Classification of Histopathological Images Using Convolutional Neural Network

Nuh Hatipoglu[1], Gokhan Bilgin[2]
[1] Trakya University Computer Technology and Programming Department, Edirne, Turkey
e-mail: nhatipoglu@trakya.edu.tr
[2] Yildiz Technical University Computer Engineering Department, Istanbul, Turkey
e-mail: gbilgin@yildiz.edu.tr

***Abstract*— In this work, classification of cellular structures in the high resolutional histopathological images and the discrimination of cellular and non-cellular structures have been investigated. The cell classification is a very exhaustive and time-consuming process for pathologists in medicine. The development of digital imaging in histopathology has enabled the generation of reasonable and effective solutions to this problem. Morever, the classification of digital data provides easier analysis of cell structures in histopathological data. Convolutional neural network (CNN), constituting the main theme of this study, has been proposed with different spatial window sizes in RGB color spaces. Hence, to improve the accuracies of classification results obtained by supervised learning methods, spatial information must also be considered. So, spatial dependencies of cell and non-cell pixels can be evaluated within different pixel neighborhoods in this study. In the experiments, the CNN performs superior than other pixel classification methods including SVM and *k*-Nearest Neighbour (*k*-NN). At the end of this paper, several possible directions for future research are also proposed.**

***Keywords*— Histopathological images, convolutional neural networks, classification, image processing**

I. Introduction

Image processing algorithms for automated analysis of histopathological images have become increasingly popular in the last decade with the remarkable growth in computational power. The advent of high-throughput scanning devices allows for computer-aided evaluation of microscopic images, resulting in a quick and unbiased image interpretation that will facilitate the clinical decision-making process. In traditional analysis methods pathologists use H&E (hematoxylin and eosin) stained tissue for cytological description of cells such as cell size, shape, distribution and cellular spaces. Computer histological image processing techniques have been developed to detect cancer automatically at a level of accuracy comparable to an expert pathologist. The long term goal of this research is to develop a computer aided diagnosis (CAD) system that can be used to help pathologists diagnose cancer with improved accuracy, reproducibility and efficiency. In present research work, we use computer image-processing techniques and machine learning algorithms for classifying cancer tissue and nonmalignant tissue in digital histological images acquired from tissue section with H&E stains. Digital color histological images were acquired using light microscopes and charge coupled device (CCD) cameras from region of interest (ROIs) marked by pathologists on tissue sections [1], [2].

Computer-aided clinical image analysis has attracted huge interest from both signal processing and medical researchers due to its potential to surmount the challenges associated with the subjective examination of microscopic images. Quantitative tools for characterization of biomedical images mitigate the effects of inter and intra-reader variability on diagnosis and complement the clinical decision by acting as a second reader. Computer-assisted diagnosis (CAD) systems can prevent pathologists from wasting their time on image regions where decisions can be made in a straightforward manner; the percentage of the benign prostate biopsies in the U.S. is around 80%, which may be handled by computerized image analysis, leaving more space for pathologists to deal with challenging cases [3]. Biologists would also need automated tools to discriminate between a large number of different cell types since the biochemical tests performed for identification of cells may be very costly. The goal of this thesis is to develop new algorithms for computer-aided analysis of biomedical images by employing image processing and machine learning techniques. The main objective is to devise automated methods for feature detection, extraction and classification that would provide robust interpretation of biological and histopathological images.

Many attributes that will be obtained from digital histopathological images are associated with cell morphology. Therefore, it constitutes one of the basic steps of the system that will classify cells independent of background. [3].

In recent years, many supervised [4] and unsupervised [5] machine learning algorithms have been used to train histopathological image classification such as support vector machines [6], neural networks [7], fuzzy and genetic algorithms [8], *k*-NN [9], kernel PCA [10], etc. Supervised image classification methods of training and testing procedures to leave class or classes of the image to be obtained which is known in advance are required.

The neural networks area are used for cancer affected micro object classification purpose. Neural network which is different from algorithmic calculation method of today's computers is developed as a new science area. Neural network is modeled as the result of inspiration from biological neural systems and it has a simpler structure compared to them. Many developed neural network systems imitate some well-known characteristics of biological neural networks like learning ca-

978-1-4799-6463-5/14/$31.00 ©2014 IEEE

pability. Some other features are developed using engineering approach instead of neural physiological approaches. Neural network which can produce new outcomes, generate new information, discover new outputs by capability of learning, is a new computer system which does not need any aid. The neural network examines the learning samples, generalizes these samples and generates a learning rule from samples. Neural network can make a decision on any of the samples which is not seen before by using learning rules.

Convolutional neural networks (CNN) have been applied with great success for high-level computer vision tasks such as object recognition [11]. Recent studies have shown that they can also be used as a general method for low-level image processing problems, such as restoration [12], denoising [13], classification [14] and mitosis detection [15]. In this study, the classification of convolutional neural network cell nucleus and the problem of recovering a labeling image of cell nuclei from observed color of H&E stained cancer tissue images are aimed. The result of the classification can be used as the input for other tasks such as the classification or diagnosis of nucleus feature extraction.

Convolutional neural networks (CNN) have been applied with great success for high-level computer vision tasks such as object recognition [11]. Recent studies have shown that they can also be used as a general method for low-level image processing problems, such as restoration [12], denoising [13], classification [14] and mitosis detection [15]. In this study, the classification of convolutional neural network cell nucleus and the problem of recovering a labeling image of cell nuclei from observed color of H&E stained cancer tissue images are aimed. The result of the classification can be used as the input for other tasks such as the classification or diagnosis of nucleus feature extraction.

II. Convolutional neural network

Convolutional neural network (CNN) was initially proposed by LeCun in the early 1980s. It was further revised by LeCun and his colleagues in the 1990s [16]. CNN is a neural network model that is based on three key architectural ideas: local receptive fields, weight sharing, and subsampling in the spatial domain. The CNN is designed mainly for the recognition of two dimensional visual patterns. Convolutional neural network has much strength: (i) feature extraction and classification are integrated into one structure and fully adaptive; (ii) the network extract 2-D image features at increasing dyadic scales; (iii) it is relatively invariant to geometric, local distortions in the image. In this paper we use development of CNN architecture based on [17] and [18].

Convolutional neural networks are designed to process two dimensional (2-D) images. A CNN consists of three main types of layers: (i) convolution layers, (ii) subsampling layers, (iii) an output layer. Network layers are arranged in a feed-forward structure: each convolution layer is followed by a sub-sampling layer, and the last convolution layer is followed by the output layer. These layers showed Figure1. The convolution and subsampling layers are considered as 2-D layers, whereas the output layer is considered as a 1-D layer. In CNN, each 2-D layer is made up from many planes. A plane consists of neurons that are arranged in a 2-D array. The output of a plane is called a feature map [17], [18]. Fig.1. shows that the diagram of the CNN.

A convolutional layer, each plane is connected to one or more feature maps of the preceding layer. A connection is associated with a convolution mask, which is a 2-D matrix of adjustable entries called weights. Each plane first computes the convolution between its 2-D inputs and its convolution masks. The convolution outputs are summed together and then added with an adjustable scalar, known as a bias term. Finally, an activation function is applied on the result to obtain the plane's output. The plane output is a 2-D matrix called a feature map; this name arises from the fact that the convolution output indicates the presence of a visual feature at a given pixel location [17], [18].

Consider a convolution layer l. l is an odd interger, $l = 1, 3, ..2a+1$. Let $r_l \times c_l$ denote the size of convolution mask for layer l. For feature map n in layer C_l, let $w^l_{m,n} = w^l_{m,n}(i,j)$ be the convolution mask form feature mask map m in layer $l-1$ to feature map n in layer l, b^l_n be the bias term associated with feature map n , V^l_n denote the list of all planes in layer $l-1$ that are connected to feature map n.

Feature map n of convolution layer l is calculate as

$$y^l_n = f_l(\sum_{m \in V^l_n} y^{l-1}_m \otimes w^l_m, n + b^l_n) \quad (1)$$

Where $\otimes$ denotes the 2D convolution operator. Suppose that the size of input featurea map y^{l-1}_m is $H_{l-1} \times W_{l-1}$ pixels, and the size of convolution masks $w^l_{m,n}$ is $r_l c_l$ pixels. Because we perform convolution without zero-padding the inputs, the size of output feature map y^l_n is $(H_{l-1} + r_l) \times (W_{l-1} - c_l + 1)$ pixels.

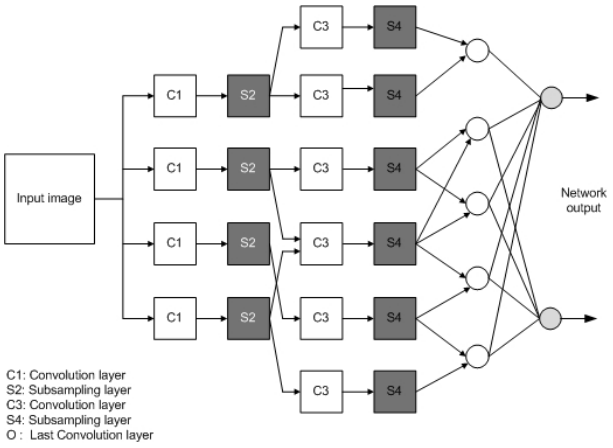


Fig. 1: Convolutional neural networks layers.

Sub-sampling layer has the same number of planes as the preceding convolution layer. A subsampling plane divides its 2-D input into non-overlapping blocks. For each block, the

sum of four pixels is calculated; this sum is multiplied by an adjustable weight before being added to a bias term. The result is passed through an activation function to produce an output for the 2x2 block. Clearly, each sub-sampling plane reduces its input size by half, along each dimension. A feature map in a sub-sampling layer is connected to one or more planes in the next convolution layer [17], [18].

Consider a subsampling layer l ,where l is an even integer $l = 2, 4, ...2a$. For feature map n in layer, w_n^l be the weight and b_n^l be the bias term. Feature map n of convolution layer $l-1$ into non-overlapping blocks of size 2x2 pixels. Let z_n^{l-1} be a matrix obtained by summing the four pixels en each block. That can be shown as

$$\begin{aligned} z_n^{l-1}(i,j) = y_n^{l-1}(2i-1, 2j-1)+ \\ y_n^{l-1}(2i-1, 2j+ \\ y_n^{l-1}(2i, 2j-1)+ \\ y_n^{l-1}(2i, 2j) \end{aligned} \tag{2}$$

Feature map n of subsampling layer l is now calculated as $y_n^l = f_l(z_n^{l-1} \times w_n^l + wb_n^l)$. A feature map y_n^l in subsampling layer $lH_l = H_{l-1}/2$ will have a size of $H_l \times W_l$, where $W_l = W_{l-1}/2$. In the last convolution layer, each plane is connected to exactly one preceding feature map. This layer uses convolution masks that have the same size as the input feature maps. Therefore, each plane in the last convolution layer will produce one scalar output. The outputs from all planes in this layer are then connected to the output layer [17], [18].

The output layer, in general, can be constructed from sigmoidal neurons or radial-basis-function (RBF) neurons. Here, we will focus on using sigmoidal neurons for the output layer. The outputs of this layer are considered as the network outputs. In applications such as visual pattern classification, these outputs indicate the category of the input image [17], [18].

Consider output layer L that consists of sigmoidal neurons. Let N_L be the number of output sigmoidal neurons. Let $w_{m,n}^L$ denote the weight from feature map m of the last convolutional layer, to neuron n of the output layer L. The output of sigmoidal neuron n is calculate as

$$y_n^L = f^L(\sum_{m=1}^{N_{L-1}} y_m^{L-1} w_{m,n}^L + b_n^L) \tag{3}$$

The output of all sigmoidal neurons from the network outputs.

III. Experimental Results

In this section, the influence of the proposed method, namely convolutional neural network (CNN), is shown using *k*-Nearest Neighbour (*k*-NN) and Support Vector Machine (SVM) on histopathological image dataset.

The *k*-NN algorithm is a well known algorithm in machine learning algorithms. It is an instance-based learning algorithm where density function is only approximated locally. The key idea is to classify a test point in the d-dimensional feature space by a majority voting of its k neighbors where k is positive integer. Majority voting refers to taking the most common class label among points from training set which are most similar to the test point. In summary, determining the class of a test point consists of calculating Euclidean distance between a test point and all points in the training set (i.e. class labels are already known for training set), selection of *k*-nearest points to test point in the training set, and finally assigning the test point to the most common class among its *k*-nearest neighbors. The algorithm needs no explicit training process to build models to reduce the feature dimensions. So, it is not needed to take all training points into account for computation [19]

In origin, SVM is a binary classifier; it can evaluate data points and assign them one of two classes. In order to perform this classification, SVMs are trained with training data and their class labels, and then assign a class to novel data. In SVM learning algorithm there are many types of general purpose kernels in the literature like linear, polynomial and Radial Basis Function (RBF) kernels. RBF kernel is one of the most popular kernel in the literature which uses a Gaussian distribution to mapping data to higher dimensional space. Namely, RBF kernel non-linearly maps samples into a higher dimensional space unlike the linear kernel [20].

Data set contains 58 Hematoxylin and Eosin (H&E) histopathology images of breast tissue from the Yale Tissue Microarray Facility. There are 26 malignant and 32 benign images respectively. Images have RGB color channels. Each image is represented in analog image as 8 bit with 896x768 binary ground truth images. These images and the associated ground truth data are provided by the Bio-image Informatics Center of UCSB, and can be downloaded as a part of UCSB Bio-Segmentation Benchmark data set [21]. Fig.2, Fig.3 represents benign and malignant histopathology image examples within a cropped window structure and their associated ground truth labels.

In histopathological H&E images, tissue components are stained as dark purple as nuclei and pink as cytoplasm and the extracellular connective tissue. ROI(regions-of-interest) extraction process from cancer cell line images is described.

First, image classification is performed to obtain cellular regions and then ROIs are selected from foreground cellular regions. Cellular regions of each image in dataset are labeled as “1” and noncellular regions are labeled as “-1”. Non cellular regions are called as “background”. Since the ”foreground” and the ”background” regions can be roughly distinguished by color of the pixels, it is natural to attempt binary classified by pixel-level classification using multispectral inputs from the color histopathology images.

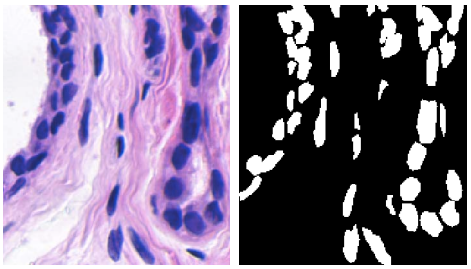
In the formation of training sets of cellular and extracellular

structures, equal number of samples has been taken from 3x3, 5x5, 7x7 and 9x9 sub-windows in different number of random samples chosen from all histopathological slides. At the end of the procedure, four training sets were created. In each group, an equal amount of random cell and non-cell samples are taken. The data sets used in this study are formed as 1/3 of the training sets and 2/3 of test sets. For example, if the 50 cellular and 50 non-cellular samples are taken, there are 2900 total samples for training and 8700 total samples are tested in the formed sub-dataset. These training sets are created with three different models using various supervised learning methods: *k*-NN, SVM and CNN. The study is implemented in all Matlab^{TM} 2009b platform. Three commonly used classifiers, including [18], SVM [22], and *k*-NN are applied to the dataset to produce the baseline performances for comparison purposes.

Flow chart of the intended study is summarized in Fig.4. 5x5 median filter has been used to eliminate the noise in images and to carry out smoothing application by preserving cell edges. Each colored image is cropped according to maximum-minimum values of ground truths. Pixels selected randomly from cell and noncell pixel areas are obtained in the original RGB images R channel. Classification results are evaluated by statistic and overall accuracy (OA). Overall accuracy is estimated by ratio between correctly classified test samples and the total number of test samples. The κ statistic is estimated by weighting the estimated accuracies.

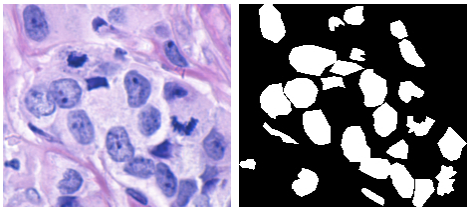
The classification accuracies are represented by CNN, SVM and *k*-NN classification methods with different sub-window sizes as shown in Table 1, Table 2, Table 3 and Table 4. In all sub-datasets, the accuracy increases when the number of training samples increases. Accuracies of CNN implemented pixels with 3x3, 5x5, 7x7 and 9x9 sub-window sizes are almost the same by slight differences.

The visual comparison of the classification results of the benign image shown in Fig.2(a) using the combination inputs is represented in Fig.5 and malignant image shown in Fig.3(a) combination inputs is represented in Fig.6. As can be seen, the convolutional networks provide superior performance to other pixel classification methods used in this paper.



(a) (b)

Fig. 2: Benign histopathology image and its ground truth label



(a) (b)

Fig. 3: Malignant histopathology image and its ground truth label

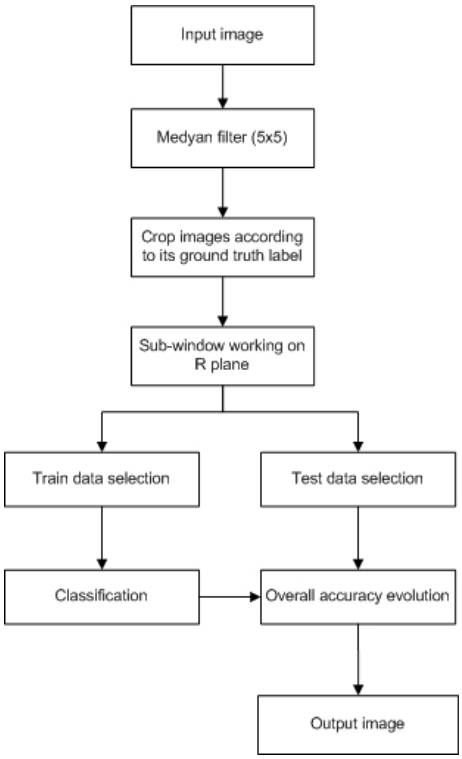


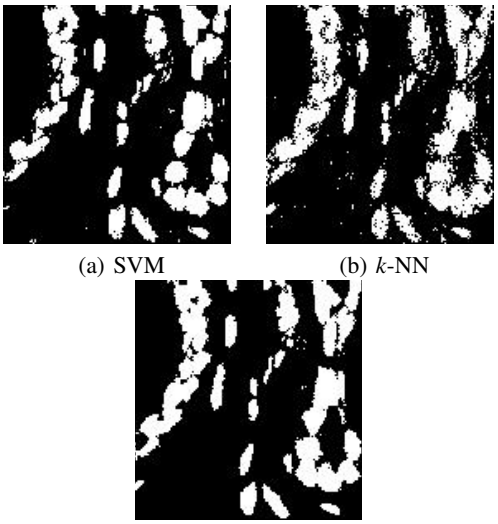
Fig. 4: Classification flow chart

Table 1: The test result of training samples 1160.

	Classification algorithm					
	CNN		SVM		*k*-NN	
	OA	κ	OA	κ	OA	κ
Sub-window 3x3	66.51	0.5177	61.51	0.52	62.54	0.55
Sub-window 5x5	66.57	0.5157	61.45	0.55	62.35	0.53
Sub-window 7x7	66.63	0.5142	61.56	0.56	62.61	0.55
Sub-window 9x9	66.75	0.5127	61.73	0.57	62.82	0.58

Table 2: The test result of training samples 2900.

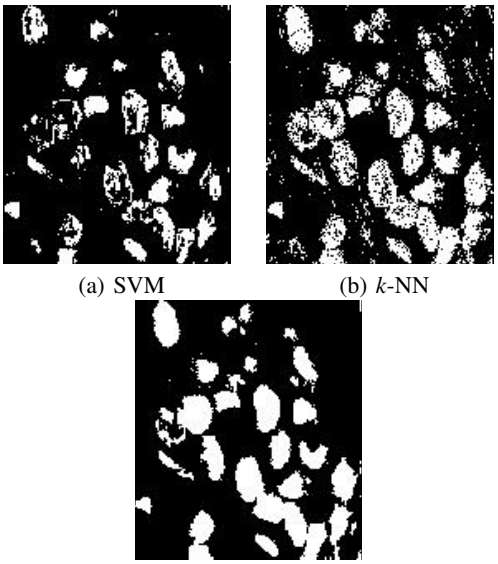
	Classification algorithm					
	CNN		SVM		*k*-NN	
	OA	κ	OA	κ	OA	κ
Sub-window 3x3	71.43	0.61	70.25	0.57	68.45	0.56
Sub-window 5x5	71.71	0.61	70.54	0.57	68.53	0.56
Sub-window 7x7	71.63	0.61	70.65	0.57	68.65	0.56
Sub-window 9x9	71.51	0.61	70.78	0.57	68.78	0.56



(a) SVM (b) k-NN

(c) CNN

Fig. 5: Virtual comparison of the benign histopathology image classification results



(a) SVM (b) k-NN

(c) CNN

Fig. 6: Virtual comparison of the malignant histopathology image classification results

Table 3: The test result of training samples 5800.

	Classification algorithm					
	CNN		SVM		k-NN	
	OA	κ	OA	κ	OA	κ
Sub-window 3x3	81.59	0.66	78.21	0.61	77.11	0.59
Sub-window 5x5	81.63	0.67	78.38	0.61	77.65	0.59
Sub-window 7x7	81.76	0.66	78.43	0.61	77.78	0.59
Sub-window 9x9	81.81	0.66	78.32	0.61	77.81	0.59

IV. Conclusion

In this study, cell classification in histopathological images has been proposed using the spatial information via CNN.

Table 4: The test result of training samples 11600.

	Classification algorithm					
	CNN		SVM		k-NN	
	OA	κ	OA	κ	OA	κ
Sub-window 3x3	86.67	0.7327	81.25	0.632	82.45	0.6245
Sub-window 5x5	86.91	0.7357	81.54	0.615	82.53	0.6253
Sub-window 7x7	86.17	0.7242	81.65	0.616	82.65	0.6265
Sub-window 9x9	86.88	0.7327	81.78	0.617	82.78	0.6278

CNN, SVM and k-NN implementation of randomly selected cell and non-cell samples with different sub-window sizes have been compared. Classification accuracies of CNN algorithm with respect to total number of samples of cell and non-cell pixels have also been evaluated for comparative purposes. According to the demonstrated results, CNN classification accuracies are better than the other classification methods. Furthermore, it is also possible to extend convolutional networks to adopt other domain-specific features (such as textural features) to further improve the classification accuracies.

References

[1] M. Dundar, S. Badve, G. Bilgin, V. Raykar, R. Jain, O. Sertel, and M. Gurcan, "Computerized classification of intraductal breast lesions using histopathological images," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 7, pp. 1977–1984, July 2011.

[2] M. Veta, J. Pluim, P. van Diest, and M. Viergever, "Breast cancer histopathology image analysis: A review," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 5, pp. 1400–1411, May 2014.

[3] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *Biomedical Engineering, IEEE Reviews in*, vol. 2, pp. 147–171, 2009.

[4] D. Onder, S. Sarioglu, and B. Karacali, "Automated labelling of cancer textures in colorectal histopathology slides using quasi-supervised learning," *Micron*, vol. 47, no. 0, pp. 33 – 42, 2013.

[5] A. GençTav, S. Aksoy, and S. Önder, "Unsupervised segmentation and classification of cervical cell images," *Pattern Recogn.*, vol. 45, no. 12, pp. 4151–4168, Dec. 2012.

[6] V. Ojansivu, N. Linder, E. Rahtu, M. Pietikainen, M. Lundin, H. Joensuu, and J. Lundin, "Automated classification of breast cancer morphology in histopathological images," *Diagnostic Pathology*, vol. 8, no. 1, pp. 1–4, 2013.

[7] D. Ciresan, A. Giusti, L. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, ser. Lecture Notes in Computer Science, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds. Springer Berlin Heidelberg, 2013, vol. 8150, pp. 411–418.

[8] L. Gan Lim, R. Naguib, E. Dadios, and J. Avila, "Implementation of gaksom and anfis in the classification of colonic histopathological images," in *TENCON 2012 - 2012 IEEE Region 10 Conference*, Nov 2012, pp. 1–5.

[9] "Computer-aided diagnosis for breast tumor classification using microscopic images of fine needle biopsy," in *Intelligent Systems in Technical and Medical Diagnostics*, ser. Advances in Intelligent Systems and Computing, J. Korbicz and M. Kowal, Eds. Springer Berlin Heidelberg, 2014, vol. 230.

[10] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J. Saltz, and M. Gurcan, "Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development," *Pattern Recognition*, vol. 42, no. 6, pp. 1093 – 1103, 2009.

[11] Y. LeCun, F. J. Huang, and L. Bottou, "Learning methods for generic object recognition with invariance to pose and lighting," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, pp. II–97–104 Vol.2.

[12] V. Jain, J. F. Murray, F. Roth, S. C. Turaga, V. P. Zhigulin, K. L. Briggman, M. Helmstaedter, W. Denk, and H. S. Seung, "Supervised

learning of image restoration with convolutional networks." IEEE, 2007, pp. 1–8.

[13] V. Jain and H. S. Seung, "Natural image denoising with convolutional networks." in *NIPS*. Curran Associates, Inc., 2008, pp. 769–776.

[14] C. Malon, M. Miller, H. C. Burger, E. Cosatto, and H. P. Graf, "Identifying histological elements with convolutional neural networks," in *Proceedings of the 5th International Conference on Soft Computing As Transdisciplinary Science and Technology*, ser. CSTST '08. New York, NY, USA: ACM, 2008, pp. 450–456.

[15] H. Wang, A. Cruz-Roa, A. Basavanhally, H. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, "Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection," pp. 90 410B–90 410B–10, 2014.

[16] L. Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Advances in Neural Information Processing Systems*. Morgan Kaufmann, 1990, pp. 396–404.

[17] S. L. Phung and A. Bouzerdoum, "A pyramidal neural network for visual pattern recognition," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 329–343, 2007.

[18] ——, "Matlab library for convolutional neural networks," ICT Research Institute,Visual and Audio Signal Processing Laboratory, University of Wollongong, Tech. Rep., 2009.

[19] Y. Gao, P. Jin-yan, and F. Gao, "Improved boosting algorithm through weighted k-nearest neighbors classifier," in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, vol. 6, July 2010, pp. 36–40.

[20] C. W. Hsu, C. C. Chang, and J. Lin, "A practical guide to support vector classification," Department of Computer Science and Information Engineering, National Taiwan University, Tech. Rep., 2003.

[21] E. Drelie Gelasca, B. Obara, D. Fedorov, K. Kvilekval, and B. Manjunath, "A biosegmentation benchmark for evaluation of bioimage analysis methods," *BMC Bioinformatics*, vol. 10, no. 1, pp. 1–12, 2009.

[22] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.