

Edge Impulse FOMO

real-time objectdetectie voor MCU's

Jan Jongboom, Edge Impulse

Wij mensen zijn sterk afhankelijk van ons gezichtsvermogen bij de uitvoering van onze dagelijkse taken – van de simpelste tot de meest ingewikkelde. Met één blik weten we of er mensen bij ons in de kamer zijn, of er een olifant in de buurt is, of hoeveel vrije parkeerplaatsen er zijn. Ondanks het belang van het gezichtsvermogen kunnen veel embedded apparaten visueel niets waarnemen. Zou het niet geweldig zijn als we al onze apparaten konden leren de wereld te zien zoals wij dat doen?

De afgelopen jaren zijn er verbazingwekkende ontwikkelingen geweest op het gebied van computer vision, die hebben bijgedragen aan de vooruitgang in bijvoorbeeld zelfrijdende auto's en biometrische immigratiepoorten (erg handig als u, zoals ik, veel reist!). Maar deze toepassingen zijn ongelooflijk rekenintensief en vereisen dure GPU's of speciale versnellers om ze uit te voeren.

Het mooie is dat niet alle computer vision-taken zoveel rekenkracht vereisen. Elke ja/nee-vraag ("Is dat een olifant?", "Zit het etiket goed op deze fles?") kan een enorme waarde toevoegen aan relatief eenvoudige embedded apparaten. Bovendien kunnen deze beeldclassificatie-problemen zelfs worden opgelost door huidige microcontrollers.

Stelt u zich voor dat we nog meer geavanceerde vision-mogelijkheden konden toevoegen aan embedded apparaten!

Classification result

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

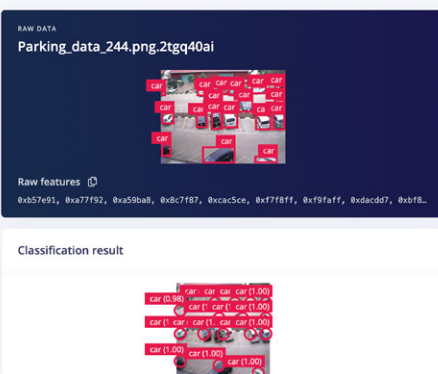
Figuur 1. FOMO-classificatie binnen Edge Impulse Studio.

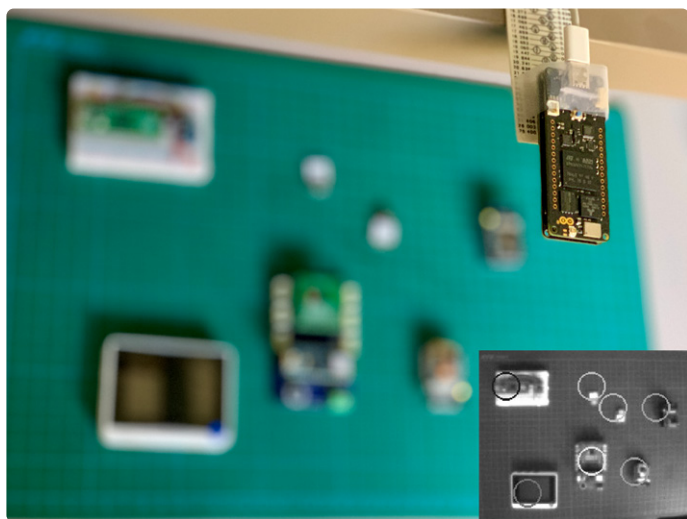
Zeg hallo tegen FOMO

Wij maken dat werkelijkheid. We hebben een nieuwe neurale netwerkarchitectuur ontwikkeld voor objectdetectie: *Faster Objects, More Objects* – kortweg FOMO (figuur 1). Het is van meet af aan ontworpen om in werkelijke tijd op microcontrollers te draaien, zodat embedded engineers niet meer bang hoeven zijn (ahem) om op het gebied van computer vision de boot te missen.

Snel, slank en flexibel

FOMO draait op een 32-bit MCU, zoals een Arm Cortex-M7, met een frame rate van 30 beelden per seconde. En wanneer u voor een Raspberry Pi 4 of vergelijkbaar kiest, is objectdetectie met een snelheid van tegen de 60 frames per seconde mogelijk. Dat is ruwweg 30 maal sneller dan MobileNet SSD of YOLOv5.





Figuur 2. Objectdetectie is mogelijk met een grote verscheidenheid van ontwikkelboards, waaronder de Arduino Portenta.



Figuur 3. Hier is een eerdere iteratie van de FOMO-benadering gebruikt om afzonderlijke bijen te tellen.

FOMO kan worden geminimaliseerd tot ongeveer 100 kilobyte in RAM, waardoor het mogelijk wordt objectdetectie in real-time uit te voeren op zo ongeveer alles, van zeer beperkte Arm Cortex-M4 cores tot krachtiger exemplaren zoals de Cortex-M7 cores op de Arduino Portenta H7 (figuur 2), de nieuwe Arduino Nicla Vision (ook een dubbele Arm Cortex-M7/M4 CPU), of zelfs gespecialiseerde DSP's zoals de Himax WE-I.

FOMO kan worden opgeschaald van de kleinste microcontrollers tot complete gateways of GPU's. Deze hoge mate van flexibiliteit maakt FOMO ook nuttig wanneer voor foutopsporing binnen een beeld uiterst minimale variaties moeten worden geïdentificeerd.

In een MCU met weinig reken- en geheugencapaciteit, verdient het aanbeveling een beeldformaat van ongeveer 96x96 pixels te gebruiken. Maar bij een grotere microcontroller is 160x160 pixels waarschijnlijk goed te doen. Het belangrijkste is dat FOMO volledig convolutioneel is, dus het werkt op elk willekeurig invoergrootte. Als u meer granulariteit, meer detail, of meer objecten nodig hebt, kunt u gewoon de invoerresolutie vergroten.

Het ziet de kleine dingen

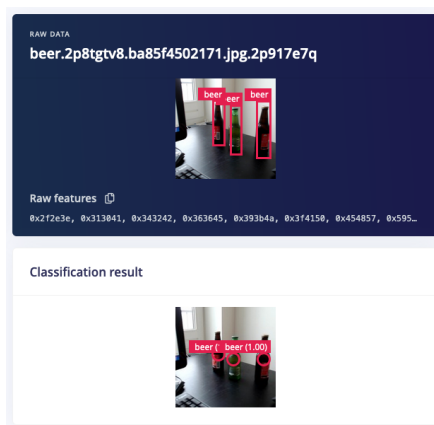
Zolang de objecten in het frame van vergelijkbare grootte zijn en elkaar niet overlappen, kan deze nieuwe architectuur zelfs heel veel kleine objecten zeer effectief

opsporen en tellen (figuur 3). Dat is iets wat MobileNet SSD en YOLOv5 niet al te goed kunnen, hoewel die groter en krachtiger zijn.

Niets meer missen

FOMO is nu beschikbaar, draait op een groot aantal computerplatforms en is compatibel met Linux-systemen, Cortex-M microcontrollers, en gespecialiseerde DSP's. Voeg een camera en Edge Impulse toe, en u kunt aan de slag.

Met FOMO kunt u snel objectdetectie toevoegen aan zo ongeveer elk apparaat dat op een camera is gebaseerd, en de angst vermijden waarmee embedded engineers nu vaak te kampen hadden: de angst om met computer vision de boot te missen (figuur 4).



Om meer over FOMO te weten te komen en met uw eigen algoritme te experimenteren, kunt u edgeimpulse.com/fomo bezoeken. 

220207-03

Over de auteur

Jan Jongboom is een embedded engineer en voorvechter van machine learning, altijd op zoek naar manieren om meer informatie uit de echte wereld te halen. Hij heeft apparaten verscheept, aan de nieuwste netwerktechnologie gewerkt, microcontrollers gesimuleerd en er staat zelfs een monument in San Francisco met zijn naam erop. Momenteel is hij medeoprichter en CTO van Edge Impulse, het toonaangevende ontwikkelplatform voor embedded machine learning met meer dan 80.000 projecten.



Figuur 4. Training op het geometrische middelpunt van bierflessen. Bovenaan de bronlabels, onderaan het inferentieresultaat.