Edition Multimédi@: Pour la reconnaissance faciale à distance ou locale, les enjeux éthiques ne sont pas les mêmes

Winston Maxwell, David Bounie

▶ **To cite this version:**

Winston Maxwell, David Bounie. Edition Multimédi@: Pour la reconnaissance faciale à distance ou locale, les enjeux éthiques ne sont pas les mêmes. 2020. hal-02501056

HAL Id: hal-02501056

https://telecom-paris.hal.science/hal-02501056

Submitted on 12 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pour la reconnaissance faciale à distance ou locale, les enjeux éthiques ne sont pas les mêmes

Identifier un visage dans une foule soulève de sérieuses questions sur les libertés individuelles. Mais il existe de nombreux autres usages de la reconnaissance faciale, notamment la validation d'identité en local. Ces utilisations ont vocation à se développer mais posent d'autres questions éthiques.

Par Winston Maxwell* et David Bounie**, Telecom Paris, Institut polytechnique de Paris

L'utilisation de la reconnaissance faciale pour l'identification à distance constitue une menace pour les libertés individuelles, car cela tend à banaliser une société de surveillance. Selon le *New York Times,* une start-up américaine Clearview AI a déjà fabriqué des gabarits d'identification de 3 milliards d'individus à partir d'images copiées sur le Web (1). N'importe quelle force de l'ordre – mais pas le grand public (2) – peut utiliser le logiciel de Clearview AI et identifier les visages dans une foule. Cependant, plusieurs villes américaines ont temporairement banni cette utilisation de la technologie par leurs autorités publiques.

Outils de surveillance généralisée

En Europe, la Commission européenne appelle à un grand débat européen sur l'utilisation de la reconnaissance faciale pour l'identification à distance. En France, le secrétaire d'Etat au numérique, Cédric O, souhaite lancer des expérimentations. Pour l'identification à distance, il faut avancer à tâtons pour trouver le bon équilibre entre les impératifs de sécurité publique et la préservation des valeurs démocratiques. Mais ce débat n'est pas différent au fond de celui qui, depuis 50 ans, entoure les technologies de surveillance des communications électroniques. La technologie utilisée pour la surveillance des communications n'a pas cessé d'évoluer : IMSI-*catchers* ou intercepteurs d'IMSI (3), boîtes noires, Deep Packet Inspection (DPI), captation à distance, ... Ces outils permettraient une surveillance généralisée de la population. Leur utilisation en France est interdite, sauf par les forces de polices et des autorités de renseignement sous le contrôle de juges et de la CNCTR (4).

En application de la jurisprudence européenne, l'utilisation de technologies invasives de surveillance par l'Etat se justifie uniquement si l'utilisation est prévue par une loi. Et ce, pour faire face à une menace particulièrement grave, la lutte contre le terrorisme par exemple, et sous le contrôle d'un juge ou d'une commission indépendante. L'utilisation de la reconnaissance faciale pour identifier les individus à distance devrait suivre la même trajectoire : interdiction, sauf pour les autorités de police ou de renseignement sous le contrôle des juges. D'ailleurs, c'est déjà ce qui est prévu par la directive européenne dite « Police-Justice » (5) de 2016, puisque la reconnaissance faciale est un traitement biométrique soumis à des règles strictes. Mais il existe un deuxième type d'utilisation, non-évoqué par la Commission européenne dans son livre blanc (6) sur l'intelligence artificielle (IA). Il s'agit de valider l'identité « en local » d'un individu en comparant sa photo « selfie » avec la photo de la pièce d'identité. Cette utilisation permet notamment d'ouvrir un compte bancaire à distance ou bien de passer plus vite dans un portique automatique à l'aéroport. Cette utilisation de la reconnaissance faciale se généralise, et elle paraît – de prime abord – moins attentatoire aux libertés individuelles : d'une part, parce que les personnes sont conscientes et consentantes de l'utilisation (ce qui n'est pas le cas pour l'identification à distance) ; d'autre part, parce qu'aucune image ni gabarit biométrique n'est stocké de manière centralisée. La vérification s'effectue en local, comme pour déverrouiller un smartphone avec l'empreinte digitale. Le système crée un gabarit biométrique à partir de la photo du passeport, analyse ensuite la photo de selfie, crée un deuxième gabarit biométrique du selfie, et compare les deux gabarits pour établir une probabilité de correspondance. Ensuite les gabarits sont détruits *(lire encadré page suivante).*

La reconnaissance faciale locale soulève néanmoins des questions éthiques et juridiques importantes : l'existence d'un consentement libre, le problème des biais, l'explicabilité des algorithmes, et la difficile articulation avec le règlement général sur la protection des données (RGPD) pour la phase d'entraînement. La reconnaissance faciale « locale » pose la question du consentement libre. Si la personne subit des conséquences négatives en refusant la reconnaissance faciale, le consentement ne sera pas libre. Il en sera de même si le consentement est demandé par une personne jouissant d'une position d'autorité, par exemple si la direction d'un lycée demandait aux élèves de consentir à l'utilisation de la reconnaissance faciale pour rentrer dans l'établissement (7).

Les biais statistiques sont inévitables

Concerne les biais cette fois, le Parlement européen a appelé le 12 février 2020 à l'utilisation d'algorithme qu'il faut entraîner avec des données « *non-biaisées* » (8). Or, une telle condition est impossible à satisfaire en pratique. Certains groupes de la population seront toujours sous-représentés dans les images d'entraînement, ce qui signifie que les biais

Notes

(1) - https://lc.cx/NYT-ClearviewAI

(2) - « *Clearview Is Not A Consumer Application* » : https://lc.cx/Clearview23-01-20

(3) - International Mobile Subscriber Identity (IMSI).

(4) - Commission nationale de contrôle des techniques de renseignement (CNCTR).

(5) - Directive européenne n° 2016/680 du 27 avril 2016, dite directive « Police-Justice » : https://lc.cx/PoliceJustice

(6) - https://lc.cx/LivreBlancIA

(7) - Voir jugement du 27-02-20 à Marseille : https://lc.cx/JugementTA-LQDN

La photocopie non autorisée est un délit. Pour les abonnements : 01 39 15 62 15 (EMG). Pour les reproductions : 01 44 07 47 70 (CFC).

statistiques seront inévitables. Cela peut conduire à des niveaux de performance inégaux selon le genre, la couleur de peau ou la situation de handicap d'une personne. Par exemple, l'algorithme pourrait avoir plus de difficulté à identifier une femme noire qu'un homme blanc au moment de la vérification de l'identité à l'aéroport. Ces biais peuvent exister sous une forme bien pire chez les humains. Mais pour un algorithme, ce genre de biais est peu acceptable. Corriger ces biais dans l'algorithme est possible, mais cela soulève d'autres questions. Par exemple, si l'algorithme a un taux d'erreur élevé pour des personnes atteintes d'une certaine maladie de la peau, devons-nous baisser artificiellement le niveau de performance pour tous les autres groupes de la population pour que le taux d'erreur soit équivalent ? Ces questions deviennent rapidement politiques : à partir de quel moment un biais algorithmique devient-il suffisamment problématique pour le corriger, ce qui affectera inévitablement la performance de l'algorithme pour les autres personnes ?

Savoir s'il y a discrimination algorithmique

Un autre aspect éthique de la reconnaissance faciale concerne l'explicabilité des algorithmes. En France, le code des relations entre le public et l'administration garantit à chaque individu le droit d'obtenir une explication lorsqu'un algorithme géré par l'Etat prend une décision à son encontre (9). Logiquement, ce droit exige que l'exploitant de l'algorithme soit en mesure d'expliquer à une personne pourquoi un système n'a pas pu vérifier son image par rapport à sa pièce d'identité. Techniquement, des solutions d'explicabilité existent, même pour des réseaux de neurones. Mais fournir une explication exige le stockage d'informations, et notamment les gabarits générés par l'algorithme. Or, le RGPD et la directive « Police-Justice » interdisent généralement ce stockage, surtout lorsqu'il s'agit de données biométriques.

Résultat : dans certains cas, il n'y aura aucune explication quant au refus du système de vérifier l'identité. Le système ne réussira pas à identifier la personne, sans que la personne puisse vérifier si elle a fait l'objet d'une discrimination algorithmique. Cette absence de transparence pose une difficulté au niveau des droits fondamentaux, comme le démontre une récente décision du tribunal de la Haye (10).

Enfin, l'entraînement des algorithmes de reconnaissance faciale est difficile à réconcilier avec le RGPD. Pour réduire les discriminations, l'Agence européenne des droits fondamentaux (FRA) souligne la nécessité d'entraîner l'algorithme sur une grande quantité d'images représentatives de la population, et notamment les personnes vulnérables (11). Or cette condition est quasiment impossible à remplir en Europe puisque le RGPD et la directive « Police-Justice » interdisent la création de grandes bases d'images, surtout lorsque les images sont étiquetées selon la couleur de peau ou la situation de handicap. Les systèmes américains et chinois bénéficient, eux, d'entraînement sur des dizaines de millions d'images, ce qui crée un avantage concurrentiel considérable. De plus, les tests de non-discrimination des algorithmes s'effectuent tous aux Etats-Unis à l'agence NIST (12), même pour les systèmes européens.

L'entraînement des algorithmes pose un problème particulier puisque le gabarit d'un visage est considéré comme une donnée biométrique. Or le RGPD interdit le traitement de données biométriques, hormis des cas limités – par exemple, le consentement explicite de la personne. Du coup, un entraînement sur des millions d'images récupérées sur Internet devient impossible par une société européenne puisque l'entraînement nécessite la création, au moins temporaire, de gabarits, une donnée biométrique. Une solution pourrait consister en l'assouplissement des conditions d'application du RGPD lorsqu'il s'agit de créer des gabarits éphémères pour l'apprentissage des algorithmes dans des environnements contrôlés, voire de considérer que ces gabarits ne sont pas des données biométriques puisque la finalité de leur traitement n'est pas l'identification d'une personne mais seulement l'entraînement de l'algorithme. Lorsque l'algorithme est mis en exploitation, les dispositions du RGPD ou de la directive « Police-Justice » sur la biométrie retrouveraient toute leur force, puisque les gabarits seraient bien utilisés pour identifier des personnes. Le consentement explicite de la personne, ou en cas d'intérêt public et de nécessité absolue, serait alors nécessaire. @

* Winston Maxwell, ancien avocat, est depuis juin 2019 directeur d'études Droit et Numérique à Telecom Paris. ** David Bounie est directeur du département Economie et Sciences sociales à Telecom Paris.

Notes

(8) - https://lc.cx/EurodéputésIA2020

(9) - Article L311-3-1 du code des relations entre le public et l'administration : https://lc.cx/L311-3-1

(10) - Le 5 février 2020, le tribunal de district de la Haye a invalidé l'utilisation par l'Etat néerlandais d'un algorithme destiné à lutter contre la fraude à la sécurité sociale. NJCM c. The Netherlands, District Court of The Hague, Case n° C-09-550982-HA ZA 18-388.

(11) - https://lc.cx/FRArecognition 27-11-19

(12) - National Institute of Standards and Technology (NIST).

Zoom

Qu'est-ce qu'un gabarit ?

Un gabarit est l'équivalent d'un code barre qui contient les mensurations uniques d'un visage. La clé du succès en matière de reconnaissance faciale est de créer un algorithme capable de générer des gabarits de qualité à partir d'images d'individus.

Un algorithme de qualité doit savoir générer le même gabarit pour l'image de Paul, quelles que soient les différences de lumière, d'angle de vue et de netteté de l'image de Paul. Pour entraîner l'algorithme, on va présenter à un réseau de neurones 100 photos d'une même personne — par exemple Angelina Jolie — récupérées sur le Web, avec des angles de vue et des lumières différents, et demander au réseau de neurones de trouver une formule mathématique qui permettra pour chaque photo d'Angelina Jolie de générer le même gabarit, quels que soient l'angle de vue ou la lumière. Les gabarits générés pendant l'apprentissage sont éphémères. Ils servent uniquement à aider l'algorithme à trouver la bonne formule mathématique. Une fois cette formule mathématique unique établie, elle peut s'appliquer à n'importe quelle nouvelle photo de passeport et générer, pour cette photo, un gabarit de qualité. L'algorithme va ensuite générer un deuxième gabarit à partir d'une deuxième photo (par exemple un selfie), et si l'algorithme est bien fait, les deux gabarits vont correspondre malgré les différences d'angle de vue et de lumière. La qualité de cet algorithme est au cœur des systèmes de reconnaissance faciale. @

La photocopie non autorisée est un délit. Pour les abonnements : 01 39 15 62 15 (EMG). Pour les reproductions : 01 44 07 47 70 (CFC).