

Whole Slide Imaging

Current Applications and
Future Directions

Anil V. Parwani
Editor

Whole Slide Imaging

Anil V. Parwani
Editor

Whole Slide Imaging

Current Applications and Future
Directions

Editor

Anil V. Parwani

Department of Pathology

The Ohio State University Wexner Medical School

Columbus, OH

USA

ISBN 978-3-030-83331-2

ISBN 978-3-030-83332-9 (eBook)

<https://doi.org/10.1007/978-3-030-83332-9>

© Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

The advance of technology is based on making it fit in so that you don't really even notice it, so it's part of everyday life. Bill Gates

When people talk about digital pathology today, they imply whole slide imaging (WSI). That is because WSI has become the dominant imaging modality for digitizing material in the pathology laboratory. WSI has played an integral part in pathology practice for more than two decades now, with some pathology labs already demonstrating success at going *fully digital* for rendering primary diagnoses using WSI instead of glass slides. Current digital pathology systems arose from computer science research projects in the 1990s. Since then, we have witnessed the commercial introduction of sophisticated WSI systems that have productively incorporated advanced optics, digital cameras, robotics, image management, software, cloud computing, and computer vision technology.

This book effectively encapsulates the entire story about WSI from the past, what the status is at present, and delves into the future. Joel Saltz, one of the pioneers in the early development of the first ever WSI scanner, provides a historical account of the field. WSI technology, however, is complex and accordingly can be intimidating for end users to understand. Therefore, readers should welcome the useful chapters by Mohanty and Parwani as well as McClintock that offer a detailed explanation of the hardware, software, and the prerequisite IT infrastructure needed to operate these systems. For WSI to be effective, this technology also needs to be integrated in the pathology lab, which Hartman eloquently lays out in his chapter on workflow.

There are numerous applications for WSI that range from clinical to non-clinical use cases. These are covered in the chapter by Parwani and Mohanty, and also addressed in much more informative detail by global experts in the field such as Singh et al on education, Treanor and Williams on primary diagnosis, McClintock and Cornish on telepathology, Lujan et al on teleconsultation, as well as Raess and Sirintrapun on quality assurance. WSI in cytopathology has been less pervasive due to technical challenges related to focusing and screening workflow. Li and Pantanowitz suitably address these obstacles in their chapter on WSI and cytopathology. The chapters written by Dangott deal with WSI for research and image analysis, two areas where this technology has perhaps had the largest footprint and continues to drive the fields of computational pathology and biomedical informatics forward. We are also on the brink of AI adoption into mainstream pathology clinical practice. This is possible due to advances in computational technology and deep

learning. The final chapter in this book by Machiraju and Parwani nicely demonstrates the synergism possible when WSI is coupled with AI.

There is no longer doubt about whether WSI is here to stay or will fade away as another novel fad in the history of pathology. WSI has ushered in a new platform that by allowing us to digitize not just an entire glass slide, but also an entire anatomical pathology lab's routine workload, has transformed the field of pathology. WSI has thereby finally untethered pathologists from their microscopes and delivered pathology care to patients who otherwise would never have benefited from access to expert diagnoses. WSI has also liberated pathology laboratories to leverage WSI in favor of more cost-efficient processes and allowed them to expand their services. Finally, WSI has additionally allowed the field of pathology to remain in the driver's seat for precision medicine and AI. I am certain that you will derive great benefit from this comprehensive book on WSI and likely find yourself returning to it time after time to refer to many of these valuable chapters.

Liron Pantanowitz, MD
University of Michigan, Pathology & Clinical Labs
Ann Arbor, MI, USA

Preface

In recent years, advances in imaging modalities and the ability to analyze these images using digital pathology and artificial intelligence software have created many new opportunities to advance patient care. The conversion of glass slides into a digital format and the electronic communication of digitized images is digital pathology. Digital pathology provides the users with the ability to transfer a microscopic image, between one pathologist and another physician (pathologist or other clinician). Digital pathology has been around for decades and continues to have many applications today in the national as well as global pathology community to be used for primary diagnosis, intraoperative consultation, second opinion consultations, research, quality reviews, tumor boards, and education.

One of the mediums used in digital pathology is whole slide imaging (WSI). WSI technology has several advantages over conventional microscopy; portability (images are often accessible anywhere and at any time), ease of sharing and retrieval of archival images, and the ability to make use of computer-aided diagnostic tools (image analysis algorithms). The automated instrument used for WSI is a scanner equipped with a robotic microscope capable of digitalizing an entire glass slide, using software to merge or stitch individually captured images into a composite digital image. The critical components of an automated WSI system include bar coded slides, hardware (scanner composed of an optical microscope and digital camera connected to a computer), software (responsible for image creation and management, viewing of images, and image analysis where applicable), and network connectivity. The last decade has seen significant technology advances in the evolution of WSI with the ability to rapidly digitize large numbers of slides automatically and at high resolution. Many applications have emerged and, as a result, WSI is increasingly being used in both clinical and research areas. Whole slide imaging technology has evolved to the point where digital slide scanners are currently capable of automatically producing high-quality, high-resolution digital images within a relatively short time – less than one minute per slide.

The focus of this book is to provide up-to-date and practical knowledge in all aspects of whole slide imaging by experts in the field. This includes a historical perspective on the evolution of this technology, technical aspects of making a great whole slide image, the various applications of whole slide imaging, and future applications using WSI for computer-aided diagnosis.

The goal is to provide practical knowledge and address knowledge gaps in this emerging field. This book is unique because it will address an emerging area in pathology for which currently there is only limited information about the practical aspects of deploying this technology. For example, there are no established selection criteria for choosing new scanners and a knowledge base with the key information. The authors of the various chapters have years of real-world experience in selecting and implementing WSI solutions in various aspects of pathology practice. This book will also provide practical tips and pearls to address the selection of a WSI vendor, technology details, implementing this technology and provide an overview of its everyday uses in all areas of pathology.

This book will also provide readers with important information on how to integrate their digital slides with the laboratory information system and streamline their “digital workflow” with the intent of saving time, saving money, reducing errors, improving efficiency and accuracy, and ultimately benefiting patient outcomes.

I am particularly excited about this book and have invited expert contributors to also focus on applications of WSI in the area of artificial intelligence and machine learning techniques such as deep neural networks which may be trained to not only recognize specific patterns on a whole slide image of an H&E slide but in addition AI tools may also help in the interpretation of features in the tissue that are predictive and/or prognostic.

This is an exciting time in pathology, and this book aims to give the readers a look at WSI with a deeper lens and also envision the future of pathology imaging as it pertains to WSI and associated digital innovations. These digital innovations have the potential to change the way clinical diagnosis occurs, with added benefits of shared images and data; increased efficiency and integrated diagnostics; modernized pathology work flows to improve patient care and safety; increased collaboration through multidisciplinary, disease-specific patient care conferences; improved accountability in the work flow; and, finally, cost savings by optimizing staff performance. The possibility of using WSI in computational pathology and artificial intelligence has the promise to open new frontiers in pathology which even I cannot fully imagine but can only dream of. The possibilities are endless, and I want to invite you to share the vision and the possibilities and take a virtual journey into the next generation of amplified and augmented pathology.

Columbus, OH, USA

Anil V. Parwani

Contents

Introduction to Digital Pathology from Historical Perspectives to Emerging Pathomics. 1

Rajarsi Gupta, Tahsin Kurc, and Joel H. Saltz

 Introduction. 1

 Origins of the Virtual Microscope. 3

 Development of Computational Frameworks and Software. 4

 Frontiers of Computational Pathology and Pathology Informatics 9

 Conclusions. 16

 References. 17

Whole Slide Imaging Hardware, Software, and Infrastructure. 23

David S. McClintock, Jacob T. Abel, and Toby C. Cornish

 Introduction. 23

 Whole Slide Imaging Technology. 24

 Whole Slide Imaging Fundamentals. 24

 WSI File Format 26

 DICOM in Whole Slide Imaging 29

 Whole Slide Imaging Systems: Hardware 30

 Whole Slide Imaging Devices. 30

 Slide Capacity. 30

 Slide Loading and Handling 31

 Slide Format 33

 Magnification and Spatial Resolution. 34

 Scanning and Focus Strategy 35

 Scanning Speed. 36

 Image Modalities 39

 Workstations 40

 Displays. 41

 Whole Slide Imaging: Software 42

 The Viewer 42

 Image Management Systems 43

 Clinical Image Management Systems 43

 Research and Education Image Management Systems 44

 Adding Value Through Software. 45

Whole Slide Imaging: Infrastructure	46
Networking	46
Storage	47
Conclusions.	51
References.	51
Whole Slide Imaging: Applications	57
Sambit K. Mohanty and Anil V. Parwani	
Introduction.	57
Regulation and Validation.	59
Applications of Whole Slide Imaging.	59
WSI in Education, Tumor Boards, Presentations, Quality Assurance, and Research	60
WSI in Primary Frozen Section/Intraoperative Consultation/Diagnosis.	62
WSI in Routine Pathological Diagnosis	63
WSI and Immunohistochemistry and Electron Microscopy.	65
WSI in Artificial Intelligence	67
Education	67
Quality Assurance.	67
Pathological Diagnosis	68
Image Analysis	69
Caveats and Challenges of WSI	70
Cost.	70
Technological Issues.	71
Professional Barriers.	71
Regulatory Issues	71
Conclusions.	72
References.	73
Developing a Clinical Workflow That Fits Your Needs	81
Douglas J. Hartman	
Remote Access	81
Slide Archival	82
Primary Diagnosis/Facilitate Clinical Sign Out	84
Image Analysis	85
Artificial Intelligence/Machine Learning	87
External Consultation	88
Education	88
Stakeholder Engagement	89
Selecting a System/Use Cases	89
Conclusion	90
References.	90

Whole Slide Imaging: Applications in Education	95
Matthew G. Hanna, Rajendra Singh, and Anil V. Parwani	
Introduction	95
Building a Workflow for Enhancing Education	97
Challenges and Barriers	98
Use of Digital Pathology for Education: Advantages	98
How Do Digital Slides Enhance Pathology Education	98
Advantages with Presentations, Conferences,	
Tumor Boards, and Social Media	99
Digital Pathology and Social Media	100
Future of Pathology Education Using Digital Slides	101
References	102
Whole Slide Imaging and Primary Diagnosis	105
Bethany Jill Williams and Darren Treanor	
Introduction	105
Benefits of Large-Scale Digitisation for Primary Diagnosis	106
Efficiency and Workflow	106
Service Quality	106
Workforce Optimisation	106
Patient Safety	107
The Evidence Base for Primary Diagnosis	107
Technical Considerations	107
IT Considerations	107
Integration Is Key	108
Barcoding and Tracking	108
IT Network and Storage Capacity	108
Pathologist Interface for Primary Diagnosis	109
Laboratory Considerations	109
Non-standard Slide Formats	109
Scanning Magnification	109
Primary Diagnosis Validation and Training	110
General and Specific Training Points	
for Primary Digital Diagnosis	111
Detection of Small Diagnostic and Prognostic Objects	111
Dysplasia	112
Mitotic Figure Counting	112
Specific Diagnostic Items and Features	113
Potential Pitfalls	113
Continuing Surveillance and Audit	113
Conclusion	115
References	116

Whole Slide Imaging and Telepathology	117
Toby C. Cornish and David S. McClintock	
Introduction to Telepathology	117
Definitions	118
Telepathology	118
Transmitting Site	119
Receiving Site	119
Store and Forward	119
Primary Diagnosis	120
Secondary Diagnosis	120
Intraoperative Diagnosis	121
A Brief History of Telepathology	121
Drivers of Adoption for Telepathology	124
Telepathology as a Use Case for WSI	125
Comparison with Other Telepathology Modalities	125
WSI Versus Static Image Telepathology	125
WSI Versus Dynamic Telepathology	126
WSI Versus Robotic Telepathology	126
Hybrid WSI Scanners	127
Technical Considerations for WSI Telepathology	128
Network Quality	128
Throughput	128
Latency	129
Network Firewalls	130
Approaches to Transmitting WSI Data for Telepathology	130
Traditional Client-Server Solutions	131
Cloud-Based Solutions	133
Telepathology and Cybersecurity	133
Telepathology Network Topology	134
Hub and Spoke Model	134
Reverse Hub and Spoke Model	134
Mixed Model	136
Regulatory and Legal Aspects of WSI Telepathology	136
Telepathology and In Vitro Device Regulation	136
Whole Slide Imaging, the FDA, and CLIA	137
Laboratory Accreditation and Telepathology	138
Telepathology Accreditation Standards	138
Laboratory Validation of Telepathology Systems	140
Laboratory Validation of Whole Slide Imaging Systems	141
Image Retention and Telepathology	142
Medical Licensure	142
Interstate Medical Licensure Compact	143
Privileging and Credentialing	144
Liability, Malpractice, and Malpractice Insurance	145
Reimbursement	145
COVID-19 and Telepathology	146

Summary	148
References	148
Whole Slide Imaging: Remote Consultations/Second Opinions	153
Giovanni Lujan, Anil V. Parwani, and Marilyn M. Bui	
Introduction	153
Importance of Consultation for Second Opinions	154
Shortage of Pathologists in General Around the World	155
Worldwide Initiatives	157
Some Examples of a Work Already in Progress	158
Digital Pathology as a Window into Computational Pathology	159
Telepathology and COVID-19 Pandemic	159
Conclusions	160
References	160
Quality Assurance and Quality Improvement Enabled by Whole Slide Imaging	163
Philipp W. Raess and S. Joseph Sirintrapun	
Introduction	163
WSI in the Pre-analytical Phase of Testing	164
WSI and Patient/Specimen Identification	164
Case Assignment and Distribution	165
Quality Assurance of Physical Glass Slides with Histology and Immunohistochemistry	166
Quality Assurance of Digital Images	167
WSI in the Analytical Phase of Testing	168
Intramural and Extramural Consultation	168
Intradepartmental Quality Assurance Programs	170
Increased Diagnostic Efficiency	171
WSI in the Post-analytical Phase of Testing	172
WSI and Quality Assurance: Future Considerations with Computational Pathology	173
References	174
Whole Slide Imaging in Cytopathology	179
Zaibo Li and Liron Pantanowitz	
Introduction	179
Challenges of WSI in Cytopathology	179
Focus	179
Magnification and Resolution	183
Cytopathology Workflow	184
Clinical Applications	185
Primary Diagnosis	185
Secondary Diagnosis	185
Rapid on-Site Evaluation (ROSE)	185
Non-clinical Applications	186

Education	186
Proficiency Testing	187
Quality Assurance	187
Cytology Slide Archiving	188
Conclusion	188
References.	188
Whole Slide Imaging and Research Applications	193
Bryan Dangott	
Analog vs Digital	193
Manual Methods	194
Clinical and Research Applications for Biomarkers.	195
Multiplexing Techniques.	196
Immunofluorescence	197
Other Techniques	197
Tissue Microarrays (TMAs)	199
Summary	201
References.	201
Whole Slide Image Analysis.	203
Bryan Dangott	
Regulatory.	204
WSI Scanner as a Medical Device	206
Economics.	207
Digital Images Dissected	208
RGB Space	208
Pixels in Digital Detectors	209
Pixels in Digital Files	210
Pixels in Digital Displays	211
WSI Images vs Other Images	211
Image Analysis Benefits	212
Image Analysis Concerns	214
Parallel Processing	215
Tiling.	215
Annotations	216
Image Analysis Tools	216
Pixel-Based Methods	217
Object-Based Methods	217
Architecture-Based Methods	218
Future Directions	219
Conclusion	221
References.	221

Whole Slide Imaging: Deep Learning and Artificial Intelligence.	223
Asmaa Aljuhani, Arunima Srivastava, James P. Cronin, Jany Chan, Raghu Machiraju, and Anil V. Parwani	
Introduction.	224
Challenges in Implementing Convolutional Neural Networks in Digital Pathology	225
Data Quality and Transformation	226
Sample Size.	226
Image Format	226
Color Space, Transformation, and Contrast	227
Tiling and Filtering the Image	228
Normalization	228
Stain Deconvolution	228
Annotation and Labeling.	229
Annotation.	229
Clinical and Histopathologic Labels.	230
Convolutional Neural Networks	230
Anatomy of a CNN.	231
Hidden Layers.	231
Further Steps for Fine-Tuning the CNN	231
Feature Identification	231
Validation and Performance	232
Applications of AI in WSI	232
Detection and Segmentation.	232
Artifact Discovery.	233
Classification (Diagnosis and Grading)	233
Summary.	233
References.	234
Index.	237

Contributors

Jacob T. Abel, MD University of Michigan, Department of Pathology, Ann Arbor, MI, USA

Asmaa Aljuhani, MS in Computer Science Department of Computer Science and Engineering, College of Engineering, The Ohio State University, Columbus, OH, USA

Marilyn M. Bui, MD, PhD Moffitt Cancer Center, Department of Pathology, Tampa, FL, USA

Jany Chan, PhD Department of Computer Science and Engineering, College of Engineering, The Ohio State University, Columbus, OH, USA

Toby C. Cornish, MD, PhD University of Colorado School of Medicine, Department of Pathology, Aurora, CO, USA

James P. Cronin, DVM Department of Veterinary Biosciences, College of Veterinary Medicine The Ohio State University, Columbus, OH, USA

Department of Internal Medicine, College of Medicine, The Ohio State University, Columbus, OH, USA

Bryan Dangott, MD Mayo Clinic – Jacksonville, Department of Laboratory Medicine and Pathology, Jacksonville, FL, USA

Rajarsi Gupta, MD, PhD Department of Biomedical Informatics, Stony Brook University, Health Sciences Center, Stony Brook, NY, USA

Matthew G. Hanna, MD Memorial Sloan Kettering Cancer Center, Department of Pathology, New York, NY, USA

Douglas J. Hartman, MD University of Pittsburgh Medical Center, Department of Pathology, Pittsburgh, PA, USA

Department of Anatomic Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Tahsin Kurc, PhD Department of Biomedical Informatics, Stony Brook University, Health Sciences Center, Stony Brook, NY, USA

Zaibo Li, MD PhD Department of Pathology, The Ohio State University Wexner Medical Center, Columbus, OH, USA

Giovanni Lujan, MD College of Medicine, The Ohio State University Wexner Medical Center, Department of Pathology, Columbus, OH, USA

Raghu Machiraju, PhD Department of Computer Science and Engineering, College of Engineering, The Ohio State University, Columbus, OH, USA

Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH, USA

Department of Pathology, College of Medicine, The Ohio State University Wexner Medical School, Columbus, OH, USA

David S. McClintock, MD University of Michigan, Department of Pathology, Ann Arbor, MI, USA

Sambit K. Mohanty, MD Department of Pathology and Laboratory Medicine, Advanced Medical Research Institute, Bhubaneswar, Odisha, India

Liron Pantanowitz, MD Department of Pathology, University of Michigan, Ann Arbor, MI, USA

Anil V. Parwani, MD, PhD, MBA Department of Pathology, College of Medicine, The Ohio State University Wexner Medical School, Columbus, OH, USA

Philipp W. Raess, MD, PhD Oregon Health & Science University, Department of Pathology, Portland, OR, USA

Joel H. Saltz, MD, PhD Department of Biomedical Informatics, Stony Brook University, Health Sciences Center, Stony Brook, NY, USA

Rajendra Singh, MD Northwell Health, Department of Pathology, New York, NY, USA

S. Joseph Sirintrapun, MD Memorial Sloan Kettering Cancer Center, Department of Pathology, New York, NY, USA

Arunima Srivastava, PhD Department of Computer Science and Engineering, College of Engineering, The Ohio State University, Columbus, OH, USA

Darren Treanor, MB, BSc(Computing), PhD, FRCPath Leeds Teaching Hospitals NHS Trust, Department of Histopathology, Leeds, UK

University of Leeds, Leeds, UK

Department of Clinical Pathology, and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden

Bethany Jill Williams, MBBS BSc Leeds Teaching Hospitals NHS Trust, Department of Histopathology, Leeds, UK



Introduction to Digital Pathology from Historical Perspectives to Emerging Pathomics

Rajarsi Gupta, Tahsin Kurc, and Joel H. Saltz

Introduction

Digital pathology became a vast new frontier in medicine and science ever since glass slide scanners emerged 20 years ago. Nowadays, high-resolution whole slide images (WSIs) of histologic tissue samples are available on demand through virtual microscopy. As the number of glass tissues slides that are converted into WSIs continues to grow, digital pathology is leading to the creation of substantial multidisciplinary research efforts comprised of physicians, scientists, and engineers who are actively collaborating across academia and industry around the world.

Whole slide imaging, virtual microscopy, and digital pathology were driven by the need for telepathology to enable pathologists with the ability to remotely view tissue samples and communicate histopathologic diagnoses. The first applications of telepathology utilized cameras to take pictures and record videos of tissue samples while using robotic light microscopes and satellite communications [1–5]. As we transitioned from the analog to the digital age of data, whole slide imaging and the internet supplanted those technologies in modern telepathology, which were supported by advances in computer hardware, frameworks, networks, and data management to support the capture and storage of high-resolution WSIs. Currently available applications of telepathology include remote microscopic examination for rapid on-site evaluation (ROSE) of cytology samples, primary diagnosis by pathologist with subspecialty expertise, and intradepartmental and outside institutional consultation for urgent, challenging, and difficult cases.

After several decades of development, high-resolution digital WSIs are routinely captured by robust and automated glass slide scanners that are easily stored, shared,

R. Gupta · T. Kurc · J. H. Saltz (✉)

Department of Biomedical Informatics, Stony Brook University, Health Sciences Center,
Stony Brook, NY, USA

e-mail: Joel.Saltz@stonybrookmedicine.edu

and readily viewed with established software systems [3, 6, 7]. The wide availability of WSIs has led to using the terminology, “digital pathology” to refer to scanning slides, archival of tissue images for tumor boards and education, and diagnostic consultation with telepathology [8]. The clinical adoption of digital pathology has been welcomed due to readily apparent opportunities that can meaningfully impact laboratory efficiency and delivering better patient care through rapid remote subspecialty consultation. Other exciting opportunities include improving diagnostic accuracy, increased review for quality assurance and control (QA/QC), and computational image analysis.

Digital pathology also represents a vast frontier for collaborative research among physicians, scientists, and engineers. For example, common steps in a typical surgical pathology research project to evaluate the prognostic and predictive value for biomarker expression may include (1) reviewing diagnostic reports to identify cohorts, (2) identifying formalin-fixed paraffin-embedded (FFPE) tissue samples of interest, (3) requesting glass slides for microscopic review, (4) preparation of non-diagnostic histologic tissue sections for research from FFPE tissue blocks, (5) performing immunohistochemical (IHC) studies, (6) obtaining clinicopathologic data from electronic health records (EHR) or tumor registries, and (7) correlative analyses with clinicopathologic and tumor registry data. In comparison, a straightforward application of digital pathology in the same setting easily saves a lot of time, costs, and resources by supporting cohort discovery via virtual access to tissue samples to ascertain the potential of pursuing a wide variety of research avenues.

Beyond the borders of laboratory medicine, digital pathology has also led to the establishment of pathomics as a result of the emergence of novel computational image analysis methodology driven by scientific and technical expertise in machine learning, artificial intelligence (AI), computer vision, and data science. Currently, sophisticated deep learning computer vision methods are being developed, implemented, and automated to routinely analyze WSIs and harvest quantitative pathomics data in order to develop advanced precision medicine applications for future clinical use. Thus, scalable Pathomics methodology is being increasingly considered for use in clinical trials and research in international academic, industrial, and pharmaceutical partnerships in order to identify patterns and relationships in embedded in massive amounts of clinical, imaging, and laboratory data to help further understand the nuances of complex human diseases.

As the integration of WSIs into clinical and research laboratories increases, it is important to be aware about various technological advances that were needed to establish digital pathology. Before having WSIs at our fingertips with broadband internet was a reality, the development of remote-controlled robotic microscopes that permitted navigation, changing magnification, and adjusting focus were critical to the development of early telepathology applications. Moreover, the fact that we can explore large collections of WSIs so seamlessly have been made possible due to significant improvements in scanning speeds, storage capacity, file compression, software, data transfer, and powerful data management resources and applications [1, 9–11]. Therefore, we provide a brief overview about the first virtual microscope, computational frameworks, and software that paved the way for whole slide imaging and digital pathology en route to the emergence of Pathomics in this chapter [1, 8, 12–18].

Origins of the Virtual Microscope

Our collective understanding of tissues and cells has been dramatically transformed by light microscopy. The fundamental goal of digital pathology is to further advance our understanding of biology and pathology in the same manner. From a historical perspective, the roots of digital pathology are rooted in optics, robotics, and computers. In this section, we focus on virtual microscopy as a critical component of digital pathology and modern telepathology for viewing WSIs in a feasible and practical manner, which could never be possible without all of the necessary technological innovations that permit us to scan large quantities of glass tissue slides and generate high-resolution WSIs. However, we focus on virtual microscopy and how the current software systems and methods for data management, query, and viewing WSIs in digital pathology arose in the 1990s during the era of spatial dataset research in computer science [19].

The core functionality of a virtual microscope emulates conventional light microscopy. Virtual microscopy enables a person to use a computer to view, pan, and zoom in and out of WSIs in the same way that a glass tissue slide is examined with a microscope. In comparison, early telepathology systems provided support for remote access with either static images or live microscopy [2–5]. Beyond basic functionality, virtual microscopes provide the capability to organize and manage a collection of tissue images for remote access and viewing by concurrent users via a client-server configuration. Implementations of the virtual microscope have made it possible to efficiently catalog WSIs, share information, and perform collaborative consultation to remotely examine tissue samples for telepathology applications, where concurrent users can access the same image or the same set of images.

The main challenge of implementing a virtual microscope in the 1990s was the difficulty of achieving interactive viewing of images that did not fit in the memory of a computer that had relatively limited memory, disk storage space, and I/O bandwidths. There were also low network bandwidths, so it was not feasible to read an entire image and transfer it to a remote client. To work around these limitations, captured images were stitched together to create multi-gigabyte WSIs to work with the first glass tissue slide scanners, which were also slow at that time [20]. Thus, it was necessary to use distributed memory computational clusters with one or more disks that were attached to each cluster node to provide a request-response capability to (1) retrieve and reduce data on the server depending on the client request and (2) send the reduced data to the client side to achieve acceptable response times. This approach required methods and tooling for the careful placement of image data across the system, as well as orchestration of I/O, data filtering, and reduction operations to minimize data retrieval overheads and latency.

The first virtual microscope system that was capable of achieving interactive viewing functionality utilized high-performance computing, which was publicly demonstrated at the American Medical Informatics Association (AMIA) conferences in 1997 and 1998 [9, 20]. The development of this virtual microscope system arose from a computer science research project that targeted the management, visualization, and analysis of large datasets from sensors [21, 22]. The novel

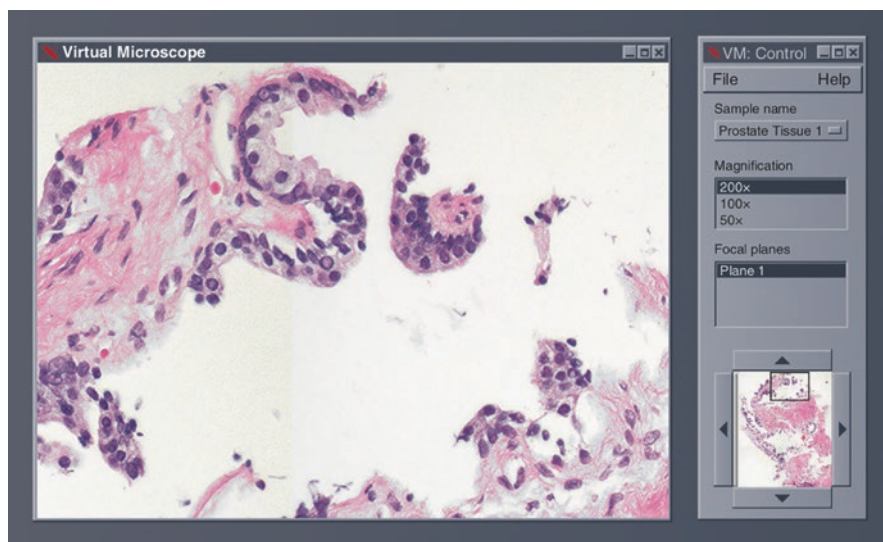


Fig. 1 1997–1998 era virtual microscope client

focus of this project was processing extremely large datasets with intensive computing to develop techniques and tools to analyze images in an extensible software platform. Until that time, the supercomputing community primarily focused on optimization to increase computational speed for data that would fit into distributed computer memory.

The first virtual microscope is shown in Fig. 1 and received the best application paper award from AMIA in 1997. By the late 1990s, it was increasingly recognized that the amount of data that was beginning to be captured by sensors on instrumentation like satellites was rapidly exceeding several orders of magnitude beyond the capacity of computer memory. The amount of data was even larger than aggregate memory on a high-performance computer with distributed memory.

Development of Computational Frameworks and Software

As the computational requirements associated with large-scale data became increasingly appreciated, the systems software group in the computer science department at the University of Maryland at College Park developed several prototype software systems to traverse datasets of images captured at multiple resolutions in the late 1990s and early 2000s. These software systems performed customized computations with sub-setting operations that rendered data for visualization and sent the output to clients to be displayed [21–23]. The scientific applications that motivated the development of these prototypes included the management and analysis of digital images from space telescopes to study changing global vegetation, seismic surveys, and subsurface oil reservoirs for Earth science. Figure 2 shows a representative screenshot from a project supported by the National Science Foundation (NSF) to

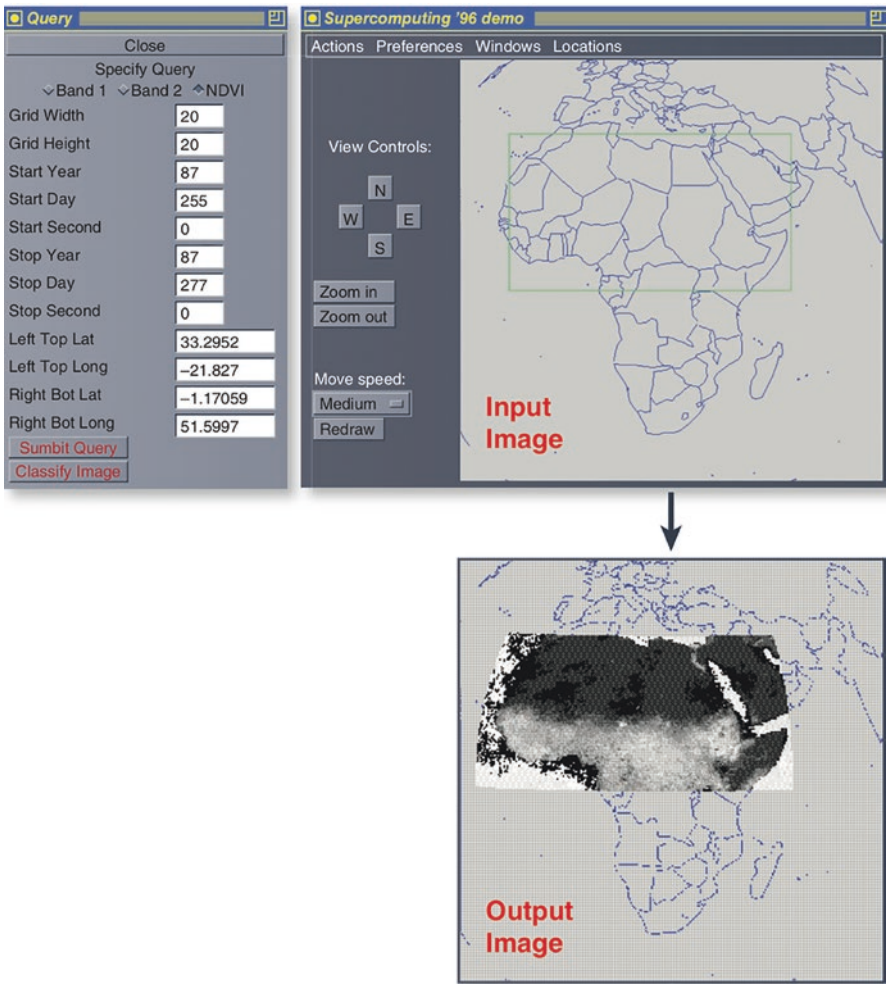


Fig. 2 NSF grand challenge in land cover dynamics

analyze the changes in global vegetation by using high-resolution satellite images by creating input data to develop models to study hydrology, carbon, and the global biogeochemical cycle.

These efforts were also driving the development of software systems and applications for digital pathology. Datasets in virtual microscopy were recognized as being very rich and complex at multiple scales of magnification for both normal and cancer-associated histology, which vary substantially across organ sites. Digital pathology was as computationally demanding as analyzing high-resolution satellite data, if not much more, due to the high degree of variability of the morphologic appearance of normal and diseased tissues and cells. Therefore, software systems for virtual microscopy were implemented to run on distributed memory computer

clusters and supercomputers with many disks and processors, where the nature of the computational framework resembled Hadoop or Apache Spark to deal with massive amounts of data [24–26].

Early software prototypes were generalized models that executed mapping operations between different multidimensional coordinate systems (e.g., three-dimensional mesh space to two-dimensional image space or between two different two-dimensional representations) and performed reduction operations on mapped data, which served as a predecessor to the popularized MapReduce model [24]. One of the innovations in data-intensive computational frameworks involved designing the capability to (1) process information from multiple coordinate systems, (2) process data captured at different levels of spatial resolution, and (3) perform computations involving multiple spatial datasets. During the design and implementation of the framework, it was also anticipated that some datasets would have non-uniform resolution that is common in both satellite imagery and virtual microscopy.

Therefore, early prototypes were able to support a broad range of methods to interpolate, upscale, downscale, warp, and render volume and other types of generalized reduction or aggregation functions. The computational frameworks were also capable of combining data sources and performing in situ data visualization, which was used for data analysis in many other types of scientific research applications besides virtual microscopy. A schematic depicting one of these early systems is shown in Fig. 3.

Even though pathology images were typically limited to manually captured photomicrographs at the time, digital pathology was an early target application domain

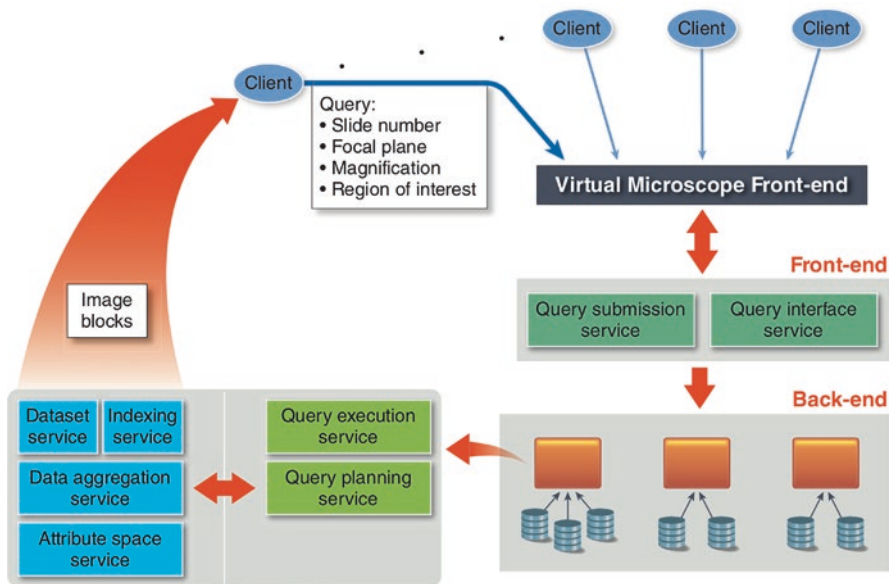


Fig. 3 Active data repository used to support whole slide digital microscopy

for this computational framework since it was clear that the widespread digitization of glass tissue slides would ultimately prove to be of great importance in medicine and science. Digital pathology was a natural application for this computational framework since processing and visualizing WSIs is very resource intensive. Basic operations such as panning and zooming were implemented with mapping and reduction operations. For example, zooming operations were viewed as reductions on image data by subsampling image pixels to fit in the viewer window at a given magnification and resolution.

Even though the first virtual microscope application was developed and described from 1997 to 2003, the fundamental concepts for computational demands and core functionality are still the same in modern whole slide imaging systems [9, 20, 27]. Users must be able to traverse WSIs by panning and zooming, overlay manual and computer-generated annotations, and collaboratively interact with the same set of WSIs. This data-intensive computational framework provided the building blocks to implement the virtual microscope system [21, 22]. However, additional optimizations were still needed to achieve high performance for this core set of tasks due to the relatively primitive nature of the hardware at that time, during when virtual microscopy was considered a relatively heroic computational effort and accomplishing these fundamental tasks to support navigating histologic images of tissues was considered monumental.

A popular high-performance computer system architecture was one in which each processor managed its own hard drives. This simplified the implementation and operating system requirements of the computer system. However, it required careful placement and management of the data and precise orchestration of I/O and computations. In order to reduce disk storage requirements and I/O retrieval costs, WSIs were partitioned into patches that were stored in compressed files in the virtual microscope system. These image patches were de-clustered across processor clusters or supercomputers to achieve computational and I/O load balance in order to achieve interactive level performance. This approach made it possible to utilize relatively inexpensive disks and aggregate I/O bandwidth from multiple storage units.

The virtual microscope software would partition a data request to view regions of WSIs at a desired zoom level into patches, overlap disk retrieval, and then assemble the image to be sent to the client in order to reduce I/O overheads. An R-tree index was implemented to quickly find the image patches that satisfied a given request [28]. In order to minimize network transfer overheads, the virtual microscope system implemented client-side caching to request regions of an image that were not in the client cache. Two versions of the virtual microscope were tested and implemented [27], where the first version was developed with the assumption that the system would be deployed on a homogeneous distributed memory system with tightly coupled nodes over a switch, whereas the second version was built on a software component architecture called DataCutter [10, 29].

In the DataCutter implementation, operations such as index lookup, data retrieval, data compression, data decompression, data subsampling, and data assembly would be implemented as individual components and loosely coupled to each other via a

streaming component framework. This implementation was based on the recognition of the emerging grid computing paradigm, where even a moderate size computing environment could consist of a heterogeneous collection of storage and computation devices. Moreover, a component-based implementation would allow new computational capabilities by either adding components and/or modifications to individual components without having to maintain a single code base. In some ways, this implementation resembled microservice architectures that have become very popular in cloud computing and distributed computing environments.

In terms of software, both open-source and commercial virtual microscopy systems have proliferated in recent years. These software systems are typically built with advanced web and cloud computing technologies that include JavaScript modules for enhanced client-side functionality and microservice-based implementation through containerization technologies. For example, the Quantitative Imaging in Pathology (QuIP) platform [30] is a fully containerized open-source system software that was developed by an academic collaboration between Emory University and Stony Brook University. Individual containers implement core functionalities, such as data management, visualization, security, and data manipulation. The containers interact with each other in a loosely coupled manner via well-defined services and interfaces to enable user interaction through internet applications. The design and implementation of QuIP has leveraged and adapted the techniques developed since the early 2000s for data management, interactive exploration of images, and viewing the results of various types of image analyses in the context of modern web and cloud computing technologies.

QuIP is one of many examples of open-source systems for virtual microscopy [31–37], as shown in Fig. 4. Other notable open-source software systems include the Digital Slide Archive [34] and Cytomine [35], which are web-based, containerized technologies and service architectures. Popular alternatives that provide desktop functionality include the Pathology Image Informatics Platform for Visualization, Analysis, and Management (PIIP) [36] and QuPath (University of Edinburgh, Edinburgh, UK) [32]. In addition, vendors of commercial slide scanners also offer their own proprietary virtual microscopy software. Alongside these options, there are also an increasing number of specialized commercial software products for viewing digital images and performing image analysis, such as HALO (Indica Labs, Corrales, New Mexico, USA), Aperio GENIE (Vista, California, USA), HistoRx AQUA Analysis (Branford, Connecticut, USA), and Visiopharm (Hoersholm, Denmark).

These open-source and commercial systems generally support a wide array of functions that surpass core image retrieval and visualization capabilities. All of the viewers are designed to give users the ability to freely explore any part of WSIs by panning and zooming to recapitulate the experience of using traditional light microscopes to examine glass tissue slides. These software applications also provide interfaces that permit viewing, organizing, and annotating large collections of WSIs. Since image analysis is beginning to play a much larger role in digital pathology, most modern software packages also support viewing computational analyses as well. Software packages that incorporate data analysis typically allow users to display the results of image analysis in an additional companion viewer or overlaid as