

Faux nus et réseaux sociaux : comment les algorithmes d'intelligence artificielle transcendent la réalité mais pas le droit

Le mois dernier, une société spécialisée en cybersécurité a publié un rapport dévoilant un réseau de partage sur le réseau social Télégram de photographies de femmes dénudées par un logiciel automatisé.

Environ 100.000 femmes se sont vues déshabillées virtuellement sur photographie par des techniques de deepfake. Ces photographies ont pour la plupart été utilisées par leurs auteurs anonymes à des fins de chantage à l'encontre des personnes concernées.

Heureusement, ces nouveaux outils techniques n'empêchent pas le recours à des mécanismes juridiques pour faire cesser la diffusion a posteriori de ces contenus.

La technologie du deepfake

Un deepfake est une technique d'intelligence artificielle permettant à l'aide de synthèse d'images de superposer des fichiers entre eux (audio, vidéo, images etc.). Le concept est populaire sur les réseaux sociaux, qui proposent pour la plupart des filtres de reconnaissance faciale permettant de mixer le visage ou la voix de deux personnes par exemple. Le deepfake est souvent utilisé pour des trucages, canular ou à des fins humoristiques par des imitateurs notamment qui prennent l'aspect physique de la personne dont la voix est imitée.

Il est apparu que cette technique est également souvent utilisée à des fins pornographiques, représentant des personnes à leur insu. Des sextapes (vidéos érotiques) ou revenge porn (pornodivulgateur) ont mis en scène le visage de personnes associées à des corps qui n'étaient pas les leur, laissant alors penser qu'elles étaient actrices de ces vidéos. La plupart du temps, ces vidéos sont publiées sur internet sans le consentement de la personne ou envoyées à la personne concernée directement pour effectuer un chantage.

Si le montage et la retouche d'images ne sont pas nouveaux, le réalisme permis par l'intelligence artificielle devient problématique. En effet, il est difficilement possible d'identifier si une image de nus est un faux et donc d'identifier si elle porte préjudice à la personne concernée.

Législation applicable

La technique de photomontages a toujours existé ainsi que le préjudice lié. Quelle que soit la technique utilisée pour y arriver et effectuer une publication, les outils juridiques sont les mêmes.

D'une part, le Code pénal prévoit une peine d'un an d'emprisonnement et de 15.000 euros d'amende le fait de publier, par quelque voie que ce soit, le montage réalisé avec les paroles ou l'image d'une personne sans son consentement, s'il n'apparaît pas à l'évidence qu'il s'agit d'un montage ou s'il n'en est pas expressément fait mention(1).

En outre, la loi de 1881 sur la liberté de la presse prévoit également une infraction de diffamation, susceptible d'engager la responsabilité pénale de son auteur, sur la base de laquelle une action civile est souvent engagée. La diffamation étant définie comme « toute allégation ou imputation d'un

fait qui porte atteinte à l'honneur ou à la considération de la personne ou du corps auquel le fait est imputé est une diffamation », une action sur ce fondement en cas de publications de photomontages pornographiques serait légitime en tant qu'atteinte au corps.

Néanmoins, les condamnations à ce titre se font rares. Pourtant elles existent, et les tribunaux ont considéré que les photomontages peuvent bien constituer une atteinte à la réputation et à l'image sur le fondement de la loi de la presse de 1881 (2).

Sur les réseaux sociaux, la publication d'images controversées et de photomontages est la plupart du temps publique. Il n'en demeure pas moins qu'en cas de publication privée de telles photographies, la diffamation pourrait être caractérisée, cette fois en tant que contravention et non en tant que délit.

D'autre part, sur le plan civil, il est également possible d'envisager une atteinte à l'image voire aux données personnelles par la diffusion sans consentement d'une photographie de la personne concernée.

Concernant les mineurs en particulier, le Règlement européen pour la protection des données personnelles (RGPD) prévoit que le consentement des parents d'un mineur âgé de moins de 16 ans est nécessaire avant de publier la photographie du mineur sur internet. La loi pour une république numérique ajoute que les mineurs disposent d'un droit à l'oubli accéléré, leur permettant de demander l'effacement de données problématiques sur des plateformes en ligne, dans les meilleurs délais.

Néanmoins, les tribunaux peinent à trouver une base légale leur permettant de sanctionner les auteurs de publication de photomontages concernant les mineurs. A titre d'exemple, il a été jugé que le fait d'adresser à un père, par courriel privé, des photomontages à caractère pornographique de sa fille mineure ne peut pas constituer une faute au sens de l'article 1382 du Code civil (nouvel article 1240). En revanche, le fait d'accompagner ce courriel de menaces constitue une faute au sens de l'article 1382 du Code civil, justifiant le dédommagement du préjudice moral du père (3).

Responsabilités et conséquences

Par principe, les auteurs de photomontages engagent leur responsabilité à la diffusion de ladite photographie, quel que soit le support. Néanmoins, en pratique, cette responsabilité peut s'avérer difficile à actionner, les auteurs agissant sur les réseaux sociaux notamment sous couvert d'anonymat.

Il est donc possible dans un second temps et en parallèle d'actionner la responsabilité de l'éditeur du site, voire même de l'hébergeur si ce dernier a été informé du contenu manifestement illicite. L'éditeur dispose d'une obligation générale de surveillance du contenu de son site, alors que l'hébergeur bénéficie d'un régime de responsabilité allégé. Si la distinction entre la définition de l'hébergeur et de l'éditeur génère de multiples débats en jurisprudence, il est à retenir que les juges considèrent la plupart du temps que les réseaux sociaux (comme Facebook ou Télégram) agissent comme hébergeurs.

Néanmoins, du fait des algorithmes de modération automatisés mis en place ces dernières années par les réseaux sociaux pour détecter des nus ou images violentes, la question du statut des réseaux sociaux pourrait se poser de nouveau, et les juges pourraient renforcer leur responsabilité pour les ériger comme éditeur.

En tout état de cause, la personne concernée pourra demander à l'auteur et/ou à l'éditeur de la publication et/ou à l'hébergeur de la publication le retrait du photomontage, ainsi qu'un

dédommagement au titre du préjudice encouru du fait de la publication.

En pratique et si l'on est victime d'une telle publication, il est fortement conseillé à la personne concernée par le deepfake, dont un photomontage aurait été publié ou partagé sur un réseau social, dans la mesure où la photographie porte atteinte à ses droits, d'en demander un retrait sans délai à l'auteur et/ou au réseau social.

1 Article 226-8 du Code pénal

2 Cour de cassation, 1ère civ., 30 Mai 2006 - n° 04-18.520

3 Cour d'appel, Versailles, 1ère chambre, 1ère section, 3 Novembre 2016 - n° 14/06867