Deep Learning for Magnification Independent Breast Cancer Histopathology Image Classification

Neslihan Bayramoglu
Center for Machine Vision
and Signal Analysis
University of Oulu, Finland
Email: nyalcinb@ee.oulu.fi

Juho Kannala
Aalto University
Department of Computer Science
Finland
Email: juho.kannala@aalto.fi

Janne Heikkilä
Center for Machine Vision
and Signal Analysis
University of Oulu, Finland
Email: jth@ee.oulu.fi

Abstract—Microscopic analysis of breast tissues is necessary for a definitive diagnosis of breast cancer which is the most common cancer among women. Pathology examination requires time consuming scanning through tissue images under different magnification levels to find clinical assessment clues to produce correct diagnoses. Advances in digital imaging techniques offers assessment of pathology images using computer vision and machine learning methods which could automate some of the tasks in the diagnostic pathology workflow. Such automation could be beneficial to obtain fast and precise quantification, reduce observer variability, and increase objectivity.

In this work, we propose to classify breast cancer histopathology images independent of their magnifications using convolutional neural networks (CNNs). We propose two different architectures; single task CNN is used to predict malignancy and multi-task CNN is used to predict both malignancy and image magnification level simultaneously. Evaluations and comparisons with previous results are carried out on BreaKHis dataset. Experimental results show that our magnification independent CNN approach improved the performance of magnification specific model. Our results in this limited set of training data are comparable with previous state-of-the-art results obtained by hand-crafted features. However, unlike previous methods, our approach has potential to directly benefit from additional training data, and such additional data could be captured with same or different magnification levels than previous data.

I. INTRODUCTION

Cancer is still one of the top leading cause of death worldwide [1]. And breast cancer is the most common cancer among women [1], [2]. A biopsy followed by microscopic analysis of breast tissue is necessary for a definitive diagnosis of breast cancer [2]. Firstly, thin sections are cut from biopsy material and then stained generally with hematoxylin and eosin (H&E). Hematoxylin highlights nuclei by binding DNA and eosin highlights other structures by binding proteins [3]. Visualization is the next step after staining. Finally, pathologists evaluate tissue biopsies under microscope under various magnifications or on the digital image with no appreciable difference in diagnoses [4]. Pathologist examines tissue patterns, textures, and morphology to find clinical assessment clues to produce correct diagnoses. They may need to pan, focus, zoom, and scan through the entire image at high magnification which may be very time consuming [5]. In addition to timely and costly process, such analysis can be hampered by pathologists different interpretations [3].
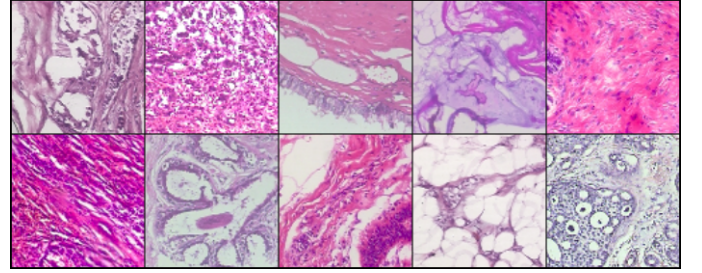

Fig. 1. Sample breast cancer histopathology images from BreaKHis database. They were all acquired at a magnification factor of $100\times$.

The advances in digital imaging techniques enables digitizing pathology images at microscopic resolution [5]. This development offers the histological assessment of hematoxylin and eosin stained sections by computer vision and machine learning methods. These methods could automate some of the tasks in the diagnostic pathology workflow which could be used to reduce observer variability and increase objectivity. In addition, fast and precise quantification could enhance the healthcare quality. However, this requires robust image analysis algorithms which is still far from clinical acceptance [4]. Nonetheless, there has been a progress in the development of image analysis algorithms for histopathological image assessment. For more information on automated image analysis methods in histology, we refer reader to [3], [4], and [6].

Appearance variability of hematoxylin and eosin stained sections is one of the major challenges in breast cancer histopathology image analysis [3] (see Figure 1). These variations are due to variability among people, differences in protocols between labs, fixation, specimen orientation in the block, human skills in tissue preparation, microscopy main-
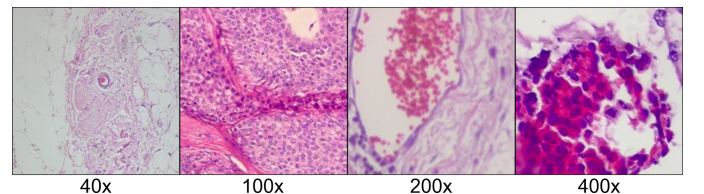

Fig. 2. A malignant breast tumor acquired from a single slide seen in different magnification factors: $40\times, 100\times, 200\times,$ and $400\times$.

tenance, and color variation due to differences in staining procedures [4]. However, the main challenge for image analysis researchers is accessing relevant images and databases. Development of more efficient data analysis methods to be applied in pathological diagnosis could be facilitated by publicly available large annotated datasets which is currently lacking. Such datasets and benchmarks enable validating and comparing algorithms to develop more robust ones.

Recently, a dataset of breast cancer histopathology images (BreaKHis) is released for this purpose [7]. It was collected from 82 patients using different magnifying factors ($40\times, 100\times, 200\times, 400\times$) (Figure 2). The database contains 2,480 benign and 5,429 malignant images. Moreover, the study in [7] provides the classification performance of several hand-crafted textural features as baselines to discriminate between benign and malignant tumors. These include state-of-the-art descriptors such as Local Binary Patterns (LBP) [8], Completed LBP (CLBP) [9], Local Phase Quantization (LPQ) [10], Grey-Level Co-occurrence Matrix (GLCM) [11], Oriented FAST and Rotated BRIEF (ORB) [12], and Threshold Adjacency Statistics (PFTAS) [13].

In this paper, we propose to classify breast cancer histopathology images independent of their magnifications. We present the classification performance of a deep learning method on the BreaKHis dataset in order to provide additional baseline. Moreover, we propose to detect the image magnification level and classify benign and malignant tumors simultaneously by utilizing a similar network architecture which is adapted for multi-task classification. Most importantly, we introduce a learning based approach which could benefit from additional labeled training data in straightforward manner, unlike the approaches based on hand-crafted features.

II. METHOD

Recently, neural networks achieved a great success in object classification [14]. Specifically, Convolutional Neural Network (CNN) based approaches showed significant improvements over state-of-the-art recognition and classification approaches. CNNs are also applied to solve various problems in biomedical image analysis research. For example, Ciresan et. al [15] detect mitosis in breast cancer histology images and won the related contest. It has been also used in cell classification [16], [24], tumor cell detection in blood samples [17] and in segmenting magnetic resonance images of the human brain into anatomical regions [18].

In breast cancer histopatholoy image analysis, convolutional neural networks are used for region of interest detection [19], segmentation [20], and also for mitosis detection [15]. On the other hand, for classification purposes, hand-crafted features are often employed [5], [21], [2], [7]. They include complex preprocessing pipeline including stain normalization, nucleus detection, and region of interest segmentation. This is mainly due to the heterogeneous structure of histopatholoy images. First, there is a wide variety of tissues and the complexity in appearance increases at lower microscopy magnifications. Various multi-cellular structures and diverse backgrounds are
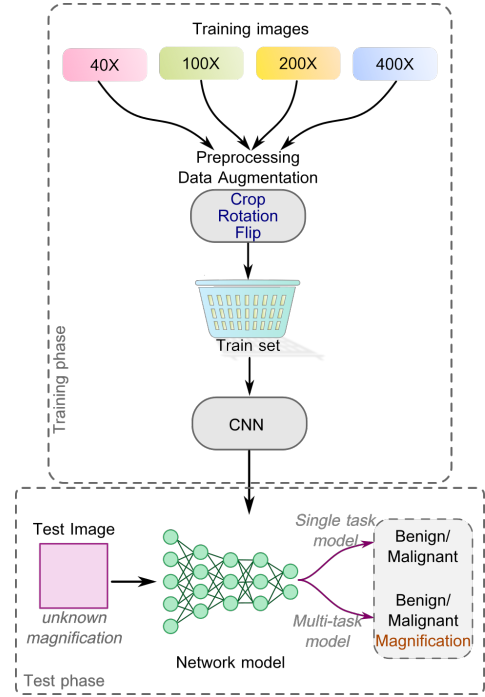

Fig. 3. Schematic presentation of our proposal for classifying breast histology images which is independent from the image magnification factor.

captured in different magnifications. Therefore, in order to obtain distinguishable characteristics, structure specific features are needed. Second, images could contain both benign and malignant regions and therefore, it is challenging to learn discriminative features globally from histology images and therefore, segmentation is needed. When the training data is small, the later issue becomes a challenge also for CNNs.

One way to reduce variability in microscopy images is to utilize images acquired at the same magnification level. Vast majority of previous studies adopt this method and employ single magnification level. Some studies ([7], [22]) utilize multiple magnifications but they use a different classifier for each magnification level. However, such approaches have practical limitations. First, multiple training stages are needed for different magnifications. Second, during test time, the magnification factor of the test image must be known and corresponding model should be used. Such information might not be available all the time. Third, the classification method might perform poorly when test images are acquired at new magnification levels. Because, during training stage, classifiers learn magnification specific features and they could not adapt themselves to unseen image features. Therefore, a classification systems which is intended to be used in a diagnostic practice should handle the diversity in microscopy images and should not depend on the device settings such as microscopy magnification. Our approach has ability to utilize additional training data from various imaging devices with different magnification factors.

Magnification independent classification(single-task CNN): We propose to employ deep learning method to handle the
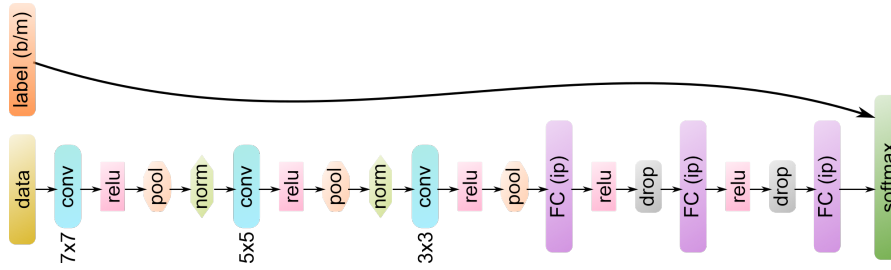
Fig. 4. Schematic presentation of our network architecture for classifying bening and malignant images (single task CNN).

diversity of breast cancer histopathology image appearance. CNN models have high capacities to represent diverse features. We used BreaKHis database [7] in our experiments. BreaKHis provides images at four different magnifications $(40\times, 100\times, 200\times, 400\times)$. They also provide magnification specific baseline results. To train our CNN model, we utilized all available data in the training set of BreaKHis independent of their magnifications. We needed only one training stage for parameter learning and we tested each image using the learned model (Figure 3).

The network architecture used in our experiments is presented in Figure 4. Previously, it was used for age and gender estimation from real-world images [23]. The network contains three convolutional layers, each one followed by a rectified linear operation and a pooling layer. Local response normalization [14] is employed after the first two layers. In the first convolutional layer, 96 filters of size $3 \times 7 \times 7$ pixels are applied to the input. This layer is followed by a rectified linear operator (ReLU), a max-pooling layer taking the maximal value of 3x3 regions with two-pixel strides, and a local response normalization layer [14]. The second convolutional layer contains 256 filters of size 5×5. ReLU, max-pooling and normalization layer is applied again after the second convolutional layer. Finally, in the third convolutional layer, a set of 384 filters of size 3×3 are used. This layer is followed by ReLU and a max-pooling layer. Finally, two fully-connected layers are added after the convolutional layers. Each fully-connected layer contains 512 neurons and each followed by a ReLU and a dropout layer with a dropout ratio of 0.5.

Preprocessing and data augmentation:
Increasing the number of training samples is useful for small training sets in CNN frameworks. We augment the training set by affine transformations. We rotate the images around their centers with angles 90°, 180°, and 270°. Rotated and original images are then flipped and added to the training set. Each image is cropped around its center to obtain a square patch. For non-square images, rotation operations introduce discontinuities at the image borders therefore, images are cropped to include only tissue pixels. Images are then all scaled to a fixed size for speed up.

Multi-task classification:
CNN models build highly non-linear mappings between the input and the output using cascaded convolutional layers. Such complex hierarchical representations are capable of capturing

features from basic to more complicated structures which are then used to predict attributes. Attributes could simply be class labels defined for a task. Different tasks could share common features. Therefore, convolutional neural networks are well-suited for multi-task learning. Based on this, we propose a multi-task framework to jointly learn classifiers for image malignancy and magnification factor. We modified the network architecture to generate two output layers by splitting the last fully connected layer into two branches. During backpropagation, the two gradients are added together in the split layer. The network is shown in Figure 5. After the network splits into two branches:

- The first branch learns the benign/malignant decision. The output is fed into a 2-way softmax, and we minimize the softmax loss which is equivalent to the cross-entropy loss in the 2-class case.
- The second branch learns the magnification factor. Again the output is fed into a 4-way softmax and the softmax loss is minimized.

Multi-task loss:
Each output layer computes a discrete probability distribution by a softmax over the outputs of a fully connected layer. During training phase, the minimization is done over the softmax loss function given by Equation 1.

$$\mathcal{L}(x, y) = -\sum_i y_i \, log \, p_i(x) \tag{1}$$

where $p_i(x)$ is the probability of input x being labeled with l_i and y is the true distribution (i.e. ground truth) where $\sum_i y_i = 1$, with the true class equal to one and the rest are zero.

We define the total cost ($\mathcal{C}$) in the multi-task setting as the weighted sum of the benign/malignancy cost and magnification cost:

$$\mathcal{C} = \omega_{bm}\mathcal{L}_{bm} + \omega_{mag}\mathcal{L}_{mag} \tag{2}$$

where ω_{bm} and ω_{mag} are the cost weights for malignancy and magnification tasks respectively. We used $\omega_{bm} = \omega_{mag} = 0.5$, but different weights might improve the results which is difficult to determine theoretically but it needs to be estimated empirically.

III. EXPERIMENTS AND RESULTS

BreaKHis database provides 7,909 histopathology images divided into benign and malignant tumors that are obtained
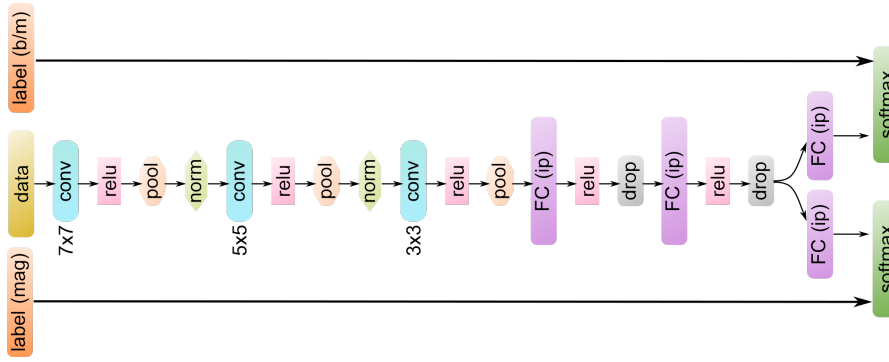
Fig. 5. Schematic presentation of our network architecture for classifying both malignancy and image magnification level (multi-task CNN).

from 82 patients. Images are of 3-channel RGB, 8-bit depth in each channel, and of size 700×460. On the average, 24 images per patient is captured from each slide using the lowest magnification factor ($40\times$). The magnification is then manually increased by the pathologist to $100\times$, $200\times$, and $400\times$ and a similar number of images is captured inside the initial region at each magnification level. Out-of-focus images are then discarded by a final visual inspection. In addition to the images, BreaKHis also provides a testing protocol. The dataset has been divided into a training (~70%) and a testing (~30%) set where the patients used to build the training set are not used for the testing set. The split protocol has been used to obtain 5-folds and results were reported based on the average of five folds.

BreaKHis also presents performances of six state-of-the-art hand-crafted features (Table IV) and four classifiers (1-Nearest Neighbor, Quadratic Linear Analysis, Support Vector Machines, Random Forests) on the database. Method performances were reported at the patient level, and not at the image level. Recognition rate is defined as follows:

$$Recognition\ Rate = \frac{\sum_P Patient\ Score_P}{Total\ number\ of\ patients} \quad (3)$$

where the patient score is defines as:

$$Patient\ Score_P = \frac{N_{rec}}{N_P} \quad (4)$$

where N_P is the number of images of patient P and N_{rec} is the number of correctly classified images.

In our study, we follow the same test protocol and also the same performance measure to make a fair comparison with previous results. In our preprocessing step, images are cropped from their centers to 460×460 sized images. Square sized patches are adopted in order to avoid image discontinues at the image borders in rotated images and also fixed size images are required in CNN frameworks. Training sets are populated by rotations and flipping operations that enlarged the sets by 8 times. In our tests, all the images are resized to 100×100 both in training phase and testing phase for speeding-up purposes. The network is trained using the minibatch stochastic gradient descent with a momentum factor of 0.9. Each iteration operates

on a minibatch of 100 images that are sampled randomly from the training set. BreaKHis is a small database compared to the appearance variation in breast cancer histopathology images therefore, image features were not reflected evenly in the training sets. This leads the optimal iterations and the learning rates vary from fold to fold. We utilized images from one magnification level ($40\times$) in the test set to tune the number of iterations and base learning rates. We applied early stopping when we noticed over-fitting. We observed that when we continue learning, the training set accuracies for folds usually reaches up to 100%. This is an indication that the network has a sufficient capacity to classify breast cancer histopathology images and, at the same time, this is an indication that we have a very limited amount of labeled training data, as often is the case. However, our approach which combines data from all magnification levels to train a single network is able to reach good performance despite the lack of magnification specific training data.

TABLE I
COMPARISON OF BENIGN/MALIGNANT CLASSIFICATION PERFORMANCE FOR MAGNIFICATION SPECIFIC AND MAGNIFICATION INDEPENDENT TRAINING (SINGLE TASK CNN)

	Patient Score (%)	
Fold 1	**Mag. Specific**	**Mag. Independent**
40×	79.40	80.97
100×	78.69	80.92
200×	83.72	83.42
400×	80.83	83.02

In our first experiment in Table I, we compare performances of our CNN framework for both magnification specific and magnification independent training for benign/malignancy classification (single task CNN). In magnification specific experiments, for each magnification level we trained a separate network starting from a random state and tested images according to their magnification factor. For the magnification independent training, we augment the training set with all the training images independent of their magnifications (Figure 3) and trained a single network (Figure 4). Images are then tested using the learned model. Comparison is done for the

first fold. Mixing images from different magnifications in the training stage does not degrade the classification performance. On the contrary, the performance is slightly increased. The increase is due to the expansion of training set with additional data imported from other magnification levels and also due to affine transformations.

Recognition rates using magnification independent training (single task CNN) for the five folds are presented in Table II. For benign/malignant classification, with our single task CNN model, an average recognition rate of 83.25% is achieved with a single training per fold. The differences in recognition rates for different folds are due to uneven distribution of the data in test and training sets. This indicates that there is a need to increase the amount of training data with additional patients in order to better cover variation in image features during training. On the other hand, the average recognition rate which is shown in last column of Table II is similar among all magnification levels. This shows that, proposed network learns balanced set of features from different levels and during test time, the model generates magnification independent predictions which we aimed in the first place.

Next, we tested our multi-task network framework (Figure 5) for predicting both magnification factor of an input image and its malignancy. Table III shows recognition rates for both tasks. An average recognition rate of 82.13% is achieved for benign/malignant classification task and an average recognition rate of 80.10% is achieved for magnification estimation task. Compared to single task CNN, performance of malignancy estimation is not degraded much in multi-task setting. Therefore, multi-task model could be more useful in applications where magnification factor of an input image needs to be estimated and/or where there is insufficient amount of training data from a single magnification level.

In Table IV, we compare our results (single task CNN (*CNN*) and multi-task CNN (*mt_CNN*)) with previous methods. We have selected the best performed classifier results for hand-crafted features. For example, for PFTAS features, best results in $40\times$ and $100\times$ magnifications are achieved with QDA classifier while SVM performed better at factors $200\times$ and $400\times$. PFTAS features are based on a three dimensional convolution operation on thresholded color channels which is similar to CNN in that sense. Therefore, PFTAS and CNN performances are similar (83.33% vs. 83.25%). On the other hand, there is a large gap between results of CNN and other hand-crafted features at all magnification levels. We believe that additional training data would enhance CNN results and therefore, deep learning is more promising as their capacities for classify training sets are shown to be perfect.

IV. Discussions and Conclusion

In this work, we have proposed a general framework based on CNNs for learning breast cancer histopathology image features. The proposed framework is independent from microscopy magnification and faster than previous methods as it requires single training. Speed and magnification independence properties are achieved without sacrificing the state-of-the-art performance. Magnification independent models are scalable, new training images from any magnification level could be utilized and trained models could easily be tuned (fine-tuning) by introducing new samples.

In this work, we have also proposed a multi-task CNN architecture to predict both the image magnification level and its benign/malignancy property simultaneously. The proposed model allows combining image data from many more resolution levels than four discrete magnification levels. In fact, magnification level prediction could be formulated as a regression problem which is not limited to a discrete set of levels. Multi-task prediction requires essentially no additional computation over single-task prediction. Besides, experimental result shows that classification performance does not degrade in multi-task network.

For the future work, stain normalization, deeper architectures, and splitting the network before the last fully-connected layer could be investigated. It would be interesting to observe task-wise early stopping in multi-task architecture. More importantly, additional data with increased number of patients should be introduced. We believe CNNs are more promising in breast cancer histopathology image classification than hand-crafted features and the data is the key issue to obtain more robust models.

References

[1] Torre LA, Siegel RL, Ward EM, Jemal A. "Global cancer incidence and mortality rates and trendsan update." Cancer Epidemiology Biomarkers & Prevention. 2016 Jan 1;25(1):16-27.

[2] Loukas C, Kostopoulos S, Tanoglidi A, Glotsos D, Sfikas C, Cavouras D. "Breast cancer characterization based on image classification of tissue sections visualized under low magnification." Computational and mathematical methods in medicine. 2013 Aug 31;2013.

[3] Veta M, Pluim JP, van Diest PJ, Viergever MA. "Breast cancer histopathology image analysis: A review." Biomedical Engineering, IEEE Transactions on. 2014 May;61(5):1400-11.

[4] McCann MT, Ozolek JA, Castro CA, Parvin B, Kovacevic J. "Automated histology analysis: Opportunities for signal processing." IEEE Signal Processing Magazine. 2015;32(1):78.

[5] M. Peikari, M. J. Gangeh, J. Zubovits, G. Clarke and A. L. Martel, "Triaging Diagnostically Relevant Regions from Pathology Whole Slides of Breast Cancer: A Texture Based Approach," in IEEE Transactions on Medical Imaging, vol. 35, no. 1, pp. 307-315, Jan. 2016.

[6] Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. "Histopathological image analysis: a review. Biomedical Engineering," IEEE Reviews in. 2009;2:147-71.

[7] F. Spanhol; L. Oliveira; C. Petitjean; L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," in IEEE Transactions on Biomedical Engineering, 2015.

[8] Ojala, T., Pietikäinen, M.,Mäenpää, T. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24, no. 7 (2002): 971-987.

[9] Guo, Zhenhua, Lei Zhang, and David Zhang."A completed modeling of local binary pattern operator for texture classification." Image Processing, IEEE Transactions on 19, no. 6 (2010): 1657-1663.

[10] Ojansivu, V., Heikkilä, J."Blur insensitive texture classification using local phase quantization." In Image and signal processing, pp. 236-243. Springer Berlin Heidelberg, 2008.

TABLE II

MAGNIFICATION INDEPENDENT BENIGN/MALIGNANT CLASSIFICATION (SINGLE TASK CNN) PERFORMANCE

	Recognition Rate (based on patient score) (%)					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
40×	80.97	82.51	81.46	84.68	85.80	83.08
100×	80.92	88.30	79.13	83.07	84.44	83.17
200×	83.42	89.19	82.03	83.84	84.66	84.63
400×	83.02	89.00	77.23	79.73	81.52	82.1

TABLE III

MULTI-TASK CNN (MT_CNN) PERFORMANCE. GIVEN A TEST IMAGE, IT'S MAGNIFICATION FACTOR (**MAG.**) AND ITS MALIGNANCY (BENIGN/MALIGNANT) (**B/M**) IS PREDICTED SIMULTANEOUSLY. RECOGNITION RATES ARE BASED ON THE PATIENT SCORE (EQUATION 3 AND 4)

	Recognition Rate (based on patient score) (%)											
	Fold1		Fold2		Fold3		Fold4		Fold5		Average	
	Mag.	B/M	Mag.	B/M	Mag.	B/M	Mag.	B/M	Mag.	B/M	Mag.	B/M
40×	79.43	79.51	84.87	82.43	70.67	78.10	91.10	85.58	84.14	83.77	82.04±7.59	81.87±3.06
100×	87.87	79.49	79.74	90.01	70.49	77.25	71.60	83.61	81.14	86.60	78.16±7.20	83.39±5.17
200×	72.63	81.04	71.72	87.48	75.51	78.00	80.86	83.79	85.73	82.50	77.29±5.91	82.56±3.49
400×	71.34	84.23	84.13	85.41	78.05	75.84	92.80	80.90	88.80	77.06	83.02±8.54	80.69±4.23

TABLE IV

COMPARISON OF CLASSIFICATION PERFORMANCE WITH PREVIOUS RESULTS: LOCAL BINARY PATTERNS (LBP) [8], COMPLETED LBP (CLBP) [9], LOCAL PHASE QUANTIZATION (LPQ) [10], GREY-LEVEL CO-OCCURRENCE MATRIX (GLCM) [11], ORIENTED FAST AND ROTATED BRIEF (ORB) [12], AND THRESHOLD ADJACENCY STATISTICS [13], MULTI-TASK CONVOLUTIONAL NEURAL NETWORKS (MT_CNN, THIS WORK), SINGLE-TASK CONVOLUTIONAL NEURAL NETWORKS (CNN, THIS WORK). RESULTS FOR HAND-CRAFTED FEATURES ARE DIRECTLY TAKEN FROM [7].

	Recognition Rate (based on patient score) (%)				
Method/ Magnification	40×	100×	200×	400×	Average
CLBP [9]	77.4±3.8	76.4±4.5	70.2±3.6	72.8±4.9	**74.20**
GLCM [11]	74.7±1.0	78.6±2.6	83.4±3.3	81.7±3.3	**79.60**
LBP [8]	75.6±2.4	73.2±3.5	72.9±2.3	73.1±5.7	**73.70**
LPQ [10]	73.8±5.0	72.8±5.0	74.3±6.3	73.7±5.7	**73.65**
ORB [12]	74.4±1.7	69.4±0.4	69.6±3.0	67.6±1.2	**70.25**
PFTAS [13]	83.8±2.0	82.1±4.9	85.1±3.1	82.3±3.8	**83.33**
mt_CNN	81.87±3.06	83.39±5.17	82.56±3.49	80.69±4.23	**82.13**
CNN	83.08±2.08	83.17±3.51	84.63±2.72	82.10±4.42	**83.25**

[11] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." Systems, Man and Cybernetics, IEEE Transactions on 6 (1973): 610-621.

[12] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G. "ORB: an efficient alternative to SIFT or SURF." In Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 2564-2571. IEEE, 2011.

[13] Hamilton, Nicholas A., Radosav S. Pantelic, Kelly Hanson, and Rohan D. Teasdale. "Fast automated cell phenotype image classification." BMC bioinformatics 8, no. 1 (2007): 110.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp. 10971105, 2012.

[15] Cireşan, D.C., Giusti, A., Gambardella, L.M. and Schmidhuber, J., 2013. "Mitosis detection in breast cancer histology images with deep neural networks." In Medical Image Computing and Computer-Assisted InterventionMICCAI 2013 (pp. 411-418). Springer Berlin Heidelberg.

[16] Bayramoglu N., Kannala J., Heikkilä J. "Human Epithelial Type 2 cell classification with convolutional neural networks". In Bioinformatics and Bioengineering (BIBE), IEEE 15th Int. Conf. on 2015 Nov 2 (pp. 1-6).

[17] Y. Mao, Z. Yin and J. M. Schober, "Iteratively training classifiers for circulating tumor cell detection," Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on, New York, NY, 2015, pp. 190-194.

[18] Brebisson, Alexander, and Giovanni Montana. "Deep Neural Networks for Anatomical Brain Segmentation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 20-28. 2015.

[19] Cruz-Roa, A., Basavanhally, A., Gonzlez, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J. and Madabhushi, A., "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks." In SPIE Medical Imaging, pp. 904103-904103. International Society for Optics and Photonics, 2014.

[20] Su H, Liu F, Xie Y, Xing F, Meyyappan S, Yang L. "Region segmentation in histopathological breast cancer images using deep convolutional neural network." In Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on 2015 Apr 16 (pp. 55-58).

[21] Wan, S., Huang, X., Lee, H.C., Fujimoto, J.G. and Zhou, C., 2015, April. "Spoke-LBP and ring-LBP: New texture features for tissue classification". In Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on (pp. 195-199). IEEE.

[22] Keskin F, Suhre A, Kose K, Ersahin T, Cetin AE, Cetin-Atalay R. "Image classification of human carcinoma cells using complex wavelet-based covariance descriptors." PloS one. 2013 Jan 16;8(1):e52807.

[23] Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 34-42. 2015.

[24] Bayramoglu N. and Heikkilä J. "Transfer Learning for Cell Nuclei Classification in Histopathology Images", In Computer Vision - ECCV 2016 Workshops - Amsterdam, the Netherlands, October 8-10 and 16.