Multi-Instance Learning for Skin Biopsy Image Features Recognition

Gang Zhang[1,2,#], Xiangyang Shu[3,#], Zhaohui Liang[4], Yunting Liang[3], Shuyi Chen[2], Jian Yin[1]
[1] School of Computer Science and Technology, SUN YAT-SEN University, Guangzhou, China, 510275
[2] School of Automation, Guangdong University of Technology, Guangzhou, China, 510006
[3] Department of Dermatology and Venerology, Third Affiliated Hospital of SUN YAT-SEN University Guangzhou, China, 510630
[4] Acupuncture Department, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, 510120
Corresponding author: Jian Yin. Email: issjyin@mail.sysu.edu.cn
[#] These authors contributed equally to this work.

***Abstract*—In this paper, a multi-instance learning framework is introduced to solve the problem of skin biopsy image features recognition. Previously reported methods for skin surface images were mostly based on color features extraction. They are incapable to be directly applied to skin biopsy image features recognition because biopsy images are often dyed and have obvious inner structures with different textures. Therefore, we regard skin biopsy images as multi-instance samples, whose instances are regions or structures captured by applying Normalized Cut. Texture feature extraction methods are used to express each region as a vectorial expression. Then two multi-instance learning algorithms reported successful in various image retrieval tasks were applied. Nine features were manually selected as target features to evaluate the proposed method on a skin disease diagnosis datasets of 6579 biopsy images from 2010 to 2011. The result showed that the proposed method is effective and medically acceptable.**

Keywords-skin biopsy image feature recognition; multi-instance learning; Citation KNN; mi-Graph; texture feature extraction

I. INTRODUCTION

The prevalence of skin disease is high around the world [1][2][3]. There are more than 3000 kinds of skin disease which are capable to be diagnosed in clinical practice, and a large proportion of them are common in public life [1][4]. Meanwhile, these skin disorders need to be carefully distinguished from some malignant cancers and sexually transmitted diseases to prevent them from transmission and threatening the health of the patients or other people around them [5][6]. Consequently, there is a great demand to seek an effective method for doctors to rapidly identify and correctly distinguish these disorders [1][3].

In general, the diagnosis of skin disease is divided into two steps [7]. The first step is skin surface inspecting. A diagnosis is roughly reached after routine physical exams including observation and physical examination of skin lesions. The second step is skin biopsy image analysis [8]. It is accepted that the second step is a complement of the first one [9][10], in case that doctor has less confidence or even can't make decision only by inspecting skin surface. Skin biopsy image analysis can provide further information about the morbid condition that under the skin surface at in a microscopic vision [9][11]. Thus its result can more accurately explain the true problem than merely the signs on the surface. Towards a medical acceptable diagnosis, it usually requires many skin biopsy image cases to identify the significant changes and differentiate them from similar skin diseases [12]. Since understanding skin biopsy images requires more professional knowledge and profound experience [13] than inspecting skin surface, it becomes a great challenge for doctor to correctly recognize huge amount of skin biopsy images.

Currently many literatures reported that with the help of image understanding technology, some kinds of skin disease images can be automatically classified by machine learning classifiers [14]. Most publicly reported work considered only skin surface images classification problem, while recognition of skin biopsy images was seldom reported. The reason behind this phenomenon may be that skin surface images of most skin diseases are well featured for recognition, while skin biopsy images are often dyed and have inner structures, and to recognize them requires much professional knowledge and experience. Dyed images mean that simple color features vector may not be an effective representation for classification. Inner structures would lead to a more general learning problem, in which only one or some structures contribute to the final diagnosis. So it requires different methods to solve the problem of skin biopsy images recognition.

Multi-instance learning was introduced by Dietterich et al. [15] to solve drug activity prediction problem. Different to single-instance learning, samples in multi-instance learning (also called bags) contain several instances while concept labels are attached to samples and labels of instances are unknown. For binary classification task, a bag is positive if and only if it contains at least one positive instance, and negative otherwise. Multi-instance learning is to predict labels of unseen bags by training a classifier with labeled bags. Since multi-instance learning is suitable for many learning tasks, it becomes a hot topic in machine learning research recently. Many multi-instance learning algorithms were proposed [16][17][18][19] and some extended multi-instance assumptions were reviewed in [20].

In image understanding, multi-instance learning based methods have been studied in many reports. Chen et al. [17] attempted to introduce multi-instance learning into image classification. They proposed to express images as multi-instance samples by a clustering based method. Then apply a multi-instance version of SVM for learning. However, their

978-1-4673-2560-8/12/$31.00 ©2012 IEEE

image cutting method is not suitable for our task, since skin tissue is consist of several different structures microscopically and doctor needs to observe them respectively to describe features of abnormal parts. Regions of skin biopsy image should be divided according to structures of skin tissue so as to come up with the feature description of each part whereas the generated points of clustering based algorithms may not be contiguous in region division. Wang et al. [21] proposed a graph-based multi-instance learning method for object-based image retrieval.

The main reason of applying multi-instance learning for skin biopsy images feature recognition is that we don't want to manually label normal and lesion regions which requires large amount of human efforts. A most recent work on skin surface images classification [14] built their training and testing dataset by manually labeled normal and lesion regions, as shown in Fig. 1. Features extraction is applied in these two labeled regions for both training and test images.

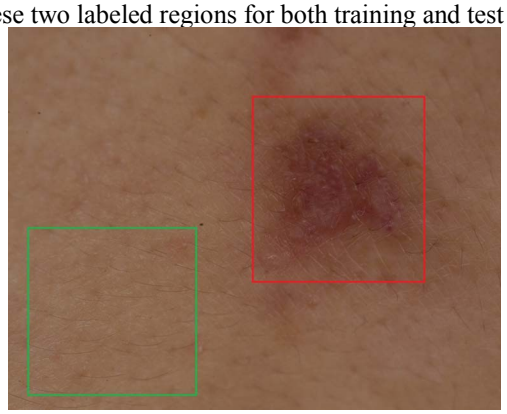


Figure 1. Manually labeled regions for skin surface image retrieval. (Red: Lesion, Green: Normal)

We argue that labeling normal and lesion regions for each image is not suitable for skin biopsy image, since it requires much more professional knowledge than labeling skin surface images. For example, in Fig. 2 there are 3 features recognized from its diagnosis record: hyperkeratosis, acanthosis and infiltration of lymphocytes. We figure out which regions of the image lead to these features, as indicated in Fig. 2, in which regions A, B, and C refer to the above three features respectively. In our case, it is not feasible to label normal and lesion regions before recognizing a biopsy image. Hence we propose to apply multi-instance learning algorithm to automatically detect whether a biopsy image contains regions with concerned features.

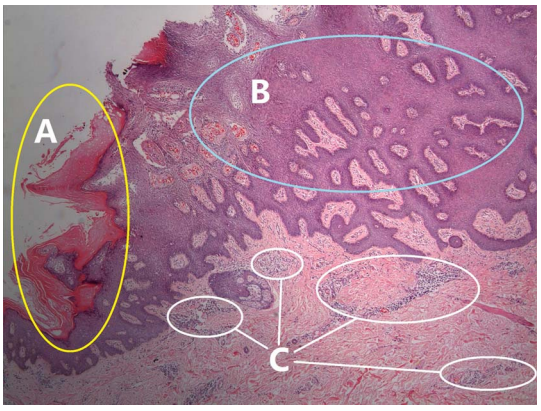


Figure 2. Lesion region of an sample biopsy image. A: *hyperkeratosis*, B: *acanthosis*, C: *infiltration of lymphocytes*

Before going further, it is worth noting that by applying multi-instance learning, meaningful regions must be effectively recognized. It is well accepted that biopsy images contain obvious inner structures and when recognizing single features, relationships between structures are seldom concerned. Hence we apply a famous image regions detection method Normalized Cut [22] to detect regions in a skin biopsy image. Fig. 3 illustrates regions generated by 7-region Normalized Cut marked with features. Note that the label *dummy* refers to regions not having any concerned features.

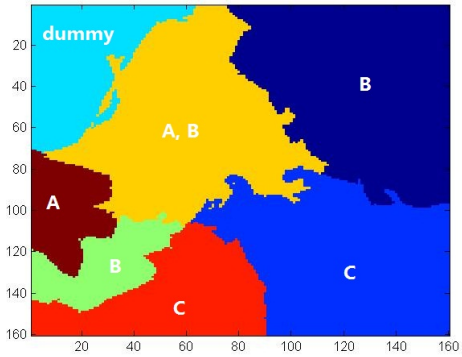


Figure 3. 7-Region normalized cut biopsy image

By extracting feature vector of each region, it is possible to express each image as a multi-instance sample (also call bag), among which regions are instances. An image has a certain feature if and only if it contains at least one region with the same feature. Moreover, in our paper features are only attached to images, but not regions. Thus we transfer the original features recognition problem into a standard multi-instance learning problem. Our method is suitable for pathological changes identification of skin biopsy image. Fig. 4 sketches the framework of this work.

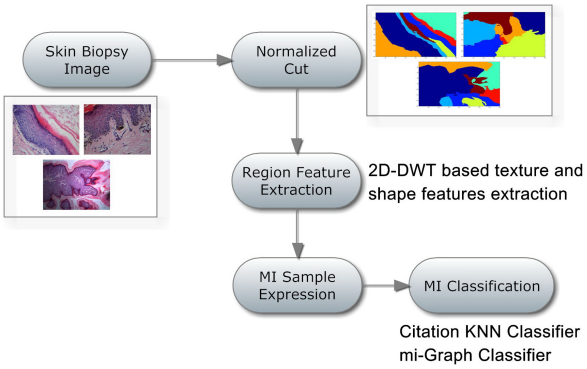


Figure 4. Main Framework

This paper is organized as following. In Section II we present Normalized Cut and the method for individual region feature extraction, which are the foundation of expressing biopsy images as multi-instance samples. In Section III we present two famous multi-instance learning algorithms, which are the main algorithms for this paper. Evaluation settings and result are presented in Section IV, followed by some discussion. Finally we conclude the paper in Section V.

II. NORMALIZED CUT AND FEATURE EXTRACTION

This section presents how to transfer skin biopsy images into multi-instance samples. As illustrated in Fig. 4, a given image is divided through Normalized Cut, and then feature extraction is applied to each generated region.

A. Normalized Cut

Normalized Cut was proposed by Shi et al. [22] in 2000, aiming at extracting perceptual grouping of a given image. Different from clustering based image segmentation algorithms, e.g. [17], Normalized Cut extracts global impression of a given image, i.e. disjoint visual grouping. To make the paper self-contained, we briefly review the main idea of Normalized Cut presented in [22].

Normalized Cut regards segmentation of an image as a graph cutting problem. It constructs a fully connected with pixels of a given image. Vertexes of the constructed graph are pixels and weights of edges are similarity between pixels. The problem of Normalized Cut is to find a cutting that minimizes in-segment similarity and maximizes cross-segment similarity. Formally, suppose there is a graph $G = (V, E)$ we aim at finding an optimal cutting to partition it into two disjoint sets A and B where $A \cap B = \emptyset$ and $A \cup B = V$. A measure is defined in Eq. 1 as optimal graph cutting.

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \quad (1)$$

where $cut(A,B) = \sum_{u \in A, v \in B} w(u,v)$, and $w(u,v)$ is weight of edge between vertexes u and v. $assoc(A,V) = \sum_{u \in A, t \in V} w(u,t)$ is sum weights of edges between vertexes in segment A and all the other vertexes in graph G. Since graph G is fully connected, a binary column vector $x_{|V| \times 1}$ is defined to indicate whether a vertex belongs to subset A. The problem of Normalized Cut is to find a cutting that minimizes $Ncut(A,B)$, as Eq. 2 shows.

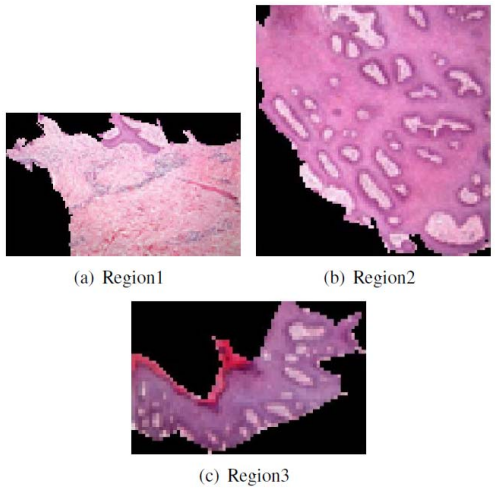
$$min_x Ncut(x) \quad (2)$$

According to [22], solution of Eq. 2 captures visual segment of the image, whose underlying idea is naturally suitable for our task. They rewrote Eq. 2 to a standard Rayleigh quotient [23] which can be solved effectively.

B. Region Feature Extraction

By applying Normalized Cut on a skin biopsy image, we can get several contiguous regions, as illustrated in Fig. 3. To further express each region as a vector for multi-instance learning, we then perform a feature extraction procedure for each region. Our method is based on the feature extracting method proposed in [17], which applies 2D wavelet transformation in LUV color space for each block ($m * m$ pixels) within a given image. Different from [17], we use Normalized Cut instead of clustering to generate regions. Algorithm 1 describes main steps of feature extraction of a single region.

It can be seen that regions obtained through Normalized Cut are irregular, i.e. not rectangles, as illustrated in Fig. 5. Each region is padded with black pixels to form a minimal rectangle. Then decompose the padding image into $m * m$ pixels blocks. In this work we set $m = 4$. If a given image is p by q, then there should be $(p/m) * (q/m)$ blocks. We remove totally black blocks to strengthen captured features of the generated regions.



(a) Region1 (b) Region2

(c) Region3

Figure 5. Cropped irregular regions with black padding

Algorithm 1 is a modified version of the method presented in [17], since it doesn't cluster on subspace $(t1, t2, t3)$ to find regions. In line 10, only L part of an input region is considered in 2D wavelet transformation. Each returned band coefficient is a 4-ary vector. All blocks' moments of LH, HL and HH bands are calculated from line 6 to line 15, whose mean, i.e. $(f4, f5, f6)$ are effective for expressing texture features [25]. The same as [24], another three features $(f7, f8, f9)$, i.e. normalized inertia of order 1 2 and 3, are computed to express shape features.

Algorithm 1 Biopsy image feature extraction

Require:
Image of padding region in RGB color space, IM;
Block size, m;

Ensure:
Feature vector for input image, F;

Divide IM_LUV into blocks of size $m * m$
$f1 = mean(IM\_LUV.L)$ //Feature1
$f2 = mean(IM\_LUV.U)$ //Feature2
$f3 = mean(IM\_LUV.V)$ //Feature3
$IM\_LUV = colorspace('RGB \Rightarrow LUV', IM)$;
for each block b in IM_LUV **do**
  if b is all black **then**
    discard this block
  end if
  $[LL, LH, HL, HH] = 2D - DWT(b.L)$
  $t1 = (1/4 * \sum_{i=1}^{4} {LH_i}^2)^{1/2}$
  $t2 = (1/4 * \sum_{i=1}^{4} {HL_i}^2)^{1/2}$
  $t3 = (1/4 * \sum_{i=1}^{4} {HH_i}^2)^{1/2}$
  $t1, t2, t3$ add to S_LH, S_HL and S_HH
end for
$f4 = mean(S\_LH)$ //Feature4
$f5 = mean(S\_HL)$ //Feature5
$f6 = mean(S\_HH)$ //Feature6
$f7, f8, f9$ are normalized inertia of order 1, 2, 3 [24]
$F = (f1, f2, f3, f4, f5, f6, f7, f8, f9)$
return F;

The motivation to propose this feature extraction method for our task is two-fold. On one hand, it effectively extracts both shape and texture features essential for features recognition within regions. On the other hand, since it performs per block wavelet transformation, blank blocks can be easily discarded so as to get more precise representations for irregular regions. We regard this feature extraction method potentially close to the intuition presented in many famous dermatology literatures [7][8][9].

III. MULTI-INSTANCE LEARNING ALGORITHM

We use two famous multi-instance learning algorithms to build our model. They are Citation KNN and mi-Graph. Citation KNN was proposed by Jun Wang et al. [26], which can be regarded as a multi-instance version of traditional KNN classifier. However, to determine a given test bag's label, it considers not only K nearest labeled bags, i.e. references, but also labeled bags that regard the given bag as a K nearest neighbor, i.e. citers. Citation KNN is well studied and has many successful applications in many learning tasks. mi-Graph was proposed by Z.-H.Zhou [27] whose idea was to consider both instances and relation between instances within a bag in training a learner. mi-Graph regards each bag as a fully connected graph, whose nodes are instances and weights of edges are similarity between pairs of instances. mi-Graph defines a kernel function between bags with graph representation and plugs this kernel into a support vector machine classifier for training and testing.

The motivation of adopting these two algorithms is that we want to show empirically that for skin biopsy image features recognition task, how well these algorithms based on different multi-instance assumptions [20] work. Roughly speaking, Citation KNN falls into original multi-instance assumption which was firstly proposed by in Dietterich [15]. A bag is positive if it contains at least one positive instance, and negative otherwise. Note that in original multi-instance assumption, relationship between instances is not considered. mi-Graph is an algorithm with consideration of bag structures, i.e. relationship between instances. These two algorithms have been reported to be well-performed in both multi-instance learning benchmark dataset [15] and image retrieval tasks.

A. Citation KNN

Citation KNN is a multi-instance learning algorithm inspired by paper citation and reference. To determine the label of a test bag X, it considers not only neighbors (references) of X, but also the bags (citers) that regard X as a neighbor. Citation KNN uses both references and citers to determine an unseen bag's concept label. The key problem is to how to evaluate distance between bags so as to find neighbors and citers. Citation KNN makes use of a modified version of Hausdorff distance to describe distance between bags. Intuitively speaking, the distance function of Citation KNN evaluates the smallest and largest distance between instances of two bags.

B. mi-Graph

mi-Graph [27] is a kernel function considering both instances representation and relationship between instances. It formulizes a bag as a fully connected graph and defined a kernel between bags. Formally, given two bags $X_i = (x_{i1}, \cdots, x_{in_i})$ and $X_j = (x_{j1}, \cdots, x_{jn_j})$, mi-Graph kernel is defined as Eq. 3.

$$k_g(X_i, X_j) = \frac{\sum_{a=1}^{n_i} \sum_{b=1}^{n_j} W_{ia} W_{jb} k(x_{ia}, x_{jb})}{\sum_{a=1}^{n_i} W_{ia} \sum_{b=1}^{n_j} W_{jb}} \quad (3)$$

where $k(\cdot, \cdot)$ is a popular kernel function defined in instances space, $W_{ia} = 1/\sum_{u=1}^{n_i} w_{au}^i$, $W_{jb} = 1/\sum_{v=1}^{n_j} w_{bv}^j$, w_{au}^i are elements of threshold Gram matrix W^i for instances in bag X_i. If the distance between instances x_{ia} and x_{iu} is larger than threshold δ, w_{au}^i is set to 0, and 1 otherwise. W^i reflects binary relationship between instances within a bag. The kernel function in Eq. 3 can be plugged in a kernel classifier, e.g. support vector machine. And in our work we use RBF kernel in constructing mi-Graph kernel and support vector machine as main classifier for our task.

IV. EVALUATION

A. Data set and settings

We evaluate the proposed method on a real skin biopsy image dataset from the 3rd affiliate hospital of SUN YAT-SEN University, Guangzhou, China. The dataset contains diagnosis data from 2010 to 2011, in which there are 2734 patient records and 6579 skin biopsy images, attached with a text-based description in Chinese with unified terms of

dermatopathology. Each description indicates the concerned features of biopsy images of a certain patient. Each image has $2048 * 1536$ pixels with 24k colors in RGB space. We extract 9 features among which some are often appeared in lesion regions and others are just for some special kinds of skin diseases. Table I lists names of these features with occurrence rate of each feature in our evaluation dataset.

TABLE I. 9 SKIN BIOPSY FEATURES TO BE RECOGNIZED

Name	Rate
hyperkeratosis	28.65%
parakeratosis	22.71%
absent granular cell layer	1.8%
acanthosis	32.15%
thin prickle cell layer	4.14%
hyperpigmentation of Basal cell layer	6.48%
Munro microabscess	2.61%
nevocytic nests	9.12%
infiltration of lymphocytes	36.99%

A binary matrix is obtained by text matching, in which each row is a 9-ary binary vector to show whether an image has these features. Based on domain knowledge, a skin biopsy image is usually composed of 4 to 8 regions. We set the number of regions k as 4, 6, 8 to run our proposed algorithm and then combine them through majority voting. Images fed to Normalized Cut are all rescaled to $200 * 150$ for effective calculation. Feature extraction algorithm is also applied to rescaled images instead of original images since rescaled images contain enough information for our task.

B. *Evaluation results*

We train each model for each concerned feature respectively, i.e. bi-classification, and use 10-fold cross validation. The evaluation dataset is constructed for each feature. Positive bags are selected according to the feature matrix and negative bags are randomly selected from the rest which may contain other different features. For each feature, we set the ratio between the sizes of positive and negative bags to $1:2$, instead of the ratios listed in Table I to simply avoid learning bias. We record mean accuracy of Citation KNN and mi-Graph with different numbers of regions k in Table II.

TABLE II. CLASSIFICATION ACCURACY FOR CITATION KNN AND MI-GRAPH

Features	Citation KNN			mi-Graph		
	$k=4$	$k=6$	$k=8$	$k=4$	$k=6$	$k=8$
F-1	82.3%	**85.3%**	81.8%	73.8%	79.9%	75.8%
F-2	84.1%	79.2%	**85.2%**	79.1%	76.34%	74.9%
F-3	78.5%	**84.4%**	79.11%	74.9%	76.2%	79.3%
F-4	84.3%	85.1%	**86.2%**	80.1%	75.2%	80.2%
F-5	78.7%	75.7%	**83.0%**	76.8%	75.8%	81.1%
F-6	72.3%	73.8%	**74.3%**	71.5%	73.6%	71.5%
F-7	**71.6%**	68.8%	69.5%	67.0%	68.2%	70.2%
F-8	76.9%	74.4%	81.2%	**84.1%**	82.0%	79.1%
F-9	**86.2%**	84.3%	81.6%	82.6%	85.2%	85.1%

We highlight the highest classification accuracy for each feature. From Table II, it can be seen that Citation KNN performs better than mi-Graph in most cases. The reason may be that for biopsy features recognition problem, relationship between regions is not an essential factor to be concerned, though doctors' experience doesn't tell us about it. It seems that number of regions k doesn't affect much to the final result. Fig. 6 shows ensembles of classifiers with different k.

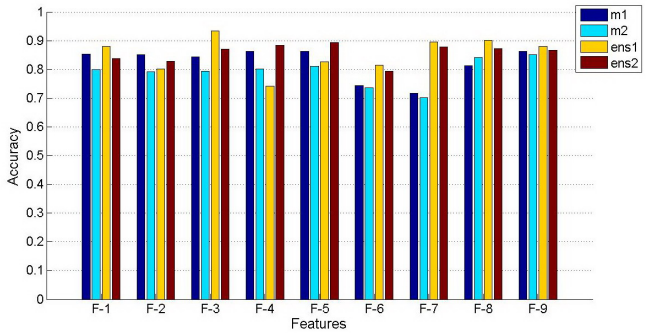


Figure 6. Ensemble Accuracy

In Fig. 6, $m1$ and $m2$ stand for the highest classification accuracy rates of different features for Citation KNN and mi-Graph respectively, and $ens1$ and $ens2$ stand for accuracy rates of ensemble classifiers with three different numbers of regions setting for these two algorithms respectively. Majority voting strategy is applied here to combine three slightly different classifiers in each case. It can be seen from Fig. 6 that in most cases ensemble classifiers outperform the best individual classifiers. When there is not sufficient prior knowledge on the number of regions, ensemble of classifiers with different regions may yield acceptable results.

C. *Discussion*

We address some issues on setting reasonable number of regions in our task. Setting number of regions as 4, 6 and 8 is directly motivated by medical experience. When inspecting skin biopsy images, small numbers of regions indicate the doctor's focus on relatively global features, while large numbers indicate more detailed features. Based on different educational background and knowledge structure, doctors' behavior may range from global to detail. As we all know, skin tissue is composed of three anatomically distinct layers which are epidermis, dermis, and subcutaneous tissue (fat). Epidermis also can be divided into four layers. Each layer has its distinctive stained color and special structures. Distinct pathological changes involving any of these whole layers such as Hyperkeratosis, Acanthosis and Hyperpigmentation of basal cell layer can be easily recognized in a small number of segmentation. Specific changes within a layer such as Munro microabscess, Nevocytic nests and Infiltration of lymphocytes can be more accurate to detect when we divide the image into more pieces. As shown in our evaluation result, the degree of global and detail view of a skin biopsy image doesn't affect much on final result, which obeys our above analysis empirically.

When it comes to consideration of relationship between regions, we mentioned that biopsy images of skin tissue are of clear inner structures and each part of the structure has its own features, so as the spatial relationships. For example, features like Hyperkeratosis and Parakeratosis can only be

found in certain region and above the feature as Acanthosis or Hyperpigmentation of basal cell layer (if such feature existing in the same image). Theoretically speaking, mi-Graph captures such relationship to some extents by regarding each image as a graph. But empirically, such relationship doesn't make positive contribution to our goal. Further work is to be done to shed light on this problem.

V. CONCLUSION

Multiple instance learning has been applied in image retrieval problems with many successful applications. In this work, we introduce it to feature recognition of skin biopsy image of dermatology for the first time. To express a given image as a multi-instance sample, we apply Normalized Cut to divide biopsy image into continuous regions and extract texture feature for each region by a 2D wavelet transformation based algorithm. The evaluation results show that the proposed method is effective for biopsy image features recognition.

In clinical, the results contribute to the development of dermatopathology. Time-consumption and expenditure will be less when computer takes over the work from pathologist. The accuracy of diagnosis would be increased when subjective factor as doctors' skill and objective factor as light are eliminated. The application accords with the developing trends of dermatopathology. With the establishment of more models and more precisely defined features, we would get more powerful tools to recognize skin biopsy images for histopathological diagnosing.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of Guangdong Province (S2011020001182), Research Foundation of Science and Technology Plan Project in Guangdong Province and Guangzhou City (2009B030801090, 2010A040303004, 11A12050914, 11A31090341, 2011Y5-00004)， Research Foundation of Sysung-Etri project (2011A091000026), and the 2012 College Student Career and Innovation Training Plan Project (No. 16).

REFERENCES

[1] S. Kalia and M. L. Haiducu, "The burden of skin disease in the united states and canada," Dermatol Clin, vol. 30, no. 1, pp. 5–18, vii, 2012.

[2] F. Dalgard, A. Svensson, J. O. Holm, and J. Sundby, "Selfreported skin morbidity among adults: associations with quality of life and general health in a norwegian survey," J Investig Dermatol Symp Proc, vol. 9, no. 2, pp. 120–5, 2004.

[3] J. Fleischer, A. B., S. R. Feldman, and S. R. Rapp, "Introduction. the magnitude of skin disease in the united states," Dermatol Clin, vol. 18, no. 2, pp. xv–xxi, 2000.

[4] M. K. Basra and M. Shahrukh, "Burden of skin diseases," Expert Rev Pharmacoecon Outcomes Res, vol. 9, no. 3, pp. 271–83, 2009.

[5] L. Brochez, K. Myny, L. Bleyen, G. De Backer, and J. M. Naeyaert, "The melanoma burden in belgium; premature morbidity and mortality make melanoma a considerable health problem," Melanoma Res, vol. 9, no. 6, pp. 614–8, 1999.

[6] C. VanderPlate and S. O. Aral, "Psychosocial aspects of genital herpes virus infection," Health Psychol, vol. 6, no. 1, pp. 57–72, 1987.

[7] L. Cerroni, Z. Argenyi, R. Cerio, F. Facchetti, H. Kittler, H. Kutzner, L. Requena, O. P. Sangueza, B. Smoller, J. Wechsler, and H. Kerl, "Influence of evaluation of clinical pictures on the histopathologic diagnosis of inflammatory skin disorders," J Am Acad Dermatol, vol. 63, no. 4, pp. 647–52, 2010.

[8] A. Fogelberg, M. Ioffreda, and K. F. Helm, "The utility of digital clinical photographs in dermatopathology," J Cutan Med Surg, vol. 8, no. 2, pp. 116–21, 2004.

[9] M. Llamas-Velasco and B. E. Paredes, "Basic concepts in skin biopsy. part i," Actas Dermosifiliogr, vol. 103, no. 1, pp. 12–20, 2012.

[10] G. Ferrara, Z. Argenyi, G. Argenziano, R. Cerio, L. Cerroni, A. Di Blasi, E. A. Feudale, C. M. Giorgio, C. Massone, O. Nappi, C. Tomasini, C. Urso, I. Zalaudek, H. Kittler, and H. P. Soyer, "The influence of clinical information in the histopathologic diagnosis of melanocytic skin neoplasms," PLoS One, vol. 4, no. 4, p. e5375, 2009.

[11] C. D. Neitzel, "Biopsy techniques for skin disease and skin cancer," Oral Maxillofac Surg Clin North Am, vol. 17, no. 2, pp. 143–6, v, 2005.

[12] W. Grayson, "Recognition of dual or multiple pathology in skin biopsies from patients with hiv/aids," Patholog Res Int, vol. 2011, p. 398546, 2011.

[13] K. Sellheyer andW. F. Bergfeld, "A retrospective biopsy study of the clinical diagnostic accuracy of common skin diseases by different specialties compared with dermatology," J Am Acad Dermatol, vol. 52, no. 5, pp. 823–30, 2005.

[14] K. Bunte, M. Biehl, M. F. Jonkman, and N. Petkov, "Learning effective color features for content based image retrieval in dermatology," Pattern Recogn., vol. 44, no. 9, pp. 1892–1902,2011.

[15] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Prez, "Solving the multiple instance problem with axis-parallel rectangles," Artif. Intell., vol. 89, no. 1-2, pp. 31–71, 1997.

[16] T. Gartner, P. A. Flach, A. Kowalczyk, and A. J. Smola, "Multi-instance kernels," pp. 179–186, 2002.

[17] Y. Chen and J. Z. Wang, "Image categorization by learning and reasoning with regions," J. Mach. Learn. Res., vol. 5, pp. 913–939, 2004.

[18] Z. Fu, A. Robles-Kelly, and J. Zhou, "Milis: Multiple instance learning with instance selection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 5, pp. 958–977, 2011.

[19] H. Blockeel, D. Page, and A. Srinivasan, "Multi-instance tree learning," pp. 57–64, 2005.

[20] J. Foulds and E. Frank, "A review of multi-instance learning assumptions," Knowl. Eng. Rev., vol. 25, no. 1, pp. 1–25, 2010.

[21] C. Wang, L. Zhang, and H.-J. Zhang, "Graph-based multiple-instance learning for object-based image retrieval," pp. 156–163, 2008.

[22] J. Shi and J. Malik, "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, Aug. 2000.

[23] G. H. Golub and C. F. Van Loan, Matrix computations, 2nd ed., ser. Johns Hopkins series in the mathematical sciences. Baltimore: Johns Hopkins University Press, 1989, 88045404 (Gene Howard), Gene H. Golub, Charles F. Van Loan. ill. ; 28 cm. Includes index. Bibliography: p. 601-634. Johns Hopkins series in the mathematical sciences ; 3.

[24] A. Gersho, "Asymptotically optimal block quantization," Information Theory, IEEE Transactions on, vol. 25, no. 4, pp. 373–380, 1979.

[25] M. Unser, "Texture classification and segmentation using wavelet frames," IEEE Transactions on Image Processing, vol. 4, no. 2, pp. 1549–1600, 1995

[26] J. Wang and J.-D. Zucker, "Solving multiple-instance problem: A lazy learning approach," pp. 1119–1125, 2000.

[27] Z.-H. Zhou, Y.-Y. Sun, and Y.-F. Li, "Multi-instance learning by treating instances as non-i.i.d. samples," pp. 1249–1256, 2009.