

**Программная платформа компьютерного зрения для  
распознавания поведения водителя  
Описание функциональных характеристик**

Москва 2023

## **Аннотация**

В настоящем программном документе представлено описание функциональных характеристик программной платформы компьютерного зрения для распознавания поведения водителя, реализующего комплексную математическую модель выявления позы водителя для классификации типа инцидента в текущей ситуации с использованием технологий искусственного интеллекта.

### **1. ОБЩИЕ СВЕДЕНИЯ**

#### **1.1. Наименование программы**

1.1.1. Программная платформа компьютерного зрения для распознавания поведения водителя (далее – Программа).

#### **1.2. Условия использования программы**

1.2.1. Программа работает на персональных компьютерах под управлением Ubuntu.

1.2.2. Для работы программы необходимо наличие установленных программных пакетов Python 3.10, Tensorflow 2.12.

1.2.3. Объем программы 1,4 Мб.

1.2.4. Текст программного обеспечения на исходном языке приведен в документе «Текст программы».

1.2.5. Объем исходного кода – 182 Кб.

#### **1.3. Классы и характеристики пользователей**

1.3.1. Для работы с Программой квалификация работников должна соответствовать уровню уверенных пользователей операционной системы Microsoft Windows, а также подразумевать наличие необходимых знаний соответствующей предметной области.

## **2. ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ**

### **2.1. Функционал программы**

2.1.1. Программа предназначена для определения позы водителя, положения рук и головы для оценки опасного поведения за рулём, приводящим к следующим возможным инцидентам:

- разговор по мобильному телефону;
- курение;
- употребление пищи/напитков;
- отвлечения от дороги;
- засыпание/усталость.

2.1.2. Программа реализует нейросетевые алгоритмы классификации типа опасного инцидента в процессе управления транспортным средством.

2.1.3. Программа предназначена для работы в составе единого интеллектуального программного обеспечения аппаратно-программного комплекса видеоаналитики (АПК «РЕЙС»).

### **2.2. Особенности функционирования программы**

2.2.1. Для оценки позы водителя реализованы процедуры обработки видеопотока.

2.2.2. Оценка положения суставов водителя в 3D измерении однозначно определяются по их 2D местоположениям на видеокадре используя модель FBI (Forward-or-Backward Information).

2.2.3. При оценке принимается соотношение длин составных частей тела, в качестве априорной информации модель выбирает отображение  $y = f(I)$  в виде трёх подфункций:

- отображение от входного изображения на 2D местоположения суставов,
- отображение от входного изображения на FBI,
- отображение от этих двух условий на 3D координаты суставов.

Все эти подфункции аппроксимируются с помощью глубоких нейронных сетей.

2.2.4. Для поддержки обучения Программы случайным образом выбираются изображения поз из набора данных для настройки модели машинного обучения (датасета).

### 3. ОПИСАНИЕ ЛОГИЧЕСКОЙ СТРУКТУРЫ

#### 3.1. Алгоритм работы программы

3.1.1. Алгоритм работы Программы показан на рисунке 1.

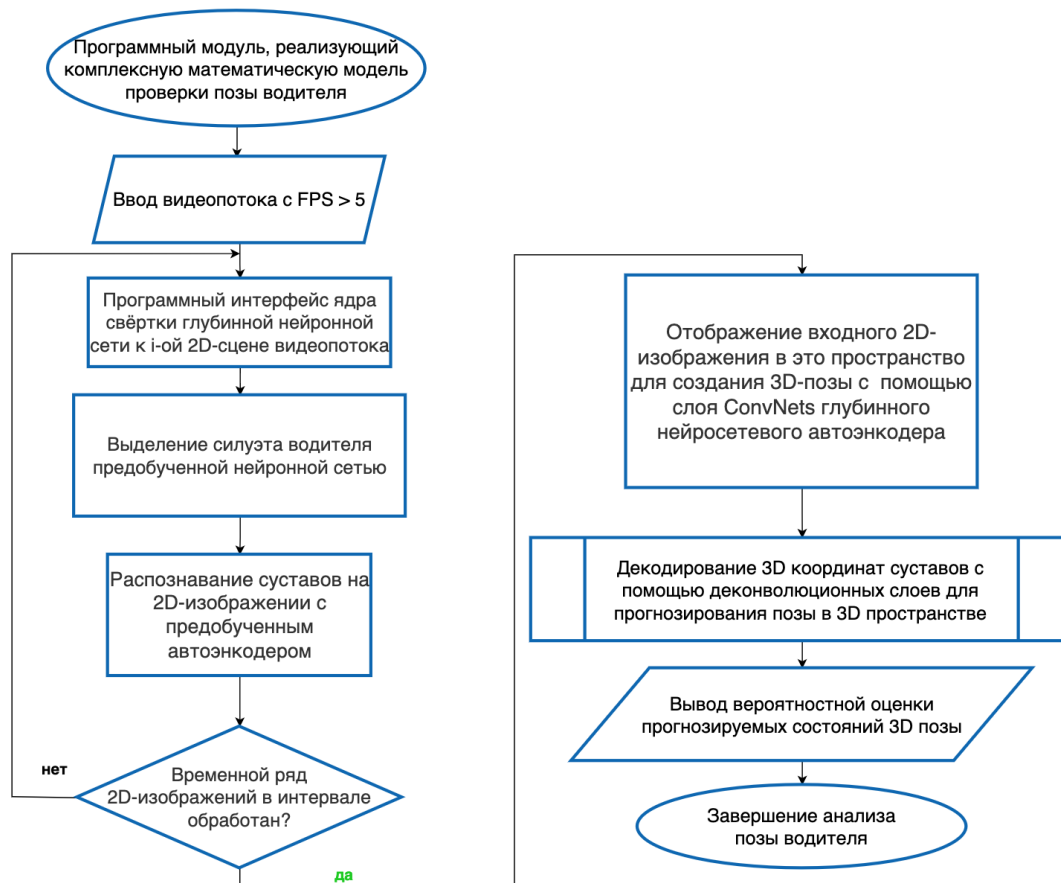


Рисунок 1 – Алгоритм работы Программы.

3.1.2. В алгоритме программы нейросетевой автоэнкодер принимает на вход сгенерированную матрицу вероятностей слоя *softmax*, которая предоставляет информацию о предобработанном видеопотоке с изображением водителя.

#### 3.2. Методы решения

3.2.1. Программа получает информацию для определения 3D позы человека из видеопотока, фиксирующего водителя.

3.2.2. Для маркировки изображения аннотаторам FVI используется бинарный выбор для каждой кости. Это позволяет использовать бесконечно большой набор обучающих данных.

3.2.3. Вначале объединяются 2D-локации и матрица вероятностей, а затем они сопоставляются с 3D-позой с помощью двух каскадных блоков.

3.2.3.1. В частности, каждый блок отображает входной признак в более высокое измерение с помощью двух полностью связанных слоев.

3.2.3.2. В нейросетевом автоэнкодере типа Unet используется 1024 нейрона, чередующихся со слоями пакетной нормализации, отсева и функцией активации ReLU.

3.2.3.3. В конце первого блока используется дополнительный линейный слой для получения грубой трехмерной позы.

3.2.3.4. Далее она перепроецируется в 1024-мерное пространство и служит частью входных данных для второго блока.

3.2.4. Остаточные связи между двумя блоками применяются для полного использования информации.

## **4. ВЫЗОВ И ЗАГРУЗКА**

4.1.1. Вызов, загрузка и завершение работы программного модуля осуществляется стандартными средствами ОС.

## **5. ВХОДНЫЕ ДАННЫЕ**

5.1.1. Входными данными программы являются 2D- изображения с камеры, направленной на водителя.

## **6. ВЫХОДНЫЕ ДАННЫЕ**

### **6.1. Выходная информация**

6.1.1. Выходными данными ПО являются.

- вероятность типа инцидента в текущей ситуации;
- изображение (маска) рук и головы водителя (рисунок 2).

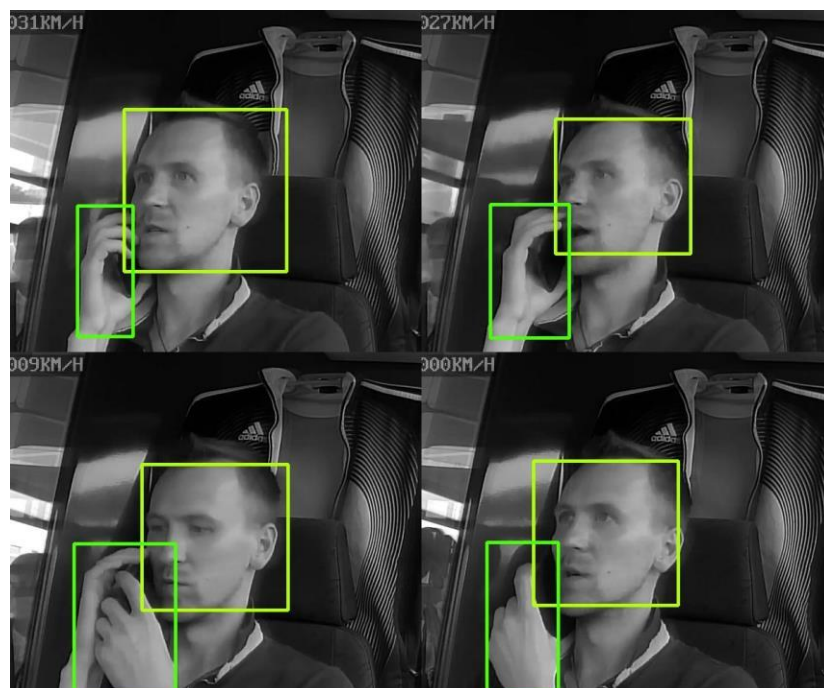


Рисунок 2 – Отображение результаты работы программы - отображение позы руки и головы водителя

### **Перечень сокращений и условных обозначений**

АПК аппаратно-программный комплекс

БД база данных

ИИ искусственный интеллект

НС нейронная сеть

ОС операционная система

ПК персональный компьютер

ПО программное обеспечение

ТЗ техническое задание

ТС транспортное средство

CNN Convolutional Neural Networks, сверточная нейронная сеть

FBI Forward-or-Backward Information

ROCAUC Receiver Operating Characteristic Area Under the Curve, метрика, используемая для оценки качества бинарной классификации модели