IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 18, 2019, accepted January 29, 2019, date of publication February 4, 2019, date of current version February 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897281

Heterogeneity-Aware Local Binary Patterns for Retrieval of Histopathology Images

HAMED ERFANKHAH[1], MEHRAN YAZDI[1], MORTEZA BABAIE[2], AND HAMID R. TIZHOOSH[2]
[1]Faculty of Electrical and Computer Engineering, Shiraz University, Shiraz 71946, Iran
[2]Kimia Lab, University of Waterloo, Waterloo, ON N2L 3G1, Canada

Corresponding author: Morteza Babaie (mbabaie@uwaterloo.ca)

This work was supported by the Ontario Research Fund-Research Excellence (ORF-RE).

ABSTRACT Histopathology images exhibit considerable variability, which can make diagnosis prone to uncertainty and errors. Using retrieval systems to locate similar images when a query image is given can assist pathologists in making more reliable decisions when diagnosing, based on accurately diagnosed past cases. Local binary patterns (LBP) have been successfully used as image descriptors for different applications. However, using LBP on histopathology images is still under investigation from different perspectives. The immense texture variability of these images and the lack of labeled training data are among the challenges that must be addressed. In this paper, we propose a new extension of LBP that is explicitly aware of the heterogeneity of local texture patterns through heterogeneity-based weighting. We use both homogeneity and the second moment (variance) of local neighborhoods to extend LBP histograms with heterogeneity information so that they better capture the polymorphism in histopathology images. We applied all concepts at multiple scales and investigated the extensions through both separated and concatenated extended histograms. We applied the proposed method on three publicly available datasets, namely, KimiaPath24, invasive ductal carcinoma (IDC), and BreakHis. The experimental results revealed that our features could retrieve and classify images with the highest accuracy. For KimiaPath24, we achieved 96.4% surpassing both LBP (91.0%) and deep features from VGG network (79.5%). For IDC, we reached the highest $F1$ measure (0.7665). Only for BreakHis, the proposed method was slightly less accurate than a customized convolutional neural network with rule fusion.

INDEX TERMS Content based image retrieval (CBIR), whole slide image (WSI) processing, heterogeneity mask, local binary pattern (LBP).

I. INTRODUCTION

In recent years, research on histopathology images has attracted a lot of interest. Histopathology is concerned with the investigation and analysis of biological cells and tissue structures by examining the thin slice of tissue under a microscope with regard to the specific diseases of interest. Histopathology slides (prepared tissue samples on a glass slide) are used for education and diagnosis purposes. In the educational domain, histopathology slides can help students learn about biological tissues. For diagnostic purposes, tissue samples (also called specimen) taken from the patient's body can be investigated to learn more about the morphological characteristics indicating the patient's condition. Generally, these daily investigations verify the presence of diseases and determine their causes. Thus, they can help pathologists in making proper treatment recommendations [1].

With the rapid increase in the availability of digital image acquisition devices, glass slides with tissue samples are captured by whole slide digital scanners and stored as digital *whole slide images* (WSIs). These WSIs are generally very large; therefore, examining them manually is very time-consuming and overwhelming for pathologists and other laboratory experts. In recent decades, many computer-aided systems have been proposed to analyze histopathology images [2]–[4]. Improvements in image processing and machine-learning algorithms with regard to big data processing, in addition to the delivery of acceptable results by the application of such algorithms to practical clinical datasets, have demonstrated the applicability of computer-aided algorithms as reliable tools enabling diagnostic accuracy [2], [5], [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Luca Cassano.

Content-based image retrieval (CBIR) and image classification can be considered as applications of computer-aided systems. In a classification task, the objective is to classify the WSIs as one of the disease categories or different lesion types [7]–[9]. In a retrieval approach, images that reflect the same visual attributes, such as color, texture, and shape, with respect to a query image, are indexed and retrieved from large image archives [10]–[12]. The assessment and identification of similar images can assist pathologists in performing diagnosis by providing a baseline for comparison against accurately diagnosed past cases.

Feature extraction is typically considered as the main stage in computer-aided systems. Texture is a low-level image feature and an important property of histopathology images that can be regarded as a similarity criterion to group images. In texture-based image retrieval, the repeatability, locality, quantity, and distinctiveness of texture patterns can improve the performance of a CBIR system [13]. The scale invariant feature transform (SIFT) [14], speeded up robust features (SURF) [15], histograms of oriented gradients (HoG) [16], and local binary patterns (LBP) [17] are all effective feature extraction methods. However, LBP may qualify better as a texture operator. Most of these methods, such as SIFT, SURF, and HOG, perform well in applications such as object detection, face recognition, or tracking [18]–[20], but perform poorly in comparison to LBP when applied to the retrieval and classification of histopathology images [21]–[23].

Representing each image in the dataset with low dimension feature vectors, or reducing the dimensionality of each feature vector with various methods such as principal component analysis [24] or minimum description length [25], may also be considered to reduce the memory requirements and response time of computer-aided systems.

In this paper, we propose a heterogeneity-aware multi-resolution LBP (hmLBP). By considering the power of rotation-invariant uniform LBP in medical domain we used that as our initial model. The intuition behind the proposed method is that the measurement of heterogeneity in multi-resolutions can leverage the texture recognition capability of LBP when dealing with immense pattern variability in histopathology scans. General (non-medical) texture patterns are usually processed in low resolutions. The extreme variability and high resolution of tissue patterns are further amplified by the different manifestations of malignancies for different tissue types. This is called polymorphism and poses a considerable challenge in diagnostic pathology that existing versions of LBP appear to miss. For computers, this is only manageable either through the preparation of large labeled datasets encompassing all variations, which is understandably impractical, or through the design of hand-crafted descriptors that place more emphasis on variability, i.e., heterogeneity.

We applied our method to the KimiaPath24, IDC, and BreakHis public histopathology datasets. The experimental results demonstrate the effectiveness of the proposed method, particularly for KimiaPath24 and IDC. For KimiaPath24, our method achieved an accuracy of 96.24%, which outperforms all results reported in the literature, including deep features.

The rest of this paper is organized as follows: In Section II, we review related studies. In section III, we introduce and explain the proposed method in details. In Section IV, we describe the datasets and their corresponding evaluation criteria. In Section V, we present and discuss the experimental results. Finally, we present the conclusions drawn from this study in the last section.

II. RELATED WORKS

A. IMAGE RETRIEVAL AND CLASSIFICATION

Feature extraction is considered as a key aspect of CBIR systems. Generally, features can be assembled by combining visual features (texture, color, shape, and segments). Moreover, features can be divided into general and domain-specific features [26]. To narrow the gap between computer similarity and human specialist similarity (semantic gap) in CBIR, domain-specific features have been applied broadly in every CBIR system, which is the most challanging problem [27].

Deep-network-based features are the main representative of this class. Deep solutions have been extensively used in histopathology image retrieval. In [28], a sparse autoencoder was applied to extract features from histopathology images. Kieffer *et al.* [29] used pre-trained deep networks to extract feature vectors. In [30], a patch-based convolutional neural network (CNN) was presented. The authors proposed an expectation-maximization (EM) method for CNN training. In another study [31], a nuclei-guided feature extraction method based on a CNN was proposed for histopathology images.

Other learning methods have also been applied to create a domain-specific feature. Zheng *et al.* [32], [33] have utilized the topic model of probabilistic latent semantic analysis (PLSA). First, block LBP, block Gabor, and nucleus-based features were extracted. Then, the PLSA model was applied to discover the high-level semantics of pathology images. In [34], multiple clustered instance learning (MCIL) was used to classify, segment, and cluster cancer cells in colon histopathology images. The MCIL method performed an image-level classification, pixel-level segmentation, and patch-level clustering simultaneously by using the SURF algorithm investigated in [35]. The discriminative feature-oriented dictionary learning (DFDL) method for learning class-specific features was used in [36] with the objective of emphasizing inter-class differences while keeping the intra-class differences small to enhance the classification performance. Zhu *et al.* [37] proposed a variant of the bag-of-visual-words with multiple dictionaries for histopathology image classification. The features extracted from the patches were clustered to form multiple cookbooks. This study used histogram intersection and support vector machines (SVM) to build multiple classifiers, and majority voting was applied to the classification of each patch.

Texture features play an important role in medical image retrieval [38]. Using visual features to retrieve

similar histopathology images was first investigated by Comaniciu *et al.* [39], who presented a content-based image retrieval system that supports decision making in the field of clinical pathology. A fast color segment method was used to extract the features of interest, namely, the shape, area, and texture of the nucleus.

Texture feature descriptors can be divided into key-point based features and dense sampling methods. SIFT, SURF, and ORB are the most well-known methods in this category. Gheisari *et al.* [40] proposed the combination of SIFT with a feature encoding algorithm to extract highly discriminative features from neuroblastoma histology images, and applied an SVM classifier to classify the images into five subtypes. In [41], medical image classification based on the PLSA-BOW model using SIFT features was proposed. The experimental results indicated an improvement in the classification results in comparison with other classification methods. Wojnar and Pinheiro [42] used the SURF descriptor and SVM classifier to classify lung images. The application of the SURF descriptor led to an improvement in the classification of lung images.

Although key-point based methods have been widely applied to general image processing, it has been repeatedly reported in the medical field that they might fail to detect useful key-points [22], [43], [44]. However, dense sampling methods (i.e. LBP) have solved this problem by sampling all of the pixels in the image and creating a histogram. Despite the fact that they may not detect the exact position of an object, they can distinguish texture types more efficiently comparing to key-point based methods. Reis *et al.* [45] applied multiscale basic image features and LBP in combination with a random decision classifier to categories, and automatically classified stromal regions according to their maturity. Camlica *et al.* [46] used an SVM classifier trained with LBP features derived from saliency image regions containing significant information to classify medical x-ray images. Song *et al.* [47] proposed a bag-of-visual-words method based on the histograms of oriented gradients (HOG) and the LBP to represent regions of interest (ROIs) in lung CT images. They applied the Max-Min posterior pseudo-probabilities (MMP) learning method to identify the category of the imaging sign contained in each ROI. The experimental results confirmed the effectiveness of their approach.

Finally, various studies have investigated the combination of high level features and domain-specific features by designing new texture descriptors or modifying common texture descriptors for a specific task. In [48], a model capable of extracting high-level features was proposed to reflect the semantic content of an image. This study mapped conventional low-level features to high-level features by implementing machine learning techniques. The Radon transform, which is used in the medical field to create X-ray, CT, and MR images, has recently attracted particular attention [49]. Tizhoosh and Babaie [22] introduced encoded local projections (ELP), based on the Radon transform, as a new dense-sampling descriptor for searching and classification in the medical field.

B. LOCAL BINARY PATTERN

LBP is a simple but powerful local texture descriptor [50] based on thresholding the neighborhood of each pixel against the center value, and considering the result as a binary code. The basic LBP values in a (P, R) neighborhood are computed as follows:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \tag{1}$$

where

$$s(x) = \begin{cases} 1 & x \geqslant 0, \\ 0 & x < 0. \end{cases} \tag{2}$$

The intensity value g_c corresponds to the intensity value of the local neighborhood's center pixel, and g_p (=0,..., P-1) corresponds to the gray values of the P equally-spaced pixels on a circle with a radius R ($R > 0$) that form a circular symmetric neighbor set [50].

In this study, we used another extension of LBP with a shorter feature vector for the rotation of invariant uniform patterns. The term **uniform** refers to the limited number of bitwise transitions from 0 to 1 or 1 to 0 in the binary LBP codes at the circular presentation. Ojala *et al.* [51] noticed that uniform patterns comprise the vast majority of texture patterns; that is, over 90% in (8, 1) neighborhoods. The rotation invariant local binary pattern in the (P, R) neighborhood is defined as follows:

$$LBP^{ri}_{P,R} = min\{ROR(LBP_{P,R}, i)|i = 0, 1, ..., P - 1\}, \tag{3}$$

where ROR(x, i) performs a circular bit-wise right shift on the P-bit number x for i times. The rotation-invariant uniform patterns $LBP^{riu2}_{P,R}$ in the (P, R) neighborhood can be computed by the following operator:

$$LBP^{riu2}_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP^{ri}_{P,R}) \leqslant 2, \\ P + 1 & \text{otherwise,} \end{cases} \tag{4}$$

where s is a sign function, and $U(LBP^{ri}_{P,R})$ refers to the uniform LBP pattern in (P, R), and can be computed as follows:

$$U(LBP^{ri}_{P,R}) = \left|s(g_{p-1} - g_c) - s(g_0 - g_c)\right| + \sum_{p=1}^{P-1} \left|s(g_p - g_c) - s(g_{p-1} - g_c)\right|. \tag{5}$$

After calculating the LBP values of the input image, the discrete occurrence histogram of the rotation invariant uniform patterns is computed over an image. Because of its computational simplicity, LBP has been used in many applications such as face detection and recognition [52], gender recognition [53], and image retrieval and classification [45], [46].

C. SUPPORT VECTOR MACHINE (SVM)

SVM is a class of supervised machine-learning algorithm which can be used for both classification or regression problems [54], [55]. SVM operates on an n-dimensional feature space with the value of each feature being the value of a particular coordinate. Subsequently, a margin-based optimization will be applied to separate these points by finding the optimal hyper-plane that maximizes the margin between classes. However, SVM was originally designed for binary classification but several algorithms such as ''one-against-one,'' ''one-against-all'' and DAGSVM [56] have been proposed to extend it to multi-class cases by combining several binary classifiers

III. PROPOSED METHOD

The main idea of this study may be expressed as a simple rule: **If** patterns exhibit high heterogeneity, **then**, the local binary patterns are counted as usual; **otherwise**, their influence on the construction of the histogram is reduced by considering their weighted sum whereas the weights should quantify the heterogeneity of the neighborhoods. Measuring the heterogeneity at multiple resolutions (manifested in using neighborhoods of different sizes-radii of the operator window-in LBP) ascertains to exploit the well-known benefits of a pyramid-approach in local neighborhoods.

This process is equivalent to capturing the pixel variability. Hence, we can use well-established statistical measures like variance, and concepts like homogeneity.

We extracted the texture features from a circularly symmetric pixel neighborhood with P members and radius R, denoted as (P, R). To exploit the benefits of multi-resolution analysis, the LBP features and heterogeneity measures were extracted for three different values of (P, R). Moreover, we investigated two distinct, albeit similar, ways for capturing heterogeneity, as follows:

1) Second moment (variance) of the neighborhood based on its average μ:

$$V = \frac{1}{p}\sum_{i=1}^{p}(g_i - \mu)^2. \tag{6}$$

The variance values are continuous and invariant to rotation.

2) Local dissimilarity can be calculated based on any notion of homogeneity that captures the uniformity of a group of elements. If all elements have the same value, then, the homogeneity of that set/neighborhood is equal to one. The local homogeneity H of a region in an image can be computed as follows [57]:

$$H = 1 - \frac{1}{L}\sqrt{\sum_i \sum_j (w_{ij} - m)^2}, \tag{7}$$

where w_{ij} are pixels in the region, m is the median or mean value of the pixels in the region, and L is the size of the region. In the homogeneity calculation of the neighborhood pixels in (P, R), we set $L = P$. To capture the dissimilarity D of a local region, we use $D = 1 - H$ for (P, R). Similar to the variance features, the dissimilarity features are also continuous and invariant to rotation.

Binning the continuous features– To create a histogram from the LBP features by heterogeneity map, we first use the extracted $LBP^{riu2}_{P,R}$ values in Eq. 4 (this matrix will hereafter be referred to as **LBP**). Subsequently, a weighted histogram is created by multiplying the heterogeneity weights $\mathbf{W}$ with each bin of the LBP histogram. To formulate the heterogeneity-aware histogram $\breve{\mathbf{h}}$, we define $\Psi_i(x, y)$ as follows:

$$\Psi_i(x, y) = \begin{cases} 1, & \text{if } LBP(x, y) = i, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where the binary matrix Ψ_i can be created for each arbitrary i from 1 to n (number of LBP bins). The coordinates x and y specify a position in the **LBP** and Ψ. By applying the Frobenius inner product $\langle \cdot, \cdot \rangle_F$, each bin of the final histogram ($\breve{\mathbf{h}}$) can be defined as follows:

$$\breve{\mathbf{h}}(i) = \langle \mathbf{W}, \Psi_i \rangle_F. \tag{9}$$

Thus, we can compute the heterogeneity matrix $\mathbf{W}$ based on the variance (Eq. 6) or dissimilarity (Eq. 7). Figure 1 shows the feature extraction process for the thumbnail of a histopathology image from the KimiaPath24 dataset. The image was divided into square patches, and then sample LBP values (with seven bins) were extracted for each block. Additionally, a sample heterogeneity matrix is shown under the LBP values. The weighted histogram is presented as the final outcome. For colored WSIs the feature extraction process is applied to each color channel space separately and then these channel histogram features are concatenated to form the final feature vector.

To provide an intuitive understanding of the proposed method, Fig. 2 shows the LBP and heterogeneity images in the first row, based on the variance and dissimilarity concepts. The corresponding heterogeneity-aware histograms for various (P, R) values can be seen in the second rows. It can be seen that the heterogeneity maps describe the non-homogeneous regions of an image. The homogeneous regions (low pixel variability regions) have low values, while the non-homogeneous regions (high pixel variability regions) have high values in the heterogeneity images (V and D images). As the P and R values increase, more details of the non-homogeneous and homogeneous regions in the original image are clarified. By comparing the V and D images, we can conclude that the heterogeneity image based on dissimilarity contains more detailed information in comparison with the image based on variance. As shown in the figure, the shape of the LBP histogram generally approximates the heterogeneity-aware histograms. However, the last LBP bin, which contains non-uniform patterns, is damped in both of the proposed methods and a little less in various first and last bins. Additionally, the height of the bins is reduced considerably by reducing the effect of the flat regions.

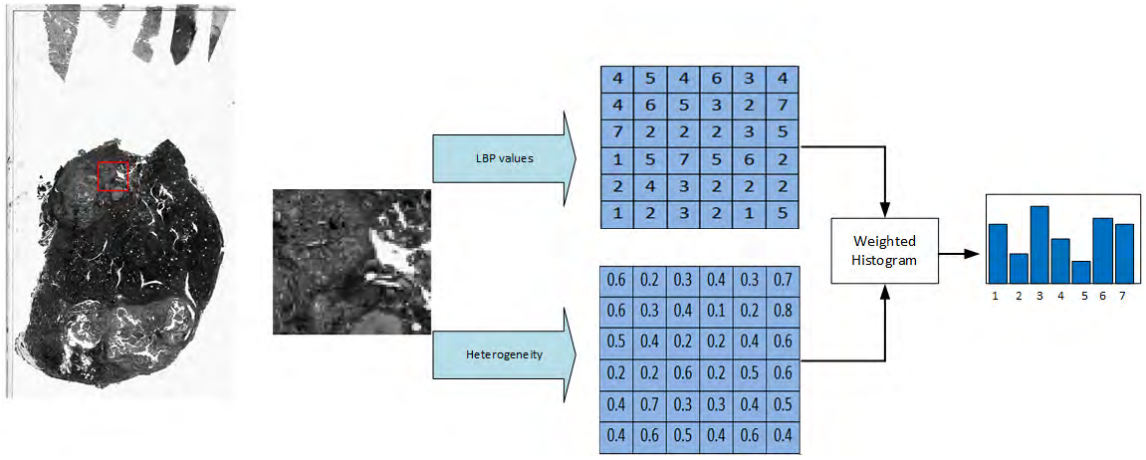
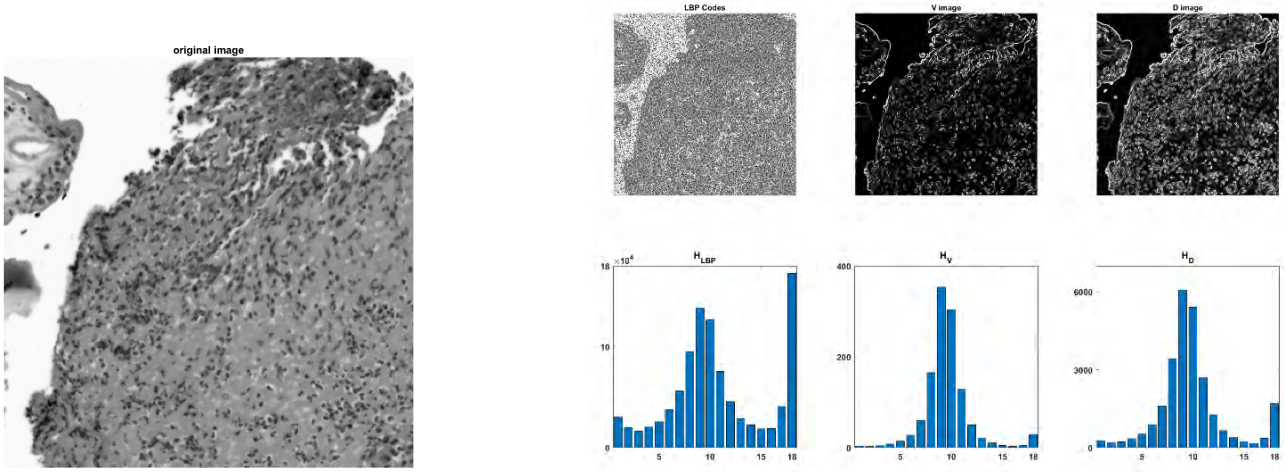


FIGURE 1. Feature extraction process for histopathology image thumbnail from KimiaPath24 dataset.



(a) (b)

FIGURE 2. LBP and heterogeneity images based on variance and dissimilarity concepts and corresponding heterogeneity-aware histograms for sample image (a) of (P, R) values for (b) P=16, R=2.

IV. DATASETS

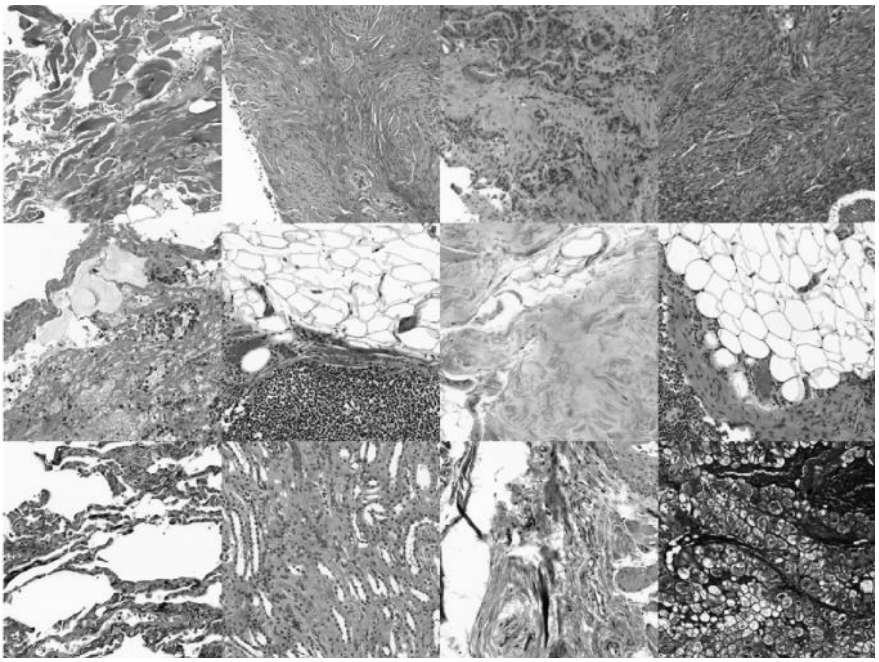
In this study, we used three publicly available datasets containing histopathology images: KimiaPath24, IDC, and BreakHis. In all experiments, as a pre-processing step, we normalized all patches in the [0, 1] interval.

KimiaPath24 dataset– This dataset contains 24greyscale whole slide images, which were manually selected from a larger pool of scans. The images show diverse body parts with texture patterns that appear different to human operators [58]. The glass slides were captured by a digital scanner (TissueScope LE 1.0) in a bright field using a 0.75 NA lens. The dataset contains 1,325 test images (patches) with a size of 1000 × 1000 pixels (0.5 mm × 0.5 mm) from all 24 cases. However, it is possible to extract a different number of training images based on the preference of the algorithm designer. Figure 3(a) shows various thumbnails colored images of this dataset.

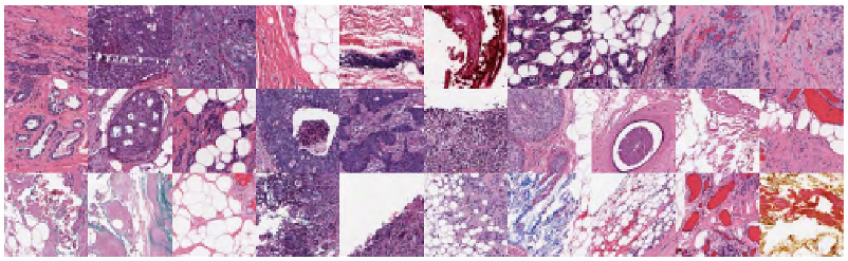
TABLE 1. Image distribution by magnification factor and class [60].

magnification	Benign	Malignant	Total
40x	625	1370	1995
100x	644	1437	2081
200x	623	1390	2013
400x	588	1232	1820
Total	2480	5429	7909
# Patients	24	58	82

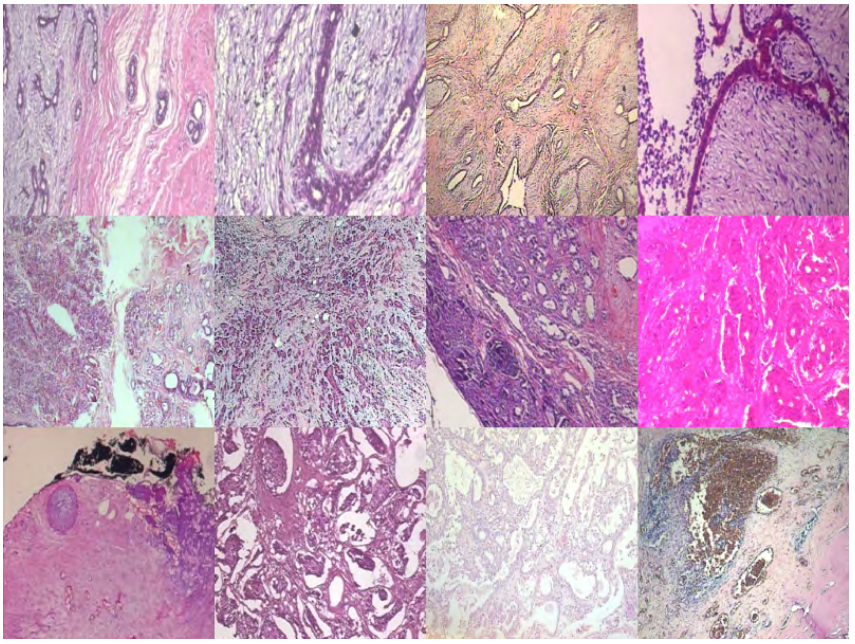
The various patches selected for training may not contain significant information because they may be primarily background patches and devoid from any substantial



(a)



(b)



(c)

FIGURE 3. Various samples from (a) KimiaPath24, (b) IDC, and (c) BreakHis datasets.

TABLE 2. Accuracy $\{\eta_p, \eta_w, \eta_{total}\}$ for different features and various combinations of these features in supervised and unsupervised approaches for KimiaPath24 dataset.

h	$\|h\|$	search			classification
		l1	l2	χ^2	svm
		$\{\eta_p, \eta_w, \eta_{total}\}$	$\{\eta_p, \eta_w, \eta_{total}\}$	$\{\eta_p, \eta_w, \eta_{total}\}$	$\{\eta_p, \eta_w, \eta_{total}\}$
$LBP_{8,1}$	10	{60.38%, 57.88%, 34.95%}	{61.51%, 58.32%, 35.87%}	{63.25%, 60.49%, 38.25%}	{72.53%, 68.33%, 49.56%}
$LBP_{16,2}$	18	{66.94%, 64.66%, 43.29%}	{65.89%, 63.63%, 41.92%}	{68.30%, 66.01%, 45.09%}	{80.30%, 78.14%, 62.74%}
$LBP_{24,3}$	26	{68.75%, 66.17%, 45.50%}	{71.17%, 69.17%, 49.23%}	{71.70%, 69.73%, 50.00%}	{80.98%, 78.59%, 63.64%}
$LBP_{8,1} + LBP_{16,2}$	28	{71.70%, 69.16%, 49.59%}	{71.40%, 69.01%, 49.27%}	{73.43%, 71.13%, 52.23%}	{87.92%, 86.19%, 75.78%}
$LBP_{8,1} + LBP_{24,3}$	36	{74.19%, 71.53%, 53.07%}	{71.70%, 69.24%, 49.64%}	{76.15%, 73.77%, 56.18%}	{88.60%, 86.98%, 77.07%}
$LBP_{16,2} + LBP_{24,3}$	44	{71.32%, 69.53%, 49.59%}	{70.04%, 67.99%, 47.62%}	{73.21%, 71.14%, 52.08%}	{88.53%, 87.00%, 77.02%}
$LBP_{P,R}$	54	{75.02%, 72.72%, 54.33%}	{73.13%, 70.74%, 51.73%}	{76.08%, 73.89%, 56.21%}	{91.02%, 89.33%, 81.31%}
$V_{8,1}$	10	{60.68%, 58.93%, 35.76%}	{59.02%, 57.89%, 34.16%}	{68.83%, 67.17%, 46.23%}	{71.17%, 69.82%, 49.69%}
$V_{16,2}$	18	{69.21%, 68.67%, 47.53%}	{66.04%, 65.32%, 43.14%}	{73.81%, 73.12%, 53.97%}	{83.02%, 82.50%, 68.49%}
$V_{24,3}$	26	{73.13%, 71.72%, 52.45%}	{69.81%, 68.70%, 47.96%}	{77.74%, 76.59%, 59.54%}	{85.96%, 84.89%, 73.05%}
$V_{8,1} + V_{16,2}$	28	{74.57%, 74.16%, 55.30%}	{71.62%, 71.49%, 51.21%}	{79.17%, 78.26%, 61.96%}	{90.11%, 89.37%, 80.54%}
$V_{8,1} + V_{24,3}$	36	{78.04%, 77.28%, 60.31%}	{76.15%, 75.05%, 57.15%}	{81.13%, 79.86%, 64.80%}	{92.60%, 91.66%, 84.88%}
$V_{16,2} + V_{24,3}$	44	{77.51%, 76.66%, 59.42%}	{75.47%, 74.82%, 56.47%}	{80.07%, 78.97%, 63.23%}	{92.91%, 92.39%, 85.83%}
$V_{P,R}$	54	{78.26%, 77.41%, 60.59%}	{75.85%, 75.32%, 57.13%}	{81.36%, 80.35%, 65.37%}	{93.81%, 93.05%, 87.29%}
$D_{8,1}$	10	{64.45%, 62.83%, 40.50%}	{65.43%, 64.08%, 41.93%}	{69.81%, 67.98%, 47.46%}	{73.74%, 71.98%, 53.08%}
$D_{16,2}$	18	{69.96%, 69.34%, 48.51%}	{67.40%, 67.16%, 45.26%}	{71.85%, 71.25%, 51.19%}	{83.40%, 82.78%, 69.03%}
$D_{24,3}$	26	{72.98%, 71.83%, 52.42%}	{72.45%, 71.67%, 51.93%}	{77.58%, 76.61%, 59.43%}	{86.19%, 84.91%, 73.18%}
$D_{8,1} + D_{16,2}$	28	{73.89%, 73.36%, 54.21%}	{72.08%, 71.29%, 51.38%}	{76.30%, 75.44%, 57.56%}	{91.55%, 90.92%, 83.24%}
$D_{8,1} + D_{24,3}$	36	{77.13%, 76.74%, 59.19%}	{76.00%, 75.69%, 57.53%}	{78.79%, 78.40%, 61.78%}	{92.23%, 91.50%, 84.39%}
$D_{16,2} + D_{24,3}$	44	{76.00%, 75.14%, 57.10%}	{73.43%, 72.54%, 53.27%}	{78.94%, 78.05%, 61.61%}	{92.15%, 91.69%, 84.49%}
$D_{P,R}$	54	{77.89%, 77.11%, 60.06%}	{75.02%, 74.69%, 56.03%}	{79.25%, 78.59%, 62.28%}	{94.72%, 93.78%, 88.83%}
$LBP_{P,R} + V_{P,R}$	108	{82.94%, 82.04%, 68.04%}	{80.08%, 78.80%, 63.10%}	{84.53%, 83.22%, 70.35%}	{96.45%, 95.23%, 91.85%}
$LBP_{P,R} + D_{P,R}$	108	{81.51%, 80.16%, 65.34%}	{79.77%, 78.31%, 62.47%}	{82.19%, 80.80%, 66.41%}	{95.85%, 94.86%, 90.92%}
$V_{P,R} + D_{P,R}$	108	{80.30%, 79.87%, 64.14%}	{78.04%, 77.94%, 60.83%}	{83.55%, 83.13%, 69.45%}	{96.38%, 95.51%, 92.05%}
$LBP_{P,R} + V_{P,R} + D_{P,R}$	162	{82.42%, 81.47%, 67.15%}	{80.53%, 79.24%, 63.81%}	{85.06%, 83.67%, 71.16%}	{96.00%, 94.68%, 90.89%}

tissue pattern. Therefore, we used a homogeneity criterion to select only patches containing a considerable part of the tissue. Patches with homogeneity over 99% were dismissed as training cases. Additionally, we extracted a total of 26,596 patches with a size of 1000 × 1000 pixels as training data. Despite the fact that KimiaPath24 is multi-class, the main limitation of this dataset is related to its classes nature. Two same tissue type (i.e. fat tissue) test patches might come from two different scans. As a result, two visually similar patches should classify in different classes.In addition, this dataset comes grayscale witch make it hard for color based approaches.

IDC dataset– This dataset contains breast cancer histopathology slides belonging to 162 patients diagnosed with invasive ductal carcinoma (IDC) at the University of Pennsylvania Hospital and The Cancer Institute of New Jersey [59]. All slides were digitized at 40x magnification. Each WSI was downsampled by a factor of 16:1 to a resolution of 4 μm/pixel. The dataset contains patches of size 50 × 50 in 3-channel RGB (Red-Green-Blue) color space and was randomly split into three different subsets by data supplier, comprising 84 training (DS1), 29 validation (DS2) cases for parameter exploration, 49 test cases as final instances (DS3), 114,235 instances for full training (DS1 and DS2), and 51,171 instances for testing (DS3). Each patient's file name was formatted as *u_xX_yY_classC.png*, where u is the patient ID, X and Y are the x-coordinate and y-coordinate of the location from where this patch was cropped, and C indicates the class where 0 is non-IDC and 1 is IDC. Examples of this dataset are shown in Fig. 3(b). Small patch size (50 × 50) and binary classification are the main limitations of IDC dataset which may limit its performance.

BreakHis dataset– This dataset contains biopsy images of different benign and malignant breast tumors. The images were collected through a clinical study from January 2014 to December 2014 [60]. The samples were generated from breast tissue biopsy slides and stained with hematoxylin and eosin (H&E). This dataset contains 7,909 images of size 700 × 460 in 3-channel RGB, categorized into benign and malignant tumors, which were collected at four different magnification factors: 40x, 100x, 200x, and 400x. Table 1 summarizes the image distribution of this dataset, and various samples are shown in Fig. 3(c). However, BreakHis offers multi-magnification patches but but providing just two classes labels (malignant versus benign) decrease its applicability.

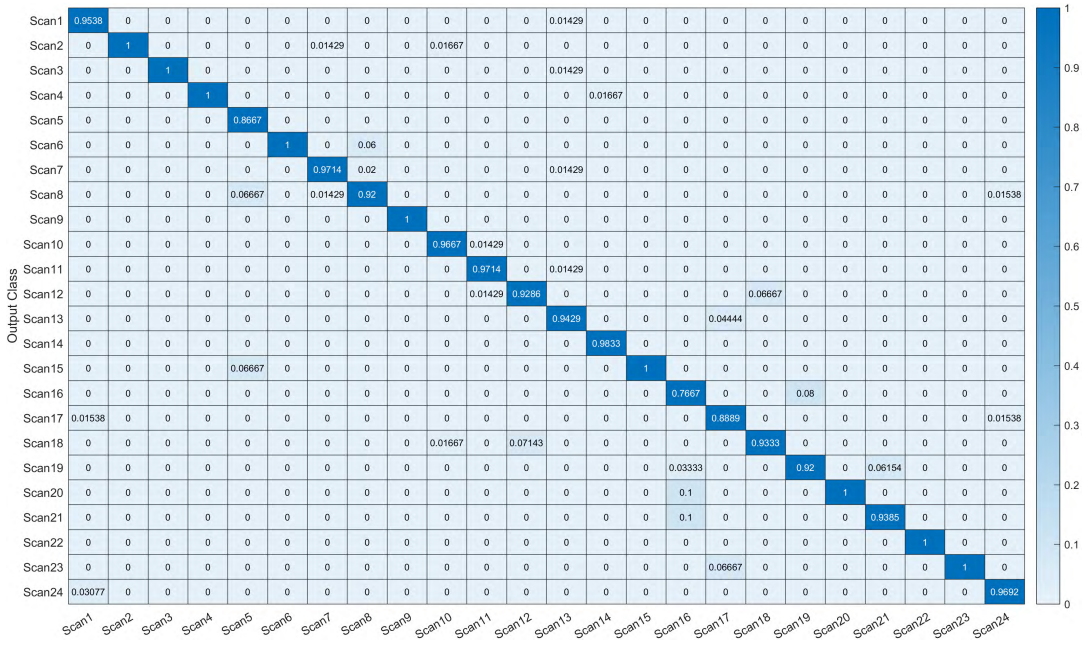


FIGURE 4. Confusion matrix for KimiaPath24 dataset.

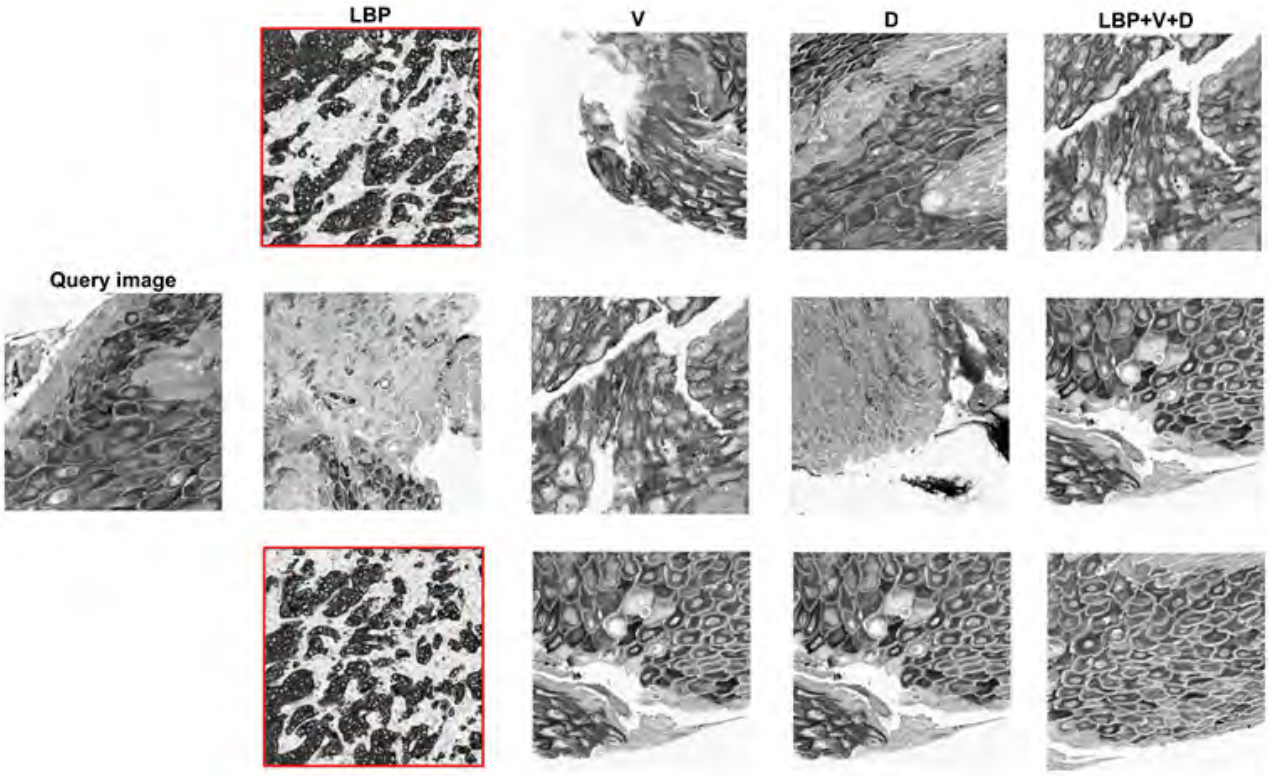


FIGURE 5. Image retrieval results based on *LBP*, *V*, *D*, and *LBP+V+D* features: left image is query image; right images are retrieved images. Images with borders are incorrect retrievals (KimiaPath24 dataset).

A. ACCURACY CALCULATION

In order to generate comparable results, we used the same evaluation metrics that have been put forward by these datasets.

KimiaPath24 accuracy measurement – There are $\eta_{tot} = 1,325$ testing patches P_s^j belonging to 24 sets $\Gamma_s = \{P_s^j | s \in S, j = 1, 2, ..., n_{\Gamma_s}\}$ with $s = 0, 1, ..., 23$ (n_{Γ_s} being the number of tests per each WSI). Considering the set of

TABLE 3. Accuracy of various methods for KimiaPath24 dataset. Results for the proposed approach are given in **bold**.

Method	Accuracy $\{\eta_p, \eta_w, \eta_{total}\}$	$\|h\|$	# training patches
$\boldsymbol{V_{P,R} + D_{P,R}}$	**{96.4%,95.5%,92.1%}**	**108**	**26596**
$\boldsymbol{D_{P,R}}$	**{94.7%,93.8%,88.8%}**	**54**	**26596**
$\boldsymbol{V_{P,R}}$	**{93.8%,93.1%,87.3%}**	**54**	**26596**
$\boldsymbol{LBP_{P,R}}$	**{91.0%,89.3%,81.3%}**	**54**	**26596**
$MBOW$ [37]	$\{89.2\%, 85.3\%, 76.9\%\}$	1600	25390
$ELP_{(10,d)}$ [22]	$\{82.7\%, 79.9\%, 66.1\%\}$	1024	27055
$ELP_{(10,m)}$ [22]	$\{82.3\%, 79.3\%, 65.3\%\}$	256	27055
VGG_{FC7} [22]	$\{79.5\%, 76.9\%, 61.1\%\}$	4096	n.a.
$LBP_{(24,2)}$ [22]	$\{77.8\%, 73.3\%, 57.0\%\}$	555	27055
VGG_{pool} [22]	$\{72.5\%, 67.2\%, 48.7\%\}$	6227	n.a.
CNN [29]	$\{65.0\%, 64.8\%, 42.1\%\}$	40513	n.a.
BoVM [29]	$\{65.0\%, 61.0\%, 39.7\%\}$	800	27055

TABLE 4. Accuracy $\{F1, BAC\}$ for IDC dataset.

h	serach			classification
	l1 $\{F1, BAC\}$	l2 $\{F1, BAC\}$	χ^2 $\{F1, BAC\}$	svm $\{F1, BAC\}$
$LBP_{P,R}$	$\{0.6284, 0.7444\}$	$\{0.6268, 0.7431\}$	$\{0.6279, 0.7438\}$	$\{0.7437, 0.8455\}$
$V_{P,R}$	$\{0.6196, 0.7418\}$	$\{0.6067, 0.7315\}$	$\{0.6165, 0.7397\}$	$\{0.7367, 0.8386\}$
$D_{P,R}$	$\{0.6277, 0.7459\}$	$\{0.6150, 0.7364\}$	$\{0.6213, 0.7413\}$	$\{0.7452, 0.8450\}$
$LBP_{P,R} + V_{P,R}$	**{0.6439, 0.7587}**	**{0.6346, 0.7514}**	**{0.6380, 0.7549}**	$\{0.7614, 0.8578\}$
$LBP_{P,R} + D_{P,R}$	$\{0.6427, 0.7560\}$	$\{0.6365, 0.7512\}$	$\{0.6355, 0.7509\}$	$\{0.7620, 0.8588\}$
$V_{P,R} + D_{P,R}$	$\{0.6264, 0.7462\}$	$\{0.6160, 0.7383\}$	$\{0.6221, 0.7434\}$	$\{0.7562, 0.8544\}$
$LBP_{P,R} + V_{P,R} + D_{P,R}$	$\{0.6404, 0.7560\}$	$\{0.6331, 0.7501\}$	$\{0.6361, 0.7531\}$	**{0.7665, 0.8613}**

retrieved images **Rt** for an experiment, the **patch-to-scan accuracy** η_p can be defined as follows [58]:

$$\eta_p = \frac{1}{\eta_{tot}} \sum_{s \in S} |Rt \cap \Gamma_s| \,. \tag{10}$$

The **whole-scan accuracy** η_w can be defined as follows:

$$\eta_w = \frac{1}{24} \sum_{s \in S} \frac{|Rt \cap \Gamma_s|}{n_{\Gamma_s}}, \tag{11}$$

where the total accuracy is defined as $\eta_{total} = \eta_p \times \eta_w$.

IDC accuracy measurement – For the IDC dataset, we used the F-measure (F1) and balanced accuracy (BAC), which are defined as follows [59]:

$$\text{F1} = \frac{2 \times Pr \times Rc}{Pr + Rc}, \tag{12}$$

$$\text{BAC} = \frac{\text{Sensitivity} + \text{Specificity}}{2}, \tag{13}$$

where Rc and Pr are the recall and precision, respectively. These numbers are obtained by calculating the true positive (TP), false positive (FP), false negative (FN), and true negative (TN) as follows:

$$Pr = \frac{TP}{TP + FP}, \tag{14}$$

$$Sensitivity = \frac{TP}{TP + FN} = Rc, \tag{15}$$

$$Specificity = \frac{TN}{TN + FP}. \tag{16}$$

BreakHis accuracy measurement – For the BreakHis dataset, we used the patient score and global recognition rate according to [60], as follows:

$$PatientScore = \frac{N_{rec}}{N_{Pa}}, \tag{17}$$

$$\text{Recognition Rate} = \frac{\sum \text{patient scores}}{\text{total number of patients}}, \tag{18}$$

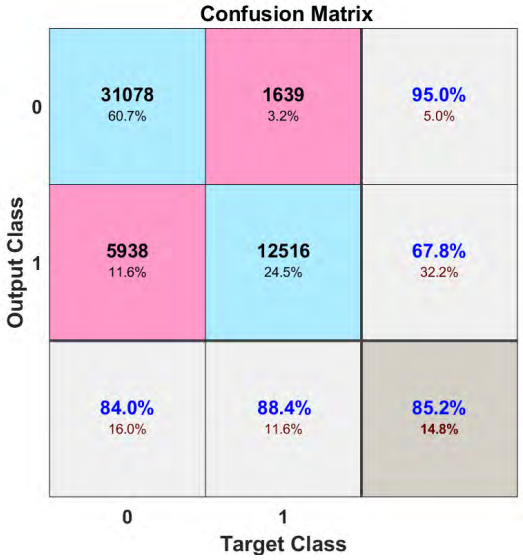


FIGURE 6. Confusion matrix for IDC dataset.

where N_{Pa} is the number of cancer images for patient Pa, and N_{rec} is the number of images classified correctly for each patient.

V. EXPERIMENTS

A. PREPROCESSING

To balance the images, all images were rescaled to [0 1] interval. The rescaled image $\bar{I}$ can be computed as follows:

$$\bar{I} = \frac{I - I_{min}}{I_{max} - I_{min}} \tag{19}$$

where I_{min} and I_{max} are the minimum and maximum values of the image. For multichannel images, each channel was normalized separately.

B. EXPERIMENT SETTINGS

The IDC, Kimiapath24, and BreakHis datasets have been split into test and train sets through the guidelines of the dataset providers. We extracted the $LBP_{P,R}$, variance V, and heterogeneity D features from the circularly symmetric neighborhoods with (P, R) values of (8, 1), (16, 2), and (24, 3). Both the retrieval and classification approaches were examined. In the retrieval approach, three different distance measures, including the L_1 and L_2 norms, and the χ^2 distance were used to retrieve the most similar patch with respect to a given test patch (query image). In the classification approach, we applied SVM to clasiffy extracted features. In this study, we used binary-class SVM for IDC and BreakHis datasets and multi-class SVM with one-against-all strategy for KimiaPath24 dataset. Kernel function was set to radial basis function and the penalization coefficient C and factor γ in the kernel function have been obtained through exhaustive experiments for each case dataset which have been provided in the results subsection. In our experiments, we used the LIBSVM package [61].

TABLE 5. F1 and BAC measures for various methods for IDC dataset (results for our approach are given in **bold**).

Method	F1	BAC
Proposed method	**0.7665**	**0.8613**
Alexnet,resize [62]	0.7648	0.8468
CNN [59]	0.7180	0.8423
Fuzzy Color Histogram [59]	0.6753	0.7874
RGB Histogram [59]	0.6664	0.7724
Gray Histogram [59]	0.6031	0.7337
JPEG Coefficient Histogram [59]	0.5758	0.7126
MPEG7 Edge Histogram [59]	0.5485	0.6979
Nuclear textural [59]	0.3915	0.6199
Nuclear Architectural [59]	0.3472	0.6009
HSV Color Histogram [59]	0.3446	0.6022

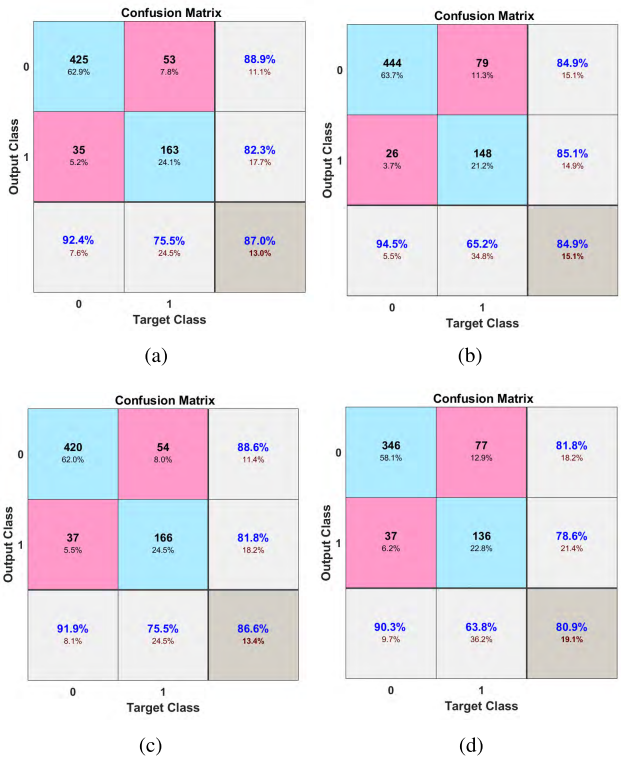


FIGURE 7. Confusion matrices for BreakHis dataset at magnification factors of (a) 40x, (b) 100x, (c) 200x and (d) 400x.

C. RESULTS AND COMPARISONS

Results for KimiaPath24 dataset– Table 2 presents the effect of combining heterogeneity information and LBP features on the performance in the retrieval and classification approaches. This table reports the results for extracted features from different values of (P, R) and various combinations

TABLE 6. Global recognition rate for different extracted features and various combinations of these features in classification and retrieval approaches for BreakHis dataset.

	Magnification factor															
	40x				100x				200x				400x			
Method	L1	L2	χ^2	SVM	L1	L2	χ^2	SVM	L1	L2	χ^2	SVM	L1	L2	χ^2	SVM
$LBP_{P,R}$	75.90%	74.25%	76.71%	85.33%	75.91%	74.01%	74.55%	84.54%	72.27%	71.27%	73.46%	82.12%	72.16%	69.81%	73.57%	76.46%
$V_{P,R}$	76.07%	72.36%	77.21%	85.79%	77.19%	76.71%	78.84%	85.55%	**80.78%**	**79.19%**	**81.65%**	83.53%	75.75%	73.84%	**78.65%**	81.95%
$D_{P,R}$	75.56%	75.30%	78.11%	86.03%	**79.95%**	**78.71%**	77.46%	86.43%	79.43%	76.15%	80.80%	84.61%	74.91%	72.99%	78.12%	**83.40%**
$LBP_{P,R} + V_{P,R}$	**78.55%**	76.19%	78.89%	85.68%	75.80%	75.37%	75.97%	88.25%	79.61%	78.80%	79.95%	**87.14%**	76.53%	71.90%	76.36%	80.92%
$LBP_{P,R} + D_{P,R}$	77.56%	75.79%	**79.15%**	84.16%	76.36%	75.27%	76.42%	86.17%	77.04%	76.02%	78.39%	83.93%	75.03%	73.21%	77.58%	79.93%
$V_{P,R} + D_{P,R}$	78.00%	74.74%	78.83%	**88.32%**	77.90%	77.03%	**79.30%**	87.51%	80.31%	78.68%	80.38%	86.34%	**76.79%**	**73.95%**	77.02%	81.29%
$LBP_{P,R} + V_{P,R} + D_{P,R}$	77.31%	**76.73%**	79.32%	85.90%	75.73%	76.02%	75.57%	**88.30%**	80.69%	78.58%	78.95%	85.93%	75.71%	72.31%	75.62%	77.75%

of these features. In this table, $|h|$ is the dimension of the feature vector, $X_{P,R}$ refers to the concatenation of all three histograms for the features with (P, R) values of (8, 1), (16, 2), and (24, 3). For example, $LBP_{P,R} = [LBP_{8,1}, LBP_{16,2}, LBP_{24,3}]$. This style was used for all three methods: $LBP_{P,R}$, $V_{P,R}$, and $D_{P,R}$. The best results are shown in bold fonts or gray color. As can be seen in the Table 2, by increasing the (P, R) values, the accuracy of the retrieval (search) and classification increased for all methods. The image searching using the χ^2 distance had higher accuracy than the L_1 and L_2 norms. The addition of $LBP_{24,3}$ or $LBP_{16,2}$ to $LBP_{8,1}$ increased the accuracy, owing to the rich information provided by the combination of features. This was also observed for the V and D features. From the combinations of individual cases, the heterogeneity-aware histogram $D_{P,R} = [D_{8,1}D_{16,2}D_{24,3}]$ had an accuracy of 94.72% in classification, while the heterogeneity-aware histogram $V_{P,R} = [V_{8,1}V_{16,2}V_{24,3}]$ had an accuracy of 81.36% in the retrieval approaches. From the combinations, an accuracy of 96.36% was achieved for classification by combining the heterogeneity-aware histograms of variance and homogeneity. The penalization coefficient C and factor γ were set to 400 and 2, respectively. Moreover, an accuracy of 85.06% was achieved for the retrieval approaches by combining the heterogeneity-aware histograms of variance and homogeneity, and the LBP histogram. Fig. 4 shows the confusion matrix. Based on this matrix, the most important source of error is associated with the false negative of scan 16 which misclassified in class 19 and 20 with the same probability of %10 each. These slides come from the same body organ (kidney) but in different staining i.e. H&E and immunohistochemistry (IHC). By considering the grayscale problem of KimiaPath24, it seems hard to separate them without the color information, just based on texture.

As an example, Fig. 5 shows three retrieved images based on LBP, V, D, and the concatenation of the $LBP + V + D$ features. The query image is shown on the left, while the retrieved and ranked images are shown on the right. The images with borders are incorrectly retrieved images. As can be seen, all of the images based on the V, D, and $LBP + V + D$ features were retrieved correctly. However, the images based on the LBP features were retrieved incorrectly. The two wrongly retrieved images appear to represent the same texture, although the sharpness of the textures is completely different. However, the heterogeneity based methods surpassed this problem considerably.

TABLE 7. Global recognition rate of various methods for BreakHis dataset. Results for our approach are given in bold.

Method	40x	100x	200x	400x
CNN & Fusion Rules [7]	90.0%	88.4%	84.6%	86.1%
Proposed method	**88.3%**	**88.3%**	**87.1%**	**83.4%**
CNN [63]	84.0%	83.9%	86.3%	82.1%
PFTAS [60]	83.8%	82.1%	85.1%	82.3%
CLBP [60]	77.4%	76.4%	70.2%	72.8%
GLCM [60]	74.7%	78.6%	83.4%	81.7%
QRB [60]	74.4%	69.4%	69.6%	67.6%
LPQ [60]	73.8%	72.8%	74.3%	73.7%

Comparison for KimiaPath24 dataset- The results obtained by the proposed method were compared with the results obtained by previous studies. Table 3 presents the accuracy measurements for various methods, and also lists the numbers of extracted training patches. In comparison with existing methods, the proposed method has lower dimensionality and higher image retrieval accuracy. As can be seen, the accuracy of different histograms would be ranked as $D > V > LBP$. These accuracy results are also higher than the accuracies reported in the literature. Moreover, the highest accuracy was achieved for the concatenated heterogeneity-aware histograms classified by the SVM method. However deep networks are extremely powerful tools for machine learning but no good results were achieved for KimiaPath24 dataset. Kieffer *et al.* [29] mentioned some of the reasons such as lack of color information, intra-class variability and scaling down a patch for use within a deep network and that's the reason of why using deep networks provide poor results on KimiaPath24 dataset.

Results for IDC dataset- We set the SVM parameters to $C = 40$ and $\gamma = 0.4$. A Python implementation of SVM was used for this dataset. We applied the proposed method to each color channel. All possible combinations of the extracted features were considered, and the accuracy of the combinations was calculated. The main results are presented in Table 4. As can be seen in Table 4, the performance improved by combining the $V_{P,R}$ and $D_{P,R}$ features with the $LBP_{P,R}$ features. The best performance, both in terms of F-measure and BAC, in the search mode was achieved by combining $V_{P,R}$ with $LBP_{P,R}$ and amounted to (0.6439, 0.7587). In the classification mode of the $LBP_{P,R}+V_{P,R}+D_{P,R}$ combination, we achieved (0.7665, 0.8613). Fig. 6 shows the confusion matrix. As we can see sensitivity, specificity and accuracy

TABLE 8. Major symbols used in this paper.

$w_{i,j}$: Pixels in the patch.	Rc: Recall.	D: Dissimilarity Image.	L_2: L2 norm distance.
P: The number of pixels on circle in LBP.	Pr: Precision.	η_p: Patch-To-Scan Accuracy.	χ^2: Chi Square distance.
R: Radius of circle in LBP.	TP: True Positive.	η_w: Whole-Scan Accuracy	N_{Pa}: The number of cancer images for patient Pa.
m: The median of the pixels in the patch.	FP: False Positive.	η_{total}: Total Accuracy	N_{rec}: The number of images classified correctly for each patient.
L: size of the region.	TN: True Negative.	Γ_s: Patches belong to scan s.	C: Penalization coefficient in SVM.
F_1: F-Measure.	V= Variance Image.	Rt: The set of retrieved images.	γ: Parameter of SVM in radial basis function.
$\langle \cdot, \cdot \rangle_F$: Frobenius inner product.	Ψ_i: Binary matrix.	L_1: L1 norm distance.	$x(p, r)$: Concatenating x(8,1), x(16,2) and x(24,3).
W: The heterogeneity weight matrix.			

are in the same range (between 84% to 88.4%). However, the specificity is the highest one by 88.4%. As a result, the probability of misclassifying a normal patch to the cancerous one is higher than to misclassify the cancerous patch to normal one.

Comparison for IDC dataset- We compared the performance of the proposed method with the performance of existing methods. Table 5 presents the comparison results. As can be seen, the proposed method has the best accuracy.

Results for BreakHis dataset- We concatenated the extracted features from each color channel for each resolution analysis. The SVM parameters are set to $\{C = 200, \gamma = 2\}$, $\{C = 200, \gamma = 1\}$, $\{C = 200, \gamma = 1\}$, and $\{C = 120, \gamma = 0.5\}$ for the 40x, 100x, 200x and 400x magnification factors, respectively. The global recognition rate was used as the evaluation criterion. Some of the main results are presented in Table 6, and the best results are printed in bold fonts. Amongst all feature vectors, for the 40x dataset, $LBP_{P,R}$+$D_{P,R}$ and $V_{P,R}$+$D_{P,R}$ achieved the best performance in the search and classification approaches with recognition rates of 79.15% and 88.32%. For the 100× dataset, $D_{P,R}$ and $LBP_{P,R} + V_{P,R}$+ $D_{P,R}$ achieved the best performance in the search and classification approaches with a recognition rate of 79.95% and 88.30%. For the 200x dataset, $D_{P,R}$ and $LBP_{P,R} + V_{P,R}$ achieved the best performance in the search and classification approaches with the recognition rate of 80.80% and 87.14%. For the 400x dataset, $V_{P,R}$ and $D_{P,R}$ achieved the best performance in the search and classification approaches with a recognition rate of 78.65% and 83.40%. Fig. 7 shows the confusion matrices at each magnification factor. The interesting point is, even though in 40x and 100x the global recognition rate are same, their accuracy is different around 2.1%. In addition, our method's specificity range between 63.8% in 400x to 75.5% in 40x is not satisfactory.

Comparison for BreakHis dataset- We compared our best results with the results obtained by previous studies. The comparisons are presented in Table 7. As can be seen, the proposed method could classify the dataset with higher accuracy than most reported methods.

VI. DISCUSSION

It seems our simple and compact method performs well in histology image search mainly based on the following reasons: 1- Multi-resolution analysis: texture features are extracted from different resolutions from histology images and then concatenated to form the final feature vector. The multi-resolution analysis provides rich texture information. 2- Extracting compact binary codes: using rotation invariant binary codes not only reduce the dimension of the feature vector which is important in multi-resolution analysis but also, comprises the vast majority of texture patterns that can describe the texture properties of the histology images well. 3- Weighting the histogram bins: LBP codes combine heterogeneity information by weighting the bins of the histogram through the counting of LBP codes. Heterogeneity highlights the non-homogeneous regions of the histology images and capturing heterogeneous information into the histogram bins of LBP codes will cause more texture information to be preserved.

VII. SUMMARY AND CONCLUSIONS

In this paper, we proposed a novel method of extending LBP histograms so that they are more aware of pattern heterogeneity. Thiswould be a particular interest to applications with highly variable texture patterns, such as applications in the field of histopathology. Additionally, we investigated heterogeneity at multiple scales to capture the pattern variability more robustly. As, in contrast to natural patterns images, histology patterns are processed at very high resolutions, hence, a multiscale approach exploits both low and high resolutional patterns. Existing LBP versions do not account for both heterogeneity and scale at the same time.

Moreover, we investigated the discrimination power of the developed descriptors with regards to search and classification, and tested them on three publicly available datasets, namely, the KimiaPath24, IDC, and BreakHis datasets. The results revealed that the extracted features could retrieve and classify the images with high accuracy. Furthermore, we tested various concatenating configurations of the extracted features. The results revealed that the concatenated features achieved better performance in modeling the histopathology texture patterns. In comparison with other proposed methods, our method has a lower dimension feature vector and higher accuracy. In future works, we need to investigate the possibility to equip the heterogeneity-awareness with learning capability to ascertain generalization for any concept drift due to introduction of new (unseen) patterns.

REFERENCES

[1] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, no. 2, pp. 147–171, Feb. 2009.

[2] C. Mosquera-Lopez, S. Agaian, A. Velez-Hoyos, and I. Thompson, "Computer-aided prostate cancer diagnosis from digitized histopathology: A review on texture-based systems," *IEEE Rev. Biomed. Eng.*, vol. 8, pp. 98–113, 2015.
[3] J. A. A. Jothi and V. M. A. Rajam, "A survey on automated cancer diagnosis from histopathology images," *Artif. Intell. Rev.*, vol. 48, no. 1, pp. 31–81, 2016.
[4] M. M. Dundar *et al.*, "Computerized classification of intraductal breast lesions using histopathological images," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 7, pp. 1977–1984, Jul. 2011.
[5] S. Huda, J. Yearwood, H. F. Jelinek, M. M. Hassan, G. Fortino, and M. Buckland, "A hybrid feature selection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis," *IEEE Access*, vol. 4, pp. 9145–9154, 2016.
[6] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 34–42, 2018.
[7] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2560–2567.
[8] M. Shah, D. Wang, C. Rubadue, D. Suster, and A. Beck, "Deep learning assessment of tumor proliferation in breast cancer histological images," in *Proc. IEEE Int. Conf. Bioinform. Biomed. (BIBM)*, Nov. 2017, pp. 600–603.
[9] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. K. A. Rao, "Histopathological image classification using discriminative feature-oriented dictionary learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 3, pp. 738–751, Mar. 2016.
[10] Z. Xue, L. R. Long, S. K. Antani, G. R. Thoma, "Pathology-based vertebral image retrieval," in *Proc. ISBI*, 2011, pp. 1893–1896.
[11] N. Mehta, R. S. Alomari, and V. Chaudhary, "Content based sub-image retrieval system for high resolution pathology images using salient interest points," in *Proc. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 3719–3722.
[12] A. Sridhar, S. Doyle, and A. Madabhushi, "Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces," *J. Pathol. Inform.*, vol. 6, p. 41, Jun. 2015.
[13] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Fundations Trends Comput. Graph. Vis.*, vol. 3, no. 3, pp. 177–280, Jan. 2008.
[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.
[16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.
[17] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[18] S. Yao, T. Wang, W. Shen, S. Pan, Y. Chong, and F. Ding, "Feature selection and pedestrian detection based on sparse representation," *PLoS ONE*, vol. 10, no. 8, p. e0134242, 2015.
[19] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "A comparative study of CFs, LBP, HOG, SIFT, SURF, and BRIEF techniques for face recognition," *Proc. SPIE*, vol. 10649, Apr. 2018, Art. no. 106490M.
[20] A. A. Micheal and K. Vani, "Comparative analysis of SIFT and surf on KLT tracker for UAV applications," in *Proc. IEEE Commun. Signal Process. (ICCSP)*, Apr. 2017, pp. 1000–1003.
[21] T. J. Alhindi, S. Kalra, K. H. Ng, A. Afrin, and H. R. Tizhoosh. (2018). "Comparing LBP, HOG and deep features for classification of histopathology images." [Online]. Available: https://arxiv.org/abs/1805.05837
[22] H. R. Tizhoosh and M. Babaie, "Representing medical images with encoded local projections," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 10, pp. 2267–2277, Oct. 2018.
[23] Ş. Öztürk and A. Bayram, "Comparison of HOG, MSER, SIFT, FAST, LBP and CANNY features for cell detection in histopathological images," *HELIX*, vol. 8, no. 3, pp. 3321–3325, 2018.
[24] U. Sinha and H. Kangarloo, "Principal component analysis for content-based image retrieval," *Radiographics*, vol. 22, no. 5, pp. 1271–1289, 2002.
[25] C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, and A. M. Aisen, "Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision," in *Proc. AAAI/IAAI*, 1999, pp. 760–767.
[26] Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Vis. Commun. Image Represent.*, vol. 10, no. 1, pp. 39–62, 1999.
[27] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognit.*, vol. 40, no. 1, pp. 262–282, 2007.
[28] A. A. Cruz-Roa, J. E. A. Ovalle, A. Madabhushi, and F. A. G. Osorio, "A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervent.* Berlin, Germany: Springer, 2013, pp. 403–410.
[29] B. Kieffer, M. Babaie, S. Kalra, and H. R. Tizhoosh. (2017). "Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks." [Online]. Available: https://arxiv.org/abs/1710.05726
[30] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2424–2433.
[31] Y. Zheng *et al.*, "Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification," *Pattern Recognit.*, vol. 71, pp. 14–25, Nov. 2017.
[32] Y. Zheng, Z. Jiang, J. Shi, and Y. Ma, "Retrieval of pathology image for breast cancer using plsa model based on texture and pathological features," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2304–2308.
[33] Y. Zheng, Z. Jiang, J. Shi, and Y. Ma, "Pathology image retrieval by block LBP based PLSA model with low-rank and sparse matrix decomposition," in *Proc. Chin. Conf. Image Graph. Technol.* Berlin, Germany: Springer, 2014, pp. 327–335.
[34] Y. Xu, J.-Y. Zhu, E. Chang, and Z. Tu, "Multiple clustered instance learning for histopathology cancer image classification, segmentation and clustering," in *Proc. CVPR*, 2012, pp. 964–971.
[35] A. Kallipolitis and I. Maglogiannis, "Content based image retrieval in digital pathology using speeded up robust features," in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov.* Cham, Switzerland: Springer, 2018, pp. 374–384.
[36] T. H. Vu, H. S. Mousavi, V. Monga, U. K. A. Rao, and G. Rao, "DFDL: Discriminative feature-oriented dictionary learning for histopathological image classification," in *Proc. IEEE Biomed. Imag. (ISBI)*, Apr. 2015, pp. 990–994.
[37] S. Zhu, Y. Li, S. Kalra, and H. Tizhoosh, "Multiple disjoint dictionaries for representation of histopathology images," *J. Vis. Commun. Image Represent.*, vol. 55, pp. 243–252, Aug. 2018.
[38] U. Avni, H. Greenspan, E. Konen, M. Sharon, and J. Goldberger, "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 733–746, Mar. 2011.
[39] D. Comaniciu, P. Meer, and D. J. Foran, "Image-guided decision support system for pathology," *Mach. Vis. Appl.*, vol. 11, no. 4, pp. 213–224, Dec. 1999.
[40] S. Gheisari, D. R. Catchpoole, A. Charlton, Z. Melegh, E. Gradhand, and P. J. Kennedy, "Computer aided classification of neuroblastoma histological images using scale invariant feature transform with feature encoding," *Diagnostics*, vol. 8, no. 3, p. 56, 2018.
[41] C. H. Cao and H. L. Cao, "The research on medical image classification algorithm based on PLSA-BOW model," *Technol. Health Care*, vol. 24, no. s2, pp. S665–S674, 2016.
[42] A. Wojnar and A. M. G. Pinheiro, "Annotation of medical images using the SURF descriptor," in *Proc. IEEE Biomed. Imag. (ISBI)*, May 2012, pp. 130–133.
[43] M. Kashif, T. M. Deserno, D. Haak, and S. Jonas, "Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment," *Comput. Biol. Med.*, vol. 68, pp. 67–75, Jan. 2016.
[44] S. W. Teng, M. T. Hossain, and G. Lu, "Multimodal image registration technique based on improved local feature descriptors," *J. Electron. Imag.*, vol. 24, no. 1, p. 013013, 2015.
[45] S. Reis *et al.*, "Automated classification of breast cancer stroma maturity from histological images," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2344–2352, Oct. 2017.

[46] Z. Çamlica, H. R. Tizhoosh, and F. Khalvati, "Medical image classification via SVM using LBP features from saliency-based folded data," in *Proc. IEEE Mach. Learn. Appl. (ICMLA)*, Dec. 2015, pp. 128–132.
[47] L. Song, X. Liu, L. Ma, C. Zhou, X. Zhao, and Y. Zhao, "Using HOG-LBP features and MMP learning to recognize imaging signs of lung lesions," in *Proc. IEEE Comput.-Based Med. Syst. (CBMS)*, Jun. 2012, pp. 1–4.
[48] J. C. Caicedo, F. A. Gonzalez, and E. Romero, "A semantic content-based retrieval method for histopathology images," in *Proc. Asia Inf. Retr. Symp.* Berlin, Germany: Springer, 2008, pp. 51–60.
[49] M. Babaie, H. R. Tizhoosh, A. Khatami, and M. E. Shiri, "Local radon descriptors for image search," in *Proc. IEEE Image Process. Theory, Tools Appl. (IPTA)*, Nov./Dec. 2017, pp. 1–5.
[50] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
[51] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
[52] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2004, pp. 469–481.
[53] N. Sun, W. Zheng, C. Sun, C. Zou, and L. Zhao, "Gender classification based on boosting local binary pattern," in *Proc. Int. Symp. Neural Netw.* Berlin, Germany: Springer, 2006, pp. 194–201.
[54] T. B. Trafalis and R. C. Gilbert, "Robust classification and regression using support vector machines," *Eur. J. Oper. Res.*, vol. 173, pp. 893–909, Sep. 2006.
[55] J. S. Nath and S. K. Shevade, "An efficient clustering scheme using support vector methods," *Pattern Recognit.*, vol. 39, pp. 1473–1480, Aug. 2006.
[56] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.
[57] A. Jurio, H. Bustince, M. Pagola, P. Couto, and W. Pedrycz, "New measures of homogeneity for image processing: An application to fingerprint segmentation," *Soft Comput.*, vol. 18, no. 6, pp. 1055–1066, 2014.
[58] M. Babaie *et al.*, "Classification and retrieval of digital pathology scans: A new dataset," in *CVMI Workshop CVPR*, 2017, pp. 1–16.
[59] A. Cruz-Roa *et al.*, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Proc. SPIE*, vol. 9041, Mar. 2014, Art. no. 904103.
[60] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.
[61] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
[62] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *J. Pathol. Inform.*, vol. 7, p. 29, Jul. 2016.
[63] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *Proc. IEEE Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 1868–1873.

HAMED ERFANKHAH received the master's degree in biomedical engineering from Shiraz University, Shiraz, Iran, in 2012, where he is currently pursuing the Ph.D. degree with the Faculty of Electrical and Computer Engineering. His major research interests include the field of medical image processing and big data analysis.

MEHRAN YAZDI received the B.Sc. degree in communication systems from the Department of Electrical and Electronic Engineering, Shiraz University, Shiraz, Iran, in 1992, and the M.Sc. and Ph.D. degrees in digital vision and image processing from the Department of Electrical Engineering, Laval University, Quebec, QC, Canada, in 1996 and 2003, respectively. He was on sabbatical leave from the MIA Lab, La Rochelle University, La Rochelle, France, from 2015 to 2016. He is currently an Associate Professor with the Department of Communications and Electronic Engineering, Shiraz University. He conducted several projects in the area of hyper-spectral image compression and denoising, CT metal artifact reduction, and video compression. His major research interests include the field of image and video processing, remote sensing, multimedia networks, multidimensional signal processing, and medical image analysis.

MORTEZA BABAIE received the Ph.D. degree in computer from the Amirkabir University of Technology, Tehran, Iran. He has had collaboration with the Kimia Lab, University of Waterloo, as a Visiting Scholar, since 2016. He is currently a Postdoctoral Fellow with the Kimia Lab, University of Waterloo. His main research interests include medical image processing, machine learning, and AI.

HAMID R. TIZHOOSH was a Research Associate with the Knowledge and Intelligence Systems Laboratory, University of Toronto, where he focused on dynamic bandwidth allocation using AI methods such as reinforcement learning. He has extensive industrial experience. He was with numerous companies. Since 2001, he has been a Professor with the Faculty of Engineering, University of Waterloo, where he also leads the Laboratory for Knowledge Inference in Medical Image Analysis. He is currently an AI Advisor with Huron Digital Pathology, St. Jacobs, ON, Canada. He has developed algorithms for medical image filtering, segmentation, and search. He has authored two books, 14 book chapters, and more than 140 journal and conference papers. He has also filed five patents in collaboration with Waterloo Commercialization Office. Since 1993, his research activities encompass artificial intelligence, computer vision, and medical imaging.