Informative retrieval framework for histopathology whole slides images based on deep hashing network

Dingyi Hu[a,b], Huining Yuan[a], Yushan Zhengu[a,b], Haopeng Zhang[a,b], Jun Shi[c], and Zhiguo Jiang[a,b]

[a]Image Processing Center, School of Astronautics, Beihang University, Beijing, China;
[b]Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing, China;
[c]School of Software, Hefei University of Technology, Hefei 230601, China

ABSTRACT

Histopathology image retrieval is an emerging application for Computer-aided cancer diagnosis. However, most of current retrieval methods have ignored the characteristic of histopathology images. It causes repeated results within similar image content from the same slide. Meanwhile, the data sets cannot be sufficiently utilized. To solve these issues, we proposed an informative retrieval framework based on deep hashing network. Specifically, a novel loss function for the hashing network and a retrieval strategy are designed, which contributes to more informative retrieval results without reducing the retrieval precision. The proposed method was verified on the ACDC-LungHP dataset and compared with the state-of-the-art method. The experimental results have demonstrated the effectiveness of our method in the retrieval of large-scale database containing histopathology while slide images.

Keywords: C omputer-aided diagnosis, retrieval, histopathology images, hashing network.

1. INTRODUCTION

The digital pathology system based on digital whole slide images (WSIs) is built and widely applied to clinical diagnosis. While the manual diagnosis with digital WSIs is still less efficient. Therefore, an increasing number of computer-aided diagnosis (CAD) methods based on digital WSIs have been developed. Content-based image retrieval(CBIR)[1–3] can recall the most similar images and the corresponding diagnosis information from the database. An increasing number of researches have been made on the histopathology image retrieval in the last decade. The previous works mainly conducted on little-scale datasets. The retrieval frameworks were proposed based on traditional manual feature extraction methods[4–6] and especially the classical feature descriptors such as SIFT ,[7,8] HOG [8] and GIST .[9] Then, the semantic analysis models[10,11] were established based on these manual features. Ma [12] designed an unsupervised hashing-based retrieval method, where the image features were compressed to topic histograms using latent Dirichlet allocation model and converted into binary codes for fast retrieval. In recently years, the retrieval methods of convolutional neural network (CNN) have been proposed, which have proven more powerful than the methods based on traditional features. Typically, Shi et al.[2] proposed a deep ranking hashing that converted the image patches into binary representations, which emphasized the distinction across different classes and within the same class. Based on the CNN, the hashing-based retrieval method is efficient and can adapt to the huge amount of data. However there is still a main limitation that the present methods are mainly developed for the patch-level retrieval. As is shown in Fig. 1, when applied to database consisting of WSIs, the returned results are usually crowded with adjacent regions from numbered WSIs (Fig. 1. (a)). The results are redundant and the retrieval information is less efficient for clinical applications. The retrieval results are expected to return from different WSIs (Fig. 1. (b)).

In this paper, we propose a novel deep hashing method for the retrieval of histopathological images from WSIs database. The framework is illustrated in Fig. 2. In the training stage, the WSIs are sampled into image patches (Fig. 2 (a)). Then, the image patches are used to train a deep hashing network(Fig. 2 (b)). In the online retrieval stage, the query image is first encoded into binary codes using the trained deep hashing network. Next, the similarities between the query image and the samples in the database are measured based

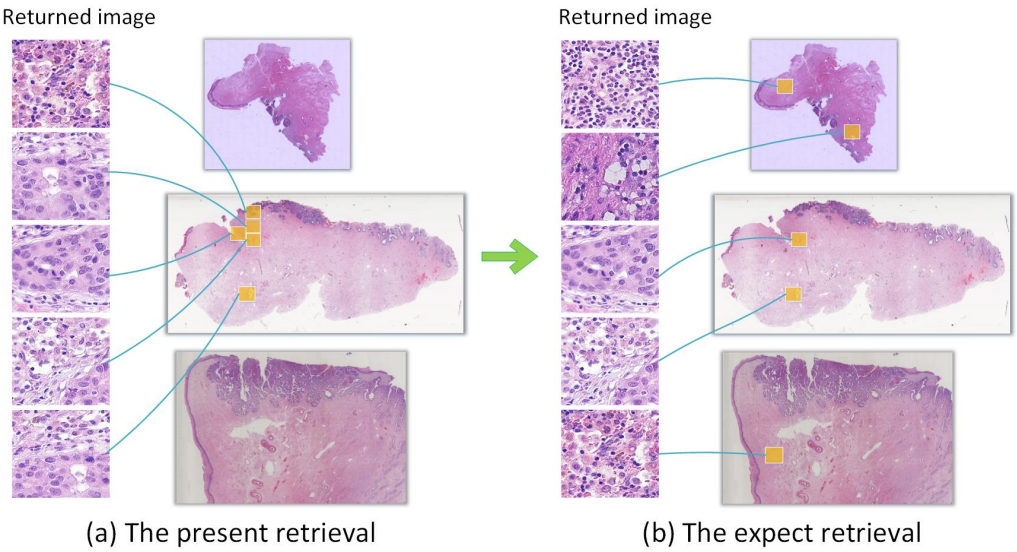

Figure 1. (a) The limitation of the present retrieval. (b) The expect retrieval results.

on hamming distance. Finally, the most similar regions are returned to pathologists after filtering the redundant results(Fig. 2 (d)). The proposed method was evaluated on ACDC-lungHP database. We compared the method with three state-of-the-art methods.[2,3,13] The result indicates that our method is effective in histopathological image retrieval from the WSI database.

The contribution of our work to the problem is three-fold:

1) We proposed a novel informative retrieval framework for histopathology whole slide images based on deep hashing network.

2) We designed a specific hashing loss function for histopathological whole slide images. The loss function can unify the region similarities inner- and intra-WSIs. Consequently, the regions from various WSIs can be returned.

3) We proposed a novel retrieval strategy based on Non-maximum Suppression (NMS) to relieve the redundancy from adjacent regions in a WSI.

2. METHOD

In this section, we first introduce the deep-CNN-based retrieval structure and then discuss the loss function. Finally, the proposed retrieval strategy is detailed.

2.1 Deep hashing network

Hashing method can leverage limited storage space to achieve large-scale retrieval. Letting $\{I_1, I_2, ..., I_N\}$ represent N image samples, the hashing function is defined as

$$b_n = sgn(\mathcal{F}_{CNN}(I_n)), n = 1, 2, ..., N,$$

where $b_n \in \{-1, 1\}^d$ is the binary code of I_n, $sgn(\cdot)$ is the sign function with $sgn(t) = 1$ if $t > 0$ and $sgn(t) = -1$ otherwise, $\mathcal{F}_{CNN}(\cdot)$ represents a feature extractor with CNN backbone that compresses the images into discriminate features.

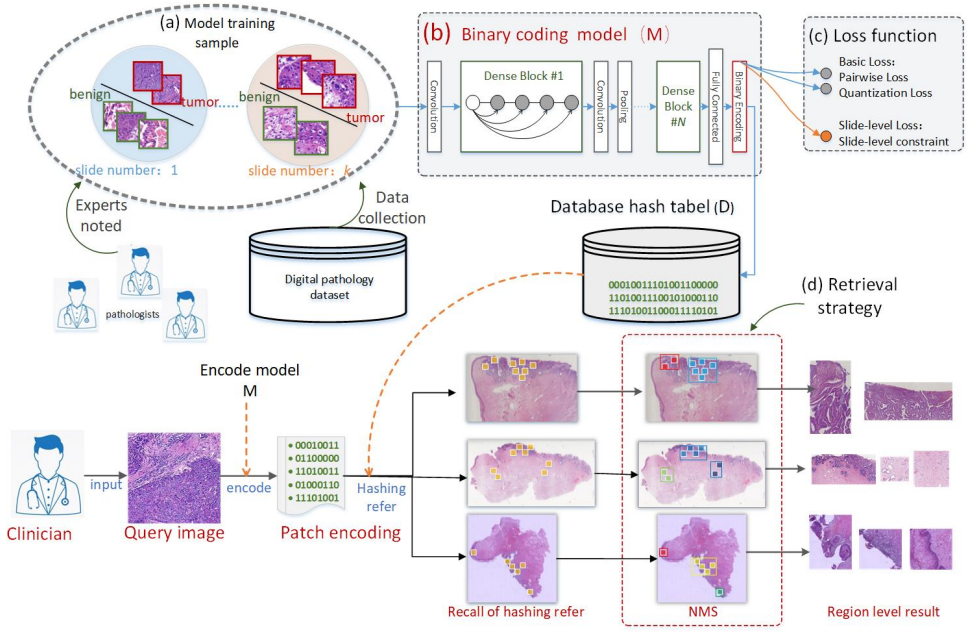

Figure 2. The overview of proposed retrieval methods.

2.2 Loss function

2.2.1 Basic loss function

Setting Γ as the RGB space, the network aims to map histopathological images from Γ into compact binary code: $\Gamma \longrightarrow b_n \in \{-1,1\}^{d\times 1}$. The network is trained to compress similar image pairs into close binary codes, and keep the dissimilar image pairs with far binary codes. To this end, the loss function is designed.

Specifically, for a pair of image $I_1, I_2 \in \Gamma$ and the corresponding binary codes $b_1, b_2 \in \{-1,1\}^{d\times 1}$, we set $x = 1$ while the two images are in the same class and $x = 0$, otherwise. The basic loss function refers to the definition in paper.[13]

$$\begin{aligned} L(c_1,c_2,x) = (1-x)max(m - \|c_1 - c_2\|^2, 0) \\ + x\|c_1 - c_2\|^2 + \beta(\||c_1| - 1\| + \||c_2| - 1\|), \end{aligned} \tag{1}$$

where $c_1 = \mathcal{F}(I_1)$ and $c_2 = \mathcal{F}(I_1)$ are CNN features of I_1 and I_2, β is a hyper-parameter, where $(\||c_1|-1\|+\||c_2|-1\|)$ is used to impel the $|c_1|, |c_2|$ close to 1.[13]

The Loss function above is designed to image patches retrieval, however it ignores the particular characteristics of histopathology images. In histopathology sections, there are large regions of homogeneous tissues. They are usually clustered in the code space and are very likely be returned together for a certain query. These regions will be counted as correct retrieval by the traditional metrics for retrieval. But, these regions are redundant information for cancer diagnosis and will occupy the queue of images returned to pathologists. It makes pathologists less efficient to obtain regions from different positions in a WSI and even from different WSIs.

To overcome this issue, we added a slide-level constraint to the basic hashing loss, to improve the diversity of WSIs involved in the retrieval result. Correspondingly, we designed a specific retrieval strategy to reduce the redundancy raised by adjacent tissue regions.

2.2.2 Slide-level constraint (SLC)

The slide-level constraint is expected to reduce the distance between the relevant regions from different slides while gain the distance between regions from the same slide. Therefore, the SLC is defined as:

$$\begin{aligned} L_s(c_1, c_2, y) = \rho(y\|c_1 - c_2\| \\ + (1-y)max(m_2 - \|c_1 - c_2\|, 0)), \end{aligned} \tag{2}$$

where $y = 1$ if the image pair is from different slides and $y = 0$, otherwise, ρ is a weight to control the proportion of slide constraint. Supposing that there are N training pairs in one batch, the loss function to minimize is formulated as:

$$L^* = \Sigma_{i=1}^{N}[L(c_{i,1}, c_{i,2}, x) + L_s(c_{i,1}, c_{i,2}, y)] \tag{3}$$

2.3 Retrieval strategy

The loss function can enhance the slide diversity to a certain extent. Meanwhile, we propose to further improving the performance through better retrieval strategy. Specifically, we performed NMS in each slide to merge the adjacent image patches retrieved by our method.

Overall, the algorithm of the proposed retrieval strategy is summarized in Algorithm 1.

Algorithm 1: The algorithm of the designed retrieval strategy.

Input:
$M \leftarrow$ The number of WSIs in the database;
$N \leftarrow$ The total number of patches involved in the WSIs;
$I_i \leftarrow$ The i-the patch;
$I_q \leftarrow$ The query image;
$D \leftarrow$ The threshold of spatial distance;
Output: $R_1, R_2, ..., R_S \leftarrow$ The S returned regions;

$b_q \leftarrow sgn(\mathcal{F}_{CNN}(I_q))$;
for $i = 1$ *to* N **do**
  $b_i = sgn(\mathcal{F}_{CNN}(I_i))$;
end
$R_1, R_2, ..., R_N \leftarrow$ Sort I_i according to $Hamming(b_i, b_q)$;
for $m = 1$ *to* M **do**
  for *pair* (I_a, I_b) *in slide* m **do**
    $R_a \leftarrow$ Fuse I_a, I_b if $d(I_a, I_b) < D$
  end
end
return $R_1, R_2, ..., R_S$;

3. EXPERIMENTS

The experiment was conducted on the ACDC-LungHP dataset, which contains 150 H&E stained biopsy samples with lung cancer under 40x magnification. Experienced pathologists have manually annotated the cancer regions on tissue level . 120 WSIs were randomly selected to train the model. The remainder 30 WSIs were used as testing data. We segmented each WSI into patches in size of 224 x 224 following sliding-window paradigm. As a result, there are 1,936,156 and 374,544 patches in training and testing dataset. We employed the DenseNet-121 structure[14] to extract the CNN features. In the query process, we defined the retrieval result is correct when the returned images sharing the same label with the query image.

Table 1. Retrieval performance in the database of WSIs

Methods	$d = 12$			$d = 24$			$d = 36$			$d = 48$		
	map10	map20	ϕ_S	map10	map20	ϕ_S	map10	map20	ϕ_S	map10	map20	ϕ_S
PDRH[2]	0.754	0.735	3.183	0.763	0.741	3.067	0.758	0.738	3.02	0.762	0.743	2.97
DCH[3]	**0.761**	0.74	3.033	0.759	0.742	2.845	0.763	0.742	2.74	0.76	0.74	2.706
DSH[13]	0.757	0.737	3.443	**0.768**	**0.745**	3.436	**0.768**	**0.746**	3.429	0.762	0.739	3.413
Ours w/o SLC	0.757	0.736	3.877	0.766	0.743	3.85	**0.768**	0.745	3.879	0.761	0.736	3.884
Ours w/o SNMS	**0.761**	**0.741**	3.456	0.767	**0.745**	3.458	0.755	0.735	3.453	**0.768**	**0.747**	3.454
Ours	**0.761**	0.739	**3.892**	0.766	0.743	**3.889**	0.754	0.733	**3.89**	0.767	0.746	**3.898**

3.1 Evaluation metric

The mean average precision (MAP) for retrieval is used as the evaluation metric. Furthermore, we designed a novel metric based on information entropy to evaluate the amount of information referring to slide-level diversity of the returned results. The metric is defined as

$$\Phi_S(K) = -\Sigma_{i=1}^{M} \frac{N_i(K)}{\Sigma_{j=1}^{M} N_j(K)} \lg \frac{N_i(K)}{\Sigma_{j=1}^{M} N_j(K)}, \tag{4}$$

where $N_i(K)$ is the number of patches from the i-th slide in the top-K-returned results, M represents the number of slides in the database. The metric $\phi_S(K)$ is positive consistent with the information of the retrieval.

3.2 Result and discussion

We compared our method with 3 state-of-the-art retrieval frameworks for histopathological image retrieval. They are pair-wise based Deep Ranking Hashing (PDRH),[2] Deep Convolutional Hashing (DCH)[3] and Deep Supervised Hashing (DSH).[13] Moreover, we conducted ablation experiments to certify the effectiveness of the proposed slide-level constraint (SLC) (in Eq. 2) and the NMS-based retrieval strategy. The results are presented in Table 1.

From Table 1, the proposed method has the best information entropy over all the compared methods, which means the method can provide more diagnosis information. The information entropy significantly decreases when the SLC or the SNMS are erased. It demonstrates that the proposed loss function and retrieval strategy is effective. Meanwhile, the MAP slightly decreases when the the SNMS is applied. It is because the strategy has removed the redundant results in the top-ranked images and thus the irrelevant regions are moved forward in the return queue. Fig. 3. presents the qualitative comparison, where the correct images are framed in green and the incorrect images are in red. It shows that the DCH,[3] PDRH,[2] DSH[13] can achieve good accuracy, however there are lots of redundant results from the same WSIs. The diagnostic information provided to the pathologists is limited. In comparison, our method has considered the spatial redundancy and slide-level diversity in the retrieval results. Therefore, the returned images are from various WSIs meanwhile are relevant to the query image.

4. CONCLUSION

In this paper, We proposed an informative retrieval framework for histopathology WSIs based on deep hashing network. There are two main contributions about this work. First, we designed a novel loss function to perform slide-level constraint to the hash encoding, which is proven effective in improving the information of the retrieval. Second, we designed a novel retrieval strategy based on NMS, which can merge the redundancy patches from adjacent regions of the WSI. The experimental results have demonstrated the proposed retrieval framework is informative and effective in histopathological WSI-database retrieval.

REFERENCES

[1] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma, H. Shi, and Y. Zhao, "Histopathological whole slide image analysis using context-based cbir," *IEEE transactions on medical imaging*, vol. 37, no. 7, pp. 1641–1652, 2018.

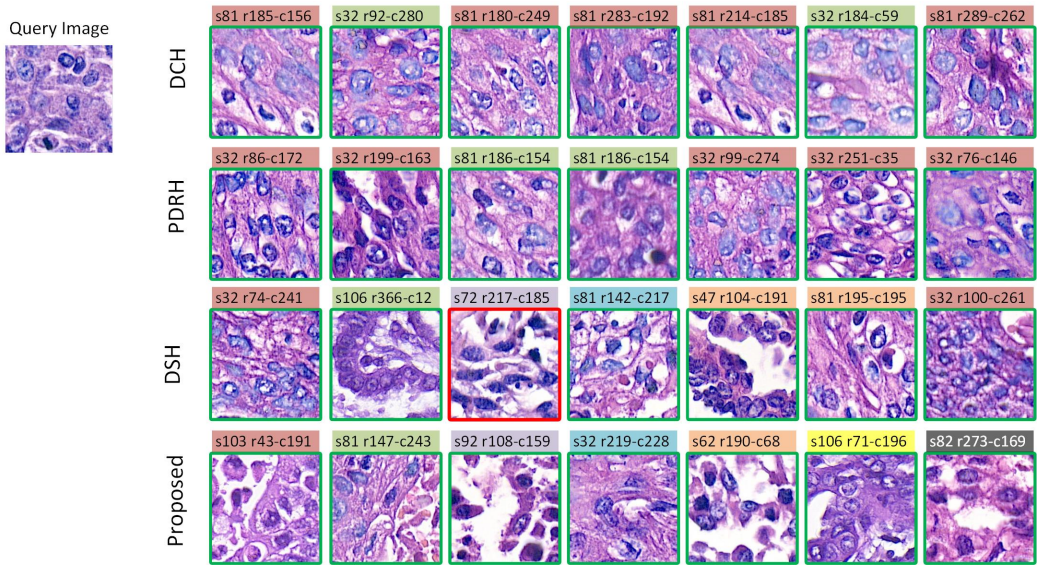

Figure 3. Examples of retrieval using different retrieval method, where the left image is query images and right images are returned top 7 images. Where s represents the WSI number, r and c represents the row and columns coordinates in WSIs, the patches returned from the same WSI are annotated in the same color.

[2] X. Shi, M. Sapkota, F. Xing, F. Liu, L. Cui, and L. Yang, "Pairwise based deep ranking hashing for histopathology image classification and retrieval," *Pattern Recognition*, vol. 81, pp. 14–22, 2018.

[3] M. Sapkota, X. Shi, F. Xing, and Y. Lin, "Deep convolutional hashing for low dimensional binary embedding of histopathological images," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2018.

[4] 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. 249–55, 2004.

[5] G. Zhou, L. Jiang, L. Luo, X. Bao, and H. Shu, "Content-based cell pathology image retrieval by combining different features," in *Medical Imaging 2004: PACS and Imaging Informatics*, vol. 5371. International Society for Optics and Photonics, 2004, pp. 326–333.

[6] N. Mehta, R. S. Alomari, and V. Chaudhary, "Content based sub-image retrieval system for high resolution pathology images using salient interest points." in *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 3719–22.

[7] X. Zhang, W. Liu, and S. Zhang, "Mining histopathological images via hashing-based scalable image retrieval," in *2014 IEEE 11th international symposium on biomedical imaging (ISBI)*. IEEE, 2014, pp. 1111–1114.

[8] M. Jiang, S. Zhang, J. Huang, L. Yang, and D. N. Metaxas, "Scalable histopathological image analysis via supervised hashing with multiple features," *Medical image analysis*, vol. 34, pp. 3–12, 2016.

[9] J. Liu, S. Zhang, W. Liu, X. Zhang, and D. N. Metaxas, "Scalable mammogram retrieval using anchor graph hashing," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2014, pp. 898–901.

[10] J. C. Caicedo, F. A. Gonzalez, and E. Romero, "A semantic content-based retrieval method for histopathology images," in *Asia Information Retrieval Conference on Information Retrieval Technology*, 2008, pp. 51–60.

[11] J. C. Caicedo and E. Izquierdo, "Combining low-level features for improved classification and retrieval of histology images," *Ibai Publishing*, vol. 2, no. 1, pp. 68–82, 2010.

[12] Y. Ma, Z. Jiang, H. Zhang, F. Xie, Y. Zheng, H. Shi, and Y. Zhao, "Breast histopathological image retrieval based on latent dirichlet allocation," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2016.

[13] H. Liu, R. Wang, S. Shan, and X. Chen, "Deep supervised hashing for fast image retrieval," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2064–2072.

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.