Immunohistochemistry Biomarkers-Guided Image Search for Histopathology

Abubakr Shafique[1,2], Morteza Babaie[1,3], Ricardo Gonzalez[4], and H.R. Tizhoosh[1,2,3]

Abstract— **Medical practitioners use a number of diagnostic tests to make a reliable diagnosis. Traditionally, Haematoxylin and Eosin (H&E) stained glass slides have been used for cancer diagnosis and tumor detection. However, recently a variety of immunohistochemistry (IHC) stained slides can be requested by pathologists to examine and confirm diagnoses for determining the subtype of a tumor when this is difficult using H&E slides only. Deep learning (DL) has received a lot of interest recently for image search engines to extract features from tissue regions, which may or may not be the target region for diagnosis. This approach generally fails to capture high-level patterns corresponding to the malignant or abnormal content of histopathology images. In this work, we are proposing a targeted image search approach, inspired by the pathologists' workflow, which may use information from multiple IHC biomarker images when available. These IHC images could be aligned, filtered, and merged together to generate a *composite biomarker image* (CBI) that could eventually be used to generate an attention map to guide the search engine for localized search. In our experiments, we observed that an IHC-guided image search engine can retrieve relevant data more accurately than a conventional (i.e., H&E-only) search engine without IHC guidance. Moreover, such engines are also able to accurately conclude the subtypes through majority votes.**

Index Terms— **Image Retrieval, Image Search, Histopathology, Immunohistochemistry, Biomarkers, Digital Pathology, Fuzzy Inference**

I. INTRODUCTION

The emergence of digital pathology (DP) has opened new horizons for histopathology [1]. Machine Learning (ML) algorithms are able to operate on digitized slides to assist pathologists with different tasks. Whereas ML-involving classification and segmentation methods have obvious benefits for image analysis, image search represents a fundamental shift in computational pathology [2]. Matching the pathology of new patients with already diagnosed and curated cases offers pathologists a new approach to improve diagnostic accuracy through visual inspection of similar cases and a computational majority vote for consensus building [2].

Pathologists examine tissue slides under a microscope on a regular basis and write diagnostic and prognostic reports based on their visual inspections. The use of immunologic research methodologies in histopathology has resulted in a significant improvement in neoplasm microscopic diagnosis [3].

Immunohistochemistry (IHC) has become a formidable tool at the pathologist's disposal, despite the fact that histological analysis of hematoxylin and eosin (H&E) stained tissue sections remain at the core of the discipline [3]–[5]. IHC is a technique for detecting specific antigens (proteins) in tissue slices using labeled antibodies that bind with the antigen [6]. The purpose of staining is to draw attention to the area of interest while also providing contrast against the 'background'. Diagnostic IHC tests have been commonly used to investigate and confirm diagnoses by determining a tumor's differentiation route [3], [7]. As oncologists demand more specific diagnoses, IHC has proven to be one of the most essential auxiliary techniques in the characterization of neoplastic disorders in humans [8], [9]. Simultaneously, archives of digital scans in pathology are progressively becoming a reality as workload increases [2].

Since the last decade, content-based image retrieval (CBIR) has been one of the intensively researched fields in computer vision [1]. Recently, the use of deep learning (DL) for image search has captured significant attention [10]. CBIR systems allow the user to search for images that are morphologically and semantically similar to one another. CBIR has numerous real-world applications, and it may be especially valuable for medical images, as linguistic features collected from medical reports may be frequently insufficient to adequately represent the content of the related cases [11], [12]. The mass medical image archives have traditionally been bundled with textual annotations classified by professionals; however, this technique does not scale well with the ever-increasing demands of DP. While CBIR systems for histopathology have received a lot of attention [13], image search and analysis for histopathology has just recently become a focus of research, due to the rise of DP and DL [1], [2], [14], [15]. Pathologists can access evidence-based knowledge from prior cases through CBIR systems designed specifically for histopathology, allowing them to make decisions more quickly and reliably [2]. Histopathology whole slide images (WSI) are gigapixel files, which generally cannot be processed as a whole and hence are often broken down into numerous "patches" [16]. However, the majority of CBIR solutions employ fundamental image features from randomly selected or softly categorized regions, which may or may not be the target region, i.e, the abnormal tissue. This approach generally fails to capture high-level patterns corresponding to the malignant content of histopathology images. The motivation for developing targeted search engines emerged from multiple observations: 1) Not many works have previously incorporated IHC biomarkers in their search engines, 2) Only H&E-stained WSIs were previously used

[1]Abubakr Shafique, Morteza Babaie, and H.R. Tizhoosh are with Kimia Lab, University of Waterloo, Waterloo, ON, Canada (email: abubakr.shafique@uwaterloo.ca)

[2]Abubakr Shafique and H.R. Tizhoosh are also affiliated with Rhazes Lab, Department of Artificial Intelligence & Informatics, Mayo Clinic, Rochester, MN, USA.

[3]Hamid Tizhoosh and Morteza Babaie are also affiliated with Vector Institute, MaRS Centre, Toronto, Canada.

[4]Ricardo Gonzalez is affiliated with Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA.
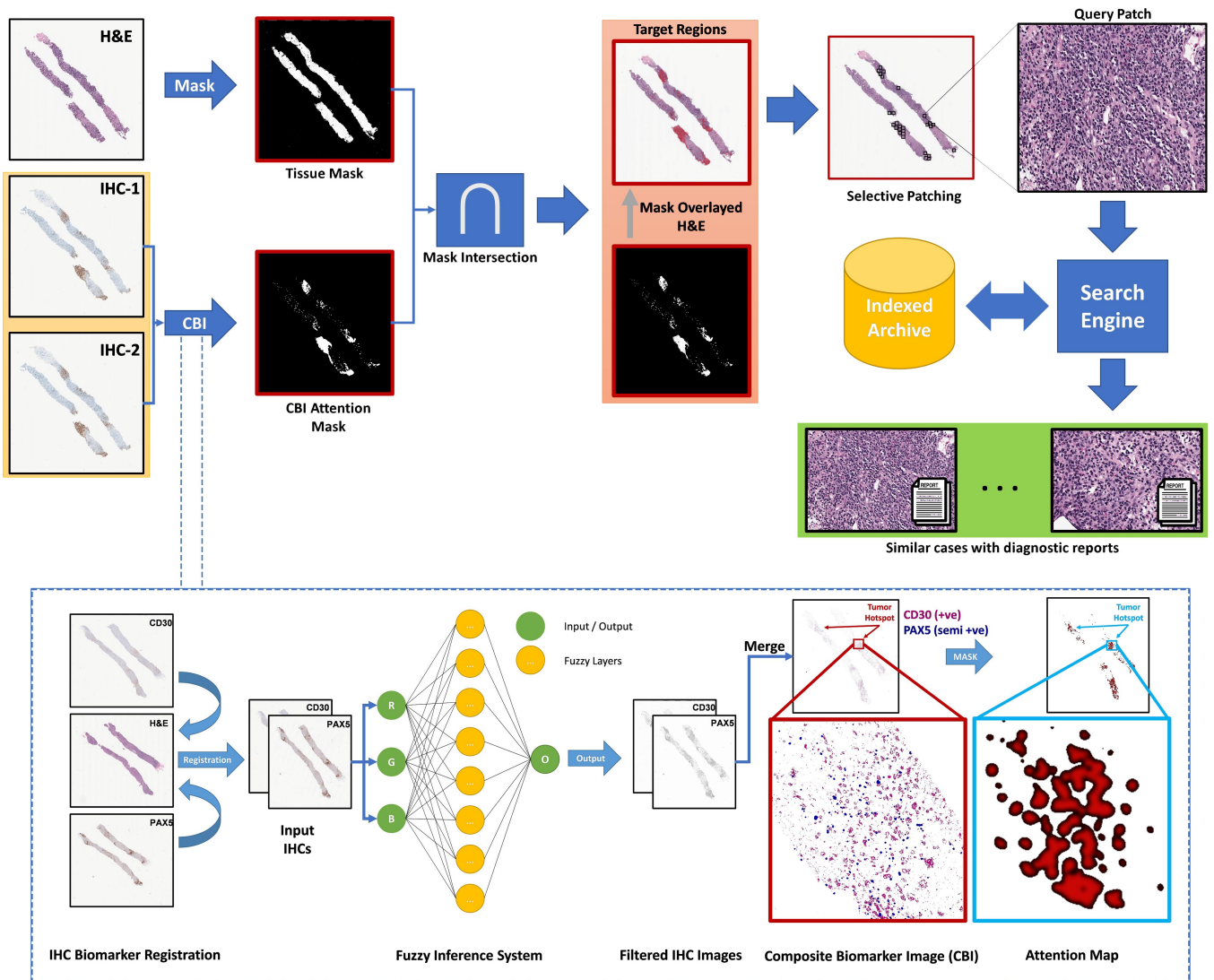
Fig. 1. The overall framework for targeted image search. Multiple IHC biomarkers (IHC-1 and IHC-2) are used to generate a CBI-based attention mask, which is used for selective patching. These patches with significant patterns are then indexed by the search engine to accommodate future queries.

with random or soft patch selection which may or may not include the embedding from the key suspected regions.

In this work, we propose a targeted image search approach based on the information from the IHC biomarker images. This approach is inspired by pathologists as they use IHC biomarkers for detailed cross-stain analysis, diagnosis, and narrowing down visual inspection to the significant regions with malignant cells. Multiple IHC biomarker images could be aligned, filtered, and merged together to generate a "composite biomarker image" (CBI) to generate an attention map to guide the search engine for better search accuracy.

This paper is organized as follows: Section II describes the algorithms behind the targeted image search and data used in the experiments. The experimental setup and results are described in Section III, which are further discussed and concluded in Section IV.

II. MATERIALS & METHODS

Targeted image search, utilizing IHC attention maps, for the histopathology could be a significant upgrade for existing search engines. To make full use of IHC images, we proposed to use CBIs to generate a malignant attention map by combining multiple IHC biomarker images into one (See Fig. 1). In ideal scenarios, having multiple medical professionals annotating tumorous regions in many images is quite desirable. But manual delineation is a time-consuming task and subject to user variability.

IHC Biomarkers & Alignment – Image alignment is the first step that is crucial for the whole procedure as the morphology and orientation of each cut on different glass slides are often different when looking at IHC and H&E images. For this reason, we have carefully designed our fully automated non-rigid image registration procedure, which accurately aligns all the biomarkers [17]. In the registration
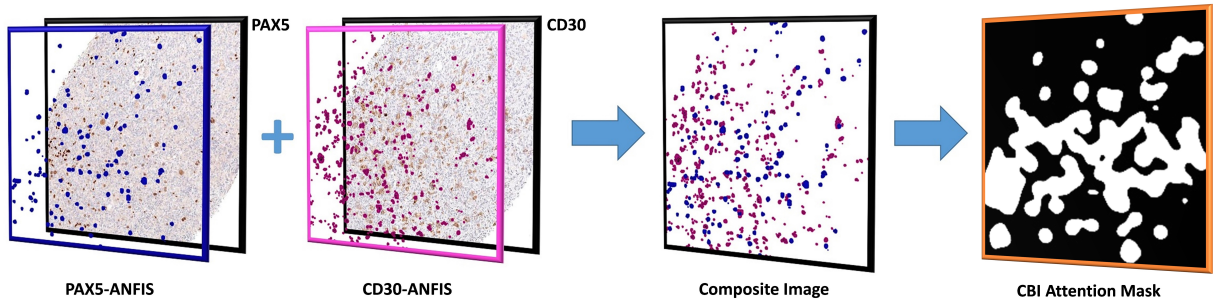
Fig. 2. CBI sample produced by the two IHCs (CD30 positive cell population, and PAX5 semi-positive cell population in the same area).

process, CD30, and PAX5 biomarker WSIs are aligned to the H&E-stained WSI. IHC biomarkers CD30 and PAX5 were selected as these are critical for Hodgkin lymphoma (HL) [18]. CD30 is a protein that can be detected in T and B lymphocytes (two types of white blood cells). It is a receptor for tumor necrosis factor, a protein that plays a role in cell development and survival [19]. On the other hand, paired box protein, PAX5, is a protein that in humans is encoded by the PAX5 gene. PAX5 is a transcription factor that is required for lymphoid progenitors to commit to the B lymphocyte lineage [20]. PAX5 has a dual function, inhibiting 'inappropriate' B lineage genes while simultaneously activating B lineage–specific genes [20]. In the case of classical Hodgkin lymphoma (CHL), it could be identified when there is the presence of CD30 (+) Hodgkin and Reed-Sternberg cells with light expression of PAX5 (B-lymphocytes) [18].

Adaptive Neuro-Fuzzy Inference System (ANFIS) – After aligning the biomarker images, ANFIS is used to classify and extract the positive or negative information from the three biomarker images. ANFIS is the combination of neural networks (NN) and FIS which uses NNs learning algorithms to self tune the rules, parameters, and structure of FIS [21], [22]. For simplicity, we assume a network with three inputs, r, g, and b, and one output, o [22], [23]. To design the ANFIS architecture, three *fuzzy if-then rules* based on a first-order Sugeno model are shown below:

Rule 1: *if r is A_1 and g is B_1 and b is C_1, then* $o_1 = j_1 r + k_1 g + l_1 b + z_1$
Rule 2: *if r is A_2 and g is B_2 and b is C_2, then o_2 = $j_2 r + k_2 g + l_2 b + z_2$*
Rule 3: *if r is A_3 and g is B_3 and b is C_3, then o_3 = $j_3 r + k_3 g + l_3 b + z_3$*

where r, g, and b are the inputs, A_i, B_i, and C_i are the fuzzy sets, o is the output of the fuzzy system, and j_i, k_i, l_i, and z_i are the design parameters that are calculated during the training process.

ANFIS is used to filter the IHC-stained WSIs at pixel level based on thee RGB color, thus taking three inputs as R, G, and B and giving one output o in the range of 0 to 255. Here, biomarker specific ANFIS models are trained, for instance with CD30 only positive regions (Class 3, 4,

and 5 from table I), and with PAX5 semi-positive tissues (Class 4 only from table I), a relevant IHC indication for Hodgkin lymphoma diagnosis. Hence, only IHC-relevant tissue regions will get some grey level and remaining image area will be regarded as irrelevant background. For the ANFIS training and testing, 30 RGB values were randomly selected by the pathologist from each class, so in total 180 x 3 (RGB) points were used in our experiment. From 30 points, 25 RGB points were used for the training, and 5 points were used for testing for each class. The range of RGB data for each class for training is shown in Table. I.

TABLE I
R, G, AND B INPUT AND OUTPUT DATA RANGES FOR ANFIS TRAINING.

Input Range			Class
R	G	B	
214 - 247	214 - 247	213 - 247	0 = Background
24 - 208	44 - 217	79 - 228	1 = Blue
63 - 221	65 - 221	77 - 226	2 = Gray
163 - 251	124 - 218	107 - 214	3 = Light Brown
136 - 192	87 - 157	70 - 147	4 = Medium Brown
37 - 98	0 - 67	0 - 61	5 = Dark Brown

Composite Image & Attention Mask – Each biomarker's grey level image with the relevant tissues are obtained via ANFIS. These images are then processed and combined to create a composite image (see Fig. 1). The morphological operations erosion and dilation are used to process all biomarker grayscale images before they are organized and superimposed. In our experiment, we arranged CD30 as the first image, and then PAX5 on top of CD30. A sample CBI image with the two biomarkers is shown in Fig. 2. The overlay color of the biomarker images can be chosen by the user, and their overlay order in the CBI image can also be adjusted according to the needs.

Although, annotation of the tumorous regions by the pathologists for training the neural network is required for supervised learning, but it is a cumbersome task subject to observer variability. CBI masks are automated tumor annotation based on the IHC biomarker images (See Fig. 2), which are supposed to be similar to a pathologists' annotation when they delineate H&E images based on their visual clues from IHC slides. The population of targeted cells are focused in IHC biomarker image(s) to get tumorous hotspots (see Fig. 1 and Fig. 2). Therefore, the network learns significant
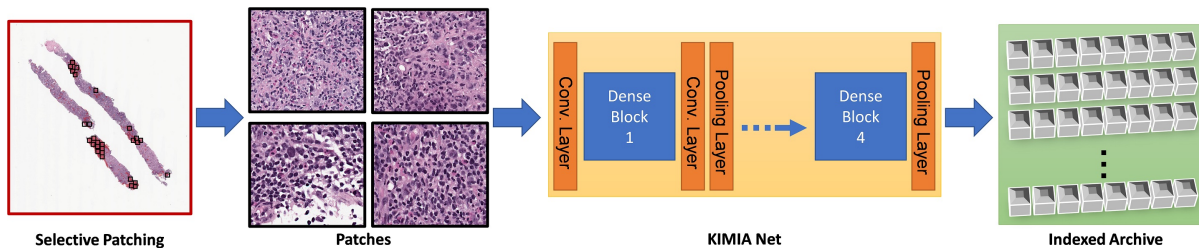
Fig. 3. The process of feature extraction and indexing using from the chosen patches using KimiaNet [24].

features/patterns more efficiently while training by relaying on targeted tumorous regions.

Targeted Image Search – Patches can be matched and retrieved after the target regions have been obtained from the CBI attention mask (See Fig. 1). All selected patches from each WSI are now pushed through a pre-trained network namely, KimiaNet (DenseNet-121 architecture), which had been specifically trained and optimized on more than 240,000 histopathological images to distinguish 30 cancer subtypes including normal tissue, benign tumors, and various types of malignant tumors [24]. The last pooling layers or the first connected layers are typically utilized as "features" to represent each patch, with the actual output of the network being ignored. In this study, we turn several patch embeddings into a single feature vector of length 1,024 per WSI as the output of the KimiaNet using the median of the minimums approach from Yottixel. (see Fig. 3) [2], [24].

Data – For the experiments, 15 cases of CHL and 8 cases of nodular lymphocyte predominant Hodgkin lymphoma (NLPHL) were employed, which included H&E and two IHC biomarkers (CD30 and PAX5). A total of 23 cases, including 69 WSIs and 9,296 patches (size 300×300 at $20\times$ magnification) from H&E-stain slides were used in this experiment. The WSIs were obtained from the Grand River Hospital (Kitchener, Ontario, Canada). Specimens were sliced into thin layers and placed onto a paraffin cassette one by one to prepare the glass slides. Furthermore, utilizing a digital scanner, the prepared glass slides were digitized into WSI data at $40\times$ magnification.

III. RESULTS

In our experiments, the performance of the targeted search (using a CBI attention mask) is compared with the conventional search (without a CBI attention mask) in histopathology. We performed the experiments on a Dell EdgeServer Ra with 2x Intel(R) Xeon(R) Gold 5118 (12 cores, 2.30GHz), 1x Tesla V100 (v-RAM 32 GB), and 394 GB RAM. The accuracy of image search is evaluated through "leave-one-patient-out" validation as the most rigorous technique for small to medium-size datasets [1]. In this experiment, the majority-n accuracy was determined; only if the majority among n search results were correct, the search considered valid (this is in contrast to the top-n approach in computer vision that regards the attempt successful if *any* of n search results is correct). Precisely, "correct" denotes that the majority of detected and retrieved instances correctly identified and
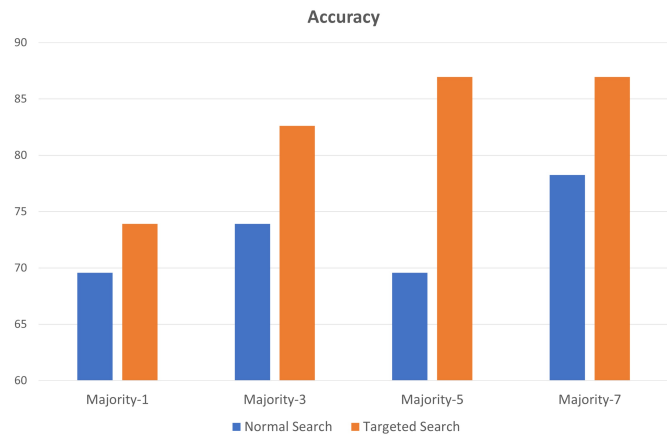

Fig. 4. The accuracy of the proposed targeted search (with selective patch selection based on CBI attention mask) and normal search (with random patch selection) according to the majority-1 (top-1), majority-3, majority-5, and majority-7 validation approach.

matched the tumor type or tumor subtype within a particular diagnostic category. We performed the experiments using H&E and IHC WSIs from two different types of Hodgkin lymphoma (CHL and NLPHL), retrieving the top one, three, five, and seven WSI patches that match the query. Fig. 4 shows the accuracy and comparison of the proposed targeted search (with CBI attention mask) and normal search (without CBI attention mask) of the top-n results.

For performance evaluation, when searching through the archive to discover the best matches, we employed the k-nearest neighbors (k-NN) approach to find the top-matched (most similar) embeddings. The best matches, or those closest to the search parameters (based on Euclidean distance), can then be retrieved. for all of the search and matching distance metrics calculations. As a result, the feature with the minimum Euclidean distances among the features from two cases (CHL and NLPHL) is the best match. The standard evaluation metrics for verifying the performance of image search and CBIR algorithms are precision, recall, and F1-score when treated as a classifier (see Table II) [2], [24].

IV. CONCLUSIONS

In this paper, we proposed an IHC biomarker-guided image search that identifies the key areas for unlabelled H&E WSIs based on the IHC staining patterns. The proposed method can automatically extract the patches from non-

TABLE II

Precision, recall, and F1-score for top-1, majority-3, majority-5, and majority-7 search results for the proposed targeted search compared with the normal search.

Search Type	Tumor	Top Selection	Precision	Recall	F1-Score
Target Search	CHL	top-1	**84.62**	73.33	**78.57**
		majority-3	**86.67**	**86.67**	**86.67**
		majority-5	**100.0**	80.00	**88.89**
		majority-7	**100.0**	80.00	**88.89**
	NLPHL	top-1	**60.00**	**75.00**	**66.67**
		majority-3	**75.00**	75.00	**75.00**
		majority-5	**72.73**	**100.0**	**84.21**
		majority-7	**72.73**	**100.0**	**84.21**
Normal Search	CHL	top-1	75.00	80.00	77.42
		majority-3	84.62	73.33	78.57
		majority-5	83.33	66.67	74.07
		majority-7	91.67	73.33	81.48
	NLPHL	top-1	57.14	50.00	53.33
		majority-3	60.00	75.00	66.67
		majority-5	54.55	75.00	63.16
		majority-7	63.64	87.50	73.68

annotated H&E WSIs guided by its corresponding IHC biomarkers WSIs. Patches from these selected regions contain abnormal/malignant tissue, from which the network can extract relevant feature vectors. Consequently, the proposed method improves the accuracy of the search engine compared to the conventional search engines, where they select the patches from WSIs either randomly or based on manual delineations. The results from our experiment confirmed that the proposed IHC-guided search has improved the matching accuracy significantly as we can see in Fig. 4 and Table. II.

Future research should focus on the solution for the potentially low availability of IHC biomarker-stained WSIs. Despite the advantages of IHC in identifying the types and subtypes of tumors, IHCs may not be readily available as compared to H&E slides. Requesting IHC may also result in delays in the diagnostic process.

References

[1] S. Kalra, H. R. Tizhoosh, S. Shah, C. Choi, S. Damaskinos, A. Safarpoor, S. Shafiei, M. Babaie, P. Diamandis, C. J. Campbell et al., "Pancancer diagnostic consensus through searching archival histopathology images using artificial intelligence," NPJ digital medicine, vol. 3, no. 1, pp. 1–15, 2020.

[2] S. Kalra, H. R. Tizhoosh, C. Choi, S. Shah, P. Diamandis, C. J. Campbell, and L. Pantanowitz, "Yottixel–an image search engine for large archives of histopathology whole slide images," Medical Image Analysis, vol. 65, p. 101757, 2020.

[3] A. Kabiraj, J. Gupta, T. Khaitan, and P. T. Bhattacharya, "Principle and techniques of immunohistochemistry—a review," Int J Biol Med Res, vol. 6, no. 3, pp. 5204–5210, 2015.

[4] R. C. Jordan, T. E. Daniels, J. S. Greenspan, and J. A. Regezi, "Advanced diagnostic methods in oral and maxillofacial pathology. part ii: Immunohistochemical and immunofluorescent methods," Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, vol. 93, no. 1, pp. 56–74, 2002.

[5] S.-W. Kim, J. Roh, and C.-S. Park, "Immunohistochemistry for pathologists: protocols, pitfalls, and tips," Journal of pathology and translational medicine, vol. 50, no. 6, p. 411, 2016.

[6] R. Walker, "Quantification of immunohistochemistry—issues concerning methods, utility and semiquantitative assessment i," Histopathology, vol. 49, no. 4, pp. 406–410, 2006.

[7] A. Gown, "Genogenic immunohistochemistry: a new era in diagnostic immunohistochemistry," Current Diagnostic Pathology, vol. 8, no. 3, pp. 193–200, 2002.

[8] J. A. Ramos-Vara, M. Kiupel, T. Baszler, L. Bliven, B. Brodersen, B. Chelack, K. West, S. Czub, F. Del Piero, S. Dial et al., "Suggested guidelines for immunohistochemical techniques in veterinary diagnostic laboratories," Journal of Veterinary Diagnostic Investigation, vol. 20, no. 4, pp. 393–413, 2008.

[9] J. Duraiyan, R. Govindarajan, K. Kaliyappan, and M. Palanisamy, "Applications of immunohistochemistry," Journal of pharmacy & bioallied sciences, vol. 4, no. Suppl 2, p. S307, 2012.

[10] B. Kieffer, M. Babaie, S. Kalra, and H. R. Tizhoosh, "Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks," in 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, 2017, pp. 1–6.

[11] S. Doyle, M. Hwang, S. Naik, M. Feldman, J. Tomaszeweski, and A. Madabhushi, "Using manifold learning for content-based image retrieval of prostate histopathology," in MICCAI 2007 Workshop on Content-based Image Retrieval for Biomedical Image Archives: Achievements, Problems, and Prospects. Citeseer, 2007, pp. 53–62.

[12] J. G. Dy, C. E. Brodley, A. Kak, L. S. Broderick, and A. M. Aisen, "Unsupervised feature selection applied to content-based retrieval of lung images," IEEE transactions on pattern analysis and machine intelligence, vol. 25, no. 3, pp. 373–378, 2003.

[13] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications—clinical benefits and future directions," International journal of medical informatics, vol. 73, no. 1, pp. 1–23, 2004.

[14] N. Hegde, J. D. Hipp, Y. Liu, M. Emmert-Buck, E. Reif, D. Smilkov, M. Terry, C. J. Cai, M. B. Amin, C. H. Mermel et al., "Similar image search for histopathology: Smily," NPJ digital medicine, vol. 2, no. 1, pp. 1–9, 2019.

[15] X. Shi, F. Xing, K. Xu, Y. Xie, H. Su, and L. Yang, "Supervised graph hashing for histopathology image retrieval and classification," Medical image analysis, vol. 42, pp. 117–128, 2017.

[16] F. Wang, T. W. Oh, C. Vergara-Niedermayr, T. Kurc, and J. Saltz, "Managing and querying whole slide images," in Medical Imaging 2012: Advanced PACS-Based Imaging Informatics and Therapeutic Applications, vol. 8319. International Society for Optics and Photonics, 2012, p. 83190J.

[17] A. Shafique, M. Babaie, M. Sajadi, A. Batten, S. Skdar, and H. Tizhoosh, "Automatic multi-stain registration of whole slide images in histopathology," in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), 2021, pp. 3622–3625.

[18] E. Vali Betts, D. M. Dwyre, H.-Y. Wang, and H. H. Rashidi, "Pax5-negative classical hodgkin lymphoma: A case report of a rare entity and review of the literature," Case Reports in Hematology, vol. 2017, p. 7531729, Oct 2017.

[19] C. Van der Weyden, S. Pileri, A. Feldman, J. Whisstock, and H. Prince, "Understanding cd30 biology and therapeutic targeting: a historical perspective providing insight into future directions," Blood cancer journal, vol. 7, no. 9, pp. e603–e603, 2017.

[20] C. Cobaleda, A. Schebesta, A. Delogu, and M. Busslinger, "Pax5: the guardian of b cell identity and function," Nature immunology, vol. 8, no. 5, pp. 463–470, 2007.

[21] M. Alizadeh, O. H. Maghsoudi, K. Sharzehi, H. Reza Hemati, A. Kamali Asl, and A. Talebpour, "Detection of small bowel tumor in wireless capsule endoscopy images using an adaptive neuro-fuzzy inference system," Journal of biomedical research, vol. 31, no. 5, pp. 419–427, Sep 2017.

[22] M. S. Hosseini and M. Zekri, "Review of medical image classification using the adaptive neuro-fuzzy inference system," Journal of medical signals and sensors, vol. 2, no. 1, pp. 49–60, Jan 2012.

[23] J. Jang, "Anfis: Adaptive-network-based fuzzy," IEEE Trans Neural Netw, vol. 3, no. 889, p. 98, 1992.

[24] A. Riasatian, M. Babaie, D. Maleki, S. Kalra, M. Valipour, S. Hemati, M. Zaveri, A. Safarpoor, S. Shafiei, M. Afshari et al., "Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides," Medical Image Analysis, vol. 70, p. 102032, 2021.