

Validatie en implementatie van commerciële AI-software in de radiologie



Kicky van Leeuwen

Dit proefschrift richt zich op de snel groeiende markt van commercieel beschikbare kunstmatige intelligentie (AI) software voor de radiologie. Het eerste deel belicht de toepassingen van AI in de radiologie en de producten die hiervoor op de markt zijn. Het tweede deel onderzoekt hoe goed deze AI-oplossingen momenteel zijn en hoe we de toegevoegde waarden van AI-producten kunnen inschatten. Tot slot is de implementatie van AI-software in de praktijk onderzocht.

Radiologie is een van de meest technologie-gedreven en snel veranderende medische disciplines. Recent heeft het radiologisch veld kennisgemaakt met een nieuwe technologische ontwikkeling, die de manier waarop radiologen werken en de manier waarop patiënten worden gediagnosticeerd drastisch kan veranderen. Deze technologie is kunstmatige intelligentie (AI). In 2019, toen ons onderzoek van start ging – ongeveer vier jaar na de doorbraak van *deep learning* – was er nog veel onzekerheid rondom deze nieuwe technologie. Er werden tal van AI-producten geïntroduceerd met veelbelovende claims, maar het was onduidelijk welke van deze producten klinische gebruikt mochten en konden worden, hoe effectief ze waren en welke waarde ze zouden toevoegen in de klinische praktijk.

Overzicht van AI-producten

Het eerste doel van ons onderzoek was om AI-toepassingen te identificeren die voldeden aan de Europese wet- en regelgeving. Dit resulteerde al snel in een lijst van honderd producten. Vanwege de snelle ontwikkelingen op de markt realiseerden we ons dat een wetenschappelijk artikel, hoewel informatief, niet de meest doeltreffende manier was om dit overzicht te delen. De website www.AlforRadiology.com was geboren. Op dit platform zijn inmiddels meer dan 220 commercieel beschikbare AI-producten te vinden, inclusief productspecificaties en wetenschappelijk bewijs voor elk product. Deze website wordt wereldwijd gebruikt om inzicht te krijgen in de huidige stand van zaken in de AI-markt voor de radiologie.

Voor deze website hebben we ook alle beschikbare studies naar de producten verzameld. Hieruit ontdekten we in 2020 dat voor twee derde van de honderd geanalyseerde producten geen enkele publicatie beschikbaar was! Gelukkig zien we inmiddels wel steeds meer retrospectieve, maar ook prospectieve studies uitkomen. Zo weten we steeds beter wat van een product te verwachten. Via onze maandelijkse nieuwsbrief houden we onze lezers op de hoogte van de meest recente onderzoeksresultaten.

Vergelijking AI-producten

Ondanks de toename van publicaties over de prestaties en toegevoegde waarde van AI-producten, blijft het uitdagend om producten onderling te vergelijken. Zo worden verschillende datasets en onderzoeksprotocollen gebruikt voor validatie. Bovendien kan het gebeuren dat op het moment van publicatie, de commercieel beschikbare algoritmes al zijn geüpdatet. Hierdoor verouderen de resultaten snel.

Project AIR

Om deze problemen aan te pakken, hebben we de Project AIR-methodologie ontwikkeld. Met Project AIR kunnen we AI-producten op een objectieve en transparante manier valideren en vergelijken.

Dit concept van directe en transparante productvergelijking is gangbaar in de consumentenmarkt, denk aan organisaties zoals de *Consumentenbond* of *Tweakers*, maar wordt zelden toegepast in de zorg en radiologie. We hebben de Project AIR-methodologie toegepast op twee *use cases*: de detectie van longnodules op conventionele longfoto's en de voorspelling van skeletleeftijd op basis van handfoto's. We gebruikten beelden van tien verschillende Nederlandse ziekenhuizen om benchmarkdatasets samen te stellen.

‘De vraag is niet langer óf AI het werk van radiologen zal veranderen, maar vooral wanneer we ons een wereld zonder AI in de radiologie niet meer kunnen voorstellen’

Hiermee hebben we zeven AI-producten voor longnodules en twee AI-producten voor skeletleeftijdbepaling met elkaar kunnen vergelijken. Tevens hebben tientallen radiologen een deel van de beelden beoordeeld ter vergelijking. We constateerden dat beide producten minstens zo goed zijn als de radiologen bij het bepalen van de skeletleeftijd, terwijl er meer variabiliteit was tussen de producten en radiologen bij het detecteren van longnodules. Vier van de zeven producten bleken significant beter te presteren dan de gemiddelde radioloog.

Stimulans voor leveranciers

Dit is de grootste directe vergelijkingsstudie van commerciële AI-producten tot

nu toe. De resultaten worden online gepubliceerd op www.grand-challenge.org. Op deze manier kunnen nieuwe software en updates van de software in de loop der tijd worden toegevoegd. Radiologieafdelingen en vergoedingsorganisaties kunnen de performance-informatie gebruiken om inkoop- of vergoedingsbeslissingen te nemen. Dit zou op zijn beurt een belangrijke stimulans creëren voor leveranciers

om deel te nemen aan de vergelijking en voortdurend hun product te blijven verbeteren.

Gebruik in Nederland

Een cruciale voorwaarde voor de klinische impact van AI is de adoptie ervan. We hebben gedurende drie jaar onderzocht welke radiologieafdelingen welke AI-producten in de dagelijkse praktijk gebrui-

ken. In samenwerking met de NVvR hebben we het AI-netwerk hiervoor opgezet. Dit netwerk bestaat uit ongeveer vijftig radiologen met interesse in AI, elk werkzaam bij een andere radiologieafdeling. Deze radiologen schreven we jaarlijks aan om informatie te verzamelen over de AI adoptie op hun afdeling. Hieruit is gebleken dat minstens een derde van de Nederlandse radiologieafdelingen nu één of meerdere commerciële AI-producten gebruikt in de klinische praktijk. Voor NVvR-leden is deze informatie beschikbaar op www.radiologen.nl. Hier is te zien welke AI-producten door collega's worden gebruikt. We hopen dat radiologen op deze manier kennis en ervaringen kunnen delen en informatie kunnen inwinnen voordat ze besluiten om een AI-product aan te schaffen.

Concluderend

Velen herinneren zich wellicht de uitspraak uit 2016 van Geoffrey Hinton, een bekende computerwetenschapper, waarin hij stelde: *'People should stop training radiologists now. It's just completely obvious that in five years deep learning [AI] is going to do better than radiologists'*. Omtrent AI in de radiologie is intussen veel gebeurd. Er zijn nu meer dan 220 AI-producten beschikbaar op de Europese markt. Het gebruik van AI in minstens een derde van de Nederlandse ziekenhuizen is inmiddels een feit. De hoge verwachtingen en scepsis rondom AI zijn genormaliseerd. De vraag is niet langer óf AI het werk van radiologen zal veranderen, maar vooral wanneer we ons een wereld zonder AI in de radiologie niet meer kunnen voorstellen. Echter, één ding blijft voorlopig ongewijzigd: de opleiding tot radioloog blijft nog wel even bestaan. ■

De promotiedag



Op 25 september had ik de eer om mijn proefschrift te verdedigen voor een indrukwekkend gezelschap van opponenten, onder anderen overgevlogen uit Engeland en Duitsland. Aangezien ze allemaal in Nijmegen waren, besloten we om er een speciaal evenement van te maken. We hebben er een debat-event aan gekoppeld. Daarin bespraken we drie stellingen over de toekomst van AI in de radiologie. Het ging er vurig aan toe tussen de duo's en het publiek van meer dan honderd mensen. Tijdens het stemmen bleek wel dat het laatste woord over de onderwerpen nog niet gezegd is.

Mijn verdediging volgde. Tijdens het vragenvuur vergat ik even dat er allemaal vrienden, familie, collega's en vakgenoten live en digitaal aanwezig waren. Hier werd ik snel aan herinnerd toen na de *hora est* een applaus losbarstte. Tot mijn grote verrassing werd me daarna het doctoraat cum laude toegekend. Na een prachtig laudatio van professor Bram van Ginneken sloten we de middag af met een borrel en vele hapjes.



Nijmegen, 25 september 2023

dr. Kicky van Leeuwen

adviseur AI inkoop en implementatie
oprichter AlforRadiology.com

Met veel dank aan mijn (co)promotoren:
prof. dr. B. van Ginneken, Radboudumc
dr. M.J.C.M. Rutten, Radboudumc en
Jeroen Bosch Ziekenhuis
dr. M. de Rooij, Radboudumc
dr. S. Schalekamp, Radboudumc

Het proefschrift is te downloaden via:
<https://tinyurl.com/kvl-phd-thesis>
De verdediging is terug te zien via:
<https://youtu.be/O6Kq12oEjz0>