ДИПФЕЙК-ВИДЕО: УГРОЗЫ, ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ ПРАВОВОГО РЕГУЛИРОВАНИЯ

Ляшенко Сергей Андреевич

студент, Иркутский национальный исследовательский технический университет, РФ, г. Иркутск

Аннотация. Современные технологии дипфейк-видео представляют собой серьезный вызов для общества. Они открывают возможности для создания фальсифицированных видеоматериалов, которые трудно отличить от оригинала, что делает их мощным инструментом манипуляции. В данной статье рассматриваются принципы работы видео-дипфейков, их применение в криминальной сфере, угрозы, а также правовые аспекты регулирования данной технологии.

Ключевые слова: дипфейк-видео, искусственный интеллект, манипуляция, правовое регулирование, цифровые технологии, киберпреступность.

Введение.

Современные технологии искусственного интеллекта (ИИ) стремительно развиваются, затрагивая медиа, науку, бизнес, безопасность и политику. Одним из самых спорных достижений ИИ стали дипфейк-видео — поддельные видеоматериалы, созданные с помощью нейронных сетей. Они позволяют менять лица, имитировать речь и эмоции, делая фальсификацию видеозаписей практически неотличимой от оригинала.

Изначально дипфейк-видео использовались в киноиндустрии, рекламе и образовании. Однако со временем они стали инструментом манипуляций, фейковых новостей, мошенничества и политических провокаций. Сегодня для их создания не нужны сложные технологии — доступные программы позволяют любому человеку сгенерировать правдоподобный дипфейк.

Проблема в том, что методы выявления дипфейк-видео пока отстают от технологий их создания, а законодательство не успевает за стремительным развитием этой сферы. В данной статье рассматриваются принципы работы дипфейк-видео, риски их использования и возможные пути правового регулирования.

Как работают дипфейк-видео.

Технология создания дипфейк-видео опирается на современные методы глубокого обучения, позволяющие синтезировать поддельные, но по качеству неотличимые от реальных, видеоматериалы [2]. Процесс можно разделить на несколько ключевых этапов:

- **Сбор и подготовка данных.** На первом этапе осуществляется сбор большого объема видеоматериалов и изображений, содержащих лица целевых объектов. Затем с помощью алгоритмов детекции извлекаются отдельные кадры и вырезаются лица. Для обеспечения единообразия изображения подвергаются нормализации: корректируется размер, яркость и контраст.
- **Обучение нейронной сети.** Автоэнкодеры сжимают входное изображение до более

компактного представления, а затем восстанавливают его, обучаясь запоминать характерные особенности лица. Генеративно-состязательные сети состоят из двух компонентов: генератора, создающего реалистичные изображения, и дискриминатора, который обучается отличать реальные кадры от синтезированных
- **Синтез изображения и адаптация к видео.** После обучения модель способна генерировать изображение лица с нужными параметрами – с учетом эмоций, ракурса и освещения. Полученное синтетическое изображение адаптируется к исходному видеоряду, синхронизируясь с движением, мимикой и динамикой оригинала.
- **Наложение и постобработка.** На завершающем этапе сгенерированное лицо интегрируется в исходное видео. Применяются методы коррекции цвета, сглаживания границ и устранения артефактов, чтобы итоговый видеоряд выглядел гармонично.

Угрозы и негативные последствия.

Несмотря на полезные применения в творчестве и образовании, дипфейк-видео несут серьезные риски для общества, репутации и экономики:

Преступное использование.

Дипфейки создают ложные видео, компрометирующие публичных лиц и сотрудников компаний, что используется для шантажа, манипуляций и вымогательства. Фальсификация способна подорвать доверие к судопроизводству и привести к ошибочным обвинениям [4].

Социально-политические последствия.

Дипфейки служат инструментом фейковых новостей, искажая политическую реальность и провоцируя волнения. Их массовое распространение снижает доверие к СМИ и усугубляет кризис информационной безопасности.

Нарушение личной репутации и безопасности.

Фальсифицированные видео могут разрушить репутацию, карьеру и личную жизнь жертв. Использование дипфейков для запугивания приводит к угрозам физической и моральной безопасности, вызывая общественное осуждение и стресс.

Экономические и корпоративные риски.

Ложные видеозаписи могут подорвать доверие к компаниям, повлиять на их финансовые показатели и вынудить обращаться в суд. Отсутствие четких правовых норм усложняет защиту репутации и ведет к дополнительным расходам.

Правовые и этические дилеммы.

Понятие дипфейков не имеет четкого правового определения, что затрудняет их регулирование. Использование технологии поднимает вопросы морали, ответственности и защиты частной жизни [3].

Психологическое воздействие и социальная нестабильность.

Дипфейки подрывают доверие к визуальной информации, создавая атмосферу неуверенности. В кризисные периоды они могут усиливать тревогу, панику и социальное напряжение.

Современное правовое регулирование и возможные решения.

В современной России отсутствуют конкретные нормы для регулирования дипфейков, однако ситуация активно обсуждается, и ведутся работы над законопроектами [1].

Разработка законопроектов:

- Защита изображения и голоса: рассматриваются поправки, которые бы запрещали

использование изображений или голосов граждан без их согласия. Законопроект №718834-8 предусматривает охрану голоса как нематериального права.
- Уголовная ответственность: Другой законопроект предлагает внести изменения в УК РФ, усиливая наказание за распространение дипфейков, наносящих ущерб репутации или использующих биометрические данные для мошенничества.

Технологическая инфраструктура:

- Разработка инструментов обнаружения: необходимо интегрировать в систему безопасности России инструменты для оперативного выявления дипфейков. Такие технологии могут быть внедрены как в правоохранительные органы, так и в частный сектор.
- Информирование и образование: Повышение уровня медийной грамотности граждан, чтобы они могли самостоятельно распознавать дипфейки, а также обучение специалистов по цифровой безопасности является ключевым направлением работы.

Заключение.

Дипфейк-видео представляют серьезные риски для информационной безопасности, правового регулирования и общественного доверия. Исследование показало, что технологии детекции пока уступают методам их создания, что требует дальнейших разработок в области ИИ. Опыт демонстрирует, что эффективная борьба с дипфейками возможна при комплексном подходе, включающем законодательные инициативы, технологические решения и повышение цифровой грамотности. В России правовая база в этой сфере только формируется, и ее развитие должно сопровождаться внедрением механизмов обнаружения дипфейков и обучением специалистов. Дальнейшие исследования должны быть направлены на совершенствование алгоритмов детекции, защиту цифровой идентичности и создание правовых норм, обеспечивающих баланс между технологическим прогрессом и информационной безопасностью.

Список литературы:

1. Правовые проблемы использования технологии Deepfake [Электронный ресурс]. – Режим доступа: https://zakon.ru/blog (дата обращения: 06.03.2025).

2. Обзор технологий создания Deepfake и методов его выявления [Электронный ресурс]. – Режим доступа: https://rdc.grfc.ru/2020/06/research-deepfake/ (дата обращения: 08.03.2025).

3. Законопроект об охране голоса [Электронный ресурс]. – Режим доступа: https://sozd.duma.gov.ru/bill/718834-8 (дата обращения: 09.03.2025).

4. Цифровой обман: какие права нарушают дипфейки и как за это наказывают [Электронный ресурс]. – Режим доступа: https://pravo.ru/story/249733 (дата обращения: 09.03.2025).