Weighted Hashing with Multiple Cues for Cell-Level Analysis of Histopathological Images

Xiaofan Zhang[1], Hai Su[2], Lin Yang[2], and Shaoting Zhang[1(✉)]

[1] Department of Computer Science, UNC Charlotte, Charlotte, NC 28223, USA
shaoting@cs.rutgers.edu
[2] J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, USA

Abstract. Recently, content-based image retrieval has been investigated for histopathological image analysis, focusing on improving the accuracy and scalability. The main motivation is to interpret a new image (i.e., query image) by searching among a potentially large-scale database of training images in real-time. Hashing methods have been employed because of their promising performance. However, most previous works apply hashing algorithms on the whole images, while the important information of histopathological images usually lies in individual cells. In addition, they usually only hash one type of features, even though it is often necessary to inspect multiple cues of cells. Therefore, we propose a probabilistic-based hashing framework to model multiple cues of cells for accurate analysis of histopathological images. Specifically, each cue of a cell is compressed as binary codes by kernelized and supervised hashing, and the importance of each hash entry is determined adaptively according to its discriminativity, which can be represented as probability scores. Given these scores, we also propose several feature fusion and selection schemes to integrate their strengths. The classification of the whole image is conducted by aggregating the results from multiple cues of all cells. We apply our algorithm on differentiating adenocarcinoma and squamous carcinoma, i.e., two types of lung cancers, using a large dataset containing thousands of lung microscopic tissue images. It achieves 90.3 % accuracy by hashing and retrieving multiple cues of half-million cells.

1 Introduction

Content-based image retrieval (CBIR) has been an effective approach in analyzing medical images [1–3]. It aims to retrieve and visualize relevant medical images with diagnosis information, which assist doctors to make consistent clinical decisions. Successful use cases include clinical pathology, mammogram analysis, and categorization of X-ray images [1,4–7]. Recently, the research focus of CBIR for medical images has been on the efficient and large-scale methods, and several benchmarks have been designed, such as Image CLEF and VISCERAL [2,8].

S. Ourselin et al. (Eds.): IPMI 2015, LNCS 9123, pp. 303–314, 2015.
DOI: 10.1007/978-3-319-19992-4_23

The motivation of leveraging large databases of training images is that they have the potential to offer abundant information to precisely interpret the new data. However, the scalability or efficiency of these algorithms is usually an issue. In this paper, we design a scalable CBIR algorithm to tackle a challenging problem, differentiating lung cancers using histopathological images. Lung cancer is one of the most common cancers [9]. There are four typical histologic types of lung cancers, including adenocarcinoma, squamous carcinoma, small cell carcinoma, and large cell carcinoma, each of which needs a different treatment [10]. Therefore, early diagnosis and differentiation of these four types is clinically important. Bronchial biopsy is one of the most effective diagnosis methods to differentiate them, with the aid of Computer Aided Diagnosis (CAD) systems [11–13]. Many previous methods emphasize the diagnosis of small cell vs. non-small cell (i.e., adenocarcinoma, squamous carcinoma, and large cell carcinoma) types of lung cancers, achieving promising accuracy. On the other hand, differentiation of the adenocarcinoma and squamous carcinoma, both of which are non-small cells, is much more difficult, while it is also clinically significant, since management protocols of these two types of cancers are different [14]. This is challenging because the difference between the adenocarcinoma and squamous carcinoma highly depends on the cell-level information, e.g., its morphology, texture and appearance, requiring to analyze multiple cues of all cells for accurate diagnosis. In fact, thoroughly examining cell-level information is necessary for many use cases of biopsy or histopathological image analysis. To this end, a potential solution proposed recently is to extract high-dimensional local texture features (e.g., SIFT [15]) that align with the cell-level information, and then compress them as binary codes via hashing-based CBIR algorithms [16] to improve the computational efficiency. Hashing methods [17–20] have been intensively investigated in machine learning communities, which enable fast approximated nearest neighbors (ANN) search for scalability. However, information loss is inevitable in such holistic approximation of cell-level information. Segmenting and hashing each cell offers a potential solution [21], and our framework also follows this strategy. Nonetheless, it only utilizes one type of features, while multiple cues should be examined for accurate classification.

Different from these previous methods, our proposed framework is able to (1) hash multiple cues of all cells with weights, and (2) accommodate new training data on-the-fly. Specifically, each cue of a cell is represented as binary codes through supervised hashing with kernels [20]. Then, the importance of such hash entry is determined adaptively according to its discriminativity for differentiating different classes, based on several measures such as probability scores. Given these probability scores, we integrate multiple features by considering importance. An additional benefit of this design is that the hashing results of multiple cues can be updated on-the-fly when handling new training samples. The classification of the whole image is conducted by aggregating the results of multiple cues of all cells. We evaluate our algorithm on this specific problem of differentiating lung cancers, using a large dataset containing 1120 lung microscopic tissue images acquired from hundreds of patients, achieving an accuracy of 90.3 % within several seconds by searching among half-million cells.

The rest of the paper is organized as follows. Section 2 presents our framework for real-time cell examination by hashing multiple cues of cells, including details of the hashing method, probabilistic-based weighting schemes, and aggregation of multiple cues. Section 3 shows the experimental results and comparisons. Section 4 draws the concluding remarks and shows future directions.

2 Methodology

2.1 Overview

The overview of our proposed framework is illustrated in Fig. 1. It includes automatic cell detection and segmentation, supervised hashing that generates binary codes from multiple types of image features, and the probabilistic-based weighting scheme that decides the importance of the hash entry for each cell. Specifically, the segmentation is based on the off-the-shelf method [22], while many other methods and systems are also applicable for this task [23,24].

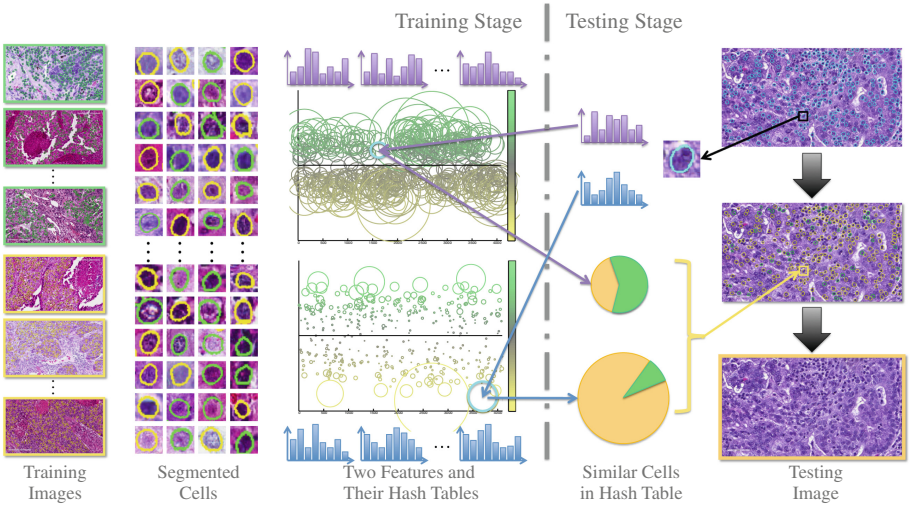

Fig. 1. Overview of our framework. In the training stage, the input is histopathological images representing two types of lung cancers. Green stands for adenocarcinoma, and yellow stands for squamous carcinoma. First, all cells are detected and segmented from these images. Second, two types of texture features are extracted and compressed as binary codes by hashing methods. These hash codes are visualized in two hash tables, representing two features, according to their ability to differentiate two categories (details in Sect. 2.3). In the testing stage, all cells are segmented from the query image, from which feature and binary codes are obtained using the same preprocess. Each cell is mapped into hash tables to search for the most similar cases, which are used to interpret the category of this unknown cell. The hash entries of two features are then integrated to enhance the accuracy. Finally, the results of all cells are aggregated to classify the testing image (Color figure online).

After segmenting all cells from the training images, a large-scale database of half-million cell patches is created. Then, two types of texture features [25,26] are extracted for each cell, within the segmented region. After that, kernelized and supervised hashing (KSH) [20] is employed as a baseline to compress these features as binary code, since it can bridge the semantic gap of image appearances and their labels, which is essential for medical image retrieval. However, different from hashing the whole image, hashing cells (i.e., sub-regions of the whole image) is more challenging, due to cells' high intra-class but low inter-class variations. Therefore, traditional hashing methods result in low-discriminative hash entries. In addition, it is necessary to integrate multiple features from each cell, so the information can be largely explored. Our solution is the probabilistic-based weighting schemes that stress discriminative hash entries, and the integration of multiple cues of cells through the probability scores. Given a testing image, the same framework is utilized to segment cells, extract their features, and hash them for real-time comparison with the training database. Each cell is assigned with multiple weights or probability scores after this matching process. Finally, the classification of the testing image is achieved by aggregating the probability scores of all its cells. In the following sections, we introduce the details of the employed hashing method and our proposed strategy for cell-level analysis.

2.2 Kernelized and Supervised Hashing for Large-Scale Image Retrieval

In this section, we briefly introduce the hashing method employed as our baseline. For each segmented cell, two features are extracted, i.e., GIST [25] and HOG [26], and both of which are hundreds of dimensions, causing issues for the computational efficiency of comparing all samples. To this end, hashing methods have been widely used to compress features into binary codes with merely tens of bits. As a result, such short binary features allow mapping into a hash table for efficient retrieval, e.g., constant-time. To improve the accuracy, the kernelized scheme [18] is usually utilized to handle practical data that is mostly linearly inseparable:

$$h = sgn\left(\sum_{j \in \text{anchors}} \left(\kappa(\mathbf{x}_j, \mathbf{y}) - \frac{1}{n}\sum_{i=1}^{n}\kappa(\mathbf{x}_j, \mathbf{x}_i)\right) a_j\right), \tag{1}$$

where $\mathbf{y}$ is the feature (e.g., GIST or HOG) to be compressed as binary code, $\mathbf{x}_i$ with i from 1 to n means all training samples, i.e., cell patches, $\mathbf{x}_j$ denotes the anchors, i.e., random samples selected from the data, h is the kernelized hashing method taking the sign value of a kernel function with kernel κ, and a_j is the coefficient determining hash functions. The resulting binary codes can be used for indexing and differentiating different categories. Although kernelized scheme well solves the linear inseparability problem of features, it is still not able to provide accurate retrieval or classification of cell images, because of their large

variations. Therefore, supervised information [20] can be leveraged to bridge the semantic gap by designing more discriminative hash functions:

$$\min_{A \in \mathbb{R}^{m \times r}} \mathcal{Q}(A) = \left\| \frac{1}{r} sgn(\bar{K}_l A)(sgn(\bar{K}_l A))^T - S \right\|_F^2 \tag{2}$$

where S is a matrix encoding the supervised information (e.g., 1 for same category and -1 for different categories, which is applicable to multi-class problems) and A is the model parameter to compute hashing code, and $\bar{K}_l = [\bar{\mathbf{k}}(\mathbf{x}_1), \cdots, \bar{\mathbf{k}}(\mathbf{x}_l)]^T \in \mathbb{R}^{l \times m}$ is the matrix form of the kernel function, in which $\bar{\mathbf{k}}(\mathbf{x}_i)$ is a kernelized vectorial map $\mathbb{R}^d \mapsto \mathbb{R}^m$, $A = [\mathbf{a}_1, \cdots, \mathbf{a}_r] \in \mathbb{R}^{m \times r}$. The optimization of $\mathcal{Q}$ is based on Spectral Relaxation [27] for convexification, which is used as a warm start, and Sigmoid Smoothing that applies standard gradient descent technique for accurate hashing.

Indexing these compressed features in a hash table, our framework can match each cell of the testing image with all cells in the training database in constant-time. The category of each cell is decided straightforwardly with the majority logic of retrieved cells, and the whole image is hence classified by aggregating results of all cells from the testing image. The whole process is very efficient and takes 1–2 s.

2.3 Weighted Hashing with Multiple Cues

Despite its efficacy in large-scale image retrieval, KSH still has several limitations when dealing with our use case, which requires to hash a large number of cell images. First of all, it builds hash functions for one type of feature, while it is preferred to model multiple cues of cells for accurate classification. Second, multiple cells can be mapped into the same hash entry using KSH, i.e., the hamming distances among them are zero. In this case, one may use majority voting to decide the label of a testing cell image having the same hash entry. However, cell images from different categories can be easily mapped into the same hash entry, due to image noise, erroneous segmentation results, and the low inter-class variations. In other words, not all hash entries are reliable for classification. Figure 2 visualizes the hash tables of two features, GIST [25] and HOG [26], representing the texture characteristics of cells. The entries in each hash table are illustrated according to the distribution of cells mapped into them, such as the ratio between two categories and the number of cells mapped into that entry. The indecisive hash entries are usually around the 0.5 ratio, indicating equal opportunity for either category. The small circles in Fig. 2 are also not reliable, since only few cells are mapped there, which can be easily affected by the image noise or erroneous segmentation. A potential solution is to identify reliable hash entries and omit indecisive one, by heuristically select or prune them via feature selection. However, this may involve tuning parameters and have difficulties in modeling multiple cues of cells because of lacking the consistent measures. Furthermore, it is hard to guarantee that the selected hash entry is sufficiently discriminative for classification. Therefore, we introduce a unified formulation to solve these problems in a principled way. First, probability scores

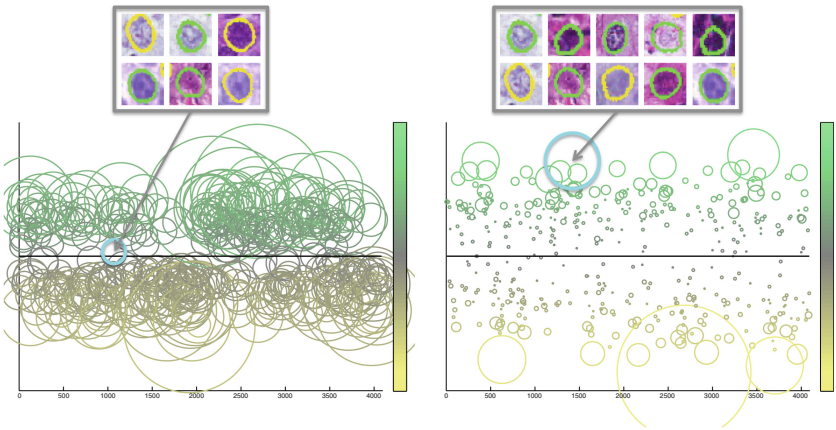

Fig. 2. We visualize hash tables and their entries according to cells mapped into them, and each circle represents one entry. The left hash table corresponds to HOG feature, and the right is GIST. The x-axis represents different hash entries corresponding to 12 bits, indicating 4096 different entries. The y-axis means the ratio between the adenocarcinoma and squamous carcinoma, ranging from 0 to 1, which is also visualized as different colors. The size of each circle denotes the number of cells mapped into that entry. As shown in the figures, different features may result in diverse cell distributions in the hash table, making it essential to explore consistent measures for feature fusion.

are assigned to each hash entry, based on its ability to differentiate different categories. Then, such probability scores can be integrated from different types of features, by emphasizing reliable hash entries of certain features. In this section, we introduce the details of our method.

Probabilistic-Based Weights for Hashing: We define three types of weights to emphasize discriminative hash entries. These weights can be consistently compared among different hash tables, representing multiple features.

- The first weight is defined as the conditional probability of a cell belonging to the ith category when its hash entry is H: $P(L_i|H) = P(L_i, H)/P(H) = |\{\text{cell} : l(\text{cell}) = L_i, \text{cell} \in S_H\}|/|S_H|$, where $S_H = \{\text{cell} : h(\text{cell}) = H\}$ is the set of cells mapped into a specific hash entry H, $|S|$ is the number of element in set S, $h(\text{cell})$ is the hash entry of this cell, $l(\text{cell})$ is the label of a cell image and L_i means the ith label or category. This represents the confidence of assigning a label to this hash entry. Instead of giving hash entry a hard label by majority voting and ignoring the minority categories, this soft assignment from probability distribution on all categories can fully utilize training data's information in the hash entry.
- The second weight is based on the information entropy E_H that is calculated from the probability distribution of each category in a hash entry: $E_H = -\sum_{i \in \text{labels}} P(L_i|H) \log(P(L_i|H))$. The entropy measures uncertainty

of a hash entry. High entropy means that it is not discriminative enough, e.g., cells mapped into the same hash entry H are evenly distributed in all categories. To reduce the importance of these non-discriminative entries, we define $W_H^E = 1 - E_H$.

– The third weight is decided according to the number of cells mapped into this entry. Entries with fewer cells are assigned lower weights, since they may be easily affected by image noise and erroneous segmentation results, i.e., less reliable compared to entries with more cells. The third weight W_H^S is defined as: $W_H^S = |S_H| / \sum_{k=0}^{2^r-1} |S_k|$, where r is the number of hash bits, representing 2^r hash values.

Combining these three weights together, we can get a probability-based score $W_{i,H} = W_H^S W_H^E P(L_i|H)$ for hash value H in the ith category. With these weights, we can utilize all training samples and reduce the influence of hash entries that are not discriminative. During the training process, $P(L_i|H)$, W_H^E and W_H^S can be computed for all hash entries. The category of a whole testing image is decided by $\arg\max_{\{i\}} \sum_{\text{cell}\in\text{query}} W_{i,H_{\text{cell}}}$, where H_{cell} is the hash value of the cell belonging to the query or testing image.

Feature Fusion and Selection for Hashing Entries: Since these weights are based on probabilities, they are comparable among multiple features. For example, hash entries that are able to differentiate different categories should be advocated in different features. Therefore, feature fusion and selection can also be designed based on the proposed framework. When there are multiple types of features, hash tables are built for each of them and the weights of every hash entries in these hash tables are calculated during the training stage. To search cells for the query image, we first extract those types of features, denoted as F_j, $j \in \{1, 2, ..., N\}$, where N is the number of features, map them into hash entries H_{F_j} in these hash tables, and calculate their weights $W_{i,H_{F_j}}$. For feature fusion, the weights can be accumulated as $W_{i,H} = \sum_{j=1}^{N} W_{i,H_{F_j}}$, indicating that all these features contribute equally to the classification. For feature selection, the maximum of the weights can be chosen, $W_{i,H} = max(W_{i,H_{F_1}}, ..., W_{i,H_{F_N}})$, meaning that the most reliable one (e.g., discriminative feature) is selected and the others are ignored. Both feature fusion and selection methods are conducted in a cell-specific fashion, instead of on the whole image. Therefore, the strengths of multiple features can be fully explored on the cell-level.

This framework is also able to accommodate new samples efficiently. This online updating scheme can be achieved by storing not only the weights but also the number of cells in each category. Given new samples, we can update the cell number in their mapped hash entries, re-calculate and update the weights based on such information. The computational overhead is negligible. To summarize, the whole framework includes cell segmentation, hashing, and retrieval. The probability scores are assigned to each hash entry, and they are aggregated within the whole image for the final classification. This process is computationally efficient, with small overhead in the aggregation of probabilities. Benefited from the thorough analysis of multiple cues from each individual cell, this framework can achieve promising accuracy without sacrificing the efficiency.

3 Experiments

In this section, we conduct extensive experiments to evaluate our weighted hashing with multiple features for cell-level analysis. Our dataset is collected from the Cancer Genome Atlas (TCGA) [28], including 57 adenocarcinoma and 55 squamous carcinoma. 10 patches with 1712×952 resolution, i.e., region-of-interests (ROIs), are cropped from each whole slide scanned pathology specimens, by consulting with certified pathologists. Generally, the ROIs mainly consist of cancer cells. The lymphocytes regions which have different visual patterns than the representative tumor regions are avoided. All the data are prepared and labeled based on the independent confirmation of the pathologists. There are around half-million cells segmented for large-scale image retrieval, including nearly 20 K adenocarcinoma cells and 30 K squamous carcinoma cells. We evaluate the efficacy of our proposed framework in terms of the classification accuracy and computational efficiency. The evaluations are conducted on a 3.40 GHz CPU with 4 cores and 16 G RAM, in MATLAB and `C++` implementation.

Table 1. We report the quantitative comparisons of the classification accuracy and efficiency, based on the mean value, standard deviation and running time. We compare with several methods which have been used for histopathological image analysis, including kNN [29], SVM [30] and KSH [20,31], using GIST [25] and HOG [26] features to represent cells' texture.

	Adeno	Squamous	Mean	Variance	Time(s)
kNN-GIST	0.567	0.933	0.750	0.076	∼2600
kNN-HOG	0.354	0.820	0.587	0.063	∼2600
SVM-GIST	0.925	0.533	0.729	0.072	∼50
SVM-HOG	0.775	0.583	0.679	0.094	∼50
KSH-GIST	0.925	0.658	0.792	0.081	1.22
KSH-HOG	0.757	0.748	0.753	0.082	1.22
Weight-GIST	0.833	0.875	0.854	0.052	1.70
Weight-HOG	0.818	0.793	0.806	0.065	1.70
Feature fusion	0.895	0.903	**0.899**	0.064	3.45
Feature selection	0.903	0.903	**0.903**	0.062	3.45

During the evaluation, 25 % patients are randomly selected as the testing data, and the remaining cases are used as the training. This procedure is repeated tens of times to get the mean and standard deviation. As shown in Table 1, we compare our algorithm with several methods that have been employed for histopathological image classification, including k-nearest neighbor (kNN) method [29] which is usually chosen as the baseline for comparison, Support Vector Machine (SVM) [30] that uses supervised information to improve the accuracy, and KSH [20,31] that is used as our baseline to generate binary codes.

For fair comparison, same features (GIST [25] and HOG [26]) are used for all compared methods, and their parameters and kernel selections are optimized by cross-validation. In general, these compared methods do not achieve very accurate results, with 75.0 % and 58.7 % for kNN using two features, 72.9 % and 67.9 % for SVM and 79.2 % and 75.3 % for KSH, even though they thoroughly analyze all segmented cells, same as our algorithm. The main reason is twofold. First, the cell images have high intra-class but low inter-class variations, and the number of two classes is not balanced. Second, same as most segmentation methods, ours is not perfectly accurate, especially for cell images with noise. These inaccurate segmentation results can adversely affect the classification accuracy. Supervised information used in SVM and KSH can alleviate this problem and improve the accuracy, while the results are still not promising.

Using our probabilistic-based weighting scheme, the accuracy is improved to 85.4 % for GIST and 80.6 % for HOG, around 6 % better than the baseline hashing method. The reason is that these weights emphasize certain hash entries that are more discriminative and have more "evidence" than others (i.e., more cells are mapped into that entry), alleviating the issue of high intra-class but low inter-class variations. Furthermore, our weighting scheme reduces the importance of unreliable features, most of which are extracted from inaccurate segmentations. Therefore, it ensures the robustness of the classification module, making it less sensitive to the segmentation accuracy or image noise. In fact, this not only benefits the classification accuracy, but also is compatible with the paradigm of cell-level analysis, given the fact that most existing cell segmentation methods are still not perfect.

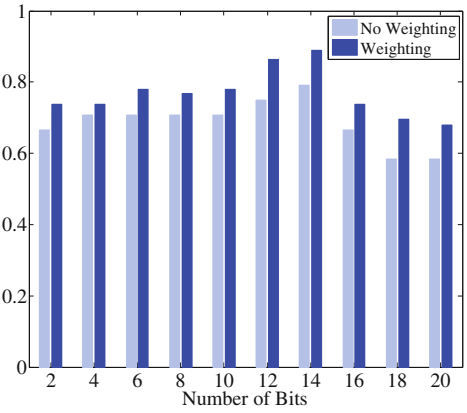

Fig. 3. Classification accuracy of KSH [20] (no weighting) and the weighted hashing applied for five rounds, with different number of hashing bits (2 to 20).

This weighted hashing framework has one important parameter, i.e., the number of hash bits. In our experiments, we have used 12 bits for classification, indicating 4096 hash entries. Theoretically, using one bit is already sufficient for binary classification, i.e., differentiation of two types of cells. However, as shown in Fig. 2, some hash entries may not be reliable and have to be pruned, due to image noise and inaccurate segmentations. Therefore, it is necessary to use many hash entries, which also enable multi-label classification. On the other hand, it is also preferred to have enough samples mapped into each hash entry, so the weight W_H^S can be effective and benefit the classification accuracy. Therefore, the number of hash bits should not be very large either. For example, using

20 hash bits can result in one million hash entries, sufficiently representing half million cells in our dataset. In addition, using a large number of hash bits (e.g., 64 bits) may reduce the computational and memory efficiency, since hash table is no longer an option owing to the memory constraint. Therefore, we have chosen 12 bits for this task, mapping half million cells to 4096 hash values and hence ensuring sound accuracy of classification without sacrificing the efficiency. This is also demonstrated by our experiments shown in Fig. 3. Note that this parameter is not that sensitive to different values, i.e., good accuracy in a certain range of values. This is critical to an automatic framework for histopathological image analysis, since tuning sensitive parameters is infeasible when conducting this large-scale and cell-level analysis. Furthermore, Fig. 3 also shows that our weighting scheme consistently improves the hashing method for classification accuracy, when using different number of hash bits.

Our feature fusion and selection schemes further improve the accuracy to around 90 %, about 11 % to 15 % higher than the KSH with single feature. Utilizing the probability scores from the weighting stage, we can naturally integrate strengths of different features. Particularly, "Feature Fusion" combines all features, and "Feature Selection" selects the best one. Despite the simplicity of these schemes, they achieve promising results, i.e., both strategies can improve the individual feature by a certain margin. Therefore, using multiple cues of cells is essential for fine-grained examination of histopathological images. Note that our fusion schemes can certainly handle more than two features, although we just employ GIST and HOG in this experiment. Other morphological features of cells may also benefit the classification accuracy, and will be investigated in the future. Furthermore, these schemes have no parameter to tune, avoiding overfitting problems that may happen for many learning-based fusion methods.

Table 1 also compares the computational efficiency of these methods, i.e., testing time. Hashing methods are always efficient, since they compress each feature into 12 bits, allowing constant time access using a hash table. Therefore, KSH achieves 1–2 s classification time, much faster than kNN and SVM. Both weighting and fusion schemes have computational overhead (i.e., 0.5 s), while it is negligible in practice. In general, the classification stage is very efficient, and can be used for large-scale and cell-level analysis. However, the segmentation and preprocessing can take tens of seconds, which are the bottleneck for real-time analysis. Currently, we have around one thousand images with half million cells. We expect to apply it on much larger databases (e.g., hundreds of millions of cells) or whole slide images in the future. In this case, parallel computing may be necessary to ensure the computational efficiency for both preprocessing and classification. Fortunately, our framework for cell-level analysis can be parallelled straightforwardly. For example, the whole slide image can be divided as multiple patches, and each patch can be processed by one node of the cluster for cell segmentation and classification independently. In general, the computational efficiency of our framework is very promising and has the potential to handle large-scale databases.

4 Conclusions

In this paper, we proposed an efficient framework for cell-level analysis of histopathological images, by conducting CBIR in a large amount of cell images. This large-scale retrieval is based on weighted hashing with multiple features, which is able to analyze multiple cues of cells and model them in hash entries. We applied this framework on the differentiation of two types of lung cancers, the adenocarcinoma and squamous carcinoma, and achieved promising accuracy and efficiency. In the future, we plan to apply our framework on larger databases and whole slides images, and investigate the correlation of database sizes and the classification accuracy. We also plan to evaluate our method on other use cases of histopathological image analysis.

References

1. Comaniciu, D., Meer, P., Foran, D.J.: Image-guided decision support system for pathology. Mach. Vis. Appl. **11**(4), 213–224 (1999)
2. Müller, H., Geissbühler, A., Ruch, P.: ImageCLEF 2004: combining image and multi-lingual search for medical image retrieval. In: Peters, C., Clough, P., Gonzalo, J., Jones, G.J.F., Kluck, M., Magnini, B. (eds.) CLEF 2004. LNCS, vol. 3491, pp. 718–727. Springer, Heidelberg (2005)
3. Syeda-Mahmood, T., Turaga, P., Beymer, D., Wang, F., Amir, A., Greenspan, H., Pohl, K.: Shape-based similarity retrieval of doppler images for clinical decision support. In: CVPR, pp. 855–862. IEEE (2010)
4. Foran, D.J., Yang, L., et al.: Imageminer: a software system for comparative analysis of tissue microarrays using content-based image retrieval, high-performance computing, and grid technology. JAMIA **18**(4), 403–415 (2011)
5. Dy, J.G., Brodley, C.E., Kak, A., Broderick, L.S., Aisen, A.M.: Unsupervised feature selection applied to content-based retrieval of lung images. TPAMI **25**(3), 373–378 (2003)
6. El-Naqa, I., Yang, Y., Galatsanos, N.P., Nishikawa, R.M., Wernick, M.N.: A similarity learning approach to content-based image retrieval: application to digital mammography. TMI **23**(10), 1233–1244 (2004)
7. Greenspan, H., Pinhas, A.T.: Medical image categorization and retrieval for PACS using the GMM-KL framework. TITB **11**(2), 190–202 (2007)
8. Langs, G., Hanbury, A., Menze, B., Müller, H.: VISCERAL: towards large data in medical imaging — challenges and directions. In: Greenspan, H., Müller, H., Syeda-Mahmood, T. (eds.) MCBR-CDS 2012. LNCS, vol. 7723, pp. 92–98. Springer, Heidelberg (2013)
9. Siegel, R., Naishadham, D., Jemal, A.: Cancer statistics, 2013. CAJC **63**(1), 11–30 (2013)
10. Freeman, D.L.: Harrison's principles of internal medicine. JAMA **286**(8), 506 (2001)
11. Kayser, G., Riede, U., Werner, M., Hufnagl, P., Kayser, K.: Towards an automated morphological classification of histological images of common lung carcinomas. Elec. J. Pathol. Histol. **8**, 022–03 (2002)

12. Thunnissen, F., Diegenbach, P., Van Hattum, A., Tolboom, J., van der Sluis, D., Schaafsma, W., Houthoff, H., Baak, J.R.: Further evaluation of quantitative nuclear image features for classification of lung carcinomas. Pathol. Res. Pract. **188**(4), 531–535 (1992)
13. Mijović, Ž., Mihailović, D., Kostov, M.: Discriminant analysis of nuclear image variables in lung carcinoma. Facta Univ. Ser. Med. Biol. **15**(1), 28–32 (2008)
14. Edwards, S., Roberts, C., McKean, M., Cockburn, J., Jeffrey, R., Kerr, K.: Preoperative histological classification of primary lung cancer: accuracy of diagnosis and use of the non-small cell category. Am. J. Clin. Path. **53**(7), 537–540 (2000)
15. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. IJCV **60**(2), 91–110 (2004)
16. Zhang, X., Yang, L., Liu, W., Su, H., Zhang, S.: Mining histopathological images via composite hashing and online learning. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 479–486. Springer, Heidelberg (2014)
17. Datar, M., Immorlica, N., Indyk, P., Mirrokni, V.S.: Locality-sensitive hashing scheme based on p-stable distributions. In: SoCG, pp. 253–262. ACM (2004)
18. Kulis, B., Grauman, K.: Kernelized locality-sensitive hashing for scalable image search. In: CVPR (2009)
19. Andoni, A., Indyk, P.: Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In: FOCS, Berkeley, CA, 21–24 October 2006
20. Liu, W., Wang, J., Ji, R., Jiang, Y.G., Chang, S.F.: Supervised hashing with kernels. In: CVPR, pp. 2074–2081 (2012)
21. Zhang, X., Su, H., Yang, L., Zhang, S.: Fine-grained histopathological image analysis via robust segmentation and large-scale retrieval. In: CVPR. IEEE (2015)
22. Xing, F., Su, H., Neltner, J., Yang, L.: Automatic ki-67 counting using robust cell detection and online dictionary learning. TBME **61**(3), 859–870 (2014)
23. Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al.: Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. **7**(10), R100 (2006)
24. Arteta, C., Lempitsky, V., Noble, J.A., Zisserman, A.: Learning to detect cells using non-overlapping extremal regions. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 348–356. Springer, Heidelberg (2012)
25. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. IJCV **42**(3), 145–175 (2001)
26. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR, vol. 1, pp. 886–893 (2005)
27. Weiss, Y., Torralba, A., Fergus, R.: Spectral hashing. In: NIPS (2008)
28. National Cancer Institute: The cancer genome atlas retrieved from https://tcga-data.nci.nih.gov (2013)
29. Tabesh, A., Teverovskiy, M., Pang, H.Y., Kumar, V.P., Verbel, D., Kotsianti, A., Saidi, O.: Multifeature prostate cancer diagnosis and gleason grading of histological images. TMI **26**(10), 1366–1378 (2007)
30. Doyle, S., Agner, S., Madabhushi, A., Feldman, M., Tomaszewski, J.: Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In: ISBI, pp. 496–499 (2008)
31. Zhang, X., Liu, W., Dundar, M., Badve, S., Zhang, S.: Towards large-scale histopathological image analysis: hashing-based image retrieval. TMI **34**(2), 496–506 (2015)