Applied Soft Computing Journal 85 (2019) 105765

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com/locate/asoc

Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images

Ümit Budak [a,*], Zafer Cömert [b], Zryan Najat Rashid [c,d], Abdulkadir Şengür [e], Musa Çıbuk [f]

[a] Department of Electrical and Electronics Engineering, Bitlis Eren University, Bitlis, Turkey
[b] Department of Software Engineering, Samsun University, Samsun, Turkey
[c] Computer Science Institute, Network Department, Sulaimani Polytechnic University, Sulaimani, Iraq
[d] Information Technology Department, Collage of Science and Technology, University of Human Development, Sulaimani, Iraq
[e] Department of Electrical and Electronics Engineering, Technology Faculty, Firat University, Elazig, Turkey
[f] Department of Computer Engineering, Bitlis Eren University, Bitlis, Turkey

ARTICLE INFO

Article history:
Received 26 May 2019
Received in revised form 2 September 2019
Accepted 5 September 2019
Available online 13 September 2019

Keywords:
Breast cancer detection
Histopathological biopsy image
Deep neural network
Combined FCN and Bi-LSTM

ABSTRACT

Breast cancer (BC) is one of the most frequent types of cancer that adult females suffer from worldwide. Many BC patients face irreversible conditions and even death due to late diagnosis and treatment. Therefore, early BC diagnosis systems based on pathological breast imagery have been in demand in recent years. In this paper, we introduce an end-to-end model based on fully convolutional network (FCN) and bidirectional long short term memory (Bi-LSTM) for BC detection. FCN is used as an encoder for high-level feature extraction. Output of the FCN is turned to a one-dimensional sequence by the flatten layer and fed into the Bi-LSTM's input. This method ensures that high-resolution images are used as direct input to the model. We conducted our experiments on the BreaKHis database, which is publicly available at http://web.inf.ufpr.br/vri/breast-cancer-database. In order to evaluate the performance of the proposed method, the accuracy metric was used by considering the five-fold cross-validation technique. Performance of the proposed method was found to be better than previously reported results.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Cancer, as a set of diseases, is where cells divide without stopping and spread into the surrounding tissue and accumulate to create a lump called a tumor or malignancy [1]. The normal cells in the human body divide in order to generate new cells as the body requires them for the purposes of repair or replacement. Normal cells perish over time, whereas cancerous cells act abnormally due to mutations in the cells that crowd out the normal cells [2,3]. One of the more commonly seen cancer types is Breast Cancer (BC) that predominately affects the adult female population. There are different BC types, depending on which cells become cancerous in the breast. The breast consists of three basic parts that are lobules, ducts, and connective tissue [4–6]. The lobules are the glands yielding the milk, while the ducts are the tubes used to transport milk to the nipple, and the connective tissue keeps everything together and forms the physical breast. Generally, cancerous cells in BC start to grow in the lobules and ducts [7]. Although the symptoms of BC vary according to the individual, there are certain significant major warning signs such as lumps found in the breast or the underarm, thickening or swelling of a part of the breast, abnormal changes in the nipple area, reddening of the breast skin tissue, or pain experienced in the breast [8]. BC can present as a painless lump or as a painful hard mass developed inside the breast.

Any suspected breast lump or growth should be immediately checked by the appropriate medical professional, in addition to regular screening tests. It is well known that the mortality rate of BC is quite high, and is a substantial public health concern since one in eight adult females face this diagnosis at least once in their lifetime [9], and one out of 37 women may die due to BC [10]. For this reason, early detection and diagnosis of BC are of vital importance. In the current clinical practices, the early diagnosis of BC can be realized by way of non-operative, non-invasive techniques such as contrast-enhanced (CE) digital mammography, breast ultrasound and magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET) [11]. It should also be noted that patient self-awareness is of considerable importance for early BC diagnosis. Today, mammograms are routinely applied in clinics and have been adopted as one of the most promising means of early BC diagnosis. MRI tests utilizing magnets and radio waves to capture the breast image are also used along with mammography in order to screen adult females who are considered at high risk of contracting BC.

* Corresponding author.
E-mail address: umtbudak@gmail.com (Ü. Budak).

https://doi.org/10.1016/j.asoc.2019.105765
1568-4946/© 2019 Elsevier B.V. All rights reserved.

In summary, BC can be evaluated as a heterogeneous disease that is examined based on clinical, radiological and pathological assessment. In clinical assessments, familial history and lifestyle-related risk factors of patients are traditionally taken into account. Mammography and ultrasound devices are used in the radiological assessment of BC detection. Mammography produces more effective results for women over 40 years old and in dense tissue breasts, whereas ultrasound is appropriate for patients of all ages. Core biopsy or fine needle aspiration cytology are applied in the pathological assessment of BC detection where an abnormality has been determined through either clinical or radiological assessment. Through biopsy, the histopathological images are divided into two types of lesion; benign and malignant. Lastly, a scoring system based on this triple assessment approach is used by specialists as the basis for reaching a diagnosis decision regarding the cells.

The diagnosis of BC based on computational approaches using histopathological microscopic images has been considered less attractive since this process takes time, and is error-prone and labor-intensive. However, histopathological images are considered the gold standard for BC diagnosis. In this regard, several studies have been carried out for this specific task. Convolutional Neural Network Improvement for Breast Cancer Classification (CNN-BCC) system using mammographic images, which are provided by Mammographic Image Analysis Society (MIAS) and include 21 benign, 17 malignant and 183 normal cases, have been presented. The model achieved a 90.50% accuracy [12]. A novel BC intelligent diagnosis approach equipped with several optimization methods was presented, tested on Wisconsin Original Breast Cancer (WBC) and Wisconsin Diagnostic Breast Cancer (WDBC) datasets, and classification success reported to be 97.57% [13]. An incremental boosting CNN was developed to recognize normal tissue, benign lesions, in situ carcinomas, and invasive carcinoma patterns. The experiments were performed on BC microscopic histology images, and the researchers achieved 96.9% accuracy using histology images with 200× zoom rate [14]. A combination of nonparametric Multiple Instance Learning (MIL) and CNN was presented for BC classification. The experiments was performed on the BreaKHis dataset. The classification rate was observed as a 92.1% at a 40× magnification factor [15]. A novel Gauss–Newton representation based algorithm (GNRBA) was introduced as a novel expert system for BC diagnosis. The key feature of the model relies upon finding optimal weight coefficients of the significant training samples. The authors' all experiments were conducted on two publicly available datasets named WBCD and WDBC. They reported an overall accuracy of 99.27% for 60–40 training and test partition [10]. A hybrid deep neural network exploiting the multilevel feature representation of the histopathological image patches was suggested. The hybrid convolutional and recurrent deep neural network exploits the richer multilevel feature representation of the histopathological image patches. The model was applied to a new publicly available BC dataset consisting of 3771 samples categorized into four classes. The average accuracy of the model was observed at 91.3% [16].

The diagnosis of BC has also been realized using Back Propagation Neural Network (BPPN) and radial basis neural networks (RBFN) in another work, with accuracy obtained as 59.0% and 70.4%, respectively [17]. A new fuzzy-logic-based system including clustering, noise removal, and classification techniques was proposed for BC classification. Expectation Maximization (EM) was utilized as a clustering method, Regression Trees (CART) was used to generate fuzzy rules, and Principal Component Analysis (PCA) was preferred to tackle the multicollinearity issue. As a result, the proposed knowledge-based system significantly improved the prediction accuracy of BC [18].

Pre-trained deep CNN architectures such as GoogLeNet, VGGNet, and ResNet have been investigated using standard benchmark datasets for BC diagnosis. The proposed framework returned an average accuracy of 97.67% [19]. Similarly, VGG-16, VGG-19, and ResNet deep architectures have been employed for the BC histology image classification task. The combination of VGG-16 and logistic regression (LR) yielded the best results with 92.60% classification success [20]. Ultrasound images have been evaluated in order to discriminate benign and malignant breast tumors by using neural network and multi-fractal dimension features in another study, with classification accuracy reported as 82.04% [21].

As mentioned in the related works, various attempts have been conducted to improve the performance of computational approaches to diagnosing BC and to ensure the development of a more robust and consistent system of diagnosis. To this aim, we came up with a novel model based on the fully convolutional network (FCN) and bidirectional long short term memory (Bi-LSTM) for BC detection. FCN is adopted to directly handle high resolution histopathological images, free of the demand to compress or enlarge the input to a fixed-size in order to keep critical information from being lost. FCN encodes the biopsy samples into a high-level representation, while the Bi-LSTM mechanism impels the remaining sublayers of the model to focus on specific regions of the input array.

The contributions of this study are listed as follows:

(1) We introduce an end-to-end model based on FCN and Bi-LSTM to diagnose types of BC. This proposed model differs from conventional deep learning methods with regard to its structure.
(2) We propose exploiting an FCN as an encoder, making it possible to process largescale images efficiently and, hence, allowing variable input sizes.
(3) Although a Bi-LSTM model is actually used in time series problems, we have shown how this model is able to be integrated into image processing applications.

The remainder of the paper is organized as follows: the proposed method is described in detail in Section 2, the experimental work and results are presented in Section 3, and concluding remarks are given in Section 4.

2. Proposed method

In this paper, we propose an end-to-end model based on fully convolutional network and bidirectional long short-term (Bi-LSTM) for BC detection. The FCN is used to encode the input image into a high-level representation. Then the output of the FCN is flattened and fed into to the Bi-LSTM model. The overall architecture is illustrated in Fig. 1.

2.1. Fully convolutional network

In recent years, CNNs have been widely used for pattern recognition, especially in the field of image recognition, and do not require classic handcraft feature extraction. Moreover, it has been also proven that CNN-based systems can provide comparable and even improved classification performance in previous BC detection works when compared to traditional systems [22,23]. The CNNs are basically composed of convolution, pooling and activation layers. A convolution layer is designed according to the number of input channels, the number of output activation maps, the filter size, and the stride. Generally, each filter size can be considered to be very small compared to the input size. For this reason, a filter operates in a local region of the input instead of the whole activation map. The spatial extent of these

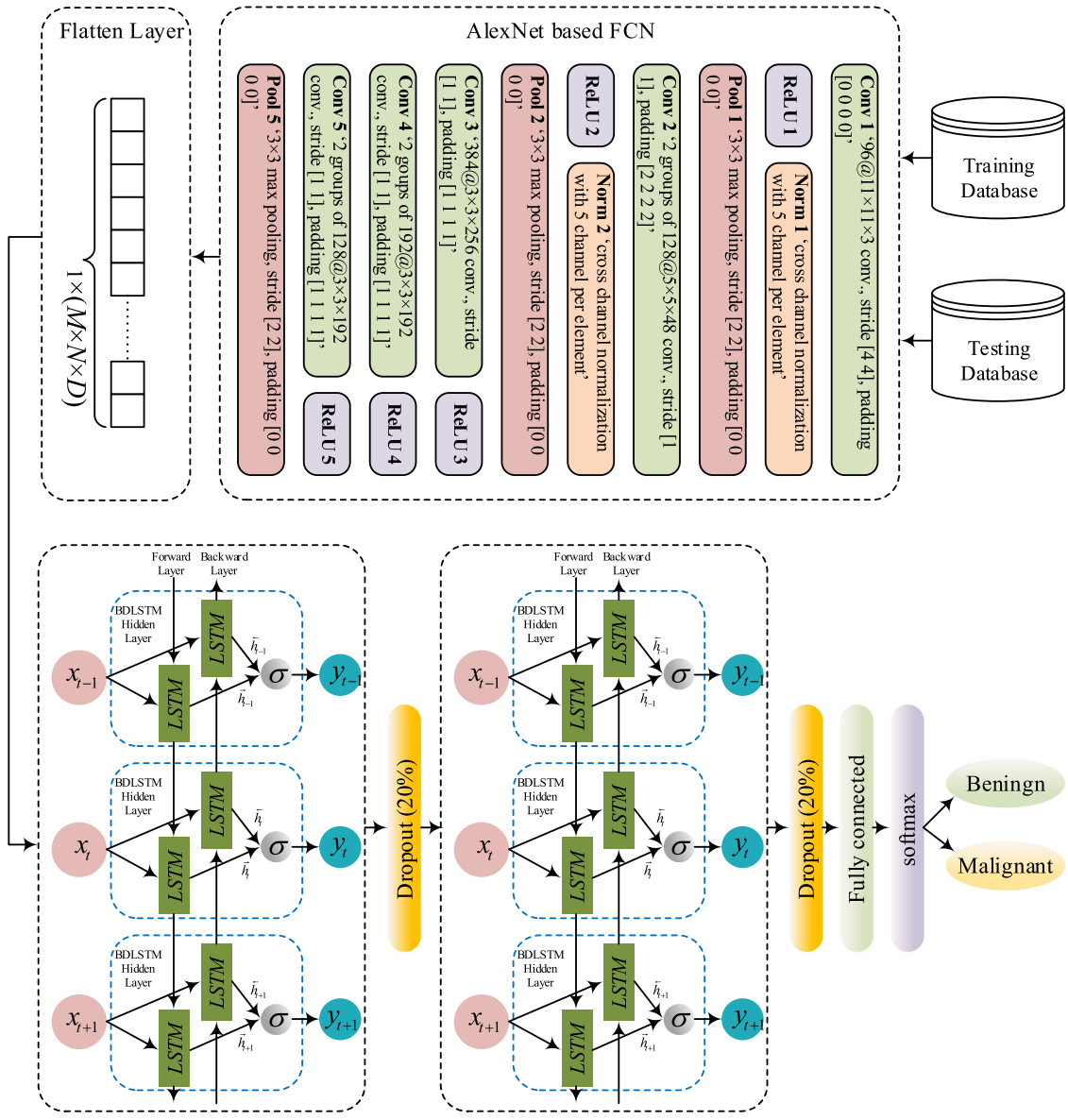

Fig. 1. Architecture of FCN and Bi-LSTM based network.

locations that connect to the higher layers is a hyperparameter known as receptive fields. On an activation map, the filter weights are shared in order to extract certain features in different locations. This also helps to reduce complexity of the network and makes computation more efficient. The pooling layer is a sample-based discretization process and conducts an average or maximum pooling operation in order to reduce dimensionality, remove noise, and to extract robust features. The activation layers operates as an element-wise nonlinear function. If CNN is used for image recognition work, then the layers that are fully connected and a softmax must also be connected in succession following these layers.

Pre-trained CNN models such as AlexNet [24], Oxford VGGNet [25], and ResNet [26] can process fixed-sized inputs. Inspired by [27], instead of extracting features from a fully connected (fc) layer [22–28], we transform AlexNet into a fully convolutional network by simply removing its "fc" layers, and then it is used as our encoder for high-level feature extraction. This allows the FCN to process the input image independently of its size. In another words, the FCN are able to handle input of arbitrary size without need to compress or enlarge it to a fixed size. As shown in Fig. 1, each convolution layer is followed by a ReLU layer, and the first two convolution layers are equipped with five-channel cross-channel normalization per element. The output of the FCN encoder is a three-dimensional array of size $M \times N \times D$, where the M and N correspond to width and height of the activation map and D is the channel size.

After using FCN encoder to extract the spatial features of an image, we turn the output of the FCN to a one-dimensional array vector by using a flatten layer. This layer can be considered as a kind of regulation layer which can be used to convert the $M \times N \times D$ dimensional feature map into a one-dimensionally $1 \times (M \times N \times D)$ suitable form necessary for the input of the Bi-LSTM model.

2.2. Bidirectional Long Short-Term Memory (Bi-LSTM)

Recently, the LSTM architecture that performs better than deep neural network systems for the classification of sound and signal categories is a unit of Recurrent Neural Networks (RNNs) [29–31]. Recurrent connections and special memory blocks in the recurrent hidden layer make them the more robust

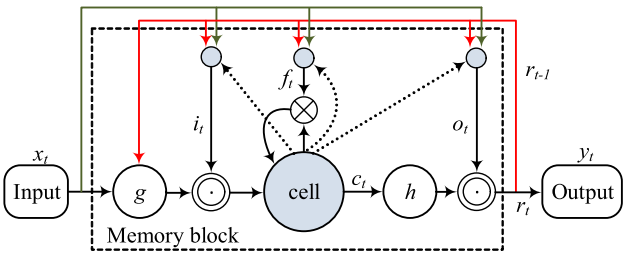

Fig. 2. The structure of the LSTM unit (for simplicity, only one memory block is shown).

structure to model sequence data than traditional RNNs and feedforward neural networks. Memory structures have memory cells with self-connections that store the temporal state of the network at each time step. In addition, the multiplication units named as gates also provide the flow of information to the memory cell (input gate) and from the cell to the other units (output gate). The forget gate is also employed in order to scale the cell's internal state prior to adding it as input to the cell by way of self-recurrence, and then resets or forgets the cell's memory when necessary. The LSTM architecture includes "peephole" connections to learn the exact timing of the outputs [31]. As seen in Fig. 2, the weighted peephole connections from cell to gates are shown in dashed lines.

The LSTM network architecture [30–32] is shown in Fig. 1. Let $x_t = (x_1, \ldots, x_T)$ be an input sequence, and $y_t = (y_1, \ldots, y_T)$ express an output sequence. The output sequence is calculated using the following mathematical sentences from the time step $t = 1$ to T, iteratively.

$$i_t = \sigma\left[(W_{ix} \times x_t) + (W_{ir} \times r_{t-1}) + (W_{ic} \times c_{t-1}) + b_i\right] \quad (1)$$

$$f_t = \sigma\left[(W_{fx} \times x_t) + (W_{fr} \times r_{t-1}) + (W_{fc} \times c_{t-1}) + b_i\right] \quad (2)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot tanh\left[(W_{cx} \times x_t) + (W_{cr} \times r_{t-1}) + b_c\right] \quad (3)$$

$$o_t = \sigma\left[(W_{ox} \times x_t) + (W_{or} \times r_{t-1}) + (W_{oc} \times c_t) + b_o\right] \quad (4)$$

$$r_t = o_t \odot \tanh(c_t) \quad (5)$$

$$y_t = \emptyset(W_{yr} \times r_t + b_y) \quad (6)$$

where W is the weight matrices. For example, W_{ix} denotes the weight's matrix from the input gate to the input. W_{ic}, W_{fc}, W_{oc} denote diagonal weight matrices for peephole connections, b is the bias vector, σ states the sigmoid function, and i, f, o, c are the input gate, forget gate, output gate, and cell activation vectors, respectively. In addition, $\odot$ and $\emptyset$ indicate element-wise multiplication and softmax activation function for the LSTM.

Although the RNNs have the advantage of coding the dependencies between inputs, it causes a state of exploding and vanishing against their gradient for long data sequences. Because both LSTM and RNN can only take information from the previous context, further improvements can be made using Bi-RNN. The Bi-RNN can handle two sources of information [32] and their combination, combined with the LSTM, creates Bi-LSTM, as illustrated in Fig. 1. Apart from the Bi-LSTM having the advantage of LSTM with feedback for the next layer, it can also handle data with dependence on long range.

3. Experimental works and results

3.1. Database

As the proposed method aims to efficiently diagnose BC from histopathological images, the BreaKHis database [33] was used during the experimental works. This database contains breast tissue biopsy samples stained with hematoxylin and eosin collected during a clinical study between January and December 2014. Within this time period, all patients referred to the P&D Laboratory in Brazil were invited to participate in the study. The 7909 samples in the database were generated from 82 volunteer participants, of which 2480 samples are labeled as benign and the remaining 5429 labeled as malignant tumors. The patient samples were divided into four sub-databases of 40×, 100×, 200×, and 400×, according to the image magnification factor. All of the images are colored and sized as 700 × 460 pixels. The database was randomly arranged into five folds. Table 1 summarizes the distribution of images in the database.

Table 1
Image distribution by magnification factor and class.

Magnification	Benign	Malignant	Total
40×	625	1370	1995
100×	644	1437	2081
200×	623	1390	2013
400×	588	1232	1820
Total	2480	5429	7909
# Patients	24	58	82

Fig. 3 shows a typical histopathological breast tissue sample, zoomed to four different magnification factors. The highlighted rectangles were added manually for illustrative purposes only in order to highlight the regions of interest selected by the pathologist for increased magnification.

3.2. Experimental details

All of the programs were conducted on a workstation equipped with an Intel(R) Xeon(R) Gold 6132 @2.60 GHz CPU with 14 core and NVIDIA Quadro P6000 GPU. The experiments were conducted in a MATLAB® environment installed on a Windows operating system.

The initial experiments were conducted with the pre-trained AlexNet model in order to assess the performance of the BC detection. Detailed information about the network structure of the AlexNet model can be viewed in [24]. All input images were initially resized to 227 × 227 pixels in terms of compatibility with the model. For fine-tuning the AlexNet model, the mini batch size and max epoch were selected as 8 and 20, respectively, and the initial learning rate was adjusted to 0.0001, and the model was trained by stochastic gradient descent with momentum. For performance evaluation, accuracy, sensitivity and specificity scores were used. Table 2 presents the performance results for the AlexNet model. As can be seen, the best average accuracy of 94.54 ± 1.01% was obtained for the 200× magnification factor. In addition, with 93.68 ± 1.19% average accuracy, the 40× magnification factor yielded the second best performance. The worst average accuracy score was 88.68 ± 7.19%, which was obtained for the 400× magnification factor. When evaluation was performed over the Folds, it was seen that the best average accuracy was obtained for Fold4, where the calculated average accuracy was 93.66 ± 1.71%. The second best performance was obtained for Fold2, where the average accuracy was 93.52 ± 1.05%. Fold3 yielded the worst performance, where the average accuracy was 90.02 ± 9.32%. In addition, when average sensitivity scores were considered, it is seen that 200× magnification factor yielded the 96.40 ± 1.13% score which is the highest one among other. The 400×, 40× and 100× magnification factors obtained the second, third and fourth highest scores, respectively. Furthermore, when the average sensitivity scores across folds were investigated, it is seen that fold3 produced the highest 96.93 ± 1.30% average sensitivity scores. 96.54 ± 1.95%, 95.40 ± 0.53%, 95.26 ± 2.93%

Table 2
Classification performance obtained with transfer learning using the AlexNet model.

	Accuracy (%)				Sensitivity (%)				Specificity (%)			
	40×	100×	200×	400×	40×	100×	200×	400×	40×	100×	200×	400×
Fold1	91.98	92.79	93.05	93.68	91.24	98.25	96.04	95.52	93.60	80.62	86.40	89.83
Fold2	94.24	93.53	94.28	92.03	95.98	94.79	95.68	95.14	90.40	90.69	91.12	85.47
Fold3	93.48	94.70	95.78	76.10	96.35	95.81	96.76	98.78	87.20	92.18	93.60	28.81
Fold4	95.24	92.33	95.02	92.03	97.44	93.75	98.20	96.76	90.40	89.14	87.90	82.05
Fold5	93.48	91.11	94.54	89.56	97.08	94.77	95.32	93.49	85.60	82.94	92.80	81.35
Mean & Std.	93.68 ± 1.19	92.89 ± 1.34	94.54 ± 1.01	88.68 ± 7.19	95.62 ± 2.51	95.47 ± 1.71	96.40 ± 1.13	95.94 ± 1.96	89.44 ± 3.11	87.11 ± 5.05	90.36 ± 3.11	73.50 ± 25.20

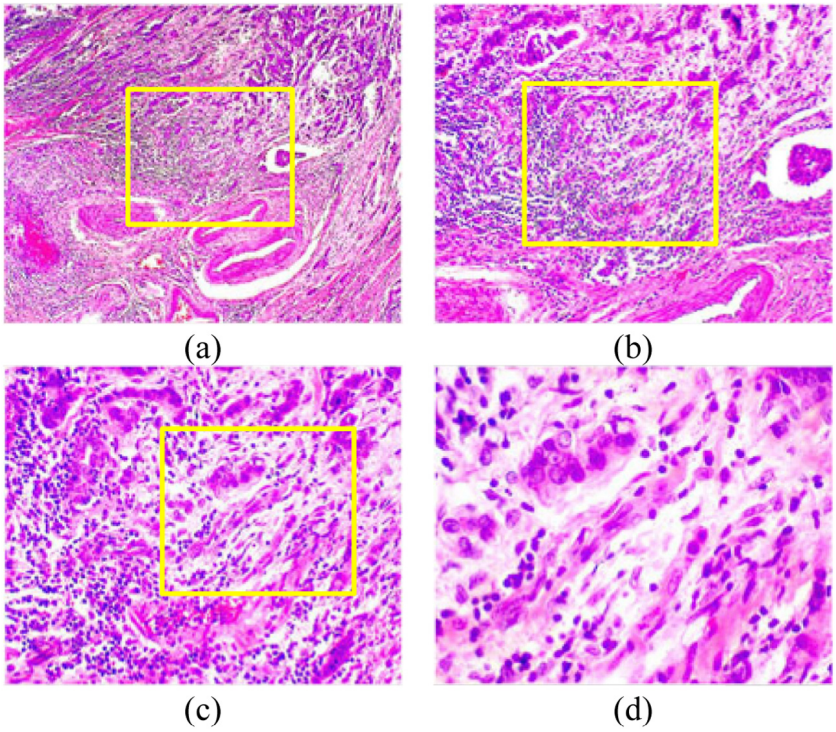

Fig. 3. Sample malignant breast tissue stained with hematoxylin at four different magnification factors: (a) 40×, (b) 100×, (c) 200×, (d) 400× extracted from [18].

and 95.17 ± 1.49% average sensitivity scores were obtained for fold4, fold2, fold1 and fold5, respectively. The calculated average specificity scores also showed that the highest average specificity score 90.36 ± 3.11% was obtained for 200× magnification factor. And the worst average specificity score 73.50 ± 25.20% was obtained for 400× magnification factor. It is also worth to mentioning that fold2 produced 89.42 ± 2.65% average specificity score which is the highest among other folds.

Then, we turned the AlexNet into the FCN form, combined a flatten and Bi-LSTM layer to create our proposed model. Unlike the AlexNet model, training of the proposed model was carried out with "*adam*" optimizer, which is an adaptive learning rate optimization algorithm. We also fed the input images directly to our proposed model without need for compression or magnification. The performance results of our proposed model are reported in Table 3. As can be seen, the best average accuracy of 96.32 ± 0.51% was obtained for magnification factor 200×. Moreover, for magnification factors 40×, 100×, 400×, the obtained average accuracies were 95.69 ± 1.78%, 93.61 ± 2.28%, and 94.29 ± 1.86%, respectively. It is worth mentioning that the highest accuracy of 97.59% was obtained for Fold3 with 100× magnification factor, and the worst accuracy of 91.76% was also obtained for Fold3 but with 400× magnification factor.

The average sensitivity and specificity scores for each magnification factor can be seen in Table 3. While the 40× magnification factor produced the highest average sensitivity score, the highest average specificity score was produced for 200× magnification factor. In addition, the worst average sensitivity and specificity scores were obtained for 400× and 100× magnification factors, respectively.

We also compared the results for the proposed model with other results previously reported in [22,23,33]. In [23], researchers applied an end-to-end CNN approach that was based on the extraction of image patches for the training of the network, and the combination of these patches for final classification. In [33], the authors exploited various classifiers, i.e., Neural Networks (NN), Quadratic Discriminant Analysis (QDA), Support Vector Machine (SVM) and Random Forest (RF) for diagnosis of BC from histopathological images. Additionally, in [22], the authors dealt with a transfer learning (TL) based approach by using a pre-trained CNN model for fine-tuning. Table 4 reports on the related comparisons. As can be seen, the proposed model achieved improved accuracy scores when compared to the other studies for each magnification factor, whilst the TL approach [22] attained the second best accuracy scores for each magnification factor. The proposed method demonstrates its robustness by achieving a difference of approximately 5 points for the 40× and 200× sub-databases, and approximately 3 points for the other sub-databases.

The proposed method was further applied on classification of the CIFAR-10 dataset []. The CIFAR-10 dataset contains 60

Table 3
Classification performance of the proposed model.

	Accuracy (%)				Sensitivity (%)				Specificity (%)			
	40×	100×	200×	400×	40×	100×	200×	400×	40×	100×	200×	400×
Fold1	95.99	93.03	95.78	96.70	98.17	98.95	96.04	96.34	91.20	79.84	95.20	97.45
Fold2	97.49	93.05	96.77	93.96	98.17	97.22	97.48	93.52	96.00	83.72	95.16	94.87
Fold3	96.49	97.59	96.77	91.76	99.63	97.21	97.84	97.15	89.60	98.43	94.40	80.50
Fold4	92.73	92.57	96.52	93.68	98.90	92.01	97.48	97.16	79.20	93.79	94.35	86.32
Fold5	95.74	91.83	95.78	95.33	95.62	94.77	97.84	94.71	96.00	85.27	91.20	96.61
Mean & Std.	95.69 ± 1.78	93.61 ± 2.28	96.32 ± 0.51	94.29 ± 1.86	98.10 ± 1.51	96.03 ± 2.69	97.33 ± 0.74	95.77 ± 1.60	90.40 ± 6.88	88.21 ± 7.65	94.06 ± 1.65	91.15 ± 7.41

Table 4
Comparison of proposed method with previously reported results.

Method	Accuracy (%)			
	40×	100×	200×	400×
NN [33]	80.9 ± 2.0	80.7 ± 2.4	81.5 ± 2.7	79.4 ± 3.9
QDA [33]	83.8 ± 4.1	82.1 ± 4.9	84.2 ± 4.1	82.0 ± 5.9
RF [33]	81.8 ± 2.0	81.3 ± 2.8	83.5 ± 2.3	81.0 ± 3.8
SVM [33]	81.6 ± 3.0	79.9 ± 5.4	85.1 ± 3.1	82.3 ± 3.8
CNN [23]	90.0 ± 6.7	88.4 ± 4.8	84.6 ± 4.2	86.1 ± 6.2
TL [22]	90.96 ± 1.59	90.58 ± 1.96	91.37 ± 1.72	91.30 ± 0.74
Proposed	**95.69 ± 1.78**	**93.61 ± 2.28**	**96.32 ± 0.51**	**94.29 ± 1.86**

Table 5
The obtained classification results on CIFAR-10 dataset.

Method	Accuracy (%)					Mean & Std.
	Fold1	Fold2	Fold3	Fold4	Fold5	
AlexNet (TL)	89.35	89.15	89.77	89.34	89.63	89.44 ± 0.24
Ours	91.56	91.88	92.12	91.72	92.23	91.90 ± 0.27

000 color images in 10 classes of size 32×32. The dataset was divided into five training and one test sets. The training sets were considered in our experiments, as they contain ground-truth class labels. Table 5 shows the obtained results on CIFAR-10 dataset. The first row of Table 5 shows the results obtained by AlexNet and the second row shows the achievements of our proposed method. As seen in Table 5, for all training folds, the proposed method outperformed. While the proposed method obtained 91.90% average accuracy score, the AlexNet model obtained 89.44% average accuracy score.

4. Conclusion

In this paper, an end-to-end CNN model based on FCN and Bi-LSTM is proposed for improved efficiency in BC detection. The proposed novel scheme is composed of FCN and Bi-LSTM layers The FCN layer is considered as an encoder for high-level feature extraction and the output of the FCN is resized to a one-dimensional sequence by the flatten layer. The output of the flatten layer is then fed into the Bi-LSTM's layer. The proposed deep scheme was applied to the BreaKHis database and the obtained results proved to be significantly positive. To the best of our knowledge, the obtained results exceed all other results obtained from the same database. In the future works, we are planning to construct new deep feature extractors and deep stacked auto-encoders for classification purposes.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.asoc.2019.105765.

References

[1] M. Sibbering, C.-A. Courtney, Management of breast cancer: basic principles, Surg. 34 (2016) 25–31, http://dx.doi.org/10.1016/j.mpsur.2015.10.005.

[2] T.-A. Moo, R. Sanford, C. Dang, M. Morrow, Overview of breast cancer therapy, PET Clin. 13 (2018) 339–354, http://dx.doi.org/10.1016/j.cpet.2018.02.006.

[3] L. Wang, Early diagnosis of breast cancer, Sensors 17 (2017) 1572, http://dx.doi.org/10.3390/s17071572.

[4] J.-L. Gonzalez-Hernandez, A.N. Recinella, S.G. Kandlikar, D. Dabydeen, L. Medeiros, P. Phatak, Technology, application and potential of dynamic breast thermography for the detection of breast cancer, Int. J. Heat Mass Transfer 131 (2019) 558–573, http://dx.doi.org/10.1016/j.ijheatmasstransfer.2018.11.089.

[5] E. Provenzano, G.A. Ulaner, S.-F. Chin, Molecular classification of breast cancer, PET Clin. 13 (2018) 325–338, http://dx.doi.org/10.1016/j.cpet.2018.02.004.

[6] H. Blumen, K. Fitch, V. Polkus, Comparison of treatment costs for breast cancer, by tumor stage and type of service, Am. Heal. Drug Benefits 9 (2016) 23–32, https://www.ncbi.nlm.nih.gov/pubmed/27066193.

[7] A. Rezaee, A. Buck, M. Raderer, W. Langsteger, M. Beheshti, Chapter 3 - breast cancer, in: M. Beheshti, W. Langsteger, A. Rezaee (Eds.), PET/CT Cancer an Interdiscip. Approach to Individ. Imaging, Elsevier, 2018, pp. 43–63, http://dx.doi.org/10.1016/B978-0-323-48567-8.00003-1.

[8] M. Sibbering, C.-A. Courtney, Management of breast cancer: basic principles, Surg. 37 (2019) 157–163, http://dx.doi.org/10.1016/j.mpsur.2019.01.004.

[9] K. McPherson, C.M. Steel, J.M. Dixon, Breast cancer-epidemiology, risk factors, and genetics, BMJ 321 (2000) 624–628, http://dx.doi.org/10.1136/bmj.321.7261.624.

[10] L. Dora, S. Agrawal, R. Panda, A. Abraham, Optimal breast cancer classification using Gauss–Newton representation based algorithm, Expert Syst. Appl. 85 (2017) 134–145, http://dx.doi.org/10.1016/j.eswa.2017.05.035.

[11] V. Jordan, M. Khan, D. Prill, Breast cancer screening: Why can't everyone agree? Prim. Care Clin. Off. Pract. 46 (2019) 97–115, http://dx.doi.org/10.1016/j.pop.2018.10.010.

[12] F.F. Ting, Y.J. Tan, K.S. Sim, Convolutional neural network improvement for breast cancer classification, Expert Syst. Appl. 120 (2019) 103–115, http://dx.doi.org/10.1016/j.eswa.2018.11.008.

[13] N. Liu, E.-S. Qi, M. Xu, B. Gao, G.-Q. Liu, A novel intelligent classification model for breast cancer diagnosis, Inf. Process. Manag. 56 (2019) 609–623, http://dx.doi.org/10.1016/j.ipm.2018.10.014.

[14] D.M. Vo, N.-Q. Nguyen, S.-W. Lee, Classification of breast cancer histology images using incremental boosting convolution networks, Inf. Sci. (Ny). 482 (2019) 123–138, http://dx.doi.org/10.1016/j.ins.2018.12.089.

[15] P.J. Sudharshan, C. Petitjean, F. Spanhol, L.E. Oliveira, L. Heutte, P. Honeine, Multiple instance learning for histopathological breast cancer image classification, Expert Syst. Appl. 117 (2019) 103–111, http://dx.doi.org/10.1016/j.eswa.2018.09.049.

[16] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, F. Zhang, Breast cancer histopathological image classification using a hybrid deep neural network, Methods (2019) http://dx.doi.org/10.1016/j.ymeth.2019.06.014.

[17] S. Kaymak, A. Helwan, D. Uzun, Breast cancer image classification using artificial neural networks, Procedia Comput. Sci. 120 (2017) 126–131, http://dx.doi.org/10.1016/j.procs.2017.11.219.

[18] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, A knowledge-based system for breast cancer classification using fuzzy logic method, Telemat. Inform. 34 (2017) 133–144, http://dx.doi.org/10.1016/j.tele.2017.01.007.

[19] S. Khan, N. Islam, Z. Jan, I. Ud Din, J.J.P.C. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, Pattern Recognit. Lett. 125 (2019) 1–6, http://dx.doi.org/10.1016/j.patrec.2019.03.022.

[20] Shallu, R. Mehra, Breast cancer histology images classification: Training from scratch or transfer learning? ICT Express 4 (2018) 247–254, http://dx.doi.org/10.1016/j.icte.2018.10.007.
[21] M.A. Mohammed, B. Al-Khateeb, A.N. Rashid, D.A. Ibrahim, M.K. Abd Ghani, S.A. Mostafa, Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images, Comput. Electr. Eng. 70 (2018) 871–882, http://dx.doi.org/10.1016/j.compeleceng.2018.01.033.
[22] E. Deniz, A. Sengür, Z. Kadiroglu, Y. Guo, V. Bajaj, Ü. Budak, Transfer learning based histopathologic image classification for breast cancer detection, Heal. Inf. Sci. Syst. 6 (2018) 18, http://dx.doi.org/10.1007/s13755-018-0057-x.
[23] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using Convolutional Neural Networks, in: 2016 Int. Jt. Conf. Neural Networks, 2016, pp. 2560–2567, http://dx.doi.org/10.1109/IJCNN.2016.7727519.
[24] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1–9, http://dx.doi.org/10.1016/j.protcy.2014.09.007.
[25] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, http://arxiv.org/abs/1409.1556, (2014).
[26] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conf. Comput. Vis. Pattern Recognit, 2016, pp. 770–778, http://dx.doi.org/10.1109/CVPR.2016.90.
[27] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: 2015 IEEE Conf. Comput. Vis. Pattern Recognit, 2015, pp. 3431–3440, http://dx.doi.org/10.1109/CVPR.2015.7298965.
[28] M. Cıbuk, U. Budak, Y. Guo, M. Cevdet Ince, A. Sengur, Efficient deep features selections and classification for flower species recognition, Measurement 137 (2019) 7–13, http://dx.doi.org/10.1016/j.measurement.2019.01.041.
[29] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780, http://dx.doi.org/10.1162/neco.1997.9.8.1735.
[30] F.A. Gers, Earning to forget: continual prediction with LSTM, in: 9th Int. Conf. Artif. Neural Networks ICANN '99, IEE, 1999, pp. 850–855, http://dx.doi.org/10.1049/cp:19991218.
[31] F.A. Gers, N.N. Schraudolph, J. Schmidhuber, Learning precise timing with LSTM recurrent networks, J. Mach. Learn. Res. (2002) 115–143, http://www.jmlr.org/papers/v3/gers02a.html.
[32] X. Li, H. Xianyu, J. Tian, W. Chen, F. Meng, M. Xu, L. Cai, A deep bidirectional long short-term memory based multi-scale approach for music dynamic emotion prediction, in: 2016 IEEE Int. Conf. Acoust. Speech Signal Process, 2016, pp. 544–548, http://dx.doi.org/10.1109/ICASSP.2016.7471734.
[33] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, IEEE Trans. Biomed. Eng. 63 (2016) 1455–1462, http://dx.doi.org/10.1109/TBME.2015.2496264.