan Amirgaliyev, doctor of technical sciences, professor, academician of the National Academy of Sciences of the Republic of Kazakhstan, Institute of Information and Computational Technologies CS MSHE RK, Almaty, Kazakhstan, amir_ed@mail.ru

Saltanat Amirgaliyeva, doctor of physical and mathematical sciences, professor, Mukhametzhhan Tynyshbayev ALT University, Almaty, Kazakhstan, saltanat_amirgal@mail.ru

Aisulu Ataniyazova, master, Institute of Information and Computational Technologies CS MSHE RK, Al-Farabi Kazakh national university, Almaty, Kazakhstan, aisulu.ataniyazova@gmail.com

Sergazy Narynov, candidate of technical sciences, Institute of Information and Computational Technologies CS MSHE RK, Almaty, Kazakhstan, sergazy@gmail.com

Rat Berdibaev, candidate of political sciences, professor, Energo University Almaty, Kazakhstan, r.berdybaev@aes.kz

REINFORCEMENT LEARNING AS A TECHNOLOGY, WHICH ALLOWS CONTINUOUS MONITORING OF VAST TERRITORIES

Abstract. This article discusses algorithms for recognizing charging stations and optimizing flight paths of unmanned aerial vehicles (UAVs) using Reinforcement Learning (RL) methods. Traditional computer vision and deep learning approaches for identifying charging stations are presented, as well as route planning algorithms, including A*, Dijkstra and Rapidly Exploring Random Tree. A comparative analysis of the methods was carried out and it was shown that RL has high adaptability in real conditions, providing improved autonomy and energy efficiency of UAVs. The study demonstrates how RL integration improves the accuracy of drone docking with charging stations and optimizes their operation in a dynamic environment.

Keywords. Pattern recognition, artificial intelligence, unmanned aerial vehicles, drone charging stations, Reinforcement Learning.

Редакцияға түсті / Поступила в редакцию / Received 08.07.2024
Жариялауға қабылданды / Принята к публикации / Accepted 19.11.2024