31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Automated Classification of Renal Cell Carcinoma Subtypes Using Scale Invariant Feature Transform

S. Hussain Raza, Yachna Sharma, Qaiser Chaudry, Andrew N. Young, May D. Wang

Abstract—The task of analyzing tissue biopsies performed by a pathologist is challenging and time consuming. It suffers from intra- and inter-user variability. Computer assisted diagnosis (CAD) helps to reduce such variations and speed up the diagnostic process. In this paper, we propose an automatic computer assisted diagnostic system for renal cell carcinoma subtype classification using scale invariant features. We capture the morphological distinctness of various subtypes and we have used them to classify a heterogeneous data set of renal cell carcinoma biopsy images. Our technique does not require color segmentation and minimizes human intervention. We circumvent user subjectivity using automated analysis and cater for intra-class heterogeneities using multiple class templates. We achieve a classification accuracy of 83% using a Bayesian classifier.

Keywords – Renal Cell Carcinoma, Computer Assisted Diagnosis, Image Classification, Scale Invariant Features.

I. Introduction

Renal cell carcinoma (RCC) accounts for approximately 3% of adult malignancies and 90-95% of neoplasms arising from the kidney [1]. It is characterized by lack of early warning signs, diverse clinical manifestations and resistance to radiation and chemotherapy. The World Health Organization (WHO) classification system has defined several subtypes of RCC [2]; the most common subtypes include Clear Cell (CC) RCC (83%), Chromophobe (CH) RCC (2%), Papillary (PA) RCC (11%) and Oncocytoma (ON) RCC (4%). Fig. 1 shows sample images of each subtype.

Clinicians treat these RCC subtypes differently; therefore, it is extremely important to identify them accurately for treatment planning. Manual categorization of subtypes is a challenging and time-consuming process, involving inter- and intra-observer variability. In addition, for greater impact and outreach of translational medicine, CAD systems should be designed to provide consistent results for image datasets acquired from different clinical labs using varying image acquisition protocols. Moreover, user involvement should be minimized to reduce bias and subjectivity. The majority of efforts on cancer subtype classification involve manual intervention for color segmentation to extract tissue morphological features from differentially stained images [3, 4]. A color segmentation based CAD algorithm may not be efficient for dynamic handling of all the variations in a new dataset. The scale invariant feature transform (SIFT) is a computer vision technique [5] that has been widely used in object recognition [6], feature tracking [7] and image registration [8]. The SIFT algorithm operates on features computed around key points in an image. These key points are consistent and recognizable under different magnifications and illumination conditions. Descriptors are calculated for each key point to encapsulate intensity variations in the neighborhood. RCC subtypes are mainly distinguished by their morphology and stain color distribution [9]. Table I summarizes distinct characteristics of the four subtypes. The location of key points depends on the morphology and the descriptors capture intensity variations around the neighborhood. Fig. 2 shows a sample papillary image with key points. Note that the key points are mainly located in the nuclear rich regions and not the homogenous lumen regions.

In this work, we develop an automatic CAD system for RCC subtype classification using SIFT features. Our methodology circumvents user involvement in an effort to

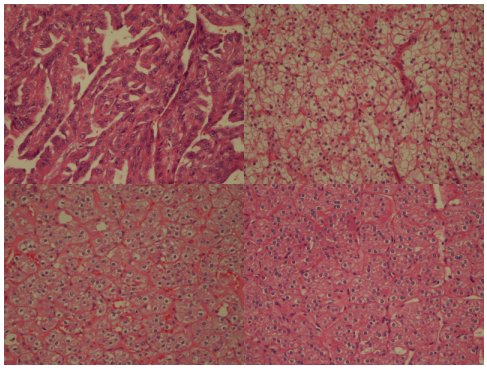

Fig. 1. (Left to right, top to bottom) Sample images of each subtype, a) Papillary (PA), b) Clear Cell (CC), c) Chromophobe (CH), d) Oncocytoma (ON).

This research has been supported by grants from National Institutes of Health (Bioengineering Research Partnership R01CA108468, P20GM072069, Center for Cancer Nanotechnology Excellence U54CA119338), Georgia Cancer Coalition (Distinguished Cancer Scholar Award to Professor Wang), Hewlett Packard, and Microsoft Research.

S. H. Raza is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: sraza3@gatech.edu).

Y. Sharma is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ysharma3@gatech.edu).

Q. Chaudry is with the Georgia Institute of Technology, Atlanta GA 30332 USA (e-mail: qaiser@gatech.edu).

A. N. Young is with Grady Health System and Emory University, Atlanta GA 30322 USA (e-mail: ayoun01@emory.edu).

M. D. Wang is with the Wallace H. Coulter Department of Biomedical Engineering at the Georgia Institute of Technology and Emory University, Atlanta, GA 30332 USA (Corresponding Author: phone: 404-274-4625; e-mail: maywang@bme.gatech.edu).

provide consistent results for highly heterogeneous datasets with a wide range of intra-class variations.

II. BACKGROUND

There have been several works pertaining to the development of CAD systems for cancer classification [3, 4, 10]. Extraction of morphological features has played an important role in biological image classification. For example, mathematical morphology has been used to classify the images as cancerous or non-cancerous [11].

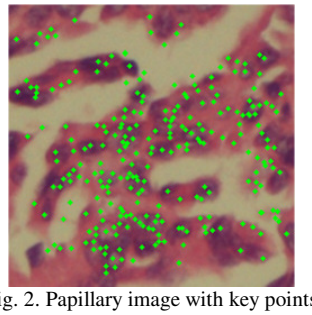

Fig. 2. Papillary image with key points

TABLE 1
Hallmarks of RCC Subtypes

Subtype	Morphology	Stain color
CC	Solid growth pattern, condensed and hyper chromatic nuclei.	Clear cytoplasm
CH	Large polygonal cells arranged in solid nests or tubules.	Clear area/halo around nucleus,
ON	Granular cytoplasm, cells arranged in nests, or trabecular patterns, round and centrally located nuclei.	Abundant pink cytoplasm
PA	Tubulo-papillary architecture	granular cytoplasm

Others have based their classification on textural features derived from ether multispectral analysis [12], a GLCM (gray level co-occurrence matrix) [4, 13], wavelet coefficients [10] or a combination of these [14]. Combined use of morphological and textural features has also been reported. For example, an improvement in the classification accuracy of colon cancer is reported by using a fractal dimension along with conventional texture analysis [15]. GLCM and wavelet based works [4, 16] have reported reasonably good classification accuracies, however, their color segmentation is done using manually seeded k-means. SIFT has been used in some works pertaining to medical image processing such as for deformable registration [17], image annotation [18] and image classification [19] with the goal of assigning keywords to medical X-ray images. X-ray images are intensity images and mainly depict bone structures. However, RCC images are heterogenously stained color images and we use scale invariant feature transform (SIFT) to automatically encapsulate the morphology (the location/distribution of key points) and texture (key point descriptor) into one descriptor. Our system minimizes manual intervention while still providing consistent subtype classification.

III. METHODOLOGY

Fig. 3 shows a flow chart of the complete methodology. First, we convert the RGB images to gray scale, then we detect key points. Thereafter, we compute a feature descriptor around each key point. Then, we match key point descriptors from input images to template images. Next, we evaluate the number of matches between each template and the test image. We use the number of matches for each sub-type as features for classification. A detailed description of each step is given below.

A. *Image Acquisition*

The image data set consists of hematoxylin and eosin (H&E) stained tissue biopsy images. We resected tissue samples for this study by total nephrectomy following standard pathological procedures to fix, section, and stain the tissue. Then, we embedded histological samples in paraffin to slice the microscopic sections and stained them with Hematoxylin and eosin. Board-certified anatomic pathologists using WHO histo-pathological criteria diagnosed all the tumors. We took photomicrographs at a total magnification of 200x and 1200x1600 pixels. We captured 48 images, 12 for each subtype: CC, CH, ON, and PA.

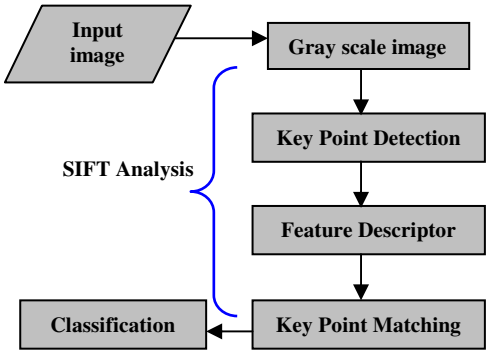

Fig. 3. Flowchart for Overall Methodology

B. *Feature Extraction*

Key Point Detection using the Harris Method: We used the Harris corner detection method [20] for key point localization. We want to select key points (i.e., corners), which will be stable under different scales. That is, if we have two images of the same object each with a different scale and view, the same key points should be detected. The basic idea is that we should easily recognize these corners by looking at the image gradient near a candidate key point. In smooth regions, the gradient is near zero. Near an edge, the gradient points strongly in one direction. Near a corner, the gradient points strongly in multiple directions.

Key point Feature Descriptor: We evaluate the feature descriptor around a 16×16 neighborhood of detected key points as described in [5]. Briefly, the gradient magnitudes in the 16×16 neighborhood are summarized in 4×4 sub regions by calculating the histogram of normalized gradient magnitudes in eight directional bins. Fig. 4 shows the

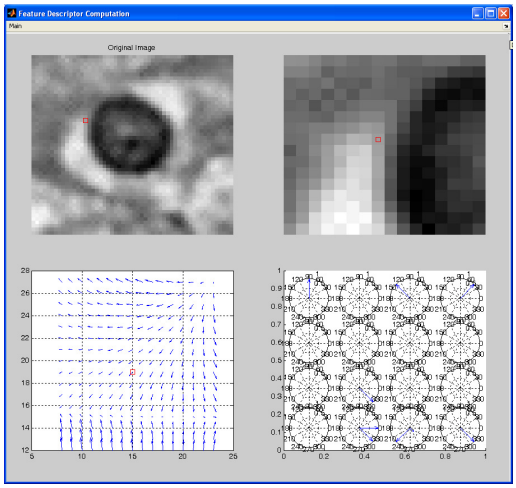

Fig. 4. Screenshot of the GUI showing a sample key point (top left), a 16x16 neighborhood around key point (top right), plot of gradient magnitudes (bottom left) and normalized image gradient descriptor summarized in 4x4 region (bottom right).

gradient values and associated descriptor for a sample key point. The final descriptor length is 128 (8 bins for 16 regions), see Fig 4(bottom right).

C. *Key point Matching*

We use minimization of sum of squared differences (SSD) to match the key points in a sample image with the key points in the template images. We perform initial analysis using a standard template image for one subtype and calculate the number of matching key points with standard templates of other subtypes. We do this to test the efficacy of SIFT matching for RCC subtype classification. Fig. 5 shows the results. It is clear from Table II that there are more matches (diagonal elements) with the same subtype class than with different classes.

Inter-subtype (non-diagonal elements) matches reflect the neighborhood similarities embedded within certain key points due to image components that are ubiquitous among most or all subtypes. For example, key points around large lumen spaces and necrotic regions may match key points from other classes that have similar neighborhood gradients. In addition, sometimes images from within classes exhibit different key point descriptors due to *intra-class* heterogeneity in the biopsy image data. These limitations add to the classification error. We use two template images for each subtype so that intra-class variations can be accommodated. We match each input image with the template images of each subtype and use the average number of matches from the two templates as a feature for classification. We normalize the number of matches by dividing by the product of the number of points in the test and template images. Table III shows the mean and standard deviation of matches among image subtypes.

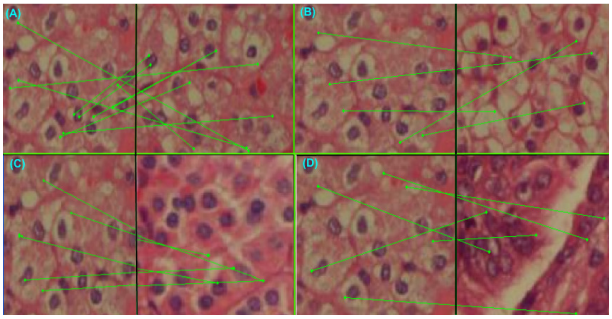

Fig. 5 Matching of Chromophobe template with each subtype (A) CH-CH (B) CH-CC (C) CH-ON (D) CH-PA SIFT matching

TABLE II
Number of SIFT Matches between Template Pairs.

	CC	CH	ON	PA
CC	18	13	5	11
CH	5	10	4	5
ON	7	16	20	10
PA	7	6	4	13

D. *Subtype Classification*

We used a Bayesian classifier to classify a set of 40 test images (10 for each subtype). We selected the templates from images that are not part of the test set. To estimate the error of classification, we used leave one out cross validation (LOO CV).

TABLE III
Mean and Standard Deviation of Number of SIFT Matches between Templates and Test Images.

	CC	CH	ON	PA
CC	.0168±.004	.0527±.004	.03±.004	.026±.003
CH	.0142±.001	.0571±.007	.0341±.003	.029±.003
ON	.0132±.002	.0547±.005	.046±.011	.031±.004
PA	.0129±.001	.0478±.003	.039±.003	.0365±.003

TABLE IV
Confusion Matrix for RCC Subtype Classification using SIFT.

	CC	CH	ON	PA
CC	8	2	0	0
CH	1	8	1	0
ON	0	1	9	0
PA	1	1	0	8

IV. RESULTS AND CONCLUSION

Table IV shows the confusion matrix for classification. Fig. 6. shows the scatter plot of each test image using three out of four features. We achieved an accuracy of 83% among the four subtypes. SIFT descriptors encapsulated the intensity variations around a key point. Lumen spaces and necrotic regions were common in all of the subtypes. Therefore, sift descriptors near lumen features will result in considerable numbers of matches with templates from other classes. In addition, intra-class heterogeneity contributes to lesser number of matches with the respective templates and more numbers of matches with templates from other classes. For example, in Fig. 7, the image on the left is a Clear Cell RCC but our system classified it incorrectly because it is very different from a typical Clear Cell image (Fig 7, right image). However, our method is automatic and attempts to

handle intra-class heterogeneities by using multiple templates.

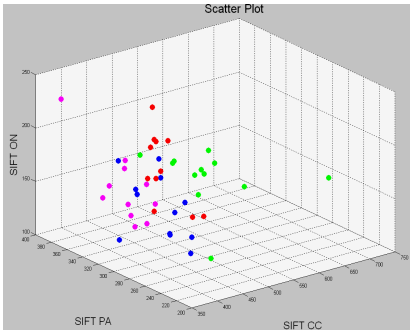

Fig. 6. Number of Matches to PA, CC, ON Templates. CC samples are red, CH are blue, ON are magenta, and PA are green.

V. Discussion and Future Work

Our classification accuracy (83%) is encouraging for a four class system since there is no manual interaction involved. Still, other works have reported higher accuracies using more user interaction. We favor automatic methods, because user interaction not only takes effort but also introduces subjectivity.

Biological images are immensely diverse and creation of an ideal template set with all possible variations is not always feasible, especially for relatively rare diseases where training images may be scarce. We have used two representative template images of each subtype in an attempt to account for possible intra class heterogeneities, but selection of more templates may be justified, depending on the expected heterogeneity of images. A technique with expert knowledge based template set selection might result in increased classification accuracy, but care should be taken to not just include ideal images, but also to include some low quality images in template sets to match all possible inputs.

In the future, we plan to extend our analysis using a variable size descriptor (proportional to cell size) which includes stain color information, (color SIFT). This would help us encapsulate tissue morphology, stain color, and texture into one descriptor. In addition, we can use the correlation of tissue features, such as cell size, with the descriptors to create a codebook to map tissue morphology back to key point descriptors. We expect that this methodology can easily extend to the classification of other medical images, including other types of cancers, or imaging modalities.

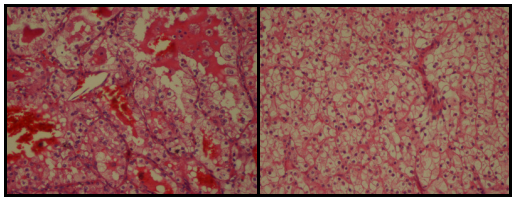

Fig. 7. CC Test Image Examples. (left) incorrectly classified, (right) correctly classified

Acknowledgment

The authors would like to thank Dr. R. Mitchell Parry and Richard Moffitt for their valuable comments during the preparation of this manuscript.

References

[1] Yin-Goen, Q., et al., Advances in molecular classification of renal neoplasms. Histology and histopathology, 2006. 21(1-3): p. 325-339.

[2] "WHO Classification of Tumours. Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs," Edited by J. N. Eble, G. Sauter, J. I. Epstein and I. A. Sesterhenn, *IARC Press*, Lyon France, February 2004, ISBN 92 8322 415 9.

[3] S. Waheed, R. A. Moffitt, Q. Chaudry, A. N. Young, M. D. Wang, "Computer Aided Histopathological Classification of Cancer," *Proc. IEEE Bioinformatics and Bioengineering*, 2007, Page(s):503 – 508.

[4] Q. Chaudry, S. H. Raza, A. N. Young, M. D. Wang, "Automated Renal Cell Carcinoma Subtype Classification Using Morphological, Textural and Wavelet-based Features," *Journal of Signal Processing: Special Issue on Biomedical Imaging*, 2008.

[5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.

[6] A. Suga, K. Fukuda, T. Takiguchi, Y. Ariki, "Object recognition and segmentation using SIFT and Graph Cuts", ICPR, 2008, pp. 1-4.

[7] S. Battiato, G. Gallo, G. Puglisi, S. Scellato, "SIFT Features Tracking for Video Stabilization", CIAP 2007, pp. 825 – 830.

[8] E. Delponte, F. Isgro, F. Odone, A.Verri, " SVD-matching Using SIFT features", Special Issue on the Vision, Video and Graphics Conference 2005, Vo. 68, Issue 5-6, 2006,pp. 415-431.

[9] Atlas of Genetics and Cytogenetics in Oncology and Haematology. URL http://AtlasGeneticsOncology.org

[10] R.F. Walker, P.T. Jackway, B. Lovell, "Classification of cervical cell nuclei using morphological segmentation and textural feature extraction," *Proc., of the 2nd Australian and New Zealand Conference on Intelligent Information Systems*, 297-301, 1994.

[11] M. A. Roula, J. Diamond, A. Bouridane, P. Miller, and A. Amira, "A multispectral computer vision system for automatic grading of prostatic neoplasia," *IEEE International Symposium on Biomedical Imaging,* Washington D. C., 2002, pp. 193- 196.

[12] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa," *IEEE Transactions on Information Technology in Biomedicine*, Volume 2, Issue 3, Sept. 1998 Page(s):197 - 203.

[13] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, "Fractal analysis in the detection of colonic cancer images," *IEEE Transactions on Information Technology in Biomedicine*, Volume 6, Issue 1, March 2002, Page(s): 54-8.

[14] M. Fern´andez and A. Mavilio, "Texture analysis of medical images using the wavelet transform," *AIP Conference Proceedings*, vol. 630, October 2002, pp. 164–168.

[15] J. Diamond, N. Anderson, P. Bartels, R. Montironi, and P. Hamilton, "The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia," *Human Pathology*, vol. 35, pp. 1121-1131, 2004.

[16] S. Arivazhagan and L. Ganesan, "Texture classification using wavelet transform," *Pattern Recogn. Lett.* , vol. 24, no. 9-10, pp. 1513–1521, 2003.

[17] Mehdi Moradi, Purang Abolmaesoumi, and Parvin Mousavi, "Deformable Registration Using Scale Space Keypoints" Proceedings of SPIE, 2006.

[18] Bo Qiu, "A Refined SVM Applied in Medical Image Annotation" pp. 690 – 693, Springer-Verlag Berlin Heidelberg 2007.

[19] Lokesh Setia,Alexandra Teynor, Alaa Halawani and Hans Burkhardt,"Image Classification using Cluster Co occurrence Matrices of Local Relational Features" ACM 2006.

[20] C. Harris and M. Stephens (1988). "A combined key point and edge detector" (PDF). Proceedings of the 4th Alvey Vision Conference: pp 147--151.