POLITECNICO DI TORINO
Repository ISTITUZIONALE

SlideInspect: From Pixel-Level Artifact Detection to Actionable Quality Metrics in Digital Pathology

Original
SlideInspect: From Pixel-Level Artifact Detection to Actionable Quality Metrics in Digital Pathology / Scotto, M., Patti, R., L'Imperio, V., Fraggetta, F., Molinari, F., Salvi, M.. - In: INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY. - ISSN 0899-9457. - 36:1(2026). [10.1002/ima.70292]

Availability:
This version is available at: 11583/3006457 since: 2026-01-11T23:16:22Z

Publisher:
Wiley

Published
DOI:10.1002/ima.70292

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

17 November 2026

Check for updates

International Journal of Imaging Systems and Technology

WILEY

RESEARCH ARTICLE OPEN ACCESS

SlideInspect: From Pixel-Level Artifact Detection to Actionable Quality Metrics in Digital Pathology

Manuela Scotto[1] | Roberta Patti[1] | Vincenzo L'imperio[2] | Filippo Fraggetta[3] | Filippo Molinari[1] | Massimo Salvi[1]

[1]Department of Electronics and Telecommunications, Biolab, PoliTo[BIO]Med Lab, Politecnico di Torino, Turin, Italy | [2]Department of Medicine and Surgery, Pathology, Fondazione IRCCS San Gerardo Dei Tintori, University of Milan-Bicocca, Monza, Italy | [3]UOC di Anatomia Patologica, Caltagirone, Italy

Correspondence: Massimo Salvi (massimo.salvi@polito.it)

Received: 23 May 2025 | **Revised:** 26 November 2025 | **Accepted:** 4 January 2026

Keywords: artifact segmentation | digital pathology | quality controls | supervised learning | whole-slide imaging

ABSTRACT

The presence of artifacts in whole slide images (WSIs), such as tissue folds, air bubbles, and out-of-focus regions, can significantly impact WSI digitization, pathologists' evaluation, and the accuracy of downstream analyses. We present SlideInspect, a novel AI-based framework for comprehensive artifact detection and quality control in digital pathology. Our system leverages deep learning techniques to segment multiple artifact types across diverse tissue types and staining methods. SlideInspect provides a hierarchical output: a color-coded slide quality indicator (green, yellow, red) with recommended actions (no action, re-scan, re-mount, re-cut) based on artifact type and extent, and pixel-level segmentation masks for detailed analysis. The system operates at multiple magnifications (1.25× for tissue segmentation, 5× for artifact detection) and also incorporates stain quality assessment for histological stain evaluation. We validated SlideInspect on a large, multi-centric, multi-scanner dataset of over 3000 WSIs, demonstrating robust performance across different tissue types, staining methods, and scanning platforms. The system achieves high segmentation accuracy for various artifacts while maintaining computational efficiency (average processing time: 72.7 s per WSI). Pathologist evaluations confirmed the clinical relevance and accuracy of SlideInspect's quality assessments. By providing actionable insights at multiple levels of granularity, SlideInspect significantly improves the efficiency and standardization of digital pathology workflows. Its vendor-agnostic design and multi-stain capability make it suitable for integration into diverse clinical and research settings.

1 | Introduction

In recent years, advanced imaging systems and Whole-Slide Image (WSI) scanners have revolutionized pathology, bringing it into the digital age [1, 2]. These systems enable rapid acquisition of high-resolution pyramidal images of entire histopathology slides at various magnifications for convenient display on computer monitors. Digital pathology offers numerous advantages over traditional systems, including centralized data organization, easy sharing of image data for telepathology, and facilitation of worldwide interspecialty communication for remote diagnosis and consultations [3].

Furthermore, digital pathology reduces the risk of physical slides being damaged or lost while creating precise, reproducible tissue-derived readouts that reduce inter-pathologist variability [1]. Digital images also enable automated algorithms for quantitative assessment of histologic features, morphologic patterns, and biologically relevant regions of interest [4], enhancing efficiency and streamlining procedures for better patient outcomes [1].

The manual slide preparation process can introduce various artifacts that compromise image quality [5]. These artifacts fall into two categories: those occurring during slide preparation

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *International Journal of Imaging Systems and Technology* published by Wiley Periodicals LLC.

and those during digitization. Slide preparation artifacts include tissue folds, mounting artifacts (bubbles, dust, pen marks, coverslip issues), and pen marks, while digitization artifacts primarily include out-of-focus regions (Figure 1) [5]. Tissue folds arise when thin tissue sections are mounted, creating thicker, overlapping regions that absorb more color. Air bubbles occur during mounting when the coverslip is placed, altering the underlying tissue's appearance. Pen marks used to indicate features of interest can mask parts of the tissue.

During digitization, out-of-focus regions occur when specimens are not properly aligned with the scanner's focal plane, resulting in blurred images affecting part or all of the slide [6]. These artifacts negatively impact diagnostic accuracy and hinder the reproducibility of results [7, 8], making their detection and correction crucial for ensuring reliable histological images. Thus, it is crucial to minimize these artifacts and develop effective strategies for artifact detection and correction to ensure high-quality, reliable histological images.

The presence of low-quality tissue causes delays in pathology reporting, potentially requiring slide re-preparation or repeat scanning [9]. Artifacts can also alter diagnostically relevant tissue regions, rendering quantitative analysis unreliable [10, 11]. A quality control (QC) system is therefore essential to ensure accurate and timely reporting of pathology results. Such a system can identify low-quality slides and, based on the type and severity of artifacts, determine whether slides need re-preparation or re-scanning. Furthermore, a QC system can identify regions to avoid during computational analysis for more reliable results, which is particularly important for artificial intelligence (AI) systems that may fail to make accurate predictions in regions with appearance-altering artifacts.

This study introduces SlideInspect, a fully automated AI-based framework providing precise and quantitative feedback for enhancing quality control in digital pathology. The main contributions of this paper are:

- Development of a multi-tissue, multi-stain framework that performs multi-class artifact segmentation in WSIs using two recent segmentation networks. These networks recognize histological tissue and artifacts in slides at different scales.
- Segmentation of multiple artifacts enables stratification of system outputs in terms of WSI quality (color codes: green, yellow, red) and actions to take on the slide (no action,

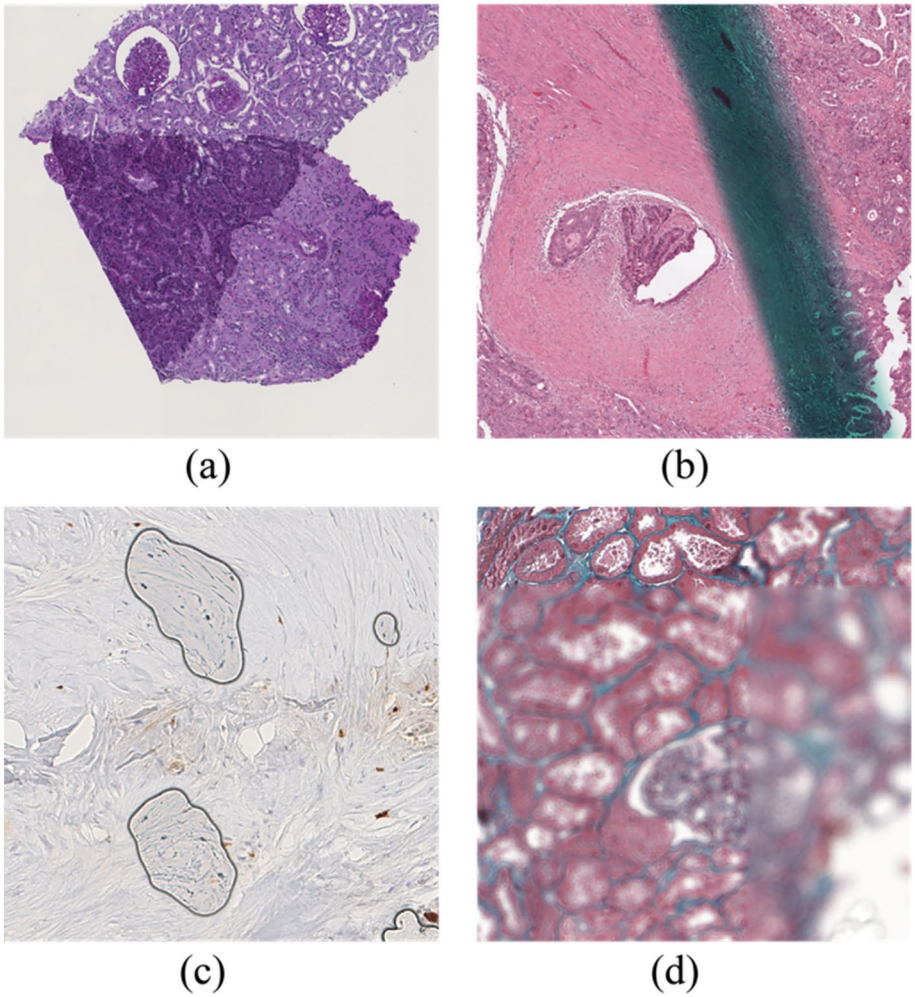


FIGURE 1 | Examples of common artifacts in Whole-Slide Images (WSIs). (a) Tissue folds: Overlapping tissue regions with excessive stain absorption; (b) Pen markers: Artificially highlighted regions obscuring underlying structures; (c) Mounting artifacts: Including bubbles, dust, pen marks, and coverslip issues that can distort tissue visualization; (d) Out-of-focus region: Blurring caused by focal plane misalignment during scanning.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

re-scan, re-mount, re-cut). SlideInspect also provides pixel-based and quantitative results for all the artifacts in the slide. The framework also evaluates the intensity/contrast of histological stains and the spatial distribution of artifacts with respect to histological tissue.

- Our framework is inherently explainable, providing a hierarchical output at different levels of complexity, starting from an immediate score (green/yellow/red) and extending to a quantitative, pixel-based analysis for each type of artifact, along with visual output/feedback.
- SlideInspect was trained and developed using a large dataset of annotated images and can accurately detect and segment artifacts in diverse image contexts (biopsies, wedges, etc.). Our method was tested on external, multi-center, multi-scanner datasets that we publicly release along with manual annotations. The results demonstrate the efficacy of our system in processing multi-tissue, multi-stain images from various sources.

The rest of this paper is organized as follows: Section 2 presents an overview of the current approaches for artifact segmentation in digital pathology; Section 3 provides an exhaustive description of the proposed method; experimental results are reported in Section 4; Section 5 and 6 discuss the work as a whole.

2 | Related Works

QC is a critical aspect of digital pathology workflow that ensures the reliability of downstream analysis. Currently, QC processes in pathology laboratories primarily involve manual inspection of prepared slides by experts to evaluate their appearance. This process is repetitive, time-consuming, prone to human error, and subject to intra- and inter-operator variability. Moreover, while artifacts may not always affect a pathologist's diagnosis, they can significantly impede the performance of automatic algorithms. To address this issue, various automated methods have been developed for artifact detection in digital pathology.

Several tools have been developed to address QC in digital pathology. HistoQC [10, 12], a non-deep learning tool, detects artifacts using pixel analysis techniques and has been used for domain shift analysis. Various research efforts [6, 9, 13–17] offer artifact detection capabilities but are often limited to specific domains or artifact types.

More recently, multi-artifact segmentation models have been developed, some of which suggest appropriate operations based on the established quality of the slide. Smit et al. [18] proposed an artifact segmentation module that identifies different types of artifacts and transforms the output into one of four actions using a decision tree. Although this approach provides actionable insights based on the detected artifacts, the study used only a small number of WSIs with limited staining and histological diversity. Haghighat et al. [19] presented PathProfiler, an AI-based tool that automates the quality assessment of prostate specimens. PathProfiler provides a qualitative output indicating the overall slide quality but does not offer pixel-level segmentation of artifacts or quantitative measures of artifact extent. Patil et al. [20] developed HistoROI, a robust lightweight classifier for segregating WSI patches into six categories, including artifacts, using a human-in-the-loop training paradigm. However, HistoROI was developed for breast tissue, and it operates as a patch-level classifier rather than a pixel-wise segmentation tool, limiting its ability to precisely delineate artifact boundaries.

Recent comprehensive approaches include GrandQC [21], which performs tissue segmentation and segmentation of tissue without artifacts. This tool allows for multi-class artifact detection and has been validated across multiple institutions and scanner systems. While GrandQC provides excellent segmentation accuracy, it does not evaluate stain quality or provide direct actionable recommendations based on detected artifacts. Another notable approach is the artifact detection pipeline by Kanwal et al. [22], which implements a mixture of experts (MoE) scheme for detecting five types of artifacts.

The only available tool that provides an assessment of the slide in terms of overall quality is magIQ [23]. This commercial QC software analyzes images and provides normalized numerical scores for sample detection and focus. The tool categorizes slides into 'To keep,' 'To review,' or 'To re-scan' based on customizable thresholds. Although it offers low computation times, magIQ performs patch-level classification rather than pixel-level segmentation and takes into account only focus artifacts.

Despite these advances, several limitations persist in automated solutions for QC in digital pathology. Many methods use only a small number of WSIs [18] with limited staining and histological diversity [9, 15, 19]. Some state-of-the-art methods work with full-resolution images, making them computationally expensive [6, 9, 13, 15, 16]. Finally, some approaches provide only a qualitative or semi-quantitative output rather than pixel-level segmentation of artifacts [6, 16, 19]. Most existing tools focus on physical artifacts but fail to evaluate stain quality and intensity, which are critical factors affecting both manual and automated analysis. Neither GrandQC [21] nor other current approaches incorporate comprehensive stain quality metrics in their evaluation framework. Additionally, current tools provide segmentation maps or artifact classifications but rarely translate these findings into specific, automated actions to be taken (re-scan, re-mount, re-cut). This gap prevents integration into automated digital pathology workflows. Table 1 reports approaches used for artifact segmentation and quality control in digital pathology.

The objective of this study is to address the limitations of current QC algorithms in digital pathology. We propose a novel multi-tissue, multi-stain algorithm that has been extensively tested on diverse datasets obtained from multiple centers. Our solution accurately segments various artifact types while operating efficiently at both low (1.25×) and medium (5×) resolutions, balancing accuracy and computational demands. The system provides hierarchical output—beginning with a simple traffic light indicator (green, yellow, red) linked to recommended actions (no action, re-scan, re-mount, re-cut), followed by detailed semantic masks and quantitative metrics on artifact extent and stain quality. We believe this simple, interpretable, and actionable feedback provided in a stepwise manner could be highly appreciated by operators and pathologists, allowing quick assessment of slide quality before diving into the details.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Comparison of artifact detection approaches in digital pathology.

Authors	# organs	Staining	Magnification	Detected artifacts
Hosseini et al. [6]	n/a	1 (H&E)	40×	OOF
Senaras et al. [9]	n/a	2 (H&E, IHC)	40×	OOF
Lopez et al. [13]	11	2 (H&E, IHC)	20×	OOF
Babaie et al. [15]	n/a	1 (H&E)	20×	TF
Kohlberger et al. [16]	> 30	2 (H&E, IHC)	40×	OOF
Smit et al. [18]	9	3 (H&E, IHC, PAS)	2.5×	TF, OOF, MA
Haghighat et al. [19]	1	1 (H&E)	5×	TF, OOF, MA
Patil et al. [20]	1	1 (H&E)	10×	TF, OOF, MA, PN
Weng et al. [21]	19	1 (H&E)	5×–10×	TF, OOF, MA, PN
Kanwal et al. [22]	1	1 (H&E)	40×	TF, OOF, MA
SlideInspect (proposed)	8	4 (H&E, IHC, PAS, TRIC)	1.25×–5×	TF, OOF, MA

Abbreviations: H&E: Hematoxylin and Eosin; IHC: Immunohistochemistry; MA: mounting artifacts; OOF: out-of-focus; PAS: Periodic acid–Schiff; PN: pen marks; TF: Tissue folds; TRIC: Trichrome.

TABLE 2 | Overall dataset composition.

	Stains				
Tissue	**H&E**	**IHC**	**PAS**	**TRIC**	***Total***
Adrenal	29	—	—	—	29
Breast	507	515	—	—	1022
Colon	340	—	—	—	340
Kidney	12	—	98	87	197
Liver	403	—	—	—	403
Lung	181	—	101	—	282
Myocardial	—	—	—	98	98
Prostate	511	154	—	—	665
Total	*1983*	*669*	*199*	*185*	*3036*

Note: A total of 3036 slides were used in this study, coming from eight tissues and four different stainings.
Abbreviations: H&E, Hematoxylin and Eosin; IHC, Immunohistochemistry; PAS, Periodic acid–Schiff; TRIC, Trichrome.

3 | Materials and Methods

3.1 | Datasets

For this study, a total of 3036 WSIs were utilized (Table 2). The staining methods included hematoxylin and eosin (H&E), immunohistochemistry (IHC) with DAB staining, periodic acid–Schiff (PAS), and trichrome (TRIC). The dataset was divided into training (2160 WSIs), validation (435 WSIs), and testing (441 WSIs) subsets, with the distribution of staining methods shown in Figure 2. Part of the training set and test set were also used to develop the classifier that recommends action to take on every slide (no action, re-scan, re-mount, re-cut).

The dataset included eight different tissue types: adrenal, breast, colon, kidney, liver, lung, myocardial, and prostate. The breakdown of the number of WSIs for each tissue and stain combination is provided in the Supporting Information (Table S1). The WSIs originated from a total of five centers, three located in Italy and two from The Cancer Genome Atlas (TCGA) and The Cancer Imaging Archive (TCIA) public repositories. Importantly, the test set contained slides prepared and digitized at centers not represented in the training and validation sets. The WSIs of our multi-centric dataset were scanned using different scanners, and the analyzed formats included Aperio (.svs), Hamamatsu (.ndpi), and 3DHistech (*.mrxs). More information on the dataset composition, including the number of WSIs per center and the distribution of tissues across the training, validation, and test subsets, is provided in the Supporting Information (Figure S1 and Table S2).

To create the manual annotations, an expert operator (M.Sc.) labeled the following classes with different labels: histological tissue, out-of-focus regions, mounting artifacts (bubbles, dust, pen marks, coverslip issues), and tissue folds. Since the WSIs originated from different vendors, multiple software tools were used for manual annotations. Aperio ImageScope (version 12.4.3.5008) was used for *.svs files, NDP.view2 (version 2.9.29) for *.ndpi files, and QuPath (version 0.3.2) for *.mrxs and .tif files.

3.2 | The SlideInspect Framework

In this paper, we present a novel fully automated method based on deep learning for multi-class segmentation of artifacts in digital pathology. This approach employs a combination of advanced techniques, including the K-Net architecture with Swin Transformer backbone [24], and a machine learning classifier. Figure 3 summarizes the overall flowchart of our approach. Our method follows a three-branch procedure: the first branch is designed for tissue segmentation at low magnification (1.25×), while the second and third branches operate at higher magnification (5×) and perform artifact segmentation and stain quality assessment, respectively. Finally, the system provides a

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

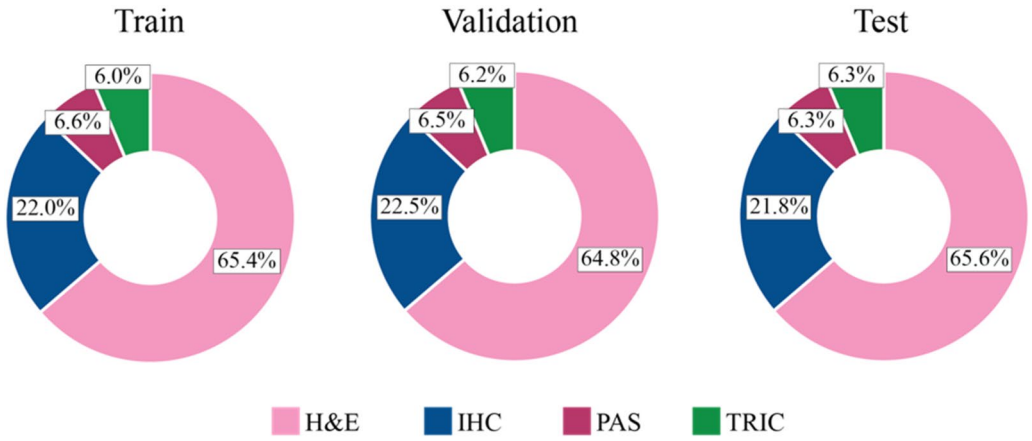


FIGURE 2 | Distribution of staining types in the training, validation, and test datasets.

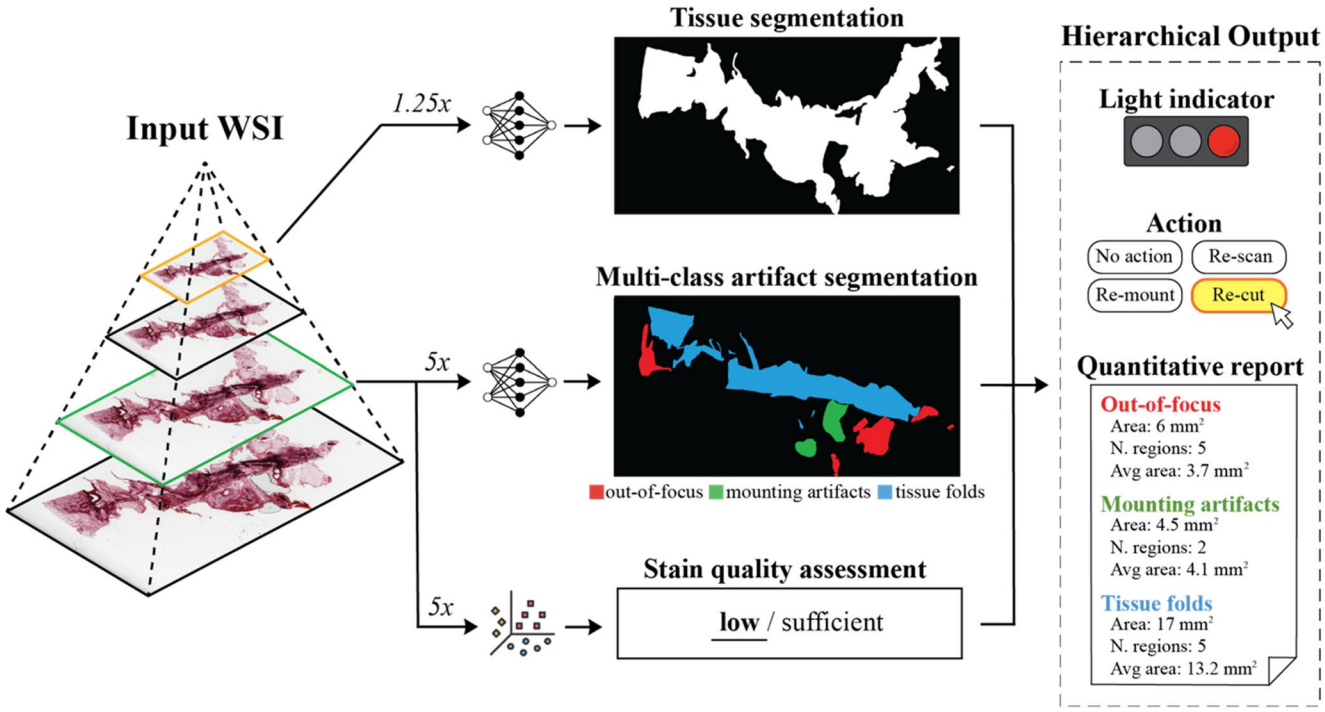


FIGURE 3 | Pipeline of the proposed SlideInspect algorithm. Two magnification levels are extracted from the pyramidal image: 1.25× and 5×. The 1.25× image is used for tissue segmentation, while the 5× image is employed for multi-class artifact detection and stain quality assessment. Features extracted from the tissue, artifacts, and stain analysis are then fed into a machine learning classifier that generates the suggested action for the current slide. The system also provides a qualitative output in the form of a light indicator (green, yellow, red) and quantitative outputs about the distribution of artifacts across the slide.

hierarchical output that includes a simple light indicator (green: acceptable quality, yellow: some artifacts detected, and red: insufficient quality), the suggested action to perform on the slide, and quantitative outputs in terms of the percentage of class-wise artifacts within the histological tissue. The system classifies each WSI into one of four levels based on the detected artifacts and their impact:

- Level 0 (no action): The slide quality is sufficient, with minimal or no artifacts affecting the tissue.
- Level 1 (re-scan): Out-of-focus areas are detected overlapping with histological tissue, requiring a new scan with proper focus points.
- Level 2 (re-mount): Mounting artifacts (bubbles, dust, pen marks, coverslip issues) are present that impair proper tissue visualization.
- Level 3 (re-cut): Severe artifacts (primarily tissue folds) are present that cannot be resolved through re-scanning or re-mounting, necessitating new slide preparation.

3.2.1 | Theoretical Background

Our multi-scale processing approach is based on the understanding that different artifacts are optimally detected at specific magnification levels. Tissue-level features are captured

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

at lower magnifications (1.25×), while fine-grained artifacts require higher resolution (5×) for accurate identification. The choice of these specific magnifications is deliberate: 1.25× provides sufficient resolution for coarse tissue segmentation while minimizing computational load, and 5× offers an optimal trade-off for artifact detection, as most artifacts remain clearly visible without requiring the processing overhead of higher magnifications (20× or 40×).

The standardized use of lower magnifications for artifact detection, regardless of native scanning magnification, serves multiple purposes: (1) ensuring consistency across different scanner configurations, (2) avoiding interpolation artifacts from aggressive downsampling, and (3) capitalizing on the intrinsic multi-scale nature of whole slide images, where quality control information exists primarily at lower resolutions. This approach prioritizes computational efficiency while maintaining detection accuracy.

3.2.2 | Segmentation Networks

SlideInspect employs a segmentation network based on the K-Net architecture [25], which uses a Swin Transformer backbone [26] combined with a UPerNet decoder [27] (Figure 4). The Swin Transformer backbone utilizes a hierarchical architecture based on shifted windows to enable efficient modeling of spatial relationships and multi-scale features. The shifted window approach allows the model to capture both local details as well as global dependencies across the image. The UPerNet decoder consists of a nested U-structure with interleaved encoding and decoding pathways, enabling robust fusion of multi-scale contextual information from the encoder with fine-grained spatial details from the decoder. More details about the network's architecture are provided in Table S3.

To efficiently annotate the large dataset, we employed an active learning approach with multiple expert validation steps. Initially, an expert operator (M.Sc.) manually annotated 500 WSIs from the training set for both tissue segmentation and multi-class artifact detection. These annotations were reviewed and corrected by a second expert (R.P.) and served as the basis for preliminary training of the segmentation networks for 20 epochs. Once the preliminary networks demonstrated reasonable performance, they were applied to the remaining training WSIs (~1660 slides). The expert operator (M.S.) then reviewed and corrected these automated annotations, substantially accelerating the annotation process while maintaining full human supervision. Following the active learning phase, a third expert (M.S.) performed a final verification of the entire annotated dataset. This approach significantly reduced the annotation time while preserving annotation quality. For validation and test sets, direct manual annotations from the expert operator were used without the active learning process to ensure unbiased evaluation of model performance. After the entire training set was annotated and verified, the final training of both networks was conducted for 50 epochs.

To train the tissue and artifact segmentation models, patches of size 1024 ×1024 pixels were extracted from the WSI at 1.25× and 5× magnifications, respectively. To prevent selecting non-informative patches, the tissue segmentation network used a white mask to evaluate the percentage of background in each patch, with the probability of patch selection being inversely proportional to the amount of white. The artifact segmentation network retained all patches with annotations and up to 50 high-quality tissue patches per WSI with no annotations. This

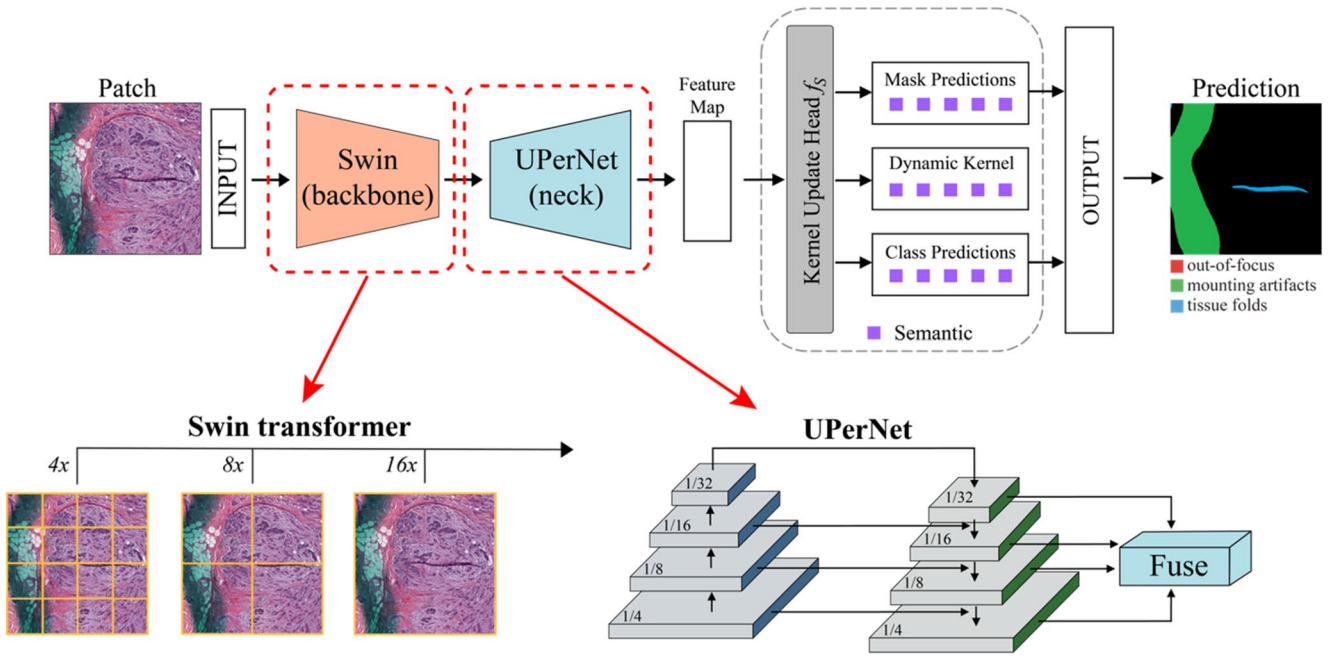


FIGURE 4 | Overall architecture of segmentation network used in this work. The Swin Transformer backbone employs a hierarchical transformer architecture to capture long-range dependencies and multi-scale features from the input image effectively. The extracted features are then fed into the UPerNet neck, which is an encoder-decoder architecture designed for semantic segmentation. The UPerNet aggregates multi-scale contextual information and preserves spatial details, enabling precise segmentation of artifacts and tissue boundaries within the digital pathology image.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

prevented bias caused by an uneven number of patches per WSI, thereby reducing dataset imbalance.

The histological tissue identification network was trained on a binary task: background vs. histological tissue. On the other hand, the artifact segmentation network was trained on a total of four classes: (i) background, (ii) out-of-focus regions, (iii) mounting artifacts (including bubbles, dust, pen marks, and coverslip issues), and (iv) tissue folds. This grouping strategy aligns with typical laboratory workflows where similar corrective actions address related artifact types. The training process was conducted for 50 epochs, with early stopping triggered if the validation loss did not improve for four consecutive epochs. Real-time data augmentation techniques were applied during training to help prevent overfitting, including random flipping and rotation of images. Photometric distortion was also utilized, with random adjustments to brightness within a range of 2 units, contrast within 0.9–1.1, saturation within 0.9–1.1, and hue within 2 units. The generalized Dice loss function [28] was employed for the decoder network as it can effectively handle highly imbalanced segmentation problems in 2D and 3D medical images. To accelerate training, transfer learning was utilized by pre-training both networks on the ADE20K dataset.

During the inference process, a straightforward reconstruction method is utilized to obtain the segmentation mask for the entire slide. This involves implementing a sliding window approach that applies the model to consecutive patches of the WSI, as detailed in our previous works [29]. For handling various proprietary WSI formats (Aperio .svs, Hamamatsu .ndpi, and 3DHistech .mrxs), we utilized large-image (version 1.31.1) and pyvips (version 2.2.3) libraries. The deep learning implementation was developed using Python 3.9.16 and the MMSegmentation framework (version 1.2.2).

3.2.3 | Stain Quality Assessment

In parallel with artifact segmentation, the stain quality assessment branch analyzes the stain intensity and contrast of the histological tissue. The assessment focuses only on valid tissue regions by first excluding pixels outside the tissue mask and within artifact areas, as well as white areas within the tissue. The stain quality assessment evaluates two key characteristics:

1. Stain intensity (SI): quantified by calculating the Euclidean distance between each RGB pixel value and the background reference color (typically white):

$$SI(p) = \sqrt{\left[(Rp - Rbg)2 + (Gp - Gbg)2 + (Bp - Bbg)2\right]} \tag{1}$$

where (Rp, Gp, Bp) represents the RGB values of pixel p, and (Rbg, Gbg, Bbg) represents the reference background color, calculated as the median RGB values from background regions

2. Contrast (C): evaluated using the magnitude of the gradient on the grayscale image:

$$C(p) = \sqrt{\left[Gx(p)^2 + Gy(p)^2\right]} \tag{2}$$

where Gx and Gy are the horizontal and vertical image gradients computed using Sobel operators.

The assessment is divided into two modules that evaluate different staining issues: Light stains assessment (Module 1) and Dark stains assessment (Module 2).

Light stains are characterized by low contrast and color intensity similar to the background. A region is classified as having light stains if the average gradient magnitude (C_{avg}) is below a global threshold equal to 30 and the average stain intensity is below 10% of the maximum possible stain intensity. If both conditions are met, the WSI's stain quality is rated as 'low'. If only one condition is met (either adequate contrast or adequate intensity), the quality is rated as 'sufficient'.

On the other hand, dark stains are characterized by high intensity but potentially poor contrast between histological structures. If the average stain intensity exceeds 50% of the maximum possible intensity, we evaluate contrast using color clustering. For this evaluation, tissue pixels are clustered into four groups using k-means clustering in RGB space. The relative differences between cluster centroids are calculated as:

$$RD_{i,j} = \| C_i - C_j \| \,/\, \max\left(\| C_i \|, \| C_j \| \right) \tag{3}$$

where C_i and C_j are cluster centroids and $\|\cdot\|$ denotes the Euclidean norm. If all pairwise relative differences are below 50% ($RD_{i,j} < 0.5$ for all i,j pairs), it indicates poor color separation, and the WSI's stain quality is rated as 'low'.

The final stain quality assessment combines the outputs from both modules. If either module rates the quality as 'low', the overall stain quality is rated as 'low'; otherwise, it is rated as 'sufficient'. Processing at 5× magnification significantly reduces computational load while preserving sufficient detail for accurate stain quality assessment. Figure 5 provides a schematic representation of this heuristic algorithm. It is important to note that by working with grayscale images and relative contrast analysis, our stain quality assessment algorithm becomes independent of the specific staining methods used in the slide preparation (H&E, IHC, TRIC, or PAS). This approach enables the algorithm to perform consistently across various staining protocols without requiring specific calibration for each stain type, thereby enhancing the system's versatility and robustness in diverse histopathological contexts.

3.2.4 | Quantitative and Qualitative Outputs

The SlideInspect system provides two types of outputs: quantitative and qualitative. Starting from the automatic multi-class segmentations of the artifacts, the area, number of regions, and average area are quantified for each artifact class. Based on the types of artifacts and their spatial localization within the slide, the WSI is labeled into four different classes:

- Level 0: no action. Artifacts, if present, have a limited extent both in area and number, not impacting diagnostic quality.
- Level 1: re-scan. A significant number of out-of-focus areas overlapping with histological tissue are detected,

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

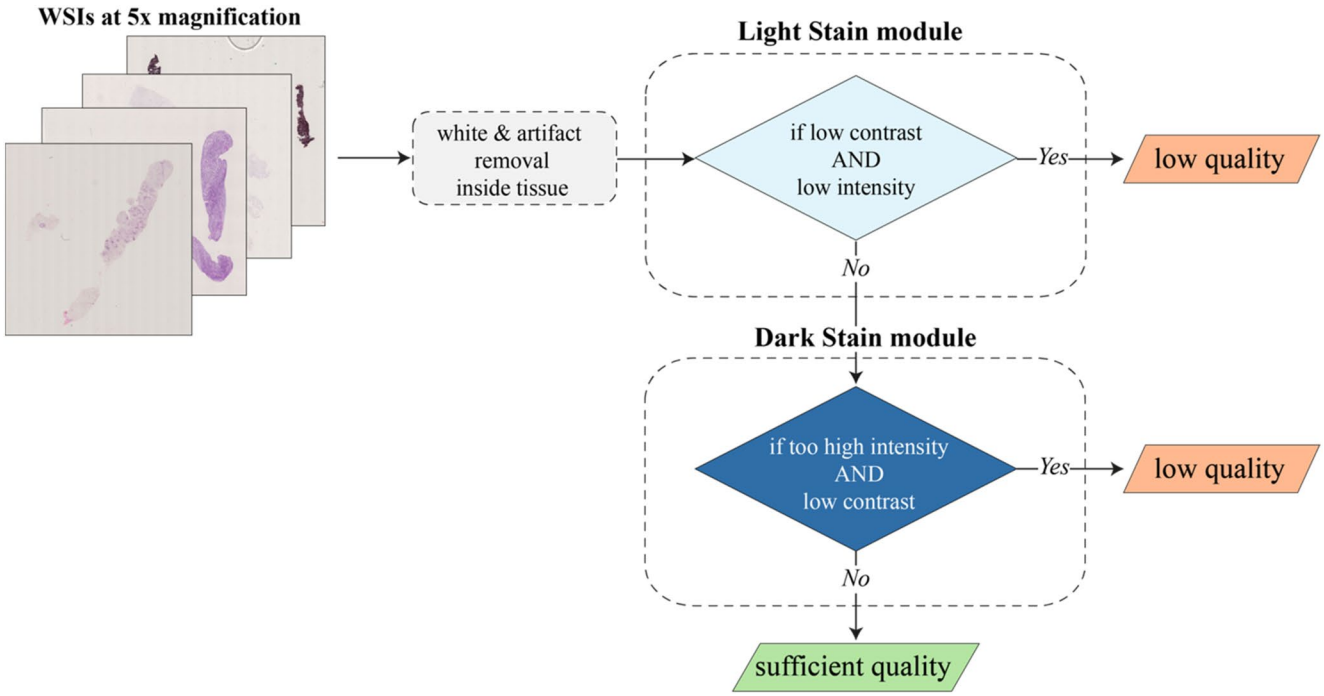


FIGURE 5 | Overview of the stain quality assessment module. This module evaluates staining intensity and contrast to identify WSIs with poor quality stains. The assessment is divided into two branches: one evaluating very light stains based on low contrast and color intensity close to the background, and one evaluating very dark stains based on high color intensity but poor contrast between clustered pixel colors.

requiring new image acquisition with optimized focus points.

- Level 2: re-mount. The presence of mounting artifacts (bubbles, dust, pen marks, and coverslip issues) prevents correct visualization of critical tissue regions. These issues can be resolved by re-mounting the coverslip.
- Level 3: re-cut. The slide presents severe artifacts (such as extensive tissue folds) that, due to their type, number, and extent, cannot be resolved with any of the previous actions, requiring new slide preparation.

To perform the classification, the first step consists of extracting features from the automatic masks of the tissue and artifacts. For each artifact class, the following parameters are calculated during feature extraction: percentage of artifact area relative to tissue area, percentage of artifact area relative to slide area, overlap of the artifact with respect to the tissue, and presence of artifacts near the edge of the tissue. The output of the stain quality assessment module ('low' or 'sufficient') is also provided as input to the classifier.

A Random Forest classifier was then employed because of its overall interpretability and resistance to overfitting. To train the classifier, 410 slides were selected from the training set and 97 from the test set. Training set and test set were stratified according to the actions. An expert pathologist (V. I.) performed the classification of the slides by recommending one of the four actions described above. These labels serve as ground truth to train the machine learning classifier. Figure 6 shows an example of the quantitative and qualitative output provided by the system.

3.3 | Performance Metrics

We employed several metrics to evaluate the performance of the segmentation task and overall slide quality assessment. For segmentation, we calculated the Dice Similarity Coefficient (DSC) [30], sensitivity, and precision on the full dataset. To ensure comprehensive evaluation, analyses were performed on the full dataset as well as stratified by staining type to understand model performances with different imaging modalities. This provided insights into how segmentation accuracy varied across sample groups.

We also quantified the absolute error in artifact level measurements. Accurate artifact quantification is essential for determining the appropriate corrective action and assessing overall slide quality.

Finally, we evaluated how well our system assigned the correct corrective action (no action, re-scan, re-mount, re-cut) compared with the ground truth provided by an expert pathologist. This directly assessed the system's ability to provide reliable recommendations for slide handling based on quality issues and artifact presence.

4 | Results

4.1 | Tissue and Artifact Segmentation

Figure 7 illustrates the segmentation performance of the histological tissue network (Figure 7a) and the multi-class artifact segmentation network (Figure 7b) on the test set. The histological tissue segmentation network achieved excellent performance with a DSC above 0.98 for all staining types. The

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

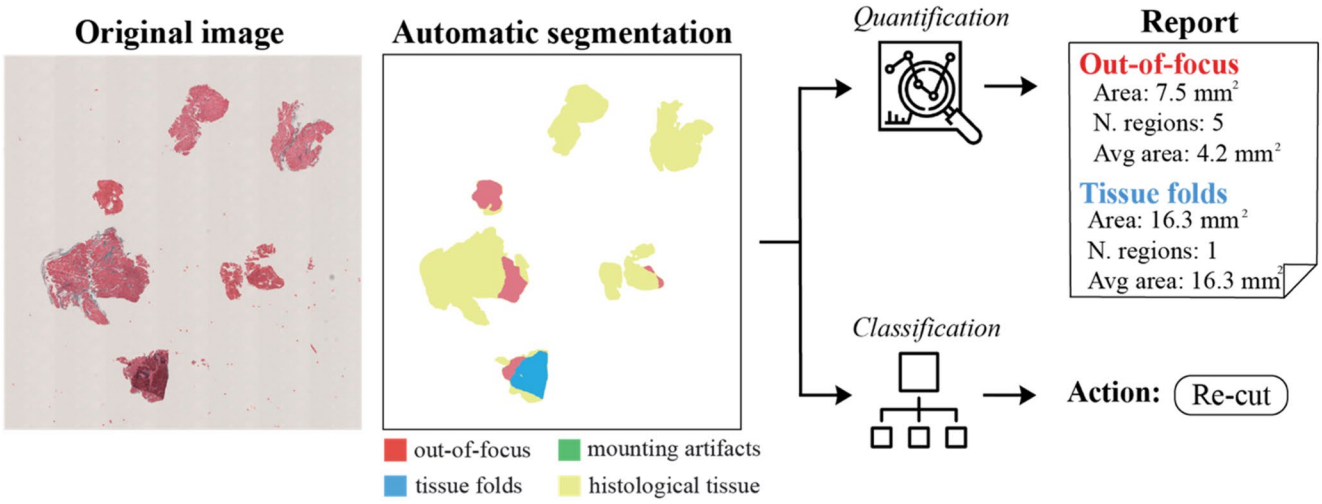


FIGURE 6 | Example of hierarchical output provided by the SlideInspect system. The slide is labeled as 'Re-cut' due to the presence of a large tissue fold. Quantitative metrics are provided for each artifact class in terms of artifact area and number of regions.

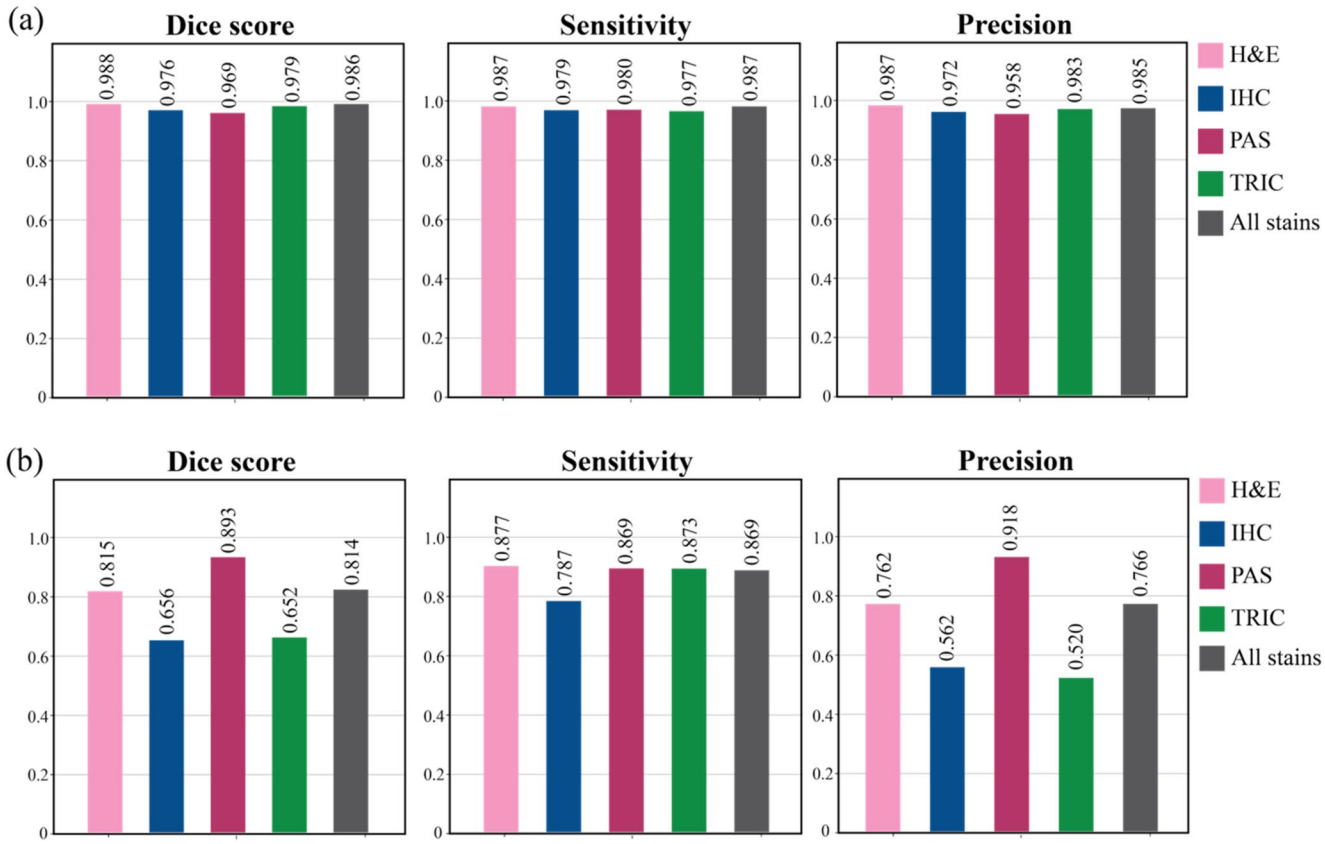


FIGURE 7 | Segmentation performance of the proposed framework stratified by staining type on the test set. (a) Histological tissue segmentation network. (b) Multi-class artifact segmentation network.

complete performance analysis across all subsets (training, validation, and test) is provided in Supplementary (Figure S2).

Regarding artifact segmentation, the network obtained good performance with a DSC of 0.814 on the test set. It is interesting to note that despite the limited number of WSIs in the TRIC and PAS staining categories, the network still achieved DSC values above 0.65 (TRIC) and 0.89 (PAS). In terms of sensitivity, the network obtained values above 0.85 for almost all combinations of staining types and subsets, indicating a high sensitivity to artifact-affected regions. Concerning precision, the network generated some false positives, resulting in a precision of 0.766 on the test set. In the Supporting Information, we have reported the segmentation performance divided for each artifact type (Figure S3) and stratified by tissue type (Table S4).

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

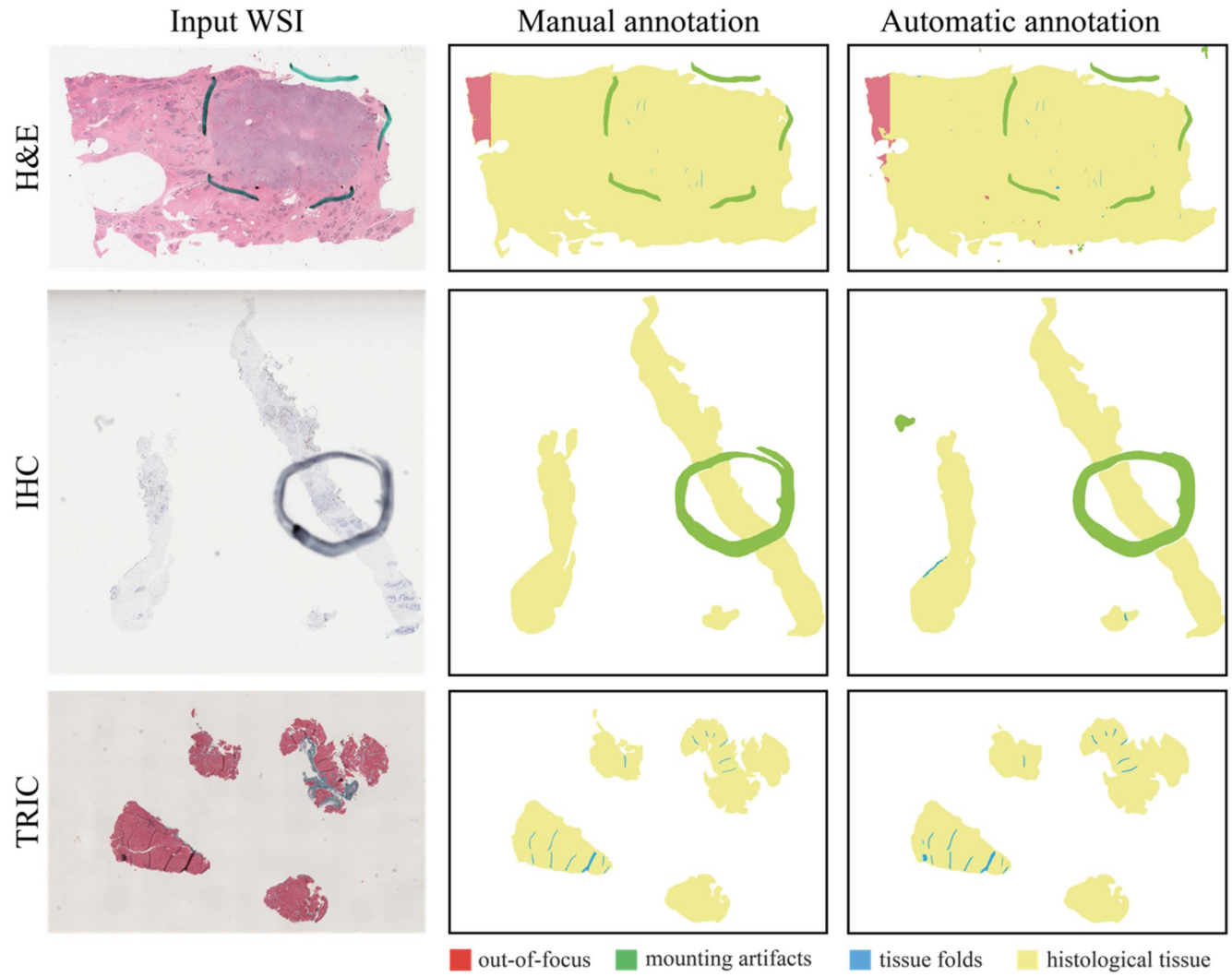


FIGURE 8 | Visual examples of histological tissue and artifact segmentation results for different staining types and sampling mechanisms (macrosections, biopsy). The manual segmentations (ground truth) are compared with the automatic segmentations generated by the proposed framework.

Figure 8 shows visual comparisons between manual and automatic segmentations generated by SlideInspect for histological tissue and artifacts across different staining types and sampling methods. The network accurately segmented histological tissue and correctly classified the artifact types present on each slide. By simultaneously segmenting both tissue and artifacts, the framework can localize artifacts in relation to tissue regions. This enables evaluation of how artifacts may impact on the overall slide quality. As seen, the network produced segmentations closely aligned with manual annotations, consistent with the numeric performance metrics reported previously.

Finally, we evaluated the error in quantifying artifact extent between manual segmentations and those generated automatically by our system. While traditional metrics like DSC provide overall segmentation performance, they may not fully represent performance across WSIs with varying annotation sizes. A WSI with small annotated artifact areas that are missed by the algorithm could result in a low DSC, despite minimal absolute error. To provide a more comprehensive evaluation, we employed Bland–Altman plots (Figure 9) for the training, validation, and test subsets. This statistical method assesses agreement between two measurement techniques by plotting measurement differences against their averages [12], allowing us to effectively summarize artifact estimation errors and identify potential trends relative to annotation size. Visually, the distributions are nearly identical across subsets, with mean errors close to zero and very low standard deviations. For the test set, the mean error between manual and automatic segmentation is $-0.08\,\mathrm{mm}^2$. This indicates our system provides highly accurate quantification of artifact areas that closely matches expert assessments. The lack of variation between subsets further suggests the artifact quantification was not overfit and maintains reliable performance on new data.

4.2 | Validation of Stain Quality Assessment

To validate our stain quality assessment module, we conducted a comprehensive evaluation involving two expert operators (M.Sc., M.S.) who independently assessed the stain quality of all images in the test set ($n = 441$). The operators were blinded to both the algorithmic output and each other's assessments and classified each WSI's stain quality as either 'sufficient' or 'low'. We then compared these manual assessments with SlideInspect's automated classification using confusion matrices (Figure 10).

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

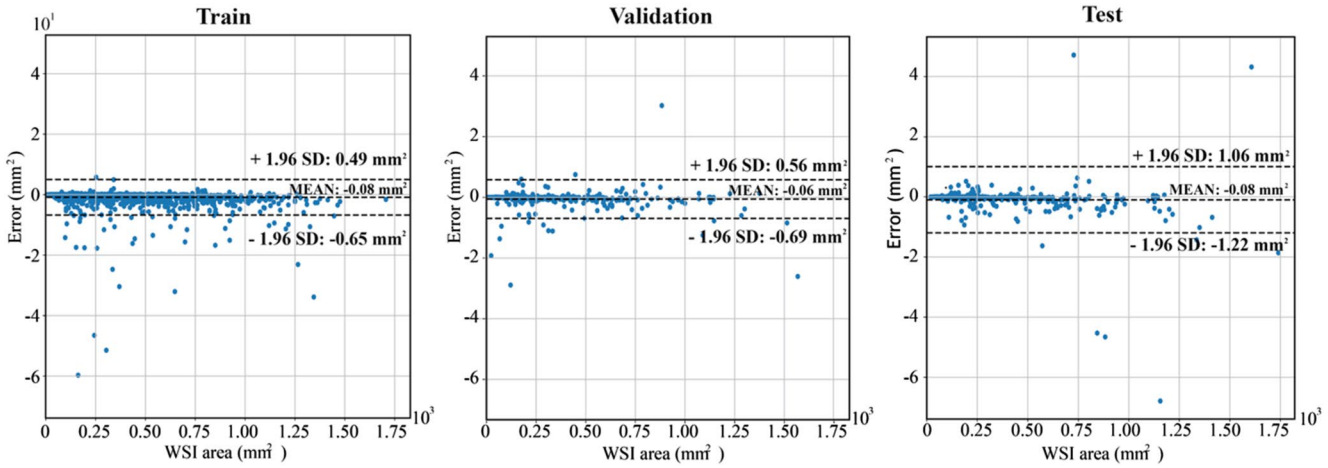


FIGURE 9 | Bland–Altman plots illustrating the agreement between manual and automatic quantification of artifact extent across the training, validation, and test subsets.

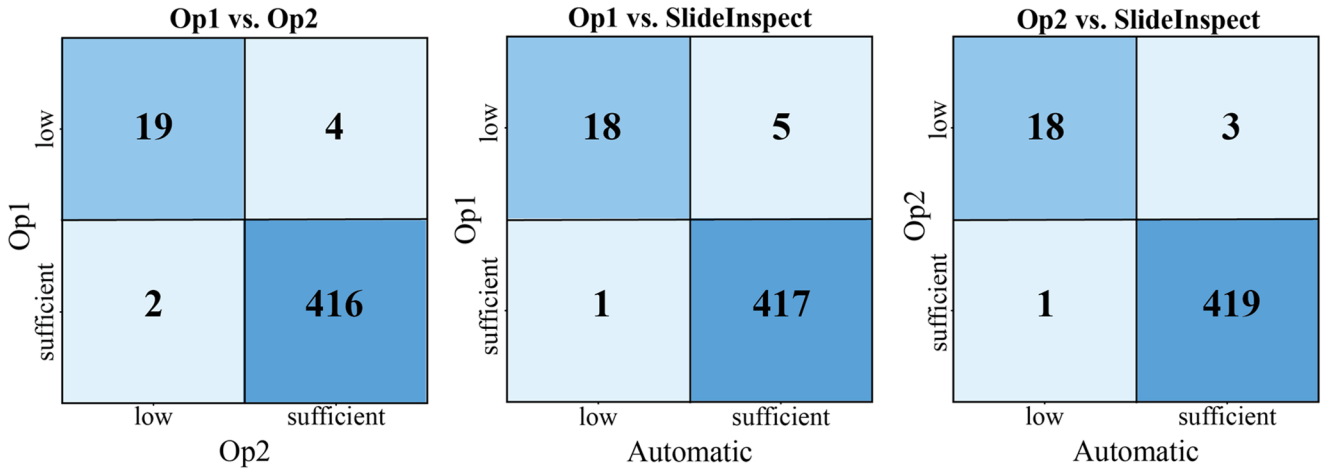


FIGURE 10 | Confusion matrices for stain quality assessment validation showing inter-operator agreement and comparisons between expert operators and the automated assessment provided by SlideInspect.

The inter-operator agreement showed excellent concordance, with an accuracy of 98.64% (435/441 cases) between the two experts. When comparing SlideInspect's automated assessment against each operator, we observed comparable performance with accuracies of 98.64% (435/441 cases) for Operator 1 and 99.09% (437/441 cases) for Operator 2. These results show how our heuristic approach to stain quality assessment achieves performance comparable to inter-operator variability.

4.3 | WSI Quality

To evaluate the performance of the system in assessing overall WSI quality and recommending appropriate actions, we trained and tested several machine learning models. The models were trained to classify the slides into one of four possible actions: no action, re-scan, re-mount, or re-cut. Table 3 presents the accuracy of these models on the training and test sets.

The Random Forest classifier, which was chosen for the SlideInspect system, was compared with other models such as Support Vector Machines (SVM) and Decision Trees. It is important to note that during the training phase, the classifiers were trained using manual annotations of the artifacts, while on the test set, the automatic outputs provided by our networks were used. This approach simulates a real-world usage scenario of the system, where the classifiers rely on the automatic segmentations of the artifacts to make their predictions.

TABLE 3 | Performance of different machine learning models for action recommendation (no action, re-scan, re-mount, re-cut).

Model	Accuracy on train set ($n = 410$)	Accuracy on test set ($n = 97$)
SVM	92.44%	88.66%
Decision Tree	87.56%	82.47%
Random Forest (proposed)	100%	92.78%

As shown in Table 3, the Random Forest classifier achieved the highest accuracy on both the training and test sets. On the training set, the Random Forest classifier achieved a perfect

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

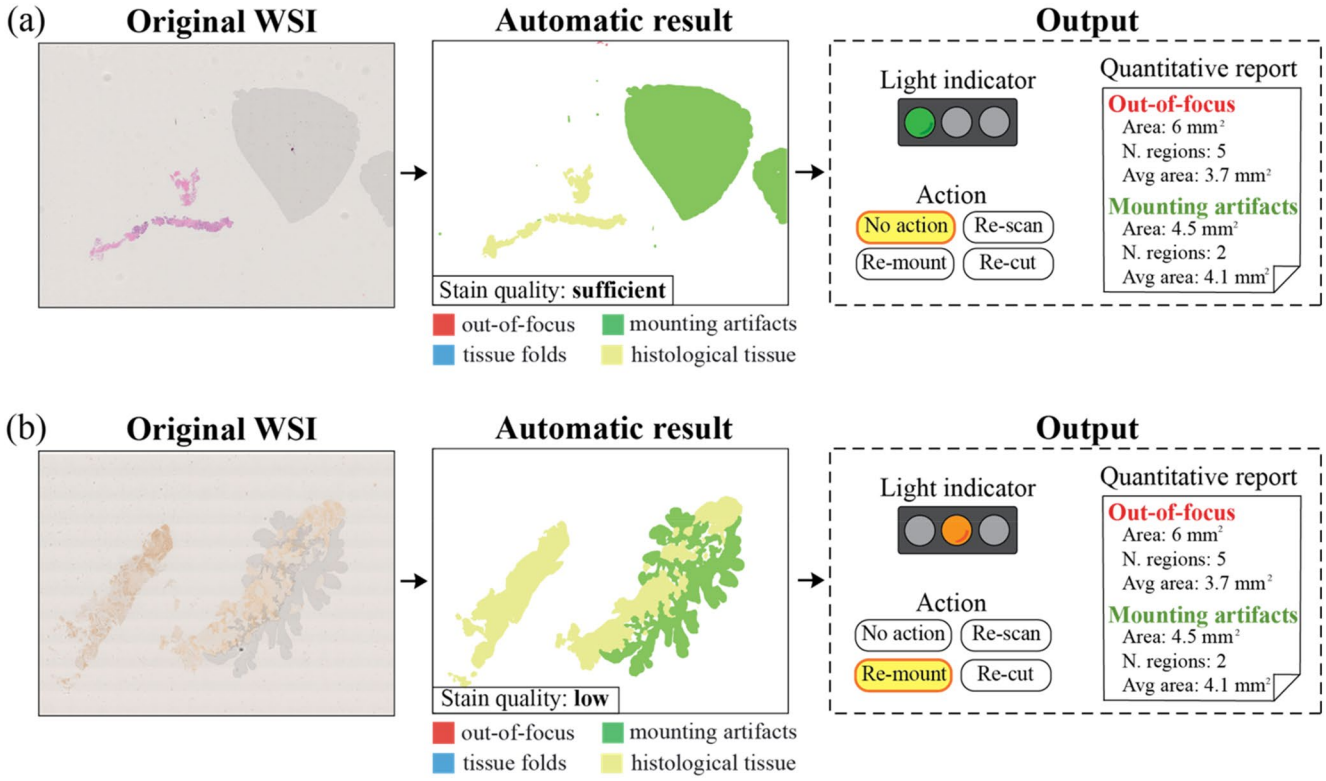


FIGURE 11 | Representative examples of the output of SlideInspect for WSIs affected by various artifacts. (a) WSI with mounting artifacts not overlapping with the histological tissue, resulting in a "no action" recommendation. (b) WSI with mounting artifacts affecting the histological tissue, leading to a "re-mount" recommendation. Computational efficiency is critical for digital pathology tools to effectively integrate into clinical workflows.

accuracy of 100%, while on the test set, it reached an accuracy of 92.78%.

Figure 11 provides an overview of SlideInspect's performance on two representative WSIs from our test set, showing both visual and quantitative outputs. The figure illustrates the automatic artifact segmentation, stain quality assessment, overall quality indicator, quantitative report, and recommended action.

The WSI in Figure 11a contains numerous mounting artifacts that do not overlap with histological tissue, thus not significantly impacting slide quality. SlideInspect assigns a "no action" recommendation, indicating sufficient quality for further analysis. This example highlights the importance of considering artifact location relative to regions of interest. By assessing overlap between artifacts and tissue, our system provides context-aware quality assessment that evaluates both artifact quantity and their spatial distribution relative to interpretable tissue regions.

Table 4 presents the computational times required by SlideInspect for each analysis phase. The average size of processed WSIs is ~80,000 ×60,000 pixels. Tissue detection and artifact segmentation, which process the entire WSI, constitute most of the computational burden, requiring 12 and 45 s per WSI, respectively. Stain quality assessment averages 13 s per WSI, while artifact quantification and action recommendation are substantially faster at 1 and 0.5 s. The entire processing pipeline completes in ~73 s per digital slide.

TABLE 4 | Breakdown of the computational times required by SlideInspect for each phase of the analysis.

Task	Computational times (seconds)
Tissue segmentation	12.8 ± 2.4
Artifact segmentation	45.3 ± 8.1
Stain quality assessment	13.0 ± 4.4
Artifact quantification	1.1 ± 0.2
Action recommendation	0.5 ± 0.1
Entire processing	*72.7 ± 9.5*

4.4 | Ablation Studies

Since artifact segmentation constitutes more than 50% of the entire pipeline's computational time, we conducted ablation studies by varying both the network architecture (from lightweight to heavier models) and magnification levels, since lower magnifications reduce inference time while higher magnifications (10×) might improve segmentation performance.

Our analyses compared various segmentation architectures and magnification levels on the test set, measuring DSC and computational time. Table 5 presents a comprehensive comparison of different models and magnification levels, including lightweight

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 | Comparison of average segmentation performance and computational time across different architectures and magnification levels on all artifacts in the test set.

Architecture	Magnification	Dice ↑	Sensitivity ↑	Precision ↑	Computational time (s) ↓
MobileNetV3	2.5×	0.631	0.685	0.589	33.2
	5×	0.692	0.725	0.662	60.8
	10×	0.718	0.748	0.691	158.4
U-Net	2.5×	0.673	0.712	0.638	40.6
	5×	0.731	0.768	0.698	67.2
	10×	0.756	0.791	0.724	143.5
Swin	2.5×	0.712	0.745	0.682	46.8
	5×	0.768	0.804	0.735	79.5
	10×	0.795	0.832	0.761	217.2
K-Net + ResNet50	2.5×	0.726	0.773	0.685	41.2
	5×	0.782	0.836	0.733	66.3
	10×	0.798	0.852	0.749	188.4
K-Net + Swin (proposed)	2.5×	0.764	0.811	0.721	43.6
	5× (chosen)	**0.814**	**0.869**	**0.766**	**72.7**
	10×	0.819	0.887	0.771	206.5

Note: The chosen configuration (architecture + working magnification) is highlighted in bold.

architectures (MobileNetV3 [31]), standard U-Net [32], K-Net with different backbones (ResNet-50) and "pure" vision transformers (Swin [26]).

Our results indicate that the chosen architecture (K-Net with Swin Transformer) outperforms other architectures in overall performance across all magnification levels. The performance improvement is particularly evident compared with lightweight networks like MobileNetV3, where K-Net + Swin achieves ~18% higher Dice score at comparable magnifications. Notably, K-Net + Swin at 5× magnification achieves performance metrics nearly identical to its 10× counterpart (0.814 vs. 0.819 Dice score) while reducing computational time by nearly 65% (72.7 s vs. 206.5 s). This significant finding guided our final implementation choice, as it represents an optimal balance between segmentation quality and computational efficiency for clinical deployment.

4.5 | Clinical Impact Analysis

We conducted a study using our test set of 441 WSIs to quantify the potential impact of our tool in a clinical setting. The analysis compared two operational scenarios: traditional workflow and SlideInspect-enabled workflow.

The traditional workflow processes slides in batches of 50, corresponding to the standard scanner tray capacity. After each batch is scanned, there is approximately a one-hour delay before the pathologist can review the batch and identify any quality issues. Following this review, slides requiring re-scanning are processed again at the end of the batch with doubled scanning time due to additional focus points, while those needing re-mounting or re-cutting are sent back for slide preparation.

The SlideInspect-enabled workflow incorporates automated quality assessment during the scanning process. The system immediately identifies slides requiring re-cutting or re-mounting, preventing unnecessary scanning of these cases. When focus-related issues requiring re-scanning are detected, the slide is immediately re-processed using additional focus points, which doubles the scanning time for that specific slide. The workflow includes an average processing time of 72.7 s per slide for the SlideInspect analysis.

Analysis of our test set revealed the following distribution of recommended actions:

- No action required: 248 WSIs (56.2%).
- Re-cut needed: 112 WSIs (25.4%).
- Re-mount needed: 46 WSIs (10.4%).
- Re-scan needed: 35 WSIs (7.9%).

Figure 12a shows the distribution of recommended actions across the test set, while Figure 12b presents the cumulative processing time comparison between the traditional and SlideInspect-enabled workflows. The simulation demonstrated significant time savings with the SlideInspect system. The total processing time with SlideInspect was 96.8 h, compared with 152.2 h for the traditional workflow, resulting in a time saving of 55.4 h (36.4%) across the test set.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

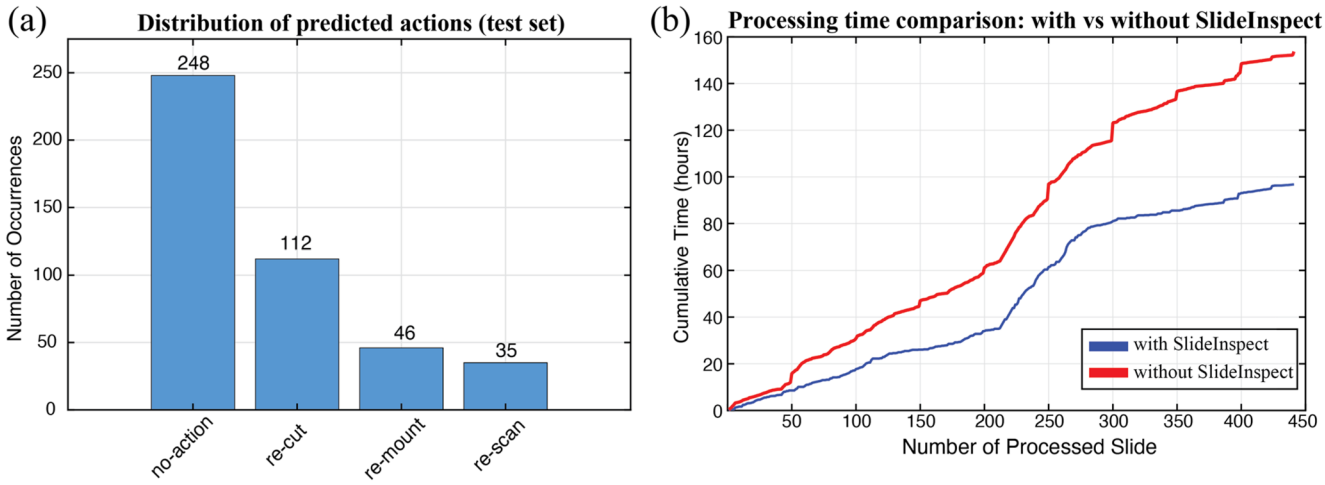


FIGURE 12 | Impact analysis of SlideInspect implementation. (a) Distribution of recommended actions across the test set. (b) Cumulative processing time comparison between traditional workflow (solid red line) and SlideInspect-enabled workflow (solid blue line).

4.6 | Comparison With Previously Published Tools

Table 6 summarizes features of recent published tools (2019–2025) for multi-artifact segmentation and quality control in digital pathology. The proposed framework is compared against these approaches in terms of the number of WSIs used, the number of organs, staining types, working magnification, and type of output provided. SlideInspect uses one of the largest datasets (3036 WSIs) from multiple centers, covering a wide range of organs (8) and staining types (H&E, IHC, PAS, TRIC). SlideInspect also operates at lower magnification levels (1.25×–5×) while still accurately detecting and segmenting tissue folds, out-of-focus regions, and mounting artifacts.

Many previous tools [10, 19, 20, 22] were developed and validated on single-organ datasets, which could limit their generalizability across different tissue types. Furthermore, all approaches except Smit et al. [18] were designed for a single stain type (predominantly H&E), which represents only a portion of slides processed in clinical pathology laboratories.

To our knowledge, we are the only method providing comprehensive outputs that include suggested actions, stain quality assessment, and quantification of artifacts across multiple staining types. Only Smit et al. [18] offer suggested actions as output, although limited to fewer stain types. Additionally, SlideInspect joins [18] as one of the few approaches that work effectively at lower magnifications (1.25×–5×), enabling more efficient processing while maintaining high accuracy.

5 | Discussion

In digital pathology, artifacts from tissue processing, staining, and imaging equipment [17, 33] can significantly impact analysis accuracy by obscuring key tissue features. Detecting and segmenting these artifacts is crucial for reliable downstream analyses [10]. Deep learning techniques have shown promise in addressing these challenges [34, 35], potentially improving workflow efficiency while reducing human error [36].

In this study, we introduced SlideInspect, a novel AI-based framework for comprehensive artifact detection and quality control in digital pathology. Our system addresses the limitations of existing methods by providing a fully automated, multi-tissue, multi-stain solution that accurately segments a wide range of artifacts at multiple magnifications. The key advantages of our approach include its ability to handle diverse tissue types (adrenal, breast, colon, kidney, liver, lung, myocardial, and prostate) and staining methods (H&E, IHC, PAS, TRIC), while providing both qualitative (color-coded slide quality indicators) and quantitative (pixel-level segmentations, artifact percentages) outputs. Moreover, SlideInspect demonstrates computational efficiency, processing a WSI in ~72.7 s, making it suitable for integration into clinical workflows.

It's important to emphasize that SlideInspect represents, to our knowledge, the only approach that implements a comprehensive assessment of stain quality and contrast in a multi-stain context. While other systems primarily focus on physical artifacts or a single staining type, our framework simultaneously evaluates both physical and colorimetric aspects of image quality, automatically adapting to different staining methods (H&E, IHC, PAS, TRIC). This feature is particularly relevant in clinical settings where various staining techniques are used to highlight different histological aspects, ensuring uniform and comparable quality control across different preparation protocols.

The color-coded output (green, yellow, red) and suggested actions are designed to be integrated into existing digital pathology workflows. This system can be implemented at various stages: immediately after slide scanning, during quality control checks, or before pathologist review. For instance, it could be integrated into the scanner software to provide immediate feedback, or into the laboratory information system (LIS) for automated quality tracking. This allows for quick intervention when quality issues are detected, potentially reducing turnaround times and improving overall slide quality. Our preliminary analysis confirms this potential, demonstrating a 36.4% reduction in total processing time when using SlideInspect compared with traditional workflows (Figure 12). This significant improvement (i.e., saving 55.4 h across 441 WSIs) is primarily achieved through immediate artifact

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Recent approaches (2019–2025) for artifact segmentation and quality control in digital pathology.

Authors	Name of the tool	# WSI	Staining	Architecture	Magnification	Output provided	
						Detected artifacts	Output type
Janowczyk et al. [10], 2019	HistoQC	450	1 (H&E)	Traditional computer vision techniques	40×	TF, OOF, MA	Segmentation masks, metrics report
Smit et al. [18], 2021	—	142	3 (H&E, IHC, PAS)	DeepLabV3+ with EfficientNet-B2	2.5×	TF, OOF, MA	Segmentation masks, actions
Haghighat et al. [19], 2022	PathProfiler	198	1 (H&E)	ResNet18	5×	TF, OOF, MA	Patch classification, usability, quality scores
Patil et al. [20], 2022	HistoROI	503	1 (H&E)	ResNet18	10×	TF, OOF, MA, PN	Patch classification, quality scores
Weng et al. [21], 2024	GrandQC	4300	1 (H&E)	UNet++ with EfficientNetB0	5×–10×	TF, OOF, MA, PN	Segmentation masks, artifact report, quality benchmark
Kanwal et al. [22], 2024	—	55	1 (H&E)	Mixture of Experts (DCNNs/ViT)	40×	TF, OOF, MA	Segmentation mask, artifact report
Scotto et al. (proposed)	SlideInspect	3036	4 (H&E, IHC, PAS, TRIC)	K-Net + Swin	1.25×–5×	TF, OOF, MA	Segmentation masks, artifact report, actions

Abbreviations: H&E, Hematoxylin and Eosin; IHC, Immunohistochemistry; MA, mounting artifacts; OOF, out-of-focus; PAS, Periodic acid–Schiff; PN, pen marks; TF, Tissue folds; TRIC, Trichrome.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

detection and elimination of batch review delays, while adding only 72.7 s of processing time per slide.

Compared with state-of-the-art methods, our system demonstrates superior performance in terms of accuracy, computational efficiency, and generalizability. Our framework achieves high segmentation accuracy across a wide range of artifacts, including tissue folds, out-of-focus regions, and mounting artifacts, outperforming existing algorithms that often focus on a single artifact type [6, 15]. SlideInspect efficiently processes WSIs at both low (1.25×) and mid-low (5×) magnifications, finding a balance between computational speed and segmentation precision. This feature is particularly valuable in clinical settings where hardware resources may be limited [37]. Furthermore, our extensive validation on a multi-centric, multi-scanner dataset showcases the generalizability of our system, a critical aspect often overlooked in previous studies [9, 15].

SlideInspect is designed to be vendor-agnostic and integrable with various digital pathology infrastructures, compatible with different scanning systems and any LIS. Additionally, its stain quality output could trigger automatic stain normalization processes [38, 39], when suboptimal staining is detected despite overall acceptable quality.

Despite the promising results, our study has several limitations that should be acknowledged. While our dataset covers a broad range of tissue types and staining methods, it may not capture all the variations encountered in clinical practice. Future work should focus on expanding the dataset to include additional organ systems, rare pathologies, and other staining protocols (e.g., Giemsa). Although our current results demonstrate robust performance across staining types, with some less common stains like PAS showing strong performance (DICE scores over 0.89), a more balanced dataset would enable even more thorough validation of the system's capabilities. This is particularly important as laboratories increasingly employ diverse staining methods for specialized diagnostic purposes. Our framework currently handles a limited set of artifact types, and while it can detect various mounting-related issues (such as pen marks, dust and ink contamination) within broader categories, it does not provide granular classification of individual artifact subtypes. This design choice prioritizes actionable feedback over detailed artifact categorization, grouping artifacts based on their required corrective actions. While this approach aligns with clinical workflow needs, extending the system to provide more detailed artifact classification and to detect additional quality issues, such as tissue damage, would further enhance its utility.

Additionally, we haven't evaluated the impact of SlideInspect on downstream tasks such as tumor segmentation [40] or biomarker quantification [11]. Although artifact detection is fundamental for ensuring the reliability of subsequent analyses, directly quantifying the improvement in algorithm performance on images 'filtered' by our system would provide important additional validation. Furthermore, while we detect pen marks as part of our mounting artifacts category, we do not provide separate classification for this specific artifact type, unlike some recent tools [20, 21]. This approach aligns with our strategy of grouping artifacts based on their required corrective actions rather than providing individual classifications.

To address these limitations and expand the capabilities of SlideInspect, several future directions can be explored. One avenue is to incorporate unsupervised or weakly supervised learning techniques to minimize the reliance on manual annotations and facilitate the inclusion of larger, more diverse datasets. A promising future development is the implementation of a 'fast' version of SlideInspect, designed to process exclusively the histological tissue while ignoring background areas. This approach could lead to a significant reduction in processing time since clinically relevant artifacts are primarily those interfering with tissue.

6 | Conclusion

In this study, we presented SlideInspect, an innovative AI-based framework for comprehensive quality control in digital pathology. Our system demonstrates several key strengths that contribute to its potential impact in the field. SlideInspect's multi-tissue, multi-stain capability makes it adaptable to various clinical and research settings, while its hierarchical output provides actionable insights at multiple levels of granularity, from simple color-coded indicators to detailed segmentations. The computational efficiency of our system, processing a single WSI in ~72.7 s, makes it suitable for integration into existing workflows without causing significant delays. Furthermore, our extensive validation on a multi-centric, multi-scanner dataset showcases the system's robustness across different imaging platforms.

Author Contributions

Manuela Scotto: Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft. Roberta Patti: Data curation, Validation, Writing – review and editing. Vincenzo L'imperio: Data curation, Formal analysis, Writing – review and editing. Filippo Fraggetta: Resources, Writing – review and editing. Filippo Molinari: Supervision, Writing – review and editing. Massimo Salvi: Conceptualization, Formal analysis, Methodology, Supervision, Visualization, Writing – original draft.

Acknowledgements

This research was supported by Fondazione Compagnia di San Paolo (Italy). Grant ID: E13C23001660007.

Funding

This work was supported by Compagnia di San Paolo.

Conflicts of Interest

F. Molinari and M. Salvi are equity holders in AEQUIP S.r.l., Turin, Italy. M. Scotto has working relationships with AEQUIP S.r.l., Turin, Italy. Remaining authors declare no competing interests regarding the publication of this article.

Data Availability Statement

The complete test set used in this work has been made publicly available on Kaggle at https://www.kaggle.com/datasets/artemis90/slideinspect. The dataset includes: (i) WSIs, (ii) manual annotations for both tissue and artifact segmentation, (iii) automated stain quality assessments, and (iv) recommended actions generated by SlideInspect.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

References

1. V. Baxi, R. Edwards, M. Montalto, and S. Saha, "Digital Pathology and Artificial Intelligence in Translational Medicine and Clinical Practice," *Modern Pathology* 35 (2022): 23–32.

2. B. Gecer, S. Aksoy, E. Mercan, L. G. Shapiro, D. L. Weaver, and J. G. Elmore, "Detection and Classification of Cancer in Whole Slide Breast Histopathology Images Using Deep Convolutional Networks," *Pattern Recognition* 84 (2018): 345–356.

3. M. G. Hanna, O. Ardon, V. E. Reuter, et al., "Integrating Digital Pathology Into Clinical Practice," *Modern Pathology* 35 (2022): 152–164.

4. Y. Nan, F. Li, P. Tang, et al., "Automatic Fine-Grained Glomerular Lesion Recognition in Kidney Pathology," *Pattern Recognition* 127 (2022): 108648.

5. N. Kanwal, F. Pérez-Bueno, A. Schmidt, K. Engan, and R. Molina, "The Devil Is in the Details: Whole Slide Image Acquisition and Processing for Artifacts Detection, Color Variation, and Data Augmentation: A Review," *IEEE Access* 10 (2022): 58821–58844.

6. M. S. Hosseini, J. A. Z. Brawley-Hayes, Y. Zhang, L. Chan, K. N. Plataniotis, and S. Damaskinos, "Focus Quality Assessment of High-Throughput Whole Slide Imaging in Digital Pathology," *IEEE Transactions on Medical Imaging* 39 (2019): 62–74.

7. G. Campanella, A. R. Rajanna, L. Corsale, P. J. Schüffler, Y. Yagi, and T. J. Fuchs, "Towards Machine Learned Quality Control: A Benchmark for Sharpness Quantification in Digital Pathology," *Computerized Medical Imaging and Graphics* 65 (2018): 142–151.

8. Y. Chen, J. Zee, A. Smith, et al., "Assessment of a Computerized Quantitative Quality Control Tool for Whole Slide Images of Kidney Biopsies," *Journal of Pathology* 253 (2021): 268–278.

9. C. Senaras, M. K. K. Niazi, G. Lozanski, and M. N. Gurcan, "DeepFocus: Detection of Out-Of-Focus Regions in Whole Slide Digital Images Using Deep Learning," *PLoS One* 13 (2018): e0205387.

10. A. Janowczyk, R. Zuo, H. Gilmore, M. Feldman, and A. Madabhushi, "HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides," *JCO Clinical Cancer Informatics* 3 (2019): 1–7.

11. A. I. Wright, C. M. Dunn, M. Hale, G. G. A. Hutchins, and D. E. Treanor, "The Effect of Quality Control on Accuracy of Digital Pathology Image Analysis," *IEEE Journal of Biomedical and Health Informatics* 25 (2020): 307–314.

12. D. Giavarina, "Understanding Bland Altman Analysis," *Biochem Med (Zagreb)* 25 (2015): 141–151.

13. X. Moles Lopez, E. D'Andrea, P. Barbot, et al., "An Automated Blur Detection Method for Histological Whole Slide Imaging," *PLoS One* 8 (2013): e82710.

14. H. Wu, J. H. Phan, A. K. Bhatia, C. A. Cundiff, B. M. Shehata, and M. D. Wang, "Detection of Blur Artifacts in Histopathological Whole-Slide Images of Endomyocardial Biopsies," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2015), 727–730.

15. M. Babaie and H. R. Tizhoosh, "Deep Features for Tissue-Fold Detection in Histopathology Images," In: Digital Pathology: 15th European Congress, ECDP 2019, Warwick, UK, April 10–13, 2019, Proceedings 15, Springer, 2019: pp. 125–132.

16. T. Kohlberger, Y. Liu, M. Moran, et al., "Whole-Slide Image Focus Quality: Automatic Assessment and Impact on Ai Cancer Detection," *Journal of Pathology Informatics* 10 (2019): 39.

17. R. Brixtel, S. Bougleux, O. Lézoray, et al., "Whole Slide Image Quality in Digital Pathology: Review and Perspectives," *IEEE Access* 10 (2022): 131005–131035.

18. G. Smit, F. Ciompi, M. Cigéhn, A. Bodén, J. Van Der Laak, and C. Mercan, "Quality Control of Whole-Slide Images Through Multi-Class Semantic Segmentation of Artifacts," in: Medical Imaging with Deep Learning (2021).

19. M. Haghighat, L. Browning, K. Sirinukunwattana, et al., "Automated Quality Assessment of Large Digitised Histology Cohorts by Artificial Intelligence," *Scientific Reports* 12 (2022): 5002.

20. A. Patil, H. Diwakar, J. Sawant, et al., "Efficient Quality Control of Whole Slide Pathology Images With Human-In-The-Loop Training," *Journal of Pathology Informatics* 14 (2023): 100306.

21. Z. Weng, A. Seper, A. Pryalukhin, et al., "GrandQC: A Comprehensive Solution to Quality Control Problem in Digital Pathology," *Nature Communications* 15 (2024): 1–12.

22. N. Kanwal, F. Khoraminia, U. Kiraz, et al., "Equipping Computational Pathology Systems With Artifact Processing Pipelines: A Showcase for Computation and Performance Trade-Offs," *BMC Medical Informatics and Decision Making* 24 (2024): 288.

23. D. Ameisen, C. Deroulers, V. Perrier, et al., "Stack or Trash? Quality Assessment of Virtual Slides," in *Diagn Pathol* (Springer, 2013), 1–5.

24. M. Salvi, N. Michielli, K. M. Meiburger, et al., "Cyto-Knet: An Instance Segmentation Approach for Multiple Myeloma Plasma Cells Using Conditional Kernels," *International Journal of Imaging Systems and Technology* 34 (2024): e22984.

25. W. Zhang, J. Pang, K. Chen, and C. C. Loy, "K-net: Towards Unified Image Segmentation," *Advances in Neural Information Processing Systems* 34 (2021): 10326–10338.

26. Z. Liu, Y. Lin, Y. Cao, et al., "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (IEEE, 2021), 10012–10022.

27. T. Xiao, Y. Liu, B. Zhou, Y. Jiang, and J. Sun, "Unified Perceptual Parsing for Scene Understanding," in *Proceedings of the European Conference on Computer Vision (ECCV)* (Springer, 2018), 418–434.

28. F. Marzola, K. M. Meiburger, and M. Salvi, "Innovative Temporal Loss Function for Segmentation of Fine Structures in Ultrasound Images," in *2023 IEEE International Ultrasonics Symposium (IUS)*, ed. IEEE (IEEE, 2023), 1–4.

29. M. Salvi, A. Mogetta, A. Gambella, et al., "Automated Assessment of Glomerulosclerosis and Tubular Atrophy Using Deep Learning," *Computerized Medical Imaging and Graphics* 90 (2021): 101930.

30. J. Bertels, T. Eelbode, M. Berman, et al., *Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22* (Springer, 2019), 92–100.

31. B. Koonce and B. Koonce, *MobileNetV3* (Image Recognition and Dataset Categorization, 2021), 125–144.

32. O. Ronneberger, P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18* (Springer, 2015), 234–241.

33. H. Shakhawat, S. Hossain, A. Kabir, S. M. H. Mahmud, M. M. M. Islam, and F. Tariq, "Review of Artifact Detection Methods for Automated Analysis and Diagnosis in Digital Pathology," in *Artificial Intelligence for Disease Diagnosis and Prognosis in Smart Healthcare* (CRC Press, 2023), 177–202.

34. M. Salvi, A. Mogetta, K. M. Meiburger, et al., "Karpinski Score Under Digital Investigation: A Fully Automated Segmentation Algorithm to Identify Vascular and Stromal Injury of Donors' Kidneys," *Electronics* 9 (2020): 1644, https://doi.org/10.3390/electronics9101644.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

35. M. Salvi, F. Molinari, S. Iussich, et al., "Histopathological Classification of Canine Cutaneous Round Cell Tumors Using Deep Learning: A Multi-Center Study," *Frontiers in Veterinary Science* 8 (2021): 640944.

36. M. K. K. Niazi, A. V. Parwani, and M. N. Gurcan, "Digital Pathology and Artificial Intelligence," *Lancet Oncology* 20 (2019): e253–e261.

37. M. Cui and D. Y. Zhang, "Artificial Intelligence and Computational Pathology," *Laboratory Investigation* 101 (2021): 412–422.

38. D. Anand, G. Ramakrishnan, and A. Sethi, *Fast GPU-Enabled Color Normalization for Digital Pathology, in: 2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, ed. IEEE (IEEE, 2019), 219–224.

39. A. Janowczyk, A. Basavanhally, and A. Madabhushi, "Stain Normalization Using Sparse Autoencoders (StaNoSA): Application to Digital Pathology," *Computerized Medical Imaging and Graphics* 57 (2017): 50–61.

40. B. Schömig-Markiefka, A. Pryalukhin, W. Hulla, et al., "Quality Control Stress Test for Deep Learning-Based Diagnostic Model in Digital Pathology," *Modern Pathology* 34 (2021): 2098–2108.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** Supporting Information.

10981098, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ima.70292 by Massimo Salvi - Politecnico Di Torino Sist. Bibl Del Polit Di Torino , Wiley Online Library on [11/01/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License