

## ТЕХНИЧЕСКОЕ ЗАДАНИЕ

**на разработку опытного образца системы искусственного интеллекта  
«Система оценки вовлеченности обучающихся в образовательный  
процесс на основе анализа статических изображений  
учебных аудиторий во время занятий»**

### 1. ОБЩИЕ СВЕДЕНИЯ

1.1. Основаниями для разработки служат реализация учебного задания «Разработка программной реализации модели искусственного интеллекта как объекта техники и технологий» по треку «Технологии искусственного интеллекта и их применение» курса «Искусственный интеллект в высшей школе», а также потребность образовательных организаций в объективном мониторинге вовлеченности и посещаемости как компонентов эффективности образовательного процесса.

1.2. Целью проекта является повышение эффективности управления образовательным процессом и качества преподавания за счет создания автоматизированного инструмента мониторинга вовлеченности и посещаемости обучающихся в учебных аудиториях на основе анализа статических изображений (далее – фотографий, кадров).

Реализация системы позволит:

- обеспечить объективность данных о посещаемости и вовлеченности студентов, исключив субъективные оценки;
- организовать регулярный мониторинг ключевых параметров учебного процесса, влияющих на успеваемость и дисциплину;
- формировать аналитические отчеты для преподавателей, кафедр и деканатов для выявления проблемных групп и анализа динамики вовлеченности;
- проводить оценку влияния педагогических приемов, структуры занятий и стиля преподавания на внимание и активность обучающихся.

1.3. Объектом разработки является опытный образец системы компьютерного зрения, способный по статическим фотографиям аудитории во время учебных занятий автоматически:

- определить количество обучающихся, находящихся в помещении (метрика посещаемости);

– классифицировать обучающихся по признакам вовлеченности в образовательный процесс на основе визуальных индикаторов (метрика вовлеченности);

– формировать агрегированные показатели по каждому кадру: общее количество студентов, число вовлеченных, число не вовлеченных, доля вовлеченных.

1.4. Область применения системы охватывает процессы высшего образования, включая:

- регулярный мониторинг посещаемости и активности на занятиях;
- оценку вовлеченности студентов в образовательный процесс;
- поддержку управленческих решений по оптимизации преподавания;
- создание аналитических сводок на уровне кафедр, факультетов и учебных программ.

1.5. Актуальность проекта определяется «Стратегией цифровой трансформации отрасли науки и высшего образования», утвержденной Минобрнауки России, и растущей потребностью образовательных организаций в объективных, автоматизированных инструментах анализа качества преподавания и включенности обучающихся в учебный процесс.

1.6. Для достижения поставленной цели будут решены следующие задачи:

- сформировать верифицированный датасет изображений учебных аудиторий с разметкой количества обучающихся и уровней их вовлеченности;
- определить визуальные признаки и поведенческие индикаторы, коррелирующие с высокой и низкой вовлеченностью;
- разработать прототип системы компьютерного зрения, реализующий детекцию и классификацию объектов на изображении;
- подготовить методические рекомендации по применению системы в образовательной аналитике.

1.7. В рабочую группу по разработке системы входят:

- Никулина Юлия Владимировна, Сочинский институт (филиал) Российской университет дружбы народов;
- Бобчинский Артем Евгеньевич, Санкт-Петербургский государственный экономический университет;
- Копышева Татьяна Николаевна, Чувашский государственный университет;
- Желтова Кристина, Национальный исследовательский университет информационных технологий, механики и оптики.

## 2. ТРЕБОВАНИЯ К СИСТЕМЕ

### 2.1. Функциональные требования:

Таблица 1 – Функциональные требования к системе

№ п/п	Наименование требования	Описание
1.	Загрузка изображений	Прием одиночного изображения или пакета файлов в популярном формате (JPEG, PNG и др.).
2.	Предобработка изображений	Приведение к стандартному размеру, нормализация, аугментации при обучении.
3.	Обработка пакета изображений	Возможность обработки сразу нескольких изображений.
4.	Детекция обучающихся (посещаемость)	Обнаружение всех объектов класса обучающийся на изображении и подсчет их общего количества.
5.	Классификация вовлеченности	Определение признаков вовлеченности у каждого найденного обучающегося (например, по направлению взгляда, положению головы и т. д.).
6.	Расчет коэффициента вовлеченности	Вычисление отношения количества вовлеченных обучающихся к общему числу обучающихся, идентифицированных на изображении.
7.	Визуализация результатов	Отображение полученных значений в таблице и графиках.
8.	Экспорт результатов	Сохранение результатов анализа в машиночитаемом формате (CSV и др.) для дальнейшей аналитики.

### 2.2. Нефункциональные требования:

Таблица 2 – Нефункциональные требования к системе

№ п/п	Наименование требования	Описание
1.	Архитектура модели	В основе системы лежит архитектура object detection, способная локализовать и классифицировать объекты класса person с последующей оценкой вовлеченности. Возможное использование моделей типа YOLOv5, SSD или Faster R-CNN.
2.	Конфигурация выходного слоя	Модель должна обеспечивать определение количества человек на изображении, а также производить постобработку для классификации каждого объекта (например, «вовлечен» / «не вовлечен») на основе пространственного внимания, позы и других признаков.
3.	Оценка качества модели	Качество модели оценивается с использованием метрик, соответствующих выбранной архитектуре:

		<ul style="list-style-type: none"> <li>– для детекции: mAP (mean Average Precision), IoU (Intersection over Union);</li> <li>– для классификации вовлеченности: accuracy, F1-score.</li> </ul>
4.	Время обработки одного изображения	Среднее время обработки одного изображения не должно превышать 1 секунды при использовании GPU и 10 секунд при использовании CPU (на разрешении $1920 \times 1080$ px).
5.	Использование GPU при обучении	Обучение модели должно выполняться на графическом процессоре для ускорения вычислений.
6.	Масштабируемость	Система должна обеспечивать корректную обработку не менее 1000 изображений за один сеанс. Возможность параллельной пакетной обработки должна быть реализована.
7.	Поддержка дообучения	Система должна предусматривать возможность переобучения модели по мере накопления новых размеченных изображений, без необходимости полной переинициализации.
8.	Устойчивость к некачественным данным	Допускается ухудшение показателей точности при снижении разрешения до $1280 \times 720$ px или наличии шума/теней/перекрытий, при этом mAP не должен снижаться более чем на 15% от базовой валидации.
9.	Этические ограничения	Обработка изображений с участием обучающихся допускается только при наличии письменного согласия.
10.	Программная реализация	Разработка выполняется на языке Python с использованием библиотек PyTorch, OpenCV, Albumentations, Pandas, Matplotlib и специализированных библиотек детекции объектов (например, Ultralytics YOLOv5).

### 3. ТРЕБОВАНИЯ К ДАННЫМ

#### 3.1. Требования к входным данным:

- Входными данными системы являются цифровые изображения учебных аудиторий, зафиксированные в процессе учебных занятий.
- Допустимые форматы: .jpg, .jpeg, .png. Использование других форматов допускается только при условии их предварительной конвертации.
- Оптимальное разрешение изображения:  $1920 \times 1080$  пикселей.
- Изображения должны обладать достаточной резкостью и контрастностью для идентификации обучающихся.

- Изображение должно быть ориентировано на максимальный охват мест размещения обучающихся.
- Рекомендуемый объем обучающего датасета: от 300 до 1000 изображений.
- Каждое изображение должно иметь уникальное имя файла.
- К каждому изображению должна быть привязана одна или несколько текстовых аннотаций, содержащих:
  - a) координаты ограничивающих рамок для каждого обучающегося;
  - b) класс объекта: «вовлечен» или «не вовлечен».
- Каждая аннотация должна иметь уникальный идентификатор, позволяющий однозначно связать ее с соответствующим изображением.
- Для увеличения объема данных допускается аугментация (повороты, изменение яркости, масштабирование и пр.).
- Каждое текстовое описание должно быть проверено минимум тремя независимыми экспертами на предмет соответствия изображению.
- Для обеспечения корректного обучения и оценки модели данные должны быть разделены на:
  - a) обучающую выборку (train) – 70% общего объема изображений;
  - b) валидационную выборку (validation) – 15% изображений;
  - c) тестовую выборку (test) – 15% изображений.
- При разделении необходимо обеспечить равномерное распределение классов и визуальных условий (освещенность, ракурсы, плотность) во всех подвыборках.

### 3.2. Требования к выходным данным:

- Для каждого изображения должно быть возвращено:
  - a) исходное имя файла;
  - b) количество обнаруженных обучающихся;
  - c) количество обучающихся, отнесенных к классу «вовлечен»;
  - d) количество обучающихся, отнесенных к классу «не вовлечен»;
  - e) коэффициент вовлеченности (доля вовлеченных от общего числа);
- Выходные данные сохраняются в формате CSV или аналогичном, пригодном для последующего анализа.
- По каждому изображению должно обеспечиваться визуальное отображение предсказаний: изображение с наложенными рамками и классами.

## 4. АРХИТЕКТУРА СИСТЕМЫ

Разрабатываемая система имеет модульную структуру, обеспечивающую последовательную обработку изображений от этапа загрузки до формирования аналитических результатов. Каждый модуль выполняет строго определенные функции, обеспечивая надежность, масштабируемость и воспроизводимость работы системы.

Состав и назначение основных модулей приведены в таблице ниже.

Таблица 3 – Архитектура системы

№ п/п	Наименование модуля	Описание
1.	Модуль загрузки изображений	Принимает одиночные изображения или пакеты файлов в поддерживаемых форматах, выполняет первичную валидацию и подготовку к дальнейшей обработке.
2.	Модуль предобработки	Осуществляет очистку и стандартизацию изображений, включая: – приведение к единому размеру; – нормализацию пиксельных значений; – аугментацию (при обучении); – преобразование изображений в формат, пригодный для обработки нейросетью (векторизация, перевод в тензор и др.).
3.	Модуль детекции объектов	Реализует сверточную нейросетевую архитектуру типа object detection (например, Faster R-CNN, YOLOv5 и др.). Выходом модели являются bounding boxes и классы объектов: «вовлечен» и «не вовлечен».
4.	Модуль постобработки предсказаний	Выполняет агрегацию результатов по каждому изображению: – подсчет общего количества обучающихся; – определение числа вовлеченных и не вовлеченных; – расчет коэффициента вовлеченности.
5.	Модуль оценки качества	Используется в процессе валидации и тестирования. Оценивает качество модели по задачам: – для детекции объектов (обучающихся): a) mAP (mean Average Precision); b) IoU (Intersection over Union); – для классификации вовлеченности: a) accuracy (доля верно классифицированных объектов); b) F1-score (сбалансированная метрика для учета точности и полноты).
6.	Модуль интерпретации результатов	Формирует итоговый отчет, включающий: – предсказания модели по каждому изображению; – визуализации предсказаний и ошибок; – экспорт результатов в машиночитаемом формате (CSV, JSON).

## 5. ОБОСНОВАНИЕ ВЫБОРА МЕТРИК КАЧЕСТВА

Для комплексной оценки эффективности работы системы используются метрики, соответствующие двум ключевым задачам:

- обнаружению обучающихся на изображении (object detection) и
- оценке их вовлеченности (classification).

5.1. Метрики для задачи детекции обучающихся (посещаемость):

– IoU (Intersection over Union). Показывает степень пересечения предсказанных и истинных ограничивающих рамок.

– mAP (mean Average Precision). Усредненная точность по всем классам и уровням IoU. Является стандартом оценки качества в задачах object detection.

5.2. Метрики для задачи классификации вовлеченности:

– Accuracy. Доля корректно классифицированных объектов (обучающихся) по признаку вовлеченности.

– F1-score. Сбалансированная метрика, учитывающая точность (precision) и полноту (recall), особенно важна при наличии несбалансированных классов.

## 6. ПОРЯДОК ВЫПОЛНЕНИЯ РАБОТ

Проект реализуется в несколько этапов, охватывающих анализ требований, формирование датасета, проектирование и обучение модели, а также подготовку итоговых материалов. Этапы приведены в таблице ниже.

Таблица 4 – Этапы работ проекта

№ п/п	Наименование этапа	Срок реализации	Описание
1.	Анализ требований и формализация задачи		<ol style="list-style-type: none"><li>1. Уточнение постановки задачи на основе подхода object detection и классификации состояния обучающихся.</li><li>2. Определение набора метрик: mAP и IoU для детекции; accuracy и F1-score для классификации вовлеченности.</li><li>3. Согласование структуры аннотаций (bounding boxes + класс объекта).</li><li>4. Разработка чек-листа критериев вовлеченности и не вовлеченности.</li><li>5. Подготовка формы информированного согласия на обработку персональных данных.</li></ol>
2.	Сбор и подготовка обучающего датасета		<ol style="list-style-type: none"><li>1. Получение согласия участников съемки в письменной форме.</li><li>2. Фотографирование обучающихся в аудиториях во время занятий.</li><li>3. Обеспечение разнообразия в составе группы (число обучающихся, поведение, позы).</li></ol>

			<p>4. Создание разметки: определение bounding boxes для каждого обучающегося и присвоение класса «вовлечен» и «не вовлечен».</p> <p>5. Проведение разметки минимум тремя независимыми экспертами.</p>
3.	Проектирование архитектуры модели		<p>1. Выбор базовой архитектуры детектора (YOLO, Faster R-CNN, SSD или аналог), способной обнаруживать людей на изображениях.</p> <p>2. Определение схемы классификации вовлеченности: либо как второй этап после детекции, либо как составную часть единой архитектуры.</p> <p>3. Разработка пайплайна предобработки изображений и аугментаций, обеспечивающего устойчивость модели к реальным условиям съемки.</p>
4.	Обучение модели, оценка качества		<p>1. Настройка набора аугментаций: масштабирование, повороты, изменение яркости и контрастности, шум.</p> <p>2. Обучение модели детекции на размеченном датасете, подбор гиперпараметров.</p> <p>3. Отдельное или совмещенное обучение классификатора вовлеченности.</p> <p>5. Оценка модели по метрикам mAP, IoU, accuracy и F1-score на валидационной выборке.</p> <p>6. Отбор лучшей модели на основании совокупной оценки качества.</p>
7.	Тестирование		<p>1. Запуск модели на тестовой выборке.</p> <p>2. Анализ ошибок детекции и классификации.</p> <p>3. Формирование набора визуализаций: изображений с нанесенными bounding boxes и классами вовлеченности.</p> <p>4. Формирование отчета о результатах работы.</p>
8.	Подготовка документации и презентации		<p>1. Составление отчета о работе модели.</p> <p>2. Подготовка технической и пользовательской документации.</p> <p>3. Разработка презентационных материалов.</p> <p>4. Публичная презентация проекта.</p>

## 7. ЭТИЧЕСКИЕ АСПЕКТЫ

Обеспечение этичности сбора и использования данных:

7.1. Сбор, хранение и обработка изображений должны осуществляться в соответствии с требованиями Федерального закона № 152-ФЗ «О персональных данных», а также иными действующими нормативными актами.

7.2. Все обучающиеся, изображенные на фотографиях, обязаны предоставить письменное информированное согласие на использование своих изображений в рамках научно-исследовательского проекта.

7.3. Использование изображений допускается исключительно в рамках данного проекта и не может быть передано третьим лицам без отдельного согласия субъектов персональных данных.

7.4. Все данные должны храниться в защищенной цифровой среде с ограниченным доступом. Должны быть реализованы меры, предотвращающие несанкционированный доступ, утечку или утрату данных.

7.5. Разрабатываемая система не предназначена для индивидуальной оценки успеваемости, дисциплины или поведения конкретных обучающихся. Идентификация личности не производится.

7.6. Все оценки, формируемые системой, носят агрегированный характер и предназначены исключительно для анализа групповой динамики и качества образовательного процесса.

## **8. ОЖИДАЕМЫЙ РЕЗУЛЬТАТ**

В результате выполнения проекта будет создан опытный образец системы компьютерного зрения, способный:

- определять общее количество обучающихся на изображении с использованием алгоритмов object detection;
- классифицировать каждого обучающегося по признакам вовлеченности или отсутствия вовлеченности;
- рассчитывать коэффициент вовлеченности как отношение числа вовлеченных обучающихся к общему числу обучающихся в кадре;
- обрабатывать изображения в пакетном режиме с выводом результатов в формате CSV;
- демонстрировать устойчивость к изменяющимся условиям съемки (разное освещение, ракурсы, плотность размещения обучающихся);
- обеспечивать интерпретируемые и логически корректные предсказания;
- достигать приемлемых значений метрик качества:
  - a) для детекции:  $\text{IoU} \geq 0.5$ ,  $\text{mAP} \geq 0.6$ ;
  - b) для классификации:  $\text{accuracy} \geq 0.75$ ,  $\text{F1-score} \geq 0.7$ .
- формировать отчеты по результатам анализа с визуализацией и возможностью последующего применения в образовательной аналитике;
- соответствовать требованиям к этической обработке изображений, включая наличие письменного согласия, анонимность и ограничение на передачу данных третьим лицам.

## **9. РИСКИ РЕАЛИЗАЦИИ**

В процессе разработки и внедрения системы могут возникнуть определенные риски, связанные как с техническими, так и с организационными аспектами проекта. В таблице ниже приведены ключевые

потенциальные риски, их вероятность, возможные последствия и предлагаемые методы управления.

Таблица 5 – Риски реализации проекта

№ п/п	Риск	Вероятность	Последствия	Методы управления
1.	Недостаточный объем фотографий	Средняя	Переобучение модели, низкая обобщаемость	Расширение выборки за счет синтетических данных или открытых источников
2.	Сложности с получением согласий	Средняя	Юридические ограничения на сбор фотографий	Замена части данных на актерские постановки или стоковые фото
3.	Высокая вариативность условий съемки	Высокая	Нестабильные предсказания	Аугментация, нормализация
4.	Ошибки в разметке	Средняя	Снижение точности модели	Множественная аннотация, экспертные оценки
5.	Нехватка вычислительных ресурсов	Средняя	Задержки при обучении	Использование Google Colab, оптимизация модели
6.	Переобучение модели	Средняя	Высокая точность на обучающем, низкая на тесте	Регуляризация, кросс-валидация