

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технологический университет «СТАНКИН»**

**(ФГБОУ ВО «МГТУ «СТАНКИН»)**

|  |  |
| --- | --- |
| **Институт**  **информационных технологий** | **Кафедра**  **информационных систем** |

|  |
| --- |
| **«УТВЕРЖДАЮ»**  И.о. заведующего кафедры ИС  к.т.н., доц. Бильчук М.В.    « » 2023 г. |

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

|  |
| --- |
| на выпускную квалификационную работу по направлению подготовки 09.03.02 «Информационные системы и технологии» |
|  |
| **Тема: «Разработка приложения для управления курсором при помощи жестов рук»** |

|  |  |
| --- | --- |
| Обучающийся  Группы ИДБ-20-05 Грошев Владимир Павлович | Руководитель  к.т.н., доцент  Волкова Ольга Рудольфовна |

Тема утверждена приказом от «\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_2023 г. № \_\_\_\_\_\_\_\_.

Срок сдачи законченной ВКР на кафедру «\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_2024 г.

**ОГЛАВЛЕНИЕ**

[**ВВЕДЕНИЕ** 3](#_Toc153646037)

[**ГЛАВА 1.** **АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ** 5](#_Toc153646038)

[**1.1.** **Методы распознавания и их классификации** 5](#_Toc153646039)

[**1.2.** **Проблемы** 6](#_Toc153646040)

[**1.2.1.** **Проблемы сегментации** 6](#_Toc153646041)

[**1.2.2.** **Проблемы определения жестов** 7](#_Toc153646042)

[**1.2.3.** **Проблемы, связанные с определением обеих рук** 8](#_Toc153646043)

[**1.3.** **Пример архитектуры с распознаванием жестов** 8](#_Toc153646044)

[**1.4.** **Компьютерное зрение** 9](#_Toc153646045)

[**1.5.** **ГЛУБОКОЕ ОБУЧЕНИЕ** 11](#_Toc153646046)

[**1.5.1.** **Применение глубокого обучения для компьютерного зрения** 11](#_Toc153646047)

[**1.5.2.** **HandNet** 12](#_Toc153646048)

[**1.5.3.** **EfficientDet** 13](#_Toc153646049)

[**1.5.4.** **CenterNet** 13](#_Toc153646050)

[**1.5.5.** **Отличия между CenterNet и EfficientDet** 14](#_Toc153646051)

[**1.5.6.** **Сравнительный анализ архитектур HandNet, CenterNet и EfficientDet для решения задачи обнаружения и сегментации рук в режиме реального времени.** 15](#_Toc153646052)

[**1.6.** **Анализ правовых и нормативно-технических документов** 15](#_Toc153646053)

[**1.6.1.** **ГОСТ Р ИСО/МЭК 25000-2021** 15](#_Toc153646054)

[**1.6.2.** **ГОСТ Р ИСО/МЭК 60447-2015** 16](#_Toc153646055)

[**1.6.3.** **ГОСТ Р ИСО/МЭК 9241-10** 17](#_Toc153646056)

[**1.6.4.** **ГОСТ Р ИСО/МЭК 70462.1-2022** 18](#_Toc153646057)

[**ЗАКЛЮЧЕНИЕ** 20](#_Toc153646058)

[**Список литературы** 22](#_Toc153646059)

# **ВВЕДЕНИЕ**

Взаимодействие человека с компьютером является важной составляющей современной информационной технологии, так как человек в независимо от дисциплины, так или иначе взаимодействует с ними. В последние годы наблюдается активное развитие новых методов управления компьютерными системами, включая технологии распознавания жестов руки. Распознавание жестов руки позволяет пользователю взаимодействовать с компьютером, используя жесты и движение рук, что может быть особенно полезно в случаях, когда традиционные методы ввода ограничены и неудобны.

Распознавание жестов рук заключается в получении необработанных данных, которые могут быть выполнены в основном с помощью двух методологий: на основе датчиков и на основе компьютерного зрения. Методология, основанная на датчиках, требует использования инструментов или датчиков, которые должны будут присоединены к руке/кисти пользователя для извлечения информации. В то время как распознавание, основанные на компьютерном зрении, требуют получения изображений или записей жестов рук с помощью видеокамеры.

Распознавание жестов руки является сложной задачей, требующей разработки эффективных алгоритмов и методов, способных точно и надежно распознавать жесты, выполненные пользователем. С развитием компьютерного зрения, машинного обучения и глубоких нейронных сетей, появилось множество новых подходов и алгоритмов.

Актуальность данной работы обусловлена растущим интересом и потребностью в разработке эффективных и точных алгоритмов распознавания жестов руки для управления курсором. Эта технология может быть особенно полезна для людей с физическими ограничениями, такими как моторные нарушения или ограниченные возможности движения, которые могут затруднять использования традиционных методов ввода. Вредные предприятия, где к примеру неудобно использовать традиционные способы ввода. Кроме того, распознавание жестов руки может быть интересно для широкого круга пользователей, включая людей, играющих на компьютере, разработчиков виртуальной и дополненной реальности, медицинских специалистов, исследователей в области человеко-компьютерного взаимодействия и других профессионалов.

Цель данной работы изучить и проанализировать существующие алгоритмы и методы распознавания жестов руки для управления курсором, а также оценить их эффективности.

Задачи:

1. Собрать, изучить и провести анализ научно-технической информации по теме исследования;
2. Рассмотреть основные понятия предметной области;
3. Исследовать различные алгоритмы и методы распознавания жестов руки;
4. Сравнить эффективность различных алгоритмов и методов распознавания жестов;
5. Определить, какие алгоритмы и методы распознавания жестов дают наилучшие результаты;
6. Оценить сложность реализации различных алгоритмов и методов распознавания по скорости, точности и простоте использования.

# **ГЛАВА 1. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ**

# **Методы распознавания и их классификации**

В последние годы значительные усилия были направлены на анализ движений тела и распознавания жестов [5]. С ростом интереса к взаимодействию человека и компьютера исследования, связанные с распознаванием жестов, стремительно развиваются. Наряду с речью, они являются очевидным выбором для естественного взаимодействия между человеком и компьютером. Человеческие жесты представляют собой распространённое и естественное средство невербальной коммуникации. Система распознавания жестов позволяет человеку вводить команды, используя естественные движения руки, головы и других частей тела (рис.1.1). А поскольку рука является наиболее широко используемой частью тела для жестикуляции, помимо лица. Жесты рук классифицируются как статические жесты или позы, и динамические или основанные на траектории жестов (рис. 1.). Динамические или основанные на траектории могут быть изолированными или непрерывными (рис. 1.1.).



Рис. 1.1. Классификация жестов

Рассматриваемая методология, основанная на использовании компьютерного зрения, подразумевает использование следующих инструментов или методов:

1. Одиночная камера – к ней относятся веб-камеры, различные типы видеокамеры и камеры смартфонов;
2. Стереокамеры и системы на основе нескольких камер. Пара стандартных цветных видео- или фотокамер захватывает несколько изображений для измерений глубины. Несколько монокулярных камер могут лучше передать трехмерную структуру объекта;
3. Методы светового кодирования – проекция света для получения 3D-структуры объекта. К таким устройствам относятся PrimeSense, Microsoft Kinect, Creative Senz-3D, Leap Motion и т.д;
4. Инвазивные методы – маркеры тела, такие как цвет руки, браслеты на запястье и маркеры пальцев. Однако термин "зрение" обычно используется для захвата изображений или видеозаписей голой руки без перчаток и/или маркеров. Подход на основе датчиков снижает необходимость в предварительной обработке и сегментации, которые необходимы в классических системы распознавания жестов на основе зрения.

# **Проблемы**

# **Проблемы сегментации**

Сегментация объектов на изображениях или видеозаписях остаётся одной из главных проблем компьютерного зрения, включая точное выделение руки или других жестикулирующих частей тела [5]. Ограничения, такие как изменение освещения, сложность фона и перекрытия, могут затруднить этот процесс. На рисунках 1.2 – 1.3. представлены проблемы сегментации изображения.



Рис. 1.2. Сложность фона



Рис. 1.3. Освещение

# **Проблемы определения жестов**

Определение жестов означает обнаружение от начала и до конца жеста в непрерывном потоке жестов [5]. Когда границы жестов разрешены, жест можно извлечь и сгруппировать. В любом случае, выявление значимых паттерном из потока жестов является исключительно сложной задачей, главных образом из-за двух проблем: неоднозначность сегментации и пространственно-временной изменчивости. Также необходимо удалить не жестовые движения. Случаи не жестовых движений включают в себя "эпентезис движения" и "жестовую коартикуляцию". Эпентезис движения происходит между двумя жестами, и на текущий жест влияет предыдущий или последующий. Коартикуляция жеста – нежелательное движение, возникающее в середине выполнения жеста. В некоторых случаях жест может быть похож на часть более длинного жеста, что называется "проблемой длинного жеста".

Проблема длинных жестов возникает, когда жест, который нужно распознать, является частью более длинного жеста. Это может сильно затруднить распознавание жеста, так как его можно неправильно интерпретировать в контексте более длинного жеста, в котором он находится. Например, жест, который означает "поднять руку", может быть частью более длинного жеста, означающего "поднять руку и помахать ей". В этом случае, если система распознавания жестов не учитывает контекст более длинного жеста, она может неправильно определить, какой жест был выполнен.

# **Проблемы, связанные с определением обеих рук**

Включение жестов с двумя руками в словарь жестов может сделать человеческо-компьютерный интерфейс более естественным и выразительным для пользователя. Это может значительном увеличить размер словарного запаса из-за различных комбинаций жестов левой и правой руки. Несмотря на свои преимущества, распознавание жестов с двумя руками сталкивается с некоторыми серьезными трудностями: вычислительная сложность, перекрытие между руками, одновременное отслеживание обеих рук.

## **Пример архитектуры с распознаванием жестов**

Одна из возможных архитектур приложения для распознавания жестов для управления курсором может включать следующие компоненты [5], использующие компьютерное зрение и глубокое обучение:

1. Захват изображения: приложение может использовать веб-камеру или другое устройство для захвата изображения жестов;
2. Предобработка изображения: входное изображение может быть подвергнуто обработке для уменьшения шума, улучшения контраста, изменения размера и других подобных преобразований;
3. Извлечение признаков: глубокие нейронные сети могут быть использованы для извлечения признаков из предварительно обработанного изображения. Например, сверточные нейронные сети могут быть использованы для обнаружения различных форм и текстур в изображении жестов;
4. Классификация жестов: полученные признаки могут быть поданы на вход другой нейронной сети, которая будет использоваться для классификации жестов. Эта нейронная сеть может быть обучена на большом наборе различных жестов с использованием методов обучения с учителем, таких как обратное распространение ошибки или градиентный спуск;
5. Управление курсором: после того, как жест был успешно классифицирован, приложение может использовать его для управления курсором на экране. Например, приложение может переводить движение жеста в определенное направление и скорость движения курсора.

# **Компьютерное зрение**

Для решения данной задачи предлагается использовать компьютерное зрение с целью обеспечения более эргономичного процесса взаимодействия между пользователем и компьютером [1]. Для достижения поставленной цели необходимо определиться с конкретной задачей, которую будет решать система компьютерного зрения. В данном случае рассматривается задача обнаружения и сегментации, которая позволит идентифицировать и различать фаланги и пальцы на руке пользователя.

Применение методов сегментации и обнаружения позволит решать данную задачу эффективнее, чем задача классификации. В процессе выполнения задачи сегментации будут сегментироваться объекты на изображении, в данном случае фаланги и пальцы на руке. Задача обнаружения позволит определить наличие или отсутствие интересующих объектов на изображении.

Однако, следует учитывать, что применение методов компьютерного зрения может потребовать значительных вычислительных ресурсов и времени. Поэтому необходимо выбрать оптимальную архитектуру нейронной сети и настроить ее гиперпараметры для обеспечения оптимальной скорости и точности обработки изображений [6].

Для управления курсором при помощи жестов руки, будет использоваться одиночная камера, так как она чаще всего встречается, например, веб-камера в ноутбуке, в телефоне и т.д. Также использование одиночной камеры, позволит обрабатывать только двухмерные изображения. Если бы было больше камер, тогда необходимо было бы синхронизировать изображение, потому что разные модели камер снимают по-разному и с разной скоростью. Также при использовании нескольких камер будет необходимо обучать камеру работать с трехмерным изображением, потому что искусственный интеллект не будет понимать похожие ли у него изображения, что дает также сложности при обучении, так как надо будет добавлять систему поиска изображений (Image retrieval) [8], то есть надо будет объяснить нейронной сети, что перед ней один и тот же объект.

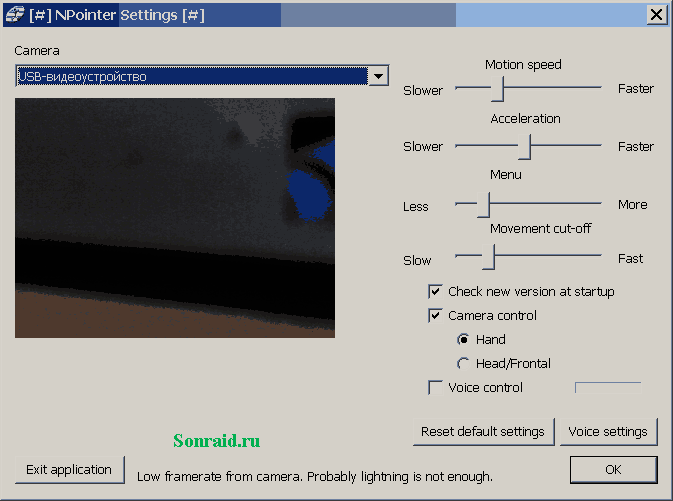
Еще одно из преимуществ использования двухмерных изображений, так это уже большое количество разработанных приложений в открытых источниках, использующие методы сегментации и обнаружения, что позволяет брать готовые примеры и улучшать существующие разработки. А также стоимость использования приложения, т.е. докупая дополнительные камеры для использования приложения, заставляет пользователя докупать дополнительное оборудование.

# **Анализ существующих решений**

В открытых источниках было найдено решение для управления курсором – Npointer 2.0 [6]. Кроме того, разработка компании Meta (запрещенная на территории РФ), в частности Meta Quest, включает в себя разработку системы управления интерфейсом с использованием жестов рук.

# **NPointer 2.0**

При установке указанного приложения пользователь сталкивается с появлением окна настроек, предоставляющего широкий спектр опций для настройки (рис. 1.4), что значительно расширяет его возможности во взаимодействии с приложением. Однако использование данного приложения сопровождается рядом проблем, включая сложности с использованием жестов и недостаточно эффективное отслеживание курсора. Например, приложение интерпретирует любое действие пользователя как перемещение курсора, что негативно сказывается на пользовательском опыте и снижает удовлетворенность использованием приложения.

  
Рис. 1.4. NPointer 2.0

Кроме того, отсутствует возможность отключения отслеживания жестами. Например, после завершения использования приложения требуется его полное выключение. Однако, если активировано управление курсором с помощью движений головы, возникают сложности с отключением, так как курсор продолжает перемещаться в стороны, усложняя процесс выхода из приложения.

# **Meta Quest**

Данная компания уже реализовала 3 версии своих решений. Данные продукты имеют встроенное отслеживание рук при взаимодействии с виртуальной и дополненной реальностью.

Для обеспечения отслеживания рук устройства Meta Quest используются различные технологии, включая встроенные датчики и камеры. На начальных этапах развития Oculus Quest была реализована технология отслеживания рук под названием Oculus Touch. Это контроллеры, которые пользователь держит в руках и которые отслеживаются в пространстве. Этот метод отслеживания позволяет пользователям взаимодействовать с виртуальным окружением, делая жесты, движения рук и даже действия, такие как захват и перемещение предметов.

Однако из-за текущего запрета на деятельность компании на территории Российской Федерации, а также в связи с наложенными санкциями, использование продуктов данной компании становится затруднительным. Для того чтобы воспользоваться их решениями, необходимо обновить драйвера и операционную систему, что требует доступа к серверам компании, что в настоящее время невозможно из-за указанных обстоятельств.

# **1.6. Глубокое обучение**

# **Применение глубокого обучения для компьютерного зрения**

Глубокое обучение необходимо использовать, когда одна компонента не несет в себе ценной информации. В данном случае изображение, которое мы можем передать на вход искусственному интеллекту, не дает полной картины, когда мы говорим только об одном пикселе с какой-то интенсивностью цвета [7].

Также проблема локальности, если поменять размер изображения, то может быть проблема с использованием алгоритмов машинного обучения [7]. У глубокого обучения уходит локальность [6].

Также глубокие нейронные сети могу решить проблему сложного фона, но минус решения данного выбора варианта — это значительные вычислительные способности, которые возрастают от использования нейронных сетей. На данный момент времени существуют видеокарты, которые позволяют благоприятно использовать нейронные сети. Однако, они находятся не в таком большом доступе из-за их высокой стоимости. С развитием технологий производства, стоимость видеокарт будет снижаться, а производительность будет расти, что позволит использовать такие нейронные сети в более обширном широком спектре приложений компьютерного зрения.

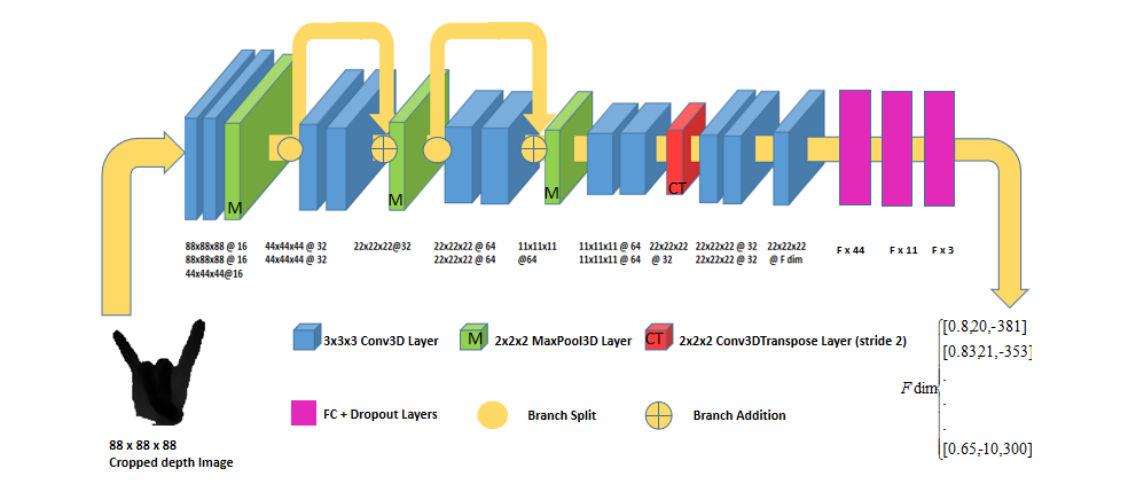
Для получения наиболее лучших результатов работы с компьютерным зрением и глубокими нейронными сетями, необходимо, выбрать правильную архитектуру нейронной сети, которая уже проверена экспертами по глубокому обучению.

Для решения данной задачи рассматривались следующие архитектуры [9]:

1. HandNet;
2. EfficientDet;
3. CenterNet.

# **HandNet**

HandNet – cверточная нейронная сеть, созданная специально для распознавания рук на изображении или видео в реальном времени [4]. На рисунке 1.4. представлена составляющие архитектуры HandNet



1.4. Архитектура HandNet

Можно выделить следующие преимущества алгоритма HandNet для задачи распознавания жестов рук:

1. Высокая точность обнаружения рук
2. Быстрое обучение
3. Доступность предобученных моделей, что упрощает начало работы над конкретной задачей

Однако, у алгоритма также имеются недостатки, среди которых:

1. Потребление большего количества ресурсов
2. Поэтому, для более эффективной работы с фоном изображений, может потребоваться:
   1. Дополнительная настройка и доработка архитектуры с использованием методов предобработки изображений, например, сегментации фона и удаления его изображения
   2. Добавление слоев, способствующих выделению идентифицирующих признаков жестов рук

# **EfficientDet**

EfficientDet применяется для обнаружения объектов. Архитектура показана на рисунке ниже. Состоит из EfficientNet в качестве основы, к которой приделан слой по работе с пирамидой признаков под названием BiFPN, за которым идет «стандартная» сеть вычисления класс/рамка объекта [3]. На рисунке 1.5. представлена составляющие архитектуры EfficientDet.

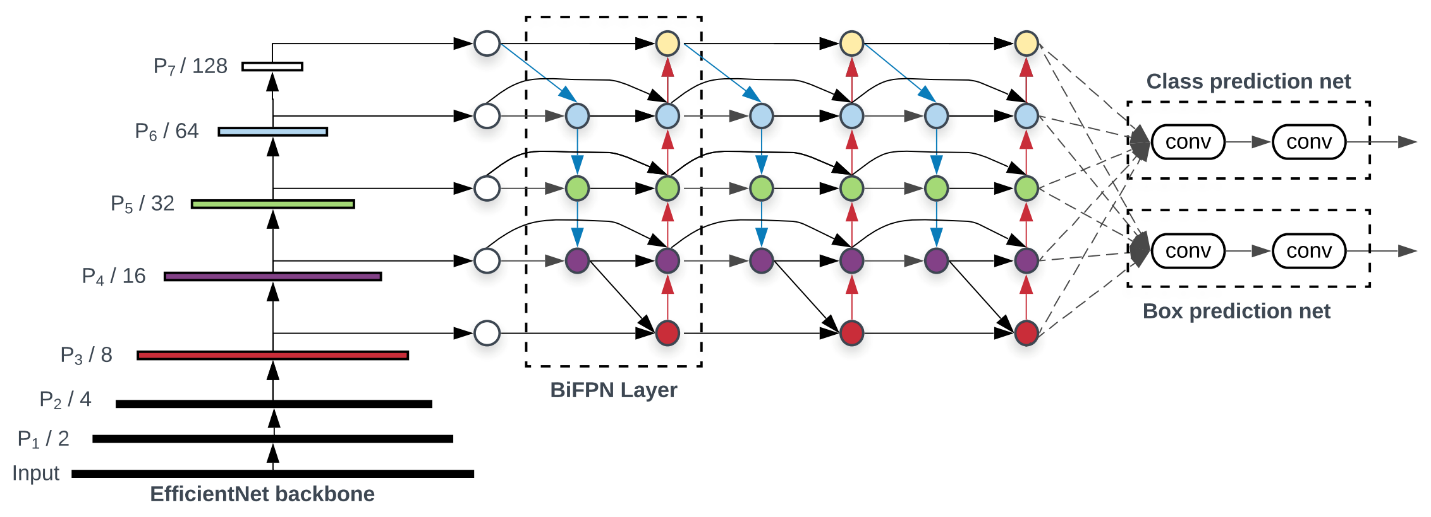


Рис. 1.5. Архитектура EfficientDet

Из преимуществ:

Высокая точность на различных датасетах

1. Относительно небольшое количество параметров, что позволяет использовать ее на устройствах с ограниченными ресурсами;
2. Масштабируемость, что позволяет использовать ее для обнаружения для объектов различных размеров;
3. Работает быстрее, чем многие архитектуры;

Из недостатков:

1. Ресурсы для обучения требуют вычислительно мощные GPU;
2. Возможно, требуются дополнительное обучение на некоторых датасетах для достижения максимальной точности.

# **CenterNet**

CenterNet и CornerNet-Lite считались на 2019 год мейнстримовыми легковесными системами обнаружения объектов в реальном времени [2]. На рисунке 1.6. представлена архитектура CenterNet.

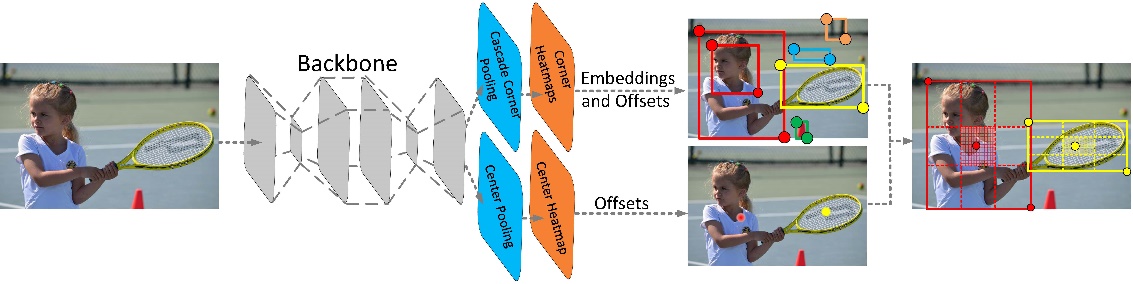


Рис. 1.6. Архитектура CenterNet

Из преимуществ:

1. Высокая точность;
2. Высокая скорость;
3. Масштабируемость.

Недостатки:

1. Ограничения по размеру объектов;
2. Требуется большое количество данных для обучения, чтобы достичь высокой точности;
3. Ресурсы для обучения требуют вычислительно мощные GPU.

# **Отличия между CenterNet и EfficientDet**

CenterNet является одним из методов обнаружения объектов, который основан на идее определять центры объектов и их ограничивающие прямоугольники в одном проходе (single-stage). То есть он одновременно предсказывает координаты центров объектов и соответствующие ограничивающие прямоугольники. Для этого он использует сверточную нейронную сеть с относительно простой архитектурой, что позволяет достигать высокой скорости обнаружения объектов. Однако, в силу относительно простой архитектуры, точность обнаружения объектов не всегда является наивысшей.

EfficientDet, с другой стороны, является более сложной архитектурой, которая основана на методе обучения с учителем, который использует автокодировщик для извлечения признаков изображений, а затем применяет эти признаки для обнаружения объектов. EfficientDet также основан на идее одного прохода и обеспечивает высокую скорость обнаружения объектов, но в то же время он может достигать более высокой точности, чем CenterNet.

# **Сравнительный анализ архитектур HandNet, CenterNet и EfficientDet для решения задачи обнаружения и сегментации рук в режиме реального времени.**

При выборе HandNet необходимо понимать, что данная архитектура была разработана специально для распознавания рук, что делает ее наиболее приоритетной для использования в связи с разработкой приложения для распознавания рук.

Если рассматривать EfficientDet и CenterNet для своей реализации HandNet, тогда необходимо будет также модифицировать архитектуру для более точных и быстрых задач.

В итоге выбор падает на архитектуру HandNet, потому что тогда задача сводится к тому, чтобы сосредоточится на разработке функционала, а именно разработка управления компьютером при помощи специальных жестов.

# **Анализ правовых и нормативно-технических документов**

# **ГОСТ Р ИСО/МЭК 25000-2021**

Данный стандарт, руководство по использованию серии стандартов SQuaRE (Требования и оценка качества систем и программных средств), предоставляет обзор содержания серии стандартов и устанавливает общие эталонные модели и определения. Цель этого руководства — обеспечить понимание положений серии стандартов SQuaRE и их применение в контексте разработки приложения для управления курсором [10].

Основной акцент делается на важности высокого качества систем и программного обеспечения в контексте широкого спектра областей применения, включая управление курсором. В разделе по качественным характеристикам освещаются особенности, которые могут быть критически важными для приложений данного типа.

Вместо использования устаревших стандартов серий ИСО/МЭК 9126 и ИСО/МЭК 14598, предлагается переходить к серии стандартов SQuaRE для более современного и всестороннего подхода к разработке и оценке качества программного обеспечения.

Это руководство может служить основой для определения соответствующих качественных характеристик при разработке приложения для управления курсором и обеспечения их оценки с использованием аттестованных или широко принятых показателей. Такой подход позволяет удовлетворять требованиям бизнеса и обеспечивать безопасность в использовании данного приложения.

# **ГОСТ Р ИСО/МЭК 60447-2015**

Использование камеры в качестве средства ввода данных открывает возможности для разработки приложения, которое реагирует на жесты и движения рук пользователя. Это поднимает важные вопросы безопасности и эффективности взаимодействия, особенно в случаях, когда управление может повлиять на работу оборудования или требует быстрого реагирования.

Стандартизация в данной области становится ключевым фактором, обеспечивая безопасность и удобство использования. Особенно важно это в ситуациях, где неправильное управление может привести к повреждению оборудования. Нормативы также актуальны в контексте возможного использования приложения неопытным персоналом.

Настоящий стандарт определяет общие принципы приведения в действие исполнительных механизмов с ручным управлением, включая их часть в человеко-машинном интерфейсе [11]. Эти принципы направлены на повышение безопасности и обеспечение удобства использования, а также на эффективное включение исполнительных механизмов в работу.

Стандарт применим в различных сценариях — от использования простых органов управления до сложных систем автоматизации. Он также устанавливает связи между функцией органа управления и его действием, что важно для разработки приложения управления курсором с использованием жестовых взаимодействий.

# **ГОСТ Р ИСО/МЭК 9241-10**

Стандарт ИСО 9241-10, также известный как "Эргономика взаимодействия человек - система", устанавливает принципы диалога между человеком и компьютерной системой. Эти принципы представляют собой основу для разработки пользовательских интерфейсов с целью обеспечения эффективности, удобства и удовлетворения пользователя [11].

Принципы диалога:

1. Пригодность для задания: Интерфейс должен быть спроектирован таким образом, чтобы пользователь мог легко выполнять свои задачи. Это включает в себя понятность структуры системы и доступность функций.
2. Само документированность: Интерфейс должен быть интуитивно понятен, чтобы пользователь мог понимать, как использовать его без необходимости обращения к дополнительным источникам информации.
3. Контролируемость: Пользователь должен иметь контроль над процессом взаимодействия. Это включает в себя возможность начала, приостановки, продолжения и завершения задач в любой момент.
4. Соответствие ожиданиям пользователя: Интерфейс должен соответствовать ожиданиям пользователей, основанным на их предыдущем опыте и знаниях.
5. Устойчивость по отношению к ошибкам: Система должна быть спроектирована так, чтобы минимизировать возможность ошибок пользователя, а также предоставлять средства для их обнаружения и исправления.
6. Пригодность для индивидуализации: Интерфейс должен предоставлять возможности для настройки под индивидуальные предпочтения пользователей.
7. Пригодность для обучения: Система должна быть легко обучаемой новым пользователям, и обучение должно быть эффективным и доступным.

Эти принципы являются общими руководящими принципами, которые могут применяться к различным типам интерфейсов, включая интерфейсы для управления курсором при помощи жестов рук. Разработчики приложений могут использовать эти принципы в процессе проектирования, чтобы создать пользовательский интерфейс, который будет интуитивно понятен, удобен и эффективен для широкого круга пользователей.

# **ГОСТ Р ИСО/МЭК 70462.1-2022**

При разработке приложения для управления курсором при помощи жестов рук, основанного на технологиях искусственного интеллекта (ИИ), важно учесть ряд свойств и принципов, включая робастность, отказоустойчивость, надежность, точность, безопасность и другие. Эти свойства играют важную роль, особенно когда применяются нейронные сети, широко используемые в сфере промышленности [13].

Принципы и характеристики, определенные в стандартах, таких как ИСО 9241-10, могут быть дополнены рекомендациями и требованиями, связанными с ИИ и нейронными сетями. Важным аспектом является робастность системы управления курсором в различных сценариях использования, а также ее способность адекватно реагировать на жесты рук пользователя.

Для обеспечения высокого качества и эффективности приложения рекомендуется:

1. Робастность: Обеспечение стабильной работы системы в различных условиях, включая разнообразные жесты рук, освещение и фон.
2. Точность: гарантировать высокую точность распознавания жестов и управления курсором для минимизации ошибок.
3. Безопасность: защитить данные пользователя и обеспечить безопасность взаимодействия с системой, особенно при использовании биометрических данных.
4. Производительность: обеспечить высокую производительность системы, реагируя на жесты пользователя с минимальной задержкой.
5. Тестирование и валидация: использовать подходы к тестированию и валидации, учитывающие специфику работы с нейронными сетями, включая статистические, формальные и эмпирические методы.
6. Внедрение стандартов: придерживаться принципов и стандартов, определенных в области разработки ИИ, чтобы обеспечить соответствие требованиям безопасности и надежности.
7. Документация: Создание подробной документации по работе системы и обоснование решений, принятых в процессе разработки.

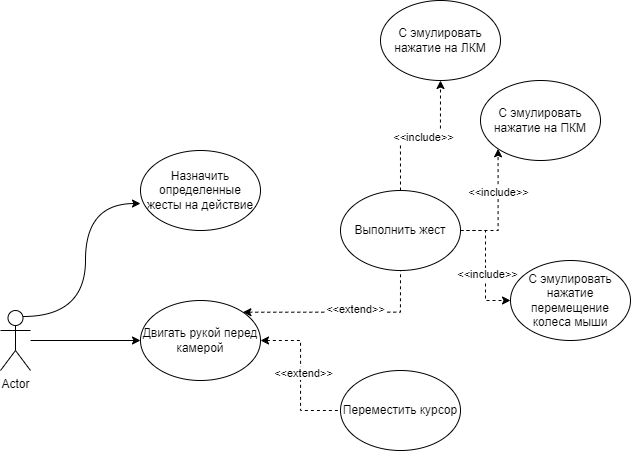
Такой подход позволит создать приложение для управления курсором при помощи жестов рук, которое соответствует высоким стандартам качества, безопасности и эффективности.

# **ГЛАВА 2. МОДЕЛИРОВАНИЕ ПРОЦЕССОВ**

# **Требования к разрабатываемой системе**

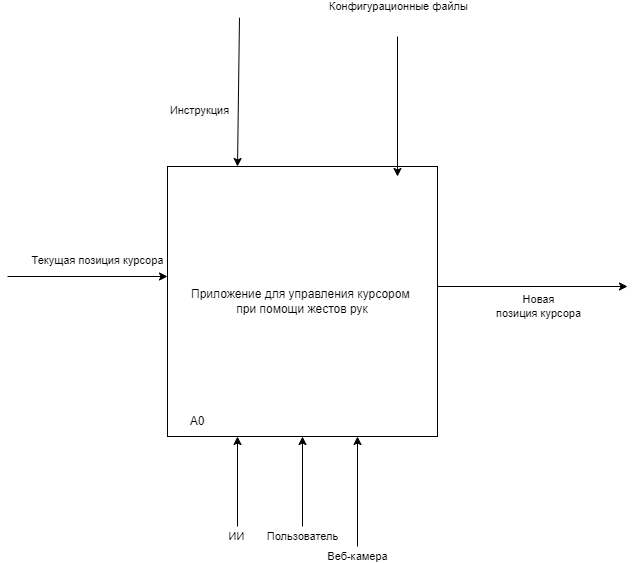
# **Функциональные требования**

На рисунке 2.1. представлены диаграмма использования программного продукта.

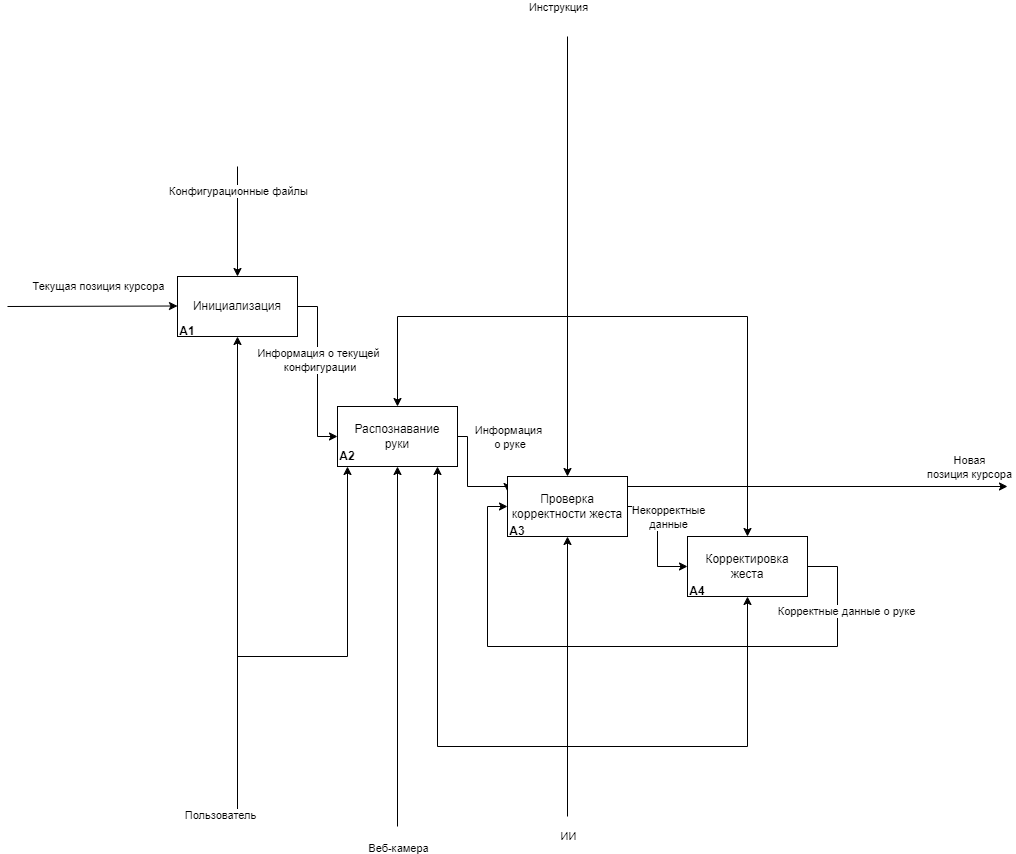


На рисунке 2.2 – 2.7 представлена диаграмма IDEF0 для визуализации взаимосвязи между различными функциями системы и их взаимодействием.

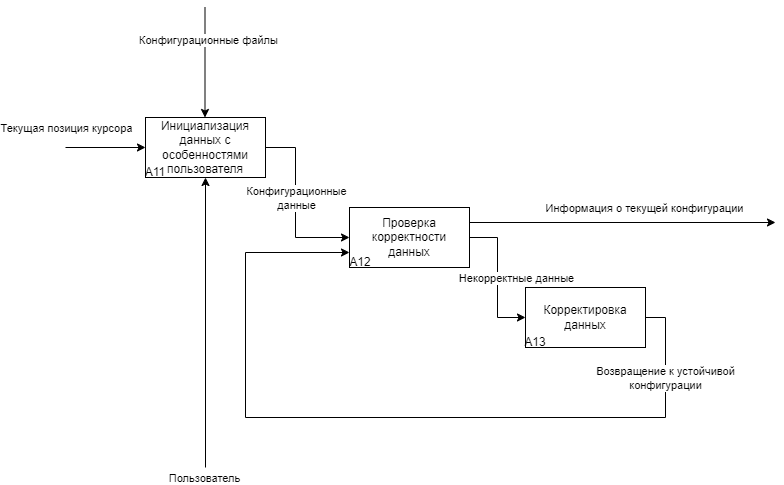
На рисунке 2.2. представлена система в виде черного ящика, у которого есть вход (текущая позиция курсора) и выход (новая позиция курсора).

  
Рис. 2.2. IDEF0 уровень A0

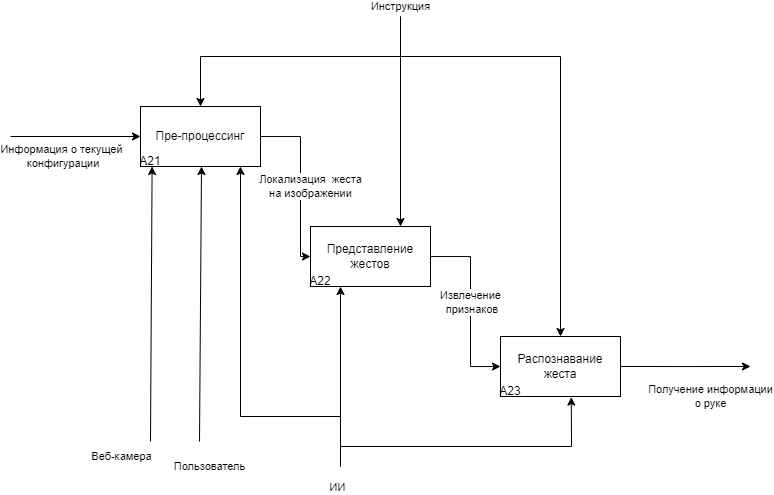
На рисунке 2.3. представлены действия системы, которые необходимо сделать для получения новой позиции курсора, такие как: Инициализация, Распознавания руки, Проверка корректности жеста, а также Корректировка жеста, при возникновении ошибки при проверке правильность жеста.

  
Рис. 2.3. IDEF0 диаграмма

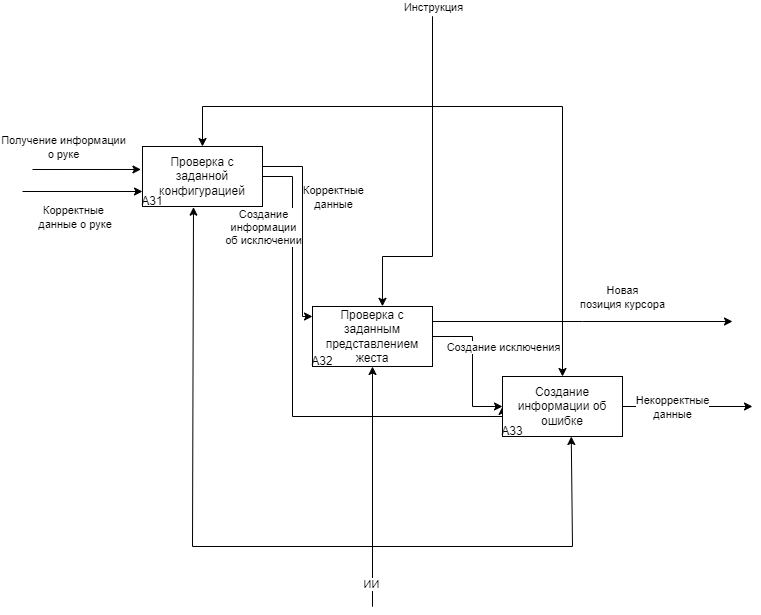
На рисунке 2.4. представлен процесс инициализации системы при начале работы системы. Пользователь может настроить систему, а также система сверяется с текущими данными, чтобы при возникновении ошибки, система не завершалась принудительно.

  
Рис. 2.4. IDEF0 диаграмма

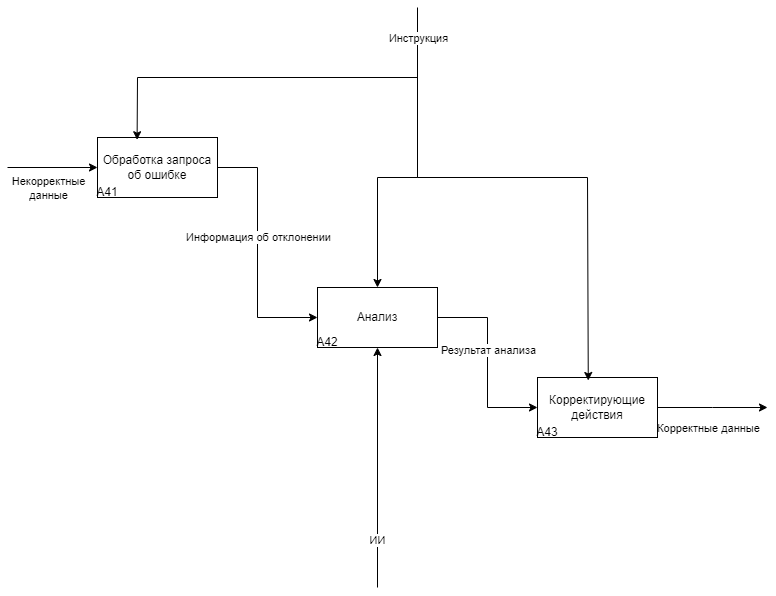
На рисунке 2.5. представлен процесс распознавания руки в данный процесс входит: пре-процессинг, представление жестов, а также распознавание жеста.

  
Рис. 2.5. IDEF0 диаграмма

На рисунке 2.6. можно увидеть проверку на правильность введенного жеста, так как система должна сравнить с тем, что имеется в системе, чтобы не создать неприятного опыта использования от приложения.

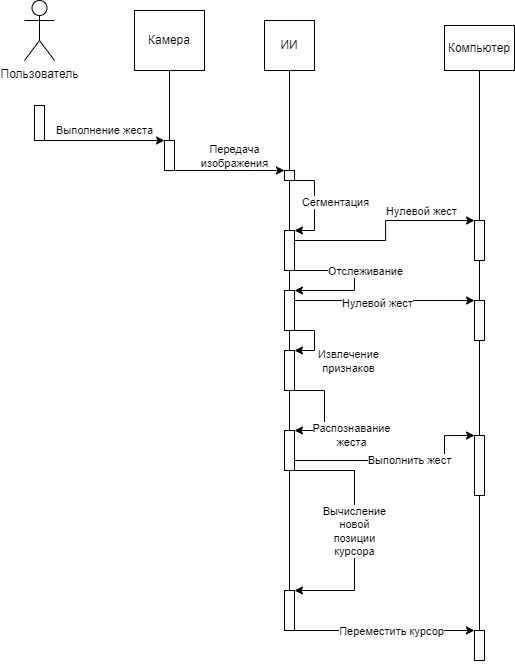
  
Рис. 2.6. IDEF0 диаграмма

На рисунке 2.7. представлена IDEF0 диаграмма, которая описывает процесс обработки неправильного жеста, так как пользователь может отвлечься или передумает использовать данный жест, и чтобы не возникло проблем, для этого будут применены корректирующие действия.

  
Рис. 2.7. IDEF0 диаграмма

# **Диаграмма последовательности**

На рисунке 2.8. представлена диаграмма последовательности, в которой можно увидеть взаимодействие пользователя с разработанным приложением в рамках определенных сценариев использования. В данной диаграмме участвует один пользователь и 3 объекта, такие как Камера, которая будет снимать действия пользователя, Искусственный интеллект (ИИ), а также компьютер, с которым взаимодействует пользователь.

  
Рис. 2.8. Диаграмма последовательности

При возникновении ошибок при таких процессах ИИ, как сегментация, отслеживание должен создаваться нулевой жест, чтобы программа не передвигала курсор или выполняла действия, которые не подразумевал пользователь.

# **ЗАКЛЮЧЕНИЕ**

Целью данной работы было исследование, а также моделирование процессов распознавания жестов рук для последующей разработки приложения по управлению курсором при помощи рук.

По итогу была совершена следующая работа:

* 1. Рассмотрена предметная область поставленной темы, углубление в способы распознавания, необходимые для исследования, проблемы, связанные с распознаванием, а также рассмотрен пример архитектуры для распознавания из чего он состоит, а также существующие разработки;
  2. Рассмотрены причины почему необходимо использовать именно компьютерное зрение, а не другие способы регистрации рук. А также почему было выбрано использование одной веб-камеры;
  3. Были рассмотрены существующие архитектуры нейронных сетей. Были рассмотрены их преимущества и недостатки по сравнению с другими. А также была выбрана архитектура, на которой будет строится будущее решение распознавания жестов рук.
  4. Были смоделированы процессы работы будущей разработки, выделены функциональные требования к системе.

Применение распознавания жестов рук для управления курсором имеет важную ценность в развитие интерфейсов между компьютером и человеком. Такие приложения позволят многим людям с ограниченными возможностями или без отклонений людей, более благоприятно использовать электронные устройства с монитором или экраном.

Однако в современном мире до сих пор не были решены некоторые проблемы для распознавания жестов рук, поэтому такие системы пока не так широко используются в повседневном мире. Также еще может быть проблемы с требовательностью ресурсов. Так как обучение нейронной сети или использования распознавания в реальном времени, к примеру 60 кадров в секунду (к примеру, 60 Гц монитор), некоторые устройства не могут обеспечить нужной производительности.

Поэтому создав решение распознавания жестов, которое будет решать одну из проблем распознавания, сделает такую систему более востребованной, а также может привести к интересу данной сферы среди других разработчиков или ученных, что в будущем могут создать более совершенную и практичную систему, что даст человеку более приятный способ взаимодействия между устройствами.

# **Список литературы**

1. Davis, Roy and Terk, Matthew. Advanced Methods and Deep Learning in Computer Vision, 2021.
2. CenterNet - https://github.com/Duankaiwen/CenterNet. Режим доступа: свободный. Дата обращения: 28.04.2023
3. EfficientDet - https://github.com/google/automl/tree/master/efficientdet
4. FastV2C-HandNet: Fast Voxel to Coordinate Hand Pose Estimation with 3D Convolutional Neural Networks [Электронный ресурс] - https://arxiv.org/pdf/1907.06327.pdf. Режим доступа: свободный. Дата обращения:30.04.2023
5. Methods, Databases and Recent Advancement of Vision-Based Hand Gesture Recognition for HCI Systems: A Review. [Электронный ресурс] // Springer. URL: https://link.springer.com/article/10.1007/s42979-021-00827-x. Режим доступа: свободный. Дата обращения: 03.03.2023
6. NPointer 2.0 - http://sonraid.ru/npointer/ - Режим доступа: свободный. Дата обращения 23.09.2023
7. Курс лекций по глубокому обучению (весна, 2022). [Электронный ресурс] URL: https://www.youtube.com/watch?v=LHlM2sf1YpU&list=PL6-BrcpR2C5Szpo6RS370Ck1\_-emlwlWc. Режим доступа: свободный. Дата обращения:15.03.2023
8. Курс лекций по компьютерному зрению с использованием DL (весна, 2022). [Электронный ресурс] URL: https://www.youtube.com/watch?v=W\_cxT4uSsac&list=PL6-BrcpR2C5T03IRUj7X\_qVbwjJHwyO-W. Режим доступа: свободный. Дата обращения: 30.03.2023
9. Новые архитектуры нейронных сетей. [Электронный ресурс] URL: https://habr.com/ru/articles/498168/. Режим доступа: свободный. Дата обращения:16.04.2023
10. Стандарт ГОСТ Р ИСО/МЭК 25000-2021. [Электронный ресурс] URL: https://docs.cntd.ru/document/1200181361. Режим доступа: свободный. Дата обращения: 12.12.2023
11. Стандарт ГОСТ Р ИСО/МЭК 60447-2015. [Электронный ресурс] URL: https://allgosts.ru/29/020/gost\_iec\_60447-2015. Режим доступа: свободный. Дата обращения: 12.12.2023
12. Стандарт ГОСТ Р ИСО/МЭК 9241-10. [Электронный ресурс] URL: http://www.zakonprost.ru/content/base/part/642508. Режим доступа: свободный. Дата обращения: 12.12.2023
13. Стандарт ГОСТ Р 70462.1-2022. [Электронный ресурс] URL: https://allgosts.ru/35/020/gost\_r\_70462.1-2022. Режим доступа: свободный. Дата обращения: 12.12.2023