



Universiteit
Leiden
The Netherlands

Aspects of the analysis of cell imagery: from shape to understanding

Li, C.

Citation

Li, C. (2024, June 27). *Aspects of the analysis of cell imagery: from shape to understanding*. Retrieved from <https://hdl.handle.net/1887/3765419>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3765419>

Note: To cite this publication please use the final published version (if applicable).

Nederlandse Samenvatting

Microscopie ondersteunt ons bij het observeren van kleine objecten zoals cellen, weefsels en micro-organismen. In het werkveld van BioImaging worden beelden opgenomen om biologische fenomenen te bestuderen. Vaak worden grote verzamelingen beelden opgenomen zodat deze beelden kunnen worden geanalyseerd om een fenomeen te begrijpen. Om deze analyses te kunnen doen zijn computationele methoden nodig waarmee we kennis uit de beelden kunnen genereren. Imaging kan, met deze werkwijze, studies in het biomedische domein versnellen. In dit proefschrift hebben we beelden van cellen bestudeerd waarbij de analyse inzicht verschaft in begrip van de inhoud van deze beelden. We hebben in het bijzonder naar beelden van pollenkorrels en immuuncellen gekeken. Bij deze cellen betrof de beeldanalyse van zowel 2D als 3D beelden. Daarenboven is gekeken naar dynamiek uit beeldreeksen. De beeldanalyse van de pollenkorrels heeft een focus op classificatie van de pollen terwijl de aandacht bij de analyse van de immuuncellen vooral gaat naar de dynamiek van neutrofielen door deze te volgen in een tijdinterval opgenomen videosequentie. De aanpak die nodig is voor deze analyses berust op technieken uit de kunstmatige intelligentie, waar computers leren om bepaalde patronen te herkennen.

Beelden van pollen worden vaak gemaakt van uit de lucht genomen monsters. In beelden van deze luchtmonsters kan een specialist verschillende soorten pollen herkennen. Echter, soms is het moeilijk, zo niet onmogelijk om sommige soorten uit elkaar te houden. In onze studie hebben we ons beziggehouden met soorten waar dit typisch het geval is. We doen daarmee een poging om allergeen pollen van Parietaria van niet-allergeen pollen van Urtica en Urtica membranacea, allen uit dezelfde Urticaceae familie, te kunnen onderscheiden. We gebruiken beelden van pollen om na te gaan of we een computer(-programma) kunnen leren de soorten uit elkaar te houden. Om voldoende detail in de beelden te verkrijgen, maken we gebruik van de mogelijkheid om de pollenkorrel in 20 parallelle optische aanzichten in beeld vast te leggen, en uit die 20 aanzichten een projectie te genereren in een beeld. Van ieder pollenbeeld verkrijgen we 3 verschillende projecties die we als invoerbeelden gebruiken voor een classificatiemodel. Langs deze weg proberen we zoveel mogelijk representatieve

kenmerken van pollen te integreren. Bovendien is gebleken dat met deze 3 types van projecties, de prestatie van de classificatie aanzienlijk verbeterd.

De hoeveelheid beelden die we beschikbaar hebben is voldoende om zogenaamde “deep learning” leermodellen te gebruiken. Eerst hebben we dit type modellen getraind om drie soorten pollen te classificeren. De modellen die we hiervoor hebben gebruikt zijn: VGG16, MobileNet V1 en MobileNet V2. Teneinde de mogelijkheden van “deep learning” verder te exploreren, hebben we onze studie uitgebreid met nog 3 andere “deep learning” modellen. Deze zijn: AlexNet, VGG19 en ResNet50. In de dagelijkse praktijk is het echter niet altijd gegarandeerd dat er een grote hoeveelheid monsters van pollen, in beeldvorm, aanwezig is. Daarom hebben we naast de “deep learning” methoden, ook de prestaties van classificatie modellen gebaseerd op traditioneel machinaal leren; zoals bijvoorbeeld de Support Vector Machine (SVM) en Random Forest (RF). In convolutionele neurale netwerken (CNN), worden kenmerken automatisch geëxtraheerd en gebruikt in een model voor training. Voor traditioneel machinaal leren moeten we zelf de kenmerken ontwerpen, zogenaamde “handcrafted” kenmerken, om het model te trainen. Methoden voor kenmerkselectie zijn overgenomen om het aantal ongerelateerde kenmerken te reduceren. Na het toepassen van kenmerkselectie zijn verschillende rechttoe-rechtaan modellen gemaakt om iedere pollencategorie te herkennen. Daarenboven is er een hiërarchische classificatie-strategie toegepast door eerst de verschillende klassen in subgroepen te categoriseren. Voor iedere subgroep worden de verschillende classificatiemodellen dan afzonderlijk getraind. Met deze experimenten hebben we laten zien dat de prestatie verbeterd door gebruik te maken van de hiërarchische strategie.

Van de implementatie van verschillende classificatie modellen, zowel gebaseerd op “deep learning” als traditioneel machinaal leren, concluderen we dat in vergelijking met machinaal leren, “deep learning” methoden een hogere nauwkeurigheid in de classificatie bereiken, ResNet 50 in het bijzonder. Onze experimentele resultaten, op zowel grote als kleine datasets, hebben de robuustheid van de “deep learning” methoden laten zien. Een ablatie studie demonstreert het succes van “deep learning” modellen door verschillende technieken, zoals data-augmentatie, “transfer learning” en “hard voting” te integreren. We hebben een werkwijze uitgewerkt en geïmplementeerd om een oplossing te vinden voor het nauwkeurig kunnen classificeren van zeer gelijkende pollencategorieën. Deze studie geeft een goed inzicht in de complexe classificatie van pollen en geeft gezondheidsdiensten de mogelijkheid om hooikoortspatiënten beter te adviseren voor hun behandeling en plannen van activiteiten buitenshuis.

Naast het verwerken van beelden van pollen, hebben we nog een compleet verschillend domein bestudeerd waar we eveneens op gemakkelijke weg beelden in meerder dimensies

kunnen verkrijgen. We bestuderen het volgen van immune cellen in hun respons op een mogelijke infectie. We kijken daarom naar live-beelden van cellen zoals ze door het lichaam bewegen; i.e. neutrofielen in de larve van een zebra vis. We monitoren deze neutrofielen door ze te volgen in een videosequentie die in vaste tijdsintervallen is opgenomen (time-lapse). Het volgen van een subject in een videosequentie wordt “tracking” genoemd. De neutrofielen zijn betrokken in de eerstelijnsverdediging bij het beschermen van een organisme tegen een pathogeen; in het geval van onze studie betreft het een verwonding aan de staartvin van de larve van een zebra vis. Het monitoren van de migratie van neutrofielen is een belangrijk onderdeel voor het verkrijgen van begrip van het gedrag van de cellen van het immuunsysteem. Echter, het volgen (tracking) van neutrofielen is een uitdaging. In onze studie hebben we tijdreeks-sequenties van neutrofielen geanalyseerd die zijn opgenomen met een Confocal Laser Scanning Microscope (CLSM). Met de tijdreeks-sequenties als uitgangspunt zijn computationele methoden en algoritmen ontworpen met het specifieke doel het “tracking” probleem in zowel het 2D als het 3D domein op te kunnen lossen.

Voor het volgen van neutrofielen in een 2D tijdreeks videosequentie, stellen we een protocol voor waarin opgenomen de segmentatie van de cel, het volgen van de bewegingen van de cel tussen opeenvolgende “frames” en het aaneenrijgen van de cel uit de tijdreeks-sequenties in een traject over de tijd. De neutrofielen zijn fluorescent gelabeld en de eerste stap in het identificeren van deze cellen door is een segmentatie. Voor deze segmentatie gebruiken we U-Net gebaseerde segmentatie modellen en een algoritme dat is gebaseerd op regels, nl. Watershed Masked Clustering (WMC). Deze twee methoden voor segmentatie zijn geïmplementeerd en hun prestaties zijn onderling vergeleken met betrekking tot accurate cel-lokalisatie voor de tracking procedure. Het is onze bedoeling het best presterende segmentatiemodel op te nemen in ons protocol. Vanuit de resultaten van de segmentatie moet, vervolgens de cel beweging kunnen worden begrepen. Daartoe hebben we een extra U-net model geconstrueerd dat is getraind om cel beweging en migratie te leren uit geannoteerde grondwaarheid data. Op deze manier kunnen we de positie van een cel in een volgend frame van de tijdreeks-sequentie voorspellen. Om vervolgens trajecten van cellen te verkrijgen moeten we ons buigen over de uitdagende taak van het ontrafelen van cellen in de tijdreeksen-beelden die zich lijken samen te voegen of op te splitsen. We hebben hier een belangrijke bijdrage geleverd die bestaat uit een uitbreiding van het “Viterbi Linkage Algoritme” waarmee de ambiguë situaties objectief kunnen worden opgelost. Met deze drie componenten levert ons protocol bevredigende prestaties voor het volgen van neutrofielen *in vivo*.

In navolging van onze resultaten in de analyse van 2D video-sequenties, hebben we het volgen van neutrofielen in 3D tijdreeks-beelden onderzocht. We hebben een protocol ontworpen waarin opgenomen 3D cel segmentatie, een gewogen kenmerk-similariteitschatting

en wederom het linken van de cel-posities in een traject (trajectory linkage). Voor de 3D celsegmentatie hebben we een 3D U-Net model geselecteerd, om zo voor ieder punt in de tijdreeks een segmentatie van een cel te krijgen. Een similariteitschatting wordt gedaan door een aantal kenmerken van de cel te extraheren: i.e. afgelegde cel afstand, bewegingsrichting van de cel en de gemiddelde afstand over het afgelegde traject. Door deze kenmerken op verschillende schaal te wegen wordt er een similariteitscore berekend. Op basis van deze similariteitscore hebben we het zogenaamde Hongaarse Algoritme aangepast om de cel-posities te linken in trajecten van de cellen. In vergelijking met bestaande moderne methoden, kunnen we concluderen dat ons protocol, gegeven onze experimentele resultaten, een goede prestatie levert.

In dit proefschrift hebben we gedemonstreerd hoe “deep learning” methoden kunnen worden toegepast in uiteenlopende taken voor de analyse van beelden van cellen en daarbij een inkijk gegeven in de potentie van deze methoden voor de toekomst. We illustreren de valkuilen maar presenteren eveneens nieuwe en efficiënte algoritmische benaderingswijzen.