

la détection d'objets en temps réel pour microcontrôleurs grâce à Edge Impulse FOMO

Jan Jongboom, Edge Impulse

Nous, les humains, dépendons largement de la vue pour accomplir de nombreuses tâches quotidiennes, des plus simples aux plus complexes. D'un seul coup d'œil, nous savons s'il y a des gens autour de nous, s'il y a un éléphant à proximité ou combien de places de parking sont disponibles. En dépit de l'importance de la vision, de nombreux appareils embarqués n'ont pas la capacité de la perception visuelle. Ne serait-il pas merveilleux de pouvoir apprendre à tous nos appareils à voir le monde comme nous le voyons ?

Ces dernières années, la vision par ordinateur a connu des développements surprenants, qui ont permis des avancées dans des domaines tels que les voitures autonomes et les systèmes d'authentification biométrique (très utiles si, comme moi, vous voyagez beaucoup !). Mais ces cas d'utilisation sont incroyablement gourmands en calcul, car ils nécessitent des processeurs graphiques coûteux ou des accélérateurs spéciaux pour fonctionner.

La bonne nouvelle est que les tâches de vision par ordinateur ne nécessitent pas toujours un calcul aussi puissant. Toute question fermée (« Est-ce que je vois un éléphant ? », « Est-ce que cette étiquette est bien collée sur la bouteille ? ») peut apporter une valeur considérable aux dispositifs embarqués limités. De plus, ces problèmes de classification d'images peuvent même

être résolus par les microcontrôleurs modernes.

Et si nous pouvons ajouter des capacités de vision encore plus avancées à chaque dispositif embarqué ?

Classification result

Summary	
Name	Parking_data_244.png.2tgq40ai
CATEGORY	COUNT
F1 score	91.43%
car	16

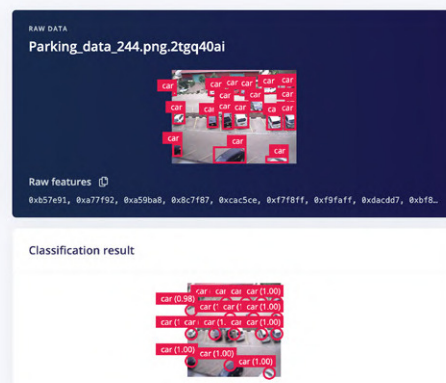
Figure 1. Classification FOMO dans Edge Impulse Studio.

Bonjour FOMO

Nous concrétisons cette idée. Nous avons développé une nouvelle architecture de réseau neuronal pour la détection d'objets, appelée Faster Objects, More Objects, ou FOMO (figure 1). Elle a été conçue dès le départ pour fonctionner en temps réel sur des microcontrôleurs, de sorte que les ingénieurs embarqués peuvent éviter la peur de manquer le coche en matière de vision par ordinateur.

Rapide, léger et flexible

FOMO est capable de fonctionner sur un microcontrôleur à 32 bits, comme un Arm Cortex-M7, avec une fréquence de 30 images par seconde. Et si vous optez pour un Raspberry Pi 4, vous serez en mesure de détecter des objets à une fréquence d'environ 60 images par seconde. C'est à peu près 30 fois plus rapide que



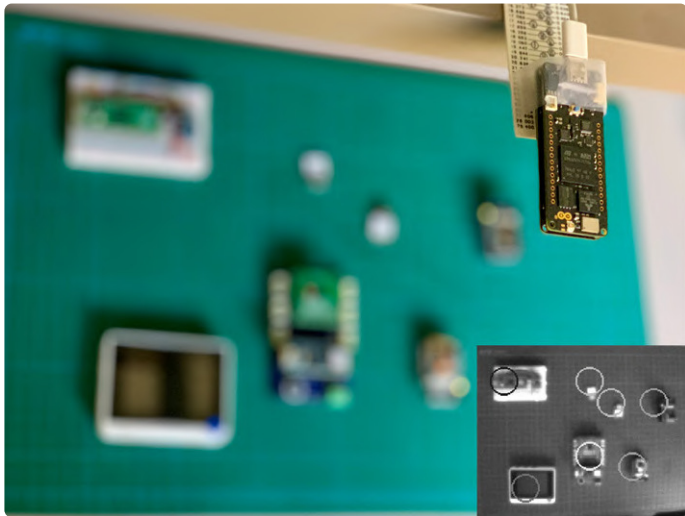
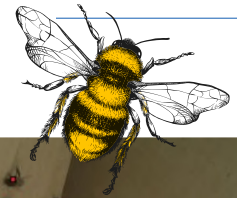


Figure 2. Exécution de la détection d'objets sur une grande variété de cartes de développement, y compris l'Arduino Portenta.



Figure 3. Voici une ancienne itération de l'approche FOMO utilisée pour compter les abeilles individuelles.

MobileNet SSD ou YOLOv5.

La capacité de FOMO est réduite à environ 100 Ko de RAM, ce qui permet d'exécuter les applications de détection d'objets en temps réel sur des cœurs Arm Cortex-M4 très contraignants, sur des cœurs plus puissants, comme les noyaux Cortex-M7 de l'Arduino Portenta H7 (figure 2), sur le nouvel Arduino Nicla Vision (un autre processeur double Arm Cortex-M7/M4), ou même sur des DSP spécialisés comme le WE-I de Himax.

FOMO peut s'adapter aux plus petits microcontrôleurs tout comme aux passerelles ou aux processeurs graphiques. Ce haut degré de flexibilité le rend également utile lorsque la détection de défauts nécessite d'identifier des variations les plus petites dans une image.

Dans un microcontrôleur dont la capacité de calcul et de mémoire est strictement limitée, il est préférable d'utiliser une taille d'image d'environ 96x96 pixels. Mais avec un microcontrôleur plus puissant, une image de 160x160 pixels est probablement adéquate. Le plus important est que FOMO est entièrement convolutif, il fonctionne donc avec n'importe quelle taille d'entrée arbitraire. Si vous souhaitez avoir une plus grande granularité, plus de détails ou plus d'objets, vous pouvez simplement augmenter la résolution d'entrée.

Il perçoit les petits détails

Cette nouvelle architecture peut même repérer et compter de nombreux objets

très petits de manière très efficace tant que les éléments présents dans le cadre sont de taille similaire et ne se chevauchent pas (figure 3). C'est quelque chose que le MobileNet SSD et le YOLOv5 ne peuvent pas réaliser très bien, malgré le fait qu'ils soient des modèles plus grands et plus performants.

Fini les ratés

FOMO est disponible dès aujourd'hui, fonctionne sur une grande variété de plateformes informatiques et est compatible avec les systèmes Linux, les microcontrôleurs Cortex-M et les DSP spécialisés. Il suffit d'ajouter une caméra et d'utiliser Edge Impulse, et le tour est joué. Avec FOMO, vous pouvez rapidement

intégrer la détection d'objets à n'importe quel appareil doté d'une caméra. Ne ratez plus ce que les ingénieurs en systèmes embarqués ont dû affronter jusqu'à présent en matière de vision par ordinateur (figure 4).

Pour en savoir plus sur FOMO et expérimenter votre propre algorithme, rendez-vous sur edgeimpulse.com/fomo. ◀

220207-04



Jan Jongboom est un ingénieur de systèmes embarqués et un défenseur de l'apprentissage automatique,

toujours en quête de moyens de collecter davantage d'informations dans le monde réel. Il a livré des appareils, travaillé sur les dernières technologies de réseau, simulé des microcontrôleurs. Il y a même un monument à San Francisco qui porte son nom. Il est actuellement cofondateur et directeur technique de Edge Impulse, la principale plateforme de développement pour l'apprentissage automatique embarqué, qui compte plus de 80 000 projets.

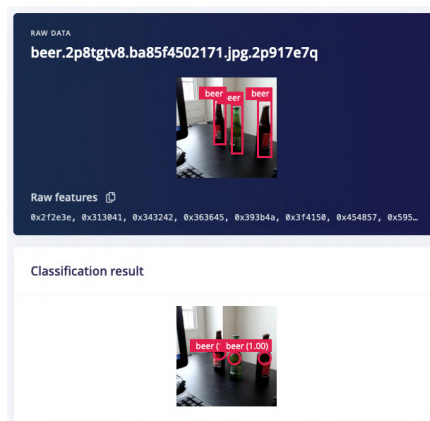


Figure 4. Formation sur les centroïdes de bouteilles de bière. En haut les étiquettes sources, en bas le résultat de l'inférence.