

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/298797075>

Medical image classification using spatial adjacent histogram based on adaptive local binary patterns

Article *in* Computers in Biology and Medicine · March 2016

DOI: 10.1016/j.combiomed.2016.03.010

CITATION

1

READS

92

6 authors, including:



[Huiling Chen](#)

Wenzhou University

67 PUBLICATIONS 765 CITATIONS

[SEE PROFILE](#)



Medical image classification using spatial adjacent histogram based on adaptive local binary patterns

Dong Liu ^{a,b}, Shengsheng Wang ^{a,b,*}, Dezhong Huang ^{a,b}, Gang Deng ^c, Fantao Zeng ^{a,b},
Huiling Chen ^d

^a Jilin University, College of Computer Science and Technology, Changchun 130012, China

^b Jilin University, Key Laboratory of Symbolic Computation and Knowledge Engineering of the Ministry of Education, Changchun 130012, China

^c Renmin Hospital of Wuhan University, Department of Neurosurgery, Wuhan 430060, China

^d College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China



ARTICLE INFO

Article history:

Received 2 September 2015

Received in revised form

16 March 2016

Accepted 16 March 2016

Keywords:

Local binary patterns

Image classification

Feature extraction

Medical images

Microscope images

ABSTRACT

Medical image recognition is an important task in both computer vision and computational biology. In the field of medical image classification, representing an image based on local binary patterns (LBP) descriptor has become popular. However, most existing LBP-based methods encode the binary patterns in a fixed neighborhood radius and ignore the spatial relationships among local patterns. The ignoring of the spatial relationships in the LBP will cause a poor performance in the process of capturing discriminative features for complex samples, such as medical images obtained by microscope. To address this problem, in this paper we propose a novel method to improve local binary patterns by assigning an adaptive neighborhood radius for each pixel. Based on these adaptive local binary patterns, we further propose a spatial adjacent histogram strategy to encode the micro-structures for image representation. An extensive set of evaluations are performed on four medical datasets which show that the proposed method significantly improves standard LBP and compares favorably with several other prevailing approaches.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Medical images have played an important role in the diagnostic workup of patients. Automated classification of medical images is a desirable tool to assign the interpretation of images, and then would help the expert in diagnosis of diseases [1–3]. Compared with general image recognition, medical image recognition is more challenging because of the higher ambiguity and complexity; most of the medical image contents are quite similar, but also different in their emphasis.

In terms of the features used for medical image recognition, it can be mainly classified into three groups: shape, color, and texture features. For example, in [4], shape features such as moment invariants and Fourier descriptor are employed to classify medical X-ray images. A color vector field is considered in [5] for improving the performance of endoscopic image classification.

The local binary patterns, first proposed by [6], are widely considered as a state-of-the-art image feature descriptor among

texture descriptors, since it can more effectively describe texture information. It has been successfully applied to many applications, such as face recognition, texture classification, scene recognition, human detection and others. LBP has several attractive advantages: it has proven to be a powerful discriminator with low computational cost, it is robust against changes in image intensity, and it can be easily implemented. Due to these merits, it makes a good choice for extracting fine features for medical images.

However, the standard LBP still suffers from several drawbacks, including limited semantic description of local patterns, sensitive to non-uniform patterns and affine transformation, and missing of efficient spatial encoding among patterns. To overcome these shortcomings, numerous works [9–11] focused on improving LBP in recent years, in terms of rotation-invariant, multi-scale, the utilization of non-uniform patterns, and so on. There are two types of LBP patterns: *uniform* and *non-uniform* patterns. Some works, such as [7], only considered uniform patterns for extracting LBP features since non-uniform patterns involve noise and high dimensionality. And the work [8] proposed a hierarchical multi-scale LBP to further utilize the information of non-uniform patterns. They also certify that, the percentage of non-uniform patterns increases as the neighborhood radius increases. To reduce the LBP dimensionality, center-symmetric local binary patterns

* Corresponding author at: Jilin University, College of Computer Science and Technology, Changchun 130012, China.

E-mail address: wss@jlu.edu.cn (S. Wang).

(CSLB) [15] is studied and applied to image recognition. Since LBP is sensitive to noise in near uniform regions, Local Ternary Pattern [14] with three value coding scheme was proposed to address this problem. The rotation invariant [6] descriptor can be obtained through the circular neighborhood definition, but in some cases the anisotropic structural information is lost. To utilize these anisotropic structural information, a novel elliptical binary pattern (EBP) [16] has been proposed for face recognition, in which elliptical neighborhood definitions are studied. Completed LBP [13] utilizes both the sign and magnitude information in the difference between the central pixel and the neighborhood pixels. In the work [18], LBP is combined with Gabor filters to achieve a better classification performance. The study [19] extracted the most frequent patterns in LBP histogram and formed a novel descriptor, which achieved better performance with this technique. Mesh-LBP [21] is novel method which computed the mesh-local binary pattern on a triangular-mesh manifold. In [22], a scale- and rotation-invariant LBP is proposed, in which the rotation-invariant is combined with a scale-adaptive textron for texture classification. SOALBP [23] constructed a novel scale- and orientation invariant LBP feature combined in a multi-resolution representation, which has been proven superior in texture classification.

The basic idea behind LBP is that it describes an image by local patterns. The existing methods have been proven to improve the LBP to some extent by reconfiguring or utilizing the patterns. However, most of existing works encode the binary patterns in a fixed neighborhood radius. This fixed neighborhood radius strategy is irrelevant to local image content and disregards micro-structure information of the multi-scale patterns. Intuitively, the micro-structures, i.e. the spatial relationships among local patterns generated by adaptive radius, provide crucial feedback in disambiguating texture information especially for complex medical images, i.e. microscope images that involve with pathological changes. This subsequently leads to improved recognition performance. To this end, our target is to design a novel LBP histogram representation for medical images to (1) compute the local binary patterns in an adaptive neighborhood radius, and (2) encode micro-structures among the multi-scale patterns. In the first stage, with the help of gradient operators, we obtain a gradient map from each original image, and the adaptive LBP neighborhood radius could be then determined for each pixel by utilizing the gradient information. As a result, our adaptive strategy will assign a relatively small radius to pixels that are located in local regions with dramatic gray variation, while assigning a relatively large radius to pixels that are located in local regions with slight gray variation. This adaptive technique will provide the image with rich micro-structure textures, which is discriminative in image representation. Then in the next stage, we propose a spatial adjacent histogram based on adaptive LBP radius to describe these discriminative micro-structure features. Finally, the adaptive LBP

radius and spatial adjacent histogram strategies produce a much more powerful LBP variant, which performs well in four benchmark medical datasets and compares favorably to other methods.

In this context, our contribution is threefold.

- 1) Using the adaptive strategy we proposed, the neighborhood radius of LBP is determined based on local image content, therefore more adaptive and useful features can be obtained.
- 2) We propose to use spatial adjacent histogram to encode the micro-structures produced by adaptive strategy, which results in convincing improvement on the standard LBP histogram.
- 3) Our approach also considers three LBP coding schemes, i.e. set the threshold T in three different ways when computing the LBP value. And we further evaluate the three LBP coding schemes in order to find which one performs more competitive in medical image classification.

The remainder of this paper is organized as follows. In Section 2, we introduce the proposed algorithm, i.e. spatial adjacent histogram based on adaptive local binary patterns for image classification. Section 3 presents an extensive set of experimental evaluations on four medical image datasets, and finally, in Section 4, we draw the conclusions.

2. Spatial adjacent histogram based on adaptive local binary patterns for image classification

In this section, we propose a novel idea using spatial adjacent histogram based on adaptive local binary patterns for medical image classification. We first present a concise review of standard local binary patterns (LBP) in subsection 2.1. Next, we explain how to determine the adaptive radius for each pixel in subsection 2.2, then in subsection 2.3, three coding schemes are introduced to compute the LBP value for each pixel. In subsection 2.4, we propose a spatial adjacent histogram technique based on adaptive LBP to represent the whole image. Finally, the proposed image classification framework using our spatial adjacent histogram based on adaptive local binary patterns is presented in subsection 2.5.

2.1. Brief review of LBP

Given an image I , the LBP is a gray-scale texture operator that characterizes the local spatial patterns of the image texture, which is calculated at each pixel by evaluating the binary differences between it and its neighbors:

$$LBP(P, R) = \sum_{i=0}^{P-1} s(g_i - T)2^i, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

where P is the number of pixels in the neighborhood, R is the

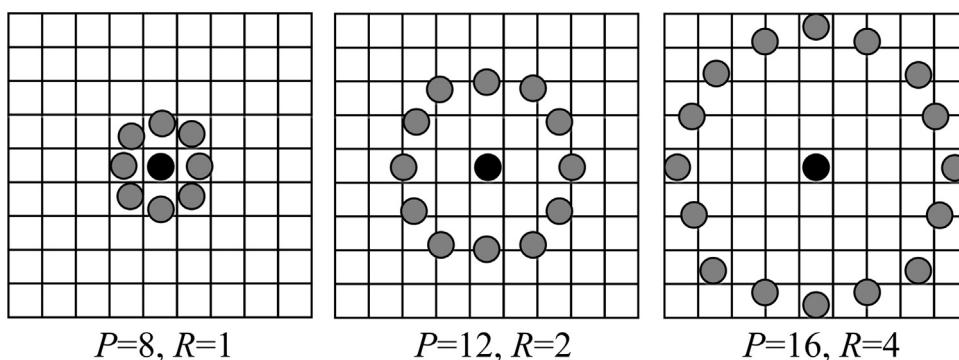


Fig. 1. Three circularly symmetric neighborhood sets for different (P, R) .

radius of the neighborhood and T is a threshold. The original LBP set T as the gray value of the central pixel and g_i is the gray value of its neighborhood. Fig. 1 shows examples of different configurations of (P, R) . A pattern is considered *uniform* if the number of transitions in the sequence between 0 and 1 is less than or equal to two. For instance, LBP pattern 00100000 is *uniform* (2 times transition) and 00101000 is *non-uniform* (4 times transition).

2.2. Determining the adaptive radius R

In order to extract micro-structures of different scales, the key idea behind this paper is to adaptively obtain the LBP radius of each pixel by analyzing the differences based on calculated gradients. Given an image I and we use $f(x,y)$ as the gray value of pixel (x, y) , then the Sobel [20] gradient magnitude $g(x, y)$ can be obtained by Eq. (2) and Eq. (3).

$$\begin{cases} g_x = [f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)] \\ \quad - [f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)] \\ g_y = [f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1)] \\ \quad - [f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)] \end{cases} \quad (2)$$

$$g(x, y) = \sqrt{g_x^2 + g_y^2} \quad (3)$$

Then, given a center pixel g_c and its neighborhood radius k , the average gradients $G_k(x, y)$ in $(2k+1) \times (2k+1)$ image blocks is defined as follows:

$$G_k(x, y) = \frac{\sum_{i=x-k}^{x+k} \sum_{j=y-k}^{y+k} g(i, j)}{(2k+1)^2}, \quad k \in \{1, 2, \dots, R_{\max}\} \quad (4)$$

where R_{\max} is the maximal search radius.

Finally, the gray variation is defined as the intensity difference among the average gradients in different sizes of image blocks. Thus, the neighborhood radius R_{gc} of the central pixel g_c can be obtained according to the maximum value:

$$R_{gc} = \arg \max_k |G_{k+1} - G_k|, \quad k \in \{1, 2, \dots, R_{\max}\} \quad (5)$$

An illustration of the proposed adaptive radius computing process for a central point in a 7×7 blocks is provided in Fig. 2. The input central point and its neighborhood of 7×7 blocks, which are described as gradient magnitude, appears on the left part of Fig. 2. Then the three scales of blocks and the average gradients are defined by Eq. (4). It is due to the fact that the most dramatic change lies between G_2 and G_3 , so the adaptive radius for the central pixel is set as 2, which are presented in the right part of Fig. 2. This means the outermost neighborhood pixels are not considered based on our adaptive radius strategy, because they adopt irrelevant gray value and may be harmful for the robustness.

Three toy examples of adaptive radius map are illustrated in Fig. 3. In each example, the original image, adaptive radius map and the percent of pixels with different radius are arranged from

left to right. In each adaptive radius map, four colors denote different LBP radii for each pixel, i.e. red denotes $R=4$, green denotes $R=3$, blue denotes $R=2$ and black denotes $R=1$. The first and second original images in Fig. 3 are taken from the two different categories (microtubules and endosome) of 2D HeLa dataset [28], respectively, in which the pixels outside the objects are not considered since it represents the black background. It is observed that Fig. 3(a) and Fig. 3(b) have much of different spatial distribution of adaptive radius, and the micro-structures produced by neighbor adaptive radius are much discriminative for two images. Another intuitive instance can be found in Fig. 3(c), in which the original image is a scene image. Therefore, one could conclude that the micro-structures of neighbor adaptive radius are much related to image contents. Furthermore, we could discover that our adaptive strategy assigned a relatively small radius to pixels that are located in local regions with dramatic gray variation, while assigning a relatively large radius to pixels that are located in local regions with slight gray variation. In the rest of this paper, we will show how to utilize these discriminative features coming from micro-structures.

2.3. LBP coding schemes based on adaptive radius

In this subsection, we formulate the process of LBP coding based on adaptive radius R . Given a center pixel g_c , the neighborhood radius R can be computed by Eq. (5). Inspired by state-of-the-art LBP coding strategy, three coding schemes are selected based on setting different thresholds to T of Eq. (1) adaptively, which are named T_1 , T_2 and T_3 for short, respectively. And LBPT_1 , LBPT_2 and LBPT_3 denote standard LBPs obtained by the three coding schemes, respectively.

We start with an original image I , let g_c be a center pixel and g_i denotes its neighborhood pixel. P denotes the total number of the neighbors, and R denotes the adaptive radius of g_c obtained during the process in subsection 2.2.

(1) The first coding scheme T_1

Set threshold T as the gray value of the central pixel, i.e. $T_1 = g_c$, thus, the LBPT₁ can be formulated as

$$\text{LBPT}_1 = \sum_{i=0}^{P-1} s(g_i - T_1) 2^i \quad (6)$$

(2) The second coding scheme T_2

Let $m_i = |g_i - g_c|$, set T_2 as the mean value of m_i , i.e.

$$T_2 = \sum_{i=0}^{P-1} m_i / P = \sum_{i=0}^{P-1} |g_i - g_c| / P \quad (7)$$

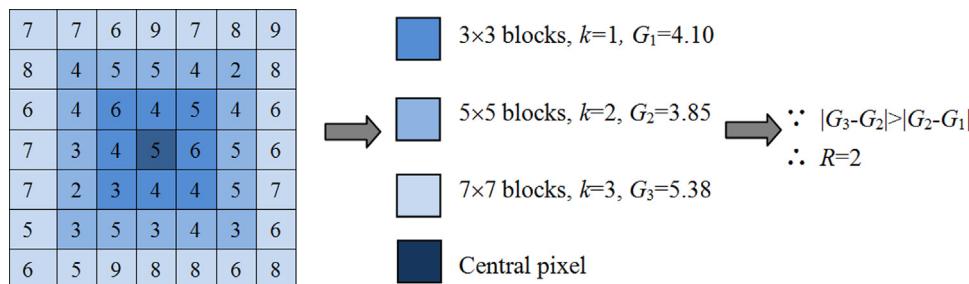


Fig. 2. An example of proposed adaptive radius computing process for determining the radius.

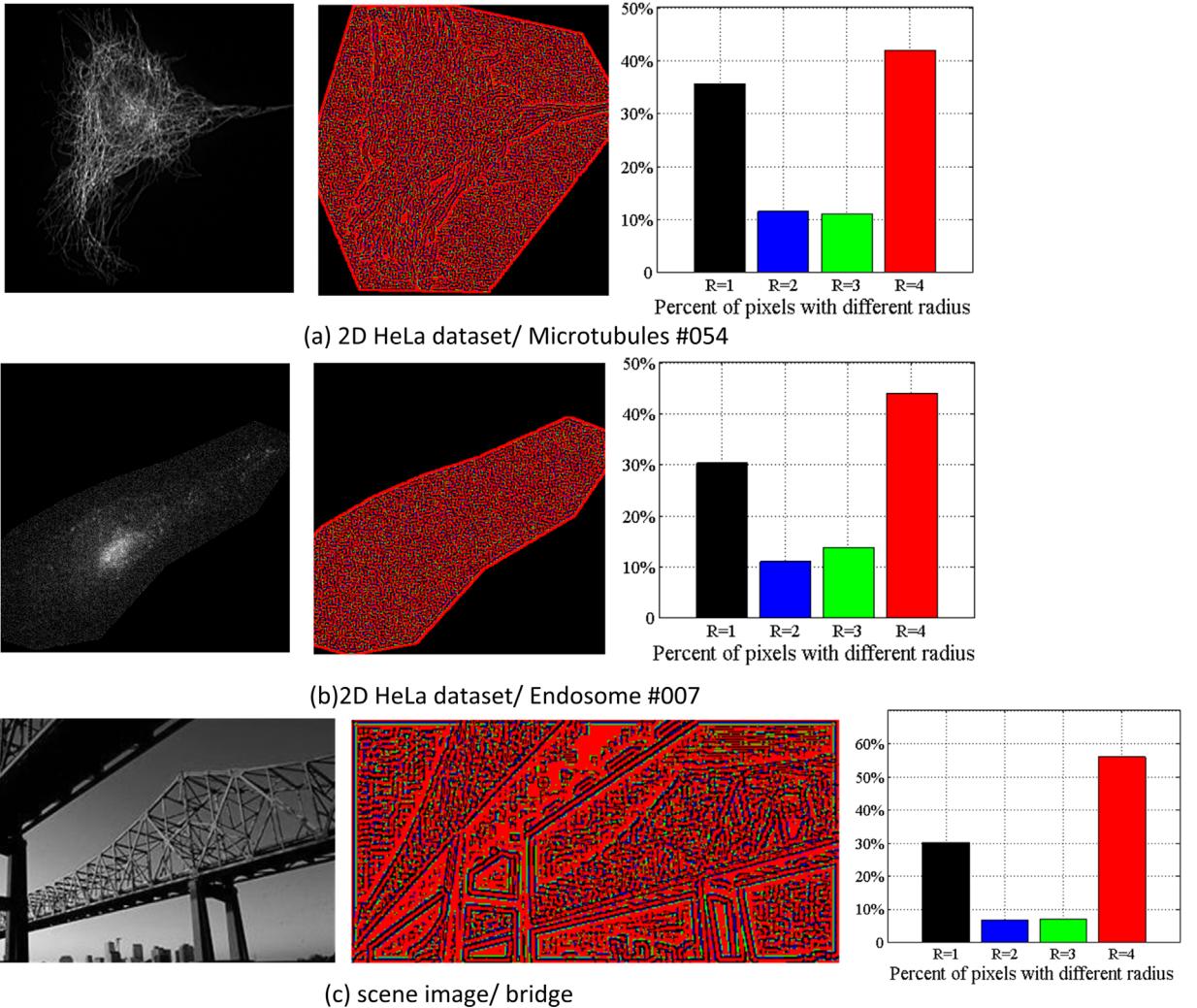


Fig. 3. Toy examples of adaptive radius map, i.e. different radius of each pixel: red denotes $R=4$, green denotes $R=3$, blue denotes $R=2$ and black denotes $R=1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Then, the $LBPT_2$ are presented as follows:

$$LBPT_2 = \sum_{i=0}^{P-1} s(m_i - T_2)2^i = \sum_{i=0}^{P-1} s(|g_i - g_c| - T_2)2^i \quad (8)$$

(3) The third coding scheme T_3

Set T_3 as the mean gray value of the $(2R+1) \times (2R+1)$ image blocks, i.e.

$$T_3 = \frac{1}{N} \sum_{i=1}^N g_i / N, \quad N = (2R+1) \times (2R+1) \quad (9)$$

Then we can define the third $LBPT_3$ as follows

$$LBPT_3 = \sum_{i=0}^{P-1} s(g_i - T_3)2^i \quad (10)$$

Fig. 4 illustrates the coding process using different thresholds for the 3×3 sample blocks. Note that the first and second coding schemes are equivalent to the standard LBP [6] and LDBP [12], respectively. For visualization, LBP transformed the images by replacing each pixel of the image with its LBP value, these are also presented. **Fig. 5** shows examples of LBP transformed images with three coding schemes. From the examples, we can see that all schemes retain the global structure of the image, but display much diversity in capturing the local structures. In the next experiments, we will evaluate performances of the three coding schemes based

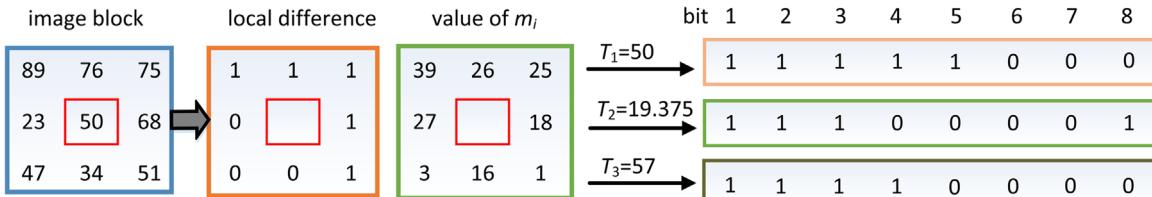


Fig. 4. Coding process of the three thresholds. (Here $R=1$, $P=8$).

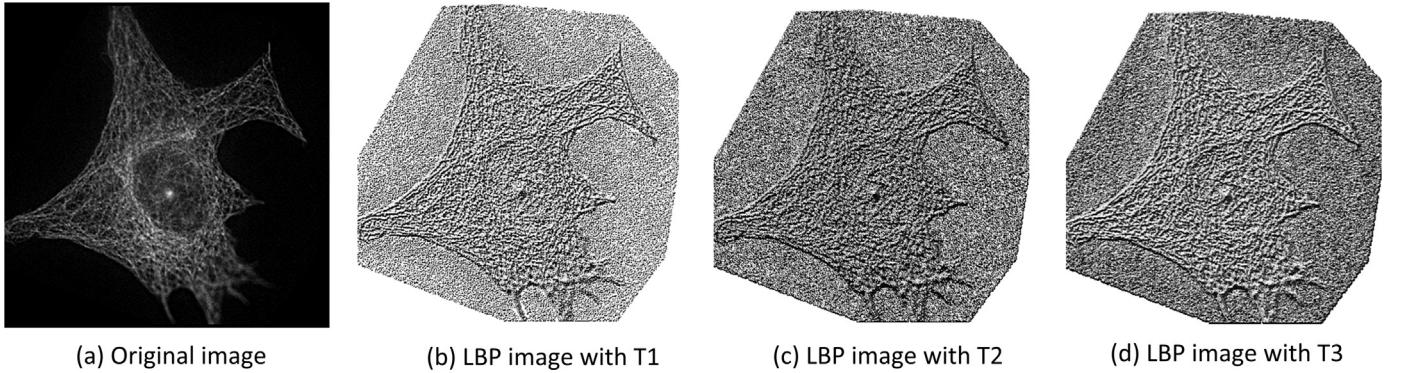


Fig. 5. Example image and its LBP transformed images with three coding schemes.

on our proposed algorithm to explore which one is more competitive in medical image classification.

2.4. Spatial adjacent histograms based on adaptive LBP

From the Fig. 3, one can see that adaptive radius map provides the image with rich micro-structure textures. These micro-structures reflect discriminative features and can be well represented by the approaches of histogram analysis. In this subsection, we explain how to construct the spatial adjacent histogram for an image when considering both micro-structures and adaptive radius patterns.

As mentioned in subsection 2.2, each pixel of an image has different neighborhood radius during our adaptive strategy, however, different configurations of (P, R) will result in $LBPT_i$ of each pixel mapping to different dimension. In this paper, we set the number of involved adaptive neighbors as 8, i.e. $P=8$, in a similar fashion to [22]. Besides, our spatial adjacent histogram encoding strategy will result in the final feature vector with $2 \times R_{\max} \times 2^P$ dimensions. This means that increasing the values for P and R_{\max} will make the feature dimension significantly large, that will further increase the time complexity. Considering the trade-off between description and performance, we set the maximal search radius R_{\max} as 4. Moreover, increasing R_{\max} means that more image pixels are involved to interpolate and form neighborhoods with fixed number $P=8$. This may bring more noise to the interpolated neighborhoods. So we think it is appropriate to set R_{\max} as 4.

Then, four histograms with 256 bins under four radii can be extracted for each image:

$$h(i)|R_k = \sum_{j \in R_k} \text{isequal}(lbp(j)=i), \quad i = 1, \dots, D; k = 1, \dots, 4 \quad (11)$$

where R_k indicates the set of pixels with radius $R=k$, and D denotes the dimension of LBP histogram (here D is equal to 256). The function $lbp(j)$ returns the LBP value of pixels j , and the function $\text{isequal}(lbp(j)=i)$ returns 1 if $lbp(j)=i$, else returns 0.

Then for image representation, an easy option is that we directly concatenate (for short, we name this histogram construction strategy as DC strategy) the four histograms under the different radii, then the concatenated histogram $DCLBP$ is as follows.

$$DCLBP^{T_i} = H^{T_i}(R_1)(+)H^{T_i}(R_2)(+)H^{T_i}(R_3)(+)H^{T_i}(R_4), \quad i \in \{1, 2, 3\} \quad (12)$$

where $(+)$ indicates the operation of concatenation, superscript T_i denotes the three coding schemes we have selected. A brief example is shown in Fig. 6(a).

Based on the preliminary experimental results, we found that the pixels which have been assigned $R=1$ are sparse but discriminative. We propose to use $R=1$ to calculate the spatial relation between them and pixels with larger radii for two reasons:

- (1) As shown in Fig. 3, the pixels with $R=1$ are mainly located in the regions with dramatic changes in gray value. The pixels with dramatic changes are also involved with ample micro-structures of different scales, i.e. spatial correlation of neighbor pixels with different LBP radii. We do not choose pixels

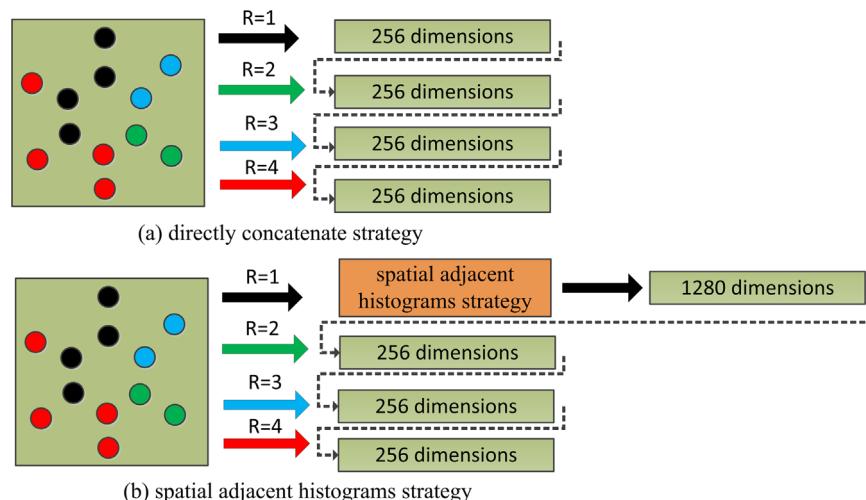


Fig. 6. The two strategies of combining sub-histograms of pixels with different radii to represent the image. (a) the directly concatenate strategy; (b) the spatial adjacent histograms strategy.

Table 1

The percentage of non-uniform patterns in medical datasets.

	$R=1, P=8\% ($	$R=2, P=8\% ($	$R=3, P=8\% ($	$R=4, P=8\% ($
2D-Hela [28]	27.84	38.24	44.21	45.13
Hep-2 [29]	10.96	18.78	21.86	26.89
PAP [30]	13.82	18.58	22.06	24.82

with larger radii for constructing the spatial adjacent histogram, such as $R=4$, since they are mainly located in regions with relative smooth change in gray value. Intuitively, the spatial correlation between pixels with dramatic changes and its neighbor pixels provide crucial feedback in disambiguating texture features.

- (2) Using $R=1$ is helpful to incorporate uniform idea. It has been validated that uniform patterns show superiority in texture classification and some non-uniform patterns involve noise [7]. Uniform patterns are also considered as the major parts of all patterns. Moreover, [8] certify that, the percentage of non-uniform patterns increases as the neighborhood radius increases. Table 1 shows the percentage of non-uniform patterns in three medical datasets and presents the same conclusion with [8]. An appropriate way of incorporating uniform idea and avoiding noise caused by some non-uniform patterns, in our case, is to choose pixels with $R=1$ (that involve the most uniform patterns) as baseline to construct the spatial adjacent histogram. Note that our goal

is not discarding the non-uniform local patterns, since some non-uniform patterns also can provide much useful features. And how to dig out the useful non-uniforms among all non-uniforms is another topic.

Therefore, instead of using original histograms of pixels with $R=1$, we propose to use spatial adjacent histogram strategy to embed the micro-structures information to $h(i)|R_1$. For each pixel $j \in R_1$ of an image, we calculate not only the number of occurrences of each $LBP(j)$ in the image, but also the number $h(i)|(R_1, R_k)$, that is the number of pixels with radius R_k which appears next to pixel j .

$$h(i)|(R_1, R_k) = \sum_{j \in R_1, w \in R_k} \sum_{w \in N(j)} \text{isequal}(LBP(j) = i), \quad k = 1, 2, 3, 4 \quad (13)$$

Where $N(j)$ is the four-neighbor system of j . Then the spatial adjacent histogram (SAH) for pixels set with radius $R=1$ can be defined as:

$$\begin{aligned} SAH(R_1) = & \{h(i)|R_1, h(i)|(R_1, R_1), h(i)|(R_1, R_2), h(i)|(R_1, R_3), \\ & h(i)|(R_1, R_4)\}, \quad i = 1, \dots, 256 \end{aligned} \quad (14)$$

It is easy to find that the dimension of $h(i)|(R_1, R_k)$ is 256, resulting in that the dimension of $SAH(R_1)$ is $256 + 256*4 = 1280$. Thus, the dimension of the histogram in Formula (14) is quintuple larger than that in Formula (11).

A toy example of explaining how to count $SAH(R_1)$ is present in Fig. 7, we design three image patches with 5×5 size, i.e. 25 pixels. To simplify the visualization, note that, we assume that the pixels

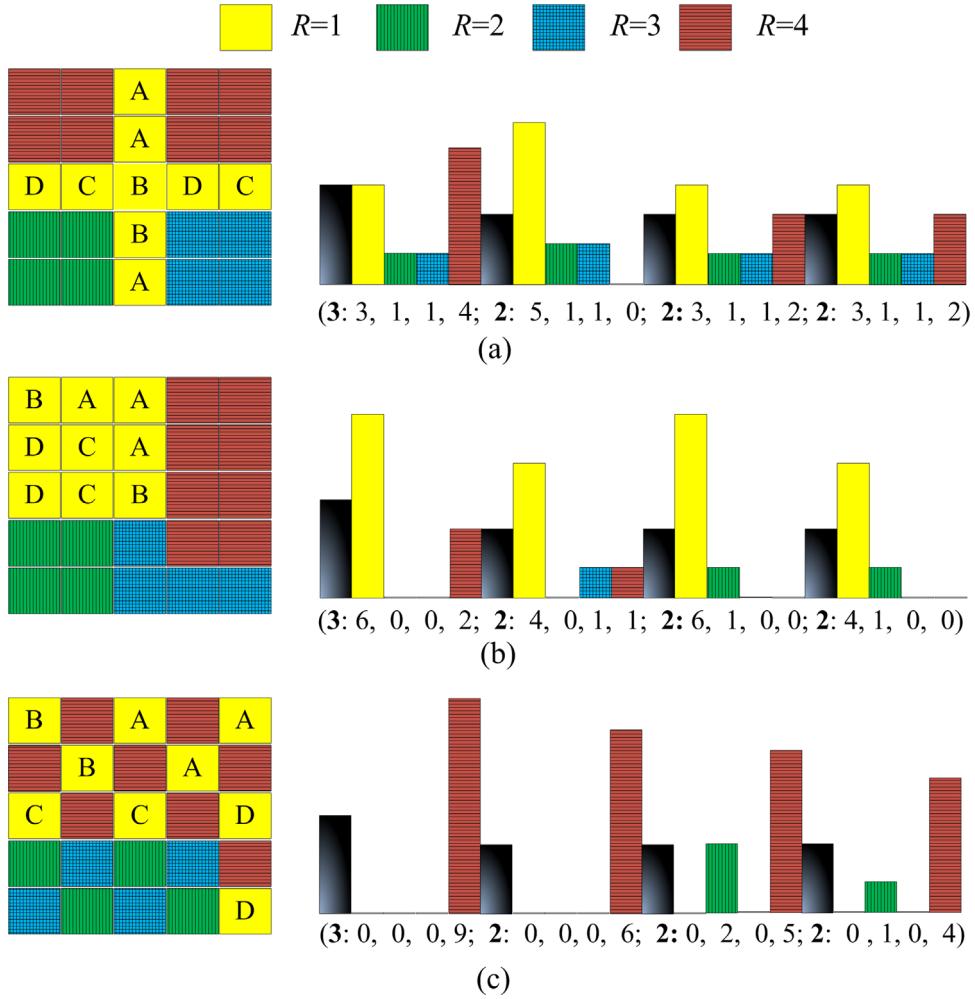


Fig. 7. Histograms for representing the pixels set with radius $R=1$.

with radius $R=1$ only be assigned to four types of LBP values, i.e. A to D ($0 < A < B < C < D < 255$). Furthermore, each pixel has an adaptive radius which is marked by different colors and textures, i.e. yellow pixel with no texture denotes $R=1$, green pixel with vertical texture denotes $R=2$, blue pixel with crossed texture denotes $R=3$, and red pixel with horizontal texture denotes $R=4$. Then the corresponding histogram and feature vectors are shown in Fig. 7. The four black bins with gradual change in each histogram denote the number of occurrences of A, B, C, and D, respectively. Note that the numbers of occurrences of A, B, C, and D are the same in the three image patches, i.e. for all image patches: $h(A)|R_1=3$, $h(B)|R_1=2$, $h(C)|R_1=2$ and $h(D)|R_1=2$. Following each black bin with gradual change, there are four bins with different textures which denote the number of four-neighbor pixels with four different radii, i.e. $h(A)|(R_1,R_1)$, $h(A)|(R_1,R_2)$, $h(A)|(R_1,R_3)$, and $h(A)|(R_1,R_4)$, respectively.

More specifically, looking at the image patches in Fig. 7(a), the first black bin with gradual change denotes the number of occurrences of A. Since there exist three yellow pixels with no texture located in the four-neighbors of all type A, then $h(A)|(R_1,R_1)=3$, drawn as the first yellow bin with no texture. Note that if two four-neighbor pixels are marked as the same type, then we count it only one time to update the histogram. For example, two pixels marked with A in Fig. 7(a) are mutual four-neighbors, but we count it only one time to update $h(A)|(R_1,R_1)$. There is only one

green pixel with vertical texture located in the four-neighbors of all type A, so $h(A)|(R_1,R_2)=1$, drawn as the first green bin with vertical texture. There is also only one blue pixel with crossed texture located in the four-neighbors of all type A, so $h(A)|(R_1,R_3)=1$, drawn as the first blue bin with crossed texture. And there are four red pixels with horizontal texture located in the four-neighbors of all type A, so $h(A)|(R_1,R_4)=4$. It is similar for counting the other non-black bins, i.e. $h(B)|(R_1,R_1)$, $h(B)|(R_1,R_2)$, $h(B)|(R_1,R_3)$, $h(B)|(R_1,R_4)$, $h(C)|(R_1,R_1)$, $h(C)|(R_1,R_2)$, $h(C)|(R_1,R_3)$, $h(C)|(R_1,R_4)$, $h(D)|(R_1,R_1)$, $h(D)|(R_1,R_2)$, $h(D)|(R_1,R_3)$ and $h(D)|(R_1,R_4)$.

Comparing the three image patches in Fig. 7, using the original LBP histogram model, the pixels set with radius $R=1$ in three image patches are represented as the same feature histogram (3,2,2,2) as shown by the black bins in each histogram. But using the histograms obtained by our spatial adjacent histogram (SAH) strategy, these image patches can be effectively distinguished from each other. This indicates that the micro-structure features have been effectively utilized in SAH strategy, that had been ignored in DC strategy.

As explained above, for an image, we can obtain four histograms, i.e. one histogram with 1280 dimensions to represent the pixels set with radius $R=1$ (see Eq. (14)), and three histograms with 256 dimensions to represent the pixels set with radii $R=2, 3, 4$, respectively (see Eq. (11)). Then, the final LBP feature vectors

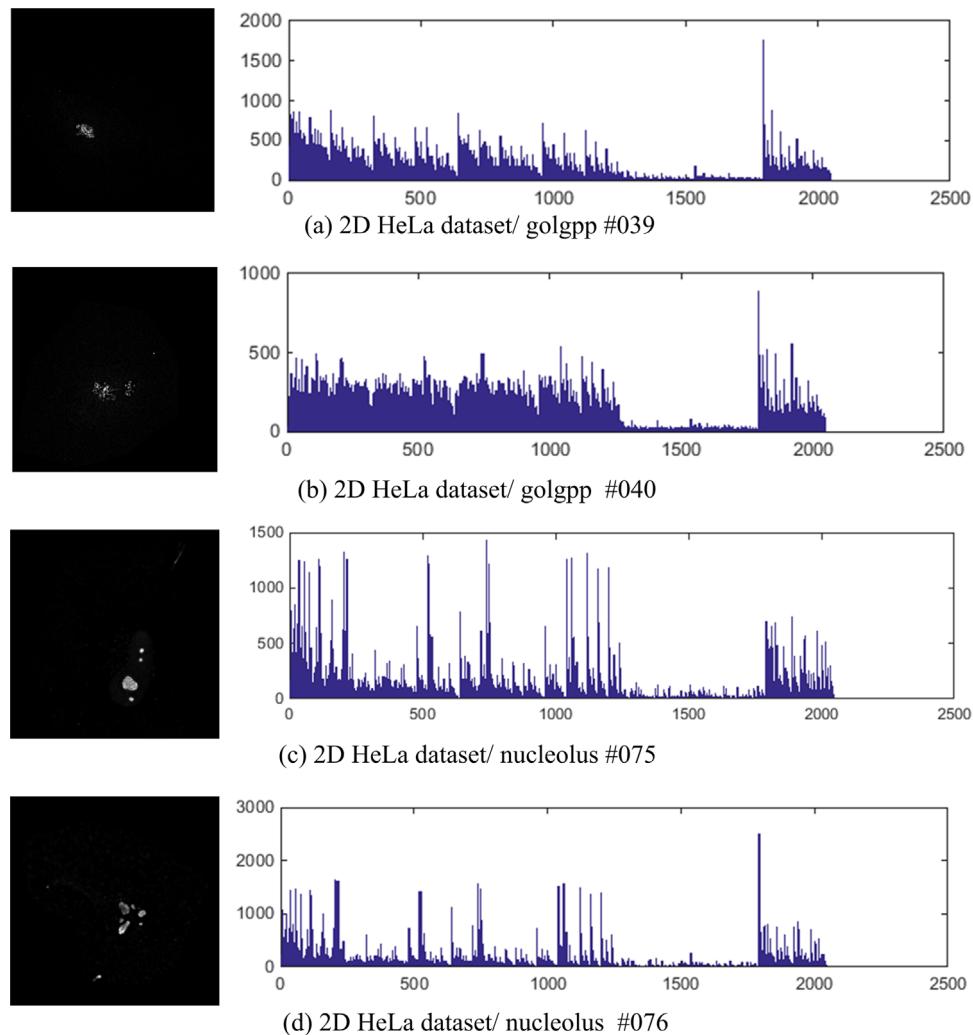


Fig. 8. Example images and its spatial adjacent histograms (the third coding scheme T_3 is employed here).

with our spatial adjacent histograms strategy can be computed as:

$$SAHLBP^{T_i} = SAH^{T_i}(R_1)(+)H^{T_i}(R_2)(+)H^{T_i}(R_3)(+)H^{T_i}(R_4), i \in \{1, 2, 3\} \quad (15)$$

where (+) indicates the operation of concatenation, superscript T_i denotes the three coding schemes above. And the dimension of $SAHLBP$ is $1280 + (256^3) = 2048$. Fig. 6 shows the comparison of the directly concatenate strategy and the spatial adjacent histograms strategy to represent the images.

In order to present an intuitive example, Fig. 8 shows four images and its $SAHLBP$ histogram features. The first two images come from the 'golgpp' category of 2D HeLa dataset, while the last two images come from the 'nucleolus' category of 2D HeLa dataset. Though it may be a challenge to distinguish the two categories with human vision, but the $SAHLBP$ histogram features can provide much discriminative information to distinguish the two different types of patterns. Therefore, our $SAHLBP$ method can be expected to perform well in medical image classification, we will give a complete evaluation in next experiments.

2.5. The proposed image classification framework

As illustrated in Fig. 9, the framework for medical image classification consists of four steps, i.e. determining the adaptive LBP radius for each pixel, computing the spatial adjacent histogram based on adaptive LBP radius, representing the image based on the spatial adjacent histograms, and training a SVM classifier for recognizing new samples. Furthermore, the first three steps can be considered as feature extraction for medical image, which is the key idea of our proposal.

3. Experimental results

3.1. Experimental datasets

The medical datasets for the evaluation of the proposed method are selected based on their variety and broadness of their

use. Experiments are performed on the following four reference datasets.

3.1.1. The 2D HeLa dataset

This dataset [28] is composed of 862 single-cell images (16 bit gray scale of size 512 by 382 pixels) from ten classes, i.e. Actin, Nucleus, Endosome, ER, Golgi Giantin, Golgi GPP130, Lysosome, Microtubules, Mitochondria, and Nucleolus. The description of the dataset in terms of samples per class is presented in Table 2, and some sample images from the ten categories are shown in Fig. 10. In our experiments, the LBP based histogram bin with the highest occurrence is discarded since it represents the black background.

3.1.2. Hep-2 cell dataset

Hep-2 cell dataset [29] consists of 28 slide images, and each image contains several cells that are segmented by specialist. The total number of cells is 1455, including training set with 721 cells and testing set with 734 cells. And each cell has been assigned to one of the six categories: centromere, homogeneous, nucleolar, coarse speckled, fine speckled, and cytoplasmatic. A summary of the Hep-2 cell dataset in terms of number of training set, testing set and classes is reported in Table 3. Some sample images are also presented in Fig. 11.

3.1.3. PAP dataset

The PAP smear dataset [30] contains 917 samples collected at the Herlev University Hospital using a digital camera and microscope. As shown in Table 4, the samples belong to seven different classes. And two skilled cyto-technicians further classified each sample into two super classes (normal and abnormal). The sample images that came from the two super classes are shown in Fig. 12.

3.1.4. The brain tumor dataset

The brain tumor dataset used throughout this experiment consists of 285 images of size 256*256, which was collected at the department of neurosurgery in Renmin Hospital of Wuhan University, P.R China. All samples are derived from real pathological images and 19 patients who suffer different types of brain tumor

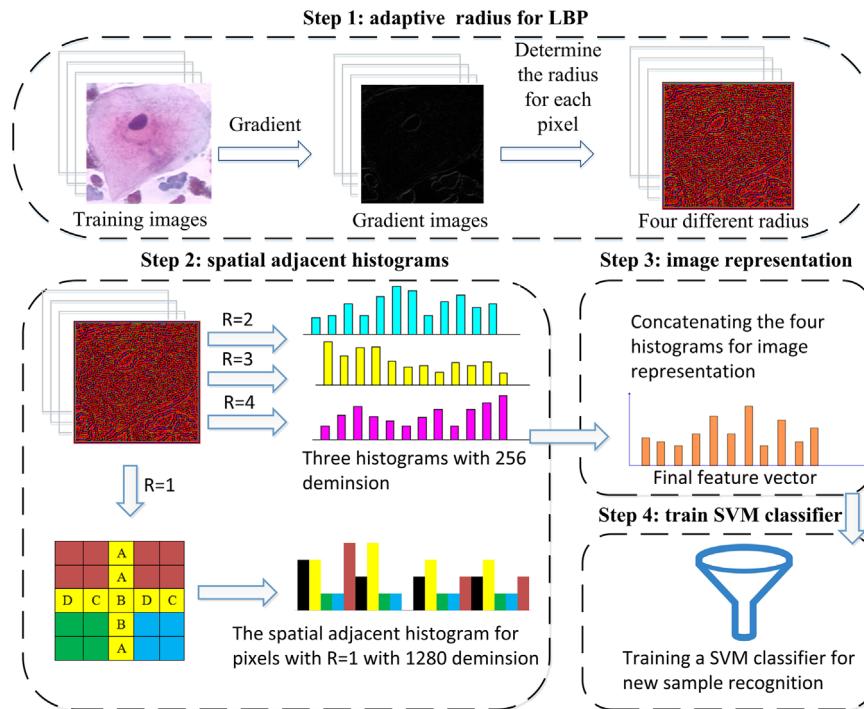


Fig. 9. Overview of the proposed image classification framework.

Table 2

A summary of the 2D HeLa dataset: classes and number of samples per class.

Class	Number
Actin	98
Nucleus	87
Endosome	91
ER	86
Golgi Giantin	87
Golgi GPP130	85
Lysosome	84
Microtubules	91
Mitochondria	73
Nucleolus	80
Total	862

are involved. In pathological analysis, the high resolution microscopy (Olympus BX51) was used to perform on the tumor lesion in order to get a histopathological diagnosis. Then five categories of microscopic images could be obtained according to the five types of brain tumor, i.e. astrocytoma, tuberculum sellae meningioma, olfactory groove meningioma, acoustic neuroma and pituitary tumor. Table 5 reports the number in each class and Fig. 13 depicts the microscopic images from each category.

Note that the original images in Hep-2 cell, PAP and brain tumor datasets are color images. In this study, the LBP variants and SIFT descriptors are extracted from the grayscale version of the corresponding images. The grayscale values are transformed by a weighted sum of R (red), G (green), B (blue) channels, i.e. Gray = 0.2990 * R + 0.5870 * G + 0.1140 * B.

3.2. Experimental setup and implementation

We implement the experiments in an iterated random splitting-scheme (consistent among all methods) for training and testing. Each dataset is randomly divided into 80% for training and 20% for testing. We use the training-split to optimize the SVM parameters and train model in a 5-fold cross validation. And then the trained model is employed for the test-split. The random splitting is repeated 30 times. For performance quantification, the

Table 3

A summary of the Hep-2 cell dataset: classes and number of samples per class.

Class	Training set	Testing set	Number of each class
Centromere	208	149	357
Homogeneous	150	180	330
Nucleolar	102	139	241
Coarse speckled	109	101	210
Fine speckled	94	114	208
Cytoplasmatic	58	51	109
Total	721	734	1455

average mean classification accuracy over 30 times iterations with its standard deviation is reported. In order to use this evaluation protocol consistently in all datasets, we merge the training and testing set for Hep-2 cell dataset, as shown in Table 3.

Since different classes may have different numbers of samples, we use stratified cross validation instead of standard cross validation to handle this imbalance of image numbers among classes. In stratified cross validation, the distribution of samples among classes in each subset is basically consistent with that in original dataset.

For the classifier, we employ the SVM with linear kernel in the next experiment. For the implementation of the SVM classifier, the public LIBSVM library [31] is employed.

3.3. Evaluation of our methods

In this section, we report our experimental results using the methods and dataset with performance indicators described in previous sections. Evaluation of the proposed model is accomplished through two different tasks: (1) evaluate the performance of three different coding schemes for medical image datasets; (2) verify that the spatial adjacent histograms is effective to increase the classification accuracy based on our adaptive radius strategy.

For the first task, we report on the classification accuracy of standard LBP with three scales ($R=1, P=8$), ($R=2, P=8$), and ($R=2, P=16$) and our method that adopted three coding schemes respectively, i.e. $LBPT_1^{(R=1, P=8)}$, $LBPT_2^{(R=1, P=8)}$, $LBPT_3^{(R=1, P=8)}$,

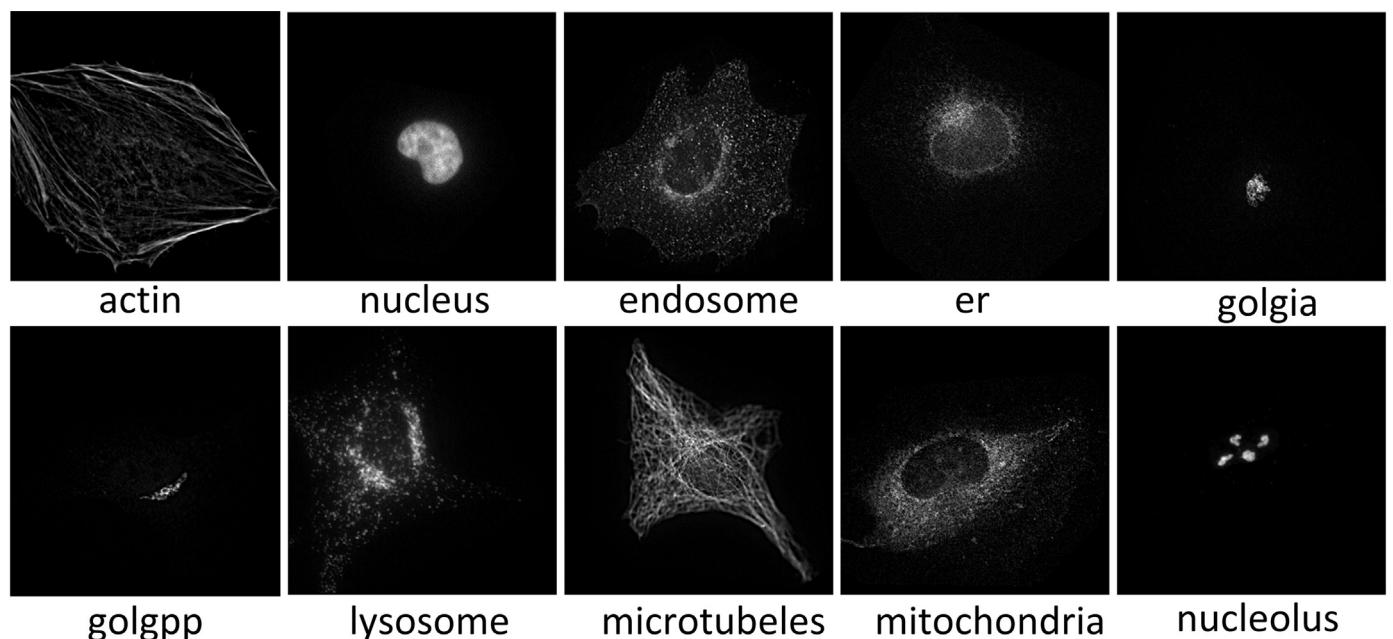


Fig. 10. Ten type of 2D-HeLa images.

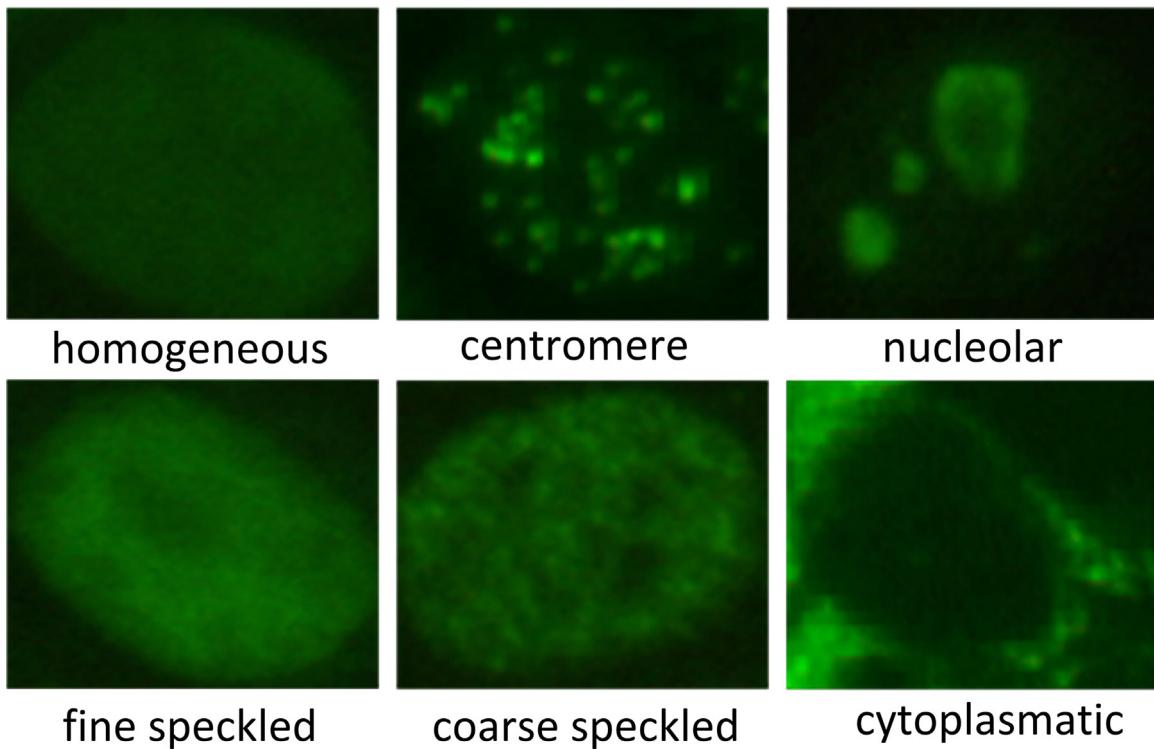


Fig. 11. Six type of Hep-2 cell image: homogeneous, centromere, nucleolar, fine speckled, coarse speckled and cytoplasmatic.

Table 4

A summary of the PAP smear dataset: classes and number of samples per class.

Class	Pap dataset	Super classes	Total
Superficial squamous epithelial	74	Normal	242
Intermediate squamous epithelial	70	Normal	
Columnnar epithelial	98	Normal	
Mild squamous non-keratinizing dysplasia	182	Abnormal	675
Moderate squamous non-keratinizing dysplasia	146	Abnormal	
Severe squamous non-keratinizing dysplasia	197	Abnormal	
Squamous cell carcinoma in situ intermediate	150	Abnormal	

$LBPT_1^{(R=2, P=8)}$, $LBPT_2^{(R=2, P=8)}$, $LBPT_3^{(R=2, P=8)}$, $LBPT_1^{(R=2, P=16)}$, $LBPT_2^{(R=2, P=16)}$, $LBPT_3^{(R=2, P=16)}$, $SAHLBPT_1$, $SAHLBPT_2$ and $SAHLBPT_3$, as shown in Table 6. In each table, the best results in each dataset are bolded.

For the second task, we compare the methods generated by spatial adjacent histograms (SAH) strategy and directly concatenate (DC) strategy under the three coding schemes, respectively, i.e. $DCLBPT_1$ and $SAHLBPT_1$, $DCLBPT_2$ and $SAHLBPT_2$, $DCLBPT_3$ and $SAHLBPT_3$. The results are also reported in Table 6.

Examining the Table 6, we can make the following conclusions:

- (1) Our $SAHLBP$ models consistently outperform standard LBP with $(R=1, P=8)$ and $(R=2, P=8)$ on the four medical datasets, no matter which coding schemes that are employed. Our $SAHLBP$ models also outperform standard LBP with $R=2, P=16$ in all cases except that perform on PAP dataset with T_2 , i.e. only $SAHLBPT_2$ is inferior to $LBPT_2^{(R=2, P=16)}$ on PAP dataset. Moreover, the dimension of our method (i.e. 2048 dimensions) is much smaller than that of $LBPT_2^{(R=2, P=16)}$ (i.e. 65536 dimensions). This is probably because standard LBP s cannot capture spatial relationships among local textures, while our $SAHLBP$ models describe micro-structures obtained by

adaptive radius and perform successfully well in capturing the valuable structural information, which is beneficial to medical image classification.

- (2) The models with the third coding scheme T_3 achieved higher accuracy, on the average, than that with the first and second coding schemes. More specifically, for the standard LBP with $R=1$ and $P=8$, the first coding scheme T_1 gained the best performance in the 2D-Hela and Tumor datasets, while T_3 gained the best performance in the PAP smear and Hep-2 cell datasets. For the standard LBP with $R=2$ and $P=16$, T_1 gained the best performance in the 2D-Hela dataset and T_3 proves superior in the remaining datasets. For $SAHLBPs$, the T_3 performed better than the other coding schemes in the PAP smear, 2D-Hela and Tumor datasets, while the T_1 achieved the best performance in the Hep-2 cell datasets.
- (3) Particularly interesting are the results of the methods with the second coding scheme T_2 , performs slightly worse than the other two schemes in most of the cases. However, the coding scheme T_2 has been proven in [12] that it describes the global structure successfully and has a good performance in natural scene image classification. From this observation, we can appreciate that the micro-structure features also remain very relevant to coding scheme.
- (4) $LBPT_i$ with $(R=2, P=16)$ performs better than $DCLBPT_i$. We conjecture that it is because the $LBPT_i$ with $(R=2, P=16)$ captures more texture information with very large feature dimension (i.e. 65536 dimensions), while the feature dimension of $DCLBPT$ is 1024. However, the $DCLBPT_i$ achieves higher accuracies than $LBPT_i$ with both $(R=1, P=8)$ and $(R=2, P=8)$ in most cases. The comparison between the $DCLBPT_i$ and $LBPT_i$ suggests that the adaptive radius is not very suitable to combine with directly concatenate strategy. We mainly compare the $DCLBPT_i$ with $SAHLBPT_i$ to show the superiority of adaptive radius combined with spatial adjacent histogram.
- (5) Our spatial adjacent histogram (SAH) based on adaptive radius is better than directly concatenate strategy for all datasets, no

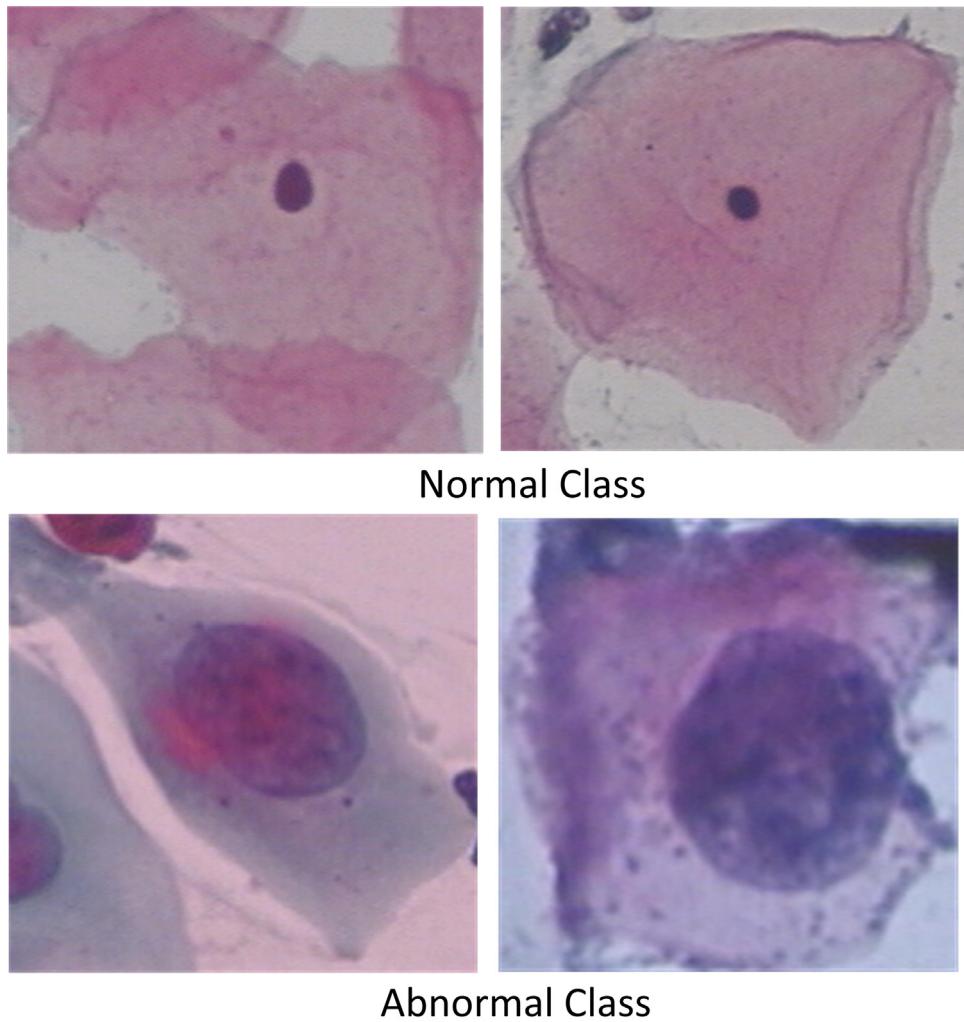


Fig. 12. Sample images from the pap smear dataset.

Table 5

A summary of the brain tumor dataset: classes and number of samples per class.

Class	Number
Astrocytoma	58
Tuberculum sellae meningioma	54
Olfactory groove meningioma	54
Acoustic neuroma	56
Pituitary tumor	63
Total	285

matter which coding schemes T adopted, i.e. $SAHLBPT_i$ is always better than $DCLBPT_i$. Therefore, one can conclude that SAH is highly suitable for representing the micro-structures which is discriminative in medical image classification.

To provide more robust results, we further perform the Wilcoxon ranksum test in combination with the iterated splitting scheme to compare the methods. More specifically, for each method in Table 6, we can obtain 30 classification results during 30 times splitting iterations for each dataset. Then for a pair of compared methods, we perform Wilcoxon ranksum test with their 30 classification results over one specific dataset.

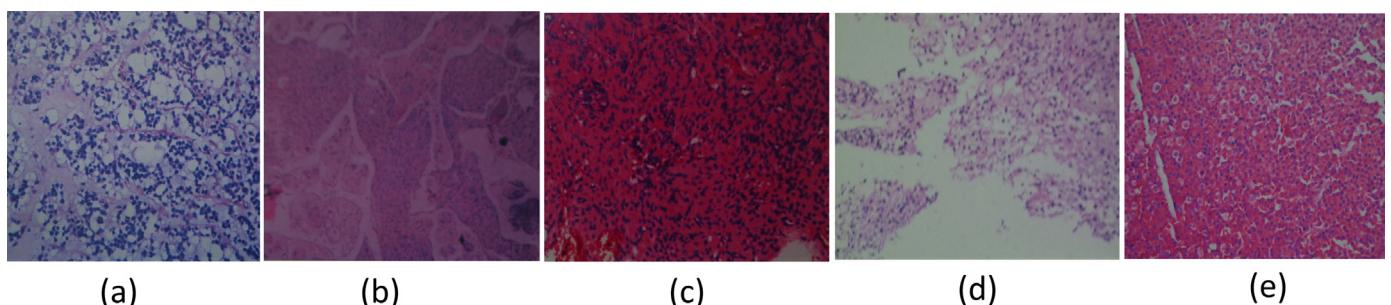


Fig. 13. Microscope images for different types of brain tumor lesion. (a) astrocytoma; (b) tuberculum sellae meningioma; (c) olfactory groove meningioma; (d) acoustic neuroma; (e) pituitary tumor.

Table 6

Classification accuracy of each method over four datasets.

Methods	Datasets			
	PAP	2D-Hela	Hep-2	Tumor
<i>LBPT</i> ₁ ^(R=1, P=8)	81.12 ± 2.3	85.24 ± 2.4	75.65 ± 1.1	75.41 ± 1.8
<i>LBPT</i> ₂ ^(R=1, P=8)	80.92 ± 1.7	83.23 ± 1.9	72.17 ± 1.2	74.45 ± 1.7
<i>LBPT</i> ₃ ^(R=1, P=8)	81.87 ± 1.5	84.31 ± 2.1	79.94 ± 1.6	75.22 ± 1.8
<i>LBPT</i> ₁ ^(R=2, P=8)	80.34 ± 2.5	86.07 ± 2.5	82.14 ± 1.4	75.04 ± 2.1
<i>LBPT</i> ₂ ^(R=2, P=8)	79.76 ± 2.5	82.59 ± 1.8	76.23 ± 1.5	73.62 ± 2.1
<i>LBPT</i> ₃ ^(R=2, P=8)	82.78 ± 2.3	85.39 ± 2.2	80.23 ± 1.1	73.78 ± 2.1
<i>LBPT</i> ₁ ^(R=2, P=16)	85.47 ± 2.2	87.32 ± 2.2	87.67 ± 1.4	79.17 ± 1.6
<i>LBPT</i> ₂ ^(R=2, P=16)	84.56 ± 2.0	86.23 ± 1.9	85.23 ± 1.0	78.82 ± 1.9
<i>LBPT</i> ₃ ^(R=2, P=16)	85.56 ± 1.9	86.76 ± 2.1	87.92 ± 1.2	79.25 ± 2.0
<i>DCLBP</i> ₁	83.96 ± 1.9	84.96 ± 1.4	81.31 ± 0.8	77.69 ± 1.9
<i>DCLBP</i> ₂	82.14 ± 1.9	85.34 ± 1.4	76.39 ± 1.0	75.22 ± 1.8
<i>DCLBP</i> ₃	83.29 ± 1.9	87.84 ± 2.2	82.22 ± 1.1	78.13 ± 1.5
<i>SAHLBPT</i> ₁	86.69 ± 2.1	89.68 ± 2.0	91.89 ± 0.8	82.91 ± 1.6
<i>SAHLBPT</i> ₂	84.03 ± 2.1	87.29 ± 1.6	88.95 ± 0.8	81.24 ± 2.1
<i>SAHLBPT</i> ₃	88.03 ± 1.7	90.06 ± 1.5	91.86 ± 0.9	84.48 ± 1.8

More specifically, for 2D-HeLa dataset, we have compared the following approaches using the Wilcoxon ranksum test:

1. *SAHLBPT*_i versus *LBPT*_i ($i=1, 2, 3$) with different scales, *SAHLBPT*_i wins against *LBPT*_i with all scales ($R=1, P=8$), ($R=2, P=8$), and ($R=2, P=16$), i.e. we reject the null hypothesis at the level of significance 0.05, and accept that *SAHLBPT*_i and *LBPT*_i have significant different performance.
2. *SAHLBPT*_i versus *DCLBP*_i ($i=1, 2, 3$), *SAHLBPT*_i wins against *DCLBP*_i.
3. *SAHLBPT*₃ versus *SAHLBPT*_i ($i=1, 2$), *SAHLBPT*₃ wins against *SAHLBPT*₂, however, *SAHLBPT*₃ and *SAHLBPT*₁ have no significant different performance in 2D-HeLa dataset.

We also do the same Wilcoxon ranksum test for the other three datasets, and most of the conclusions remain valid. Very few different conclusions are marked as the following: (1) For PAP and brain tumor datasets, *SAHLBPT*₃ wins against both *SAHLBPT*₁ and *SAHLBPT*₂. (2) For PAP dataset, *SAHLBPT*₂ and *LBPT*₂ (with $R=2, P=16$) have no significant different performance. The p-values for all tests are presented in Appendix A.

Finally, we analyze the confusion matrices for *SAHLBPT*₃ which display the best results selected from the 30 times iterations for each dataset. The confusion matrices for the three multiclass datasets, i.e. 2D-HeLa, Hep-2 cell, and brain tumor datasets, are presented in Appendix A. From the results, we can confirm that our method is very useful for some particular patterns, such as actin, dna, golgia, golgpp, nucleolus patterns in 2D-HeLa dataset and nucleolar pattern in Hep-2 cell dataset.

From the analysis of the above experimental results, it is clear the advantage of our spatial adjacent histogram strategy employed the third coding scheme T_3 , i.e. the *SAHLBPT*₃. Hence, to achieve a tradeoff between description and performance, we have chosen the *SAHLBPT*₃ in the remainder of the experiments.

3.4. Comparison of the LBP-based methods

The tests in this subsection are aimed at comparing our method with the recent LBP variants reported in the literature. In particular, the following methods are evaluated which are often used for medical image classification:

- 1) LBP-r [6], the standard LBP with rotation invariant;
- 2) EBP [24], the elliptical binary pattern variant;
- 3) LBPsu2[22], scale-adaptive and subuniform-based rotation invariant LBP;

Table 7

Classification accuracy of compared methods over four datasets.

Methods	PAP	2D-Hela	Hep-2	Tumor
LBP-r	75.5 ± 1.2	81.21 ± 1.9	75.23 ± 0.9	73.71 ± 1.4
ELBP	80.02 ± 1.5	85.11 ± 1.7	80.17 ± 1.1	75.82 ± 1.5
LBPsu2	82.42 ± 1.3	84.82 ± 1.5	83.23 ± 0.6	76.72 ± 0.6
LBPnr	83.14 ± 1.8	85.76 ± 2.1	82.76 ± 1.3	78.29 ± 2.0
LBP-HF	84.13 ± 1.5	85.14 ± 2.0	87.23 ± 1.0	77.75 ± 1.7
LDBP	81.26 ± 1.7	82.5 ± 1.9	79.58 ± 0.7	75.11 ± 2.0
CLBP	87.21 ± 1.4	88.91 ± 1.1	87.91 ± 0.9	81.12 ± 1.8
CoALBP	87.02 ± 1.9	87.72 ± 2.1	92.76 ± 1.1	81.33 ± 1.9
<i>SAHLBPT</i> ₃	88.03 ± 1.7	90.06 ± 1.5	91.86 ± 0.9	84.48 ± 1.8

- 4) LBPnr[22], scale-adaptive LBP without rotation invariant, that employs the LBPsu2 approach by using non-rotation invariant instead of rotation invariant.
- 5) LBP-HF[25], local binary pattern histogram Fourier features;
- 6) LDBP[12], the local difference magnitude binary pattern variant;
- 7) CLBP[26], in which three types of features are combined to form the completed LBP features;
- 8) CoALBP [27], in which the co-occurrence among adjacent LBPs has been considered. The effectiveness of CoALBP for medical images can be confirmed since it won the first prize in the 2012 HEP-2 cells classification contest.

In Table 7, we report the classification accuracies using the datasets and the compared methods described above. In each table, the best result in each datasets is bolded.

From Table 7, we can make the following findings.

First, *SAHLBPT*₃ is one of the best variants proposed in this work. It achieves, on average, a better result than the compared methods. Also, *SAHLBPT*₃ achieves the highest accuracy in PAP, 2D-HeLa, and Brain tumor datasets. We conjecture that it is because our method enhanced its performance by exploiting the spatial information of the adaptive radius with different scales, which produced more discriminative features for medical images.

Second, CoALBP works better than *SAHLBPT*₃ in Hep-2 cell dataset, and *SAHLBPT*₃ outperforms the CoALBP for the other three datasets. However, our *SAHLBPT*₃ has one advantage over the CoALBP: its feature dimension is much smaller. The experimental feature sizes of *SAHLBPT*₃ and CoALBP are 2048 and 3072, respectively.

Finally, *SAHLBPT*₃ works much better than the methods with rotation invariant, such as LBP-r and LBPsu2, even though the rotation invariant features are not involved in our methods. However, the rotation invariant contains additional discriminative information. How to utilize this rotation invariant feature based on our adaptive strategy will be explored in our future research.

From Table 7, it can be observed that CLBP, CoALBP, and *SAHLBPT*₃ achieve similar classification accuracies. To provide another basis for comparison of the three methods, we have also reported the macro-averaged F1-scores as shown in Table 8. Note that the reported classification accuracy in Table 7 is defined as the number of samples correctly classified divided by the total number of samples in the test set. It is equivalent to the micro-averaged F1 score in the single-label classification task. Comparing the results

Table 8

Macro-averaged F1 scores of the three compared methods over four datasets.

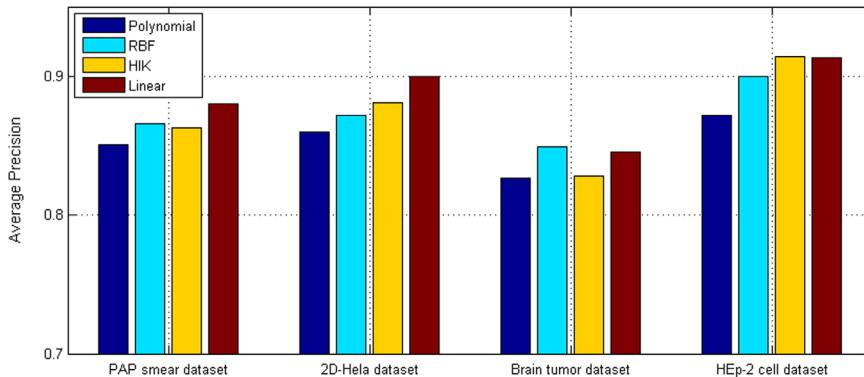
Methods	PAP	2D-Hela	Hep-2	Tumor
CLBP	82.87 ± 2.1	87.12 ± 1.9	86.41 ± 1.5	78.65 ± 2.1
CoALBP	81.45 ± 1.9	86.23 ± 1.7	91.89 ± 1.7	79.23 ± 2.0
<i>SAHLBPT</i> ₃	84.23 ± 1.9	88.76 ± 1.5	91.29 ± 1.5	83.81 ± 1.9

Table 9

Statistical significance of differences in classification accuracies for different methods.

Datasets	Compared methods	p-values	Corrected α	Result
PAP	SAHLBPT ₃ vs. CLBP	0.1435	0.00625	NS
	SAHLBPT ₃ vs. CoALBP	0.0122	0.00625	NS
2D-Hela	SAHLBPT ₃ vs. CLBP	0.0028	0.00625	*
	SAHLBPT ₃ vs. CoALBP	< 0.0001	0.00625	*
Hep-2	SAHLBPT ₃ vs. CLBP	< 0.0001	0.00625	*
	SAHLBPT ₃ vs. CoALBP	0.1723	0.00625	NS
Tumor	SAHLBPT ₃ vs. CLBP	< 0.0001	0.00625	*
	SAHLBPT ₃ vs. CoALBP	< 0.0001	0.00625	*

the computational time for feature extraction (for a single image), training and testing on 2D-Hela dataset, as shown in **Table 10**. The results show that SAHLBPT₃ is slower than standard LBP (with both $R=1$, $P=8$ and $R=2$, $P=16$) and CoALBP, which is caused by the increased time demand time for searching the adaptive radius and computing spatial adjacent histogram. It is worth stating that, in training and testing stage, our method is faster than LBP (with $R=2$, $P=16$) and CoALBP. This is because the feature dimension of our method (i.e. 2048 dimensions) is shorter than LBP($R=2$, $P=16$) (i.e. 65536 dimensions) and CoALBP (i.e. 3072 dimensions). Based on consideration of both time and performance, we infer that SAHLBPT₃ is suitable for medical image classification.

**Fig. 14.** Classification accuracies of our method on four database for different SVM kernels.

of **Table 7** and **Table 8**, we can find out that the micro- and macro-averaged results do not differ significantly and the previous findings based on accuracy would not change considering the macro-averaged F1-score.

In order to further perform statistical analysis for comparing our SAHLBPT₃ with the other two representative methods, i.e. CLBP and CoALBP, we employ Wilcoxon ranksum test and the Bonferroni method [36] for multiple pair-wise comparisons. In this case, the level of significance has been corrected as $\alpha=0.05/8=0.00625$ according to the Bonferroni method. The significant p-values and corrected α are reported in **Table 9**, where * denotes that the two methods have significant different performance and NS denotes that the two methods have no significant different performance.

In summary, we can conclude that our method is indeed a superior algorithm with a clear motivation, convincing improvements, and an easy implementation. The advantage of our method with respect to other variants makes it a good choice for medical image classification.

3.5. Runtime performance analysis

To study the computational demand of the proposed method, we analyze the required time of our model and some representative methods. The MATLAB R2014a are employed, running on an Intel i5-2400 processor at 3.1 GHz with 4 GB RAM. We report

Furthermore, The computational time for feature extraction mainly depends on the descriptor and the image size, while the training and testing are mainly dependant on the classifier, feature dimension and the scale of the dataset. Our method has took 0.823 s in the feature extraction stage per image with 512 × 382 size and 6.553 s in the training stage for 862 samples with 3072 feature dimensions. We think that our computational demand is an adequate trade-off. Viewed from this perspective, we believe our method can be extended to large-scale datasets. Due to the limitations of the scale of medical image datasets, we will evaluate our method on more general datasets with larger scale in our future research.

3.6. Influence of different SVM kernels

Since the choice of the SVM kernel may have an impact on the classification rates, we also conducted an experiment by using several SVM kernels. Though abundantly new kernels have been proposed, the most frequently used kernel functions [32,33] in image classification are linear, polynomial, Radial Basis Function (RBF), and histogram intersection (HI) kernels [34], formulated as in Eq.(16)–Eq.(19), respectively.

$$(1) \text{ Linear : } K(x_i, x_j) = x_i^T x_j \quad (16)$$

$$(2) \text{ Polynomial : } K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (17)$$

$$(3) \text{ Radial basis function(RBF)}$$

$$: K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (18)$$

$$(4) \text{ Histogram intersection(HI) : } K(x_i, x_j) = \sum_{i=1}^n \min(x_i, x_j) \quad (19)$$

Here, γ , r , and d are kernel parameters.

In our test, the default parameter values are used for different kernels. **Fig. 14** shows the average classification accuracies we achieved with different kernels on the four datasets. Note that we

Table 10

Computation time for feature extraction (single image), training and testing on 2D-Hela dataset.

Methods	Feature extraction	Training	Testing
LBP($R=1$, $P=8$)	0.103 s	1.423 s	0.056 s
LBP($R=2$, $P=16$)	0.197 s	814.27 s	44.305 s
CoALBP	0.562 s	12.37 s	0.798 s
SAHLBPT ₃	0.823 s	6.553 s	0.493 s

Table 11

Classification accuracy of compared descriptors in dense BoVW framework over four datasets.

		PAP	2D-Hela	Hep-2	Tumor
BoVW framework	SIFT	84.03 ± 2.3	83.79 ± 2.5	80.88 ± 1.5	78.23 ± 1.7
	LBP	81.43 ± 2.1	81.47 ± 2.1	77.52 ± 1.9	75.21 ± 2.1
	SAHLBPT ₃	86.21 ± 2.0	84.89 ± 2.2	85.06 ± 2.0	81.45 ± 1.7
	SIFT +	87.63 ± 2.1	86.21 ± 2.5	83.47 ± 1.8	83.34 ± 1.9
Proposed framework	SAHLBPT ₃	88.03 ± 1.7	90.06 ± 1.5	91.86 ± 0.9	84.48 ± 1.8

Table 12

Macro-averaged F1 score of compared descriptors in dense BoVW framework over four datasets.

		PAP	2D-Hela	Hep-2	Tumor
BoVW framework	SIFT	77.97 ± 2.5	82.29 ± 2.4	79.76 ± 1.8	77.07 ± 1.9
	LBP	73.17 ± 2.5	80.13 ± 2.1	76.19 ± 1.7	74.01 ± 2.1
	SAHLBPT ₃	81.58 ± 2.4	83.75 ± 2.0	84.05 ± 1.9	80.13 ± 1.7
	SIFT +	83.27 ± 2.4	85.35 ± 2.3	82.31 ± 2.1	82.41 ± 1.9
	SAHLBPT ₃				

have not optimized SVM kernel parameters in Fig. 14. Particularly noteworthy, in our case is that, the parameter optimization performed for SVM with RBF kernel is time-consuming, while the parameter optimization performed for SVM with linear kernel has a little influence on classification performance. This means the classification result of our method has little change if we set different cost factor for SVM with linear kernel. This may be due to the fact that the feature vectors of our method are rather sparse with high dimensionality. Based on the average results on four datasets, linear SVM kernels should be preferred in this case, since computation time in this case varies linearly with the size of the training data.

3.7. Evaluating the performance of the SAHLBPT in BoVW framework

In this section, for the sake of completeness, we further test the performance of our texture descriptor within the bag of visual words (BoVW) framework. The BoVW model is one of the most popular algorithms for image representation by using visual words formed by vector-quantizing local features with a clustering method (such as K-means). The standard LBP and our SAHLBPT₃ method can also be used as local texture descriptors for BoVW. In the experiment, we use dense sampling strategy for the BoVW model. More specifically, the local features are extracted from 16 × 16 image patches on a regular grid spaced at 8 pixels for all images. Since the SIFT descriptor is probably the most widely used descriptor for describing the local patches in BoVW model, we mainly compare our method with dense SIFT descriptor. For dense SIFT descriptor, the two parameters patch size and grid spacing are set as 16 and 8, respectively, to allow for dense sampling. Moreover, the spatial pyramid matching (SPM) [37] settling is employed for SIFT descriptor. The SPM is calculated at three levels. At the first level, the original image is considered as a sub-region. At the second and the third level, the original image is divided into 2 × 2 and 4 × 4 sub-regions, respectively. Then all BoVW histograms of 21 sub-regions are concatenated to form the SPM representation of the image. Additionally, single feature may fail to capture the rich information within local image patches, as such, it is reasonable to extract multiple features for compensation. Therefore, we also use multiple features that consist of SIFT and SAHLBPT₃ to describe local patches and then apply it to construct the visual vocabulary (named SIFT + SAHLBPT₃ for short).

Table 13

Statistical significance of differences in classification accuracies for different local descriptors in BoVW model.

datasets	Compared methods	p-values	Corrected α	Result
PAP	SAHLBPT ₃ vs. SIFT	0.0031	0.00417	*
	SAHLBPT ₃ vs. LBP	< 0.0001	0.00417	*
	SIFT + SAHLBPT ₃ vs. SAHLBPT ₃	0.0053	0.00417	NS
	SAHLBPT ₃ vs. SIFT	0.0079	0.00417	NS
2D-Hela	SAHLBPT ₃ vs. LBP	< 0.0001	0.00417	*
	SIFT + SAHLBPT ₃ vs. SAHLBPT ₃	0.0370	0.00417	NS
	SAHLBPT ₃ vs. SIFT	< 0.0001	0.00417	*
	SAHLBPT ₃ vs. LBP	< 0.0001	0.00417	*
Hep-2	SIFT + SAHLBPT ₃ vs. SAHLBPT ₃	0.0208	0.00417	NS
	SAHLBPT ₃ vs. SIFT	< 0.0001	0.00417	*
	SAHLBPT ₃ vs. LBP	< 0.0001	0.00417	*
	SIFT + SAHLBPT ₃ vs. SAHLBPT ₃	0.0013	0.00417	*
Tumor	SAHLBPT ₃ vs. SIFT	< 0.0001	0.00417	*
	SAHLBPT ₃ vs. LBP	< 0.0001	0.00417	*
	SIFT + SAHLBPT ₃ vs. SAHLBPT ₃	0.0013	0.00417	*

The classification accuracies and the macro-averaged F1 scores for different local descriptors used in BoVW model are reported in Table 11 and Table 12, respectively. Here the visual vocabulary size is 200 for all methods.

To perform statistical analysis, we also compare the SAHLBPT₃ with SIFT, LBP and SIFT + SAHLBPT₃ within BoVW framework tested by Wilcoxon ranksum test, the test protocol is the same as that in subsection 3.4. In this test, the level of significance has been corrected as $\alpha=0.05/12=0.00417$ according to the Bonferroni method. The p-values and corrected α are reported in Table 13, where * denotes that the two methods have significant different performance and NS denotes that the two methods have no significant different performance.

From Table 11, Table 12 and Table 13, we can make the following conclusions.

- (1) From Table 11 and Table 12, it can be observed that the SAHLBPT₃ consistently outperforms both SIFT and LBP on the four datasets when they are applied to BoVW framework. However, the LBP is slightly weaker than SIFT as local descriptors in BoVW model. Looking at the significance of differences in Table 13, we can find out that the improvements of SAHLBPT₃ over LBP are significant on the four datasets. Moreover, SAHLBPT₃ and SIFT have significant different performance in PAP, Hep-2 cell and Tumor datasets.
- (2) The BoVW model using the multiple features, i.e. SIFT + SAHLBPT₃, excels that using a single feature on PAP, 2D-Hela, and Tumor datasets based on the accuracy and the macro-averaged F1 score. However, the SAHLBPT₃ works better than SIFT + SAHLBPT₃ on Hep-2 cell dataset, which means that SIFT and SAHLBPT₃ descriptors do not provide very useful compensation for Hep-2 cell dataset. Looking at Table 13, SIFT + SAHLBPT₃ and SAHLBPT₃ have significant different performance in Tumor datasets, but they have no significant different performance in PAP, 2D-Hela and Hep-2 cell datasets. This means that SIFT and our SAHLBPT₃ are complementary features for Tumor datasets in BoVW model.
- (3) To discuss, we return to the classification performance of the SAHLBPT₃ in our framework and also re-record the results in Table 7. We can easily find that SAHLBPT₃ in our framework achieves better classification results than that in the BoVW framework. This is because our SAHLBPT₃ aims to describe the micro-structures of the whole image and these micro-structures tend to play the largest role with a globally encoding mode. Encoding the micro-structures in local image patches would slightly decrease the performance of SAHLBPT₃. Moreover, the run-time for clustering in BoVW framework is time-consuming. Based on both the classification result and

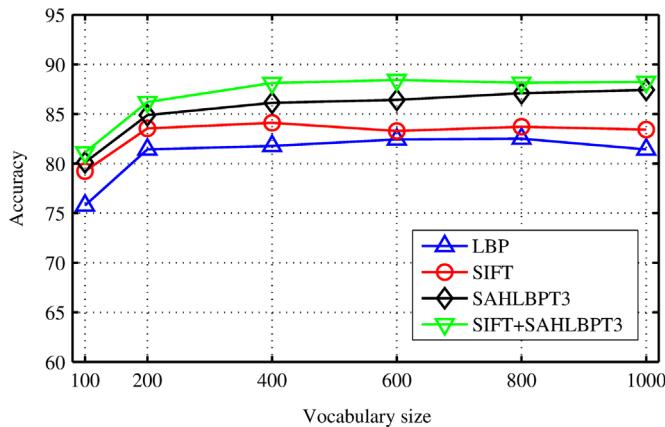


Fig. 15. Comparison of classification accuracies on 2D-Hela datasets using different vocabulary sizes for BoVW model.

Table 14
P-values of statistical tests for compared methods on four datasets.

Compared methods	PAP	2D-Hela	Hep-2	Tumor
SAHLBPT ₁ vs. LBPT ₁ ^(R=1,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₂ vs. LBPT ₂ ^(R=1,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₃ vs. LBPT ₃ ^(R=1,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₁ vs. LBPT ₁ ^(R=2,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₂ vs. LBPT ₂ ^(R=2,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₃ vs. LBPT ₃ ^(R=2,P=8)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₁ vs. LBPT ₁ ^(R=2,P=16)	0.0213	0.0021	< 0.0001	< 0.0001
SAHLBPT ₂ vs. LBPT ₂ ^(R=2,P=16)	0.072	0.0033	< 0.0001	0.0015
SAHLBPT ₃ vs. LBPT ₃ ^(R=2,P=16)	0.0019	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₁ vs. DCLBPT ₁	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₂ vs. DCLBPT ₂	< 0.0001	0.0018	< 0.0001	< 0.0001
SAHLBPT ₃ vs. DCLBPT ₃	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SAHLBPT ₃ vs. SAHLBPT ₁	0.0181	0.0823	0.1823	0.0109
SAHLBPT ₃ vs. SAHLBPT ₂	< 0.0001	< 0.0001	0.0013	< 0.0001

time complexity, the SAHLBPT₃ involved in our framework should be more preferred in medical image classification task.

Finally, we then report the classification accuracies of the methods using different vocabulary sizes. In Fig. 15, the classification accuracies have a similar trend as the visual vocabulary size increases, i.e. the performance saturates when the vocabulary size larger than 200. Moreover, SAHLBPT₃ consistently outperforms both SIFT and LBP using different vocabulary sizes.

4. Conclusion and future work

In this paper, we proposed a novel framework based on our adaptive local binary patterns and spatial adjacent histogram for medical image classification. We first employed gradient operator to determine the adaptive neighborhood radius of each pixel. Then three coding schemes based on adaptive radius were introduced for processing the LBP histograms. To capture the discriminative micro-structures features produced by the adaptive radius, we proposed to using spatial adjacent histogram strategy for image representation. Finally, a SVM classifier is learned for medical image classification task. We also evaluated four SVM kernels into our algorithm to enhance its classification power. Based on the experiments on PAP smear, 2D-Hela, Hep-2 cell and Brain tumor datasets, we can draw a conclusion that our proposed method achieved better performance compared with other LBP-based approaches.

C1	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C2	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C3	0.00	0.00	0.72	0.00	0.17	0.00	0.11	0.00	0.00	0.00
C4	0.00	0.00	0.00	0.94	0.00	0.00	0.00	0.00	0.06	0.00
C5	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
C6	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
C7	0.00	0.00	0.13	0.00	0.00	0.00	0.88	0.00	0.00	0.00
C8	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.94	0.00	0.00
C9	0.00	0.00	0.07	0.07	0.00	0.07	0.00	0.00	0.79	0.00
C10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	

Fig. 16. Confusion matrices of the proposed method for 2D-HeLa dataset. C1:actin; C2:dna; C3:endosome; C4: er; C5:golgia; C6:golgpp; C7:lysosome; C8:mircrotubules; C9:mitochondria; C10: nucleolus.

C1	0.96	0.00	0.00	0.03	0.01	0.00
C2	0.00	0.90	0.00	0.10	0.00	0.00
C3	0.05	0.10	0.86	0.00	0.00	0.00
C4	0.05	0.02	0.00	0.80	0.12	0.00
C5	0.00	0.02	0.00	0.00	0.98	0.00
C6	0.00	0.00	0.00	0.00	0.00	1.00
C1	C2	C3	C4	C5	C6	

Fig. 17. Confusion matrices of the proposed method for Hep-2 cell dataset. C1: centromere; C2:coarse speckled; C3:cytoplasmatic; C4: fine speckled; C5:homogeneous; C6:nucleolar.

Despite the well achieved accuracy, there are still some open problems in our framework.

- (1) The dimension of our model is much larger than that of the standard LBP, and this will result in the computational time increasing linearly as training data grow in number. The scale of medical image datasets is generally not large, however, our method will cost more time when encountering general datasets with large scale. The feature selection techniques could be employed to solve this problem.
- (2) In our proposal, we focus on improving LBP in terms of encoding the micro-structures features, and we did not design patterns with rotation invariant. However, the rotation invariants are expected to further improve the performance of our model.
- (3) We mainly employ SVM with linear kernel and we also report the performance of different kernels with default parameter values. However, a number of optimization algorithms could be employed to further improve the classification performance, such as swarm intelligence optimization algorithm.

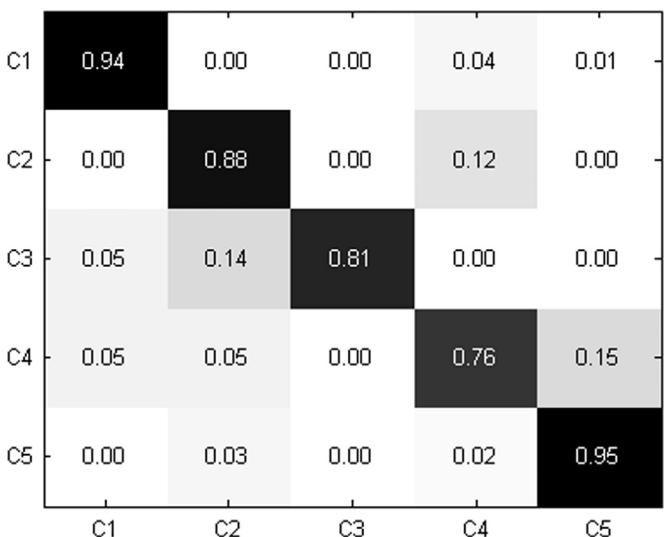


Fig. 18. Confusion matrices of the proposed method for brain tumor dataset. C1: astrocytoma; C2: tuberculum sellae meningioma; C3: olfactory groove meningioma; C4: acoustic neuroma; C5: pituitary tumor.

We will explore these problems in the future towards a more powerful medical image classification model.

Conflict of interest statement

None declared.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61472161, 61133011, 61402195, 61502198, 61303132, 61202308), Science and Technology Development Project of Jilin Province (20140101201JC), the Science and Technology Plan Project of Wenzhou of China (G20140048) and the Program of China Scholarships Council (No. 201406170116).

Appendix A. Supporting information

The *p*-values of statistical tests for compared methods mentioned in Section 3.3 are presented in Table 14.

The confusion matrices for 2D-HeLa, Hep-2 cell, and brain tumor datasets, are presented in Fig. 16, Fig. 17, and Fig. 18, respectively. For each confusion matrix, the average classification accuracies for individual classes are listed along the diagonal, and the entry in the *i*th row and *j*th column is the percentage of images from the class *i* that are misidentified as class *j*.

References

- [1] K. Doi, Computer-aided diagnosis in medical imaging: historical review, current status and future potential, *Comput. Med. Imaging Graph.* 31 (4) (2007) 198–211.
- [2] F. Lalys, L. Riffaud, X. Morandi, P. Jannin, Automatic Phases Recognition in Pituitary Surgeries by Microscope Images Classification, In: *Information Processing in Computer-assisted Interventions*, Springer Berlin Heidelberg (2010), p. 34–44.
- [3] T.F. Cootes., C.J. Taylor., Statistical models of appearance for medical image analysis and computer vision. *Medical imaging 2001, Int. Soc. Opt. Photon.* (2001) 236–248.
- [4] H. Pourhassem, H. Ghassemian, Content-based medical image classification using a new hierarchical merging scheme, *Comput. Med. Imaging Graph.* 32 (2008) 651–661.
- [5] M. Häfnera, M. Liedlgruber, A. Uhl, A. Vécsei, F. Wrba, Color treatment in endoscopic image classification using multi-scale local color vector patterns, *Medical Image Analysis*, 16, pp. 75–86.
- [6] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary pattern, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7) (2002) 971–987.
- [7] H. Zhou, R. Wang, C. Wang, A novel extended local binary pattern operator for texture analysis, *Inf. Sci.* 22 (2008) 4314–4325.
- [8] Z.H. Guo, L. Zhang, D. Zhang, X.Q. Mou, Hierarchical multiscale lbp for face and palmprint recognition, in: *Proceedings of the 17th IEEE International Conference Image Processing (ICIP 2010)*, 2010, pp. 4521–4524.
- [9] L. Nanni, A. Lumini, S. Brahnam, Survey on LBP based texture descriptors for image classification, *Expert Syst. Appl.* 39 (2012) 3634–3641.
- [10] J.F. Ren, X.D. Jiang, J.S. Yuan, Noise-resistant local binary pattern with an embedded error-correction mechanism, *IEEE Trans. Image Process.* 22 (10) (2013) 4049–4060.
- [11] Jun Shang, et al., Robust image region descriptor using local derivative ordinal binary pattern, *J. Electron. Imaging* 24 (3) (2015) 033009.
- [12] X.L. Meng, Z.Z. Wang, L.Z. Wu, Building global image features for scene recognition, *Pattern Recognit.* 5 (2012) 373–380.
- [13] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Trans. Image Process.* 19 (2010) (1675–1683).
- [14] X. Tan., B. Triggs., Enhanced local texture feature sets for face recognition under difficult lighting conditions, *IEEE Trans. Image Process.* 19 (6) (2010) 1635–1650.
- [15] C. Zhu, C.E. Bichot, L. Chen, Multi-scale color local binary patterns for visual object classes recognition, in: *Proceedings of the International Conference on Pattern Recognition (ICPR)*, 2010, pp. 3065–3068.
- [16] S. Liao, A.C.S. Chung, In: *Face Recognition by Using Elongated Local Binary Patterns with Average Maximum Distance Gradient Magnitude*, Asian Conference on Computer Vision, 2007.
- [17] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local Gabor binary pattern Histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition, *Int. Conf. Comput. Vision.* 1 (2005) 786–791.
- [18] S. Liao, M.W.K. Law, A.C.S. Chung, Dominant local binary patterns for texture classification, *IEEE Trans. Image Process.* 18 (2009) 1107–1118.
- [19] N. Senthilkumar, R. Rajesh., Edge detection techniques for image segmentation—a survey of soft computing approaches, *Int. J. Recent Trends Eng.* 1 (2) (2009).
- [20] N. Werghi, S. Berretti, A.D. Bimbo, The mesh-LBP: a framework for extracting local binary patterns from discrete manifolds, *IEEE Trans. Image Process.* 24 (1) (2015).
- [21] Z. Li, G. Liu, Y. Yang, et al., Scale-and rotation-invariant local binary pattern using scale-adaptive texton and subuniform-based circular shift, *Image Process. IEEE Trans.* 21 (4) (2012) 2130–2140.
- [22] S. Hegenbart, A. Uhl, A scale-and orientation-adaptive extension of Local Binary Patterns for texture classification, *Pattern Recognit.* 48 (8) (2015) 2633–2644.
- [23] S. Liao, A.C.S. Chung, Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude, *Asian Conf. Comput. Vision* (2007) 627–629.
- [24] T. Ahonen, J. Matas, C. He, M. Pietikäinen, Rotation invariant image description with local binary pattern histogram Fourier features, *Image Anal.* (2009) 61–70.
- [25] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Trans. Image Process.* 19 (2010) 1657–1663.
- [26] Ryusuke Nosaka, Yasuhiro Ohkawa, Kazuhiro Fukui., Feature Extraction Based on Co-occurrence of Adjacent Local Binary Patterns. *Advances in Image and Video Technology*, Springer Berlin Heidelberg (2012), p. 82–91.
- [27] A. Chebira, Y. Barbotin, C. Jackson, T. Merryman, G. Srinivasa, R.F. Murphy, J. Kovačević, A multiresolution approach to automated classification of protein subcellular location images, *BMC Bioinform.* 8 (1) (2007) 210.
- [28] P. Foggia, G. Percannella, P. Soda, M. Vento, Benchmarking HEp-2 cells classification methods, *Med. Imaging IEEE Trans.* 32 (10) (2013) 1878–1889.
- [29] J. Jantzen, J. Norup, G. Dounias, B. Bjerregaard, Pap-smear benchmark data for pattern classification, *Nat. Inspired Smart Inf. Syst.* (2005) 1–9.
- [30] C.C. Chang, C.J. Lin, LIBSVM: A Library for Support Vector Machines, 2001. Software available from: <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>.
- [31] H. Byun, S.W. Lee, Applications of Support Vector Machines for Pattern Recognition: A Survey[ml]/Pattern Recognition with Support Vector Machines, Springer Berlin Heidelberg (2002), p. 213–236.
- [32] G.B. Huang, D.H. Wang, Y. Lan, Extreme learning machines: a survey, *Int. J. Mach. Learn. Cybern.* 2 (2) (2011) 107–122.
- [33] S. Maji, A.C. Berg, J. Malik, Classification using intersection kernel support vector machines is efficient[C]/Computer Vision and Pattern Recognition, 2008, CVPR 2008, IEEE Conference on IEEE, 2008: pp. 1–8.
- [34] J.M. Bland, D.G. Altman, Multiple significance tests: the Bonferroni method, *BMJ* 310 (6973) (1995) 170.
- [35] S. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, *Comput. Vision. Pattern Recognit. (CVPR)* (2006).