Novel Image Features for Categorizing Biomedical Images

Jianqiang Sheng*, Songhua Xu†, Weicai Deng*, Xiaonan Luo*

*National Engineering Research Center of Digital Life
*State-Province Joint Laboratory of Digital Home & Interactive Applications
*School of Information Science and Technology, Sun Yat-sen University, Guangzhou, 510006, China

†Oak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge, Tenneessee, USA, 37830

Emails: shengjianqiang@163.com, xus1@ornl.gov, dengweic@mail2.sysu.edu.cn, lnslxn@mail.sysu.edu.cn

***Abstract*—Images embedded in biomedical publications are richly informative. For example, they often concisely summarize key hypotheses, illustrate new methods, and highlight major experimental findings in a research article. Prior studies [1] suggested that images embedded in biomedical publications offer effective clues for retrieving and mining their source documents. To facilitate accessing such valuable imagery resources, image categorization can be helpful. Like many other image processing tasks, extracting discriminative image features is critical for the success of image categorization. For biomedical images, we notice that many of them are embedded with abundant annotation text. Observing this property, we introduce a set of novel image features that exploit the spatial distribution of text information inside an image as essential clues for categorizing biomedical images. Through results of our evaluation experiments, this paper demonstrates the effectiveness of the proposed novel features—compared with conventional image features, our new features can help categorize biomedical images with superior performance using a standard supervised learning based approach.**[1]

***Keywords*-image categorization, novel image features, spatial distribution of text information.**

I. Introduction

Biomedical researchers tend to include carefully composed images in their publications to visualize study subjects (such as a specific type of cell, tissue, or organ concerned in the research), present study design, explain research approaches, and report experimental findings. For example, we randomly sampled 500 recent biomedical articles published in the *Proceedings of the National Academy of Sciences* (PNAS) and found that each article on average contains four images. We also surveyed articles published in the past two years of the medical journal *The Lancet* and discovered that around 40% of these articles contain at least one image. It is commonly acknowledged that images embedded in the biomedical literature can help medical professionals quickly glance over key contents of an article for more effectively accessing knowledge conveyed in the document. Therefore, within the field of biomedical information retrieval, an important research topic is to categorize images embedded in the literature for efficient navigation, retrieval, and mining of visual and textual contents. An abundant collection of prior work has been dedicated to the problem of automatic categorization of biomedical images, e.g [2], [3]. Researchers have also investigated how to classify medical images to assist clinical diagnosis. For instance, Zhang et al. [4] designed a clinical diagnosis auxiliary tool for medical decision making that applies a number of image classification methods for 3D brain images acquired through functional magnetic resonance. Balasubramanyam et al. [5] studied a similar problem of optical tomography image processing. They trained support vector machines (SVMs) to categorize images of healthy joints from unhealthy ones.

To more accurately categorize images embedded in biomedical publications, in this paper we propose a set of novel features that exploit the spatial distribution of text information inside an image. To the best of our knowledge, none of these features has been previously explored in the literature. Using these new image features, we can categorize biomedical images more accurately, as verified by our experimental results. The image categorization procedure used in our experiments works as follows. First, we transform an input image into its binary version. We then detect all text regions inside the image using the algorithm proposed in [6], whose software implementation was publicly released by the authors on the Internet. Next, we derive the proposed image features according to the spatial distribution of text information in the image. Finally, using the extracted image features, we apply a supervised learning based method to categorize biomedical images.

The rest of this paper is organized as follows. We first

[1]J. Sheng and S. Xu are equal first authors.

†Correspondence author: S. Xu (xus1@ornl.gov), Oak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge, Tenneessee, USA, 37830.

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

overview some closely related work in Section II. Next, in Section III, we introduce the image taxonomy used in our image categorization work as well as the necessary image preprocessing step prior to image feature extraction. After that, in Section IV, we introduce the definition of our novel image features, followed by the experimental results of biomedical image categorization using these new features in Section V. Lastly, in Section VI, we conclude this paper.

II. Related Work

Recently a large number of researchers explored biomedical image processing with encouraging progress. In this section, we briefly overview two aspects of prior work that closely relate to our study here, i.e. image representation and categorization methods.

Image Representation Methods Representing an image's content in a computer understandable way is a fundamental problem in image processing. A vast collection of strategies have been proposed for extracting image features. The most common ones include image pixel values [7], colors [8], and color histograms [9]. People have also derived image features according to shapes of image edges [10] and characteristics of an image's texture [11]. High-level, semantically-oriented image features have also been exploited, such as features extracted from image regions around points of interests and various bag-of-features [12]. For example, Tommasi et al. [13] proposed fusing multiple local and global descriptors of an image into a uniform image feature representation.

Image Categorization Methods Chen et al. [14] introduced a new learning method based on Multiple-Instance Learning (MIL) for region-based image categorization. Shatkay et al. [15] introduced an approach for biomedical image categorization by using image features extracted from gray-level histogram statistics and edge direction histograms of an image. To our best knowledge, no published work has previously pursued the idea of using text distribution in an image as a type of novel image features for biomedical image categorization.

III. Image Taxonomy and Preprocessing

A. Image Taxonomy

To categorize biomedical images, we first need to introduce an image taxonomy. As pointed out by Murphy et al. [16], no uniform standards exist regarding the composition of images to use in biomedical publications. In this study, we use the image taxonomy (shown in Fig. 1), which is defined according to feedbacks collectively received from a group of biomedical researchers at our institution. Such a taxonomy consists of a two-level hierarchy. On the top level, images are first categorized into five classes, including: 1) *flow charts*, 2) *experimental images*, 3) *graph images*, 4) *mixed images*, and 5) *others*. In particular, the category of *mixed images* refers to the type of images that consist of sub-images belonging to more than a single image category. Images that do not belong to any of the first four classes are then categorized into the class of *others*. On the second level of our image taxonomy, we further categorize images into eight sub-classes, which are exemplified in Fig. 2(a) to Fig. 2(h). More concretely, the class of *experimental images* on the top level are further categorized into *microscopy images* and *gel electrophoresis images*; the class of *graph images* on the top level are further categorized into *line charts*, *bar charts*, *spot charts*, and *tables*.

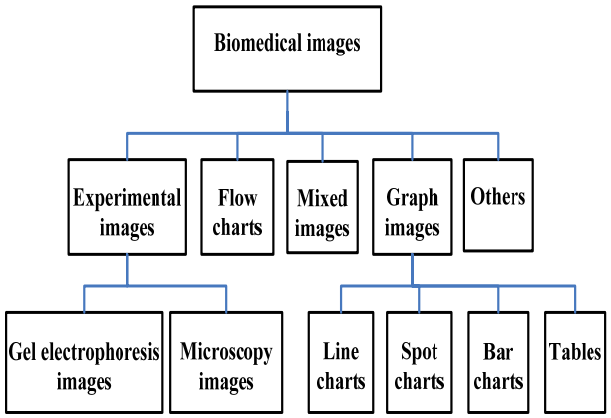
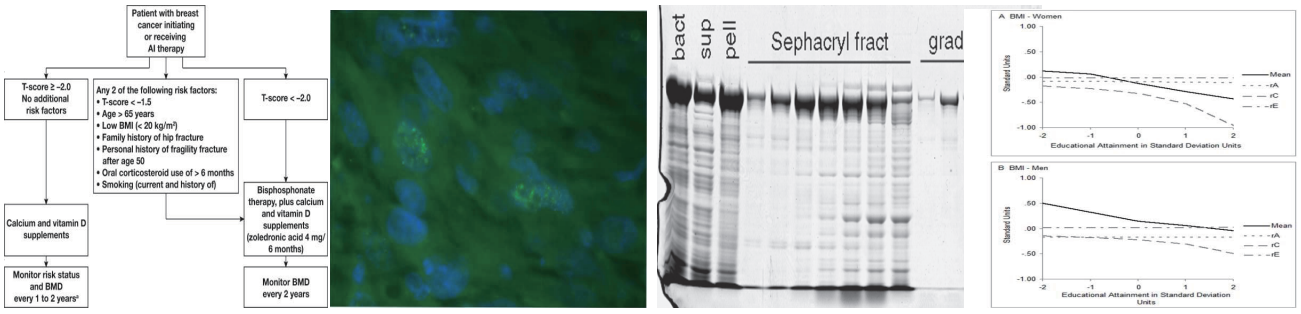


Figure 1. Image taxonomy used in our work.

B. Image Segmentation

To achieve a good image categorization performance, we first segment an input image into a set of sub-images where each sub-image corresponds to a panel in the original input image. In particular, many images belonging to the class of *experimental images* or *others* are composed of multiple sub-images. Our image segmentation algorithm operates as follows. First, we convert an input image S into its binary version. We then count the number of foreground pixels on each horizontal scanning line of the image. If a horizontal scanning line encounters less than 5 foreground pixels, we consider the scanning line a candidate cutting line for segmenting the image. The threshold of 5 pixels is determined empricially by searching for the optimal threshold value in the range of 1 to 20 pixels against the image set used in our experiments. For vertical scanning lines, we conduct the same procedure to detect candidate vertical cutting lines for segmentating the image. Using all the candidate horizontal and vertical image cutting lines detected, the input image is partitioned into multiple rectangular grid cell regions, i.e. $S = \{S_i\}$. For each resulting grid cell region S_i in the segmented image, we calculate its area, $Area(S_i)$, and consider it a valid sub-image region if $Area(S_i) \geq \frac{1}{20} Area(S)$. Figure 3 shows a sample image that consists of several sub-images and its segmentation results as detected by the procedure above.



(a) flow chart (b) microscopy image (c) gel electrophoresis (d) line chart

Clip application	Group A (easy level)			Group B (hard level)		
	Task 1	Task 5	Task 10	Task 1	Task 5	Task 10
Loss of blood (ml)	0 (0-470)	0 (0-110)	0 (0-570)	2000 (910-570)	570 (200-680)	460 (150-1940)
Dropped clips (n)	0 (0-9)	0 (0-2)	0 (0-2)	0 (0-1)	1 (1-2)	0 (0-1)
Badly placed clips (n)	0 (0-2)	0 (0-3)	0 (0-11)	0 (0-2)	1 (0-2)	1 (0-1)
Incomplete target areas (n)	0 (0)	0 (0-2)	0 (0-12)	2 (1-2)	2 (1-2)	0 (0-1)
Stretch-damage in %	59 (4-100)	24 (0-100)	52 (0-100)	100 (100-100)	100 (100-100)	100 (100-100)

(e) table (f) spot chart (g) bar chart (h) others

Figure 2. Examples of eight (sub-)classes of images in our image categorization taxonomy.

IV. Novel Image Features for Categorizing Biomedical Images

As mentioned earlier, in this work we introduce a set of novel image features that exploit the spatial distribution of annotation text inside an image for image categorization. In the following, we look at the definitions of these novel image features.

Since our novel image features are all concerned with the spatial distribution of text information inside an image, to define these features, we first need to detect the presence and locations of text information residing in a biomedical image. For this purpose, we use the pivoting text region detection algorithm [6], which is specially proposed for detecting text regions inside biomedical images. Each text region detected by the algorithm is indicated by a rectangular bounding box. Based on the distributions of these detected text regions, we can derive the following novel image features for categorizing biomedical images.

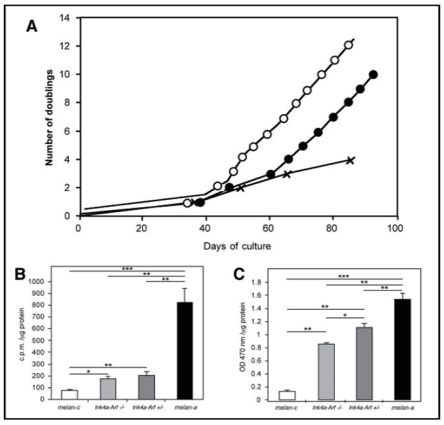
A. Entropy Distribution of Text Regions

For each vertical and horizontal scanning line l_i of an input image $\mathcal{I}$, we derive the entropy of its encountered foreground pixels, i.e. $E(l_i) = -p\ln(p)$ where p is the percentage of foreground pixels lying on the scanning line. Based on the entropy values derived from all horizontal scanning lines in $\mathcal{I}$, we construct a histogram that equally divides the value range of the maximum and minimum entropy values of all horizontal scanning lines of $\mathcal{I}$ into ten bins. Let $H_{\text{h}}^{\text{entropy}}(\mathcal{I})$ be a ten dimensional vector where its j-th component, $H_{\text{h},j}^{\text{entropy}}(\mathcal{I})$, denotes the percentage of horizontal scanning lines whose entropy values fall into the j-th value sub-range of the aforementioned histogram. In the same way, we can prepare another ten dimensional vector $H_{\text{v}}^{\text{entropy}}(\mathcal{I})$ through analyzing the entropy distribution of all vertical scanning lines of $\mathcal{I}$.

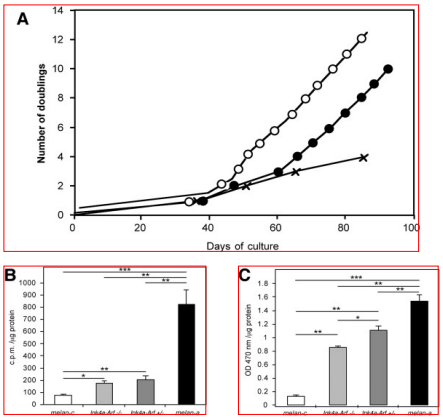
B. Structural Patterns of Text Regions

We observe that among many biomedical images, the distribution of their embedded text regions exhibits interesting spatial structures, such as images of *tables* and *flow charts*. Characterizing structural patterns of text region distribution allows us to harvest discriminative features for image categorization. In our work, we extract three categories of such features.

For a horizontal scanning line $l_i \in \mathcal{I}$, we mark pixels encountered by the scanning line that lie inside at least one text region detected from $\mathcal{I}$ as 1 and otherwise 0. We denote the resultant binary label sequence associated with the scanning line l_i as $B(l_i)$. Let $w(\mathcal{I})$ be the width of the image $\mathcal{I}$. For each $B(l_i)$, we then construct an infinite signal $S_{l_i}(x)$ where $S_{l_i}(x) = B(x)$ if $x \in [0, w(\mathcal{I})]$ and $S_{l_i}(x) = B(x - \kappa w(\mathcal{I}))$ if $x \in [\kappa w(\mathcal{I}), (\kappa+1)w(\mathcal{I})]$ in which κ is an integer. For the resultant signal $S_{l_i}(x)$, we then apply the fast Fourier transform and record the first five Fourier coefficients. Let $f_j(l_i)$ be the j-th heading Fourier coefficient computed of $S_{l_i}(x)$. Next, we formulate a set $\psi_{\text{freq},j}$ based on the j-th Fourier coefficients derived



(a) A sample image



(b) subfigures after segmentation

Figure 3. (a) A sample input image [17], (b) Sub-images detected from our image segmentation.

of all horizontal scanning lines $l_i \in \mathcal{I}$. Let $LQ(\psi)$, $Mean(\psi)$, $Median(\psi)$, and $HQ(\psi)$ respectively be the lower quartile, mean, median, and higher quartile values of a number set ψ. Using these notations, we construct five four-dimensional vectors $V_{\text{freq,h},j}$ $(j = 1, \cdots, 5)$ where $V_{\text{freq,h},j} = \{LQ(\psi_{\text{freq},j}), Mean(\psi_{\text{freq},j}), Median(\psi_{\text{freq},j}), HQ(\psi_{\text{freq},j})\}$. Similarity, we also construct another five four-dimensional vector $V_{\text{freq,v},j}$ $(j = 1, \cdots, 5)$ by analyzing the signals prepared from the binary label sequences of all vertical scanning lines of $\mathcal{I}$.

For each horizontal scanning line l_i, we merge all adjacent pixels lying on the line with label value 1 into continuous intervals. We number the resulting intervals from left to right respectively as $I^h_{i,1}, I^h_{i,2}, \cdots, I^h_{i,m_i}$, assuming there are m_i such intervals in total. The left and right boundaries of an internal $I^h_{i,j}$ are respectively denoted as $left(I^h_{i,j})$ and $right(I^h_{i,j})$. By definition, no two text intervals overlap. The distance between a pair of text intervals $I^h_{i,1}$ and $I^h_{i,2}$, denoted as $d(I^h_{i,1}, I^h_{i,2})$, is computed as $\min\{left(I^h_{i,1}) - right(I^h_{i,2}), left(I^h_{i,2}) - right(I^h_{i,1})\}$. We then compute all the distances between any two text intervals, leading to $\frac{m_i \times (m_i - 1)}{2}$ results of calculated distances. Given a threshold τ, we can cluster these $\frac{m_i \times (m_i - 1)}{2}$ distance values into several clusters using the standard hierarchical clustering algorithm [18] where the iterative cluster merging process will terminate if no two clusters' inter-cluster distance is within the threshold τ. Here we define the inter-cluster distance between two clusters as the absolute difference between the mean values of the two number sets. Assuming such a hierarchic clustering procedure leads to k clusters in the end where the j-th resulting cluster contains the elements $d_{j,1}, d_{j,2}, \cdots, d_{j,n_j}$. We then introduce the metric $\vartheta_h(l_i, \tau)$ to measure the regularity of the spatial structure lying on the horizontal scanning line l_i with respect to the threshold τ, as follows:

$$\vartheta_h(l_i, \tau) = \frac{\sqrt{\frac{m_i \times (m_i - 1)}{2}}}{k\sqrt{\sum_{x=1}^{k} \sum_{s=1}^{n_x} (d_{x,s} - \frac{1}{n_x} \sum_{s'=1}^{n_x} d_{x,s'})^2}}. \quad (1)$$

Formula (1) expresses the fact that the more structurally regularly the text intervals distribute on the scanning line l_i, the fewer number of clusters will yield at the end of the hierarchic clustering procedure and the deviation of each resultant cluster will also be very small. Both factors help contribute to a large $\vartheta_h(l_i, \tau)$ value. Based on $\vartheta_h(l_i, \tau)$, we can further measure the significance of structural distribution exhibited by text intervals on the horizontal scanning line l_i, denoted as $\vartheta_h(l_i)$, as follows:

$$\vartheta_h(l_i) = \max_{\tau} \vartheta_h(l_i, \tau). \quad (2)$$

In our implementation, to look for an optimal τ that can maximize the metric $\vartheta_h(l_i, \tau)$, we search the value range of $[0, \frac{w(\mathcal{I})}{2}]$ with an incremental searching step of $\frac{w(\mathcal{I})}{20}$.

For each horizontal scanning line $l_i \in \mathcal{I}$, we derive its $\vartheta_h(l_i)$ following a similar procedure. We then further construct a ten dimensional vector $H^{\text{structure}}_{\text{h}}(\mathcal{I})$ by computing a ten-bin histogram that equally divides the value range of $\vartheta_h(l_i)$ for $l_i \in \mathcal{I}$ into ten equal bins. The i-th component $H^{\text{structure}}_{\text{h},i}(\mathcal{I})$ of the histogram vector $H^{\text{structure}}_{\text{h},i}(\mathcal{I})$ indicates the percentage of horizontal scanning lines that fall into the bin. Following a similar procedure delineated in the above, we can also derive another ten dimensional histogram vector $H^{\text{structure}}_{\text{v}}(\mathcal{I})$ by analyzing the structural distribution of text intervals on vertical scanning lines of the input image $\mathcal{I}$.

C. Distance Distribution of Text Regions to Their Closest Neighbors

For a pair of text regions Rec_i and $Rec_j \in \mathcal{I}$, we define their mutual distance as the shortest distance between an arbitrary pixel on the boundary of Rec_i and another arbitrary pixel on the boundary of Rec_j. Let $d_{\max}(\mathcal{I})$ be the largest distance between any pair of text regions detected from within $\mathcal{I}$. We then create a

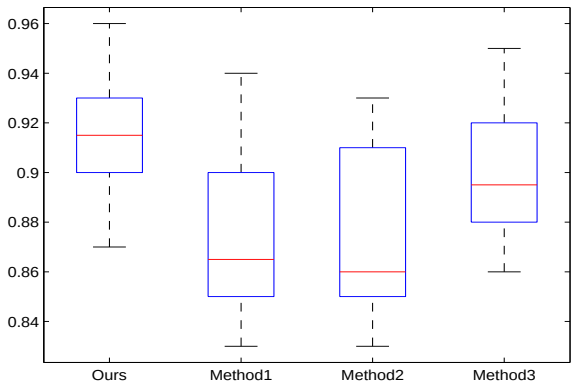


Figure 4. Comparison of accuracy between our method and the peer methods for image categorization. We illustrate the accuracy of different methods (Method 1 for [14], Method 2 for [15], and Method 3 for [20]) measured in ten rounds of our experiments, using the ten-fold cross validation scheme.

Table I
CONFUSION MATRIX FOR OUR IMAGE CATEGORIZATION RESULTS

True category	Predicted category				
	flow chart	experiment	graph	mix	others
flow chart	134	1	8	0	7
experiment	0	78	0	15	7
graph	8	3	194	25	20
mix	0	5	7	84	4
others	2	2	3	15	57

Table II
PERFORMANCE OF IMAGE CATEGORIZATION USING OUR NEWLY PROPOSED IMAGE FEATURES

Category	TP	FP	FN	Precision	Recall	F-score
flow chart	134	10	16	0.9306	0.8933	0.9116
experiment	78	12	22	0.8667	0.7800	0.8211
graph	194	18	56	0.9151	0.7760	0.8398
mixed	84	55	16	0.6043	0.8400	0.7029
others	57	38	23	0.6000	0.7125	0.6514

five-band histogram where the j-th bin of the histogram corresponds to the value range of $[\frac{d_{\max}(\mathcal{I})j}{5}, \frac{d_{\max}(\mathcal{I})(j+1)}{5}]$ $(j = 0, \cdots, 4)$. For every detected text region $Rec_i \in \mathcal{I}$, we compute its distances to its k closest neighboring text regions. We then construct the vector $H_{\text{dis}}(Rec_i)$ where its j-th component, denoted as $H_{\text{dis},j}(Rec_i)$, records the percentage of distances of Rec_i to its k nearest neighboring text regions that fall into the j-th value bin of the above histogram. Accordingly, we can create five sets $\psi_{\text{dist,j}}$ $(j = 0, \cdots, 4)$ where $\psi_{\text{dist,j}} \triangleq \{H_{\text{dis},j}(Rec_i) | Rec_i \in \mathcal{I}\}$. Further, we derive the vector $V_{\text{dis},j} \triangleq \big(LQ(\psi_{\text{dist,j}}), Mean(\psi_{\text{dist,j}}), Median(\psi_{\text{dist,j}}), HQ(\psi_{\text{dist,j}})\big)$. Let n_{Rec} be the total number of text regions detected from $\mathcal{I}$. In our experiments, we empirically found out that $k = \frac{n_{\text{Rec}}}{3}$ provides a set of most discriminative features, $V_{\text{dis},j}$ $(j = 0, \cdots, 4)$, for our image categorization purpose.

V. EXPERIMENTATION

We employed a supervised learning based method for image categorization using the image features proposed in the above. This learning method is based on the Support Vector Machine implementation offered by the free MATLAB package "LIBSVM" [19] because of its popularity and satisfying performance as widely reported in the literature. To acquire images for use in our experiments, we downloaded all open access materials released by PMC (ftp://ftp.ncbi.nlm.nih.gov/pub/pmc) until the end of year 2011. Among all the downloaded images, we randomly selected 990 JPG format images to use in our experiments. Fig. 3(a) shows a sample image used in our study.

We first segmented all our experiment images following the procedure described in Sec. III(B). For each resultant sub-image, we then manually labeled its category according to the two-level hierarchical image taxonomy introduced earlier in Sec. III.

To measure the image categorization performance of the aforementioned SVM categorizer using our image features, we conducted a set of comparative studies in our experiments. For performance benchmarking, we used recall, precision, and F-score as our metrics for evaluating image categorization performance. Table I shows the confusion matrix of our image categorization results using the newly proposed image features in this paper. Table II reports the performance of image categorization results for each of the five most prevalent image categories where TP, FP, and FN respectively stand for true positive, false positive, and false negative rates.

Table III reports the performance of image categorization, in terms of precision, recall, and F-score, for a run of our experiments using conventional image features alone versus another run of our experiments using both conventional image features and the novel image features proposed in this paper. In this comparative study, we extracted the edge-direction histogram [21] and Haralick's texture-features [22] as the conventional image features. The results of our comparative experiments clearly show the improved performance of image categorization when our novel image features were leveraged.

Figure 4 shows the accuracy of our method in comparison with that of three peer methods. This time, we also used the annotated corpus of 990 images in our experiments and adopted a ten-fold cross validation scheme to measure the performance of an image categorization method. As shown

Table III
IMAGE CATEGORIZATION PERFORMANCE USING THE CONVENTIONAL IMAGE FEATURES ALONE VERSUS WITH OUR NOVEL IMAGE FEATURES

Features used	Precision	Recall	F-score
conventional image features	0.500	0.480	0.489
conventional image features+ **our novel image features**	0.703	0.740	0.725

in the figure, the performance of our method is evidently superior to that of all three peer methods due to the use of our new image features.

VI. Conclusion

We proposed a set of novel image features that exploit the distribution of text information inside a biomedical image for image categorization. By integrating these new image features within conventional image features, we can significantly boost the performance of biomedical image categorization, whose effectiveness is verified by the results of all our experiments.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (NSFC), grant No. 60903132, as well as NSFC-Guangdong Joint Fund (U0935004, U1135003), National Key Technology R & D Program (2011BAH27B01, 2011BHA16B08), and NSFC (No. 61103162). Xu performed this research as a Eugene P. Wigner Fellow and staff member at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

References

[1] S. Xu, J. McCusker, and M. Krauthammer, "Yale image finder (YIF): a new search engine for retrieving biomedical images," *Bioinformatics*, vol. 24, no. 17, pp. 1968–1970, 2008.

[2] T. Lehmann, M. Güld, T. Deselaers, D. Keysers, H. Schubert, K. Spitzer, H. Ney, and 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.

[3] M. Giuld, M. Kohnen, D. Keysers, H. Schubert, B. Wein, J. Bredno, and T. Lehmann, "Quality of dicom header information for image categorization," in *Proc. SPIE*, vol. 4685, 2002, pp. 280–287.

[4] L. Zhang, D. Samaras, D. Tomasi, N. Volkow, and R. Goldstein, "Machine learning for clinical diagnosis from functional magnetic resonance imaging," in *IEEE Transactions on Computer Vision and Pattern Recognition*, vol. 1. IEEE, 2005, pp. 1211–1217.

[5] V. Balasubramanyam and A. Hielscher, "Classification of optical tomographic images of rheumatoid finger joints with support vector machines," *Proc. SPIE Advanced Biomedical and Clinical Diagnostic Systems III*, vol. 5692, pp. 37–43, 2005.

[6] S. Xu and M. Krauthammer, "A new pivoting and iterative text detection algorithm for biomedical images," *Journal of Biomedical Informatics*, vol. 43, no. 6, pp. 924–931, 2010.

[7] T. Deselaers, T. Weyand, and H. Ney, "Image retrieval and annotation using maximum entropy," *Evaluation of Multilingual and Multi-modal Information Retrieval*, pp. 725–734, 2007.

[8] B. Funt and G. Finlayson, "Color constant color indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 522–529, 1995.

[9] E. Hadjidemetriou, M. Grossberg, and S. Nayar, "Spatial information in multiresolution histograms," *Computer Vision and Pattern Recognition. (CVPR 2001)*, vol. 1, pp. 695–702, 2001.

[10] C. Liu, J. Ma, and G. Ye, "Medical image segmentation by geodesic active contour incorporating region statistical information," *Fuzzy Systems and Knowledge Discovery*, vol. 3, pp. 63–67, 2007.

[11] T. Deselaers, T. Weyand, D. Keysers, W. Macherey, and H. Ney, "Fire in imageclef 2005: Combining content-based image retrieval with textual information retrieval," *Accessing Multilingual Information Repositories*, pp. 652–661, 2006.

[12] T. Tommasi, F. Orabona, and B. Caputo, "Discriminative cue integration for medical image annotation," *Pattern Recognition Letters*, vol. 29, no. 15, pp. 1996–2002, 2008.

[13] ——, "An svm confidence-based approach to medical image annotation," *Evaluating Systems for Multilingual and Multimodal Information Access*, pp. 696–703, 2009.

[14] Y. Chen and J. Wang, "Image categorization by learning and reasoning with regions," *The Journal of Machine Learning Research*, vol. 5, pp. 913–939, 2004.

[15] H. Shatkay, N. Chen, and D. Blostein, "Integrating image data into biomedical text categorization," *Bioinformatics*, vol. 22, no. 14, pp. e446–e453, 2006.

[16] R. Murphy, Z. Kou, J. Hua, M. Joffe, and W. Cohen, "Extracting and structuring subcellular location information from on-line journal articles: The subcellular location image finder," *Proceedings of the IASTED International Conference on Knowledge Sharing and Collaborative Engineering*, pp. 109–114, 2004.

[17] A. Lavado, A. Matheu, M. Serrano, and L. Montoliu, "A strategy to study tyrosinase transgenes in mouse melanocytes," *BMC cell biology*, vol. 6, no. 1, pp. 18–27, 2005.

[18] Y. Zhao, G. Karypis, and U. Fayyad, "Hierarchical clustering algorithms for document datasets," *Data Mining and Knowledge Discovery*, vol. 10, no. 2, pp. 141–168, 2005.

[19] C. Chang and C. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, pp. 144–152, 2011.

[20] J. Shotton, M. Johnson, and R. Cipolla, "Semantic texton forests for image categorization and segmentation," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008.

[21] A. Jain and A. Vailaya, "Shape-based retrieval: A case study with trademark image databases," *Pattern Recognition*, vol. 31, no. 9, pp. 1369–1390, 1998.

[22] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.