УДК 004

Бурцева Татьяна Алексеевна, профессор кафедры менеджмента Московский финансово-юридический университет МФЮА, г. Москва

Юсупова Анастасия Владимировна, магистрант, Московский финансово-юридический университет МФЮА, г. Москва

DEEPFAKE DETECTION ОТ VISIONLABS КАК ЭЛЕМЕНТ СИСТЕМЫ ПРОТИВОДЕЙСТВИЯ БИОМЕТРИЧЕСКОМУ МОШЕННИЧЕСТВУ

Аннотация

В настоящее время борьба с дипфейками становится наиболее актуальной. По прогнозам аналитиков, с каждым годом возрастает возможность кибератак с помощью дипфейков, так как большинство материалов находится в открытом доступе в Интернете. Мошенники и преступники трансформируют техники вымогательства денежных средств у предполагаемых жертв. Вследствие этого, растет спрос и потребность своевременно предотвращать потенциальные преступления. Для этого используют различные технологии, в том числе применяют и ИТ-решения. В работе представлен анализ технологии Deepfake Detection. Решение от VisionLabs находит своё применение в различных отраслях: безопасный город, финансы и промышленность. Основной функцией продукта является определение реального человека, а не дипфейка: система сможет идентифицировать, сгенерированное ли изображение перед камерой. В статье показана роль влияния дипфейков на общество, выявлены негативные последствия и будущее развитие технологии.

Annotation

Currently, the fight against deepfakes is becoming the most urgent. Analysts predict that the possibility of cyber attacks using deepfakes is increasing every year, as most of the materials are publicly available on the Internet. Fraudsters and criminals are transforming the techniques of extorting money from alleged victims. As a result, there is a growing demand and need to prevent potential crimes in a timely manner. Various technologies are used for this, including IT solutions.

The paper presents an analysis of Deepfake Detection technology. The VisionLabs solution finds its application in various industries: safe city, finance and industry. The main function of the product is to identify a real person, not a deepfake: the system will be able to identify whether the generated image is in front of the camera. The article shows the role of the influence of deepfakes on society, identifies the negative consequences and the future development of technology.

Ключевые слова: дипфейки, цифровые технологии, нейросеть, цифровизация, ИТ-решение, система.

Keywords: Deepfakes, digital technologies, neural network, digitalization, IT solution, system.

С развитием глобальной сети Интернет, киберпреступность становится приобретает обширное явление. Злоумышленники используют различные ресурсы для получения информации о жертве. Как правило основным источником сбора сведений становятся социальные сети. На основе полученных фотографий, видеоматериалов и аудиоматериалов, появляется возможность сгенерировать дипфейк с помощью Искусственного Интеллекта с участием реального человека. Под угрозу создания дипфейков попадают не только открытые аккаунты, доступные в сети, но и закрытые аккаунты, которые содержат в себе информацию о

личности человека. Поэтому подвержены риску большинство людей, использующих социальные сети и выкладывающие различный контент.

Угроза получения дипфейка присутствует не только у пользователей, но и у родственников и друзей жертвы. Они становятся основной целью для мошеннических схем. Представляясь личностью другого человека, злоумышленники создают фейковые аккаунты. В основном, преступники просят о материальной помощи. Возможны аудио-атаки с использованием настроенного сгенерированного голоса, также видеоматериалы, используемые как доказательства возникнувшей проблемы. Если вовремя не обратить внимание на то, что аккаунт пользователя не является настоящим, то достаточно легко поверить в то, что кто-то из твоего близкого круга находится в трудном положении.

Другой вид мошенничества затрагивает финансовую отрасль, в частности банковский сектор. Множество денежных операций проходят через онлайн-платформы. На сегодняшний день все банки используют цифровую валюту. Удобство использования банковских приложений, где возможно провести любую операцию, связанную с сбережениями, становится реальностью для пользователей по всему миру. Банки могут проводить консультации онлайн, выдавать кредиты и заключать сделки. Несмотря на современные методы защиты, мошенники с каждым годом находят новые возможности, как пройти банковскую верификацию для того, чтобы завладеть финансами жертвы. Подобная практика становится основой преступлений финансового характера. Принцип мошенничества остается аналогичным: создается дипфейк с помощью нейросетей, где генерируется фотографии, видео и аудиоматериалы, на основе собранных фотографий, видео и записей голоса. Далее, представляясь клиентом банка или сотрудником финансовой организации, «цифровая копия» проводит попытку операций со счетами для получения сбережений.

Данная преступная схема наиболее опасна, так как человек не сможет определить сгенерированное ли перед ним изображение или реальный человек. Это представляет угрозу не только для пользователя, но и для безопасности банка. Для предотвращения мошенничества, используют многофакторную аутентификацию, где задействована технология проверки по биометрии. Deepfake Detection как продукт предлагает решение для банков, которое позволяет с высокой точностью определить подмену личности.

Татьяна Дешкина, заместитель директора по продуктам VisionLabs, отмечает, что преступники используют мошеннические схемы при видеозвонке от банка - заменяют видеопоток с фронтальной камеры дипфейком - и подобные сценарии применяют для подтверждения перевода выше определенной суммы. Если за проверку личности отвечает человек, он с высокой вероятностью не распознает подмену. Показательным примером является видео с реальной и синтезированной девушкой, только 40–60% аудитории правильно определяют подмену. Таким образом, проблема использования дипфейков остро стоит в различных сферах общества. VisionLabs стала первой компанией, которая создала систему распознавания дипфейков.

Генеральный директор VisionLabs Дмитрий Марков подчеркивает, что в связи с развитием технологии Deepfake появилось очень много мошенников, которые атакуют с фальшивым лицом. Со своей стороны, VisionLabs представило эту технологию, не имея средств и специальных разработок. Важность создания продукта Deepfake Detection была отмечена после его презентации – высокая точность определения дипфейков, с помощью продуктов от VisionLabs: Luna Pass и Luna Platform. Встроенные платформы помогают считывать реального человека на основе биометрических данных, что помогало предотвратить большое количество

кибератак. Вследствие этого, клиенты компании проявили большой интерес к новому решению от VisioLabs.

На сегодняшний день практически все банки Российской Федерации и часть СНГ используют технологию Deepfake Detection для распознавания дипфейков при финансовых операциях. Так, на примере одного из банков возможно отследить количество атак и примерную стоимость ущерба от использования дипфейков злоумышленниками: происходит 7-8 атак в неделю, стоимость каждой атаки примерно равна 6 миллионам долларов в год. На основании представленных материалов, продукт Deepfake Detection необходим для использования не только пользователям банковских приложений, но финансовому сектору, так как сокращает кибератаки и повышает безопасность онлайн использования.

14 мая 2025 года VisionLabs представила запуск новой технологии Deepfake Detection, с возможностью распознавания дипфейки с точностью 99,3%. МТС ID KYC - платформа управления идентификации для удаленной проверки документов - первая структура, которая внедрила дипфейк детекции внутрь программы. Компания применила продукт от VisionLabs для защиты от цифрового мошенничества. Решение Deepfake Detection способно определять самые разнообразные дипфейки. В основе дипфейка лежит: замена и генерация лица, перенос выражения лица, синхронизация губ. Благодаря постоянному обновлению системы, алгоритмы способны обучаться новым видам дипфейков, созданным с помощью современных технологий, а также определять подмену при плохом качестве видео.

Александр Паркин, руководитель исследовательских проектов в VisionLabs, поясняет как создаются Deepfake с помощью видео: выносится отдельно видео и аудио, происходит клонирование голоса, текст сгенерировался голосом из представленного аудио и после ещё одной генеративной сетью синхронизировались губы с аудио. Подобное видео

создается в течении 5 минут. Также отмечается, что программа по созданию дипфейков может быть встроена в камеру и использоваться непосредственно во время видеоконференций. Между настоящей веб-камерой и виртуальной может стоять видеокарта, которая поможет сгенерировать дипфейк и во время видеоконференции сможет повторить действия за человеком: поворот головы, кивок и т.д. Мошенник сможет наложить лицо жертвы на своё, тем самым подменяя личность. Вследствие этого, ИТ-решение рассматривается как возможность предотвратить использование «цифровых копий» в момент видео конференций.

Генеральный директор VisionLabs Дмитрий Марков отмечает доступность сервисов для создания дипфейков и сложность определения человеческому глазу. Защита от мошенничества особенно важна для видеоконференцсвязи, где нужно понимать, что все собеседники являются реальными людьми. Видеозвонки сегодня применяются повсеместно: при медицинских консультациях, на онлайн-экзаменах, для подтверждения личности при банковских операциях. Необходима повышенная защита от цифровых атак, и VisionLabs предоставляет свои продукты для повышения уровня безопасности. Уже сегодня большое количество компаний используют ИТ-решения VisionLabs.

Дипфейки представляют собой признанную угрозу для человека и компаний. На основе вышесказанного, синтетический контент нуждается в обработке специализированной обученной программой, способной выявить цифровую подмену. Большинство людей не сможет с первого раза определить находится ли перед ними реальный человек, так как человеческое зрение не сможет распознать сигналы подмены, в отличии от специально обученной компьютерной программы. Наиболее остро встает вопрос цифровой гигиены: как возможно уберечь свои сбережения от атаки дипфейков и не поддаться на провокации от злоумышленников.

Важнейшим условием для обеспечения цифровой безопасности считается внедрение верификации для любых финансовых операций, осуществляемых даже со знакомыми и близкими. С её помощью возможно определить реального человека, который просит о материальной помощи. На сегодняшний день не все платформы оборудованы достаточной идентификацией пользователей. Одним из простых способов проверки подлинности просьбы является «второй канал связи»: при получении видеоматериала с просьбой о переводе денежных средств в мессенджере, необходимо перезвонить получателю по номеру для подтверждения.

Помимо этого, требуется глобальное внедрение технологических решений, таких как продукт Deepfake Detection от VisionLabs, способные выявить на начальной стадии коммуникации «цифрового двойника», тем самым предотвратив возможность киберпреступления. Анализ выражения лица, неестественного моргания и артефакты генерации сообщат пользователю о том, что перед ним на видео дипфейк, сгенерированный нейросетью. Подобные инструменты должны стать основой верификацией на любых платформах.

Более сложным направлением является правовая защита при атаке дипфейков. Существующее правовое регулирование не до конца адоптировано под новые вызовы общества, что вызывает сложности в наказании за использование дипфейков. На сегодняшний момент, в Российской Федерации, в январе 2026 года в Госдуме был внесен законопроект об уголовной ответственности за дипфейки от партии «Справедливая Россия». С каждым годом подобная инициатива привлекает большое внимание среди законодателей. В настоящее время за незаконное использование информации предусмотрена статья 272.1 УК РФ «Незаконные использование и (или) передача, сбор и (или) хранение компьютерной информации, содержащей персональные данные, а равно создание и (или) обеспечение функционирования информационных

ресурсов, предназначенных для ее незаконных хранения и (или) распространения». За противоправные действия предусмотрено разного рода наказания: от штрафа до лишения свободы сроком до 4 лет. При этом статья не регулирует вопросы создания и использования дипфейков в мошеннических действиях, что упрощает возможность мошенников уйти от правосудия. Поэтому необходимость обновления законодательства с учетом новых цифровых вызовов становится первостепенной задачей для государства с учетом актуальных проблем.

Современный мир с каждым годом стремительно становится цифровым. Распространяется использование ИИ-моделей в различных сферах общества, в том числе и нелегальных. На сегодняшний день кибератаки с помощью ИИ-контента являются одним из популярных способов мошенничества, в частности дипфейки. Технологии обнаружения дипфейков помогают бороться с злоумышленниками. Deepfake Detection от VisionLabs – продукт, с помощью которого компании и платформы смогут предотвратить кибератаки с помощью дипфейков и сохранить безопасность использования для пользователей. Таким образом, именно сегодня закладываются стандарты кибербезопасности XXI века, где мгновенное обнаружение дипфейков становится главным инструментом в противостоянии против преступности в глобальной сети Интернет.

Список источников и литературы

1. УК РФ Статья 272.1. «Незаконные использование и (или) передача, сбор и (или) хранение компьютерной информации, содержащей персональные данные, а равно создание и (или) обеспечение функционирования информационных ресурсов, предназначенных для ее незаконных хранения и (или) распространения от 07.12.2011 года №421-ФЗ // Справочная система «КонсультантПлюс». – URL:

https://www.consultant.ru/document/cons_doc_LAW_10699/deefead19003ba8266e85fbf42fc31f60ed3c698/ (дата обращения: 01.04.2026).
2. Федеральный закон «Об информации, информационных технологиях и о защите информации» от 27 июля 2006 г. № 149-ФЗ (ред. от 24.06.2025) // Собрание законодательства Российской Федерации, 31.07.2006, № 31 (1 ч.), ст. 3448.
3. "Актуальный репортаж": в РФ создадут ПО для распознавания видеофейков // Москва 24 URL: https://www.m24.ru/videos/tehnologii/12012023/540060 (дата обращения: 01.04.2026).
4. В Госдуму внесут проект об уголовной ответственности за дипфейки // РИА Новости URL: https://ria.ru/20260127/gosduma-2070454584.html (дата обращения: 06.04.2026).
5. В России внедрят технологию распознавания дипфейков в решениях для видеоконференцсвязи // CNews URL: https://www.cnews.ru/news/line/2025-06-05_v_rossii_vnedryat_tehnologiyu (дата обращения: 01.04.2026).
6. Зелёный переход: ИИ-трансформация устойчивого развития. Сфера безопасности // Россия 24 URL: https://smotrim.ru/video/2972748 (дата обращения: 01.04.2026).
7. Как защитить от дипфейков клиентов и бизнес // PlusWorld URL: https://plusworld.ru/articles/67370/ (дата обращения: 01.04.2026).
8. Специальный репортаж: «дипфейки» // Москва 24 URL: https://www.m24.ru/shows1/4/811151 (дата обращения: 01.04.2026).

References

1. Article 272.1 of the Criminal Code of the Russian Federation. "Illegal use and (or) transfer, collection and (or) storage of computer information containing

personal data, as well as the creation and (or) maintenance of information resources intended for its illegal storage and (or) distribution of 07.12.2011 №421-FZ // ConsultantPlus help system. – URL: https://www.consultant.ru/document/cons_doc_LAW_10699/deefead19003ba8266e85fbf42fc31f60ed3c698 / (date of access: 04/01/2026).

2. Federal Law "On Information, Information Technologies and Information Protection" dated July 27, 2006 No. 149-FZ (as amended on 06/24/2025) // Sobranie zakonodatelstva Rossiyskoy Federatsii, 07/31/2006, No. 31 (1 part), art. 3448.

3. "Current reportage": the Russian Federation will create software for recognition of video images // Moscow 24 URL: https://www.m24.ru/videos/tehnologii/12012023/540060 (date of appeal: 04/01/2026).

4. A draft law on criminal liability for deepfakes will be submitted to the State Duma // RIA Novosti URL: https://ria.ru/20260127/gosduma-2070454584.html (date of request: 04/06/2026).

5. Russia will introduce deepfake recognition technology in video conferencing solutions // CNews URL: https://www.cnews.ru/news/line/2025-06-05_v_rossii_vnedryat_tehnologiyu (accessed: 04/01/2026).

6. Green transition: AI-transformation of sustainable development. Security sphere // Russia 24 URL: https://smotrim.ru/video/2972748 (date of access: 04/01/2026).

7. How to protect clients and businesses from deepfakes // PlusWorld URL: https://plusworld.ru/articles/67370 / (date of access: 04/01/2026).

8. Special report: "deepfakes" // Moscow 24 URL: https://www.m24.ru/shows1/4/811151 (date of request: 04/01/2026).