

УДК 004.93'1

*К.Б. Мухамадиева*

*Старший преподаватель кафедры «ИТС»*

*Profy university*

*Узбекистан. Ташкент*

**КОМПЬЮТЕРНОЕ ЗРЕНИЕ ДЛЯ АНАЛИЗА ПРОЦЕССА  
ОБУЧЕНИЯ  
АННОТАЦИЯ**

Автоматизированная аналитика обучения становится важной темой в образовательной сфере, которая нуждается в эффективных системах мониторинга процесса обучения и обеспечения обратной связи с преподавателем. Последние достижения в области визуальных датчиков и методов компьютерного зрения позволяют автоматически отслеживать поведение и эмоциональные состояния учащихся на разных уровнях, от университета до дошкольного учреждения. Целью данного исследования разработка структуры системы, которая позволила бы преподавателям фиксировать и составлять краткое описание поведения учащихся в классе как части сбора данных для принятия решений процесс. Система записывает весь сеанс и определяет, когда учащиеся проявляют внимание в классе, а затем сообщает об этом в учебные отделы.

Ключевые слова: поведение учащихся, визуальное внимание, распознавание лиц, оценка взгляда, классификация, машинное обучение.

*К.В. Mukhamadieva*

*Senior Lecturer at the Department of Information Technology and  
Systems*

*Profy university*

*Uzbekistan. Tashkent*

# **COMPUTER VISION FOR LEARNING PROCESS ANALYSIS**

## **ABSTRACT**

Automated learning analytics is becoming an important topic in the educational field, which needs effective systems to monitor the learning process and provide feedback to the teacher. Recent advances in visual sensors and computer vision techniques make it possible to automatically monitor the behaviour and emotional states of learners at different levels, from university to preschool. The aim of this research is to develop a system structure that would allow teachers to record and produce a summary of student behaviour in the classroom as part of the data collection for decision making process. The system records the entire session and identifies when students are showing attention in the classroom and then reports this to the educational institutions.

Keywords: student behavior, visual attention, face recognition, gaze estimation, classification, machine learning.

## **ВВЕДЕНИЕ**

Успеваемость студента зависит от множества факторов, включая качество преподавания, образовательные программы, условия обучения, количество учебных часов, академическую инфраструктуру, институциональный климат и финансовые проблемы. Важно создать эффективный инструмент, который позволит преподавателям и специалистам собирать данные о поведении учащихся с минимальными затратами человеческих ресурсов. Такой инструмент поможет разрабатывать стратегии поддержки учащихся, что, в свою очередь, повысит их академическую успеваемость.

## **ОБЩИЕ СВЕДЕНИЯ**

Многие исследователи изучали влияние поведения студентов на их успеваемость[1-3]. Было выявлено, что учащиеся, отвлекающиеся на просмотр видео, игры или отправку сообщений во время лекций на цифровых устройствах, испытывают сложности с долговременной памятью. Это, в свою очередь, негативно сказывается на их экзаменационных результатах, даже если кратковременная память не страдает. Кроме того, постоянное переключение между задачами приводит к когнитивным затратам: хотя эти затраты могут быть минимальными, в условиях многозадачности они накапливаются и существенно снижают эффективность обучения. Если учащиеся уделяют больше внимания занятиям, вероятность достижения ими высоких результатов значительно возрастает [2].

Для мониторинга поведения учащихся могут применяться два основных подхода: опросы и викторины. Однако эти методы не всегда объективны, так как студенты могут не помнить свои действия или исказить информацию. Современные технологии, в частности компьютерное зрение, позволяют автоматизировать процесс сбора и анализа данных о поведении студентов в режиме реального времени. Это дает возможность систематически анализировать образовательные потребности учащихся и разрабатывать индивидуальные стратегии обучения на основе собранных данных.

## **СУЩЕСТВУЮЩИЕ И ПРЕДЛАГАЕМАЯ СИСТЕМА**

Сегодня академическая аналитика позволяет фиксировать данные в реальном времени и применять предиктивное моделирование для прогнозирования вероятных результатов на основе наблюдаемых моделей поведения. В перспективе такие данные могут использоваться преподавателями для пересмотра учебных курсов, улучшения методов оценки и коммуникации со студентами. Например, если анализ поведения студентов

покажет низкий уровень вовлеченности, преподаватели смогут изменить содержание уроков, делая их более интересными и интерактивными. Кроме того, на основании анализа поведения можно проводить индивидуальные консультации с учащимися, у которых наблюдаются проблемы с концентрацией внимания.

Целью данного исследования является разработка алгоритма программной системы, основанной на технологиях компьютерного зрения, для автоматического распознавания поведения учащихся в классе. Система использует камеры, установленные в подходящих местах аудитории, в качестве модуля сбора данных. Распознавание лиц и движений тела позволяет анализировать посещаемость и уровень вовлеченности студентов в процесс обучения.

В данной системе применяется каскадное распознавание лиц методом Хаара для детекции лиц студентов, а также методы распознавания ориентиров лица и анализ выражений. Для анализа телодвижений используются каскадные классификаторы и гистограммы ориентированных градиентов (HOG). Определяются несколько ключевых состояний учащихся: активное участие в уроке, отсутствие концентрации и намерение покинуть аудиторию.

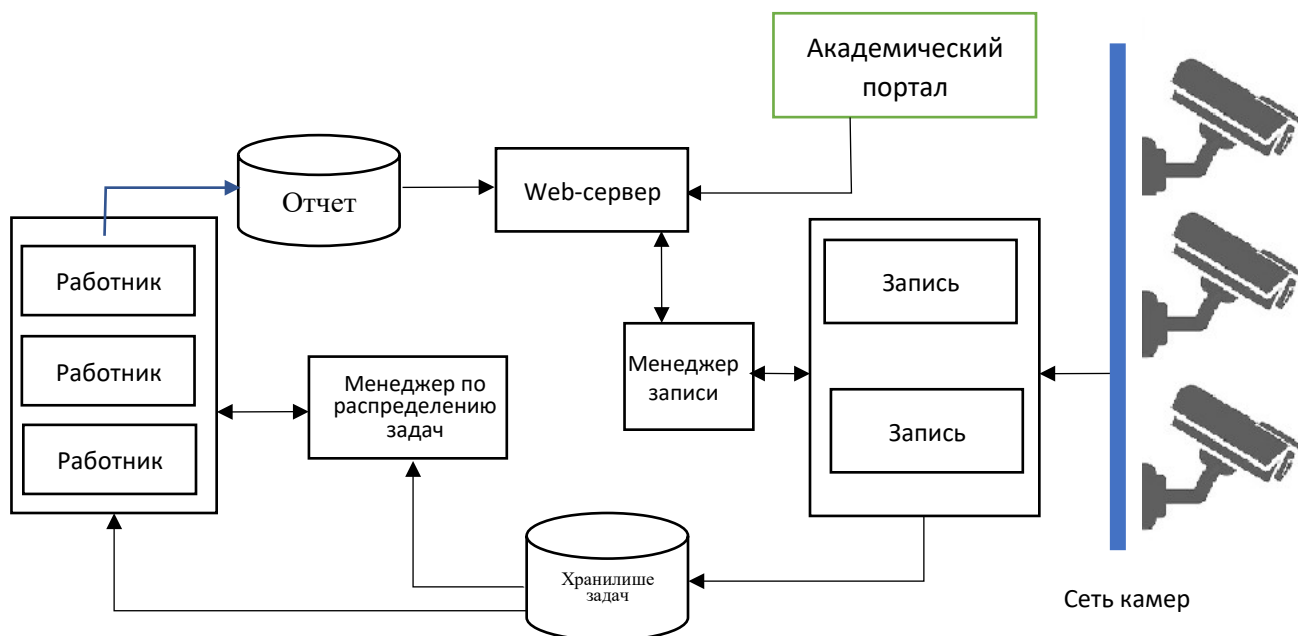
Специализированные устройства, такие как Microsoft Kinect, ранее использовались для анализа поведения учащихся. Однако существовали некоторые ограничения: оценки внимания, проводимые наблюдателями, были недостаточно точными, объем данных для обучения был ограничен, а устройства допускали ошибки в распознавании. Кроме того, семи признаков, вычисленных на основе данных Kinect, было недостаточно для полной оценки различий в поведении студентов, например, система не могла корректно распознавать процесс письма[1].

Современные технологии классификации поведения учащихся основаны на анализе их эмоций, варьирующихся от негативных до положительных. Также система сверяет лица студентов с базой данных учебного заведения, чтобы фиксировать посещаемость, и может оценивать состояние учащегося, например, выявлять признаки усталости или стресса. Хотя результаты большинства подобных исследований пока не опубликованы, они демонстрируют перспективность применения технологий распознавания лиц и анализа поведения для образовательных целей [2].

Основным вкладом данного исследования является разработка комплексного алгоритма для автоматизированного мониторинга поведения учащихся. Разработанная система представляет собой инструмент для сбора и анализа данных, помогающий принимать обоснованные решения в образовательном процессе. В системе применяются передовые методы компьютерного зрения, включая распознавание лиц, определение ориентиров на лице, встраивание лица, классификацию эмоций и анализ направления взгляда. Также реализован алгоритм определения местоположения в 3D-пространстве, позволяющий учитывать особенности планировки аудиторий и объединять данные с нескольких камер. Дополнительно используется статистический анализ предыдущих наблюдений для повышения точности алгоритмов компьютерного зрения. Внедренный алгоритм обобщения данных позволяет повысить эффективность анализа и сделать систему более адаптивной к различным условиям [3].

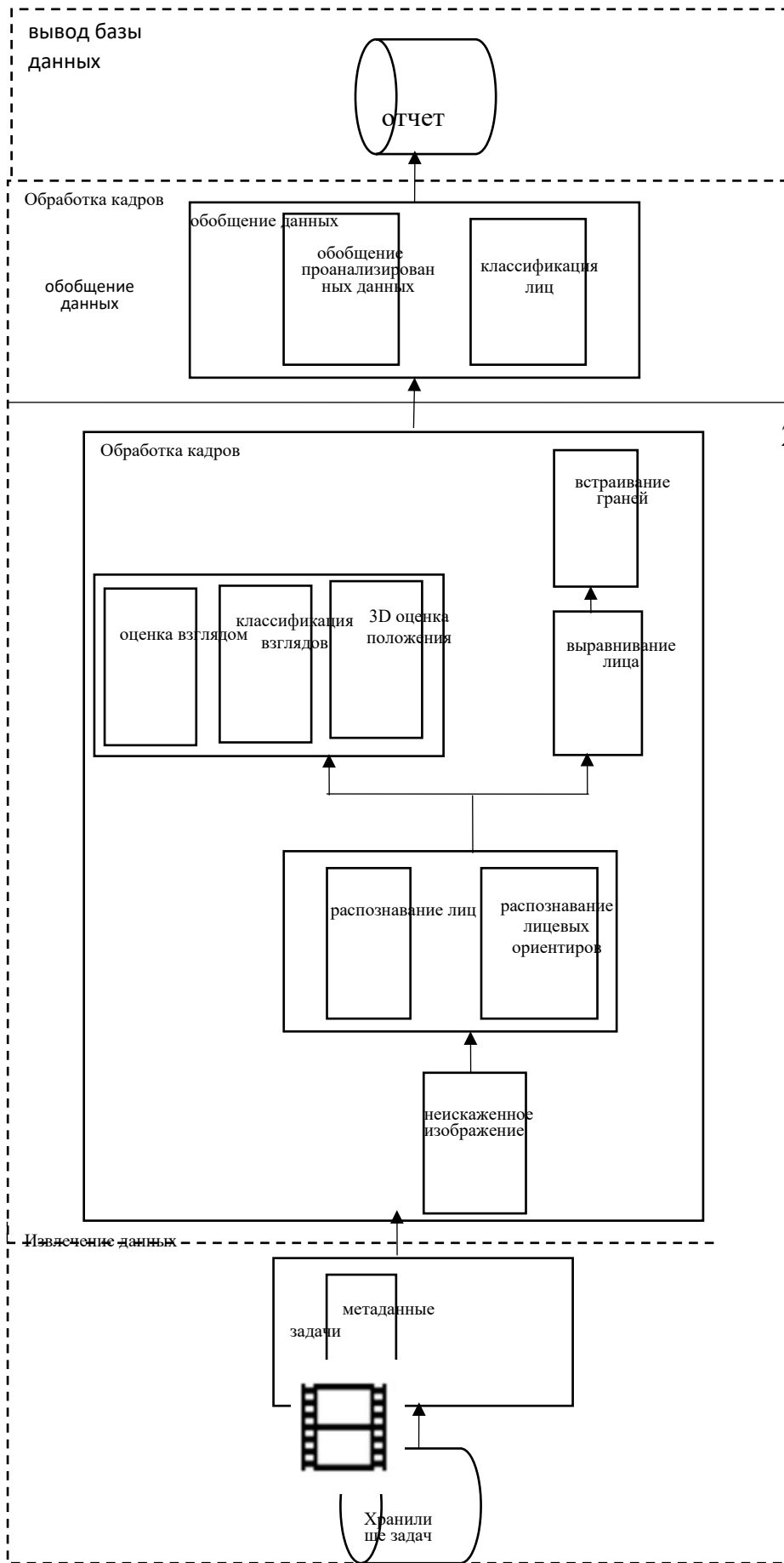
Система мониторинга поведения учащихся напрямую должна быть подключена к сети камер и академическому порталу для получения подробного расписания и уменьшения масштабов распознавания учащихся. На рисунке 1 представлена упрощенная схема системы. Она содержит семь основных компонентов: регистратор, контроллер регистратора, хранилище

задач, менеджер назначения задач, рабочий, отчет и веб-сервер. Видеорегистратор (или медиарегистратор мультимедиа) отвечает за запись видео с камеры. Контроллер видеорегистратора отвечает за распределение задач записи. Это означает, что необходимо указать, какой видеорегистратор будет записывать с какой камеры, поскольку процесс видеозаписи выполняется вручную. Между тем, сигнал для запуска/остановки процесса записи контролируется веб-сервером. Хранилище задач - это хранилище для хранения записанных видеороликов и их метаданных (класс в видео, список учащихся, конфигурации камер и т.д.). Менеджер по распределению задач отвечает за автоматическое получение расписаний и распределение задач для работника (worker). Worker который содержит модуль анализа данных (или ядро искусственного интеллекта), должен обрабатывать задачи, назначенные менеджером по распределению задач, и записывать их в базу данных отчетов в качестве результатов. Веб-сервер визуализирует данные из базы данных отчетов и управляет процессом записи. Можно видеть, что основой системы является модуль искусственного интеллекта, который находится внутри модуля «Работник», который можно разделить на четыре этапа: извлечение данных, обработка кадров, обобщение и вывод в базу данных, как показано на рисунке 1.



1 рисунок. Система сбора для обработки данных

На этом рисунке показан обзор системы, которая содержит семь компонентов: запись, контроллер записи, хранилище задач, диспетчер назначения задач, рабочий, а также отчет и веб-сервер.



2 рисунок. Система работы с данными



На этапе извлечения данных извлекается вся информация о задаче (записанное видео, список учащихся, конфигурации камер и т.д.), предоставленная менеджером по распределению задач. Данные, содержащиеся в видео, извлекаются, а метаданные передаются на этап обработки кадров. На этапе обработки кадров обрабатывается каждый видеокادر и выводятся ограничивающие рамки лица, ориентиры лица, встраивание лица, используя только изображение видеокадра, вектор взгляда, классификацию взгляда и оценку положения в 3D. Кроме того, эти выходные данные извлекаются с использованием изображения видеокадра и конфигураций камеры (положение в трехмерном пространстве, векторы поворота и т.д.), которые были получены из метаданных задачи. Этап подведения итогов отвечает за обобщение всех данных, полученных на этапе 2, для записи в базу данных. Он содержит два компонента: обобщение проанализированных данных и классификация лиц. Компонент обобщения проанализированных данных отвечает за обобщение данных предыдущего этапа. Компонент face classification использует список учащихся из метаданных задачи и данные о лицах для выполнения окончательной классификации после обобщения данных. Наконец, все результаты предыдущего этапа записываются в базу данных [5].

Распознавание лиц - это процесс обнаружения лиц, которые появляются в данной сцене. Это неотъемлемая часть большинства систем идентификации личности. Распознавание лиц - это ответвление общего обнаружения объектов. Общее обнаружение объектов делится на два типа: двухэтапное и одноступенчатое обнаружение объектов. Общее обнаружение объектов можно использовать для задачи распознавания лиц, обучив его на наборах данных о лицах. Двухэтапный алгоритм обнаружения делит процесс обнаружения на два этапа: сканирование в поисках интересных областей и классификация этих областей. Существует несколько популярных двухэтапных алгоритмов обнаружения, таких как R-CNN,

Быстрый R-CNN, более быстрый R-CNN, R-FCN и маскирующий R-CNN. Существуют также специальные алгоритмы для распознавания лиц, такие как MTCNN. В отличие от двухэтапного распознавания, одноступенчатый алгоритм непосредственно сопоставляет пиксели входного изображения с координатами ограничивающего прямоугольника и вероятностями классов. К числу последних одноступенчатых алгоритмов обнаружения относятся YOLO, YOLOv2, YOLOv3, SSD, RetinaNet, SSH детектор лиц и RetinaFace детектор лиц. За распознавание лицевых ориентиров отвечает функция выравнивания лица. Существуют отдельные детекторы лицевых ориентиров, такие как OpenCV landmark detector и DLib landmark detector, а также встроенные детекторы лицевых ориентиров, такие как MTCNN и RetinaFace [4].

Распознавание лиц, позволяющее идентифицировать конкретного учащегося, позволяет лицу, принимающему решения, понять реальную ситуацию в отношении каждого отдельного человека. Вместо того, чтобы просто следить за поведением всего класса или неустановленных лиц, создается профиль каждого учащегося. Встраивание лица - это процесс представления изображения лица в виде вектора чисел. Встраивание лица играет роль выделения признаков в системе распознавания лиц. Алгоритмы встраивания граней можно разделить на три типа в зависимости от их показателей потерь:

- Метрика, основанная на евклидовом расстоянии.
- Метрика, основанная на угловых/косинусных полях.
- Softmax и его варианты.

Существует несколько популярных, недавно разработанных алгоритмов встраивания лиц, таких как DeepFace, FaceNet, VGGFace, SphereFace и современный ArcFace, который обеспечивает точность 99,83% для помеченных лиц в условиях дикой природы (LFW) набор данных. Следующим шагом встраивания лиц является классификация лиц, которая является частью системы

распознавания лиц. Алгоритм классификации лиц использует векторы встраивания из алгоритма встраивания лиц и выводит классы идентификаторов (или удостоверения личности) заданных векторов встраивания [5]. Наиболее часто используемым методом для решения задач классификации лиц является определение ближайшего соседа (NN) с заданной метрикой и методом опорных векторов (SVM).

Как описано во многих работах, взгляд и выражение лица имеют большое значение важность для оценки когнитивной вовлеченности или невнимательности учащихся. Были разработаны некоторые методы для оценки взгляда глаз. Однако возможность извлечения взгляда может быть ограничена из-за размытых изображений, разрешения камеры и т.д. В связи с этим возникает необходимость использования положения головы в качестве альтернативного подхода. Для оценки положения головы 3DDFA и KEPLER определяют ориентиры на лице, а затем сопоставляют их с помощью сверточной нейронной сети (CNN) или ее модифицированной версии. Однако использование ориентиров может быть недостатком. Как и в случае с изображениями с низким разрешением, неправильное определение ориентиров может привести к ухудшению результатов. Norenet объединил ResNet50 с архитектурой с несколькими потерями. Каждая потеря содержала классификацию положения и регрессию, соответствующую рысканию, тангажу и крену по отдельности. Это показало, что он может напрямую предсказывать поворот головы и значительно превосходит методы привязки ориентира к позе, используя самые современные методы определения ориентиров. FSA-Net уделяет особое внимание оценке позы и даже оказался небольшим улучшением по сравнению с Norenet.

Оценка местоположения играет немаловажную роль в данной задаче. У двух учеников могут быть два одинаковых направления обзора, но объекты наблюдения могут быть разными. Это зависит от положения учеников. Например,

два соседа по столу (левый и правый) смотрят налево, но один ученик может смотреть на доску, в то время как другой смотрит за дверь. Таким образом, определение взаимного расположения учеников напрямую влияет на прогноз того, откуда они наблюдают. Распознавание лиц и определение поведения, по-видимому, слабо взаимосвязаны. Однако важно отслеживать поведение идентифицированного учащегося. Это обеспечивает многоуровневую детализацию при построении системы поддержки принятия решений. При этом обязательным требованием является тестирование в реальном времени для получения максимально эффективного результата.

## **ЗАКЛЮЧЕНИЕ**

В заключение можно отметить, что внедрение технологий компьютерного зрения в образовательный процесс открывает новые возможности для повышения качества обучения. Автоматизированный мониторинг поведения учащихся позволит преподавателям оперативно адаптировать учебные материалы и методы преподавания, а также обеспечит более индивидуальный подход к обучению. Таким образом, использование систем компьютерного зрения может стать важным шагом на пути к созданию более эффективных образовательных технологий.

## **СПИСОК ЛИТЕРАТУР:**

1. Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 4690–4699.
2. Glass, A.L.; Kang, M. Dividing attention in the classroom reduces exam performance. *Educ. Psychol.* 2019, 39, 395–408.
3. Yang, T.-Y.; Chen, Y.-T.; Lin, Y.-Y.; Chuang, Y.-Y. FSA-Net: Learning fine-grained structure aggregation for head pose estimation from a single image. In

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 1087–1096.

4. Мухамадиева К. Б. Сравнительный анализ алгоритмов распознавания лиц //Современные материалы, техника и технологии. – 2017. – №. 7 (15). – С. 58-63.

5. Мухамадиева К. Б. Анализ исследований по применению искусственного интеллекта в высшем образовании //Образование и проблемы развития общества. – 2020. – №. 2 (11). – С. 119-124.