

JOUER AUX MOTS CROISÉS

• DEEPFAKES

Sur Internet, la prolifération de « deepfakes » pornographiques embarrasse professionnels et régulateurs

En raison des nouveaux abus sexuels en ligne que la technologie « deepfake » permet, les entreprises se remettent aujourd'hui en question. Par ailleurs, les législateurs se penchent sur le sujet.

Par Aurélien Defer

Publié hier à 20h00 • Lecture 5 min.

Article réservé aux abonnés



D'après une étude néerlandaise, 96 % des deepfakes en circulation sur Internet sont des détournements pornographiques. SOLÈNE REVENÉY POUR « LE MONDE »

Au 21^e jour de la guerre en Ukraine, une vidéo, montée de toutes pièces et diffusée massivement en ligne, montrait Volodymyr Zelensky appelant ses troupes à « *rendre les armes* ». Si elle a été très vite démentie par le président ukrainien, elle a mis un nouveau coup de projecteur sur le *deepfake*, une technologie exploitant l'intelligence artificielle (IA) pour, par exemple, substituer dans une vidéo un visage par un autre. Cette technologie intéresse désormais les pouvoirs publics, qui redoutent la propagation virale de fausses nouvelles de plus en plus difficiles à repérer.

Pourtant, cet usage est largement minoritaire. Il faut, pour le comprendre, se pencher sur une autre tendance en ligne liée au deepfake : le remplacement, dans une vidéo pornographique, du visage de l'actrice par celui d'une autre femme. Cette pratique, parfois appelée « *deep-porn* », a gagné en popularité depuis son émergence, en 2017, et correspondait deux ans plus tard à 96 % des deepfakes trouvables sur Internet, d'après l'étude d'une entreprise néerlandaise de cybersécurité.

Une technologie intrinsèquement liée à la pornographie

DeepFaceLab est un outil de création de deepfakes. Ses concepteurs, qui ont publié en libre accès leur code sur la plate-forme GitHub, estiment que la quasi-totalité des vidéos deepfakes créées le sont aujourd'hui avec ce logiciel. Sur leur page, on y trouve de nombreux exemples : le visage du milliardaire Elon Musk intégré au film *Interstellar*, l'acteur Arnold Schwarzenegger rajeuni, le chef de l'Etat nord-coréen Kim Jong-un en plein laïus sur la préservation de la démocratie... Mais aussi, plus bas, un lien vers MrDeepFakes, un site pornographique.

On y trouve des milliers de vidéos et de photos détournées, représentant, pour la plupart, des femmes influentes, comme la comédienne britannique Emma Watson ou l'actrice israélienne Gal Gadot. Celle-ci, interprète de *Wonder Woman*, a été parmi les premières victimes de cette pratique, à laquelle elle n'a évidemment jamais consenti. Autre cible de choix, l'actrice américaine Scarlett Johansson avait réagi en 2018 auprès du *Washington Post* pour dénoncer un trucage humiliant et sexiste face auquel beaucoup de femmes, parfois sans notoriété publique, sont impuissantes : « *Rien ne peut empêcher quiconque de découper et de coller mon image sur celle de quelqu'un d'autre et de la rendre étrangement réaliste.* »

Lire le récit : [On a essayé de fabriquer un deepfake \(et on est passé à autre chose\)](#)

La forte fréquentation de ce site (13,6 millions de visites en juin, d'après l'outil SimilarWeb) et de dizaines d'autres espaces en ligne dédiés à ces fausses vidéos montre à quel point la technologie deepfake est liée à la pornographie. Comme l'a démontré une longue enquête du média américain *Vice* publiée le 17 mai, le logiciel DeepFaceLab et MrDeepFakes se renvoient régulièrement la balle : le premier encourage à s'entraîner à la manipulation de son code en détournant des films pornographiques, tandis que l'autre propose un guide détaillé d'utilisation du logiciel. Un article scientifique publié en 2020 sur la plate-forme de prépublication arXiv a même été cosigné à la fois par les développeurs de l'outil gratuit et par l'un des gestionnaires du site pornographique.

Les entrepreneurs cherchent à faire bonne figure

La timide médiatisation de ce type de cyberviolence a permis aux entrepreneurs qui investissaient dans le deepfake de passer jusqu'ici entre les gouttes. Récemment, la société de production audiovisuelle Metaphysic, à l'origine notamment des très innocents et très populaires détournements de Tom Cruise, a dû s'expliquer sur sa proximité avec le logiciel DeepFaceLab, et donc avec les contenus pornographiques qui pullulent sur Internet. De manière générale, alors que se multiplient à travers le monde les projets de loi pour réguler la technologie, nombre de chefs d'entreprise regrettent aujourd'hui que cette mauvaise réputation entache dans son entier une industrie en plein essor.

C'est le cas de Dima Shvets, à la tête d'une entreprise spécialisée dans la production de vidéos à l'aide de l'IA. Il a lancé en 2020 l'application Reface, qui permet à ses utilisateurs de remplacer un visage par un autre de manière réaliste. « *Nous investissons beaucoup de temps et d'efforts pour montrer que ces technologies peuvent être utilisées avec de bonnes intentions, et il est certain que cela serait plus facile s'il n'y avait pas autant de deepfakes pornographiques, de désinformation et de manipulation* », concède le directeur de l'application mobile auprès du *Monde*.

Même constat en Inde, où la production de deepfakes pornographiques a explosé ces dernières années. Ashray Malhotra, patron de Rephrase, une entreprise qui développe un outil de création

vidéo par l'IA, affirme que l'image sulfureuse de la technologie n'est plus un problème depuis quelques mois seulement, grâce à un long travail de sensibilisation et d'autorégulation. Conscient du succès de son application et d'une potentielle utilisation malveillante, Dima Shvets dit avoir employé dès le départ une modération humaine et un outil de vérification des contenus.

Lire notre enquête : [« Deepfakes » : faut-il avoir peur de ces vidéos trafiquées ?](#)

Certaines start-up établissent également des règles strictes pour les utilisateurs. Victor Riparbelli, spécialisé dans ce qu'il appelle les « médias synthétiques », a établi une liste de principes réglementaires pour son entreprise, Synthesia. « *Nous n'utilisons jamais l'image de quelqu'un sans son consentement, et notre logiciel ne peut pas être utilisé pour du porno* », insiste-t-il auprès du Monde. Selon lui, les mauvais usages du deepfake s'expliquent d'ailleurs en partie par la facilité avec laquelle toutes sortes d'utilisateurs y ont accès : « *Je suis pour l'open source dans la plupart des cas, mais pas pour cette technologie.* »

Vers une régulation plus sévère

Dernièrement, Google et Reddit ont discrètement pris leurs distances avec les deepfakes, y compris non pornographiques. Colaboratory, le service de code de l'entreprise de Larry Page, a annoncé début juin ne plus autoriser la production de telles images, tandis que Reddit a tout simplement fermé l'un de ses derniers forums communautaires sur le sujet. Ces revirements s'inscrivent dans un effort général de régulation de ces contenus, repris à l'échelon légal par certaines juridictions. Avec son Digital Services Act, l'Union européenne souhaite bientôt punir les plates-formes numériques qui ne modèrent pas les deepfakes trompeurs.

Lire aussi : [Vidéos truquées ou détournées : Facebook durcit ses règles de modération](#)

Le Royaume-Uni s'intéresse aussi à la question, cette fois en ciblant précisément les images pornographiques. Professeure de droit au King's College de Londres et membre de la commission des lois pour l'Angleterre et le Pays de Galles, Penney Lewis a travaillé sur le sujet de la criminalisation des abus sexuels en ligne depuis deux ans et a émis avec ses collègues un rapport et une série de recommandations le 7 juillet. « *Sur le deepfake, il y a trois choses à considérer : l'action de prendre l'image d'une personne sans qu'elle soit au courant ou consentante, l'action de créer des deepfakes sexuels, et l'action de partager ces images sans consentement ou de menacer de les partager* », analyse-t-elle auprès du Monde. Selon elle, c'est le dernier point qui doit être enregistré comme une infraction, la diffusion de ces images étant le plus violent pour les victimes. Mis en pause depuis la démission du premier ministre Boris Johnson, les débats sur le texte de loi Online Safety Bill, censé améliorer la sécurité des Britanniques sur Internet, pourraient intégrer d'ici à la fin de l'année ces recommandations sur les deepfakes, espère Penney Lewis.

Lire la tribune : [« Les avancées de l'intelligence artificielle pourraient rendre les "deepfakes" encore plus redoutables »](#)

Côté français, un tel arsenal légal n'est pas à l'ordre du jour, estime la députée Caroline Janvier (Renaissance). « *La crise sanitaire liée à la Covid-19 a quelque peu rebattu les cartes des urgences politiques* », justifie celle qui avait pourtant interpellé en 2019 le secrétaire d'Etat au numérique Mounir Mahjoubi au sujet des deepfakes, notamment pornographiques. Informée de l'ampleur du phénomène et des conséquences néfastes sur les victimes de ces abus, elle considère aujourd'hui que c'est à l'Union européenne de se saisir du dossier. Et à la France de suivre : « *Les mesures devront naturellement être traduites dans le droit français, et la législation en matière de harcèlement, d'usurpation d'identité ou encore de diffamation aura bien sûr vocation à être amendée pour prendre en compte ces nouveaux contenus.* »