



# Improved bag-of-features using grey relational analysis for classification of histology images

Raju Pal<sup>1</sup> · Mukesh Saraswat<sup>1</sup> · Himanshu Mittal<sup>1</sup>

Received: 20 April 2020 / Accepted: 13 January 2021 / Published online: 8 February 2021  
© The Author(s) 2021

## Abstract

An efficient classification method to categorize histopathological images is a challenging research problem. In this paper, an improved bag-of-features approach is presented as an efficient image classification method. In bag-of-features, a large number of keypoints are extracted from histopathological images that increases the computational cost of the codebook construction step. Therefore, to select the a relevant subset of keypoints, a new keypoints selection method is introduced in the bag-of-features method. To validate the performance of the proposed method, an extensive experimental analysis is conducted on two standard histopathological image datasets, namely ADL and Blue histology datasets. The proposed keypoint selection method reduces the extracted high dimensional features by 95% and 68% from the ADL and Blue histology datasets respectively with less computational time. Moreover, the enhanced bag-of-features method increases classification accuracy by from other considered classification methods.

**Keywords** Keypoint selection · Grey relational analysis · Bag-of-features · Histopathological image analysis

## Introduction

Histopathology involves a microscopic investigation of diseased tissues for examining the pathogological and biological structures. For histopathological analysis, tissue slides are prepared by taking a tissue samples from the diseased body and stained them with different methods for better visualization of different tissue structures [17]. To convert a tissue slide into digital image, whole slide imaging (WSI) scanners are widely used [49]. The pathology labs are using digital tissue slides for the investigations which helped them in making the decisions accurately for disease diagnosis [50].

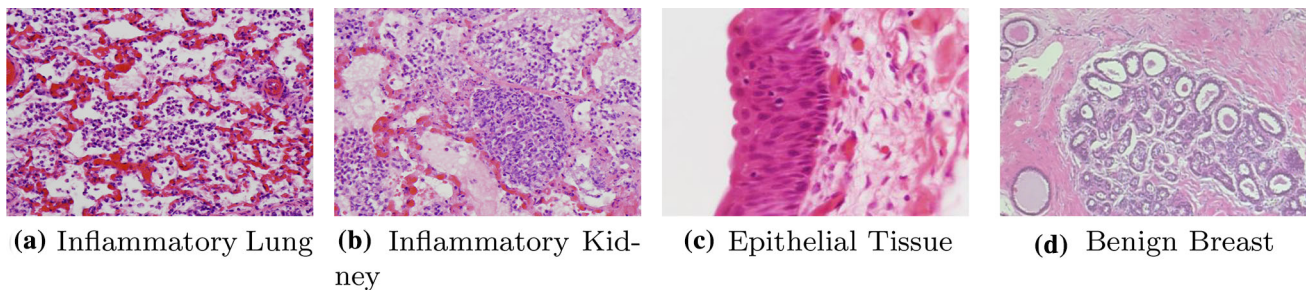
In recent years, there has been an huge growth of digital tissue images over the Internet and these images need to be well organized for better analysis and retrieval processes. Therefore, an automated system for the classification of histopathological images can be useful [32]. However, due to the complexity of histopathological images, it is a complicated task to design an automated image classification system.

Figure 1 depicts some images of tissues to illustrate their structural complexities. Generally, pathologists examine certain visual features from the histopathological images to classify them into their respective categories. To automate the classification process, such visual features are extracted by feature extraction methods but it is hard to extract features due to diverse and complex disease-specific tissue structures [31].

In the literature, several automated histopathological image categorization methods exist which are based on the approaches like graph algorithms [5], hashing [56], bag-of-features [9], and deep neural networks [24]. Song et al. [47] resolved the issue of variances within class and between classes for the given categories of histopathological images by proposing a sub-categorization-based model known as LMLE (large margin local estimate). The model is further extended for interstitial lung disease that is based on the locality constrained sub-cluster representation of an image [48]. Besides, Nayak et al. [32] developed an automated dictionary-based feature learning method to classify various morphometric regions in the whole slide image. Vu et al. [52] determined the discriminative features from images and used dictionary learning for classifying the histopathological images. Orlov et al. [33] presented a multipurpose automated image classifier, known as WND-CHARM, by extracting

✉ Raju Pal  
raju3131.pal@gmail.com

<sup>1</sup> Jaypee Institute of Information Technology, Noida,  
Uttar Pradesh, India



**Fig. 1** Structures of various H&E stained histopathological tissue images [19,29]

a large number of various types of features such as texture, a polynomial decomposition, and high contrast features. WND-CHARM is analyzed and tested on two applications, namely face recognition and biomedical image classification. Tang et al. [51] presented the I-Browse system to automatically classify the histopathological images by using visual features and semantic properties along with the contextual knowledge. A broad review of various computer-assisted diagnosis algorithms in medical imaging has been presented by Gurcan et al. [17]. Díaz et al. [12] presented a method to select and describe the local patches from histopathological images based on their staining components. These local patches are further used in the probabilistic latent semantic analysis framework (pLSA) for the classification. Moreover, Srinivas et al. [49] represented histological images using a multi-channel sparsity model, having specified channel-wise constraints, with a linear arrangement of training samples. Saraswat and Arya [44] presented and discussed different techniques for classification, segmentation, and feature detection of nuclei in histopathological images. Fondon et al. [15] provided an automated tissue classification method to diagnose the breast cancer carcinoma having four malignancy levels, namely invasive, in-situ, benign, and normal. The method considered three different types of features, i.e., texture, nuclei, and color regions to train the support vector machine classifier. Lichtblau and Stoean [26] proposed an automated classification method for cancer diagnosis which is based on the weighted outcome of six classifiers. The optimal weights are found using the differential evolution approach with error minimization as the objective function.

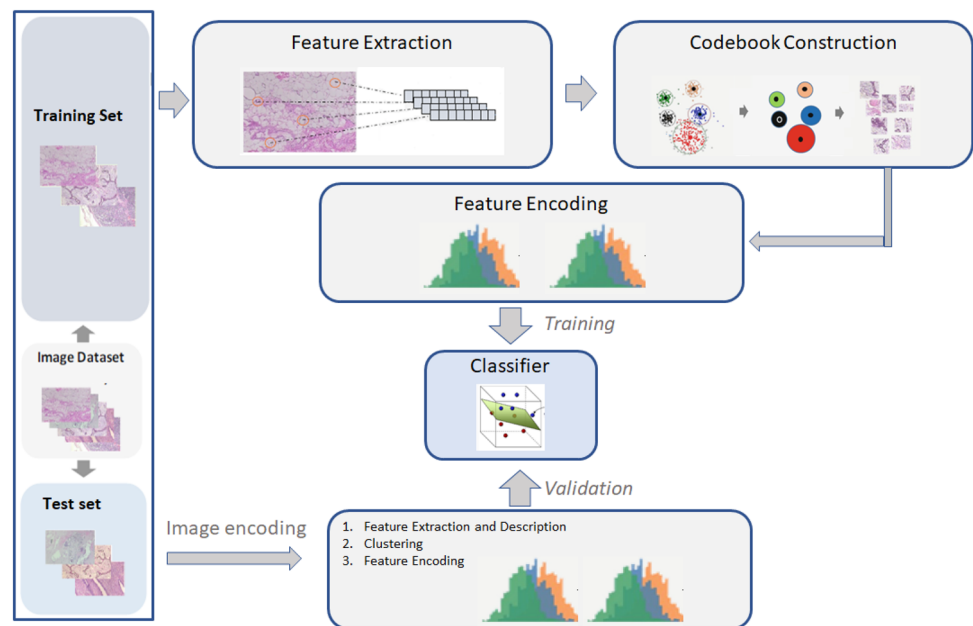
The above-discussed classification methods are based on local features that consider the features such as color, shape, texture, and distribution of the nuclei for the representation of the histopathological images. However, these features are not adequate for images, having complex and unbalanced visual structures [18]. Moreover, for better medical image representation, learning-based methods are used which automatically extract the features from the images and represent the complex morphological structures in a more meaningful way [36,46]. However, these methods are not computationally efficient. Therefore, to achieve better image representation,

mid-level features are used in medical image representation [40]. The bag-of-features (BOF) method [9] is one of the popular mid-level image representation methods. This concept is inherited from the bag-of-words (BoW) which is used for textual document analysis in natural language processing [25,41,42].

Recently, the BOF-based classification methods have been proved effective over the existing ones in terms of computational resources and efficiency for histopathological image analysis [24] [16]. Caicedo et al. [9] categorizes the histopathological images using the BOF method. Cruz et al. [11] represent the histopathological images in form of histogram of visual words and found the correlation between these visual patterns. To mitigate the rotation and scale invariance problem of image classification, Raza et al. [42] studied and analyzed the effect of both in renal cell carcinoma images and found that rotation invariance is more effective but by combining both better classification accuracy can be achieved. Moreover, the dictionary representation of the visual words enhances the performance of the BOF method. The efficiency of the BOF method is dependent on the codebook constructed using the K-means algorithm. However, the K-means clustering method sometimes sticks into local optima when applied on a large feature set [34]. To overcome this, Mittal and Saraswat [28] modified the codebook construction phase of the BOF method by generating optimal visual words using gravitational search algorithm for the categorization of tissue images. Furthermore, Pal and Saraswat [38] used biogeography-based optimization [35] for the codebook construction phase and tested the proposed method on ICIAR breast cancer dataset. However, the meta-heuristic-based codebook construction is a computationally expensive method [40].

The standard BOF method generally consists of four phases, namely feature extraction, codebook construction, feature encoding, and classification. The features are extracted in form of keypoints from the local regions of the images using any local feature descriptor like histograms of oriented gradients (HOG) [3], speeded up robust features (SURF) [6], and fast retina keypoints (FREAK) [2]. Furthermore, K-means clustering is used to form the vocabulary of visual

**Fig. 2** Bag-of-feature approach for histopathological image classification



words from the extracted keypoints and each image is then converted into a histogram of these visual words. The histograms along with the labels are used to train the classifier.

However, due to the complexity of histopathological images, the feature extraction phase may generate a large number of keypoint descriptors which makes the codebook construction phase computationally inefficient [57].

Various methods have been proposed in the literature to select the relevant keypoints [8,13]. Dorko and Schmid [13] divide the descriptor vector into groups using the Gaussian mixture model (GMM) and apply SVM to the most relevant group to improve the classification accuracy. Lin et al. [27] introduced two methods for keypoints selection using two different approaches (IKS1 and IKS2) to eliminate similar key points. The Euclidean distance is used as the similarity measure. The IKS1 and IKS2 methods show good performance on Caltech datasets.

On the other hand, due to the complex structural morphology of histopathological images, a large number of keypoints are extracted and there is no method exists in the literature to select the relevant keypoints. Therefore, in this paper, a new keypoints selection technique is introduced which uses the Grey relational analysis (GRA) to find the similarity between the keypoints.

The main contribution of this paper has three folds, (i) a new computationally efficient keypoints selection technique is proposed based on the GRA, (ii) the proposed method is introduced in the BOF method for finding the relevant keypoints, and (iii) the modified BOF method is used to automatically classify the histopathological images. To conduct the experimental analysis, two histopathological image datasets are considered, namely the Blue histology dataset

of tissue images and the animal diagnostic laboratory (ADL) histopathological image dataset. These dataset contains less number of images and the proposed method is specifically designed for the medical datasets having less number of images available.

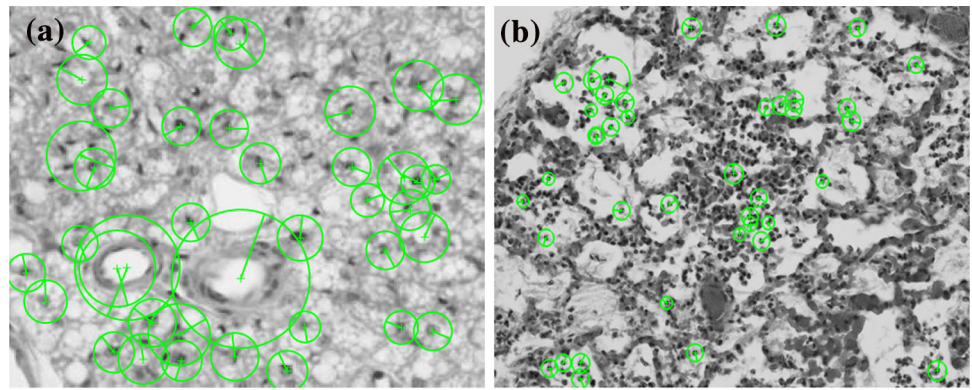
The rest of the paper contains a description of the standard BOF method in “Bag-of-features method” section followed by the description of the modified Grey relational analysis-based BOF method in “Proposed grey relational analysis-based bag-of-feature method” section. The result analysis and discussion on the considered real-world datasets are presented in “Experimental results” section. Finally, “Conclusion” section concludes the paper with some future work.

## Bag-of-features method

The BOF method is one of the convenient mechanisms for histopathological image classification. It generally consists of four phases as shown in Fig. 2: (i) Extract the texture features or keypoints using feature extraction method, (ii) Cluster the keypoints to generate the visual words, (iii) Encode each image as the histogram of visual words, and (iv) Train the classifier using these histograms and corresponding image labels. Finally, the images from the test set are fed to the trained classifier without a label to predict their labels. Mathematically, the BOF method can be described as follows:

Consider a set  $C = \{c_1, c_2, \dots, c_i, \dots, c_n\}$  of  $n$  classes. Each class  $c_i$  is associated with a set of images. The image dataset is divided into two parts. One is a training set on which

**Fig. 3** The keypoints detected by SURF in **a** connective tissue image and **b** inflamed lung tissue images



the classifier is trained and the other is a test set, which is used to validate the trained classifier. The training set of  $N$  images is prepared by randomly selecting  $M_i$  images from each class  $c_i$  which is also given by Eq. (1). The remaining images of the classes are considered as a part of the test set.

$$N = \sum_{i=1}^n M_i \quad (1)$$

- 1. Feature extraction:** Extract the keypoints from all  $N$  images of training set using a feature extraction method like, SURF [6], FREAK [2], SIFT [30], and HOG [3]. Let  $X$  is a set of keypoints, defined as Eq. (2).

$$X = [F_1, F_2, \dots, F_N]^T \quad (2)$$

where  $F_i$  is a matrix of  $P$  keypoints for the  $i^{th}$  image, defined over  $d$ -dimensional space and is given by Eq (3).

$$F_i = \begin{Bmatrix} f_{11} & f_{12} & f_{13} & \dots & f_{1d} \\ f_{21} & f_{22} & f_{23} & \dots & f_{2d} \\ \dots & \dots & \dots & \dots & \dots \\ f_{P1} & f_{P2} & f_{P3} & \dots & f_{Pd} \end{Bmatrix} \quad (3)$$

Figure 3 shows representative keypoints detected by one of the feature extraction method i.e., SURF from two images, randomly taken from the two considered histopathological image datasets. Each image is first converted to grayscale then SURF detector is used to find the predefined number of keypoints from these images. In the figure, only 40 strong keypoints are depicted for simplicity and visualization.

- 2. Code-book construction:** Create visual words by repeatedly grouping the extracted descriptor vector  $X$  into  $n$ -mutually exclusive clusters. Each cluster can have any number of keypoints based on the similitude of the intensity values of pixels in an image with the extracted keypoints. For the same,  $K$ -means clustering algorithm is used and the cluster centers returned by  $K$ -means are represented as visual words.

- 3. Encoding:** Encode each image into a histogram ( $H_j = \{H_{j1}, H_{j2}, \dots, H_{jn}\}$  for  $j = 1, 2, \dots, N$ ), representing the visual word occurrences in each image which is given by Eq. (4).

$$H_{jk} = \sum_{i=1}^P \mu_{ik}(j) \text{ where } k = 1, 2, \dots, n \quad (4)$$

$$\mu_{ik} = \begin{cases} 1 & \text{if } \|v_k - f_i\| \leq \|v_s - f_i\| \text{ for } s = 1, 2, \dots, n, \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where  $P$  represents the number of keypoints and  $\mu_{ik}(j)$  is 1 when any visual word ( $v_k$ ) is close to any keypoint  $f_i$  in the image. This method is also known as vector quantization.

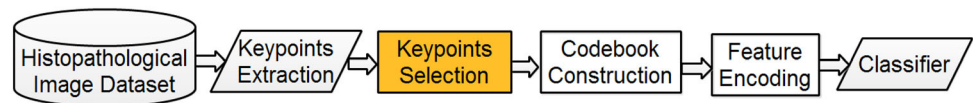
- 4. Classification:** Each histogram ( $H_j$ ) along with its annotation is used to train the classifier for the image classification task. Once the classifier is trained, it is tested to predict the label of images provided in the **test** set. Each **test** image is represented as the histogram as discussed above and fed to the classifier without a label. Based on the returned label by the classifier, its accuracy is measured.

## Proposed grey relational analysis-based bag-of-feature method

In the feature extraction phase of the BOF method, a feature detection and representation method is used to find the keypoints in the images. These keypoints are then represented as the descriptor vectors which are further used for code-book construction. Out of many feature extraction methods, SURF is the fastest method because it uses box filters for the convolution of images and converts each image as the integral image. It extracts the texture features from the images [6]. Moreover, SURF is a resolution invariant feature detector, hence images of different resolutions do not have any impact on the classification performance. This property of



**Fig. 4** Flowchart of the enhanced BOF method



SURF helps to analyze the histopathological images, having different resolutions (e.g., 10x, 20x, 40x) [55]. The interest points in the images are detected using the Hessian matrix approximation. SURF also shows good performance over other alternatives like SIFT [21]. Therefore, in the proposed method, the SURF feature detector is used to extract a set of keypoints ( $X$ ) from  $N$  training images.

Generally, SURF extracts a large number of keypoints due to the complex texture of histopathological images. This reduces the efficiency of visual vocabulary generation [27]. Furthermore, all of the detected keypoints are not necessary for image classification and annotation [27]. Hence, an efficient keypoints selection method is required for the acquisition of relevant keypoints that can improve the speed and efficiency of the BOF method. Some of the popular keypoints selection techniques are IB3 (**instance-based learning**) [1] and iterative keypoints selection (IKS1, IKS2) [27]. IB3 is an efficient instance selection method with high space complexity, while IKS1 and IKS2 are the keypoints selection methods that are used to find representative keypoints from the images. IKS1 and IKS2 are differed by their initial representative keypoints selection methods. In IKS1, representative keypoints are selected randomly while in IKS2, cluster centers are considered as representative keypoints. The remaining keypoints are eliminated based on their Euclidean distances from the selected representative keypoints. However, Euclidean distance similarity measure is computationally expensive for high-dimensional data. Chang et al. [10] **has shown that computational cost of Grey relational analysis (GRA)** [22]-based similarity measure is better than the Euclidean distance-based similarity. Therefore, in this work, a new GRA-based keypoints selection (GKS) method is introduced to reduce the number of keypoints before feeding them into the next phase of the BOF method i.e., codebook construction. The modified flow of the BOF method is depicted in Fig. 4. Moreover, the next subsection provides a detailed description of the Grey relation analysis-based keypoints selection method.

**Grey relational analysis-based keypoints selection**

The GKS method uses the concept of Grey relational analysis for finding the similarity between the keypoints. GRA [22] is a part of Grey system theory which is used to examine the similarity between data tuples based on geometrical mathematics [43]. It conforms to four basic principles in the dataset, i.e., proximity, normality, symmetry, and entirety [53]. In GRA, the similarity between a reference tuple and

the remaining tuples for a given data is computed by Grey relational grades (GRGs) whose value lies between 0 and 1. For any data tuple, if GRG is close to 1, then it is highly similar to the reference tuple while the dissimilarity will be signified if GRG value is close to 0 [10].

Therefore, the new keypoints selection method uses GRA to eliminate similar keypoints from the feature descriptor, generated by SURF. The new GKS method has the following steps:

1. Cluster the keypoints into  $K$  clusters using approximate K-means (AKM) algorithm [54]. AKM is used due to its less computational complexity.
2. Make the cluster centers as the member of selected keypoints set and also consider them as reference points for the computation of GRGs for the remaining keypoints.
3. Compute the GRG values between the reference point and the keypoints lying within the corresponding cluster. The mathematical formulation of GRG computation is described below.

Let  $X_o = X_{r1}, X_{r2}, \dots, X_{ri}, \dots, X_{rn}$  be a set of  $n$  reference points. The elements in  $X_o$  are of the form  $X_{ri} = \langle X_{ri}(1), X_{ri}(2), \dots, X_{ri}(u) \rangle$ , where  $u$  corresponds to the dimension of the extracted keypoint. Similarly, let  $X_c = X_{c1}, X_{c2}, \dots, X_{cm}$  be a set of  $m = P - n$  remaining keypoints considered as comparative keypoints where, each element in  $X_c$  can be denoted as  $X_{cj} = \langle X_{cj}(1), X_{cj}(2), \dots, X_{cj}(u) \rangle$ . **Here,  $P$  represents total number of keypoints.** The GRG value of each keypoint in  $X_c$  is given by Eq. (6) [10].

$$GRG(X_{oi}, X_{cj}) = \sum_{t=1}^u [\alpha_i(t) \cdot GRC(X_{oi}(t), X_{cj}(t))] \tag{6}$$

where GRC is the Grey relational coefficients and  $\alpha_i(t) = \frac{1}{u}$  is the weighting factor of GRC. The GRC value, between  $i^{th}$  keypoint of  $X_o$  and  $j^{th}$  keypoint of  $X_c$  at  $u^{th}$  datum, belonging to the  $i^{th}$  cluster only is given by Eq. (7) [10].

$$GRC(X_{oi}(u), X_{cj}(u)) = \frac{\min_{ij} \Delta_{ij}(u) + \xi \max_{ij} \Delta_{ij}(u)}{\Delta_{ij}(u) + \xi \max_{ij} \Delta_{ij}(u)}, \tag{7}$$

where  $\xi \in (0, 1]$  is a random number to control the constancy between  $\max_{ij} \Delta_{ij}(u)$  and  $\min_{ij} \Delta_{ij}(u)$ .  $\Delta_{ij}(u)$

is computed by  $|X_{oi}(u) - X_{cj}(u)|$  for  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, c$ .

4. In every cluster, the above computation is performed to find the highly similar points with cluster center and eliminate  $s\%$  of the keypoints from each cluster whose GRG values are higher, in their corresponding cluster. Here,  $s$  is termed as shrinking threshold.
5. Repeat the steps 1–4 till the remaining keypoints are greater than  $K$  and add the last set (having  $K$  points only) of cluster centers to the selected keypoints set.
6. Use the selected keypoints set as input to the next phase of BOF i.e., codebook construction.

After finding the optimum keypoints from the new GKS method, the codebook construction phase of BOF (as described in “Bag-of-features method” section) is performed which uses K-means clustering to generate various visual words. Furthermore, the frequencies of each visual word in the images are represented by histograms. These histograms along with the corresponding image labels are given to SVM for training which is further used for image classification.

## Experimental results

The experimental analysis has been conducted on MATLAB 2017a. The computer system includes Intel Core i5-2120 having 8 GB of RAM. The performance of the proposed method is analyzed in three phases on two histopathological image datasets. First, the proposed keypoints selection method (GKS) is compared with the state-of-the-art keypoints selection methods in “Performance analysis of proposed keypoint selection method” section. Second, the results of the GKS-based BOF method for classifying histopathological images are depicted in “Classification results of the GKS-based BOF Method” section. In the third phase, the performance of the proposed classification method has been analyzed against the state-of-the-art classification methods as well as some deep learning-based classification methods in Sects. 4.4 and 4.5, respectively.

### Datasets

Two standard histopathological image datasets are considered for the classification task, namely ADL histopathological image dataset and Blue histology image dataset which are described below.

- *ADL histopathological image dataset* [50]: This dataset is generated by Animal Diagnostics Lab at Pennsylvania State University which contains histopathological images of three different organs of animals namely Kidney, Lung, and Spleen. Each organ has healthy and inflamed tis-

sue images. Some of the images from these categories for each organ are depicted in Fig. 5. The hematoxylin and eosin (H&E) dye have been used for staining. The inflamed images can be identified by counting some specific white blood cells such as neutrophils and lymphocytes cells. These cells represent different types of infections in tissue images such as allergic infections, bacteria, parasites, and many others. The inflamed organ images depicted in Fig. 5 have uncleared alveoli which are permeated with bluish infected cells. These cells generally indicate the transferable disease. The dataset contains a total of 963 images of three organs. There are 335, 308, and 320 images of kidney, lung, and spleen, respectively.

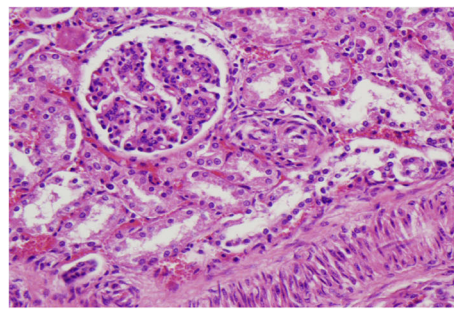
- *Blue histology image dataset* [19]: Every animal contains four types of tissues, namely connective tissue, nervous tissue, epithelial tissue, and muscle tissue. The connective tissues are comprised of various protein fibers like collagen or elastin. These protein fibers along with some ground substances create an extracellular matrix that provides shape to the organs and made them connected. These tissues can be found in palmar skin, adipose tissue, hyaline cartilage, and bone tissue slides. Nervous tissues are specific tissues that are constituted by the mind, nerves, and spinal cord [14]. These tissues generally contain two types of cells, namely neurons and glial. Neurons are used for communication between cells and glial cells provide support to nervous tissues. Muscle tissues contain muscle fibers which are elongated cells and used for contraction. Actin and myosin are the two proteins that are used to shorten the cells. The muscle tissues are responsible for movement within internal organs. Epithelium tissues provide the layer between the internal and external environment of the organ which is used to protect the organ from fluid loss, microbe, and laceration. The tissue cells are tightly connected with each other via cellular junction to provide the fence. Figure 6 shows sample images taken from each type of tissue images. Each image category contains 101 tissue images.

### Performance analysis of proposed keypoint selection method

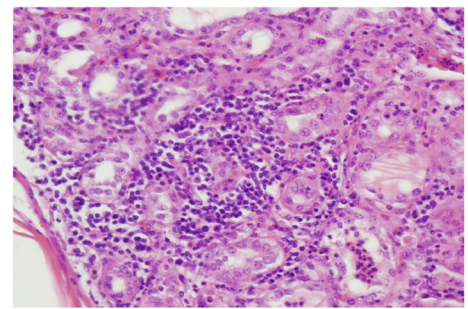
The performance of the GKS method is evaluated against three other methods, namely IB3 [1], IKS1 [27], and IKS2 [27]. IB3 is very old method but due to its simplicity, it has been treated as a baseline algorithm for the analysis of the new GKS. The other two methods, IKS1 and IKS2, find the representative keypoints from the images using iterative keypoints selection method. These methods have different initialization procedure. IKS1 selects the initial keypoints randomly while IKS2 uses cluster centroids returned by K-means algorithm as initial keypoints. After the initialization



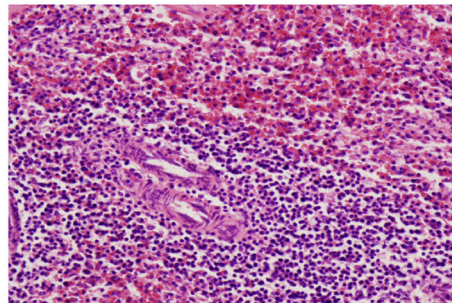
**Fig. 5** Ground truth labels for healthy and inflammatory tissues of Kidney, Lung, and Spleen [50]



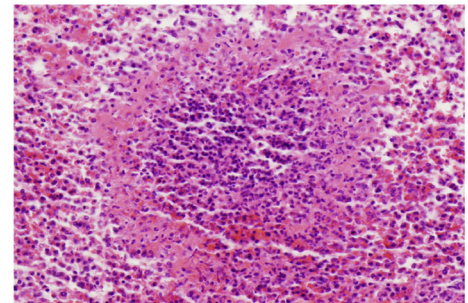
**(a)** Healthy Kidney



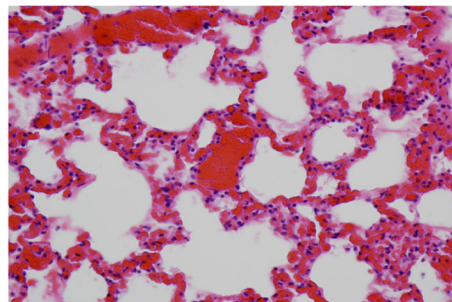
**(b)** Inflamed Kidney



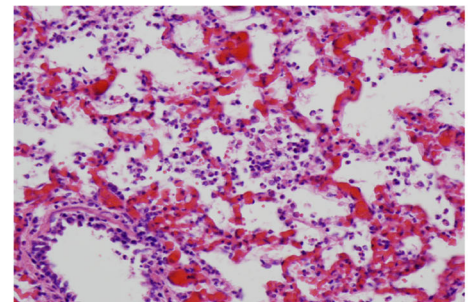
**(c)** Healthy Spleen



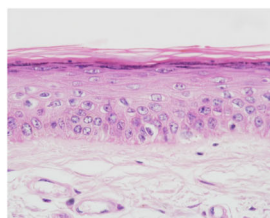
**(d)** Inflamed Spleen



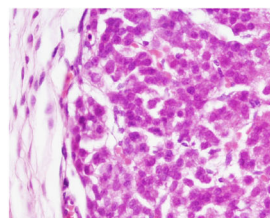
**(e)** Healthy Lung



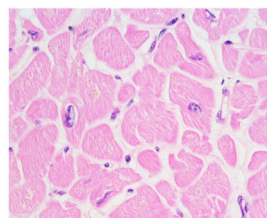
**(f)** Inflamed Lung



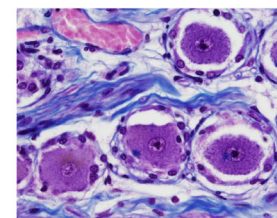
**(a)** ET



**(b)** CT



**(c)** MT



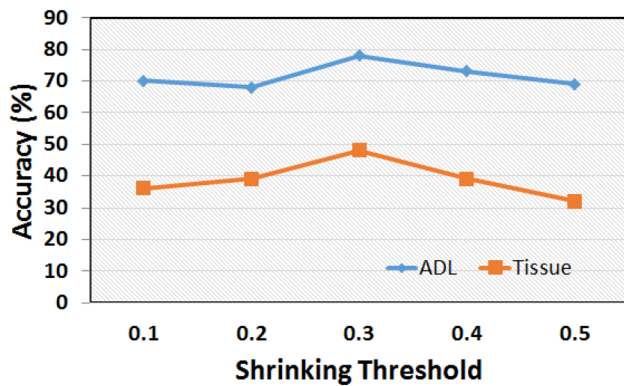
**(d)** NT

**Fig. 6** Representative animal tissues from blue histology dataset at 40× magnification level [19]. Here, *CT* connective tissue, *ET* epithelial tissue, *MT* muscle tissue, and *NT* nervous tissue

of keypoints, the other keypoints are not selected if their euclidean distances from representative keypoints are less than a predefined threshold. The parameter settings for all

the considered algorithms are taken from their respective literature [1,27].

Moreover, the GKS method uses a shrinking threshold to eliminate similar points from the clusters. In this paper, its



**Fig. 7** Classification accuracy on validation set using the GKS method with different shrinking threshold values

value is empirically set to 0.3 using its effect on classification accuracy on test images. To visualize the same, Fig. 7 shows the classification accuracy on the test images of two considered datasets for different shrinking threshold values. It can be observed from the figure that the classification accuracy on ADL and Blue histology datasets are higher at the shrinking threshold (s) of 0.3. It means that 30% highly similar keypoints, based on their GRG values, are eliminated in each iteration of the proposed GKS method. However, when this elimination rate increases to 40% or 50%, the classification accuracy is reduced. It may happen due to the fact that high elimination rate may delete some relevant keypoints required for better classification process. The other parameter in the proposed GKS method is the number of clusters for approximate k-means which is also set empirically to 1000.

The performance of the GKS method has been evaluated in terms of the number of selected keypoints and average computation time taken by the considered methods. Table 1 depicts the total number of extracted keypoints from SURF and selected keypoints by the GKS and the considered methods over two datasets. The percentage of the eliminated keypoints is also mentioned for each algorithm on the different datasets in parenthesis. Table 1 also depicts the average computational time taken by the different algorithms. From the table, it can be observed that the IB3 algorithm eliminates 85% and 64% keypoints from ADL and Blue histology

datasets, respectively. However, it consumes more computational cost as its complexity is  $O(n^2 \log_2 n)$  [1]. From Table 1, it can be observed that IKS1 and IKS2 methods eliminate almost similar amount of keypoints (41% and 44%) for Blue histology dataset. However, for ADL dataset, the reduction rate of IKS1 (74%) is higher than IKS2. As far as time complexities are a concern, both the methods take lesser time than IB3. However, the time complexity of IKS2 is  $O(n \log_k n)$  which is better than IKS1 whose complexity is  $O(n^2)$ , where  $k$  is the number of clusters. As compared to the algorithms mentioned above, the new GKS method shows the best reduction rate along with an efficient computational cost. The GKS method eliminates 95% and 68% keypoints from ADL and Blue histology datasets respectively. The time complexity of the GKS method is similar to IKS2 i.e.,  $O(n \log_k n)$ . However, the GKS method uses approximate K-means and GRA which take lesser time than K-means and Euclidean distance similarity measure, used by IKS2, respectively. This difference can be visualized from the average time consumed, as mentioned in Table 1.

### Classification Results of the GKS-based BOF Method

In this section, the efficiency of GKS for keypoints selection is validated through the BOF method for classifying the histopathological images. For the classification task, 30 images per category are randomly selected for the training set and the remaining images in that category are used for the validation set.

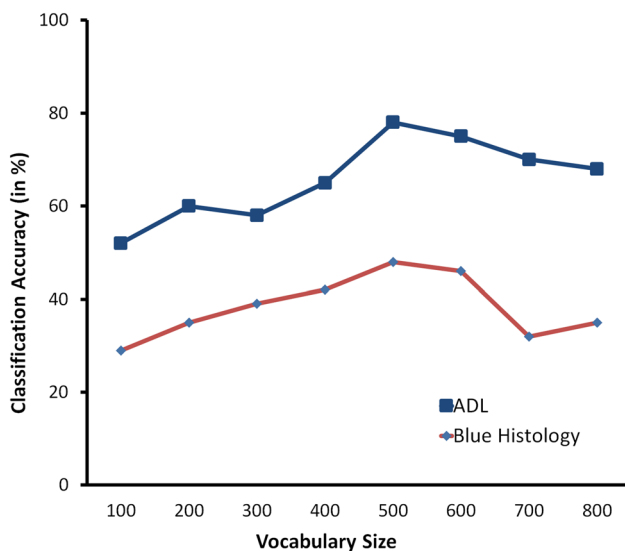
In BOF method, after keypoint selection, codebook construction phase is applied to find the visual words. The size of the codebook is very important for the classification performance. If large-sized codebooks generate flattened histograms which result in less classification accuracy. Similarly, small sized codebook is responsible for biased histograms for which the classifier may not generate good results. Figure 8 shows the classification accuracy for different codebook sizes starting from 100 to 800. It can be visualized that for codebook (or vocabulary) size 500 both of the considered datasets return higher accuracy. Therefore, the codebook size is set to 500 for visual word generation.

**Table 1** Number of selected keypoints along with their average computational cost

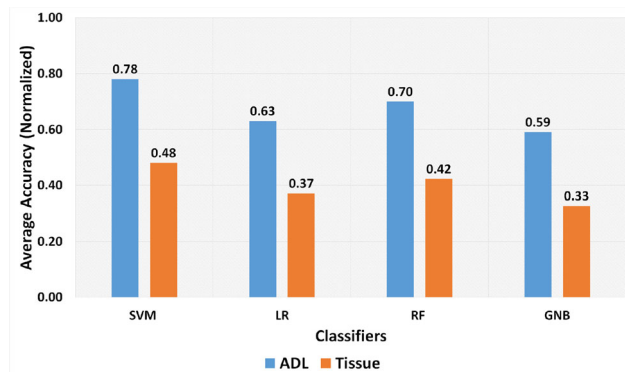
Methods	Number of keypoints		Average computational time (in hours)	
	ADL	Blue histology	ADL	Blue histology
NKS*	4177920	158720		
IB3	626688 (85%)	57139 (64%)	85.71	15
IKS1	1100000 (74%)	65948 (59%)	40	7
IKS2	2003000 (52%)	70312 (56%)	8	3.5
GKS	<b>203000 (95%)</b>	<b>51000 (68%)</b>	<b>4</b>	<b>1</b>

\*Non Keypoint selection





**Fig. 8** Classification accuracy of the GKS-based BOF method with different codebook sizes



**Fig. 9** Classification accuracy on validation set using the GKS method with different classifier on ADL and Blue histology datasets

Moreover, the performance of the GKS-based BOF method is analyzed using four different classifiers, namely support vector machine (SVM), logistic regression (LR), random forest (RF), and Gaussian naive Bayes (GNN) classifiers. Figure 9 shows the classification accuracy returned by the proposed method with different classifiers on ADL and Blue histology datasets. From the figure, it can be visualized that the proposed method performs better when the SVM classifier is used. Hence, for further analysis, SVM is used as the classifier in the proposed BOF method.

For the classification of images using histograms, the SVM classifier using error correcting codes (ECOC) [4] is used. ECOC is an efficient method to handle multi-class classification problems and is based on aggregating the binary classifiers. Each considered binary classifier is independent. Efficient selection of kernel function is also desirable for better classification results. In this paper, the  $\chi^2$ -kernel function is used instead of linear-kernel function due to its higher per-

formance [20]. Moreover, 10 fold cross-validation is used to prevent the over-fitting problem. The random search is also used for hyperparameter tuning which uses uniformly distributed random values and finds the optimal combination in the parameter space.

Figures 10 and 11 show the confusion matrices, generated by each considered method over ADL and Blue histology datasets, respectively. The confusion matrices for ADL dataset show that IB3-based BOF method does not perform well on any of the classes, although it eliminates a significant amount of keypoints as shown in Table 1. That means, it does not select the prominent keypoints. The performance of both the IKS1 and IKS2 methods is far better than IB3 for the ADL dataset. However, IKS2 is somewhat more reliable than IKS1 for recognizing kidney inflamed (KI) class and spleen normal (SN) class images. The performance of the new GKS-based BOF method is enormous in identifying the inflamed images of all the classes more accurately. Likewise, Fig. 11 shows the confusion matrices for the Blue histology methods, returned by IB3, IKS1, IKS2, and GKS based BOF methods. It can be seen from the figure that a classification accuracy of 75% is returned by the GKS-based BOF method for connective and muscle tissues which is better as compared to other methods. For epithelial tissue, IKS1 shows slightly better classification accuracy than the new method. For nervous tissue, IKS1 and IKS2-based methods outperform GKS and classify the images with equal accuracy. Similar to ADL dataset, IB3 does not perform well for the Blue histology dataset too.

To analyze the results of confusion matrices quantitatively, recall, precision, F1-measure, specificity, and average accuracies are measured and depicted in Tables 2 and 3 for ADL and Blue histology datasets respectively.

From Table 2, it can be stated that GKS outperforms the other methods for almost all the parameters. Furthermore, the average classification accuracy of GKS on ADL dataset is 78% which is higher than other considered state-of-the-art methods, i.e., IB3, IKS1, and IKS2 which give 27%, 68%, and 69% accuracy, respectively. Likewise, the new method also shows the best performance for all the tissue classes of Blue histology dataset with F1-measures equals to 65%, 50%, and 42% for muscle, epithelial, and connective, respectively, except nervous tissue where IKS2 shows better results. Moreover, the overall accuracy of the new method for the Blue histology dataset is 48% while IB3, IKS1, and IKS2 return 17%, 36%, and 43% accuracy respectively. However, the accuracy on the Blue histology dataset is not up to the mark due to the lots of staining variations available in the images of the Blue histology dataset as depicted in Table 3. Especially, in the nervous tissue, LFC staining images are very much different from nervous tissue images. Therefore, its performance is degraded in all the methods.

**Fig. 10** The confusion matrices for the ADL dataset, generated by **a** IB3, **b** IKS1, **c** IKS2, and **d** GKS-based classification methods. Here, *KI* Kidney-Inflamed, *KN* Kidney-Normal, *LI* Lung-Inflamed, *LN* Lung-Normal, *SI* Spleen-Inflamed, and *SN* Spleen-Normal

IB3		Predicted					
		KI	KN	LI	LN	SI	SN
Actual	KI	0.35	0.15	0.05	0	0	0.5
	KN	0.8	0.1	0	0	0	0.25
	LI	0.1	0.25	0.4	0.15	0	0
	LN	0	0.55	0.1	0.5	0.2	0
	SI	0	0.1	0	0.5	0.1	0.65
	SN	0	0	0	0	0.05	0.2

(a) IB3

IKS1		Predicted					
		KI	KN	LI	LN	SI	SN
Actual	KI	0.6	0.35	0	0.05	0	0
	KN	0.15	0.6	0	0.1	0.05	0.1
	LI	0.05	0.15	0.4	0.3	0.1	0
	LN	0	0	0	1	0	0
	SI	0	0	0	0	0.6	0.4
	SN	0	0.05	0.05	0	0.05	0.85

(b) IKS1

IKS2		Predicted					
		KI	KN	LI	LN	SI	SN
Actual	KI	0.8	0.15	0.05	0	0	0
	KN	0.35	0.65	0	0	0	0
	LI	0.1	0.25	0.5	0.15	0	0
	LN	0	0	0	1	0	0
	SI	0	0.1	0	0	0.25	0.65
	SN	0	0	0	0	0.05	0.95

(c) IKS2

GKS		Predicted					
		KI	KN	LI	LN	SI	SN
Actual	KI	0.85	0.15	0	0	0	0
	KN	0.05	0.8	0	0	0.15	0
	LI	0.2	0.05	0.5	0.1	0.1	0.05
	LN	0	0	0	1	0	0
	SI	0	0.1	0	0	0.7	0.2
	SN	0	0.05	0	0	0.1	0.85

(d) GKS

**Fig. 11** The confusion matrices for the Blue histology dataset, generated by **a** IB3, **b** IKS1, **c** IKS2, and **d** GKS-based classification methods

IB3		Predicted			
		CT	ET	MT	NT
Actual	CT	0.25	0.3	0.4	0.05
	ET	0.25	0.2	0.35	0.2
	MT	0.35	0.1	0.1	0.25
	NT	0.4	0.3	0.3	0.1

(a) IB3

IKS2		Predicted			
		CT	ET	MT	NT
Actual	CT	0.5	0.2	0.3	0
	ET	0.35	0.3	0.25	0.1
	MT	0.15	0.05	0.7	0.1
	NT	0.5	0.1	0.2	0.2

(b) IKS2

IKS1		Predicted			
		CT	ET	MT	NT
Actual	CT	0.45	0.1	0.4	0.05
	ET	0.05	0.5	0.35	0.1
	MT	0.35	0.1	0.3	0.25
	NT	0.4	0.3	0.1	0.2

(c) IKS1

GKS		Predicted			
		CT	ET	MT	NT
Actual	CT	0.75	0.1	0.15	0
	ET	0.45	0.35	0.2	0
	MT	0.2	0.05	0.75	0
	NT	0.6	0.15	0.2	0.05

(d) GKS

**Table 2** Comparative analysis of the new GKS-based BOF method with other considered methods for ADL dataset

Category	Parameters	IB3	IKS1	IKS2	GKS
Kidney inflammation	Recall	33	60	80	<b>85</b>
	Precision	28	75	64	<b>77</b>
	F-measure	30	67	71	<b>81</b>
Kidney normal	Specificity	82	<b>96</b>	91	95
	Recall	9	60	65	<b>80</b>
	Precision	9	52	57	<b>70</b>
Lung inflammation	F-measure	9	56	60	<b>74</b>
	Specificity	79	89	90	<b>93</b>
	Recall	44	40	50	<b>50</b>
Lung normal	Precision	73	89	91	<b>100</b>
	F-measure	55	55	65	<b>67</b>
	Specificity	97	99	99	<b>100</b>
Spleen inflammation	Recall	37	99	99	<b>100</b>
	Precision	43	69	87	<b>91</b>
	F-measure	40	82	93	<b>95</b>
Spleen normal	Specificity	86	91	97	<b>98</b>
	Recall	7	60	25	<b>70</b>
	Precision	29	75	83	<b>67</b>
Spleen inflammation	F-measure	12	67	38	<b>68</b>
	Specificity	95	96	<b>99</b>	93
	Recall	80	85	<b>95</b>	85
Spleen normal	Precision	13	63	59	<b>77</b>
	F-measure	22	72	73	<b>81</b>
	Specificity	76	90	87	<b>95</b>
Average Accuracy (%)		27	68	69	<b>78</b>

**Table 3** Comparative analysis of the new GKS-based BOF method with other considered methods on blue histology tissue image dataset

Category	Parameters	IB3	IKS1	IKS2	GKS
Muscle Tissue	Recall	13	30	70	<b>75</b>
	Precision	9	26	48	<b>58</b>
	F-measure	10	28	57	<b>65</b>
Connective Tissue	Specificity	66	72	75	<b>82</b>
	Recall	25	45	50	<b>75</b>
	Precision	20	36	33	<b>38</b>
Epithelial Tissue	F-measure	22	40	40	<b>50</b>
	Specificity	66	<b>73</b>	67	58
	Recall	20	34	30	<b>35</b>
Nervous Tissue	Precision	22	50	46	<b>54</b>
	F-measure	21	40	36	<b>42</b>
	Specificity	75	83	88	<b>90</b>
Nervous Tissue	Recall	9	18	<b>20</b>	5
	Precision	17	33	50	<b>100</b>
	F-measure	12	25	<b>29</b>	10
Average Accuracy (%)		17	36	43	<b>48</b>

From the results, it can be stated that the classification accuracy of the GKS-based BOF method is better than the other considered methods. The baseline algorithm (IB3) gives poor performance in all scenarios as it filters out a large number of keypoints including the relevant ones. This reduces the size but also degrades the classification performance. IKS2 performs better than IKS1 as it starts with multiple reference points together and applies the reduction phase cluster-wise to reduce the overall training set. Therefore, IKS2 is fast and efficient than IKS1. In the new GKS method, the use of Grey relational analysis-based similarity measure and approximate K-means make it faster and efficient. As the number of keypoints is reduced, the number of visual words is also reduced in the GKS-based BOF method.

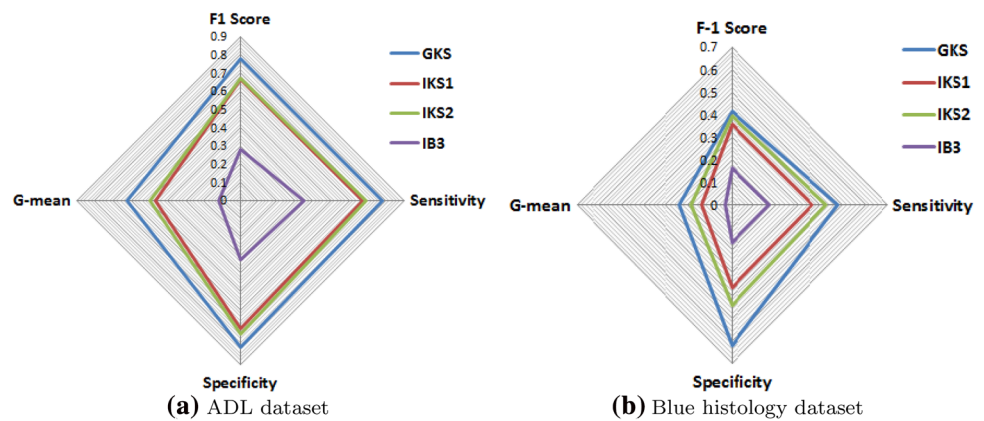
However, accuracy may not only be the suitable criteria to measure the performance if images with normal class label are lesser than images with inflamed class label. Let us consider there are 100 images in which there are 95 normal images and 5 inflamed images. If any classification method correctly identified all inflamed images, then it returns 95% of accuracy. However, the method does not recognize normal cases at all. Therefore, a metric is required which considers both true positives (TP) and true negative (TN) cases. G-mean is a metric ( $\sqrt{TP * TN}$ ) which considers the both TN and TP. Furthermore, the performance of IB3, IKS1, IKS2, and GKS methods are also analyzed using radar charts which are shown in Fig. 12 that depict four evaluation criteria, namely F1 score, sensitivity, specificity, and G-mean which resulting in four-sided shape. The method with a maximum area and symmetrical shape perform better than others. From the figure, it can be observed that the GKS-based BOF method achieves better results among all the other methods in four considered measures. Therefore, it can be stated that the new keypoints selection in the BOF method outperforms the other keypoints selection methods and may be applied for histopathological image classification.

**Comparative analysis of GKS-based BOF with state-of-the-art methods**

The performance of the GKS-based BOF method is also compared with the three state-of-the-art methods for ADL histopathological image dataset, namely WND-CHRM [45], SRC [50], and SHIRC [49] in terms of recall, specificity, precision, false negative rate (FNR), average accuracy, and F1-score. Shamir et al. [45] introduced a method for the analysis of biological images in which image content features are detected from the raw images and selected informative feature descriptors are used to train the classifier. In the **sparse representation-based classification** (SRC) method [50], RGB images are represented by a single luminance channel and this representation is used to train the classifier. Moreover, this work is further extended to three color



**Fig. 12** Radar charts for average results obtained for SVM classifier on **a** ADL dataset and **b** Blue histology image dataset by considering F-1 score, sensitivity, specificity, and G-mean



**Table 4** Classification performance of GKS-based BOF method with other state-of-the-art methods

Organ	Algorithms	Recall	Specificity	Precision	FNR	F1-score	Avg. accuracy (%)
Kidney	WND-CHARM	0.690	0.720	0.710	0.280	0.700	71.0
	SRC	0.875	0.750	0.778	0.250	0.825	81.3
	SHIRC	0.825	0.833	0.832	0.167	0.828	82.9
	BOF	0.870	0.650	0.731	0.350	0.826	80.0
	GKS	<b>0.950</b>	<b>0.890</b>	<b>0.888</b>	<b>0.110</b>	<b>0.879</b>	<b>88.0</b>
Lung	WND-CHARM	0.725	0.626	0.705	0.374	0.791	75.7
	SRC	0.880	0.765	0.750	0.235	0.737	74.5
	SHIRC	0.750	0.850	0.833	0.150	0.791	80.0
	BOF	0.730	0.750	0.745	0.250	0.737	74.0
	GKS	<b>0.888</b>	<b>0.860</b>	<b>0.863</b>	<b>0.140</b>	<b>0.871</b>	<b>87.0</b>
Spleen	WND-CHARM	0.512	0.873	0.800	0.128	0.640	69.2
	SRC	0.708	0.792	0.773	0.208	0.740	75.0
	SHIRC	0.650	0.883	0.848	<b>0.117</b>	0.742	76.7
	BOF	0.880	0.530	0.652	0.470	0.749	70.5
	GKS	<b>0.750</b>	<b>0.880</b>	<b>0.862</b>	0.120	<b>0.804</b>	<b>81.5</b>

channels and known as a **multi-channel simultaneous sparsity model** (SHIRC) [49]. This method is also analyzed and validated on ADL histopathological images. Table 4 shows the results of each considered method on the various performance parameters, namely recall, specificity, precision, false negative rate (FNR), and F1 score to identify the inflamed images of each organ in ADL dataset.

Recall and specificity are the two key statistics to validate the performance of classification in medical diagnosis. Recall is the probability to identify diseased images correctly, while specificity returns the probability of identifying the healthy images correctly. In histopathological image analysis, it is always important to identify inflamed images with higher accuracy. From Table 4, it can be noticed that the new GKS method has high recall values of 95%, 88.8%, and 75% for Kidney, Lung, and Spleen organs respectively. Moreover, the true negative rates returned by the GKS method are 89%, 86%, and 88% for Kidney, Lung, and Spleen organs, respectively. Hence, it can be stated that the GKS method also

identifies healthy images more accurately as compared to the other considered methods. Furthermore, the GKS-based BOF method also attains high average accuracy, precision, and F-1 score. The results have also been analyzed on the FNR which can be defined as the rate of identifying inflamed images as healthy images. It is very dangerous in medical diagnosis and it should be minimized. The GKS method has the lowest FNR of 11% and 14% on Kidney and Lung organ images respectively. However, for Spleen organ images, SHIRC outperforms the GKS method in terms of FNR.

### Comparative analysis of GKS-based BOF with deep learning-based methods

In recent years, it has been observed that deep learning models are performing very well in case of image classification. These methods are commonly known as convolutional neural networks (CNNs). However, various articles have depicted that these CNN models do not perform well for

histopathological images due to limited training set. To verify the same, the proposed method is also being compared with two CNN-based methods, proposed by Bayramoglu et al. [7], CNN-IBBO-BOF [37] and AlexNet [23]. Bayramoglu et al. [7] proposed a CNN model whose architecture consists of three convolutional layers and two fully connected layers. After each convolution layer, a rectified linear unit (ReLU) and a max-pooling layer with filter size  $3 \times 3$  and stride size two, are encountered. The first convolutional layer uses 96 filters of size  $3 \times 7 \times 7$ . The second and third convolutional layers contain 256 filters of size  $5 \times 5$  and 384 filters of size  $3 \times 3$ , respectively. At the output end, two fully-connected layers are used with 512 neurons along with a dropout layer. Furthermore, in the CNN-IBBO-BOF method, a pre-trained CNN model, known as AlexNet [23], is used to extract the features from histopathological images. These features are used by the IBBO-based BOF method for the classification of histopathological images.

The above-discussed deep learning-based methods are applied to the considered histopathological image datasets, namely ADL and Blue histology. As the number of images are very less in these datasets, the transfer learning approach is used. For classification 10-fold cross validation approach is applied. Table 5 shows the comparison of all the proposed methods and mentioned deep learning-based methods over considered histopathological datasets. From the table, it can be observed that the method of Bayramoglu et al. [7] returns 52.72% and 28.12% accuracy for the ADL and Blue histology datasets respectively. Similarly, AlexNet returns 51.30% and 29.68% accuracy for the ADL and Blue histology datasets respectively. On the other hand, CNN-IBBO-BOF method gives the accuracy of 79.66% and 52% for the ADL and Blue histology datasets respectively. The major difference between the accuracy of these two deep learning based methods are due to the use of pre-trained CNN in CNN-IBBO-BOF method while Bayramoglu et al. [7] is trained with the available datasets. This signifies the requirement of large dataset for deep learning-based models. Furthermore, the GRA-based keypoint selection method enhances the performance of the BOF methods and works well for small datasets also. This validates that the proposed system outperforms the existing methods for histopathological image classification.

## Conclusion

In this paper, a new method of keypoints selection has been proposed which improves the efficiency of the bag-of-features method. The method uses Grey relational analysis and approximate k-means for the elimination of irrelevant

**Table 5** The performance comparison of the GKS-BOF method with deep-learning-based image classification methods

S. No.	BOF approach	ADL dataset	Blue Histology dataset
1.	GKS-BOF	78	48
2.	Bayramoglu et al. [7]	52.72	28.12
3.	CNN-IBBO-BOF [39]	75.66	43
4.	AlexNet [23]	51.30	29.68

and similar keypoints. Furthermore, the proposed keypoint selection method has been incorporated in the BOF method to reduce the computational complexity of its codebook construction phase. Moreover, the support vector machine with error correcting output code is used to train and classify the images. The proposed method is tested on two histopathological image datasets, namely ADL and Blue histology. The GKS method reduces the extracted high-dimensional keypoint descriptors by 95% and 68% from the ADL and Blue histology datasets, respectively. Moreover, the GKS-based BOF method increases the respected classification accuracy by 13% and 11% from IKS2-based BOF method. Moreover, the GKS-based BOF method also outperforms transfer learning-based considered deep learning models.

However, the following issues can be considered for future research. First, the optimal value of the shrinking threshold can be computed by the use of meta-heuristic methods to enhance the selection rate. Second, data augmentation methods can be used to increase the training sample for better training of deep learning methods. Finally, the proposed method can be analyzed and tested on non-medical image datasets.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Aha DW, Kibler D, Albert MK (1991) Instance-based learning algorithms. *Mach Learn* 6(1):37–66

2. Alahi A, Ortiz R, Vandergheynst P (2012) Freak: Fast retina keypoint. Computer vision and pattern recognition (CVPR). IEEE conference on, Ieee, pp 510–517
3. Albiol A, Monzo D, Martin A, Sastre J, Albiol A (2008) Face recognition using hog-ebgm. *Pattern Recogn Lett* 29(10):1537–1543
4. Ali Bagheri M, Montazer GA, Escalera S (2012) Error correcting output codes for multiclass classification: application to two image vision problems. In: *Artificial Intelligence and Signal Processing (AISP)*, 2012 16th CSI International Symposium on, IEEE, pp 508–513
5. Basavanthally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, Bhanot G, Madabhushi A (2009) Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology. *IEEE Trans Biomed Eng* 57(3):642–653
6. Bay H, Tuytelaars T, Van Gool L (2006) Surf: Speeded up robust features. In: *European conference on computer vision*, Springer, pp 404–417
7. Bayramoglu N, Kannala J, Heikkilä J (2016) Deep learning for magnification independent breast cancer histopathology image classification. In: *Proceedings of international conference on pattern recognition*, Cancun, Mexico, pp 2440–2445
8. Brighton H, Mellish C (2002) Advances in instance selection for instance-based learning algorithms. *Data Min Knowl Disc* 6(2):153–172
9. Caicedo JC, Cruz A, Gonzalez FA (2009) Histopathology image classification using bag of features and kernel functions. In: *Conference on artificial intelligence in medicine in Europe*, Springer, pp 126–135
10. Chang K, Lee R, Wen C, Yeh M (2005) Comparison of similarity measures for clustering electrocardiogram complexes. In: *Computers in cardiology*, 2005, IEEE, pp 759–762
11. Cruz-Roa A, Caicedo JC, González FA (2011) Visual pattern mining in histology image collections using bag of features. *Artif Intell Med* 52(2):91–106
12. Díaz G, Romero E (2010) Histopathological image classification using stain component features on a pls model. In: *Iberoamerican congress on pattern recognition*, Springer, pp 55–62
13. Dorkó G, Schmid C (2003) Selection of scale-invariant parts for object class recognition. In: *Proceedings of international conference on computer vision*, Beijing, China, pp 634–640
14. Eurell JA, Frappier BL (2013) *Dellmann's textbook of veterinary histology*. Wiley, Amsterdam
15. Fondón I, Sarmiento A, García AI, Silvestre M, Eloy C, Polónia A, Aguiar P (2018) Automatic classification of tissue malignancy for breast carcinoma diagnosis. *Comput Biol Med* 96:41–51
16. Gangeh MJ, Sørensen L, Shaker SB, Kamel MS, De Bruijne M, Loog M (2010) A texton-based approach for the classification of lung parenchyma in ct images. In: *International conference on medical image computing and computer-assisted intervention*, Springer, pp 595–602
17. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B (2009) Histopathological image analysis: a review. *IEEE Rev Biomed Eng* 2:147–171
18. Gutiérrez R, Rueda A, Romero E (2013) Learning semantic histopathological representation for basal cell carcinoma classification. In: *Medical imaging 2013: digital pathology*, international society for optics and photonics, vol 8676, p 86760U
19. *Histology B* (2017) <http://www.lab.anhb.uwa.edu.au/mb140/>
20. Jiang YG, Yang J, Ngo CW, Hauptmann AG (2010) Representations of keypoint-based semantic concept detection: a comprehensive study. *IEEE Trans Multimed* 12(1):42–53
21. Juan L, Gwon O (2009) A comparison of sift, pca-sift and surf. *Int J Image Process (IJIP)* 3(4):143–152
22. Julong D (1989) Introduction to grey system theory. *J Grey Syst* 1(1):1–24
23. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp 1097–1105
24. Kumar MD, Babaie M, Zhu S, Kalra S, Tizhoosh HR (2017) A comparative study of cnn, bow and lbp for classification of histopathological images. *Computational Intelligence (SSCI). IEEE Symposium Series on, IEEE*, pp 1–7
25. Li T, Mei T, Kweon IS, Hua XS (2011) Contextual bag-of-words for visual categorization. *IEEE Trans Circ Syst Video Technol* 21(4):381–392
26. Lichtblau D, Stoean C (2019) Cancer diagnosis through a tandem of classifiers for digitized histopathological slides. *PLoS One* 14(1):e0209274
27. Lin WC, Tsai CF, Chen ZY, Ke SW (2016) Keypoint selection for efficient bag-of-words feature generation and effective image classification. *Inf Sci* 329:33–51
28. Mittal H, Saraswat M (2019) Classification of histopathological images through bag-of-visual-words and gravitational search algorithm. In: *Soft computing for problem solving*, Springer, pp 231–241
29. Monga V (2018) Adl data set. <http://signal.ee.psu.edu/histimg2.html>
30. Morel JM, Yu G (2011) Is sift scale invariant? *Inverse Probl Image* 5(1):115–136
31. Mousavi HS, Monga V, Rao G, Rao AU (2015) Automated discrimination of lower and higher grade gliomas based on histopathological image analysis. *J Pathol Inf* 6
32. Nayak N, Chang H, Borowsky A, Spellman P, Parvin B (2013) Classification of tumor histopathology via sparse feature learning. In: *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on, IEEE*, pp 410–413
33. Orlov N, Shamir L, Macura T, Johnston J, Eckley DM, Goldberg IG (2008) Wnd-charm: multi-purpose image classification using compound image transforms. *Pattern Recognit Lett* 29(11):1684–1693
34. Pal R, Saraswat M (2017a) Data clustering using enhanced biogeography-based optimization. In: *2017 Tenth international conference on contemporary computing (IC3)*, IEEE. <https://doi.org/10.1109/ic3.2017.8284305>
35. Pal R, Saraswat M (2017b) Improved biogeography-based optimization. *International Journal of Advanced Intelligence Paradigms*
36. Pal R, Saraswat M (2018a) Enhanced bag of features using alexnet and improved biogeography-based optimization for histopathological image analysis. In: *2018 Eleventh International Conference on Contemporary Computing (IC3)*, IEEE, pp 1–6
37. Pal R, Saraswat M (2018b) Enhanced bag of features using alexnet and improved biogeography-based optimization for histopathological image analysis. In: *Proceedings of Eleventh international conference on contemporary computing*, Noida, India, IEEE, pp 1–6
38. Pal R, Saraswat M (2018c) A new bag-of-features method using biogeography-based optimization for categorization of histology images. *Int J Inf Syst Manag Sci* 1(2)
39. Pal R, Saraswat M (2019a) Grey relational analysis based keypoint selection in bag-of-features for histopathological image classification. *Recent Patents Comput Sci* 12:1–9
40. Pal R, Saraswat M (2019b) Histopathological image classification using enhanced bag-of-feature with spiral biogeography-based optimization. *Appl Intell* pp 1–19
41. Raza SH, Parry RM, Sharma Y, Chaudry Q, Moffitt RA, Young A, Wang MD (2010) Automated classification of renal cell carcinoma subtypes using bag-of-features. In: *Engineering in medicine and*



- biology society (EMBC), 2010 annual international conference of the IEEE, IEEE, pp 6749–6752
42. Raza SH, Parry RM, Moffitt RA, Young AN, Wang MD (2011) An analysis of scale and rotation invariance in the bag-of-features method for histopathological image classification. In: International conference on medical image computing and computer-assisted intervention, Springer, pp 66–74
  43. Sallehuddin R, Shamsuddin SMH, Hashim SZM (2008) Application of grey relational analysis for multivariate time series. In: Intelligent systems design and applications, 2008. ISDA'08. Eighth international conference on, IEEE, vol 2, pp 432–437
  44. Saraswat M, Arya K (2014) Automated microscopic image analysis for leukocytes identification: a survey. *Micron* 65:20–33
  45. Shamir L, Orlov N, Eckley DM, Macura T, Johnston J, Goldberg IG (2008) Wndchrm-an open source utility for biological image analysis. *Source Code Biol Med* 3(1):13
  46. Sirinukunwattana K, Raza SEA, Tsang YW, Snead DR, Cree IA, Rajpoot NM (2016) Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans Med Image* 35:1196–1206
  47. Song Y, Cai W, Huang H, Zhou Y, Feng DD, Wang Y, Fulham MJ, Chen M (2015a) Large margin local estimate with applications to medical image classification. *IEEE Trans Med Image* 34(6):1362–1377
  48. Song Y, Cai W, Huang H, Zhou Y, Wang Y, Feng DD (2015b) Locality-constrained subcluster representation ensemble for lung image classification. *Med Image Anal* 22(1):102–113
  49. Srinivas U, Mousavi H, Jeon C, Monga V, Hattel A, Jayarao B (2013) Shirc: A simultaneous sparsity model for histopathological image representation and classification. In: Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on, IEEE, pp 1118–1121
  50. Srinivas U, Mousavi HS, Monga V, Hattel A, Jayarao B (2014) Simultaneous sparsity model for histopathological image representation and classification. *IEEE Trans Med Image* 33(5):1163–1179
  51. Tang HL, Hanka R, Ip HHS (2003) Histological image retrieval based on semantic content analysis. *IEEE Trans Inf Technol Biomed* 7(1):26–36
  52. Vu TH, Mousavi HS, Monga V, Rao G, Rao UA (2016) Histopathological image classification using discriminative feature-oriented dictionary learning. *IEEE Trans Med Image* 35(3):738–751
  53. Wang C, Chen SF, Yuen MMF (2001) Fuzzy part family formation based on grey relational analysis. *Int J Adv Manuf Technol* 18(2):128–132
  54. Wang J, Wang J, Ke Q, Zeng G, Li S (2015) Fast approximate  $k$   $k$ -means via cluster closures. In: Multimedia data mining and analytics, Springer, pp 373–395
  55. Xu ZG, Chen C, Liu XH (2013) An efficient view-point invariant detector and descriptor. *Adv Mater Res Trans Tech Publ* 659:143–148
  56. Zhang X, Liu W, Dundar M, Badve S, Zhang S (2014) Towards large-scale histopathological image analysis: hashing-based image retrieval. *IEEE Trans Med Image* 34(2):496–506
  57. Zheng Y, Jiang Z, Xie F, Zhang H, Ma Y, Shi H, Zhao Y (2017) Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification. *Pattern Recogn* 71:14–25

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.