

DIFFUSION IMMEDIATE

Un million d'images à vérifier manuellement

Pour diminuer les biais de la reconnaissance faciale, il est nécessaire de travailler avec des banques d'images représentatives. Afin de s'assurer de la cohérence de la démarche, l'Idiap participe de façon éthique à la création de tels stocks d'images.

Si vous êtes un homme blanc, les algorithmes seront généralement performants pour reconnaître votre visage. En revanche, les taux de réussite diminuent dès lors que vous êtes une femme ou une personne à la couleur de peau différente. Ces biais ne signifient pas que l'intelligence artificielle est programmée pour être discriminante. Ils viennent du fait que les programmes utilisés sont entraînés à reconnaître des visages grâce à des bases de données d'exemples qui ne sont pas représentatifs. Pour pallier cette inégalité et améliorer la sécurité des systèmes de reconnaissance faciale, le groupe de recherche Biometrics Security and Privacy crée une nouvelle base de données plus fiable. Cette démarche s'inscrit dans le cadre d'un partenariat avec une entreprise active dans le domaine de la sécurité.

Ethique et confidentialité

« Habituellement, ce genre de travail d'annotation d'images – pour produire une métadonnée telle que par exemple le genre ou couleur des yeux – est sous-traité sur des plateformes en ligne de partage de travail, comme Amazon Mechanical Turk (MTurk), » explique Sébastien Marcel, responsable du groupe de recherche. « Chacun peut y réaliser une petite part du travail contre une petite rémunération. Outre l'uberisation du travail qui en résulte, cela pose des problèmes de confidentialité. » En effet, la base de données créée à l'Idiap est réalisée à l'occasion d'un projet avec un partenaire industriel. Pour des questions de sécurité, l'accord conclu spécifie que les données ne peuvent pas être distribuées et doivent rester en interne.

La fiabilité contre les biais

Dans un bureau à l'Idiap, ils sont quatre devant leurs écrans à comparer des séries d'images pour les valider. « Le plus difficile, c'est de rester concentré, » lance Magali. « Pour y parvenir, il faut souvent faire des pauses, » ajoute Josselin. « Chaque heure, » précise Oriane. A raison de huit heures par jour, durant une ou deux semaines, chacun apporte sa pierre à l'édifice. Un travail de longue haleine, très demandant. « C'est parfois difficile avec certaines photos ou certain types de personnes, » explique Léo.

Le coût d'un tel travail est important, de l'ordre de 20'000 francs. Les jeunes découvrent ainsi l'Idiap et les enjeux de l'annotation d'images tout en étant rémunérés à 20 francs de l'heure. Le même travail proposé par des sociétés étrangères privées peut être jusqu'à 20 fois moins cher. « Le fait de réaliser nous-mêmes ce travail nous permet aussi de vérifier de façon plus rapide la qualité de notre base de données, » précise Sébastien Marcel. « Plus une base de données est "propre", plus le résultat de l'apprentissage automatique du programme sera fiable. »

Plus d'informations

- Groupe de recherche Biometrics Security and Privacy: <https://www.idiap.ch/en/scientific-research/biometrics-security-and-privacy>
- Swiss Center for Biometrics Research and Testing: <https://www.biometrics-center.ch/>

Contacts

Sébastien MARCEL, responsable du Swiss Center for Biometrics Research and Testing de l'Idiap Research Institute, sebastien.marcel@idiap.ch, +41 27 721 77 27