Deepfake : Microsoft veut démasquer les fausses vidéos

La capacité de fournir de fausses images existe depuis un certain temps maintenant. Le film de 2016 « Rogue One: A Star Wars Story », par exemple, a vu les studios hollywoodiens redonner vie à Peter Cushing. Les films Star Wars suivants ont utilisé la technologie deepfake pour la princesse Leia après le décès de Carrie Fisher.

Mais au cours des années suivantes, les cas de deepfake sont devenus plus sinistres, des images de l'ancien président américain Barak Obama et du président américain Donald Trump étant utilisées dans diverses vidéos trompeuses.

Pour lutter contre les campagnes de désinformation, Microsoft propose un logiciel pour identifier les photos et vidéos deepfake avant les élections américaines.

Lutte contre la désinformation

Tom Burt, vice President « Customer Security & Trust » et Eric Horvitz, « Chief Scientific Officer » expliquent dans un billet de blog. «La recherche du professeur Jacob Shapiro de Princeton que nous avons soutenue , mise à jour ce mois-ci, a répertorié 96 campagnes d'influence étrangères distinctes ciblant 30 pays entre 2013 et 2019.».
Et de poursuivre : «Ces campagnes, menées sur les réseaux sociaux, visaient à diffamer des notables, à persuader le public ou à polariser les débats». «Des rapports récents montrent également que de la désinformation a été diffusée sur la pandémie de Covid-19, entraînant des décès et des hospitalisations de personnes recherchant des traitements supposés qui sont en fait dangereux.»

Pour lutter contre ces campagnes de désinformation, Microsoft veut protéger le vote via ElectionGuard et aider à sécuriser les campagnes et autres personnes impliquées dans le processus démocratique via AccountGuard.

Video Authenticator a été créé à l'aide d'un ensemble de données public de Face Forensic ++. Cette technologie a été développée par Microsoft Research et Microsoft Azure en partenariat avec le programme Defending Democracy. Il alimentera une initiative récemment annoncée par la BBC appelée Project Origin.

> *#Deepfakes no more. Behold, the Microsoft Video Authenticator, a tool that can analyze a still photo or video and provide a percentage chance that the media is artificially manipulated. (1/2) pic.twitter.com/llNud4lWmE*
>
> *— Microsoft On the Issues (@MSFTIssues) September 1, 2020*

Evaluer la véracité des vidéos en temps réel

«La désinformation prend de nombreuses formes et aucune technologie ne résoudra le défi d'aider les gens à déchiffrer ce qui est vrai et exact», écrit Microsoft. «Un problème majeur concerne les deepfakes qui sont des photos, des vidéos ou des fichiers audio manipulés par l'intelligence artificielle (IA) de manière difficile à détecter.» poursuit l'éditeur.

«Aujourd'hui, nous annonçons Microsoft Video Authenticator qui peut analyser une photo ou une vidéo fixe pour fournir un pourcentage de chance, ou un score de confiance, que le média soit manipulé artificiellement. Dans le cas d'une vidéo, il peut fournir ce pourcentage en temps réel sur chaque image pendant la lecture de la vidéo. » explique-t-il.

Microsoft explique que le logiciel fonctionne en détectant la limite de fusion des éléments profonds et subtils de décoloration ou d'échelle de gris qui pourraient ne pas être détectables par l'œil humain.

«Nous prévoyons que les méthodes de génération de supports synthétiques continueront de gagner en sophistication», a averti Redmond. «Comme toutes les méthodes de détection de l'IA ont des taux d'échec, nous devons comprendre et être prêts à réagir aux deepfakes qui passent à travers les méthodes de détection. Ainsi, à plus long terme, nous devons rechercher des méthodes plus solides pour maintenir et certifier l'authenticité des articles de presse et autres médias. »

Tom Jowitt, Silicon.co.uk