

Deepfakes : des atouts négligés

Anna Tranchida, Christel Espinasse, Doriane Rosset

Etudiant-e-s en ingénierie des médias, 1^{ère} année, HEIG-VD

Les *deepfakes*, souvent associés à une technologie dangereuse, suscitent des inquiétudes. En effet, l'intelligence artificielle facilite la production de vidéos falsifiées, souvent employée à des fins malveillantes telles que la manipulation de l'opinion publique, la cyberintimidation et la diffusion de désinformation à l'échelle mondiale.

Pourtant, les *deepfakes* peuvent être utilisés de manière très ludique. N'avez-vous jamais rêvé d'apprendre les mathématiques grâce à Donald Trump¹ ? Sur Instagram, des comptes dédiés ont été créés et vous y trouverez quantité de personnalités publiques qui sont en réalité des hypertrucages. Alors, plutôt que de condamner cette technologie, ne devrions-nous pas explorer comment l'utiliser de manière responsable et éthique ?

I. QUAND LE RÉEL DEVIENT FAUX

Les *deepfakes*, fusion de *deep learning* et *fake*, ou *hypertrucage*, désignent des médias synthétiques manipulés numériquement pour substituer de manière convaincante l'image d'une personne par celle d'une autre, voire créer des visages humains n'existant pas dans la réalité [1].

L'histoire des *deepfakes* remonte au développement de la manipulation de photos à partir du XIXe siècle avec par exemple le portrait truqué d'Abraham Lincoln [2], progressant ensuite avec l'avènement du cinéma à la fin du même siècle. Au fil du temps, les *deepfakes* ont bénéficié des avancées en vision par ordinateur et en intelligence artificielle devenant ainsi de plus en plus convaincants et accessibles [3]. Le terme a été utilisé dès 2017 lorsqu'un modérateur de *Reddit* a commencé à publier des vidéos pornographiques falsifiées de célébrités [4].

II. CONTEXTE JURIDIQUE

A. En Suisse

Au niveau légal, les auteurs de *deepfakes* peuvent parfaitement faire l'objet de poursuites pénales. Les infractions contre l'intégrité sexuelle, la diffamation, la calomnie, l'injure sont autant d'atteintes à l'honneur ou à la personnalité, passibles d'amendes ou de peines de prison [5]. L'usurpation d'identité, entrée en vigueur dans le Code pénal en 2023 est également punie par la loi. Elle s'applique à toute personne qui utiliserait l'image d'un tiers afin de nuire à sa personne [6], tant physiquement que virtuellement [7].

B. Dans le monde

L'Europe est soumise à l'*AI Act* (Législation sur l'Intelligence Artificielle) qui réglemente les applications de l'IA, selon leur niveau de risque pour la société [8]. Comme exemple de risque très élevé, on retrouve la surveillance étatique pour observer le comportement des citoyens, qui obtiennent des avantages s'ils sont irréprochables (situation en Chine). Les *deepfakes* quant à eux se trouvent dans la catégorie des risques limités [9].

La plupart des nations se penche sur la question, à l'exemple des États-Unis où certains états ont adopté des lois qui restreignent l'utilisation de l'hypertrucage. En règle générale, on limite sans interdire [10].

III. USAGES COURANTS

Il existe plusieurs manières de créer des *deepfakes*, mais toutes se basent sur des IAs de type *deep learning* nommés GANs (pour *Generative Adversarial Networks*).

Les GANs sont des modèles d'apprentissage automatique non supervisés qui utilisent des réseaux de neurones pour générer de nouvelles données réalistes [11].

Ces derniers sont donc « nourris » d'images vidéo. Ils analysent les gestes et les mimiques de personnes sources, seconde par seconde, pour les « apprendre ». L'étape suivante consiste à fournir des images cibles à modifier, pour y superposer, reproduire ou coller les résultats de l'analyse. Plus le volume d'informations saisi dans le modèle est important, plus le *deepfake* sera réaliste [11].

Le *face swap* permet de remplacer un visage par un autre, que ce soit celui d'une personne à proximité ou celui d'une photo de votre galerie. Des applications grand public comme Snapchat ont démocratisé cette technique avec des filtres applicables en temps réel [12].

Le *face to face*, consiste à prendre une vidéo existante, extraire le visage de la cible, puis le combiner avec les expressions faciales d'un imitateur [13]. Cette technique a par exemple été utilisée pour créer une vidéo en 2019 où Barack Obama, dans un discours fictif, insultait Donald Trump. Les mouvements de bouche étaient en fait ceux de l'acteur Jordan Peele, une mise en scène montée avec *BuzzFeed* pour sensibiliser aux dangers des *deepfakes* [14].

Enfin, le *lip sync* modifie directement les mouvements des lèvres pour les faire correspondre à un audio donné [15]. A l'exemple d'un Mark Zuckerberg, fondateur de Facebook, qui semblait prêter allégeance à une organisation criminelle fictive

¹ <https://www.instagram.com/calculuswithai/>

dans une vidéo. La supercherie a été rapidement démasquée, car sa voix ne correspondait pas [16].

L'univers des *deepfakes* est en plein essor. Autrefois réservés aux professionnels et chercheurs, ils s'adressent désormais au grand public grâce à une variété d'outils gratuits en ligne (voir ci-dessous). Ces plateformes simplifient la création de *deepfakes* avec des interfaces intuitives, rendant la technologie accessible même aux débutants. Cependant, pour des fonctionnalités avancées et un soutien technique personnalisé, des options payantes sont disponibles, offrant des algorithmes plus sophistiqués, une qualité de rendu améliorée et une assistance spécialisée.

IV. DES APPLICATIONS POSITIVES

A. Culture et divertissement

La montée en puissance des *deepfakes* promet de révolutionner le monde du cinéma et des arts de manière tangible et concrète. A Hollywood, cette technologie ouvre de nouvelles portes en permettant la résurrection virtuelle d'acteurs disparus ou en recréant des versions plus jeunes de stars iconiques. Par exemple, dans *Rogue One : A Star Wars Story*, la recréation numérique de Peter Cushing en Grand Moff Tarkin montre le potentiel de la technologie à raviver des personnages emblématiques. De même, la possibilité d'utiliser des *deepfakes* pour remplacer des acteurs dans des scènes post-production offre un gain d'efficacité immense. Des films comme *The Irishman* ont adopté cette approche pour rajeunir numériquement leurs acteurs, démontrant ainsi les avantages pratiques de la technologie [17].

Dans le monde de l'art, les *deepfakes* sont devenus un outil de création révolutionnaire, permettant aux artistes de manipuler et de réinventer des œuvres d'art célèbres. Des chercheurs du *Samsung AI Center* de Moscou et de l'*Institut des sciences et technologies* de Skolkovo ont donné vie à des chefs-d'œuvre tels que la *Joconde* de Léonard de Vinci. En utilisant des algorithmes avancés, ces œuvres ont été animées de manière réaliste, offrant une nouvelle dimension à des peintures emblématiques et suscitant un intérêt renouvelé pour l'art classique. En découle un tout nouveau courant artistique avec des artistes contemporains qui font usage de ces outils technologiques pour les intégrer à leurs futures et récentes œuvres [11].

De même, les *deepfakes* offrent de nouvelles possibilités d'engagement du public dans les institutions culturelles, Le musée *Dali* de Saint-Petersbourg en Floride, a récemment lancé une exposition intitulée *Dali Lives*, où une reconstitution grandeur nature de Salvator Dali a été créée à partir de milliers d'heures de séquences vidéo d'archives. Cette version numérique réaliste de Dali accueille les visiteurs à l'entrée du musée, leur racontant des histoires sur sa vie et offrant une expérience immersive et interactive nouvelle [18].

Plus proche de chez nous, l'EPFL a réinventé l'exposition muséale en créant *Deep Fakes: Art and its Double* qui

questionne les notions traditionnelles d'authenticité et de provenance des œuvres d'art. Des pièces constitutives de l'art du monde entier sont rendues tangibles grâce aux technologies les plus avancées de l'imagerie et de l'immersion interactive [19].

B. Une vie après la mort ?

Les *deepfakes* offrent une nouvelle dimension aux hommages rendus aux personnalités et proches décédés. En utilisant des enregistrements audio et vidéo existants, des modèles d'IA peuvent générer des réponses réalistes aux questions posées par la personne en deuil. Ils permettent de les voir partager des anecdotes, des messages inspirants ou simplement converser avec eux. Cette technologie permet aussi de recréer des moments précieux : un discours historique, une performance musicale ou une conversation intime.

C'est ce qu'a fait un jeune chinois en « ressuscitant sa mère » à l'aide de quelques photos et vidéos. Il a pu créer un avatar virtuel de sa mère qui a son apparence, sa voix et jusqu'à même son accent [20].

C. Et encore ?

L'IA évolue très rapidement. Dans un futur plus ou moins proche, les GANs - qui sont à l'origine des *deepfakes* - pourraient s'avérer très utiles dans le domaine médical, notamment pour repérer des signes de cancer. La plus grande promesse serait de pouvoir détecter les tumeurs cancéreuses à un stade précoce que les médecins ne sauraient voir. Le problème souvent relevé est la confidentialité. Il pourrait ne pas y avoir assez de données pour former correctement cette IA [21].

Toutefois, à Taipei, des chercheurs ont mis au point une manière d'anonymiser les patients afin de pouvoir analyser leurs émotions. Ils modifient les visages tout en conservant les traits faciaux. Ainsi, les médecins pourraient s'entraîner à reconnaître et à interpréter les expressions faciales émotionnelles des malades de manière plus précise. Un environnement empathique serait alors créé pour qu'ils se sentent plus compris et soutenus [22].

Nous pouvons également relever l'utilisation de *deepfakes* dans le domaine militaire, même si cela est vu d'un très mauvais œil. Dans un contexte de guerre, une chaîne ukrainienne a été piratée et une vidéo montrant Volodymyr Zelensky ordonnant à son pays de rendre les armes avait été diffusée nationalement [23].

D. A vous de jouer ?

Quoi de mieux que d'essayer par vous-même ? Une façon simple et gratuite, bien que limitée, de créer des *deepfakes* consiste tout d'abord à visiter des sites qui vous permettent de superposer votre visage sur celui d'une autre personne, le *face swap*². Une fois cette fusion réalisée, vous pouvez créer une vidéo en utilisant des sites qui donnent vie à vos photos en créant un avatar et en le faisant parler et dire ce que vous souhaitez, utilisant notamment le *lip sync*³. C'est une façon ludique d'explorer les possibilités des *deepfakes*, même si ces

² <https://www.deepswap.ai/fr>

³ <https://www.vidnoz.com/ai-solutions/ai-lip-sync.html>

outils ont leurs limites en termes de réalisme et de précision.

RÉFÉRENCES

- [1] « Deepfake - Wikipedia », *Wikipedia*. 2 avril 2024. Consulté le: 8 avril 2024. [En ligne]. Disponible sur: <https://en.wikipedia.org/w/index.php?title=Deepfake&oldid=1216789047>
- [2] A. Savoye, « Bien avant Photoshop, Lincoln faisait retoucher ses photos pour paraître beau gosse », *Konbini - Musique, cinéma, sport, food, news : le meilleur de la pop culture*. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://www.konbini.com/arts/bien-avant-photoshop-lincoln-faisait-retoucher-ses-photos-pour-paraître-beau-gosse/>
- [3] geraldholubowicz, « La courte histoire des deepfakes », *Mediapart*. Consulté le: 8 avril 2024. [En ligne]. Disponible sur: <https://blogs.mediapart.fr/geraldholubowicz/blog/230520/la-courte-histoire-des-deepfakes>
- [4] « Deepfake | History & Facts | Britannica ». Consulté le: 29 avril 2024. [En ligne]. Disponible sur: <https://www.britannica.com/technology/deepfake>
- [5] Q. Jacquemin, « Deepfakes pornographiques : Le droit pénal suisse prend-il en compte cette nouvelle menace ? », *LexTech Institute*. Consulté le: 22 avril 2024. [En ligne]. Disponible sur: <https://www.lextechinstitute.ch/deepfakes-pornographiques-le-droit-penal-suisse-prend-il-en-compte-cette-nouvelle-menace/>
- [6] « RS 311.0 - Code pénal suisse du 21 décembre 1937 | Fedlex ». Consulté le: 22 avril 2024. [En ligne]. Disponible sur: https://www.fedlex.admin.ch/eli/cc/54/757_781_799/fr
- [7] « Les Deepfakes : des fausses vidéos plus vraies que nature », *Bilan*. Consulté le: 8 avril 2024. [En ligne]. Disponible sur: <https://www.bilan.ch/opinions/steven-meyer/les-deepfakes-des-fausses-vidéos-plus-vraies-que-nature>
- [8] Onfido, « Getting EU AI regulation right », *Onfido*. Consulté le: 22 avril 2024. [En ligne]. Disponible sur: <https://onfido.com/blog/getting-eu-ai-regulation-right/>
- [9] « Loi sur l'intelligence artificielle de l'UE - Développements et analyses actualisés de la loi sur l'intelligence artificielle de l'UE ». Consulté le: 29 avril 2024. [En ligne]. Disponible sur: <https://artificialintelligenceact.eu/fr/>
- [10] C. Bowyer, « Deepfake laws: is AI outpacing legislation? », *Onfido*. Consulté le: 22 avril 2024. [En ligne]. Disponible sur: <https://onfido.com/blog/deepfake-law/>
- [11] A. Kohli et A. Gupta, « Detecting DeepFake, FaceSwap and Face2Face facial forgeries using frequency CNN », *Multimed. Tools Appl.*, vol. 80, n° 12, p. 18461-18478, mai 2021, doi: 10.1007/s11042-020-10420-8.
- [12] A. GENTEZ, « Que veut dire "faire un Face Swap" ? » Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://www.programme-tv.net/news/people/97079-que-veut-dire-faire-un-face-swap/>
- [13] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, et M. Nießner, « Face2Face: Real-Time Face Capture and Reenactment of RGB Videos », in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, juin 2016, p. 2387-2395. doi: 10.1109/CVPR.2016.262.
- [14] « Fake news – Un Obama plus vrai que nature insulte Trump », *Tribune de Genève*. Consulté le: 29 avril 2024. [En ligne]. Disponible sur: <https://www.tdg.ch/un-obama-plus-vrai-que-nature-insulte-trump-553836746599>
- [15] « Lip-syncing Obama: New tools turn audio clips into realistic video », *UW News*. Consulté le: 29 avril 2024. [En ligne]. Disponible sur: <https://www.washington.edu/news/2017/07/11/lip-syncing-obama-new-tools-turn-audio-clips-into-realistic-video/>
- [16] « Mark Zuckerberg pris au piège des "deepfakes" », *France 24*. Consulté le: 29 avril 2024. [En ligne]. Disponible sur: <https://www.france24.com/fr/20190612-facebook-mark-zuckerberg-instagram-deepfakes-pelosi-video>
- [17] C. Hood, « How Deepfake Technology Can Change The Movie Industry », *ScreenRant*. Consulté le: 21 avril 2024. [En ligne]. Disponible sur: <https://screenrant.com/movies-deepfake-technology-change-hollywood-how/>
- [18] « 3 Things You Need to Know About "Deep Fakes" in Art & Culture », *Cuseum*. Consulté le: 8 avril 2024. [En ligne]. Disponible sur: <https://cuseum.com/blog/2019/12/17/3-things-you-need-to-know-about-ai-powered-deep-fakes-in-art-amp-culture>
- [19] « Deep Fakes – EPFL Pavilions ». Consulté le: 21 avril 2024. [En ligne]. Disponible sur: <https://epfl-pavilions.ch/fr/exhibitions/deep-fakes-art-and-its-double>
- [20] « Immortalité numérique : la création de deepfakes de proches décédés est en plein essor sur le marché chinois, les utilisateurs se tournent vers les avatars générés par IA pour surmonter leur peine », *Developpez.com*. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://intelligence-artificielle.developpez.com/actu/357614/Immortalite-numerique-la-creation-de-deepfakes-de-proches-decedes-est-en-plein-essor-sur-le-marche-chinois-les-utilisateurs-se-tournent-vers-les-avatars-generes-par-IA-pour-surmonter-leur-peine/>
- [21] A. Vera, « Et si les deepfakes avaient les capacités d'aider la médecine ? », *Siècle Digital*. Consulté le: 22 avril 2024. [En ligne]. Disponible sur: <https://siecledigital.fr/2019/07/05/et-si-les-deepfakes-avaient-les-capacites-daider-la-medecine/>
- [22] « Can deepfakes be beneficial in healthcare? » Consulté le: 22 avril 2024. [En ligne]. Disponible sur: <https://www.paubox.com/blog/can-deepfakes-be-beneficial-in-healthcare>
- [23] « Piratée, une chaîne d'information ukrainienne diffuse un "deepfake" de Volodymyr Zelensky », *BFMTV*. Consulté le: 22 avril 2024. [En ligne]. Disponible sur: https://www.bfmtv.com/tech/piratee-une-chaîne-d-information-ukrainienne-diffuse-un-deepfake-de-volodymyr-zelensky_AN-202203170296.html