Machinevisiesysteem leert zichzelf objecten herkennen

19 oktober 2018, 02:00

Peter Paulissen
Jan Kempeneers

Zelflerende robots die in staat zijn 'bewuster' hun omgeving waar te nemen kunnen veelzijdiger worden ingezet en op termijn magazijn en productievloer overstijgen.

Wanneer robotvisie binnen de nauwe grenzen waarvoor het ontworpen is wordt gebruikt, is die vandaag heel goed. Dit volstaat voor machines die een bepaalde beweging keer op keer moeten uitvoeren: denk maar aan een voorwerp van een productielijn opnemen en het in een bak plaatsen. Willen robots echter voor meer worden ingezet dan alleen maar voor dergelijke specifieke taken in fabrieken en magazijnen, dan moet hun beperkte zicht aangepakt worden. Onderzoekers aan het Computer Science and Artificial Intelligence Laboratory (CSAIL) van MIT ontwikkelden hiervoor het 'DON'-systeem (Dense Object Nets).

Taakspecifiek vs. algemeen

Twee frequent gebruikte manieren om robots te laten grijpen, zijn gebaseerd op taakspecifiek leren of de creatie van een algemeen grijpalgoritme. Deze technieken hebben te kampen met enkele hindernissen: taakspecifieke methoden zijn moeilijk uit te breiden naar andere taken en het algemeen grijpen is niet specifiek genoeg om te kunnen omgaan met de nuances van specifieke taken, zoals objecten op bepaalde plaatsen zetten.

Vandaag is het beste wat je van een computervisiesysteem kan verwachten objectdetectie. Het volgende niveau is pixel-labeling. Deze twee staan grotendeels voor waar zelfrijdende auto's gebruik van maken. Wil je echter interactie aangaan met het beoogde object, zoals het grijpen, dan zijn de pixels van dat object niet voldoende.

Visuele roadmaps

DON is een nieuwe vorm van machinevisie die visuele roadmaps genereren, wat neerkomt op verzamelingen visuele datapunten gerangschikt als coördinaten. Het systeem zal elk van deze individuele coördinatensets aan elkaar hangen tot een grotere coördinatenset en visualiseert zo een 3D-vorm van een object, vergelijkbaar met hoe een digitaal fototoestel of smartphone verschillende foto's kan samenbrengen tot één panoramafoto. Dit laat het systeem toe een beter en meer intuïtief begrip te hebben van de vorm van een willekeurig object en hoe het zich gedraagt in de context van zijn omgeving.

Het DON-systeem draait rondom de details binnen het object, het soort informatie dat nodig is om geavanceerde behandelingstaken te kunnen uitvoeren. Het systeem zal een robot toelaten naar bijvoorbeeld een kop koffie te kijken, zichzelf oriënteren om deze te hanteren en tot het besluit komen dat de bodem van de kop naar beneden moet wijzen wanneer deze opgenomen wordt, om te voorkomen dat van de inhoud gemorst wordt. Het systeem zal zo ook een robot toelaten een specifiek voorwerp uit een stapel gelijkaardige objecten te pikken.

Heel wat benaderingen voor manipulatie laten de identificatie van specifieke onderdelen van een object niet toe over de vele oriëntaties waaruit een object kan benaderd worden heen. Bestaande algoritmen zijn bijvoorbeeld niet in staat een kop bij het oor te grijpen, vooral niet als de kop op verschillende manieren georiënteerd wordt, zoals rechtop of op zijn zijde.

Zelflerend

Het DON-systeem is gebaseerd op een RGB-D-sensor in combinatie met een RGB-dieptecamera, en is in staat zichzelf te trainen, zonder menselijke tussenkomst, ook voor nieuwe taken of objecten die het nooit eerder gezien heeft. Het is dus niet nodig de AI honderden tot duizenden beelden van een object aan te reiken om de DON te laten leren. Wil je dat het systeem bijvoorbeeld een bruine laars herkent, plaats je de robot een tijdje in een kamer met een bruine laars. Het systeem zal automatisch rond de laars gaan, referentiefoto's nemen die het gebruikt om coördinatiepunten te genereren, waarna het zichzelf zal trainen op basis van wat het gezien heeft. Het volledige proces neemt minder dan een uur in beslag.

Op de productievloer moeten robots gebruik maken van complexe aanvoersystemen voor onderdelen om betrouwbaar werk te leveren. Een systeem als dit kan de oriëntering van voorwerpen begrijpen door een foto te nemen en in staat zijn ze te grijpen en het object op een passende manier te manipuleren.



(Bron: MIT)

Geslaagde testen

In een testenreeks met een zacht stuk speelgoed was een robotarm uitgerust met DON in staat het speelgoed bij het rechteroor te grijpen vanuit verschillende configuraties. Het systeem kon dus het onderscheid maken tussen links en rechts bij symmetrische objecten.

Bij testen op een doos met verschillende baseballpetjes kon het systeem een specifiek hoofddeksel oppikken, ondanks dat alle petjes vergelijkbare vormen hadden en het nooit eerder foto's gezien had van de petjes in trainingdata van de dag voordien.

De technologie staat nog in de kinderschoenen, maar op termijn hopen de onderzoekers robots met verbeterde visie en coördinatie te kunnen ontwikkelen die worden ingezet voor specifieke taken waarvoor een meer diepgaand begrip nodig is en op diverse plaatsen, waaronder in woningen en kantoren.
Bekijk het filmpje:

Bronnen

- http://news.mit.edu
- https://www.engadget.com
- https://arxiv.org

Auteurs



Peter Paulissen



Jan Kempeneers