

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Advancing open-source visual analytics in digital pathology: A systematic review of tools, trends, and clinical applications**

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/2078353> since 2025-06-03T10:33:28Z

*Published version:*

DOI:10.1016/j.jpi.2025.100454

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

## Journal Pre-proof

Advancing open-source visual analytics in digital pathology: A systematic review of tools, trends, and clinical applications

Zahoor Ahmad, Mahmood Alzubaidi, Khaled Al-Thelaya, Corrado Cal'i, Sabri Boughorbel, Jens Schneider, Marco Agus



PII: S2153-3539(25)00039-2

DOI: <https://doi.org/10.1016/j.jpi.2025.100454>

Reference: JPI 100454

To appear in:

Received date: 14 March 2025

Revised date: 16 May 2025

Accepted date: 18 May 2025

Please cite this article as: Z. Ahmad, M. Alzubaidi, K. Al-Thelaya, et al., Advancing open-source visual analytics in digital pathology: A systematic review of tools, trends, and clinical applications, (2025), <https://doi.org/10.1016/j.jpi.2025.100454>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2025 The Author(s). Published by Elsevier Inc. on behalf of Association for Pathology Informatics.

# Advancing Open-Source Visual Analytics in Digital Pathology: A Systematic Review of Tools, Trends, and Clinical Applications

Zahoor Ahmad<sup>1,1</sup>, Mahmood Alzubaidi<sup>1,\*1</sup>, Khaled Al-Thelaya<sup>1,1</sup>,  
Corrado Calì<sup>2,3,1</sup>, Sabri Boughorbel<sup>4,1</sup>, Jens Schneider<sup>1,1</sup>, Marco  
Agus<sup>1,\*1</sup>

<sup>1</sup>College of Science and Engineering, Hamad Bin Khalifa University,  
Qatar Foundation, 34110, Doha, Qatar

<sup>2</sup>Department of Neuroscience, University of Turin,  
Corso Massimo D'Azeglio 52, 10126, Torino, Italy

<sup>3</sup>Neuroscience Institute Cavalieri Ottolenghi,  
Regione Gonzole 10, 10143, Orbassano, Italy

<sup>4</sup>Qatar Computing Research Institute, Hamad Bin Khalifa University,  
Qatar Foundation, 34110, Doha, Qatar

\*Corresponding author(s).

E-mail(s): mahmood.phd@ieee.org; magus@hbku.edu.qa

Contributing authors: zaah20848@hbku.edu.qa; khalthelaya@hbku.edu.qa;  
corrado.cali@unito.it; sboughorbel@hbku.edu.qa; jeschneider@hbku.edu.qa

†

## Abstract

Histopathology is critical for disease diagnosis, and digital pathology has transformed traditional workflows by digitizing slides, enabling remote consultations, and enhancing analysis through computational methods. In this systematic review, we evaluated open-source visual analytics abilities in digital pathology by screening 254 studies and including 52 that met predefined criteria. Our analysis reveals that these solutions—comprising abilities (n=29), software (n=13), and frameworks (n=10)—are predominantly applied in cancer research (e.g., breast, colon, ovarian, and prostate cancers) and primarily utilize Whole Slide Images (WSIs). Key contributions include advanced image analysis capabilities (as demonstrated by platforms such as QuPath and CellProfiler) and the integration of machine learning for diagnostic support, treatment planning, automated tissue segmentation, and collaborative research. Despite these promising advancements, challenges such as high computational demands, limited external validation, and difficulties integrating into clinical workflows remain. Future research should focus on establishing standardized validation frameworks, aligning with regulatory requirements, and enhancing user-centric designs to promote robust, interoperable solutions for clinical adoption.

---

<sup>1</sup> These authors contributed equally to this work.

**Keywords:** Open-source, Visual analytics, Digital pathology, Histopathology, Computational pathology, Whole Slide Imaging, Machine Learning in Pathology, Cancer Research, Image Analysis, Clinical Implementation

Journal Pre-proof



ELSEVIER

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: [www.elsevier.com/locate/jpi](http://www.elsevier.com/locate/jpi)



Advanci

Authors removed for double-blind review.<sup>a</sup>

<sup>a</sup>Authors removed for double-blind review.

---

ARTICLE INFO

ABSTRACT

---

Journal Pre-proof

*Article history:*

Received DD MMM YYYY

Received in final form DD MMM YYYY

Accepted DD MMM YYYY

Available online DD MMM YYYY

---

*Keywords:* Open-source, Visual analytics, Digital pathology, Histopathology, Computational pathology, Whole Slide Imaging, Machine Learning in Pathology, Cancer Research, Image Analysis, Clinical Implementation,

Histopathology is critical for disease diagnosis, and digital pathology has transformed traditional workflows by digitizing slides, enabling remote consultations, and enhancing analysis through computational methods. In this systematic review, we evaluated open-source visual analytics tools in digital pathology by screening 254 studies and including 52 that met predefined criteria. Our analysis reveals that these solutions—comprising tools (n=29), software (n=13), and frameworks (n=10)—are predominantly applied in cancer research (e.g., breast, colon, ovarian, and prostate cancers) and primarily utilize Whole Slide Images (WSIs). Key contributions include advanced image analysis capabilities (as demonstrated by platforms such as QuPath and CellProfiler) and the integration of machine learning for diagnostic support, treatment planning, automated tissue segmentation, and collaborative research. Despite these promising advancements, challenges such as high computational demands, limited external validation, and difficulties integrating into clinical workflows remain. Future research should focus on establishing standardized validation frameworks, aligning with regulatory requirements, and enhancing user-centric designs to promote robust, interoperable solutions for clinical adoption.

© 2025 Elsevier B. V. All rights reserved.

---

\*Corresponding author: Tel.: +0-000-000-0000; fax: +0-000-000-0000;

<sup>1</sup>Authors removed for double-blind review.

## 1. Introduction

Histopathology is a key part of medical diagnosis, involves studying disease manifestations through microscopic tissue examination [1]. It plays a critical role in medicine, particularly for cancer detection, classification, and staging [2]. Traditionally, histopathologists examine thin, stained (e.g., Hematoxylin and Eosin, H&E) tissue sections on glass slides under a microscope [3], requiring extensive training to identify subtle pathological changes in cell morphology, tissue architecture, and specific markers [4]. Beyond diagnosis, histopathology informs treatment decisions, provides prognostic information, and aids in monitoring disease progression and therapy response [4]. The field has continually evolved, integrating immunohistochemistry, molecular pathology, and, more recently, digital imaging for increasingly powerful disease characterization [5].

Histopathology is undergoing a significant transformation with the advent of digital pathology (DP), shifting traditional microscopy towards digital systems [6]. Whole slide imaging (WSI) is fundamental to this shift, representing a necessary first step for integrating diverse digital tools [7]. WSI systems digitize entire glass slides using scanners with complex optics and cameras, creating high-resolution gigapixel images [8, 1, 3, 7]. These "virtual slides" allow pathologists to view and interact with tissue images on computer monitors, akin to navigating online maps [7]. The process involves sophisticated scanning techniques (e.g., tile-based, line-scanning) and image stitching, with file sizes often necessitating compression (e.g., JPEG 2000) and multi-resolution pyramid storage formats for efficient viewing and navigation [7]. This digitization enables numerous benefits: robust digital slide archival, remote consultation and telepathology (facilitating subspecialty access and international collaboration), enhanced multidisciplinary tumor board presentations, standardized educational materials, and powerful platforms for research, including quantitative biomarker analysis [1, 2, 7]. Crucially, studies have shown that diagnostic interpretations using WSI can be equivalent to traditional light microscopy for specific applications, such as breast needle biopsies [9]. Digital pathology thus opens the door for computer-aided diagnosis, applying artificial intelligence (AI) to assist with tasks like automated region detection and quantitative analysis, potentially improving diagnostic efficiency and consistency [8, 4, 5, 7]. Visual analytics, which synthesizes automated analysis derived from AI and machine learning (ML) with interactive visualizations, is particularly suited to supporting reasoning on the large, complex datasets generated in digital pathology [10]. AI, particularly deep learning (DL) using architectures like CNNs, excels at extracting complex features from raw image data, mirroring capabilities seen across medical imaging and beyond [11]. Visual analytic approaches leverage these AI/ML capabilities (e.g., for classification or segmentation) and present the results interactively, enabling pathologists to explore and interpret the vast information within WSIs [12, 11]. These tools aim to enhance pathologists' abilities by highlighting salient information, facilitating navigation, providing quantitative measurements, and allowing integration with other data sources (e.g., genomic, clinical) for comprehensive disease analysis [13, 14, 15].

Significant interest has focused recently on developing visual analytics techniques and AI applications for digital pathology workflows. AI algorithms are being developed for diverse tasks, including screening slides, automated cancer grading (e.g., Gleason grading in prostate cancer), identifying prognostically significant features like perineural invasion, and even performing *in silico* staining [16, 11]. Various studies propose interactive visual analytics methods to support WSI analysis, including systems for tissue microarray study, visualizing multivariate features, and spatial overviews [17, 15, 14]. Despite these advancements and the potential demonstrated by AI tools (some achieving regulatory approval for specific tasks like prostate pathology assistance [16]), widespread adoption of DP for routine clinical diagnosis remains limited [18]. A recent systematic review found only 4% of DP publications focused on routine diagnostic applications, with most covering technical development, AI research, or education [18]. Many proposed solutions remain primarily research-focused, facing significant hurdles to clinical implementation [10, 5, 18]. Key challenges include the technical complexity of WSI systems, the need for robust IT infrastructure (high-capacity networks, storage solutions, high-quality calibrated monitors), substantial initial investment and unclear reimbursement pathways, lack of familiarity among staff requiring extensive training, ensuring data security and privacy, seamless integration with existing laboratory information systems (LIS), and the critical need for rigorous validation and adherence to regulatory requirements (though pathways like the FDA's Class II designation for WSI are simplifying this) [19, 20, 21, 22, 7, 18]. Slower adoption is also noted in specific areas like cytopathology due to the challenges of digitizing 3D cell structures [18].

While commercial platforms exist, the availability of open-source solutions is paramount in this field [23]. Open source promotes transparency and reproducibility via accessible source code and fosters collaborative development [24].

It allows researchers and practitioners to build upon existing work rather than starting anew, adapting tools for specific workflows. Furthermore, open-source democratizes access to advanced analytical tools, especially for institutions with limited resources, spurs innovation through global contributions, and facilitates the standardization and interoperability crucial for widespread digital pathology adoption [25, 23, 24, 25].

Some key open-source software tools for analysis and visualization of histopathology images are shown in Fig. 1. These range from general-purpose platforms like QuPath [26] (bioimage analysis) and CellProfiler [27] (cell image analysis) to more specialized ones like Ilastik [28] (interactive segmentation), ICY [29] (cell detection/classification), and Cytomine [30] (web-based collaborative analysis of multi-gigapixel images). Our interest lies particularly in tools suited for visual analytics tasks beyond basic classification or segmentation. However, the impact of open-source solutions is sometimes limited as many innovative approaches described in the literature remain commercial or unshared. This creates a gap between cutting-edge

Journal Pre-proof

research and practical tools readily available to the pathology community. Promoting the development and sharing of open-source visual analytics tools is essential to bridge this gap, ensuring wider adoption and refinement of innovations.

Supplementary Table 11, provides a detailed comparison of major open-source bioimage analysis tools currently available for digital pathology. The comparison spans across four key dimensions: core capabilities, performance characteristics, user features, and application focus. QuPath and Cytomine stand out for their robust whole slide imaging support and ability to handle large images (>40GB), while Ilastik, CellProfiler, and ICY excel in cell biology applications. Notably, Cytomine is the only platform offering comprehensive web-based collaborative features, making it particularly suitable for multi-user environments. All tools support machine learning integration and plugin development, though they differ in their specific strengths and primary applications. This comparison highlights how different tools address various needs in the digital pathology workflow, from basic image analysis to sophisticated collaborative research platforms.

To ensure a systematic and objective evaluation of the open-source visual analytics tools, we assessed each based on its developmental stage and a set of predefined criteria designed to compare maturity and practical applicability. We classified tools into three categories:

- **Frameworks:** Foundational pipelines that require additional integration and customization.
- **Tools:** Interactive solutions that perform specific tasks with a focus on usability.
- **Software:** Comprehensive, user-friendly applications with integrated interfaces for multiple tasks. Furthermore, our

evaluation considered the following aspects:

- **Core Functionalities:** Visualization methods, machine learning integration, and interactivity.
- **Performance Characteristics:** Ability to handle large datasets (e.g., whole slide images), processing speed, and scalability.
- **Usability Factors:** Ease of integration into clinical workflows, user interface design, and overall user acceptance. This combined

approach allowed for a systematic comparison of each solution's readiness for practical use.

While existing surveys of digital pathology and visual analytics offer valuable insights into computational methods, they often lack comprehensive coverage of available open-source tools and specific visual analytics techniques. This can leave practitioners uncertain about how to effectively implement advanced VA in their workflows. Our systematic review addresses these gaps. As highlighted by the comparison with recent literature (2021-2024) in Table 1, we provide a broader analysis across multiple dimensions, including publication trends, disease spectrum, image modalities, dataset accessibility, visualization methods, user interaction, and clinical implementation readiness. Crucially, our review places a unique emphasis on open-source solutions, specifically evaluating their developmental maturity (categorized as frameworks, tools, or software) to assess their practical availability and applicability for bridging the gap between research and practice.

Despite significant advancements, a systematic evaluation synthesizing the capabilities, limitations, maturity, and practical applications of open-source visual analytics tools for digital pathology has been missing. This review critically analyzes the current landscape of these open-source tools to highlight their utility, assess their developmental stages, and identify key areas for improvement, thereby aiming to facilitate their integration into clinical and research workflows for enhanced diagnostic accuracy and patient care. Specifically, our analysis addresses the following primary research questions:

- **RQ1:** What are the prevalent open-source visual analytics techniques in histopathology imaging, and what stages of development do they encompass (frameworks, tools, software)?
- **RQ2:** How do visual analytics methods enhance the interpretation of histopathological images, and what is their comparative efficacy?
- **RQ3:** What are the existing challenges in visual analytics for histopathology, and what are the practical considerations for clinical implementation?

The remainder of this paper is organized as follows: Section 2 details the Methods employed in our systematic review, including the Inclusion and exclusion criteria, the process for Study selection and data extraction, and the approach to Data synthesis. Section 3 presents the Results, starting with Study Selection and Bibliometric Analysis. It then delves into Data Characteristics and Availability, covering Tool Development Trends, Disease Focus, Histopathology Image Modalities, and Dataset Characteristics and Availability. This is followed by an analysis of Technical Approaches and Methodologies, including Data Preprocessing Techniques, Algorithms and Models, and Performance Metrics and Evaluation Methods. Next, we examine Visualization and Interaction Techniques, detailing specific Visualization Techniques, User Interaction Features, a Discussion of Key Open-Source Tools, and their Technical Contributions. The results section concludes with

Clinical Applications and Impact, discussing the Strengths and Limitations of Open Source Visual Analytics Solutions, along with their Clinical Applications and Real-World Implementation. Section 4 provides the Discussion, summarizing the Key Findings and outlining Implications and Future Directions. Finally, Section 5 offers the Conclusion, summarizing the review's main insights and contributions.

Journal Pre-proof

**Table 1. Comparison of Our Review with Other Histopathology Image Analysis Reviews & Surveys from 2021 to 2024**

Comparative Dimensions	Our Review	[17]	[15]	[14]	[31]	[13]	[12]
Temporal Distribution <sup>1</sup>	✓					✓	✓
Journal Profile <sup>1</sup>	✓						
Citation Metrics <sup>2</sup>	✓	✓	✓	✓	✓	✓	✓
Disease Spectrum <sup>2</sup>	✓	✓	✓	✓	✓	✓	✓
Image Modalities <sup>2</sup>	✓						
Dataset Accessibility <sup>3</sup>	✓	✓	✓		✓	✓	✓
Visualization Methodologies <sup>3</sup>	✓			✓			
UI & Interaction <sup>3</sup>	✓						✓
Clinical Implementation <sup>3</sup>	✓	✓		✓			

<sup>1</sup> Temporal Distribution: Temporal Distribution of Publications; Journal Profile: Journal Profile Analysis.

<sup>2</sup> Citation Metrics: Citation Metrics; Disease Spectrum: Histopathological Disease Spectrum; Image Modalities: Histopathological Image Modalities.

<sup>3</sup> Dataset Accessibility: Dataset Accessibility; Visualization Methodologies: Visualization Methodologies; UI & Interaction: User Interface and Interaction; Clinical Implementation: Clinical Implementation.

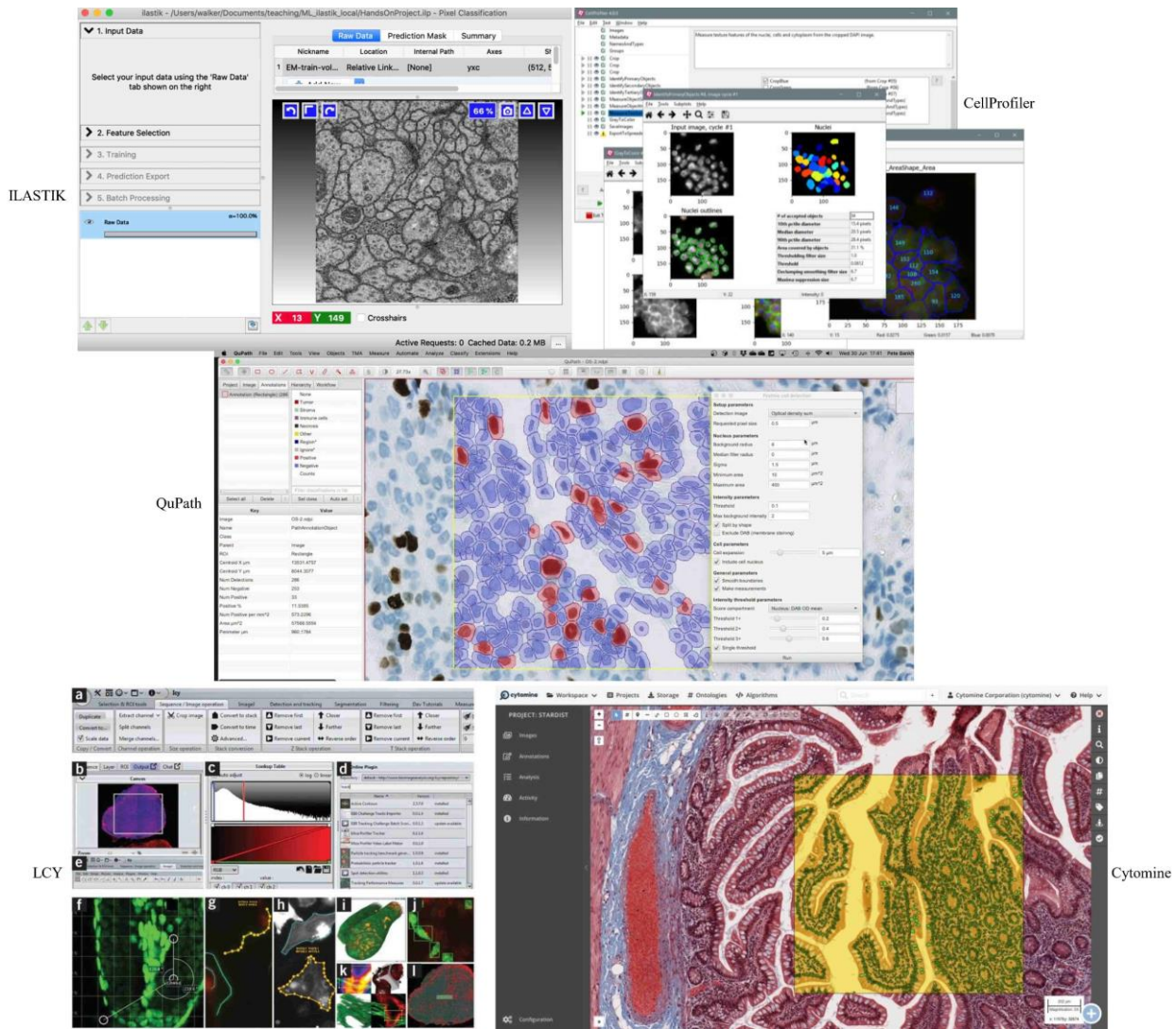


Fig. 1. Representative Open Source Software Tools for Visual Analytics in Digital Pathology. This figure illustrates five prominent tools: QuPath [26], Ilastik [28], CellProfiler [27], ICY [29], and Cytomine [30], showcasing the diverse capabilities available in open source digital pathology software.

## 2. Methods

This systematic review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [32] (details in Supplementary Table 1). We conducted a comprehensive search across Scopus, PubMed, IEEE Xplore, ScienceDirect, ACM, Eurographics DL, and Google Scholar for studies published between January 2000 and July 2024. PubMed includes MEDLINE citations, so a separate MEDLINE search was unnecessary. For Google Scholar, the initial 200 relevant results were reviewed. Reference lists of included studies were also screened.

Search terms, identified through prior literature and domain expertise, focused on studies applying visual analytics to histopathology images. The core search string was: ("Visual Analytics" OR "Data Visualization" OR "Visual Annotation" OR "Image Visualization" OR "Interactive" OR "Visual Data") AND ("Histopathology" OR "Digital Pathology" OR "Microscopic Imaging" OR "Slide Image" OR "Tissue Analysis"). This string was adapted for each database; the complete strategy is detailed in Supplementary Table 2.

### 2.1. Inclusion and exclusion criteria

The authorship team collaboratively defined selection criteria. Included studies must: (1) utilize histopathology modalities; (2) primarily use visual analytics for interacting with and visualizing histopathology images (as per search terms in Supplementary Table 2); (3) focus on developing an interactive tool, framework, or software usable without extensive coding; (4) be

open source and publicly available; and (5) be original, peer-reviewed research or conference proceedings in English.

Journal Pre-proof

Exclusions comprised studies performing only classical analysis (e.g., regression), simple segmentation/classification/diagnosis tasks, those unrelated to histopathology, non-open source research, non-English publications, and non-original research formats (reviews, abstracts, proposals, editorials, commentaries, letters).

To ensure PRISMA compliance and transparency, a list of all 254 studies screened at the full-text stage, with explicit exclusion reasons, is provided in Supplementary Table 13. This aligns with the process shown in the PRISMA flow diagram (Fig. 2) and supports reproducibility.

## 2.2. Study selection and data extraction

The Rayyan web tool [33] facilitated screening and selection. One reviewer (Z.A.) performed the search. After deduplication, titles/abstracts were screened, followed by full-text screening to finalize inclusions. Two reviewers (Z.A., M.AL) independently conducted selection and data extraction, resolving discrepancies through discussion or consultation with a third author (M.A.). A standardized data extraction form, piloted on five studies, captured key information: Paper Title, Authors, Publication Year, Journal Name, Research Objective, Study Design, Methodology, Data Source, Data Preprocessing, Algorithm/Model, Functionality, Performance Metrics, Results, Visualization Techniques, Tool/Application Developed, User Interaction, Histopathology Modalities, Evaluation Methods, Strengths, Limitations, Clinical Application, Open-Source Status, Category, Nickname, Country, Citation, and Histopathology Diseases (details in Supplementary Table 3).

## 2.3. Data synthesis

We employed a narrative synthesis of the extracted data to identify and analyze open-source visual analytics tools for histopathology images. Given the heterogeneity in modalities, methods, implementations, data, and evaluations, our analysis covered multiple dimensions.

We first examined study characteristics (demographics, aims, design) for context. Then, we summarized data types and sources used. Technical aspects explored included implementation level, ease of use, citation counts (via Google Scholar) as a proxy for impact, and verification of source code accessibility.

Based on this analysis, studies were categorized as:

- **Frameworks:** Foundational pipelines requiring integration, offering flexibility but potentially needing more technical expertise.
- **Tools:** Interactive interfaces for specific tasks or limited modalities, possibly plugins; offering targeted functionality.
- **Software:** Comprehensive, user-friendly applications with integrated interfaces for multiple tasks and modalities; offering the easiest user experience.

Our synthesis also covered dataset availability, prevalent visualization techniques, reported performance metrics, and underlying algorithms/methods. Finally, we assessed each study's strengths, limitations, and potential clinical applicability to connect research findings with practical healthcare implementation. This comprehensive approach provides a detailed overview of the current state, trends, gaps, and future opportunities in open-source visual analytics for histopathology.

## 3. Results

This section synthesizes key findings from the systematic review, organized thematically. We begin with study selection characteristics and bibliometrics, followed by data analysis (modalities, diseases, dataset access), a review of methodologies (pre-processing, algorithms, evaluation), visualization techniques and user interaction, technical contributions, and finally, clinical applications, strengths, and limitations. This structure progresses from a general overview to specific technical and practical aspects of open-source visual analytics in histopathology.

### 3.1. Study Selection and Bibliometric Analysis

Initial title/abstract screening yielded 254 candidate studies. Following rigorous evaluation against inclusion criteria, 219 articles were excluded. Forward and backward reference screening identified an additional 17 studies. Ultimately, 52 studies were included for comprehensive analysis. The selection process is depicted in the PRISMA diagram (Fig. 2A). Detailed study characteristics are available in Supplementary Tables 4-12.

The corpus comprises 48 journal papers and 4 conference publications, spanning 2006 to 2024. A notable concentration ( $n = 31$ ) occurred in the last four years, indicating accelerating interest. Research contributions are global, led by the USA ( $n = 25$ ), followed by China ( $n = 5$ ), the UK ( $n = 4$ ), and India ( $n = 3$ ), involving 19 countries overall.

The Sankey diagram (Fig. 3) illustrates the distribution of studies by type (tool, software, framework), publication year, and country. It visually confirms the increase in publications over time, particularly since 2019, with peaks in 2022-2023, contrasting with sparser contributions in earlier years (2005-2014). Tools ( $n=29$ ) represent the largest category [34, 35, 36, 37, 38, 39, 40,

Journal Pre-proof

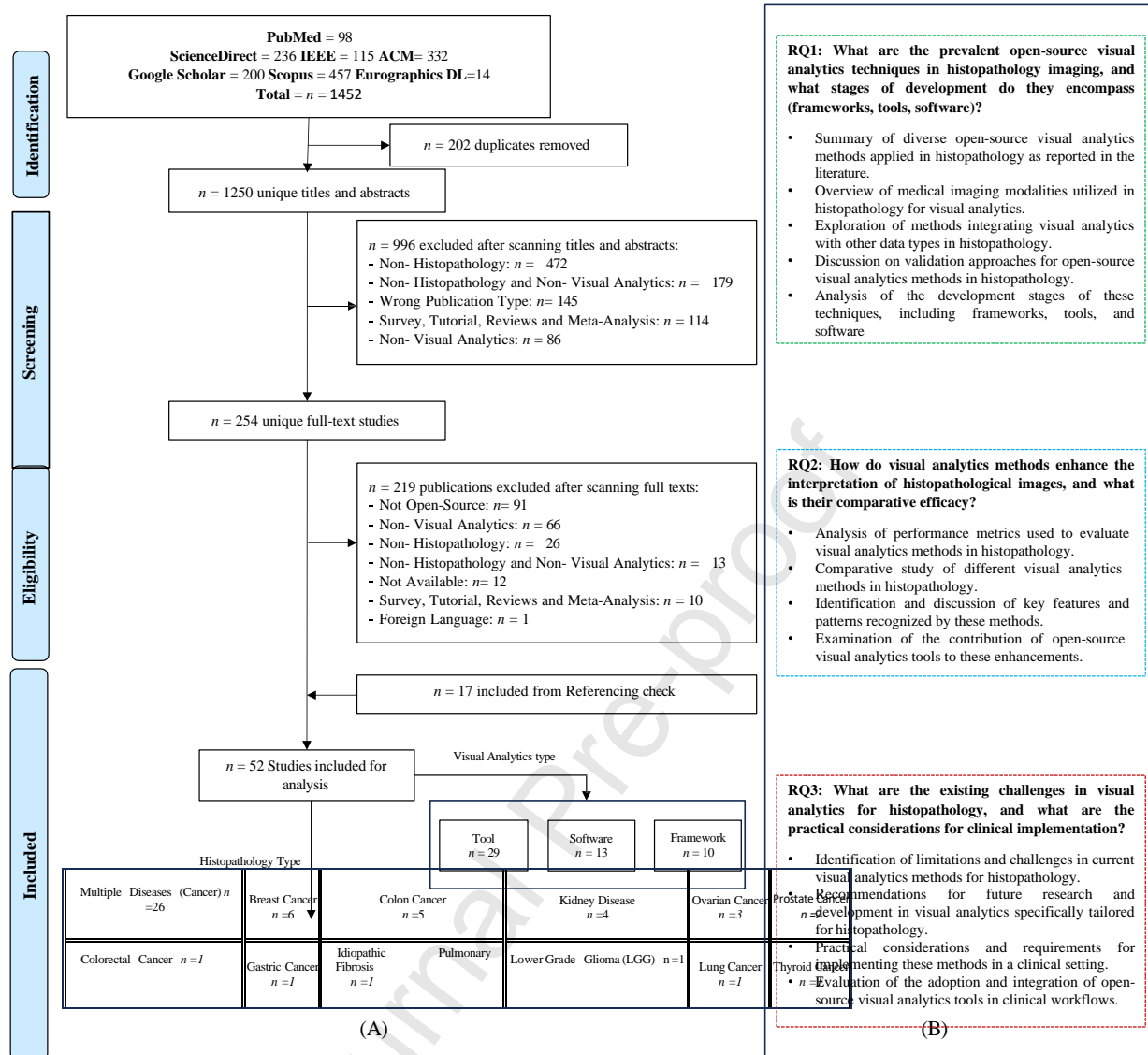


Fig. 2. Overview of the study selection process and research questions. A PRISMA Flow Diagram of Study Selection Process for Open Source Visual Analytics in Histopathology. B Research questions posed.

41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62], followed by software (n=13) [63, 64, 26, 65, 66, 67, 68, 69, 30, 28, 70, 27, 29] and frameworks (n=10) [71, 72, 73, 74, 75, 76, 77, 78, 79, 80], suggesting a focus on practical, user-oriented solutions. Supplementary Table 10 provides links and setup guidelines for these open-source resources. The data indicate a trend shifting from foundational framework development towards more applied tools and software, reflecting a maturation of the field and demand for clinically integrable solutions. The prevalence of tools suggests a problem-driven approach, highlighting the need for specialized solutions.

It is important to note that while our review focuses on open-source visual analytics, there is a significant emerging trend in histopathology involving large language models (LLMs) and multimodal AI systems. Recent non-open-source studies, outside our scope, demonstrate advancements using vision-language models for diagnostics and analytics [81, 82]. For example, PathChat integrated pathology-specific vision encoders with LLMs [81], and PathAsst emphasized data-driven training [82]. Newer systems like GPT-4V [83] and Claude 3 [84] show promise in histopathology tasks, while frameworks like MLLM4PUE

[85] and MMed-RAG [86] aim to improve multimodal embeddings and factuality. These developments highlight critical future directions, including generalization (e.g., AIDA [87]), ethics, and human-AI collaboration, even though they fall outside our specific inclusion criteria.

Our bibliometric analysis (Fig. 4) examines citation trends for frameworks, software, and tools. Fig. 4b (Software)

shows CellProfiler [27] (2006, 5595 citations) maintaining long-term impact, while QuPath [26] (2017, 5077 citations) demonstrates

Journal Pre-proof

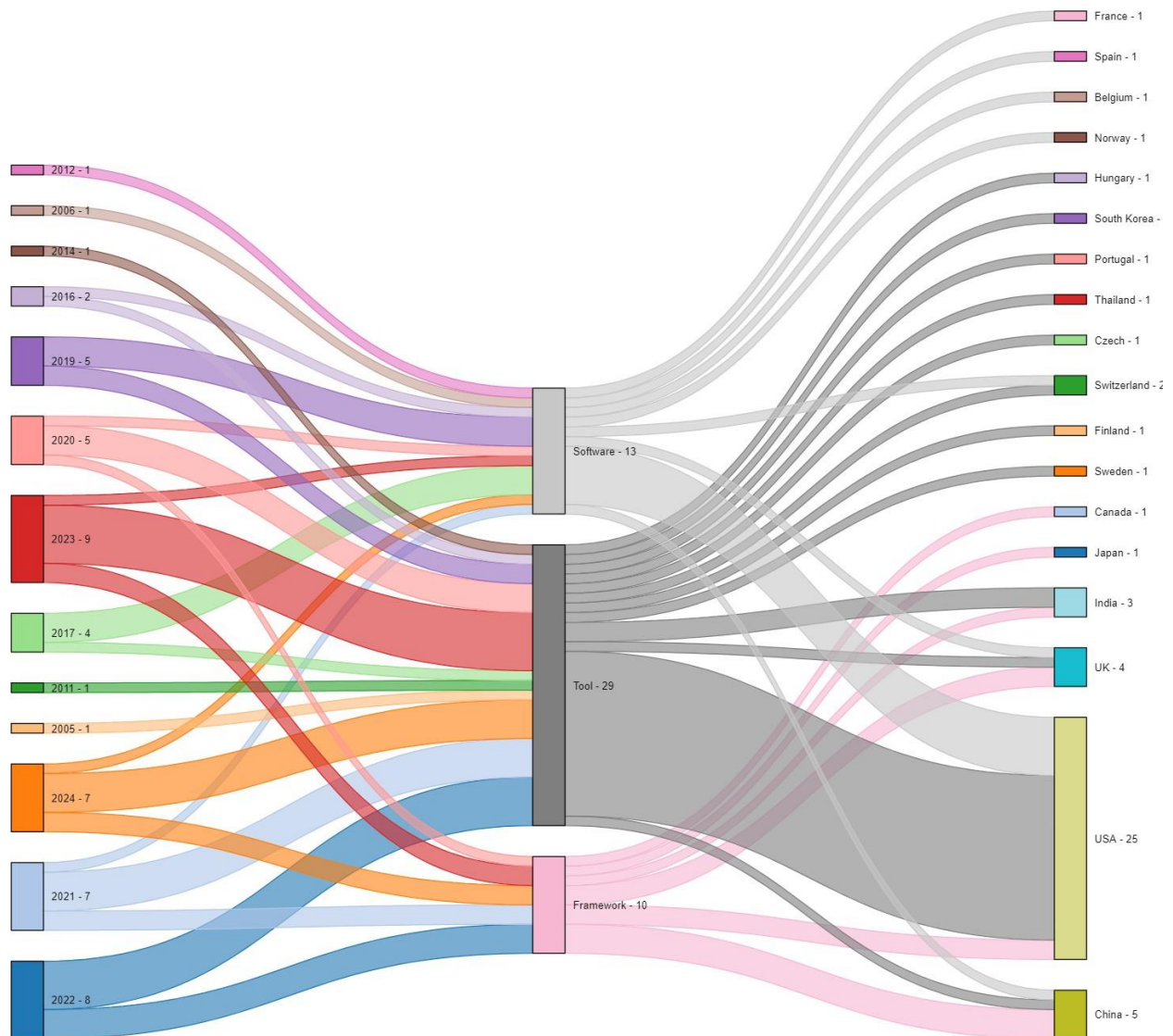


Fig. 3. Distribution of open source visual analytics studies in histopathology by type, year (2006-2024), and country, showing trends and geographical research efforts.

rapid recent adoption. ICY [29] (2012, 1550 citations) is also notable. Supplementary Table 11 details comparisons of major tools like QuPath, Cytomine, Ilastik, CellProfiler, and ICY across capabilities, performance, user features, and application focus, highlighting their respective strengths (e.g., WSI support in QuPath/Cytomine, collaborative features in Cytomine, cell biology focus in others). Fig. 4c (Frameworks) features NuClick [74] (2020, 132 citations) and SSL CR Histo [79] (2022, 99 citations), indicating newer frameworks can gain traction quickly. Fig. 4a (Tools) highlights IHC Profiler [59] (2014, 1097 citations) for sustained importance and CLAM [61] (2020, 1002 citations) for rapid recent relevance. Comparing categories, software like CellProfiler [27] and QuPath [26] lead in citations, followed by tools like IHC Profiler [59], reflecting their foundational or pivotal roles. Overall, the analysis shows both established and innovative solutions contribute significantly, reflecting a dynamic research landscape.

The distribution of publications by venue (Fig. 5) reveals significant growth and diversification over the past two decades, accelerating notably since 2018. Medical Image Analysis (n=6) and Journal of Pathology Informatics (n=5) are leading journals, followed by Scientific Reports (n=3). Publications in high-impact journals like Nature Biomedical Engineering, Nature Methods, and Nature Communications (n=2 each) underscore the quality of research. The wide range of journals and conferences (including CVPR, ICBBE) reflects the field's interdisciplinary nature, bridging medical imaging, pathology, and computer science. The presence of open-access platforms (PLoS ONE, arXiv) suggests a trend towards accessible dissemination. This diverse publication landscape, combining specialized and general venues, fosters

cross-pollination of ideas. The recent surge and venue diversity indicate a rapidly evolving field gaining momentum and recognition, poised for further integration and potentially transformative advancements.

Journal Pre-proof

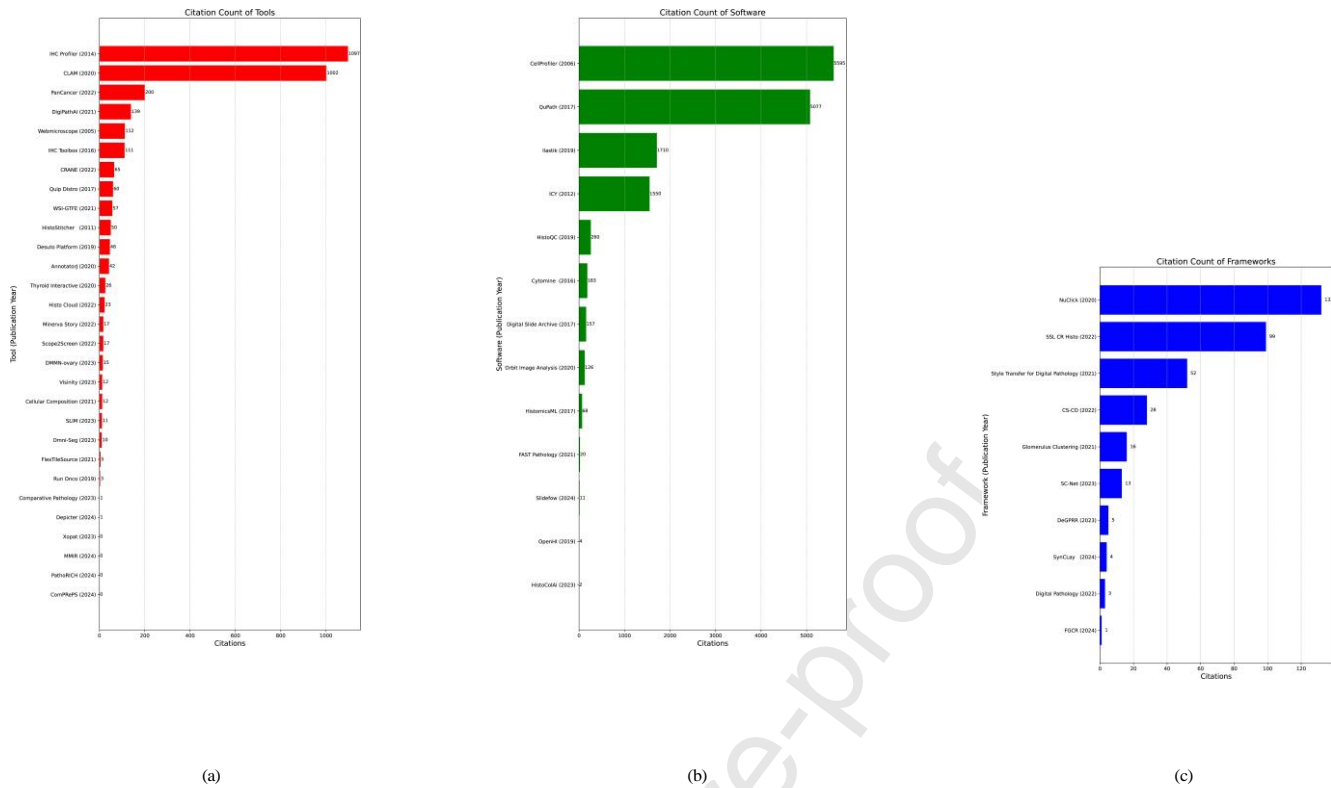


Fig. 4. Bibliometric analysis of Tools (A), software (B), and Framework (C). The y-axis indicates the year and name, while the x-axis shows the number of citations.

### 3.2. Data Characteristics and Availability

Our systematic review of visual analytics in histopathology imaging revealed significant trends in tool development, disease focus, and data modalities. These findings provide crucial insights into the current landscape of visual analytics and highlight areas for future research and development. We detailed this in Supplementary Tables 5.

#### Tool Development Trends

The analysis of visual analytics tools, as defined in our methodology, revealed a clear preference for standalone tools ( $n=29$ ) over independent software ( $n=13$ ) and frameworks ( $n=10$ ). This distribution, detailed in Table 2, suggests a trend towards more specialized, task-specific solutions in histopathology image analysis. The prevalence of standalone tools may indicate a need for flexible, modular approaches that can be easily integrated into existing workflows or customized for specific research questions.

#### Disease Focus

Our review uncovered a broad spectrum of disease foci, with a notable emphasis on cancer research. As illustrated in Table 2, the majority of studies ( $n=26$ ) addressed multiple cancer types, demonstrating the versatility and wide applicability of visual analytics in oncology. This trend aligns with the increasing focus on pan-cancer analyses in the field of visual analytics.

Specific cancer types, including breast cancer ( $n=6$ ), colon cancer ( $n=5$ ), ovarian cancer ( $n=3$ ), and prostate cancer ( $n=2$ ), also received significant attention. The focus on these cancer types likely reflects their high incidence rates and the availability of large-scale datasets. Interestingly, kidney diseases ( $n=4$ ) emerged as a notable non-cancer focus, indicating the potential of visual analytics in nephropathology.

The diversity of disease focus, including less common cancers and non-neoplastic conditions, underscores the expanding scope of visual analytics in histopathology. This breadth suggests a growing recognition of the technique's potential to address a wide range of pathological conditions.

### *Histopathology Image Modalities*

We investigated the use of different histopathology image data types and categorized them into four main modalities:

Journal Pre-proof

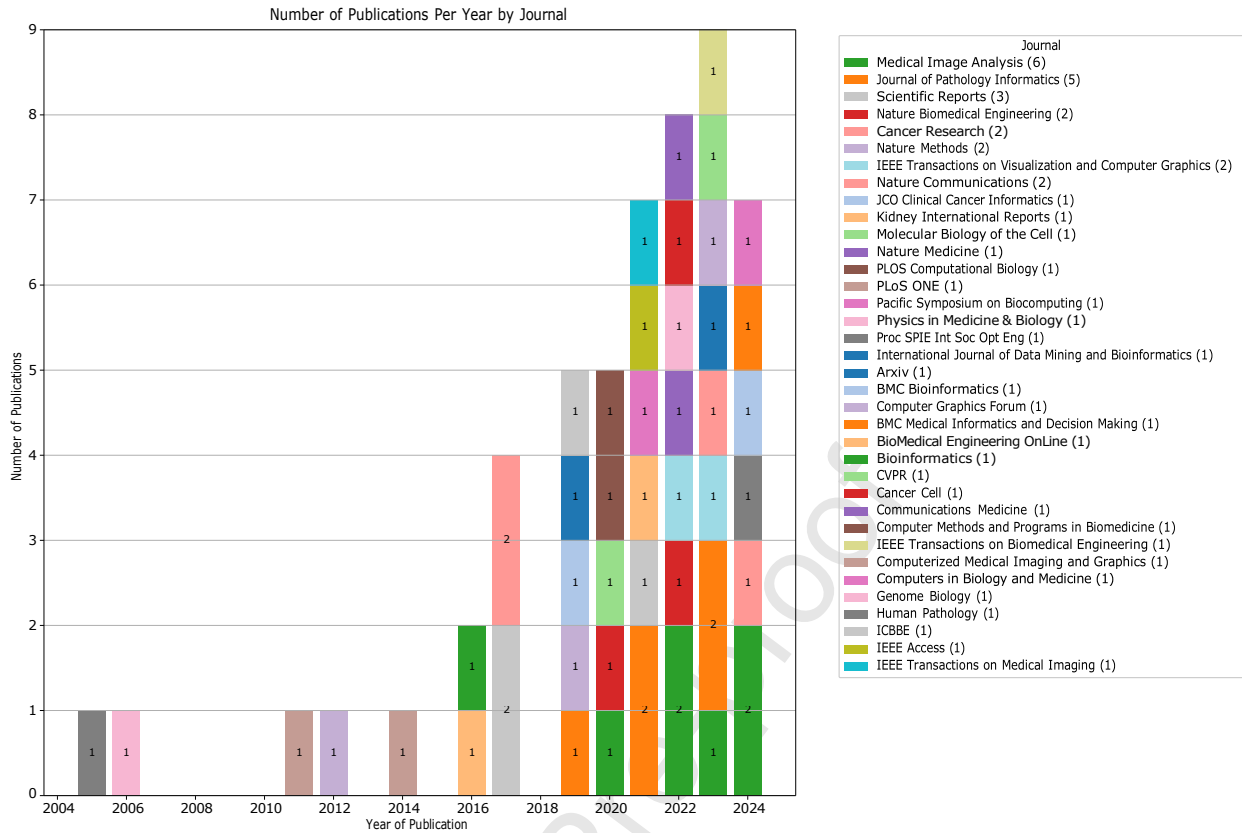


Fig. 5. Temporal Distribution of Publications in Open Source Visual Analytics for Histopathology by Journal and Conference (2004-2024).

- **Whole Slide Images (WSIs):** High-resolution digital scans representing entire histological sections from glass slides. These are typically gigapixel-sized images that allow for interactive panning and zooming, forming the basis of most modern digital pathology applications.
- **Multiplexed Tissue Images:** Images generated using techniques that visualize multiple biomarkers simultaneously on the same tissue section (e.g., cyclic immunofluorescence, mass cytometry imaging, multiplex immunohistochemistry), often resulting in multi-channel image data.
- **Microscopy Images:** This category refers specifically to traditional static images captured directly through a conventional microscope, typically representing a single field-of-view at a specific magnification. Unlike WSIs, these are not scans of the entire slide and lack the inherent navigability of virtual slides. Examples include standard light microscopy snapshots, fluorescence microscopy images of specific regions, or electron micrographs. These represent a fundamentally different data acquisition approach compared to WSI, often resulting in smaller file sizes but lacking the comprehensive context of the entire tissue section.
- **Immunohistochemistry (IHC) Images:** Images specifically showing tissue sections stained using antibodies to detect specific antigens. While often captured as WSIs in modern workflows, this category can also include traditional static microscopy images focused on IHC staining results.

Our analysis revealed a clear predominance of WSIs as the primary data modality, used in 46 out of 52 studies. This prevalence reflects the increasing digitization of pathology workflows and the rich, comprehensive information content captured by scanning entire slides. The widespread use of WSIs also indicates their potential as a standardized format for large-scale visual analytics studies.

Other modalities, including traditional microscopy images (n=15), immunohistochemistry (IHC) images (n=9), and multi-plexed tissue images (n=6), were also represented, although less frequently. The inclusion of these diverse modalities demonstrates the adaptability of visual analytics techniques to various data types and acquisition methods in histopathology.

It's noteworthy that several studies utilized multiple imaging modalities, as detailed in Table 2. This multi-modal approach suggests a trend towards more comprehensive and integrative analyses in visual analytics.

The datasets used (presented in Table 3) show several key trends:

**Table 2. Distribution of visual analytics studies by type, disease focus, and imaging modality**

Category	No. of Studies	Study References
Visual Analytics type: Tool	29	[34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62]
Visual Analytics type: Software	13	[63, 64, 26, 65, 66, 67, 68, 69, 30, 28, 70, 27, 29]
Visual Analytics type: Framework	10	[71, 72, 73, 74, 75, 76, 77, 78, 79, 80]
Disease focus: Multiple Diseases (Cancer)	26	[35, 63, 41, 72, 42, 73, 44, 74, 45, 47, 50, 51, 77, 53, 54, 68, 55, 69, 59, 60, 62, 30, 28, 70, 27, 29]
Disease focus: Breast Cancer	6	[36, 65, 66, 79, 80, 61]
Disease focus: Colon Cancer	5	[34, 26, 43, 75, 52]
Disease focus: Kidney Disease	4	[46, 76, 57, 58]
Disease focus: Ovarian Cancer	3	[40, 48, 56]
Disease focus: Prostate Cancer	2	[37, 38]
Disease focus: Other specific cancers/diseases	6	[71, 39, 64, 49, 67, 78]
Histopathology Image modalities: WSI	46	[71, 34, 35, 36, 37, 38, 63, 64, 26, 65, 40, 41, 72, 42, 73, 43, 44, 66, 75, 45, 46, 47, 48, 49, 50, 76, 51, 52, 67, 77, 53, 68, 55, 78, 56, 69, 57, 58, 79, 80, 61, 62, 30, 28, 70, 29]
Histopathology Image modalities: Microscopy Images	15	[71, 26, 74, 75, 45, 46, 76, 77, 69, 59, 60, 30, 28, 27, 29]
Histopathology Image modalities: Immunohistochemistry (IHC)	9	[26, 52, 67, 69, 59, 60, 30, 28, 29]
Histopathology Image modalities: Multiplexed Tissue Imaging	6	[35, 39, 26, 54, 55, 29]

- The Cancer Genome Atlas (TCGA) was the most frequent dataset (11 studies), reflecting a preference for large-scale, multi-disease resources suitable for pan-cancer analyses.
- Specialized datasets focusing on specific diseases or tasks (e.g., CAMELYON, DigestPath) were also common.
- Datasets integrating multiple data types (imaging, genomic, proteomic), such as the Human Cell Atlas and HTAN, indicate a move towards multi-modal analysis.
- Most datasets are publicly available with accessible links, supporting reproducibility and open science principles.
- Datasets employing advanced imaging (e.g., CyCIF for multiplexed imaging) signal growing interest in high-dimensional tissue analysis.

The diversity and accessibility of these datasets highlight collaborative research efforts and the importance of data sharing.

In summary, this analysis of data characteristics reveals trends towards specialized tools, a strong focus on cancer (particularly pan-cancer studies), the dominance of WSIs alongside growing multi-modal approaches, and increasing use of large, publicly available datasets. These insights can guide future research and development in histopathology visual analytics.

#### *Dataset Characteristics and Availability*

The datasets used (presented in Table 3) show several key trends:

- The Cancer Genome Atlas (TCGA) was the most frequent dataset (11 studies), reflecting a preference for large-scale, multi-disease resources suitable for pan-cancer analyses.
- Specialized datasets focusing on specific diseases or tasks (e.g., CAMELYON, DigestPath) were also common.
- Datasets integrating multiple data types (imaging, genomic, proteomic), such as the Human Cell Atlas and HTAN, indicate a move towards multi-modal analysis.
- Most datasets are publicly available with accessible links, supporting reproducibility and open science principles.
- Datasets employing advanced imaging (e.g., CyCIF for multiplexed imaging) signal growing interest in high-dimensional tissue analysis.

**Table 3. Overview of datasets used in histopathology visual analytics research**

Dataset Name	Description	Data Link	Study Ref.
TCGA	Comprehensive collection of genomic changes in various cancer types.	<a href="https://portal.gdc.cancer.gov/">https://portal.gdc.cancer.gov/</a>	[71, 63, 41, 66, 45, 67, 68, 56, 69, 80, 61]
CAMELYON	Breast cancer metastasis detection in lymph nodes.	<a href="https://camelyon17.grand-challenge.org/">https://camelyon17.grand-challenge.org/</a>	[36, 43, 79, 80, 61]
NCT-CRC-HE-100K	Large-scale dataset for colorectal cancer classification.	<a href="https://zenodo.org/record/1214456">https://zenodo.org/record/1214456</a>	[71]
Human Cell Atlas	Comprehensive reference map of all human cells.	<a href="https://www.humancellatlas.org/">https://www.humancellatlas.org/</a>	[35]
TCGA-HCC	Hepatocellular Carcinoma dataset from TCGA.	<a href="https://portal.gdc.cancer.gov/projects/TCGA-HCC">https://portal.gdc.cancer.gov/projects/TCGA-HCC</a>	[71]
HTAN	3D atlases of human cancers.	<a href="https://www.htanetwork.org/">https://www.htanetwork.org/</a>	[35, 45]
DigestPath	Colon cancer diagnosis and gland segmentation.	<a href="https://digestpath2019.grand-challenge.org/">https://digestpath2019.grand-challenge.org/</a>	[36, 43]
CDSA	Images associated with TCGA specimens.	<a href="https://cancer.digitalslidearchive.org/">https://cancer.digitalslidearchive.org/</a>	[63]
CyCIF	Multiplexed tissue imaging data.	<a href="https://www.cycif.org/">https://www.cycif.org/</a>	[39, 54]
BACH	Breast cancer histology image classification.	<a href="https://digitalslidearchive.org/">https://digitalslidearchive.org/</a>	[65]
PanNuke	Pan-cancer histology dataset for nuclei segmentation.	<a href="https://pannuke.org/">https://pannuke.org/</a>	[65, 72]
CONIC	Nuclei instance segmentation and classification.	<a href="https://conic-challenge.grand-challenge.org/">https://conic-challenge.grand-challenge.org/</a>	[72]
MoNuSeg	Multi-organ nuclei segmentation dataset.	<a href="https://monuseg.grand-challenge.org/">https://monuseg.grand-challenge.org/</a>	[73, 74]
PAIP	Liver cancer analysis dataset.	<a href="https://paip.grand-challenge.org/">https://paip.grand-challenge.org/</a>	[43]
GlaS	Gland segmentation in histological dataset.	<a href="https://glas.grand-challenge.org/">https://glas.grand-challenge.org/</a>	[74]
CoNSeP	Cell nuclei segmentation in pathology images.	<a href="https://warwick.ac.uk/fac/cross_fac/tia/data/hovernet/">https://warwick.ac.uk/fac/cross_fac/tia/data/hovernet/</a>	[75]
CPTAC	Comprehensive proteogenomic analysis of cancer types.	<a href="https://cptac-data-portal.georgetown.edu/">https://cptac-data-portal.georgetown.edu/</a>	[45]
IDR	Public repository of image datasets.	<a href="https://idr.openmicroscopy.org/">https://idr.openmicroscopy.org/</a>	[50]
OMERO	Software platform for biological microscopy data.	<a href="https://www.openmicroscopy.org/omero/">https://www.openmicroscopy.org/omero/</a>	[50]
SCEA	Single-cell gene expression database.	<a href="https://www.ebi.ac.uk/gxa/sc/experiments">https://www.ebi.ac.uk/gxa/sc/experiments</a>	[50]
SEER	Cancer statistics in the United States.	<a href="https://seer.cancer.gov/">https://seer.cancer.gov/</a>	[51]
KPMP	Dataset focused on kidney diseases.	<a href="https://www.kpmp.org/">https://www.kpmp.org/</a>	[57]
HuBMAP	Molecular and cellular atlas of the human body.	<a href="https://hubmapconsortium.org/">https://hubmapconsortium.org/</a>	[57]
BreastPathQ	Tumor cellularity assessment in breast cancer.	<a href="https://breastpathq.grand-challenge.org/">https://breastpathq.grand-challenge.org/</a>	[79]
Human Protein Atlas	Protein expression patterns in human tissues.	<a href="https://www.proteinatlas.org/">https://www.proteinatlas.org/</a>	[59]

The diversity and accessibility of these datasets highlight collaborative research efforts and the importance of data sharing.

In summary, this analysis of data characteristics reveals trends towards specialized tools, a strong focus on cancer (particularly pan-cancer studies), the dominance of WSIs alongside growing multi-modal approaches, and increasing use of large, publicly available datasets. These insights can guide future research and development in histopathology visual analytics.

### 3.3. Technical Approaches and Methodologies

#### *Data Preprocessing Techniques*

Effective data preprocessing is vital for histopathology image analysis. Table 4 details the techniques employed across the reviewed studies (further details in Supplementary Tables 5-7). Image normalization was the most common technique (n=23 studies, e.g., [71, 65, 40, 41]), highlighting the need for consistent intensity values across diverse datasets. Patch extraction/sampling (n=15, e.g., [36, 42, 73, 75]) was also frequent, reflecting the standard practice of analyzing large WSIs in manageable sub-regions.

Background removal (n=11, e.g., [40, 41, 43]) and segmentation (n=10, e.g., [65, 48, 51]) are crucial for focusing analysis on relevant tissue and separating structures. Techniques addressing staining variability (stain normalization/separation) and computational load (image resizing/downsampling) were each used in (n=6) studies. Data augmentation and feature extraction (n=5 each) were applied to improve model generalization and capture relevant image characteristics. Less common techniques like wavelet compression, SLIC superpixels, and style transfer (n=1 each) represent potential areas for further exploration. The observed diversity reflects the complexity of histopathology images and varied analytical requirements.

#### *Algorithms and Models*

The algorithms and models used in the reviewed studies are diverse, reflecting the complexity of histopathology analysis tasks (overview in Table 5). CNNs were the dominant approach (n=20 studies, e.g., [71, 34, 36, 65]), underscoring DL's effectiveness in capturing intricate spatial patterns. Traditional image processing techniques remain relevant (n=10, e.g., [37, 38, 64, 26]).

Among CNNs, ResNet architectures were popular (n=9, e.g., [71, 34, 36, 73]), likely due to their ability to train deep networks effectively on complex datasets. Segmentation algorithms (n=8, e.g., [65, 72, 43, 74]) and U-Net variants (n=7, e.g., [65, 73, 43, 74]) highlight the importance of precise tissue/cellular delineation. The use of attention-based models (n=5, e.g., [41, 78, 61, 62]) and Multiple Instance Learning (MIL) (n=4, e.g., [78, 56, 61, 62]) indicates growing interest in focusing on relevant regions and handling weakly labeled data.

Advanced techniques like Graph Neural Networks (GNNs) and Generative Adversarial Networks (GANs) (n=2 each) are less common, suggesting areas for future exploration. The single use of Natural Language Processing (NLP) [78] points towards emerging interdisciplinary approaches. This variety, from traditional machine learning to advanced DL, reflects the field's multifaceted nature and rapid evolution in addressing histopathology data challenges.

#### *Performance Metrics and Evaluation Methods*

A diverse range of performance metrics and evaluation methods are used to quantify and validate algorithms in this field (overview in Table 6). Accuracy was the most frequent metric (n=24 studies, e.g., [71, 64, 65, 73]), reflecting concern for overall correctness. Area Under the ROC Curve (AUC) was also common (n=17, e.g., [34, 76, 51, 52]), providing a comprehensive measure of discriminative ability. F1-score (n=14, e.g., [71, 34, 65, 73]), often reported alongside precision and recall (n=12 each), indicates the importance of balancing the identification of relevant instances while minimizing false positives. Domain-specific metrics like Dice coefficient (n=8, e.g., [36, 64, 73]) and Intersection over Union (IoU) (n=6, e.g., [40, 73, 58]) are frequently used, reflecting the significance of accurate segmentation.

Regarding evaluation, cross-validation was the most common method (n=15 studies, e.g., [71, 34, 48, 49]), emphasizing the need to assess model generalizability. Comparison with baselines/other models (n=12, e.g., [71, 34, 72]) and visual assessment by experts (n=13, e.g., [64, 26, 41]) highlight the dual importance of quantitative benchmarking and qualitative expert validation.

Ablation studies (n=6, e.g., [71, 72, 46]) and statistical analysis (n=7, e.g., [71, 26, 41]) provide deeper insights into model behavior and result significance. The use of external validation (n=5, e.g., [46, 56, 69]) and user feedback/testing (n=6, e.g., [35, 39, 44]) underscores the focus on real-world applicability and usability.

#### 3.4. *Visualization and Interaction Techniques*

This subsection analyzes the visualization and user interaction features crucial for effective visual analytics in histopathology (details in Supplementary Tables 8).

Journal Pre-proof

Table 4:

Technique	No. of Studies	Study Ref.
Image normalization	23	[71, 65, 40, 41, 43, 75, 76, 51, 52, 67, 55, 69, 57, 58, 79, 60, 80, 61, 62, 30, 70, 27]
Patch extraction/sampling	15	[36, 42, 73, 75, 46, 49, 76, 78, 56, 57, 58, 61, 62, 28, 70]
Background removal	11	[40, 41, 43, 76, 55, 69, 57, 58, 61, 62, 30]
Segmentation	10	[65, 48, 51, 68, 56, 61, 62, 28, 70, 27]
Stain normalization/separation	6	[71, 36, 76, 67, 69, 79]
Image resizing/downsampling	6	[75, 46, 76, 80, 70, 27]
Data augmentation	5	[73, 43, 74, 79, 80]
Feature extraction	5	[64, 51, 67, 78, 61, 62]
Conversion to pyramidal format	4	[42, 77, 55, 30]
Color space transformation	4	[43, 52, 79, 80]
Image annotation	3	[35, 48, 54]
Otsu's thresholding	3	[43, 69, 70]
Image stitching	2	[37, 54]
Tile-based processing	2	[64, 69]
Watershed algorithm	2	[48, 67]
Morphological operations	2	[43, 67]
Wavelet compression	1	[37]
NPY format conversion	1	[34]
Image sharpening	1	[37]
TMA dearraying	1	[26]
Nuclei object extraction	1	[72]
DICOM format conversion	1	[45]
SLIC Superpixel algorithm	1	[68]
Threshold segmentation	1	[78]
Style transfer	1	[80]
Illumination correction	1	[27]

Overview of data preprocessing techniques used in histopathology visual analytics research

Table 5. Overview of algorithms and models used in histopathology visual analytics research

Algorithm/Model	No. of Studies	Study Ref.
CNNs	20	[71, 34, 36, 65, 40, 41, 72, 73, 43, 75, 49, 76, 78, 56, 69, 58, 80, 61, 62, 70]
U-Net and variants	7	[65, 73, 43, 74, 46, 56, 60]

ResNet and variants	9	[71, 34, 36, 73, 43, 75, 49, 56, 79]
Multiple Instance Learning (MIL)	4	[78, 56, 61, 62]
Attention-based models	5	[41, 78, 61, 62, 70]
Graph Neural Networks (GNN)	2	[34, 72]
Generative Adversarial Networks (GANs)	2	[72, 69]
Support Vector Machine (SVM)	3	[64, 48, 49]
Random Forest	4	[49, 67, 30, 28]
Active Learning	2	[67, 70]
Self-supervised Learning	3	[71, 56, 79]
Semi-supervised Learning	3	[46, 79, 70]
Clustering algorithms	5	[73, 76, 53, 61, 62]
Image processing techniques	10	[37, 38, 64, 26, 66, 45, 51, 52, 59, 27]
Segmentation algorithms	8	[65, 72, 43, 74, 48, 51, 68, 57]
Feature extraction methods	5	[64, 49, 76, 78, 27]
Dimension reduction techniques	2	[34, 76]
Visualization techniques	5	[35, 44, 76, 54, 70]
Web-based frameworks	7	[35, 63, 44, 47, 50, 77, 30]
DL frameworks	3	[65, 69, 57]
Statistical models	2	[48, 52]
Natural Language Processing	1	[78]

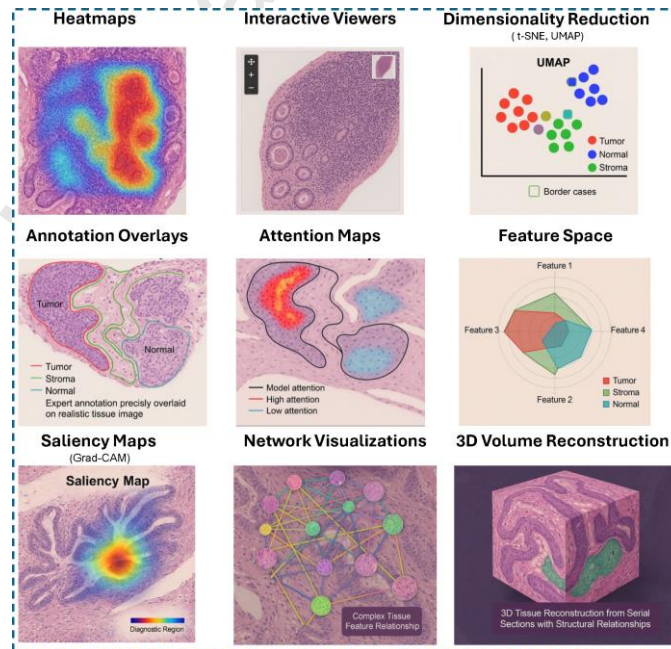
**Table 6. Overview of Performance Metrics and Evaluation Methods in histopathology visual analytics research**

Performance Metrics	No. of Stud- ies	Study ref.	Evaluation Methods	No. of Stud- ies	Study ref.
Accuracy	24	[71, 64, 65, 73, 66, 48, 49, 76, 51, 52, 56, 69, 57, 58, 80, 61, 62, 30, 70, 27]	Cross-validation	15	[71, 34, 48, 49, 56, 69, 58, 79, 80, 61, 62, 70, 41, 76, 52]
AUROC/AUC	17	[34, 76, 51, 52, 67, 56, 69, 58, 79, 80, 61, 62, 70, 27, 41, 72, 57]	Comparison with base- line/other models	12	[71, 34, 72, 43, 47, 52, 78, 56, 69, 27, 26, 41]
F1-score	14	[71, 34, 65, 73, 74, 46, 49, 51, 57, 70, 56, 69, 58, 30]	Visual assessment by experts	13	[64, 26, 41, 72, 51, 67, 55, 69, 57, 30, 73, 48, 56]
Precision	12	[65, 40, 73, 75, 49, 51, 56, 57, 30, 58, 80, 70]	Ablation studies	6	[71, 72, 46, 78, 56, 79]
Recall	12	[65, 40, 73, 75, 49, 51, 56, 57, 30, 58, 80, 70]	Statistical analysis	7	[71, 26, 41, 76, 59, 67, 56]
Dice coefficient/score	8	[36, 64, 73, 43, 74, 46, 48, 56]	External validation	5	[46, 56, 69, 61, 70]
Mean Average Precision (MAP)	5	[42, 75, 49, 78, 57]	User feedback/testing	6	[35, 39, 44, 53, 68, 30]
Intersection over Union (IoU)	6	[40, 73, 58, 60, 57, 80]	Case studies	4	[44, 50, 54, 29]
Cohen's kappa	4	[36, 43, 58, 57]	Kaplan-Meier analysis	4	[26, 41, 56, 69]
Hausdorff distance	4	[74, 46, 55, 57]	Runtime analysis	4	[39, 65, 68, 80]

### Visualization Techniques

A comprehensive overview of visualization techniques employed in the reviewed studies is presented in Table 7, with representative visual examples showcased in Figure 6. Heatmaps were the most prevalent method (n=17 studies), effective for showing spatial distributions of features or prediction confidences. Interactive viewers (zoomable/tile-based interfaces) were second most common (n=12), essential for navigating large WSIs at multiple resolutions, often incorporating annotation overlays (n=10) to link expert knowledge directly onto images. Dimensionality reduction plots (scatter plots, UMAP, t-SNE) were used in (n=11) studies, vital for visualizing high-dimensional data comprehensively.

Techniques like attention maps (n=8) and feature space visualizations (n=6) reflect the adoption of advanced machine learning, offering insights into model decision-making. Thumbnails (n=5) provide quick overviews, while parallel coordinate plots, network visualizations, saliency maps, and bar charts/histograms (n=4 each) offer various ways to represent complex data relationships. A wide array of more specialized techniques were also employed (e.g., Kaplan-Meier curves, 3D reconstruction, various CAM methods, specific rendering types, interactive UI elements), each used in one or two studies. The frequent use of multiple techniques within single studies indicates a trend towards comprehensive platforms catering to diverse analytical needs. This diversity underscores the complexity of histopathology data and the field's rapid evolution, integrating advanced computational methods with traditional visualization to enhance interpretation, as illustrated in Figure 6.



**Fig. 6.** This figure illustrates common approaches reported in Table 7, showcasing nine distinct visualization methods: Heatmaps: displaying spatial distribution of cancer probability across tissue structures. Interactive Viewers: enabling multi-resolution whole slide image exploration with navigation tools. Dimensionality Reduction techniques (t-SNE, UMAP): revealing tissue type clustering patterns. Annotation Overlays: precisely marking tumor, stroma, and immune components. Attention Maps: highlighting cellular structures of model focus. Feature Space Visualization: comparing multi-dimensional tissue characteristics. Saliency Maps (including Grad-CAM): identifying diagnostically relevant regions. Network Visualizations: displaying weighted relationships between tissue components. 3D Volume Reconstruction: showing structural relationships across tissue sections. These complementary visualization approaches support interpretation of complex histopathological data.

## User Interaction Features

User interaction features employed are described in Table 8. Interactive navigation (panning/zooming) was most frequent (n=17 studies), aligning with the common use of interactive viewers and highlighting the need for seamless WSI exploration. Annotation tools were second (n=16), indispensable for pathologists to delineate regions, label structures, and provide ground truth. Data filtering/sorting (n=12) and brushing/linking across visualizations (n=9) enable deeper data exploration and pattern discovery. Interactive exploration via heatmaps/scatter plots was also noted (n=9).

Web-based interfaces (n=8) indicate a shift towards cloud-based, collaborative solutions, supported by features like session sharing/storytelling (n=5). Efforts to improve usability include manual annotation refinement (n=6) and graphical interfaces for non-developers or programming (n=3 each). Various specialized features (tangible interactions, voice narration, lensing, drag-and-drop, model training/visualization tools) were also implemented, reflecting the need for multifaceted interaction approaches. The common integration of multiple features suggests a move towards comprehensive, user-friendly interfaces supporting the full spectrum of histopathology analysis tasks. Supplementary Table 10 provides practical guidelines (links, installation, usage) for relevant tools, aiming to facilitate their adoption.

**Table 7. Visualization Techniques in Histopathology Visual Analytics**

Visualization Technique	No of Studies	Study References
Heatmaps	17	[71, 26, 65, 40, 41, 73, 43, 44, 66, 51, 67, 53, 78, 56, 69, 58, 70]
Interactive viewers (e.g., zoomable, tile-based)	12	[35, 63, 39, 64, 65, 44, 45, 47, 77, 57, 30, 29]
Scatter plots / UMAP / t-SNE	11	[34, 35, 36, 39, 26, 49, 51, 56, 69, 58, 61]
Annotation overlays	10	[64, 65, 50, 68, 57, 58, 30, 28, 70, 29]
Attention maps	8	[40, 41, 76, 78, 56, 69, 80, 61]
Feature space visualization	6	[36, 49, 56, 69, 61, 62]
Thumbnails	5	[39, 66, 50, 67, 54]
Parallel coordinate plots	4	[39, 66, 58, 29]
Network visualizations	4	[72, 53, 56, 29]
Saliency maps	4	[76, 69, 80, 61]
Bar charts / Histograms	4	[39, 48, 54, 59]
Kaplan-Meier curves	2	[40, 41]
3D volume reconstruction	2	[77, 29]
Doughnut charts	1	[72]
Radial charts	1	[54]
Chessboard view	1	[55]
Mosaic maps	1	[69]
ROC curves	1	[52]
Score-weighted Class Activation Mapping (Score-CAM)	1	[76]
Gradient-weighted Class Activation Mapping (Grad-CAM)	1	[76]
Guided backpropagation	1	[76]
Channel-based rendering	1	[54]
Cell-based rendering	1	[54]
Segmentation with color-coding	1	[54]
Sliding-window search algorithm	1	[54]
Virtual magnification indicator	1	[68]
Line matching view	1	[55]
Warp image view	1	[55]
Annotation view	1	[55]
Gaussian mixture models (GMM)	1	[56]
Conditional GANs	1	[69]
Color-coded annotation layers	1	[57]
Interactive UI for toggling visualization layers	1	[57]
Contour assist mode	1	[60]
U-Net presegmentation	1	[60]
Integrated gradients	1	[80]
PCA for feature space visualization	1	[61]
Annotation galleries	1	[30]
Textual search engine	1	[30]
Proofreading tools	1	[30]
Interactive segmentation	1	[28]
Probability maps	1	[28]
Boundary maps	1	[28]
Region labeling	1	[70]
Integrated visual programming framework	1	[29]

## Discussion of Key Open-Source Tools

To illustrate the current landscape, we discuss five prominent open-source tools: QuPath [26], Ilastik [28], CellProfiler [27], ICY [29], and Cytomine [30], representing diverse approaches to scalability and interactivity.

**QuPath** [26] is highly popular for WSI analysis, offering robust annotation, segmentation, and an extendable plugin architecture for ML integration. Its interactive interface is optimized for gigapixel images but requires significant computational resources for large datasets.

**Ilastik** [28] excels in user interactivity with an intuitive brush-based interface for rapid classifier training (segmentation/classification), allowing quick refinement. However, scalability can be limited for extremely large WSIs without preprocessing.

**CellProfiler** [27] is known for robust quantitative analysis and high-throughput cell image processing via a modular design, valuable for reproducible research measurements. It offers less interactive visualization compared to newer tools,

potentially making real-time exploration less intuitive.

**Table 8. User Interaction Features in**

User Interaction Feature	Number of Studies	References
Interactive navigation (e.g., panning, zooming)	17	[35, 63, 39, 65, 44, 45, 47, 77, 54, 68, 57, 30, 70, 27, 29]
Annotation tools (e.g., drawing, labeling)	16	[36, 64, 26, 65, 74, 48, 49, 67, 77, 54, 57, 60, 30, 70, 27]
Filtering and sorting of data	12	[39, 72, 66, 51, 52, 69, 58, 30, 70, 27, 29]
Brushing and linking across visualizations	9	[39, 69, 58, 30, 70, 27, 29, 56, 61]
Interactive visualization of results (e.g., heatmaps, scatter plots)	9	[40, 41, 51, 69, 61, 62, 30, 70, 27]
Web-based interface for data management and analysis	8	[63, 53, 54, 68, 57, 30, 70, 29]
Manual refinement of annotations	6	[36, 74, 60, 30, 70, 27]
Collaborative features (e.g., session sharing, storytelling)	5	[44, 70, 27, 29, 30]
Point-and-click interface for non-programmers	3	[27, 29, 70]
Visual programming interface	3	[27, 29, 70]
Tangible interactions (e.g., multi-touch tabletop)	2	[72, 56]
Voice narration	1	[35]
Lensing feature for close-up inspections	1	[54]
Virtual magnification indicator	1	[68]
Drag-and-drop functionality	1	[50]
Menu options for editing cells, adding comments	1	[50]
Linking images	1	[50]
Access to external resources (e.g., OMERO image viewer)	1	[50]
Interactive selection of color regions	1	[52]
Model training and detection phases	1	[52]
Labeling objects by clicking on them in the image	1	[67]
Reviewing predictions and adjusting classification rules	1	[67]
Exploring images at multiple resolutions	1	[67]
Interactive project management	1	[55]
Algorithm management	1	[55]
Visualization of registration results	1	[55]
Interactive retrieval using image or text inputs	1	[78]
Real-time predictions and heatmap generation	1	[69]
Interactive zoom and inspection of mosaic maps	1	[69]
Editing FTU mask boundaries	1	[57]
Choice of color deconvolution	1	[59]
Threshold setting	1	[59]
Batch processing	1	[28]
Save and share analysis pipelines	1	[27]
Develop, share, and extend image analysis protocols	1	[29]

**Histopathology Visual Analytics**

**ICY** [29] provides a versatile platform with many plugins for cell detection/classification. While offering strong analytical capabilities, its interface may have a steeper learning curve than newer platforms, but its flexibility is valuable for advanced tasks.

**Cytomine** [30] features a web-based architecture supporting collaborative annotation and analysis of multi-gigapixel WSIs. Designed for scalability and concurrent use, it addresses remote collaboration and large data volumes, integrating ML tools. Its reliance on web infrastructure necessitates network stability and server support.

In summary, these tools represent a spectrum: QuPath and Cytomine offer scalable, interactive platforms for large datasets and collaboration; Ilastik provides rapid, user-friendly interactivity; CellProfiler and ICY deliver robust quantitative analysis, sometimes with trade-offs in interactive visualization. Tool selection depends on specific study needs regarding data volume, interactivity requirements, and analytical tasks.

### 3.5. Technical Contributions

The open-source visual analytics solutions for histopathology reviewed in this study demonstrate a rich tapestry of mathematical and algorithmic innovations. These contributions span a wide spectrum of computer vision, machine learning, and image processing techniques, reflecting the multidisciplinary nature of visual analytics. For a comprehensive overview of these contributions, the reader is directed to Supplement Table 6.

Several solutions introduce novel network architectures and learning paradigms that push the boundaries of DL in histopathology. CS-CO [71] presents a unified framework that ingeniously combines generative (cross-stain prediction) and discriminative (contrastive learning) methods, leveraging stain separation and vector perturbation to enhance learning. This approach addresses the critical challenge of stain variability in histopathological images while simultaneously improving feature representation. DMMN-ovary [40] introduces a Deep Multi-Magnification Network with a sophisticated multi-encoder, multi-decoder, multi-concatenation architecture for segmentation, effectively capturing multi-scale information crucial for accurate tissue analysis.

Advanced segmentation techniques feature prominently among the contributions. NuClick [74] incorporates multi-scale convolutional blocks to capture features at various scales, enabling effective segmentation of both small and large objects. It employs a weighted hybrid loss function, combining soft dice loss and weighted cross-entropy, to address class imbalance and penalize false segmentation. This approach is particularly valuable in histopathology, where object sizes can vary significantly. Omni-Seg [46] introduces a scale-aware controller that extends dynamic neural networks from single-scale to multi-scale segmentation, adapting to the inherent multi-scale nature of histological structures.

Several solutions focus on improving classification and feature extraction. WSI-GTFE [34] integrates Topological Data Analysis (TDA) techniques, specifically the Mapper algorithm, to summarize graph neural network embeddings into interpretable topological structures. This novel approach bridges the gap between the high-dimensional feature spaces of DL models and the need for interpretable results in clinical settings. PanCancer [41] introduces the use of Kronecker Product for pairwise fea-

ture interactions, enhancing the interpretability of multimodal data, which is crucial in integrating diverse data types in cancer research.

Innovative approaches to data representation and processing are evident in several solutions. OpenHI [68] introduces methods for calculating the number of superpixels based on desired sub-region size and total WSI resolution, as well as a method for calculating virtual magnification. These contributions address the computational challenges of working with large-scale whole slide images. Slideflow [69] presents techniques for efficient tile extraction and real-time stain normalization, crucial for processing large datasets efficiently.

Domain adaptation and generalization, critical challenges in deploying AI models in diverse clinical settings, are addressed by solutions like Style Transfer for Digital Pathology [80], which introduces STRAP for better domain generalization. HistoColAi [70] presents a teacher-student model paradigm for pseudo-supervised learning, enhancing tumor region classification and segmentation accuracy with limited labeled data.

Active learning and user interaction are focus areas for solutions like HistomicsML [67], which utilizes instance-based and heatmap-based active learning methods to improve classifier training efficiency. It also introduces novel metrics like the Clustering Index (CI) and Hypertrophy Index (HI) for quantifying spatial patterns and morphological changes in nuclei.

Several solutions introduce novel loss functions and optimization techniques. FGCR [78] defines anchor-based and prompt-based representations for semantic alignment and introduces multivariate cross-modal loss functions. SSL CR Histo [79] presents a novel self-supervised pretext task (Resolution Sequence Prediction) and a semi-supervised consistency training paradigm, addressing the challenge of limited labeled data in histopathology.

Spatial attention mechanisms are leveraged in solutions like PathoRICH [56] and Histo Cloud [58], which introduce Spatial Transformer Networks (STN) to improve classification and segmentation by focusing attention on relevant sub-areas using affine transformation matrices.

Other notable contributions include the use of Gaussian-based kernel density estimation (KDE) for multi-scale characterization of tumor-stroma interactions [48], the development of robust image analysis algorithms for nuclear segmentation and feature extraction [51], and the introduction of statistical modeling of color detection and removal of luminance channel [52].

This diverse contributions show the fast growing innovation of open source visual analytics for histopathology. They stress the developments of the field towards more precise, faster, and explainable analysis of histopathological images. The existence of a wide number of methods demonstrates that histopathological image analysis is a challenging task and current research tries to tackle issues like stain heterogeneity, multi-scale analysis, lack of labeled data, and the requirement of explainable solutions for clinical applications.

To see a list of specific technical contributions made by each solution as well as corresponding references, the readers are referred to Supplement Table 6. As such, this comprehensive overview of visual analytics highlights the need to advance the studies in this area with more work to be done in this field to drive the advancements of visual analytics as well as its use in clinical settings.

### 3.6. Clinical Applications and Impact

This subsection discusses the strengths, limitations, and potential clinical applications of the reviewed open-source visual analytics solutions for histopathology, drawing upon details in Supplementary Tables 8 and 10.

#### *Strengths of Open Source Visual Analytics Solutions*

The reviewed solutions exhibit several key strengths (summarized in Table 9). User-friendly interfaces were frequently highlighted (n=22 studies), facilitating adoption. High classification accuracy (n=11) and integration with AI/ML techniques (n=11) demonstrate advanced analytical capabilities. Scalability for large datasets and efficient processing (n=9 each) are crucial for handling typical histopathology data volumes. Interpretable results (n=8) and support for collaborative analysis (n=8) enhance clinical trust and teamwork. Reduced need for manual annotation (n=7), interoperability (n=7), real-time visualization (n=6), and potential for enhancing learning/standardization (n=5) were also noted strengths, collectively indicating potential to improve efficiency, accuracy, and collaboration in histopathology.

### *Limitations of Open Source Visual Analytics Solutions*

Despite strengths, these solutions face limitations (outlined in Table 10). High computational resource requirements were the most common issue (n=19 studies), potentially hindering adoption in resource-limited settings. Many solutions are limited to specific histopathology types or stains (n=14), restricting generalizability. Limited user studies or usability evaluations (n=12) suggest a need for more real-world clinical assessment.

Regulatory approval significantly influences clinical adoption. While the FDA has cleared some commercial digital pathology platforms [88], open-source tools require rigorous validation to meet regulatory standards for diagnostic use, as demonstrated by QuPath's accreditation in a UK lab [89]. Most current open-source tools (e.g., CellProfiler, Fiji, Icy) are primarily validated for

Journal Pre-proof

Table 9. Summary of Strengths of Open-Source Visual Analytics Solutions for Histopathology

Strengths	Study References
High classification accuracy	[71, 34, 41, 48, 51, 52, 56, 57, 58, 26, 28]
User-friendly interface	[35, 36, 63, 39, 64, 26, 65, 40, 41, 66, 50, 51, 52, 54, 68, 55, 57, 58, 28, 70, 27, 29]
Scalability for large datasets	[63, 39, 64, 26, 51, 57, 27, 44, 65]
Interpretable results	[34, 41, 48, 76, 61, 62, 26, 28]
Efficient processing and performance	[38, 26, 65, 43, 47, 69, 28, 44, 27]
Integration with AI/ML techniques	[64, 65, 42, 75, 46, 49, 60, 70, 26, 28, 27]
Supports collaborative analysis	[72, 44, 50, 68, 30, 70, 26, 63]
Minimal manual annotation required	[40, 73, 74, 67, 78, 79, 28]
Interoperability with various systems/datasets	[35, 63, 45, 53, 55, 26, 27]
Real-time visualization capabilities	[65, 44, 54, 69, 26, 63]
Enhances learning and standardization	[35, 37, 49, 26, 63]
Robust across different datasets/inputs	[71, 74, 52, 79, 80, 26, 28]
Supports 3D reconstruction	[77, 26, 27]
Data-efficient	[61, 62, 26, 28]
Facilitates reproducible research	[29, 26, 27]
High annotation precision	[68, 26, 28]
Handles large-scale WSIs	[78, 26, 63, 65]
High predictive performance	[56, 26, 28]
Comprehensive and flexible	[69, 26, 27, 28]
Adaptable for different tissue types and stains	[58, 26, 27]
Reduces observer bias	[59, 26, 28]
Accelerates annotation tasks	[60, 26, 28]
Improves generalizability	[80, 26, 28]

research [88]. Transitioning to clinical diagnostics necessitates extensive validation following frameworks like FDA guidelines for software as a medical device, including documented analytical validation, standard operating procedures, quality assessment participation, and version control [90, 91]. Our analysis focuses on research applications, acknowledging that clinical implementation requires additional regulatory validation specific to jurisdictional requirements, especially if the software is classified as a medical device based on intended use [88].

Other limitations include scalability issues (n=4), potential for bias/inter-observer variations (n=5), limited support for 3D/time-lapse analysis (n=2), complexity with diverse image types (n=2), interpretability challenges (n=2), and software/hardware dependencies (n=3). These highlight areas for future development.

#### *Clinical Applications and Real-World Implementation*

Potential clinical applications of open-source visual analytics solutions are diverse and impactful (summarized in Table 11). Diagnostic support and decision-making were most frequently cited (n=25 studies), highlighting a key role in assisting pathologists. Cancer research and biomarker discovery (n=18), along with general image analysis/visual analytics (n=16), are significant applications. Prognostic prediction/risk stratification (n=13), education/training (n=11), and treatment planning/personalized medicine (n=11) are also prominent. Other applications include automated tissue segmentation/quantification (n=10), collaborative research/telepathology (n=9), quality control/standardization (n=8), clinical trial support (n=6), tumor subtyping (n=5), and various other specialized uses (n=1 to n=5 each).

Case studies exemplify the potential real-world impact. QuPath [26] has been used in clinical workflows for breast cancer analysis, improving diagnostic accuracy and supporting treatment decisions through biomarker quantification [89]. Cytomine [30] facilitates multi-center telepathology, enabling remote consultations and standardized analysis. Ilastik [28] and CellProfiler [27] support clinical research through automated segmentation and quantitative analysis for biomarker discovery and risk stratification in trials. These examples underscore the potential for open-source tools to enhance precision, reproducibility, and accessibility in histopathology.

Successful implementation in clinical environments validates practical utility. Specific examples of real-world clinical deployments include:

- **QuPath Implementation:** Deployed in hospitals like Belfast Health and Social Care Trust (meeting UK accreditation standards [26]) and Leeds Teaching Hospitals NHS Trust [92]. Used for routine breast cancer biomarker quantification, showing high concordance with manual assessment and reduced inter-observer variability in Belfast [93]. Integrated into Leeds' digital pathology initiative, improving diagnostic consistency and reducing quantitative assessment time by 30% [94]. Leeds has published validation guidelines [92].
- **Cytomine in Collaborative Diagnostics:** Deployed in European hospital networks for remote consultations and standardized analysis. At University Hospital of Liège, it facilitated telepathology for regional hospitals, achieving 98.2% diagnostic

Table 10. Summary of Limitations of Open Source Visual Analytics Solutions for Histopathology

Limitations	Study References
High computational resource requirements	[71, 36, 38, 64, 26, 65, 41, 42, 73, 66, 75, 51, 57, 58, 80, 61, 62, 70, 27]
Limited to specific histopathology types or stains	[71, 26, 41, 42, 66, 46, 48, 49, 51, 52, 78, 57, 58, 70]
Limited user studies or usability evaluation	[71, 39, 26, 65, 73, 74, 66, 75, 49, 68, 57, 58]
Requires manual annotation or expert input	[34, 35, 42, 48, 59, 60]
Limited validation or need for further testing	[71, 26, 40, 41, 51, 56]
Dependence on data quality or completeness	[35, 64, 30]
Scalability issues with large datasets	[63, 67, 53, 68]
Potential for bias or inter-observer variations	[37, 43, 76, 52, 68]
Limited support for 3D or time-lapse analysis	[39, 27]
Complexity in handling diverse image types	[55, 27]
Interpretability challenges	[41, 75]
Dependence on specific software or hardware	[65, 77, 69]
Limited annotation features	[65, 54]
Requires internet connectivity	[50]
Limited to predefined workflows	[28]
Initial setup complexity	[70, 27, 29]
Dependence on community contributions	[29]
Performance variability with different inputs	[36, 47]
Limited to specific tasks or analyses	[79]
Requires large datasets for training	[80, 61, 62]

concordance between digital and conventional microscopy for primary diagnosis [30].

- **CellProfiler/Ilastik in Clinical Research:** While primarily research tools, CellProfiler was validated at Massachusetts General Hospital for quantifying immune cell infiltration in colorectal cancer specimens for clinical trials, showing 92% concordance with manual assessment and reducing analysis time >75% [95].

Common barriers to clinical implementation include integration with Laboratory Information Systems (LIS), computational resource requirements, and the need for comprehensive validation protocols [96]. Institutions like Leeds address these challenges with phased implementation strategies and thorough training/competency assessments, proving effective in overcoming adoption resistance [97]. Regulatory frameworks (e.g., FDA) guide the validation process; while open-source tools themselves may not be classified as medical devices, their use in clinical diagnosis requires adherence to these standards [98].

These examples demonstrate that successful integration is achievable, offering tangible benefits in accuracy, efficiency, and collaboration, and providing valuable roadmaps for wider adoption in hospital settings.

Table 11. Summary of Clinical Applications of Open Source Visual Analytics Solutions for Histopathology

Clinical Application	Number of Studies	References
Diagnostic support and decision-making	25	[36, 37, 64, 26, 65, 40, 41, 42, 73, 43, 44, 49, 51, 67, 78, 56, 69, 57, 61, 62, 70, 80, 35, 39, 72]
Cancer research and biomarker discovery	18	[71, 39, 26, 41, 72, 51, 53, 54, 55, 59, 70, 63, 65, 68, 79, 40, 42, 44]
Image analysis and visual analytics	16	[71, 63, 65, 74, 66, 45, 68, 79, 38, 64, 73, 58, 30, 27, 38, 46]
Prognostic prediction and risk stratification	13	[34, 65, 40, 41, 76, 56, 69, 26, 67, 59, 60, 48, 77]
Education and training	11	[35, 37, 42, 44, 50, 63, 49, 77, 29, 36, 64]
Treatment planning and personalized medicine	11	[35, 40, 41, 43, 53, 56, 26, 57, 58, 67, 59]
Automated tissue segmentation and quantification	10	[46, 52, 30, 28, 64, 66, 48, 55, 74, 45]
Collaborative research and telepathology	9	[63, 44, 61, 62, 50, 53, 29, 35, 42]
Quality control and standardization	8	[37, 66, 76, 63, 68, 79, 48, 77]
Clinical trials support	6	[26, 51, 53, 57, 40, 56]
Tumor subtyping and classification	5	[71, 40, 41, 72, 70]
Patient stratification and risk assessment	5	[34, 65, 40, 41, 56]
Prescreening and diagnosis	4	[65, 42, 66, 48]
Epidemiological studies	3	[51, 53, 59]
Precision medicine and personalized diagnosis	3	[53, 56, 59]
Multimodal image analysis	2	[38, 55]
Computational model training	2	[74, 68]
Tissue-based investigations	1	[77]
Histological biomarker quantification	1	[52]
Renal disease pathology research	1	[58]

#### 4. Discussion

Our comprehensive systematic review illuminates the rapidly evolving landscape of open-source visual analytics in histopathology, situated at the dynamic intersection of AI, computer vision, and the ongoing digital transformation of pathology. The findings underscore the substantial potential of these tools to revolutionize histopathological analysis while simultaneously highlighting critical challenges and areas demanding future research and development. The field has witnessed

significant growth, evidenced by the surge in publications over the past four years (2020-2024). This momentum reflects the increasing recognition of open-source solutions' value in advancing visual analytics, aligning with the broader movement towards open science, reproducibility, and collaborative innovation in biomedicine. While contributions are globally distributed, indicating diverse perspectives, there's a noticeable concentration in the United States. Concurrently, our analysis reveals a maturation trend, marked by a clear preference for practical, user-oriented standalone tools (n=29) over more foundational software (n=13) or frameworks (n=10). This suggests a shift from exploratory development towards creating applications readily integrable into specific pathology workflows, particularly focused on high-impact areas like cancer research (especially breast, colon, and ovarian cancers). Technologically, the field leverages a rich tapestry of innovations. The dominance of Whole Slide Images (WSIs) as the primary data modality (46/52 studies) mirrors pathology's digital shift, enabling advanced AI applications and facilitating remote collaboration. Algorithmically, CNNs, particularly ResNet variants, are prevalent, consistent with broader AI trends in medical imaging. However, the emergence of more sophisticated techniques like attention mechanisms, graph neural networks, and self-supervised learning signals a push towards more nuanced and efficient data analysis. Notably, novel approaches are tackling key issues, such as CS-CO's [71] framework for stain variability and WSI-GTFE's [34] use of topological data analysis for enhancing model interpretability – a critical need in clinical settings.

A growing emphasis on user-centric design is evident in the visualization and interaction techniques employed. The common use of heatmaps for spatial insights, interactive viewers for navigating large WSIs, dimensionality reduction plots for high-dimensional data, and annotation overlays reflects efforts to present complex results intuitively and integrate pathologist expertise. Advanced features like attention maps and real-time visualization further suggest a move towards more dynamic and collaborative digital pathology workflows.

#### Significant Challenges to Clinical Translation:

- **Computational Demands:** High computational resource requirements (reported in 19/52 studies) are a major barrier, particularly in resource-constrained settings. The gigapixel nature of WSIs demands substantial processing power, high-end GPUs, and large memory capacities for storage, real-time visualization, and analysis, often leading to processing times unsuitable for immediate diagnostic needs.
- **Generalizability and Robustness:** Many tools exhibit limitations, often being specific to certain histopathology types or stains (14 studies). Achieving robust performance across diverse tissue types, staining protocols, and scanning variations remains a challenge, necessitating the development of more generalizable solutions.
- **Clinical Integration and Validation:** Integrating these tools into established clinical workflows presents considerable difficulties. Seamless interoperability with Laboratory Information Systems (LIS), adherence to stringent regulatory standards (e.g., FDA, College of American Pathologists (CAP)), and robust data security and privacy protocols are essential but often lacking in open-source solutions. Furthermore, a critical gap exists in rigorous clinical validation; the scarcity of external validation studies (only a few identified) and comprehensive user/usability evaluations (reported in only 12 studies) hinders assessment of real-world performance and pathologist acceptance. Addressing these integration and validation challenges is paramount for translating potential benefits into routine practice.

Nevertheless, the potential clinical applications are vast. Diagnostic support and decision-making were the most cited applications (22 studies), highlighting the potential to augment pathologist accuracy and efficiency. Applications in cancer research, biomarker discovery, and prognostic prediction further underscore the role these tools can play in advancing personalized medicine.

#### 4.1. Key Findings

Our review uncovered several important findings that illuminate the current state and future directions of open-source visual analytics in histopathology:

- **Growth and Tool Preference:** Substantial growth (2020-2024) with a clear preference for standalone tools (n=29), suggesting demand for specialized, integrable solutions.
- **Data Modality and Algorithms:** WSIs dominate (46/52 studies); CNNs are the most prevalent algorithm (20 studies).
- **Visualization and Interaction:** Emphasis on user-friendly interfaces (heatmaps, interactive viewers, annotation tools).
- **Clinical Applications:** Wide potential, led by diagnostic support (22 studies).
- **Performance Reporting Gaps:** Variable reporting quality, limited calibration/multi-metric reporting hinders generalizability assessment.
- **Limited External Validation:** Scarcity raises concerns about model generalizability and practical implementation.

- **Interpretability Needs:** Lack of focus on model interpretability/explainability is a critical gap for clinical trust.

Journal Pre-proof

#### 4.2. Implications and Future Directions

Based on our findings, we propose several implications and directions for future research:

- **Flexible, Interoperable Solutions:** Focus on modular, easily integrated tools compatible with LIS/digital pathology platforms.
- **Multimodal Integration:** Explore models integrating WSIs with other data (genomics, clinical) for holistic patient views.
- **Explainable AI (XAI):** Prioritize interpretable models and clear explanations using histopathology-adapted XAI methods.
- **User-Centric Design and Validation:** Conduct extensive user studies and clinical validations to bridge the technology-adoption gap.
- **Rigorous External Validation:** Prioritize external validation in diverse settings and populations for clinical readiness.
- **Standardized Reporting:** Adopt standardized evaluation metrics (accuracy, precision, recall, calibration) for consistent comparison.
- **Collaboration:** Foster collaboration between developers, pathologists, and IT specialists.
- **Education and Training:** Invest in programs to equip pathologists with skills for utilizing these tools.

Looking ahead, the evolution of open-source visual analytics must embrace emerging technologies. Integrating multimodal data (genomics, proteomics, clinical) with imaging can enable more comprehensive disease models, improved biomarker discovery, and precise patient stratification. Concurrently, advancements in AI, particularly Large Language Models (LLMs), offer potential for bridging image analysis and natural language processing – enabling automated report generation, integration of unstructured notes, and context-aware insights. Future research should focus on robust multimodal fusion pipelines and exploring LLM utility.

Furthermore, establishing rigorous standardization and validation frameworks is crucial for clinical adoption. Adherence to standards like DICOM for digital pathology and OME-TIFF [99] for bioimaging data ensures interoperability. Aligning with regulatory guidelines (e.g., U.S. FDA [100], CAP) and implementing comprehensive clinical validation (external validation, multi-center trials, SOPs) are essential for building trust and meeting benchmarks.

Finally, careful consideration of ethical and regulatory challenges is paramount. Ensuring data privacy, security, algorithmic transparency, accountability, and governance, as highlighted by recent studies [101, 102, 103, 104], is fundamental for the safe, effective, and trustworthy deployment of AI tools in clinical workflows. Addressing these multifaceted challenges is key to realizing the full potential of open-source visual analytics in transforming pathology practice.

#### 5. Conclusion

This systematic review highlights the transformative potential of open-source visual analytics in histopathology, enhancing diagnostic accuracy, efficiency, and research capabilities. The recent publication surge underscores their growing importance. Key findings show a preference for flexible standalone tools, the dominance of WSIs reflecting digital transformation, and the effectiveness of CNNs. User-centric design focusing on advanced visualization and interaction is critical. Significant challenges persist, including high computational demands, limited generalizability, insufficient external validation, and the need for comprehensive user evaluations. Future research must prioritize developing robust, generalizable, and interpretable solutions, alongside rigorous clinical validation and standardized evaluation. Fostering collaboration and investing in education are vital. Emerging trends like integrating multimodal data and leveraging LLMs/LMMs offer promising avenues for creating more powerful and intuitive diagnostic tools, representing important future directions. Realizing the full potential requires balancing technological innovation with practical clinical applicability, ensuring these tools genuinely enhance pathology practice and improve patient care.

#### Data Availability

All data generated during this study is provided as Supplementary materials.

#### Code Availability

This study did not involve the utilization of any custom code or mathematical algorithm.

Journal Pre-proof

## Acknowledgement

The project was partially funded by Hamad Bin Khalifa Office of Vice Presidency for Research under the grant HBKU-OVPR- TG-Cycle-2-19 with the Award Number HBKU-INT-VPR-TG-02-04 and the title *Large-Scale Human-in-the-Loop Histopathology Dataset Creation and AI Models for Clinical Pediatrics Applications*. The open access is funded by Qatar National Library.

## Author Contributions

Z.A. and M.AL. contributed to conceptualization. M.A. searched the electronic databases and conducted backward and forward reference list checking. Z.A. and M.AL. performed screening, study selection, and data extraction. K.A. performed data synthesis and contributed to writing—the original draft. Z.A., M.AL., M.A. and C.C. performed writing—review, and editing. M.A., J.S and C.C. supervised the study. All authors approved the manuscript for publication and agreed to be accountable for all aspects of the work.

## Competing Interests

The authors declare no competing interests.

## Supplementary material

The online version contains supplementary material available at

## Declaration of generative AI and AI-assisted technologies in the writing process.

The manuscript was edited for English language and consistency using generative AI tools, specifically ChatGPT-4o and Claude Sonnet 3.5. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

## References

- [1] M. Titford, A short history of histopathology technique, *Journal of Histotechnology* 29 (2) (2006) 99–110. arXiv:<https://doi.org/10.1179/his.2006.29.2.99>, doi:10.1179/his.2006.29.2.99. URL <https://doi.org/10.1179/his.2006.29.2.99>
- [2] J. van der Laak, G. Litjens, F. Ciompi, Deep learning in histopathology: the path to the clinic, *Nature Medicine* 27 (5) (2021) 775–784. doi:10.1038/s41591-021-01343-4. URL <https://doi.org/10.1038/s41591-021-01343-4>
- [3] S. Al-Janabi, A. Huisman, P. J. Van Diest, Digital pathology: current status and future perspectives, *Histopathology* 61 (1) (2012) 1–9. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2559.2011.03814.x>, doi:<https://doi.org/10.1111/j.1365-2559.2011.03814.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2559.2011.03814.x>
- [4] S. Banerji, S. Mitra, Deep learning in histopathology: A review, *WIREs Data Mining and Knowledge Discovery* 12 (1) (2022) e1439. arXiv:<https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1439>, doi:<https://doi.org/10.1002/widm.1439>. URL <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1439>
- [5] C. N. Heinz, A. Echle, S. Foersch, A. Bychkov, J. N. Kather, The future of artificial intelligence in digital pathology—results of a survey across stakeholder groups, *Histopathology* 80 (7) (2022) 1121–1127. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/his.14659>, doi:<https://doi.org/10.1111/his.14659>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/his.14659>
- [6] K. Al-Thelaya, F. Joad, N. U. Gilal, W. Mifsud, G. Pintore, E. Gobetti, M. Agus, J. Schneider, HistoContours: a Framework for Visual Annotation of Histopathology Whole Slide Images, in: R. G. Raidou, B. Sommer, T. W. Kuhlen, M. Krone, T. Schultz, H.-Y. Wu (Eds.), *Eurographics Workshop on Visual Computing for Biology and Medicine*, The Eurographics Association, 2022. doi:10.2312/vcbm.20221192.
- [7] M. D. Zarella, D. Bowman, F. Aeffner, N. Farahani, A. Xthona, S. F. Absar, et al., A practical guide to whole slide imaging: A white paper from the digital pathology association, *Archives of pathology & laboratory medicine* 143 (2) (2019) 222–234. doi:10.5858/arpa.2018-0343-ra. URL <https://www.archivesofpathology.org/doi/pdf/10.5858/arpa.2018-0343-RA>
- [8] K. Al-Thelaya, M. Agus, N. U. Gilal, Y. Yang, G. Pintore, E. Gobetti, et al., Inshade: Invariant shape descriptors for visual 2d and 3d cellular and nuclear shape analysis and classification, *Computers & Graphics* 98 (2021) 105–125. doi:<https://doi.org/10.1016/j.cag.2021.04.037>. URL <https://www.sciencedirect.com/science/article/pii/S0097849321000790>
- [9] W. S. Campbell, S. H. Hinrichs, S. M. Lele, J. J. Baker, A. J. Lazenby, G. A. Talmon, et al., Whole slide imaging diagnostic concordance with light microscopy for breast needle biopsies, *Human pathology* 45 (8) (2014) 1713–1721. doi:10.1016/j.humpath.2014.04.007. URL <https://doi.org/10.1016/j.humpath.2014.04.007>
- [10] A. Corvo, M. Westenberg, R. Wimberger-Friedl, S. Fromme, M. Peeters, M. van Driel, J. van Wijk, Visual analytics in digital pathology: challenges and opportunities, in: B. Kozlíková, L. Linsen, P.-P. Vázquez, K. Lawonn, R. Raidou (Eds.), *Eurographics Workshop on Visual Computing for Biology and Medicine*, VCBM 2019, Eurographics Association, 2019, pp. 129–143, 2019 Eurographics Workshop on Visual Computing for Biology and Medicine, VCBM 2019, EG VCBM ; Conference date: 04-09-2019 Through 06-09-2019. doi:10.2312/vcbm.20191240. URL <https://conferences.eg.org/vcbm2019/>

- [11] O. Koteluk, A. Wartecki, S. Mazurek, I. Kołodziejczak, A. Mackiewicz, How do machines learn? artificial intelligence as a new era in medicine, *Journal of personalized medicine* 11 (1) (2021) 32. doi:10.3390/jpm11010032.  
URL <https://europepmc.org/articles/PMC7825660>
- [12] C. L. Srinidhi, O. Ciga, A. L. Martel, Deep neural network models for computational histopathology: A survey, *Medical Image Analysis* 67 (2021) 101813.  
doi:<https://doi.org/10.1016/j.media.2020.101813>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841520301778>
- [13] J. P. M. Rodriguez, R. Rodriguez, V. W. K. Silva, F. C. Kitamura, G. C. A. Corradi, A. C. B. de Marchi, R. Rieder, Artificial intelligence as a tool for diagnosis in digital pathology whole slide images: A systematic review, *Journal of Pathology Informatics* 13 (2022) 100138. doi:<https://doi.org/10.1016/j.jpi.2022.100138>.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922007325>
- [14] K. Al-Thelaya, N. U. Gilal, M. Alzubaidi, F. Majeed, M. Agus, J. Schneider, M. Househ, Applications of discriminative and deep learning feature extraction methods for whole slide image analysis: A survey, *Journal of Pathology Informatics* 14 (2023) 100335. doi:<https://doi.org/10.1016/j.jpi.2023.100335>.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353923001499>
- [15] D. Petriková, I. Cimrák, Survey of recent deep neural networks with strong annotated supervision in histopathology, *Computation* 11 (4) (2023). doi:10.3390/computation11040081.  
URL <https://www.mdpi.com/2079-3197/11/4/81>
- [16] S. Marletta, A. Eccher, F. M. Martelli, N. Santonicco, I. Girolami, A. Scarpa, *et al.*, Artificial intelligence-based algorithms for the diagnosis of prostate cancer: A systematic review, *American journal of clinical pathology* 161 (6) (2024) 526–534. doi:10.1093/ajcp/aqad182.  
URL <https://doi.org/10.1093/ajcp/aqad182>
- [17] M. S. Hosseini, B. E. Bejnordi, V. Q.-H. Trinh, L. Chan, D. Hasan, X. Li, *et al.*, Computational pathology: A survey review and the way forward, *Journal of Pathology Informatics* 15 (2024) 100357. doi:<https://doi.org/10.1016/j.jpi.2023.100357>.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353923001712>
- [18] P. C. Rizzo, A. Caputo, E. Maddalena, N. Caldonazzi, I. Girolami, A. P. Dei Tos, *et al.*, Digital pathology world tour, *Digital health* 9 (2023) 20552076231194551. doi:10.1177/20552076231194551.  
URL <https://europepmc.org/articles/PMC10467307>
- [19] N. Zerbe, L. O. Schwen, C. Geißler, K. Wiesemann, T. Bisson, P. Boor, *et al.*, Joining forces for pathology diagnostics with ai assistance: The empaia initiative, *Journal of Pathology Informatics* 15 (2024) 100387. doi:<https://doi.org/10.1016/j.jpi.2024.100387>.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353924000269>
- [20] P. Nagabhushana, C. Rüttsche, C. Nakas, A. B. Leichtle, Deployment of an automated method verification-graphical user interface (mv-gui) software, *BioMedInformatics* 3 (3) (2023) 632–648. doi:10.3390/biomedinformatics3030043.  
URL <https://www.mdpi.com/2673-7426/3/3/43>
- [21] L. Qiu, A. Khormali, K. Liu, Deep biological pathway informed pathology-genomic multimodal survival prediction (2023). arXiv:2301.02383.  
URL <https://arxiv.org/abs/2301.02383>
- [22] K. Ding, M. Zhou, D. N. Metaxas, S. Zhang, A. Madabhushi, P. Mousavi, *et al.*, Pathology-and-genomics multimodal transformer for survival outcome prediction, in: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, Springer Nature Switzerland, Cham, 2023, pp. 622–631.
- [23] R. Escobar Díaz Guerrero, L. Carvalho, T. Bocklitz, J. Popp, J. L. Oliveira, Software tools and platforms in digital pathology: a review for clinicians and computer scientists, *Journal of Pathology Informatics* 13 (2022) 100103. doi:<https://doi.org/10.1016/j.jpi.2022.100103>.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922006976>
- [24] M. Plass, M. Kargl, T.-R. Kiehl, P. Regimig, C. Geißler, T. Evans, *et al.*, Explainability and causability in digital pathology, *The Journal of Pathology: Clinical Research* 9 (4) (2023) 251–260. arXiv:<https://pathsocjournals.onlinelibrary.wiley.com/doi/pdf/10.1002/cjp2.322>. doi:<https://doi.org/10.1002/cjp2.322>.  
URL <https://pathsocjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cjp2.322>
- [25] J. Pallua, A. Brunner, B. Zelger, M. Schirmer, J. Haybaeck, The future of pathology is digital, *Pathology - Research and Practice* 216 (9) (2020) 153040. doi:<https://doi.org/10.1016/j.prp.2020.153040>.  
URL <https://www.sciencedirect.com/science/article/pii/S0344033819330596>
- [26] P. Bankhead, M. B. Loughrey, J. A. Fernández, Y. Dombrowski, D. G. McArt, P. D. Dunne, *et al.*, Qupath: Open source software for digital pathology image analysis, *Scientific Reports* 7 (1) (2017) 16878. doi:10.1038/s41598-017-17204-5.  
URL <https://doi.org/10.1038/s41598-017-17204-5>
- [27] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, *et al.*, Cellprofiler: image analysis software for identifying and quantifying cell phenotypes, *Genome Biology* 7 (10) (2006) R100. doi:10.1186/gb-2006-7-10-r100.  
URL <https://doi.org/10.1186/gb-2006-7-10-r100>
- [28] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, *et al.*, ilastik: interactive machine learning for (bio)image analysis, *Nature Methods* 16 (12) (2019) 1226–1232. doi:10.1038/s41592-019-0582-9.  
URL <https://doi.org/10.1038/s41592-019-0582-9>
- [29] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, *et al.*, Icy: an open bioimage informatics platform for extended reproducible research, *Nature Methods* 9 (7) (2012) 690–696. doi:10.1038/nmeth.2075.  
URL <https://doi.org/10.1038/nmeth.2075>
- [30] R. Marée, L. Rollus, B. Stévens, R. Hoyoux, G. Louppe, R. Vandaele, *et al.*, Collaborative analysis of multi-gigapixel imaging data using cytomine, *Bioinformatics* 32 (9) (2016) 1395–1401. doi:10.1093/bioinformatics/btw013.  
URL <https://doi.org/10.1093/bioinformatics/btw013>
- [31] M. M. Abdelsamea, U. Zidan, Z. Senousy, M. M. Gaber, E. Rakha, M. Ilyas, A survey on artificial intelligence in histopathology image analysis, *WIREs Data Mining and Knowledge Discovery* 12 (6) (2022) e1474. arXiv:<https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1474>. doi:<https://doi.org/10.1002/widm.1474>.  
URL <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1474>
- [32] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, *et al.*, The prisma 2020 statement: an updated guideline for reporting systematic reviews, *BMJ* 372 (2021). arXiv:<https://www.bmj.com/content/372/bmj.n71.full.pdf>. doi:10.1136/bmj.n71.  
URL <https://www.bmj.com/content/372/bmj.n71>
- [33] M. Ouzzani, H. Hammady, Z. Fedorowicz, A. Elmagarmid, Rayyan—a web and mobile app for systematic reviews, *Systematic Reviews* 5 (1) (2016) 210. doi:10.1186/s13643-016-0384-4.  
URL <https://doi.org/10.1186/s13643-016-0384-4>
- [34] J. Levy, C. Haudenschild, C. Barwick, B. Christensen, L. Vaickus, Topological Feature Extraction and Visualization of Whole Slide Images using Graph

Neural Networks, World Scientific, 2021, pp. 285–296. doi:10.1142/9789811232701\_0027.

- [35] R. Rashid, Y.-A. Chen, J. Hoffer, J. L. Muhlich, J.-R. Lin, R. Krueger, et al., Narrative online guides for the interpretation of digital-pathology images and tissue-atlas data, *Nature Biomedical Engineering* 6 (5) (2022) 515–526. doi:10.1038/s41551-021-00789-8.
- [36] E. Chelebian, C. Avenel, F. Ciompi, C. Wählby, Depicter: Deep representation clustering for histology annotation, *Computers in Biology and Medicine* 170 (2024) 108026. doi:https://doi.org/10.1016/j.compbio.2024.108026.  
URL <https://www.sciencedirect.com/science/article/pii/S0010482524001100>
- [37] H. Helin, M. Lundin, J. Lundin, P. Martikainen, T. Tammela, H. Helin, et al., Web-based virtual microscopy in teaching and standardizing gleason grading, *Human Pathology* 36 (4) (2005) 381–386. doi:https://doi.org/10.1016/j.humpath.2005.01.020.  
URL <https://www.sciencedirect.com/science/article/pii/S0046817705000481>
- [38] J. Chappelow, J. E. Tomaszewski, M. Feldman, N. Shih, A. Madabhushi, Histostitcher®: An interactive program for accurate and rapid reconstruction of digitized whole histological sections from tissue fragments, *Computerized Medical Imaging and Graphics* 35 (7) (2011) 557–567, whole Slide Image Process. doi:https://doi.org/10.1016/j.compmedimag.2011.01.010.  
URL <https://www.sciencedirect.com/science/article/pii/S0895611111000218>
- [39] S. Warchol, R. Krueger, A. J. Nirmal, G. Gaglia, J. Jessup, C. C. Ritch, et al., Visinity: Visual spatial neighborhood analysis for multiplexed tissue imaging data, *IEEE Transactions on Visualization and Computer Graphics* 29 (1) (2023) 106–116. doi:10.1109/TVCG.2022.3209378.
- [40] D. J. Ho, M. H. Chui, C. M. Vanderbilt, J. Jung, M. E. Robson, C.-S. Park, et al., Deep interactive learning-based ovarian cancer segmentation of h&e- stained whole slide images to study morphological patterns of brca mutation, *Journal of Pathology Informatics* 14 (2023) 100160. doi:https://doi.org/10.1016/j.jpi.2022.100160.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922007544>
- [41] R. J. Chen, M. Y. Lu, D. F. Williamson, T. Y. Chen, J. Lipkova, Z. Noor, et al., Pan-cancer integrative histology-genomic analysis via multimodal deep learning, *Cancer Cell* 40 (8) (2022) 865–878.e6. doi:10.1016/j.ccell.2022.07.004.  
URL <https://doi.org/10.1016/j.ccell.2022.07.004>
- [42] R. Schaer, S. Otálora, O. J. del Toro, M. Atzori, H. Müller, Deep learning-based retrieval system for gigapixel histopathology cases and the open access literature, *Journal of Pathology Informatics* 10 (1) (2019) 19. doi:https://doi.org/10.4103/jpi.jpi\_88\_18.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922003820>
- [43] M. Khened, A. Kori, H. Rajkumar, G. Krishnamurthi, B. Srinivasan, A generalized deep learning framework for whole-slide image segmentation and analysis, *Scientific Reports* 11 (1) (2021) 11579. doi:10.1038/s41598-021-90444-8.  
URL <https://doi.org/10.1038/s41598-021-90444-8>
- [44] J. Horák, K. Furmanová, B. Kozlíková, T. Brázdil, P. Holub, M. Kac’enga, et al., xopat: explainable open pathology analysis tool, *Computer Graphics Forum* 42 (3) (2023) 63–73. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14812, doi:https://doi.org/10.1111/cgf.14812.  
URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.14812>
- [45] C. Gorman, D. Punzo, I. Octaviano, S. Pieper, W. J. R. Longabaugh, D. A. Clunie, et al., Interoperable slide microscopy viewer and annotation tool for imaging data science and computational pathology, *Nature Communications* 14 (1) (2023) 1572. doi:10.1038/s41467-023-37224-2.  
URL <https://doi.org/10.1038/s41467-023-37224-2>
- [46] R. Deng, Q. Liu, C. Cui, T. Yao, J. Long, Z. Asad, et al., Omni-seg: A scale-aware dynamic network for renal pathological image segmentation, *IEEE Transactions on Biomedical Engineering* 70 (9) (2023) 2636–2644. doi:10.1109/TBME.2023.3260739.
- [47] P. J. Schüffler, G. G. Ozcan, H. Al-Ahmadie, T. J. Fuchs, Flexitilesource: An openseadragon extension for efficient whole-slide image visualization, *Journal of Pathology Informatics* 12 (1) (2021) 31. doi:https://doi.org/10.4103/jpi.jpi\_13\_21.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922001535>
- [48] J. Jiang, B. Tekin, R. Guo, H. Liu, Y. Huang, C. Wang, Digital pathology-based study of cell- and tissue-level morphologic features in serous borderline ovarian tumor and high-grade serous ovarian cancer, *Journal of Pathology Informatics* 12 (1) (2021) 24. doi:https://doi.org/10.4103/jpi.jpi\_76\_20.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353922001468>
- [49] P. Chen, X. Shi, Y. Liang, Y. Li, L. Yang, P. D. Gader, Interactive thyroid whole slide image diagnostic system using deep representation, *Computer Methods and Programs in Biomedicine* 195 (2020) 105630. doi:https://doi.org/10.1016/j.cmpb.2020.105630.  
URL <https://www.sciencedirect.com/science/article/pii/S0169260720314632>
- [50] M. N. Wicks, M. Glinka, B. Hill, D. Houghton, M. Sharghi, I. Ferreira, et al., The comparative pathology workbench: Interactive visual analytics for biomedical data, *Journal of Pathology Informatics* 14 (2023) 100328. doi:https://doi.org/10.1016/j.jpi.2023.100328.  
URL <https://www.sciencedirect.com/science/article/pii/S2153353923001426>
- [51] J. Saltz, A. Sharma, G. Iyer, E. Bremer, F. Wang, A. Jasniewski, et al., A Containerized Software System for Generation, Management, and Exploration of Features from Whole Slide Tissue Images, *Cancer Research* 77 (21) (2017) e79–e82. arXiv:https://aacrjournals.org/cancerres/article-pdf/77/21/e79/2934706/e79.pdf, doi:10.1158/0008-5472.CAN-17-0316.  
URL <https://doi.org/10.1158/0008-5472.CAN-17-0316>
- [52] J. Shu, G. E. Dolman, J. Duan, G. Qiu, M. Ilyas, Statistical colour models: an automated digital image analysis method for quantification of histological biomarkers, *BioMedical Engineering OnLine* 15 (1) (2016) 46. doi:10.1186/s12938-016-0161-6.  
URL <https://doi.org/10.1186/s12938-016-0161-6>
- [53] N. Peyrone, D. Wichadakul, Run-onco: A highly extensible software platform for cancer precision medicine (2020). doi:10.1145/3375923.3375928.  
URL <https://doi.org/10.1145/3375923.3375928>
- [54] J. Jessup, R. Krueger, S. Warchol, J. Hoffer, J. Muhlich, C. C. Ritch, et al., Scope2screen: Focus+context techniques for pathology tumor assessment in multivariate image data, *IEEE Transactions on Visualization and Computer Graphics* 28 (1) (2022) 259–269. doi:10.1109/TVCG.2021.3114786.
- [55] R. Escobar Díaz Guerrero, J. L. Oliveira, J. Popp, T. Bocklitz, Mmir: an open-source software for the registration of multimodal histological images, *BMC Medical Informatics and Decision Making* 24 (1) (2024) 65. doi:10.1186/s12911-024-02424-3.  
URL <https://doi.org/10.1186/s12911-024-02424-3>
- [56] B. Ahn, D. Moon, H.-S. Kim, C. Lee, N. H. Cho, H.-K. Choi, et al., Histopathologic image-based deep learning classifier for predicting platinum-based treatment responses in high-grade serous ovarian cancer, *Nature Communications* 15 (1) (2024) 4253. doi:10.1038/s41467-024-48667-6.  
URL <https://doi.org/10.1038/s41467-024-48667-6>
- [57] S. Mimar, A. S. Paul, N. Lucarelli, S. Border, A. Naglah, L. Barisoni, et al., CompRePS: an automated cloud-based image analysis tool to democratize AI in digital pathology, in: J. E. Tomaszewski, A. D. Ward (Eds.), *Medical Imaging 2024: Digital and Computational Pathology*, Vol. 12933, International Society for Optics and Photonics, SPIE, 2024, p. 129330Z. doi:10.1117/12.3008469.  
URL <https://doi.org/10.1117/12.3008469>
- [58] B. Lutnick, D. Manthey, J. U. Becker, B. Ginley, K. Moos, J. E. Zuckerman, et al., A user-friendly tool for cloud-based whole slide image segmentation with examples from renal histopathology, *Communications Medicine* 2 (1) (2022) 105. doi:10.1038/s43856-022-00138-z.

URL <https://doi.org/10.1038/s43856-022-00138-z>

- [59] F. Varghese, A. B. Bukhari, R. Malhotra, A. De, Ihc profiler: An open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples, *PLOS ONE* 9 (5) (2014) 1–11. doi:10.1371/journal.pone.0096801.  
URL <https://doi.org/10.1371/journal.pone.0096801>
- [60] R. Hollandi, Á. Diósi, G. Hollandi, N. Moshkov, P. Horváth, Annotatorj: an imagej plugin to ease hand annotation of cellular compartments, *Molecular Biology of the Cell* 31 (20) (2020) 2179–2186. doi:10.1091/mbc.E20-02-0156.  
URL <https://doi.org/10.1091/mbc.E20-02-0156>
- [61] M. Y. Lu, D. F. K. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, F. Mahmood, Data-efficient and weakly supervised computational pathology on whole-slide images, *Nature Biomedical Engineering* 5 (6) (2021) 555–570. doi:10.1038/s41551-020-00682-w.  
URL <https://doi.org/10.1038/s41551-020-00682-w>
- [62] J. Lipkova, T. Y. Chen, M. Y. Lu, R. J. Chen, M. Shady, M. Williams, et al., Deep learning-enabled assessment of cardiac allograft rejection from endomyocardial biopsies, *Nature Medicine* 28 (3) (2022) 575–582. doi:10.1038/s41591-022-01709-2.  
URL <https://doi.org/10.1038/s41591-022-01709-2>
- [63] D. A. Gutman, M. Khalilia, S. Lee, M. Nalishnik, Z. Mullen, J. Beezley, et al., The Digital Slide Archive: A Software Platform for Management, Integration, and Analysis of Histology for Cancer Research, *Cancer Research* 77 (21) (2017) e75–e78. doi:10.1158/0008-5472.CAN-17-0629.
- [64] M. Stritt, A. K. Stalder, E. Vezzali, Orbit image analysis: An open-source whole slide image analysis tool, *PLOS Computational Biology* 16 (2) (2020) 1–19. doi:10.1371/journal.pcbi.1007313.  
URL <https://doi.org/10.1371/journal.pcbi.1007313>
- [65] A. Pedersen, M. Valla, A. M. Bofin, J. P. De Frutos, I. Reinertsen, E. Smistad, Fastpathology: An open-source platform for deep learning-based research and decision support in digital pathology, *IEEE Access* 9 (2021) 58216–58229. doi:10.1109/ACCESS.2021.3072231.
- [66] A. Janowczyk, R. Zuo, H. Gilmore, M. Feldman, A. Madabhushi, Histoqc: An open-source quality control tool for digital pathology slides, *JCO Clinical Cancer Informatics* (3) (2019) 1–7, PMID: 30990737. arXiv:<https://doi.org/10.1200/CCI.18.00157>, doi:10.1200/CCI.18.00157.  
URL <https://doi.org/10.1200/CCI.18.00157>
- [67] M. Nalishnik, M. Amgad, S. Lee, S. H. Halani, J. E. Velazquez Vega, D. J. Brat, et al., Interactive phenotyping of large-scale histology imaging data with histomicsml, *Scientific Reports* 7 (1) (2017) 14588. doi:10.1038/s41598-017-15092-3.  
URL <https://doi.org/10.1038/s41598-017-15092-3>
- [68] P. Puttapirat, H. Zhang, J. Deng, Y. Dong, J. Shi, P. Lou, et al., Openhi: open platform for histopathological image annotation, *International Journal of Data Mining and Bioinformatics* 22 (4) (2019) 328–349. arXiv:<https://www.inderscienceonline.com/doi/pdf/10.1504/IJDMB.2019.10139>, doi:10.1504/IJDMB.2019.10139.  
URL <https://www.inderscienceonline.com/doi/abs/10.1504/IJDMB.2019.10139>
- [69] J. M. Dolezal, S. Kochanny, E. Dyer, S. Ramesh, A. Srisuwananukorn, M. Sacco, et al., Slideflow: deep learning for digital histopathology with real-time whole-slide visualization, *BMC Bioinformatics* 25 (1) (2024) 134. doi:10.1186/s12859-024-05758-x.  
URL <https://doi.org/10.1186/s12859-024-05758-x>
- [70] C. C. Pulgarín-Ospina, R. del Amor, A. Colomera, J. Silva-Rodríguez, V. Naranjo, HistoColai: An open-source web platform for collaborative digital histology image annotation with ai-driven predictive integration (2023). arXiv:2307.07525, doi:<https://arxiv.org/abs/2307.07525>.
- [71] P. Yang, X. Yin, H. Lu, Z. Hu, X. Zhang, R. Jiang, H. Lv, Cs-co: A hybrid self-supervised visual representation learning method for h&e-stained histopathological images, *Medical Image Analysis* 81 (2022) 102539. doi:<https://doi.org/10.1016/j.media.2022.102539>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841522001864>
- [72] S. Deshpande, M. Dawood, F. Minhas, N. Rajpoot, Synclay: Interactive synthesis of histology images from bespoke cellular layouts, *Medical Image Analysis* 91 (2024) 102995. doi:<https://doi.org/10.1016/j.media.2023.102995>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841523002554>
- [73] Y. Lin, Z. Qu, H. Chen, Z. Gao, Y. Li, L. Xia, et al., Nuclei segmentation with point annotations from pathology images via self-supervised learning and co-training, *Medical Image Analysis* 89 (2023) 102933. doi:<https://doi.org/10.1016/j.media.2023.102933>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841523001937>
- [74] N. Alemi Koohbanani, M. Jahanifar, N. Zamani Tajadin, N. Rajpoot, Nuclick: A deep learning framework for interactive segmentation of microscopic images, *Medical Image Analysis* 65 (2020) 101771. doi:<https://doi.org/10.1016/j.media.2020.101771>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841520301353>
- [75] A. K. Tyagi, C. Mohapatra, P. Das, G. Makharia, L. Mehra, P. AP, Mausam, Degpr: Deep guided posterior regularization for multi-class cell detection and counting (June 2023).
- [76] N. Sato, E. Uchino, R. Kojima, M. Sakuragi, S. Hiragi, S. Minamiguchi, et al., Evaluation of kidney histological images using unsupervised deep learning, *Kidney International Reports* 6 (9) (2021) 2445–2454. doi:<https://doi.org/10.1016/j.ekir.2021.06.008>.  
URL <https://www.sciencedirect.com/science/article/pii/S2468024921012341>
- [77] A. Shen, F. Wang, S. Paul, D. Bhuvanapalli, J. Alayof, A. B. Farris, et al., An integrative web-based software tool for multi-dimensional pathology whole-slide image analytics, *Physics in Medicine & Biology* 67 (22) (2022) 224001. doi:10.1088/1361-6560/ac8fde.  
URL <https://dx.doi.org/10.1088/1361-6560/ac8fde>
- [78] D. Hu, Z. Jiang, J. Shi, F. Xie, K. Wu, K. Tang, et al., Histopathology language-image representation learning for fine-grained digital pathology cross-modal retrieval, *Medical Image Analysis* 95 (2024) 103163. doi:<https://doi.org/10.1016/j.media.2024.103163>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841524000884>
- [79] C. L. Srinidhi, S. W. Kim, F.-D. Chen, A. L. Martel, Self-supervised driven consistency training for annotation efficient histopathology image analysis, *Medical Image Analysis* 75 (2022) 102256. doi:<https://doi.org/10.1016/j.media.2021.102256>.  
URL <https://www.sciencedirect.com/science/article/pii/S1361841521003017>
- [80] R. Yamashita, J. Long, S. Banda, J. Shen, D. L. Rubin, Learning domain-agnostic visual representation for computational pathology using medically-irrelevant style transfer augmentation, *IEEE Transactions on Medical Imaging* 40 (12) (2021) 3945–3954. doi:10.1109/TMI.2021.3101985.
- [81] M. Y. Lu, B. Chen, D. F. K. Williamson, R. J. Chen, M. Zhao, A. K. Chow, et al., A multimodal generative ai copilot for human pathology, *Nature* (Jun 2024). doi:10.1038/s41586-024-07618-3.  
URL <https://doi.org/10.1038/s41586-024-07618-3>
- [82] Y. Sun, C. Zhu, S. Zheng, K. Zhang, L. Sun, Z. Shui, et al., Pathasst: A generative foundation ai assistant towards artificial general intelligence of pathology, *Proceedings of the AAAI Conference on Artificial Intelligence* 38 (5) (2024) 5034–5042. doi:10.1609/aaai.v38i5.28308.  
URL <https://ojs.aaai.org/index.php/AAAI/article/view/28308>
- [83] D. Ferber, G. Wölflein, I. C. Wiest, M. Liger, S. Sainath, N. Ghaffari Laleh, et al., In-context learning enables multimodal large language models to classify cancer pathology images, *Nature Communications* 15 (1) (2024) 10104. doi:10.1038/s41467-024-51465-9.  
URL <https://doi.org/10.1038/s41467-024-51465-9>

- [84] R. Kaczmarczyk, T. I. Wilhelm, R. Martin, J. Roos, Evaluating multimodal ai in medical diagnostics, *npj Digital Medicine* 7 (1) (2024) 205. doi:10.1038/s41746-024-01208-3. URL <https://doi.org/10.1038/s41746-024-01208-3>
- [85] Q. Zhou, T. M. Dang, W. Zhong, Y. Guo, H. Ma, S. Na, J. Huang, Mllm4pue: Toward universal embeddings in computational pathology through multimodal llms (2025). arXiv:2502.07221. URL <https://arxiv.org/abs/2502.07221>
- [86] P. Xia, K. Zhu, H. Li, T. Wang, W. Shi, S. Wang, et al., Mmed-rag: Versatile multimodal rag system for medical vision language models (2024). arXiv: 2410.13085. URL <https://arxiv.org/abs/2410.13085>
- [87] M. Asadi-Aghbolaghi, A. Darbandsari, A. Zhang, A. Contreras-Sanz, J. Boschman, P. Ahmadvand, et al., Learning generalizable ai models for multi-center histopathology image classification, *npj Precision Oncology* 8 (1) (2024) 151. doi:10.1038/s41698-024-00652-4. URL <https://doi.org/10.1038/s41698-024-00652-4>
- [88] G. A. Matthews, C. McGenity, D. Bansal, D. Treanor, Public evidence on ai products for digital pathology, *npj Digital Medicine* 7 (1) (2024) 300. doi: 10.1038/s41746-024-01294-3. URL <https://doi.org/10.1038/s41746-024-01294-3>
- [89] P. Bankhead, J. A. Fernández, D. G. McArt, D. P. Boyle, G. Li, M. B. Loughrey, et al., Integrated tumor identification and automated scoring minimizes pathologist involvement and provides new insights to key biomarkers in breast cancer, *Laboratory investigation* 98 (1) (2018) 15–26.
- [90] J. C. Goldsack, A. Coravos, J. P. Bakker, B. Bent, A. V. Dowling, C. Fitzer-Attas, A. Godfrey, et al., Verification, analytical validation, and clinical validation (v3): the foundation of determining fit-for-purpose for biometric monitoring technologies (biomets), *npj digital Medicine* 3 (1) (2020) 55.
- [91] A. E. Lottes, K. J. Cavanaugh, Y. Y.-F. Chan, V. J. Devlin, C. J. Goergen, R. Jean, et al., Navigating the regulatory pathway for medical devices—a conversation with the fda, clinicians, researchers, and industry experts, *Journal of Cardiovascular Translational Research* 15 (5) (2022) 927–943.
- [92] D. Treanor, B. Williams, The leeds guide to digital pathology, Leeds: The Leeds Teaching Hospitals NHS, University of Leeds (2019).
- [93] M. Humphries, P. Maxwell, M. Salto-Tellez, Qupath: The global impact of an open source digital pathology system, *Computational and structural biotechnology journal* 19 (2021) 852–859.
- [94] B. J. Williams, A. Hanby, R. Millican-Slater, A. Nijhawan, E. Verghese, D. Treanor, Digital pathology for the primary diagnosis of breast histopathological specimens: an innovative validation and concordance study on digital pathology validation and training, *Histopathology* 72 (4) (2018) 662–671.
- [95] P. D. Caie, R. E. Walls, A. Ingleston-Orme, S. Daya, T. Houslay, R. Eagle, et al., High-content phenotypic profiling of drug response signatures across distinct cancer cells, *Molecular cancer therapeutics* 9 (6) (2010) 1913–1926.
- [96] D. G. Pinto, A. Bychkov, N. Tsuyama, J. Fukuoka, C. Eloy, Real-world implementation of digital pathology: results from an intercontinental survey, *Laboratory Investigation* 103 (12) (2023) 100261.
- [97] B. J. Williams, D. Bottoms, D. Treanor, Future-proofing pathology: the case for clinical adoption of digital pathology, *Journal of clinical pathology* 70 (12) (2017) 1010–1018.
- [98] S. Cross, P. Furness, L. Igali, D. Snead, D. Treanor, Best practice recommendations for implementing digital pathology january 2018, The Royal College of Pathologists web site (2018).
- [99] S. Besson, R. Leigh, M. Linkert, C. Allan, J.-M. Burel, M. Carroll, D. Gault, R. Gozim, S. Li, D. Lindner, et al., Bringing open data to whole slide imaging, in: *European Congress on Digital Pathology*, Springer, 2019, pp. 3–10.
- [100] E. Abels, L. Pantanowitz, Current state of the regulatory trajectory for whole slide imaging devices in the usa, *Journal of pathology informatics* 8 (1) (2017) 23.
- [101] K. Reinhardt, Trust and trustworthiness in ai ethics, *AI and Ethics* 3 (3) (2023) 735–744. doi:10.1007/s43681-022-00200-5. URL <https://doi.org/10.1007/s43681-022-00200-5>
- [102] C. Chauhan, R. R. Gullapalli, Chapter 9 - ethics of ai in pathology: Current paradigms and emerging issues, in: C. Chauhan, S. Cohen (Eds.), *Artificial Intelligence in Pathology (Second Edition)*, second edition Edition, Elsevier, 2025, pp. 159–180. doi:<https://doi.org/10.1016/B978-0-323-95359-7.00009-1>. URL <https://www.sciencedirect.com/science/article/pii/B9780323953597000091>
- [103] N. Kazimova, Data governance and privacy challenges in the digital healthcare revolution, SSRN Available at SSRN: <https://ssrn.com/abstract=4626179> or <http://dx.doi.org/10.2139/ssrn.4626179> (2023).
- [104] S. Shafi, A. V. Parwani, Artificial intelligence in diagnostic pathology, *Diagnostic Pathology* 18 (1) (2023) 109. doi:10.1186/s13000-023-01375-z. URL <https://doi.org/10.1186/s13000-023-01375-z>

**Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof