|  |  |
| --- | --- |
|  | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ **Информатика и системы управления**

КАФЕДРА **Компьютерные системы и сети (ИУ6)**

НАПРАВЛЕНИЕ ПОДГОТОВКИ **09.04.01 Информатика и вычислительная техника**

МАГИСТЕРСКАЯ ПРОГРАММА **09.04.01/07 Интеллектуальные системы анализа, обработки и интерпретации больших данных**

**Отчет**

|  |  |
| --- | --- |
| **по рубежному контролю №** | 2 |

**Название:** Проектирование системы определения дипфейков

**Дисциплина:** Искусство системного инжиниринга и менеджмента организаций

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Студент | ИУ6-12М |  |  | С.В. Астахов | |
|  | (Группа) |  | (Подпись, дата) | | (И.О. Фамилия) |
|  |  |  |  | |  |
| Преподаватель |  |  |  | | Д.В. Березкин |
|  |  |  | (Подпись, дата) | | (И.О. Фамилия) |

Москва, 2023

**СОДЕРЖАНИЕ**

[Введение 3](#_Toc3)

[1. Методы на основе сверточных и капсульных нейросетей 4](#_Toc4)

[2. Методы на основе временной согласованности 7](#_Toc5)

[2.1. Комбинация сверточной и рекуррентной нейронных сетей 7](#_Toc6)

[2.2. Улучшенные системы анализа временной согласованности 10](#_Toc7)

[3. Методы на основе визуальных артефактов 13](#_Toc8)

[Заключение 15](#_Toc9)

# Введение

В наше время социальные сети и мессенджеры стали неотъемлемой частью жизни многих людей. Однако, вместе с возможностью общения и обмена информацией, появилась и проблема фейковых новостей и фотографий.

Дипфейки - это фотографии или видео, созданные с помощью искусственного интеллекта, которые могут быть использованы для создания фальшивых новостей или дезинформации.

В связи с этим, возникает необходимость в разработке методов определения дипфейков, которые позволят бороться с распространением фальшивой информации. В данном реферате рассмотрены основные методы определения дипфейков.

# **Х. Определение зон интереса на изображении (/подготовка данных)**

Исходя из разработанного сценария обработки изображение, очевидно, что первым этапом на пути обработки изображения является выделение зон интереса. На основе характеристик изображения в этих зонах и будет производиться дальнейшее определение признаков дипфейка.

Для определения зон интереса подпрограмма предварительной обработки данных определяет положение 68 типовых точек лица с помощью библиотеки dlib [1].

Для определения положения точек библиотека использует каскад регрессионных моделей.

Прежде всего, использование каскадов решает такую проблему, как влияние освещения и артефактов на точность определения точек. Изначальная дилемма состоит в том, что при определении признаков регрессорами, признаки могут быть искажены. Для борьбы с этим используется каскадная модель, которая итеративно уточняет форму лица и значения признаков в ее рамках.

Второй проблемой является тот факт, что алгоритм должен оценить форму, вектор высокой размерности, который наилучшим образом согласуется с данными изображения и нашей моделью формы. Успешные алгоритмы решают эту проблему, предполагая, что предполагаемая форма должна лежать в линейном подпространстве, которое может быть обнаружено, например, путем нахождения основных компонентов форм [2].

# **Заключение**

Проблема фейковых новостей, видео и фотографий становится все более актуальной в нашем информационном обществе. Дипфейки могут нанести серьезный ущерб как отдельным людям, так и обществу в целом.

Однако, существуют различные методы определения дипфейков, которые позволяют бороться с распространением фальшивой информации.

В данном реферате были рассмотрены основные методы определения дипфейков, такие как анализ пикселей, анализ движения, анализ голоса и другие.

Каждый из этих методов имеет свои преимущества и недостатки, и их эффективность зависит от конкретной ситуации. В целом, разработка и совершенствование методов определения дипфейков является важной задачей для обеспечения достоверности информации в нашем информационном обществе.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. <http://dlib.net/train_shape_predictor.py.html>
2. https://www.cv-foundation.org/openaccess/content\_cvpr\_2014/papers/Kazemi\_One\_Millisecond\_Face\_2014\_CVPR\_paper.pdf