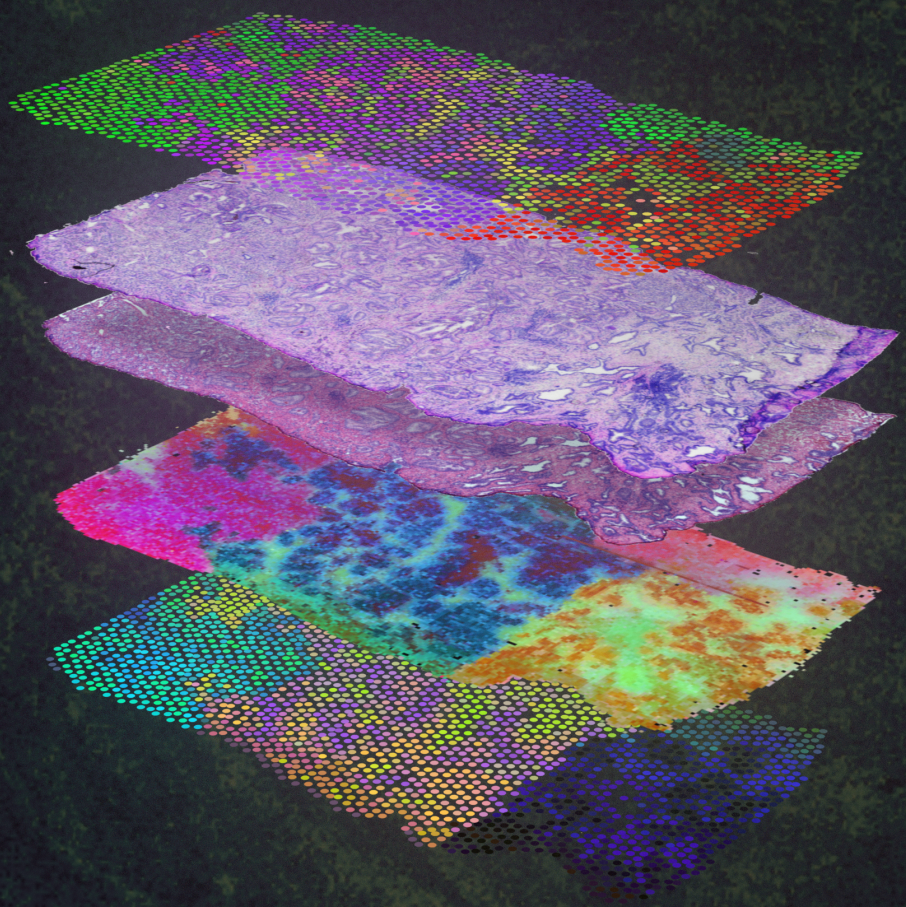
KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Machine Learning and Multimodality in Mass Spectrometry Imaging



Wanqiu Zhang | 张琬秋

August 2023

Machine Learning and Multimodality in Mass Spectrometry Imaging

Towards Spatial Multi-Omics

Wanqiu ZHANG 张琬秋

Examination committee:
Prof. dr. ir. Hendrik Van Brussel, chair
Prof. dr. ir. Bart De Moor, supervisor
Dr. ir. Marc Claesen, co-supervisor
(Aspect Analytics NV, Genk, Belgium)
Prof. dr. ir. Johan Suykens
Prof. dr. Jesse Davis
Prof. dr. Martina Marchetti-Deschmann
(TU Wien, Austria)
Prof. dr. Ron Heeren
(Maastricht University, Netherlands)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

August 2023

© 2023 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Wanqiu Zhang, Kasteelpark Arenberg 10 box 2446 , B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

故天將降大任於是人也，必先苦其心志，勞其筋骨，餓其體膚，空乏其身，行拂亂其所爲，所以動心忍性，曾益其所不能。

- 孟子

When the one is given a great mission, her willpower will be frustrated; her muscles and bones will be exhausted; she will be exposed to starvation and poverty; she will be harassed by troubles and setbacks; so that her spirit can be stimulated; her nature can be toughened; and eventually her abilities will be enhanced.

Mencius (372BC - 289BC)

When I was 9 years old, I was selected to the team of table tennis in my primary school, just because I was taller than my other classmates. Every day after school, I went to the ping-pong club of my school and stayed there till 6pm. My coach Miss Wang was a young, beautiful, strong but very tough lady. One time, she got a big bucket full of the ping-pang balls, and she asked us to catch as many of the balls that she played. She said "if you miss one ball from me, you have to run one lap around the playground (200m)". Many kids were nervous and missed 4 or 5 balls. And when it was my turn, I was too stressful to focus. I missed 14 balls! I broke the record in the club.

Actually, I started to cry when I missed 7 balls. My coach was disappointed and asked me to start running. I ran and ran until all the other kids left. It was getting late and dark. My mother arrived after her work and wanted to pick me up. The coach explained to her and asked her to stay a bit longer so that I can finish 14 laps. My mother was a bit afraid of my coach as well and she was standing there watching me. There were only three of us in the school

at that moment. It was so quiet and I could hear myself crying so loud. My coach kept saying "keep breathing, do some stretching, then continue. Don't stop!". My mother and the coach were counting laps for me.

"Perseverance is victory!" my mother was cheering for me, and she kept repeating that. That's all I heard when I was running alone in the school at that night. I don't remember when I finished the running, but I remembered that my mother and coach were very proud of me in the end. I still remember this story, and I know the experience in the ping-pong club shaped me when I was growing up.

The 16th of September 2015 was the day I arrived in Belgium for the first time. It was also my first time to take a flight and travel out of my home country China. I just became 21 years old and got an opportunity to do my bachelor thesis in the lab of prof. Luc Claesen (UHasselt). I made a phone call from Luc's office to my family in China immediately when I arrived. My parents were happy but also very worried for me. Cultural difference and language barrier were really difficult for me at that moment. Missing home and food were even harder. Gradually, I noticed that I really enjoyed doing research.

Luc once asked me "what do you want to do in the future?", I had no answer for that question at that moment. He said "I think you could be a very good researcher in the future! Why not do a PhD?" I was so shocked, as I have never thought about this and nobody told me that before. Then after 2 years, I started my PhD training in the group of prof. Bart De Moor (ESAT - KULeuven). Professor Bart invited me to join his group after I did my master thesis defense. I am very grateful for professor Bart to give me this opportunity.

A look back on these 5 years of my PhD,

I was doubting myself all the time;

I was frustrated and stressful about the progress of different projects;

I was so impatient and angry to myself when I was stuck or progressing too slow.

Yet in the mean time,

I got the excitement and goosebumps when I read a great paper;

I felt so accomplished when I figured out a solution and it worked;

I got so much joy when I communicated with other researchers interested in my work;

I felt so motivated when I knew that my work will be impactful for the society.

I found what I really enjoy doing - pursuing science and technology to help cure human disease.

There are a couple of things I learned from this journey: (1) I should be more patient to myself; (2) 'to let it go' is sometimes more difficult than 'keep holding on to it'; (3) It is not the end of the world, as professor Bart often told me "Wanqiu, don't worry, no panic!" To me, a PhD degree is a stubbornness degree. A PhD student will fail years and years, again and again. But eventually you accomplish it. "Perseverance is victory!" as my mother said to me when I was 9 years old.

However, perseverance does not always lead to the victory, you'll need a bit of luck. I am very lucky that I met so many great people in my life. They believed in me; they encouraged me; they helped me. I wouldn't be able to finish and enjoy this PhD without those people.

Firstly, I want to thank my promoter prof. Bart De Moor for providing me this great opportunity to do PhD in his group. Thank you for introducing me to some very interesting and challenging projects. Thank you for supporting me when I was stuck in the projects and telling me "don't worry about it". I still remember that you sent me messages of worry when I was on holiday in China during the start of the covid-19 pandemic.

I also want to thank all my jury members for the highly constructive comments on my thesis. I am very lucky to have a 'all-star' examination committee. Thank you prof. Johan Suykens and prof. Jesse Davis, as the giants in machine learning, you took the time and gave me valuable suggestions and guidance through my PhD. Thank you prof. Ron Heeren and prof. Martina Marchetti-Deschmann, as the pioneer in the mass spectrometry imaging community, you gave me so much new insights and encouragement.

I wouldn't finish this journey without Dr. Marc Claesen. Marc, you are my mentor, my friend, my big brother. You were my daily supervisor when I did my master thesis, then you became my co-promoter when I did my Phd. You taught me how to become a better scientist. I've learned so much from you. You helped me get my own PhD funding and introduced me to so many exciting (but tough!) projects. You always believed in me and supported me even when I didn't believe in myself. You were always so patient and encouraging to me. You told me that I should be proud of myself and I am working on that! Thank you for trusting me and hiring me! I am looking forward to working as a real 'Aspectos'!

I want to thank Dr. Nico Verbeeck for the great support during my PhD. Nico, you are my mentor and my ginger brother. I learned from you to always try to

be careful in every detail of my research. You are always so kind and thoughtful to everyone. When I was panicking, you calmed me down. Thank you for the encouragement and comfort. Thank you for driving me from Leuven to Genk for almost a year. I always enjoyed talking to you.

I want to thank everyone in the STADIUS group, especially Ida, John and Elsy for all the kind help. I want to thank my previous colleague Dr. Xi Shi for giving me so many valuable suggestions and helping me through some difficult times. I am glad to have so many nice colleagues: Tina, Bram, Willem, Christof, Oliver, Mauricio, Katrien, Hannes, Giulia, Melanie, Lola... Thank you all for giving me such nice time in BDM's group!

I want to thank prof. Dr. Maarten Naesens and Dr. Elisabet Van Loon for one of the best collaborating experience that I had with UZLeuven. You were so devoted to the project and gave me so much time and patience, especially when explaining medical knowledge to me. You always gave me valuable suggestions and encouraged me. You respected my work and were never shy with giving me positive comments. I am so happy and proud to see that our collaboration led to a nice publication on JAMA Network.

I will certainly not forget the great AA team. I am so lucky to work with those great, kind colleagues. Dr. Alice Ly, thank you for giving me so many useful suggestions and always supporting me. Thank you for sharing the burden on my shoulders when I needed it. It is a great pleasure to work with you and you inspired me a lot. Thank you Thomas, Peter, Els, Maria, Heath, Baruch... for all the encouragement and nice working experience. I am very happy to join this team and keep working with you all in the future!

I want to thank prof. Luc Claesen for being my first promoter of my research journey. Thank you for giving me the opportunity to come to Belgium to study, encouraging me to start the research journey, writing me the best recommendation letter and accidentally introducing the love of my life to me (mijn beste Belgische cadeau).

Finally I want to thank all my friends and family. It was certainly not easy to do a PhD in a foreign country. Fortunately, I have made some great friends in Belgium and they gave me such a great time. I am lucky to have my dear parents in-law, mama Julitt and papa prof. Dr. Paul Dendale by my side. Bedankt dat jullie me in je familie hebben opgenomen; bedankt dat jullie voor me zorgen; bedankt dat jullie me begrijpen; bedankt dat jullie trots op me zijn en voor me juichen. Bedankt familie Dendale en familie Camps dat ik een deel van jullie familie mag zijn.

I wouldn't achieve any of those without my parents. During covid-19, I was not able to visit them (for 3 and half years) and I am so glad that I went back

home to visit them again last month.

我要感谢我的父母，他们是最平凡却又是最伟大的人。感谢你们为我提供了你们所能给予的一切。感谢你们一直以来对我的无条件的爱与包容。老爸，你用你自身的励志故事激励着我，你以身作则让我知道了学习的机会来之不易。老妈，感谢你让我去飞，去闯，我知道你对我的思念之情如此之深。以后我会多陪伴在你们身边。

In the end, I want to thank my personal technical support, my best friend, my love of the life, my husband Thomas. Thomasje, thank you for being my rock for the past great 6 years of my life; thank you for making me laugh by your stupid jokes everyday; thank you for cooking delicious food for me when I was buried with stressful work and deadlines; thank you for cleaning the dishes and the house when I wanted to take a break from the work; thank you for learning Chinese for me and my family; thank you for understanding everything and always being supportive. I could not do it without you.

Wanqiu Zhang

张琬秋

Leuven, Belgium

August, 2023

Abstract

Mass spectrometry imaging (MSI) is a promising, label-free molecular imaging technology, which enables mapping the spatial distribution of hundreds to thousands of biomolecules in a tissue within a single experiment. Additionally, as it is a label-free and untargeted molecular imaging technology, it does not require prior knowledge about the tissue or target-specific reagents. Due to those outstanding characteristics, MSI is becoming increasingly popular and rapidly adopted for various applications such as biomarker discovery, clinical diagnostics, and drug delivery studies.

The generated data from a single experiment is complex, large and high-dimensional, posing great challenges in computational analysis in MSI. Therefore, this thesis focuses on the development and application of computational methods to MSI data. More specifically, various advanced machine learning algorithms and multimodal strategies were developed and applied, revealing significant improvement to current computational analysis in MSI. The advent of machine learning and deep learning has greatly facilitated complex spatial, high-dimensional data analysis. Thus, those deep learning and machine learning based models were applied in this thesis to improve the performance of unsupervised learning (clustering) and supervised learning (classification) tasks in MSI. Additionally, an advanced multimodal integration pipeline was proposed, enabling the direct correlation and comparison of spatial multi-omics datasets.

Firstly, we focused on the unsupervised learning methods, which supports exploring the underlying patterns within the complex MSI data. We applied a pre-trained neural network to extract high-level features from ion images in MSI data, which then were used to cluster ion images in MSI data based on the spatial expressions. We compared the proposed strategy with the standard clustering pipeline, which uses regular ion images as input of the clustering tasks. The results show the improvement of the proposed deep learning based clustering strategy in more fine-grained clusters, and greater consistency in cluster assignment. Additionally, we introduced the relative isotope ratio

metric to quantitatively evaluate clustering quality and benchmarked our results on public data. The results show the potential of resulting neural network interpretation of the ion images, which can be extended into any MSI-focused unsupervised or supervised machine learning pipeline.

One of the promising applications in MSI is clinical decision support and, ultimately, diagnostics, in which MSI data is used to build a classifier to improve the pathologists' diagnosis and tumor typing. Most MSI-focused studies collect microscopy data in tandem with MSI data, meaning that this imaging modality (microscopy) is generally omitted in the downstream analysis. However, there is very little study about using microscopy data together with MSI for clinical diagnosis, namely, the classification task. Therefore, we proposed a multimodal MSI classification pipeline in the melanoma study that uses a pre-trained neural network to extract the morphological features from microscopy data. The resulting morphological features were integrated with MSI data for improving downstream classification tasks, therefore, improving melanoma diagnosis. The multimodal pipeline was compared with other unimodal strategies, resulting in significant improvement based on nested-cross-validated ROC-AUCs. More importantly, as a pre-trained neural network requires no training nor labels, this multimodal pipeline is extremely efficient and flexible. Thus, it can be readily applied in other experimental settings where microscopy is acquired in tandem with MSI.

Finally, we extended the multimodal MSI pipeline to spatial multi-omics applications. We obtained valuable spatial multi-omics datasets from high-risk prostate cancer patients, including in situ lipidomics via mass spectrometry imaging (MSI) and spatial transcriptomics (ST) data. The proposed multimodal pipeline successfully registered and aligned these two complex and high-dimensional datasets from the serial sections, enabling direct comparison and correlation of lipidomics and transcriptomics. Thus, the proposed spatial multi-omics integration pipeline is a key stepping stone for any other downstream analysis of the integrated datasets, empowering the mapping of the molecular heterogeneity and the understanding of their intricate connections for gaining better understanding of cancer biology and a better patient stratification.

Beknopte samenvatting

Massaspectrometrie-beeldvorming (MSI) is een veelbelovende, labelvrije moleculaire beeldvormingstechnologie. Het laat toe om de ruimtelijke verdeling van honderden tot duizenden biomoleculen in een weefsel binnen een enkel experiment in kaart te brengen. Bovendien is er, aangezien het een labelvrije en ongerichte moleculaire beeldvormingstechnologie is, geen voorafgaande kennis over het weefsel of doelspecifieke reagentia vereist. Vanwege die uitstekende eigenschappen wordt MSI steeds populairder en wordt het snel toegepast voor verschillende toepassingen, zoals het ontdekken van biomarkers, klinische diagnostiek en onderzoeken naar medicijnafgifte.

De gegenereerde gegevens van een enkel experiment zijn complex, groot en hoog-dimensionaal, wat grote uitdagingen vormt bij computationele analyse in MSI. Daarom richt dit proefschrift zich op de ontwikkeling en toepassing van computationele methoden op MSI-gegevens. Meer specifiek werden verschillende geavanceerde algoritmen voor machinaal leren en multimodale strategieën ontwikkeld en toegepast, die een aanzienlijke verbetering van de huidige computationele analyse in MSI aan het licht brachten. De komst van machinaal leren en "deep learning" heeft complexe ruimtelijke, hoog-dimensionale data-analyse enorm vergemakkelijkt. Die op deep learning en machine learning gebaseerde modellen werden dus in dit proefschrift toegepast om de prestaties van "unsupervised learning" (clustering) en "supervised learning" (classificatie) taken in MSI te verbeteren. Daarnaast werd een geavanceerde multimodale integratie pijplijn voorgesteld, die de directe correlatie en vergelijking van ruimtelijke multi-omics datasets mogelijk maakt.

Ten eerste hebben we ons gericht op de "unsupervised learning" methodes (de leermethoden zonder toezicht), die het verkennen van de onderliggende patronen binnen de complexe MSI-gegevens ondersteunen. We hebben een vooraf getraind neuraal netwerk toegepast om functies op hoog niveau te extraheren uit ionenbeelden in MSI-gegevens, die vervolgens werden gebruikt om ionenbeelden in MSI-gegevens te clusteren op basis van de ruimtelijke

uitdrukkingen. We hebben de voorgestelde strategie vergeleken met de standaard clustering pijplijn, die reguliere ionenbeelden gebruikt als input voor de clustering taken. De resultaten tonen de verbetering van de voorgestelde, op "deep learning" gebaseerde, clustering strategie in meer fijnmazige clusters en een grotere consistentie in clustertoewijzing. Daarnaast hebben we een relatieve isotopenverhouding geïntroduceerd om de kwaliteit van de clustering kwantitatief te evalueren en hebben we onze resultaten vergeleken met openbare gegevens. De resultaten tonen het potentieel van de resulterende neurale netwerkinterpretatie van de ionenbeelden, die kan worden uitgebreid naar elke MSI-gerichte "unsupervised" of "supervised" pijplijn.

Een van de veelbelovende toepassingen in MSI is klinische diagnose, waarbij MSI-gegevens worden gebruikt om een classificator te bouwen om de diagnose en tumortypering van de patholoog te verbeteren. De meeste op MSI gerichte onderzoeken verzamelen microscopiegegevens samen met MSI-gegevens, wat betekent dat deze beeldvormende modaliteit (microscopie) over het algemeen wordt weggelaten in de stroomafwaartse analyse. Er is echter zeer weinig onderzoek gedaan naar het gebruik van microscopiegegevens samen met MSI voor klinische diagnose, namelijk de classificatietaak. Daarom hebben we een multimodale MSI-classificatiepijplijn voorgesteld in de melanoomstudie die een vooraf getraind neuraal netwerk gebruikt om de morfologische kenmerken uit microscopiegegevens te extraheren. De resulterende morfologische kenmerken werden geïntegreerd met MSI-gegevens voor het verbeteren van stroomafwaartse classificatietaken, waardoor de melanoomdiagnose werd verbeterd. De multimodale pijplijn werd vergeleken met andere unimodale strategieën, wat resulteerde in een aanzienlijke verbetering op basis van "nested-cross"-gevalideerde ROC-AUC's. Wat nog belangrijker is, aangezien een vooraf getraind neuraal netwerk geen training of labels vereist, is deze multimodale pijplijn uiterst efficiënt en flexibel. Het kan dus gemakkelijk worden toegepast in andere experimentele omgevingen waar microscopie wordt verkregen in combinatie met MSI.

Ten slotte hebben we de multimodale MSI-pijplijn uitgebreid naar ruimtelijke multi-omics-toepassingen. We hebben waardevolle ruimtelijke multi-omics-datasets verkregen van patiënten met een hoog risico op prostaatkanker, waaronder in situ lipidomics via MSI en ruimtelijke transcriptomics (ST)-gegevens. De voorgestelde multimodale pijplijn heeft deze twee complexe en hoog-dimensionale datasets van de seriële secties met succes geregistreerd en uitgelijnd, waardoor directe vergelijking en correlatie van lipidomics en transcriptomics mogelijk is. De voorgestelde ruimtelijke multi-omics-integratie-pijplijn is dus een belangrijke opstap voor elke andere stroomafwaartse analyse van de geïntegreerde datasets, waardoor het in kaart brengen van de moleculaire heterogeniteit en het begrip van hun ingewikkelde verbindingen mogelijk

wordt gemaakt, en het begrip van kankerbiologie en patiënt stratificatie wordt verbeterd.

List of Abbreviations

AI Artificial Intelligence. 8

ANNs artificial neural networks. 33

AUC area under the curve. 71

CODEX co-detection by indexing. 19

DESI desorption electrospray ionisation. 6

FPR false positive rate. 71

FTICR Fourier transformation ion cyclotron resonance. 6

GlcNAc N-acetylglucosamine. 20

H&E haematoxylin and eosin stain. 23

HDBSCAN Hierarchical Density-based Spatial Clustering for Applications with Noise. 16

IMC imaging mass cytometry. 19

LAESI laser ablation electrospray ionisation. 6

MALDI matrix-assisted laser desorption-ionisation. 2

MIBI multiplexed ion beam imaging. 19

ML machine learning. 13

MS mass spectrometry. 1

nanoDESI nanospray desorption electrospray ionisation. 6

NMF nonnegative matrix factorization. 14

PCA principal component analysis. 14

PLS partial least-squares. 26

PR precision-recall. 74

RF random forests. 13

RIR relative isotope ratio. 42

ROC receiver operating characteristic. 71

scRNA-Seq single-cell RNA sequencing. 2

SHAP Shapley additive explanations. 13

SIMS secondary ion mass spectrometry. 6

STARmap spatially-resolved transcript amplicon readout mapping. 19

SVM support vector machine. 13

t-SNE t-distributed stochastic neighbour embedding. 15

TIC total ion count. 11

TOF time of flight. 6

TPR true positive rate. 71

UMAP uniform manifold approximation and projection. 15

VEM Vector Embedding of Morphology. 70

Contents

Abstract vii

Beknopte samenvatting ix

List of Abbreviations xiv

List of Symbols xv

Contents xv

List of Figures xix

List of Tables xxi

1 Introduction 1
- 1.1 Mass Spectrometry Imaging technology . . . 1
- 1.2 Key advantages of MSI . . . 4
- 1.3 Flavors of MSI and their impact on data . . . 5
- 1.4 Trade-offs and challenges in MSI . . . 8
- 1.5 Aim of this thesis . . . 8
- 1.6 Data analysis in MSI . . . 9
  - 1.6.1 Signal processing (Pre-processing) . . . 11
  - 1.6.2 Supervised analysis in MSI . . . 12
  - 1.6.3 Unsupervised analysis in MSI . . . 14
- 1.7 Spatial Multi-Omics . . . 18
  - 1.7.1 Current applications in MALDI-MSI and challenges . . . 19
  - 1.7.2 Data acquisition . . . 23
  - 1.7.3 Image registration . . . 23
  - 1.7.4 Measurements matching . . . 24
  - 1.7.5 Downstream multimodal data analysis . . . 26
- 1.8 Structure of the thesis . . . 28

2 Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning 29
2.1 Abstract . . . 31
2.2 Introduction . . . 31
2.3 Methods . . . 36
2.3.1 Materials . . . 36
2.3.2 Ion Image Clustering . . . 37
2.3.3 Evaluation of Clustering Results . . . 42
2.3.4 Visualization of Clustering Results . . . 43
2.4 Results and Discussion . . . 43
2.4.1 Generation of neural ion images . . . 44
2.4.2 Human lymph node dataset . . . 45
2.4.3 Mouse kidney data . . . 48
2.4.4 Benchmark on ColocML data . . . 54
2.5 Conclusions . . . 57
2.6 Acknowledgements . . . 58

3 Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma 59
3.1 Abstract . . . 60
3.2 Introduction . . . 60
3.3 Materials and Methods . . . 63
3.3.1 Tissue Collection . . . 63
3.3.2 Histopathology Sample Preparation and Data Acquisition 63
3.3.3 MALDI IMS Sample Preparation and Data Acquisition 65
3.3.4 Data Processing and Sample Selection . . . 65
3.3.5 Data Analysis . . . 66
3.4 Results . . . 69
3.4.1 UMAP visualization of intermediate data . . . 70
3.4.2 Classification performance . . . 71
3.5 Discussion . . . 74
3.6 Conclusion . . . 76
3.7 Acknowledgement . . . 76

4 Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer 78
4.1 Abstract . . . 79
4.2 Introduction . . . 79
4.3 Results . . . 81
4.4 Discussion . . . 98
4.5 Methods . . . 100
4.5.1 Data acquisition . . . 100
4.5.2 Data preprocessing . . . 104

4.6 Acknowledgement 108

5 Details of applied methods 110
5.1 Pre-trained neural networks 110
5.1.1 Xception model 110
5.1.2 SimCLR model 112
5.1.3 Benchmark on the Camelyon dataset 114
5.1.4 Discussion 117
5.2 Downstream classification model 118
5.2.1 Linear SVM model 119
5.2.2 Model selection results 120
5.3 Uniform Manifold Approximation and Projection 124
5.3.1 Method 124
5.3.2 Parameters 126
5.3.3 Limitations 128

6 Conclusions and future perspectives 129
6.1 Contributions 129
6.2 Remaining computational challenges in MSI 131
6.3 Future outlooks 133
6.3.1 It is all about *good data* 133
6.3.2 Development of computational tools for spatial multi-omics 134

A Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning 137
A.1 Relative isotope ratio (RIR) metric examples 137
A.2 Used data sets from ColocML 140
A.3 Examples from ColocML data 141

B Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer 145
B.1 Clinical data for the eight patients used in the SMOx study . . 145
B.2 Averaged MALDI-2 spectra 147
B.3 Correlation heatmaps comparing pathology annotations versus spatial-omics expression data 147
B.4 Pathological annotations of all samples 149

Bibliography 153

About the author 185

List of publications 187

List of Figures

1.1 Workflow for MALDI MSI experiment . . . 3
1.2 Workflow for different ionization sources in MSI . . . 7
1.3 Current opinion in biotechnology about MSI methods . . . 9
1.4 MSI dataset visualization . . . 10
1.5 Steps of a typical data analysis workflow in MSI . . . 11
1.6 Unsupervised machine learning approaches in MSI . . . 17
1.7 Descriptions of different spatial molecular profiling technologies. 19
1.8 Multi-view representation learning schemes. . . . 20
1.9 Example of the challenging image registration case . . . 22
1.10 Schematic of challenges in multimodal image registration . . . 25
1.11 Typical processing and analysis workflow for multimodal MALDI-MSI . . . 27

2.1 Overview of the ion image clustering workflow . . . 38
2.2 Patch generation example on an ion image . . . 45
2.3 UMAP 3d scatter plots of clustering results . . . 49
2.4 Ion image clustering results of the lymph node dataset . . . 50
2.5 Overlay visualization of clusters from lymph node data . . . 51
2.6 Ion image clustering results of the mouse kidney dataset . . . 53
2.7 Overlay visualization of clusters from mouse kidney data . . . 54
2.8 Benchmark results on open MSI datasets . . . 56

3.1 Overview of the histology-guided IMS sample preparation and processing workflow . . . 64
3.2 Multimodal data analysis workflow on an example tissue . . . 67
3.3 UMAP 2-D visualization of each input data across pipelines . . 72
3.4 Classification results on training and held-out test sets . . . 73
3.5 Misclassified cases by unimodal algorithms from the held-out test set . . . 75

4.1 Study design and overview of the spatial multi-omics integration pipeline (SMOx). 84
4.2 Integration of spatially resolved multi-omics reveals spatially correlated lipid-gene pairs in a prostate cancer sample. 87
4.3 Correlation between histological and molecular profiling. 89
4.4 Spot deconvolution reveals cell type specific transcriptional and lipidomic profiles. 91
4.5 Integration of spatially resolved omics unveils molecular heterogeneity in PIN. 93
4.6 Integrated spatial multi-omics supporting histological annotation of ambiguous tissue areas. 95
4.7 Implementation of the spatial multi-omic workflow across the entire prostate cancer sample cohort. 97

5.1 Comparison of available models in the Keras package 111
5.2 SimCLR model architecture 114
5.3 Comparison of pre-trained models 115
5.4 Camelyon 16 challenge data 117
5.5 UMAP embeddings of the outputs generated from pre-trained models 118
5.6 UMAP 3-d embeddings with different hyperparameter settings 127

A.1 Detected as isotope in lymph node data 138
A.2 Example ion images (belongs to the same cluster but not the isotope) 138
A.3 Detect as isotope in Mouse kidney data 139
A.4 Example ion images (belongs to the same cluster but not the isotope) 139
A.5 datasetId = '2017-01-26_09h08m49s' 141
A.6 datasetId = '2017-07-18_17h21m08s' set 1 142
A.7 datasetId = '2017-07-18_17h21m08s' set 2 143
A.8 datasetId = '2017-07-18_17h21m08s' set 3 144

B.1 Clinical data for the eight patients used in the SMOx study . . 146
B.2 Averaged MALDI-2 spectra. 147
B.3 Correlation heatmaps comparing pathology annotations versus spatial omics expression data. 148
B.4 Different quality pathological annotations 150
B.5 Different quality pathological annotations 151
B.6 Different quality pathological annotations 152

List of Tables

2.1 Overview of experimental results per configuration 47

5.1 Cross-validated classification results for benchmarked pre-trained models 119
5.2 Nested cross-validation results of unimodal MSI pipeline . . . 122
5.3 Nested cross-validation results of unimodal microscopy pipeline 122
5.4 Unimodal microscopy pipeline of multimodal MSI and microscopy pipeline 123

Chapter 1

Introduction

Mass Spectrometry Imaging Takes Off

'*What if it were possible to image the locations of multitudes of molecules in a tissue or even a single cell and to determine their chemical identities without using any fluorescent labels or antibodies?*' – **Mass Spectrometry Imaging** was mentioned as one of the *'Methods to watch'* in Nature Method (2018) [80].

1.1 Mass Spectrometry Imaging technology

Mass spectrometry (MS) is a widely and successfully used technology to detect various analytes, including metabolites, lipids, peptides, and proteins in tissues and cells, that has been developed and matured over the past decades [263]. The advent of soft ionization techniques such as matrix assisted laser desorption ionization (MALDI) [131, 243] and electrospray ionization (ESI) [93] enabled MS to achieve fast, sensitive performance, triggering increasing relevance of MS in biochemical and medical applications [260]. It is also recognized as the most widespread and powerful tool for bulk metabolomics [77].

Although MS provides an excellent combination of molecular sensitivity, specificity and coverage, it lacks spatial information. Spatial information is crucial for being able to understand the association of overall molecular changes with spatially localized aberrations, for example tumors, as tissues and organs are groups of cells that are highly spatially organized to facilitate specific functions. The importance of spatial information is also increasingly recognized in the fields of transcriptomics and proteomics. Single-cell RNA sequencing

(scRNA-Seq) enables the identification of rare cell populations and recognition of genes of interest related to certain cell types. However, their spatial information is lost in these experiments [280, 258]. In the study of cancer, the lack of the information about spatial localization and context prevents us from obtaining a holistic understanding of underlying tumor biology. Bulk analysis fails to provide the origin of the data readout, limiting the biological interpretation, especially for those spatially complex samples (e.g., tumor microenvironments). Without the spatial information, one can not assign the detected molecular information to cells and/or their surrounding regions [263]. Fortunately, spatial resolved transcriptomics were later developed and selected as 'the method of the Year 2020', as it can not only extract transcriptomic data but also the positional context of those cells in a tissue [173].

In 1997, prof. Richard Caprioli and coworkers applied the technique of matrix-assisted laser desorption-ionization (MALDI) based mass spectrometry in particular to localize proteins in tissue sections [54]. In their breakthrough publication, they demonstrated the spatial distribution of hundreds of proteins in a single experiment based on a single section. Since then, the focus of this technique has shifted toward the detection of different analytes, such as lipids, metabolites, small molecules and glycans in various applications. This promising technology has been successfully applied in many fields, such as clinical diagnosis, predicting prognosis, biomarker discovery and many pharmacological studies e.g., pharmacokinetics or pharmacodynamics, toxicology.

Mass spectrometry imaging measures the spatial distribution of molecules through systematically locally ionizing a tissue sample and then analyzing the resulting ions via mass spectrometry. In a typical MSI experiment, as shown in Figure 1.1, the surface of a tissue sample is divided into a virtual grid of pixels. In each pixel, molecules are desorbed and ionized via a laser or other ionization techniques, followed by the generation of a mass spectrum. Each of the peaks in the mass spectrum can be interpreted as showing the relative intensity of a molecule within that pixel. One can plot its spatial distribution and abundance in the sample by extracting the intensities of selected m/z from each mass spectrum and arranging them into an image based on the pixel location of the respective mass spectrum. This resulting image is referred to as an m/z image or an ion image [226].

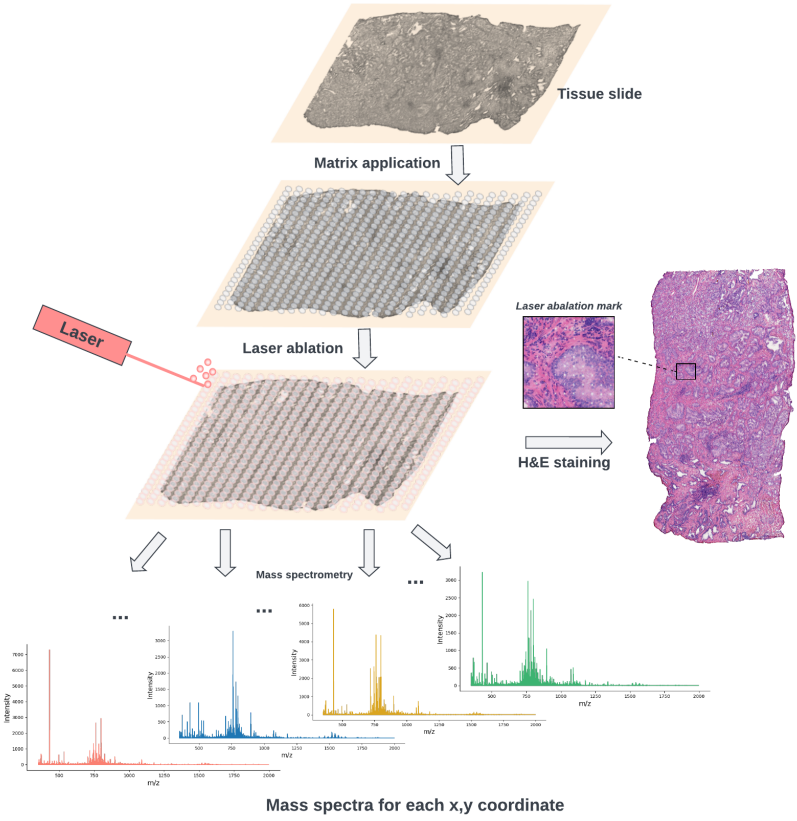


Figure 1.1: Workflow for MALDI MSI experiment. A typical workflow for fresh frozen tissue samples. After cutting and mounting the tissue section on a conductive target, the matrix is applied in an ordered array across the section. Then each mass spectrum is acquired at each (x,y) grid point on the tissue. Since MALDI MSI keeps tissue sections largely intact throughout the experiment, the same tissue can be stained after MS for pathological study. The figure is inspired by [226].

1.2 Key advantages of MSI

Analyzing the spatial distribution of analytes within a tissue

The discovery of spatial molecular distribution in a disease enables improved evaluation of, for example, the aggressiveness of the disease in the distinct areas of the tissue, and the outcome predictions of the patient. With this strength, MSI has become popular and widely applied in many clinical applications with long-term clinical potential, as it provides deeper and more fundamental insights of molecular activities including proteins, lipids and other biological compounds in the development and progression of cancer.

Applications in the field of cancer study [102, 226] can be divided into 3 categories: diagnosis [12, 205, 135], prognosis [172, 99, 161], and drug response [214]. In diagnostic studies, different tissue types are often compared in order to help pathological evaluations, by identifying new molecular markers that can distinguish different phenotypes. While in prognostic studies, the goal is to uncover molecular markers associated with specific patient outcomes (e.g., survival time, time to relapse). For drug response applications, the goal is to identify markers that correlate with the effectiveness of a therapeutic regime for a patient (e.g., finding a molecular signature that can be used to discriminate between responders and non-responders).

Given its roots in mass spectrometry, MSI also provides access to analytes that cannot be measured using most other platforms, especially in a highly multiplexed sense. As an example, MSI has become widely accepted within preclinical pharmaceutical research to localize drug metabolites in tissue, which is critical to understand whether said drug metabolites (a) get formed at all and (b) reach the right locations in tissue [58].

Enabling multimodal analysis on the very same sample

As the histology of tissues remain largely intact throughout the MSI experiment, the neighboring section or even the very same section can be stained afterwards and coregistered with MSI data. Manual annotations can be done on the stained sample, i.e., microscopy data, in use of further analysis, enabling multimodal analysis. Incorporating histological and morphological information with MSI is crucial for biomedical research [225], and it is particularly important and popular in the current MSI field which will be addressed in more detail in the section 1.7.

Enabling unbiased, label-free molecular imaging

Another key advantage of MSI is that it enables unbiased analysis of tissue sections without requiring prior knowledge about the tissue or target-specific reagents – a true label-free, untargeted molecular imaging technology. This invaluable untargeted characteristic of MSI makes itself suitable for exploratory studies. For example, novel biological compounds that are associated with specific tissue types or disease phenotypes may be found and identified, which may lead to the discovery of diagnostic and prognostic biomarkers of a certain type of cancer and the determination of effective therapies. On the other hand, this untargeted, label-free molecular imaging technology generates larger and more complex datasets, posing great challenges in computational analysis.

These challenges include, but are not limited to:

- Extremely high plexity within the context of spatial omics. Depending on instrumentation, individual raw spectra may have 10,000s to millions of peaks, many of which do not contain relevant signals. Such high dimensionality presents numerous challenges in terms of computational performance, data interpretation and overall downstream analysis.
- Depending on instrumentation, additional analytical separations may be available (e.g., ion mobility [79] or tandem mass spectrometry (IMS and MS/MS [1], respectively)), which further increase the complexity and mathematical dimensionality of the resulting data, by adding additional physical axes to the data.
- MSI's label-free nature is a big asset, but often induces the need for explicit identification, i.e., to map specific spectral fingerprints into identified molecules. This step may or may not use information from additional experiments such as liquid chromatography (LC), tandem mass spectrometry (MS/MS), ion mobility spectrometry (IMS) or combinations thereof. Performing such identifications requires specific data analysis approaches and searches within molecular databases.

1.3 Flavors of MSI and their impact on data

A mass spectrometer consists of (1) ion source: where the analyte molecules are desorbed from their medium, and ionized; (2) mass analyzer: where ions are separated based on their mass-to-charge ratio; (3) ion detector: where ions are detected and measured. The most popular ion source type used for imaging is matrix-assisted laser desorption/ionization (MALDI), where the sample is

mixed with an energy-absorbing chemical matrix and then irradiated by a laser. Its popularity owes to its versatility regarding spatial resolution and molecular coverage.

However, the drawback of the MALDI method is that the sample preparation requires expert knowledge, can be time consuming and may induce artifacts [144, 74]. Due to this drawback, other imaging-compatible ionization techniques have been developed. For example, other well-known source types include the desorption electrospray ionization (DESI) [66], where the desorption is made by a solvent spray directed towards the sample instead of using a laser as MALDI, resulting in no requirement of a matrix thus generally requiring little to no sample preparation.

There are many other ionization sources in MSI, such as secondary ion mass spectrometry (SIMS) [94, 201], nanospray desorption electrospray ionization (nanoDESI) [145], laser ablation electrospray ionization (LAESI) [190], etc. Brief introduction and comparison of those different techniques are illustrated in Figure 1.2.

In this thesis, we focused on MALDI based imaging data, which is the most widely-used ionization method in MSI.

After analytes are ionized from the sample, the mass analyzer characterizes them based on their molecular weight before they are detected. The most commonly used analyzer is time of flight (TOF), as it offers a wide mass range, relatively fast operation speed and high throughput. As the name suggests, TOF mass spectrometer characterizes and detects ions based on the time of their 'flight' along a certain distance. A popular alternative, called the Orbitrap [291], is an ion trap mass spectrometer where ions are trapped electrostatically and travel in an orbital motion around a central spindle-like electrode. Once the image current of the trapped ions is detected in the detector plates, they are converted to a mass spectrum by using the Fourier transform of the time-domain current signal. Compared with TOF, Orbitrap is capable of higher mass resolution, enabling separation of analytes with similar m/z values. Fourier transformation ion cyclotron resonance (FTICR) [171], as another popular mass analyzer, determines the m/z values of ions based on cyclotron frequency of the ions in a fixed magnetic field. This instrument currently is capable of providing the highest mass resolution (>1,000,000) [223].

The choice of the ion source, mass analyzer, and sample preparation in MSI will have a considerable impact on the data – They will predetermine which molecular species will be desorbed and thus substantially influence the biochemical content of the data. Ion source can affect the spatial resolution, for example, SIMS achieves at spatial resolution < 1 um in the subcellular scale

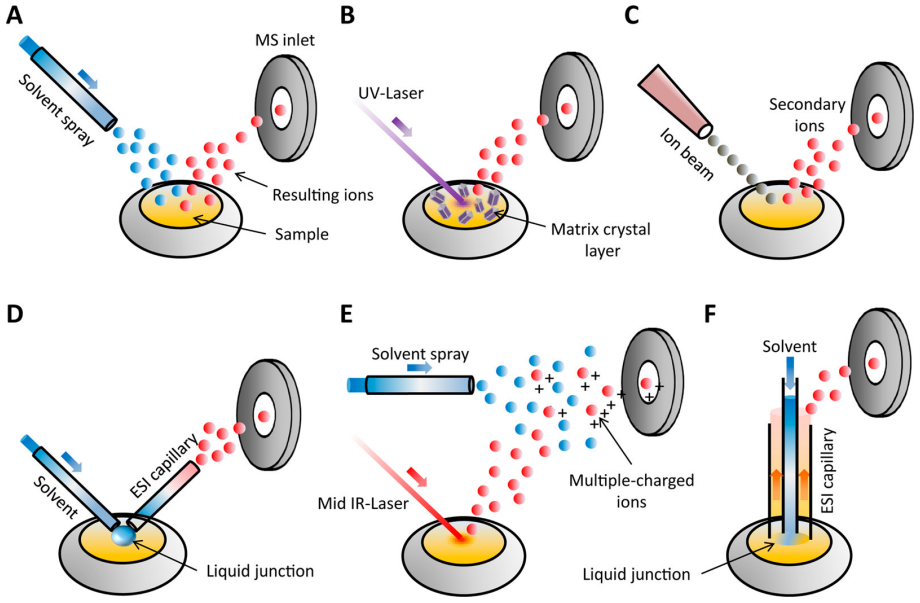


Figure 1.2: Workflow for different ionization sources in MSI. The upper row shows the established ionization sources: desorption electrospray ionization (DESI, A), matrix-assisted laser desorption/ionization (MALDI, B), and secondary ion mass spectrometry (SIMS, C). The lower row shows the emerging ionization technologies: nano-DESI (D), laser ablation electrospray ionization mass spectrometry (LAESI, E), and liquid extraction surface analysis (LESA, F). The figure is reprinted with permission from [11] Copyright 2015 American Chemical Society.

[232], 20-50um for MALDI [286, 139] and 100-200 um for DESI [283] routinely. The analyzer type (e.g., TOF, Orbitrap, FTICR) substantially affects the spectral properties, i.e., the mass-resolving power. This power pretermines how detailed the spectra are and whether two molecular species with similar m/z values can be distinguished from each other. For example, for TOF analyzer, the dimensionality of a (non-peak-picked) mass spectrum is usually on the order of 10^4, and over 10^5 for FTICR or Orbitrap.

1.4 Trade-offs and challenges in MSI

The ideal scenario for MSI would satisfy the '4S-criteria for performance' [223] as shown in Figure 1.3: speed, specificity (resolving power and mass accuracy), spatial resolution, sensitivity, in addition to mass range (range of investigable molecular class). However, there are many trade-offs. For example, high spatial resolution will induce smaller ablation areas thus the sensitivity will be limited; measurement speed per tissue needs to be balanced with specificity, spatial resolution and sensitivity. It would be beneficial if high spatial resolution could combine with high molecular specificity, as it would empower better insights of the construction of a cellular system from the sample [263]. However, it would also pose great challenges in computational analysis of MSI, as the data size will be even larger and the complexity of the data will be dramatically increased. There is still much room for MSI technology to improve those 5 criteria, especially the sensitivity (ability to detect low-intensity molecules) and molecule identification (particularly of metabolites) [80].

In parallel with the trend to further improve MSI instrumentation, more efforts need to be made in enhancing multi-site standardization [223]. As discussed in *'Why the metabolism field risks missing out on the AI revolution'* [83], the lack of high-quality data, and the inadequacy in multi-site & -platform standardization, are the crucial parameters that slows down the application of machine learning/ Artificial Intelligence (AI) in the metabolomics field. Standardized protocols have to be involved in all steps, from 'wet' lab e.g., sample preparation, data acquisition to 'dry' lab e.g., data analysis. In this way, a high level of reproducibility can be achieved [104]. Therefore, a unified effort at all levels is required, including from individual researchers, labs, institutes, universities for improving metadata completion, pre-cleaning data, data management and storage, etc. In this way, we can apply machine learning algorithms more effectively and take full advantage of AI in aiding science and human health.

1.5 Aim of this thesis

A key limitation within spatial omics in general and MSI in particular is the lack of easily accessible, scalable tools for data analysis. Despite the fact that the need for good data analysis tools is arguably the highest for MSI as the technology generates very complex data (i.e., label free, high plexity), available solutions have major limitations. While several software solutions exist, both commercial and open-source, they present significant practical constraints in terms of user friendliness (particularly for end-users without a computational

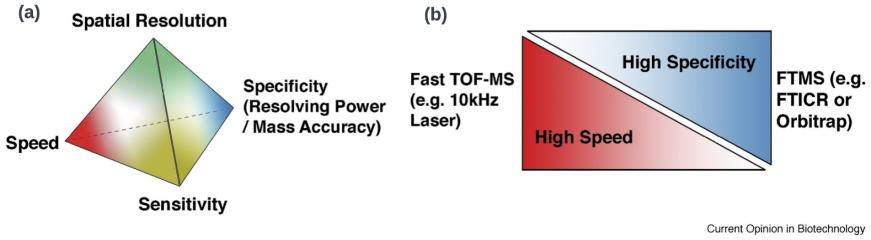


Figure 1.3: Current opinion in biotechnology: (a) MSI methods need to compromise between the '4S-criteria for performance' (speed, specificity, spatial resolution and sensitivity). (b) High-specificity FTICR MS (FTMS)-instruments, currently the best choice in pharmaceutical MSI, are complemented by much faster TOF instruments. This figure is reprinted from [223] with permission. Copyright Creative Commons CC-BY license.

background) and scalability (both in terms of data size as well as operational efficiency in high-throughput contexts).

This thesis aims to overcome some of these challenges among others. The methods we have designed throughout this work make a trade-off between depth of analysis and ease of use for end users, i.e., few tuning parameters to be set by experts, drop-in components to improve data analysis, etc. Therefore, we will first give an overview of computational methods that are currently used in MSI below.

1.6 Data analysis in MSI

The prepared tissue section is placed in the mass spectrometer and analyzed by an array of positions [14]. The resulting mass spectrometry imaging data typically consists of **a collection of mass spectra, each of which associates with a pixel location**. MSI data can be interpreted as a hyperspectral image with multiple ion channels, each of which corresponds to a certain mass-to-charge (m/z), and represents relative intensities of this m/z value. A mass-to-charge value can be seen as a molecule mass in a MSI experiment, as ions with a charge (z) of +1 prevail. The intensity of the m/z value represents the relative abundance of a compound with that m/z value. By selecting an m/z value across all measured spectra, the intensity of this m/z value across the tissue can be visualized as a heat-map image [14], see Figure 1.4. This heat-map image is coloured based on the intensity of this m/z of each pixel. When

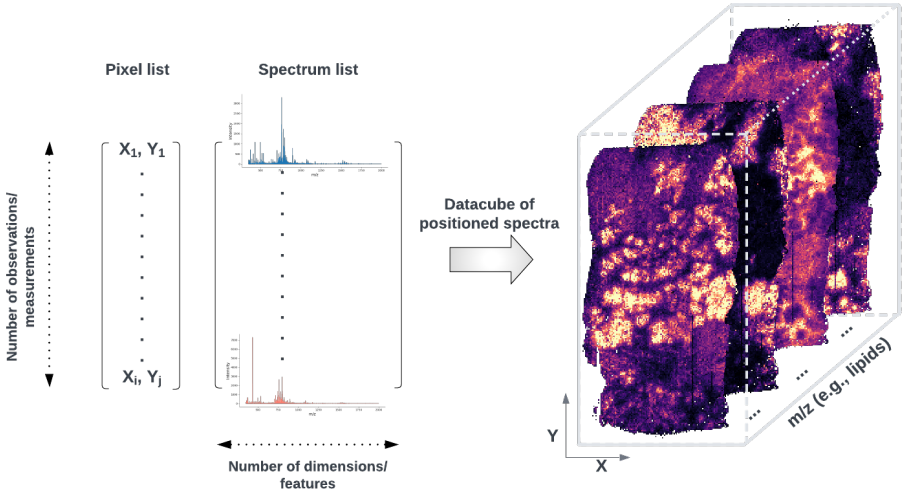


Figure 1.4: Mass spectrometry imaging experiments generate a collection of mass spectra acquired from a grid of pixels as representation of a tissue section surface. With a selected mass-to-charge (m/z) value or a peak, one can display its intensities from each acquired mass spectrum into an image, based on their corresponding pixel locations. Thus a MSI datacube is generated, with a number of pixels (= number of observations) and their associated mass spectra (features/dimensions).

the mass resolution is high enough, this m/z bin represents a single molecular ion species, of which the abundance can be visualized throughout the tissue. Thus, this m/z image is also referred to as an ion image.

The number of observations in an MSI experiment are the number of pixels, and the variables are the m/z values. A mass spectrum from a pixel is treated as an independent measurement/observation in most data analysis algorithms [14]. As reported by Alexandrov [15], there are usually between 10,000 and 100,000 pixels/observations in an MSI experiment. This is often limited by the operating speed/ throughput as 1-50Hz, where 1 Hz represents 1 spectrum/s. Based on the current progress of the instrument, over a million pixels generated in one experiment routinely will be possible in the next year or two. The dimensionality of the data or the number of m/z values depends on the mass spectrometer. For example, a TOF analyzer usually generates the number of m/z values on the order of 10^4 while over 10^5 for the Fourier transformation (FT) mass spectrometer (the non-peak-picked data).

As mentioned in a recent review of MSI [15], a typical data analysis

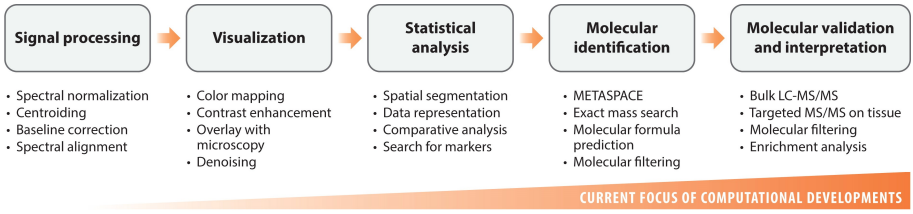


Figure 1.5: Steps of a typical data analysis workflow in MSI: signal processing, visualization, statistical analysis, molecular identification, molecular validation,and interpretation. Figure reprinted with permission from [15] Copyright (2020) Annual Reviews.

workflow in MSI includes signal processing, which is also named as pre-processing, visualization, statistical analysis, molecular identification and molecular validation and interpretation, see Figure 1.5. Typically, after data analysis, a list of m/z's of interest is determined, and which will be matched to a database of known molecules or analyzed via tandem MS (MS/MS) for identification.

1.6.1 Signal processing (Pre-processing)

Signal processing or pre-processing steps is a fundamental step for any data analysis. There are a number of typical pre-processing steps generally applied in MSI experiments, including spectral normalization, baseline correction, spectral smoothing, spectra recalibration and spectra centroiding [14].

Spectral normalization is an important part of MSI processing, and scales each spectrum to a factor in order to better compare the intensities between different spectra. The widely used and most reported method is total ion count (TIC) normalization, where the sum of all intensities (TIC) of a mass spectrum is calculated and all spectrum intensities are divided by the TIC value. There are still some discussions about this topic and other normalization methods are explored and implemented [74, 95, 209]. The choice of normalization method is still an open question as there is a lack of ground truth data which provides a known distribution of molecules and can be utilized to compare existing normalization algorithms objectively [15].

As mentioned in [15, 83, 104], except the dearth of ground truth data, the current challenge that hinders the computational developments also lies in a lack of a certain degree of standardization of MSI data pre-processing in

multiple labs/groups. Usually, scientists in different 'wet' biology labs choose different pre-processing pipelines based on their own empirical experiences, which prevent the comparison of the data between different labs.

1.6.2 Supervised analysis in MSI

One of the most important goals of processing a MSI dataset is to discover a list of m/z values which relates to the ions of interest. Those ions of interest may co-localize with a certain biomarker or a region of interest, e.g., an anatomical region. Ions of interest may also express differently in different regions of a single sample or in different samples, e.g., different expression in tumor region from the normal region in a sample, or even within a single heterogeneous tumor [14].

The first decision to be made in data analysis is: perform a **supervised** or an **unsupervised** analysis? This depends on whether there is available prior knowledge about the data, e.g., manually annotated labels, and whether they are used in the analysis. The core difference between these two analyses is that the former utilizes labeled input and output data, whereas the latter aims at revealing unknown latent variables in data without labels.

In MSI data analysis, the prior knowledge of the supervised analysis is usually from the pathological annotation on the same tissue section or serial section by the experts. Then the mass spectra from the pixels, which are localized in the annotated region by the pathologists, are labeled to that specific histopathological entity, which is named *histology-defined* analysis strategy in MSI. It should be noted that histology-defined data analysis is different from *histology-defined profiling* [67], as the former acquires mass spectra from all over the tissue section whereas the latter only obtains data from annotated regions of interest. In this way, histology-defined profiling speeds up the experiments and concentrates more on the pathological question [129].

Biomarker discovery

As the nature of MSI technology is non-targeted and label-free, a wide range of molecules can be detected in one sample, which is helpful for biomarker discovery. The most common use of a supervised analysis is to differentiate at least two groups of spectra that distinguish between pathological regions, e.g., by detecting m/z values that differentiate spectra of tumor areas from spectra from normal regions. This is a crucial step towards biomarker discovery. One of the strategies is **class comparison**, which seeks to

detect differentially abundant ions between different regions through dedicated experimental designs based on statistical principles [35].

From a machine learning perspective, the biomarker discovery task can be translated into a **feature importance ranking** problem. Recently, machine learning (ML) algorithms have been proposed to automatically detect the differentially expressed molecular species e.g., an interpretable ML workflow [251] via Shapley additive explanations (SHAP) [166, 167]. Tideman et al. [251] constructed a supervised ML model to classify mass spectra from different annotated regions, i.e., pixels are assigned to different classes as well as their corresponding mass spectra, and then applied an interpretable ML algorithm SHAP to determine the discriminative relevance and therefore potential biomarkers, i.e., find out which m/z values are important to the decision making of the classification model.

Classification - Diagnosis applications

Machine learning (ML) is an approach aiming to understand and construct models that 'learn' by leveraging data, thereby improving performance on different tasks [187]. Another popular supervised analysis in MSI via ML algorithm is classification, which is often applied to improve clinical diagnosis.

In addition to pathological annotations, the prior knowledge sometimes may also cover the patient outcome or the response to treatment, and supervised classification algorithms can be applied and help the diagnosis and prognostic markers identification. Classification methods like support vector machines (SVM) and random forests (RF) are commonly used in MSI supervised data analysis. Many studies have reported that these ML classification algorithms can accurately and reliably e.g., support the pathologist's diagnosis [12]; distinguish various cancer types [98, 183]; identify survival and recurrence-associated protein profiles in such as melanoma metastases [113] and sarcomas [162]; characterize tumor margins and microenvironment for more thorough tumor removal [195].

Inherited information leakage problem

Despite the fact that supervised ML algorithms applications play an important role in MSI, the proper model training and validation should be taken good care of by all the data scientists. In classification tasks, data should be carefully split into training, validation and independent held-out test datasets to avoid overfitting and information leakage.

Given a MSI dataset, different regions of pixels or groups of corresponding mass spectra are labeled as different classes such as tumor and normal. One should note that, due to the spatial nature of MSI data, pixels or **observations in a MSI data are not independent from each other**, and spatial correlation between these data points is expected to occur [57]. Additionally, one must be extra cautious when splitting the dataset, especially when the dataset contain observations from the same 'group', e.g., the same annotated region or the same patient, which should not be represented in both testing and training sets.

1.6.3 Unsupervised analysis in MSI

In contrast to supervised analysis, unsupervised analysis does not require any prior knowledge and seeks to reveal unknown underlying features within an MSI data. Instead of having a specific ion of interest or spatial region of interest in mind, these unsupervised methods are applied in order to provide an open-ended exploratory view on the data [265]. Most of the unsupervised learning methods in MSI can be characterized into dimensionality reduction, clustering of pixels or mass spectra (also known as *'spatial segmentation'*), and clustering of ion images or m/z images [15]. Figure 1.6 gives an overview of these unsupervised approaches, with an application to a MALDI Orbitrap MSI data acquired from a high-risk prostate cancer (PCa) patient.

Dimensionality reduction

Dimensionality reduction has primarily been applied in the early days through using component analysis or factorization methods, e.g., principal component analysis (PCA) and nonnegative matrix factorization (NMF), as they can visualize both spatial and associated spectra patterns. PCA has been widely used in MSI research, and its goal is to describe the dataset with a lower number of variables yet keeping as much of the original variation as possible [127]. The newly generated variables, named principal components (PCs), are the linear combinations of original variables, each of which is uncorrelated to the other ones. While this strongly depends on the dataset, examining the first 20-30 PCs usually gives good insight into the spatial and molecular trends that are in the dataset. This is extremely favorable for MSI data analysis as humans can avoid exhaustively examining thousands of or even more ion images one by one manually [265].

However, the interpretation of the loadings and scores generated by PCA is not always straightforward, and it is generally not advised to link biological interpretation directly to this pseudospectrum [273]. For example,

pseudospectra generated from PCA often contain negative values, which is difficult to interpret as original MSI data only contains positive values, namely ion counts. Nonnegative matrix factorization (NMF) methods have been adopted by researchers to avoid this issue, as it enforces the non-negativity property in the resulting components. Nonlinear dimensionality reduction methods, such as autoencoder models, have been applied in the MSI field since 2017 by Thomas et al. [248]. This strategy was extended by Abdelmoula et al. [9] in 2021 through a fully connected variational autoencoder on 3D MSI data. Their study claims that deep learning based strategy provides the opportunity to learn the underlying nonlinear spectral features for the molecular patterns identification and visualization from this high-dimensional dataset, without pre-processing steps. Nonlinear manifold learning methods, such as t-distributed stochastic neighbor embedding (t-SNE) [259, 96] and uniform manifold approximation (UMAP) [181, 229], are gaining popularity in MSI study as it enables integration of nonlinearities into the dimensionality reduction.

Clustering

Another important unsupervised learning method applied in MSI is clustering. Clustering means the grouping together of pixels based on the similarity of their associated mass spectra, therefore chemical information, enables partitioning a sample into distinct spatial regions, which is one of the most popular strategies for data representation in MSI. As each pixel is labeled with a color based on its assigned cluster, a false color image is generated to provide an overview of the high-dimensional MSI data. This process is often named *spatial segmentation*, which is particularly interesting for pathology-directed and clinical applications [265]. In a histology-independent analysis, the clustering results can be then compared with the tissue's histology, enabling to complement well-established histological methods by discovering differences in molecular level which were previously unknown [129]. Due to the high-dimensionality nature of MSI data, dimensionality reduction methods are often suggested to be applied prior to the clustering analysis. Popular clustering methods such as hierarchical clustering [7, 75, 72, 85, 177], k-means clustering [199, 130, 48], are commonly used in MSI. Additionally, several approaches have been proposed to address the inter-pixel variation issue through incorporating spatial information between pixels [112, 111, 179, 16, 18, 277].

Alternatively, rather than grouping together pixels based on the mass spectra profiles, clustering techniques can be also applied along the spatial domain - grouping together m/z bins based on the spatial distribution of their m/z images/ion images. Thus, the visualization of cluster centers results in a mean

ion image for each cluster, empowering straightforward insight into which m/z values illustrate comparable spatial expression [265]. Not only enabling a quick overview of how all m/z images look like, clustering ion images can also reveal molecular colocalizations. Previous studies applied image vector-based strategy to measure the similarity through clustering algorithms such as k-means [140], Gaussian mixture model [17], Hierarchical Density-based Spatial Clustering for Applications with Noise (HDBSCAN) [230]. More sophisticated models like deep convolutional neural networks have been utilized as an embedding function for ion images prior to the clustering methods [288, 197, 123], which reported a certain degree of improvement over the aforementioned image-vector based method after benchmarking. To address one of the biggest challenges in MSI - lack of annotations, transfer learning [288, 37] and self-supervised learning [123, 62] methods are proposed, where deep learning models can effectively learn high-level spatial features without manual labels.

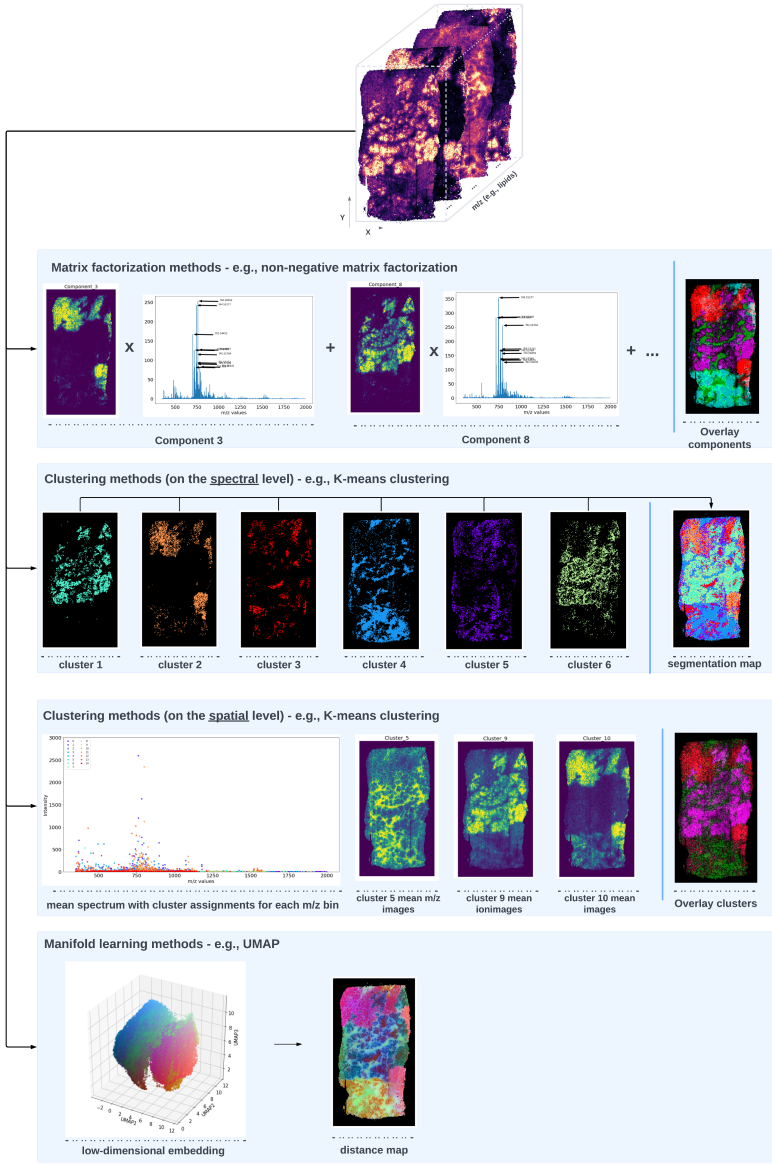


Figure 1.6: Unsupervised machine learning approaches for exploratory data analysis in MSI. (Top) Matrix factorization, where nonnegative matrix factorization as a representative example. (Middle) Clustering method on the spectral level (i.e., segmentation map) and on the spatial level (i.e., m/z images clustering), where standard K-means clustering as a representative example. (Bottom) Manifold learning, with UMAP as a representative example. This figure is inspired from Verbeeck et al. [265].

1.7 Spatial Multi-Omics

Spatial multi-omics was listed as one of *'Seven technologies to watch in 2022'* by Nature [84]. In June 2022, Ravi et al. [213] employed spatially resolved multi-omcis to characterize glioblastomas which involves integration of spatially resolved transcriptomics (10X Visium [220]), metabolomics (MALDI-FTICR-MSI), and proteomics (Imaging mass cytometry). The breakthrough of this spatial multi-omics study shows the potential of multimodality, but also stresses the importance of the development of advanced computational methods in this field.

In the last couple of years, spatially resolved techniques for profiling RNA (Spatial Transcritomics) and protein expression (Spatial Proteomics) are rapidly evolving, and the integration of those enables a more comprehensive understanding of molecular composition of cells in their full spatial context. However, spatial profiles of molecules generated by those techniques vary in resolution and number of molecular features. Different resolutions lead to different captured information, ranging from single molecules up to dozens of cells i.e., in spot-based protocols like spatial transcriptomics [235] and mass spectrometry imaging. The number of molecular features can vary from different techniques, ranging from 10 to 10,000 [151, 119, 198]. As shown in Figure 1.7, various techniques are grouped together based on their molecular entity capture capability and the spatial resolution they manage to profile. Therefore, in order to maximize the biochemical understanding from those exciting powerful technologies i.e., link MSI data with other modalities generated by other techniques, key biological questions need to be discussed and more advanced computational methods need to be developed [198].

From the perspective of machine learning, this topic falls into *multi-view representation learning*, which recently has gained attention in machine learning and data mining areas [153]. Two mainstream categories in multi-view representation learning are: (1) multi-view representation alignment, which captures relationship between different views through feature alignment, e.g., canonical correlation analysis (CCA) [121, 45] and its kernel extensions; (2) multi-view representation fusion, which aims to fuse features from different views into a single compact representation, e.g., neural network-based methods, multimodal autoencoders [191, 186]. Aforementioned strategies, illustrated in Figure 1.8, both seek to exploit the complementary knowledge from multiple views to represent the data in a more comprehensive way.

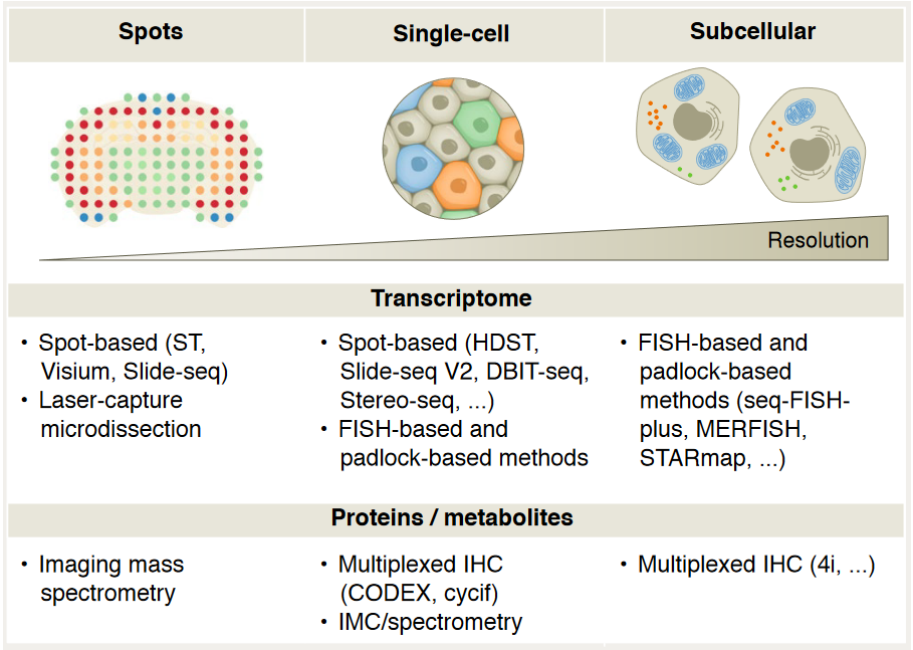


Figure 1.7: Descriptions of different spatial molecular profiling technologies. (1) Transcriptome: (a) spot-based spatially barcoded microarrays that capture transcriptome-wide gene-expression profiles: from low resolution (e.g., spatial transcriptomics, Visium) to high resolution (e.g., Slide-seqV2). (b) In situ transcriptome profiles with barcoded oligos: from low throughput (e.g., Barista-seq, osmFISH) to high-throughput (e.g., MERFISH, NanoString), sequencing by synthesis in situ (STARmap) and others. (2) Proteins: IMC (MIBI–TOF) and multiplexed IHC (e.g., CODEX). (3) Metabolites: MSI. Note: currently, only MSI can analyze metabolites (also lipids, glycans, etc.), which makes MSI is extremely complementary to other methods. Figure is reproduced from [198] with permission from Springer Nature.

1.7.1 Current applications in MALDI-MSI and challenges

Multimodal imaging is a powerful tool to integrate information from various imaging modalities, which involves areas in data acquisition, processing, integration and interpretation of images. In this introduction, the focus is on multimodal applications in MALDI-MSI, with which other imaging modalities couple with. In fact, a recent study reported by Seeley et al. [90] demonstrated 8 different imaging MS experiments from the same tissue section after careful

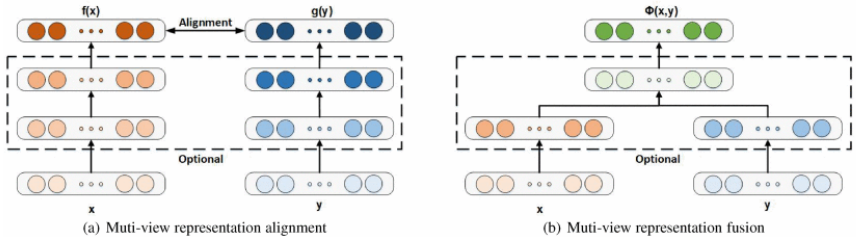


Figure 1.8: The multi-view representation alignment scheme is shown in (a) where the representations from different views are enforced the alignment through certain metrics, such as similarity and distance measurement. The multi-view representation fusion scheme (b) aims to integrate the multi-view inputs into a single and compact representation. Figure is reprinted with permission from [153] IEEE © 2018

sample preparations. This advanced method enables sequential analysis of metabolites, lipids, N-linked glycans, GlcNAc, small proteins, and tryptic peptides from a single tissue section, followed by histological staining.

Despite the fact that MALDI-MSI excels in rich chemical content without prior molecular tagging, it is not readily informative about the morphology in the tissue, as its ion images are not directly indicative of the histology. Additionally, specific sample preparation is required during the experiment and its spatial resolution is relatively lower than other imaging modalities such as microscopy, other MSI techniques (e.g., SIMS). Considering the advantages and limitations of MALDI-MSI, other imaging modalities are increasingly applied to compensate for this technique, such as with other MSI techniques, multidimensional imaging techniques (e.g., vibrational spectroscopy, immunofluorescence microscopy, Imaging mass cytometry), and even in vivo techniques like magnetic resonance imaging [252, 146, 253, 175, 176].

All of these techniques can augment the knowledge from MALDI-MSI, yet there remain many challenges and open questions in this field. Careful data acquisition, dedicated processing, registration are required prior to any downstream data analyses. As complementary modalities contain different views of the information, processing workflows are usually tailored to specific needs thus with limited generalizability. Multimodal analysis is certainly not trivial, however, once achieved, it opens new opportunities for data mining, enabling new understanding, discoveries in medicine and biochemistry [253, 35].

One important application of multimodality in MALDI-MSI other than yielding an integrated multimodal datasets is acquisition guidance. For example, MALDI-MSI data acquisition can be guided with microscopy data, enabling reduced imaged area, thus data size, and acquisition time [203, 43]. On the other hand, MALDI-MSI can also provide guidance for other modality e.g., laser-capture microdissection, for in-situ characterization of tissue subtypes [78], as described in the SpatialOMx workflow (Bruker Daltonik, Bremen, Germany).

Once the samples are prepared, integration of multimodal datasets usually involves three steps: data acquisition, (image) registration / spatial alignment, (optionally measurement matching), and data analysis. As one of the key machine learning challenges, (image) registration, which will be described in detail later, poses extra complexity throughout the analysis, especially when using serial sections. First of all, serial sections may cause complex tissue deformation, thus a more advanced registration algorithm needs to be applied, e.g., non-rigid transformation. As illustrated nicely in Figure 1.9, two serial sections from mouse kidney were used for the measurement of MSI data and histology (i.e., H&E stained microscopy data), respectively. Left figure shows 3 representative ion images of MSI data (as 3 different RGB color channels), and its transformed representation after the non-rigid registration is presented in the right figure. To show the severe level of deformation between the serial sections, the transformed ion images of MSI data is shown on the top of the microscopy image (in the right figure). The border line (in dark grey) shows a clear deformation of the serial sections and the challenge of the registration, as it was the border line in the left figure.

Secondly, data generated from various modalities usually contain different scales, often referred to as the *multi-scale data problem*. Additionally, imaging modality such as microscopy, bring significant challenges to making registration performance **at full resolution**. Finally, there is usually no immediately obvious overlap in features between MSI and other imaging modalities, e.g., microscopy images, as shown in Figure 1.10.

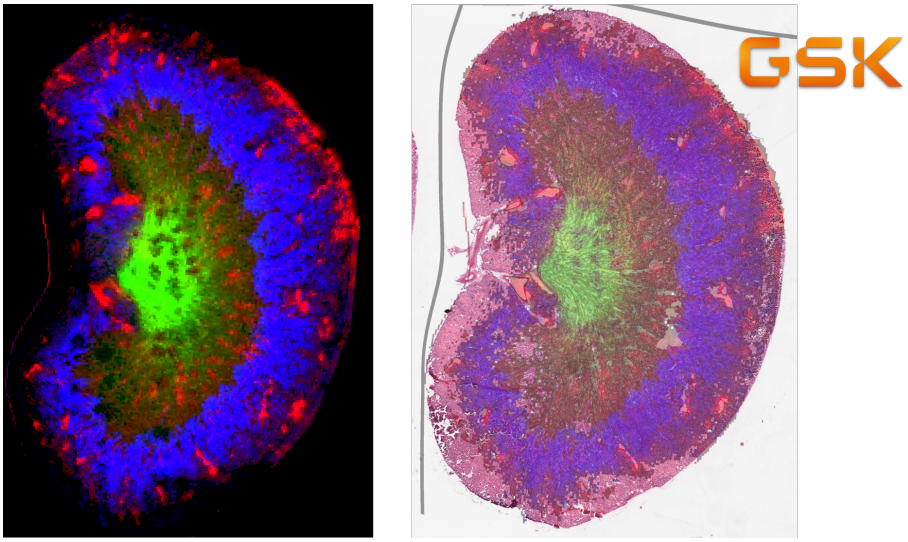


Figure 1.9: Example of the challenging image registration case: two serial sections from mouse kidney were used for the measurement of MSI data and histology, which were provided by GlaxoSmithKline (Collegeville, PA 19426, USA). Left figure displays 3 representative ion images of MSI data (as 3 different RGB color channels). Its transformation after the non-rigid registration is shown in the right figure, which lies on the top of the H&E stained microscopy image. The border line (in dark grey) in the right figure shows the clear, severe deformation between serial sections and the challenge of image registrations.

1.7.2 Data acquisition

Multimodal acquisitions involve several critical aspects to be considered [253], (1) whether generating from the same section or the neighboring section, (2) choosing which level of destructiveness of the complementary modalities, (3) in which order to operate the acquisitions. For the first aspect, using the same section is suggested so that cellular or subcellular correlations can be established, however, it is barely achievable when all modalities are destructive or when separating sample preparation processes are needed. Therefore, an adjacent section has to be used and the resulting integrated data interpretation is not reliable. A significant question *"how much can we assume these modalities are describing the same 'view'?"* has to be proposed and considered carefully. Those aforementioned aspects in multimodal acquisition are extremely crucial and can affect all downstream analyses.

1.7.3 Image registration

Another important step for multimodal analysis and recently gaining increasing attention is registration. The goal of registration is to align multimodal images to each other spatially, thus transferring them to a common coordinate system. Therefore, each pixel location or observation can describe the same 'view' and be directly linked or compared [47, 69] across multimodal datasets.

To enhance the registration, a couple of representative images from each modality are selected typically by manual selection or dimensionality reduction methods such as NMF, UMAP. To determine which representative images to be used for registration is not always trivial, especially when the material and signal are distinct. As discussed in [252] those highly spatially specific signals from one imaging modality, that may only present in a small portion of the image, yet can be useful for image registration. However, those signals may not even be present in the corresponding modality. Figure 1.10A illustrates this dilemma, where MSI data of the rat kidney displays narrow, localized images in the kidney sub-regions but H&E microscopy shows the broad staining across the kidney [252]. Dimensionality reduction methods e.g., UMAP can be very helpful for this, as it can be used to summarize high-dimensional MSI data in a fast and scalable manner by searching for a low dimensional projection that has the closest possible equivalent fuzzy topological structure. Details of UMAP are discussed in Chapter 5 section 5.3.

Then a certain registration method is applied to estimate the transformation, mapping a moving image to a fixed reference image. The transformation can be linear models [160, 44] such as rigid, affine, and projective. These methods

are suggested when modalities are acquired from the same section. Nonlinear methods are preferable for more advanced registration, when modalities are captured from a serial section and the morphology changes are not uniform across the tissue [82, 218, 267].

However, registration is not always trivial as (1) different modalities can lead to the multi-scale data problem, and (2) quantitative evaluation of registration performance is not always straightforward [252].

1.7.4 Measurements matching

Multi-scale data problem occurs when modalities share very distinct characteristics i.e., one modality sampled at 1 um/pixel whereas the other sampled at 100 im/pixel. Thus each 'pixel' from one modality can describe a 'footprint'/'view', which is represented by many hundreds of pixels or observations in another [252]. For example, in Figure 1.10B illustrates an H&E microscopy is sampled at 1 um/pixel and how a single pixel varies from 10 to 200 um/pixel. When it is at 10 um/pixel, 1-2 cells are visible whereas at 200 um/pixel, hundreds of cells are covered. Basically, one observation in one modality can represent multiple observations in another, and many studies try to find a proper aggregation strategy. Many studies used strategies such as interpolation and averaging, whereas more advanced methods are proposed e.g., Gaussian smoothing [257]. For example, in order to match the measurements acquired from MSI and microscopy data, Van de Plas et al. applied a 2-D Gaussian bell curve centered on a pixel/measurement of MSI data to approximate MSI laser sampling experiment, where the closer the pixels from microscopy (modality with higher spatial resolution) to the center of a MSI pixel, the more importance will be given.

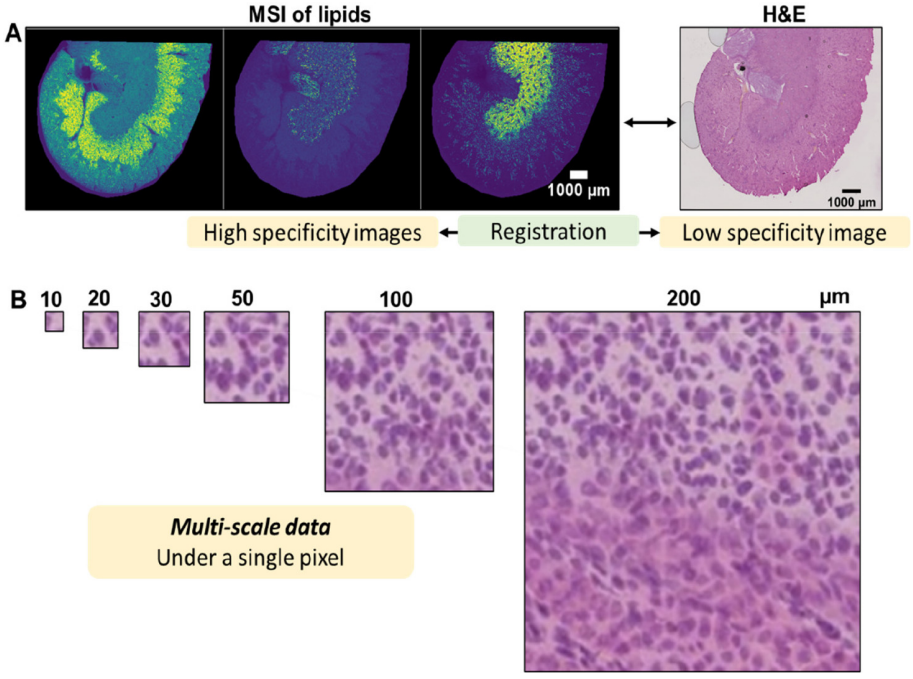


Figure 1.10: Illustration of some challenges in multimodal image registration by using a lipid MALDI MSI and H&E microscopy example: (A) A representative image must be extracted from both modalities and it is especially challenging when they have very distinct spatial specificity. Thus, deriving multiple representative images from each modality is suggested. (B) Multiscale/multi-resolution data is very common in chemical imaging, and the context within a single pixel can be very different. Figure is reprinted with permission from [252]. Copyright 2021 American Chemical Society.

1.7.5 Downstream multimodal data analysis

The objective of data analysis in multimodal MALDI-MSI [253] is to highlight shared and complementary information between imaging modalities, by adopting strategies characterized by Bemis et al. [35], see Figure 1.11:

Class Discovery: reveal shared features from multimodal datasets through unsupervised learning methods such as dimensionality reduction and clustering [10, 219]. This strategy falls into **multi-view representation alignment** in the context of multiview representation learning [153].

Class Prediction: generate an enriched multimodal dataset, therefore a single representation, for class prediction via supervised learning algorithms like classification and regression. The integrated representation can be used for predicting experts' annotations (e.g., pathological outcomes) or information from complementary modality. This strategy falls into **multi-view representation fusion** method [153]. For example, one of the most iconic studies in MALDI-MSI cross-modality application, proposed by Van de Plas et al. [257], trained a partial least-squares (PLS) regression model to predict ion distributions from microscopy at higher spatial resolution, as well as predict distributions out of the sample. Additionally, several studies reported the fused data integrated from MALDI-MSI and microscopy enables improved classification performance in cancer study [289, 207].

Class Comparison: find differentially-abundant ions in different regions of interest via various statistical methods. Regions of interest can be acquired from another modality such as microscopy images (H&E and fluorescence) and transferred to MALDI-MSI data for ion differential expression analysis [266, 203].

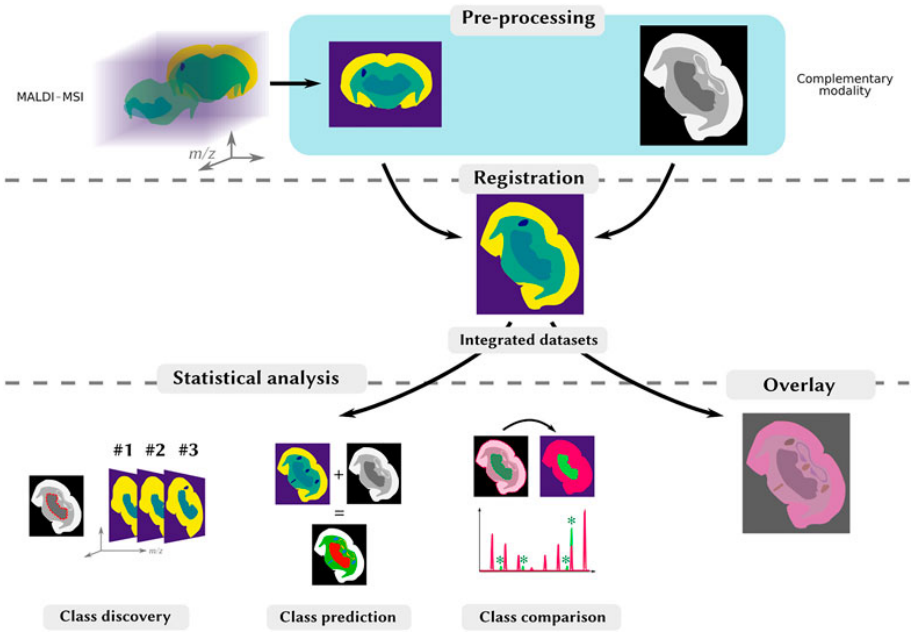


Figure 1.11: Typical processing and analysis workflow for multimodal MALDI-MSI. **Top-row**: MALDI-MSI data on the left and the complementary dataset on the right are processed, in order to be able to extract representative images for later registration step. **Middle-row**: representative images from each modality are spatial aligned and registered. **Bottom-row**: after the modalities are registered, their overlays can be visualized so that similar or distinct spatial distributions can be revealed (on the right). Patterns can be explored by a joint statistical analysis to, e.g., find spatial correlations, by investigating spatial clusters (class discovery), generate an enriched dataset from both modalities (class prediction) or seek region-specific ions in two different regions of interest (class comparison). Figure is reprinted from [253] with permission. Copyright Creative Commons CC-BY license.

1.8 Structure of the thesis

In this thesis, we will focus on three elements in particular, with each having its own chapter dedicated to it:

- **In Chapter 2** we focus on unsupervised learning for exploratory study on MSI data. Particularly we focus on clustering methods on the spatial level of MSI data, namely clustering m/z or ion images. More specifically, we applied a pre-trained neural network to extract morphological embeddings from ion images, and the resulting embeddings are used for downstream clustering instead of the original ion images. The clustering results show more fine-grained clusters and greater consistency in the spatial structures assigned to individual clusters.
- **In Chapter 3** the focus is on supervised learning approaches, namely classification tasks, in melanoma diagnostic application. We developed a multimodal pipeline to incorporate the morphological features from H&E microscopy images with MSI data for improved classification results. Two unimodal classification pipelines were built and compared with the proposed multimodal pipeline, revealing the strength of multimodality in melanoma diagnosis.
- **In Chapter 4** we explore the spatial multi-omics approach in prostate cancer study, where we integrated in situ lipidomics via MSI with spatial transcriptomics data, enabling the direct combination and correlation of spatial lipidomics and transcriptomics, and therefore, gaining new insights into the cellular and molecular heterogeneity of high-risk prostate cancer.
- **In Chapter 5**, we will discuss specific challenges that were tackled throughout my work from a machine learning perspective. We refer back to the earlier chapters that are effectively based on publications, to dive deeper into their associated machine learning challenges and how these were addressed in my work.
- **In Chapter 6** we conclude the work of this thesis with future perspectives.

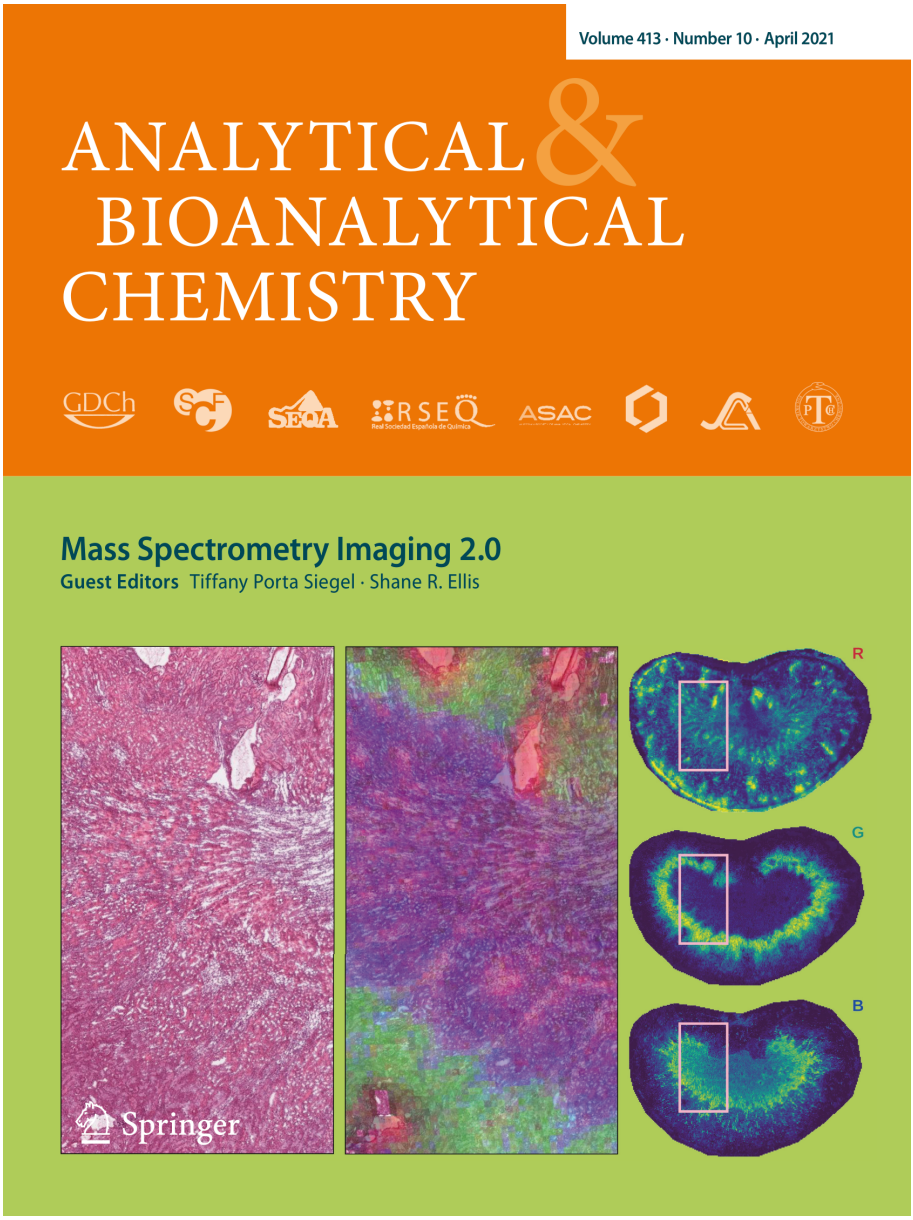
Chapter 2

Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning

This chapter was previously published as:

Zhang, W., Claesen, M., Moerman, T., Groseclose, M. R., Waelkens, E., De Moor, B., Verbeeck, N. (2021). **Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning**. *Analytical and bioanalytical chemistry*, 413, 2803-2819. Copyright Creative Commons CC-BY license.

This work was featured as **front cover** for issue *Mass spectrometry imaging 2.0*, edited by Tiffany Porta Siegel and Shane R. Ellis. https://link.springer.com/journal/216/volumes-and-issues/413-10


Volume 413 · Number 10 · April 2021
ANALYTICAL & BIOANALYTICAL CHEMISTRY
GDCh
SCF
SEQA
RSEQ
ASAC
Mass Spectrometry Imaging 2.0
Guest Editors Tiffany Porta Siegel · Shane R. Ellis
R
G
B
Springer

2.1 Abstract

Computational analysis is crucial to capitalize on the wealth of spatio-molecular information generated by mass spectrometry imaging (MSI) experiments. Currently, the spatial information available in MSI data is often under-utilized, due to the challenges of in-depth spatial pattern extraction.

The advent of deep learning has greatly facilitated such complex spatial analysis. In this work, we use a pre-trained neural network to extract high-level features from ion images in MSI data, and test whether this improves downstream data analysis. The resulting neural network interpretation of ion images, coined *neural ion images*, are used to cluster ion images based on spatial expressions.

We evaluate the impact of neural ion images on two ion image clustering pipelines, namely DBSCAN clustering, combined with UMAP-based dimensionality reduction, and k-means clustering. In both pipelines, we compare regular and neural ion images from two different MSI datasets. All tested pipelines could extract underlying spatial patterns, but the neural network-based pipelines provided better assignment of ion images, with more fine-grained clusters, and greater consistency in the spatial structures assigned to individual clusters.

Additionally, we introduce the Relative Isotope Ratio metric to quantitatively evaluate clustering quality. The resulting scores show that isotopical m/z values are more often clustered together in the neural network-based pipeline, indicating improved clustering outcomes.

The usefulness of neural ion images extends beyond clustering towards a generic framework to incorporate spatial information into any MSI-focused machine learning pipeline, both supervised and unsupervised.

2.2 Introduction

Mass Spectrometry Imaging (MSI) is a powerful, label-free molecular imaging technology that enables mapping the spatial distribution of thousands of biomolecules in a tissue section, based on a single experiment [54]. Due to its ability to combine rich biochemical characterisation with spatial information, MSI technology is rapidly being adopted for a panoply of applications, including biomarker discovery, clinical diagnostics and drug delivery studies [227, 224, 110]. While a number of different MSI variants exist [50], in general MSI operates by first overlaying the tissue with a virtual rectangular grid,

and then collecting a mass spectrum at each grid location. Each of these collected mass spectra is a histogram of biomolecular ions counts, partitioned by their mass-to-charge values (m/z), within a target mass-to-charge range. MSI experiments result in a three dimensional data cube, with spatial coordinates (x and y) and a m/z axis containing the spectral information. Ion images are constructed by plotting the intensities for a single mass bin (m/z value) for each acquired pixel, i.e., (x, y) grid location in the tissue.

Technological advancements continue to push the capabilities of MSI, leading to improvements in specificity, sensitivity and speed, which have translated into the detection of larger numbers of biomolecular species at ever-increasing spatial resolution. A corollary of these technological improvements is that MSI datasets have grown substantially over the years. At the time of writing, a single experiment commonly generates tens of gigabytes of raw data, but this can even range into terabytes. Moreover, MSI data analysis is encumbered by the large number of variables measured, in terms of pixels, m/z bins or both.

Unless prior targets of interest are known, it is infeasible for a researcher to manually investigate and find molecular differences in these large datasets, and, as a result, computational approaches have become indispensable to support data analysis. A wide variety of unsupervised and supervised machine learning methods have been used in a broad range of applications. In this paper, we focus on unsupervised learning methods, which support exploring the underlying patterns within the MSI data, providing open-ended exploratory insights in the data. These methods include factorization methods (i.e. principal component analysis [256], nonnegative matrix factorization [244]) and clustering methods [178, 200], e.g. to group spectra or pixels with similar chemical expression into the same cluster [265]. Supervised learning methods, on the other hand, are used when prior knowledge or annotations are available, such as target m/z's, or regions of interest, which can be used to guide the analysis, and, for example, extract differentially expressed molecular ions between regions in the tissue.

Using appropriate *distance measures* can substantially improve the outcomes of both supervised and unsupervised data analysis. A distance measure formalizes how to quantify similarity between data instances, e.g., between spectra originating from different pixels or between ion images associated to different m/z values. A recurring task in MSI data analysis is identifying which ions are co-localized, which translates into the problem of clustering *similar* ion images. As mentioned previously, such tasks are ideally tackled in a largely automated fashion to deal with the large amount of information in each MSI experiment. In order to answer such questions, a human expert

would implicitly rely on his or her excellent visual pattern recognition abilities, and will use morphology, anatomical structure and saliency to assess the ions' expression.

Unfortunately, it is difficult to translate the extensive spatial pattern recognition that is so easily performed by humans into robust, automated data analysis pipelines. To this day, most data analysis pipelines in MSI forgo the inclusion of available spatial information altogether [265]. Pipelines that do account for spatial information typically use measures such as spatial correlation or cosine distance to quantify similarity between ion images [237, 180]. As these are global measures, they are prone to miss relevant, localized differences between ion images. There have been a number of MSI machine learning methods that have explicitly focused on including spatial information into the analysis. Most of these are found in the task of performing spatial segmentation of a MSI dataset, i.e. grouping together pixels by their chemical content, where both Alexandrov et al. [19] and Bemis et al. [36] have developed methods to integrate local pixel neighborhood information directly into the clustering process, in order to improve clustering results. Others have incorporated spatial information into the analysis by combining MSI data with anatomical atlases [268, 8].

In this work, we propose a simple yet powerful approach to replicating human visual pattern matching. Owing to its simplicity, our approach can be incorporated in a wide range of data analysis pipelines, enabling them to exploit spatial information with few to no modifications. Our approach is rooted in deep learning, which is a class of models and associated learning methods based on artificial neural networks (ANNs). ANNs consist of layers of neurons, which are simple mathematical functions capable of doing basic, non-linear transformations of their inputs. By composing and stacking many layers, a *deep* network becomes able to recognize complex patterns in data, despite consisting of simple building blocks. This layered information processing resembles the workings of the brain, starting from low-level signals in initial layers up to high-level, domain-specific feature recognition in deeper layers.

In the case of computer vision tasks, the initial layers of the network act as image feature extractors, whereas deeper layers make higher level abstractions over those detected features [242]. In recent years, deep learning has shown tremendous potential in visual recognition tasks, having outperformed the state of the art machine learning algorithms in many computer vision tasks, and even human experts in certain biomedical tasks [116]. The strength of deep learning methods is their ability to learn and combine both low-level and

high-level features and abstractions from image data. Deep learning has also previously been used in the context of MSI without the inclusion of spatial information. Thomas et al. [249] and Inglese et al. [125] have used a neural network autoencoder approach for the dimensionality reduction of MSI data, whereas Behrmann et al. [31] have used a supervised deep learning pipeline for the classification of tumor cores in tissue microarrays.

While deep neural networks are very powerful, the disadvantage is that they are complex, and have a high number of parameters to be learned, and as such require a huge amount of - ideally labeled - data to learn powerful, high level abstractions. Directly learning those parameters from scratch every time for new datasets poses a big computational challenge, and is prone to overfitting in a setting such as MSI, where a lot of data is available on individual tissues, but the number of different tissues measured in a study is often limited, which is required to achieve good generalization. While it is possible to mitigate the issue of overfitting to a certain extent using data augmentation, for example by applying transformations to the original data [142] or by using generative adversarial networks [105, 210] to automatically generate new representative data.

Moreover, deep learning is highly conducive to transfer learning, in which models trained for a certain task are re-purposed for other, related tasks [37]. In transfer learning, the requirement that the training data must have the same distribution as the test data is relaxed, i.e., a model is trained on one task, and the learned model is (fully or partially) transferred to another task. The more similar the task is, the better this approach generally works. Model-based transfer learning is particularly powerful for computer vision applications [141], because a lot of implicit knowledge, i.e., general image feature recognition, is shared amongst most practical tasks. Oquab et al. [196] have shown that the front layers of a convolutional neural network, that has been pre-trained on a large-scale annotated dataset, can be efficiently transferred to another computer vision task where limited data is available. Furthermore, the researchers noted that the way the neural networks interpreted the images was reminiscent of a human operator. A number of other studies have shown similar potential for re-purposing of neural networks for various downstream tasks [158, 132].

In this work we will use the well-known Xception network [64], which has been trained on the ImageNet database [142] consisting of over a million images. Due to the large variety of the images included in this dataset, networks that perform well on this dataset, perform well on many other computer vision tasks [141],

and as such can be used as general purpose, image feature extractors. Xception, for example, has previously been successfully used for feature extraction in various life science applications such as microscopy and electron microscopy [211, 271]. In the context of MSI, we will pass each ion image in the dataset through the network Xception, which generates a vector representation of each image, that encodes the high-level features within the image.

Mathematically, we use the neural network as an embedding function for ion images. The aim of this strategy is to map images with a high visual similarity close together, largely in line with how a human observer would interpret spatial resemblance between the original images. We will call these vector representations *neural ion images*, to highlight their origin and clearly contrast them with regular ion images. Note that these neural ion images are abstract, high-dimensional representations without any direct visual interpretation.

In their recent work, Ovchinnikova et al. [197] have also used Xception for evaluation of MSI ion image co-localization (abbreviated as ColocML). Here, the Xception model was used to compare the visual ranking of ion image similarity by human experts to those produced by various distance measures and algorithms, such as the Xception network. Here, however, the Xception network was used in a supervised way, i.e it was further trained using labels provided by human experts. The algorithm showed good performance in this task, albeit similar in performance to standard cosine distance for their use case and tests (assessing the algorithm over different datasets). In contrast, our approach uses the pre-trained Xception model in a completely unsupervised way, without any domain transfer or fine tuning, and as such we show that it is directly usable for a wide breadth of applications in MSI, regardless of how much domain-specific data or computational resources are available. The authors of ColocML have made their data and rankings available to the MSI commnunity, and as such are a valuable resource that will be used to benchmark our method.

It is important to stress that the resulting neural ion images can conceptually replace regular ion images as inputs for most downstream machine learning and statistical analyses, both supervised and unsupervised. Using neural ion images enables improving existing data analysis pipelines to better incorporate spatial expression patterns with minimal change to the full pipeline.

In this work, we focus on assessing the merits of neural ion images for the unsupervised task of clustering, i.e. grouping together, ion images based on

their spatial expression. Similar to how factorization is often used in the context of MSI, the clustering of ion images reveals the different spatial patterns that are present in a MSI dataset. This is done by visualizing the mean ion image for each of the different clusters (or groups) of ion images that are found in the dataset. Furthermore, this method provides direct insight into which m/z values exhibit comparable spatial expressions, compared to factorization where this relationship is not always straightforward [265]. Alexandrov et al. [17] previously applied a probabilistic clustering algorithm (a Gaussian mixture model) to cluster ion images in MALDI MSI datasets based on their spatial similarity. Similarly, by Konicek et al. [140] used k-means to cluster together ion images in a TOF-SIMS dataset.
Here, our goal is to improve on this work by integrating the pre-trained Xception model into the clustering pipeline, and as such find clusters with a higher specificity, based on localized spatial features, and with a greater consistency in the images assigned to the same cluster.

2.3 Methods

2.3.1 Materials

We demonstrate our proposed method on two MSI datasets, namely one from human lymph node and one from a mouse kidney tissue.

The first dataset is collected from a human lymph node tissue sample using the Bruker rapifleX MALDI Tissuetyper. The tissue sample was removed and snap frozen in liquid nitrogen (-80 $°C$). Tissue sections of 7 μm thickness were acquired using a cryostat and thaw-mounted onto ITO-coated glass slides, after which 2,5-DHB matrix was deposited by sublimation. Experiments focused on the 620 to 1200 m/z range, using a sampling resolution of 10 μm, collecting roughly 500.000 pixels and 8.000 ion images for this tissue. The lymphoma sample was acquired from the Leuven University Hospital Biobank and usage was approved by the Ethics Committee Research UZ / KU Leuven (S-55498).

The second dataset is collected from a mouse kidney tissue section. The tissue sample was snap frozen in liquid nitrogen at necropsy. Tissue sections were collected in a cryostat (CM3050S, Leica, Buffalo Grove, IL; chamber temp. $-20°C$ and object temp. $-18°C$) at 10 μm thickness and thaw-mounted onto ITO-coated glass slides. 2,5-DHB matrix prepared at 25 mg/mL in methanol:water (1:1, v:v)(0.5% TFA) was applied to the tissue sections using a TM Sprayer (HTX Technologies, Chapel Hill, NC) automated spray device.

The following parameters were used to achieve a target matrix density of 0.4938 mg/cm^2: flow rate 0.05 ml/min; N_2 Pressure 10 psi; Spray temperature 70 $°C$; velocity 1350 mm/min; track spacing 3 mm; track offset 1.5 mm; 16 passes. MALDI MSI was conducted at a sampling resolution of 50 μm on a Bruker scimaX MRMS. Fullscan data was acquired from m/z 200-1200 using 250 laser shots per pixel at a frequency of 2 kHz and an estimated resolving power of 66K at m/z 400. The mouse kidney tissue was collected internally at GSK. This study was conducted in accordance with the GSK Policy on the Care, Welfare and Treatment of Laboratory Animals and was reviewed by the Institutional Animal Care and Use Committee either at GSK or by the ethical review process at the institution where the work was performed.

2.3.2 Ion Image Clustering

In order to assess the merits of our deep learning approach, and to demonstrate how it can be easily integrated into any data analysis pipeline, we will investigate two different, well-known clustering methods, both with and without the use of neural ion images. Figure 2.1 gives an overview of the ion image clustering workflow, and shows three separate processing blocks, which will be discussed in greater detail below. In the first block, "Shared preprocessing", all images of the MSI dataset receive a common preprocessing treatment. The second block, "Generation of neural ion images", contains the neural network steps, and either passes the ion images through the pre-trained neural network to generate "neural ion images", or leaves the images unaltered ("regular ion images"). In the third and final block, "Clustering steps", the clustering of the images is done using k-means or a combination of UMAP and DBSCAN (U-D for brevity), and takes as an input either neural or regular ion images. This results in a total of four different ion image clustering pipelines, which we will refer to as regular U-D, neural U-D, regular k-means, and neural k-means.

Shared preprocessing

Before applying any clustering pipeline, we perform basic preprocessing on the raw spectra and, afterwards, outlier removal and standardization on the resulting ion images. These procedures are done separately for each dataset.

1. *Spectral preprocessing* The lymph node data is normalized based on Total Ion Current (TIC) and baseline corrected using median filtering. The mouse kidney data is normalized based on TIC and peak-picked.

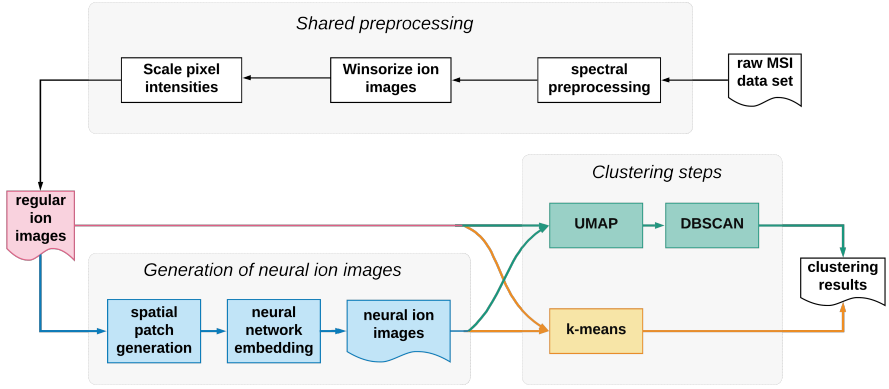


Figure 2.1: Pipelines: the top panel contains a series shared preprocessing steps. The lower left panel shows the additional steps (in blue) from our proposed method compared with regular clustering pipelines (in red). The additional steps from our proposed model include spatial patch generation, neural network embedding and neural ion images access. Finally two different clustering methods are conducted, which are DBSCAN (after UMAP dimentionality reduction) and k-means.

Peak picking was conducted using a S/N threshold of 5 and the Gibbs, harmonic, and magnetron peaks were filtered during acquisition. The SQLite peak list files were processed in MATLAB using custom software to recalibrate each spectrum using a linear best fit model and then bin corresponding peaks across all spectra in the dataset to a common m/z. The binned data was then converted to imzML for further processing.

2. *Ion image standardization* The ion images for both datasets are individually winsorized and then standardized to a common range prior to clustering.

Winsorizing is a well known robust statistical estimation technique which limits extreme values to reduce the effect of potential outliers [114, 275]. Winsorizing involves clipping extreme values, usually symmetrically on both extremities of a distribution, before computing location statistic(s) of interest such as the mean. In our case, we only winsorized above the 95th percentile to reduce the impact of high intensity outliers commonly found in MSI data. Winsorizing the lower extremity is not necessary for spectral data since there is an explicit limit at 0 intensity. After winsorizing, the resulting pixel intensities for each image are scaled to the range $[0, 1]$. The resulting preprocessed images will

henceforth be referred to as *regular ion images*, and serve as input for the next stage of the pipeline.

Generation of neural ion images

Our proposed approach involves a couple of extra steps to convert regular ion images into neural ion images. Generating neural images from individual regular ion images is done as follows:

1. *Patch generation* A neural network typically requires images of certain shape as inputs. Since ion images can be of any shape, we first define a set of overlapping patches across the ion images, which decouples our pipeline from any specific shape of ion images.
2. *Generating patch embeddings* The pixel data for each patch is fed into the neural network as a single data instance, which results into an associated neural embedding per patch.
3. *Aggregating patch-level embeddings* The neural ion image is constructed by aggregating the embeddings of individual patches into a single vector representation for each regular ion image.

We discuss each of the aforementioned steps in more detail below.

1. Patch Generation Ion images tend to be larger than the default input size expected by neural networks. For example, the Xception network we used [64], accepts square patches of 299×299 pixels by default, with a minimum of 71×71 pixels. These shape mismatches can be fixed via patch generation, which is especially common when using convolutional neural networks to analyze high-resolution images. Patch generation means that, rather than providing the full image to the network, the image is instead split up into smaller "patches" that are then given as an input to the network, thus preserving all the details in the original image in one or several patches. An example of patch generation is shown in Figure 2.2. The size of the patches is important as it determines the context that the network "sees" in one go. Therefore, we tested different patch sizes, as described in the Results section.

2. Generating patch embeddings A neural network is used to generate a vector representation (embedding) of each patch, which ideally captures all relevant morphological information present in the patch. The network itself is pluggable in our pipeline, so other networks could also be used. We leveraged the widely used Xception model because it has shown strong performance on

recent benchmarks [64], despite being relatively small compared to other models with only 36 convolutional layers. Owing to its small size, its inference speed is higher than that of many other models with comparable accuracy [42]. Finally, by using a pre-trained network, we avoid the demanding task of designing and training the network, which would require a lot of data as well as computational resources. We use the Xception network as is available in the Keras Python package `https://keras.io/`.

3. Aggregating patch-level embeddings Finally, after obtaining an embedding associated to each patch, these must be integrated into a final vector representation that constitutes the neural ion image. We opted to use max-pooling because the neural ion image then essentially captures whether a spatial pattern (single output feature in all patch embeddings) was detected in *any* of the patches. This optimizes the neural ion images' sensitivity to detect localized features.

Clustering steps

We use two different clustering approaches in this work, namely DBSCAN, combined with UMAP, and k-means clustering. As both of these clustering pipelines expect vector inputs, regular ion images were reshaped to vectors (neural ion images are already vectors). In this work, we purposefully selected two fundamentally different clustering methods to properly assess the merits of using neural ion images instead of regular ion images.

UMAP-DBSCAN The first approach, DBSCAN, which stands for Density-Based Spatial Clustering of Applications with Noise [92], is a popular density-based clustering algorithm that inspired several extensions (e.g., HDBSCAN [53, 231]). DBSCAN aims to find high density regions in feature space where samples are closely packed, i.e. samples that have a lot of other closely resembling samples (neighbors). Based on a user-defined radius, the algorithm estimates a minimum density level (i.e. minimum number of neighbors within a radius). Clusters are then regions where groups of samples exceed this density threshold, whereas points that are in low density regions are put in a common *noise cluster*. DBSCAN is widely available in different implementations, here we use the scikit-learn implementation [204].

Given that we are clustering ion images, each pixel can be seen as a separate feature, or measured variable, and, as such, the number of features per image is very high (e.g. $\sim$ 500.000 pixels per regular ion image for the lymph node dataset, thousands of variables per neural ion image). High-dimensional feature spaces are inherently sparse due to the curse of dimensionality [33], which encumbers the process of identifying clusters based on density [265].

It is therefore advisable to first do a dimensionality reduction step prior to performing DBSCAN, which we did using Uniform Manifold Approximation and Projection (UMAP) [181]. Briefly, UMAP aims to form a topological representation of the original high-dimensional data by finding local manifold approximates using their local fuzzy simplicial set representation. Prior work has shown promising results of applying UMAP on spectra in MSI data [229] (as opposed to ion images as we are doing here).

In this paper we used UMAP to reduce the dimensionality of regular and neural ion images to three prior to applying DBSCAN. Mapping to three dimensions enables visualizing the resulting embedding, which helps in interpreting the results, and is sufficiently low-dimensional to effectively support DBSCAN. Different parameters were manually optimized for UMAP and DBSCAN for the different models because the inputs for each pipeline are very different. To facilitate a fair comparison, we did a range of experiments and selected the most promising model with selected parameters for each approach. For the regular ion images, each ion image was reshaped to a vector of size one-by-'number of pixels', and cosine was used as the distance measure between the resulting image vectors in the UMAP algorithm. For the neural ion images, similarly the cosine distance was used as the distance measure between the neural ion image vectors. The Python code by UMAP's creator was used [182]. For brevity, we will refer to combination of UMAP-DBSCAN as U-D.

k-means The second clustering algorithm used in this paper is the well-known k-means clustering [41], which is available through scikit-learn [204]. k-means is a highly popular clustering algorithm used in a wide variety of applications owing to its simplicity and speed [26]. Conceptually, it tries to separate objects in a number of groups with low intra-group variance by minimizing the squared distance between samples and the prototype vector (mean) of the cluster they are associated to. The Euclidean distance (standard in k-means) was used for all experiments.

Practical differences between U-D and k-means In contrast to DBSCAN, k-means requires the number of clusters to be decided in advance, as we will discuss further below. The ability to explicitly reject certain samples from clustering grants DBSCAN increased robustness to noise and outliers compared to k-means, which always forces each sample into a cluster. Conceptually, owing to its nonlinear nature, U-D is expected to identify more subtle trends within data, but requires more expertise to optimize its various tuning parameters. From a computational perspective, k-means is significantly faster.

2.3.3 Evaluation of Clustering Results

We proposed different methods to evaluate the quality of the resulting clusters from the different pipelines.

Number of Clusters

After clustering, we first check the number of meaningful clusters. Meaningful clusters should not include the noise cluster as defined by DBSCAN (as previously discussed) or clusters that contain a single image when using k-means. Furthermore, we also consider clusters that only contain noisy images to be non-meaningful clusters. This is especially relevant in the lymph node data, where no peak-picking was done, and as a result a lot of noisy images are expected in the dataset. We assume that, the more meaningful clusters are extracted, with distinct mean images, the more likely it is that more underlying structures are detected from the dataset.

Relative isotope ratio (RIR)

In order to objectively compare the performance of different clustering pipelines, we propose a metric that quantifies how well a clustering groups ion images stemming from isotopic m/z values. Two m/z values are considered isotopes if and only if the following criteria are met:

1. The spectral distance equals 1.003m/z$\pm\delta$, where $\delta \neq 0$ allows for small deviations in mass-to-charge values. The value of δ depends on the mass accuracy of the MSI data. We set δ equal to $0.1Da$ and $0.01Da$ for the lymphoma and kidney data respectively. This spectral distance is primarily intended to identify carbon isotopes, however if δ is set large enough this will capture other isotopes as well.
2. The corresponding ion images should exhibit a clearly similar spatial expression, which we assess using Pearson correlation. Specifically, the correlation between both ion images should be at least 0.85.

To assess the quality of isotope grouping for a given clustering, we count the number of isotopes that are correctly grouped in the same clusters and divide this count by the total count of isotopes in the dataset. This fraction $F \in [0, 1]$ captures how many of the isotopes in the dataset are clustered together (higher is better). However, this fraction favors large clusters, because then the probability of clustering isotopes together increases. This can intuitively

be seen by the trivial edge case of having a single cluster, in which all isotopes are obviously clustered together, thus yielding a seemingly perfect score.

As such, because the *expected* value of F depends on the size of the clusters in a clustering, we benchmark the isotope fraction of a given clustering (F_{clust}) to that of a simulated random clustering with the same cluster sizes (F_{random}). This random clustering is simulated in a bootstrap fashion to get a consistent estimate of the expected isotope fraction. Our *relative isotope ratio* metric $\mathcal{R}$ is the relative increase in isotope grouping for a given clustering compared to its random baseline, i.e., $\mathcal{R} = F_{clust}/F_{random} \in [0, \infty)$, where higher is better. Any clustering that is better than random should yield $\mathcal{R} \gg 1$.

2.3.4 Visualization of Clustering Results

In order to visualize the spatial results, we compute the mean ion image for each cluster by first winsorizing each ion image from all clusters, followed by stratified averaging over all m/z images per cluster. Those mean ion images then show the various spatial patterns detected by different clustering pipelines. We also generate a mean spectrum for each dataset with clusters assignment for each m/z value.

Furthermore, for U-D clustering, we visualize the 3D embeddings after UMAP dimensionality reduction, which results in a 3D scatter plot where each m/z image is a point with clusters assignment. This scatter plot illustrates how all m/z images are split and clustered in different groups, the distance between different clusters, and the tightness of each cluster.

Finally, we utilize histological information from concomitant microscopy images to check whether the clusters are biologically relevant. These microscopy images were obtained on serial sections to the MSI data. To assess biological relevance, we construct a false color image from the mean ion images of 3 different clusters, where each image uses a separate color channel (RGB), which is then overlaid onto the microscopic image. If the overlay shows a clear co-localization of cluster images and morphological patterns in the associated microscopic images, it can be assumed that the clusters are biologically relevant.

2.4 Results and Discussion

We start by preprocessing the data, as per the "shared preprocessing" block described in Figure 2.1. Spectral preprocessing differs for both datasets: the lymph node dataset is TIC normalized and baseline corrected, whereas the

mouse kidney dataset is TIC normalized and peak picked. Next, each individual ion image is winsorized, substituting pixels with outlying intensity values, after which it is scaled so that pixel intensities lie between 0 to 1. After the shared preprocessing, we start one of four different clustering pipelines, namely UMAP-DBSCAN on regular ion images (regular U-D), UMAP-DBSCAN on neural ion images (neural U-D), k-means on regular ion images (regular k-means), and k-means on neural ion images (neural k-means). Each of the four clustering pipelines is used to process the human lymph node and mouse kidney datasets separately.

2.4.1 Generation of neural ion images

Once the MSI data is preprocessed, the regular ion images can be used to create neural ion images using the pre-trained neural network. First, we need to generate image patches with a size appropriate for the pre-trained Xception network. As explained previously, the size of the patches is important, as this determines the context that the network receives as an input. We therefore aimed to select a patch size with a physical size that is large enough to capture morphological patterns and anatomical structures, testing various patch sizes.

We found that patches with sides of a physical size of around 1-2 mm work well in practice. For the lymph node dataset, at a 10 μm sampling resolution, this results in a patch size of 150×150 pixels, in a total image size of $1,183 \times 696$ pixels. For the mouse kidney dataset, we tried to get patches with a similar size, however, given the 50 μm resolution, a patch with the minimum pixel size of 71×71 for the Xception network would still be too large (~ 3.5 mm). We therefore upsample the ion images for mouse kidney data with a factor of two, so that we obtain more, smaller patches with a more appropriate physical size. Figure 2.2 shows a patch generation example from mouse kidney dataset, with the original patch sizes highlighted in blue and the patches after upsampling indicated in pink. Furthermore, it is important to note that we use overlapping patches to avoid edge artifacts, e.g., missing anatomical structure because it was located at the edge of a patch. In our patch generation procedure, we ensured a 40% to 50% spatial overlap between patches.

Using this procedure, we generate image patches for each ion image in the MSI dataset. The image patches are batched per ion image and are sequentially fed to the pre-trained network, which analyzes each patch and extracts low- and high-level image features from them. This translates to a vector per patch, that scores the various features detected by the network. The results for the individual patches are then combined in a max-pooling layer at the end of the

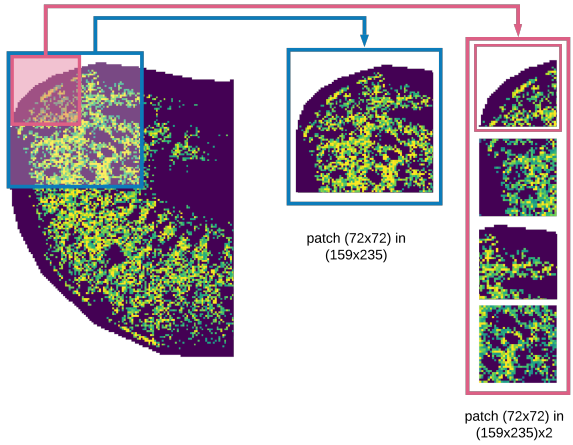


Figure 2.2: From left to right: a regular ion image from the mouse kidney dataset, generated patches with different sizes. In order to get the ideal size of patches, we up-sample the full ion image before splitting into multiple patches. All results were obtained using up-sampled patches, as indicated in pink, which have a physical size of roughly 1.85×1.85 mm.

network, which registers whether a spatial pattern is detected in any of the patches, to increase sensitivity for localized features. The result is a single vector (of size 2048) per ion image, which captures the "interpretation" of each ion image by the neural network, i.e., the neural ion image.

With both the regular and neural ion images available, we can use these as input for the clustering algorithms, in order retrieve the different spatial ion expressions in the dataset.

2.4.2 Human lymph node dataset

UMAP-DBSCAN clustering results

We apply each of the four clustering pipelines to the human lymph node dataset, starting with the U-D pipelines. First we perform UMAP dimensionality reduction on both the regular and neural images, which maps the images from the original 500.000 (regular) and 2048 dimensional (neural) space respectively to a three dimensional embedding space. The goal here is to map similar

images close together in this new embedding space. The resulting embeddings are shown in the scatterplots in Figure 2.3, where each point represents a m/z image. This dimensionality reduction serves a double purpose, firstmost it serves as an important preprocessing step to facilitate subsequent clustering with DBSCAN, and second, mapping to 3 dimensions allows us to visualize the embeddings and get some insight into what is happening. We then use DBSCAN to find regions of high density (i.e. groups of images that are highly similar in the embedding space) in the data, retrieving 21 and 22 clusters for the regular and neural pipelines respectively, as show in table 2.1. The different clusters are assigned different colors in the scatter plot, and we see clear clusters of similar m/z images in the embeddings.

Overall, the neural pipeline shows more localized clusters than the regular one. This is what we would expect, as the neural pipeline should make a better distinction of images based on localized extracted features by the neural network, than the more global cosine distance used in the regular pipeline. We also see a relatively large localized purple cluster in the neural U-D case, which is the noise cluster that DBSCAN defines (517 images); in the regular U-D pipeline this cluster is larger ($1,671$ images, around 20% of total images) and much more spread out in the embedding space.

On inspection, the noise cluster in the regular pipeline shows a great number of images that still contain spatial structure, and were thus wrongly classified as noise. While the neural pipeline was not perfect, the noise cluster contains far less images with spatial structure. We note that the spatial resolution of this dataset is very high, but as a trade-off the mass resolution is relatively low. As such, even though the dataset is not peak-picked, many of the m/z bins contain spatial signal.

Overall, when comparing cluster assignments for the neural and regular pipelines, the neural pipeline showed a much better assignment of ion images, and a greater consistency in the spatial structure of images assigned to individual cluster, resulting in better, more representative mean cluster images. Figure 2.4 shows examples of the clustering of the lymph node dataset for the regular and neural clusterings. The bottom of the Figure shows the full mean spectrum with colors indicating assignments of m/z values to the clusters for neural and regular pipelines (top and bottom, respectively). A zoom-in for part of the spectrum is shown with detailed cluster assignments, showing marked ion images at the top. By their relative spectral distance of ~ 1 Da, and similar spatial distribution, we can assume that these marked images are probably isotopes, and as such we would expect all of these images to be allocated to the same cluster. In the left panel, in purple, we see assignments for each of

Table 2.1: Overview of experimental results per configuration. RIR indicates relative isotope ratio (higher is better, best per dataset marked in bold). Noise images shows the size of the *noise cluster* identified by DBSCAN.

Dataset	Clustering	Ion images	Clusters	RIR	Noise Images
lymph node	UMAP-DBSCAN	neural	22	**66.6**	517
lymph node	UMAP-DBSCAN	regular	21	57.7	1671
lymph node	k-means	neural	20	64.5	/
lymph node	k-means	regular	20	56.2	/
mouse kidney	UMAP-DBSCAN	neural	40	35.4	401
mouse kidney	UMAP-DBSCAN	regular	21	20.4	63
mouse kidney	k-means	neural	60	**35.6**	/
mouse kidney	k-means	regular	20	23.4	/

these images for the regular pipeline. The images shown are the mean images of the cluster to which the different marked images are assigned. Firstly, we see that only two of the isotope images are assigned to the same cluster, whereas the other two images are assigned to different clusters, and secondly, we see that the mean images of the clusters that the images are assigned to, do not closely resemble the original isotope images. The neural pipeline, on the other hand, uses the features extracted by the neural network to correctly assign all images to the same cluster, with a mean image that closely resembles that of the original ion images, thus performing much better at this task. Furthermore, the mean cluster images of the neural pipeline show better definition of spatial structure, compared to those of the regular pipeline, which are generally much more blurry, indicating a more diverse collection of underlying images.

This trend is observable in a large number of the clustered images, as is confirmed when we look at the Relative Isotope Ratio (RIR), in Table 2.1. This ratio measures the overall number of isotopes in a cluster compared to a randomized clustering, as explained in Section 2.3.3. The higher this number, the more isotope ion images are assigned to the same cluster, which is the expected behaviour. Looking at the table, we can see that the neural pipeline scores much better at this metric than the regular one, demonstrating the added benefit of adding a neural network-based interpretation layer to the pipeline.

Finally, we verify the clustering results by comparing the mean cluster ions with the stained H&E microscopy image of a neigboring tissue section. Figure 2.5 shows three example mean cluster images, originating from the neural U-D pipeline, on the right, highlighting different salient structures in the lymph node tissue. The center image shows a composite image overlaid on the microscopy, where each mean cluster image is assigned a different color channel, namely red,

green and blue, that shows a clear overlap of the anatomical structure observed in microscopy with the biochemical information obtained from the MSI data. The green cluster image clearly shows the Germinal Centers (balloon shaped structures) found in lymphoid tissue. These are important areas for humoral immunity as, at these sites, activated B cells (B lymphocytes) accumulate and undergo further processing. The high-resolution, non-rigid registration was performed using Aspect Analytics' proprietary registration pipeline.

k-means clustering results

Next, we apply k-means clustering directly on both the regular and neural ion images, as a baseline comparison, and an example of how neural images can be readily plugged in to existing algorithms. We experimented with different numbers of clusters, and found that, in line with the DBSCAN experiments, 20 clusters provided the best results in terms of detected meaningful clusters and RIR (see table 2.1). For the regular k-means pipeline, this results in 6 clusters which contain only a single image that had no clear spatial structure (and which we thus consider to be noise clusters). This is probably due to the high dimensionality of the data (500.000 pixels - features), and the sensitivity of k-means to noise. The neural pipeline did not show such single image clusters, even when increasing the number of clusters, probably due to the lower dimensionality of the neural images (2048) and a better grouping of the neural ion images due to the feature recognition by the neural network. When comparing the RIR, we see a similar trend to the U-D pipeline, namely an upregulation in the numbers of isotopes that are captured by the same cluster, when using the neural images over the regular ion images. This indicates that clustering results are better with the neural network input than without, which we also saw in manual inspection of the clusters.

2.4.3 Mouse kidney data

In order to assess the general applicability of our methods, we also apply our clustering pipelines on a high mass resolution dataset obtained in mouse kidney, that has a significantly lower spatial resolution than the lymph node dataset, respectively 50 vs. 10 μ m.

As the spatial resolution affects the fine structure observed in the ion images, we want to see whether the neural network approach can improve clustering results similarly to the high spatial resolution lymph node dataset. As discussed above, we upsample the image data, so that patches are generated with a similar physical dimension to those in the lymph node data. We then pass

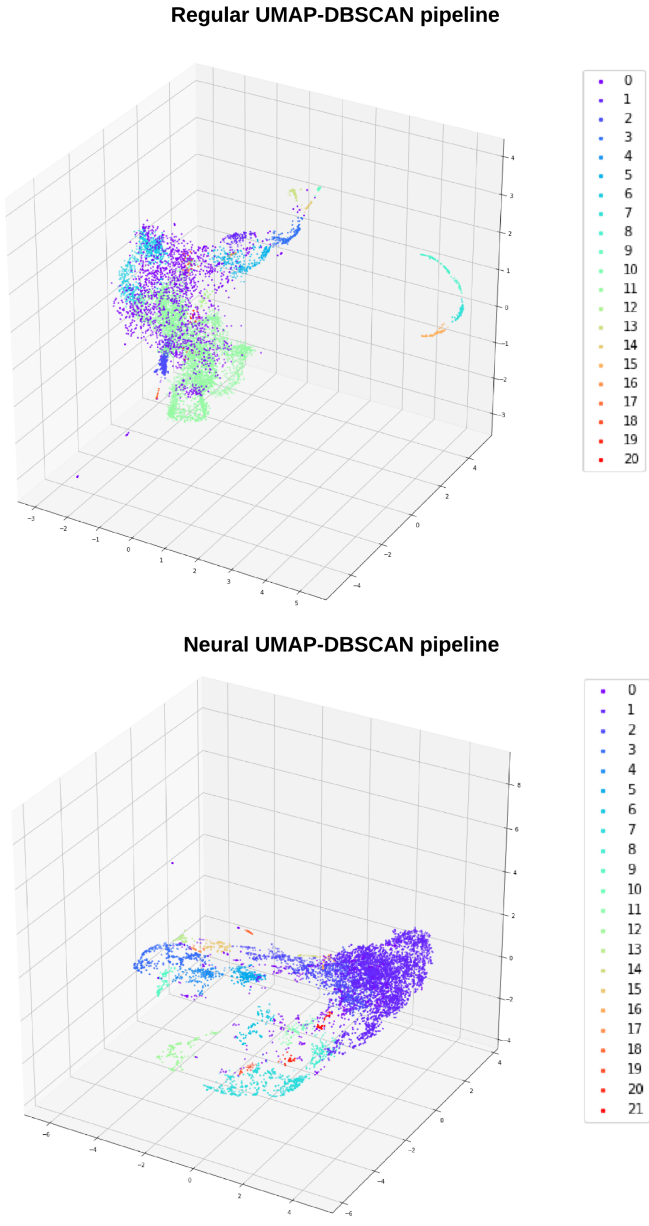


Figure 2.3: Scatter plots of resulting embeddings in 3 dimensional space after applying non-linear UMAP dimensionality reduction and DBSCAN clustering on the lymph node dataset. Each point represents a m/z image, with its colors showing cluster assignment as determined by DBSCAN. The top and bottom panels show clusterings on the regular and neural ion images respectively. The Neural U-D pipeline shows more localized and clearly defined clusters of m/z images than the regular U-D pipeline

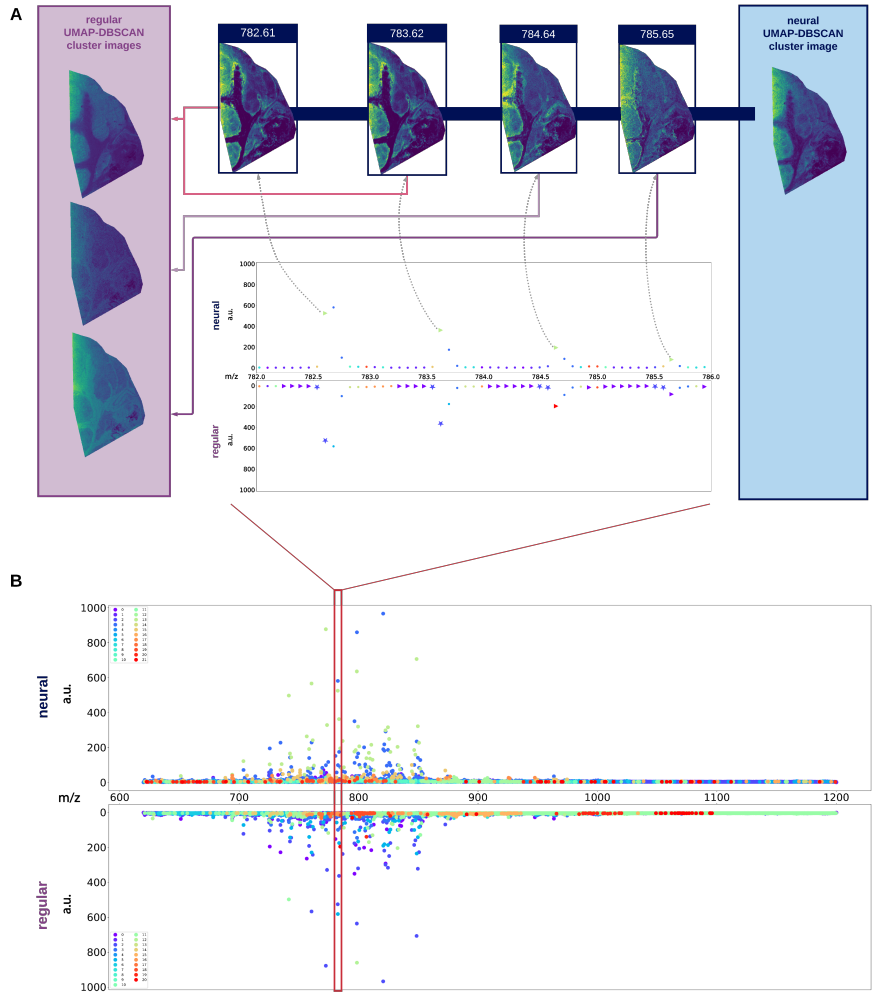


Figure 2.4: Ion image clustering results of the lymph node dataset. Panel B shows the full mean spectrum with cluster assignments for each m/z bin, for the neural (top) and regular (bottom) U-D pipelines. Panel A shows a zoom-in of the mean spectrum, highlighting an isotopic distribution, and four associated ion images. The purple (regular) and blue (neural) panels show the clusters (shown as mean images) to which the associated ion images are assigned. The neural pipeline shows better cluster assignments than the regular pipeline as (i) all isotope images are assigned to one and the same cluster, and (ii) its mean cluster image matches the associated isotopic ion images more closely than the regular pipeline.

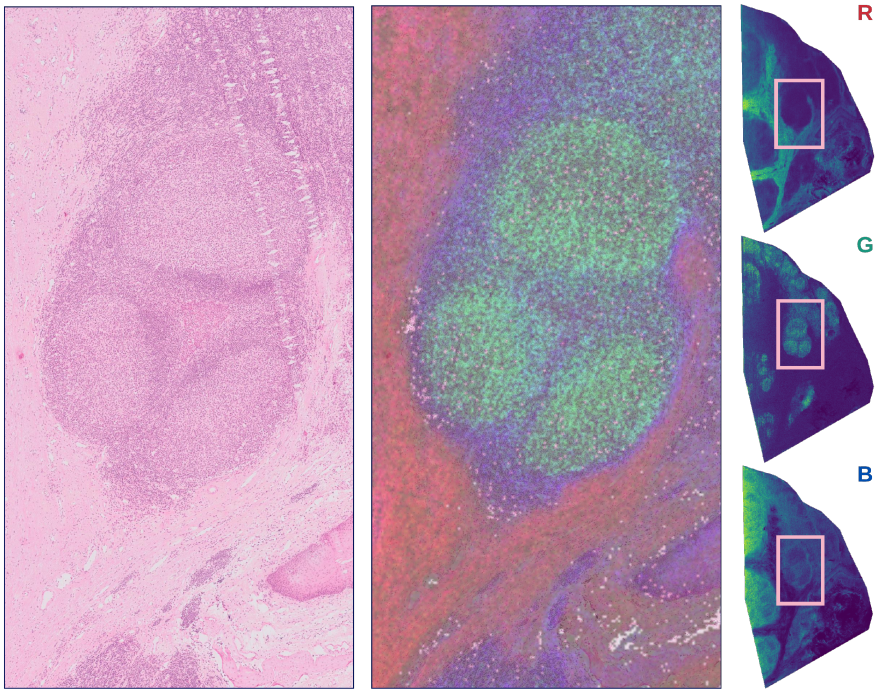


Figure 2.5: (Left) H&E stained microscopy image of neighboring lymph node tissue section. (Right) Three example mean cluster images for the neural U-D pipeline. (Center) Composite image combining the three mean cluster ion images overlaid onto the microscopy, showing clear co-localization of anatomical structures (via morphology) with the biochemical information obtained through clustering of the ion images from MSI.

the ion images through the neural network to perform feature extraction and generate the neural ion images, exactly as we did in the lymph node data.

UMAP-DBSCAN clustering results

We first apply the U-D pipeline to the regular and neural ion images. As can be seen in Table 2.1, DBSCAN finds significantly more clusters in the neural (40) pipeline than the regular (21) one. Furthermore, similarly to the lymph node data, the RIR is significantly higher for the neural pipeline than the regular one, indicating that the retrieved clusters in the neural pipeline succeed in clustering together more isotopic ion images than the regular one, meaning that we not

only get more clusters, but that these are also more relevant. When looking at the noise clusters, we see that this time the noise cluster in the neural pipeline is larger than the regular one, however the size is only 8.2% of the total number of ion images. We note that this time the data is peak-picked, and thus the noise cluster is much smaller.

The quality of the clustering is illustrated by the example shown in Figure 2.6, which shows three sets of ion images corresponding to three different isotopical distributions. Again, the blue panel shows mean cluster images for the neural pipeline, whereas the purple panel shows those for the regular pipeline. Similarly to the lymph node dataset, the neural pipeline correctly assigns images of the same isotope distribution to the same cluster. Contrary to the lymph node data, in this example, the regular pipeline also assigns the isotopic images to the same cluster. However, when comparing the mean cluster images to which the isotopic images are assigned, we see that, while not perfect, those of the neural pipeline match the isotopic images much closer than those of the regular pipeline. The neural pipeline succeeds in creating clusters with a higher specificity, which allows for distinguishing of the salient spatial distributions in the isotopic images at m/z 856.63 and 856.66. In contrast, the regular pipeline groups these patterns together, likely because it uses a more "global" distance measure that is prone to ignoring localized structure. More generally, we see a lot of these clusters that capture fine spatial structure, which are missing in the regular pipeline. Figure 2.7 shows how different mean ion images obtained with the neural U-D pipeline clearly differentiate distinct anatomical regions in the kidney, overlaying these on the H&E of a neighboring tissue section.

k-means clustering results

Finally, we apply k-means to the regular and neural ion images. When testing different numbers of clusters on the regular ion images, we run into the same issues as we did with clustering the regular ion images in the lymph node dataset when we increase the number of clusters, namely that we get a lot of clusters with a single ion image, due to the high sparsity of the data in the original feature space. For the best results in terms of detected meaningful clusters and RIR, we obtained around 20 clusters, similar to the regular U-D pipeline, which collects the most meaningful clusters and where there are no trivial clusters containing single ion images. Again, we do not encounter this issue with the neural pipeline, presumably due to a better packing of the data through neural feature extraction, and the lower dimensionality of the data ($37,365$ pixels per regular ion image versus $2,048$ dimensions in the neural ion images). As such we are able to explore more clusters via the neural pipeline.

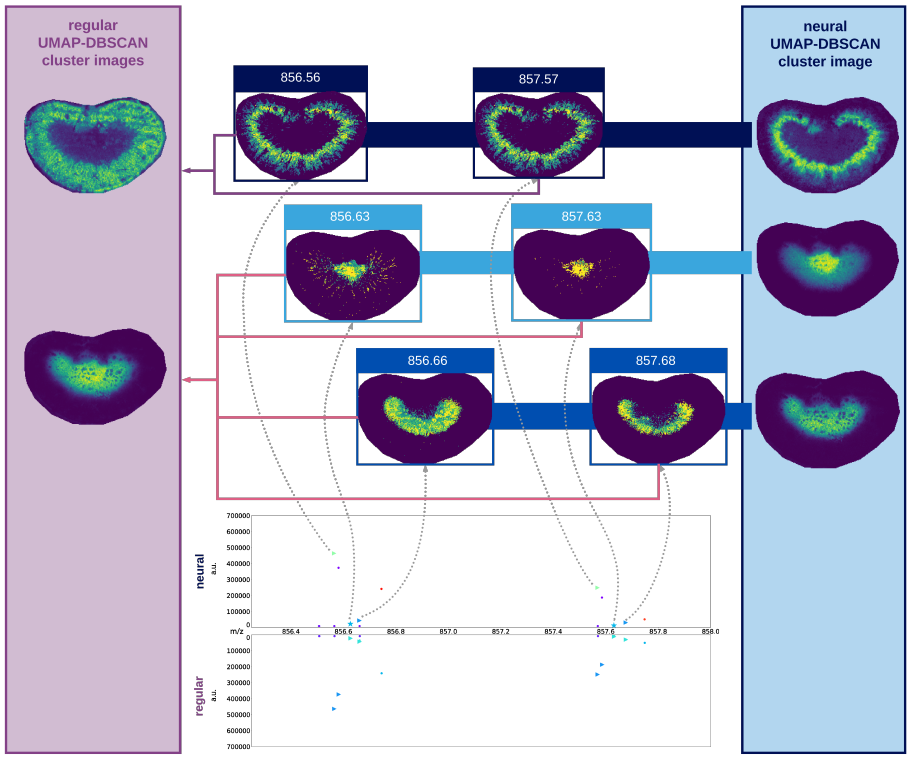


Figure 2.6: Ion image clustering results of the mouse kidney dataset. The Figure shows a zoom-in of the mean spectrum, highlighting an isotopic distribution, and marked associated ion images. Cluster assignments in the mean spectrum are shown for each m/z bin, for the neural (top) and regular (bottom) UMAP-DBSCAN pipelines. The purple (regular) and blue (neural) panels show the mean images for the clusters to which the marked ion images are assigned. Similar to the lymph node dataset, the neural pipeline shows better cluster assignments than the regular pipeline, with mean cluster images that show a closer match to the assigned cluster images.

Similarly to the U-D results, we find clusters that better highlight localized structure in the neural pipeline, however, interestingly, the RIR for k-means is similar to that of U-D, showing that the relatively simple k-means performs surprisingly well on this dataset, compared to the much more involved U-D pipeline.

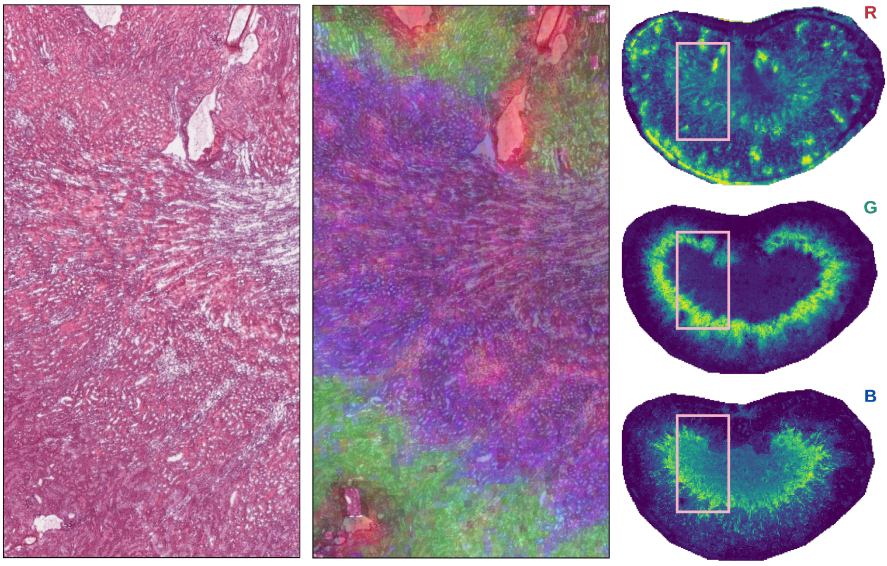


Figure 2.7: (Left) H&E stained microscopy image of neighboring kidney section. (Right) Three example mean cluster images for the neural U-D pipeline. (Center) Composite image combining the three mean cluster ions overlaid onto the microscopy. The co-localization of morphological patterns in the microscopic image and the spatial expressions in these mean cluster images corroborates the biological relevance of the clustering results.

2.4.4 Benchmark on ColocML data

Finally, we benchmark our neural ion image pipeline with the ColocML data by Ovchinnikova et al. [197]. In this dataset, human experts scored a set of ion images from 0 to 10 based on similarity to a given target ion image (0 and 10 meaning identical and nothing alike, respectively). Ovchinnikova et al. calculated the distance between the target image and the set of example ion images for a number of image-based distance metrics, including a distance calculated using an Xception-based pipeline. The similarities for each distance metric were then compared with the expert scores using correlation. From their comparison, the cosine distance showed the best overall performance, outperforming the Xception-based metric by a small margin. The authors indicated some surprise in this finding, given the proven track record of deep learning for such applications.

First, we repeat the benchmark using our variation on the Xception pipeline.

As previously discussed, we found that physical pixel size was of significant importance in generation of the neural ion images. As such, we could unfortunately only use part of the ColocML data, for which this information was available ($n = 23$ out of 182 datasets, listed in SI).

Following the ColocML workflow, we first calculate the neural ion images for the target ion image and the example ion images. We then determine similarity based on cosine similarity for neural ion images, and compare this to the similarity obtained through calculating the cosine similarity directly on regular ion images (using the ColocML methodology).

Using the comparison proposed by [197], we indeed find that cosine distance outperforms the neural ion images, as was also observed in their work.

While the applied methodology of comparing rankings between various algorithms and human experts is conceptually sensible, we propose an adjustment in methodology to make the results less opaque. Upon inspection of expert scores and their associated ion images, we found that images with expert scores less than or equal to 1.0 are highly similar, and up to 2.0 are reasonably similar. Images with a score of 2.0 and higher, however, often differ from the target image in such a way, that scores become subjective, and difficult to interpret, in the authors' opinion (see supplementary info in AppendixA for more details and some examples). Additionally, for many images with scores above 2.0, their relative similarity is often highly debatable. To avoid influencing the benchmark heavily by these fundamentally dissimilar images, we believe it is best to focus on expert scores below 1.0 or 2.0.

Rather than calculating a global score across all images, we propose to break it down into segments according to given expert score, as shown in Figure 2.8. We compared both distance metrics against the human expert scoring at a given cutoff for each segment for each set of ion images, using the following procedure:

1. Select the set of ground truth comparison images $\mathcal{S}_{truth}$ with a score less than or equal to the cutoff, yielding N images.
2. Select the set $\mathcal{S}_{rank}$ of N top ranked images based on the similarity metric (neural ion images / cosine distance).
3. Count how many of the ground truth images are in the top rank via the similarity metric, i.e., $|\mathcal{S}_{truth} \cap \mathcal{S}_{rank}| \in [0, N]$ (higher is better).

From Figure 2.8, we see that neural ion images outperform cosine distance at identifying highly similar images (expert score below 1.0), but its advantage is lost as ion images become more dissimilar. This supports our initial hypothesis

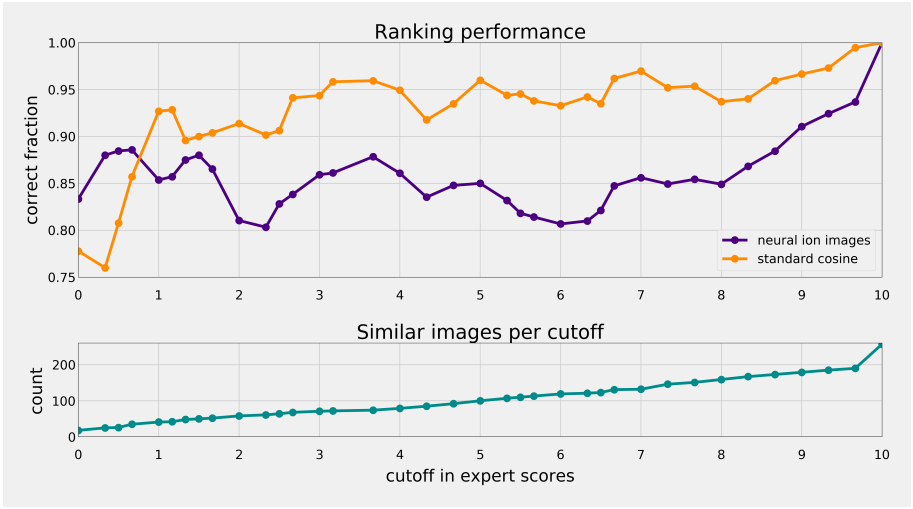


Figure 2.8: Performance comparison between using neural ion images as a distance metric versus cosine distance on raw ion images, compared to human expert similarity scores as determined in [197]. (top) ranking performance for both distance metrics at given cutoff in human expert scoring, displaying the fraction of total ground truth images correctly ranked in the top by each distance metric as outlined above (higher is better). (bottom) the total amount of ground truth ion images per cutoff in expert score.

that the neural ion images should be able to distinguish highly localized features, whereas cosine distance provides a more global similarity measure.

When considering improvements in the application of clustering ion images, as studied in this work, the key is improving the similarity metric for reasonably similar images, rather than images that are dissimilar. Especially in a density-based clustering approach, an improved scoring for highly similar images is a considerable advantage, as this will lead to more condensed clusters.

It is the authors' opinion that this is true for many practical applications, since ranking images based on similarity is only a useful conceptual task if the images are somewhat similar. The more dissimilar a set of images becomes from a given target, the lower the agreement of human experts in their relative ranking (see supplementary info in AppendixA for examples). Conceptually, ranking dissimilar images is comparable to determining whether a cat or dog is more similar to an apple, since the only pertinent conclusion is that neither are similar.

2.5 Conclusions

We set out this work with the observation that the spatial information available in MSI is often under-utilized in its computational analysis, in part due to the fact that it is non-trivial to translate the complex spatial pattern recognition that humans perform on a daily basis into simple algorithms. The advent of deep learning, particularly convolutional neural networks and their derivatives, has significantly advanced the state of the art in computer vision, making it possible to capture high-level abstractions, learned through training complex models based on millions of images.

Moreover, a key strength of neural networks is the fact they can often be transferred between related tasks, such as object detection in computer vision. This enables the direct use of pre-trained neural networks to detect complex spatial patterns without having to train such a model from scratch for each application, which would often be intractable due to the required amount of data and computational resources. In this work we have used such a general-purpose pre-trained model as a general feature extractor to find spatial similarities between ion images in MSI data. Each ion image was fed through the network, and the resulting network interpretation, the neural ion image, was used for subsequent clustering of ion images.

We evaluated the clustering results of two different ion image clustering pipelines, namely the density-based DBSCAN clustering algorithm, combined with non-linear dimensionality reduction using UMAP, and the pervasive k-means clustering algorithm. In both pipelines, we compared regular ion images and neural ion images from two different MSI datasets, namely a human lymph node and a mouse kidney dataset.

All of the tested pipelines allowed for extraction of underlying spatial patterns in both datasets, showing insight in the underlying data. However, in all our experiments, the neural pipelines provided a better assignments of ion images, with more fine-grained clusters, and greater consistency in the spatial structures assigned to individual clusters, resulting in more representative mean cluster images. Specifically, our experiments indicated that using neural ion images enables the subsequent clustering pipelines to better account for localized, salient differences between ion images to obtain a fine clustering result, in contrast to the baseline methods which were less effective in identifying clusters with localized differences.

To quantify our results, we introduced a new metric called Relative Isotope Ratio, which measures the rate at which ion images of the same isotope are assigned to the same cluster, thus capturing biological relevance to a certain degree based on both spatial and spectral information. The newly introduced

RIR metric quantitatively confirmed the merits of using neural ion images, thus corroborating our previously mentioned qualitative observations regarding improved clustering specificity.

Finally, we benchmarked the neural ion images pipeline on the ColocML dataset, and showed that it outperforms cosine distance at identifying highly similar ion images.

The proposed methodology can be readily used to incorporate an advanced form of spatial information into any MSI-focused machine learning pipeline, both supervised and unsupervised, and this without the need for large amounts of training data, or high computational needs. Furthermore, towards the future, a neural network could be trained specifically on MSI data, which we expect will further improve performance of our proposed pipeline.

2.6 Acknowledgements

The authors thank prof. Thomas Tousseyn (Department of Pathology, University Hospitals KU Leuven) for providing the lymphoma tissue sample, and Arndt Asperger (Bruker Daltonik, Bremen, Germany) for his help in the MSI measurements. Wanqiu Zhang was supported by a Baekeland PhD grant from Flanders Innovation and Entrepreneurship (VLAIO). KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C32/16/013, C24/18/022), Industrial Research Fund (Fellowship 13-0260) and several Leuven Research and Development bilateral industrial projects, Flemish Government Agencies: FWO (EOS Project no 30468160 (SeLMA), SBO project S005319N, Infrastructure project I013218N, TBM Project T001919N; PhD Grants (SB/1SA1319N, SB/1S93918, SB/151622)), This research received funding from the Flemish Government (AI Research Program). Bart De Moor and Wanqiu Zhang are affiliated to Leuven.AI - KU Leuven institute for AI, B-3000, Leuven, Belgium. VLAIO (City of Things (COT.2018.018), Innovation mandate (HBC.2019.2209), Industrial Projects (HBC.2018.0405)), European Commission: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 885682); (EU H2020-SC1-2016-2017 Grant Agreement No.727721: MIDAS)

Conflict of interest

The authors declare that they have no conflict of interest.

Chapter 3

Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma

This chapter has been submitted as:

Zhang, W., Patterson, N.H., Verbeeck, N., Moore, J.L., Ly, A., Caprioli, R.M., De Moor, B., Norris, J.L. and Claesen, M., 2022. **Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma**.

Preprint on medRxiv doi: https://doi.org/10.1101/2022.11.29.22282868

3.1 Abstract

Imaging mass spectrometry (IMS) provides promising avenues to augment histopathological investigation with rich spatio-molecular information. We have previously developed a classification model to differentiate melanoma from nevi lesions based on IMS protein data, a task that is challenging solely by histopathologic evaluation.

Most IMS-focused studies collect microscopy in tandem with IMS data, but this microscopy data is generally omitted in downstream data analysis. Microscopy, nevertheless, forms the basis for traditional histopathology and thus contains invaluable morphological information. In this work, we developed a multimodal classification pipeline that uses deep learning, in the form of a pre-trained artificial neural network, to extract the meaningful morphological features from histopathological images, and combine it with the IMS data.

To test whether this deep learning-based classification strategy can improve on our previous results in classification of melanocytic neoplasia, we utilized MALDI IMS data with collected serial H&E stained sections for 331 patients, and compared this multimodal classification pipeline to classifiers using either exclusively microscopy or IMS data. The multimodal pipeline achieved the best performance, with ROC-AUCs of 0.968 vs. 0.938 vs. 0.931 for the multimodal, unimodal microscopy and unimodal IMS pipelines respectively. Due to the use of a pre-trained network to perform the morphological feature extraction, this pipeline does not require any training on large amounts of microscopy data. As such, this framework can be readily applied to improve classification performance in other experimental settings where microscopy data is acquired in tandem with IMS experiments.

3.2 Introduction

Worldwide, skin cancer is the most common and the most deadly cancer in the white-skinned population [2, 150]. Currently, the diagnosis of melanoma is most frequently reliant upon histopathologic evaluation by dermatopathologists, which is inherently challenging. Several studies have reported significant discordance among experts in the diagnosis of challenging melanocytic lesions [100, 264]. For some challenging lesions, even experienced dermatopathologists had difficulty precisely or reproducibly classifying them into a certain category. Therefore, melanocytic lesions are the most common type of lesions that require a second opinion in pathology [287]. However, second opinions can be problematic, and often do not yield the needed accuracy.

One strategy to assist with melanoma diagnosis is by incorporating molecular technologies, to move to a more objective, evidence-based assessment. Imaging Mass Spectrometry (IMS) has been employed increasingly in clinical research and integrated into molecular pathology to complement routine histopathology evaluation [206]. The success of IMS is owed to its ability to characterize tissue microenvironments and provide a crucial link between tissue morphology and its molecular physiology. Many studies using IMS on the skin have already shown the potential to diagnose melanoma [56, 12, 147, 246, 71, 109, 148, 21, 245].

In a previous study [12], we utilized matrix-assisted laser desorption ionization imaging mass spectrometry (MALDI IMS) technology to classify melanoma and nevi lesions, with the results showing high concordance (97.6% sensitivity and 96.4% specificity) with triconcordant histopathological evaluation by board-certified dermatopathologists. MALDI IMS is a beneficial alternative technology for skin cancer diagnosis as it uses tissue sections prepared in a similar manner as standard histopathology with minimal sample material for *in situ* analysis (a single 6 μm tissue section) and can a provide spatial information when using a serial section for histopathological annotation [12].

In our previous study, the pipeline we constructed solely used MALDI IMS data to predict the sample classification. However, for each of the IMS experiments, the pathologists used an H&E stain of the neighboring tissue section to guide the histopathological annotations. These H&E images provide crucial morphological information that pathologists use to diagnose melanoma lesions. By targeting the analytical measurement to regions with the highest probability of malignancy, we strengthen the classification algorithm. We hypothesize that a classification model that combines the H&E imaging data, which is the basis of histopathological diagnosis, with MALDI IMS data, an objective but highly accurate means of melanoma classification, that the overall classification accuracy could be further improved.

Multimodal imaging is a powerful tool which integrates information from different imaging techniques, and the resulting integrated datasets may open new opportunities for data mining. This can be beneficial for MALDI IMS data, as they contain a substantial amount of chemical information but not enough readily available morphological knowledge from the tissue [253]. A multimodal MALDI IMS strategy often involves careful data acquisition, registration and computational data analysis to get the most/best of the different modalities [253]. In this way, complementarities and correlations from both modalities can be explored, providing better evidence towards biological understanding.

Currently, there are very few studies combining both spatial and chemical modalities, i.e., histopathological data and IMS data, to improve the down-

stream classification task (i.e., diagnosis). At present, there is one study that reports the synergy of metabolomics data from MALDI IMS and morphometry from histopathological data to improve the classification of kidney cancer [207]. This study was performed on a large patient cohort (n = 853) using FFPE tissue samples; when the classifier was trained on combined or multimodal datasets (morphometric and metabolite), it outperformed the classifiers trained on the individual data for each tumor subtype. This study shows the potential of using both histopathology and IMS data to improve the downstream diagnosis and tumor subtyping.

In comparison to the work by Prade et al. [207] which used a computer-assisted morphometry image analysis algorithm [91], we applied a pre-trained deep learning-based model to extract the morphological features from the histopathology dataset in this current study. Specifically, we adapted the pre-trained contrastive self-supervised learning model SimCLR [62], pre-trained on 57 histopathology datasets without any labels [65], to extract the high-level morphological features of the histopathological images previously only used to guide the MALDI IMS data acquisition [12].

In the machine learning field, the number of labeled data is found to have a positive correlation with task performance. However, the most tedious and time-consuming step is to annotate the histological data using medical image analysis applications. As labeled data is often not available, many approaches have been proposed to mitigate this bottleneck. For example, unsupervised learning [265] and its subclass self-supervised learning do not require manual annotations of the data, whereas semi-supervised learning uses partially labeled data. Interestingly, self-supervised approaches obtain 'labels' from the data itself and predict part of the data from other parts. Instead of focusing on finding specific patterns from the data like unsupervised learning often does, e.g., via clustering, self-supervised learning methods concentrate on recovering whole or part of some features from its original input, while still fitting in the type of supervised settings [156].

In this paper we adopted the pre-trained contrastive self-supervised learning model from Ciga et al. [65] and applied it on our previously collected 331 patients histopathology datasets [12]. Whereas in Al-Rohil et al. [12], the histopathological images were used to only guide MALDI-IMS data acquisition, here we also used them in the multimodal data analysis to improve the downstream classification task. Our method successfully extracts relevant and highly qualitative morphological features from the histological images, without requiring any training nor labels. Finally, we combine the extracted morphological feature with MALDI IMS data in the downstream classification task to distinguish melanoma from nevus. We compare the classification performance of the unimodal and multimodal strategies. Our multimodal

approach shows improvement over the unimodal MALDI IMS classification results, as reported in our previous work [12]. Our approach shows the potential of the synergism of spatial proteomics and morphology in (melanoma) cancer diagnosis.

3.3 Materials and Methods

3.3.1 Tissue Collection

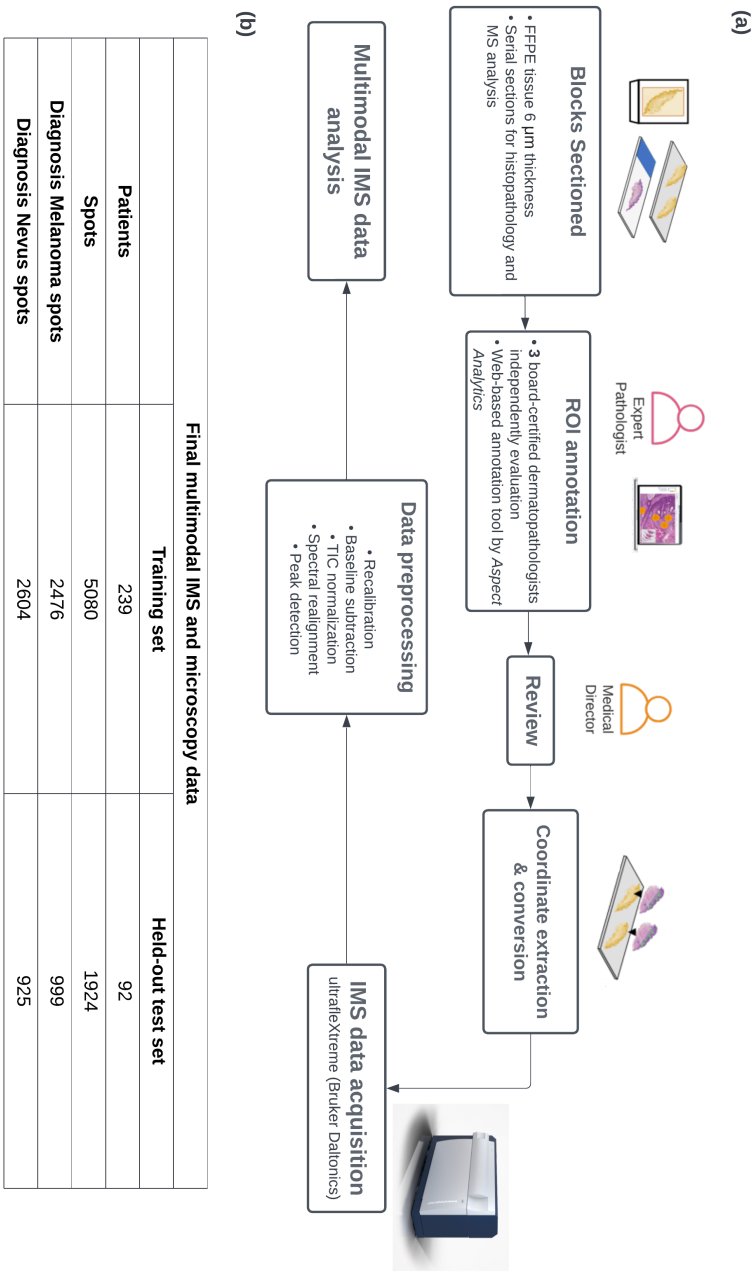
The tissue samples used in this study were the same cohort as used in Al-Rohil et al. [12]. Deidentified FFPE skin biopsies from patients diagnosed with invasive melanoma and nevi were provided by the co-operating academic institutions and private practices. All samples were collected between 2010 and 2016.

The review boards of Indiana University (IRB 1910325060), Vanderbilt University (IRB 030220), Duke University (IRB Pro00102363) and Sterling Institutional Review Board (Atlanta, GA) determined this study exempt from IRB review under the terms of the US Department of Health and Human Services Policy for Protection of Human Research Subjects at 45 CFR §46.104(d).

3.3.2 Histopathology Sample Preparation and Data Acquisition

Serial sections (6 μm) were cut from deidentified patient biopsy FFPE tissue blocks and used for histopathology and mass spectrometry analysis. Figure 3.1(a) illustrates the workflow of histology-guided IMS sample preparation and processing.

Sections used for histopathological analysis were prepared and H&E stained according to standardized institutional practices and scanned at 20x magnification using a Leica SCN-400 digital slide scanner (Leica Biosystems, Buffalo Grove, IL, USA). The digitized sections were independently annotated by three board-certified dermatopathologists in a blinded fashion using web-based annotation tools: PIMS [194] and Annotation studio (Aspect Analytics NV, Genk, Belgium). Each sample was provided to the pathologist with the patient demographic and with the goal of classifying the tissue as either melanoma or nevus.



(b)

Final multimodal IMS and microscopy data		
	Training set	**Held-out test set**
Patients	239	92
Spots	5080	1924
Diagnosis Melanoma spots	2476	999
Diagnosis Nevus spots	2604	925

Figure 3.1: Overview of the histology-guided IMS sample preparation and processing workflow (a) and information about the final included number of samples and spots (b). The workflow (a) leads into multimodal data analysis and classification, incorporating both chemical information from MALDI IMS and morphological information from H&E microscopy.

Using the annotation software, the dermatopathologists annotated regions of interest from each sample/slide with a 50 μm diameter. While 6 different annotation labels were placed (melanoma *in situ*, invasive melanoma, junctional nevus, intradermal nevus, uninvolved epidermis and dermal stroma), this study focused only on building a binary classifier to distinguish the melanoma class from nevus class.

3.3.3 MALDI IMS Sample Preparation and Data Acquisition

The sample preparation is the same as that described by Al-Rohil et al. [12]. All sections for MALDI IMS were mounted on ITO slides (Delta Technologies, Loveland, CO). Briefly, sections were incubated in an oven at 55°C for 1 hour, deparaffinized (xylene x 2), rehydrated in ethanol (100% x 2, 95%, 70%), washed in ultrapure water (x2) and dried at 37°C. Antigen retrieval was conducted in EDTA buffer (pH 8.5) in an Instant Pot Electric Pressure cooker (IP-Duo60 V3) for 10 min on high pressure mode. For *in situ* tryptic digestion, sections were coated with 0.074 μg/mL mass spectrometry-grade trypsin (Sigma Aldrich) in a 100 mM ammonium bicarbonate buffer solution using a M5-Sprayer (HTX Technologies, Chapel Hill, NC), and incubated at 37 °C overnight. After digestion, the samples were coated in 5 mg/mL α-cyano-4-hydroxycinnamic acid matrix in 90% acetonitrile/deionized water and 0.2% TFA using the M5-Sprayer.

The dermatopathologist annotations were used to generate coordinates to guide the MALDI IMS acquisition from the same region/s on the unstained serial sections. MALDI IMS was performed on an ultrafleXtreme MALDI-TOF MS (Bruker Daltonics, Billerica, MA, USA) fitted with a SmartBeam Laser Nd:YAG laser following previously published parameters [12]. Data was collected in positive reflector mode over the range of 700 to 3500 Daltons, at 500 laser shots per each profile spectrum for each region of interest.

3.3.4 Data Processing and Sample Selection

A series of preprocessing methods were applied on the acquired IMS data, including resampling of the data to a common m/z axis, baseline subtraction, TIC normalization, spectral realignment and peak detection [12].

For the multi-modal image analysis, we only included tri-concordant samples (n = 331), meaning all dermatopathologists agreed with the diagnosis. Samples that were either diagnostically challenging, or did not meet the minimal mass spectral quality standards were excluded. Cases where only one modality is

available (n=2) were also excluded from this study. Other than that, we used the same samples as the previously reported study by Al-Rohil et al. [12], and can compare the multimodal results with the unimodal IMS classification reported therein.

The included samples contained 21 collected spots on average. As we focused on the binary classification task between melanoma and nevus, spots with annotations as melanoma (i.e., invasive melanoma and melanoma *in situ*) and nevus (i.e., junctional nevus and intradermal nevus) were included (n=7004). It must be noted that, in this study, we measured the classification performance only at the spot level and not at the aggregated patients/samples level. Our spot-level results can be directly compared to those reported in Al-Rohil et al. [12].

The final multimodal IMS and microscopy dataset was randomly split into a training and a held-out test set at the patient level in order to prevent information leaks. Specifically, we designed our splits such that training and held-out test data spots originate from different patients, to ensure we are accurately assessing generalization performance on unseen slides, both during cross-validation fold generation (in the training data) as well as in determining the training vs held-out sets. The held-out test set in this study means that the data was never used during the whole model training part (during the parameter tuning, model selection). Figure 3.1(b) shows the information of the final included number of patients and the collected spots.

3.3.5 Data Analysis

In this study, we built 3 different classification pipelines for melanoma diagnosis, namely the *unimodal MALDI IMS*, the *unimodal microscopy*, and the *multimodal pipeline*. As illustrated in Figure 3.2, the downstream classification part is the same across all compared pipelines but the inputs are different. For unimodal pipelines, only a single modality, either IMS data or microscopy, is utilized as the input for the downstream classification task, whereas in the multimodal setting, both modalities are used as the input for classification model training. Finally, these 3 pipelines are compared based on their classification performance.

Unimodal IMS: Prior to feeding to the classification model, IMS data was preprocessed and peak picked. Each mass spectrum was generated from a single spot, which has $50\,\mu$m. After peak picking, a total number of 5558 peaks remained. Then we feed the preprocessed IMS data to the downstream classification model. This pipeline was reproduced from our previous study [12] and compared with the multimodal pipeline in this study.

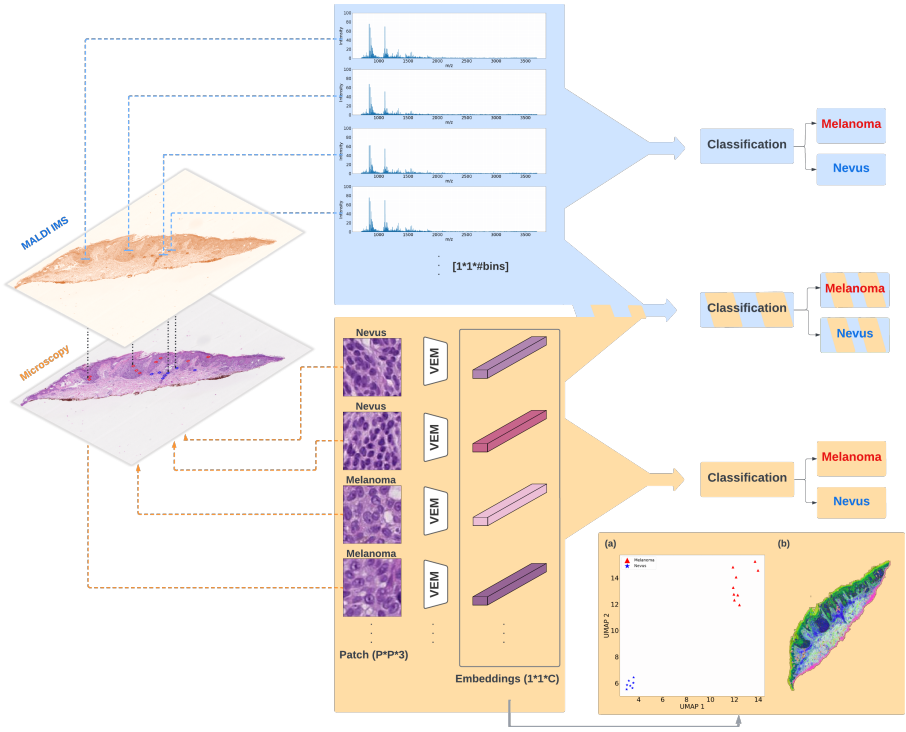


Figure 3.2: Data analysis workflow on an example tissue: **top blue panel** represents the unimodal IMS classification pipeline, where only IMS data are used for the downstream classification task; **bottom orange panel** shows the unimodal microscopy classification pipeline, where only morphology features are included for the melanoma diagnosis; In addition, intermediate embeddings from unimodal microscopy model were visualized via UMAP method: (a) 2-D embeddings of 16 patches (measured spots) from this example tissue with assigned colors based on pathologists' annotations; (b) hyperspectral visualization of 3-D embeddings from all patches across the whole example tissue; **middle fused panel** is the multimodal strategy, where both IMS data and morphology features are used to distinguish melanoma from nevus. *#bins=5558, P=96, C=512, VEM: vector embedding morphology model.*

Unimodal Microscopy: We applied a pre-trained deep learning-based, self-supervised model to extract the features from microscopy data[65]. Each 3-color-channel histopathological image is transformed to a single dense vector and then fed to the downstream classification task. This pre-trained model was adapted originally from Chen et al. [62], where the model learns the feature from data by maximizing the agreement between two stochastically augmented views of the same image via a contrastive loss function. Augmentation plays an essential role in contrastive learning. In the context of the histopathology dataset, some augmentation algorithms from the original paper [62] may not be suitable here, as it was applied on natural-scene images like ImageNet [76]. Thus, Ciga et al. [65] conducted a series of experiments and some minor modifications on data augmentation. In the end, the augmentation algorithms included in the study are listed as following: randomly resized crops, 90 degrees rotations, horizontal and vertical flips, color jittering and Gaussian blurring.

Ciga et al. [65] pre-trained their model on 57 multiple multi-organ histopathology datasets without any labels, as it is a self-supervised learning model. Those datasets contain various types of staining and resolutions, which helps the model learn better quality features. As reported in Ciga et al. [65], the features learned from the model that was pre-trained on those histopathology datasets, are better than the features that were learned from networks pre-trained on ImageNet. In total, more than 4 millions of patches (in pixel size of 224*224) are extracted from histopathology datasets and then fed to the model training.

In this study, we adapted this self-supervised model, pre-trained on a large number of histopathologic images, to the previously collected H&E stained histopathological images from our previous study [12]. Each spot generates a single unique mass spectrum, and has a diameter of 50 μm. Thus, each patch in the histopathology dataset has a physical size of 48 μm*48 μm. The collection of histopathological patches consists of mainly of 20x resolution samples (96*96 pixel size), and a small portion of 40x resolution samples (192*192 pixel size). We applied some preprocessing steps on images following the suggestions in [65], including resizing, and normalization. Finally, each microscopy patch in the dataset is embedded into a vector representation via this pre-trained deep learning based self-supervised learning model, referred to as a *'Vector Embedding Morphology'* (VEM) model in this paper. The resulting embeddings contain 512 dimensions. In order to visualize high-dimensional abstract features, we applied the method Uniform Manifold Approximation (UMAP) [181] to project the data into lower dimensions. UMAP is a dimensionality reduction method, which seeks to search for a low-dimensional representation of the data that has the closest possible fuzzy topological representation of the original high-dimensional data. Previously, Smets et al. [229] evaluated the performance of the UMAP method on IMS data and compared it with

other popular dimensionality reduction algorithms like t-distributed stochastic neighbor embedding (t-SNE) and principal component analysis (PCA).

Multimodal IMS and Microscopy: In the multimodal setting, we represent the data for each spot with a single vector, which is created by concatenating the preprocessed IMS data for a spot together with the extracted morphological features from microscopy data for that same spot, with equal weights. In this way we achieve a fusing of the two modalities, which serves as an input for the classification algorithm.

Classification model: We applied Support Vector Machines (SVMs) [68, 168] to build the binary classifiers for each pipeline. As a popular supervised learning model, SVMs try to find a hyperplane as a decision boundary to separate data points with different classes, where this hyperplane has the maximal distance between data with different labels. To implement the SVMs algorithm, we used the Python package Scikit-learn [204] with a linear kernel.

Model selection and evaluation: During the parameter tuning for SVMs models, we used a grid search strategy, where the model exhaustively generates candidates from a grid of parameter values that are predefined by the user. Furthermore, we leveraged *nested cross-validation* in order to estimate the generalization error underlying each model including the performance variability induced by the hyperparameter optimization process itself. Additionally, in order to further prevent information leaking, spots measured from the same patient/tissue were grouped in the same fold during the whole nest cross-validation process. In summary, we first split the total samples into a training set and a held-out test set on a patient level. Then we applied a grouped-nested-cross-validation strategy on the training set for model selection and evaluation. The held-out test set was not used during the whole training procedure. Thus the whole framework was designed carefully during the model selection and evaluation, so that we can fairly compare classification performance from unimodal and multimodal pipelines.

3.4 Results

In this study, we built 3 different pipelines namely the unimodal microscopy, the unimodal IMS and the multimodal classification pipelines. Unimodal pipelines only use one modality i.e., either IMS or microscopy data for the downstream melanoma diagnosis, while multimodal strategy utilized both modalities. In order to compare those 3 pipelines, we conducted various measurements to assess their classification performance. In addition, we also

applied a dimensionality reduction method UMAP to visualize the intermediate data in lower dimensions.

3.4.1 UMAP visualization of intermediate data

In order to provide more intuition for these highly abstract features, we used UMAP to project and visualize the extracted morphology feature from the unimodal microscopy pipeline. As shown in Figure 3.2(a), after the feature extraction, each of those 16 patches (measured spots from the example tissue) is transformed into a single vector and shown in a 2-d scatter plot. Each dot in the scatter plot represents a vector representation of a microscopy patch and is assigned a red or blue color based on the pathology annotations either melanoma or nevus, respectively. Clearly, all melanoma patches are grouped together and far away from the cluster of nevus patches. Additionally, we collected all patches across the whole slide of this example tissue and used UMAP to project their 512-d transformed morphological embeddings into 3-d feature space. The resulting 3-d UMAP embeddings are visualized via translating each location of the patch to the RGB color scheme, which is referred to as hyperspectral visualization and shown in Figure 2(b). Regions where share similar colors represent their morphology are similar with each other. These results show the potential of the morphology features learned from *'Vector Embedding Morphology'* (VEM) used in the unimodal microscopy pipeline, and provide more interpretation of these abstract embeddings. Importantly, this feature extraction process requires no training nor labels.

Furthermore, we performed UMAP on each input data from unimodal and multimodal pipelines. Three different inputs are collected and compared in this study, namely, (1) preprocessed IMS data (n_dimension = 5558), (2) morphological features extracted from microscopy data via a pre-trained self-supervised learning model (n_dimension = 512), (3) the combination of normalized (1) and (2) with equal weight (n_dimension = 5558+512 = 6070). UMAP was conducted to map those high-dimensional inputs to 2 dimensions in an unsupervised manner. Additionally, we also provide the labels for each data point in order to have a better intuition of the representations from those different input data.

As shown in Figure 3.3, each dot in the scatter plot represents a mass spectrum or a vector representation of a histopathological image or the combination of both, with its color showing the annotation provided by 3 dermatopathologists. Each column shows the UMAP embeddings of features from microscopy data, IMS and the combined data, respectively from left to right. The first and second rows represent the input data from the training dataset and the independent

held-out test set respectively. We applied UMAP via Python implementation [182] with cosine distance as metric for each type of input data.

3.4.2 Classification performance

In this study we applied various classification performance measurements such as, Receiver operating characteristic - Area under the curve (ROC-AUC), F1, precision and recall scores. The range of those metrics is typically from 0 to 1, where 1 represents the perfect classification performance. ROC curves are widely used to measure performance in binary classification tasks, as it plots the true positive rate (TPR) and the false positive rates (FPR) with different thresholds. By computing the area under the ROC curve, the performance across the full range of classification thresholds can be summarized in a single number.

As mentioned earlier, we carefully performed the *grouped-nest-cross-validation* strategy throughout the modeling process in order to avoid any possible information leaking. We applied this strategy on the training dataset and fit the optimized classification model for each pipeline on an independent held-out test set. The results are shown in Figure 3.4.

We reproduced the unimodal MALDI IMS classification results from our previous study[12], and obtained comparable ROC-AUC results (0.910 vs. 0.915 on training set, 0.933 vs. 0.938 on the held-out test set, for reported spots-level results from Al-Rohil et al. [12] and our reproduced results, respectively). Interestingly, the unimodal microscopy pipeline achieves higher mean ROC-AUC scores than the unimodal IMS (0.937 vs. 0.915) on the training dataset, while achieving comparable ROC-AUC results on the held-out test set. The result shows the advantage of using a pre-trained deep learning model for extracting important morphological features from microscopy data. However, it must be noted that the annotations used to train and evaluate the classification model are fully based on microscopy data by dermatopathologists.

By combining information originating from the two available modalities in a single classifier, i.e. the multimodal pipeline, we obtain the best classification performance, (ROC-AUC as 0.968 on training and held-out test set). Additionally, in order to check whether the improvement presented by our proposed multimodal pipeline was significant compared to the previous study [12] (i.e., unimodal IMS pipeline), we applied the DeLong test to compare two AUC results obtained from the unimodal IMS and multimodal pipelines on the independent held-out test set. The resulting p-value ($<$2.2e-16) confirms the significance of the improvement using the multimodal pipeline. This demonstrates the synergy in combining observations in the biochemical domain,

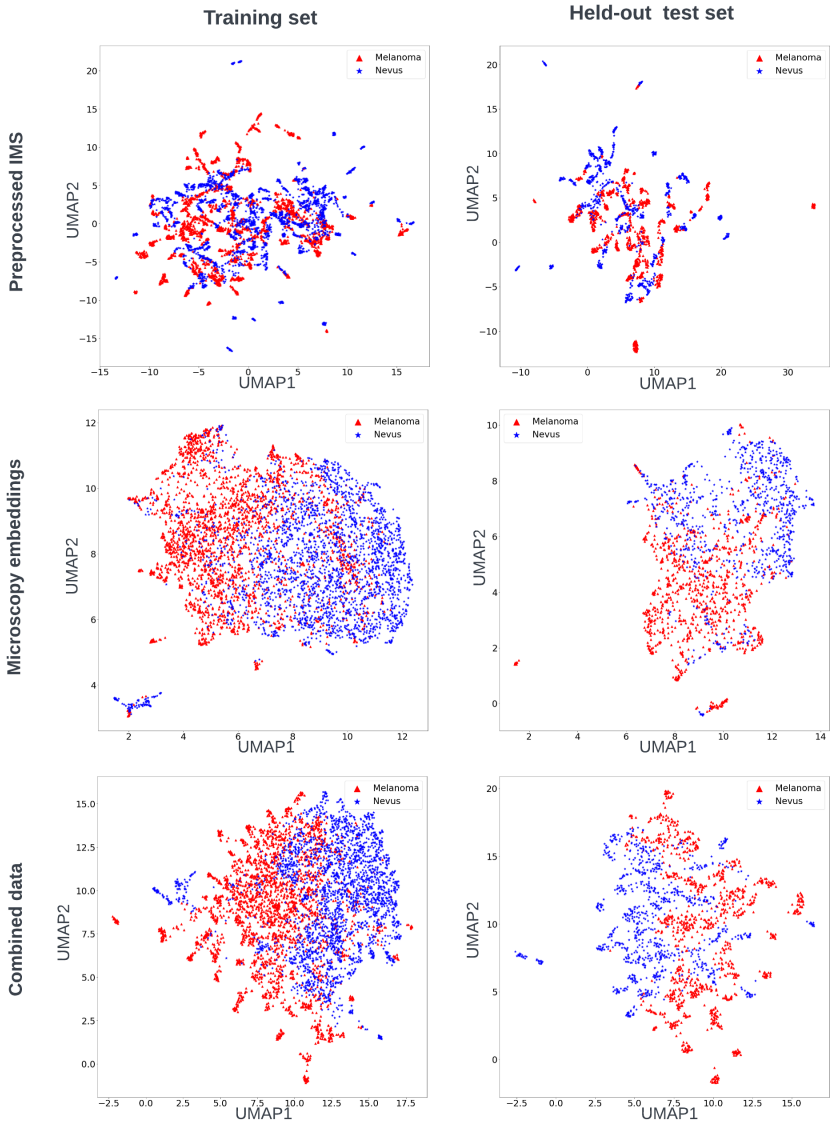


Figure 3.3: UMAP 2-D visualization of each input data across pipelines (unsupervised). From top to bottom rows are preprocessed IMS data, extracted morphological features from microscopy data and an equally weighted combination of both. From left to right columns are from the training dataset and the held-out test dataset. The assigned color of each spot is decided by its diagnosis from 3 dermatopathologists.

(a) Classification performance from training and held-out test dataset

Model	ROC-AUC		F1 score		Precision		Recall	
	Training set	Held-out test set	Training set	Held-out test set	Training set	Held-out test set	Training set	Held-out test set
Unimodal IMS	0.915 (0.049)	0.931	0.823 (0.067)	0.856	0.846 (0.085)	0.828	0.815 (0.11)	0.886
Unimodal microscopy	0.937 (0.03)	0.938	0.82 (0.056)	0.861	0.857 (0.104)	0.871	0.805 (0.111)	0.851
Multimodal	**0.968** (0.023)	**0.968**	**0.866** (0.056)	**0.91**	**0.920** (0.051)	**0.924**	**0.83** (0.112)	**0.90**

Figure 3.4: Classification results on training and held-out test sets. All results are based on **spot-level**. (a) shows the classification performance of each pipeline on training and held-out test dataset. The results from the training set represent the mean value of the performance (with their standard deviation) after the nested-cross-validation process. (b)and(c) show the precision-recall and ROC cures plots of each pipeline on the **held-out test set**.

obtained through IMS, with those in the morphology domain, which have been traditionally used in melanoma diagnosis.

The unimodal IMS and unimodal microscopy pipelines show comparable F1 scores across both training and held-out test sets. Unimodal IMS outperforms unimodal microscopy on both training and held-out test sets based on the recall scores (0.815 vs. 0.805, 0.886 vs. 0.851). On the other hand, unimodal microscopy achieves higher precision scores than unimodal IMS on training and held-out test sets. These slight variations in terms of which approach is best are attributable to differences in classification thresholds, thus emphasizing the importance of using an overall metric like ROC-AUC that accounts for the classifiers' full range.

We also computed the standard deviation for each score based on the training data, and the multimodal pipeline achieves the lowest values almost across all metrics. For example, standard deviations from ROC-AUC: 0.023 (multimodal) vs 0.03 (unimodal microscopy) vs 0.049 (unimodal IMS). This indicates higher stability of the multimodal model, which additionally reinforces the value of combining both modalities.

Additionally, we plotted the Precision-Recall (PR) curve and ROC curves from the independent held-out test set in Figure 3.4(b) and (c) respectively. PR curves illustrate the trade-off between precision score and recall score with different thresholds. Similarly to ROC-AUC, higher area under the PR curve (PR-AUC) represents higher recall and precision across all classification thresholds. PR and ROC curves are typically used to assess the binary classification performance. As shown in Figure 3.4(b)(c), the multimodal pipeline achieves the highest PR-AUC and ROC-AUC scores (0.969, 0.968).

Furthermore, we encountered some interesting misclassified cases on the held-out test dataset. As shown in Figure 3.5, the multimodal strategy shows its power in accurate classification of melanoma and nevus, even when both unimodal strategies fail to yield correct predictions. This highlights that our multimodal classification model is greater than the sum of its (unimodal) parts.

3.5 Discussion

This study extends and improves the previous melanoma diagnosis study by Al-Rohil et al. [12] by incorporating the morphological features from microscopy data in the classification model along with MALDI IMS data. The spot-level ROC-AUC results from the classification model have been largely improved from previous study strategy namely unimodal MALDI IMS (0.915 vs. 0.968,

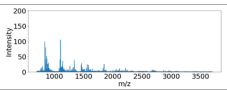
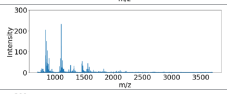
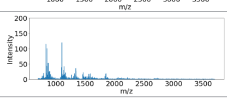
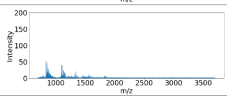
Microscopy data	Imaging mass spectrometry data	Unimodal microscopy prediction	Unimodal IMS prediction	Multimodal prediction	Experts' evaluation
		Nevus	Melanoma	Melanoma	Melanoma
		Nevus	Melanoma	Nevus	Nevus
		Melanoma	Melanoma	Nevus	Nevus
		Nevus	Nevus	Melanoma	Melanoma

Figure 3.5: Misclassified cases by unimodal algorithms from the held-out test set. Red and green text indicate model results that are incongruent and congruent with expert pathology assessment, respectively.

training data), as well as the other scores such as F1 score, precision and recall. Additionally, the classification model training became more stable and robust with the help of multimodal strategy, evidenced by lower variability in model performance.

At present, most IMS classification studies largely focus on only using histopathology data to make annotations that guide the IMS data collection or restrict which spectra are used in order to build more accurate classifiers[73, 103]. Secondly, multimodal MALDI IMS analysis requires careful data acquisition, dedicated registration before the downstream data analysis, which is not always trivial. Last but not the least, learning high-quality morphological features from microscopy data could require a large amount of manpower for dedicated annotations and also introduce computational challenges.

To our knowledge, there is only one study [207] utilizing the information from histopathology data, via computer-aided morphometric-based image analysis, in combination with MALDI IMS to improve renal carcinoma subtyping. In comparison with their work, instead of using morphometric-based algorithms, we applied a deep learning-based self-surprised model (pre-trained on a large histopathology dataset) to extract the high-quality morphological features from microscopy data in a fully unsupervised manner. This model is highly efficient as there is no training required during the feature extraction. However, the extracted features are highly abstract, which is a common challenge in the deep learning field. Thus the resulting features from microscopy data are not

as interpretable as the ones learned from the morphometric-based method as used by Prade et al. [207]. To improve the interpretability of the learned features, we applied the dimensionality reduction method like UMAP on those high-dimensional features (512) and projected them into 2 & 3-d feature space.

Finally, we built a linear classifier on the top of IMS features and/or morphological features to distinguish melanoma from nevus. The classification results show that a multimodal strategy, with a relatively simple equal weighted combination method, can outperform other unimodal workflows in both prediction accuracy and robustness. Microscopy images are acquired in tandem in a majority of IMS experiments, and as such this methodology provides excellent opportunities to improve classification and downstream machine learning at very little additional expense. Due to the use of pre-trained networks, which do not require additional training, this additional feature extraction step from microscopy images is near 'plug-and-play', and could be readily incorporated with IMS experiments on other types of tissues in the future.

3.6 Conclusion

With the increased prevalence of multimodal studies, especially those integrating traditional histological or cellular stains with molecular analyses, it is important to develop computational approaches that effectively combine the data to go beyond qualitative comparisons. A critical area of such combinations is clinical assays where the increase in information from modality combinations has the potential to improve health outcomes. In this work, we demonstrate the potential of a multimodal strategy on a relatively large patient cohort (n=331) which combines morphological features extracted from standard clinical microscopy data and MALDI IMS molecular data the melanoma diagnosis. In future work, this strategy can be further validated on other types of cancer tissues, more advanced fusion strategies could be developed, and the approach could be applied to other combinations of morphological optical images and molecular data.

3.7 Acknowledgement

The authors thank Ahmed Alomari, Rami Al-Rohil, Jason Robbins for for providing pathological analysis, and Thomas Moerman for software development. This work was supported by Small Business Innovative Research

Grant No. R44CA228897-02, provided by the National Cancer Institute of the National Institutes of Health. Wanqiu Zhang was supported by a Baekeland PhD grant from Flanders Innovation and Entrepreneurship (VLAIO). KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117, C3I-21-00316), Industrial Research Fund (Fellowships 13-0260, IOFm/16/004) and several Leuven Research and Development bilateral industrial projects; Flemish Government Agencies: FWO: EOS Project no G0F6718N (SeLMA), SBO project S005319N, Infrastructure project I013218N, TBM Project T001919N; PhD Grant (SB/1SA1319N), EWI: the Flanders AI Research Program, VLAIO: CSBO (HBC.2021.0076) Baekeland PhD (HBC.20192204), European Commission: European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant grant agreement No 885682); Other funding: Foundation 'Kom op tegen Kanker', CM (Christelijke Mutualiteit).

Chapter 4

Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer

This chapter is in preparation of submission as:

Zhang W.*, Spotbeen X.*(co-first-author), Vanuytven S.**, Kint S**.(co-second-author), Sarretto T., Socciarelli F., Vandereyken K., Dehairs J., Idkowiak J., Wouters D., Alvira Larizgoitia J., Partel G., Ly A., Mantas M., Gevaert T., Devlies W., Mah C., Butler L., Loda M., Joniau S., De Moor B., Sifrim A., Ellis S., Voet T., Claesen M., Verbeeck N., Swinnen J., 2023. **Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer**.

Preprint on bioRxiv doi: https://doi.org/10.1101/2023.08.28.555056

Part of this work has been submitted to the Belgian Patent Office: Zhang, W. , Claesen, M., Verbeeck, N., 2023. Method for vertical integration and analysis of spatial multi-omics data. No. BE2023/5080. Feb 2023.

4.1 Abstract

Recent advances in spatial omics methods are revolutionising biomedical research by enabling detailed molecular analyses of cells and their interactions in their native state. As most technologies capture only a specific type of molecules, there is an unmet need to enable integration of multiple spatial-omics datasets. This, however, presents several challenges as these analyses typically operate on separate tissue sections at disparate spatial resolutions. Here, we established a spatial multi-omics integration pipeline enabling co-registration and granularity matching, and applied it to integrate spatial transcriptomics, mass spectrometry-based lipidomics, single nucleus RNA-seq and histomorphological information from human prostate cancer patient samples. This approach revealed unique correlations between lipids and gene expression profiles that are linked to distinct cell populations and histopathological disease states and uncovered molecularly different subregions not discernible by morphology alone. By its ability to correlate datasets that span across the biomolecular and spatial scale, the application of this novel spatial multi-omics integration pipeline provides unprecedented insight into the intricate interplay between different classes of molecules in a tissue context. In addition, it has unique hypothesis-generating potential, and holds promise for applications in molecular pathology, biomarker and target discovery and other tissue-based research fields.

4.2 Introduction

Recent developments in spatial-omics technologies have revolutionised tissue-based research, enabling a deeper understanding of the cellular architecture of tissues, cell-cell interactions as well as the intricate relationship between molecular identity and tissue structure. These technologies find broad applications in studying normal tissue development [27], as well as in deciphering complex diseases like neurodegenerative disorders [63, 149], and different types of cancer [39, 24, 126, 279, 213, 238]. To investigate the various classes of molecules (RNA, proteins, metabolites, lipids) in a spatial context, different technologies are applied, including in situ hybridization, in situ sequencing and in situ RNA capturing followed by ex situ sequencing for spatial transcriptomics (ST) [235, 185, 60, 216, 157, 134, 61, 163, 212, 165], histochemistry using fluorescently labelled antibodies [165, 254, 119] or genetically encoded fluorescent protein tags [25, 290] for spatial proteomics, and mass spectrometry imaging (MSI) to study the distribution of peptides, metabolites, and lipids [54, 274, 78, 188, 87, 22].

The biological questions addressed by such platforms are often complex and involve molecular alterations at multiple levels of the gene expression cascade. As each of the spatial-omics technologies often captures only a partial aspect of a single molecular layer in a tissue, integration of spatial multi-omics (SMO) datasets that span across the biomolecular scale is essential to provide insight in the complex interplay between the different molecular layers and to better understand the intricacy of these biological processes, and their molecular drivers. Pipelines aiming to align and correlate these datasets encounter several obstacles, with the limited coverage of the molecular landscape being a prominent example. Moreover, distinct spatial technologies may not be compatible on a single tissue section, making the use of neighbouring tissue sections necessary [262]. As the structural organisation of tissues can differ between consecutive sections, and the experimental handling for the various platforms may result in tissue deformation, precise co-registration of adjacent datasets has proved difficult but is essential in order to align corresponding regions and prevent misinterpretation. This operation may be additionally complicated by the fact that many spatial technologies acquire data at different spatial resolutions, ranging from sub-cellular for microscopy-based technologies [61, 89] to 55 μm for the Visium ST platform [3] and 1-100 μm for commercial MSI platforms [206]. Therefore, a granularity matching step is required to ensure that the spatial positioning of data points from the different modalities is reasonably equivalent. Finally, in spatial-omics studies, sensitivity, especially for detecting low-abundance molecules and achieving quantitative accuracy, poses further challenges. To address these limitations, spatial omics approaches can be complemented with bulk or single-cell omics technologies. Bulk-omics methods offer high coverage and sensitivity, enabling a more comprehensive and quantitative analysis of molecular profiles [138], while single-cell omics provides insights into cellular heterogeneity within tissues, enhancing the resolution and sensitivity of spatial data [262, 23].

Recent studies have used a variety of different omics integration strategies, including co-registration and/or the use of a weighted nearest neighbour strategy to partially address the issues related to the combination of multiple spatial omics technologies [213, 238, 106, 269]. Here, we have applied a novel spatial multi-omics pipeline (SMOx), which involves multi-point co-registration and Gaussian weighted granularity matching steps for more accurate data point matching, resulting in improved integration of various spatial omics modalities. Using this pipeline, we combined Visium ST (further referred to as ST) with MALDI-MSI coupled with post-ionisation (MALDI-2), which has increased sensitivity for detecting lipids compared to conventional MALDI [233, 86], and single-nucleus RNA sequencing (snRNA-seq) to perform cell-type deconvolution of ST data [239] and high resolution pathology annotations. To explore the potential of this pipeline, we applied it

to study human prostate cancer (PCa), which in view of its highly inter- and intra-tumoural heterogeneous nature [55] is an ideal case study for the SMOx integration pipeline. In fact, PCa often exhibits multiple morphological disease states within a single tissue section, ranging from prostatic intraepithelial neoplasia (PIN) to various histological malignant states (ISUP grades) and cribriform, each representing distinct disease subtypes with unique biological characteristics and clinical outcomes [261]. Moreover, PCa development involves molecular changes across multiple omics classes. Besides alterations in genes and transcripts, prominent changes also occur in lipids, offering unique potential for novel therapeutic interventions [51, 189]. Using PCa as a paradigm, we showcase the potential of our novel spatial multi-omics pipeline to uncover lipid-gene expression pairs that are linked to different cell populations and histopathological disease states not discernible by morphology alone and can aid in the annotation of histologically ambiguous regions. These findings illustrate the potential application of our data integration pipeline to diverse fields of tissue-based research, including molecular pathology and for the generation of novel research hypotheses.

4.3 Results

Establishment of a novel spatial multi-omics data integration pipeline (SMOx) and its application in prostate cancer

With the aim to more comprehensively map the molecular heterogeneity within tissues and to better understand the interplay between different classes of molecules in a tissue context, we established a spatial multi-omics pipeline (SMOx) and applied it to the field of PCa (Fig. 4.1A). After collecting tumour and matching non-malignant biopsies from men that underwent a radical prostatectomy (n=8, Supplementary Table B.1), 10 μm thick frozen tissue sections were cut. Neighbouring tissue sections were subjected to ST analysis using the Visium Spatial Gene Expression platform (n=16) and MALDI-2-MSI based lipidomics (n=16). From the ST analysis, we obtained curated expression data of 18,950 genes across 42,475 individual capture spots on the ST array slides (summarised across all samples). Only spots under the tissue sections were considered for downstream computational analysis. From the MSI data, peak picking identified 12,510 peaks across the samples. The average overall spectra of cancer and non-cancer samples are presented in Supplementary Fig. B.2. High-resolution digital H&E images of the MSI sections were annotated by two independent expert uro-pathologists. Additionally, we collected three 100 μm thick cryo-sections and isolated nuclei for snRNA-seq (n=13) from the

remaining tissue directly adjacent to the cryo-sections used for ST and MSI analysis, to enable the identification and mapping of distinct cell populations across the tissue by applying the 'cell2location' deconvolution algorithm [136].

To integrate the data from the ST and MSI spatial-omics modalities, which are generated on separate (neighbouring) tissue sections and are acquired at different spatial resolutions, the SMOx pipeline was designed to allow co-registration and granularity matching (Fig. 4.1B-C). As both ST and MSI routinely include H&E staining of tissue sections in their workflows [220, 272, 133], we used the H&E-stained images to align the ST and MSI data. To facilitate cross-modality comparison and uncover inherent molecular patterns, the data from the different modalities were independently subjected to dimensionality reduction using uniform manifold approximation and projection (UMAP) [181]. The UMAP representation (i.e., mapped 3 dimensional embeddings) of the MSI data [229] was used to guide the co-registration with the respective post-MSI H&E image. Subsequently, the MSI H&E image was co-registered to the H&E image of the ST Visium slide. Alignment of ST data with the respective H&E histological images was done automatically using the Space Ranger pipeline (10X Genomics), eliminating the need for manual co-registration [4]. This strategy of using H&E-stained images to align ST and MSI data reduces the complex multimodal registration to a mono-modal one between their respective H&E-stained images, to which each modality has previously been co-registered, and removes the selection effect of registering to a template image. Subsequently, a shared spatial coordinate system was created by linking the MSI data with the ST data. This approach produces a readout that can be considered a "virtual section", between the sections used for ST and MSI analysis (Fig. 4.1B). Following co-registering of MSI and ST data, we addressed the disparity in spatial scales of data acquisition points through granularity matching. In this study, MSI was conducted at a spatial resolution (pixel size) of 30 μm, whereas the Visium assay has spot diameters of 55 μm with a 100 μm centre-to-centre distance between spots [3]. Consequently, each Visium spot corresponds to multiple MSI pixels with uneven amounts of coverage (Fig. 4.1C). To match the granularity of the spatial readouts, we employed Gaussian weighting to the MSI pixels close to each ST spot after registration [257]. This procedure resulted in the generation of a comprehensive aggregated MSI feature spectrum linked to each ST spot, ensuring synchronised spatial readouts between MSI and ST data (Fig. 4.1C). The resulting co-registered and matched MSI-ST data structure can be interpreted as a data frame consisting of matched readouts for each Visium spot. Within this data frame, each row represents a single matched spot, while the columns encompass distinct dataset information at that specific spot, such as gene expression, lipid content, pathology annotations, deconvoluted cell types, etc. Thus, this integrated spatial multi-omics pipeline could link transcript-lipid pairs with

tissue features.

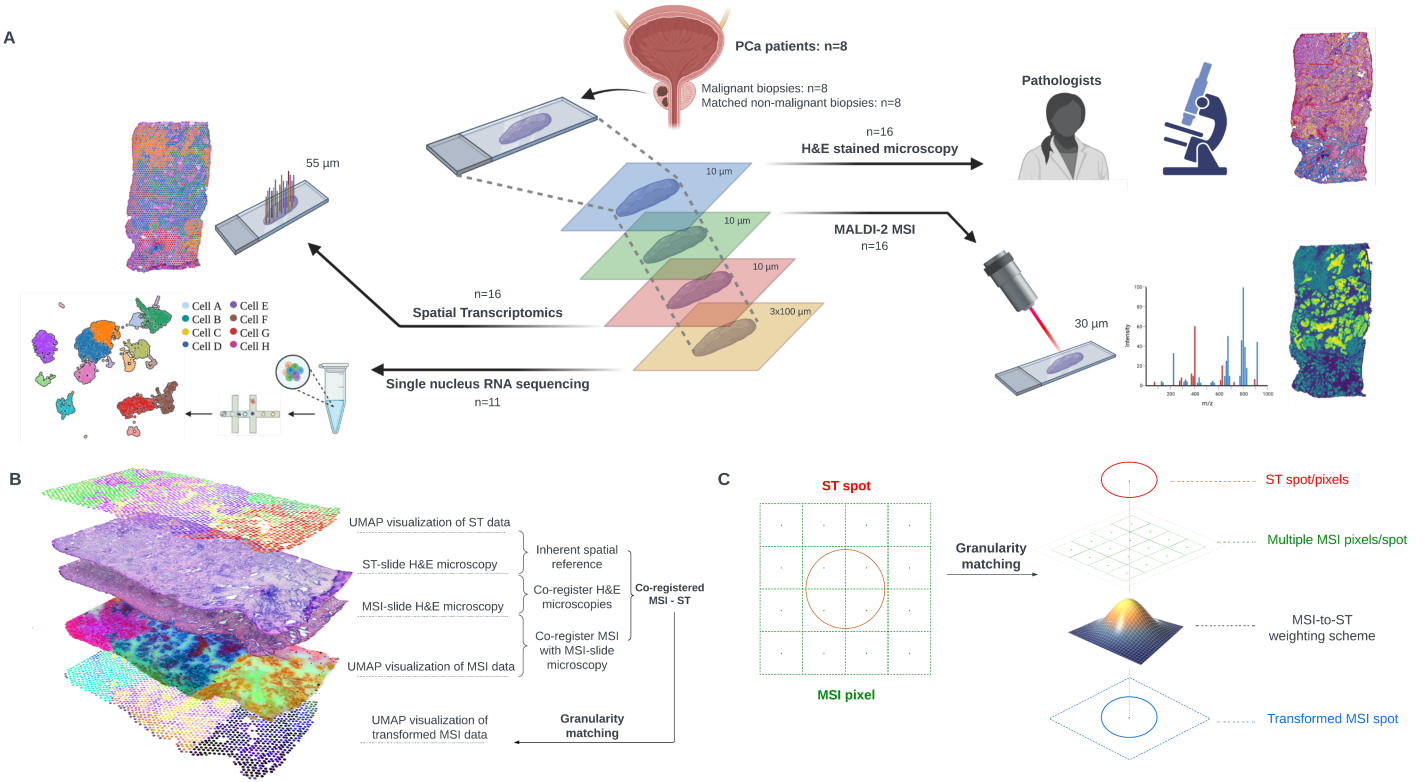


Figure 4.1: Study design and overview of the spatial multi-omics integration pipeline (SMOx). A. Schematic representation of the various datasets generated on prostate cancer samples and the applied integrated spatial multi-omics workflow. B. Strategy for co-registration and integration of different spatial multi-omics data layers generated on consecutive tissue sections. To link ST data with MSI data, UMAP visualizations of different modalities were co-registered in two steps based on the haematoxylin and eosin (H&E) stainings of the respective slides and MSI data, and a UMAP visualisation of transformed MSI data was generated after granularity matching. C. Overview of the granularity matching step to bridge the scale gap between MSI and ST data. Spatial readouts were aligned, and Gaussian weighting was applied to generate a transformed MSI spot.

Application of the SMOx pipeline to identify correlated transcript-lipid pairs linked with histological annotations

Using our spatial multi-omics integration strategy, we explored the spatial correlation of transcripts and lipids in a PCa tissue section of a single patient (patient 929) as a proof of concept. The level of detail in the annotations created by the histopathologist on the digital H&E image, along with the transformation of annotations in the shared coordinate system, is depicted in Fig. 4.2A. In this tissue section, five histologically distinct regions could be discerned: benign epithelium, stroma, PIN, tumour ISUP 3 and immune cells. When UMAP embeddings are projected onto tissue locations, similar patterns emerge among the distinct spatial modalities, indicating that despite the diverse nature of spatial data readouts (genes, lipids, etc.) (Fig. 4.2B), they all contain relevant molecular information reflecting the distinct histopathological features in the sample.

To link the different levels of molecular information, we conducted correlation analysis assessing the relationship between individual transcripts and MSI features (most of which are lipids), along with their spatial localization in the tissue (Fig. 4.2C-D, Supplementary Fig. B.3). Among the positively correlated lipid-transcript pairs (correlation coefficient over 0.65), we observed PC 34:2 [M+H]+ and TMEFF2 (a transmembrane proteoglycan with multiple roles in cancer [174]) as one of the most highly correlated pairs in regions annotated as tumour ISUP3. PIN included PE 40:6 [M+H]+ and AMACR (an established PCa marker [217]) as a highly correlated pair. LPE 16:0 [M+H]+ and NEFH (a neurofilament-encoding gene maintaining cell integrity) were identified in the benign epithelium region. In contrast, a notable spatial negative correlation (-0.631) was observed between the lipid SM 34:1;O2 [M+H]+ and the AGR2 gene (a disulphide isomerase gene family implicated in tumour metastasis [49]), with their expression detected in tumour ISUP 3 and PIN regions, respectively. Interestingly, several of the identified transcripts listed in Fig. 4.2C encode enzymes that are directly involved in lipid metabolism, including DEGS1 (a sphingolipid desaturase [270]), HPGD (a 15-hydroxyprostaglandin dehydrogenase associated with PCa risk [208]), AMACR (an enzyme involved in peroxisomal beta-oxidation of branched-chain fatty acids and one of the most widely used histological markers upregulated in PCa [217]), and ACSM1 (a medium chain Acyl-CoA Synthetase that is overexpressed in PCa [108, 228]). Importantly, these genes are up-regulated in areas annotated by pathologists as PIN and tumour ISUP3, in comparison to the regions annotated as benign epithelium. Many other transcripts are not directly implicated in lipid metabolism but are known to have a cell type-specific expression pattern and are differentially expressed upon cancer development. These include transcripts of genes TMEFF2, KLK3 (encoding the prostate-specific antigen biomarker),

PCA3, AGR2, and NEFH. These observations suggest that transcript-lipid correlations are associated with disease state and specific cell populations, and may reveal functional interactions.

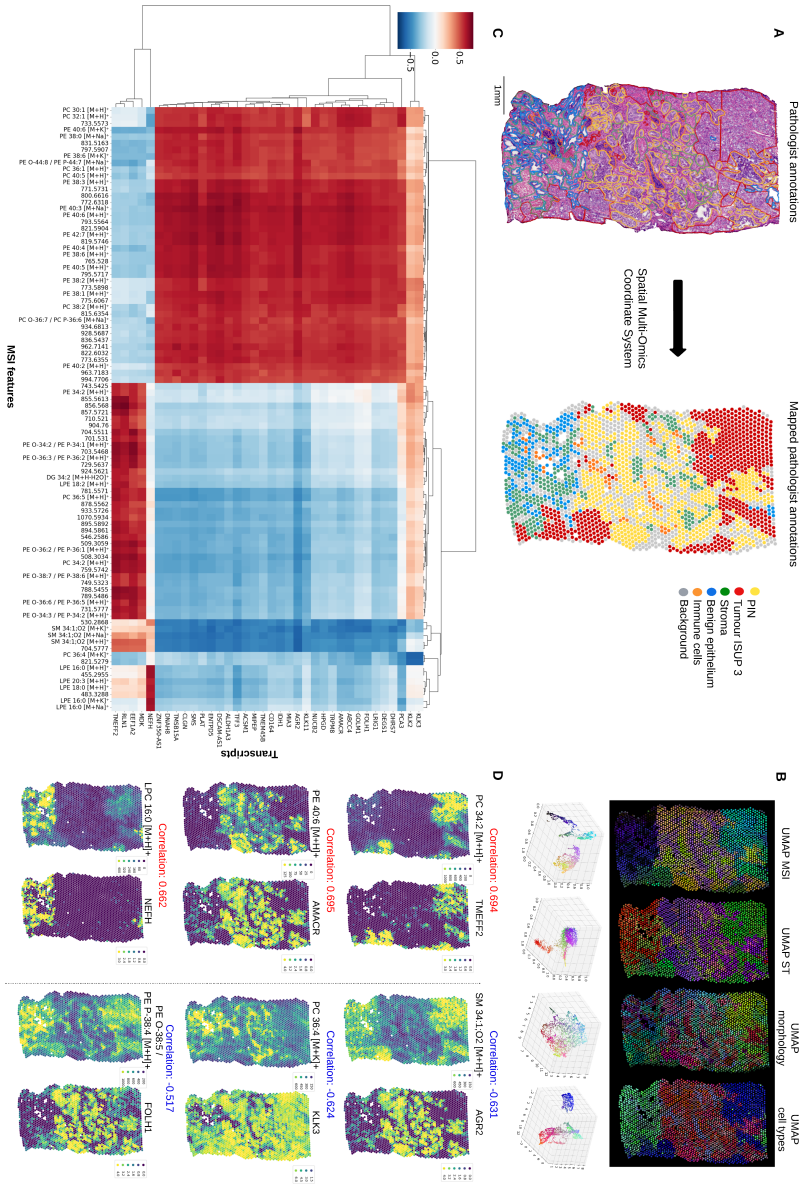


Figure 4.2: Integration of spatially resolved multi-omics reveals spatially correlated lipid-gene pairs in a prostate cancer sample. A. Highly detailed pathologist annotations of a representative prostate cancer sample (929_cancer) on the post-MSI H&E stained tissue section and their spatial distribution projected on the shared coordinate system. B. UMAP representations of individual modalities including MS imaging data, ST data, morphological features extracted from pathology annotations, and cell types derived from snRNA-seq data after spot deconvolution. The top row represents a hyperspectral visualisation of the UMAP data on the shared coordinate system using the RGB colour scheme. The bottom row shows the scatter plots of the resulting 3-dimensional UMAP embeddings for each modality. C. Hierarchically clustered heatmap of the correlation matrix between MSI features and ST data. D. Spatial visualisation of representative positively and negatively correlated lipid-transcript pairs in the 929_cancer sample.

Identification of molecular expression patterns associated with histological annotations

To further explore the alignment of MSI data and ST expression patterns with pathology annotations, we applied non-negative matrix factorisation (NMF) to extract correlated features from the individual modalities of the integrated spatial multi-omics dataset of sample 929_cancer. The resulting correlation heatmap between MSI NMF components and pathology annotations demonstrated significant positive correlations for each pathological annotation (PIN, benign epithelium, tumour IUSP3 and stroma) except for immune cells (likely due to their greater cell type variability) (Fig. 4.3A top panel). The spatial distribution of these NMF components, along with pathology annotations within the common coordinate system are illustrated in Fig. 4.3B, accentuating the congruence between lipidome and histology. Similarly, significant positive correlations were observed between ST NMF components and the different annotated regions (including immune cells) (Fig. 4.3A bottom panel and Fig. 4.3B). For MSI NMF component 9, which highly correlates with PIN, the pseudo-spectrum of the biomolecular ion peaks that are involved in the expression of the MSI NMF component's spatial patterns is shown, along with the gene expression profile of the PIN-associated ST NMF component 13 (Fig 4.3A). Integration of both spatial-omics modalities within the shared coordinate system and subsequent hierarchical clustering revealed strong correlations between specific NMF components of both modalities (Fig. 4.3C). Differentially expressed lipids and genes that contribute significantly to the molecular characteristics of the different annotated areas are listed in Fig. 4.3D. The high level of correlation provides further credence to a spatial relationship between readouts of the ST and MSI data, and overall agreement with the pathology annotations.

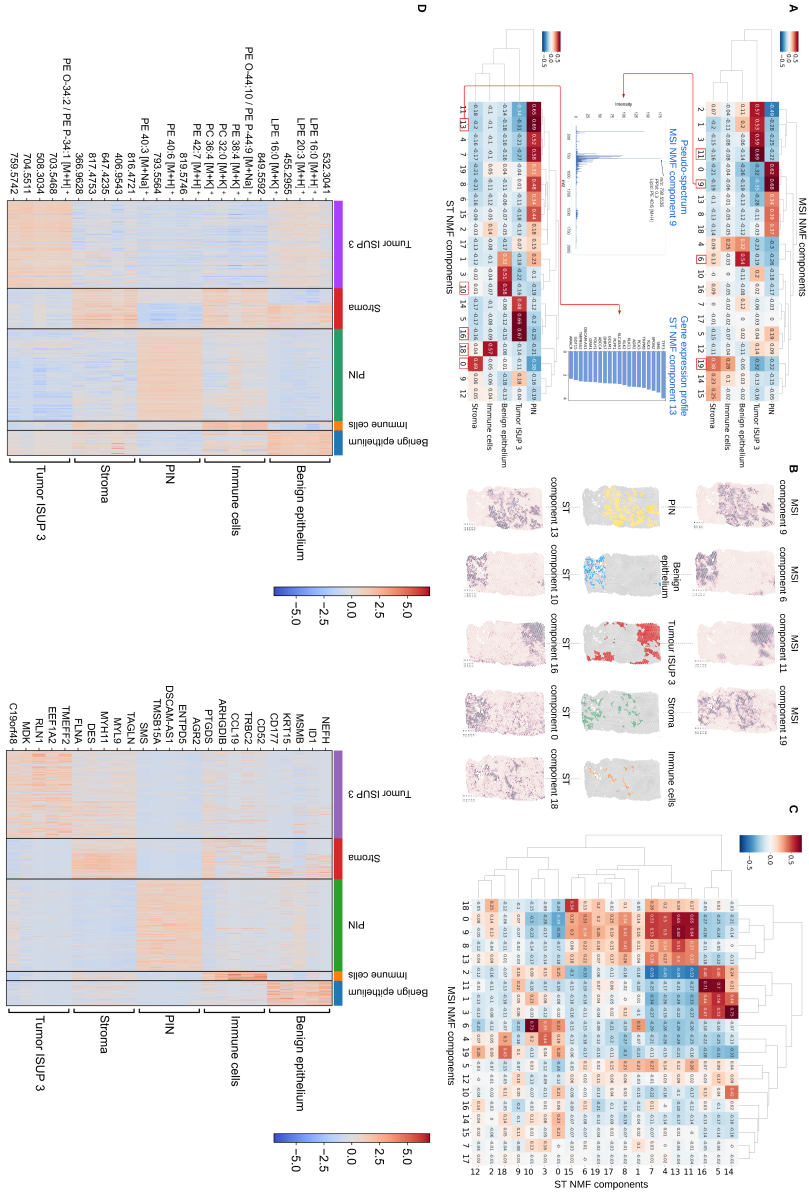


Figure 4.3: Correlation between histological and molecular profiling. A. Heatmap showing correlations between MSI NMF components and the five histologically distinct regions from sample 929_cancer. A representative pseudospectrum for MSI NMF component 9 and the most enriched genes of the correlating ST component 13 are shown. B. Spatial visualisation of representative MSI NMF components (top panels) correlating with ST NMF components (bottom panels) for each histological feature (middle panel). C. Hierarchically clustered heatmap of the overall correlations between MSI NMF and ST NMF components. Numerical values correspond to the respective correlation coefficient of each MSI NMF – ST NMF pair. D. Heatmaps illustrating MSI features (mostly lipids) (left panel) and genes (right panel) exhibiting the most significant differential expression within each of the five histological regions in the 929_cancer sample.

Linking lipids and transcripts to specific cell types in the spatial coordinate system

Due to the granularity of the matched MSI-ST data set with a spatial resolution defined by the spatial-omics analysis techniques used (ST: 55 μm resolution; MSI 30 μm resolution), each spot in the spatial coordinate system corresponds to multiple cells, hampering the assignment of lipids and transcripts to specific cell types. To address this issue, we applied snRNA-seq-based spot deconvolution using the 'cell2location' deconvolution algorithm [136], based on snRNA-seq data from a set of 100 μm thick serial tissue sections neighbouring those utilised for ST. The resulting UMAP clusters were annotated based on curated marker genes [117, 234, 120], leading to the identification of 14 distinct cell types (Fig. 4.4A). Of note, the TC1 (Tumour Cell 1) cell type was mainly enriched in both the tumour ISUP 3 and PIN region, whereas TC2 (Tumour Cell 2) was linked mainly to the PIN region only (Fig. 4.4B-C). Interestingly, cells with a club cell-specific gene signature, which have recently been associated with PCa carcinogenesis [234], localised to sub-regions within the PIN-annotated areas. Correlation analysis between MSI and ST data from spots enriched in specific cell types revealed unique lipid profiles associated with a cell type along with marker genes and enzymes involved in lipid metabolism, as illustrated or TC1 and TC2 cells, club cells and macrophages (Fig. 4.4D). This analysis revealed that ISUP 3-associated TC1 cells and PIN-associated TC2 cells correlated with long chain polyunsaturated phospholipids (mainly PE), with an additional enrichment for hexosylceramides (HexCer) in the TC2 cell component. Cholesteryl esters (CE), mainly identified as [2M+H]+ ions [46] and known to accumulate in lipid droplets, were associated with club cells. Macrophage marker genes, like HLA-D genes [169], LYZ [13] and CXCL5 [215], correlated best with ether lipids, CE and long polyunsaturated phosphatidylcholines (PC). Complementary gene signature analysis of lipid metabolism pathways revealed an overall high expression of genes linked to lipid metabolism pathways in luminal and TC2 cells, along with cell type-specific pathway activations (Fig. 4.4E). Examples include the upregulation of fatty acid biosynthesis and elongation in cancer cells (TC1 and/or TC2), steroid biosynthesis in luminal cells, and arachidonic acid metabolism in macrophages and mast cells, in line with expectations and data available in the literature [51, 52, 282]. Projection of these pathways in the UMAP representations confirmed these links and provided an additional layer of granularity, for instance revealing a specific subset of cells within the TC1 and TC2 clusters displaying elevated expression of genes involved in linoleic acid metabolism (Fig. 4.4F). These findings illustrate the benefit of linking lipids and transcripts to specific cell types in the spatial coordinate system.

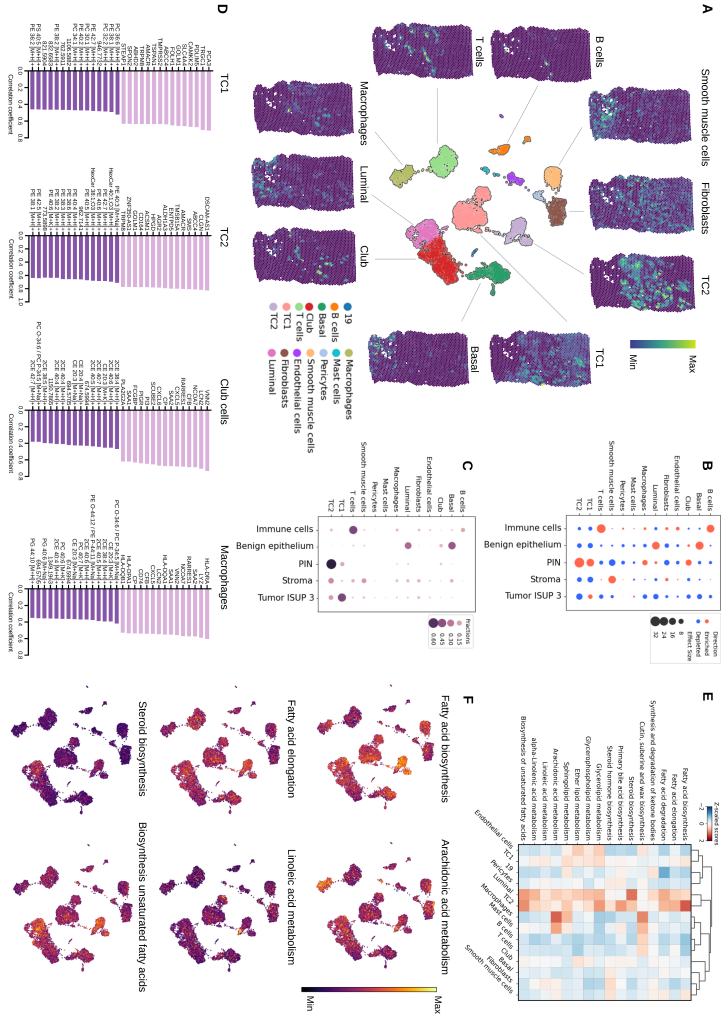


Figure 4.4: Spot deconvolution reveals cell type specific transcriptional and lipidomic profiles. A. UMAP representation of annotated snRNA-seq clusters and spatial localisation of the main cell types in the spatial transcriptomics data after spot deconvolution using cell2location42. TC1: tumour cluster 1, TC2: tumour cluster 2. B. Dot plot displaying the enrichment (orange) and depletion (blue) of cell types within the regions demarcated by the pathologist. C. Dot plot illustrating the relative abundance of a specific cell type within each pathologically annotated region. D. MSI features (pink) and transcripts (purple) with the highest positive correlations with TC1, TC2, club cells and macrophages. E. KEGG enrichment of differentially expressed genes involved in lipid metabolism per annotated cell type based on the snRNA-seq data. Colours indicate mean up- (red) or down- (blue) regulation of the pathway, while colour intensity indicates significance level. F. UMAP representation of the enrichment of different lipid metabolism pathways per cell type.

Integrated data analysis reveals molecular heterogeneity in neoplastic disease areas

Whereas the spatial distributions of MSI NMF components and ST NMF components in sample 929_cancer overall matched well with the histological annotations, the region annotated as PIN by the pathologists, was captured by two NMF components in both modalities – MSI NMF component 9 and 18 and ST NMF components 13 and 15 (Fig. 4.5A-B). Additionally, snRNA-seq-based spot deconvolution revealed that one of the components of each modality (MSI NMF 9 and ST NMF 13) corresponded well with the spatial distribution of TC2 cells (Fig. 4.5C). The other component (MSI NMF component 18 and ST NMF component 15) aligned with a specific subregion of PIN corresponding to the spatial distribution of epithelial club cells and macrophages. Examination of the H&E stainings of these latter regions confirmed the presence of a neoplastic gland-like structure. These observations underscore the potential of the integrated data analysis approach to reveal molecular heterogeneity in neoplastic disease areas that might not be discernible through morphological examination alone.

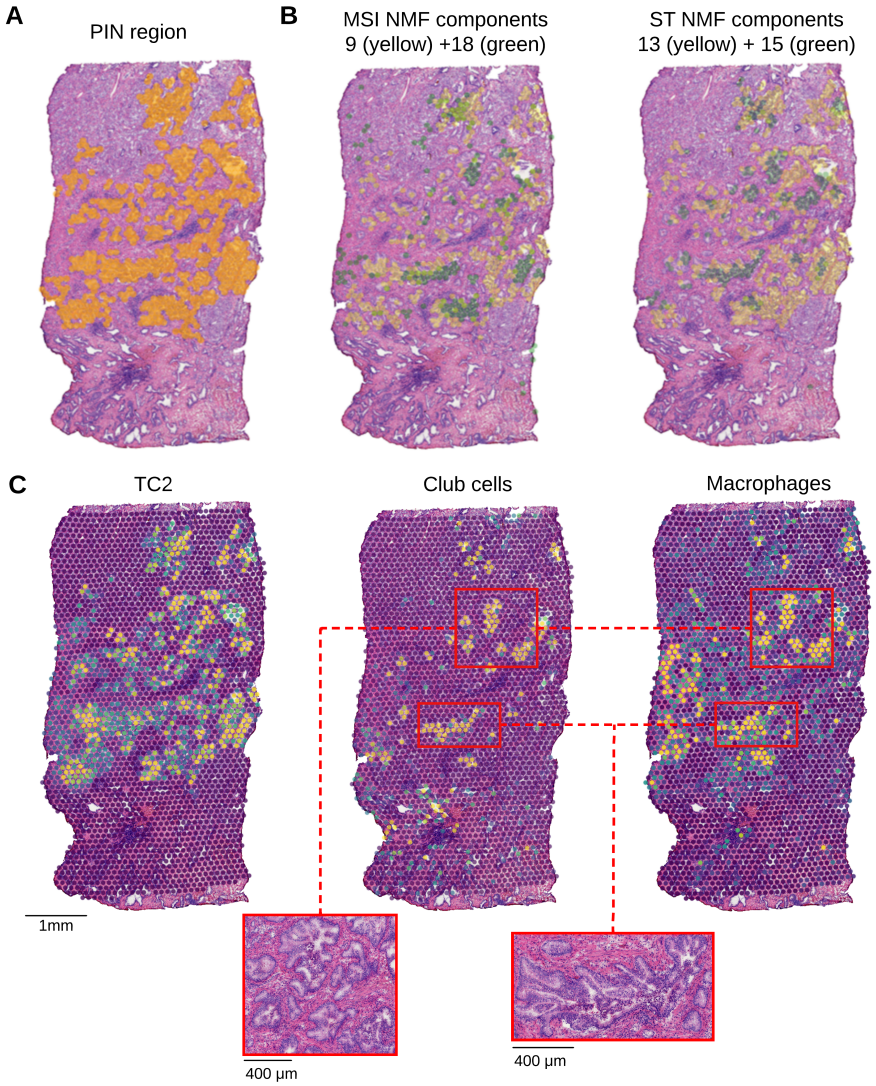


Figure 4.5: Integration of spatially resolved omics unveils molecular heterogeneity in PIN. A. Pathologist annotation of PIN in sample 929_cancer (highlighted in yellow). B MSI and ST NMF components correlating with the annotated PIN region in panel A. C. Localisation of TC2, club cells and macrophages based on snRNA-seq-deconvoluted ST data. Inserts show H&E stainings of the indicated areas.

Integrated spatial multi-omics aids in the histological annotation of morphologically ambiguous tissue areas

Integrated SMOx analysis of the entire set of 8 PCa samples overall revealed a strong concordance with the pathologist's annotations. Nevertheless, in several instances our integrated data-analysis pipeline (based on ST and MSI NMF components) suggested alternative histopathological assignments for areas in the tissues. This was for instance the case in sample 941_normal in which a region initially labelled as PIN exhibited closer molecular concordance with tumour ISUP4 (Fig. 4.6A-B). This observation was further supported by the snRNA-seq data, which indicated the presence of a population of TC1 cells. Furthermore, we observed a match between the top correlated MSI features, predominantly lipids, and genes of the ISUP 4-associated component with the top correlated MSI features and genes of the TC1 cell population (Fig. 4.6C). Re-examination of this region by the pathologist confirmed the presence of a micro-area of infiltrating prostate adenocarcinoma cells (lower inset box, Fig. 4.6A-B). Another ambiguous tissue area was identified in sample 931_cancer, which, based on the H&E staining of the frozen tissue section was annotated by the pathologist as "other" (Fig. 4.6D). Based on the ST and MSI NMF components (the latter one showing the strongest spatial correlations), our integrated approach identified this region as a subcluster within the ISUP5 tumour region. snRNA-seq data deconvolution further revealed the presence of two distinct tumour clusters, TC1 highly correlated with ISUP5 and TC2 correlated with the ambiguous region (Fig. 4.6E-F). Histopathological re-evaluation of this region assigned it as cribriform. These results highlight the potential of spatial multi-omics to aid in histopathological annotation of ambiguous tissue areas.

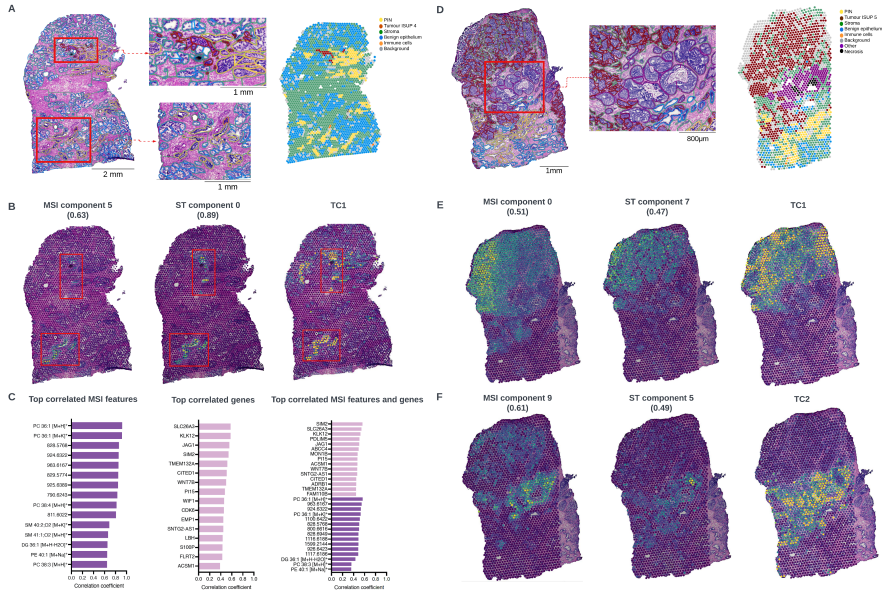


Figure 4.6: Integrated spatial multi-omics supporting histological annotation of ambiguous tissue areas. A. Highly detailed pathologist annotations of a non-malignant prostate sample (941_normal) on the post-MSI H&E-stained tissue section and their spatial distribution projected on the shared coordinate system. Inserts show H&E stainings of tumour ISUP 4 and a morphologically ambiguous region. B. Distributions of MSI NMF component 5, ST NMF component 0 and TC1 cell population based on snRNA-seq-deconvoluted ST data. Numerical values (0.63 and 0.89) correspond to the respective correlation coefficient of the MSI and ST NMF with the region annotated by the pathologists as tumour ISUP 4. C. Top correlated MSI features and genes of MSI NMF component 5 and ST NMF component 0, respectively. D. Highly detailed pathologist annotations of a high-grade prostate cancer sample (931_cancer) on the post-MSI H&E-stained tissue section and their spatial distribution projected on the shared coordinate system. Inserts show H&E stainings of an area annotated by the pathologist as "other". E. Distributions of MSI NMF component 0, ST NMF component 7 and TC1 cell population based on snRNA-seq-deconvoluted ST data. Numerical values (0.51 and 0.47) correspond to the respective correlation coefficient of the MSI and ST NMF with the region annotated by the pathologists as tumour ISUP 5. F. Distributions of MSI NMF component 9, ST NMF component 5 and TC2 cell population based on snRNA-seq-deconvoluted ST data. Numerical values (0.61 and 0.49) correspond to the respective correlation coefficient of the MSI and ST NMF with the region annotated by the pathologists as "other".

Implementation of the SMOx workflow across the entire prostate cancer sample cohort reveals tumour-specific correlations between transcripts and lipids [1]

Having demonstrated the ability of the integrated SMOx pipeline to extract relevant information from multi-layered molecular expression patterns within a tissue sample, we expanded our analysis to the entire PCa cohort with additional Pearson correlation analysis. Examination of pairs of transcripts and MSI features with correlation coefficients exceeding 0.4 across all 16 biopsies (comprising 8 cancer samples and 8 matching non-malignant samples), allowed us to identify pairs of transcripts and MSI features with a spatial distribution that either correlated or anti-correlated throughout the cohort (Fig. 4.7A). These included well-established links i.e., between FASN expression and PC34:1 [51], confirming the validity of this approach, as well as several novel links illustrating the hypothesis-generating potential of the SMOx pipeline. Next, Cohen's d values were calculated between tissue types i.e., tumour versus non-tumour tissue (as annotated by pathologists on the H&E microscopy), aiming to detect transcript-MSI feature correlations that significantly changed between these tissue types. One key example is the MSMB - PE 40:2 [M+H]+ pair displaying an associated Cohen's d value of 1.54, indicating a substantial effect size [222], which in this case implies a considerable difference in spatial co-expression between tumour and non-tumour samples. In fact, as illustrated by a line plot of correlation coefficients for the pair across the 16 biopsies (Fig. 4.7B) and the actual spatial distributions of the transcript and the lipid in the shared coordinate system of both tumour and non-tumour samples (Fig. 4.7C), these markers exhibited a clear spatial co-expression in non-tumour samples, which was absent in the tumour samples (quantified with average correlation coefficients of approximately 0.5 and 0, respectively). Specifically, in the non-tumour samples a striking alignment was observed between the gene MSMB and lipid PE 40:2 [M+H]+, whereas in tumour samples, while MSMB continued to exhibit a robust spatial correlation, the lipid PE 40:2 [M+H]+ lost its correlation with the benign epithelial cells (highlighted in green). This spatial correlation between PE 40:2 with tumour regions was observed in the majority of tumour samples. While the precise function of this lipid remains elusive, MSMB has been previously shown to exhibit a higher expression in normal prostate tissue than in cancerous tissue [70, 40]. These observations highlight the hypothesis-generating potential of our SMOx approach.

[1] This part of the work has been submitted to the Belgian Patent Office: Zhang, W. , Claesen, M., Verbeeck, N., 2023. Method for vertical integration and analysis of spatial multi-omics data. No. BE2023/5080. Feb 2023.

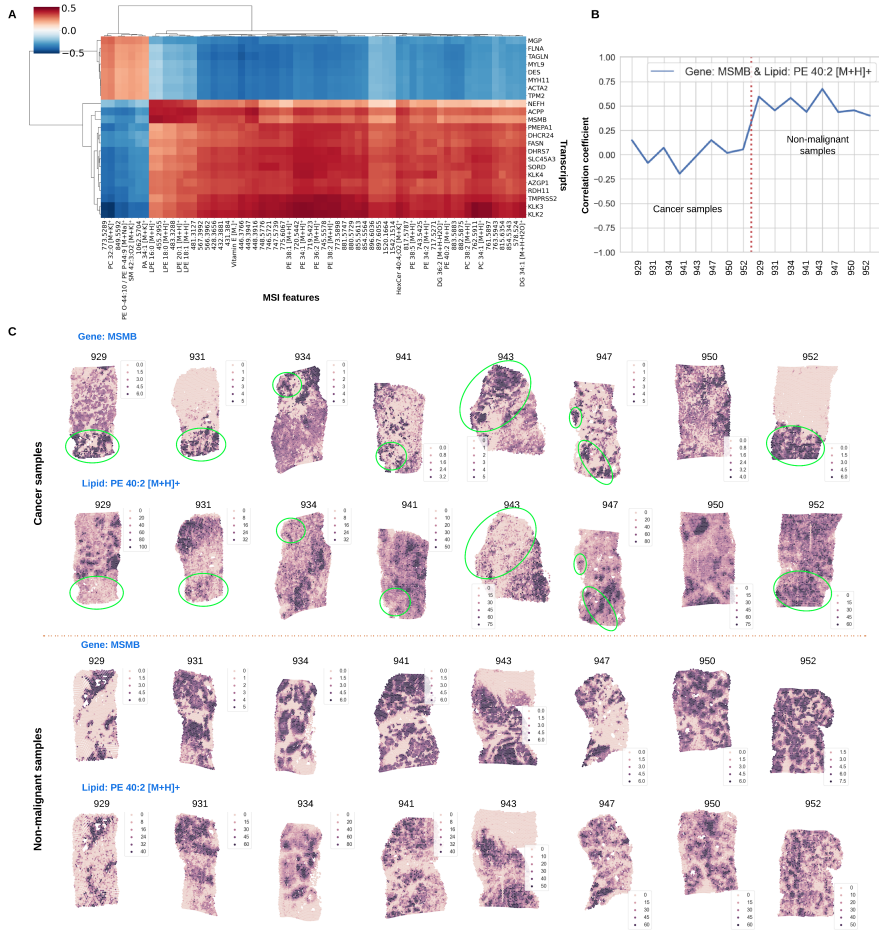


Figure 4.7: Implementation of the spatial multi-omic workflow across the entire prostate cancer sample cohort. A. Heatmap illustrating correlations between genes and MSI features, mostly lipids, where correlation coefficients greater than 0.4 are observed across the entirety of the sample cohort consisting of 16 biopsies (8 cancer samples and 8 matching non-malignant samples). B. Line plot depicting the correlation coefficients for the gene/lipid pair MSMB/PE 40:2 [M+H]+ across the entirety of the sample cohort. C. Spatial distribution of the gene/lipid pair MSMB/PE 40:2 [M+H]+ observed across the eight cancer biopsies (upper section) and the corresponding non-malignant samples (lower section). Green polygons indicate regions predominantly composed of benign epithelial cells within the cancer samples.

4.4 Discussion

Spatial multi-omics is progressively establishing itself as a powerful tool to better understand molecular events and its interactions in a native tissue context. One of the key challenges related to this approach is the integration of multiple layers of molecular information that are generally acquired using a range of technologies to optimally capture the molecular information in each layer. Although the ability to conduct diverse spatial approaches, including ST and MSI analyses on a single slide has recently been reported [269], hitherto, most spatial multi-omics studies are performed on neighbouring tissue sections and generate spatial molecular maps with different spatial resolutions. These factors are further complicated by different tissue deformations due to technology-specific tissue processing, different inter-spot distances and offset of the spots. Our study, which integrates ST and MSI analysis aimed to overcome these hurdles by establishing a data-processing pipeline involving spatial co-registration and granularity matching. Co-registration refers to the mapping of pixels between modalities and is a crucial step in multimodal imaging analysis. Registration of imaging datasets is typically contingent on shared anatomical structures but can be especially challenging when modalities have distinct characteristics, or if measurements are performed at vastly different spatial scales [252]. The former challenge is typically addressed by generating a number of representative images for each modality e.g., by manually selecting from the imaging datasets or by generating images that capture spatial trends in the data by dimensionality reduction methods. The latter, multi-scale challenge refers to the fact that a spatial data point, typically referred to as a pixel or voxel, in one modality can cover a different surface area from that in another modality. These areas can range in the order of magnitude from square nanometres to hundreds of square micrometres. This problem is receiving increasing attention in the spatial biology field, due to the increasing number and complexity of spatial omics applications [198]. A granularity matching step is required to make data points reasonably equivalent between different modalities, typically via an aggregation strategy such as interpolation [213], averaging, or more advanced methods e.g., Gaussian smoothing [257, 128]. In a recent spatial multi-omics study, combining MSI at 15 μm pixel size with multispectral immunofluorescence microscopy [106], a two-step co-registration was used reporting subcellular accuracy and ROIs from the microscopy projected onto the MSI data, however no granularity matching step was reported. Another study combining MSI and ST using the Visium platform used a spatial resolution of 100 μm, corresponding to the largest modality used, when integrating their data, but do not report whether aggregation was used to account for capture areas not being distributed equally across three modalities [238]. Another study used a weighted nearest neighbour

strategy that mechanistically paired Visium spots in the centre of one 50 μm MSI pixel [213]. In contrast to these studies, we used a Gaussian approach that accounts for sample coverage per data acquisition area (Visium spot/MSI pixel). Together with our two-step co-registration tool, a full data analysis pipeline was established that has the ability to deal with the mentioned hurdles in spatial multi-omics data integration.

When applied to a set of clinical PCa samples, downstream integrated analysis revealed a strong spatial concordance amongst the individual modalities as seen in UMAP and NMF analysis. Importantly, agreement between the modalities in this combined approach allows more confidence in defining tissue profiles in spatial atlas approaches, assignment of molecularly heterogeneous regions not discernible by morphology, and potential uses for identifying missed cancerous regions e.g., in a diagnostic context. Using our approach to combine ST, MSI and snRNA-seq analysis on PCa samples, we were able to find lipid and transcript profiles that correlated to specific histological states and to specific cell populations. Based on these correlations between lipids and transcript profiles, we were able to assign unique lipid profiles to specific cell types and disease states and to generate unprecedented insight into the heterogeneity of PCa pathology in the context of a native tissue.

One striking example of the power of our approach was the discovery of a molecularly distinct profile in a PIN region based on both ST and MSI profiling, providing more confidence in the pathological characterisation. Integrated snRNA-seq-based deconvolution revealed the presence of club cells, which recently have been linked to the pathogenesis of PCa [234]. Although the clinical significance of this finding will have to confirmed, our observation of the presence of club cells in a specific pre-malignant histological structure, may point towards that region's future potential to develop into adenocarcinoma and warrants further investigation.

When applied to the entire cohort of PCa and matched benign samples, we found that spatial transcriptomics and spatial lipidomics data largely followed histomorphological patterns. In addition, combining the two analyses allowed us to identify genetic and lipid expression profiles corresponding to different cell populations and tissue states. Our integrated spatial multi-omics approach was able to identify transcript – MSI feature pairs that correlate or anti-correlate throughout the cohort. Besides well-known associations, including the correlation between FASN and PC34:1 [51], which provides validation for our approach, several new correlations were found that warrant future investigation. One striking example was the loss of co-expression of specific lipid-gene pairs such as PE 40:2 [M+H]+ and MSMB in cancer samples. While the role of MSMB as a secreted protein with decreased expression in cancerous tissue is well established [70], the implications of a change in

the abundance of correlating lipid currently remains unknown. Based on transcript – MSI feature pairs and snRNA-seq-based deconvolution of the entire cohort of samples our approach also discovered the presence of tumour cells in a histological region previously assigned as PIN and in a normal tissue section, demonstrating the ability of this approach to aid in pathological annotations. Taken together, these observations illustrate the potential of integrated spatial multi-omics approaches to reveal tissue heterogeneity and to generate new hypotheses as potential starting points for further investigation. Drawing biological conclusions from this concept study should be taken with care in view of the small number of samples that were measured and were selected based on morphological heterogeneity rather than for testing a specific hypothesis. One issue to be further resolved is the presence of batch effects [29], restricting the comparison of multiple samples to correlation analysis on each sample individually rather than collectively. While various strategies have been proposed to alleviate this issue in both ST and MSI [184], further method development is required to determine how to mitigate its effects.

As spatial technologies are rapidly evolving and are continuously gaining sensitivity, molecular coverage and spatial resolution, flexible approaches to integrate different layers of spatial and bulk molecular information will become indispensable. Integration approaches, such as the one presented here, will be crucial to better understand the complex interactions between different classes of molecules in a spatial cellular context and to apply this knowledge in diverse fields such as biomarker or target discovery, molecular pathology, and automated tissue annotation.

4.5 Methods

4.5.1 Data acquisition

Sample collection

Matched prostate tumour and normal samples were obtained from treatment naive high-risk PCa patients (advanced clinical stage and/or biopsy Gleason 8-10 and/or PSA $\geq$ 20 ng/mL) who underwent radical prostatectomy at University Hospitals Leuven, Belgium. This study was approved by the local Ethics Committee (approval number S54424) and followed the Code of Conduct for Responsible Use of Human Tissue for Research. All patients provided informed, signed consent. After surgery, 6 mm biopsy cores were extracted from the primary tumour site and a matching non-malignant biopsy was taken from the opposite side. The selection of biopsy cores was guided by a urologist and

magnetic resonance imaging reports. Immediately after surgery, these biopsy cores were embedded in 3% carboxymethyl cellulose (CMC) (Sigma-Aldrich, cat.no C4888), snap-frozen in isopentane pre-chilled in liquid nitrogen, then stored at -80°C until further processing.

For each prostate biopsy, 10 μm thick cryo-sections were obtained at -20°C using a Microm HM525 NX cryostat (Thermo Scientific), followed by H&E staining. Subsequently, two independent expert pathologists performed a thorough examination to verify the presence of tumours in the tumour samples. In this study, a cohort of 8 patients was selected based on detailed pathological information including classification, ISUP grade and morphological features present (see Supplementary Table B.1 for patient-related information). Tissue blocks from the selected patients were subsequently checked for RNA quality. RNA was isolated from ten 10 μm sections from the selected tissue blocks using Direct-zol RNA Miniprep (Zymo Research, cat.no. R2050), and RIN score was determined using a Bioanalyzer RNA 6000 Nano kit (Agilent, cat.no 5067-1511) in combination with the Agilent Bioanalyzer software. All samples had a RIN $>$ 7. 10 μm thick sections were collected on ITO-coated glass slides (Delta Technologies, Loveland, USA) to enable matrix-assisted laser desorption/ionisation (MALDI) MSI analysis, and adjacent tissue sections of 10 μm was mounted onto pre-equilibrated 10x Genomics Visium Gene Expression slides (10x Genomics, Pleasanton, CA, USA) for ST analysis. Finally, three to five tissue sections with a thickness of 100 μm were collected in a pre-cooled centrifuge tube for subsequent single-nuclei RNA sequencing (snRNA-seq) analysis.

Spatial transcriptomics sample preparation and data acquisition

Tissue preparation for Visium ST analysis was performed according to the Tissue Preparation Guide (CG000240 rev C, 10x Genomics), except for embedding in 3% CMC which was conducted to stabilise the tissue for both ST and MSI analysis. Methanol fixation, H&E staining and imaging were performed according to the recommended protocol (CG000160 rev B, 10x Genomics), and subsequent cDNA and library preparation was performed according to the Visium Spatial Gene Expression Reagent Kit User guide (CG000239 rev D, 10x Genomics). The optimal tissue permeabilization time was determined upfront to be 20 minutes, following the Visium Spatial Gene Expression Reagent Kit - Tissue Optimization User guide (CG000238 rev A, 10x Genomics). Final libraries were sequenced to at least 50.000 reads/covered spot using a NextSeq2000 platform (Illumina, San Diego, CA, US).

Mass spectrometry imaging sample preparation and data acquisitions

Tissue sections were coated in 2,5-dihydroxybenzoic acid (DHB, AK Scientific, CA, USA) matrix using a sprayer robot (TM-Sprayer, HTX Technologies, LLC, Chapel Hill, NC, USA). DHB was dissolved in a 1:1 mixture of methanol:chloroform (v:v) at a concentration of 20 mg/mL. 12 layers of matrix solution were deposited using a flow rate of 0.12 mL/min, 10 psi of N2 pressure, nozzle temperature of 30 °C, and a 1200 mm/min velocity.

All data was acquired using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific GmbH, Bremen, Germany) coupled to an intermediate pressure MALDI source (Spectroglyph LLC, WA, USA) as described previously [34]. A frequency tripled Nd:YLF laser (Explorer One, Spectra Physics, Mountain View, CA) emitting at 349 nm operating at 300 Hz was used for MALDI. The laser was operated using a diode current of 1.80 A and pulse energy fine-tuned using an external attenuator (PowerXP, Altechna, Vilnius, Lithuania) positioned immediately in front of the laser to yield a post-attenuation pulse energy of 1.2 μJ. A frequency quadrupled laser Nd:YAG laser emitting at 266 nm and operating at 300 Hz (NanoDPSS, Litron Lasers, Warwickshire, England) was used for laser post-ionisation (MALDI-2). This setup has been described in detail previously [221]. MALDI-2 laser pulse energy was adjusted using the internal attenuator to be 600 μJ just prior to entering the ion source. Laser pulse energy measured using a calibrated energy sensor (QE12HR-H-MB-D0, Gentec-EO, Quebec, Canada). Both lasers were externally triggered using a pulse/delay generator (QC9200, Quantum Composers, Bozeman, Montana) such that the time delay between the MALDI and MALDI-2 laser pulse was 20 μs. The mass spectrometer was operated in positive ion mode using an ion injection time of 250 ms, automatic gain control turned off, m/z range 350-2000 and a nominal mass resolution of 120,000 @ m/z 400. The pixel size was 30 μm. All raw files were first recalibrated with Thermo Scientific Xcalibur RecalOffline using ion signals of [Vitamin E]+ at m/z 430.38053, [PC(34:1)+K]+ at m/z 798.54096 and a background PDMS peak at m/z 371.10124. Data was then converted to imzML format using Image Insight software (Spectroglyph LLC, Kennewick, WA, USA).

Tissue Staining and Histopathology Annotations

Tissue staining using H&E was conducted on tissue sections following MSI analysis. The matrix was first removed from the slide by rinsing in methanol. For the H&E staining, the glass slides were immersed in 95% ethanol, 70% ethanol and H_2O for 2 minutes each and then immersed in haematoxylin for 3 minutes. Slides were then washed with running tap water for 3 minutes and

immersed in eosin for 30 seconds following another washing step with water for 3 minutes and immersion in 100% ethanol for 1 minute and then in xylene for 30 seconds. The tissues were air-dried and mounted with coverslips. Optical images were acquired using a ZEISS Axio Scan.Z1 Slide Scanner (Carl Zeiss AG, Jena, Germany) with 40x objective, pixel size 0.22 μm* 0.22 μm. The files were exported as .TIFF images.

The digitalized H&E images were annotated and verified by prostate cancer specialists from the Department of Pathology, University Hospitals Leuven and Weill Cornell Medical College, in a blinded fashion using the web-based annotation tool 'Annotation studio' (Aspect Analytics NV, Genk, Belgium). Pathologists annotated the whole slides in detail by placing 14 different labels 'Inflammatory infiltrate', 'Normal epithelium', 'Stroma', 'Necrosis', 'Atrophy', 'Cribriform', 'PIN', 'Nerve', 'Tumour ISUP 5', 'Tumour ISUP 4', 'Tumour ISUP 3', 'Tumour ISUP 2', 'Tumour ISUP 1', and 'Other'.

Single cell RNA sequencing sample preparation and data acquisition

Tissue sections (2-3x100 μm) adjacent to those used for spatial transcriptomics and lipidomics sections were collected in a microcentrifuge tube and subjected to single nuclei dissociation. For this, 0.3mL Tween with Salts and Tris (TST) buffer (96mM NaCl (Sigma-Aldrich, cat.no. 59222C), 10mM Tris-HCl pH 7.5 (ThermoFisher, cat.no. 15567027), 1mM CaCl2 (Merck, cat.no. 21115), 21mM MgCl2 (Sigma-Aldrich, cat.no M1028), 0.03% Tween-20 (Biorad, cat.no 1662404), 0.01% BSA (New England Biolabs, cat.no B9000S) and 0.2U/μl RNAsin (Promega, cat.no N2615)) were added to the sections and the tissue was cut into small pieces using scissors. The tissue was subsequently transferred to a KIMBLE Dounce tissue grinder (Sigma-Aldrich, cat.no D8938), together with 0.7mL TST that was used to wash the original tube. The tissue was homogenised by 10 strokes of the large clearance pestle and 10 strokes of the small clearance pestle. The homogenate was subsequently filtered using a 40μm cell strainer (pluriSelect, cat.no. 43-50040-51) and the tissue grinder and filter were washed with Salts and Tris (ST) buffer (96mM NaCl (Sigma- Aldrich, cat.no 59222C), 10mM Tris-HCl pH 7.5 (ThermoFisher, cat.no 15567027), 1mM CaCl2 (Merck, cat.no. 21115), 21mM MgCl2 (Sigma-Aldrich, cat.no M1028) and 0.2U/μl RNAsin (Promega, cat.no. N2615)). Filtered nuclei were pelleted twice for 5 minutes at 500g at 4ºC, first followed by resuspension in 500 μl ST buffer, and second followed by resuspension in 50μl of Chromium loading (CL) buffer (PBS + 0.04% BSA + 1U/μl Sigma Protector RNase Inhibitor (Millipore Sigma, cat.no. 3335399001). Nuclei were counted using a PhotonSlide™ (Logos Biosystems, South Korea) with Acridine Orange & Propidium Iodide Cell Viability assay (Westburg, cat.no. LB F23001) on a

LUNA-FL™ Dual Fluorescence Cell Counter (Logos Biosystems). Nuclei were counted as the 'dead' cell fraction, and the suspension was diluted to 1.000.000 nuclei/mL in CL buffer. Single nuclei suspensions were used for single-nuclei gene expression analysis using the Chromium Next GEM Single Cell 3' Reagent Kits v3.1, according to the user guide (CG000315 Rev C). 7000-8000 nuclei were loaded per sample, and final libraries were sequenced to at least 25.000 reads/cell using a NextSeq2000 platform (Illumina, San Diego, CA, US).

4.5.2 Data preprocessing

Spatial transcriptomics data pre-processing

FASTQ-files were mapped to the GRCh38 human reference genome and spatially projected using Space Ranger v1.1.0 (10x Genomics). Before merging the datasets, ambient RNA was removed using an adjusted version of SpotClean v1.1.1 where all relevant genes are maintained [192]. Next Scanpy v.1.9.1 was used for quality control and processing [278]. Spots with less than 700 genes and 1000 reads were excluded, while genes were filtered for expression in more than 100 spots. The Scanpy toolkit was used to perform downstream processing per sample, including normalisation with the Scran R package (v1.26.1), log transformation and variable gene detection. The number of significant principal components for Leiden clustering and UMAP dimensionality reduction was determined using bootstrapping with a maximum of 30 components. Clustering resolutions ranging from 0.1-1.0 were assessed for stability with steps of 0.1 using clustree v0.4.2 [285].

Spatial lipidomics data preprocessing

The mass spectrometry imaging data was preprocessed by performing TIC normalization, peak picking and subsequent rebinning. Peak picking was performed on the mean spectrum across all datasets. Peaks in the mean spectrum with an intensity below 0.005% of the base peak were discarded, resulting in a total of 12510 retained peaks. Intensities were binned per spectrum using a 5 ppm window around the selected peaks.

Spatial multi-omics integration pipeline

As the ST and MSI data were acquired from two serial sections, the spatial lipidomics and transcriptomics data were indirectly registered to each other

by their respective H&E microscopy images. This strategy of co-registering multiple sections via one proxy section (i.e., microscopy) is inevitable in applications where multiple sections are involved. Before co-registration, representative images from the MSI data were generated for guiding the later registration step. These representative images were created via applying dimensionality reduction methods, such as NMF and UMAP. The MSI data was co-registered with its respective H&E microscopy images via a non-rigid registration workflow [267] (Aspect Analytics NV). Next, the MSI-slide H&E was co-registered to the H&E from the ST Visium slide. No additional co-registration is required to link the ST data and ST-slide microscopy as their spatial reference is inherent to the Visium ST data format. Finally, the MSI was directly linked with the ST data via a shared spatial coordinate system. MSI data was acquired at a spatial resolution (pixel size) of 30 μm, whereas the ST data was obtained with spot diameters of 55 μm. Multiple MSI pixels were aggregated into a single representative spectrum to match the ST spot via a Gaussian weighting algorithm [257].

Morphological feature extraction from microscopy data

Morphological features from high resolution H&E microscopy image of the MSI section were extracted using a pre-trained model that was adapted originally from the SimCLR model introduced by Chen et al. [62]. The model tries to maximize the agreement between two stochastically augmented views of the same image via a contrastive loss function. Ciga et al. later applied this SimCLR model with minor modifications and pre-trained it on 57 multi-organ histopathology datasets without any labels [65]. These 57 histopathology datasets contained microscopy images with various types of staining and resolutions, which helped the model learn features with better quality. We applied this pre-trained model on high-resolution H&E-stained microscopy data, acquired at a 40x magnification, and extracted the morphological features using overlapping windows of 512 $\times$ 512-pixel patches of the image centred on each spot. Thus, each patch covers approximately 128$\times$128 μm^2, which is similar in size to the patches extracted from histopathology datasets that the model was pre-trained on by Ciga et al. [65].

Dimensionality reduction via non-negative matrix factorization

We applied NMF to reduce the high dimensionality of our MSI and ST data. NMF has been widely used in both fields [38, 88, 265, 193] as it has the non-negativity constraint and can generate interpretable parts-based representation. We used NMF to get a more in-depth understanding of each

modality in a fully unsupervised manner, based on their resulting distinct spatial distributions and associated spectral or gene expressions. In short, NMF factorizes the matrix X into two non-negative matrices $X = WH$, where W represents the loading matrix and H is the score matrix which represents the data in lower dimensions. For example, in the case of MSI data, the lipid spectral signature of each pixel can be approximated by the additive linear combination of the column vector (spectral basis element) from W, weighted by H, which describes the contribution of each basis element to the spectrum at each pixel. We implemented NMF by the Python package *sklearn.decomposition.NMF*. As both MSI and ST data are known to follow Poisson distribution, we used Kullback–Leibler divergence (KL-NMF) as the cost function to incorporate Poisson noise. Multiplicative Update was used as the solver. The number of components was selected as 20.

Correlation analysis

We applied Pearson correlation by Python package numpy.corrcoef and scipy.stats.pearsonr to measure the relationship between two datasets. The output correlation coefficient ranges from -1 to 1 with 0 implying no correlation. Correlations of -1 or 1 represent an exact linear relationship (negative or positive).

Cohen's d

We calculated the Cohen's d value for gene-lipid correlations across different tissue types. Cohen's d is defined as the difference between two means divided by the pooled standard deviation for the data. The resulting Cohen's d values were ranked to find the most different gene-lipid correlations between normal and tumour tissue types. Using this approach, a set of changes in correlation were identified. We firstly calculated the Cohen's d value for gene-lipid correlations across different tissue types via computing their mean correlations in 8 tumour samples and 8 'normal' samples separately.

Processing of the single-nuclei RNA-sequencing data and cell type deconvolution

FASTQ-files were mapped to the GRCh38 human reference genome with Cell Ranger v6.1.2 (10x Genomics). Before merging the datasets, ambient RNA was removed using SoupX [284], while scDblFinder [101] was used to indicate doubles. Next Scanpy v.1.9.1 was used for quality control and

processing [278]. Only cells with the following criteria were considered for further analysis: more than 500 uniquely expressed genes, less than 10% of the UMI counts mapping to mitochondrial sequences and less than 5% of the UMI counts assigned to ribosomal genes. All genes that were not expressed in at least 100 cells were not considered in the downstream analysis. The Scanpy toolkit was used to perform downstream processing per sample, including normalisation, log transformation and variable gene detection. The number of significant principal components for Leiden clustering and UMAP dimensionality reduction was determined using bootstrapping with a maximum of 60 components. Clustering resolutions ranging from 0.1-1.0 were assessed for stability with steps of 0.1 using clustree v0.4.2 [285]. Cell types and states were assigned to the clusters using marker genes obtained from the PanglaoDB [97] database (version of 27/03/2020), the canonical markers from Song et al. [234] and the NMF signatures obtained from the analysis of matching ST samples. Copy-number landscapes were inferred from the transcriptomic data with inferCNV [202] v.1.3.3 to validate clusters annotated as tumorigenic. Integration of the samples was performed in a semi-supervised manner using scANVI [281]. First the transcriptome data was integrated using scVI [159] (v.0.14.5, n_hidden=128, n_latent=50, n_layers=2, dispersion='gene-batch') with correcting batch effects and removing unwanted source of variations (total counts, percent mitochondrial genes, and percent ribosomal genes for 'continuous_covariate_keys'). After training the model for 400 epoch, scANVI using the cluster labels obtained at single-cell level, was trained for an additional 10 epochs. The integrated latent embedding generated by scANVI was used for downstream analysis (clustering and visualisation). For gene-level analyses, uncorrected counts were used. Spot deconvolution was performed using the integrated single-nuclei data per patient and Cell2Location [136]. The lipid metabolism gene sets of the KEGG pathway database [**?**] and the 'score genes' function of scanpy were used to asses pathway activities.

Lipid Annotation/Identification

Lipids were annotated with mass accuracy of 3 ppm based on the LipidMaps guidelines on shorthand notation for lipid structures published by Liebisch et al. [155, 154] For the glycerophospholipids (GPL), the following shorthand notations of lipid classes were used: PC – phosphatidylcholines, LPC –lysophosphatidylcholines, PC O – phosphatidylcholine ethers, PC P – phosphatidylcholine plasmalogens, PE – phosphatidylethanolamines, LPE –lysophosphatidylethanolamines, PE O – phosphatidylethanolamine ethers, PE P – phosphatidylethanolamine plasmalogens, PS – phosphatidylserines, PI – phosphatidylinositols. For the sphingolipids (SPL), the following class name abbreviations were used: Cer – ceramides, SM – sphingomyelins, and

Hex(n)Cer – hexosyl ceramides, where n refers to the number of hexosyl units. Cholesterol is referred to as 'Chol', cholesteryl esters as CE, acylcarnitines as CAR, monoacylglycerols as MG, and diacylglycerols as DG. Lipids are annotated at the lipid species level, namely the lipid class abbreviation is followed by a total number of carbons in fatty acyl(s) rest(s), colon, and total number of double bonds in fatty acyl rest(s). For example, PC 36:3 refers to phosphatidylcholine with a total number of carbons in both fatty acyls equal to 36 and total number of double bonds equal to 3. The same principle was used for annotation of signals corresponding to sphingolipids, but additionally the total number of oxygen was indicated after semi-colon, for example, SM 36:2;O2 refers to a sphingomyelin with sum composition of carbon atoms in N-linked fatty acyl rest and sphingoid base of 36, the number of double bonds equal to 2, and total number of O-atoms equal to 2. CE 16:0 means cholesteryl ester with palmitic acid attached to cholesterol via an ester bond. In MALDI-2 mode cholesteryl esters form dimers as showed recently by Bowman et al. [46]. For m/z values annotated as dimers of cholesteryl esters, the following shorthand notations were used: 2CE for cholesteryl ester dimer, followed by total number of carbons in both fatty acyls, colon, and total number of double bonds in both fatty acyls e.g., 2CE 38:4.

4.6 Acknowledgement

We thank all patients that participated in this study. WZ was supported by a Baekeland PhD grant from Flanders Innovation and Entrepreneurship (VLAIO) - HBC.20192204. XS is a recipient of a PhD fellowship (SB/1S49218N) from the Research Foundation – Flanders (FWO) and was in part supported by Opening the Future (OtF). SRE acknowledges support from the Australian Research Council (FT190100082). TS acknowledges support through the Australian Government Research Training Program Scholarship. SV and VdL were sponsored by an FWO PhD Followship strategic basic research (1S93320N and 1157918N, respectively). KV, SK and TV are supported by KU Leuven (IDN/19/039, Opening the Future (OtF)). This work was supported by a grant from the Belgian Foundation Against Cancer (STK), a Focus Group grant from the LKI Fund for Innovative Cancer Research (FIKO), the KU Leuven Opening the Future campaign, KU Leuven Incubation financing Lipometrix, a grant PRISMO from the Flemish Resilience Plan, FWO-EOS projects 30837538 (DECODE), G0F6718N (SeLMA), FWO-SBO projects S001623N (LIPOMACS) and S005319N, KU Leuven Research Fund projects C14/21/095, BOF/23/064, C16/15/059, C3/19/053, C24/18/022, C3/20/117, C3I-21-00316), US-DOD project W81XWH-19-1-0566, Industrial Research Fund (Fellowships 13-0260, IOFm/16/004) and several Leu-

ven Research and Development bilateral industrial projects, Infrastructure projects TOP/23/014, IO12220N and I013218N, TBM Project T001919N; PhD Grant (SB/1SA1319N), the Flanders AI Research Program, VLAIO CSBO (HBC.2021.0076), Baekeland PhD (HBC.20192204), European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant grant agreement No 885682) and EU Interreg EMR23 EURLIPIDS, Stand up to Cancer / Kom op tegen Kanker - the Flemish cancer society, CM (Christelijke Mutualiteit), and private donations. Figure 4.1A was created with BioRender.com.

Chapter 5

Details of applied methods

In this Chapter, we provide more details about the methods that were applied throughout previous Chapters.

5.1 Pre-trained neural networks

In Chapter 2 and 3, we applied pre-trained neural network models based on *Xception* [64] and *SimCLR* [62], in order to extract the imaging features from ion images (of MSI data) and histopathological data, respectively. The Xception model, utilized in Chapter 2, was pre-trained on a large manually labeled natural-scene database ImageNet [76]. While in Chapter 3, the SimCLR model was pre-trained on 57 multiple multi-organ histopathology datasets. In the next few sections, we will discuss in more detail about these models and their comparisons with other alternatives.

5.1.1 Xception model

Xception, which stands for *'Extreme Inception'*, is a convolutional neural network model based entirely on depthwise separable convolutional layers. Xception was inspired by Inception architecture [241], which factors convolutions into multiple branches operating successively on channels. In this way, the learning process gets easier and more efficient, i.e., learning richer representations with less parameters. Xception replaced Inception modules with depthwise separable convolutions, together with residual connections,

Model	Size (MB)	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth	Time (ms) per inference step (CPU)	Time (ms) per inference step (GPU)
Xception	88	79.0%	94.5%	22.9M	81	109.4	8.1
VGG16	528	71.3%	90.1%	138.4M	16	69.5	4.2
VGG19	549	71.3%	90.0%	143.7M	19	84.8	4.4
ResNet50	98	74.9%	92.1%	25.6M	107	58.2	4.6
ResNet50V2	98	76.0%	93.0%	25.6M	103	45.6	4.4
ResNet101	171	76.4%	92.8%	44.7M	209	89.6	5.2
ResNet101V2	171	77.2%	93.8%	44.7M	205	72.7	5.4
ResNet152	232	76.6%	93.1%	60.4M	311	127.4	6.5
ResNet152V2	232	78.0%	94.2%	60.4M	307	107.5	6.6
InceptionV3	92	77.9%	93.7%	23.9M	189	42.2	6.9
InceptionResNetV2	215	80.3%	95.3%	55.9M	449	130.2	10.0
MobileNet	16	70.4%	89.5%	4.3M	55	22.6	3.4
MobileNetV2	14	71.3%	90.1%	3.5M	105	25.9	3.8
DenseNet121	33	75.0%	92.3%	8.1M	242	77.1	5.4
DenseNet169	57	76.2%	93.2%	14.3M	338	96.4	6.3
DenseNet201	80	77.3%	93.6%	20.2M	402	127.2	6.7

Figure 5.1: Comparison of available models in Keras package: The top-1 and top-5 accuracy refers to the model's performance on the ImageNet validation dataset. Depth refers to the topological depth of the network. This includes activation layers, batch normalization layers etc. This figure is adopted from `https://keras.io/api/applications/`.

making the architecture easy to define and modify. The feature extraction part of the model consists of 36 convolutional layers, which are structured into 14 modules. All of those modules have linear residual connections around them, except the first and the last ones.

At the time of writing the publication from Chapter 2 (year 2020), Xception model achieved the most promising classification performance on ImageNet [64]. Additionally, comparing with other alternative models (i.e., VGG, ResNet, Inception), Xception has a smaller size and less parameters. Figure 5.1 shows the comparison of various available models in the Keras package [1].

We applied the pre-trained Xception model to extract the image features from ion images and it was implemented via the Keras Python pack-

[1] `https://keras.io/api/applications/`

age: *tf.keras.applications.Xception*. Parameters: weights='imagenet', pooling='max', include_top=False. The dimensions of the generated embeddings are 2048.

5.1.2 SimCLR model

SimCLR, which stands for 'a simple framework for contrastive learning of visual representations' was proposed by Chen et al. [62] in 2020. Contrastive learning is an important category in **self-supervised learning**, which has drawn massive attention for its data efficiency and generalization ability. Self-supervised learning obtains 'labels' from the data itself through a 'semi-automatic' manner and predicts part of the data from other parts, which could be incomplete, transformed or corrupted through data augmentation [156]. Thus, it does not require manual labels.

Contrastive learning seeks to train an encoder to encode input x into an explicit vector z to measure similarity and is found to be useful for almost all visual classification tasks. This is probably because the contrastive object is modeling the class-invariance between different image instances. As mentioned in the review of self-supervised learning by Liu et al. [156], *'a properly designed training objective, which is related to downstream tasks, could turn randomly initialized models into excellent pre-trained feature extrators.'*

In Chapter 3, we utilized the contrastive learning model - SimCLR [62], which was pre-trained on a large histopathology database by Ciga et al. [65]. The framework of the SimCLR model is illustrated in Figure 5.2A, which aims to learn representations of the data by maximizing the agreement between two stochastically augmented views of the same data instance via a contrastive loss in the latent space. A given data x is transformed randomly through a stochastic data augmentation module, resulting in two correlated views of the same instance: $\tilde{x_i}$ and $\tilde{x_j}$ (consider as a positive pair). As a neural network base encoder, $f(\cdot)$ extracts representation vectors from augmented data. Commonly used ResNet architecture was adopted to obtain $h_i = f(\tilde{x_i}) = \text{ResNet}(\tilde{x_i})$, where $h_i \in \mathbb{R}^d$ is the output after the average pooling layer. Then a small neural network projection head $g(\cdot)$ maps the representations to the latent space where the contrastive loss is calculated. A MLP with one hidden layer was used to obtain $z_i = g(h_i) = W^{(2)}\sigma(W^{(1)}h_i)$, where σ is ReLU nonlinearity.

Given a set $\{\tilde{x_k}\}$ including a positive pair of example $\tilde{x_i}$ and $\tilde{x_j}$, the contrastive prediction task is to identify $\tilde{x_j}$ in $\{\tilde{x_k}\}_{k\neq i}$ for a given $\tilde{x_i}$. A minibatch of N examples are randomly sampled, from which pairs of augmented examples are derived. Let similarity$(u, v) = u^T v/\|u\|\|v\|$ denoted the dot product between ℓ_2 normalized u and v (i.e., cosine similarity). The loss function for a positive

pair of examples (i, j) is defined as

$$\ell_i, j = -\log \frac{\exp(\text{similarity}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\text{similarity}(z_i, z_j)/\tau)}, \tag{5.1}$$

where $\mathbb{1}_{[k \neq i]} \in \{1, 0\}$ is an indicator function, which outputs 1 when $k \neq i$ and 0 otherwise. τ is a temperature parameter and for convenience, this loss is termed *NT-Xent* (the normalized temperature-scaled cross entropy loss) [62].

Ciga et al. pre-trained the SimCLR model on a large (unlabeled) histopathology database i.e., 57 multi-organ histopathology datasets. For each pre-training step with a batch size of $2N$, each augmented image has a positive and $2(N-1)$ negative samples. The pre-training framework by Ciga et al. is illustrated in Figure 5.2B. It is challenging to obtain visually diverse patched for digital pathology. Since multiple unlabeled datasets were used, images with various staining, resolution and tissue types will lead to a more diverse dataset. That can significantly improve the learned representations by the self-supervised learning model. In the work of Ciga et al., some valuable insights were provided: (1) pre-training on a large, varied database resulted in better validation performance than pre-training on a single dataset; (2) The features learned from the model that was pre-trained on those histopathology datasets, are better than the features that were learned from networks pre-trained on ImageNet, indicating the method has learned useful domain-specific features.

The pre-trained model was implemented with Pytorch and adopted from Ciga et al [65] [2].

Comparison of alternative pre-trained models

In the beginning of this work, we conducted some experiments to compare the learned features from various pre-trained models, including ResNet [115] and MobileNet [122]. MobileNet is investigated for computational performance reasons, as shown in Figure 5.1. As this work will be used in a real application (collaboration with Aspect Analytics NV.), we are not only considering the quality of the learned embeddings from a pre-trained model, but also considering its computational performances, e.g., size of the model, inference time. UMAP was applied to reduce the high-dimensional features extracted from the pre-trained models. Figure 5.3 shows the 2 dimensional UMAP embeddings of features learned from pre-trained SimCLR, ResNet and MobileNet model. This dataset was from the same database that was used in Chapter 3. However, here we included multiple classes that were

[2] `https://github.com/ozanciga/self-supervised-histopathology`

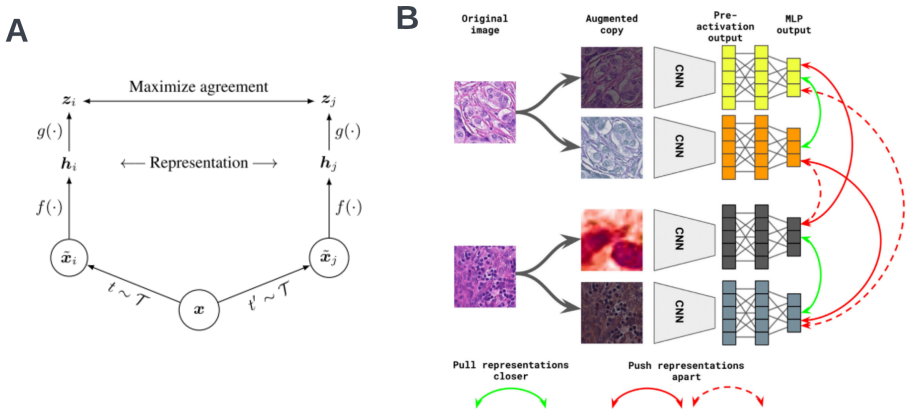


Figure 5.2: Architecture of the SimCLR model: Panel A shows the simple framework of the original work, which is reused from by Chen et al. [62] with permission (Creative Commons Attribution 4.0 International License). Panel B is reused from Ciga et al. [65] with permission (Creative Commons CC-BY license). Each image patch is augmented twice. Features are extracted via CNN, and transformed into a lower embeddings space at the multi-layer perceptron (MLP) output. Augmented image patches from different sources are pushed apart using the NT-Xent loss function (both solid and dashed red arrows represent an identical push operation.

labeled by various experts. As shown in the figure, UMAP embeddings of the representations learned from the pre-trained SimCLR model by Ciga et al. are more separable across all classes than learned from other pre-trained models (based on ImageNet). Interestingly, the pre-trained SimCLR model based on histopathology datasets learned more separable representations than pre-trained based on ImageNet.

5.1.3 Benchmark on the Camelyon dataset

Before the work of Ciga et al. was published (year 2022), we've tested different pre-trained models for extracting morphological features from microscopy data. We used open, annotated microscopy data from the Camelyon 16 challenge [32], which contains H&E stained whole-slide images of lymph nodes of women with breast cancer from 2 centers in the Netherlands. A panel of 11 pathologists did the annotations within time constraints.

We have compared 3 different models based on this Camelyon dataset: Bidirectional Generative Adversarial Network (BiGAN) [81], ResNet50 [115] and MobileNet [122]. Tellez et al. pre-trained the BiGAN model on 3 different

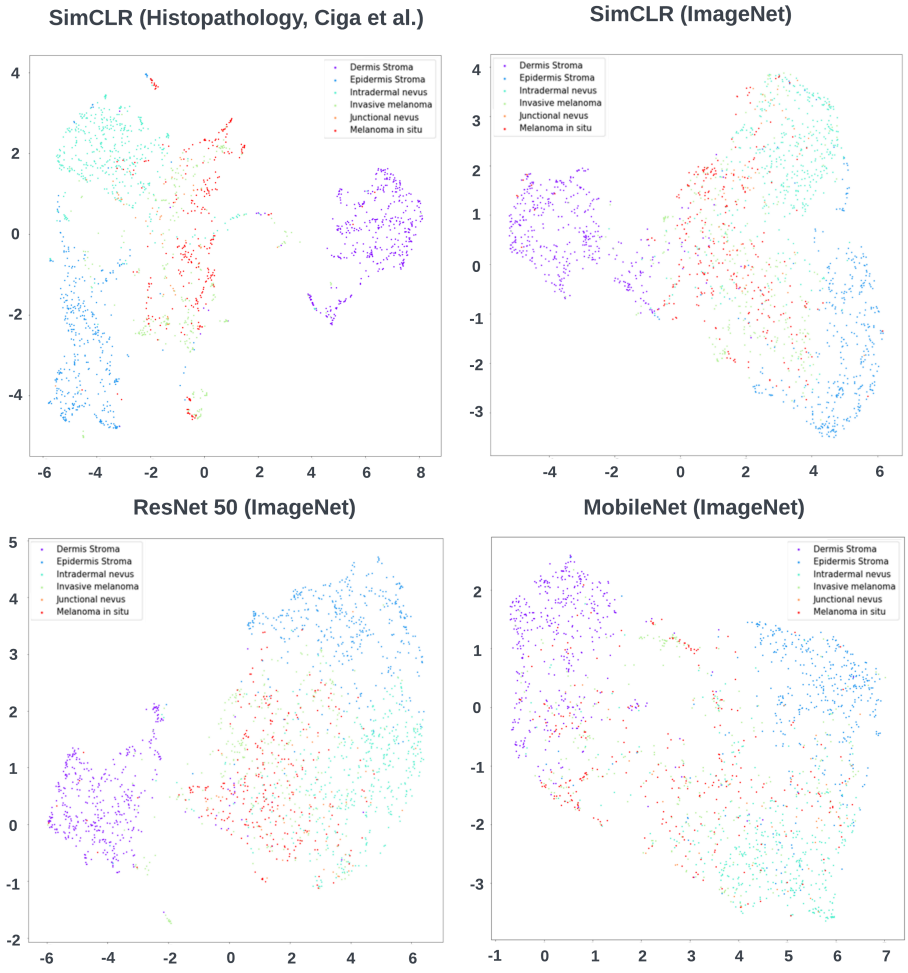


Figure 5.3: Comparison of different pre-trained models: top two scatter plots illustrate UMAP embeddings of outputs generated from pre-trained SimCLR model. The left on is from the pre-trained model based on 57 histopathology datasets, and the right one is from the pre-trained model based on ImageNet database. Bottom 2 scatter plots show the UMAP embeddings of outputs generated from pre-trained ResNet50 (left) and MobileNet (right), based on ImageNet database.

histopathological datasets and it achieved the best classification performance over the other two models, i.e., variational autoencoder and contrastive learning models [247]. We selected MobileNet for benchmarking, as it achieves promising classification performance (on ImageNet) considering its small size and high inference speed, as shown in Figure 5.1.

Thus, we applied 3 pre-trained neural network models, i.e., BiGAN (3 histopatholgoy datasets, Tellez et al.), ResNet50 (ImageNet) and MobileNet (ImageNet) on the Camelyon dataset to compare their learned morphological features. 4 whole-slide images were randomly selected from the Camelyon dataset, together with its detailed annotations. Sample Tumor_047 is shown in Figure 5.4A, where the left is the H&E stained whole-slide image and the right is its overlay with expert's annotations of tumor regions (in yellow). Firstly, a patch-extraction model was built, consisting of the following steps: automatically detect the region of interest, crop the patch by certain size and preprocess the patches. We extracted each patch with patch size as 128 and shift size as 64. Since the resolution of the image is $0.5\,\mu$m, each patch covers a region of $64\,\mu$m by $64\,\mu$m. Example patches are shown in Figure 5.4B and each whole-slide image contains **around 100k patches**.

In order to visually compare the learned representations for those pre-trained models, we first applied UMAP on their extracted high-dimensional features to reduce the dimensionality to 3. The UMAP mapping was performed with a Python implementation [182], with parameters: *n_neighbors* = 10, *n_components* = 3, *min_dist* = 0.1, *metric* = 'cosine'. Resulting 3 dimensional UMAP embeddings are shown in Figure 5.5, where representations learned from MobileNet (ImageNet) shows a clear separation of tumor regions (the green cluster in the scatter plot) from the rest of the tissue region. A tiny tumor region (red arrow) was also detected nicely by MobileNet (ImageNet) and ResNet50 (ImageNet).

Furthermore, we built a classifier based on the linear SVM model. Each whole-slide image was split into 5 folds for generating cross-validated results, as shown in Figure 5.4C. The SVM model was implemented with the Python sklearn package *sklearn.linear_model.SGDClassifier*, with parameters *penalty* = '*l*2', *loss* = 'hinge', *max_iter* = 450, *tol* = 1e-4, *l1_ratio* = 0.5. The Cross-validated results over 4 different whole-slide images from the Camelyon dataset are shown in Table 5.1. MobileNet achieves the best performance across all images. Although ResNet50 achieves comparable performance as MobileNet, its size is larger and the inference time is slower than MobileNet. BiGAN achieves the lowest classification performance, and this may be caused by the limited pre-training data, i.e., only 3 public histopathology datasets.

Surprisingly, MobileNet did not just achieve the best classification performance,

but also has the most lightweight network and fastest inference time. It was built by depth-wise separable convolution. Thus, it has been designed to effectively maximize accuracy while being mindful of the restricted resources of a mobile device or embedded application, which can be an important factor in an industrial setting.

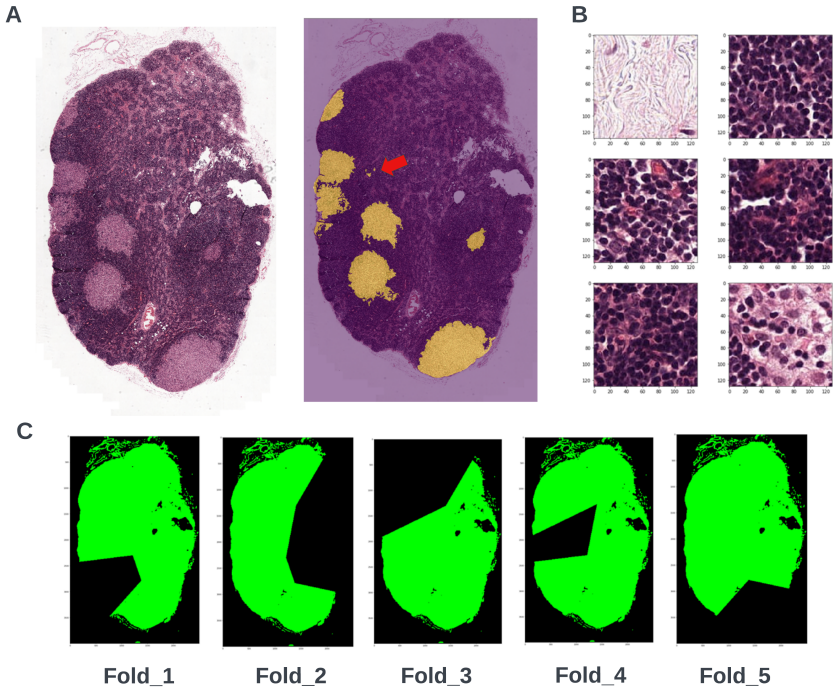


Figure 5.4: Camelyon 16 challenge data Tumor_047: Panel A shows the H&E stained whole-slide image (left), and its overlay with expert's annotations (in yellow) of tumor regions (right). Panel B shows the extracted patches from this sample. 5 folds of training dataset are displayed in panel C.

5.1.4 Discussion

In both Chapter 2 and 3, we constructed this kind of framework: extract features via a pre-trained model, then apply clustering or classification algorithms to those learned features for the downstream data analysis. Especially in Chapter 3, we only trained the classifier without changing or fine-tuning the feature extraction part. This was deliberately chosen. This way, we can show to the reader that this feature-extraction module can be

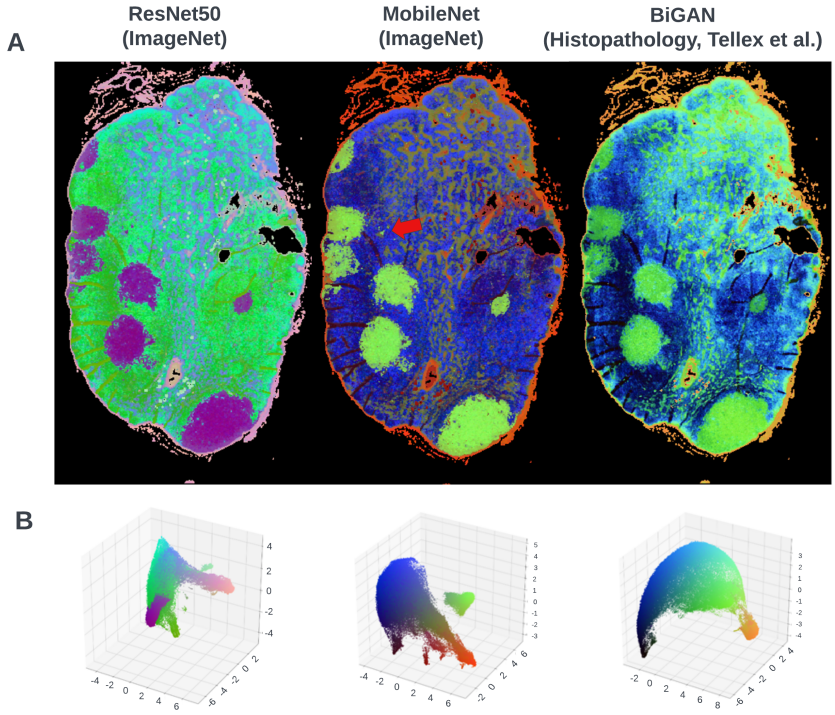


Figure 5.5: 3-d UMAP embeddings of the outputs generated from benchmarked pre-trained models. UMAP maps the learned high-dimensional features from pre-trained models into 3 dimensions, which is shown in panel B as scatter plots. Panel A displays the 'UMAP images' of learned features by mapping embeddings to the pixel locations.

easily plugged into any data analysis pipeline in a flexible way. Additionally, this feature-extraction module is 'cost-free' and highly efficient as there is no training required for building the framework. As many readers in the field of MSI are not always very familiar with machine learning, this type of framework is more user-friendly and efficient.

5.2 Downstream classification model

In Chapter 3, we applied a pre-trained neural network model to extract the morphological features from microscopy images, then a classifier was built to distinguish the nevus from melanoma. We started to use the linear support vector machine (SVM) model as the classifier, due to its simplicity and

Data	Model	fold_1	fold_2	fold_3	fold_4	fold_5	**Mean**
Tumor_031	BiGAN	0.86	0.92	0.7	0.8	0.93	0.84
	ResNet50	0.94	0.96	0.95	0.92	0.97	**0.95**
	MobileNet	0.94	0.96	0.95	0.94	0.96	**0.95**
Tumor_026	BiGAN	0.92	0.69	0.7	0.88	0.9	0.82
	ResNet50	0.97	0.96	0.92	0.94	0.97	0.95
	MobileNet	0.98	0.96	0.94	0.95	0.98	**0.96**
Tumor_047	BiGAN	0.89	0.66	0.63	0.91	0.88	0.79
	ResNet50	0.95	0.95	0.91	0.98	0.97	**0.95**
	MobileNet	0.96	0.93	0.91	0.97	0.97	**0.95**
Tumor_005	BiGAN	0.69	0.96	0.89	0.78	0.96	0.85
	ResNet50	0.91	0.95	0.96	0.91	0.99	**0.94**
	MobileNet	0.91	0.94	0.99	0.89	0.98	**0.94**

Table 5.1: Cross-validated classification results for benchmarked pre-trained models

efficiency. Based on the nested cross-validated results shown in Table 5.2 - 5.4, the linear SVM model achieves promising classification performance (i.e., AUROC = 0.968). Thus, we chose this simple yet excellent classification model for this work.

5.2.1 Linear SVM model

The SVM model aims to construct a separating hyperplane to distinguish classes. By maximizing the margin between instances of each class and the hyperplane, an optimal separating hyperplane can be obtained. For most real-life problems, not all the instances of the training set will be correctly classified, when taking a linear classifier. Thus, misclassifications will be allowed in contrast to hard-margin classifiers. Consider a given training set:

$$\{(x_i, y_i)|x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n \tag{5.2}$$

Where x_i is the input vector for instance i and y_i is its associated label. p is the input dimensionality and n is the amount of training samples. A common formulation for soft-margin SVM is the support vector classifier (SVC) by

Cortes and Vapnik [68]:

$$\begin{aligned} \min_{w,\xi,b} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \\ \text{s.t.} \quad & y_i(< w, \varphi(x_i) > +b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, \quad \forall_i \end{aligned} \tag{5.3}$$

The misclassification cost C is a positive regularization constant. It is a user-specified tuning parameter, explicating the trade-off between maximizing the margin and rejecting misclassifications. The slack variables ξ are used to tolerate misclassifications in non-separable problems (soft-margin). SVMs map the p-dimensional input vector x from the input space to the P-dimensional feature space via a (non)linear function $\varphi(\cdot) : \mathbb{R}^p \to \mathbb{R}^P$. Typically, the range of φ is higher dimensional than its domain: $P > p$. The main goal of the map φ is to convert nonlinear relations into linear ones, which can be discovered in feature space efficiently.

Surprisingly, the data points are mapped into an inner product space does not need to compute their images explicitly. An explicit mapping to feature space can be avoid via so-called *kernel trick*. This states that there exists a mapping φ and an expansion

$$\kappa(x_i, x_j) =< \varphi(x_i), \varphi(x_j) >, \tag{5.4}$$

if and only if, for all square-integrable functions $g(x)$ has

$$\iint \kappa(x, y)g(x)g(y)dxdy \geq 0 \tag{5.5}$$

A kernel function κ computes the inner products directly, without first computing the images of data in feature space, enabling the use of high dimensional feature spaces without computation problems. Some widely used kernel functions are:

- linear SVM: $K(x_i, x_j) = x_i^T x_j$,
- polynomial SVM of degree d: $K(x_i, x_j) = (\tau + x_i^T x_j)^d, \tau \geq 0$,
- radial basis function (RBF): $K(x_i, x_j) = \exp(\|x_i - x_j\|_2^2/\sigma^2)$

5.2.2 Model selection results

In this work, we started using the linear kernel. The Python package *sklearn.linear_model.SGDClassifier* was used to implement the linear SVM

model. Parameters *loss*='hinge' (gives linear SVM), *alpha* is the regularization constant, *l1_ratio* is the Elastic Net mixing parameter, with $0 \leq l1_ratio \leq 1$. When $l1_ratio = 0$ corresponds to L2 penalty (standard regularizer for linear SVM models), $l1_ratio = 1$ to L1.

During the parameter tuning for the SVMs model, we used a grid search strategy, where the model exhaustively generates candidates from a grid of parameter values that are predefined by the user.

Furthermore, we leveraged nested cross-validation in order to estimate the generalization error underlying each model, including the performance variability induced by the hyperparameter optimization process itself. Additionally, in order to further prevent information leaking, spots measured from the same patient/tissue were grouped in the same fold during the whole nest cross-validation process.

In summary, we first split the total samples into a training set and a held-out test set on patient level. Then we applied a grouped-nested-cross-validation strategy on the training set for model selection and evaluation. The held-out test set was not used during the whole training procedure. Thus the whole framework was designed carefully during the model selection and evaluation, so that we can fairly compare classification performance from unimodal and multimodal pipelines.

To efficiently keep track of the intermediate results during the nested cross-validation, we applied an open source platform *MLflow* [3] to manage the machine learning life-cycle. Specifically, we applied *mlflow.sklearn.autolog* for logging and loading scikit-learn models. Results of the nested cross-validation on the training dataset are shown in Table 5.2 - 5.4. The results on the held-out dataset are illustrated in Chapter 3.

[3] `https://github.com/mlflow/mlflow`

Outer fold	best alpha	best l1 ratio	best inner cv score-AUC	outer cv test mean AUC	outer cv test mean F1	outer cv test mean Precision	outer cv test mean Recall
1	1.00E-05	0.5	0.895	0.940	0.902	0.959	0.852
2	1.00E-09	1	0.913	0.807	0.679	0.853	0.563
3	1.00E-05	0.5	0.897	0.953	0.844	0.874	0.815
4	1.00E-05	0.5	0.905	0.911	0.832	0.826	0.837
5	1.00E-09	1	0.910	0.956	0.893	0.922	0.866
6	1.00E-05	0	0.915	0.911	0.784	0.933	0.676
7	1.00E-09	0.5	0.910	0.963	0.856	0.794	0.929
8	1.00E-05	0	0.912	0.899	0.800	0.755	0.850
9	1.00E-09	0.5	0.908	0.864	0.773	0.682	0.890
10	1.00E-05	0	0.910	0.948	0.864	0.863	0.866
	Mean		0.907	0.915	0.823	0.846	0.815
	St.d		0.007	0.049	0.067	0.085	0.110

Table 5.2: Nested cross-validation results of unimodal MSI pipeline

Outer fold	best alpha	best l1 ratio	best inner cv score-AUC	outer cv test mean AUC	outer cv test mean F1	outer cv test mean Precision	outer cv test mean Recall
1	0.00001	0	0.932	0.962	0.910	0.968	0.859
2	0.00001	0	0.941	0.887	0.839	0.807	0.873
3	0.00001	0	0.938	0.902	0.758	0.778	0.738
4	0.00001	0.5	0.935	0.933	0.815	0.882	0.758
5	0.00001	0	0.931	0.959	0.766	0.943	0.645
6	0.00001	0	0.941	0.915	0.797	0.919	0.703
7	0.00001	0	0.927	0.977	0.757	0.609	1.000
8	0.000001	0	0.933	0.925	0.833	0.897	0.777
9	0.00001	0.5	0.936	0.951	0.812	0.883	0.752
10	0.00001	0	0.928	0.963	0.913	0.882	0.945
	Mean		0.934	0.937	0.820	0.857	0.805
	St.d		0.005	0.030	0.056	0.104	0.111

Table 5.3: Nested cross-validation results of unimodal microscopy pipeline

Outer fold	**best alpha**	**best l1 ratio**	**best inner cv score-AUC**	**outer cv test mean AUC**	**outer cv test mean F1**	**outer cv test mean Precision**	**outer cv test mean Recall**
1	1.00E-06	0	0.954	0.980	0.731	1.000	0.576
2	1.00E-05	0	0.967	0.919	0.829	0.886	0.780
3	1.00E-05	0	0.962	0.964	0.850	0.875	0.826
4	1.00E-05	0	0.960	0.977	0.887	0.974	0.815
5	1.00E-05	0	0.960	0.987	0.922	0.954	0.892
6	1.00E-05	0	0.971	0.939	0.850	0.945	0.772
7	1.00E-09	1	0.958	0.996	0.919	0.850	1.000
8	1.00E-09	1	0.962	0.968	0.888	0.863	0.915
9	1.00E-05	0	0.953	0.975	0.888	0.906	0.871
10	1.00E-09	0.5	0.962	0.976	0.898	0.952	0.850
	Mean		0.961	0.968	0.866	0.920	0.830
	St.d		0.005	0.023	0.056	0.051	0.112

Table 5.4: Unimodal microscopy pipeline of multimodal MSI and microscopy pipeline

5.3 Uniform Manifold Approximation and Projection

Uniform Manifold Approximation and Projection (UMAP) [181] has been applied throughout the previous Chapters as a dimensionality reduction technique. In Chapter 2, we used UMAP to visualize the high dimensional MSI data in a low dimensional space (i.e., 3 dimensions) and utilized its embeddings for the downstream clustering task. In Chapter 3 and 4, we applied UMAP purely for the low dimensional visualization but **not** including its embeddings for any downstream tasks. Comparing with the embeddings from other nonlinear dimensionality reduction, such as t-SNE [259], UMAP provides better preservation of global structure in the final projection. Additionally, UMAP is much faster than t-SNE when handling large, high dimensional data.

Although UMAP offers a number of great advantages, it takes some care when interpreting the results. Hyperparameters are very important and it is not easy to choose good values. One needs to have good understanding of the data and UMAP algorithm. Thus, it is suggested to run UMAP multiple times with a variety of hyperparameters.

5.3.1 Method

The UMAP algorithm can be constructed in 2 phases: (1) constructing a fuzzy topological representation of the data; (2) optimizing the low dimensional representation to closely match the constructed fuzzy topological representation from phase 1.

The first phase aims to capture the topology of the manifold underlying the data via constructing a fuzzy topological representation. Each point in the dataset is given an edge to each of its k-nearest neighbors. The weights on the edge represent the likelihood that two points are connected. To check their connectedness, UMAP extends a radius outwards from each point to connect points until those radii overlap [5]. It is not required to connect the manifold as a whole - it can be made up of many connected components. However, any point on the manifold should have "some sufficiently small neighborhood" of the point that is connected. Here, "some sufficiently small neighborhood" is what the "local" part means. In practice, as we only have a finite approximation of the manifold, **no point is allowed to be completely isolated**. This is important for the high dimensional dataset, because distances in high dimensions tend to be larger yet also similar to one another, a phenomenon commonly referred to as the curse of dimensionality. The local connectivity constraint encourages the

model to focus on the difference in distances among nearest neighbours instead of the absolute distance, so that the local structure is preserved in balance with global structure [6].

Once we obtain the fuzzy topological representation of the data, it is converted into a low dimensional representation during the algorithm's second phase. In the ideal scenario, we want this low dimensional representation to have as similar a fuzzy topological structure as possible. This can be turned into an optimization problem. Basically, we construct a weighted graph in the first phase, and we interpret the weights as the probability of the simplex[4] existing. As both topological structures to be compared share the same 0-simplices (i.e., a point), we can imagine that we are actually comparing the two vectors of probabilities indexed by the 1-simplices (i.e., a line segment, between two 0-simplices). Thus, the choice for the loss function is the cross entropy, as these are Bernoulli variables (i.e., whether the simplex exist or not).

E is the set of all possible 1-simplices, and the weight functions $w_h(e)$ and $w_l(e)$ are the weight of 1-simplex e in the high dimensional and the low dimensional space, respectively. We can define the binary cross entropy between the two topological representations as:

$$\sum_{e \in E} w_h\left(e\right) \log\left(\frac{w_h(e)}{w_l(e)}\right) + (1 - w_h(e)) \log\left(\frac{1 - w_h(e)}{1 - w_l(e)}\right) \tag{5.6}$$

This can be interpreted similar to a force-directed graph layout algorithm. The left term $w_h\left(e\right) \log\left(\frac{w_h(e)}{w_l(e)}\right)$ provides an attractive force between the points e spans, when a large weight is associated to the high dimensional case. This is because in order to minimize this term, $w_l\left(e\right)$ needs to be as large as possible, which will happen when the distance between the points is as small as possible. On the contrary, the right term $(1 - w_h(e)) \log\left(\frac{1-w_h(e)}{1-w_l(e)}\right)$ provides a repulsive force to the ends of e when $w_h(e)$ is small, as this term will be minimized when $w_l(e)$ is as small as possible (i.e., the distance between the points is as large as possible).

This balancing process of pull and push, which is mediated through the weights on edges of the topological representation of the high dimensional data, will enable the low dimensional representation to settle into a state that represents the overall topology of the original data relatively accurately.

[4] In geometry, a simplex describes a triangle or tetrahedron to arbitrary dimensions

5.3.2 Parameters

Two most commonly used parameters, (*n_neighbors* and *min_dist*), control the balance between local and global structure in the final projection. *n_neighbors* is the number of approximate nearest neighbors used to construct the high dimensional graph. This means that when *n_neighbors* is small, UMAP will concentrate on very local structure of the data. On the contrary, when *n_neighbors* is large, UMAP will be pushed to look at larger neighbourhoods of each point when constructing the topological representation of the data. This encourages UMAP to represent the big picture of the data at the cost of losing fine detail.

Parameter *min_dist* refers to the minimum distance between points in low dimension space. This means that low values of *min_dist* will result in more tightly packed embeddings, which is interesting for clustering or obtaining finer topological structure. While larger values of *min_dist* will keep UMAP from packing points together but focus on preserving the broad topological structure.

In order to show their effect on the real dataset, we applied different settings for these two parameters on the MSI data from Chapter 4, and showed the embeddings in Figure 5.6. The top panel shows the 3-dimensional embeddings with parameter $min_dist = 0.5$, while $min_dist = 0.02$ is displayed in the bottom panel. The top left figure shows the annotations by the pathologists and the labels are presented underneath. Then in the columns, from left to right, the *n_neighbors* range from 25, 15 to 5. As illustrated in the figure, *min_dist* controls the tightness of the embeddings and *n_neighbors* balances the local and global structure in the final project. For example, when choosing a smaller *n_neighbors* value of 5 and a *min_dist* of 0.02, the embeddings look more separable, which are located in the 'Tumor ISUP 3', 'PIN'+'Stroma'+'Immune cells', and 'Benign epithelium' regions.

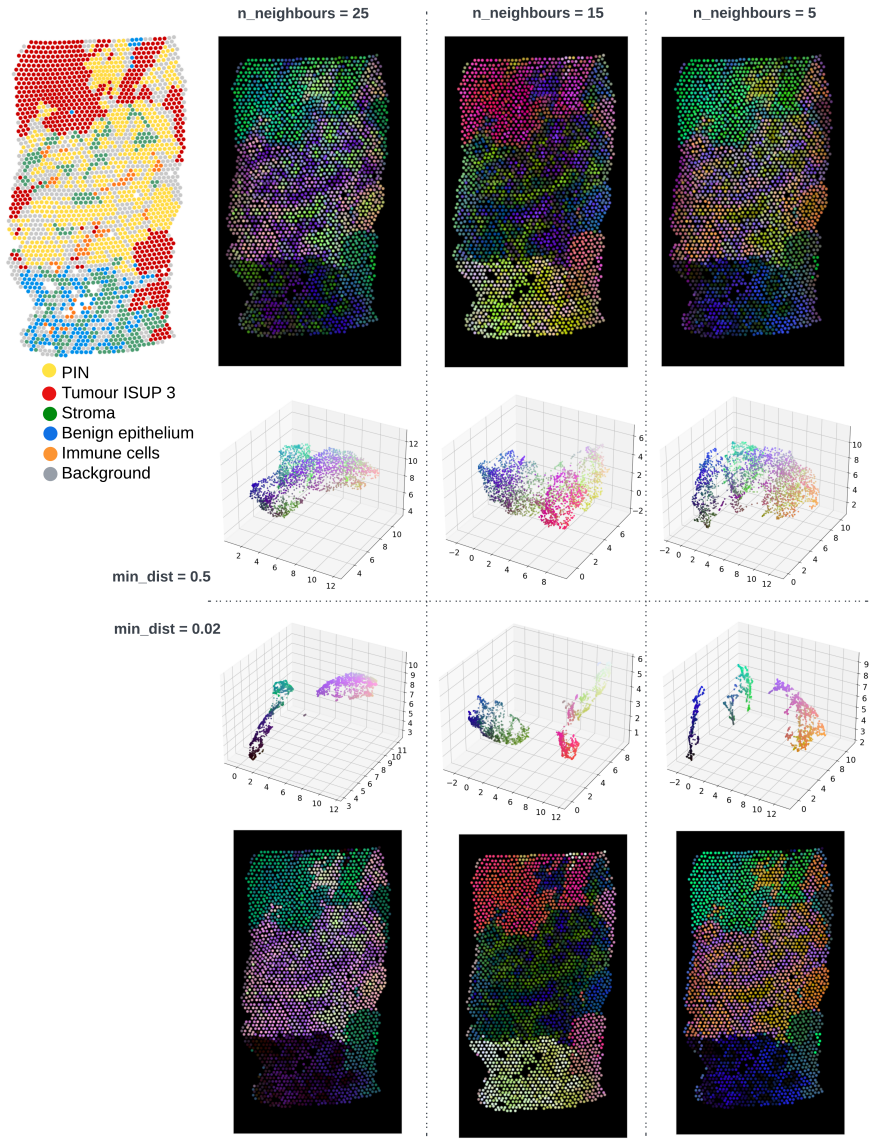


Figure 5.6: UMAP 3-d embeddings with different hyperparameter settings. UMAP is applied on the MSI dataset from Chapter 4 (human high-risk prostate cancer tissue) to reduce its dimensions to 3. Each setting of the hyperparameters generates a 3-d embeddings scatter plot, together with its translation to its corresponding pixel locations. The top left figure shows the annotations done by the pathologists, with their labels underneath. The top panel shows the embeddings when parameter $min_dist = 0.5$, and in the bottom panel displays $min_dist = 0.02$. From the left to the right column, the $n_neighbors$ range from 25, 15 to 5.

5.3.3 Limitations

Recently, UMAP and t-SNE have been widely applied in the field of biology, especially in the single-cell data [30, 137], to provide such human-parsable 'all-in-one' visuals (in 2 or 3 dimensions), which sometimes are used for the downstream qualitative and quantitative exploratory analysis. First of all, the goal of such visuals is to recapitulate local and global data point to point relationships. However, as reported by Chari et al. [59], this extreme dimension reduction (e.g., from 10k to 2 or 3) inevitably induces significant distortion of high-dimensional datasets, leading misinterpretation of the relationships between data points. Secondly, the choice of the hyperparameters is very sensitive to those embeddings, and no well-defined criteria for an 'optimal' choice is provided, which may introduce the confirmation bias [107, 137].

As such, it is recommended to use dimensionality reduction as an initial data-analysis step that requires further investigation. Conclusions should not be derived from the visual inspections of embeddings only, but should be combined with quantitative assessments, e.g., validation against the high-dimensional data [59, 118, 107]. As there are not really better alternatives, these methods are a standard in the field despite their flaws. Embeddings plots can be useful to support researchers, as long as we don't draw conclusions directly from those plots, but rather use them as tools to identify potentially interesting certain populations which are then investigated using the full data in a proper statistical sense.

Chapter 6

Conclusions and future perspectives

6.1 Contributions

In this thesis, we focused on augmenting computational analysis in MSI via advanced machine learning algorithms and multimodality. Throughout my research, I have strived to design, develop and benchmark novel MSI and spatial omics data analysis and integration strategies that are widely applicable, conceptually sound and easy to interpret.

At the outset of my work, the key objective was to develop a methodology that enables the joint analysis of MSI and microscopy data. In the second half of my project, emboldened by the initial successes in multimodal data integration, combined with the rise of spatial transcriptomics technologies, we decided to further increase the level of ambition by also incorporating spatial transcriptomics to create a unified computational framework for spatial multi-omics. The resulting framework enables the integration and downstream analysis of spatial omics data across multiple omics families, acquired from a collection of serial sections, measured using fundamentally different platforms, at different spatial resolutions and probe sizes, and incorporating ground-truth information provided by expert pathologists.

Specifically, I have focused on the following key questions and challenges:

1. How to summarize the main spatial trends in highly multiplexed

molecular imaging data, such as MSI? Which recurring spatial patterns are embedded in the thousands to hundreds of thousands of ion images within an MSI data set?

2. How to leverage routinely available histology information to further improve the classification performance of a state-of-the-art classification model based on MSI data, in order to eventually improve clinical diagnosis?

3. How to integrate MSI data with complementary high-plex imaging technologies, in order to obtain a more in-depth understanding of cancer biology?

To address the first question, in Chapter 2 we utilized a pre-trained neural network to detect complex spatial patterns of each ion image without having to train the model from scratch for each application. The resulting features were used to find spatial similarities between ion images, and group them together. This is one of the unsupervised analysis studies in MSI data, namely clustering ion images. By inspecting the resulting representative mean cluster images, one can have an overview of all ion images in this complex and high-dimensional MSI data. Compared to the standard clustering methods in MSI, our proposed pipeline provides better assignments of ion images, with more fine-grained clusters, resulting in more representative mean cluster images, quantified via our introduced relative isotope ratio metric in benchmarking results on public data.

In Chapter 3 we addressed the second question by incorporating morphological features from histopathological data with MSI. The proposed multimodal pipeline again utilized a pre-trained neural network to extract the features from microscopy data, which is often collected in tandem with MSI data. However, few studies focus on combining microscopy modality with MSI data to improve the downstream classification task. In fact, our work is the first which creates a true joint representation of chemical information obtained via MSI and morphological information derived from microscopy. Thus, we applied the multimodal MSI and microscopy pipeline on 331 patients' data for melanoma diagnosis, resulting in significant improvements over other unimodal pipelines, as quantified in an unbiased way via nested-cross-validated ROC-AUCs. The design of this multimodal strategy enables its use across a wide range of applications, both supervised and unsupervised.

Both Chapter 2 & 3 applied pre-trained neural networks to extract high-level features from either ion images or microscopy images, resulting in drop-in computational pipelines as there are no labels nor training required. The lack of ground truth data is one of the most challenging problems in computational

analysis of MSI, which will be discussed in more detail below. The feature extraction step by the pre-trained neural network model is near 'plug-and-play', which can be readily and flexibly incorporated in other machine learning applications of MSI, both supervised and unsupervised.

In Chapter 4 we developed a spatial-multi-omics integration pipeline and successfully applied it to in situ lipidomics and transcriptomics datasets from the serial sections of high-risk prostate cancer patients' tissue samples for the first time. The integrated datasets share the same coordinate systems and equivalent spot areas, thus enabling direct correlation and comparison between lipidomics and transcriptomics in prostate cancer study. Additionally, we integrated the morphological features extracted from the microscopy data into the spatial multi-omics dataset via a pre-trained neural network, which is the same as used in Chapter 3. High-quality detailed pathological annotations were obtained in the study as well, and similarly transformed to the shared coordinate system of spatial multi-omics data. Preliminary results from the integrated data have revealed the potential of the proposed pipeline, such as its ability to map molecular heterogeneity and to provide more in-depth understanding in prostate cancer biology.

Finally, in Chapter 5 we provided model details about the machine learning methods we applied throughout the previous Chapters and discussed their challenges and limitations in this field.

At the time of writing, we are among a select few world-wide that can truly integrate data from fundamentally different analytical platforms into a single unified data structure, which enables a large variety of downstream analysis. Moreover, our approach was established fully independently from other work, which is now starting to become available in literature.

6.2 Remaining computational challenges in MSI

Undoubtedly, MSI has become a promising molecular imaging technology with a proven track record in cancer research. However, there are still remaining challenges which need to be addressed.

One of the major challenges in MALDI-MSI experiments is the low reproducibility. For example, during the sample preparation, variance could be introduced and technical variability occurs across different labs, slowing down long-term MSI studies in e.g., longitudinal spatial metabolic profiling of cancer tissues [170]. After which, various pre-processing steps are demonstrated, for example quality check, feature selections and normalization. Currently, different labs

and scientists determine, based on their own empirical experiences, whether or when to apply a certain pre-processing step (e.g., normalization) hindering the comparison of data between labs, which may even lead to artifacts when generating new types of samples [15]. **Those are the examples of how a lack of high-quality data can block computational developments, specifically for machine learning algorithms.**

As discussed in Nature Metabolism about *why the metabolism field risks missing out on the AI revolution?* [83], one important reason is the lack of high-quality data. This is not so unfamiliar to the spatial transcriptomics (ST) field. In ST, with the rapid development of various technologies, a large variety of raw file formats and data structures are generated, making data and protocol standardization and exchange a challenge. To be able to develop computational methods (e.g., machine learning algorithms) it not only requires high-quality data but also their desired labels - the ground truth or the golden standard data. For example, in MSI, the ground truth data can be data from known molecules or experts annotations (i.e., pathologists' annotations). The ground truth data can be used not only for training the machine learning model but also for the evaluation and comparison across various computational methods.

With the rapid development of computational methods, it is urgent to conduct benchmarking studies on their performances, enabling the community to provide guidelines to determine appropriate algorithms for specific datasets. The lack of ground truth data has already hindered the computational progress in MSI, as numerous methods such as segmentation (i.e., cluster spectra data), have been developed but remain difficult to compare objectively, simply due to the lack of ground truth data [15]. This challenge leads to arbitrary and empiric choices of algorithms, slowing down the computational developments by preventing training and evaluating machine learning models, while also delaying the adoption of novel methods as they can not be objectively compared and evaluated.

Additionally, an increasing number of applications in multimodal MALDI-MSI offer a better understanding and new opportunities in biology, through enabling direct correlation and comparison between cellular and molecular information. However, it also brings challenges as it requires a diligent pipeline involving data acquisition, registration, pixel alignment, computational data analysis, result interpretation, and identification. Different pipelines were proposed, but current comparisons are either incomplete or using a limited set of benchmarking datasets, with inconsistent metrics in different studies, making it more difficult especially in the unsupervised learning based integration task where commonly the golden standards are unknown [236].

There is no doubt that a collaborative effort is required at all levels,

including by each scientist, research lab, publisher and funding organization towards the curation of high-quality, annotated, diverse data for the mass spectrometry imaging community to leverage the full potential of machine learning algorithms.

6.3 Future outlooks

Here we conclude this work by giving an overview of some important directions that will require more attention in the near future.

6.3.1 It is all about *good data*

Standardization of protocols and data processing

To address the aforementioned challenges, the adoption of standardized experimental protocols, quality control measurements are crucial for the development of a common language for technological process in MSI. Similarly, this is also called for in the field of spatial transcriptomics and high-plex spatial imaging [255, 250]. It will be incredibly important that each dataset generated be both usable and reusable. Additionally, some future efforts can be put in developing systematized pipelines for the data processing. For example, this is already happening in the field of image-based transcriptomics: *Starfish* [28].

(Meta)Data management and storage

Metadata are generated at multiple steps during the MSI experiments, so as other spatial biology techniques. Managing to record and store all of the experimental metadata properly is the key to promote interpretation, data reuse, reproduction and comparisons. Future effort should also be put on regulating and mandating proper metadata management and storage in line with the Findable, Accessible, Interoperable, and Reusable (FAIR standards [276]). With the volume of data increasing daily for each study, and multimodal strategies being more frequently applied, it gets extremely challenging to keep track on all the metadata manually. Additionally, the expanding of the datasets, which can easily take a few terabytes of disk space, poses a great challenge of data storage and processing. Thus, computational software tools should be developed for (meta)data management, storage, processing and sharing across the collaborating labs in more efficient manner.

High-quality benchmarking datasets

After the data being carefully acquired, processed, cleaned and documented by following the standardized and systematized pipelines, they can be effectively used by machine learning algorithms to yield meaningful results. In order to further improve the computational development, creating standardized high-quality benchmarking datasets is essential. Such datasets would require a well-defined ground truth, e.g., high-quality annotations by experts and prior knowledge supported by the literature. In the filed of single-cell genomics, some benchmarking studies have been done [152, 164, 124], and some benchmarking datasets have been created [240, 143]. Fortunately, in MSI, Alexandrov team at EMBL developed METASPACE, the largest collection of public MSI datasets [20]. It currently includes over 3500 public MSI datasets, and provides a free cloud engine for metabolite annotation. These efforts of pushing making standardized, high-quality, diverse data publicly available empowers to leverage the full potential of machine learning algorithms, and motivate more novel computational methods developments.

6.3.2 Development of computational tools for spatial multi-omics

Many studies have proved that the multimodal studies in MSI raise new questions and initiate new studies, which can not be achieved by a single modality [253]. The advent of cutting-edge spatial omics technologies enables to profile complementary modality information in the same tissue or serial tissue sections, such as RNA, protein. These technologies empower to analyze the correlation between various molecular modalities and provide mechanistic insights, and yet they also pose an additional important challenge of computational methods and software tools.

Computational tools in spatial multi-omics can be further improved by:

- incorporating the morphology information from microscopy data via advanced machine learning algorithms such as pre-trained neural networks;
- providing more user-friendly, efficient, smart annotation tools for the experts to annotate the data in a high-quality standard;
- building more comprehensive, flexible benchmarking pipelines for accommodating the increasing number of integration methods;
- supporting more interactive visualization toolboxes for better interpretation of the large results generated from the rich datasets;

Computational tools will be crucial for not only gaining biological insights in a more efficient and in-depth way, but also informing the design of spatial multi-omics experiments.

It is only the beginning...

It is clear that we are at the beginning of an exciting new era of spatial biology research. New technologies, new discoveries, new methodologies, new questions in medicine and biochemistry are being proposed and discussed everyday in this fascinating, challenging, and interdisciplinary field. As one of the key components in spatial multi-omics, MSI has not yet unlock its full potential due to the challenges in computational analysis and interpretation of generated complex and high-dimensional data. Once those challenges are addressed, with a collective effort at all levels, MSI and its multimodal applications will open novel avenues to understanding biology at a completely new level.

Appendix A

Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning

A.1 Relative isotope ratio (RIR) metric examples

We conducted some experiments on both data sets to show that the RIR metric can reasonably determine whether two ion images are an isotope by assessing their spectral and spatial expression. In the following examples using the lymph node and mouse kidney data, the RIR metric detects 3 and 2 ion images respectively as an isotope. Besides, we also show some example ion images from the same cluster, but they are not detected as isotope by the RIR metric.

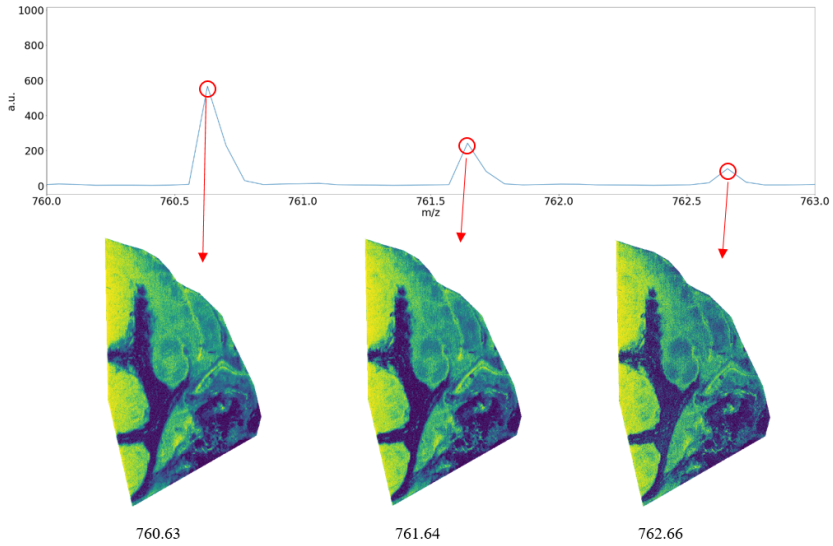


Figure A.1: Detected as isotope in lymph node data: (760.6285939640838, 761.6440294571262), (761.6440294571262, 762.6594649501685)

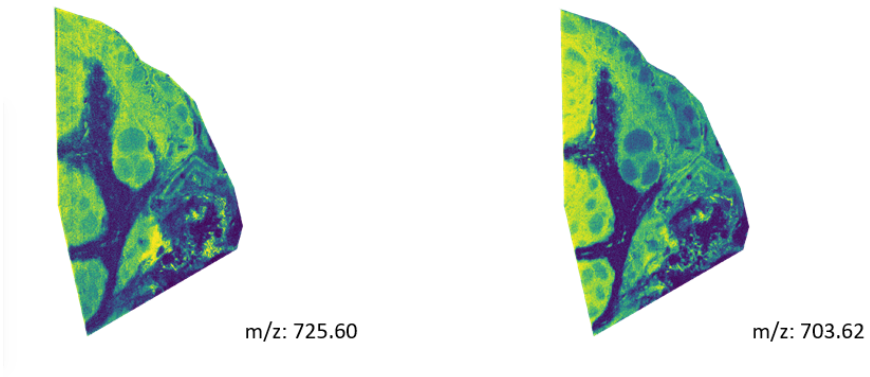


Figure A.2: Example ion images belongs to the same cluster but not the isotope in lymph node data

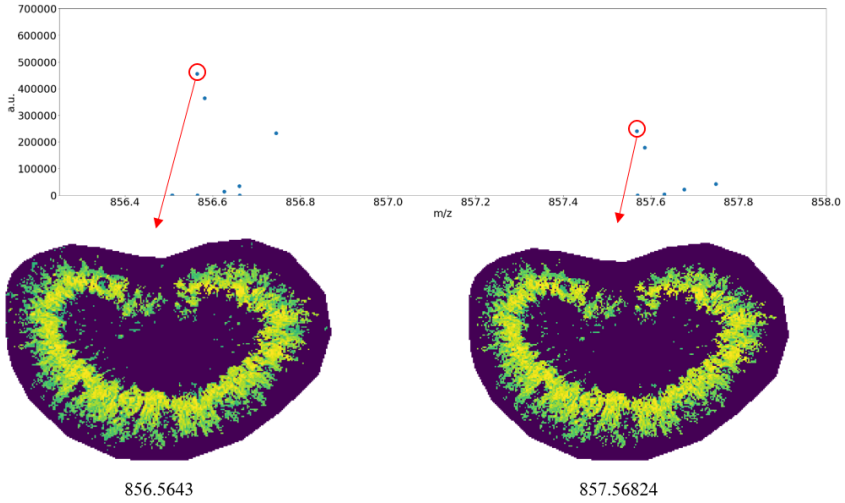


Figure A.3: Detect as isotope in Mouse kidney data: (856.5643, 857.56824)

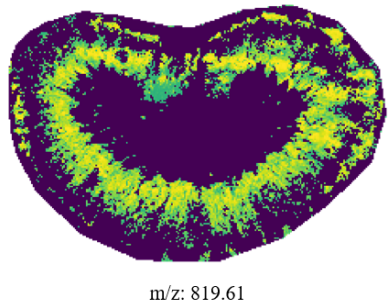


Figure A.4: Example ion images belongs to the same cluster but not the isotope in mouse kidney data

A.2 Used data sets from ColocML

Those are the dataset_Id of the data sets we used. Only those sets contain the spatial information and experts golden standards:

2017-01-26_09h08m49s

2017-07-18_17h21m08s (3 pairs of sets)

2018-09-14_22h35m17s

2018-11-28_14h47m01s

2018-12-04_14h23m49s

2018-12-12_12h51m09s

2018-12-18_12h33m17s

2019-01-01_01h09m28s

2019-01-04_00h50m49s

2019-01-04_01h25m24s

2019-01-04_20h32m07s

2019-01-05_02h22m18s

2019-01-17_19h54m30s (2 pairs of sets)

2019-01-21_22h44m45s

2019-01-22_03h04m17s

2019-01-23_00h34m56s

2019-01-25_11h22m48s

2019-01-25_11h24m31s (2 pairs of sets)

2019-01-25_11h26m55s

2019-03-12_08h57m03s

2019-03-12_10h14m51s

2019-03-19_17h21m15s (2 pairs of sets)

2019-03-22_10h16m54s

A.3 Examples from ColocML data



Figure A.5: datasetId = '2017-01-26_09h08m49s'

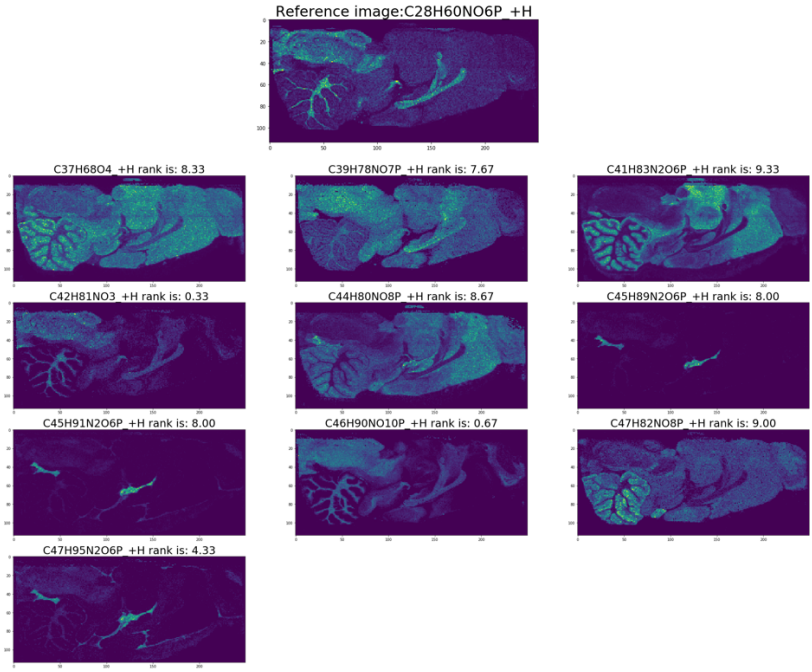


Figure A.6: datasetId = '2017-07-18_17h21m08s' set 1

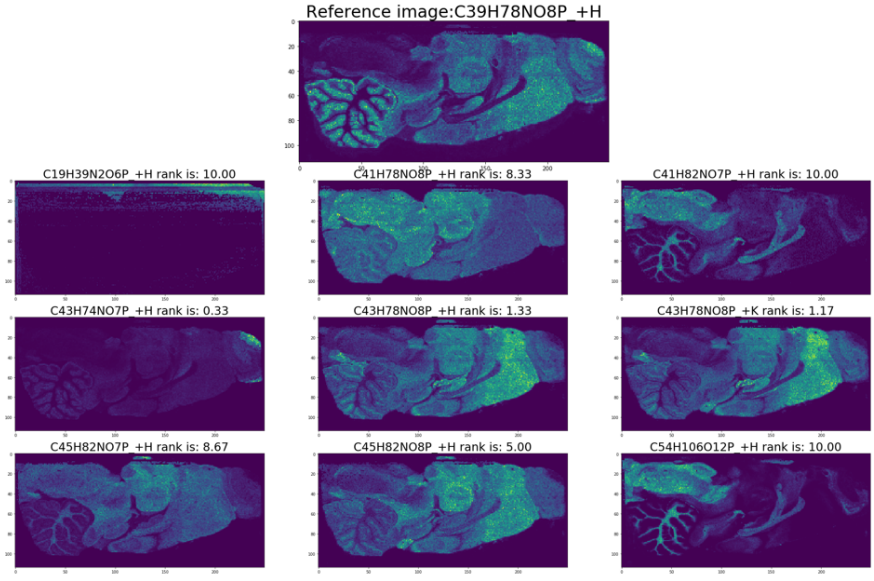


Figure A.7: datasetId = '2017-07-18_17h21m08s' set 2

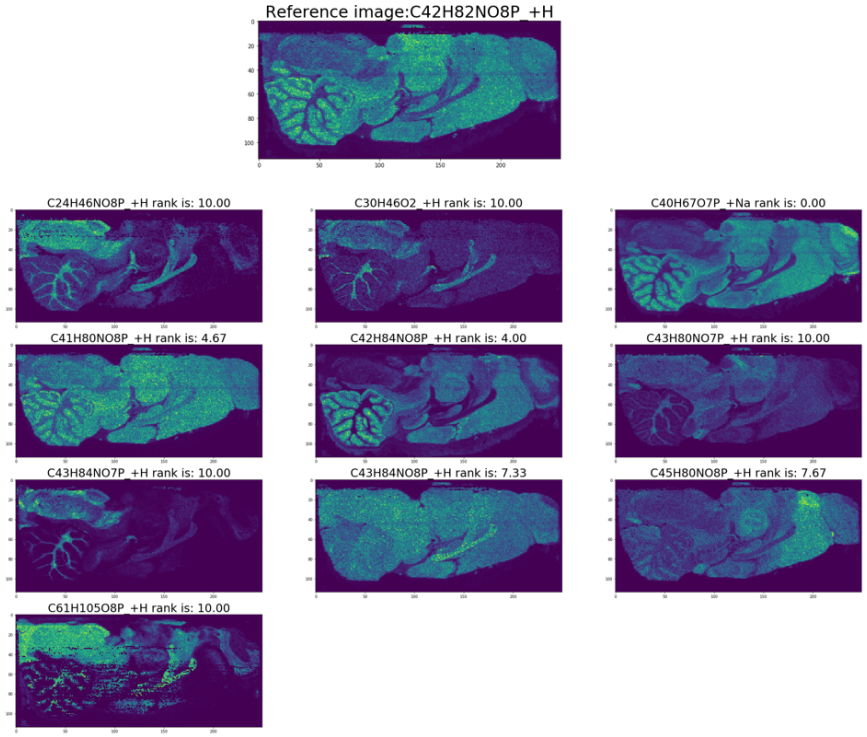


Figure A.8: datasetId = '2017-07-18_17h21m08s' set 3

Appendix B

Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer

B.1 Clinical data for the eight patients used in the SMOx study

Clinical data for the eight patients used in the SMOx study

Patient characteristics	Patient ID	929	931	934	941	943	947	950	952
	Age at surgery	73	61	72	64	72	62	72	69
	BMI	26,21	25,14	27,77	27,7	28,37	21,13	36,56	30,55
	date surgery	07/09/2020	17/09/2020	18/11/2020	08/02/2021	11/02/2021	04/03/2021	11/03/2021	13/04/2021
	Medication with an influence on lipids	1	0	1	1	1	0	0	0
	Which?	Crestor 10mg	NA	Crestor 20mg	Crestor 20mg + Ezetimibe 10mg	Atorvastatine 10mg	NA	NA	NA
Pre-OP characteristics	iPSA	8,7	16,2	19	14,6	4,81	8,04	24	3,79
	cT-stage	1	2b	1	3a	1	1	3a	2a
	Biopsy GS	4+4=8	4+4=8	4+4=8	4+3=7	3+4=7	3+5=8	4+5=9	3+4=7
	Biopsy ISUP stage	4	4	4	3	2	4	5	2
	Neoadjvuant therapy	0	0	0	0	0	0	0	0
Pathology of specimen	pT-stage	2c	3b	3a	3b	2c	2c	3a	3a
	Lymph node invasion	0	1	0	1	NA	0	0	0
	Lymphovascular invasion	0	1	1	0	0	1	0	0
	Perineural invasion	1	1	1	1	1	1	1	1
	Section margins	0	0	0	1	0	0	0	0
	pGS	3+4=7	4+5=9	4+3=7	4+5=9	4+3=7	3+4(+5)=7	4+3(+5)=7	4+3=7
	pISUP	2	5	3	5	3	2	3	3
Follow-up	Undetectable post-OP PSA	1	0	0	0	1	0	0	1
	Adj RT	0	1	1	1	0	1	0	0
	RT because of BCR or rising PSA	NA	0	1	0	NA	1	NA	NA
	Time from surgery till RT	NA	4,8	13,43	3,33	NA	26,73	NA	NA
	PSA nadir	NA	0	0	0	NA	NA	NA	NA
	Adj ADT	0	1	1	1	0	1	0	0
	Time from surgery till ADT	NA	4,8	13,43	3,33	NA	26,73	NA	NA
	BCR (PSA ≥ 0,2)	0	0	0	0	0	1	0	0
	Time from surgery till BCR	NA	NA	NA	NA	NA	24,7	NA	NA
	Clinical recurrence?	0	0	0	0	0	1	0	0
	Time from surgery till CR	NA	NA	NA	NA	NA	24	NA	NA
	Last PSA	0	0,15	0	0	0	0,23	0,1	0
	Date last PSA	26/04/2023	25/04/2023	31/05/2023	09/01/2023	13/03/2023	24/03/2023	04/04/2023	22/03/2023
	Follow-up (months)	31,67	31,3	30,43	23,07	25,1	26,73	24,8	23,33
	Alive with no evidence of disease	1	0	1	1	1	0	0	1
	Remarks?	/	No T level	Testosterone 0,133 nmol/L	Testosterone 12 ng/dL	/	RT still going	PSA rising	/

Figure B.1: Clinical data for the eight patients used in the SMOx study

B.2 Averaged MALDI-2 spectra

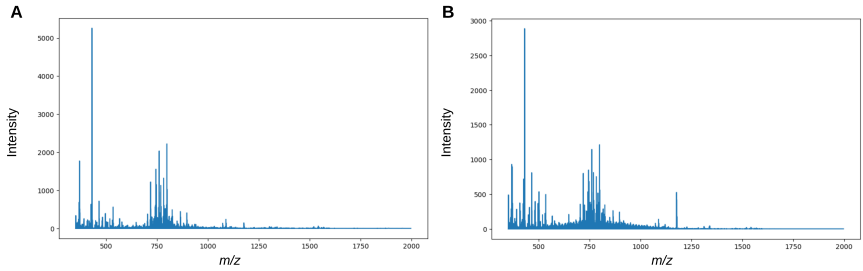


Figure B.2: Averaged MALDI-2 spectra. A. Averaged MALDI-2 spectrum of all cancer samples. B. Averaged MALDI-2 spectrum of all matching non-malignant samples.

B.3 Correlation heatmaps comparing pathology annotations versus spatial-omics expression data

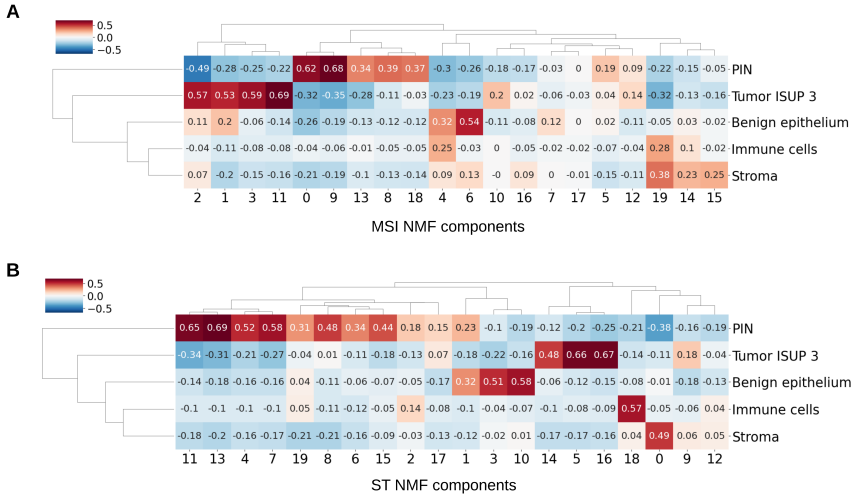


Figure B.3: Correlation heatmaps comparing pathology annotations versus spatial omics expression data. A. Correlation between MSI NMF components and the five histologically distinct regions from sample 929_cancer. B. Correlation between ST NMF components and the five histologically distinct regions from sample 929_cancer.

B.4 Pathological annotations of all samples

In this study we have collected a cohort of 16 matched prostate tumor and normal samples from 8 high-risk PCa patients. Two sets of pathological annotations were collected from two pathology groups. Two sets of annotations for all 16 samples by two groups of pathologists are shown and compared. We received JSON files from annotation tools that pathologists used, and transformed their coordinates from the microscopy data to the shared coordinate system of spatial multi-omics data. Annotations were made on both MSI-slide microscopy and ST-slide microscopy. High-quality annotations represent that more than one pathologists were annotating the whole slide in detail, taking more than 1 hour per sample. Standard-quality annotations represent that pathologist were annotating the whole slide in a quick glance, taking around 5 minutes per sample.

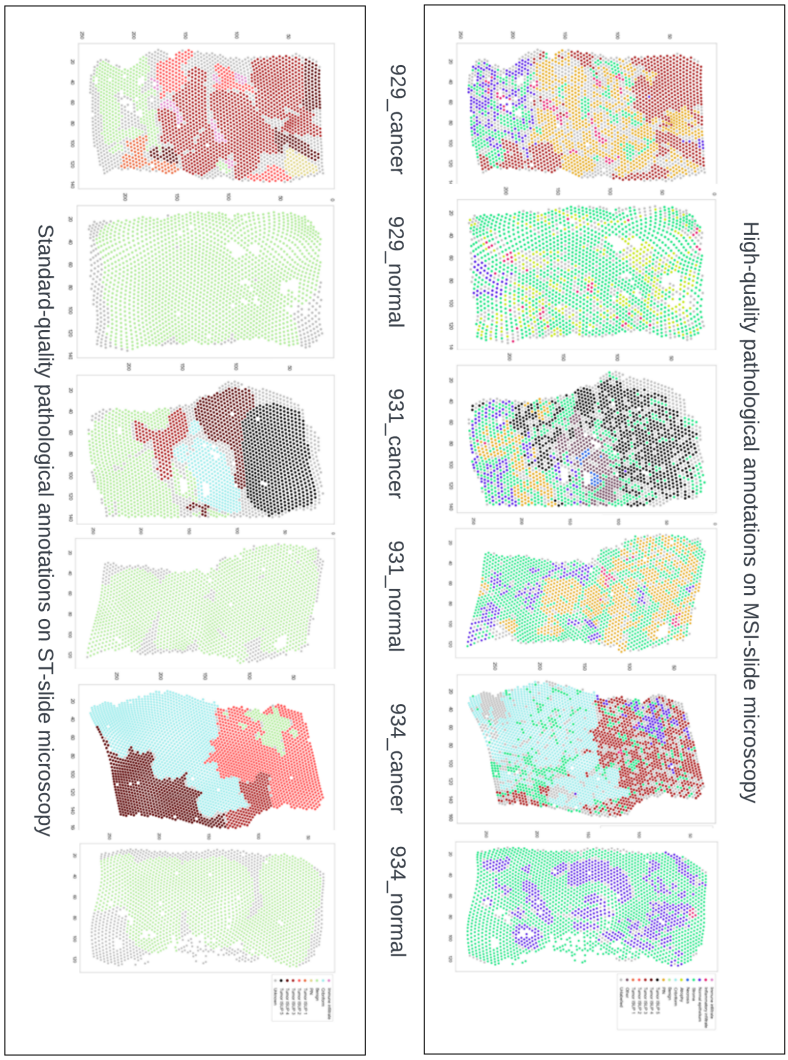


Figure B.4: Comparison of high-quality and standard quality of pathological annotations.

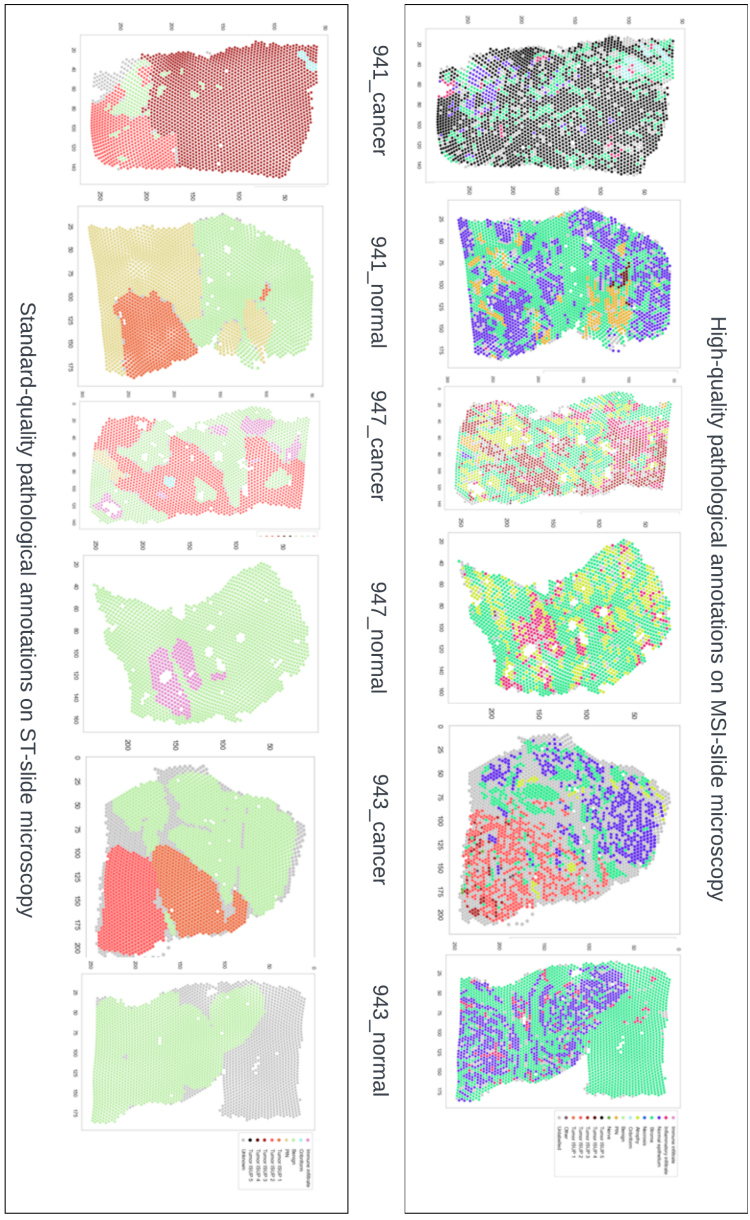


Figure B.5: Comparison of high-quality and standard quality of pathological annotations.

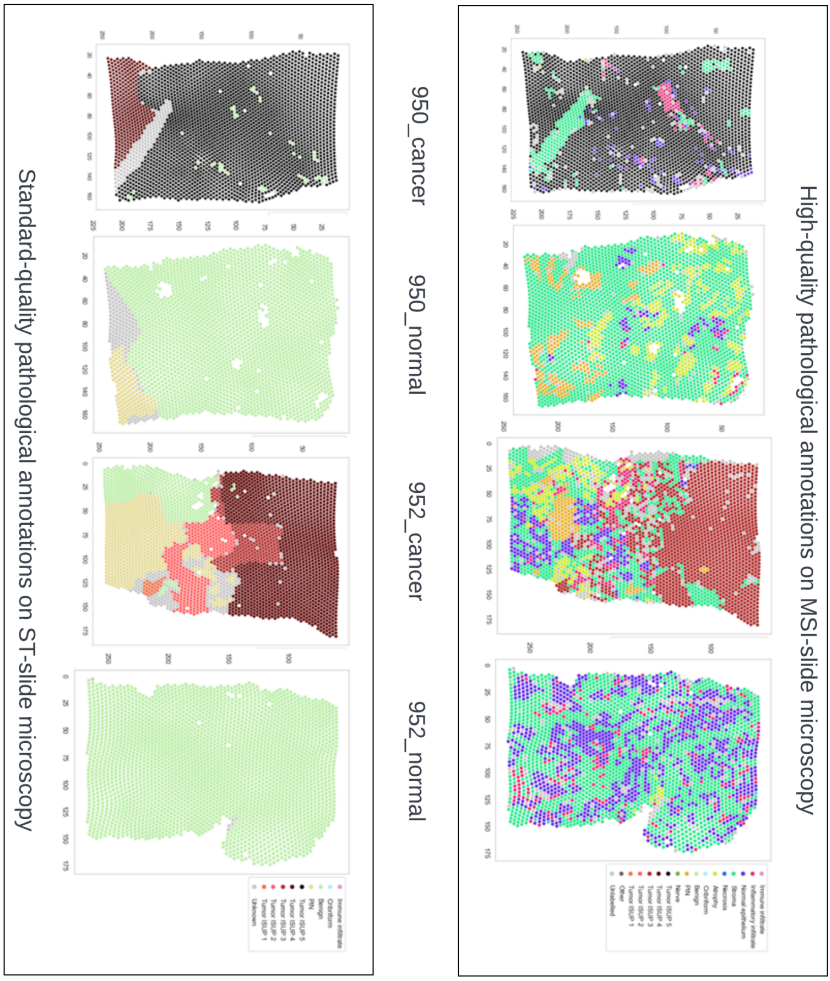


Figure B.6: Comparison of high-quality and standard quality of pathological annotations.

Bibliography

[1] https://en.wikipedia.org/wiki/Tandem_mass_spectrometry. [Online; accessed August-2023].

[2] https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html.

[3] https://www.10xgenomics.com/products/spatial-gene-expression. [Online; accessed August-2023].

[4] https://support.10xgenomics.com/spatial-gene-expression/software/pipelines/latest/what-is-space-ranger. [Online; accessed August-2023].

[5] https://pair-code.github.io/understanding-umap/. [Online; accessed August-2023].

[6] https://umap-learn.readthedocs.io/en/latest/how_umap_works.html. [Online; accessed August-2023].

[7] Abbassi-Ghadi, N., Jones, E. A., Veselkov, K. A., Huang, J., Kumar, S., Strittmatter, N., Golf, O., Kudo, H., Goldin, R. D., Hanna, G. B., et al. Repeatability and reproducibility of desorption electrospray ionization-mass spectrometry (desi-ms) for the imaging analysis of human cancer tissue: a gateway for clinical applications. *Analytical Methods 7*, 1 (2015), 71–80.

[8] Abdelmoula, W. M., Carreira, R. J., Shyti, R., Balluff, B., van Zeijl, R. J., Tolner, E. A., Lelieveldt, B. F., van den Maagdenberg, A. M., McDonnell, L. A., and Dijkstra, J. Automatic registration of mass spectrometry imaging data sets to the allen brain atlas. *Anal. Chem. 86*, 8 (2014), 3947–3954.

[9] Abdelmoula, W. M., Lopez, B. G.-C., Randall, E. C., Kapur, T., Sarkaria, J. N., White, F. M., Agar, J. N., Wells, W. M., and Agar, N. Y. Peak learning of mass spectrometry imaging data using artificial neural networks. *Nature communications 12*, 1 (2021), 5544.

[10] Abdelmoula, W. M., Regan, M. S., Lopez, B. G., Randall, E. C., Lawler, S., Mladek, A. C., Nowicki, M. O., Marin, B. M., Agar, J. N., Swanson, K. R., et al. Automatic 3d nonlinear registration of mass spectrometry imaging and magnetic resonance imaging data. *Analytical chemistry 91*, 9 (2019), 6206–6216.

[11] Addie, R. D., Balluff, B., Bovée, J. V., Morreau, H., and McDonnell, L. A. Current state and future challenges of mass spectrometry imaging for clinical research, 2015.

[12] Al-Rohil, R. N., Moore, J. L., Patterson, N. H., Nicholson, S., Verbeeck, N., Claesen, M., Muhammad, J. Z., Caprioli, R. M., Norris, J. L., Kantrow, S., et al. Diagnosis of melanoma by imaging mass spectrometry: Development and validation of a melanoma prediction model. *Journal of Cutaneous Pathology 48*, 12 (2021), 1455–1462.

[13] Alasoo, K., Martinez, F. O., Hale, C., Gordon, S., Powrie, F., Dougan, G., Mukhopadhyay, S., and Gaffney, D. J. Transcriptional profiling of macrophages derived from monocytes and ips cells identifies a conserved response to lps and novel alternative transcription. *Scientific reports 5*, 1 (2015), 12524.

[14] Alexandrov, T. Maldi imaging mass spectrometry: statistical data analysis and current computational challenges. *BMC bioinformatics 13*, Suppl 16 (2012), S11.

[15] Alexandrov, T. Spatial metabolomics and imaging mass spectrometry in the age of artificial intelligence. *Annual review of biomedical data science 3* (2020), 61–87.

[16] Alexandrov, T., Becker, M., Deininger, S.-o., Ernst, G., Wehder, L., Grasmair, M., Von Eggeling, F., Thiele, H., and Maass, P. Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering. *Journal of proteome research 9*, 12 (2010), 6535–6546.

[17] Alexandrov, T., Chernyavsky, I., Becker, M., von Eggeling, F., and Nikolenko, S. Analysis and interpretation of imaging mass spectrometry data by clustering mass-to-charge images according to their spatial similarity. *Analytical chemistry 85*, 23 (2013), 11189–11195.

[18] Alexandrov, T., and Kobarg, J. H. Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering. *Bioinformatics 27*, 13 (2011), i230–i238.

[19] Alexandrov, T., and Kobarg, J. H. Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering. *Bioinformatics 27*, 13 (06 2011), i230–i238.

[20] Alexandrov, T., Ovchinnikova, K., Palmer, A., Kovalev, V., Tarasov, A., Stuart, L., Nigmetzianov, R., Fay, D., contributors, K. M., Gaudin, M., et al. Metaspace: A community-populated knowledge base of spatial metabolomes in health and disease. *BioRxiv* (2019), 539478.

[21] Alomari, A. K., Glusac, E. J., Choi, J., Hui, P., Seeley, E. H., Caprioli, R. M., Watsky, K. L., Urban, J., and Lazova, R. Congenital nevi versus metastatic melanoma in a newborn to a mother with malignant melanoma–diagnosis supported by sex chromosome analysis and imaging mass spectrometry. *Journal of Cutaneous Pathology 42*, 10 (2015), 757–764.

[22] Anderson, D. M., Messinger, J. D., Patterson, N. H., Rivera, E. S., Kotnala, A., Spraggins, J. M., Caprioli, R. M., Curcio, C. A., and Schey, K. L. Lipid landscape of the human retina and supporting tissues revealed by high-resolution imaging mass spectrometry. *Journal of the American Society for Mass Spectrometry 31*, 12 (2020), 2426–2436.

[23] Andersson, A., Bergenstråhle, J., Asp, M., Bergenstråhle, L., Jurek, A., Fernández Navarro, J., and Lundeberg, J. Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Communications biology 3*, 1 (2020), 565.

[24] Andersson, A., Larsson, L., Stenbeck, L., Salmén, F., Ehinger, A., Wu, S. Z., Al-Eryani, G., Roden, D., Swarbrick, A., Borg, Å., et al. Spatial deconvolution of her2-positive breast cancer delineates tumor-associated cell type interactions. *Nature communications 12*, 1 (2021), 6012.

[25] Arsić, A., Hagemann, C., Stajković, N., Schubert, T., and Nikić-Spiegel, I. Minimal genetically encoded tags for fluorescent protein labeling in living neurons. *Nature Communications 13*, 1 (2022), 314.

[26] Arthur, D., and Vassilvitskii, S. k-means++: The advantages of careful seeding. Tech. rep., Stanford, 2006.

[27] Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell 179*, 7 (2019), 1647–1660.

[28] Axelrod, S., Cai, M., Carr, A. J., Freeman, J., Ganguli, D., Kiggins, J. T., Long, B., Tung, T., and Yamauchi, K. A. starfish: scalable pipelines for image-based transcriptomics. *Journal of Open Source Software 6*, 61 (2021), 2440.

[29] Balluff, B., Hopf, C., Porta Siegel, T., Grabsch, H. I., and Heeren, R. M. A. Batch effects in maldi mass spectrometry imaging. *Journal of the American Society for Mass Spectrometry 32*, 3 (3 2021), 628–635.

[30] Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W., Ng, L. G., Ginhoux, F., and Newell, E. W. Dimensionality reduction for visualizing single-cell data using umap. *Nature biotechnology 37*, 1 (2019), 38–44.

[31] Behrmann, J., Etmann, C., Boskamp, T., Casadonte, R., Kriegsmann, J., and Maass, P. Deep learning for tumor classification in imaging mass spectrometry. *Bioinformatics 34*, 7 (11 2017), 1215–1223.

[32] Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J. A., Hermsen, M., Manson, Q. F., Balkenhol, M., et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama 318*, 22 (2017), 2199–2210.

[33] Bellman, R. Dynamic programming and lagrange multipliers. *PNAS 42*, 10 (1956), 767–769.

[34] Belov, M. E., Ellis, S. R., Dilillo, M., Paine, M. R., Danielson, W. F., Anderson, G. A., de Graaf, E. L., Eijkel, G. B., Heeren, R. M., and McDonnell, L. A. Design and performance of a novel interface for combined matrix-assisted laser desorption ionization at elevated pressure and electrospray ionization with orbitrap mass spectrometry. *Analytical chemistry 89*, 14 (2017), 7493–7501.

[35] Bemis, K. A., Guo, D., Harry, A. J., Thomas, M., Lanekoff, I., Stenzel-Poore, M. P., Stevens, S. L., Laskin, J., and Vitek, O. Statistical detection of differentially abundant ions in mass spectrometry-based imaging experiments with complex designs. *International Journal of Mass Spectrometry 437* (2019), 49–57.

[36] Bemis, K. D., Harry, A., Eberlin, L. S., Ferreira, C. R., van de Ven, S. M., Mallick, P., Stolowitz, M., and Vitek, O. Probabilistic segmentation of mass spectrometry (ms) images helps select important ions and characterize confidence in the resulting segments. *Mol. Cell. Proteomics 15*, 5 (2016), 1761–1772.

[37] Bengio, Y. Deep learning of representations for unsupervised and transfer learning. In *Proceedings of ICML workshop on unsupervised and transfer learning* (2012), JMLR Workshop and Conference Proceedings, pp. 17–36.

[38] Bergenstråhle, J., Larsson, L., and Lundeberg, J. Seamless integration of image and molecular analysis for spatial transcriptomics workflows. *BMC genomics 21*, 1 (2020), 1–7.

[39] Berglund, E., Maaskola, J., Schultz, N., Friedrich, S., Marklund, M., Bergenstråhle, J., Tarish, F., Tanoglidi, A., Vickovic, S., Larsson, L., et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nature communications 9*, 1 (2018), 2419.

[40] Bergström, S. H., Järemo, H., Nilsson, M., Adamo, H. H., and Bergh, A. Prostate tumors downregulate microseminoprotein-beta (msmb) in the surrounding benign prostate epithelium and this response is associated with tumor aggressiveness. *The Prostate 78*, 4 (3 2018), 257–265. PMID: 29250809.

[41] Berkhin, P. A survey of clustering data mining techniques. In *Grouping multidimensional data*. Springer, 2006, pp. 25–71.

[42] Bianco, S., Cadene, R., Celona, L., and Napoletano, P. Benchmark analysis of representative deep neural network architectures. *IEEE Access 6* (2018), 64270–64277.

[43] Blanc, L., Daudelin, I. B., Podell, B. K., Chen, P.-Y., Zimmerman, M., Martinot, A. J., Savic, R. M., Prideaux, B., and Dartois, V. High-resolution mapping of fluoroquinolones in tb rabbit lesions reveals specific distribution in immune cell types. *Elife 7* (2018), e41115.

[44] Blanc, L., Lenaerts, A., Dartois, V., and Prideaux, B. Visualization of mycobacterial biomarkers and tuberculosis drugs in infected tissue by maldi-ms imaging. *Analytical chemistry 90*, 10 (2018), 6275–6282.

[45] Borodinov, N., Lorenz, M., King, S. T., Ievlev, A. V., and Ovchinnikova, O. S. Toward nanoscale molecular mass spectrometry imaging via physically constrained machine learning on co-registered multimodal data. *npj Computational Materials 6*, 1 (2020), 83.

[46] Bowman, A. P., Bogie, J. F., Hendriks, J. J., Haidar, M., Belov, M., Heeren, R. M., and Ellis, S. R. Evaluation of lipid coverage and high spatial resolution maldi-imaging capabilities of oversampling combined with laser post-ionisation. *Analytical and bioanalytical chemistry 412* (2020), 2277–2289.

[47] Brown, L. G. A survey of image registration techniques. *ACM computing surveys (CSUR) 24*, 4 (1992), 325–376.

[48] Brulet, M., Seyer, A., Edelman, A., Brunelle, A., Fritsch, J., Ollero, M., and Laprévote, O. Lipid mapping of colonic mucosa by cluster tof-sims imaging and multivariate analysis in cftr knockout mice [s]. *Journal of lipid research 51*, 10 (2010), 3034–3045.

[49] Bu, H., Bormann, S., Schäfer, G., Horninger, W., Massoner, P., Neeb, A., Lakshmanan, V.-K., Maddalo, D., Nestl, A., Sültmann, H., et al. The anterior gradient 2 (agr2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. *The Prostate 71*, 6 (2011), 575–587.

[50] Buchberger, A. R., DeLaney, K., Johnson, J., and Li, L. Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal. Chem. 90*, 1 (Dec. 2017), 240–265.

[51] Butler, L. M., Mah, C. Y., Machiels, J., Vincent, A. D., Irani, S., Mutuku, S. M., Spotbeen, X., Bagadi, M., Waltregny, D., Moldovan, M., et al. Lipidomic profiling of clinical prostate cancer reveals targetable alterations in membrane lipid composition. *Cancer research 81*, 19 (2021), 4981–4993.

[52] Cai, C., Chen, S., Ng, P., Bubley, G. J., Nelson, P. S., Mostaghel, E. A., Marck, B., Matsumoto, A. M., Simon, N. I., Wang, H., et al. Intratumoral de novo steroid synthesis activates androgen receptor in castration-resistant prostate cancer and is upregulated by treatment with cyp17a1 inhibitors. *Cancer research 71*, 20 (2011), 6503–6513.

[53] Campello, R. J., Moulavi, D., and Sander, J. Density-based clustering based on hierarchical density estimates. In *Pacific-Asia conference on knowledge discovery and data mining* (2013), Springer, pp. 160–172.

[54] Caprioli, R. M., Farmer, T. B., and Gile, J. Molecular imaging of biological samples: localization of peptides and proteins using maldi-tof ms. *Analytical chemistry 69*, 23 (1997), 4751–4760.

[55] Carm, K. T., Hoff, A. M., Bakken, A. C., Axcrona, U., Axcrona, K., Lothe, R. A., Skotheim, R. I., and Løvf, M. Interfocal heterogeneity challenges the clinical usefulness of molecular classification of primary prostate cancer. *Scientific reports 9*, 1 (2019), 13579.

[56] Casadonte, R., Kriegsmann, M., Kriegsmann, K., Hauk, I., Meliss, R. R., Müller, C. S., and Kriegsmann, J. Imaging mass spectrometry-based proteomic analysis to differentiate melanocytic nevi and malignant melanoma. *Cancers 13*, 13 (2021), 3197.

[57] Cassese, A., Ellis, S. R., Ogrinc Potocnik, N., Burgermeister, E., Ebert, M., Walch, A., Van Den Maagdenberg, A. M., McDonnell, L. A., Heeren, R. M., and Balluff, B. Spatial autocorrelation in mass spectrometry imaging. *Analytical chemistry 88*, 11 (2016), 5871–5878.

[58] Castellino, S., Lareau, N. M., and Groseclose, M. R. The emergence of imaging mass spectrometry in drug discovery and development: Making a difference by driving decision making. *Journal of Mass Spectrometry 56*, 8 (2021), e4717.

[59] Chari, T., and Pachter, L. The specious art of single-cell genomics. *PLOS Computational Biology 19*, 8 (2023), e1011288.

[60] Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using dna nanoball-patterned arrays. *Cell 185*, 10 (2022), 1777–1792.

[61] Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., and Zhuang, X. Spatially resolved, highly multiplexed rna profiling in single cells. *Science 348*, 6233 (2015), aaa6090.

[62] Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In *International conference on machine learning* (2020), PMLR, pp. 1597–1607.

[63] Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Frigerio, C. S., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I., et al. Spatial transcriptomics and in situ sequencing to study alzheimer's disease. *Cell 182*, 4 (2020), 976–991.

[64] Chollet, F. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 1251–1258.

[65] Ciga, O., Xu, T., and Martel, A. L. Self supervised contrastive learning for digital histopathology. *Machine Learning with Applications 7* (2022), 100198.

[66] Cooks, R. G., Ouyang, Z., Takats, Z., and Wiseman, J. M. Ambient mass spectrometry. *Science 311*, 5767 (2006), 1566–1570.

[67] Cornett, D. S., Mobley, J. A., Dias, E. C., Andersson, M., Arteaga, C. L., Sanders, M. E., and Caprioli, R. M. A novel histology-directed strategy for maldi-ms tissue profiling that improves throughput and cellular specificity in human breast cancer. *Molecular & Cellular Proteomics 5*, 10 (2006), 1975–1983.

[68] Cortes, C., and Vapnik, V. Support-vector networks. *Machine learning 20*, 3 (1995), 273–297.

[69] Crum, W. R., Hartkens, T., and Hill, D. Non-rigid image registration: theory and practice. *The British journal of radiology 77*, suppl_2 (2004), S140–S153.

[70] Dahlman, A., Rexhepaj, E., Brennan, D. J., Gallagher, W. M., Gaber, A., Lindgren, A., Jirström, K., and Bjartell, A. Evaluation of the prognostic significance of msmb and crisp3 in prostate cancer using automated image analysis. *Modern pathology 24*, 5 (2011), 708–719.

[71] de Macedo, C. S., Anderson, D. M., and Schey, K. L. Maldi (matrix assisted laser desorption ionization) imaging mass spectrometry (ims) of skin: aspects of sample preparation. *Talanta 174* (2017), 325–335.

[72] Deininger, S.-O., Becker, M., and Suckau, D. Tutorial: multivariate statistical treatment of imaging data for clinical biomarker discovery. *Mass spectrometry imaging: Principles and protocols* (2010), 385–403.

[73] Deininger, S.-O., Bollwein, C., Casadonte, R., Wandernoth, P., Gonçalves, J. P. L., Kriegsmann, K., Kriegsmann, M., Boskamp, T., Kriegsmann, J., Weichert, W., et al. Multicenter evaluation of tissue classification by matrix-assisted laser desorption/ionization mass spectrometry imaging. *Analytical Chemistry* (2022).

[74] Deininger, S.-O., Cornett, D. S., Paape, R., Becker, M., Pineau, C., Rauser, S., Walch, A., and Wolski, E. Normalization in maldi-tof imaging datasets of proteins: practical considerations. *Analytical and bioanalytical chemistry 401* (2011), 167–181.

[75] Deininger, S.-O., Ebert, M. P., Futterer, A., Gerhard, M., and Rocken, C. Maldi imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *Journal of proteome research 7*, 12 (2008), 5230–5236.

[76] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (2009), Ieee, pp. 248–255.

[77] Dettmer, K., Aronov, P. A., and Hammock, B. D. Mass spectrometry-based metabolomics. *Mass spectrometry reviews 26*, 1 (2007), 51–78.

[78] Dewez, F., Oejten, J., Henkel, C., Hebeler, R., Neuweger, H., De Pauw, E., Heeren, R. M., and Balluff, B. Ms imaging-guided microproteomics for spatial omics on a single instrument. *Proteomics 20*, 23 (2020), 1900369.

[79] Dodds, J. N., and Baker, E. S. Ion mobility spectrometry: fundamental concepts, instrumentation, applications, and the road ahead. *Journal of the American Society for Mass Spectrometry 30*, 11 (2019), 2185–2195.

[80] Doerr, A. Mass spectrometry imaging takes off. *Nature Methods 15*, 1 (2018), 32–32.

[81] Donahue, J., Krähenbühl, P., and Darrell, T. Adversarial feature learning. *arXiv preprint arXiv:1605.09782* (2016).

[82] Eastman, R. D., Le Moigne, J., and Netanyahu, N. S. Research issues in image registration for remote sensing. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (2007), IEEE, pp. 1–8.

[83] Editorial. Why the metabolism field risks missing out on the ai revolution. *Nature Metabolism 1* (2019), 929–930.

[84] Eisenstein, M. Seven technologies to watch in 2022. *Nature 601*, 7894 (2022), 658–661.

[85] El Ayed, M., Bonnel, D., Longuespée, R., Castellier, C., Franck, J., Vergara, D., Desmons, A., Tasiemski, A., Kenani, A., Vinatier, D., et al. Maldi imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers. *Medical science monitor: international medical journal of experimental and clinical research 16*, 8 (2010).

[86] Ellis, S., Soltwisch, J., Paine, M., Dreisewerd, K., and Heeren, R. Laser post-ionisation combined with a high resolving power orbitrap mass spectrometer for enhanced maldi-ms imaging of lipids. *Chemical Communications 53*, 53 (2017), 7246–7249.

[87] Ellis, S. R., Paine, M. R., Eijkel, G. B., Pauling, J. K., Husen, P., Jervelund, M. W., Hermansson, M., Ejsing, C. S., and Heeren, R. M. Automated, parallel mass spectrometry imaging and structural identification of lipids. *Nature methods 15*, 7 (2018), 515–518.

[88] Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I., and Heyn, H. Spotlight: seeded nmf regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic acids research 49*, 9 (2021), e50–e50.

[89] Eng, C.-H. L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., et al. Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. *Nature 568*, 7751 (2019), 235–239.

[90] Erin H Seeley, Edwin Escobar, J. S. B. Maximizing data coverage with sequential imaging of a single tissue section. In *Proceedings of the 71st ASMS Conference on Mass Spectrometry and Allied Topics* (Houston, TX, June 4-8, 2023), no. 315490.

[91] Erlmeier, F., Feuchtinger, A., Borgmann, D., Rudelius, M., Autenrieth, M., Walch, A. K., and Weirich, G. Supremacy of modern morphometry in typing renal oncocytoma and malignant look-alikes. *Histochemistry and cell biology 144*, 2 (2015), 147–156.

[92] Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* (1996), vol. 96, pp. 226–231.

[93] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science 246*, 4926 (1989), 64–71.

[94] Fletcher, J. S., Kotze, H. L., Armitage, E. G., Lockyer, N. P., and Vickerman, J. C. Evaluating the challenges associated with time-of-fight secondary ion mass spectrometry for metabolomics using pure and mixed metabolites. *Metabolomics 9* (2013), 535–544.

[95] Fonville, J. M., Carter, C., Cloarec, O., Nicholson, J. K., Lindon, J. C., Bunch, J., and Holmes, E. Robust data processing and normalization strategy for maldi mass spectrometric imaging. *Analytical chemistry 84*, 3 (2012), 1310–1319.

[96] Fonville, J. M., Carter, C. L., Pizarro, L., Steven, R. T., Palmer, A. D., Griffiths, R. L., Lalor, P. F., Lindon, J. C., Nicholson, J. K., Holmes, E., et al. Hyperspectral visualization of mass spectrometry imaging data. *Analytical chemistry 85*, 3 (2013), 1415–1423.

[97] Franzén, O., Gan, L.-M., and Björkegren, J. L. Panglaodb: a web server for exploration of mouse and human single-cell rna sequencing data. *Database 2019* (2019), baz046.

[98] Galli, M., Zoppis, I., De Sio, G., Chinello, C., Pagni, F., Magni, F., and Mauri, G. A support vector machine classification of thyroid bioptic specimens using maldi-msi data. *Advances in bioinformatics 2016* (2016).

[99] Gawin, M., Kurczyk, A., Niemiec, J., Stanek-Widera, A., Grela-Wojewoda, A., Adamczyk, A., Biskup-Frużyńska, M., Polańska, J., and Widłak, P. Intra-tumor heterogeneity revealed by mass spectrometry imaging is associated with the prognosis of breast cancer. *Cancers 13*, 17 (2021), 4349.

[100] Gerami, P., Jewell, S. S., Morrison, L. E., Blondin, B., Schulz, J., Ruffalo, T., Matushek IV, P., Legator, M., Jacobson, K., Dalton, S. R., et al. Fluorescence in situ hybridization (fish) as an ancillary diagnostic tool in the diagnosis of melanoma. *The American journal of surgical pathology 33*, 8 (2009), 1146–1156.

[101] Germain, P.-L., Lun, A., Meixide, C. G., Macnair, W., and Robinson, M. D. Doublet identification in single-cell sequencing data using scdblfinder. *F1000Research 10* (2021).

[102] Gonçalves, J. P., Bollwein, C., and Schwamborn, K. Mass spectrometry imaging spatial tissue analysis toward personalized medicine. *Life 12*, 7 (2022), 1037.

[103] Gonçalves, J. P. L., Bollwein, C., Schlitter, A. M., Martin, B., Märkl, B., Utpatel, K., Weichert, W., and Schwamborn, K. The impact of histological annotations for accurate tissue classification using mass spectrometry imaging. *Metabolites 11*, 11 (2021), 752.

[104] Gonçalves, J. P. L., Bollwein, C., Weichert, W., and Schwamborn, K. Implementation of mass spectrometry imaging in pathology: Advances and challenges. *Clinics in Laboratory Medicine 41*, 2 (2021), 173–184.

[105] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in neural information processing systems* (2014), pp. 2672–2680.

[106] Goossens, P., Lu, C., Cao, J., Gijbels, M. J., Karel, J. M., Wijnands, E., Claes, B. S., Fazzi, G. E., Hendriks, T. F., Wouters, K., et al. Integrating multiplex immunofluorescent and mass spectrometry imaging to map myeloid heterogeneity in its metabolic and cellular context. *Cell Metabolism 34*, 8 (2022), 1214–1225.

[107] Gorin, G., Fang, M., Chari, T., and Pachter, L. Rna velocity unraveled. *PLOS Computational Biology 18*, 9 (2022), e1010492.

[108] Guo, Y., Ren, C., Huang, W., Yang, W., and Bao, Y. Oncogenic acsm1 in prostate cancer is through metabolic and extracellular matrix-receptor interaction signaling pathways. *American Journal of Cancer Research 12*, 4 (2022), 1824.

[109] Guran, R., Vanickova, L., Horak, V., Krizkova, S., Michalek, P., Heger, Z., Zitka, O., and Adam, V. Maldi msi of melim melanoma: Searching for differences in protein profiles. *PLoS One 12*, 12 (2017), e0189305.

[110] Hanrieder, J., Phan, N. T. N., Kurczy, M. E., and Ewing, A. G. Imaging Mass Spectrometry in Neuroscience. *ACS Chem. Neurosci. 4*, 5 (2013), 666–79.

[111] Hanselmann, M., Kothe, U., Kirchner, M., Renard, B. Y., Amstalden, E. R., Glunde, K., Heeren, R. M., and Hamprecht, F. A. Toward digital staining using imaging mass spectrometry and random forests. *Journal of proteome research 8*, 7 (2009), 3558–3567.

[112] Hanselmann, M., Köthe, U., Renard, B. Y., Kirchner, M., Heeren, R. M., and Hamprecht, F. A. Multivariate watershed segmentation of compositional data. In *Discrete Geometry for Computer*

Imagery: 15th IAPR International Conference, DGCI 2009, Montréal, Canada, September 30-October 2, 2009. Proceedings 15 (2009), Springer, pp. 180–192.

[113] Hardesty, W. M., Kelley, M. C., Mi, D., Low, R. L., and Caprioli, R. M. Protein signatures for survival and recurrence in metastatic melanoma. *Journal of proteomics 74*, 7 (2011), 1002–1014.

[114] Hastings, C., Mosteller, F., Tukey, J. W., Winsor, C. P., et al. Low moments for small samples: a comparative study of order statistics. *Ann. Math. Stat. 18*, 3 (1947), 413–426.

[115] He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 770–778.

[116] Hekler, A., Utikal, J. S., Enk, A. H., Solass, W., Schmitt, M., Klode, J., Schadendorf, D., Sondermann, W., Franklin, C., Bestvater, F., et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur. J. Cancer 118* (2019), 91–96.

[117] Henry, G. H., Malewska, A., Joseph, D. B., Malladi, V. S., Lee, J., Torrealba, J., Mauck, R. J., Gahan, J. C., Raj, G. V., Roehrborn, C. G., et al. A cellular anatomy of the normal adult human prostate and prostatic urethra. *Cell reports 25*, 12 (2018), 3530–3542.

[118] Heumos, L., Schaar, A. C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M. D., Strobl, D. C., Henao, J., Curion, F., et al. Best practices for single-cell analysis across modalities. *Nature Reviews Genetics* (2023), 1–23.

[119] Hickey, J. W., Neumann, E. K., Radtke, A. J., Camarillo, J. M., Beuschel, R. T., Albanese, A., McDonough, E., Hatler, J., Wiblin, A. E., Fisher, J., et al. Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nature methods 19*, 3 (2022), 284–295.

[120] Hirz, T., Mei, S., Sarkar, H., Kfoury, Y., Wu, S., Verhoeven, B. M., Subtelny, A. O., Zlatev, D. V., Wszolek, M. W., Salari, K., et al. Dissecting the immune suppressive human prostate tumor microenvironment via integrated single-cell and spatial transcriptomic analyses. *Nature Communications 14*, 1 (2023), 663.

[121] Hotelling, H. Relations between two sets of variates. In *Breakthroughs in statistics: methodology and distribution*. Springer, 1992, pp. 162–190.

[122] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* (2017).

[123] Hu, H., Bindu, J. P., and Laskin, J. Self-supervised clustering of mass spectrometry imaging data using contrastive learning. *Chemical Science 13*, 1 (2022), 90–98.

[124] Huang, Q., Liu, Y., Du, Y., and Garmire, L. X. Evaluation of cell type annotation r packages on single-cell rna-seq data. *Genomics, proteomics & bioinformatics 19*, 2 (2021), 267–281.

[125] Inglese, P., McKenzie, J. S., Mroz, A., Kinross, J., Veselkov, K., Holmes, E., Takats, Z., Nicholson, J. K., and Glen, R. C. Deep learning and 3d-desi imaging reveal the hidden metabolic heterogeneity of cancer. *Chem. Sci. 8* (2017), 3500–3511.

[126] Ji, A. L., Rubin, A. J., Thrane, K., Jiang, S., Reynolds, D. L., Meyers, R. M., Guo, M. G., George, B. M., Mollbrink, A., Bergenstråhle, J., et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell 182*, 2 (2020), 497–514.

[127] Jolliffe, I. T. Principal component analysis, 2nd edn new york. *NY: Springer.[Google Scholar]* (2002).

[128] Jones, A., Townes, F. W., Li, D., and Engelhardt, B. E. Alignment of spatial genomics data using deep gaussian processes. *Nature Methods* (2023), 1–9.

[129] Jones, E. A., Deininger, S.-O., Hogendoorn, P. C., Deelder, A. M., and McDonnell, L. A. Imaging mass spectrometry statistical analysis. *Journal of proteomics 75*, 16 (2012), 4962–4989.

[130] Jones, E. A., van Remoortere, A., van Zeijl, R. J., Hogendoorn, P. C., Bovée, J. V., Deelder, A. M., and McDonnell, L. A. Multiple statistical analysis techniques corroborate intratumor heterogeneity in imaging mass spectrometry datasets of myxofibrosarcoma. *PloS one 6*, 9 (2011), e24913.

[131] Karas, M., and Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry 60*, 20 (1988), 2299–2301.

[132] Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A., and Deters, R. Breast cancer diagnosis with transfer learning and global pooling. *arXiv preprint arXiv:1909.11839* (2019).

[133] Kaya, I., Michno, W., Brinet, D., Iacone, Y., Zanni, G., Blennow, K., Zetterberg, H., and Hanrieder, J. Histology-compatible maldi mass spectrometry based imaging of neuronal lipids for subsequent immunofluorescent staining. *Analytical chemistry 89*, 8 (2017), 4685–4694.

[134] Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., and Nilsson, M. In situ sequencing for rna analysis in preserved tissue and cells. *Nature methods 10*, 9 (2013), 857–860.

[135] Klein, O., Kanter, F., Kulbe, H., Jank, P., Denkert, C., Nebrich, G., Schmitt, W. D., Wu, Z., Kunze, C. A., Sehouli, J., et al. Maldi-imaging for classification of epithelial ovarian cancer histotypes from a tissue microarray using machine learning methods. *PROTEOMICS–Clinical Applications 13*, 1 (2019), 1700181.

[136] Kleshchevnikov, V., Shmatko, A., Dann, E., Aivazidis, A., King, H. W., Li, T., Elmentaite, R., Lomakin, A., Kedlian, V., Gayoso, A., et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nature biotechnology 40*, 5 (2022), 661–671.

[137] Kobak, D., and Berens, P. The art of using t-sne for single-cell transcriptomics. *Nature communications 10*, 1 (2019), 5416.

[138] Köfeler, H. C., Ahrends, R., Baker, E. S., Ekroos, K., Han, X., Hoffmann, N., Holčapek, M., Wenk, M. R., and Liebisch, G. Recommendations for good practice in ms-based lipidomics. *Journal of lipid research 62* (2021).

[139] Kompauer, M., Heiles, S., and Spengler, B. Atmospheric pressure maldi mass spectrometry imaging of tissues and cells at 1.4-μm lateral resolution. *Nature methods 14*, 1 (2017), 90–96.

[140] Konicek, A. R., Lefman, J., and Szakal, C. Automated correlation and classification of secondary ion mass spectrometry images using ak-means cluster method. *Analyst 137*, 15 (2012), 3479–3487.

[141] Kornblith, S., Shlens, J., and Le, Q. V. Do better imagenet models transfer better? In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2019), pp. 2661–2671.

[142] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (2012), pp. 1097–1105.

[143] Lance, C., Luecken, M. D., Burkhardt, D. B., Cannoodt, R., Rautenstrauch, P., Laddach, A. C., Ubingazhibov, A., Cao, Z.-J., Deng, K., Khan, S., et al. Multimodal single cell data integration challenge: results and lessons learned. *bioRxiv* (2022), 2022–04.

[144] Lanekoff, I., Stevens, S. L., Stenzel-Poore, M. P., and Laskin, J. Matrix effects in biological mass spectrometry imaging: identification and compensation. *Analyst 139*, 14 (2014), 3528–3532.

[145] Laskin, J., Heath, B. S., Roach, P. J., Cazares, L., and Semmes, O. J. Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. *Analytical chemistry 84*, 1 (2012), 141–148.

[146] Lauwerends, L., Abbasi, H., Bakker Schut, T., Van Driel, P., Hardillo, J., Santos, I., Barroso, E., Koljenović, S., Vahrmeijer, A., Baatenburg de Jong, R., et al. The complementary value of intraoperative fluorescence imaging and raman spectroscopy for cancer surgery: combining the incompatibles. *European Journal of Nuclear Medicine and Molecular Imaging 49*, 7 (2022), 2364–2376.

[147] Lazova, R., and Seeley, E. H. Proteomic mass spectrometry imaging for skin cancer diagnosis. *Dermatologic Clinics 35*, 4 (2017), 513–519.

[148] Lazova, R., Smoot, K., Anderson, H., Powell, M. J., Rosenberg, A. S., Rongioletti, F., Pilloni, L., D'Hallewin, S., Gueorguieva, R., Tantcheva-Poór, I., et al. Histopathology-guided mass spectrometry differentiates benign nevi from malignant melanoma. *Journal of Cutaneous Pathology 47*, 3 (2020), 226–240.

[149] Lee, S., Devanney, N. A., Golden, L. R., Smith, C. T., Schwartz, J. L., Walsh, A. E., Clarke, H. A., Goulding, D. S., Allenger, E. J., Morillo-Segovia, G., et al. Apoe modulates microglial immunometabolism in response to age, amyloid pathology, and inflammatory challenge. *Cell reports 42*, 3 (2023).

[150] Leiter, U., Keim, U., and Garbe, C. Epidemiology of skin cancer: update 2019. *Sunlight, Vitamin D and Skin Cancer* (2020), 123–139.

[151] Lewis, S. M., Asselin-Labat, M.-L., Nguyen, Q., Berthelet, J., Tan, X., Wimmer, V. C., Merino, D., Rogers, K. L., and Naik, S. H. Spatial omics and multiplexed imaging to explore cancer biology. *Nature methods 18*, 9 (2021), 997–1012.

[152] Li, Y., Stanojevic, S., He, B., Jing, Z., Huang, Q., Kang, J., and Garmire, L. X. Benchmarking computational integration methods for spatial transcriptomics data. *bioRxiv* (2021), 2021–08.

[153] Li, Y., Yang, M., and Zhang, Z. A survey of multi-view representation learning. *IEEE transactions on knowledge and data engineering 31*, 10 (2018), 1863–1883.

[154] Liebisch, G., Fahy, E., Aoki, J., Dennis, E. A., Durand, T., Ejsing, C. S., Fedorova, M., Feussner, I., Griffiths, W. J., Köfeler, H., et al. Update on lipid maps classification, nomenclature, and shorthand notation for ms-derived lipid structures. *Journal of lipid research 61*, 12 (2020), 1539–1555.

[155] Liebisch, G., Vizcaíno, J. A., Köfeler, H., Trötzmüller, M., Griffiths, W. J., Schmitz, G., Spener, F., and Wakelam, M. J. Shorthand notation for lipid structures derived from mass spectrometry. *Journal of lipid research 54*, 6 (2013), 1523–1530.

[156] Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., and Tang, J. Self-supervised learning: Generative or contrastive. *IEEE Transactions on Knowledge and Data Engineering* (2021).

[157] Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C. C., Tebaldi, T., Zhang, D., Kim, D., Bai, Z., et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell 183*, 6 (2020), 1665–1681.

[158] Long, M., Zhu, H., Wang, J., and Jordan, M. I. Unsupervised domain adaptation with residual transfer networks. In *Advances in neural information processing systems* (2016), pp. 136–144.

[159] Lopez, R., Regier, J., Cole, M. B., Jordan, M. I., and Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nature methods 15*, 12 (2018), 1053–1058.

[160] Lotz, J. M., Hoffmann, F., Lotz, J., Heldmann, S., Trede, D., Oetjen, J., Becker, M., Ernst, G., Maas, P., Alexandrov, T., et al. Integration of 3d multimodal imaging data of a head and neck cancer and advanced feature recognition. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1865*, 7 (2017), 946–956.

[161] Lou, S., Balluff, B., Cleven, A. H., Bovée, J. V., and McDonnell, L. A. Prognostic metabolite biomarkers for soft tissue sarcomas discovered by mass spectrometry imaging. *Journal of the American Society for Mass Spectrometry 28*, 2 (2016), 376–383.

[162] Lou, S., Balluff, B., Cleven, A. H., Bovée, J. V., and McDonnell, L. A. An experimental guideline for the analysis of histologically heterogeneous tumors by maldi-tof mass spectrometry imaging. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1865*, 7 (2017), 957–966.

[163] Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M., and Cai, L. Single-cell in situ rna profiling by sequential hybridization. *Nature Methods 11*, 4 (4 2014), 360–361. number: 4 publisher: Nature Publishing Group.

[164] Luecken, M. D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Müller, M. F., Strobl, D. C., Zappia, L., Dugas, M., Colomé-Tatché, M., et al. Benchmarking atlas-level data integration in single-cell genomics. *Nature methods 19*, 1 (2022), 41–50.

[165] Lundberg, E., and Borner, G. H. H. Spatial proteomics: a powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology 20*, 5 (5 2019), 285–302. number: 5 publisher: Nature Publishing Group.

[166] Lundberg, S., and Lee, S.-I. An unexpected unity among methods for interpreting model predictions. *arXiv preprint arXiv:1611.07478* (2016).

[167] Lundberg, S. M., and Lee, S.-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems 30* (2017).

[168] Luts, J., Ojeda, F., Van de Plas, R., De Moor, B., Van Huffel, S., and Suykens, J. A. A tutorial on support vector machine-based methods for classification problems in chemometrics. *Analytica chimica acta 665*, 2 (2010), 129–145.

[169] Ma, R.-Y., Black, A., and Qian, B.-Z. Macrophage diversity in cancer revisited in the era of single-cell omics. *Trends in Immunology 43*, 7 (7 2022), 546–563. PMID: 35690521.

[170] Ma, X., and Fernández, F. M. Advances in mass spectrometry imaging for spatial cancer metabolomics. *Mass Spectrometry Reviews* (2022), e21804.

[171] Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass spectrometry reviews 17*, 1 (1998), 1–35.

[172] Martin, B., Gonçalves, J. P., Bollwein, C., Sommer, F., Schenkirsch, G., Jacob, A., Seibert, A., Weichert, W., Märkl, B., and Schwamborn, K. A mass spectrometry imaging based approach for prognosis prediction in uicc stage i/ii colon cancer. *Cancers 13*, 21 (2021), 5371.

[173] Marx, V. Method of the year: spatially resolved transcriptomics. *Nature methods 18*, 1 (2021), 9–14.

[174] Masood, M., Grimm, S., El-Bahrawy, M., and Yagüe, E. Tmeff2: A transmembrane proteoglycan with multifaceted actions in cancer and disease. *Cancers 12*, 12 (12 2020). publisher: Multidisciplinary Digital Publishing Institute (MDPI) PMID: 33371267.

[175] Masyuko, R., Lanni, E. J., Sweedler, J. V., and Bohn, P. W. Correlated imaging–a grand challenge in chemical analysis. *Analyst 138*, 7 (2013), 1924–1939.

[176] Matsumoto, K.-i., Mitchell, J. B., and Krishna, M. C. Multimodal functional imaging for cancer/tumor microenvironments based on mri, epri, and pet. *Molecules 26*, 6 (2021), 1614.

[177] McCombie, G., Staab, D., Stoeckli, M., and Knochenmuss, R. Spatial and spectral correlations in maldi mass spectrometry images by clustering and multivariate analysis. *Analytical chemistry 77*, 19 (2005), 6118–6124.

[178] McCombie, G., Staab, D., Stoeckli, M., and Knochenmuss, R. Spatial and spectral correlations in maldi mass spectrometry images by clustering and multivariate analysis. *Anal. Chem. 77*, 19 (2005), 6118–6124.

[179] McDonnell, L. A., van Remoortere, A., van Zeijl, R. J., and Deelder, A. M. Mass spectrometry image correlation: quantifying colocalization. *Journal of proteome research 7*, 8 (2008), 3619–3627.

[180] McDonnell, L. A., van Remoortere, A., van Zeijl, R. J., and Deelder, A. M. Mass spectrometry image correlation: quantifying colocalization. *J. Proteome Res. 7*, 8 (2008), 3619–3627.

[181] McInnes, L., Healy, J., and Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

[182] McInnes, L., Healy, J., Saul, N., and Grossberger, L. Umap: Uniform manifold approximation and projection. *J. Open Source Softw. 3*, 29 (2018), 861.

[183] Meding, S., Nitsche, U., Balluff, B., Elsner, M., Rauser, S., Schone, C., Nipp, M., Maak, M., Feith, M., Ebert, M. P., et al. Tumor classification of six common cancer types based on proteomic profiling by maldi imaging. *Journal of proteome research 11*, 3 (2012), 1996–2003.

[184] Melo Ferreira, R., Freije, B. J., and Eadon, M. T. Deconvolution tactics and normalization in renal spatial transcriptomics. *Frontiers in Physiology 12* (2022), 812947.

[185] Merritt, C. R., Ong, G. T., Church, S. E., Barker, K., Danaher, P., Geiss, G., Hoang, M., Jung, J., Liang, Y., McKay-Fleisch, J., et al. Multiplex digital spatial profiling of proteins and rna in fixed tissue. *Nature biotechnology 38*, 5 (2020), 586–599.

[186] Minoura, K., Abe, K., Nam, H., Nishikawa, H., and Shimamura, T. A mixture-of-experts deep generative model for integrated analysis of single-cell multiomics data. *Cell reports methods 1*, 5 (2021).

[187] Mitchell, T. M., and Mitchell, T. M. *Machine learning*, vol. 1. McGraw-hill New York, 1997.

[188] Mund, A., Coscia, F., Kriston, A., Hollandi, R., Kovács, F., Brunner, A.-D., Migh, E., Schweizer, L., Santos, A., Bzorek, M., et al. Deep visual proteomics defines single-cell identity and heterogeneity. *Nature Biotechnology 40*, 8 (2022), 1231–1240.

[189] Mutuku, S. M., Spotbeen, X., Trim, P. J., Snel, M. F., Butler, L. M., and Swinnen, J. V. Unravelling prostate cancer heterogeneity using spatial approaches to lipidomics and transcriptomics. *Cancers 14*, 7 (3 2022), 1702. PMID: 35406474 PMCID: PMC8997139.

[190] Nemes, P., and Vertes, A. Laser ablation electrospray ionization for atmospheric pressure, in vivo, and imaging mass spectrometry. *Analytical chemistry 79*, 21 (2007), 8098–8106.

[191] Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., and Ng, A. Y. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (2011), pp. 689–696.

[192] Ni, Z., Prasad, A., Chen, S., Halberg, R. B., Arkin, L. M., Drolet, B. A., Newton, M. A., and Kendziorski, C. Spotclean adjusts for spot swapping in spatial transcriptomics data. *Nature Communications 13*, 1 (5 2022), 2971. number: 1 publisher: Nature Publishing Group.

[193] Nijs, M., Smets, T., Waelkens, E., and De Moor, B. A mathematical comparison of non-negative matrix factorization related methods with practical implications for the analysis of mass spectrometry imaging data. *Rapid communications in mass spectrometry: RCM 35*, 21 (11 2021), e9181. PMID: 34374141 PMCID: PMC9285509.

[194] Norris, J. L., Tsui, T., Gutierrez, D. B., and Caprioli, R. M. Pathology interface for the molecular analysis of tissue by mass spectrometry. *Journal of Pathology Informatics 7*, 1 (2016), 13.

[195] Oppenheimer, S. R., Mi, D., Sanders, M. E., and Caprioli, R. M. Molecular analysis of tumor margins by maldi mass spectrometry in renal carcinoma. *Journal of proteome research 9*, 5 (2010), 2182–2190.

[196] Oquab, M., Bottou, L., Laptev, I., and Sivic, J. Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2014), pp. 1717–1724.

[197] Ovchinnikova, K., Stuart, L., Rakhlin, A., Nikolenko, S., and Alexandrov, T. Colocml: machine learning quantifies co-localization between mass spectrometry images. *Bioinformatics 36*, 10 (2020), 3215–3224.

[198] Palla, G., Fischer, D. S., Regev, A., and Theis, F. J. Spatial components of molecular tissue biology. *Nature Biotechnology 40*, 3 (2022), 308–318.

[199] Palmer, A. D., Bunch, J., and Styles, I. B. The use of random projections for the analysis of mass spectrometry imaging data. *Journal of The American Society for Mass Spectrometry 26*, 2 (2014), 315–322.

[200] Palmer, A. D., Bunch, J., and Styles, I. B. The use of random projections for the analysis of mass spectrometry imaging data. *J. Am. Soc. Mass Spectrom. 26*, 2 (2014), 315–322.

[201] Passarelli, M. K., Pirkl, A., Moellers, R., Grinfeld, D., Kollmer, F., Havelund, R., Newman, C. F., Marshall, P. S., Arlinghaus, H., Alexander, M. R., et al. The 3d orbisims—label-free metabolic imaging with subcellular lateral resolution and high mass-resolving power. *Nature methods 14*, 12 (2017), 1175–1183.

[202] Patel, A. P., Tirosh, I., Trombetta, J. J., Shalek, A. K., Gillespie, S. M., Wakimoto, H., Cahill, D. P., Nahed, B. V., Curry, W. T., Martuza, R. L., et al. Single-cell rna-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science 344*, 6190 (2014), 1396–1401.

[203] Patterson, N. H., Tuck, M., Lewis, A., Kaushansky, A., Norris, J. L., Van de Plas, R., and Caprioli, R. M. Next generation histology-directed imaging mass spectrometry driven by autofluorescence microscopy. *Analytical chemistry 90*, 21 (2018), 12404–12413.

[204] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res. 12* (2011), 2825–2830.

[205] Pike, J. A., Styles, I. B., Rappoport, J. Z., and Heath, J. K. Quantifying receptor trafficking and colocalization with confocal microscopy. *Methods 115* (2017), 42–54.

[206] Porta Siegel, T., Hamm, G., Bunch, J., Cappell, J., Fletcher, J. S., and Schwamborn, K. Mass spectrometry imaging and integration with other imaging modalities for greater molecular understanding of biological tissues. *Molecular Imaging and Biology 20*, 6 (2018), 888–901.

[207] Prade, V. M., Sun, N., Shen, J., Feuchtinger, A., Kunzke, T., Buck, A., Schraml, P., Moch, H., Schwamborn, K., Autenrieth, M., et al. The synergism of spatial metabolomics and morphometry improves machine learning-based renal tumour subtype classification. *Clinical and translational medicine 12*, 2 (2022).

[208] Qi, X., Wang, Y., Hou, J., and Huang, Y. A single nucleotide polymorphism in *HPGD* gene is associated with prostate cancer risk. *Journal of Cancer 8*, 19 (10 2017), 4083–4086. publisher: Ivyspring International Publisher.

[209] Race, A. M., Palmer, A. D., Dexter, A., Steven, R. T., Styles, I. B., and Bunch, J. Spectralanalysis: software for the masses. *Analytical chemistry 88*, 19 (2016), 9451–9458.

[210] Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434* (2015).

[211] Rajaraman, S., Antani, S. K., Poostchi, M., Silamut, K., Hossain, M. A., Maude, R. J., Jaeger, S., and Thoma, G. R. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ 6* (2018), e4568.

[212] Rao, A., Barkley, D., França, G. S., and Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature 596*, 7871 (2021), 211–220.

[213] Ravi, V. M., Will, P., Kueckelhaus, J., Sun, N., Joseph, K., Salié, H., Vollmer, L., Kuliesiute, U., von Ehr, J., Benotmane, J. K., et al. Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell 40*, 6 (2022), 639–655.

[214] Reyzer, M. L., Caldwell, R. L., Dugger, T. C., Forbes, J. T., Ritter, C. A., Guix, M., Arteaga, C. L., and Caprioli, R. M. Early changes in protein expression detected by mass spectrometry predict tumor response to molecular therapeutics. *Cancer research 64*, 24 (2004), 9093–9100.

[215] Roca, H., Jones, J. D., Purica, M. C., Weidner, S., Koh, A. J., Kuo, R., Wilkinson, J. E., Wang, Y., Daignault-Newton, S., Pienta, K. J., et al. Apoptosis-induced cxcl5 accelerates inflammation and growth of prostate tumor metastases in bone. *The Journal of clinical investigation 128*, 1 (2018), 248–266.

[216] Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., Welch, J., Chen, L. M., Chen, F., and Macosko, E. Z. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (New York, N.Y.) 363*, 6434 (3 2019), 1463–1467. PMID: 30923225 PMCID: PMC6927209.

[217] Rubin, M. A., Zhou, M., Dhanasekaran, S. M., Varambally, S., Barrette, T. R., Sanda, M. G., Pienta, K. J., Ghosh, D., and Chinnaiyan, A. M. α-methylacyl coenzyme a racemase as a tissue biomarker for prostate cancer. *Jama 287*, 13 (2002), 1662–1670.

[218] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., and Hawkes, D. J. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging 18*, 8 (1999), 712–721.

[219] Ryabchykov, O., Popp, J., and Bocklitz, T. Fusion of maldi spectrometric imaging and raman spectroscopic data for the analysis of biological samples. *Frontiers in Chemistry 6* (2018), 257.

[220] Salmén, F., Ståhl, P. L., Mollbrink, A., Navarro, J. F., Vickovic, S., Frisen, J., and Lundeberg, J. Barcoded solid-phase rna capture for spatial transcriptomics profiling in mammalian tissue sections. *Nature protocols 13*, 11 (2018), 2501–2534.

[221] Sarretto, T., Spotbeen, X., Gevaert, T., Joniau, S., Swinnen, J. V., Trevitt, A. J., and Ellis, S. R. Selective mass spectrometry

imaging of aromatic antioxidants using sequential matrix-assisted laser desorption and resonant photoionisation. *Analysis & Sensing 2*, 1 (2022), e202100052.

[222] Sawilowsky, S. S. New effect size rules of thumb. *Journal of modern applied statistical methods 8*, 2 (2009), 26.

[223] Schulz, S., Becker, M., Groseclose, M. R., Schadt, S., and Hopf, C. Advanced maldi mass spectrometry imaging in pharmaceutical research and drug development. *Current opinion in biotechnology 55* (2019), 51–59.

[224] Schulz, S., Becker, M., Groseclose, M. R., Schadt, S., and Hopf, C. Advanced MALDI mass spectrometry imaging in pharmaceutical research and drug development, 2019.

[225] Schwamborn, K. The importance of histology and pathology in mass spectrometry imaging. *Advances in cancer research 134* (2017), 1–26.

[226] Schwamborn, K., and Caprioli, R. M. Molecular imaging by mass spectrometry—looking beyond classical histology. *Nature Reviews Cancer 10*, 9 (2010), 639–646.

[227] Schwamborn, K., Kriegsmann, M., and Weichert, W. MALDI imaging mass spectrometry — From bench to bedside, jul 2017.

[228] Shrestha, R., Nassar, Z. D., Hanson, A. R., Iggo, R., Townley, S. L., Dehairs, J., Mah, C. Y., Helm, M., Alizadeh-Ghodsi, M., Pickering, M., et al. Acsm1 and acsm3 regulate prostate cancer fatty acid metabolism to promote tumour growth and constrain ferroptosis. *bioRxiv* (2022), 2022–10.

[229] Smets, T., Verbeeck, N., Claesen, M., Asperger, A., Griffioen, G., Tousseyn, T., Waelput, W., Waelkens, E., and De Moor, B. Evaluation of distance metrics and spatial autocorrelation in uniform manifold approximation and projection applied to mass spectrometry imaging data. *Analytical chemistry 91*, 9 (2019), 5706–5714.

[230] Smets, T., Waelkens, E., and De Moor, B. Prioritization of m/z-values in mass spectrometry imaging profiles obtained using uniform manifold approximation and projection for dimensionality reduction. *Analytical chemistry 92*, 7 (2020), 5240–5248.

[231] Smets, T., Waelkens, E., and De Moor, B. Prioritization of m/z-values in mass spectrometry imaging profiles obtained using uniform manifold approximation and projection for dimensionality reduction. *Anal. Chem. 92*, 7 (April 2020), 5240—5248.

[232] Sodhi, R. N. Time-of-flight secondary ion mass spectrometry (tof-sims):—versatility in chemical and imaging surface analysis. *Analyst 129*, 6 (2004), 483–487.

[233] Soltwisch, J., Kettling, H., Vens-Cappell, S., Wiegelmann, M., Müthing, J., and Dreisewerd, K. Mass spectrometry imaging with laser-induced postionization. *Science 348*, 6231 (4 2015), 211–215. publisher: American Association for the Advancement of Science.

[234] Song, H., Weinstein, H. N., Allegakoen, P., Wadsworth, M. H., Xie, J., Yang, H., Castro, E. A., Lu, K. L., Stohr, B. A., Feng, F. Y., et al. Single-cell analysis of human primary prostate cancer reveals the heterogeneity of tumor-associated epithelial cell states. *Nature communications 13*, 1 (2022), 141.

[235] Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science 353*, 6294 (2016), 78–82.

[236] Stanojevic, S., Li, Y., Ristivojevic, A., and Garmire, L. X. Computational methods for single-cell multi-omics integration and alignment. *Genomics, Proteomics & Bioinformatics* (2022).

[237] Suits, F., Fehniger, T. E., Végvári, A., Marko-Varga, G., and Horvatovich, P. Correlation queries for mass spectrometry imaging. *Anal. Chem. 85*, 9 (2013), 4398–4404.

[238] Sun, C., Wang, A., Zhou, Y., Chen, P., Wang, X., Huang, J., Gao, J., Wang, X., Shu, L., Lu, J., Dai, W., Bu, Z., Ji, J., and He, J. Spatially resolved multi-omics highlights cell-specific metabolic remodeling and interactions in gastric cancer. *Nature Communications 14*, 1 (5 2023), 2692. number: 1 publisher: Nature Publishing Group.

[239] Sutton, G. J., Poppe, D., Simmons, R. K., Walsh, K., Nawaz, U., Lister, R., Gagnon-Bartsch, J. A., and Voineagu, I. Comprehensive evaluation of deconvolution methods for human brain gene expression. *Nature Communications 13*, 1 (3 2022), 1358. number: 1 publisher: Nature Publishing Group.

[240] Swechha, Mendonca, D., Focsa, O., Díaz-Mejía, J. J., and Cooper, S. scmark an 'mnist'like benchmark to evaluate and optimize models for unifying scrna data. *bioRxiv* (2021), 2021–12.

[241] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. Going deeper

with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 1–9.

[242] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. A survey on deep transfer learning. In *International conference on artificial neural networks* (2018), Springer, pp. 270–279.

[243] Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid communications in mass spectrometry 2*, 8 (1988), 151–153.

[244] Tata, A., Woolman, M., Ventura, M., Bernards, N., Ganguly, M., Gribble, A., Shrestha, B., Bluemke, E., Ginsberg, H. J., Vitkin, A., et al. Rapid detection of necrosis in breast cancer with desorption electrospray ionization mass spectrometry. *Sci. Rep. 6* (2016), 35374.

[245] Taverna, D., Boraldi, F., De Santis, G., Caprioli, R. M., and Quaglino, D. Histology-directed and imaging mass spectrometry: An emerging technology in ectopic calcification. *Bone 74* (2015), 83–94.

[246] Taverna, D., Nanney, L. B., Pollins, A. C., Sindona, G., and Caprioli, R. Spatial mapping by imaging mass spectrometry offers advancements for rapid definition of human skin proteomic signatures. *Experimental dermatology 20*, 8 (2011), 642–647.

[247] Tellez, D., Litjens, G., van der Laak, J., and Ciompi, F. Neural image compression for gigapixel histopathology image analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* (2019).

[248] Thomas, S. A., Race, A. M., Steven, R. T., Gilmore, I. S., and Bunch, J. Dimensionality reduction of mass spectrometry imaging data using autoencoders. In *2016 IEEE symposium series on computational intelligence (SSCI)* (2016), IEEE, pp. 1–7.

[249] Thomas, S. A., Race, A. M., Steven, R. T., Gilmore, I. S., and Bunch, J. Dimensionality reduction of mass spectrometry imaging data using autoencoders. *IEEE SSCI 2016* (2016), 1–7.

[250] Tian, L., Chen, F., and Macosko, E. Z. The expanding vistas of spatial transcriptomics. *Nature Biotechnology* (2022), 1–10.

[251] Tideman, L. E., Migas, L. G., Djambazova, K. V., Patterson, N. H., Caprioli, R. M., Spraggins, J. M., and Van de Plas, R. Automated biomarker candidate discovery in imaging mass spectrometry

data through spatially localized shapley additive explanations. *Analytica Chimica Acta 1177* (2021), 338522.

[252] Tuck, M., Blanc, L., Touti, R., Patterson, N. H., Van Nuffel, S., Villette, S., Taveau, J.-C., Rompp, A., Brunelle, A., Lecomte, S., et al. Multimodal imaging based on vibrational spectroscopies and mass spectrometry imaging applied to biological tissue: a multiscale and multiomics review. *Analytical chemistry 93*, 1 (2020), 445–477.

[253] Tuck, M., Grélard, F., Blanc, L., and Desbenoit, N. Maldi-msi towards multimodal imaging: challenges and perspectives. *Frontiers in Chemistry 10* (2022).

[254] Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. Tissue-based map of the human proteome. *Science 347*, 6220 (2015), 1260419.

[255] van Dam, S., Baars, M. J., and Vercoulen, Y. Multiplex tissue imaging: Spatial revelations in the tumor microenvironment. *Cancers 14*, 13 (2022), 3170.

[256] Van de Plas, R., De Moor, B., and Waelkens, E. Imaging mass spectrometry based exploration of biochemical tissue composition using peak intensity weighted pca. In *2007 IEEE/NIH Life Science Systems and Applications Workshop* (2007), IEEE, pp. 209–212.

[257] Van de Plas, R., Yang, J., Spraggins, J., and Caprioli, R. M. Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nature methods 12*, 4 (2015), 366–372.

[258] van den Brink, S. C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., Robin, C., and Van Oudenaarden, A. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature methods 14*, 10 (2017), 935–936.

[259] Van der Maaten, L., and Hinton, G. Visualizing data using t-sne. *Journal of machine learning research 9*, 11 (2008).

[260] van der Merwe, D.-E., Oikonomopoulou, K., Marshall, J., and Diamandis, E. P. Mass spectrometry: uncovering the cancer proteome for diagnostics. *Advances in cancer research 96* (2006), 23–50.

[261] Van Leenders, G. J., Van Der Kwast, T. H., Grignon, D. J., Evans, A. J., Kristiansen, G., Kweldam, C. F., Litjens, G., McKenney, J. K., Melamed, J., Mottet, N., et al. The 2019 international society of urological pathology (isup) consensus conference on grading of prostatic carcinoma. *The American journal of surgical pathology 44*, 8 (2020), e87.

[262] Vandereyken, K., Sifrim, A., Thienpont, B., and Voet, T. Methods and applications for single-cell and spatial multi-omics. *Nature Reviews Genetics 24*, 8 (8 2023), 494–515. number: 8 publisher: Nature Publishing Group.

[263] Vaysse, P.-M., Heeren, R. M., Porta, T., and Balluff, B. Mass spectrometry imaging for clinical research–latest developments, applications, and current limitations. *Analyst 142*, 15 (2017), 2690–2712.

[264] Veenhuizen, K. C., De Wit, P. E., Mooi, W. J., Scheffer, E., Verbeek, A. L., and Ruiter, D. J. Quality assessment by expert opinion in melanoma pathology: experience of the pathology panel of the dutch melanoma working party. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland 182*, 3 (1997), 266–272.

[265] Verbeeck, N., Caprioli, R. M., and Van de Plas, R. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. *Mass spectrometry reviews 39*, 3 (2020), 245–291.

[266] Verbeeck, N., Spraggins, J. M., Murphy, M. J., Wang, H.-d., Deutch, A. Y., Caprioli, R. M., and Van de Plas, R. Connecting imaging mass spectrometry and magnetic resonance imaging-based anatomical atlases for automated anatomical interpretation and differential analysis. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1865*, 7 (2017), 967–977.

[267] Verbeeck, N., Yang, J., De Moor, B., Caprioli, R. M., Waelkens, E., and Van de Plas, R. Automated anatomical interpretation of ion distributions in tissue: linking imaging mass spectrometry to curated atlases. *Analytical chemistry 86*, 18 (2014), 8974–8982.

[268] Verbeeck, N., Yang, J., De Moor, B., Caprioli, R. M., Waelkens, E., and Van de Plas, R. Automated anatomical interpretation of ion distributions in tissue: Linking imaging mass spectrometry to curated atlases. *Anal. Chem. 86*, 18 (2014), 8974–8982. PMID: 25153352.

[269] Vicari, M., Mirzazadeh, R., Nilsson, A., Shariatgorji, R., Bjarterot, P., Larsson, L., Lee, H., Nilsson, M., Foyer, J.,

Ekvall, M., et al. Spatial multimodal analysis of transcriptomes and metabolomes in tissues. *bioRxiv* (2023), 2023–01.

[270] Voelkel-Johnson, C., Norris, J. S., and White-Gilbertson, S. Interdiction of sphingolipid metabolism revisited: Focus on prostate cancer. *Advances in cancer research 140* (2018), 265–293. PMID: 30060812 PMCID: PMC6460930.

[271] Wahid, M. F., Hasan, M. J., and Alom, M. S. Deep convolutional neural network for microscopic bacteria image classification. In *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)* (2019), IEEE, pp. 866–869.

[272] Walch, A., Rauser, S., Deininger, S.-O., and Hofler, H. Maldi imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochemistry and cell biology 130* (2008), 421–434.

[273] Wall, M. E., Rechtsteiner, A., and Rocha, L. M. Singular value decomposition and principal component analysis. *A practical approach to microarray data analysis* (2003), 91–109.

[274] Wang, G., Heijs, B., Kostidis, S., Mahfouz, A., Rietjens, R. G. J., Bijkerk, R., Koudijs, A., van der Pluijm, L. A. K., van den Berg, C. W., Dumas, S. J., Carmeliet, P., Giera, M., van den Berg, B. M., and Rabelink, T. J. Analyzing cell-type-specific dynamics of metabolism in kidney repair. *Nature Metabolism 4*, 9 (9 2022), 1109–1118. PMID: 36008550 PMCID: PMC9499864.

[275] Wilcox, R. R. *Introduction to robust estimation and hypothesis testing*. Academic press, 2011.

[276] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. The fair guiding principles for scientific data management and stewardship. *Scientific data 3*, 1 (2016), 1–9.

[277] Winderbaum, L. J., Koch, I., Gustafsson, O. J., Meding, S., and Hoffmann, P. Feature extraction for proteomics imaging mass spectrometry data. *Ann. Appl. Stat. 9(4)* (2015), 1973–1996.

[278] Wolf, F. A., Angerer, P., and Theis, F. J. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biology 19*, 1 (2 2018), 15.

[279] Wu, S. Z., Al-Eryani, G., Roden, D. L., Junankar, S., Harvey, K., Andersson, A., Thennavan, A., Wang, C., Torpy, J. R., Bartonicek, N., et al. A single-cell and spatially resolved atlas of human breast cancers. *Nature genetics 53*, 9 (2021), 1334–1347.

[280] Wu, X., Yang, B., Udo-Inyang, I., Ji, S., Ozog, D., Zhou, L., and Mi, Q.-S. Research techniques made simple: single-cell rna sequencing and its applications in dermatology. *Journal of Investigative Dermatology 138*, 5 (2018), 1004–1009.

[281] Xu, C., Lopez, R., Mehlman, E., Regier, J., Jordan, M. I., and Yosef, N. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Molecular Systems Biology 17*, 1 (1 2021), e9620. PMID: 33491336 PMCID: PMC7829634.

[282] Xu, M., Wang, X., Li, Y., Geng, X., Jia, X., Zhang, L., and Yang, H. Arachidonic acid metabolism controls macrophage alternative activation through regulating oxidative phosphorylation in pparγ dependent manner. *Frontiers in Immunology 12* (2021), 618501.

[283] Yin, R., Burnum-Johnson, K. E., Sun, X., Dey, S. K., and Laskin, J. High spatial resolution imaging of biological tissues using nanospray desorption electrospray ionization mass spectrometry. *Nature Protocols 14*, 12 (2019), 3445–3470.

[284] Young, M. D., and Behjati, S. Soupx removes ambient rna contamination from droplet-based single-cell rna sequencing data. *GigaScience 9*, 12 (12 2020), giaa151. PMID: 33367645 PMCID: PMC7763177.

[285] Zappia, L., and Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *GigaScience 7*, 7 (7 2018), giy083. PMID: 30010766 PMCID: PMC6057528.

[286] Zavalin, A., Yang, J., Hayden, K., Vestal, M., and Caprioli, R. M. Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry maldi tof ms. *Analytical and bioanalytical chemistry 407* (2015), 2337–2342.

[287] Zembowicz, A., and Scolyer, R. A. Nevus/melanocytoma/melanoma: an emerging paradigm for classification of melanocytic neoplasms? *Archives of pathology & laboratory medicine 135*, 3 (2011), 300–306.

[288] Zhang, W., Claesen, M., Moerman, T., Groseclose, M. R., Waelkens, E., De Moor, B., and Verbeeck, N. Spatially aware

clustering of ion images in mass spectrometry imaging data using deep learning. *Analytical and bioanalytical chemistry 413* (2021), 2803–2819.

[289] Zhang, W., Patterson, N. H., Verbeeck, N., Moore, J. L., Ly, A., Caprioli, R. M., De Moor, B., Norris, J. L., and Claesen, M. Multimodal maldi imaging mass spectrometry for improved diagnosis of melanoma. *medRxiv* (2022), 2022–11.

[290] Zhao, N., Kamijo, K., Fox, P. D., Oda, H., Morisaki, T., Sato, Y., Kimura, H., and Stasevich, T. J. A genetically encoded probe for imaging nascent and mature ha-tagged proteins in vivo. *Nature Communications 10*, 1 (7 2019), 2947. number: 1 publisher: Nature Publishing Group.

[291] Zubarev, R. A., and Makarov, A. Orbitrap mass spectrometry, 2013.

About the author



Wanqiu Zhang, 张琬秋, was born on 7 August 1994, Nanjing, China. In 2012, she studied Telecommunication Engineering in the department of Computer Science, YangZhou University (扬州大学), China.

In 2015, she went to Hasselt University, Belgium for exchange. In UHasselt, she finished her bachelor thesis *"Low-Latency Lossless Video Compression Methods for Multi-camera Systems"*, under the supervision of Prof.dr.ir. Luc Claesen.

In 2016, she studied Master of Artificial Intelligence in the department of Computer Science, KU Leuven, Belgium. She finished her master thesis *"Transfer Learning on Mass Spectrometry Imaging"* under the supervision of Prof.dr.ir. Bart De Moor, together with Dr.ir. Marc Claesen and Dr.ir. Nico Verbeeck.

In 2018, she started her PhD training in the department of Electrical Engineering, KU Leuven, under the supervision of Prof.dr.ir. Bart De Moor. In 2019, she received her 4-year PhD funding from the Flanders Innovation & Entrepreneurship Agency (VLAIO) and Aspect Analytics NV via a Baekeland mandate. Her granted PhD proposal focuses on *"Augmented Mass Spectrometry Imaging Data Analysis through Incorporation of Multimodality and Transfer Learning"*. Since then, she is co-supervised by Dr.ir. Marc Claesen (Aspect Analytics NV).

In 2022, she married Thomas in the city hall of Leuven, Belgium. They are currently living in Leuven, Belgium.

List of publications

International journal papers (published)

- Zhang, W., Claesen, M., Moerman, T., Groseclose, M.R., Waelkens, E., De Moor, B. and Verbeeck, N., 2021. *Spatially aware clustering of ion images in mass spectrometry imaging data using deep learning.* **Analytical and Bioanalytical Chemistry**, 413, pp.2803-2819.
- Van Loon, E.*, Zhang, W.* (co-first-author), Coemans, M., De Vos, M., Emonds, M.P., Scheffner, I., Gwinner, W., Kuypers, D., Senev, A., Tinel, C. and Van Craenenbroeck, A.H., De Moor, B., Naesens, M., 2021. *Forecasting of patient-specific kidney transplant function with a sequence-to-sequence deep learning model.* **JAMA Network Open**, 4(12), pp.e2141617-e2141617.

International journal paper (submitted)

- Zhang, W., Patterson, N.H., Verbeeck, N., Moore, J.L., Ly, A., Caprioli, R.M., De Moor, B., Norris, J.L. and Claesen, M., 2022. *Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma.* **medRxiv** (2022): 2022-11. doi: https://doi.org/10.1101/2022.11.29.22282868.

International journal paper (In preparation of submission)

- Zhang, W.*, Spotbeen X.*(co-first-author), Vanuytven S.**, Kint S**.(co-second-author), Sarretto T., Socciarelli F., Vandereyken K., Dehairs J., Idkowiak J., Wouters D., Alvira Larizgoitia J., Partel G., Ly A., Mantas

M., Gevaert T., Devlies W., Mah C., Butler L., Loda M., Joniau S., De Moor B., Sifrim A., Ellis S., Voet T., Claesen M., Verbeeck N., Swinnen J., 2023. *Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer.* **bioRxiv** (2023): 2023-08. doi: https://doi.org/10.1101/2023.08.28.555056

- Geysels, A., Garofalo, G., Zhang, W., Timmerman, D., Froyman, W., De Moor,B., 2023. *Multimodal Deep Learning for Ultrasound Ovarian Cancer Detection: a prospective, multicenter study.*

Patent (submitted)

- Zhang, W., Claesen, M., Verbeeck, N., 2023. *Method for vertical integration and analysis of spatial multi-omics data.* Submitted to Belgian Patent Office No. BE2023/5080. Feb 2023.

Oral presentations at international conferences

- Zhang, W., Patterson, N.H., Verbeeck, N., Moore, J., Ly, A., Caprioli, R.M., De Moor, B., Norris, J.L., Claesen, M. (2022). *Incorporating morphology via deep learning improves classification performance of MALDI imaging for skin lesions.* Presented at the 70 th Conference on Mass Spectrometry and Allied Topics, American Society for Mass Spectrometry (ASMS), Minneapolis, US, 4-9 Jun 2022.
- Zhang, W., Claesen, M., Moerman, T., Groseclose, R.M., Waelkens, E., De Moor, B., Verbeeck, N. (2020). *Deep learning enables spatially-aware clustering of ion images in MSI data.* Presented at the 24 hours conference of international MSI (IMSI), Online, 18-20 Nov, 2020.

Poster presentations at international conferences

- Zhang, W., Spotbeen, X., Vanuytven, S., Kint, S, Sarretto, T., Socciarelli, F., Alvira Larizgoitia, J., Wouters, D., Partel, G., Mantas, M., Gevaert, T., Devlies, W., Vandereyken, K., Joniau, S., Loda, M., De Moor, B., Voet, T., Sifrim, A., Ellis, S., Claesen, M., Verbeeck, N., Swinnen, J., 2023. *Pipeline for spatial multi-omics data integration reveals transcriptomic and lipid heterogeneity in prostate cancer.* Presented at the 1st IMSIS conference - Annual Conference on Mass Spectrometry

Imaging and Integrated Topics, Montreal, QC, Canada, 23-25 October, 2023.

- Zhang, W., Spotbeen, X., Sarretto, T., Socciarelli, F., Vanuytven, S., Alvira Larizgoitia, J., Kint, S, Wouters, D., Partel, G., Mantas, M., Gevaert, T., Devlies, W., Vandereyken, K., Joniau, S., Loda, M., De Moor, B., Voet, T., Sifrim, A., Ellis, S., Claesen, M., Verbeeck, N., Swinnen, J., 2022. *Spatial Multi-Omics Investigation of Prostate Cancer Heterogeneity using integrated spatially resolved lipidomics and transcriptomics.* Presented at the 71st Conference on Mass Spectrometry and Allied Topics, American Society for Mass Spectrometry (ASMS), Houston, Texas, USA, 4-8 June, 2023.

- Zhang, W., Patterson, N.H., Verbeeck, N., Moore, J., Ly, A., Caprioil, R.M., De Moor, B., Norris, J.L., Claesen, M. (2022). *A deep learning approach to incorporate microscopy and MALDI imaging enhances classification of skin lesions.* Presented at the International Mass Spectrometry Conference (IMSC 2022), Maastricht, The Netherlands, 27 Aug 2022-02 Sep 2022.

- Zhang, W., Claesen, M., Moerman, T., Groseclose, R.M., Waelkens, E., De Moor, B., Verbeeck, N. (2020). *Spatially-aware clustering of ion images in mass spectrometry imaging data through use of pre-trained neural networks.* Presented at the 68th Conference on Mass Spectrometry and Allied Topics, American Society for Mass Spectrometry (ASMS), Texas, USA (Online), 1-12 June, 2020.

- Zhang, W., Claesen, M., Verbeeck, N., Waelkens, E., De Moor, B. (2018). *Transfer Learning on Mass Spectrometry Imaging Data.* Presented at the IMSS II & OURCON VI joint Conference, Charleston, SC, USA, 11-14 November, 2018.

- van Loon, E., Zhang, W., van Craenenbroeck, A., de Moor, B., Naesens, M. (2021). *Dynamic forecasting of patient-specific kidney transplant function with a sequence-to-sequence deep-learning Model.* Transplant International: vol. 34, (62-62). 20th Biennial European Society for Organ Transplantation (ESOT) Congress, Milan, Italy, 29 Aug – 1 Sep 2021.



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
ESAT - STADIUS CENTER FOR DYNAMICAL SYSTEMS, SIGNAL PROCESSING, AND DATA ANALYTICS
Kasteelpark Arenberg 10 box 2446
B-3001 Leuven
wanqiu.zhang@esat.kuleuven.be