ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 3, April 2016

TOWARDS LARGE SCALE HISTOPATHOLOGICAL IMAGE ANALISIS: HASHING BASED IMAGE RETRIEVAL

K.P.Banupriya,S.Dharani
Electronics And Communication Engineering
SACS MAVMM Engineering College
Madurai,India
Email ID:banupriyakp11@gmail.com
smdharani1995@gmail.com

Dr.A.Swarnambiga M.E.,Ph.d.Associate professor and Head of the Department
Electronics And Communication Engineering
SACS MAVMM Engineering College
Madurai,India
Email ID:aswarnambiga@gmail.com.

Abstract— Histopathological image is used in leveraging computational image processing Methods and modern machine learning techniques. For diagnosis, disease detection, Computer Aided diagnosis (CAD) and content based image retrieval(CBIR)systems are used. To cope with histopathological images we developed scalable image retrieval technique. To use kernel hashing technique which leverages a small amount of information in learning to compress a 10000 dimensional image feature vector into only tens of binary bits with the informative signatures preserved. From large database the real time image retrieval is done from a hash table that indexfrom binary codes. Scalable image retrieval framework based on hashing technique accepts its performance on histopathological images acquired from breast microscopic tissues. Further classification is based on retrieval tests. Our framework is time efficient with 88.1%Classification accuracy. For example 0.01 seconds the framework can execute about 800 queries.

Index Terms—Histopathological image analysis, breast lesion, large-scale image retrieval, high dimension, hashing, supervised learning.

I. INTRODUCTION

Histopathology refers to microscopic examination of tissue in order to study the manifestation of diseases. Over the past decade, computational power& improvement in image analysis algorithm have allowed the development of computer assisted analytical approaches to radiological data. Currently histopathological tissue analysis by a pathologist represents the only definitive method for confirmation of presence or absence of disease and disease grading, or the measurement of disease progression.

Image retrieval is the application of computer vision techniques for searching digital images in large data bases. Content based means that search of images based on the content of images. The term content refers to shapes,colors,textures,derived from the image itself.

Hashing is the transformation of string of characters into usually shorter fixed length value or key, that represents the originalstring. Hashing is used to retrieve items in a data base. Hashing is the technique used for performing almost constant time search in case of insection detection and find operation. Hash table is a data structure which uses hash function to generate key corresponding to associate value.

For histopathological image retrieval we have developed a scalable image retrieval framework. Hashing technique is employeed to achieve efficient image retrieval. Our framework include image guided diagnosis,decision,support,education and efficient data management.

II. RELATED WORK

Classifier based CAD for Histopathological images:

CAD systems contains image preprossing,feature extraction, detection and/or segmentation, post processing and machine learning bnased classification methods. This paper exployed adaptive thresholding and segment cells using morphological operation and represent different type of nuclei having high density areas. They were then classified as linear discriminant analysis(LDA). LDA is a

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 3, April 2016

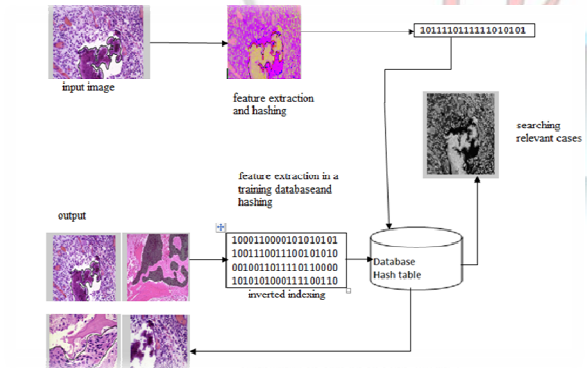
generalization of fishers linear discriminant , a method used I statistics pattern recognition and machine learning to find a linear combination. SIFT to detect and extract local descriptors for obtaining bag of words and they are classified using or support vector machine(SVM) with Kernel function. In machine learning kernel methods are a class of algorithms for pattern analysis, whose best known member is the support vector machine(SVM). The general task of pattern analyusis is to find and study general types of relations example: clusters, rankings, principal components, correlations ,classifications.

CBIR systems for medical image analysis:

Content based image retrieval also known as query by image content (QBIC) and content based visual information retrieval is the application of computer vision techniques to the imahe retrieval problem, that is, the problem of searching for digital images in large databases. In medical image analysis CBIR is important in analyzing and diagnosing relevant cases. CBIR system supports decision making in clinical pathology where fast color segmenter and central module are used inorder to extract features like shape, texture and area of nucleus. Scalability is the important factor in CBIR for the purpose of analyzing medical images.

Hashing methods for image retrieval:

In machine learning and computer vision community for image retrieval hashing methods are widely used. Few methods are weakly supervised hashing in kernel space, semi supervised hashing, supervised hashing and compact kernel hashing. Among this kernelised and supervised hasing(KSH) is generally used because of efficiency. One of the limitations of KSH is very time consuming. To overcome this further improve the original KSH with an optimization scheme is used .



III. METHODOLOGY

Scalable image retrieval:-

Scalable image retrieval includes offline learning and real time search. In offline learning we first extract visual features from histopathological images. This extraction is bnased on SIFT and are quantized with a bag of words. SIFT used the difference of Gaussian (DoG) detection. It is used in both histopathological image analysis and computer vision task. This features are used to measure the similarity in images in large database. In runtime high dimensional features are extracted from query images are proceeded to binary codes.

Kernelized and supervised hashing:

In this topic we discuss about kernelized and supervised hashing method for histopathological image retrieval.

Hashing method:

The aim of hashing method is to find group of proer hash function which generates a single hash nit to preserve the similarity of original features. Hashing methods are different from other dimensionality requirement techniques. The most fundamental requirement of hashing is map similar feature with high probability. But hashing methods need to ensure that they should have balanced and uncorrelated bit distributions. Which leads to minimum redundancy among all bits.

	LBP&LTP	GENETIC	OURS
P@10	0.3010	0.3511	0.8200
P@20	0.3122	0.3520	0.8510
P@30	0.3250	0.3550	0.8571

Kernelized hashing:

Practical data can be handled by kernel methods, which are most linearly in seperable. Linear inseperablility is an important factor in histopathological images. A kernel function is denoted as k R^d X R^d → R. The prediction function f :R^d →R with kernel k plugged in is defined as ,

$$f(x)= \sum_{j=1}^{m} (k(x(j),x) - (1/n) \sum_{i=1}^{n} k(x(j),xi$$

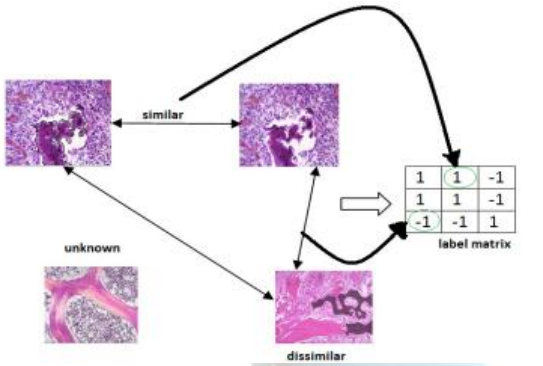
where x(1),……..x(m) are m(m<<n) feature vectors randomly selected from x, a_j € R is the coefficient and b € R is the bias. In case of natural images this scheme work well due to large

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 3, April 2016

differences their appearance. In histopathological images identifying difference between bening and actionable requires cytoplasmic texture or nuclear appearance

supervised hashing:

Hashing method minimizes the hamming distance of neighboring image pairs. Neighboring means whether the two images belong to same category or not, supervised information is encoded as similar and dissimilar pairs.



SUPERVISED INFORMATION IS ENCODED IN THE LABEL MATRIX S.

Assumptions :

Label 1: both are benign or actionable.
Label -1: one is benign and other is actionable.

IV.EXPERIMENTS

Setup:

The images were collected from web databases. A database of histopathological images were created with 1000 images.

Evaluation of image classification:

Genetic algorithm was considered as Heuristic search algorithm which undergoes 5 terminologies to obtain the optimal feature subset. Genetic algorithm is considered to solve the optimization problem which comes under the evolutionary criterion. This algorithm is mainly preferred for the search technique. This algorithm is efficiently applicable for searching technique in large state space, n dimensional state space or multi model state space and also applicable even the inputs are slightly changed. Genetic algorithm offers significant searching process among many optimization technique as depth first breath first and heuristic. Genetic algorithm pretend to survive the fittest among the individual over consecutive algorithm. Comparing to computational speed of the hashing based optimization technique the genetic algorithm will improve the complexity.

Local binary pattern is widely used in 2d texture analysis. The LBP operator is a non parametric 3x3 kernel which describes the local spatial structure of an image.LBP is defined as an ordered set of binary comparisons of pixel intensities between the centre pixel and its eight surrounding pixels. The decimal values of the resulting 8bit word leads to 8 possible combination, which ate called local binary pattern. LBP is define as an ordered set of binary comparisons of pixel and its eight surrounding pixels.

Local Ternery Pattern Analysis are an extension of LBP. Unlike LBP, it does not threshold the pixels into 0 and 1,rather it uses a threshold constant to threshold pixels into three values. Considering k as the threshold constant, c as the value of the center pixel, a neighboring pixel P. Christo Ananth et al. [15] proposed a work, in this work, a framework of feature distribution scheme is proposed for object matching. In this approach, information is distributed in such a way that each individual node maintains only a small amount of information about the objects seen by the network. Nevertheless, this amount is sufficient to efficiently route queries through the network without any degradation of the matching performance. Digital image processing approaches have been investigated to reconstruct a high resolution image from aliased low resolution images. The accurate registrations between low resolution images are very important to the reconstruction of a high resolution image. The proposed feature distribution scheme results in far lower network traffic load. To achieve the maximum performance as with the full distribution of feature vectors, a set of requirements regarding abstraction, storage space, similarity metric and convergence has been proposed to implement this work in C++ and QT. Online updation scheme provide better scalability and efficiency while comparing to offline scheme.

	LBP&LTP	GENETIC	OURS
P@10	0.3010	0.3511	0.8200
P@20	0.3122	0.3520	0.8510
P@30	0.3250	0.3550	0.8571

Evaluation of image retrieval:

The precision of our method is 6% better than others except NCA. Our method is around 3.5%

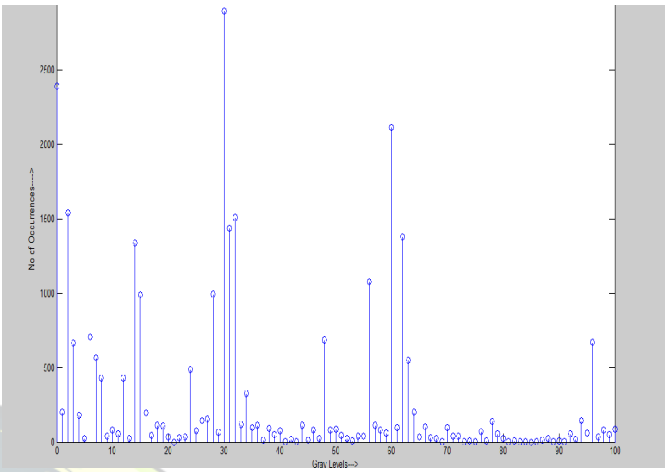
ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 3, April 2016

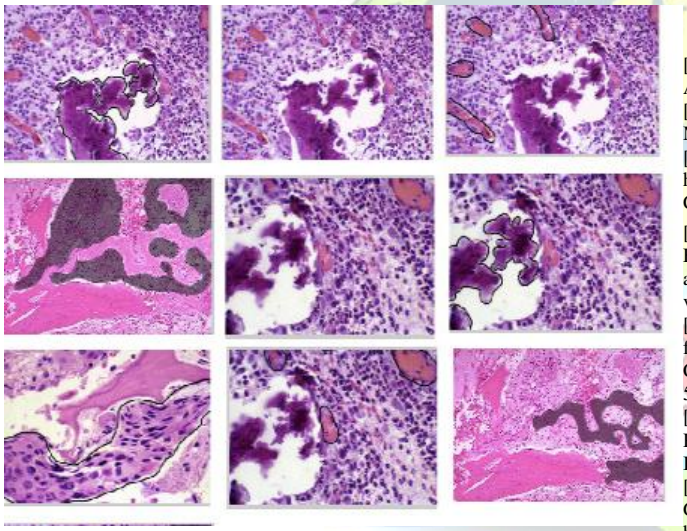
better than NCA on benign case and 14% better than NCA on actionable case. Our method tables only constant time when using hash table which is independent of number of samples and number of feature dimensions.

In our method in 46%better than local binary pattern and local ternary pattern and 40%better than genetic algorithm.

V.RESULT OF IMAGE RETRIEVAL

In the output image they are listed based on relevancy. Our accurate result shows the efficiency of our proposed method. Supervised information also helps us in improving retrieval precision. The retrieved images are relevant to input image and hence useful for decision support.





V.CONCLUSION

In this paper for histopathological image analysis we have developed scalable image retrieval framework. For efficient image retrieval we used hashing and presented an improved kernelized and supervised hashing apprioach which are used in real time image retrieval. Applications of our paper include image guided diagnosis, ecison support, education and efficient data management. In future we may examine features from segnmentation and architectures and develop further more feature fusion techniques to design hybrid hashing.

References

[1] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2013," CA: A Cancer Journal for Clinicians, vol. 63, no. 1, pp. 11–30, 2013.
[2] D. C. Sgroi, "Preinvasive breast cancer," Annual Review of Pathological Mechanical Disease, vol. 5, pp. 193–221, 2010.
[3] D. L. Page, W. D. Dupont, L. W. Rogers, and M. S. Rados, "Atypical hyperplastic lesions of the female breast. a long-term follow-up study," Cancer, vol. 55, no. 11, pp. 2698–2708, 1985.
[4] B. Weyn, G. van de Wouwer, A. van Daele, P. Scheunders, D. vanDyck, E. van Marck, and W. Jacob, "Automated breast tumor diagnosis and grading based on wavelet chromatin texture description," Cytometry, vol. 33, no. 1, pp. 32–40, 1998.
[5] C. Bilgin, C. Demir, C. Nagi, and B. Yener, "Cell-graph mining for breast tissue modeling and classification," in IEEE International Conference on Engineering in Medicine and Biology Society, 2007, pp. 5311–5314.
[6] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," IEEE Reviews in Biomedical Engineering, vol. 2, pp. 147–171, 2009.
[7] J. Kong, O. Sertel, H. Shimada, K. L. Boyer, J. H. Saltz, and M. N. Gurcan, "Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation," Pattern Recognition, vol. 42, no. 6, pp. 1080–1092, 2009.
[8] P.-W. Huang and C.-H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," IEEE Transactions on Medical Imaging, vol. 28, no. 7, pp. 1037–1050, 2009.
[9] P.-W. Huang and Y.-H. Lai, "Effective segmentation and classification for HCC biopsy images," Pattern Recognition, vol. 43, no. 4, pp. Bahlmann, P. Maday, A. Kamen, S. Gibbs-Strauss, E. M.
Genega, and J. V. Frangioni, "Computer-aided gleason grading of prostate cancer histopathological images using texton forests," in IEEE International Symposium on Biomedical Imaging, 2010, pp. 636–639.
[11] M. Veta, J. Pluim, P. van Diest, and M. Viergever, "Breast cancer histopathology image analysis: A review," IEEE Transactions on Biomedical Engineering, vol. 61, no. 5, pp. 1400–1411, May 2014.
[12] D. Comaniciu, P. Meer, and D. J. Foran, "Image-guided decision support system for pathology," Machine Vision and Applications, vol. 11, no. 4, pp. 213–224, 1999.
[13] L. Zheng, A. W. Wetzel, J. Gilbertson, and M. J. Becich, "Design and analysis of a content-based pathology image retrieval system," IEEE Transactions on Information Technology in BioMedicine, vol. 7, no. 4, pp. 249–255, 2003.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 3, April 2016

[14] H. M¨uller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications-clinical benefits and future directions," International Journal of Medical Informatics, vol. 73, no. 1, pp. 1–23, 2004.
[15] Christo Ananth, R.Nikitha, C.K.Sankavi, H.Mehnaz, N.Rajalakshmi, "High Resolution Image Reconstruction with Smart Camera Network", International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST), Volume 1,Issue 4,July 2015, pp:1-5
[16] A. Kumar, J. Kim, W. Cai, M. Fulham, and D. Feng, "Content-based medical image retrieval: A survey of applications to multidimensional and multimodality data," Journal of Digital Imaging, vol. 26, no. 6, pp. 1025–1039, 2013.
[17] T. M. Lehmann, M. O. G¨uld, T. Deselaers, D. Keysers, H. Schubert, K. Spitzer, H. Ney, and B. B.Wein, "Automatic categorization of medical images for content-based retrieval and data mining," Computerized Medical Imaging and Graphics, vol. 29, no. 2, pp. 143–155, 2005.
[18] L. Yang, W. Chen, P. Meer, G. Salaru, L. A. Goodell, V. Berstis, and D. J. Foran, "Virtual microscopy and grid-enabled decision support for large-scale analysis of imaged pathology specimens," IEEE Transactions