



Universiteit
Leiden

The Netherlands

Visual analytics for spatially resolved omics data at single cell resolution: methods & applications

Somarakis, A.

Citation

Somarakis, A. (2022, January 20). *Visual analytics for spatially resolved omics data at single cell resolution: methods & applications*. Retrieved from <https://hdl.handle.net/1887/3250550>

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/1887/3250550>

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

NEDERLANDSE SAMENVATTING

Een dieper inzicht in de pathologie van een organisme is belangrijk voor de ontwikkeling van behandelingen. In de loop van eeuwen van systematisch onderzoek hebben klinische onderzoekers aangetoond dat hoe meer informatie zij verwerven over de ceileigenschappen en hun organisatie in het weefsel, hoe beter zij de functionaliteit van een organisme en de ziekteprogressie kunnen begrijpen. De laatste jaren heeft de komst van hoge-resolutie beeldvormingstechnieken onderzoekers voorzien van nieuwe informatie over individuele cellen, die onderzoekers in staat stellen om de cellen nauwkeurig te karakteriseren en te onderzoeken hoe ze zijn verdeeld in het weefsel. Echter, de extractie van bruikbare biologische inzichten uit de analyse van zulke nieuwe en complexe data, waar experts de intrinsieke karakteristieken van de data niet kennen noch de patronen die ze willen identificeren, vereist een exploratieve data-analyse aanpak. Het doel van dit proefschrift is dan ook de ontwikkeling van een end-to-end pijplijn voor de analyse van deze hoogdimensionale cellulaire beelden; van de voorbereiding van de ruwe data tot de exploratie van cellulaire patronen en hun associatie met klinische karakteristieken. In Hoofdstuk 1 motiveren we het belang en de noodzaak van onze studie, presenteren we de belangrijkste kenmerken van de data, bespreken we de voorbereidingsalgoritmen en introduceren we de visual analytics benadering voor de exploratieve analyse van complexe data. Vervolgens beschrijven we een nieuwe methodologie voor de voorbereiding van deze hoogdimensionale cellulaire beelden (Hoofdstuk 2). In Hoofdstuk 2.1 beschrijven we de semi-supervised workflow die we hebben ontwikkeld voor de normalisatie van Imaging Mass Cytometry data tussen verschillende monsters en dimensies. Daarnaast laten we zien hoe de binaire classificatie van een beeldpixel in voorgrond en achtergrond niet-biologische variatie tussen weefselmonsters kan elimineren en de traceerbaarheid van cellen met laag-geëxprimeerde eiwitniveaus kan verbeteren. In hoofdstuk 2.2 presenteren we een segmentatie-algoritme, gericht op de precieze identificatie van microglia, zoals deze is vastgelegd met de Vectra multiplexed immunofluorescentie modaliteit. In het bijzonder wordt een combinatie van traditionele segmentatie-algoritmen (Level-Set, Watershed) en morfologische operaties gebruikt om een geautomatiseerd algoritme te creëren voor de precieze identificatie van microglia's cellulaire grenzen. In de Hoofdstukken 3 en 4 stellen we een Visual Analytics aanpak voor voor de analyse van deze zeer complexe ruimtelijke cellulaire gegevens, die twee data-gedreven interactieve tools omvat, ImaCytE en SpaCeCo. ImaCytE (Hoofdstuk 3) stelt de experts in staat om meerdere weefselmonsters te onderzoeken en nieuwe celtypen te identificeren, om op meerdere niveaus de ruimtelijke patronen die zij vormen te onderzoeken en om specifieke celtypen te stratificeren op basis van hun micro-omgeving kenmerken. SpaCeCo (Hoofdstuk 4) is een aanvulling op ImaCytE en stelt deskundigen in staat om te bepalen welke cellulaire kenmerken twee cohorten monsters met verschillende klinische kenmerken van elkaar onderscheiden. De workflow voor de vergelijking is verdeeld in twee stappen. De eerste stap omvat de vergelijking van cohorten op basis van de abundantie van de verschillende celtypen en de tweede stap op basis van de ruimtelijke pa-

tronen die de verschillende celtypes vormen. De vergelijking vindt plaats op twee niveaus, het cohort- en het monsterniveau, zodat de deskundige in staat is eventuele uitschieters in elk cohort te identificeren. Zowel in ImaCytE als in SpaCeCo kan de gebruiker elke bevinding op het weefsel lokaliseren om ze te verifiëren en in haar biologische context te plaatsen. In Hoofdstuk 5 laten we zien hoe delen van onze pijplijn zijn gebruikt voor een studie naar de ziekte van Alzheimer. Het segmentatie algoritme voor de efficiënte identificatie van microglia, de motief glyphs, voor het eerst gepresenteerd in ImaCytE, voor de exploratie van cellulaire patronen in het weefsel en SpaCeCo voor de vergelijking van Alzheimer patiënten met gezonde individuen. Tot slot hebben we Visual Analytics oplossingen ontwikkeld om de experts in staat te stellen diepgaande ruimtelijke cellulaire gegevens te verkennen en hypothesen te genereren voor de oorsprong van de functionaliteit van het weefsel in een zieke of een gezonde toestand. Naast de kerntaken op het gebied van exploratie, stellen we de experts in staat om hun gegevens te pre-processen, waardoor een end-to-end pijplijn ontstaat voor de analyse van multiplexed weefselbeelden op cellulaire resolutie. Bij elkaar beschouwd vormt het werk in dit proefschrift de basis voor een Visual Analytics-benadering voor de analyse van hoogdimensionale beeldvorming om cellen en weefsels in groot detail te kunnen bestuderen.