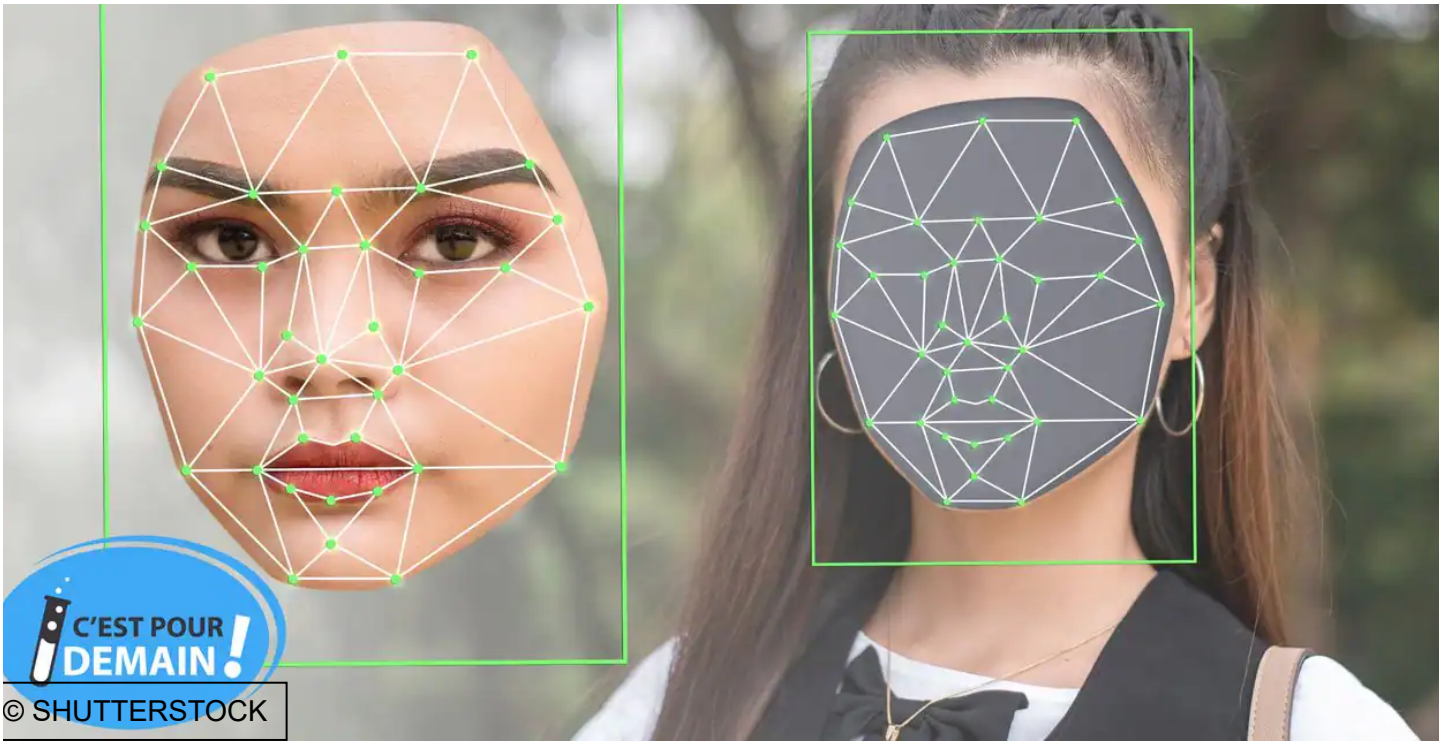
"On est face à un sérieux problème": comment les deepfakes risquent de bouleverser notre quotidien

Digital (/economie/digital)

Abonnés Publié le 24-10-21 à 12h10 - Mis à jour le 04-11-21 à 14h29

Des photos réalisées entièrement par ordinateur aux visages incrustés dans des vidéos, en passant par des voix recréées à la quasi-perfection... Les technologies permettant de créer des deepfakes s'améliorent de jour en jour. Comment distinguer le vrai du faux ? Sera-t-il bientôt possible de créer des clones autonomes ? Le point sur un futur où fiction et réalité s'entremêlent.



© SHUTTERSTOCK

Les outils de partage sont désactivés en mode preview

Cet été, le photographe Jonas Bendiksen s'est fait remarquer lors de Visa pour l'image, l'un des plus grands festivals de photojournalisme au monde. Le photographe norvégien y présentait les photos de son livre "The Book of Veles (https://www.jonasbendiksen.com/books/the-book-of-veles)", qui plonge les lecteurs dans l'histoire de la ville de Velès (en Macédoine du Nord), véritable usine à fake news réputée pour les nombreux sites frauduleux pro-Trump qui y ont été créés lors de l'élection présidentielle américaine de 2016.

Sorti en juillet 2021, le livre est un succès. Plusieurs personnes contactent le photographe pour le remercier de s'être intéressé à la problématique des fake news. Sauf que, si le photographe s'était bel et bien rendu sur place pour réaliser la série, il n'y avait rencontré personne. Tous les individus présents sur ses photos sont en réalité des personnages créés en 3D, incrustés ensuite par ordinateur. "Mon intention était que les gens le découvrent... Le problème est que cela n'est pas arrivé", explique le Norvégien dans une interview donnée une fois la supercherie révélée.

Cette tromperie, même si elle n'a pas plu à tout le monde, suscite un vif débat au sein de la profession. Pour Joris Bendiksen, qui souhaitait alerter sur les dangers des fake news, l'objectif est atteint. "*Les images générées par ordinateur sont devenues si faciles à utiliser, n'importe qui peut générer des informations artificielles désormais*", explique le photographe dans l'interview révélant la tromperie (https://www.magnumphotos.com/newsroom/society/book-veles-jonas-bendiksen-hoodwinked-photography-industry/?utm_source=twitter&utm_medium=social&utm_campaign=editorial). "*Je crois qu'on est face à un sérieux problème. J'espère que ce que je viens de faire va nous faire prendre conscience qu'on est déjà entré dans cette ère-là...*"

Visages créés de toutes pièces, clone vocal et compagnie

Si les images créées par ordinateurs étaient considérées comme "trop parfaites" par le passé, celles-ci sont désormais presque indétectables à l'œil nu. "*Grâce à l'intelligence artificielle, on sait simuler ce qu'aurait filmé une caméra. On arrive donc à un résultat assez parfait*", explique Jean-Luc Dugelay, professeur spécialiste des traitements d'image au département de sécurité numérique de l'EURECOM.

C'est par exemple le cas des portraits hyperréalistes ci-dessous. Tous ont été générés par une intelligence artificielle, via un site internet (https://thispersondoesnotexist.com/), et représentent des personnes qui n'existent pas dans la réalité. "*La machine apprend ce qu'est un visage humain puis en génère des nouveaux. On génère donc du faux à partir de l'apprentissage de beaucoup de vrai"*, détaille l'expert, précisant que, pour lui, il s'agit plutôt de visages dérivés de la réalité que de créations au sens strict du terme.

Portraits générés via une intelligence artificielle © thispersondoesnotexist.com

Et grâce aux avancées technologiques de ces dernières années, les trucages ne se limitent plus aux photos. Les "deepfakes" sont des vidéos ou des sons modifiés à l'aide d'une intelligence artificielle. Il est par exemple possible de changer le visage d'une personne apparaissant dans une vidéo. Et si la technologie n'est popularisée qu'en 2017, il devient de plus en plus facile de l'utiliser. Il est désormais possible de créer des vidéos "deepfakes" en quelques instants. Des dizaines d'applications mobiles proposent par exemple d'incruster son visage dans le dernier Marvel ou de faire chanter un tube de l'été à la Joconde, avec un résultat, il faut bien l'avouer, souvent moyen.

Il faut distinguer ces deepfakes réalisés quasiment en temps réel des vidéos où l'incrustation se fait en postproduction par des professionnels des effets spéciaux. "*Les deepfakes de qualité ne se font pas encore en temps réel"*, rappelle le professeur, prenant l'exemple des vidéos de Tom Cruise publiées en mars 2021 sur TikTok. Réalisés par le belge Chris Umé (https://www.lalibre.be/lifestyle/people/2021/03/05/les-videos-deepfake-

de-tom-cruise-sur-tik-tok-sont-loeuvre-dun-belge-M2WCHTTFQ5GYFKBRPT6JE4N67U/), ces deepfakes au résultat impeccable avaient fait couler beaucoup d'encre à l'époque tant le résultat était impressionnant.

"*Les gens doivent comprendre que c'est le travail de professionnels. J'ai formé cet acteur pendant deux mois, puis j'ai eu besoin de deux jours complets par film. Et puis encore 24 heures de postproduction"*, expliquait à ce sujet l'homme à l'origine des vidéos, avant d'ajouter que "*quiconque ne peut pas faire ça comme ça, comme certains le pensent*".

Depuis peu, certains services proposent même de cloner des voix humaines. "*On prend toutes les caractéristiques de votre voix et on arrive à la reproduire. Mais il y a quand même un côté synthétique qui reste détectable, peut-être plus facilement que pour l'image*", explique Jean-Luc Dugelay.

Et si cette technologie est encore loin d'être au point, il est possible d'arriver à des résultats assez convaincants pour duper certaines personnes. En 2019, un deepfake vocal avait été utilisé pour arnaquer le CEO d'une entreprise basée au Royaume-Uni (https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402) de près de 220.000€. Début 2020, lors du second cas de ce type rendu public, des fraudeurs avaient réussi à dérober plus de 35 millions de dollars (https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/?sh=7d3b13627559) en clonant la voix d'un directeur d'entreprise. Dans les deux cas, les personnes piégées connaissaient la personne dont la voix avait été clonée… Et n'avaient rien remarqué.

Distinguer le vrai du faux

Face à ces innovations, à quoi peut-on encore se fier ? "*L'expression 'Il faut le voir pour le croire', c'est fini*", confirme Jean-Luc Dugelay. "*Pour d'autres éléments, on ne se pose pas forcément la question. Quand quelqu'un signe, par exemple, vous savez que quelqu'un a pu imiter la signature. Désormais, quand vous verrez la vidéo ou que vous entendrez la voix de quelqu'un, il faudra faire attention car ça ne sera pas forcément la personne réelle."*

Et le professeur met en garde : "*Plus ça va, moins on distinguera le vrai de l'artificiel. Et ce sera même difficile de définir la limite parce que même l'artificiel, souvent ultra-réaliste, est dérivé du vrai.*"

Si la création de deepfakes peut sembler amusante au premier abord, l'utilisation de ces technologies s'avère parfois problématique. Rappelons qu'avant de se populariser, la technologie était principalement utilisée pour créer de faux films pornographiques en utilisant le visage de célébrités. De nombreux experts internationaux craignent également que les deepfakes ne représentent une menace pour la démocratie (https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy). Car si la majorité des deepfakes créés actuellement le sont dans le but de réaliser une prouesse technique ou de se moquer de certains hommes politiques, cela pourrait bien changer et la technique pourrait être employée au service de fake news.

"*Le prochain danger, c'est d'avoir des deepfakes dans lesquels les gens essayent vraiment de mélanger le faux et la réalité"*, explique le professeur Dugelay. "*Faire une performance technique non plus pour impressionner, mais pour duper les gens et créer la confusion. On aura des discours qui vont être travaillés, c'est-à-dire par exemple le discours du président Macron, mais juste un tout petit peu altéré. Avec la phrase de trop, mais où tout le reste paraît cohérent*", prévient le spécialiste. "*Ça va arriver et ce sera beaucoup plus compliqué à gérer.*"

Duper l'humain ou la machine ?

Et s'il va devenir de plus en plus dur pour les humains de distinguer le vrai du faux, le spécialiste des traitements d'image explique que la détection de vidéos truquées est devenue un business relativement opaque. "*Vous pouvez acheter des produits de détecteur de trucages d'images ou de deepfakes, mais vous allez avoir beaucoup de mal à connaître leur réelle efficacité*", avertit Jean-Luc Dugelay.

C'est le "Deepfake Detection Challenge" (DFDC) (https://ai.facebook.com/datasets/dfdc/), un concours de détection de deepfakes ayant eu lieu en juin 2020, qui sert de référence en la matière. Le DFDC était organisé, entre autres, par Amazon Web Service, Facebook et Google. Les 2.000 participants, parmi lesquels figuraient des experts reconnus mondialement, devaient mettre au point un modèle de détection sur base de près de 100.000 vidéos.

"*La meilleure méthode a obtenu 82% de détection. C'est bien, mais ça veut dire quand même qu'elle se trompe une fois sur 5*", relève le professeur Dugelay. "*Ce qui est un peu gênant, c'est que si on prend ce détecteur et qu'on l'utilise sur des deepfakes créés avec*

une méthode qui n'a pas été apprise, ça tombe à 65%", soit à peine plus qu'en essayant au hasard.

Malgré les recherches en la matière, Jean-Luc Dugelay estime qu'il va devenir de plus en plus difficile de créer des vidéos indétectables par les machines. "*Les deepfakes fabriqués il y a un an ou deux sont presque toutes détectées. Et plus on aura des détecteurs performants, plus il faudra mettre la barre haut pour créer des deepfakes crédibles*", insiste le spécialiste des traitements d'image. "*Si les moyens à mettre en jeu sont trop importants, peut-être que cela ne vaudra plus le coup ?*", s'interroge-t-il.

Bientôt tous remplacés par des clones ?

Alors, que nous réserve l'avenir ? "*La dernière étape, c'est le temps réel*", estime le professeur Dugelay. "*Pour l'image, on va y arriver."* Reste ensuite le problème des interactions, des réactions et du comportement facial des gens. "*Le gros frein qu'il y aura dans les années à venir, c'est si vous connaissez la personne*", explique l'expert. En effet, pour tromper quelqu'un qui connaît la personne imitée, il faut que la machine soit en mesure de répliquer la façon de parler, les intonations et, surtout, la connaissance de la vie de tous les jours. La machine ne peut pas faire mieux que ce qu'elle a déjà appris.

"*Je pense que l'on pourra faire des choses avec la complicité des gens*", estime le professeur. Selon lui, il sera bientôt possible de créer son propre clone, à qui l'on racontera toute notre vie. *"Il va s'améliorer, connaître toutes vos habitudes et tout votre réseau de connaissances. Et il y a un moment où il pourra vraiment vous remplacer.*" Et cet avenir n'est pas si lointain. "*Dans 5 ou 10 ans, on pourra avoir cette conversation, mais ça ne sera pas vraiment moi*", affirme Jean-Luc Dugelay. "*Il faudra avoir une procédure pour être certain que vous me parliez à moi, et non pas à un deepfake de moi qui a appris ma façon de parler et qui vous raconte ce que je suis en train de dire.*"

Et après ? "*Si on arrive à faire ça, la prochaine étape c'est de le faire en 3D, en vrai. Et dans ce cas-là, c'est un humanoïde*", conclut le spécialiste. "*On n'y est pas encore*", rassure tout de même Jean-Luc Dugelay, qui s'interroge déjà sur les possibilités futures. "*Est-ce que l'on pourra aussi créer des personnes de toutes pièces ? Pour le moment, ils créent de faux visages, mais pourraient-ils créer des fausses vidéos de fausses personnes?"*

Sur le même sujet

Google alimente la recherche contre les vidéos truquées hyper réalistes	(/economie/digital/google-alimente-la-recherche-contre-les-videos-truquees-hyper-realistes-5d8c524ed8ad5878fd64bb5d)
5 défis de cybersécurité à surmonter en 2020	(/economie/digital/5-defis-de-cybersecurite-a-surmonter-en-2020-5def77dff20d5a0c46f3563e)
Le patron de Google plaide pour une "approche proportionnée" pour réglementer l'intelligence artificielle	(/economie/entreprises-startup/le-patron-de-google-sundar-pichai-plaide-a-bruxelles-pour-une-approche-proportionnee-pour-reglementer-l-intelligence-artificielle-5e25d7baf20d5a3dbb86c6bc)
Facebook va créer 1000 emplois pour lutter contre les contenus "dangereux"	(/economie/digital/facebook-va-creer-1000-emplois-pour-lutter-contre-les-contenus-dangereux-5e26e465f20d5a3dbb86c792)
L'intelligence artificielle pour rendre les salons encore plus utiles	(/economie/entreprises-startup/l-intelligence-artificielle-pour-rendre-les-salons-encore-plus-utiles-61714e6ad8ad587c1ba41934)