MINISTERUL EDUCAŢIEI, CULTURII ŞI CERCETĂRII AL REPUBLICII MOLDOVA
Universitatea Tehnică a Moldovei
Facultatea Calculatoare Informatică şi Microelectronică
Departamentul Ingineria Software și Automatică

Admis la uşţinere
Şef de departament:
Fiodorov I. dr., conf.univ.

„___" _____ 2022

Анализ использования технологии deepfake

Teza de master

Student: **_____ Drelinschi Sofia, TI-201M**

Coordonator: **_____ Leahu Alexei, prof. univ.,dr.**

Consultant: **_____ Cojocaru Svetlana, lector univ.**

Chişinău, 2022

Аннотация

Данный дипломный проект посвящен изучению технологии deepfake, прежде всего, в контексте актуальности ее концепции, с учетом развития технологических процессов, создания аудио- и видеоконтента, а также механизма борьбы с поддельными изображениями и фальшивыми видео. Понимание механизмов данной технологии позволяет значительно сократить издержки, затрачиваемые на борьбу с поддельным контентом, что позволит добиваться более высокой эффективности работы различных предприятий, организаций, СМИ.

В данной дипломной работе представлен процесс развития данной технологии, а также определены механизмы, позволяющие максимально точно и быстро определять признаки использования фальшивого контента.

Целью данной работы развитие процессов использования данной технологии, всестороннего учета и понимания позитивных и негативных аспектов deepfake для достижения максимальной эффективности при ее использовании в политике, электоральном процессе, медийных и игровых программных продуктах. Данная работа расписана в трех главах, каждая из которых описывает определённый раздел функционирования и использования данной технологии.

Adnotarea

Acest proiect de diplomă este dedicat studiului tehnologiei deepfake, în primul rând în contextul relevanței conceptului său, luând în considerare dezvoltarea proceselor tehnologice, crearea de conținut audio și video, precum și mecanismul de combatere a imaginilor și videoclipurilor false. Înțelegerea mecanismelor acestei tehnologii poate reduce în mod semnificativ costul conținutului contrafăcut, ceea ce va permite obținerea unei mai mari eficiențe în diverse întreprinderi, organizații și mass-media.

Această teză prezintă procesul de dezvoltare a acestei tehnologii și, de asemenea, identifică mecanismele de detectare a semnelor de utilizare a conținutului fals cât mai precis și mai rapid posibil.

Scopul acestei teze este de a dezvolta procesele de utilizare a acestei tehnologii, luând în considerare și înțelegând aspectele pozitive și negative ale deepfake pentru o eficacitate maximă atunci când se utilizează în politică, în procesul electoral, în mass-media și în software-ul de joc. Această lucrare este structurată în trei capitole, fiecare dintre acestea descriind o secțiune specifică privind funcționarea și utilizarea tehnologiei.

Abstract

This thesis project is devoted to the study of deepfake technology, primarily in the context of the relevance of its concept, taking into account the development of technological processes, the creation of audio and video content, as well as the mechanism of combating fake images and fake videos. Understanding the mechanisms of this technology can significantly reduce the costs of counterfeit content, which will allow to achieve greater efficiency in the work of various enterprises, organizations and media.

This thesis presents the process of developing this technology, as well as identifying the mechanisms that allow the most accurate and rapid identification of signs of fake content.

The aim of this work is to develop the processes of using this technology, the full consideration and understanding of the positive and negative aspects of deepfake to achieve maximum efficiency in its use in politics, electoral process, media and game software products. This paper is written in three chapters, each of which describes a particular section of the functioning and use of this technology.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

В последние годы наблюдается огромный рост исследований с использованием новейших технологий в области машинного обучения. Особенно интенсивно в сфере IT транслируются сообщения о разработке алгоритмов «deepfake», помогающих трансформировать контент. В этой связи можно отметить их разнообразие: среди них, в частности, технология синтеза голосов, написание текстов на основе заданных исходных алгоритмов, генерирование нейросетями самого разнообразного видеоконтента, а также появление все более усложненных возможностей для переноса цифровых копий человеческих лиц и тел на различных изображениях.

Значимый толчок развитию алгоритмов «deepfake» придают все ускоряющиеся процессы глобализации, которые постоянно открывают новые возможности для совершенствования и усложнения технологий в этой области IT.

Именно такой является и современная константа алгоритмов «deepfake», при помощи которых все увереннее осваивается процесс создания программных продуктов, оптимизирующих производство контента, впоследствии становясь компонентами успешной организации различных стартапов. За период с начала применения алгоритмов «deepfake» и по настоящий момент эволюция данных технологий оказала значительное влияние на самые разные производственные и общественные процессы – от политики и экономики до медийного рынка, избирательных кампаний и шоу-бизнеса. При этом необходимо отметить, что возможности технологии «deepfake» носят как позитивный характер, но также обладают и серьезными репутационными рисками, став, фактически, и деструктивной платформой для изготовления фальшивых изображений и компрометирующего контента.

Тема данной магистерской работы — «Анализ использования технологии deepfake» — позволяет всесторонне исследовать современные концепции и эволюционные процессы технологий «deepfake», показать существующие архитектуры алгоритмов – и самых распространенных и только начинающих использоваться, а также процессы машинного обучения, позволяющие создавать креативный контент.

Так как технологии «deepfake» могут создавать все более реалистичные изображения и видеоконтент, то это позволяет определить и стандартные модели создания развлекательных приложений или же использования подобных алгоритмов в игровой индустрии, а также логику формирования брендированного контента, позволяющего улучшить монетизацию используемых технологий.

Также в процессе исследования проблематики были проанализированы доступные решения создания «deepfake» и фундаментальные возможности самой логики подобных

решений. Подобный подход позволяет масштабировать как недостатки методов создания «deepfake», так и их сильные стороны, позволяя разрабатывать технологии детекции поддельных видео и ложных новостей, а также создавать антидоты для распознавания сгенерированного контента.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Изучение deepfake технологии. Режим доступа: https://recfaces.com/articles/what-is-deepfake
2. Deepfake в бизнесе. Режим доступа: https://www.businessinsider.com/what-is-deepfake
3. Изучение принципа работы deepfake. Режим доступа: https://spectrum.ieee.org/what-is-deepfake#toggle-gdpr
4. Christoph Bregler, Michele Covell, Malcolm Slaney — Video Rewrite: Driving Visual Speech with Audio
5. SUPASORN SUWAJANAKORN, STEVEN M. SEITZ, and IRA KEMELMACHER-SHLIZERMAN, University of Washington — «Synthesizing Obama»: Learning Lip Sync from
6. Agarwal, S., Farid, H., Gu, Y., He, M., Nagano, K., and Li, H. (2019, June). Protecting world leaders against deep fakes. In Computer Vision and Pattern Recognition Workshops (pp. 38-45).
7. Lin, J., Li, Y., & Yang, G. (2021). FPGAN: Face de- identification method with generative adversarial networks for social robots. Neural Networks, 133, 132-147.
8. Liu, M. Y., Huang, X., Yu, J., Wang, T. C., & Mallya, A. (2021). Generative adversarial networks for image and video synthesis: Algorithms and applications. Proceedings of the IEEE, DOI: 10.1109/JPROC.2021.3049196.
9. Bloomberg (2018, September 11). How faking videos became easy and why that's so scary. Режим доступа: https://fortune.com/2018/09/11/deep-fakes-obama-video/
10. Chesney, R., and Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. Foreign Affairs, 98, 147.
11. Hwang,T.(2020).Deepfakes: A Grounded Threat Assessment. Centre for Security and Emerging Technologies, Georgetown University.
12. Zhou, X., and Zafarani, R. (2020). A survey of fake news: fundamental theories, detection methods, and opportunities. ACM Computing Surveys (CSUR), DOI: https://doi.org/10.1145/3395046.
13. Kaliyar, R. K., Goswami, A., and Narang, P. (2020). Deepfake: improving fake news detection using tensor decomposition based deep neural network. Journal of Supercomputing, DOI: https://doi.org/10.1007/s11227-020-03294-y.
14. Cheng, Z., Sun, H., Takeuchi, M., and Katto, J. (2019). Energy compaction-based image compression using convolu- tional autoencoder. IEEE Transactions on Multimedia. DOI: 10.1109/TMM.2019.2938345.

15. Chorowski, J., Weiss, R. J., Bengio, S., and Oord, A. V. D. (2019). Unsupervised speech representation learning using wavenet autoencoders. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 27(12), pp. 2041-2053.
16. Faceswap: Deepfakes software for all. Режим доступа:https://github.com/deepfakes/faceswap
17. FakeApp 2.2.0. Режим доступа: https://www.malavida.com/en/soft/fakeapp/
18. DeepFaceLab. Режим доступа: https://github.com/iperov/DeepFaceLab
19. DFaker. Режим доступа: https://github.com/dfaker/df
20. DeepFake tf: Deepfake based on tensorflow. Режим доступа: https://github.com/StromWine/DeepFake tf
21. Keras-VGGFace: VGGFace implementation with Keras framework. Режим доступа: https://github.com/rcmalli/keras-vggface
22. Faceswap-GAN. Режим доступа: ttps://github.com/shaoanlu/faceswap-GAN.
23. CycleGAN. Режим доступа: https://github.com/junyanz/pytorch- CycleGAN-and-pix2pix.
24. DeepFaceLab: Explained and usage tutorial. Режим доступа: https://mrdeepfakes.com/forums/thread-deepfacelab- explained-and-usage-tutorial.
25. Korshunov, P., and Marcel, S. (2018, September). Speaker inconsistency detection in tampered video. In 2018 26th Eu- ropean Signal Processing Conference (EUSIPCO) (pp. 2375- 2379). IEEE.