Breast Histopathological Image Feature Extraction with Convolutional Neural Networks for Classification †

Kevin Kiambe [a,*]

[a] Computer Engineering, Selçuk University, Konya, Turkey

Submitted: 25 Jun. 2018 — Accepted: 27 Jun. 2018 — Published: 30 Jun. 2018

Abstract—Recently, Convolutional Neural Networks (CNNs) have become a preferred deep learning artificial neural network of choice for computer assisted medical image analysis. These models are structured as a series of multiple hierarchical processing layers that can automatically learn feature representations from raw images. CNNs have in the past not been in common use, especially in the medical imaging field, due to issues such as insufficient image datasets. The revolution in CNN models has been attributed to powerful parallel processing hardware architectures, increasing number of image datasets and improved training strategies. Utilizing these deep learning techniques are enabling medical experts such as pathologists to utilize artificial intelligence to transform the world of medicine for faster and more accurate diagnoses. In this paper, a two stage model for classifying breast histopathological images is proposed. In the first stage, a CNN is used for extracting features from the images through a feature learning process. The extracted features are then used in the second stage to training classical machine learning models that include the Support Vector Machine (SVM), *k*-Nearest Neighbor (*k*-NN) and Logistic Regression (LR) models. The SVM classifier performs best with accuracies of up to 99.84%.

Keywords—Biomedical Imaging; Convolutional Neural Network; Deep Learning; Histopathological Images; Support Vector Machine.

I. INTRODUCTION

BREAST cancer is one of the most common types of cancer in the world that is more common in women than in men. According to the World Health Organization (WHO), about 15 percent of cancer incident cases are usually related to breast cancer [1]. Early detection can greatly help to reduce the mortality and morbidity rates resulting not only from these cancer type but also cancer affecting other body parts of the human being.

Medical imaging techniques such as histopathological, mammography, ultrasound and magnetic resonance imaging are used for the early detection of breast cancer [2]. Histopathological imaging produces some of the most accurate and more reliable results in the detection and staging of breast cancer. This technique is done by obtaining breast tissue biopsies from the patient. The tissue is then stained and placed under a microscope that allows the pathologists to histologically assess the microscopic structure and elements of the tissue [3]. The staining process can be done using hematoxylin and eosin (H&E). H&E staining provides permanency of the specimen and helps the pathologists to differentiate the tissue components [4]. Stained tissue from the glass slides can then be digitized by using high resolution image scanners into whole slide images [5], [6]. The histopathological images will therefore be available for long-term storage and further analysis by a computer-aided diagnosis system. Some of the manual techniques such as histological diagnosis, tumor size and axillary lymph node metasis sometimes fail to classify accurately the observed breast tumors [7], [8].

The diagnoses of the histopathological images using computer-aided tools require utilization of machine learning techniques. In the past, classification of such images would require feature engineering techniques to extract features that were supplied to a classical machine learning classifier. However, advances in machine learning such as those observed in deep learning are providing computers with the ability to automatically extract these features. Consequently, the burden of feature engineering especially in medical image analysis is being taken away from humans by computers for more accurate results [9].

Deep learning models have the ability to automatically extract features from high dimensional natural raw images for a suitable internal representation. Multiple levels of representation in deep neural networks permit representational learning, allowing hierarchical feature representations from the non-linear modules that transform the representation at one level starting with the raw input into a representation at a higher more abstract level [10], [11]. The automatically extracted features can then be used by a classifier to recognize images that are supplied to the network.

CNN is one of the deep learning models that can be used in a feature extraction and classification task. In a feature extraction task, the CNN is trained in a fully supervised

* Corresponding Author: Kevin Kiambe
E-mail: kevinkiambe@yahoo.com
† Editors' Choice

setting before using the trained network parameters to automatically extract features from images. The extracted features can then be supplied to a different classifier such as logistic regression (LR), *k*-NN and SVM for a classification task [12], [13].

In this paper, a two stage model for breast histopathology image classification is proposed. In the first stage, A CNN model is trained using histopathological images. In the second stage, the trained model is used to extract features for a classical machine learning classifier. A large number of images are required to effectively train a CNN model. Consequently, data augmentation is applied to the image dataset to artificially increase the number of images in the dataset.

II. Related Studies

Brook et al. [14], proposed a machine learning approach for breast cancer diagnosis using microscopic biopsy images. Generic features and statistical learning algorithms were used to extract features from the images. The extracted features were used to train a Support Vector Machine (SVM) for a 3-class classification task. The histopathological images were classified as healthy tissue (normal), *in situ* or invasive carcinoma breast cancer types. The authors recorded an average error rate between 6.6% with 0.8% standard error of the mean.

Zang [15] presented a 3-class classification of the breast cancer microscopic biopsy image diagnosis. The author combines a Local Binary Pattern (LBP) feature description with Curvelet Transform (CT) for texture analysis in the images. Training and testing were implemented in 2 ensembles. In the first ensemble, the extracted features were supplied to an SVM. In the second ensemble, a Multi-Layer Perceptron (MLP) was used to focus on the rejected samples from the first ensemble. The author was able to achieved high accuracies of up to 97 percent accuracy with a rejection rate of 0.8 percent.

Zhi, W., et al. [16] proposed a CNN architecture by implementing the transfer learning technique for classifying breast histopathology images. This greatly reduced model creation time instead of building one from scratch. The authors were able to compare their performance using other off-the-shelf deep learning models such as VGGNet.

Araújo, T., et al. [3] proposed a hybrid model for the classification of H&E stained breast cancer histopathological images. The images are preprocessed, using optical density colors conversion, before feeding them to the machine learning models. The authors create 70,000 patches from 250 images with each patch labeled with the same class label. A CNN model is used to extract features from breast cancer images. The extracted features are trained using an SVM for classification. The authors achieved accuracies of up to 77.8% in a four-class classification. The images were classified as normal tissue, benign lesion, in situ carcinoma or invasive carcinoma.

Golatkar et al. [17] proposed a deep learning model for classification of the H&E stained breast cancer histopathological images from the ICIAR BACH image dataset. Their model is built through transfer learning techniques where the Inception-v3 CNN model is fine tuned for the classification task. The authors record average accuracies of up to 85% in a four-class – (i) normal tissue, (ii) benign lesion, (iii) *in situ* carcinoma and (iv) invasive carcinoma) – classification task.

III. Dataset

The histopathological images have been obtained from 2018 ICIAR BACH (2018 International Conference on Image Analysis and Recognition Breast Cancer Histology) [18] image dataset. This dataset contains 400 H&E stained histopathological images that are labeled as normal, benign, *in situ* carcinoma or invasive carcinoma (identically, each class contains 100 images). These images were labeled by two pathologists depending on the predominant breast cancer type in each image. The images are released as uncompressed high-resolution of equal dimensions measuring 2040 × 1536 pixels that were acquired under the same conditions. These conditions include a magnification of 200× and pixel size of 0.42μm × 0.42μm [19]. Samples of images from the 2018 ICIAR BACH dataset are shown in Fig. 1.

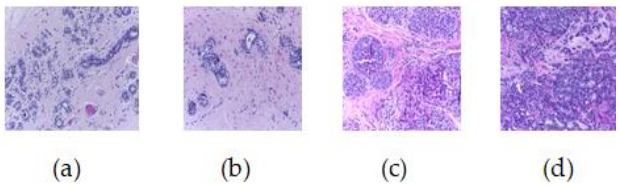

Figure 1. Sample images of (a) normal, (b) benign, (c) *in situ* carcinoma and (d) invasive carcinoma. H&E stained breast cancer histopathological images [18]

IV. Data Augmentation

The insufficient number of images in the medical domain to train accurate and robust CNN models can result to over-fitting. Over-fitting causes a trained network to perform poorly on the test data [20]. However, data augmentation techniques can be used to reduce this over-fitting problem resulting to more accurate and robust CNN models.

This technique provides an artificial way of enlarging the image dataset by generating image translations and altering the intensities of the RGB channels in the original images [21]. Data augmentation can be achieved by techniques such as cropping, rotations, mirroring and scaling.

V. Convolutional Neural Networks

A CNN (or Convnet) is a deep multilayer feed-forward neural network machine learning algorithm that resemble the functioning of a human being's visual cortex. The CNN's architecture is typically made up of one input layer, a feature

extraction layer, a classification layer and an output layer. The feature extraction layer is made up of convolution and pooling layers while the classification layer is typically build using an MLP. The convolution, pooling and fully connected (in the MLP) layers found between the input and output layers are non-linear hidden layers. The deeper the layers, the more the complexity of the learned image properties increase [22].

Deep learning architectures such as the CNN have the ability of sparse representations to uncover semantic information from naturally very high dimensional raw image data. CNN combine three architectural ideas that include local receptive fields, shared weights and spatial sub-sampling. The output of each layer, in the feature extraction module, is a feature map. The feature map is created by convolving the filter matrix by a specified number of steps for the whole input image. A convolutional operation can be represented by equation (1).

$$(I * K)_{xy} = \sum_{i=1}^{h} \sum_{j=1}^{g} K_{ij} \cdot I_{x+i-1,y+j-1} \quad (1)$$

where *I,* is a two-dimensional input image *K,* is a set of filters in the convolutional layer, *h* is the height of the image, *g* is the width of the image. A convolutional operation of an image is represented by *I * K*. If a feature is found, the responsible unit or units generate large activations, whose values can be used later by the classification module [23], [24]. The values in the feature maps are the subjected to an element-wise transfer function, such as the Rectifier Linear Unit (ReLU), hyperbolic tangents and the sigmoid functions, to improve the non-linearity of the decision function. Sub-sampling (pooling) layers are optionally added to feature maps to handle a shift (such as translation or rotation) and scale invariance thus reducing sensitivity of the output and improving generalization of the CNN model.

From 2012, AlexNet [25], Microsoft ResNet [26], VGGNet [27], GoogleNet [28] and Faster RCNN [29] are some of the deep architectures that have presented state-of-the-art performance for several generic computer vision tasks [30]. These network models have also been used in medical image analysis [31], [32]. Techniques such as increasing the model complexity, increasing the training samples and improving training strategies have been proposed to improve the performance of these deep neural network models [33].

These training strategies include stochastic gradient descent with momentum (SGDM), Initialization, Rectifying neurons, batch normalization and dropout techniques. SGDM provides the networks with a good initialization and training mechanism. Good initialization provides better activations and gradients across layers and causes faster convergence during training [34]. Rectifying neurons can cause a deep neural network to reach their best performance on purely supervised tasks [35] while Batch Normalization enables a network to achieve the same accuracy with fewer training steps [36]. Dropout technique works by randomly dropping units from the neural network to prevent the units from too much co-adapting [11].

A. Support Vector Machine

Support Vector Machine (SVM) model was introduced by Vladimir Naumovich Vapnik as a potential alternative to Artificial Neural Networks (ANN) [37]. The SVM is a binary classification method that attempts to establish a decision boundary between any two classes within the input data. An SVM creates a linear decision surface using special features within a high dimensional space that make the model highly generalizable. This classifier is one of the core machine learning methodologies and have been successfully applied to several machine learning classification tasks.

B. k Nearest Neighbor

k-Nearest Neighbor (*k*-NN) is a lazy algorithm that is based on feature similarity. In this case, feature similarity refers to how closely features resemble the training set determines how new data points are classified. *k*-NN can be used for both classification and regression problems. The input of the *k*-NN algorithm is a set of k closest training examples. The output is an object that is classified by a majority vote of its neighbors. Different distance metrics such as the Euclid, Manhattan and Minkowski can be used to calculate the closeness of objects [38].

C. Linear Regression

Linear Regression (LR) model works by constructs mathematical models that describe relationship between variables that exist in the observed data. For example, if two variables are considered, then one variable is considered the explanatory variable while the other is considered a dependent variable. A linear regression line can therefore be constructed using an equation (2).

$$y = a + bx \quad (2)$$

where y is the dependent variable, x is the explanatory variable, *b* is the slope *a* is the y-intercept [39].

VI. EXPERIMENTS

The histopathological images are collected with each image in its own class labeled folder. A lossless image conversion technique is used to convert the large capacity tiff images to jpg images that are more manageable. Clarity of the images is improved to make the edges of the objects in the tissue extract more distinct.

Cropping is done on each image with a 1:1 ratio of 1536 x 1536 pixels at (0, 0), (256, 0) and (512, 0) (x, y) co-ordinates as shown in Fig. 2. This size is large enough to select the regions of interest (ROIs) from the original images. Geometric transformations are carried out by rotating images (at 0^o, 90^o, 180^o and 270^o) and flipping the images both horizontally and vertically.

The images are then resized to 256 x 256 pixels. This results to an artificially enlarged dataset from 400 images in the original dataset to 7200 images that can be used to train the CNN model (identically, each class augmented to contain 1800 images).

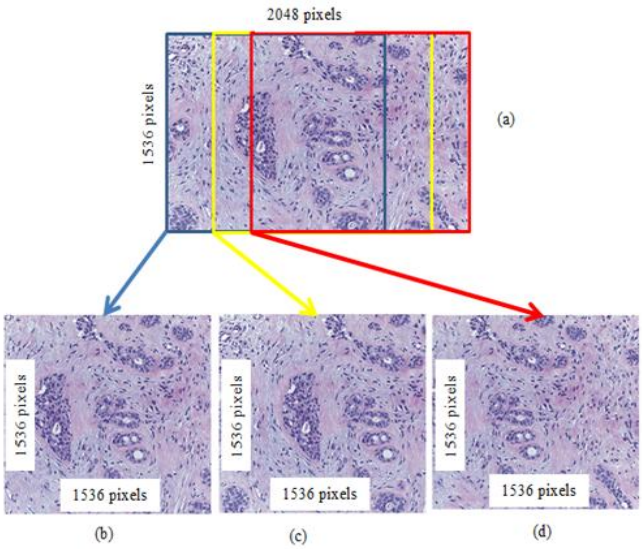

Figure 2. Training progress of the CNN model with patches

VII. The Proposed Architecture

The CNN architecture contains 5 convolutional layers and 4 fully connected layers as shown in Table I. AlexNet's feature extraction module is adopted so as to save time that could have been spent on fine tuning the CNN model. The network is trained using the histopathological images in Matlab R2017b on a Single NVIDIA GTX 680 GPU with compute capability of 3.0.

Table I
Proposed CNN Model for the Classification of Breast Histology Images

Layer	Characteristics
Input	256x256x3
Convolution	96 7x7x3, Stride 3 Padding 0 ReLU
Pooling	Max Pooling 3x3 Stride 2
Convolution	128 5x5x96 Stride 1 Padding 2 ReLU
Convolution	384 3x3x128 Stride 1, Padding 1 ReLU
Convolution	192 3x3x384, Stride 1, Padding 1 ReLU
Convolution	128 3x3x192, Stride 1, Padding 1 ReLU
Pooling	Max Pooling 3x3 Stride 2
Fully Connected	Neurons - 4096 ReLU and Dropout (0.5)
Fully Connected	Neurons - 4096 ReLU and Dropout (0.5)
Fully Connected	Neurons - 4096 ReLU and Dropout (0.5)
Fully Connected	Neurons - 4
Output	4 (Benign, Invasive, *In situ*, Normal)

The CNN model is trained with a minimum batch size of 20, epoch size is set to a maximum of not more than 20 epochs, dropout of 50 percent and learning rate of 0.001. The training progress is shown in the Fig. 3.

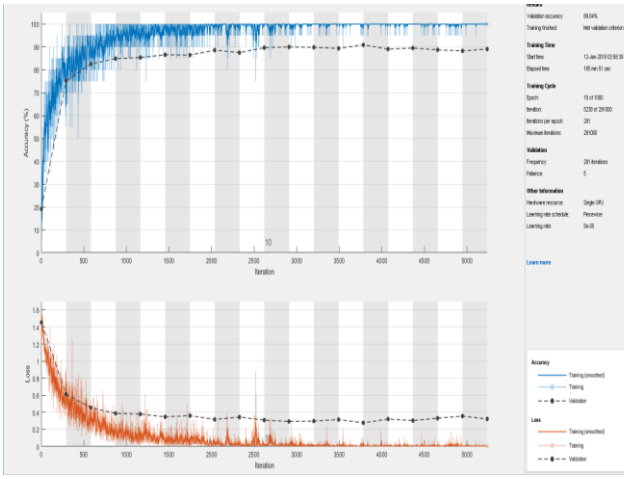

Figure 3. Sample training progress of the CNN model

VIII. Results

4096 image features are automatically extracted using the CNN model at different fully connected feed-forward layers of the MLP. The extracted features from the raw images are used in the second phase for training a classical machine learning classifier. The performance evaluations are recorded as presented in Tables II-XI. The performance of CNN model is evaluated at 0.1, 0.2, 0.3 and 0.4 image dataset training splits. The results are recorded in Tables II-V.

Table II
Confusion Matrix for Model Evaluation at 0.1 Training Split

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9883	0.0099	0.0000	0.0019
	In situ	0.0080	0.9586	0.0185	0.0148
	Invasive	0.0000	0.0000	1.0000	0.0000
	Normal	0.0074	0.0006	0.0031	0.9889

Table III
Confusion Matrix for Model Evaluation at 0.2 Training Split

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9958	0.0028	0.0000	0.0014
	In situ	0.0007	0.9848	0.0146	0.0007
	Invasive	0.0000	0.0021	0.9979	0.0000
	Normal	0.0014	0.0056	0.0028	0.9903

Table IV
Confusion Matrix for Model Evaluation at 0.3 Training Split

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9944	0.0024	0.0000	0.0032
	In situ	0.0008	0.9841	0.0079	0.0071
	Invasive	0.0000	0.0000	1.0000	0.0000
	Normal	0.0000	0.0000	0.0000	1.0000

TABLE V
CONFUSION MATRIX FOR MODEL EVALUATION AT 0.4 TRAINING SPLIT

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9972	0.0028	0.0000	0.0000
	In situ	0.0000	0.9981	0.0019	0.0000
	Invasive	0.0000	0.0074	0.9926	0.0000
	Normal	0.0009	0.0056	0.0009	0.9926

The CNN model is then used to extract features from the images. These features are supplied to the k-NN, SVM and LR classical machine learning models for classification. The results are recorded in Tables VI - VIII.

TABLE VI
CONFUSION MATRIX FOR MODEL EVALUATION AT SPLIT =0.3 USING THE KNN CLASSIFIER

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9992	0.0008	0.0000	0.0000
	In situ	0.0000	0.9984	0.0079	0.0000
	Invasive	0.0000	0.0008	0.9992	0.0000
	Normal	0.0000	0.0016	0.0032	0.9952

TABLE VII
CONFUSION MATRIX FOR MODEL EVALUATION AT SPLIT =0.3 USING THE SVM CLASSIFIER

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9976	0.0016	0.0000	0.0008
	In situ	0.0008	0.9833	0.0151	0.0008
	Invasive	0.0000	0.0008	0.9992	0.0000
	Normal	0.0000	0.0008	0.0024	0.9968

TABLE VIII
CONFUSION MATRIX FOR MODEL EVALUATION AT SPLIT =0.3 USING THE LOGISTIC REGRESSION CLASSIFIER

Split=0.2		Predicted Class			
		Benign	In situ	Invasive	Normal
True Class	Benign	0.9889	0.0087	0.0008	0.0016
	In situ	0.0063	0.9611	0.0167	0.0159
	Invasive	0.0000	0.0000	1.0000	0.0000
	Normal	0.0008	0.0000	0.0024	0.9968

IX. DISCUSSION

Digitization of histopathological images allow for computer aided diagnosis of the breast biopsies [40]. Computer-aided analysis tools are therefore made available to assist the pathology more accurate perform analysis on a larger set of histopathological images.

Deep learning is a machine learning artificial neural network technique that resembles the multi-layered human cognition system [10]. One of the successful deep learning network models is the CNN. CNNs' overall learning process simulate the organization of an human being's visual cortex such that, when successfully trained CNN can compose hierarchical information suitable for tasks such as feature extraction and classification [41].

The images dataset is successfully enlarged to enable accurate training of the CNN model. This is achieved through data augmentation techniques that create a new image set. This is achieved by cropping, rotations, mirroring and scaling of the original images. The newly created image dataset consists of 7200 images that are then randomly split into training and test sets. After training is complete, the performance of the system is measured on the test set of images. This evaluation is used to examine the generalization ability of the system.

In this task, the CNN was used in the extraction of deep features from raw images that were used for training the SVM, *k*-NN and LR classifiers. A sample of features extracted from an image can be graphically represented as shown in Fig. 4.

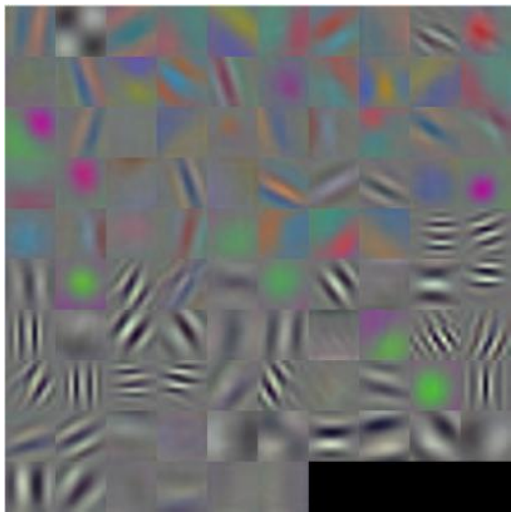

Figure 4. Graphical representation of features extracted from the images using the CNN model

A precision, sensitivity and specificity analysis is performed to examine the models' reliability on the test data. Precision analysis test whether the models produce what it reliable measures. This measure is sometimes referred to as the positive predictive value (PPV) and can be calculated using equation (3).

$$PPV = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (3)$$

Sensitivity also known as true positive rate (TPR) is a measure of actual positives that are correctly tested by a certain model. Sensitivity can be calculated using equation (4).

$$TPR = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (4)$$

Specificity is a measure that enables a model to correctly exclude the negatives from test data. Specificity is also known as the true negative rate (TNR) and can be calculated using equation (5).

$$TNR = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \quad (5)$$

The PPV, TPR and TNR analysis is carried out to compare the performance at 0.1, 0.2, 0.3 and 0.4 training splits of the model. These results are recorded in Fig. 5 - Fig. 8.

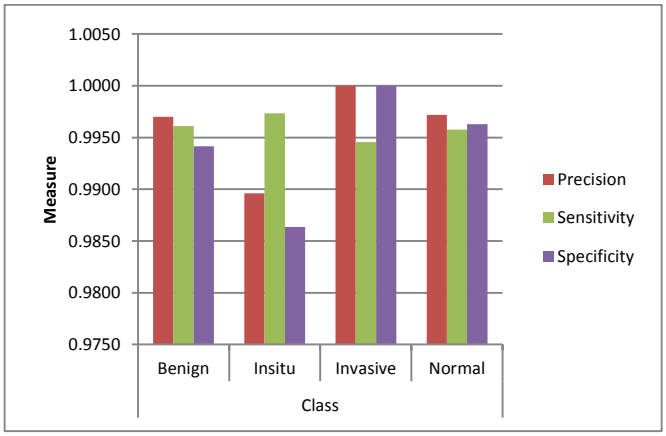

Figure 5. Precision, sensitivity and specificity analysis of the CNN model at 0.1 training split

From Fig. 5, it is observed that the highest precision and specificity values are recorded in the invasive class. The highest sensitivity measure is recorded in the *in situ* class at a training split of 0.1.

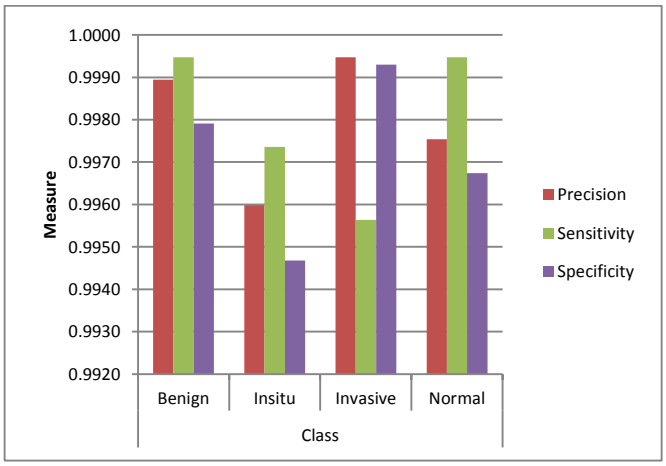

Figure 6. Precision, sensitivity and specificity analysis of the CNN model at 0.2 training split

From Fig. 6, it is observed that the highest precision and specificity values are recorded in the invasive class. The highest sensitivity measure is recorded in the normal class at a training split of 0.2.

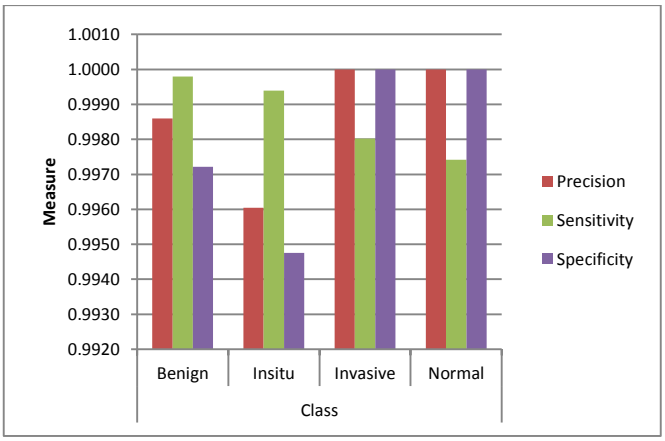

Figure 7. Precision, sensitivity and specificity analysis of the CNN model at 0.3 training split

From Fig. 7, it is observed that the highest precision and specificity values are recorded in the invasive and normal classes. The highest sensitivity measure is recorded in the benign class at a training split of 0.3.

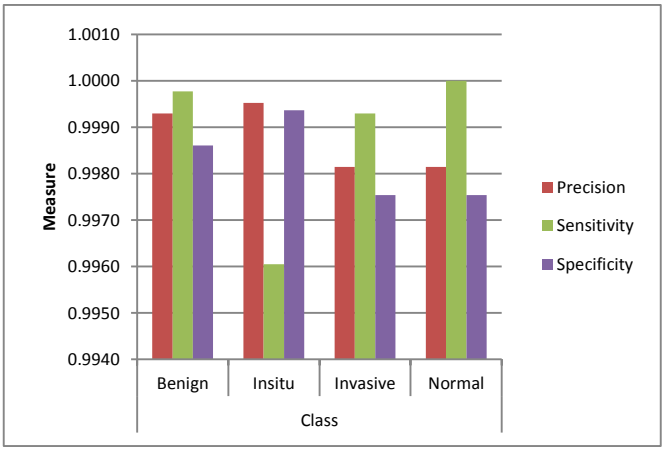

Figure 8. Precision, sensitivity and specificity analysis of the CNN model at 0.3 training split

From Fig. 8, it is observed that the highest precision and specificity values are recorded in the *in situ* class. The highest sensitivity measure is recorded in the normal class at a training split of 0.4.

The PPV, TPR and TNR analysis is further carried out to compare the performance of the classical machine learning algorithms. The results are recorded in Fig. 9 – Fig. 11.

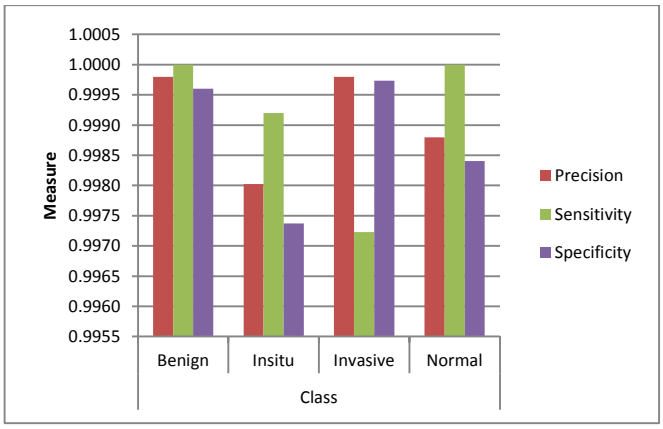

Figure 9. Precision, sensitivity and specificity analysis of the CNN model at *k*-NN classifier

From Fig. 9, it is observed that the *k*-NN classifier presents the highest precision, sensitivity and specificity values in the benign and invasive classes.

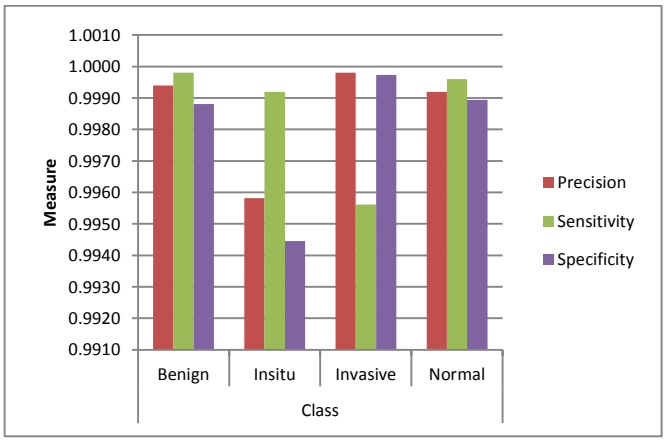

Figure 10. Precision, sensitivity and specificity analysis of the CNN model at SVM classifier

From Fig. 10, it is observed that the SVM classifier presents the highest precision and specificity values are recorded in the invasive class. The highest sensitivity measure is recorded in the benign class.

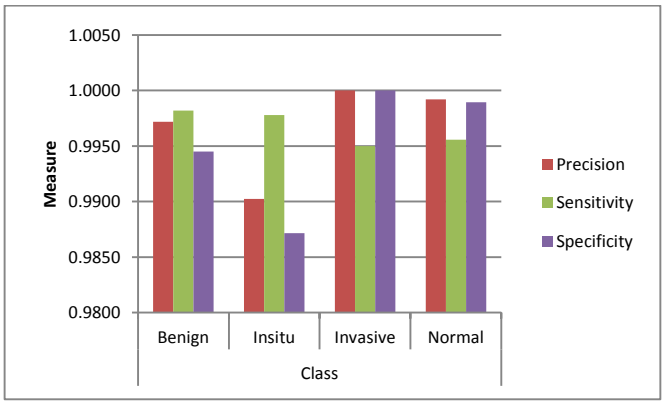

Figure 11. Precision, sensitivity and specificity analysis of the CNN model at LR classifier

From Fig. 11, it is observed that the LR classifier presents the highest precision and specificity values are recorded in the invasive class. The highest sensitivity measure is recorded in the benign class. The PPV, TPR and TNR values are summarized in Table IX.

TABLE IX
K-NN, SVM AND LR CLASSIFIERS' PPV, TPR AND TNR VALUES

		Class			
		Benign	*In situ*	Invasive	Normal
k-NN	PPV	0.9998	0.9980	0.9998	0.9988
	TPR	1.0000	0.9992	0.9972	1.0000
	TNR	0.9996	0.9974	0.9997	0.9984
SVM	PPV	0.9994	0.9958	0.9998	0.9992
	TPR	0.9998	0.9992	0.9956	0.9996
	TNR	0.9988	0.9945	0.9997	0.9989
LR	PPV	0.9972	0.9902	1.0000	0.9992
	TPR	0.9982	0.9978	0.9950	0.9956
	TNR	0.9945	0.9872	1.0000	0.9989

From Table IX it is observed that highest PPV value is recorded by the LR classifier for the invasive class. The highest TPR value is recorded by the *k*-NN classifier in the benign and normal tissue classes. The highest TNR value is recorded by the LR classifier in the invasive tissue class.

The performance of the *k*-NN, SVM and LR models are compared to examine which models measures accurately on the test data. The results are recorded in Table X.

TABLE X
PERFORMANCE COMPARISON OF THE SVM, LR AND *K*-NN CLASSIFIERS

Classifier Algorithm	Accuracy (%)	PPV	TPR	TNR
SVM	**99.84**	**0.9991**	**0.9991**	**0.9988**
LR	98.67	0.9986	0.9986	0.9980
k-NN	99.80	0.9967	0.9967	0.9952

From the results in Table X it is observed that the SVM classical machine learning model classifies the images more accurately than the *k*-NN and LR models with accuracies of up to 99.84. The *k*-NN algorithm and LR model achieves testing accuracies of up to 99.80% and 98.67%, respectively, and all of them are better than the CNN itself with 4 fully-connected layers where the achieved accuracy is up to 99.51%.

X. CONCLUSION

In conclusion, digital pathology now enables automation of a pathologist's tasks by utilizing technologies in artificial intelligence. In this paper, a novel deep learning CNN model for automatic feature extraction and classification H&E stained breast histology images is proposed. Our results yield evidence that a hybrid model can accurately detect the breast cancer digital breast histology microscopy images. These results have been successfully obtained by applying data augmentation techniques to enlarge the dataset for better automatic feature extraction. The model performs best when features are extracted from the third fully connected feed-

forward deep network layer. The features are then supplied to the classical machine learning classifier that records accuracy of up to 99.84%.

ACKNOWLEDGMENT

The author would like to thank Asst. Prof. Ahmet Babalik at the computer engineering department in Selçuk University and Selçuk University computer research center (BİLMER) for their technical support and comments.

REFERENCES

[1] WHO. "Cancer: Breast Cancer," 7/12/2017, 2017; http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/.

[2] A. C. Society. "Breast Cancer Facts & Figures 2017-2018," 7/12/2017, 2017; https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2017-2018.pdf.

[3] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using Convolutional Neural Networks," *PLOS ONE*, vol. 12, no. 6, pp. e0177544, 2017.

[4] R. Dougall Lillie, *Histopathologic technic and practical histochemistry / R. D. Lillie, Harold M. Fullmer*, 2017.

[5] D. Hartman, L. Pantanowitz, J. McHugh, A. Piccoli, M. OLeary, and G. Lauro, "Enterprise Implementation of Digital Pathology: Feasibility, Challenges, and Opportunities," *Journal of Digital Imaging*, vol. 30, no. 5, pp. 555-560, October 01, 2017.

[6] S. Al-Janabi, A. Huisman, S. M. Willems, and P. J. Van Diest, "Digital slide images for primary diagnostics in breast pathology: a feasibility study," *Human Pathology*, vol. 43, no. 12, pp. 2318-2325, 2012/12/01/, 2012.

[7] L. J. Van't Veer, H. Dai, M. J. Van De Vijver, Y. D. He, A. A. Hart, M. Mao, H. L. Peterse, K. Van Der Kooy, M. J. Marton, and A. T. Witteveen, "Gene expression profiling predicts clinical outcome of breast cancer," *nature*, vol. 415, no. 6871, pp. 530-536, 2002.

[8] W. L. McGuire, "Breast cancer prognostic factors: evaluation guidelines," Oxford University Press, 1991.

[9] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 221-248, 2017.

[10] Y. LeCun, Y. Bengio, and G. Hinton, *Deep Learning*, 2015.

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

[12] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition." pp. 647-655.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks." pp. 1097-1105.

[14] A. Brook, R. El-Yaniv, E. Isler, R. Kimmel, R. Meir, and D. Peleg, *Breast Cancer Diagnosis From Biopsy Images Using Generic Features and SVMs*, 2007.

[15] B. Zhang, "Breast cancer diagnosis from biopsy images by serial fusion of Random Subspace ensembles." pp. 180-186.

[16] W. Zhi, H. W. F. Yueng, Z. Chen, S. M. Zandavi, Z. Lu, and Y. Y. Chung, "Using Transfer Learning with Convolutional Neural Networks to Diagnose Breast Cancer from Histopathological Images." pp. 669-676.

[17] A. Golatkar, D. Anand, and A. Sethi, "Classification of Breast Cancer Histology Using Deep Learning," *Image Analysis and Recognition*. pp. 837-844.

[18] C. f. O. M. I. Computing, "ICIAR 2018 Grand Challenge on Breast Cancer Histology Images," Consortium for Open Medical Image Computing, 2017.

[19] T. Araújo, G. Aresta, C. Eloy, A. Polónia, and P. Aguiar, "ICIAR 2018 Grand Challenge on Breast Cancer Histology Images," Consortium for Open Medical Image Computing, 2017.

[20] A. Fawzi, H. Samulowitz, D. Turaga, and P. Frossard, "Adaptive data augmentation for image classification." pp. 3688-3692.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, Lake Tahoe, Nevada, 2012, pp. 1097-1105.

[22] W. Zhao, and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544-4554, 2016.

[23] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.

[24] Y. LeCun, and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, pp. 1995, 1995.

[25] A. Krizhevsky, *Learning Multiple Layers of Features from Tiny Images*, 2012.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition." pp. 770-778.

[27] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions." pp. 1-9.

[29] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks." pp. 91-99.

[30] A. Karpathy, and L. Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 664-676, 2017.

[31] Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, S. Zhang, D. N. Metaxas, and X. S. Zhou, "Multi-Instance Deep Learning: Discover Discriminative Local Anatomies for Bodypart Recognition," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1332-1343, 2016.

[32] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. S. Kruthiventi, and R. V. Babu, "Chapter 2 - An Introduction to Deep Convolutional Neural Nets for Computer Vision," *Deep Learning for Medical Image Analysis*, pp. 25-52: Academic Press, 2017.

[33] W. Shi, Y. Gong, X. Tao, J. Wang, and N. Zheng, "Improving CNN Performance Accuracies With Min-Max Objective," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, no. 99, pp. 1-14, 2017.

[34] X. Glorot, and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, 2010, pp. 249--256.

[35] X. Glorot, A. Bordes, and Y. Bengio, "Deep Sparse Rectifier Neural Networks," in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, 2011, pp. 315--323.

[36] S. Ioffe, and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in Proceedings of the 32nd International Conference on Machine Learning, Proceedings of Machine Learning Research, 2015, pp. 448--456.

[37] E. Byvatov, U. Fechner, J. Sadowski, and G. Schneider, "Comparison of support vector machine and artificial neural network systems for drug/nondrug classification," *Journal of chemical information and computer sciences*, vol. 43, no. 6, pp. 1882-1889, 2003.

[38] A. S. Arefin, C. Riveros, R. Berretta, and P. Moscato, "Gpu-fs-knn: A software tool for fast and scalable knn computation using gpus," *PloS one*, vol. 7, no. 8, pp. e44000, 2012.

[39] G. A. Seber, and A. J. Lee, *Linear regression analysis*: John Wiley & Sons, 2012.

[40] V. Della Mea, and R. Mencarelli, "2012 – The beginning of a new world for Digital Pathology," *Diagnostic Pathology*, vol. 8, no. 1, pp. S1, September 30, 2013.

[41] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, "Deep Learning in Medical Imaging: General Overview," *Korean Journal of Radiology*, vol. 18, no. 4, pp. 570-584, 2017.

Kevin Kiambe was born in Siaya, Kenya in 1990. He received the BSc degree in computer science from the University of Kabianga (Formerly Moi University – Kabianga Campus), Kericho, Kenya, in 2013. He is currently working towards an MSc degree in computer engineering, Selçuk University, Konya, Turkey. His research interests include artificial intelligence, embedded systems, computer vision and deep learning.