Segmentation is better when shared: a review of public H&E histological images datasets

Anna Clara Medina Roissmann
Institute of Science and Technology
Federal University of São Paulo
São José dos Campos, Brazil
Email: anna.medina28@unifesp.br

Thaína A. Azevedo Tosta
Institute of Science and Technology
Federal University of São Paulo
São José dos Campos, Brazil
Email: tosta.thaina@gmail.com

***Abstract*—The evaluation of histological images is a key step in cancer diagnosis, but it is a time-consuming and subjective process. To overcome these challenges, computer-aided diagnosis systems have emerged to offer a faster and more accurate analysis. Among the steps of these systems, image segmentation plays a crucial role by isolating regions of interest for further examination. In this context, this systematic review investigates the use of publicly available datasets in histological image segmentation analysis using Hematoxylin-Eosin (H&E) staining. The review addresses 15 guiding questions, covering various aspects, including the most common segmentation techniques, evaluation metrics, and existing limitations in the literature.**

I. Introduction

Cancer diagnosis involves the analysis of tissue samples. However, this manual process can be time-consuming and prone to subjective interpretation by pathologists [1]. The development of computer-aided diagnosis (CAD) systems represents a powerful alternative to deal with these challenges. Among the CAD systems processing steps, image segmentation and detection are crucial techniques, aiming to identify regions of interest (ROIs) [2], [3]. However, these algorithms require a significant number of annotated images, especially for training deep learning (DL) methods [4]. An alternative to having this kind of image volume is to use public images, which can help develop more robust algorithms for different tissue preparation protocols.

Visualization and analyses of tissue structures are possible through different staining, such as Periodic Acid-Schiff (PAS), Periodic Acid-Methenamine Silver (PAMS), Trichrome (TRI), and Hematoxylin-Eosin (H&E). Among them, cancer diagnoses can be confirmed by H&E histological images, which is a research topic of studies, including segmentation and detection of ROIs by computational methods [5], [6].

In this context, this systematic review presents different aspects of publicly available datasets and the methods used to process H&E histological images. Our contributions include insights about the segmentation and detection algorithms, their quantitative results and analyses, and the image datasets used to validate them. This is valuable for two key aspects: the development and dissemination of existing and new datasets, and the identification of future research directions by recognizing the most widely used segmentation and detection techniques applied to public images—a challenging validation context.

II. Materials and methods

This systematic review consists of articles dealing with the segmentation of publicly available H&E histological images of cancer and precancerous lesions. The articles' search considered the following keywords, combined into queries in digital libraries (Scopus, Web of Science, and PubMed): CAD systems, deep learning, image processing, classification, segmentation, machine learning, image analysis, algorithms, computer-aided diagnosis, H&E, histo* images, tissue, medical imaging, digital pathology, cancer* images, cancer diagnosis, cancer detection, data availability, available datasets, public datasets, challenge, contest, available images, benchmark.

The search stage was filtered using inclusion criteria and the following exclusion criteria: (i) removal of duplicates, (ii) articles with fewer than four pages, (iii) retracted papers, and (iv) use of regression techniques, prediction of survival and recurrence, treatment, and diagnosis. A total of 590 articles were initially identified, from which only 70 studies were selected to address the proposed research questions[1]. They were formulated by Computer Scientists, experts in H&E histological image processing, since the main topic of this review is the application of computational techniques on these images.

III. Results and discussion

GQ1 What are the main future works identified in the literature regarding the segmentation of public H&E histological images? The most cited future research included investigation of different tissues and datasets [2]–[4], [7]–[23], in addition to the use of segmented ROIs in other computational pathology tasks [14], [15], [19], [24]–[35], such as the analysis of morphological features for cancer progression prediction [36]. It was also proposed to investigate methods robust to color and scale [21], [37]–[39], tissue preparations protocols [3], [4], ROIs [9], [33], [40], [41], histopathology images staining (such as IHC) [9], [35], [41], [42], and magnifications [12]. Overlapped nuclei and false positives were observed limitations [35], which promote algorithms for border detection [43], investigation of loss functions [44], models incorporated with nuclei shapes [18],

[1]GQ stands for general questions, and SQ corresponds to specific questions.

[45], and instance segmentation [21], [46]. Barriers by the limited amount of images [1], [25], [47], [48] or specific ROIs [27], with unbalance [17], [38], reduce models generalization and increase overfitting, with possible use of scalable crowdsourcing for data annotation, generation of additional images by GANs [42], [47] and investigation of data augmentation [7], [13], [25], [38]. Open issues include optimizations to reduce inference and training time [3], [38], [39], [46], such as pruned models [49], and investigations of paralel processing [2], [34], [48]. More precise annotations [9], [35], by more than one pathologist [25], [30] or just datasets with more annotations [3], [50] were recognized as future demands. Methodologies investigations encompass individual model analyses in comparison to ensemble [51], integration with fuzzy segmentation [48], [52], transformers [44] or 3D models [20], [40]. Investigating structural similarities and boundary displacement [52], multi-scale features to deal with nuclei with different sizes [3], [38], in addition to other ROIs features [35], [42], [53] are also research possibilities, as well as application of pre- [10], [17], [21], [22] or post-processing [54]. Analyses of SSL training [46], [55], or backbones for DL [3], [20] can also be performed. Finally, new evaluation metrics can be proposed for the representation of shape, size, and proximity between nuclei [3].

GQ2 Are there published studies proposing public images for segmentation analyses? Among the analyzed articles, four of the proposed datasets are still available for download: CryoNuSeg [4], KMC-liver [21], LynSeC [56], and MthH [57].

GQ3 What is the most commonly used image processing technique for the segmentation of ROIs in histological images? This answer required the identification of the ROIs segmented in each study. Only DL techniques were investigated for segmentation of benign and malignant epithelium, breast tubules, cell detection, cell types, duct detection, epithelial and stromal nuclei, glandular epithelium, hepatocellular carcinoma, invasive breast cancer region, oral squamous cell carcinoma tumor regions, tumor-infiltrating lymphocytes, and tissue regions. Modifications of U-Net were explored [10], [13], [18], [23], [25], as well as transformers [23], [38] and various CNNs [9], [14], [24], [39]–[41], [51], [55]. For malignant nuclei segmentation, DL [58] and Otsu [59] were investigated. Considering mitoses (by cells, instances and nuclei), DL was widely used [15], [17], [19], [30], [33], [60], [61], in addition to neutrosophic sets [62], thresholding [31], [60], [61] and SVM [31], [60]. Glands were segmented by DL [10], [16], [26], [28], [51], [53], [54], Otsu [63], symbol pressure function-level set [26] and triangle membership [26]. Nuclei segmentation had a predominant use of DL [1], [3], [4], [7], [8], [11], [12], [20], [21], [27], [32], [35]–[37], [42]–[44], [46], [47], [49], [56], [57], [64]–[71], also including canny edge detection [50], [72], fuzzy clustering [2], Gaussian mixture models [45], graphcut [45], K-means [46], [50], [69], MLP [22], Otsu [29], [34], [50] and superpixels [34], [48], [52].

SQ1 Which segmentation techniques obtained the best quantitative results? This answer required a joint analysis of ROIs and evaluation metrics, disregarding the evaluated datasets. For cell type detection, the best AUC (0.99) was reached by [14]. In gland segmentation, the best Dice (0.923), F1-score (0.901), and Hausdorff distance (44.125) were obtained by [51], [26], and [26], respectively. Glandular epithelium had the best Dice (0.9119) in [23]. Mitoses obtained the highest values of F-score (0.767), precision (0.828), and recall (0.728), all in [33]. In mitosis cells, [30] reached the best precision (0.912) and recall (0.893). In nuclei segmentation, the best results are presented in the following: accuracy (0.9669) [3], Dice (0.914) [27], DQ (0.784) [71], F1-score (0.9579) [35], FN (4.4) [48], [52], FP (10.2) [48], [52], IoU (0.8911) [36], Jaccard index (0.963) [34], multi-class PQ (0.5290) [65], PQ (0.755) [56], precision (0.992) [27], recall (0.934) [27], SQ (0.768) [71] and TP (38.5) [52]. Mitotic nuclei obtained the best accuracy (88.43), sensitivity (90.13), and specificity (86.74) in [61]. Finally, the best Dice (0.84) for tumor-infiltrating lymphocytes was reached in [13].

SQ2 Where are the segmentation-based databases from? Most of the segmentation datasets came from unspecified regions, and three are from multiple centers. The UK and France have three datasets each. China, Germany, Italy, the Netherlands, and the USA have two datasets each. The ones that contributed to only one dataset each were Brazil, Canada, India, Japan, Portugal, and South Korea.

SQ3 What was the year in which most databases were released for segmentation? In 2021, the KMC [73] and CryoNuSeg [74] datasets were introduced. The most recent ones are MthH [75], proposed in 2022, and LyNSeC [76], proposed in 2024.

SQ4 How many databases are available for download? Among the analyzed studies, 35 datasets are still available for download (summarized in Table I).

SQ5 Which ROI has more public images available for segmentation analyses? The ROI with the largest number of publicly available images is the nuclei, with over 200,000 images from different tissues.

SQ6 Which cancer type offers the most public images available for segmentation? Considering that we couldn't take into account the types of images (i.e. WSIs vs patches, for instance), the cancer type with the highest number of images is the renal clear cell adenocarcinoma, with 192,581 publicly available images from TCGA-KIRC [97].

SQ7 What available public database has more ROIs or images? The criterion used in SQ6 was also applied to this question. Therefore, the highest number of images is in TCGA-KIRC. In terms of ROIs, the dataset with the highest quantity is PanNuke (with over 200,000 labeled nuclei) [96]. It is important to note that some articles did not specify the number of ROIs available.

SQ8 What were the evaluation metrics used for performance analyses of segmentation algorithms? This answer considered the different segmented ROIs. Although IoU is equivalent to the Jaccard index, and the Dice score can also be referred to as F1-score [108], we used the exact terms

TABLE I
PUBLICLY AVAILABLE DATASETS USED IN THE REVIEWED STUDIES, IN 2025.

ROIs	Datasets	Lesions	Number of images	Magnifications	Origin of samples
Duct detection	TCGA-BRCA [77]	Breast cancer	1,126 slides	20×	Unspecified
Glands	MICCAI GlaS [78]	Colorectal cancer	165 images	20×	UK
Hepatocellular carcinoma	PAIP [79]	Hepatocellular carcinoma	100 images	20×	South Korea
Invasive breast cancer regions	UHCMC&CWRU [80]	Breast cancer	110 imagens	20×	USA
Malignant cells	BreastPathQ [81]	Breast cancer	96 WSIs	20×	Canada
Mitotic cells	AMIDA 13 [82]	Breast cancer	606 HPFs	40×	The Netherlands
Mitotic cells	ICPR12 [83]	Breast cancer	50 HPFs	40×	France
Mitotic cells	MITOS & ATYPIA [84]	Breast cancer	1,420 frames	40×	France
Mitotic cells	TUPAC [85]	Breast cancer	500 images	40×	Unspecified
Mitosis	CCMCT-MEL [86]	Canine cutaneous mast cell tumor	32 WSIs	40×	Germany
Mitosis	MIDOG [87]	Breast cancer	150 samples	40×	Germany The Netherlands
Nuclei	BACH [88]	Breast cancer	40 images	-	Portugal
Nuclei	CoNIC [89]	Colorectal cancer	4,981 patches	20×	UK
Nuclei	CoNSep [90]	Colorectal adenocarcinoma	41 WSIs	40×	UK
Nuclei	CPM-17 [91]	Diverse tissues	32 images	20× and 40×	Unspecified
Nuclei	CryoNuSeg [74]	Diverse tissues	30 images	40×	Multiple centers
Nuclei	KMC [73]	Liver cancer	80 images	40×	India
Nuclei	LyNSeC [76]	Lymphoma	320 images	40×	Unspecified
Nuclei	MoNuSAC [92]	Diverse tissues	-	40×	Multiple centers
Nuclei	MoNuSeg [93]	Diverse tissues	30 images	40×	18 hospitals (USA, Canada, Germany, Australia)
Nuclei	MthH [75]	Thymic carcinoma	36,000 images	20× and 40×	China and Japan
Nuclei	NuCLS [94]	Breast cancer	1,744 FOVs	40×	Unspecified
Nuclei	NulnsSeg [95]	Diverse tissues	665 images	40×	Italy
Nuclei	PanNuke [96]	Diverse tissues	-	20× and 40×	Unspecified
Nuclei	TCGA-KIRC [97]	Renal cell carcinoma	192,581 images	40×	Unspecified
Nuclei	TNBC [98]	Diverse tissues	50 images	40×	France
Nuclei	UCSB [99]	Breast cancer	58 images	-	USA
Nuclei	[100]	Breast cancer	143 images	40×	Unspecified
Nuclei	[101]	Colorectal cancer	19 images	40×	Italy
Oral squamous cell carcinoma tumor regions	OCDC [102]	Oral cancer	15 WSIs	20×	Brazil
Oral squamous cell carcinoma tumor regions	ORCA [103]	Oral carcinoma	200 TMA	20× and 40×	Unspecified
Tumor infiltrating lymphocytes	BCa-lym [104]	Breast cancer	100 images	20×	Unspecified
Tumor infiltrating lymphocytes	Post-NAT-BRCA [105]	Breast cancer	96 images	20×	Unspecified
Tumor, stroma, lymphocytic infiltrate, necrosis, other	BCSS [106]	Breast cancer	151 WSIs	-	Unspecified
Tumor, stroma normal Tissue	LUAD-HistSeg [107]	Lung adenocarcinoma	54 WSIs	10×	China

of the authors to avoid any bias. For benign and malignant epithelium segmentation, only the Dice score was used [10]. Dice score, F1-score, mean IoU, precision, and recall evaluated breast tubules segmentation [39]. Cell detection and its types were most analyzed by AUC [14], [24], precision and recall [24], [40], in addition to accuracy [24], F1-score [24], [40], MAE [14], mean and standard deviation of the couting error and detection distance error [40], and SCC [14]. FROC, maximum sensitivity, and mAP were applied for duct detection [55]. Segmented glands were mostly evaluated through the Dice [10], [16], [23], [26], [28], [41], [51], [54], but also by accuracy [28], F1-score [16], [26], [41], [53], [54], Hausdorff distance [16], [26], [41], [54], IoU [23], Jaccard index [51], overlap [28], PPV [28], precision [23], recall/sensitivity [23], [28], specificity [28], and even with no quantitative evaluation [63]. Hepatocellular carcinoma and invasive breast cancer regions were both evaluated by Dice score and Jaccard index [51]. Segmentation of mitoses (by cells or instances) widely used F-score, precision, and recall [17], [19], [30], [33], [60], [62]. AUC [33], [62], accuracy [33], DMR and FDR [60] were also applied. The most widely used metrics to evaluate nuclei segmentation, including epitelial, stromal, mitotic, and malignant, were F1-score [1], [3], [7], [8], [11], [12], [15], [20]–[22], [29], [35], [42], [44], [46], [47], [50], [56], [59],

[61], [64], [66], [67], [71], [72], Jaccard index [4], [8], [12], [21], [29], [34], [35], [42]–[44], [47], [56], [57], [64], [66], [68]–[70], precision [1], [3], [7], [8], [11], [15], [20], [22], [27], [29], [34], [45], [47], [61], [67], [68], [71], accuracy [1]–[3], [11], [12], [29], [31], [36], [44], [46], [47], [50], [57], [61], [68], [72], recall [1], [3], [7], [8], [11], [15], [20], [22], [27], [29], [34], [45], [47], [67], [71] and the Dice score [2]–[4], [18], [27], [46], [48], [49], [56], [57], [64], [69], [70]. Other metrics for nuclei segmentation evaluation were AUC [18], aHD [42], binary PQ [71], boundary F1 [36], DQ [46], [71], F-measure [68], FN [48], [52], FP [48], [52], IGD [52], IoU [3], [7], [36], [44], [50], [68], HD [18], Kappa score [50], MS [52], MAE [57], MOS [72], MSE [50], multi-class PQ [65], [71], nuclei class evaluation [70], ODI [37], OHD [37], PQ [4], [46], [49], [56], [65], [70], [71], RMSE [50], SQ [46], [71], sensitivity [31], [59], [61], SP [52], specificity [11], [31], [61], SSIM [50], and TP [48], [52], [59]. The study [32] did not quantitatively evaluate this segmentation. Different tumor regions were also segmented, but mostly evaluated by Dice [9], [13], [38], F1-score [9], [25], IoU [25], [38] and precision [25], [38], allied to accuracy [25], PQ [9], recall [38], sensitivity [25] and specificity [25].

SQ9 What was the most-used segmentation database? The MoNuSeg dataset, used by 21 studies [3], [4], [8], [12], [18], [21], [27], [34]–[36], [42]–[44], [46], [47], [49], [57], [64], [69]–[71].

SQ10 What are the magnifications of the available images? The majority of the images are available at a magnification of 40×, in 22 databases [73]–[76], [82]–[87], [90]–[98], [100], [101], [103]. There are 13 datasets with images at 20× [75], [77]–[81], [89], [91], [96], [102]–[105] and one with images at 10× [107].

SQ11 Have preprocessing techniques been used with segmentation algorithms? Not every analyzed study applied a preprocessing technique, which opens a new research field for this investigation. In contrast, images used for gland segmentations were preprocessed by histogram equalization [63], U-Net for stain separation [26], [28], a statistical color detection model with a maximum likelihood ratio, dilation and histogram enhancement [53], and color deconvolution [41]. Mitoses used color normalization [15], [17], [19], [60] and CNN [60]. Nuclei segmentation, including malignant and mitotic, was preprocessed by anisotropic diffusion filter with K-SVD and Batch-OMP [1], using the b channel from CIELAB [34], color deconvolution [12], [20] also with smoothing median filter [59], color transformation [45], DL [57], intensity normalization [43], [49], Log-Base2-G Kernel with Gaussian blur filter [72], power law transformation with bilateral filtering [22], spectral [42] and stain normalization [36], [44], [46], [64], and HSV conversion with median filtering for noise removal [31]. Tumor regions only used color normalization [13] and DL [9].

SQ12 Are there databases with evaluations by more than one pathologist? The datasets of MIDOG and UHCMC&CWRU were reviewed by three specialists, while BACH, BreastPathQ, CCMCT-MEL, CryoNuSeg, ICPR12, and LynSeC were annotated by two. KMC-liver and OCDC were annotated by a single expert, and PAIP was reviewed by more than one. According to [18], [42]–[44], the MoNuSeg dataset was annotated by a single pathologist. The TNBC dataset contains annotations by one expert [47] and by four [7], [37], in divergence of information. The same was observed for MITOS&ATYPIA14, with annotations performed by three pathologists in [33], two in [30] and one in [62]. The studies [60] and [27] mention that annotations on TUPAC and NuCLS, respectively, were carried out by more than one pathologist.

IV. Conclusions

Advances in CAD systems are possible due to large volumes of H&E histological images, especially in the current context of DL methodologies. To this end, the use of public images represents an opportunity for the development of new algorithms, exposing them to wide variations in image characteristics, such as colors and magnifications, and different cancer types. This scenario becomes even more relevant when considering segmentation and detection methods, which must be robust to such variations. Thus, this systematic review presents an analysis of public H&E histological image datasets for the evaluation of segmentation and detection techniques applied to different ROIs. It is worth noting that this review does not include an experimental or critical analysis of the computational techniques investigated; however, the collected information is expected to serve as a foundation for the proposal of new image datasets and new techniques for these processing steps.

Through this review, it was notable that the currently available public images have limited magnifications, with no 100×, 200×, or 400×. Considering the specified origins of the samples, we have samples from Asia (∼19%), Europe (∼54%), North America (∼19%), South America (∼4%), and Oceania (∼4%), with clear underrepresentation that can impact the image quality and application of the algorithms. Considering the best quantitative results, DL was observed in almost 85% of scenarios, in addition to superpixels. Despite the diverse tissues, breast lesions (∼54%) had predominance, followed by colorectal (∼14%) cases.

Although these insights are valuable, this review is limited to studies that clearly cited the use of public datasets and publications up to May 2024. In addition to that, it was not possible to present the correlation between segmentation methods and datasets. Even though different studies investigated the same dataset, they did not always use the same images. Therefore, it was not possible to perform a fair comparison from this perspective without the reproduction and implementation of the segmentation techniques. Other limitations include inconsistent information about the datasets, and a rare definition of the number of pathologists involved in the annotations, except for the information available in SQ12. Finally, some public images were segmented but not made available after this process.

ACKNOWLEDGMENTS

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brazil. Process Numbers #2023/18404-2 and #2022/03020-1.

REFERENCES

[1] X. Pan, L. Li, H. Yang, Z. Liu, J. Yang, L. Zhao, and Y. Fan, "Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks," *Neurocomputing*, vol. 229, pp. 88–99, 2017.

[2] S. Vishnoi, A. K. Jain, and P. K. Sharma, "A nuclei segmentation method based on whale optimization algorithm fuzzy clustering in histopathological images," in *2019 4th International Conference on Information Systems and Computer Networks (ISCON)*. IEEE, 2019, pp. 728–732.

[3] M. R. Prusty, R. Dinesh, H. S. Kumar Sheth, A. L. Viswanath, and S. K. Satapathy, "Nuclei segmentation in histopathology images using structure-preserving color normalization based ensemble deep learning frameworks." *Computers, Materials & Continua*, vol. 77, no. 3, 2023.

[4] A. Mahbod, G. Schaefer, B. Bancher, C. Löw, G. Dorffner, R. Ecker, and I. Ellinger, "Cryonuseg: A dataset for nuclei instance segmentation of cryosectioned h&e-stained histological images," *Computers in biology and medicine*, vol. 132, p. 104349, 2021.

[5] M. Dabass, R. Vig, and S. Vashisth, "Review of histopathological image segmentation via current deep learning approaches," in *2018 4th International Conference on Computing Communication and Automation (ICCCA)*. IEEE, 2018, pp. 1–6.

[6] R. Krithiga and P. Geetha, "Breast cancer detection, segmentation and classification on histopathology images analysis: a systematic review," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2607–2619, 2021.

[7] W. R. Drioua, N. Benamrane, and L. Sais, "Breast cancer detection from histopathology images based on yolov5," in *2022 7th International Conference on Frontiers of Signal Processing (ICFSP)*. IEEE, 2022, pp. 30–34.

[8] I. Ahmad, Y. Xia, H. Cui, and Z. U. Islam, "Dan-nucnet: A dual attention based framework for nuclei segmentation in cancer histology images under wild clinical conditions," *Expert Systems with Applications*, vol. 213, p. 118945, 2023.

[9] X. Zhang, X. Zhu, K. Tang, Y. Zhao, Z. Lu, and Q. Feng, "Ddtnet: A dense dual-task network for tumor-infiltrating lymphocyte detection and segmentation in histopathological images of breast cancer," *Medical image analysis*, vol. 78, p. 102415, 2022.

[10] J.-M. Bokhorst, I. D. Nagtegaal, F. Fraggetta, S. Vatrano, W. Mesker, M. Vieth, J. van der Laak, and F. Ciompi, "Deep learning for multi-class semantic segmentation enables colorectal cancer detection and classification in digital pathology images," *Scientific Reports*, vol. 13, no. 1, p. 8398, 2023.

[11] M. Gour, S. Jain, and R. Agrawal, "Deeprnnetseg: deep residual neural network for nuclei segmentation on breast cancer histopathological images," in *International Conference on Computer Vision and Image Processing*. Springer, 2019, pp. 243–253.

[12] I. Kiran, B. Raza, A. Ijaz, and M. A. Khan, "Denseres-unet: Segmentation of overlapped/clustered nuclei from multi organ histopathology images," *Computers in Biology and Medicine*, vol. 143, p. 105267, 2022.

[13] A. Arab, V. Garcia, S. Guan, B. D. Gallas, B. Sahiner, N. Petrick, and W. Chen, "Effect of color-normalization on deep learning segmentation models for tumor-infiltrating lymphocytes scoring using breast cancer histopathology images," in *Medical Imaging 2023: Digital and Computational Pathology*, vol. 12471. SPIE, 2023, pp. 390–394.

[14] M. Dawood, K. Branson, N. M. Rajpoot, and F. Minhas, "Albrt: Cellular composition prediction in routine histology images," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 664–673.

[15] B. Wu, T. Kausar, Q. Xiao, M. Wang, W. Wang, B. Fan, and D. Sun, "Ff-cnn: An efficient deep neural network for mitosis detection in breast cancer histological images," in *Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21*. Springer, 2017, pp. 249–260.

[16] Z. Yildirim, R. Samet, E. Hancer, N. Nemati, and M. T. Mali, "Gland segmentation in h&e histopathological images using u-net with attention module," in *2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, 2023, pp. 1–6.

[17] L. S. Nair, R. Prabhu, G. Sugathan, K. V. Gireesh, and A. S. Nair, "Mitotic nuclei detection in breast histopathology images using yolov4," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2021, pp. 1–5.

[18] J. Li and X. Li, "Miu-net: Mix-attention and inception u-net for histopathology image nuclei segmentation," *Applied Sciences*, vol. 13, no. 8, p. 4842, 2023.

[19] T. Kausar, M. Wang, B. Wu, M. Idrees, and B. Kanwal, "Multi-scale deep neural network for mitosis detection in histological images," in *2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, vol. 3. IEEE, 2018, pp. 47–51.

[20] Y. Zhou, H. Chang, K. E. Barner, and B. Parvin, "Nuclei segmentation via sparsity constrained convolutional regression," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2015, pp. 1284–1287.

[21] S. Lal, D. Das, K. Alabhya, A. Kanfade, A. Kumar, and J. Kini, "Nucleisegnet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images," *Computers in Biology and Medicine*, vol. 128, p. 104075, 2021.

[22] J. K. L. Domoguen, J. J. P. Suarez, and P. C. Naval, "Tag: Nucleus detection in colorectal adenocarcinomas histology images using local texture, appearance, and gradient features," in *2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC)*. IEEE, 2019, pp. 155–159.

[23] L.-W. Liu, Z. Huang, K.-Y. Lu, Z.-X. Wang, Y.-M. Liang, S.-Y. Lin, and Y.-H. Ji, "Ujat-net: A u-net combined joint-attention and transformer for breast tubule segmentation in h&e stained images," *IEEE Access*, 2024.

[24] H. M. AlGhamdiă, N. A. Koohbanani, N. Rajpoot, and S. E. A. Raza, "A novel cell map representation for weakly supervised prediction of er & pr status from h&e wsis," in *MICCAI Workshop on Computational Pathology*. PMLR, 2021, pp. 10–19.

[25] D. F. Dos Santos, P. R. de Faria, B. A. Travençolo, and M. Z. do Nascimento, "Influence of data augmentation strategies on the segmentation of oral histological images using fully convolutional neural networks," *Journal of Digital Imaging*, vol. 36, no. 4, pp. 1608–1623, 2023.

[26] K. Zhang, J. Fu, L. Hua, P. Zhang, Y. Shao, S. Xu, H. Zhou, L. Chen, and J. Wang, "Multiple morphological constraints-based complex gland segmentation in colorectal cancer pathology image analysis," *Complexity*, vol. 2020, no. 1, p. 6180457, 2020.

[27] N. Altini, A. Brunetti, E. Puro, M. G. Taccogna, C. Saponaro, F. A. Zito, S. De Summa, and V. Bevilacqua, "Ndg-cam: Nuclei detection in histopathology images with semantic segmentation networks and grad-cam," *Bioengineering*, vol. 9, no. 9, p. 475, 2022.

[28] J. Fu, K. Zhang, and P. Zhang, "Poorly differentiated colorectal gland segmentation approach based on internal and external stress in histology images," in *2020 5th International Conference on Computer and Communication Systems (ICCCS)*. IEEE, 2020, pp. 338–342.

[29] M. Abdolhoseini, M. G. Kluge, F. R. Walker, and S. J. Johnson, "Segmentation of heavily clustered nuclei from histopathological images," *Scientific reports*, vol. 9, no. 1, p. 4551, 2019.

[30] T. Kausar, M. Wang, M. A. Ashraf, and A. Kausar, "Smallmitosis: Small size mitotic cells detection in breast histopathology images," *IEEE Access*, vol. 9, pp. 905–922, 2020.

[31] H. Amitha, I. Selvamani, and D. A. S. Dhas, "Developement of computer aided system for detection and classification of mitosis using svm," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*. IEEE, 2017, pp. 954–958.

[32] S. Graham, M. Jahanifar, A. Azam, M. Nimir, Y.-W. Tsang, K. Dodd, E. Hero, H. Sahota, A. Tank, K. Benes *et al.*, "Lizard: a large-scale dataset for colonic nuclear instance segmentation and classification," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 684–693.

[33] C. Fernandez-Martín, J. Silva-Rodriguez, U. Kiraz, S. Morales, E. A. Janssen, and V. Naranjo, "Uninformed teacher-student for hard-samples distillation in weakly supervised mitosis localization," *Computerized Medical Imaging and Graphics*, vol. 112, p. 102328, 2024.

[34] O. C. Linares, A. A. Soriano-Vargas, B. S. Faiçal, B. Hamann, A. T. Fabro, and A. J. Traina, "Efficient segmentation of cell nuclei in histopathological images," in *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2020, pp. 47–52.

[35] A. K. Chanchal, S. Lal, and J. Kini, "Deep structured residual encoder-decoder network with a novel loss function for nuclei segmentation of kidney and breast histopathology images," *Multimedia Tools and Applications*, vol. 81, no. 7, pp. 9201–9224, 2022.

[36] R. KC Khatri, B. J Caseria, Y. Lou, G. Xiao, and Y. Cao, "Automatic extraction of cell nuclei using dilated convolutional network," *Inverse Problems & Imaging*, vol. 15, no. 1, 2021.

[37] L. Putzu and G. Fumera, "An empirical evaluation of nuclei segmentation from h&e images in a real application scenario," *Applied Sciences*, vol. 10, no. 22, p. 7982, 2020.

[38] P. He, A. Qu, S. Xiao, and M. Ding, "Detisseg: A dual-encoder network for tissue semantic segmentation of histopathology image," *Biomedical Signal Processing and Control*, vol. 87, p. 105544, 2024.

[39] Y. Chen, Y. Zhou, G. Chen, Y. Guo, Y. Lv, M. Ma, Z. Pei, and Z. Sun, "Segmentation of breast tubules in h&e images based on a dks-doubleu-net model," *BioMed Research International*, vol. 2022, no. 1, p. 2961610, 2022.

[40] Y. Xie, F. Xing, X. Shi, X. Kong, H. Su, and L. Yang, "Efficient and robust cell detection: A structured regression approach," *Medical image analysis*, vol. 44, pp. 245–254, 2018.

[41] Y.-R. Van Eycke, C. Balsat, L. Verset, O. Debeir, I. Salmon, and C. Decaestecker, "Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise ihc biomarker quantification: A deep learning approach," *Medical image analysis*, vol. 49, pp. 35–45, 2018.

[42] F. Mahmood, D. Borders, R. J. Chen, G. N. McKay, K. J. Salimian, A. Baras, and N. J. Durr, "Deep adversarial training for multi-organ nuclei segmentation in histopathology images," *IEEE transactions on medical imaging*, vol. 39, no. 11, pp. 3257–3267, 2019.

[43] K. Roy, D. Banik, G. K. Chan, O. Krejcar, and D. Bhattacharjee, "2pclpr: A two-phase clump profiler for segmentation of cancer cells in fluorescence microscopic images," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–14, 2023.

[44] E. Hancer, M. Traoré, R. Samet, Z. Yıldırım, and N. Nemati, "An imbalance-aware nuclei segmentation methodology for h&e stained histopathology images," *Biomedical Signal Processing and Control*, vol. 83, p. 104720, 2023.

[45] H. Chang, L. A. Loss, P. T. Spellman, A. Borowsky, and B. Parvin, "Batch-invariant nuclear segmentation in whole mount histology sections," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2012, pp. 856–859.

[46] Y. Lin, Z. Qu, H. Chen, Z. Gao, Y. Li, L. Xia, K. Ma, Y. Zheng, and K.-T. Cheng, "Nuclei segmentation with point annotations from pathology images via self-supervised learning and co-training," *Medical Image Analysis*, vol. 89, p. 102933, 2023.

[47] S. Kasturi, W. T. Tran, and A. Shenfield, "Accurate nuclei segmentation in breast cancer tumour biopsies," in *2022 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, 2022, pp. 1–8.

[48] H. Mittal and M. Saraswat, "An automatic nuclei segmentation method using intelligent gravitational search algorithm based superpixel clustering," *Swarm and Evolutionary Computation*, vol. 45, pp. 15–32, 2019.

[49] A. Mahbod, G. Schaefer, G. Dorffner, S. Hatamikia, R. Ecker, and I. Ellinger, "A dual decoder u-net-based model for nuclei instance segmentation in hematoxylin and eosin-stained histological images," *Frontiers in Medicine*, vol. 9, p. 978146, 2022.

[50] A. Kumar and M. Prateek, "Localization of nuclei in breast cancer using whole slide imaging system supported by morphological features and shape formulas," *Cancer management and research*, pp. 4573–4583, 2020.

[51] X. Wang, Y. Fang, S. Yang, D. Zhu, M. Wang, J. Zhang, K.-y. Tong, and X. Han, "A hybrid network for automatic hepatocellular carcinoma segmentation in h&e-stained whole slide images," *Medical Image Analysis*, vol. 68, p. 101914, 2021.

[52] R. Sharma and K. Sharma, "An optimal nuclei segmentation method based on enhanced multi-objective gwo," *Complex & Intelligent Systems*, pp. 1–14, 2022.

[53] J. Shu, J. Lei, Q. Gao, and Q. Zhang, "Combing colour detection and neural networks for gland detection," in *Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition*, 2019, pp. 33–36.

[54] P. Rastogi, K. Khanna, and V. Singh, "Gland segmentation in colorectal cancer histopathological images using u-net inspired convolutional network," *Neural Computing and Applications*, vol. 34, no. 7, pp. 5383–5395, 2022.

[55] S. Doyle, F. Dal Canton, J. Wesseling, C. I. Sánchez, and J. Teuwen, "Mammary duct detection using self-supervised encoders," in *Medical Imaging 2022: Computer-Aided Diagnosis*, vol. 12033. SPIE, 2022, pp. 229–232.

[56] H. Naji, L. Sancere, A. Simon, R. Büttner, M.-L. Eich, P. Lohneis, and K. Bożek, "Holy-net: Segmentation of histological images of diffuse large b-cell lymphoma," *Computers in Biology and Medicine*, vol. 170, p. 107978, 2024.

[57] H. Zhang, J. Liu, P. Wang, Z. Yu, W. Liu, and H. Chen, "Cross-boosted multi-target domain adaptation for multi-modality histopathology image translation and segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3197–3208, 2022.

[58] D. R. Chambers, B. B. Brimhall, D. R. Poole Jr, and E. A. Medina, "Cancer cell segmentation for cellularity prediction via a weakly labeled/strongly labeled hybrid convolutional neural network," in *Medical Imaging 2022: Digital and Computational Pathology*, vol. 12039. SPIE, 2022, pp. 47–54.

[59] A. Husham, M. Hazim Alkawaz, T. Saba, A. Rehman, and J. Saleh Alghamdi, "Automated nuclei segmentation of malignant using level sets," *Microscopy research and technique*, vol. 79, no. 10, pp. 993–997, 2016.

[60] R. Nateghi, H. Danyali, and M. S. Helfroush, "A deep learning approach for mitosis detection: application in tumor proliferation prediction from whole slide images," *Artificial intelligence in medicine*, vol. 114, p. 102048, 2021.

[61] K. S. Beevi, M. S. Nair, and G. Bindu, "Automatic mitosis detection in breast histopathology images using convolutional neural network based deep transfer learning," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 214–223, 2019.

[62] G. I. Sayed and A. E. Hassanien, "Moth-flame swarm optimization with neutrosophic sets for automatic mitosis detection in breast cancer histology images," *Applied Intelligence*, vol. 47, pp. 397–408, 2017.

[63] S. F. H. Naqvi, S. Ayubi, A. Nasim, and Z. Zafar, "Automated gland segmentation leading to cancer detection for colorectal biopsy images," in *Advances in Information and Communication: Proceedings of the 2019 Future of Information and Communication Conference (FICC), Volume 2*. Springer, 2020, pp. 75–83.

[64] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, Ö. Smedby, and C. Wang, "A two-stage u-net algorithm for segmentation of nuclei in h&e-stained tissues," in *Digital Pathology: 15th European Congress, ECDP 2019, Warwick, UK, April 10–13, 2019, Proceedings 15*. Springer, 2019, pp. 75–82.

[65] H. Ahn, Y. Hong, and K.-M. Kim, "Class-controlling copy-paste augmentation for nuclear segmentation," in *2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)*. IEEE, 2022, pp. 1–4.

[66] A. A. Aatresh, R. P. Yatgiri, A. K. Chanchal, A. Kumar, A. Ravi, D. Das, R. Bs, S. Lal, and J. Kini, "Efficient deep learning architecture with dimension-wise pyramid pooling for nuclei segmentation of histopathology images," *Computerized Medical Imaging and Graphics*, vol. 93, p. 101975, 2021.

[67] Z. Liu, Y. Cai, and Q. Tang, "Nuclei detection in breast histopathology images with iterative correction," *Medical & Biological Engineering & Computing*, vol. 62, no. 2, pp. 465–478, 2024.

[68] H. Wang, J. Yang, R. Katayama, M. Matsusaki, T. Miyao, and J. Zhou, "Nuclseg: nuclei segmentation using semi-supervised stain deconvolution," in *Proceedings of the 5th ACM International Conference on Multimedia in Asia*, 2023, pp. 1–6.

[69] W. Lou, H. Li, G. Li, X. Han, and X. Wan, "Which pixel to annotate: a label-efficient nuclei segmentation framework," *IEEE Transactions on Medical Imaging*, vol. 42, no. 4, pp. 947–958, 2022.

[70] G. M. Dogar, M. M. Fraz, and S. Javed, "Feature attention network for simultaneous nuclei instance segmentation and classification in histology images," in *2021 International conference on digital futures and transformative technologies (ICoDT2)*. IEEE, 2021, pp. 1–6.

[71] K. Yao, K. Huang, J. Sun, and A. Hussain, "Pointnu-net: Keypoint-assisted convolutional neural network for simultaneous multi-tissue histology nuclei segmentation and classification," *IEEE Transactions*

on Emerging Topics in Computational Intelligence, vol. 8, no. 1, pp. 802–813, 2023.
[72] S. Roy, R. Khurana, and V. Jain, "Log-base2 of gaussian kernel for nuclei segmentation from colorectal cancer h and e-stained histopathology images," in *Fifteenth International Conference on Machine Vision (ICMV 2022)*, vol. 12701. SPIE, 2023, pp. 81–88.
[73] "Kmc liver dataset," https://drive.google.com/drive/u/0/folders/1_lLVLKIqkpQa2YBC_76RUOXLLtdDIBoE, 2021, [Online; accessed 03-June-2025].
[74] "Cryonuseg dataset," https://github.com/masih4/CryoNuSeg, 2021, [Online; accessed 30-May-2025].
[75] "Mthh dataset," https://pan.baidu.com/s/1fgaR13lj6BvhlbSV2AzNbQ#list/path=%2F, 2022, [Online; accessed 01-June-2025].
[76] "Lynsec dataset," https://zenodo.org/records/8065174, 2024, [Online; accessed 02-June-2025].
[77] "Tcga-brca dataset," https://www.cancerimagingarchive.net/collection/tcga-brca/, [Online; accessed 10-June-2025].
[78] "Miccai glas challenge database," https://www.kaggle.com/datasets/sani84/glasmiccai2015-gland-segmentation, 2022, [Online; accessed 01-June-2025].
[79] "Paip dataset," https://www.kaggle.com/datasets/atishekkumar/paip-2019?resource=download, 2019, [Online; accessed 14-June-2025].
[80] "Uhcmc&cwru dataset," https://datadryad.org/dataset/doi:10.5061/dryad.1g2nt41, 2018, [Online; accessed 01-June-2025].
[81] "Breastpathq dataset," https://breastpathq.grand-challenge.org/Data/, 2019, [Online; accessed 01-June-2025].
[82] "Amida 13 dataset," https://www.kaggle.com/datasets/medkad19/amida13-augmented/data, [Online; accessed 15-June-2025].
[83] "Icpr12 dataset," http://ludo17.free.fr/mitos_2012/download.html, [Online; accessed 15-June-2025].
[84] "Mitos & atypia 14 database," https://mitos-atypia-14.grand-challenge.org/Donwload/, 2014, [Online; accessed 31-May-2025].
[85] "Tupac database," https://tupac.grand-challenge.org/Dataset/, 2019, [Online; accessed 30-May-2025].
[86] "Mitos_wsi_ccmct dataset," https://github.com/DeepMicroscopy/MITOS_WSI_CCMCT, [Online; accessed 09-June-2025].
[87] "Midog21 dataset," https://imig.science/midog2021/download-dataset/, 2021, [Online; accessed 09-June-2025].
[88] "Bach iciar," https://www.kaggle.com/datasets/truthisneverlinear/bach-breast-cancer-histology-images, 2019, [Online; accessed 15-June-2025].
[89] "Conic challenge dataset," https://www.kaggle.com/datasets/aadimator/conic-challenge-dataset, 2022, [Online; accessed 01-June-2025].
[90] "Consep dataset," https://www.kaggle.com/datasets/rftexas/tiled-consep-224x224px, [Online; accessed 15-June-2025].
[91] "Cpm-17 dataset," https://drive.google.com/drive/folders/1sJ4nmkif6j4s2FOGj8j6i_Ye7z9w0TfA, 2019, [Online; accessed 01-June-2025].
[92] "Monusac dataset," https://monusac-2020.grand-challenge.org/Data/, 2020, [Online; accessed 30-May-2025].
[93] "Monuseg dataset," https://monuseg.grand-challenge.org/Data/, 2017, [Online; accessed 30-May-2025].
[94] "Nucls dataset," https://sites.google.com/view/nucls/, [Online; accessed 02-June-2025].
[95] "Nulnsseg dataset," https://www.kaggle.com/datasets/ipateam/nuinsseg, 2022, [Online; accessed 30-May-2025].
[96] "Pannuke dataset," https://warwick.ac.uk/fac/cross_fac/tia/data/pannuke, [Online; accessed 01-June-2025].
[97] "Tcga-kirc," https://www.cancerimagingarchive.net/collection/tcga-kirc/, [Online; accessed 01-June-2025].
[98] "Tnbc dataset," https://zenodo.org/records/1175282, 2018, [Online; accessed 30-May-2025].
[99] "Ucsb database," https://www.kaggle.com/datasets/andrewmvd/breast-cancer-cell-segmentation?resource=download, 2020, [Online; accessed 01-June-2025].
[100] "Benchmark dataset," https://andrewjanowczyk.com/use-case-1-nuclei-segmentation/, [Online; accessed 01-June-2025].
[101] "Local crc dataset," https://zenodo.org/records/4785131, [Online; accessed 15-June-2025].
[102] "Ocdc dataset," https://data.mendeley.com/datasets/9bsc36jyrt/1, 2022, [Online; accessed 03-June-2025].
[103] "Orca dataset," https://sites.google.com/unibas.it/orca, 2020, [Online; accessed 03-June-2025].
[104] "Bca-lym dataset," https://andrewjanowczyk.com/use-case-4-lymphocyte-detection/, [Online; accessed 01-June-2025].
[105] "Post-nat-brca dataset," https://www.cancerimagingarchive.net/collection/post-nat-brca/, [Online; accessed 01-June-2025].
[106] "Bcss dataset," https://drive.google.com/drive/folders/1zqbdkQF8i5cEmZOGmbdQm-EP8dRYtvss, [Online; accessed 01-June-2025].
[107] "Luad-histoseg dataset," https://drive.google.com/drive/folders/1E3Yei3Or3xJXukHIybZAgochxfn6FJpr, [Online; accessed 01-June-2025].
[108] D. Müller, I. Soto-Rey, and F. Kramer, "Towards a guideline for evaluation metrics in medical image segmentation," *BMC Research Notes*, vol. 15, no. 1, p. 210, 2022.