Бубличенко Никита Сергеевич, магистрант,
ФГБОУ ВО «Уфимский государственный
нефтяной технический университет»

РАЗРАБОТКА СИСТЕМЫ ИНТЕЛЛЕКТУАЛЬНОГО ВИДЕОАНАЛИЗА ДЛЯ КОНТРОЛЯ ЗАГРУЖЕННОСТИ ПЛАВАТЕЛЬНЫХ БАССЕЙНОВ НА ОСНОВЕ ТЕХНОЛОГИЙ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Аннотация. Статья посвящена системе, которая по видеопотоку с камер над чашей бассейна оценивает число людей в воде и их распределение по дорожкам. Опираясь на публикации 2024–2026 гг., автор обосновывает выбор детектора и трекера, описывает архитектуру с обработкой кадров на месте и предлагает показатель загруженности для управления сеансами.

Ключевые слова: Компьютерное зрение, видеоаналитика, детекция людей, YOLO, многообъектное сопровождение, загруженность бассейна.

В обзоре 42 работ о применении искусственного интеллекта в плавании, вышедшем в 2026 г., больше 80 % исследований касаются стиля, поворотов или утомления пловца [1]. У массового бассейна задача приземленнее: знать, сколько людей в воде прямо сейчас и как они распределились по дорожкам. Турникет такого ответа не дает. Посетитель, миновавший его, может еще четверть часа провести в раздевалке и душе, так что разница между «вошли» и «вышли» описывает загрузку здания, а не чаши. Инструктор оценивает обстановку на глаз, и эта оценка, субъективная по своей природе, нигде не сохраняется. Камеры же над ванной есть почти везде – изображение с них, как правило, лишь пишется в архив.

Цель работы – обосновать архитектуру системы видеоанализа, определяющей в реальном времени число людей в чаше и их распределение по зонам на серийном оборудовании. Для этого решались следующие задачи: проанализировать публикации 2024–2026 гг. о детекции и сопровождении людей, в том числе в бассейнах; выбрать детектор и трекер с учетом точности и вычислительных ограничений; разработать структуру системы и показатели загруженности; определить методику проверки и ограничения применения.

Подсчет людей по изображению идет двумя путями: перечислением рамок, найденных детектором, или регрессией карты плотности, интеграл которой дает численность [2]. Карты плотности нужны там, где в кадре сотни перекрытых фигур. В бассейне людей редко больше нескольких десятков, а администратору важна позиция каждого относительно дорожек, поэтому, по моему мнению, здесь предпочтительна связка «детектор + трекер».

Среди детекторов тон задают модели YOLO и трансформеры. YOLOv10 отказалась от подавления немаксимумов (NMS), что сократило задержку вывода [3]; линейка YOLO11 – от nano до extra-large – позволяет подобрать модель и под сервер, и под периферийное устройство [4]; RT-DETR-R50 дает 53,1 % AP на COCO при 108 кадрах/с [5]. На воду эти цифры переносятся с потерями. Одна и та же YOLOv8 в фитнес-зале находила занимающихся в 96,36 % случаев, а в бассейне – в 91,94 % [6]: от пловца видна голова, иногда плечо, остальное скрыто бликами и пеной. Корейские авторы, сравнив YOLOv12 и RT-DETR, получили любопытный результат [7]. Лучшую локализацию дал небольшой собственный набор из 1986 кадров одного бассейна, тогда как его объединение с публичными данными выигрыша не принесло. Легкие модели справляются и со смежной задачей: детектор утопления на YOLOv8 показал точность 90,1 % против 88,5 % у YOLOv5 [8], а компактная YOLOv8n, обученная на 69 512 вручную размеченных подводных кадрах, – 98,3 % при 22 мс на кадр [9].

Трекеры работают по схеме «сопровождение через детекцию»: рамки сопоставляются с траекториями, а фильтр Калмана удерживает объект, временно пропавший из виду [10]. BoostTrack, дополнивший меру сходства расстоянием Махаланобиса и сходством формы рамок, в версии BoostTrack+ занял первые позиции по HOTA и IDF1 на MOT17 и MOT20 [10]. Сравнение в «Компьютерной оптике» показало, что быстрее всех пара ByteTrack + YOLOv8s, хотя по качеству она уступает OC-SORT + YOLOv6l [11]. Основные цифры собраны в таблице 1.

Таблица 1

Результаты работ 2024–2026 гг., значимые для задачи

Источник	Задача	Модель	Результат
[6]	Подсчет занимающихся в бассейне	YOLOv8	91,94 % (в фитнес-зале – 96,36 %)
[7]	Детекция людей в бассейне	YOLOv12 / RT-DETR	mAP@.5:.95 0,8704 / 0,7890; 4,82 / 7,1 мс на кадр
[9]	Детекция пловцов под водой	YOLOv8n	precision 98,3 %, 22 мс на кадр
[11]	Сопровождение людей	ByteTrack + YOLOv8s	62 кадра/с на MOT17
[12]	Подсчет пассажиров на Jetson Nano	YOLOv5s	96,85 %, около 12 кадров/с

Задачи разные, но вывод общий: компактная модель, дообученная на своей сцене, точна настолько, что тяжелая архитектура уже не окупается.

Архитектура системы

Система – конвейер из пяти модулей на вычислителе в здании бассейна; наружу уходят только числа. Модуль захвата принимает поток RTSP от одной-двух IP-камер. Наклонный ракурс охранной камеры под потолком у торцевой стены заставляет дальние фигуры наслаиваться на ближние, вид сверху это перекрытие сводит к минимуму – недаром системы, снимающие пловцов с дрона, восстанавливают позу даже при брызгах [13]. Если повесить камеру над серединой ванны нельзя, выручает калибровка: по углам чаши и дорожкам, размеры которых известны, вычисляется гомография, переводящая пиксели в метры на плане. Маска области интереса отсекает бортики и трибуну, блики гасятся поляризационным фильтром и выравниванием гистограммы (CLAHE). Детектор – YOLO11s, дообученная на кадрах конкретного бассейна с классом «человек в воде», рамка которого охватывает голову и плечи. Кадры для дообучения собираются при разном освещении и на разных сеансах, включая аквааэробику; судя по опыту [7], несколько тысяч размеченных кадров своего объекта полезнее, чем наращивание публичных данных. Трекер (ByteTrack, при запасе мощности – BoostTrack) присваивает пловцу анонимный номер. Без него счетчик бы «прыгал»: нырнувший пропадает на две-три секунды и, посчитанный заново, дал бы ложное колебание. Поэтому численность – это активные треки плюс потерянные не более 3–5 с назад. Аналитический модуль относит каждый трек к дорожке или зоне и раз в секунду отправляет показатели по MQTT на панель администратора и экран у входа.

Сколько кадров в секунду нужно? Пловец-любитель движется со скоростью 1–2 м/с, и при 10 кадрах/с смещается между кадрами не более чем на 0,2 м. Плате Jetson Nano, выдающей около 12 кадров/с [12], это по силам.

Показатели и методика проверки

Двадцать человек на шести дорожках – нагрузка комфортная, те же двадцать на двух дорожках свободного плавания – уже очередь на развороте. Поэтому основной показатель – коэффициент загрузки зоны:

$$K_i(t) = \tilde{N}_i(t) \cdot s_н / S_i, \quad (1)$$

где $\tilde{N}_i(t)$ – медиана численности в зоне i за окно 30 с; $s_н$ – нормативная площадь воды на человека, м²; S_i – площадь зоны, м². Медиана, в отличие от среднего, не реагирует на секундные сбои детектора. При K выше 0,8 зона на панели становится желтой, выше 1 – красной. Дополняют картину средняя длительность пребывания в воде, рассчитываемая по трекам, и недельный почасовой профиль, по которому удобно планировать сеансы и расстановку инструкторов. Проверку предлагается проводить по случайной выборке кадров (порядка 500), которые два разметчика считают независимо, с расчетом средней абсолютной ошибки (MAE); трекер оценивается отдельно, по HOTA и IDF1 на нескольких полностью размеченных двухминутных фрагментах. Результаты раскладываются по условиям – день, вечер, аквааэробика, детский сеанс, – иначе среднее скроет провал в трудном случае. Критерий приемки – MAE не более одного человека при численности до 30.

Ограничения

Камера снимает людей в купальниках, и авторы [7] называют это серьезным риском для приватности. Ответ заложен в архитектуру: кадры не записываются и не передаются, лица не распознаются, номера треков обнуляются после сеанса. Правовую оценку, по моему мнению, все же стоит согласовать с юристом учреждения. Есть и технические пределы: на плотной аквааэробике головы почти соприкасаются, ребенок рядом со взрослым частично скрыт, а в солнечные часы окна дают блики, которые фильтр гасит не полностью. А если камера уже висит неудачно? Вторая камера у противоположного торца, сведенная с первой через общую гомографию, обычно обходится дешевле недель дообучения. И еще одно: система контролирует загруженность, но спасателя не заменяет.

Таким образом, для контроля загруженности бассейна практичнее всего связка из дообученного детектора YOLO малого размера, легкого трекера и гомографии на план чаши. Качество в воде, как выявил обзор публикаций 2024–2026 гг., определяется прежде всего собственными данными объекта: две тысячи кадров одного бассейна дали лучшую локализацию, чем объединенный набор почти вшестеро большего объема. Коэффициент загрузки со сглаживанием медианой превращает выход детектора в понятный администратору сигнал, а обработка на периферийном вычислителе без записи кадров снимает основную часть вопросов о приватности. Следующий шаг – сбор данных на действующем бассейне и проверка по описанной методике.

Список литературы:

1. Patil M., Goudar R. H., Hukkeri G. S. AI for swimming recommendation systems: exploring the current landscape and research opportunities // Discover Applied Sciences. – 2026. – Vol. 8. – Art. 156.

2. Deng L., Zhou Q., Wang S., Górriz J. M., Zhang Y. Deep learning in crowd counting: a survey // CAAI Transactions on Intelligence Technology. – 2024. – Vol. 9, No. 5. – P. 1043–1077.

3. Wang A., Chen H., Liu L., Chen K., Lin Z., Han J., Ding G. YOLOv10: real-time end-to-end object detection // Advances in Neural Information Processing Systems (NeurIPS 2024). – 2024.

4. Khanam R., Hussain M. YOLOv11: an overview of the key architectural enhancements. – 2024.

5. Zhao Y., Lv W., Xu S., Wei J., Wang G., Dang Q., Liu Y., Chen J. DETRs beat YOLOs on real-time object detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 16965–16974.

6. Raksarson S., Teeravech K., Kupongpan P. Development of artificial intelligence system to support stadium access innovation for exercise // Rambhai Barni Research Journal. – 2025. – Vol. 19, No. 2.

7. Hong S.-S., Jeon H., Park C., Kim H. AI-based object recognition and risk detection technology for swimming pool safety prediction // Sensors and Materials. – 2026. – Vol. 38, No. 1. – P. 295–310.

8. Alharbi N., Aljohani L., Alqasir A., Alahmadi T., Alhasiri R., Aldajan D. Improved automatic drowning detection approach with YOLOv8 // Engineering, Technology & Applied Science Research. – 2024. – Vol. 14, No. 6. – P. 18070–18076.

9. Alzaabi H., Alzaabi S., Kohail S. Multi-swimmer drowning detection using a custom annotated underwater dataset and real-time AI // Lecture Notes in Computer Science. – Cham: Springer, 2026. – Vol. 16168. – P. 574–585.

10. Stanojević V. D., Todorović B. T. BoostTrack: boosting the similarity measure and detection confidence for improved multiple object tracking // Machine Vision and Applications. – 2024. – Vol. 35. – Art. 53.

11. Quan H., Ma G., Weichen Y., Bohush R., Zuo F., Ablameyko S. People tracking accuracy improvement in video by matching relevant trackers and YOLO family detectors // Компьютерная оптика. – 2024. – Т. 48, № 5. – С. 734–744.

12. Diaz-Santos S., Caballero-Gil P., Caballero-Gil C. Real-time passenger flow analysis in tram stations using YOLO-based computer vision and edge AI on Jetson Nano // Computers. – 2025. – Vol. 14, No. 11. – Art. 476.

13. Tran T., Joseph H. A., Lee K., Choo K. T. W., Ma D., Foong S., Kandappu T., Ko J., Balan R. Dronaquatics: real-time swimming analytics using drone captured imagery // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). – 2026.