

Recognizing Magnification Levels in Microscopic Snapshots

Manit Zaveri¹, Shivam Kalra^{1,2}, Morteza Babaie¹, Sultaan Shah², Savvas Damskinos²,
Hany Kashani¹, H.R. Tizhoosh^{1,2}

¹*Kimia Lab, University of Waterloo, Canada, <https://kimialab.uwaterloo.ca/>*

²*Huron Digital Pathology, St. Jacobs, ON, Canada*

Abstract—Recent advances in digital imaging has transformed computer vision and machine learning to new tools for analyzing pathology images. This trend could automate some of the tasks in the diagnostic pathology and elevate the pathologist workload. The final step of any cancer diagnosis procedure is performed by the expert pathologist. These experts use microscopes with high level of optical magnification to observe minute characteristics of the tissue acquired through biopsy and fixed on glass slides. Switching between different magnifications, and finding the magnification level at which they identify the presence or absence of malignant tissues is important. As the majority of pathologists still use light microscopy, compared to digital scanners, in many instance a mounted camera on the microscope is used to capture snapshots from significant field-of-views. Repositories of such snapshots usually do not contain the magnification information. In this paper, we extract deep features of the images available on TCGA dataset with known magnification to train a classifier for magnification recognition. We compared the results with LBP, a well-known handcrafted feature extraction method. The proposed approach achieved a mean accuracy of 96% when a multi-layer perceptron was trained as a classifier.

Index Terms—Microscope, pathology, magnification, deep learning

I. INTRODUCTION

A biopsy followed by specimen preparation in laboratory and a subsequent microscopic examination by a trained pathologist is necessary for a definitive diagnosis of any type of cancer and many other diseases. Hematoxylin and eosin (H&E) staining is applied to thin cuts of the biopsy sample to visualize the structural patterns and any distortion of the tissue. To differentiate benign from malignant cells and to extract the distinctive cell features, the pathologist generally observes the tissue at several magnification levels to gain a more comprehensive understanding of the specimen. The size of the area under observation decreases with increase in magnification, allowing experts to view the enlarged tissue and observe the minute characteristics relevant for diagnosis. Some microscopes are equipped with a mounted camera that is used to capture snapshots from the glass slide. Pathologists create snapshots of the tissues and save them for future reference in reports or for research purposes. These snapshots usually miss some crucial information such as the magnification level which is required for many classification tasks. Therefore, the use of these snapshots for future research is quite limited particularly if there are large and

diverse repositories of such snapshots of different organs and diseases. This limitation is the main motivation for this work [1], [2].

The trade-off between the superior performance and time complexity in computer vision research is mainly noticeable when it comes to feature extraction. There are several algorithms for feature extraction like SIFT, HOG and LBP which are based on handcrafted design. In contrast, there are also deep learning methods using convolutional neural networks (CNN) like VGG, ResNet and Inception which are based on learning from data. The modern AI algorithms have been intensively investigated recently [3]. Deep networks can learn distinctive image features, while traditional algorithms implement a series of functions on the image to extract important characteristics. These features are being deployed for various classification and content-based image retrieval tasks [4]. Both techniques, handcrafted and learning-based, have proven to perform well for different applications although deep features are reported to generally be much more expressive [5], [6]. Recent publications have investigated this problem to be generic in histopathology domain. Several approaches for classification of malignancy in breast cancer images, for instance, have been performed using various techniques. Bayramoglu et al. [7] utilized deep features while Gupta et al. [8] used color texture features and evaluated the influence of magnification on classification model to identify malignancy. Otalora et al. [6] focused only on magnification level and implemented a CNN-based regression to find the magnification level using multiple open access datasets.

This paper is organized as follows: The feature extractors and classification algorithms are briefly discussed in section II. The dataset is introduced in section III. section IV explains the methodology used in the study. Experiments and results are discussed in section V. Finally section VI concludes the paper.

II. BACKGROUND

A. Feature Extraction

1) *Local Binary Pattern (LBP)*: LBP is a conventional feature extraction algorithm. Ojala and Pietikainen's research in multi-resolution approach [9] proved LBP's useful application in text identification [10] and histopathology classification [11], [12]. A combined approach of LBP descriptor

with histogram of Oriented Gradients (HOG) descriptor was used by Wang [13] to improve the detection performance.

There are many LBP variations available which can address the effect of rotational invariance uniformity on neighborhood pixels [14]. A brief description of most recent applications of LBP are discussed in a survey [15].

2) *Pre-Trained Deep Nets*: Pre-trained deep networks can be employed to transfer knowledge from one domain to another domain. For instance, DenseNet121 is a deep convolutional neural network proposed by Huang et al. [16] that has been applied to many different problems. Recent publications emphasize the capabilities of artificial neural networks and their performance in several classification and search-based tasks [17], [18].

B. Classifiers

There is a large body of literature on different algorithms for data classification. We restrict ourselves to the following two approaches mainly because of our previous experience with these methods.

1) *Artificial Neural Networks (ANNs)*: ANNs are classifiers with the capability to simulate the function of the human brain at a very small scale and for a given task. They are commonly used as robust classifiers in machine learning when a large body of labelled data is available for training. The network mainly consists of three primary layers: the first layer represents input neurons; the last layer represents output neurons; and hidden layers consist of a series of weighted layers which minimize the error between actual output and predicted output. It is difficult to interpret the trajectory of ANNs toward output to understand the rational behind the decision.

When we talk about ANNs we generally mean shallow networks (less than 5 layers), in contrast to convolutional neural networks like DenseNet, ResNet and VGG-19 that are deep networks with many hidden layers.

2) *Random Forests (RF)*: Random Forest (RF) classifiers are multi-way decision trees with some randomization used to grow each tree as a potential solution. The leaf nodes represent the posterior distribution of each image class. Internal nodes contain a test split based on the maximum information gained from the feature subspace. Bosch et al. implemented image classification using random forests beating state-of-the-art results [19].

III. DATASET

We used publicly available data from The Cancer Genome Atlas (TCGA). An important characteristic of this dataset is the available objective power of the whole slide images, which represents different magnification levels. Other publicly available datasets like PMC do not contain magnification information, making them unsuitable for this study. Each image was stored at various magnification levels and in a pyramid structure. The subset from the original dataset was created using an indexing algorithm, which would randomly index one whole slide image (WSI) at a time. If the objective

power(i.e., the magnification of base layer) of 40x or 20x was absent, we discarded that WSI. In total, we gathered 29,596 WSIs for creating our magnification specific dataset. For these WSI files we randomly selected the coordinates of 5 points, read the image at respective magnifications from that specific coordinates and took a snapshot at each point. This yielded us the total of 693,518 patches, consisting of 147,477 patches at 2.5x, 5x, 10x and 20x and 103,611 at 40x. A region of a sample snapshot at different magnification levels is represented in Figure 1.

IV. METHODOLOGY

For classifying the magnification, two models are trained based on different features, one CNN-based approach using DenseNet121 and one conventional algorithm, namely Local Binary Patterns (LBP). The vectors obtained by each feature extractor is considered as an input for classification models. These features are independently fed into ANNs and RF to train them for magnification recognition. The performance of classification is evaluated using a 5-fold cross-validation with 80%-20% split for training and testing, respectively. The accuracy provides a performance index for correct classification, kappa score provides empirical probability of agreement associated to each label and the F1-score provides the arithmetic mean of precision and recall [6], [20].

A. Feature Extraction with LBP

The feature vectors obtained through the application of LBP are conserved in histograms. We used rotationally invariant LBP with three parameter settings for radius r and neighbors n . For radius $r = 1$ and neighbours $n = 8$ we got features with 10 dimensions (bins). Radius $r = 2$ and neighbours $n = 16$ resulted in features with 16 bins, and features with 26 bins were the result of radius $r = 3$ and neighbours $n = 24$.

B. Deep Features using DenseNet121

Using pre-trained features for histopathology images has been the focus of attention in recent literature [21]. Image features were extracted from the last average pooling layer in DenseNet121. Before passing through the feature extractor, we pre-processed all images by making the mean zero. The input tensor flows through convolutional feature layer to capture low-level information consisting of shapes from histopathology images. This was followed by a series of 4 dense blocks. These dense blocks have 6, 12, 16 and 24 layers of batch normalization, reLu and 2D-convolutional layers stacked alternatively. The feature vector from the last average pooling layer contained 1,024 dimensions which were passed to the classifier to represent the corresponding image magnification.

In total, we had 4 individual feature sets for our dataset. One may argue that using ANNs is considered a fine-tuning technique, but for all our experiments we treated classification methods separate from feature extraction models.

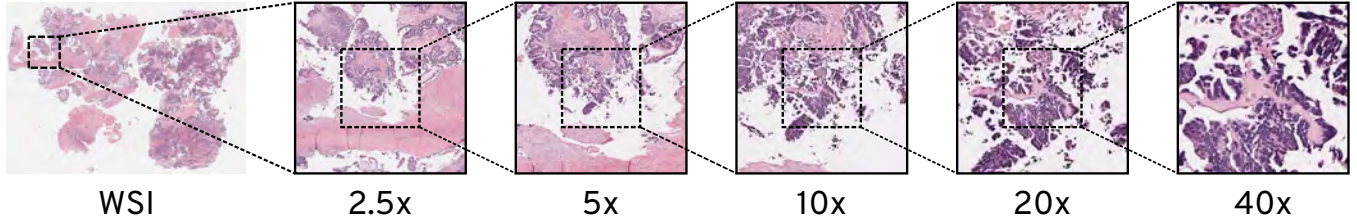


Fig. 1: Illustration of patches from different magnification levels extracted from a WSI. All the patches are centered around the same coordinates (images re-scaled for convenient visualization).

V. EXPERIMENTS AND RESULTS

A. ANNs

The shallow classifier that used deep features excelled in performance. The accuracy was observed at 96.1% which was achieved using 2 hidden layers and one fully connected layer between input and output layers. The hidden layers had 512 neurons with batch normalization and a dropout rate of 0.5 followed by other fully connected layer of 256 neurons and ReLU activation. This finally was connected to output layer of 5 neurons which was activated using softmax. For every ANN we matched the input layer neurons with the dimensions of the input feature vector. For LBP features, ANNs achieved an accuracy of 68.4%, kappa score of 0.603 and f1-score of 0.676 for one split. From Table I we can easily identify the higher performance of deep features compared to handcrafted LBP features. However, there is a trade-off between real-time high performance and time complexity plus the computational power utilized for extracting deep features. The confusion matrix in Figure 2 shows that images at a magnification level of 5x are mostly confused with class of images at 2.5x and 10x magnification. The individual class accuracy was 91% for 5x images. While images at 2.5x, 20x and 40x were separated excellently with misclassification rate of less than 1%.

B. Random Forests

We searched for best parameters using random searchCV and grid searchCV from Scikit Learn library [20]. We used 1000 estimators for training the classifier, with a maximum depth of 50 splits for individual trees. The final label assignment was determined through majority voting from each tree. The reported accuracy for classification using deep features was 96% with kappa score of 0.95 and F1-score of 0.96 while using LBP, the best accuracy obtained by RF classification model was 68.4% with kappa score of 0.60 and F1-score of 0.67. All results using RF classifier are listed in Table I. Comparing the confusion matrices in Figure 2 a and b, we can see that classifiers using deep features are more prone to confusing 5x images compared to shallow classifiers. While classes 20x and 40x are also more prone to misclassification when RFs are implemented.

VI. CONCLUSIONS

In this paper, two feature extraction methods were used to train two classifiers for discrete classification of magni-

2.5x	0.98	0.016	0.00044	0.00024	0.0001
5x	0.054	0.91	0.03	0.00074	0.0002
10x	0.0044	0.016	0.97	0.011	0.0003
20x	0.00051	0.00054	0.004	0.99	0.0076
40x	0.00011	0.00032	0.00018	0.0046	0.99
	2.5x	5x	10x	20x	40x

(a) ANN Classifier

2.5x	0.96	0.043	0.00085	0.00017	0.00014
5x	0.13	0.82	0.051	0.00068	0.00024
10x	0.011	0.053	0.92	0.017	0.00051
20x	0.0031	0.0044	0.013	0.97	0.012
40x	0.0014	0.0027	0.00058	0.048	0.95
	2.5x	5x	10x	20x	40x

(b) RF Classifier

Fig. 2: Confusion matrix of the two classifiers.

fication levels in microscopic images. We used the TCGA dataset to construct a large dataset for training and testing. The feature extraction was done from raw images without any pre-processing or segmentation. A shallow classifier with deep features obtained from DenseNet121 had the best performance in terms of accuracy (96%), kappa (0.95) and

Features	Folds	Random Forests			Shallow Classifier		
		Acc	Kappa	F1-Score	Acc	Kappa	F1-Score
DF	Fold1	0.881	0.849	0.881	0.935	0.917	0.935
	Fold2	0.879	0.849	0.879	0.96	0.95	0.96
	Fold3	0.916	0.849	0.915	0.971	0.963	0.971
	Fold4	0.924	0.905	0.924	0.972	0.966	0.973
	Fold5	0.921	0.901	0.921	0.969	0.962	0.969
	All Folds	0.904 ± 0.01	0.879 ± 0.025	0.904 ± 0.01	0.961 ± 0.01	0.951 ± 0.01	0.961 ± 0.01
LBP ₁	Fold1	0.520	0.392	0.509	0.514	0.382	0.501
	Fold2	0.612	0.513	0.604	0.601	0.499	0.588
	Fold3	0.649	0.560	0.643	0.598	0.496	0.583
	Fold4	0.684	0.605	0.679	0.646	0.557	0.631
	Fold5	0.684	0.605	0.678	0.655	0.568	0.646
	All Folds	0.629 ± 0.06	0.535 ± 0.07	0.622 ± 0.06	0.602 ± 0.05	0.50 ± 0.06	0.589 ± 0.05
LBP ₂	Fold1	0.516	0.389	0.508	0.471	0.339	0.458
	Fold2	0.582	0.476	0.571	0.565	0.453	0.556
	Fold3	0.626	0.532	0.618	0.587	0.480	0.578
	Fold4	0.656	0.569	0.649	0.613	0.516	0.607
	Fold5	0.654	0.567	0.647	0.628	0.534	0.618
	All Folds	0.606 ± 0.05	0.506 ± 0.06	0.598 ± 0.05	0.572 ± 0.05	0.464 ± 0.06	0.563 ± 0.05
LBP ₃	Fold1	0.546	0.425	0.536	0.507	0.375	0.493
	Fold2	0.609	0.510	0.599	0.582	0.476	0.574
	Fold3	0.649	0.561	0.643	0.634	0.542	0.622
	Fold4	0.682	0.603	0.676	0.668	0.584	0.656
	Fold5	0.676	0.595	0.671	0.640	0.551	0.631
	All Folds	0.632 ± 0.05	0.538 ± 0.06	0.624 ± 0.05	0.606 ± 0.05	0.505 ± 0.07	0.595 ± 0.05

TABLE I: Results for deep features (DF) and LBP with 3 parameter settings.

F1-score (0.96). The score achieved by our model outperformed existing state-of-the-art regression method with 11% improved kappa score. Because the TCGA dataset consists of various categories of images and several organs, we assume the model offers good generalization for magnification recognition providing a basic pre-processing step for all digital pathology image processing to attain better performance.

Microscopic snapshots may have a different image quality than patches from whole slide images. Investigating this difference will be subject of future work.

REFERENCES

- [1] P. P. Clark, D. E. Miller, D. A. Mulford, J. C. Ostrowski, Microscope camera, uS Patent 5,315,344 (May 24 1994).
- [2] J. Winterot, J. Knoblich, T. Kaufhold, H. Tielebier, G. Osten, Microscope camera, uS Patent App. 10/993,660 (May 26 2005).
- [3] H. R. Tizhoosh, L. Pantanowitz, Artificial intelligence and digital pathology: challenges and opportunities, *Journal of pathology informatics* 9 (2018).
- [4] S. Kalra, C. Choi, S. Shah, L. Pantanowitz, H. Tizhoosh, Yottixel—an image search engine for large archives of histopathology whole slide images, *arXiv preprint arXiv:1911.08748* (2019).
- [5] M. Dinesh Kumar, M. Babaie, S. Zhu, S. Kalra, H. R. Tizhoosh, A comparative study of cnn, bovw and lbp for classification of histopathological images, in: 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017, pp. 1–7 (Nov 2017).
- [6] S. Otálora, M. Atzori, V. Andrearczyk, H. Müller, Image magnification regression using densenet for exploiting histopathology open access content, in: *Computational pathology and ophthalmic medical image analysis*, Springer, 2018, pp. 148–155 (2018).
- [7] N. Bayramoglu, J. Kannala, J. Heikkilä, Deep learning for magnification independent breast cancer histopathology image classification, in: 2016 23rd International conference on pattern recognition (ICPR), IEEE, 2016, pp. 2440–2445 (2016).
- [8] V. Gupta, A. Bhavsar, Breast cancer histopathological image classification: is magnification important?, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 17–24 (2017).
- [9] T. Ojala, M. Pietikainen, T. Maenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on pattern analysis and machine intelligence* 24 (7) (2002) 971–987 (2002).
- [10] I. Jung, I.-S. Oh, Local binary pattern-based features for text identification of web images, in: 2010 20th International Conference on Pattern Recognition, IEEE, 2010, pp. 4320–4323 (2010).
- [11] J. C. Caicedo, A. Cruz, F. A. Gonzalez, Histopathology image classification using bag of features and kernel functions, in: *Conference on Artificial Intelligence in Medicine in Europe*, Springer, 2009, pp. 126–135 (2009).
- [12] H. Erfankhah, M. Yazdi, M. Babaie, H. R. Tizhoosh, Heterogeneity-aware local binary patterns for retrieval of histopathology images, *IEEE Access* 7 (2019) 18354–18367 (2019).
- [13] X. Wang, T. X. Han, S. Yan, An hog-lbp human detector with partial occlusion handling, in: 2009 IEEE 12th international conference on computer vision, IEEE, 2009, pp. 32–39 (2009).
- [14] S. Nigam, R. Singh, A. Misra, Local binary patterns based facial expression recognition for efficient smart applications, in: *Security in Smart Cities: Models, Applications, and Challenges*, Springer, 2019, pp. 297–322 (2019).
- [15] M. Kas, Y. El Merabet, Y. Ruichek, R. Messoussi, Survey on local binary pattern descriptors for face recognition, in: *Proceedings of the New Challenges in Data Sciences: Acts of the Second Conference of the Moroccan Classification Society*, ACM, 2019, p. 5 (2019).
- [16] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: *IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708 (2017).
- [17] C. L. Srinidhi, O. Ciga, A. L. Martel, Deep neural network models for computational histopathology: A survey, *arXiv preprint arXiv:1912.12378* (2019).
- [18] A. Gordo, J. Almazán, J. Revaud, D. Larlus, Deep image retrieval: Learning global representations for image search, in: *European conference on computer vision*, Springer, 2016, pp. 241–257 (2016).
- [19] A. Bosch, A. Zisserman, X. Munoz, Image classification using random forests and ferns, in: 2007 IEEE 11th international conference on computer vision, IEEE, 2007, pp. 1–8 (2007).
- [20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830 (2011).
- [21] B. Kieffer, M. Babaie, S. Kalra, H. R. Tizhoosh, Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks, in: 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), IEEE, 2017, pp. 1–6 (2017).