



# Moeten dermatologen vrezen voor hun baan?

M.W. Bekkenk

Artificiële Intelligentie (AI) zorgde in de jaren negentig voor het eerst voor een opzienbarend resultaat: in 1997 werd de wereldkampioen schaken, Gary Kasparov, verslagen door Deep Blue (een IBM supercomputer). Door de enorme rekenkracht van Deep Blue waarbij alle mogelijke zetten in minimale tijd werden doorgerekend was de computer de mens de baas met schaak. Dit duurde veel langer bij de Oosterse denksport Go, maar uiteindelijk wist Google's AlphaGo-programma ook bij deze uitdaging de mens te verslaan. Dit programma maakte gebruik van zogenoemde neurale netwerken, waardoor menselijke besluitvorming kan worden nagebootst. Dit soort programma's wordt steeds meer gebruikt voor allerlei toepassingen, zoals het herkennen van gezichten en het onderscheiden van verschillende hondenrassen aan de hand van een foto. Een AI-programma kijkt heel anders naar een foto dan wij doen, het beschouwt een plaatje als een 3D-object aangezien het alle dimensies van het rood-blauw-groen (RGB)-spectrum los bekijkt en al deze informatielagen als het ware op elkaar stapelt.

## CONVOLUTIONELE NEURALE NETWERKEN

Een zelfde type programma met convolutionele neurale netwerken (CNN) zorgde vorig jaar voor een schok in de dermatologische wereld. In *Nature* publiceerde een groep uit Stanford een artikel [1] waar twee groepen werden vergeleken: verruca seborrhoeica versus keratinocytkankers en benigne naevi versus melanoom. De computer nam het op tegen 21 dermatologen en bleek het onderscheid tussen de bovengenoemde aandoeningen vergelijkbaar goed (of deels zelfs beter) te kunnen maken. In geval van dermatoscopie zelfs wat beter dan de gemiddelde dermatoloog. Was de dermatoloog nu ook verslagen door de computer? Belangrijk om te melden is dat het onderzoek ging om puur het beoordelen van een plaatje. Er werd geen klinische informatie verstrekt en alleen de beoordeling benigne of maligne in twee aparte groepen foto's werd beoordeeld. Ondanks deze vrij beperkte vergelijking werden de resultaten sterk genoeg geacht om te publiceren in *Nature*. [1]

Om een goed CNN te maken voor dermatologie heb je veel, heel veel, en goede resolutieplaatjes nodig waarvan de juiste diagnose bekend is. Het zelflerende van het systeem moet vrij letterlijk genomen worden. Na analyse van het eerste plaatje wordt als het ware gekokt, waarna het juiste antwoord wordt gegeven. Doordat het programma bij wijze van spreken aan een enorm aantal knoppen kan draaien binnen het systeem van de eerder genoemde 3D-interpretatie van een plaatje zal het bij elk fout antwoord zijn methode bijstellen en dus steeds beter worden. Hoe meer informatie het systeem krijgt des te beter het wordt. Kortom meer plaatjes van goede kwaliteit met de juiste informatie zorgen voor betere antwoorden. Eigenlijk vergelijkbaar met de leercurve van een aio die steeds meer beelden heeft gezien en deze zal gaan herkennen onder

supervisie en na feedback van een goede leermeester. De leercurve van een CNN blijft alleen steeds doorgroeien, terwijl een arts op een gegeven moment met pensioen gaat.

Ondanks de overtuigende data zijn er wel wat dingen aan te merken op de studie. Zo liet Robert Novoa, een van de coauteurs van het *Nature*-artikel, in een interview weten dat je niet zo maar klakkeloos dit soort systemen kan doorvoeren in de diagnostiek. Een algoritme kan namelijk ook voor de gek gehouden worden. Bij analyse van de argumenten die het systeem gebruikte om te oordelen of het maligne of benigne was bleek dat de aanwezigheid van een meetlat op de foto een sterke indicator was voor maligniteit. Foto's aangeleverd door dermatologen bevatten namelijk vaker een meetlat op de foto en in deze groep foto's werden veel vaker maligne laesies aangetroffen dan in de door leken en huisartsen aangeleverde foto's. Een CNN bekijkt het hele plaatje en kan dus door informatie buiten de laesie om in de war worden gebracht. In feite klopte het dat dit een voorspellende waarde was voor een maligniteit, maar bij testen van het algoritme in de klinische praktijk zal de meetlat er niet meer op staan en zal het systeem dus minder goed werken. Dit is een groot nadeel van CNN-systemen, het systeem is zo zelfstandig lerend dat men niet meer weet waar op gescoord wordt en hoe de uiteindelijke beslissing tot stand komt. Het nemen van een foto vanuit een iets andere hoek of verschillende lichtinval kan in theorie veel verschil maken, dit moet dus allemaal goed gestandaardiseerd worden. Daarnaast is de methode voor validatie die het artikel kiest, namelijk alleen een keuze tussen benigne en maligne en met een dataset met alleen seborroische wratten, basaalcelcarcinomen en plaveiselcelcarcinomen in de ene en moeder-vlekken en melanomen in de andere groep, wel erg eenvoudig.

Dermatoloog, afdeling Dermatologie, Amsterdam UMC, locatie AMC, Amsterdam, namens de ICT commissie van de NVDV



Kasparov versus Deep Blue.

Het computeralgoritme neemt het nu als het ware op tegen een half geblinddoekte dermatoloog met zijn handen op de rug gebonden. Niet helemaal een faire vergelijking dus. Een dermatoloog krijgt bij een consult immers veel meer informatie dan alleen een plaatje. Het beoordelen van de rest van het lichaam en anamnestiche informatie kan onze beslissing (soms zonder dat we ons dat bewust zijn) beïnvloeden. Bovendien zijn er binnen de dermatologie veel complexere keuzes te maken dan slechts de keuze benigne of maligne en dan ook nog binnen een afgebakende groep huidafwijkingen. Wat dit artikel dus vooral laat zien is dat een computerprogramma in staat is even goed een plaatje van een vooraf gekozen huidafwijking te beoordelen op de aanwezigheid van een maligniteit dan een dermatoloog, niets minder, maar ook (nog) niets meer.

### CNN BINNEN DE GENEESKUNDE

Toch is het een belangrijke doorbraak en is deze technologie niet iets waar we als dermatologen onze ogen voor moeten sluiten. De grote vraag is waar CNN-systemen binnen de geneeskunde als eerste grote waarde gaan hebben. Binnen voor de hand liggende vakken als radiologie en pathologie worden al uitgebreid testen gedaan. Begin 2018 was al in het nieuws dat bij het onderzoeken van kankercellen in lymfeklieren de computer het beter deed dan de patholoog [2] en dat door analyse van een oogscan een voorspelling kan worden gegeven op de kans op hart- en vaatziekten. [3]

### CNN BINNEN DE DERMATOLOGIE

Binnen de dermatologie wordt met CNN voorlopig vooral op huidkanker ingezet. Het breder beoordelen van huidafwijkingen blijkt aanzienlijk lastiger en op dit moment is de voorspellende waarde te slecht om hier al mee te starten. Een logische eerste rol zou er kunnen zijn voor een systeem dat de huisarts of de dermatoloog kan ondersteunen bij de diagnostiek bij lastige melanocytair laesies. Het toevoegen van klinische informatie is mogelijk bij dit soort systemen en het toevoegen

van duizenden foto's via een systeem zou een snelle leercurve betekenen voor een CNN dat huidafwijkingen kan beoordelen bij verdenking van maligniteit. Het gebruik van mobiele telefoons kan het gebruik vergemakkelijken, maar de vraag blijft of dit systeem uiteindelijk beter werkt dan het huidige systeem. Een perfecte diagnose door welk systeem dan ook is een utopie, omdat er een te grote spreiding zit in de klinische presentatie en de uiteindelijke diagnose. Bovendien moeten we ons afvragen wat het dan moet kunnen: nooit een melanoom missen of eenvoudige benigne laesies herkennen? Daarnaast is de plaats voor deze technologie ook belangrijk, hoort dit bij de dermatoloog, de huisarts of de consument?

Een mogelijk antwoord hierop is om dit type technologie in te zetten voor de consument en de technologie te gebruiken om geen melanoom meer te missen. Het resultaat zal dan waarschijnlijk zijn dat door de scherpe instellingen de mazen van het net als het ware te klein worden en er heel veel mensen met 'verdachte plekken' naar de dermatoloog gaan. Resultaat zal dan zijn dat het eerder voller dan leger wordt in de wachtkamer van de dermatoloog.

### LITERATUUR

1. Esteve A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-8.
2. Liu Y, Gadepalli K, Norouzi M, et al. Detecting cancer metastases on gigapixel pathology images, open access file. Te bekijken op <https://drive.google.com/file/d/0B1T58bZ5vYa-QLRoQUtazdPWVvK/view>.
3. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2:158-64.

### CORRESPONDENTIEADRES

Marcel Bekkenk

E-mail: [m.w.bekkenk@amc.uva.nl](mailto:m.w.bekkenk@amc.uva.nl)