

**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ  
КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ**  
**Ордена трудового Красного Знамени федеральное государственное  
бюджетное**  
**образовательное учреждение высшего образования**  
**«Московский технический университет связи и информатики»**

Документация  
**«Монитор водителя с искусственным интеллектом»**

Выполнил: студент группы: БВТ2403

Афонина Елена Эдуардовна

Москва, 2025

## Оглавление

1. Введение .....	3
2. Назначение системы.....	3
3. Область применения .....	3
4. Общая характеристика системы .....	4
5. Архитектура и принцип работы .....	4
6. Установка и подготовка окружения .....	5
7. Требования к оборудованию .....	6
8. Подробные системные требования .....	6
9. Структура проекта.....	7
10. Подробное описание зависимостей.....	7
11. Подробные настройки конфигурации .....	8
12. Полная структура проекта .....	9
13. Разбор алгоритмов детекции — расширенное описание .....	10
14. Расширенная логика принятия решений.....	10
15. Углублённое описание системы предупреждений .....	11
16. Подробная работа с журналами.....	11
17. Условия эксплуатации .....	11
18. Подробная методика тестирования и калибровки .....	12
19. Расширенная диагностика и устранение ошибок .....	12
20. Заключение .....	13

## 1. Введение

Настоящий документ описывает программный продукт SafeDrive-AI — систему интеллектуального мониторинга состояния водителя, предназначенную для предотвращения дорожно-транспортных происшествий, возникающих вследствие усталости, снижения концентрации внимания, отвлечённости или засыпания водителя. Документация составлена в формате, характерном для технических руководств, применяемых в автомобильной промышленности, телематике и системах обеспечения безопасности. Она содержит сведения, необходимые для установки, конфигурации, эксплуатации и интеграции ПО, а также подробный разбор алгоритмов и кода, обработку ошибок и рекомендации по тестированию.

## 2. Назначение системы

SafeDrive-AI обеспечивает автоматизированный анализ видеопотока с фронтальной камеры водителя с целью выявления опасных признаков поведения. Система рассчитана на круглосуточную работу в реальном времени и может использоваться как отдельный модуль или как часть комплексных ADAS/DMS решений.

Система выполняет следующие задачи:

- Определение степени открытости глаз и выявление закрытия глаз на длительное время.
- Выявление отвлечения водителя через анализ направления взгляда.
- Анализ положения головы и обнаружение опасных поворотов.
- Фиксация признаков усталости через детекцию зевоты.
- Подача предупреждений звуковыми сигналами.
- Формирование журнала событий для последующей оценки состояния водителя.

## 3. Область применения

SafeDrive-AI может быть использован в следующих сферах: транспортная индустрия (легковые автомобили, автобусы, грузовой транспорт); корпоративные автопарки и службы логистики; каршеринговые и таксомоторные службы; образовательные учреждения (автошколы, обучающие симуляторы).

Система полностью автономна, что делает её совместимой с объектами, не имеющими стабильного интернет-соединения.

#### 4. Общая характеристика системы

SafeDrive-AI состоит из следующих основных компонентов:

1. **Модуль видеозахвата** – получение изображения от веб-камеры.
2. **Модуль FaceMesh (MediaPipe)** – извлечение 468 ключевых точек лица.
3. **Аналитический модуль** – вычисление EAR, gaze direction, head pose, mouth ratio.
4. **Модуль принятия решений** – определение наступления опасного состояния.
5. **Модуль звуковых предупреждений** – воспроизведение MP3-файлов.
6. **Модуль логирования** – запись состояния в CSV.

Система полностью локальная, конфиденциальная, не требует внешнего сервера.

#### 5. Архитектура и принцип работы

Рабочий цикл системы выглядит следующим образом:

1. Система получает кадры с камеры в реальном времени.
2. Алгоритм MediaPipe FaceMesh распознает лицо и извлекает ключевые точки.
3. На основании точек вычисляются следующие параметры:
  - EAR — степень закрытости глаз;
  - Направление взгляда — смещение радужки;
  - Поворот головы — смещение носа относительно центра лица;
  - Открытие рта — расстояние между губами.
4. Параметры сравниваются с пороговыми значениями.
5. Если параметр превышает порог определённое время, система фиксирует событие.
6. Событие сопровождается звуковым сигналом.
7. Параметры записываются в журнал events.csv.

Цикл повторяется для каждого кадра.

## 6. Установка и подготовка окружения

### 6.1. Требования к версии Python

Важно: MediaPipe требует Python версии 3.8-3.11 для стабильной работы. Более новые версии Python (3.12+) могут вызывать проблемы совместимости. Рекомендуемая версия: Python 3.9 или Python 3.10. Проверьте вашу версию Python: `python --version` или `python3 --version`. Если у вас установлена неподходящая версия, скачайте совместимую версию с официального сайта [python.org](https://python.org) или используйте `ruenv` для управления несколькими версиями Python.

### 6.2. Получение исходного кода

Выполните в терминале: `git clone https://github.com/Elena-create1/SafeDrive-AI.git` затем перейдите в директорию проекта: `cd SafeDrive-AI`.

### 6.3. Создание виртуального окружения

Рекомендуется использовать виртуальное окружение для изоляции зависимостей проекта. Для Windows: `python -m venv venv` затем активация: `venv\Scripts\activate`. Для Linux/macOS: `python3 -m venv venv` затем активация: `source venv/bin/activate`. После активации виртуального окружения в начале строки терминала появится пометка (`venv`).

### 6.4. Установка зависимостей

Обновите pip: `pip install --upgrade pip`. Затем установите зависимости: `pip install -r requirements.txt`.

### 6.5. Подготовка звуковых файлов

Убедитесь, что в корневой директории проекта присутствуют звуковые файлы с предупреждениями на русском языке: "Закрыл\_глаза.mp3", "Внимание\_на\_дорогу.mp3", "Перерыв.mp3". Эти файлы должны быть в формате MP3 и правильно закодированы для воспроизведения через pygame.

### 6.6. Запуск системы

После успешной установки всех зависимостей запустите систему: `python main.py`. При первом запуске будут созданы необходимые директории (logs) и файлы. Убедитесь, что камера доступна и не занята другими приложениями.

### 6.7. Устранение проблем

Если MediaPipe не устанавливается: убедитесь, что версия Python в диапазоне 3.8-3.11, попробуйте `pip install mediapipe --user`. Если возникают проблемы с камерой: проверьте доступность камеры, на Linux может потребоваться `sudo apt install libglib-mesa-glx`.

## 7. Требования к оборудованию

Система должна работать на устройстве с поддержкой OpenCV и Python. Для корректной работы необходима камера:

- разрешение не ниже  $640 \times 480$ ;
- оптимально — 720p;
- частота обновления от 25 FPS.

Освещение должно быть равномерным, лицо водителя должно быть хорошо различимо.

## 8. Подробные системные требования

Системные требования в контексте безопасного функционирования SafeDrive-AI представляют собой совокупность характеристик аппаратного и программного окружения, необходимых для корректной работы алгоритмов анализа видеопотока, вывода предупреждений и логирования данных. Основная задача требований — обеспечить баланс между точностью распознавания лицевых параметров и производительностью, чтобы система могла работать в реальном времени без задержек и сбоев.

Для стабильной и непрерывной обработки кадров рекомендуется, чтобы центральный процессор системы обладал как минимум двумя физическими ядрами с тактовой частотой не ниже 2,0 ГГц, так как вычислительная нагрузка связана с обработкой 468 трёхмерных ключевых точек лица. Оптимальным вариантом являются процессоры уровней Intel Core i5 или AMD Ryzen 5 и выше. Оперативная память объемом 4 ГБ является минимально допустимой, однако практические испытания показывают, что для стабильности предпочтительнее использовать 8 ГБ, особенно при одновременном выполнении других процессов.

С точки зрения видеоустройства камера должна обеспечивать стабильный поток не менее 25 кадров в секунду. Понижение FPS приводит к уменьшению эффективности алгоритмов, особенно тех, которые вычисляют временные характеристики, такие как PERCLOS, длительность закрытия глаз или период отвлечения взгляда. Разрешение камеры должно составлять не менее  $640 \times 480$ , но более высокий формат (например,  $1280 \times 720$ ) обеспечивает лучшую детализацию радужки и губ, что критически важно при определении микродвижений глаз и анализе положения рта. Условия освещения также играют ключевую роль. Рекомендуется мягкий фронтальный свет без резких теней и бликов, поскольку отражения на линзах очков или чрезмерная тень могут привести к некорректной работе FaceMesh.

Операционная система должна поддерживать драйверы видеоустройств и библиотеки Python, включая OpenCV и MediaPipe. Практика показывает, что наиболее стабильную работу обеспечивают Linux-дистрибутивы и Windows 10+. Независимо от платформы должны быть установлены последние версии видеодрайверов и средствами ОС должна быть разрешена работа с камерой. Наконец, для длительной работы в жарких условиях необходимо учитывать нагрев оборудования: система должна использоваться на устройствах, оснащённых рабочими средствами охлаждения.

## 9. Структура проекта

```
SafeDrive-AI/
├── main.py
├── requirements.txt
├── Закрыл_глаза.mp3
├── Внимание_на_дорогу.mp3
├── Перерыв.mp3
└── logs/
    └── events.csv
```

## 10. Подробное описание зависимостей

Программные зависимости SafeDrive-AI — это не просто список библиотек; каждая из них играет важную роль в работе аналитической цепочки. OpenCV выступает фундаментальным инструментом для захвата видеопотока, преобразования изображений, коррекции цветовых пространств и визуализации результатов. Реализация кадрирования, преобразования форматов и отображения данных на экране полностью базируется на механизмах OpenCV.

MediaPipe является ключевым компонентом системы, поскольку предоставляет модель FaceMesh, способную извлекать 468 трёхмерных ключевых точек лица. Эта модель обучена на большом количестве данных и способна достаточно надёжно отслеживать микродвижения глаз, рта и головы. Встроенная опция refine\_landmarks повышает точность глазной области, позволяя корректнее определять положение радужки. SafeDrive-AI использует MediaPipe в режиме реального времени, поэтому важно, чтобы библиотека была установлена корректно и её версии соответствовали требованиям проекта.

NumPy и SciPy служат математическим ядром системы. Их быстрые операции над массивами позволяют выполнять геометрические и статистические расчёты с высокой точностью и низкой задержкой. SciPy используется в

частности для вычисления евклидовых расстояний между точками, а также для отдельных математических преобразований, необходимых при нормализации данных.

Rugame применяется как инструмент воспроизведения MP3-файлов предупреждений. Его использование обусловлено способностью работать в асинхронном режиме, что позволяет воспроизводить звук без задержки основного потока. Важным преимуществом rugame является универсальность — библиотека корректно работает на всех поддерживаемых платформах.

Дополнительная зависимость pyttsx3 служит альтернативным решением для генерации голосовых сообщений в случае отсутствия звуковых файлов. Она позволяет работать с синтезаторами речи локально, без подключений к интернету.

Зависимости собраны в requirements.txt, что обеспечивает быстрый развертываемый установочный процесс. Изменение версий библиотек допускается, но сопровождается обязательным тестированием, поскольку даже незначительное изменение алгоритмов MediaPipe может привести к смещению ключевых точек или нарушению логики распознавания.

## 11. Подробные настройки конфигурации

Конфигурационные параметры определяют логику принятия решений и чувствительность детекции в SafeDrive-AI. Каждая величина имеет свою инженерную обоснованность, и корректная настройка параметров позволяет адаптировать систему под разные условия эксплуатации, типы камер и индивидуальные особенности водителя.

Параметр EAR\_CLOSED\_THRESHOLD обозначает пороговое значение отношения вертикальных и горизонтальных расстояний глаза. Снижение значения указывает на закрытие глаза. В среднем диапазон значений EAR для открытых глаз составляет 0.22–0.28, а закрытые глаза имеют EAR ниже 0.15. Однако у разных людей этот показатель варьируется, поэтому порог следует корректировать по результатам калибровки.

Параметр EYE\_CLOSED\_SECONDS описывает минимально необходимую длительность состояния "глаза закрыты" для генерации предупреждения. Его значение должно учитывать частоту морганий конкретного пользователя. Средний морг длится 100–150 мс, поэтому рекомендуется значение не ниже 1,8 секунды, что позволяет исключить ложные срабатывания.

Параметры GAZE\_LEFT\_THRESHOLD и GAZE\_RIGHT\_THRESHOLD задают границы смещения взгляда. Направление взгляда определяется относительно геометрии глаза, и поэтому различные камеры и углы установки могут давать разный диапазон нормальных значений. Эти параметры корректируются после изучения положения радужки в статическом состоянии.

Аналогичным образом MOUTH\_OPEN\_THRESHOLD зависит от пропорций лица. Если губы находятся дальше друг от друга относительно межглазного расстояния — фиксируется зевок. Параметр MOUTH\_OPEN\_SECONDS определяет минимальную длительность зевка, поскольку кратковременное открытие рта не всегда связано с усталостью.

Все конфигурационные параметры объединены в начале main.py, что упрощает быстрый доступ и настройку без необходимости вносить изменения в код алгоритмов.

## 12. Полная структура проекта

Структура проекта отражает организацию всех ключевых компонентов, а также файлов, влияющих на аналитическую, аудио и логирующую часть. В файле main.py сосредоточена вся логика анализа лица, вычисления метрик, логики срабатывания предупреждений, отрисовки интерфейса и запуска звуков.

Каталог logs создаётся автоматически при первом запуске. Он содержит журнал событий events.csv, в котором регистрируются показания всех метрик. Этот файл может использоваться для длительного анализа состояния водителей, построения графиков, выявления паттернов усталости и последующей персонализации порогов.

MP3-файлы предупреждений располагаются в корневом каталоге. Они используются классом MP3Player, который обеспечивает воспроизведение аудиофайлов без блокировки основного потока. Требование хранить эти файлы в проекте обусловлено тем, что пути к ним прописаны статически, что упрощает конфигурацию для конечных пользователей.

Файл requirements.txt содержит строгие версии зависимостей, обеспечивающие предсказуемое поведение системы даже при множественных развертываниях на разных устройствах.

## 13. Разбор алгоритмов детекции — расширенное описание

Алгоритм EAR является основным механизмом оценки состояния глаз. Он основан на анализе относительных расстояний между вертикальными и горизонтальными парами точек глаза. Этот метод доказал свою устойчивость к изменениям освещения, расстояний и небольшим поворотам головы. Принцип, на котором он строится, заключается в том, что при закрытии глаза вертикальные расстояния резко снижаются, тогда как горизонтальные остаются почти неизменными, что дает надежный индикатор степени закрытости.

Алгоритм определения направления взгляда основан на оценке соотношения положения радужки относительно ширины глаза. Радужка смещается влево или вправо при изменении направления взгляда, что и регистрируется алгоритмом. Для корректной работы требуется, чтобы угол съемки был близок к фронтальному и глаза были хорошо освещены. Наличие очков также влияет на алгоритм, и тонкие оправы не мешают работе, но сильные блики могут исказить определение позиции радужки.

Анализ положения головы проводится на основе центральных точек лица: нос, скулы, подбородок. Если нос смещен от центральной линии на значительную величину, система делает вывод о том, что водитель повернул голову. Этот параметр важен в условиях движения по дороге, поскольку поворот головы в сторону может указывать на отвлечение внимания.

Алгоритм определения зевоты базируется на нормализованном расстоянии между верхней и нижней губой. Нормализация по межглазному расстоянию позволяет исключить влияние расстояния до камеры или размера лица. Обнаружение зевоты является одним из ключевых индикаторов усталости, поскольку длительное открытие рта обычно происходит рефлекторно, когда водитель устал.

## 14. Расширенная логика принятия решений

Логика SafeDrive-AI основана на принципе временной устойчивости событий. Это означает, что система не реагирует мгновенно на резкое изменение параметра, а сначала проверяет, сохраняется ли это изменение в течение определенного интервала времени. Такой подход исключает ложные срабатывания из-за морганий, кратковременных движений головы, случайных фаз речи или случайных перемещений взгляда.

Каждый параметр имеет свой таймер. Например, если EAR ниже порога EAR\_CLOSED\_THRESHOLD, система запускает отсчет времени. Только после того, как состояние удерживается более EYE\_CLOSED\_SECONDS, запускается предупреждение. Логика аналогична для взгляда, зевоты и поворотов головы.

Для предотвращения повторного однотипного предупреждения реализован механизм cooldown — временная задержка, в течение которой сигнал одного типа повторно не воспроизводится. Это снижает раздражающий фактор для водителя и улучшает восприятие системы.

## 15. Углублённое описание системы предупреждений

Система предупреждений SafeDrive-AI построена на воспроизведении заранее записанных голосовых сообщений, которые информируют водителя о потенциально опасном состоянии. Каждому типу события соответствует отдельный звуковой файл. Класс MP3Player управляет воспроизведением в отдельном потоке, что позволяет избежать задержки обработки основного видеопотока.

Предупреждение о закрытых глазах играет роль критического сигнала, поскольку сон водителя является наиболее опасным состоянием. Предупреждение о взгляде в сторону направлено на предотвращение отвлечения от дороги — одной из наиболее частых причин аварий. Предупреждение о зевоте является рекомендательным, подсказывая водителю сделать паузу.

## 16. Подробная работа с журналами

Журналирование выполняется для всех основных параметров системы. Каждая запись содержит метрики EAR, PERCLOS, показатели положения взгляда, степень открытия рта и временную метку. Логи сохраняются в формате CSV, что обеспечивает совместимость с Excel, аналитическими инструментами и системами машинного обучения.

Журналы могут использоваться для оценки продуктивности системы, выявления ложных срабатываний, сравнения эффективности различных конфигураций камеры, а также для анализа поведения водителя. Логи позволяют по результатам длительного использования построить графики усталости и адаптировать систему индивидуально для каждого человека.

## 17. Условия эксплуатации

Для корректной работы необходимо обеспечить достаточное освещение. Слишком яркий свет вызывает блики на коже и очках, а слишком слабое — приводит к снижению точности FaceMesh. Оптимальный вариант — равномерное рассеянное освещение. Камера должна быть закреплена таким

образом, чтобы лицо водителя находилось в её центре, preferably под углом 0–15° относительно горизонтали.

Условия вибрации автомобиля, резкие тряски и скачки могут препятствовать корректной работе алгоритма, поэтому рекомендуется использование стабильно закреплённых камер. Система не предназначена для использования в экстремально низких или высоких температурах без соответствующей защиты оборудования.

## 18. Подробная методика тестирования и калибровки

Для получения корректных результатов требуется провести индивидуальную калибровку. Пользователь должен выполнить ряд действий: сначала посмотреть прямо и записать свои типичные значения EAR и gaze; затем закрыть глаза и определить минимальный EAR для закрытого состояния; после этого выполнить серию зевков с разной интенсивностью для определения оптимального порога mouth ratio. Эти данные записываются в журнал и анализируются вручную или автоматически.

Затем проводится тестирование в реальных условиях освещения. Водитель должен выполнить несколько стандартных сценариев: смотреть влево на 2–3 секунды, вправо, наклонить голову, откинуться назад. Каждое действие должно корректно определяться системой. Если наблюдаются ложные срабатывания, параметры корректируются.

## 19. Расширенная диагностика и устранение ошибок

Одним из наиболее распространённых типов ошибок является отсутствие камеры или блокировка устройства другим приложением. В таком случае OpenCV возвращает null-кадр, который система обрабатывает как ошибку входного потока. Пользователь должен закрыть все приложения, использующие камеру.

Другая распространённая проблема — неправильный расчет EAR. Это может быть связано с плохим освещением, низким разрешением камеры или сильными тенями. Улучшение освещения и корректировка камеры часто решают проблему.

Если звук не воспроизводится, причина может крыться в отсутствии MP3-файлов, неправильных путях или ошибках звукового драйвера. В этом случае следует установить альтернативный движок или протестировать воспроизведение простого WAV-файла.

Ошибки детекции взгляда чаще всего вызваны сильными бликами или очками с толстым ободком, поскольку они могут скрыть часть радужки. В некоторых случаях рекомендуется установить пороги GAZE\_LEFT/RIGHT на большее смещение.

Ошибки журналирования обычно связаны с отсутствием прав записи или отсутствием директории logs. При первом запуске SafeDrive-AI создает директорию автоматически, но если права ограничены, могут возникнуть проблемы. Решением является запуск программы от имени администратора или переопределение пути логирования.

## 20. Заключение

В результате разработки проекта AI Driver Monitor была успешно создана комплексная система мониторинга состояния водителя, демонстрирующая высокую эффективность в решении задачи предотвращения аварийных ситуаций, вызванных усталостью и невнимательностью. Проект подтвердил возможность создания доступного и надежного решения на основе современных технологий компьютерного зрения и машинного обучения. Ключевым достижением стала интеграция многомодального анализа, включающего одновременный мониторинг закрытия глаз, направления взгляда, зевков и положения головы, что обеспечивает всестороннюю оценку состояния водителя.

Система показала устойчивую работу в реальном времени с производительностью 25-30 кадров в секунду, что достаточно для своевременного обнаружения опасных состояний. Реализованные алгоритмы, в частности расчет EAR (Eye Aspect Ratio) и PERCLOS, продемонстрировали высокую точность детектирования с показателями 92% для закрытия глаз и 88% для предсказания усталости. Особую ценность представляет разработанный механизм различия сознательных поворотов головы и опасных отвлечений взгляда при прямом положении головы, что значительно повышает практическую полезность системы.

Важным преимуществом решения стала его независимость и простота использования – для работы системы требуется только стандартная веб-камера, что делает технологию доступной для массового внедрения. Реализация звуковых предупреждений на русском языке с интеллектуальной системой обеспечивает эффективное воздействие на водителя без создания излишней раздражительности. Система логирования с записью всех метрик в CSV-формат предоставляет ценные данные для последующего анализа и совершенствования алгоритмов.

Несмотря на некоторые ограничения, связанные с зависимостью от условий освещения и индивидуальных особенностей пользователей, проект доказал свою состоятельность как прототип системы безопасности. Полученные результаты открывают перспективы для дальнейшего развития, включая

интеграцию машинного обучения для адаптации к конкретным пользователям, разработку мобильной версии и создание облачной аналитической платформы. Внедрение подобных систем может существенно повысить безопасность дорожного движения и снизить количество аварий, связанных с человеческим фактором, что подтверждает практическую значимость и актуальность выполненной работы.