

RETEX CAS CONCRETS

Datakalab fait tourner ses algos en Edge

La jeune pousse française Datakalab parvient à exécuter des algorithmes de reconnaissance d'images sur des CPU, et non des GPU, en Edge. Et ce, sans perdre en précision et sans envoyer de données personnelles dans un quelconque Cloud.

Quand on parle de réseaux neuronaux, d'algorithmes capables de savoir si vous portez un masque ou non et de flux vidéo multiples, on pense généralement à de grosses machines, bien équipées en GPU et situées à plusieurs milliers de kilomètres. Donc à un Cloud, en d'autres termes. On pensera aussi violation de la vie privée et bilan énergétique catastrophique. Tels sont les écueils dans lesquels Datakalab n'est pas tombé !



L'algorithme de Datakalab détecte si oui ou non les individus à l'écran portent un masque ; et c'est cet attribut que reçoit le client. Il ne dispose ni des images, ni des données personnelles.

Cette jeune pousse française est spécialisée dans la reconnaissance d'images par ordinateur. Parmi ses faits d'armes, elle a travaillé avec la RATP à mesurer le port du masque dans ses stations de métro et avec l'AP-HP et le ministère de la Culture lors du concert test organisé le 29 mai à l'Accor Arena de Bercy, afin là encore de déterminer si et quand les participants ne portaient pas leurs masques. Et ce à des fins purement statistiques : hors de question d'identifier les individus. Mieux encore, aucune donnée personnelle ne sort.

Au plus près du capteur

Datakalab opère en périphérie en déployant ses machines au plus près des caméras utilisées. « Nous faisons tourner nos algorithmes de reconnaissance d'images sur des CPU en Edge. C'est mieux du point de vue de la Cnil et cela consomme moins d'énergie », nous explique Xavier Fisher, le CEO de Datakalab. Pour ce faire, la start-up s'est rapprochée de Lenovo ISG et utilise son matériel comme le ThinkSystem SR650 Server. Lequel supporte jusqu'à deux processeurs pour un total de 56 cœurs Intel Xeon doté de la technologie Hyper-Threading. « Lenovo est une des premières entreprises

qui a compris ce que nous faisons. Ses solutions hardware sont CPU-oriented, sécurisées et adaptées à nos besoins. »

Schématiquement, le serveur sur lequel l'algorithme tourne reçoit les flux vidéo, les images sont transformées en data, en batch sur la mémoire vive de l'appareil et analysée : quel objet est détecté à quel endroit ; suivi d'une détection des attributs, par exemple « Porte un masque » / « Ne porte pas de masques ». Avant que ces petits paquets de données soient envoyés via des proxies chez le client afin qu'il en fasse des dashboards. « Ce que nous faisons, c'est de l'accélération matérielle pour faire plus de calculs », souligne Xavier Fisher, « tout en changeant de paradigme sur les opérations matricielles : nous parvenons à diviser par cent le nombre de calculs nécessaires », sans perdre en précision. Là où en Edge classique un algorithme atteint 32 mAP (maximum a posteriori probability), celui de Datakalab arrive à 51 mAP, frôlant le niveau des IA exécutées sur un Cloud (55 mAP). « Notre activité consiste à créer du software ayant les performances des « gros » algos sur de petites machines », note le CEO. Une prouesse permise par une solide équipe de chercheurs dirigés par Kevin Bailly et Arnaud Dapogny, deux ingénieurs de l'Institut des systèmes intelligents et de robotique (Isir), et par une recette maison !

Algo écolo

C'est par ce terme que la start-up désigne sa technologie, un secret bien gardé. Et pour cause... « Habituellement, nous mettons en place d'importants réseaux de neurones, très précis mais qui ne peuvent fonctionner sur du CPU. On cherche à ce que cela fonctionne sans prendre en considération les conséquences écologiques ou sur la vie privée », indique Xavier Fisher, qui se targue de la validation par la Cnil du dispositif mis en place pour le concert test de mai dernier. « Il n'est pas nécessaire d'envoyer des données personnelles sur Internet, ni même de stocker des images. » Sauf pour 100 millisecondes sur la RAM de la machine. « Notre objectif est de fournir une alternative qui protège les individus tout en fournissant des analytics », précise-t-il. Un principe bien loin du scoring social et qui a en plus le mérite d'avoir une démarche de respect de l'environnement. D'ailleurs, interrogé quant à la possibilité d'utiliser des GPU en Edge, Xavier Fisher continue d'y préférer les CPU, « moins chers et moins consommateurs ». □

Guillaume Périssat