POLITECNICO DI TORINO
Repository ISTITUZIONALE

Stain normalization in digital pathology: Clinical multi-center evaluation of image quality

Original
Stain normalization in digital pathology: Clinical multi-center evaluation of image quality / Michielli, N., Caputo, A., Scotto, M., Mogetta, A., Pennisi, O.A.M., Molinari, F., Balmativola, D., Bosco, M., Gambella, A., Metovic, J., Tota, D., Carpenito, L., Gasparri, P., Salvi, M.. - In: JOURNAL OF PATHOLOGY INFORMATICS. - ISSN 2153-3539. - ELETTRONICO. - 13:100145(2022). [10.1016/j.jpi.2022.100145]

Availability:
This version is available at: 11583/2971960 since: 2022-10-01T17:10:12Z

Publisher:
Elsevier

Published
DOI:10.1016/j.jpi.2022.100145

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)

19 November 2026

Journal of Pathology Informatics 13 (2022) 100145

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Stain normalization in digital pathology: Clinical multi-center evaluation of image quality

Nicola Michielli [a], Alessandro Caputo [b], Manuela Scotto [a], Alessandro Mogetta [a], Orazio Antonino Maria Pennisi [c], Filippo Molinari [a], Davide Balmativola [d], Martino Bosco [e], Alessandro Gambella [f], Jasna Metovic [f], Daniele Tota [f], Laura Carpenito [g,h], Paolo Gasparri [i], Massimo Salvi [a,*]

[a] Biolab, Polito[BIO]Med Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[b] Department of Medicine and Surgery, University Hospital of Salerno, Salerno, Italy
[c] Technology Transfer and Industrial Liaison Department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[d] Pathology Unit, Humanitas Gradenigo Hospital, Corso Regina Margherita 8, 10153 Turin, Italy
[e] Department of Pathology, Michele and Pietro Ferrero Hospital, 12060 Verduno, Italy
[f] Pathology Unit, Department of Medical Sciences, University of Turin, Via Santena 7, 10126 Turin, Italy
[g] Department of Pathology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[h] University of Milan, Milan, Italy
[i] UOC di Anatomia Patologica, ASP Catania P.O. "Gravina", Caltagirone, Italy

ARTICLE INFO

Keywords:
Stain normalization
Image quality
Qualitative score
Digital pathology
H&E staining

ABSTRACT

In digital pathology, the final appearance of digitized images is affected by several factors, resulting in stain color and intensity variation. Stain normalization is an innovative solution to overcome stain variability. However, the validation of color normalization tools has been assessed only from a quantitative perspective, through the computation of similarity metrics between the original and normalized images. To the best of our knowledge, no works investigate the impact of normalization on the pathologist's evaluation.

The objective of this paper is to propose a multi-tissue (i.e., breast, colon, liver, lung, and prostate) and multi-center qualitative analysis of a stain normalization tool with the involvement of pathologists with different years of experience. Two qualitative studies were carried out for this purpose: (i) a first study focused on the analysis of the perceived image quality and absence of significant image artifacts after the normalization process; (ii) a second study focused on the clinical score of the normalized image with respect to the original one.

The results of the first study prove the high quality of the normalized image with a low impact artifact generation, while the second study demonstrates the superiority of the normalized image with respect to the original one in clinical practice.

The normalization process can help both to reduce variability due to tissue staining procedures and facilitate the pathologist in the histological examination. The experimental results obtained in this work are encouraging and can justify the use of a stain normalization tool in clinical routine.

Introduction

Digital pathology (DP) is quickly gaining traction and widespread adoption, thanks to a number of studies that are progressively demonstrating its merits in the diagnostic setting while explaining, contextualizing, and offsetting its upfront economic costs.[1–3] DP encompasses all the technologies that leverage digital slides to allow improvements and innovations in workflow (i.e., laboratory information system - LIS, workflow management, digital image analysis, labelling, and tracking). The advent of DP enabled pathologists to review digital tissue slides, share for telepathology and second opinion consultation, and store tissue samples in a high-resolution format in order to create a large digital database of histopathological images.[4,5] One of the major benefits of DP is the ability to conveniently and effortlessly utilize the whole-slide images (WSIs) for computer-aided diagnosis (CAD) and artificial intelligence (AI)-based tools. No explicit scanning step is required since in a fully digital laboratory the WSIs have already been scanned for primary diagnosis.[6] However, despite standardized procedures, the histopathology workflow is complex and mostly human- and device-dependent, hence numerous factors can affect the final appearance of the stained tissue. The histopathology process has several discrete steps requiring manual intervention and the device-related artifacts may have an impact on the final slide quality. All steps

* Corresponding author at: Biolab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy; Corso Duca degli Abruzzi 24, 10129 Turin, Italy.
E-mail address: massimo.salvi@polito.it (M. Salvi).

http://dx.doi.org/10.1016/j.jpi.2022.100145
Received 26 August 2022; Received in revised form 14 September 2022; Accepted 22 September 2022
Available online 24 September 2022
2153-3539/© 2022 The Author(s). Published by Elsevier Inc. on behalf of Association for Pathology Informatics. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

can affect the final appearance of tissue in the WSI, from surgical removal, transport to the lab, and fixation, up to staining, coverslipping, and scanning.[7] This represents a problem since a minor variation in the final WSI, which represents no problem for the human eye, can be enough to significantly affect the performance of AI/CAD tools.[8,9] As the histological laboratories are increasingly automated and the procedures become standardized, a significative caveat that still burden the routine diagnostic workup is represented by slide stain quality. Indeed, several factors can affect the final appearance of slides staining, resulting in color and intensity variation in the histopathological images. The most relevant phase that contributes to stain variability, intended as different colors and concentrations of staining dyes perceived on the digital scanned image, is undoubtedly related to the preanalytical phase: protocols variability depending on laboratories expertise and facilities availability, techniques and timing of tissue fixation, and imaging scanners.[7,10,11] The increasing use of WSIs and DP tools recently paved the road to an innovative solution to overcome the stain variability, that is the introduction of stain normalization strategies.[12–16] Indeed, digitally driven stain normalization process allows to standardize the stain color appearance of a source image with respect to a reference image (also denoted as the target image) with no further intervention in laboratory preanalytical phases and regardless procedures protocols, specific expertise, or laboratory facilities. Stain normalization should be done in such a way that the processed image maintains good contrast between cellular structures and preserves all the source information. The normalization process could be applied to digitized frozen section (FS) slides and histochemical or immunohistochemical staining. The development of WSIs enabled the use of a telepathology network for remote diagnosis of FS images which can be remotely reviewed at high resolution comparable to a traditional light microscope, connecting different laboratories at any time. In addition, the time required by a pathologist for WSI interpretation was proved to be shorter than that for conventional light microscopy. Besides the primary diagnosis, the digitized FS are the only available option for a second opinion teleconsultation and especially for intraoperative diagnoses in the field of transplant surgery due to urgent time constraints. In this context, the standardization of WSIs performed by a stain normalization process, is placed in the perspective of developing a telepathology network with the aim of improving the quality of slides and preserving the integrity of the tissue also for testing biomarkers or carrying out further investigation.[17,18] In Fig. 1, the process of stain normalization referred to the most frequently used stain in clinical routine i.e., hematoxylin and eosin (H&E), is reported.

The aim of stain normalization is twofold: firstly, the standardization of stain color appearance in digital pathology could improve the pathologist's work in the diagnosis of biological diseases, avoid the manual re-staining process and reduce the intra- and inter-operator variability; secondly this method could be used as a preprocessing step for CAD systems for accurate cellular structure segmentation and classification based on AI techniques. Several studies show that stain normalization is useful for AI tools[9,19–21] but, to the best of our knowledge, no works investigate the impact of normalization on the pathologist's evaluation. Quality control issues in digital pathology were tested in terms of how image artifacts may negatively influence the classification performance of machine or deep learning (DL)-based models,[22,23] but there is a lack of studies which investigate the image quality of multi-tissue slides from the pathologist's point of view in clinical practice.

In this paper, we assess the normalization process performed by STAINS - STAndardIzation & Normalization of histological Slides - tool (AEQUIP S.r.l., Turin, Italy), an improved version of a previously published algorithm.[13] To evaluate the clinical impact of the stain normalization, a multicentric qualitative analysis focused on image quality and clinical score was carried out with the involvement of pathologists from different institutions. The intended clinical benefit is expressed in terms of improvement of the quality of processed images available for the diagnosis to prove the need for a stain normalization tool for the pathologist.

The remainder of this paper is organized as follows: in next section, the description of materials and methods is provided, the experimental results are reported and discussed in the last two sections.

Materials and methods

The image quality and clinical score were qualitatively assessed on the normalization of digitized tissue slides, performed by STAINS tool. The images analyzed are derived from formalin-fixed and paraffin-embedded samples stained with H&E, taken from the following 5 organs: breast, colon, liver, lung, and prostate. The qualitative comparison was performed on image fields with at least 60% of tissue, extracted from the WSIs. Each tissue image field has a fixed dimension of 2000 × 2000 pixels, at 20 × magnification, covering a physical area of 1 mm × 1 mm. Two qualitative studies were carried out for this purpose, involving the end-users, i.e., the pathologists. The first study was focused on the analysis of the perceived image quality and absence of significant image artifacts in the normalized image obtained using STAINS tool, in order to investigate the robustness of the normalization process in the worst-case scenario, referring to digitized image fields with a reduced or non-optimal image quality in the starting original acquisition. The second study was focused on the clinical score of the normalized image with respect to the original one, performed on images with quality variability similar to ones used in clinical routine.

First study: Perceived image quality in the worst-case scenario

In the first study, 10 pathologists with different years of experience in anatomic pathology (experience: 9.1 ± 5.8 years, range: 4–20 years, 7 males and 3 females) and from different clinical centers were invited to inspect the normalized images. The list of the pathologists involved in this study is summarized in Table 1.

For each of the 5 tissues (i.e., breast, colon, liver, lung, and prostate), 10 H&E image fields were selected for the analysis, and the same 50 images were scored by each pathologist. For the qualitative assessment, the worst cases were selected for the analysis, intended as image fields where the original WSI did not have a good quality in terms of stain concentration, with a too high/weak color intensity or low contrast image. These images were eligible for the normalization process, which reached a compromise between reproducing the target stain colors and creating image artifacts. Both the original and normalized images were presented to each

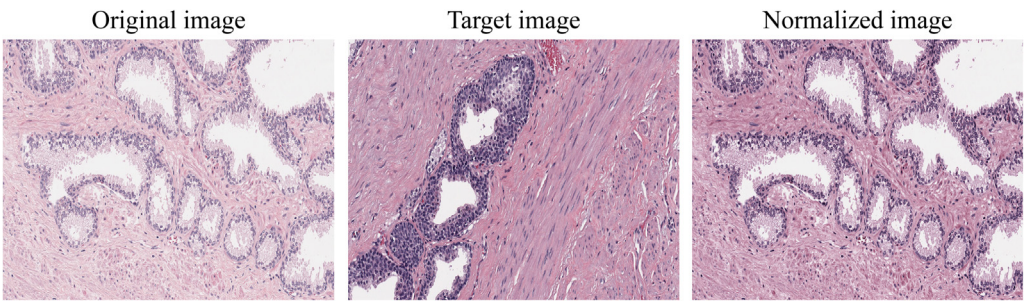


Fig. 1. Stain normalization of a H&E-stained tissue slide: original image (left), target image (center), and normalized image (right).

Table 1
List of the pathologists involved in the study.

ID pathologist	Initials	Clinical center - Affiliation	Years of experience
P1	L.C.	Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy	4
P2	J.M.	Molinette Hospital, Turin, Italy	4
P3	A.G.	Molinette Hospital, Turin, Italy	5
P4	D.T.	Molinette Hospital, Turin, Italy	5
P5	A.C.	University Hospital of Salerno, Salerno, Italy	6
P6	P.G.	Gravina Hospital, Caltagirone, Italy	7
P7	M.G.	Molinette Hospital, Turin, Italy	10
P8	D.B.	Humanitas Gradenigo Hospital, Turin, Italy	12
P9	A.F.	San Luigi Gonzaga Hospital, Orbassano, Italy	18
P10	M.B.	Michele and Pietro Ferrero Hospital, Verduno, Italy	20

pathologist and the normalized images were evaluated in terms of the image quality and the absence of structure or color artifacts. For each normalized image, each pathologist was asked to score:

- the image quality, intended as the loss of contrast between hematoxylin nuclear and eosin cytoplasmatic/stromal structures,[24] graded as "1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent";
- the presence of image artifacts, intended as any significant structure or color variations, created by the normalization process, that may have an impact on the clinical path[25,26];
- the impact of artifacts (if present) on the clinical evaluation, graded as "1: irrelevant", "2: acceptable," and "3: negative".

The significant structure or color variations include the unwanted color generation inside the white or uncolored regions (e.g., background, gland lumen, etc.), the generation of unrealistic colors inside the tissue regions or other chromatic variations which may have a negative impact on the overall clinical assessment. The workflow employed for the first experiment, i.e., the evaluation of the perceived image quality in the worst-case scenario, is summarized in Fig. 2.

Second study: Clinical score between the original and normalized image

In the second study, 3 experts were recruited among the pathologists involved in the first study, as listed in Table 1, specifically pathologists P5 (A.C.), P8 (D.B.), and P10 (M.B.) with different levels of experience (6, 12, and 20 years of experience, respectively). They were invited to inspect both the original image and the image processed by STAINS tool, after having selected 1 specific target image for each of the 5 tissues. The choice of the target image was based on the pathologist's opinion according to his/her clinical expertise and with the objective of improving the diagnosis.[13,24] The optimal target image was selected by each pathologist from a batch of 4 target image fields (at 20x magnification, i.e., 0.5 μm/pixel) for each of the 5 tissues, with different stain protocols of H&E to cover most of the stain variability. For each tissue, 3 image fields were scored independently by the pathologists (for a total number of 15 images). A clinical score in the range of 1–5 ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") was assigned by the pathologist for each original image and the corresponding image normalized by STAINS with respect to the selected target.[12,27] Fig. 3 shows the workflow adopted in this second study, i.e., the evaluation of the clinical score between the original and normalized image.

Results

First study: Perceived image quality in the worst-case scenario

In the first study, each normalized image was evaluated in terms of the perceived image quality and the creation of image artifacts in the worst-case scenario, as previously described. For each of the 5 tissues, 10 image fields were scored by 10 pathologists with different years of experience, so for each image we collected 10 evaluations, 1 for each pathologist. The average quality score for the normalized image is 4.266 ± 0.856. The average values with the standard error of the image quality score, are reported in Fig. 4 with respect to histological tissues and in Fig. 5 with respect to the pathologists involved in the study. The average value of the image quality score for the normalized image is higher than 4.3 for all expert pathologists with 10 or more years of experience in anatomic pathology.

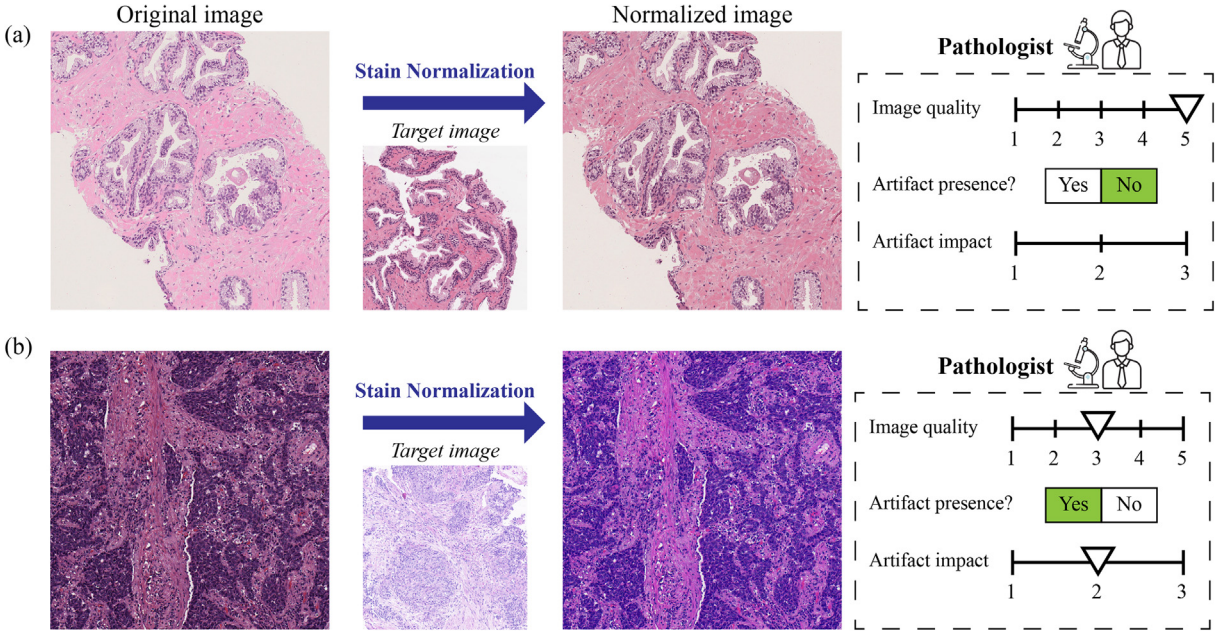


Fig. 2. Workflow adopted for the first study. H&E prostate (a) and lung (b) image field used for the first experiment. The normalized image, with respect to a specific target image, was evaluated by each pathologist in terms of image quality score ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") and impact score ("1: irrelevant", "2: acceptable," and "3: negative") of artifacts, if present in the normalized image.

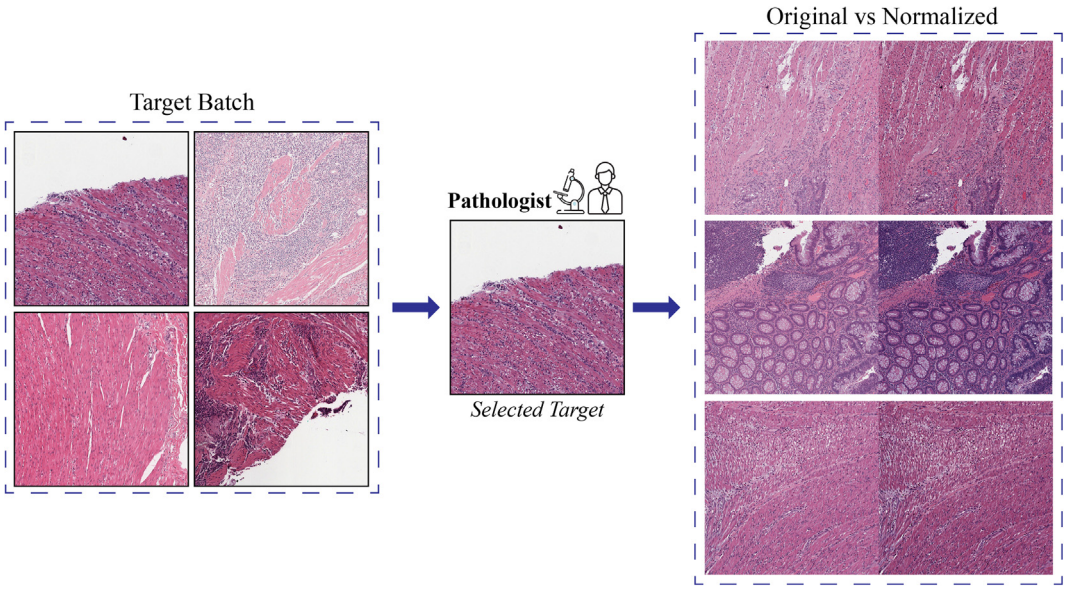


Fig. 3. Workflow adopted for the second study related to colon tissue. Each pathologist was asked to select a target image from a batch of 4 image fields, and both the original image (left, in the last column) and normalized image (right, in the last column) with respect to the selected target were evaluated in terms of clinical score ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent").

In addition, any significant color or structure artifacts created by the normalization process were assessed. In Table 2, the number of image artifacts revealed by each pathologist in the normalized image is reported for each tissue. The overall percentage of image artifacts with 3 different levels of impact score (irrelevant, acceptable, and negative) is 19.4%. Only 1.8% of artifacts (i.e., 9 out of 500 cases) were observed by expert pathologists with 10 or more years of experience in anatomic pathology.

In Fig. 6, the percentages of cases which presented image artifacts, are reported for each tissue, according to the clinical impact score of the artifacts ("1: irrelevant", "2: acceptable," and "3: negative") assigned by the pathologists.

An example of image artifacts created by the normalization process scored by 2 pathologists as acceptable for the clinical evaluation, is reported in Fig. 7.

After the normalization process, only 8 cases out of 500 (1.6%) included image artifacts with a negative impact score on the clinical evaluation. Nevertheless, the 8 impactful artifacts identified during the evaluation were reported by only 3 pathologists out of 10 that participated in the study. In addition, all observations including the clinically significant artifacts were related to images where only one pathologist out of 10 experts identified an impactful artifact, except for the observation related to a single image of the liver tissue, which was reported as including an artifact by 2 pathologists. Regarding the 8 impactful artifacts, 6 observations originated from 6 different images while 2 observations originated from the same image labeled as artifactual with a negative impact by 2 out of 10 pathologists. Fig. 8 shows this H&E-stained liver tissue image: the negative artifact reported by 2 pathologists (i.e., P2 and P6) was related to the decrease in image contrast and the darkening of staining after the normalization process, as shown in the zoomed-in view of Fig. 8. The same image was denoted with artifacts by 4 out of 10 pathologists with a different impact score (i.e., P1: acceptable, P2: negative, P3: irrelevant, P6: negative), while the remaining 6 pathologists involved in the study, had observed no significant artifacts for this image.

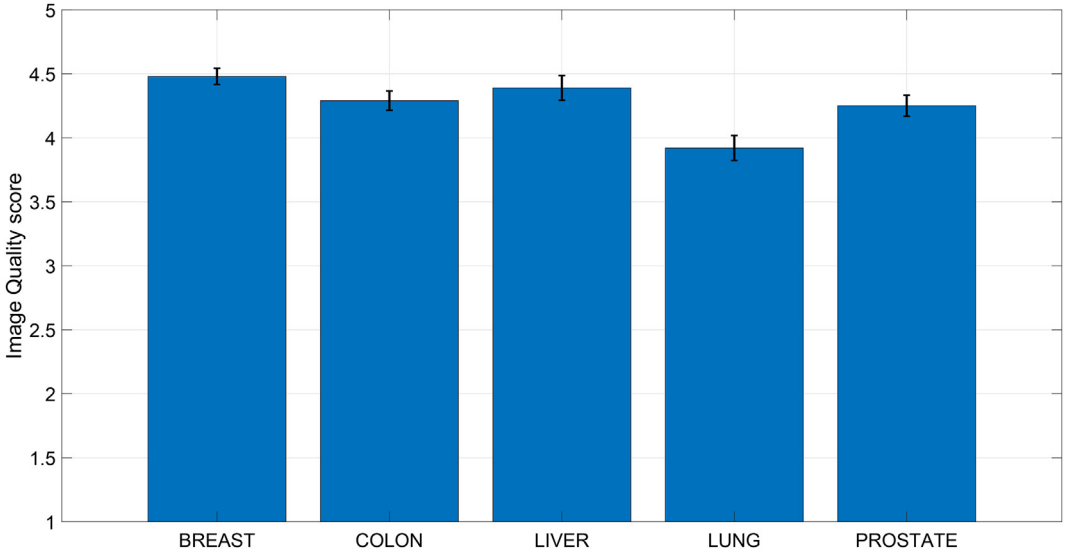


Fig. 4. Quality score for the normalized image in the range 1–5 ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") expressed as mean value with standard error bars, for each of the 5 tissues analyzed.

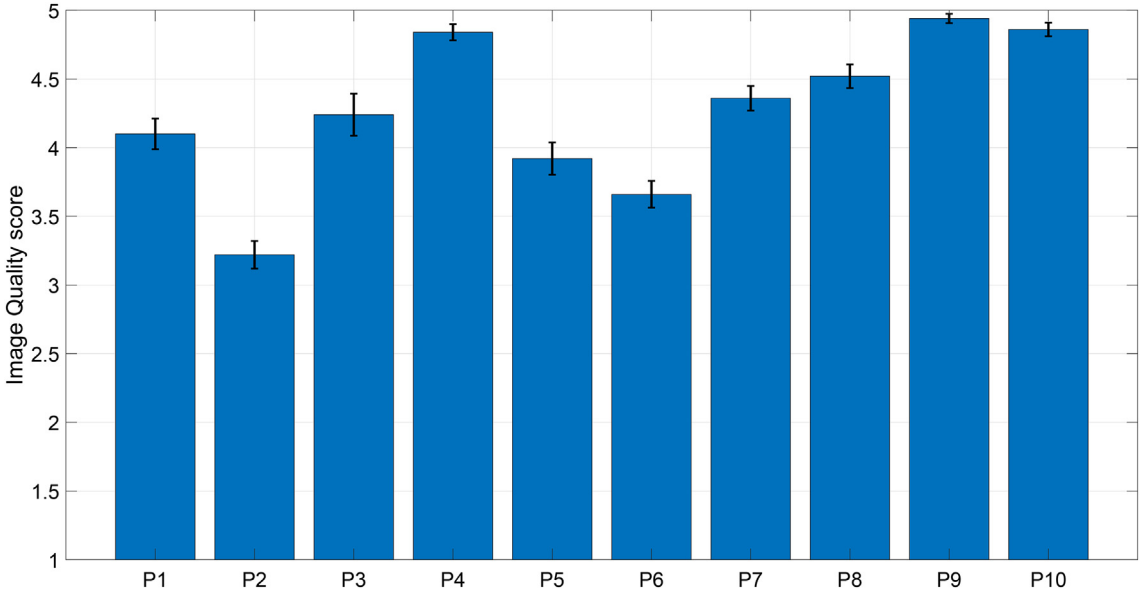


Fig. 5. Quality score for the normalized image in the range 1–5 ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") expressed as mean value with standard error bars, for each of the 10 pathologists involved in the study.

Table 2
Number of artifacts in the normalized images revealed by each pathologist for each of the 5 tissues.

Tissue	# Artifacts revealed by each pathologist										
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
Breast	2	3	0	1	0	6	2	0	0	0	14
Colon	7	4	0	1	0	8	4	2	0	1	27
Liver	2	9	2	2	0	5	0	0	0	0	20
Lung	7	9	1	0	0	1	0	0	0	0	18
Prostate	6	2	0	0	0	10	0	0	0	0	18
Total	24	27	3	4	0	30	6	2	0	1	97 **(19.4%)**

Second study: Clinical score between the original and normalized image

In the second study, 3 image fields for each of the 5 tissues (for a total of 15 images) were inspected by 3 pathologists who scored the original image and the image normalized by STAINS, with respect to the target selected by each pathologist independently for each tissue.

The average clinical scores for all cases are 4.333 ± 0.739 and 4.867 ± 0.344 for original and normalized images, respectively. The average clinical score in the range 1–5 ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") and the comparison between the original and normalized images are reported in Table 3.

The percentage of cases where the clinical score of the normalized image is higher or equal than the original one, over the total number of cases, is 88.9%. An example of a normalized image considered better, equal, and worse with respect to the original one, by the same pathologist, is reported in Fig. 9.

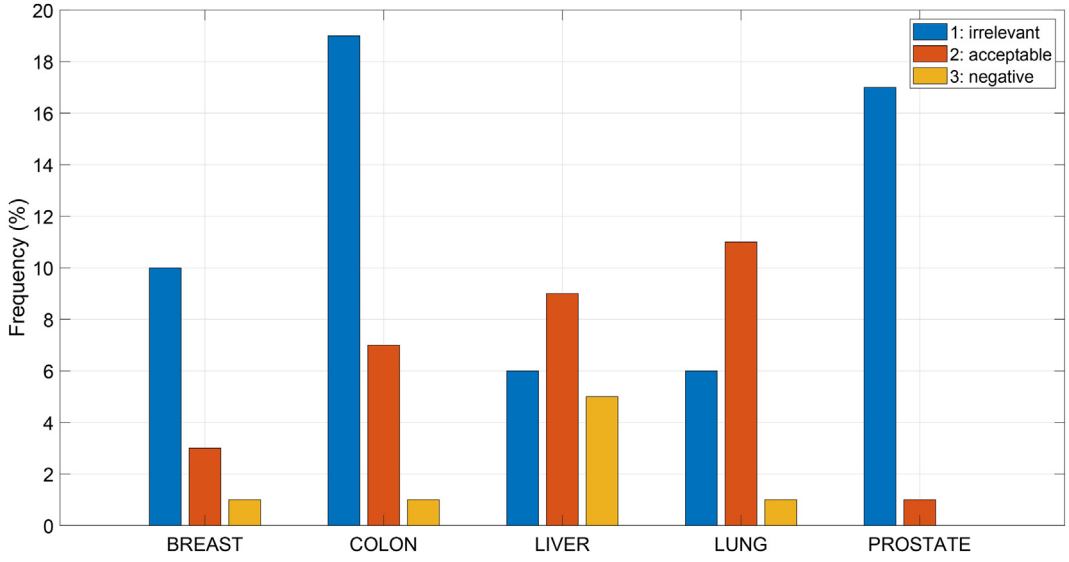


Fig. 6. Histograms of the impact score of artifacts on the clinical evaluation, in the range 1–3 ("1: irrelevant", "2: acceptable," and "3: negative") for each of the 5 tissues analyzed.

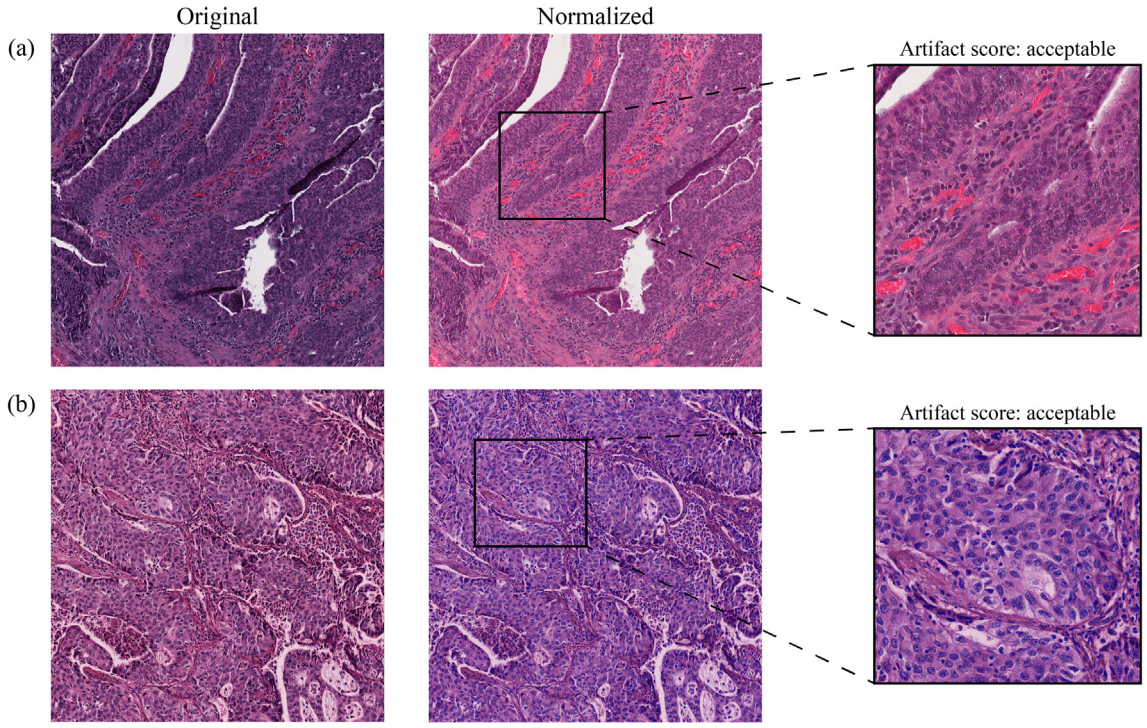


Fig. 7. Image artifacts generated by the normalization process with an impact score of "2: acceptable" on the clinical evaluation. (a) Colon tissue field: inadequate color contrast between nuclei and cytoplasm, difficult discrimination between stroma and epithelium, nuclei with internal structure not easily recognizable. (b) Lung tissue field: hematoxylin nuclei color too emphasized, loss of chromatin and nucleoli details.

Discussion and conclusions

Despite the great development of immunohistochemical and biomolecular methods in recent years, still today the pathological diagnosis is mainly based on the study of H&E-stained slides. On the other hand, random or systematic differences in the intensity, saturation, and color contrast in sample preparations from different laboratories may reduce the diagnostic reproducibility, in particular when the evaluation of the nuclear texture is crucial, such as in the grading of preinvasive colorectal or breast lesions.[28,29]

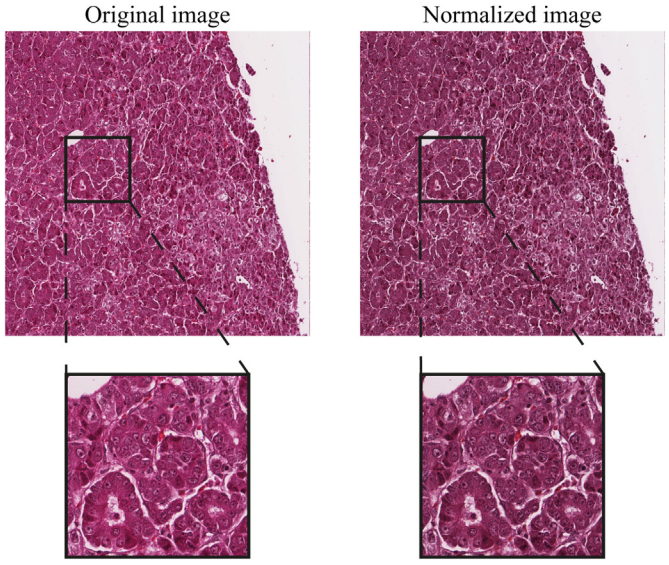


Fig. 8. H&E liver tissue original (left) and normalized (right) image labeled as artifactual by 4 out of 10 pathologists (2 of which assigned a negative artifact impact score) with a zoomed-in view showing the decrease of image contrast and the darkening of the H&E staining.

This problem results more relevant today as the exchange of digitized histological slides between laboratories has become a common practice.

The standardization of preanalytical laboratory procedures is a challenging objective complex to achieve due to several factors such as different equipment, different staining dyes, and environmental variability. The digitally driven standardization performed by automated tools as STAINS, would basically allow to overcome these difficulties without going to change the preclinical equipment and expertise. Therefore, stain normalization can be considered an automated method to reduce stain variability and an additional aid in the process of standardizing laboratory slides. In addition, the stain normalization benefits can be appreciated in the field of telepathology, in terms of reducing the number of tissue slides asked to the lab for second opinion consultation with the opportunity of reducing time of diagnosis.[17,18]

The validation of color normalization tools has been assessed in literature only from a quantitative perspective, through the computation of similarity metrics between the original and normalized images,[30–32] but there is a lack in the pathologist's assessment of image quality. To the best of our knowledge, this is the first work that proposes a multi-tissue and multi-center qualitative analysis of a stain normalization tool with the involvement of the end-users, i.e., pathologists with different years of experience in anatomic pathology. Currently, in clinical practice, the quality control of histopathological images suffers from the high inconsistency of

Table 3
Average clinical score and number of images, for each pathologist, where the score of normalized (NORM) image is better, equal, or worse than original (ORIG) one.

			Clinical score NORM vs. ORIG		
ID pathologist	Clinical score for ORIG	Clinical score for NORM	# Images where NORM > ORIG	# Images where NORM = ORIG	# Images where NORM < ORIG
P5 (A.C.)	4.333	4.867	7	6	2
P8 (D.B.)	4.267	4.733	7	5	3
P10 (M.B.)	4.400	5.000	8	7	0

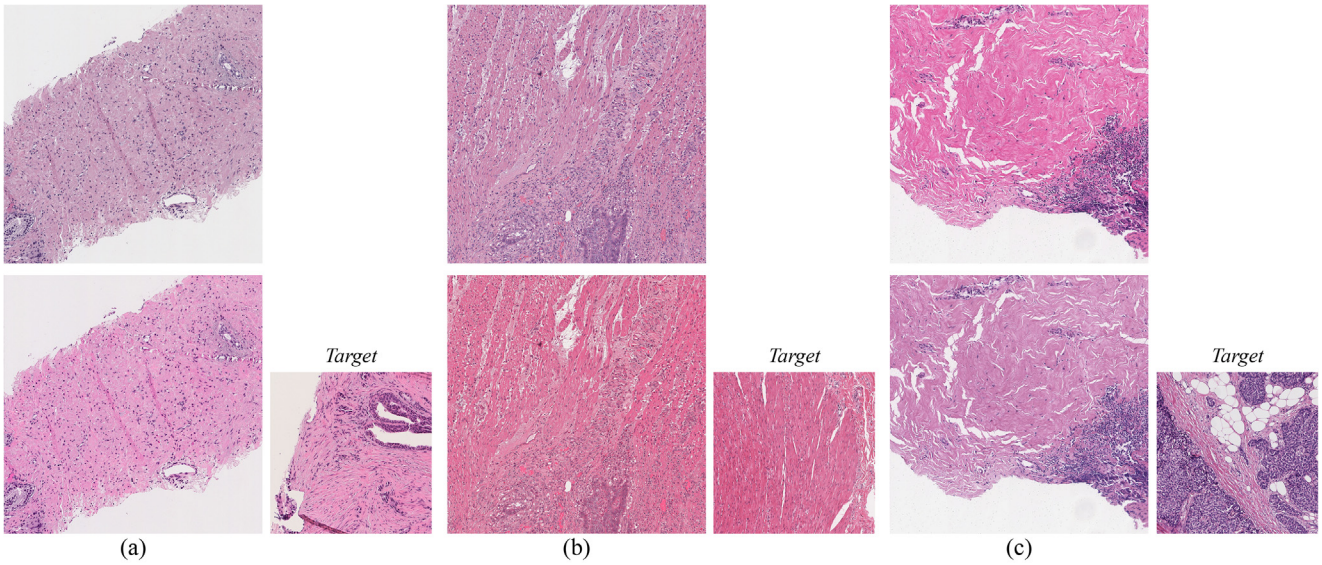


Fig. 9. Normalized images with respect to the selected target (bottom row), classified as better (a: prostate tissue), equal (b: colon tissue), and worse (c: breast tissue) than the corresponding original ones (top row), by the same pathologist.

the subjective quantification of stain quality and the inter-laboratories variability of procedures in staining tissue slides.[33] To solve this problem, a stain normalization tool which standardizes the staining variability with respect to optimally selected target colors may help the pathologist to improve the diagnosis.

The main findings of this work are related to the perceived image quality and clinical score of the standardization of histological H&E-stained tissue slides. Results of the first study indicate that the normalized image provided by a fully automated stain normalization tool, named STAINS, obtained an average image quality score of 4.266 on a 5-point qualitative scale ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent"), despite the original images presented a reduced or non-optimal image quality. All analyzed tissues (i.e., breast, colon, liver, lung, and prostate) except lung tissue, reached an average quality score higher than 4.2 and all expert pathologists with 10 or more years of experience assigned an average score higher than 4.3 for the normalized images, as reported in Fig. 4 and Fig. 5. In addition, the structure or color artifacts generation in the normalized image was inspected by the pathologists and the results show that the 19.4% of cases were labeled as artifactual with a 3-point scale of impact score ("1: irrelevant", "2: acceptable," and "3: negative"). More specifically, only in 1.8% of cases, artifacts were observed by experts with 10 or more years of experience and only 1.6% were considered negative for the diagnosis, as summarized in Table 2 and Fig. 6. The high inter-operator variability was proved by the fact that all observations including the clinically significant artifacts were related to images where only 1 pathologist out of 10 experts identified an impactful artifact except for 1 image labeled with negative artifacts with an agreement of 2 out of 10 pathologists, reported in Fig. 8. For this image, other 2 pathologists out of 10 experts observed an artifact with different impact scores and the remaining 6 out of 10 revealed no artifacts presence for that image. The results obtained in the first study prove the high quality of the normalized image with a low impact artifact generation, especially in the worst-case scenario. Moving on to the second study, the average clinical scores on a 5-point qualitative scale ("1: bad", "2: poor", "3: acceptable", "4: good," and "5: excellent") are 4.867 and 4.333 for the normalized and original image, respectively, with a lower standard deviation for the normalized configuration (0.344 vs. 0.739). In addition, the percentage of cases where the clinical score of the normalized image is higher or equal than the original one is 88.9%, as shown in Table 3. The results are in accordance with the expected subjective evaluation of the image quality depending on the original starting image, the selected target image, and the different experiences of the involved pathologists. The results obtained in the second study prove the superiority of the normalized image with respect to the original one in the clinical practice.

In conclusion, the benefit of color normalization to the quality of digital tissue slides results in an increase of perceived image quality with a low impact artifacts generation. This could be extended, in future, to the diagnostic process. Indeed, diagnosis could be adversely affected by staining variability. These color variations often impose obstacles to clinical diagnosis and prognosis performed by humans, as well as machines.[4] In addition, each pathologist, as proved in this study, has different staining preferences. Considering these factors, the normalization process can help both to reduce variability due to tissue staining procedures and to facilitate the pathologist in the histological examination.

The proposed study has some limitations. The analysis of a larger number of images is advisable to further validate the procedure and further efforts are needed to extend the quality analysis to WSIs.[34] In this study, the artifacts generated by the normalization process were evaluated by the pathologists on 2000 x 2000 pixels image fields. In future, the most common artifacts generated in routine histology, such as dust or dark spots, synthetic threads or tissue folds, scratches, and unfocused regions,[23] could be clinically assessed. Despite these limits, the experimental results obtained in this work are encouraging and can justify the use of stain normalization tools similar to STAINS in clinical routine. Future studies can be carried out with the aim of evaluating other biological tissues and other histochemical stains, such as periodic acid-Schiff (PAS) and trichrome staining (e.g., Mallory's and Masson's trichrome), or immunohistochemical biomarkers.

CRediT authorship contribution statement

Nicola Michielli: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Alessandro Caputo:** Data curation, Supervision, Writing – original draft, Writing – review & editing. **Manuela Scotto:** Formal analysis, Methodology, Writing – review & editing. **Alessandro Mogetta:** Conceptualization, Formal analysis, Methodology, Writing – review & editing. **Orazio Antonino Maria Pennisi:** Conceptualization, Supervision, Writing – review & editing. **Filippo Molinari:** Supervision, Writing – review & editing. **Davide Balmativola:** Data curation, Writing – review & editing. **Martino Bosco:** Conceptualization, Data curation, Writing – review & editing. **Alessandro Gambella:** Data curation, Writing – review & editing. **Jasna Metovic:** Data curation, Writing – review & editing. **Daniele Tota:** Data curation, Writing – review & editing. **Laura Carpenito:** Data curation, Writing – review & editing.

Paolo Gasparri: Data curation, Writing – review & editing. **Massimo Salvi:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

M. Salvi, O.A.M. Pennisi and F. Molinari are equity holders in AEQUIP S. r.l, Turin, Italy. Remaining authors declare no competing interests regarding the publication of this article.

Acknowledgements

The authors would like to acknowledge Dr. Alessandro Fornari (San Luigi Gonzaga Hospital, Orbassano, Italy) and Dr. Margherita Goia (A.O.U. Città della Salute e della Scienza Hospital, Turin, Italy) for their valuable assistance in the quality evaluation of the histological normalized images.

References

1. Fraggetta F, L'Imperio V, Ameisen D, Carvalho R, Leh S, Kiehi T, et al. Best practice recommendations for the implementation of a digital pathology workflow in the anatomic pathology laboratory by the European Society of Digital and Integrative Pathology (ESDIP). Diagnostics 2021;11(11). https://doi.org/10.3390/diagnostics11112167.
2. Baidoshvili A, Stathonikos N, Freling G, Bart J, Hart N, Laak J, et al. Validation of a whole-slide image-based teleconsultation network. Histopathology 2018;73(5):777–783. https://doi.org/10.1111/his.13673.
3. Retamero JA, Aneiros-Fernandez J, Del Moral RG. Complete digital pathology for routine histopathology diagnosis in a multicenter hospital network. Arch Pathol Lab Med 2020;144(2):221–228. https://doi.org/10.5858/arpa.2018-0541-OA.
4. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20(5):e253–e261. https://doi.org/10.1016/S1470-2045(19)30154-8.
5. Rodriguez JPM, Rodriguez R, Silva VWK, Kitamura FC, Corradi GCA, De Marchi ACB, et al. Artificial intelligence as a tool for diagnosis in digital pathology whole slide images: a systematic review. J Pathol Inform 2022:100138. https://doi.org/10.1016/j.jpi.2022.100138.Published online.
6. Fraggetta F, Caputo A, Guglielmino R, Pellegrino MG, Runza G, L'Imperio V. A survival guide for the rapid transition to a fully digital workflow: the "caltagirone example.". Diagnostics 2021;11(10). https://doi.org/10.3390/diagnostics11101916.
7. Salehi P, Chalechale A. Pix2pix-based stain-to-stain translation: a solution for robust stain normalization in histopathology images analysis. 2020 International Conference on Machine Vision and Image Processing (MVIP). IEEE; 2020. p. 1–7.
8. Khan A, Janowczyk A, Müller F, Blank A, Nguyen HG, Abbet C, et al. Impact of scanner variability on lymph node segmentation in computational pathology. J Pathol Inform 2022;13, 100127. https://doi.org/10.1016/j.jpi.2022.100127.
9. Swiderska-Chadaj Z, de Bel T, Blanchet L, Baidoshvili A, Vossen D, Laak JVD, et al. Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. Sci Rep 2020;10(1):14398. https://doi.org/10.1038/s41598-020-71420-0.
10. Sapino A, Annaratone L, Marchiò C. Current projects in pre-analytics: where to go?. Pre-Analytics of Pathological Specimens in Oncology. Springer International Publishing; 2015. p. 65–70. https://doi.org/10.1007/978-3-319-13957-9_7.
11. Annaratone L, Marchiò C, Sapino A. Tissues under-vacuum to overcome suboptimal preservation. Nat Biotechnol 2019;52:104–109. https://doi.org/10.1016/j.nbt.2019.05.007.
12. Vahadane A, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, et al. Structure-preserving color normalization and sparse stain separation for histological images. IEEE Trans Med Imaging 2016;35(8):1962–1971. https://doi.org/10.1109/TMI.2016.2529665.
13. Salvi M, Michielli N, Molinari F. Stain Color Adaptive Normalization (SCAN) algorithm: separation and standardization of histological stains in digital pathology. Comput Methods Programs Biomed 2020;193, 105506. https://doi.org/10.1016/j.cmpb.2020.105506.
14. Liang H, Plataniotis KN, Li X. Stain style transfer of histopathology images via structure-preserved generative learning. Machine Learning for Medical Image Reconstruction. Springer International Publishing; 2020. p. 153–162.
15. Vijh S, Saraswat M, Kumar S. A new complete color normalization method for H&E stained histopatholgical images. Appl Intell 2021;51(11):7735–7748. https://doi.org/10.1007/s10489-021-02231-7.
16. Boschman J, Farahani H, Darbandsari A, Ahmadvand P, Spankeren AV, Farnell D, et al. The utility of color normalization for AI-based diagnosis of hematoxylin and eosin-stained pathology images. J Pathol 2022;256(1):15–24. https://doi.org/10.1002/path.5797.
17. Cima L, Brunelli M, Parwani A, Girolami I, Ciangherotti A, Riva G, et al. Validation of remote digital frozen sections for cancer and transplant intraoperative services. J Pathol Inform 2018;9:34. https://doi.org/10.4103/jpi.jpi_52_18.
18. Eccher A, Girolami I, Brunelli M, Novelli L, Mescoli C, Malvi D, et al. Digital pathology for second opinion consultation and donor assessment during organ procurement: review of the literature and guidance for deployment in transplant practice. Transplant Rev (Orlando) 2020;34(4), 100562. https://doi.org/10.1016/j.tre.2020.100562.
19. Salvi M, Acharya UR, Molinari F, Meiburger KM. The impact of pre- and post-image processing techniques on deep learning frameworks: a comprehensive review for digital pathology image analysis. Comput Biol Med 2021;128, 104129. https://doi.org/10.1016/j.compbiomed.2020.104129.
20. Li B, Keikhosravi A, Loeffler AG, Eliceiri KW. Single image super-resolution for whole slide image using convolutional neural networks and self-supervised color normalization. Med Image Anal 2021;68, 101938. https://doi.org/10.1016/j.media.2020.101938.
21. Pontalba JT, Gwynne-Timothy T, David E, Jakate K, Androutsos D, Khademi A. Assessing the impact of color normalization in convolutional neural network-based nuclei segmentation frameworks. Front Bioeng Biotechnol 2019;7. https://doi.org/10.3389/fbioe.2019.00300.
22. Wright AI, Dunn CM, Hale M, Hutchins GGA, Treanor DE. The effect of quality control on accuracy of digital pathology image analysis. IEEE J Biomed Heal Informatics 2021;25(2):307–314. https://doi.org/10.1109/JBHI.2020.3046094.
23. Schömig-Markiefka B, Pryalukhin A, Hulla W, Bychkov A, Fukuoka J, Madabhushi A, et al. Quality control stress test for deep learning-based diagnostic model in digital pathology. Mod Pathol 2021;34(12):2098–2108. https://doi.org/10.1038/s41379-021-00859-x.
24. Bejnordi BE, Litjens G, Timofeeva N, Otte-Holler I, Homeyer A, Karssemeijer N, et al. Stain specific standardization of whole-slide histopathological images. IEEE Trans Med Imaging 2016;35(2):404–415. https://doi.org/10.1109/TMI.2015.2476509.
25. Zheng Y, Jiang Z, Zhang H, Xie F, Shi J, Xue C. Adaptive color deconvolution for histological WSI normalization. Comput Methods Programs Biomed 2019;170:107–120. https://doi.org/10.1016/J.CMPB.2019.01.008.
26. Mc Grath O, Sarfraz MW, Gupta A, Yang Y, Aslam T. Clinical utility of artificial intelligence algorithms to enhance wide-field optical coherence tomography angiography images. J Imaging 2021;7(2). https://doi.org/10.3390/jimaging7020032.
27. Slone RM, Van Metter R, Senol E, Muka E, Pilgram TK. Effect of exposure variation on the clinical utility of chest radiographs. Radiology 1996;199(2):497–504. https://doi.org/10.1148/radiology.199.2.8668802.
28. Kumar N, Verma R, Chen C, Lu C, Fu P, Wills J, et al. Computer-extracted features of nuclear morphology in hematoxylin and eosin images distinguish stage II and IV colon tumors. J Pathol 2022;257(1):17–28. https://doi.org/10.1002/path.5864.
29. Chen JM, Li Y, Xu J, Gong L, Wang LW, Liu WL, et al. Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: a review. Tumour Biol J Int Soc Oncodevelopmental Biol Med 2017;39(3). https://doi.org/10.1177/1010428317694550.1010428317694550.
30. Sethi A, Sha L, Vahadane AR, Deaton RJ, Kumar N, Macias V, et al. Empirical comparison of color normalization methods for epithelial-stromal classification in H and E images. J Pathol Inform 2016;7(1):17. https://doi.org/10.4103/2153-3539.179984.
31. Roy S, kumar Jain A, Lal S, Kini J.. A study about color normalization methods for histopathology images. Micron 2018;114:42–61. https://doi.org/10.1016/j.micron.2018.07.005.
32. Yee WC, Jian TX, Ab Rahman KS, Hoe TL, Min LJ, Hang QY, et al. Performance Analysis of Color Normalization Methods in Histopathology Images. 2022 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS). IEEE; 2022. p. 147–151.
33. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Trans Biomed Eng 2014;61(6):1729–1738. https://doi.org/10.1109/TBME.2014.2303294.
34. Ashman K, Zhuge H, Shanley E, Fox S, Halat S, Sholl A, et al. Whole slide image data utilization informed by digital diagnosis patterns. J Pathol Inform 2022;13, 100113. https://doi.org/10.1016/j.jpi.2022.100113.