

## РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЧЕСКОГО УЧЁТА УЧЕБНОЙ АКТИВНОСТИ СТУДЕНТОВ НА ОСНОВЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Деменева Ю.А.<sup>1</sup>

Научный руководитель — к.т.н., старший научный сотрудник, Ходненко И.В.<sup>2</sup>

<sup>1</sup>ГБОУ школа № 703

<sup>2</sup>Университет ИТМО

[demeneva.08@bk.ru](mailto:demeneva.08@bk.ru)

### Введение

Активное участие студентов в образовательном процессе является одним из ключевых факторов, влияющих на глубину усвоения материала и общую учебную мотивацию. Содержательные высказывания, инициативные вопросы, аргументированные ответы и участие в дискуссиях традиционно рассматриваются как важные индикаторы вовлечённости. В условиях потокового обучения, характерного для крупных технических университетов (в том числе Университета ИТМО), реальная оценка индивидуальной активности каждого учащегося преподавателем становится затруднительной. Педагог при желании поощрения инициатив может столкнуться с рядом проблем, таких как: вероятность пропуска значимых проявлений заинтересованности, временные затраты, субъективность оценки. В результате ручные методы учёта активности часто приводят к формальности оценки и снижению мотивации студентов. В связи с этим актуальной задачей является создание автоматизированной системы регистрации учебной активности на основе методов компьютерного зрения, которая позволит повысить объективность, сократить нагрузку на преподавателей и усилить мотивацию учащихся. В отличие от российских решений, таких как система «Computer Vision» Университета ИТМО (общая статистика присутствия), данная система фиксирует поднятие руки и считает персонально активность каждого студента.

### Основная часть

Целью работы являлась разработка системы автоматического учёта учебной активности студентов на аудиторных занятиях с перспективой интеграции в образовательный процесс университета.

На подготовительном этапе были собраны датасеты, отражающие реальные условия университетских аудиторий. С использованием инструмента разметки CVAT создан специализированный размеченный датасет объёмом около 12 тысяч аннотаций для дальнейшего дообучения моделей [1], учитывающий разные ракурсы съёмки, освещение и перекрытие участников занятия [2].

Проведён сравнительный анализ большого количества современных моделей компьютерного зрения. Особое внимание было уделено семейству YOLO [3], в силу их компактности и скорости инференса. Были последовательно протестированы и сравнены все версии YOLO-pose (YOLO11: n, s, m, l, x), а также все варианты моделей YOLO, предназначенных для детекции лиц, — от YOLOv8 до YOLOv12 включительно, каждая в размерах n, s, m, l. По результатам количественного и качественного анализа результатов моделей, для основного алгоритма были выбраны модели YOLOv11m-face и YOLOv11x-pose. Эта же модель использовалась для фиксирования события «поднятая рука», с помощью использования анализа угла подъёма руки в сочетании с информацией о позе и лице.

Для задачи идентификации студентов между кадрами были исследованы и сравнивались современные библиотеки и модели извлечения эмбеддингов: DeepFace,

SFace, ArcFace, Facenet, dlib, а также различные детекторы и выравниватели лиц (MTCNN, face-alignment и др.). Особое внимание уделялось качеству получаемого латентного пространства, в котором расстояние между векторами должно максимально точно отражать степень сходства лиц. В итоговой реализации была использована связка DeepFace (с приоритетом моделей SFace, ArcFace, Facenet в качестве fallback), а также FAISS — библиотека для эффективной организации эмбедингов учащихся [4]. На этапе сопоставления объектов между кадрами протестированы и сравнены методы: сопоставление по евклидову расстоянию между эмбедингами лиц, расширение рамки лица (bounding box expansion) с учётом траектории движения, а также комбинация этих подходов.

### **Выводы**

Разработана и протестирована система автоматической регистрации учебной активности студентов. Для разработки системы было выполнено сравнительное исследование более 30 архитектур нейронных сетей. На основе произведенного анализа были выбраны оптимальные модели для детекции позы и детекции лиц. Помимо этого был произведен анализ моделей для построения латентного пространства обучающихся. Выбранные модели и инструменты были объединены в единый сервис для обеспечения свободного доступа к разработанной системе.

Для тестирования системы использовалось 6 видеозаписей с разной обстановкой. Количество ошибочного распознавания руки составило 57 из 1984. За ошибку принималось не опознавание поднятой руки или лишнее опознавание поднятой руки. Большинство ошибок связано с перекрытиями объектов, их значительным расстоянием от камеры или засветами в кадре. Модель YOLO не смогла надёжно детектировать людей на дальних рядах, поскольку их размер в кадре был небольшим, а качество видео — недостаточным. Итоговая точность составила 97,13%.

Перспективными направлениями дальнейших исследований являются внедрение модуля анализа речевой активности, повышение устойчивости алгоритмов к неблагоприятным условиям съёмки, а также проведение контролируемого педагогического эксперимента для количественной оценки влияния системы на мотивацию и академическую успеваемость.

### **Литература**

1. Гужов С.В., Варшавский П.Р., Башлыков М.С., Тороп Д.В. Основы обучения и тестирование нейросетей моделей [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/722628/> (Дата обращения: 24.01.2024).
2. Magnus Tech Искусство аугментации: как улучшить модели компьютерного зрения без сбора новых данных [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/magnus-tech/articles/850070/> (Дата обращения: 20.03.2025).
3. Nuances of Programming Алгоритм YOLO простым языком [Электронный ресурс]. – Режим доступа: <https://dzen.ru/a/YNPrT1sjpDzRmG72> (Дата обращения: 07.02.2025).
4. ОККАМ Group FAISS: Быстрый поиск лиц и клонов на многомиллионных данных [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/okkamgroup/articles/509204/> (Дата обращения: 03.07.2025).