Transfer learning using Alexnet with Support Vector Machine for Breast Cancer Detection

Sema Nizam Abdulghani¹*, Ahmed Freidoon Fadhil¹, Seyfettin Sinan Gültekin²

¹ University of Kirkuk, Faculty of Engineering, Department of Electrical Engineering, Kirkuk, Iraq
² Konya Technical University, Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Konya, Turkey

(1st International Conference on Computer, Electrical and Electronic Sciences ICCEES 2020 – 8-10 October 2020)

DOI: 10.31590/ejosat.806679

ATIF/REFERENCE: Nizam Abdulghani, S., Fadhil, A. F. & Gültekin, S. S. (2020). Transfer learning using Alexnet with Support Vector Machine for Breast Cancer Detection. European Journal of Science and Technology, (Special Issue), 423-430.

Abstract

Breast cancer is one of the leading causes of women death worldwide currently. Developing a computer-aided diagnosis system for breast cancer detection became an interesting problem for many researchers in recent years. Researchers focused on deep learning techniques for classification problems, including Convolutional Neural Networks (CNNs), which achieved great success. CNN is a particular type of deep, feedforward network that has gained attention from the research community and achieved great successes, especially in biomedical image processing. In this paper, transfer learning and deep feature extraction methods are used which adapt a pre-trained CNN model to classify breast cancer histopathological images from the publically available (BreakHis dataset). The data set includes both benign and malignant images with four different magnification factors. A patch strategy method proposed based on the extraction of image patches for training the CNN and the combination of these patches for final classification. AlexNet model is considered in this work with patch strategy, and pre-trained AlexNet is used for further fine-tuning. The obtained features are then classified by using support vector machines (SVM). The evaluation results show that the pre-trained Alexnet with SVM classification and patch strategy yields the best accuracy. Accuracy between 92% and 96% was achieved using five-fold cross-validation technique for different magnification factors.

Keywords: Breast Cancer; Convolutional Neural Network; Alexnet; Transfer Learning; and Support Vector Machine

Meme Kanseri Tespi t İçin Destek Vektör Makinesi ile Alexnet Kullanarak Transfer Öğrenimi

Öz

Meme kanseri, şu anda dünya çapında kadın ölümüün ön happy nedenlerinden biridir. Meme kanseri teşhis için bilgisayar destekli teşhis sistemlerini geliştirme, son yıllarda birçok araştırmacı için ilgi çekici bir sorun haline geldi. Araştırmaçılık, büyük bir başara elde eden Evrişimli Sinir Ağları (CNN'ler) dâhil olmak üzere sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüşü ile uygulamaları da dahil olmak üzere sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntülerin işlenmesi ve sınıflandırma problemlerini için derin Öğrenme tekniklerine odaklandı. CNN'ler, özellikle biyomedikal görüntül...
1. Introduction

Nowadays, cancer is a massive public health problem around the world. According to the International Agency for Research on Cancer (IARC), there will be 27 million deaths caused by cancer are expected to occur until 2030 (Boyle and Levin 2008). Breast cancer (BC) is the second most common for women, among the cancer types. Besides, the mortality of BC is very high when compared to other types of cancer. Despite the recent advances in molecular biology of Breast Cancer progression, the histopathological analysis remains the most widely used method for BC diagnosis (Lakhani, Ellis et al. 2012). Although significant progress reached by diagnostic imaging technologies, the final BC diagnosis, including grading and staging, continues being done by pathologists applying visual inspection of histological samples under the microscope.

Recent advances in machine learning techniques allow building Computer-Aided Detection/Diagnosis (CAD) systems that can assist pathologists to be more productive, objective and consistent in diagnosis. The main challenge of such systems is dealing with the inherent complexity of histopathological images. The automatic imaging processing for cancer diagnosis has been explored as a topic of research for more than 40 years (Stenkvist, Westman-Naeser et al. 1978). However, it is still challenging due to the complexity of the images to analyze.

In literature, most of the works on BC histopathology image analysis were carried out on small datasets until. (Spanhol, Oliveira et al. 2016) introduced a dataset composed of 7,909 breast histopathological images acquired on 82 patients. The authors evaluated six different textural descriptors and different classifiers and reported an accuracy ranging from 80% to 85.

Among the different approaches, the Convolutional Neural Network (CNN) has been widely used to achieve state-of-the-art results in different pattern recognition problems (Krizhevsky, Sutskever et al. 2012, Niu and Suen 2012, Hafemann, Oliveira et al. 2014) have shown, for images of microscopic and macroscopic texture, that CNN can surpass traditional textural descriptors. In (Spanhol, Oliveira et al. 2016), the authors have evaluated the deep learning approach by combining different CNNs using simple fusion rules and achieved an improvement in classification accuracy. The authors in (Spanhol, Oliveira et al. 2017) mentioned that previously trained CNN reuse is used as a feature vector, and DeCAF features are extracted. Then, the DeCAF feature is used as an input to the classifier trained for the new classification task. It achieved an average of 84% accuracy on breast cancer case images.

In (Deniz, Şengür et al. 2018), transfer learning and pre-trained CNN AlexNet and VGG16 models are considered for feature extraction and classified by support vector machines (SVM). An accuracy ranging from 90.5% to 91.4% at image levels were achieved. Five pre-trained Deep Convolutional Neural Network (DCNN) architectures are employed as feature extractors, namely InceptionV3, Inception, ResNetV2, Xception and two VGG Net models were studied in (Kassani, Kassani et al. 2019). The authors used employed different data augmentation techniques to boost the performance of classification. The pre-trained Xception model yields the best average classification accuracy of 92.50% among all the other DCNN models. Other deep learning convolutional neural networks like Resnet, Alexnet, and VGG16 were employed in (Wenzhong, Huanlan et al. 2020) on breakHis data set. Also, a novel classification method DeepBC was proposed for classifying the pathological images. The evaluation results showed that DeepBC achieved 92% and 96.43% accuracy rates in classifying patients and images, respectively.

Currently, automatic classifications of histopathological images for breast cancer still a challenging problem due to the sudden increase of CNN parameters which leads to over-fitting of the system. Although increasing the data set of images overcomes the over-fitting problem, it increases the complexity of the system and consumes time. In this paper, a CAD system based on deep convolutional neural networks (CNN) is developed to help the pathologists classify the histopathological breast cancer images as benign and malignant. At first, the Alexnet CNN architecture was used and trained from scratch. Then, the pre-trained neural network structure of Alexnet was tested using transfer learning. Finally, Deep feature extraction was done for classification using the pre-trained AlexNet network. The deep features that are considered in this work were extracted from the fully connected (FC) layers of the pre-trained AlexNet model. The FC6 and FC7 layers are considered in this paper, which produces 4096-dimensional feature vectors. Then, the feature vectors were used with SVM classifier to augment the efficiency of the proposed work.

The remaining of the paper will be organized as follows: Section 2 explains the theory and structure of Alexnet, Transfer Learning, SVM, and Cross-validation. In section 3, the Transfer Learning Using Alexnet with SVM classifier model was proposed and explained in detail. Section 4 presents experimental results and section 5 discussion and analysis. Finally, the conclusion of the presented work made in section 6.

2. Theory and Structure

2.1 The Alexnet

Alexnet is a convolutional neural network designed by (Krizhevsky, Sutskever et al. 2012) which contains eight layers: five convolutional layers and three fully-connected layers. The original structure is designed to classify 1000 class labels. The first convolutional layer filter is fed by (227 x 227 x 3) input image with 96 kernels of size (11 x 11 x 3) and stride of 4 pixels. The second convolutional layer takes as input the response-normalized and pooled output of the first convolutional layer and filters it with 256 kernels of size (5x5x48). The third, fourth, and fifth convolutional layers are connected without any intervening pooling or normalization layers. The third convolutional layer has 384 kernels of size (3x3x256) connected to the (normalized, pooled) outputs.
of the second convolutional layer. The fourth convolutional layer has 384 kernels of size (3x3x192), and the fifth convolutional layer has 256 kernels of size (3x3x192). The fully-connected layers have 4096 neurons each (Krizhevsky, Sutskever et al. 2012).

2.2 Transfer Learning

Transfer learning is a deep learning approach in which a model that has been trained for one task is used as a starting point to train a model for a similar task. Fine-tuning a network with transfer learning is usually much faster and easier than training a network from scratch. The approach is commonly used for object detection, image recognition, speech recognition, and other applications (Abd Almisreb, Jamil et al. 2018). Facing the problem of collecting enough training data, transfer learning aims to transfer knowledge from a large dataset known as source domain to a smaller dataset named target domain. In the field of medical images processing where the data-poor exists as well, transfer learning is an effective method when employing CNNs to medical image classification with the help of sufficient annotated natural images (Huang, Pan et al. 2017).

2.3 The Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a classification technique which tries to find an effective separable hyperplane to separate two-class vectors (features from both classes) by maximizing the separable distance between that two class vectors (Duda, Hart et al. 2012). The classifier can be linear or non-linear.

Given a training set of the form \((x_i, y_i), \ i = 1, 2, 3, ..., n\), for dimension \(d\), where \(x_i \in R^d\) and \(y_i \in \{1, -1\}\), and \(x_i\) are the feature vectors and \(y_i\) are the two classes. For linear separable hyperplane, the kernel function will be

\[
K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle
\]

For non-linear separable hyperplane, the kernel function will be different. An example is the radial basis function

\[
K(x_i, x_j) = \exp(- \frac{\langle x_i-x_j \rangle, (x_i - x_j) \rangle}{2\sigma^2})
\]

where \(\sigma\) is a positive integer. In general, the class will be determined by using the equation:

\[
\text{class}(z) = \text{sign}(\sum_{i=1}^{N_s} \alpha_i y_i K(S_i, z) + b)
\]

where \(S_i\) are training instances \(z_i\) (support vectors) with \(\alpha_i > 0\) and \(N_s\) is the number of support vectors (Fadhil 2014).

2.4 Cross-validation Method

Cross-validation is a statistical method used to estimate the skill of machine learning models. Cross-validation is similar to the repeated random subsampling method. However, the sampling is done in such a way that no two test sets overlap (Berrar and Biology 2019). In k-fold cross-validation, the available learning set is partitioned into k disjoint subsets of approximately equal size. Here, fold refers to the number of resulting subsets. This partitioning is performed by randomly sampling cases from the learning set without replacement. The model is trained using k-1 subsets, which, together, represent the training set. Then, the model is applied to the remaining subset, which is denoted as the validation set to measure the performance. This procedure is repeated until each of the k subsets has served as the validation set. The average of the k performance measurements on the k validation sets is the cross-validated performance (Berrar and Biology 2019).

3. The proposed Systems

3.1 Transfer Learning Using Alexnet

AlexNet has been trained on approximately 1.2 million images from the ImageNet Dataset (http://image-net.org/index). The model has eight layers and can classify images into 1000 object (such as a keyboard, coffee mug, pencil, and many animals). As a result, the network has learned rich feature representations for a wide range of images. The earlier layers of the pre-trained AlexNet CNN are retained as a fixed feature extractor for the BreakHis dataset, while the last three layers were replaced with a set of layers that can classify two classes only (Abd Alnisreb, Jamil et al. 2018). The Stochastic Gradient Descent (SGD) method (Bottou 2012) with backpropagation was used to update the network’s parameters. A complete pass of the algorithm over the entire training set is
called an epoch. The mini-batch size is the subset of the training dataset used by the SGD to update network parameters. In contrast, the rate of adjusting the weights of the network to the gradient is called the learning rate.

3.2 Transfer Learning Using Alexnet with SVM classifier

Deep feature extraction can be seen as another type of transfer learning. Instead of fine-tuning a pre-trained CNN model, the activation layers of the CNN model can be used to extract representative feature vectors. The activations of the earlier layers provide representations similar to low-level image features such as edges. In comparison, the deeper layers provide higher-level features salient for image classification. For example, in both ImageNets, the activations of the first and second fully connected layers (FC6 and FC7) are commonly used as a feature representation for image recognition tasks. The resulting feature vectors contain 4096 attributes. These feature vectors are classified by SVM classifier to determine the class label of the input images (Deniz, Şengür et al. 2018).

4. Results

In literature (Spanhol, Oliveira et al. 2016, Spanhol, Oliveira et al. 2017, Deniz, Şengür et al. 2018, Kassani, Kassani et al. 2019, Wenzhong, Huanlan et al. 2020), the recognition rate presented by the researches was evaluated at the image or the patient level. At the patient level, the patient score needs to be calculated first by:

\[ \text{Patient Score} = \frac{N_{\text{rec}}}{N_P} \quad (4) \]

and the global patient recognition rate as:

\[ \text{Patient Recognition Rate} = \frac{\sum \text{Patient Score}}{\text{Total Number of Patients}} \quad (5) \]

where \( N_P \) is the number of cancer images of patient \( P \), and \( N_{\text{rec}} \) is the correctly classified cancer images for each patient. On the other hand, the image level provides the simple image classification accuracy of the CNN. The recognition rate is computed at the image level by:

\[ \text{Image Recognition Rate} = \frac{N_{\text{rec}}}{N_{\text{all}}} \quad (6) \]

where \( N_{\text{all}} \) is the number of cancer images of the test set, and \( N_{\text{rec}} \) is the number of correctly classified cancer images.

The experiment platform is configured using MacBook Pro with an Intel Core i7-7820 CPU and 32 GB memory. The BreaKHis database (Spanhol, Oliveira et al. 2016) contains 7,909 microscopic biopsy images of benign and malignant breast tumors. All images are colored and of size 700 \times 460 pixels. The images are collected from 82 patients with magnifying factors of 40×, 100×, 200× and 400×. Experiments were evaluated using five-fold cross-validation at the patient level. The dataset has been divided into five splits for cross-validation, and each split contains 80% of images as training and 20% of images as testing sets. The dataset split patient wise to guarantee that the classifier generalizes to unseen patients. Previously, all input images were initially resized to size 227 x 227 for the sake of convenience with Alexnet structure (Deniz, Şengür et al. 2018, Wenzhong, Huanlan et al. 2020). Since the original size of images of BreaKHis dataset is 700 x 460, the proposed method suggests extracting image patches of size 227 x 227 from the original images instead of resizing it. As a result, six different patch images of size 227 x 227 were extracted from each image of the BreakHis dataset.

The initial experiments were carried out using Alexnet CNN architecture by training it from scratch. Table 1 shows the results of training the systems using a mini-batch size of (10) and the initial learning rate of \( (10^{-2}) \). The initial learning rate was chosen large enough to increase the learning of the network from scratch. The maximum epoch number was set to 5, and the CNN model was trained by stochastic gradient descent with momentum. Table 1 shows the average accuracy results of Alexnet model for all magnification factors using five-fold cross-validation method.
Table 1: The Average Accuracy for Alexnet Algorithm

| MF  | Recognition Rate |        |        |
|-----|------------------|--------|--------|
|     |                  | Image  | Patient|
| 40X | 68.02            | 69.98  |        |
| 100X| 69.44            | 69.98  |        |
| 200X| 69.44            | 69.98  |        |
| 400X| 68.74            | 69.98  |        |

In Table 2 and Table 3, different learning rate values were tested using a maximum of 5 epochs and a mini-batch size of 100. Table 2 presents the average accuracy results for Pre-trained Alexnet method using different initial learning rate. In contrast, Table 3 presents the average accuracy results for Pre-trained Alexnet with SVM method using different initial learning rate.

Table 2: The Average Accuracy for Pre-trained Alexnet method (different learning rate)

| MF  | Learning Rate | Recognition Rate |        |        |
|-----|---------------|------------------|--------|--------|
|     |               |                  | Image  | Patient|
| 40X | $10^{-2}$     | 68.02            | 69.98  |        |
|     | $10^{-3}$     | 86.94            | 89.95  |        |
|     | $10^{-4}$     | **89.30**        | **92.04** |        |
|     | $10^{-5}$     | 85.32            | 87.99  |        |
| 100X| $10^{-2}$     | 69.44            | 69.98  |        |
|     | $10^{-3}$     | 90.52            | 94.00  |        |
|     | $10^{-4}$     | **91.84**        | **90.07** |        |
|     | $10^{-5}$     | 90.13            | 89.95  |        |
| 200X| $10^{-2}$     | 68.74            | 69.98  |        |
|     | $10^{-3}$     | 88.95            | 87.99  |        |
|     | $10^{-4}$     | **91.68**        | **92.04** |        |
|     | $10^{-5}$     | 90.07            | 92.04  |        |
| 400X| $10^{-2}$     | 68.37            | 69.98  |        |
|     | $10^{-3}$     | 86.96            | 90.07  |        |
|     | $10^{-4}$     | **88.92**        | **92.04** |        |
|     | $10^{-5}$     | 88.43            | 89.95  |        |

Table 3: The Average Accuracy for Pre-trained Alexnet with SVM method (different learning rate)

| MF  | Learning Rate | Recognition Rate |        |        |
|-----|---------------|------------------|--------|--------|
|     |               |                  | Image  | Patient|
| 40X | $10^{-2}$     | 68.02            | 69.98  |        |
|     | $10^{-3}$     | 88.79            | 95.96  |        |
|     | $10^{-4}$     | **90.36**        | **92.04** |        |
|     | $10^{-5}$     | 85.69            | 87.99  |        |
| 100X| $10^{-2}$     | 69.44            | 69.98  |        |
|     | $10^{-3}$     | 91.04            | 94.00  |        |
|     | $10^{-4}$     | **91.08**        | **92.04** |        |
|     | $10^{-5}$     | 90.81            | 94.00  |        |
| 200X| $10^{-2}$     | 68.74            | 69.98  |        |
|     | $10^{-3}$     | 90.73            | 92.04  |        |
|     | $10^{-4}$     | **91.82**        | **92.04** |        |
|     | $10^{-5}$     | 90.31            | 92.04  |        |
| 400X| $10^{-2}$     | 68.37            | 69.98  |        |
|     | $10^{-3}$     | 85.66            | 84.19  |        |
|     | $10^{-4}$     | **89.92**        | **94.00** |        |
|     | $10^{-5}$     | 88.30            | 89.95  |        |
Next, the experiments were performed using the Stochastic Gradient Descent (SGD) method with a mini-batch size of 100, a learning rate of \((10^{-4})\), momentum term of 0.9, a weight decay of \((4^{-5})\), and at a maximum of 20 epochs. Table 4 reports the best average accuracy of the proposed pre-trained Alexnet method at image and patient levels for 40X, 100X, 200X and 400X magnification factors using five-fold cross-validation. While Table 5 reports the best average accuracy of the proposed pre-trained Alexnet with SVM method at image and patient levels for 40X, 100X, 200X and 400X magnification factors using five-fold cross-validation.

### Table 4: The Average Accuracy for Pre-trained Alexnet method \( (\text{learning rate} = 10^{-4}) \)

| MF  | Max Epoch | Recognition Rate |   |   |
|-----|------------|------------------|---|---|
|     |            | Image            | Patient |   |
| 40X | 5          | 89.30            | 92.04    |   |
|     | 10         | 90.78            | 94.00    |   |
|     | 15         | 91.05            | 92.04    |   |
|     | 20         | **91.74**        | **92.04**|   |
| 100X| 5          | 91.84            | 90.07    |   |
|     | 10         | **91.92**        | **92.03**|   |
|     | 15         | 91.77            | 95.96    |   |
|     | 20         | 90.95            | 94.00    |   |
| 200X| 5          | 91.68            | 92.04    |   |
|     | 10         | 90.42            | 89.95    |   |
|     | 15         | 91.07            | 94.00    |   |
|     | 20         | **91.68**        | **92.04**|   |
| 400X| 5          | 89.92            | 94.00    |   |
|     | 10         | **90.24**        | **94.12**|   |
|     | 15         | 89.65            | 94.12    |   |
|     | 20         | 89.95            | 94.12    |   |

### Table 5: The Average Accuracy for Pre-trained Alexnet with SVM method \( (\text{learning rate} = 10^{-4}) \)

| MF  | Max Epoch | Recognition Rate |   |   |
|-----|------------|------------------|---|---|
|     |            | Image            | Patient |   |
| 40X | 5          | 90.36            | 92.04    |   |
|     | 10         | 92.57            | 94.00    |   |
|     | 15         | 91.97            | 95.96    |   |
|     | 20         | **92.05**        | **94.00**|   |
| 100X| 5          | 91.08            | 92.04    |   |
|     | 10         | **91.90**        | **94.00**|   |
|     | 15         | 91.70            | 90.08    |   |
|     | 20         | 91.14            | 92.04    |   |
| 200X| 5          | 91.82            | 92.04    |   |
|     | 10         | 91.71            | 92.04    |   |
|     | 15         | **91.96**        | **92.04**|   |
|     | 20         | 91.75            | 89.95    |   |
| 400X| 5          | 89.92            | 94.00    |   |
|     | 10         | **90.62**        | **96.08**|   |
|     | 15         | 89.95            | 92.04    |   |
|     | 20         | 89.41            | 94.12    |   |

### 5. Discussion

The results from Table 1 show that the accuracy of using Alexnet with training from scratch yields poor results. In-depth investigation of the results shows that the systems always classify the data as a malignant class. The reason for the system failure is that the data was not sufficient for building a robust system. To overcome this problem, many hundred epochs should be used with GPU support to reach better results, as in (Kassani, Kassani et al. 2019, Wenzhong, Huanlan et al. 2020). In this paper, a patch strategy will be used, resulting in six patches per image from the data set. Also, instead of training the CNN from scratch, transfer learning will be used, which makes the system converge faster. Then, the Support Vector Machine (SVM) will be used in the final layers of the Alexnet structure to improve the classification process.
The learning rate is a vital hyper-parameter when configuring the neural network. Therefore, it is essential to investigate the effects of the learning rate on model performance. Most of the previous studies showed that changing the learning rate affects system performance. For this reason, the proposed systems were trained using different learning rates to select the better initial learning rate. Results from Table 2 and Table 3 indicate that using an initial learning rate of \(10^{-4}\) yields better accuracy results. These results are reasonable since using a higher learning rate may lead to divergence, while using smaller values will slow down learning in the transferred layers and converges faster.

As seen in Table 4 and 5, the best classification accuracy of 92.05% at image level and 94% at patient level was produced for 40X magnification factor using pre-trained Alexnet with SVM with maximum epochs of 20. The accuracy score for 100X magnification factor using pre-trained Alexnet with SVM was 91.90% at image level and 94% at patient level with only 10 epochs which is better than the pre-trained Alexnet method. The other magnification factor 200X produced similar accuracy at the patient level for both proposed methods, and the classification accuracy at image levels was better for pre-trained Alexnet with SVM compared to the pre-trained Alexnet. Additionally, the accuracy score for 400X magnification factor using pre-trained Alexnet with SVM was 90.62% at image level and 96% at patient level with only 10 epochs which is better than the pre-trained Alexnet method.

Additionally, the proposed approaches (pre-trained Alexnet) and (pre-trained Alexnet with SVM) that uses six patches from each of the images were compared with the results obtained in (Spanhol, Oliveira et al. 2016, Spanhol, Oliveira et al. 2017, Deniz, Şengür et al. 2018, Wenzhong, Huanlan et al. 2020) at both patient and image levels. The experiments were performed using the Stochastic Gradient Descent (SGD) method with a mini-batch size of 100, a learning rate of \(10^{-4}\), momentum term of 0.9, a weight decay of \(4^{-5}\), and at a maximum of 20 epochs. Table 6 and Table 7 reports the best average accuracy of the proposed method for 40X, 100X, 200X and 400X magnification factors using five-fold cross-validation at the image and the patient levels, respectively.

### Table 6: Mean Accuracy rates for different strategies at the patient level

| Strategy                  | Magnification Factors |
|---------------------------|-----------------------|
|                           | 40X       | 100X      | 200X      | 400X      |
| CNN (Sum) (2016) (Spanhol, Oliveira et al. 2016) | 88.4   | 88.4   | 83.8   | 85.3   |
| CNN (Product) (2016) (Spanhol, Oliveira et al. 2016) | 89.2   | 88.4   | 83.8   | 85.3   |
| CNN (Max) (2016) (Spanhol, Oliveira et al. 2016) | 90.0   | 88.4   | 84.6   | 86.1   |
| CNN + DeCAF (2017) (Spanhol, Oliveira et al. 2016) | 88.5   | 88.5   | 90.3   | 87.1   |
| VGG16 (2020) (Wenzhong, Huanlan et al. 2020) | 50.00 | 83.33   | 33.33 | 57.14   |
| AlexNet (2020) (Wenzhong, Huanlan et al. 2020) | 83.33   | 83.33   | 83.33   | 85.71   |
| Pre-trained Alexnet       | 92.0   | 92.0   | 92.0   | 94.1   |
| Pre-trained Alexnet with SVM | 94.0   | 94.0   | 92.0   | 96.1   |

### Table 7: Mean Accuracy rates for different strategies at the image level

| Strategy                  | Magnification Factors |
|---------------------------|-----------------------|
|                           | 40X       | 100X      | 200X      | 400X      |
| CNN (Sum) (2016) (Spanhol, Oliveira et al. 2016) | 85.4   | 83.3   | 83.1   | 80.8   |
| CNN (Product) (2016) (Spanhol, Oliveira et al. 2016) | 85.5   | 83.4   | 83.0   | 80.8   |
| CNN (Max) (2016) (Spanhol, Oliveira et al. 2016) | 85.6   | 83.5   | 82.7   | 80.7   |
| CNN + DeCAF (2017) (Spanhol, Oliveira et al. 2016) | 88.0   | 88.8   | 88.7   | 86.7   |
| AlexNet-fc6 + Vgg16-fc6 (2018) (Deniz, Şengür et al. 2018) | 84.87   | 89.21   | 88.65   | 86.75   |
| AlexNet-fc7 + Vgg16-fc7 (2018) (Deniz, Şengür et al. 2018) | 84.58   | 89.03   | 88.31   | 86.00   |
| Fine-tuned AlexNet (2018) (Deniz, Şengür et al. 2018) | 90.96   | 90.58   | 91.37   | 91.30   |
| VGG16 (2020) (Wenzhong, Huanlan et al. 2020) | 84.77   | 86.82   | 87.81   | 86.96   |
| AlexNet (2020) (Wenzhong, Huanlan et al. 2020) | 85.48   | 86.64   | 87.63   | 89.53   |
| Pre-trained Alexnet       | 91.7   | 91.9   | 91.7   | 90.2   |
| Pre-trained Alexnet with SVM | 92.0   | 91.9   | 92.0   | 91.6   |
As seen in Table 6, the proposed (Pre-trained Alexnet with SVM) method achieves accuracy between 92.0% and 96.1% at the patient level, which outperforms the approaches proposed in the literature in terms of accuracy. In Table 7, the proposed method achieves an accuracy of approximately 92% at image level for magnification factors 40X, 100X, and 200X. The proposed (Pre-trained Alexnet with SVM) method scored the highest average accuracy results in both image and patient levels for all magnification factors 40X, 100X, 200X and 400X.

6. Conclusion

This paper compares the classification efforts of transfer learning and deep feature extraction on breast cancer detection based on the histopathological images. For deep feature extraction, the famous deep CNN architecture, namely AlexNet model, is used with transfer learning. The BreaKHis dataset is preferred in the experimental works due to the huge number of sample images. Two different experimental works are considered. In the first one, the pre-trained Alexnet that uses six patches from each of the images of the dataset was performed. In the second one, the features vectors from FC6 and FC7 layers of pre-trained Alexnet network were fed to the SVM classifier. The proposed (Pre-trained Alexnet with SVM) method scored the highest average accuracy results in both image and patient levels for all magnification factors 40X, 100X, 200X and 400X.

References

Abd Almisreb, A., et al. (2018). Utilizing AlexNet deep transfer learning for ear recognition. 2018. Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), IEEE.

Berrar, D. J. E. o. B. and C. Biology (2019). "Cross-validation." 1: 542-545.

Bottou, L. (2012). Stochastic gradient descent tricks. Neural networks: Tricks of the trade, Springer: 421-436.

Boyle, P. and B. Levin (2008). World cancer report 2008, IARC Press, International Agency for Research on Cancer.

Deniz, E., et al. (2018). "Transfer learning based histopathologic image classification for breast cancer detection." 6(1): 18.

Duda, R. O., et al. (2012). Pattern classification, John Wiley & Sons.

Fadhil, A. F. (2014). Formulation of detection strategies in images, Southern Illinois University at Carbondale.

Hafemann, L. G., et al. (2014). Forest species recognition using deep convolutional neural networks. 2014 22nd International Conference on Pattern Recognition, IEEE.

Huang, Z., et al. (2017). "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data." 9(9): 907.

Kassani, S. H., et al. (2019). "Breast cancer diagnosis with transfer learning and global pooling.

Krizhevsky, A., et al. (2012). Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems.

Lakhani, S., et al. (2012). "WHO Classification of Tumours of the Breast, ed 4, Lyon."

Niu, X.-X. and C. Y. J. P. R. Suen (2012). "A novel hybrid CNN–SVM classifier for recognizing handwritten digits." 45(4): 1318-1325.

Spanhol, F. A., et al. (2017). Deep features for breast cancer histopathological image classification. 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE.

Spanhol, F. A., et al. (2016). Breast cancer histopathological image classification using convolutional neural networks. 2016 international joint conference on neural networks (IJCNN), IEEE.

Spanhol, F. A., et al. (2016). "A dataset for breast cancer histopathological image classification." 63(7): 1455-1462.

Stenkvist, B., et al. (1978). "Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations." 38(12): 4688-4697.

Wenzhong, L., et al. (2020). "Classifications of Breast Cancer Images by Deep Learning."