Convolutional Neural Network for Classification of Histopathology Images for Breast Cancer Detection

Barath Narayanan Narayanan*, Vignesh Krishnaraja+ and Redha Ali±
Department of Electrical and Computer Engineering
University of Dayton, Dayton, OH 45469, USA
Email: {narayananb1*, krishnarajav1+, almahdir1±}@udayton.edu

Abstract - **One of the most common subtypes of all breast cancers is Invasive Ductal Carcinoma (IDC). Pathologists typically focus on regions which contain IDC to determine whether a patient suffers from breast cancer or not. We make use of publicly available Breast Histopathology Images dataset provided at the Kaggle for classification. In this dataset, images are delineated to extract the exact regions of IDC. This dataset contains 277,524 patches among which 198,738 belong to the negative class and 78,786 belong to the positive class (images with IDC). We present a novel deep convolutional neural network architecture for classification. Our performance in terms of area under the receiver operating characteristic curve for the detection of IDC is 0.94 on a set of 27,553 test images thereby setting a new benchmark for future research efforts.**

Index Terms – **Breast Cancer, Histopathology, Convolutional Neural Networks,**

I. INTRODUCTION

According to International Agency for Research on Cancer (IARC), breast cancer is one of the most frequently diagnosed cancers and cause of cancer deaths among women [1]. Invasive Ductal Carcinoma (IDC) also known as infiltrating ductal carcinoma is the most common subtype of cancer and it accounts to about 80% of all breast cancer cases [2-3]. IDC starts growing in a milk duct and invades the fibrous or the fatty tissue of the breast outside the duct. It can even spread to lymph nodes and other parts of body. The overall 5-year relative survival rate for people with breast cancer is 89% [4]. The stage of the breast cancer at diagnosis is one of the most important factors in predicting breast cancer survival rates [4]. The 5-year relative survival rate for stage 0-1 is close to 100% but only 61% of all breast cancers are diagnosed at this stage [4]. At stage 4, the 5-year relative survival rate drastically decreases to 22% [4]. It is important to detect cancer at the early stages. IDC cannot be easily characterized by specific morphological features that characterize the "special subtypes" [1]. Detection of IDC is a challenging and time-consuming process as pathologists visually inspect the histological samples under the microscope. Automating the detection of this cancer type would be valuable to pathologists by speeding up the diagnosis and reduce error.

Automated detection in the field of medical imaging has been a research area attracting great interest in the past decade [5-22]. Feature extractor methods such as the Scalar Invariant Fourier Transform (SIFT), Speeded-up Robust Features (SURF) and Local Binary Pattern (LBP) are utilized to classify the breast histopathological images as benign or malignant. In [12-14], Convolutional Neural Networks (CNNs) [24] are utilized to classify the histopathological images as benign or malignant. Breast cancer images have a wide inter-class variation as well as complex geometrical shape thereby making it challenging for its classification. In [19], a histopathological database of microscopic breast tumor images with two stages of benign and malignant under four different magnifications are presented. However, most of the literature work deals with classifying images as benign or malignant and not much research has been done on classifying images as cancer or not cancerous except [16-17].

In [20-23], computer aided detection methods to detect lung nodules in chest radiographs and computed tomography are presented. In the work presented in [20-23], an image preprocessing method is followed by a classification architecture to classify the candidates as nodules or non-nodules. We adopt the same architecture in this work. In [16], performance comparison of fuzzy color histogram and RGB histogram based feature engineering approaches is presented for classification of histopathological images. In [16], a CNN architecture is also presented and it significantly outperformed the histogram methods. In [17], a 6-layer CNN architecture is presented and they achieved an overall Area Under the Receiver Operating Characteristic Curve (AUC) value of 0.902. We believe this is an important area to investigate. We emphasize the same in this paper. We present a novel 5-layer CNN architecture with no data augmentation for classifying publicly available histopathological images as cancerous or not cancerous thereby setting a benchmark for future research efforts. Our proposed architecture achieves an overall AUC value of 0.935 for detecting IDC on a set of 27,553 test images.

The remainder of this paper is organized as follows. Section 2 provides a brief description about the databases that are

employed in this research. Section 3 presents the deep CNN architecture proposed in this research. Section 4 presents the experimental results obtained using the proposed methods. Finally, conclusions are offered in Section 5.

II. Materials

Aforementioned, we utilize a publicly available database [18] to study the performance of our proposed algorithms. The dataset is provided in Kaggle. It consists of 162 whole mount slide images of breast cancer specimens scanned at 40x. From that, 277,524 patches/images of size 50 × 50 were extracted (198,738 IDC negative and 78,786 IDC positive). Each image is associated with a patient ID and label marked by a trained doctor as positive (indicating the presence of IDC) or negative (indicating the absence of IDC) [18]. For our algorithm purposes, we split the dataset into groups of 81%, 9% and 10% for training, validation and testing purposes respectively. Table 1 presents the distribution of the training, validation and testing datasets utilized in this research.

Table 1: Distribution of Train, Validation and Test Datasets

Dataset	Label	
	Negative	Positive
Train	160978	63816
Validation	17886	7091
Test	19874	7879

Figures 1 and 2 show the example patches of positive and negative images present in the dataset. In these two figures, we can observe that intra-class variation is very less between the cancerous and non-cancerous images thereby making it challenging for its classification.

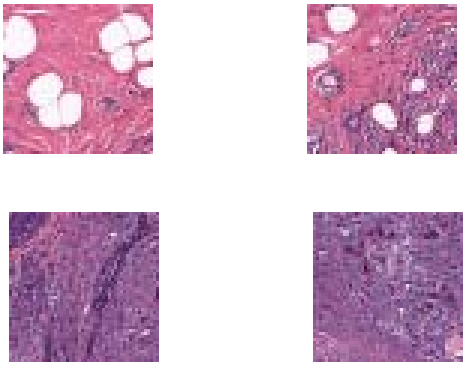

Figure 1: Examples of Histopathology images marked as Positive

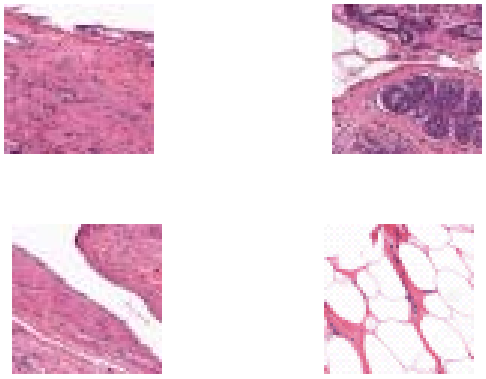

Figure 2: Examples of Histopathology images marked as Negative

III. CNN architecture

For this research, the images are first preprocessed using color constancy technique [25]. A human observer is able to recognize the color of objects irrespective of the light used to illuminate them. This is called color constancy [25]. In addition, we also study the performance of algorithms by converting images into grayscale and enhancing them using histogram equalization. Histogram equalization is a method to adjust the contrast using the image's histogram. Figure 3 shows an example of these two preprocessing techniques.

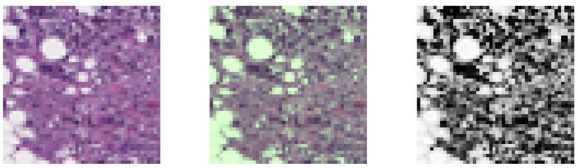

Figure 3: Example of images after preprocessing: (a) Original Image, (b) Image obtained using Color Constancy, (c) Image obtained after histogram equalization on grayscale image

The 50×50 patch/image is resized to 48×48. It is then preprocessed using the color constancy technique/histogram equalization of grayscale images to serve as the input to the convolutional layers. The CNN architecture has the following components: convolutional layer, activation using Rectified Linear Unit (ReLU) layer, batch normalization layer, pooling layer, dropout layer, fully connected layer and softmax layer [26-28]. Figure 4 presents the architecture proposed by us. Our proposed architecture has 5 convolutional layers which is then followed by a fully connected and a softmax layer for classification.

The convolutional layer will determine the output of neurons which are connected to local regions of the input throughout the

calculation of the scalar product between their weights and the region connected to the input volume [26]. A number of filters of small size (eg. 3×3, 5×5, 7×7) convolute with the input data for extracting different features. In our architecture, each convolutional layer contains 3×3 convolution filters along with a ReLU, a batch normalization layer, and maximum pooling layer. We are using a 3×3 convolution filter to make sure we capture the features of IDC. Also, the convolutional layers have same padding meaning the convoluted output retains the same size of the input. Table 2 shows the number of convolution filters present in each layer. Outputs of the convolution operations are passed to a ReLU layer.

The activated values are normalized by Batch normalization layer. Batch normalization is a technique to normalize activations in intermediate layers of deep neural networks. Its tendency to improve accuracy and speedup training shows its effectiveness [27].

The pooling layer will then simply perform down sampling along the spatial dimensionality of the given input, further reducing the number of parameters within that activation [26]. It aims to gradually reduce the dimensionality of the representation, and thus further reduce the number of parameters and the computational complexity of the model. In our CNN architecture, we are using a max pooling layer of size 2×2 with a stride of 2.

In the last convolutional layer, i.e. 5th convolutional layer we use a dropout layer. Dropout is a technique for addressing overfitting problem. The key idea is to randomly drop units (along with their connections) from the neural network during training [28]. This prevents units from co-adapting too much. [28]. Dropout significantly reduces overfitting and gives major improvements over other regularization methods [28]. In our architecture, we use a remove probability of 0.25 for the dropout layer.

The fully connected layer contains neurons of which are directly connected to the neurons in the two adjacent layers, without being connected to any layers within them [26]. Our fully connected layer has 128 neuron units.

It is then followed by a softmax layer with 2 neuron units. This layer is merely a fully connected layer with softmax function applied to the output values. The softmax function normalizes the output values into probability distribution.

These layers are implemented with the help of built-in MATLAB functions *convolution2dLayer, reluLayer, batchNormalizationLayer, maxPooling2dLayer, dropoutLayer fullyConnectedLayer, softmaxLayer,* and *classificationLaye*r [29].

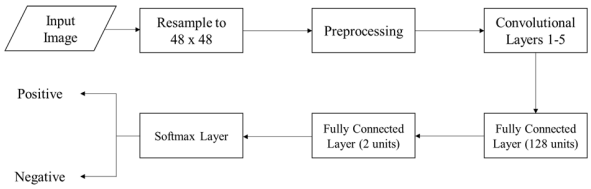

Figure 4: Proposed CNN architecture

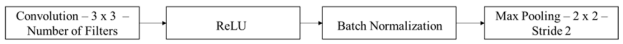

Figure 5: Convolutional Layer Structure

Table 2: Number of convolution filters present in each convolutional layer

Convolutional Layer #	Number of Filters
Layer #1	32
Layer #2	64
Layer #3	64
Layer #4	128
Layer #5	256

IV. Experimental Results

In this section, we present the results obtained using two different approaches of image preprocessing and the proposed deep CNN architecture. We train our network using Adam optimization technique and validation frequency of 200 iterations. It took less than 2 epochs to train our classifier. We also present the performance in terms of AUC for detecting IDC. Results are obtained using the Testing dataset as mentioned in Table 1. Figure 6 presents the Receiver Operating Characteristic (ROC) curve for both the approaches. Table 3 summarizes the AUC values obtained using these approaches.

Table 3: AUC using CNN with different techniques

Methodology	**AUC**
Histogram Equalization with Proposed CNN	0.876
Color Constancy with Proposed CNN	**0.935**
Existing Benchmark [17]	0.902

AUC values obtained using our proposed CNN using color constancy and histogram equalization techniques are **0.935** and **0.876** respectively which significantly outperforms the existing benchmark [17].

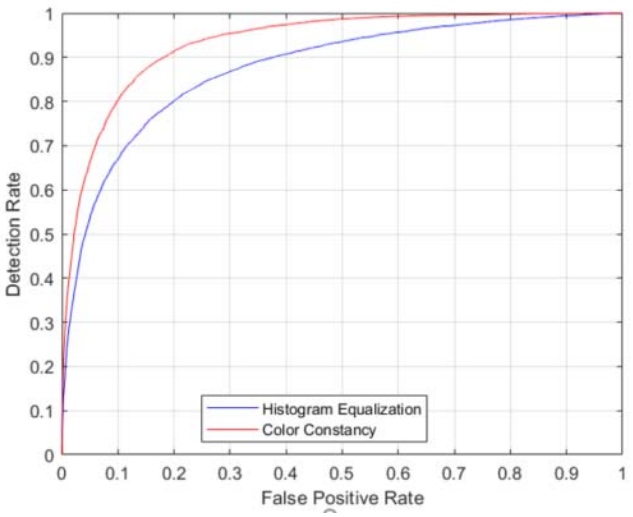

Figure 6: ROC Curves obtained using Color Constancy and Histogram equalization

There is a significant increase in performance when we used color constancy technique instead of histogram equalization technique of grayscale images. Also, from Figures 1 and 2 it is evident that color plays an important role in classification of cancerous and non-cancerous images. In [16-17], the authors converted the images from RGB to YUV color space before feeding it as input to the convolutional layers. On the other hand, we preprocessed the images with color constancy technique which outperformed the results obtained in [17]. The histopathology images could have been taken under different illuminations which affects the way we perceive color. This could have affected the performance when they converted the images from RGB to YUV color space in [16-17]. The color constancy technique performed better because we are bringing the RGB values to a constant so that the color of these images is the same when seen under different illuminations. It also helps in maintaining the contrast within each image and across all images present in the dataset.

From Table 1 we can see there is a class imbalance in the dataset provided. Overall the dataset contains 198,738 IDC negative and 78,786 IDC positive images. In [17], data augmentation is performed on just the IDC positive images to fix the class imbalance issue. Our CNN architecture is able to outperform the existing benchmark [16-17] roughly by 3% without data augmentation even with class imbalance.

V. Conclusions

In this paper, we presented a novel deep learning architecture for IDC detection. Results show that CNN is effective for detecting IDC. An AUC value of 0.935 is obtained on a set of 27,753 images (out of which 7879 contained IDC) and it sets a new benchmark for future research efforts. Results suggest that our proposed approach outperforms the existing benchmark for the dataset. Results indicate that the color constancy technique outperforms the histogram equalization technique. This might be due to the fact that color constancy helps in maintaining the contrast within each image and across all images present in the dataset. It also extracts essential spatial information to aid the CNN architecture for classification. Class imbalance also affects the performance of our classifier. Potentially we could use data augmentation to improve our performance in the future. Automating the detection of this cancer type (IDC) assists the pathologists with the diagnosis and provide a valuable second opinion.

References

[1] Stewart, B. W. and Wild, C. P., Eds., *World cancer report*. Lyon: IARC, 2014.

[2] "Invasive Ductal Carcinoma: Diagnosis, Treatment, and More," *Breastcancer.org*. [Online]. Available: https://www.breastcancer.org/symptoms/types/idc. [Accessed: 16-Mar-2019].

[3] "Invasive Breast Cancer (IDC/ILC) | Types of Invasive Breast Carcinoma," *American Cancer Society*. [Online]. Available: https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/types-of-breast-cancer/invasive-breast-cancer.html. [Accessed: 16-Mar-2019].

[4] Dresden, D., "Living with breast cancer: Statistics on survival rates by stage," *Medical News Today*. [Online]. Available: https://www.medicalnewstoday.com/articles/316867.php. [Accessed: 26-Apr-2019].

[5] Johra, F. and Shuvo, M. H. "Detection of Breast Cancer from Histopathology image and Classifying Benign and Malignant State Using Fuzzy Logic," Electr. Eng. Inf. Commun. Technol. (ICEEICT), 2016 3rd Int. Conf., no. 1, pp. 2–6, 2016.

[6] Tambasco Bruno, D. O., Do Nascimento, M. Z., Ramos, R. P., Batista, V. R., Neves, L. A., and Martins, A. S. "LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues," Expert Syst. Appl., vol. 55, pp. 329–340, 2016.

[7] Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F. "Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015: 18th International Conference Munich, Germany, October 5-9, 2015 proceedings, part III," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9351, pp. 366–373, 2015.

[8] Veta, M., Pluim, J. P. W., Van Diest, P. J., and Viergever, M. A. "Breast cancer histopathology image analysis: A review," IEEE Trans. Biomed. Eng., vol. 61, no. 5, pp. 1400–1411, 2014.

[9] George, Y. M., Zayed, H. L., Roushdy, M. I., and Elbagoury, B. M., "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," IEEE Systems Journal, vol. 8, no. 3, pp.949–964, 2014.

[10] Zhang, Y., Zhang, B., Coenen, F., and Lu, W., "Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles," Machine Vision and Applications, vol. 24, no. 7, pp. 1405–1420, 2013.

[11] Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D. R. J., Cree, I. A., and Rajpoot, N. M. "Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1196–1206, 2016.

[12] Bayramoglu, N., Kannala, J., and Heikkilä, J., "Deep learning for magnification independent breast cancer histopathology image classification," 2016 23rd Int. Conf. Pattern Recognit., pp. 2440–2445, 2016.

[13] Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. "Breast Cancer Histopathological Image Classification using Convolutional Neural Networks," Int. Jt. Conf. Neural Networks (IJCNN 2016), 2016.

[14] Lévy, D. and Jain, A. "Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks," arXiv Prepr. arXiv1612.00542, no. Nips, 2016.

[15] Nejad, E. M., Affendey, L. S., Latip, R. B., Ishak, and I. B. "Classification of Histopathology Images of Breast into Benign and Malignant using a Single-layer Convolutional Neural Network," Proceedings of the International Conference on Imaging, Signal Processing and Communication - ICISPC 2017, Jul. 2017.

[16] Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Medical Imaging 2014: Digital Pathology*, vol. 9041, no. 1, pp. 904103–1-904103–15, 2014.

[17] Cruz-Roa, A., Gilmore, H., Basavanhally, A., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., González, F., and Madabhushi, A., "Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent," *Scientific Reports*, vol. 7, no. 1, 2017.

[18] Kaggle.com. (2019). *Breast Histopathology Images*. [online]
Available at:
https://www.kaggle.com/paultimothymooney/breast-histopathology-images#IDC_regular_ps50_idx5.zip [Accessed 26 March 2019].

[19] Magna, G., Jayaraman, S. V., Casti, P., Mencattini, A., Di Natale, C., & Martinelli, E. (2015, February). "Adaptive classification model based on artificial immune system for breast cancer detection". In *2015 XVIII AISEM Annual Conference* (pp. 1-4). IEEE.

[20] Narayanan, B. N., Hardie, R. C., & Kebede, T. M. (2018). "Performance analysis of a computer-aided detection system for lung nodules in CT at different slice thicknesses". *Journal of Medical Imaging*, *5*(1), 014504.

[21] Narayanan, B. N., Hardie, R. C., Kebede, T. M., & Sprague, M. J. (2019). "Optimized feature selection-based clustering approach for computer-aided detection of lung nodules in different modalities". *Pattern Analysis and Applications*, 22(2), pp. 559-571.

[22] Narayanan, B. N., Hardie, R. C., & Kebede, T. M. (2016, July). "Analysis of various classification techniques for computer aided detection system of pulmonary nodules in CT". In *Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), 2016 IEEE National* (pp. 88-93). IEEE.

[23] Narayanan, B. N., Hardie, R. C., & Kebede, T. M. (2018, July). "Performance Analysis of Feature Selection Techniques for Support Vector Machine and its Application for Lung Nodule Detection". In *NAECON 2018-IEEE National Aerospace and Electronics Conference* (pp. 262-266). IEEE.

[24] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

[25] Ebner, M. *Color constancy: Marc Ebner*. Chichester: J. Wiley, 2007.

[26] O'Shea, K., Nash, R. "An Introduction to Convolutional Neural Networks", ArXiv e-prints, 2015.

[27] Bjorck, J., Gomes, C., and Selman, B. "Understanding Batch Normalization," *NeurIPS*, 2018.

[28] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, Jun. 2014.

[29] Mathworks.com. (2019). *MATLABDocumentation*. [online]
Available at: https://www.mathworks.com/help/matlab/, Accessed 26 Mar. 2019].