Histopathological Image Classification Using Random Binary Hashing Based PCANet and Bilinear Classifier

Jinjie Wu, Jun Shi
School of Communication and Information Engineering
Shanghai University
Shanghai, China
junshi@staff.shu.edu.cn

Yan Li
College of Computer Science and Software Engineering
Shenzhen University
Shenzhen, China
liyan@szu.edu.cn

Jingfeng Suo, Qi Zhang
School of Communication and Information Engineering
Shanghai University
Shanghai, China
zhangq@shu.edu.cn

***Abstract*—The computer-aided histopathological image diagnosis has attracted considerable attention. Principal component analysis network (PCANet) is a novel deep learning algorithm with a simple network architecture and parameters. In this work, we propose a random binary hashing (RBH) based PCANet (RBH-PCANet), which can generate multiple randomly encoded binary codes to provide more information. Moreover, we rearrange the local features derived from PCANet to the matrix-form features in order to reduce feature dimensionality, and then we apply the low-rank bilinear classifier (LRBC) to perform effective classification for matrix features. The proposed classification framework using RBH-PCANet and LRBC (RBH-PCANet-LRBC) is adopted for histopathological image classification. The experimental results on both a hepatocellular carcinoma image dataset and a breast cancer image dataset show that the RBH-PCANet-LRBC algorithm achieves best performance compared with other unsupervised deep learning algorithms.**

Keywords-PCANet; Random Binary Hashing; Low Rank Bilinear Classifier; Histopathological Image

I. INTRODUCTION

The histopathology based diagnosis remains the "gold standard" for a large number of diseases, including almost all types of cancer. The computer-aided diagnosis (CAD) of histopathological images has attracted much attention.

Feature representation is a critical factor for histopathological image analysis. In recent years, deep learning (DL) has achieved great success for medical images, also including histopathological images [1][2].

Principal component analysis network (PCANet) is a novel DL algorithm, which consists of only three basic components: cascaded PCA as a deep network, binary hashing, and block-wise histograms [3]. Although PCANet has a very simple network architecture and parameters, it indeed achieves excellent performance for image classification compared with other state-of-the-art DL algorithms [3][4]. Moreover, several improved PCANet algorithms have been proposed, such as DLANet [5], SRDANet [6], and SPCANet [7]. However, these variants mainly focus on applying different filters instead of the PCA filter for further improving representation performance.

In PCANet and its variants, the binary hash approach just simply encodes the quantized binary codes according to the sequence of principal components (PC). In other words, the first PC is assigned the most significant bit. However, this simple hashing method cannot provide information rich enough. In fact, the way of binary coding in PCANet also affects representation performance, but it has not been deeply investigated yet.

Random theory based methods, such as the random subspace and random projection algorithms, have been widely used in machine learning. In this work, motivated by the success of random methods, we propose performing random binary hashing (RBH) in PCANet without considering the sequence of PCs, which then generates multiple random binary codes with more information.

Furthermore, the local histogram features extracted from PCANet are usually concatenated to form a feature vector, which has a very high dimensionality resulting in high computational complexity. One solution is to rearrange these local features to a matrix-form, which is then fed to the matrix feature based classifier [8]. The low-rank bilinear classifier (LRBC) is a newly proposed classifier for matrix-form features with outstanding performance [8]. Therefore, it has the potential to generate matrix-form features from PCANet for more effective classification.

In this work, we propose a new classification framework for histopathological images with RBH-based PCANet

This work is supported by the National Natural Science Foundation of China (61471231, 61401267, 61471245, U1201256, and 61201042), the Innovation Program of Shanghai Municipal Education Commission (13YZ016), and the Natural Science Research Projects of Shenzhen City (JCYJ20140418091413514 and ZDSYS20140508141148477).

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

(RBH-PCANet) and LRBC. The main contributions are twofold. First, we propose the RBH-PCANet algorithm to generate more representative information. Second, we rearrange local histogram features derived from RBH-PCANet to matrix-form features, which are then fed to the LRBC algorithm to further improve classification performance for histopathological image based diagnosis.

II. Methods

Figure 1 shows the proposed classification framework, including RBH-PCANet, spatial pyramid matching (SPM), and LRBC. The RBH-PCANet algorithm is applied to histopathological images to generate local histogram features, which are then pooled and hierarchically re-organized by SPM, because SPM can integrate the spatial information in an image from local block histograms to generate more efficient and compact image feature representation instead of simply concatenated local features [9]. Furthermore, the features from SPM are rearranged to a matrix-form in a new matrix feature space, which are finally fed to LRBC for classification.

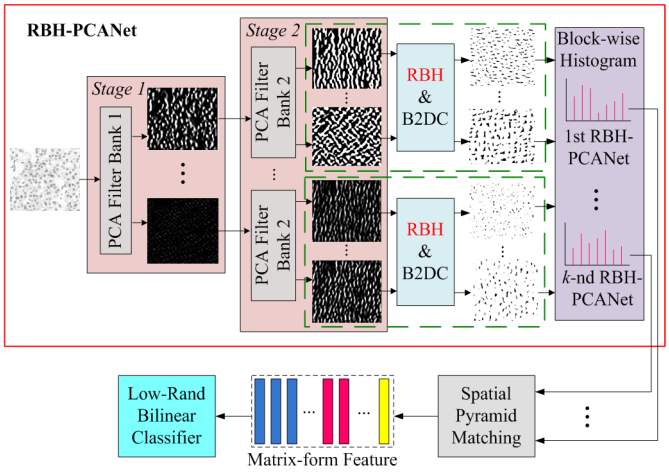


Figure 1. The flow diagram of RBH-PCANet, SPM and LRBC based classification framework.

A. Cascaded PCA Network

PCANet consists of three basic components: cascaded PCA, binary hashing, and block-wise histograms after binary to decimal conversion (B2DC) [3]. In the first stage of PCA network, a patch with a size of $m \times n$ is taken for each pixel in N training images $\{I_i\}_{i=1}^{N}$. PCA can be regarded as minimizing the reconstruction error within a family of orthonormal filters, namely

$$\min_{V \in R^{mn \times L_1}} \|X - VV^T X\|_F^2, \quad \text{s.t.} \quad V^T V = I_{L_1} \tag{1}$$

where X is the patch image training set. Then we can achieve L_1 leading components q_l $(l = 1,2 \ldots, L_1)$ for patches around each pixel in the first layer PCA network. The PCA filters of the first stage are expressed as

$$W_l^1 = mat_{m,n}(q_l) \in R^{m \times n}, \quad l = 1,2 \ldots, L_1 \tag{2}$$

where $mat_{m,n}(q)$ denotes a function that maps $q \in R^{mn}$ to a matrix $W \in R^{m \times n}$, and q_l is the l-th leading eigenvector.

The l-th filter output of the first stage is given by

$$\tilde{I}_1^l = I_1 * W_l^1, \quad i = 1,2, \ldots, \tag{3}$$

where $*$ denotes 2D convolution.

Almost repeating the same process in first stage, the next layer network can be easily built. The PCA filters of the second stage are then obtained as

$$W_l^2 = mat_{m,n}(q_l^2) \in R^{m \times n}, \quad l = 1,2 \ldots, L_2 \tag{4}$$

where L_2 is the number of leading components in the second layer PCA network. For each input I_i^l of the second stage, L_2 outputs are generated and each convolves $\tilde{I}_i^l$ with W_l^2

$$\Theta_i^l = \{\tilde{I}_i^l * W_l^2\}_{l=1}^{L_2}, \quad l = 1,2, \ldots, L_2 \tag{5}$$

After filtering by the two-stage cascaded PCA networks, there are $L_1 \times L_2$ real-valued output images in total. It is worth noting that more PCA stages can be built by simply repeating the above process.

B. Random binary hashing and block-wise histograms

In the original PCANet, a simple binary quantization (hashing) operator is performed on the $L_1 \cdot L_2$ output images by $\{H(\tilde{I}_i^l * W_l^2)\}_{l=1}^{L_2}$, where $H(\cdot)$ is a function whose value is one for positive entries and zero otherwise [3]. The binary pixel values at the same location are then regarded as an L_2-bit vector according to the sequence of PCs.

It is worth noting that the current binary hashing operator is one of the $(L_2)^2$ binary codes for L_2-bit vector; therefore, only limited information is obtained, which affects classification performance.

Motivated by the success of random methods in machine learning, we propose to randomly encode the L_2 binary pixel values k times without considering the sequence of PCs, which then generate k L_2-bit vectors. As a consequence, the k vectors provide more useful information to represent images.

B2DC is then performed after RBH, namely, each L_2-bit vector is then converted back into a new single integer-valued image:

$$\tilde{T}_i^l = \sum_{l=1}^{L_2} 2^{l-1} H(\tilde{I}_i^l * W_l^2) \tag{6}$$

Each of the final $k \cdot L_1$ images, it is partitioned into multiple blocks with a size of $m \times n$, and the decimal values are then counted in each block to yield local histogram features.

C. Spatial Pyramid Matching

An input image generates totally $k \cdot L_1$ images by RBH-PCANet. For the local histogram features, the SPM algorithm is then used for each image to effectively integrate spatial information in image by hierarchical pooling. For detailed algorithm about SPM, readers can refer to [9].

As shown in Fig. 2, the multi-scale pooled local features in different layers of SPM are rearranged to a matrix-form. It should be noted that for L_1 images belonging to the same RBH codes, their local histogram features at the same position are concatenated to form a vector with a dimensionality of $256L_1$, which is then processed by SPM to build a 2D matrix of

features with a size of $256L_1 \times Ns$, where Ns is the total number of spatial bins in the multi-level pyramid. In Fig. 2 for example, the first, second and third levels of the pyramid have 16, 4, and 1 bins respectively, and thus $Ns = 21$. Finally, k feature matrices are generated from k RBH encoding sequences, each with a size of $256L_1 \times Ns$, and they are connected to form a large matrix with a size of $256L_1 \times (Ns \cdot k)$, which serves as the input of LRBC. On the contrast, for a traditional vector-based classifier, features from Ns bins and k RBH sequences are concatenated to form a large vector of features (Fig. 2).

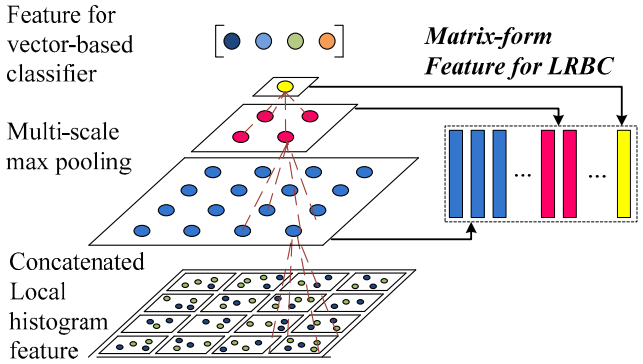


Figure 2. Rearrangement of matrix-form features from SPM.

D. Low-Rank Bilinear Classifier

The matrix-form features are then fed to LRBC for classification, which is introduced as follows [8].

Let X be an $h \times w$ feature matrix ($X \in R^{h \times w}$). The multi-rank bilinear classifier is given by

$$\hat{y} = tr(M_h^T X M_w) + b = tr(M^T X) + b \quad (7)$$

where $tr(\cdot)$ is the trace of a matrix, b is the bias, and $M = M_h M_w^T \in R^{h \times w}$ is the classifier matrix ($M_h \in R^{h \times r}, M_w \in R^{w \times r}$ where $r \leq \min[w, h]$).

In the maximum-margin framework, the margin of bilinear classifier by matrix trace norm is measured to minimize the matrix rank, which results in the following optimization problem for a binary classification problem

$$\min_{M,b} \|M\|_\epsilon + C \sum_{i=1}^{n} \max[0, 1 - y_i\{tr(M^T X_i) + b\}] \quad (8)$$

where $\|M\|_\epsilon$ indicates the trace norm of M for n sample pairs $\{X_i, y_i\}$.

The smoothing regularization is then used, which is expressed by the quardratic form derived from Laplacian of the weight with the formulation

$$\min_{M_w, M_h, b} \{\frac{1}{2} tr(M_w M_w^T) + \frac{1}{2} tr(M_h M_h^T) + \frac{1}{2} C_w tr(M_w^T L_w M_w) + \frac{1}{2} C_h tr(M_h^T L_h M_h) + C \sum_i \max[0, 1 - y_i\{tr(M_h^T X_i M_w) + b\}]\} \quad (9)$$

where L_w, L_h are the matrices to measure the smoothness, and C_w, C_h are regularization parameters.

In this work, C_w is set by $C_w = \frac{1}{\|L_w\|_s}$, to equally balance the two spectral matrix norms of I and L_w with $C_h = 1/\|L_h\|_s$ for L_h, where $\|L_w\|_s$ indicates the spectral matrix norm of L_w. Eq. (9) is then rewritten as

$$\min_{\bar{M}_w, \bar{M}_h, b} \{\frac{1}{2} tr(\bar{M}_w \bar{M}_w^T) + \frac{1}{2} tr(\bar{M}_h \bar{M}_h^T) + C \sum_i \max[0, 1 - y_i\{tr(\bar{M}_h^T \bar{X}_i \bar{M}_w) + b\}]\} \quad (10)$$

where $\bar{X}_i = (I + C_h L_h)^{-\frac{1}{2}} X_i (I + C_w L_w)^{-\frac{1}{2}}$.

Eq. (10) is the final objective function of the LRBC algorithm, from which we get the smoothed classifier weights $M_h = (I + C_h L_h)^{-\frac{1}{2}} \bar{M}_h$ and $M_w = (I + C_w L_w)^{-\frac{1}{2}} \bar{M}_w$. More detailed optimization solutions and other information about LRBC can be found in [8].

III. Experiments And Results

A. Experiments

Two histopathological image datasets, namely a hepatocellular carcinoma (HCC) image dataset and a breast cancer image dataset, are used in this work. There are 66 HCC images with a size of 1024×768, including 21 well differentiated HCC, 23 moderately differentiated HCC and 22 poorly differentiated HCC [11]. The breast cancer image dataset consists of 20 ductal carcinoma in situ (DCIS) images and 31 usual ductal hyperplasia (UDH) images from the Beth Israel Deaconess Medical Center dataset with a size of 1444×901 [12]. Typical examples are shown in Fig. 3. Here the color images are transformed into grayscale images for feature extraction and image classification.

RBH-PCANet is compared with PCANet, sparse coding (SC) [10], and two commonly used unsupervised DL algorithms, namely stacked auto-encoder (SAE) and deep belief network (DBN) [13]. For fair comparison, both SAE and DBN are set with three-layer networks. The parameters in all the DL algorithms are selected to achieve the best performance. The LRBC algorithm is compared with the linear support vector machine (SVM) classifier.

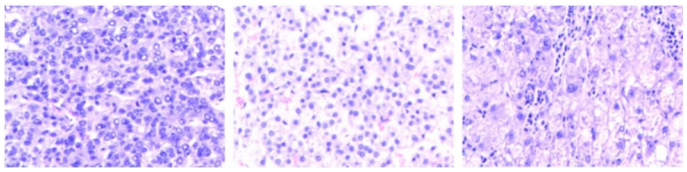
For each dataset, 10000 patches are randomly sampled from the training set to learn feature representation models for all algorithms. The patch sizes are 17×17 and 11×11 pixels for HCC images and breast images, respectively. The local patches are densely sampled without overlapping for SPM in each image. The leave-one-out strategy is used to evaluate the classification performance. We repeat the above experiments five times to calculate the averaged results.

In RBH-PCANet, the parameters are empirically set as follows: $k = 4$, $L_1 = 8$, $L_2 = 8$ and $Ns = 21$.

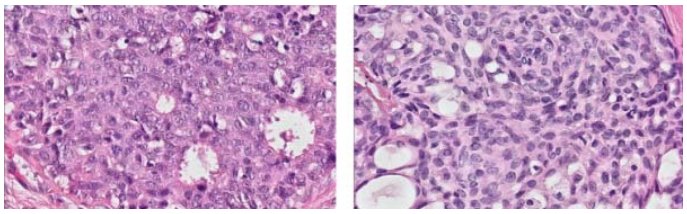
B. Results

Table 1 shows the classification results of different algorithms on HCC image dataset. It can be seen that RBH-PCANet is significantly superior to PCANet (p-value < 0.05), while PCANet outperforms SC, DBN and SAE. Moreover, the LRBC algorithm also significantly improves the classification performance compared with the feature-vector

based SVM (p-value < 0.05). The proposed RBH-PCANet-LRBC algorithm achieves the best results, whose classification accuracy (ACC), sensitivity (SEN) and specificity (SPE) are 95.45±1.07%, 97.38±1.02% and 98.04±0.67%, respectively.



(a)



(b)

Figure 3. Typical histopathological images. (a) Images of poorly, moderately and well differentiated hepatocellular carcinoma from left to right; (b) Images of ductal carcinoma in situ and ductal hyperplasia images from left to right.

TABLE I. CLASSIFICATION RESULTS OF DIFFERENT ALGORITHMS ON A HEPATOCELLULAR CARCINOMA IMAGE DATASET (UNIT: %)

	ACC	SEN	SPE
SC-SVM	85.45±2.30	85.11±2.21	92.73±1.14
DBN-SVM	82.12±2.71	82.22±2.75	91.12±1.38
SAE-SVM	85.15±2.71	84.96±2.78	92.57±1.36
PCANet-SVM	89.09±0.68	88.79±0.71	94.49±0.34
RBH-PCANet-SVM	93.03±0.83	92.89±0.82	96.50±0.42
PCANet-LRBC	94.54±2.03	94.52±2.09	97.29±1.03
RBH-PCANet-LRBC	**95.45±1.07**	**97.38±1.02**	**98.04±0.67**

TABLE II. CLASSIFICATION RESULTS OF DIFFERENT ALGORITHMS ON A BREAST CANCER IMAGE DATASET (UNIT: %)

	ACC	SEN	SPE
SC-SVM	72.94±2.56	42.00±5.70	92.90±2.70
DBN-SVM	70.20±4.25	52.00±4.47	83.23±1.44
SAE-SVM	76.47±3.10	63.00±2.73	85.16±4.89
PCANet-SVM	74.90±2.97	54.00±4.47	88.39±3.95
RBH-PCANet-SVM	75.29±4.06	47.00±8.37	**93.55±3.23**
PCANet-LRBC	76.08±5.65	66.00±5.71	82.59±8.22
RBH-PCANet-LRBC	**78.46±3.92**	**71.00±4.18**	83.23±5.30

Table 2 enumerates the classification results on the breast cancer image dataset, which shows a similar tendency to those in Table 1. The RBH-PCANet-LRBC algorithm is again superior to all other algorithms with a classification accuracy of 78.46±3.92% and a sensitivity of 71.00±4.18%. Our method RBH-PCANet-LRBC has significant improvement compared with SC, DBN and SAE (p-value < 0.05).

IV. CONCLUSIONS

In this study, we propose a RBH-based PCANet (RBH-PCANet) algorithm to effectively learn feature representation for histopathological images. RBH-PCANet is significantly superior to the original PCANet. Moreover, we rearrange the histogram features generated by SPM to form a matrix, which is fed to the matrix feature based LRBC algorithm for further improving the classification performance. The proposed RBH-PCANet-LRBC algorithm achieves the best performance for diagnosis of histopathological images, which suggests that RBH-PCANet-LRBC has the potential for histopathology-based CAD systems.

The computational load of RBH-PCANet-LRBC for the binary hashing and block-wise histograms is k times that of PCANet. However, the convolution operation in both methods is the most time-consuming, and its computing time is the same for both methods. Hence in total, RBH-PCANet-LRBC is barely inferior to PCANet in terms of computational complexity.

We transform color histopathological images to grayscale images for feature extraction and image classification. Indeed, color information is also important for the histopathology based diagnosis. Future work will be focused on effectively exploring intra-correlations among multiple color channels and fusing information from multi-channels.

REFERENCES

[1] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in the 16th Int. Conf. Med. Image Comput. Comput. Assist Interv., vol. 16, pp. 411-418, 2013.

[2] M. Veta, J. P. W. Pluim, P. J. van Diest, and M. A. Viergever. "Breast cancer histopathology image analysis: a review," IEEE Trans. Biomed. Eng., vol. 61, pp. 1400-1411, 2014.

[3] T. H. Chan, K. Jia, S. H. Gao, J. W. Lu, Z. N. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification," IEEE Trans. Image Process., vol. 24, pp. 5017-5032, 2015.

[4] S. K. Wang, L. Chen, Z. X. Zhou, X. Sun, and J. Y. Dong, "Human fall detection in surveillance video based on PCANet," Multimed. Tools Appl., pp. 1-11, 2015.

[5] Z. Y. Feng, L. W. Jin, D. P. Tao, and S. P. Huang, "DLANet: a manifold-learning-based discriminative feature learning network for scene classification," Neurocomputing, vol. 157, pp. 11-21, 2015.

[6] L. Tian, C.X. Fan, Y. Ming, and J.k. Shi, "SRDANet: an efficient deep learning algorithm for face analysis," Int. Robot. Appl., vol. 9244, pp. 499-510, 2015.

[7] L. Tian, C.X. Fan, Y. Ming, and Y. Jin, "Stacked PCA network (SPCANet): an effective deep learning for face recognition," in Int. Conf. Digit. Signal Process. Proc., pp. 1039-1043, 2015.

[8] T. Kobayashi, and N. Otsu, "Efficient optimization for low-rank integrated bilinear classifiers," in the 12th Eur. Conf. Comput. Vis., pp. 474-487, 2012.

[9] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern. Recognit., pp. 2169-2178, 2006.

[10] J. C. Yang, K. Yu, Y. H. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in Proc. IEEE

Comput. Soc. Conf. Comput. Vis. Pattern. Recognit., pp.1794-1801, 2009.

[11] J. Shi, Y. Li, J. Zhu, H.J. Sun, and Y. Cai, "Joint sparse coding based spatial pyramid matching for classification of color medical image," Comput. Med. Imag. Grap., vol.41, pp. 61-66, 2015.

[12] F. Dong, H. Irshad, and E.Y. Oh., "Computational pathology to discriminate benign from malignant intraductal proliferations of the breast," PLOS One, 0114885, 2014.

[13] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," IEEE Trans. Pattern. Anal. Mach. Intell., vol. 35, pp. 1798-1828, 2013.