Automatic Breast cancer ultrasound lesions classification using fusion of Convolutional Neural Networks.

Farzam Kharajinezhadian, Saeid Rashidi, Mojtaba Banifakhr, Mohammadali Kavousi

Automatic Breast cancer ultrasound lesions classification using fusion of Convolutional Neural Networks.

Farzam Kharajinezhadian

Faculty of biomedical engineering, Islamic Azad university, science and research branch Tehran, Iran

Farzam.kharaji@gmail.com

Saeid Rashidi

Faculty of biomedical engineering, Islamic Azad university, science and research branch Tehran, Iran

Rashidi.saeid@gmail.com

Mojtaba Banifakhr

Faculty of electrical engineering, Yazd university

Yazd, Iran

Banifakhr.mojtaba@stu.yazd.ac.ir

Abstract

Basics: In recent years, breast cancer is the most common cancer among women and the average age of its incidence is decreasing. Due to the dense breast tissue in younger women (less than 49 years old), which reduces the sensitivity of mammography in the diagnosis of carcinoma, the use of high-resolution ultrasound as a complement to mammography is very useful in its diagnosis. Most of common methods in Automated breast ultrasound lesions detection use classical machine learning method for classification of lesions that have big recognition lack in complex images. In this paper we propose a novel method for breast cancer lesions images classification, with significant result in complex breast cancer images .

Methods: In this paper we propose a method based on fusion of Convolutional neural networks .

Results: The mean age of patients was 46 years with an age range of 32-76 years. Out of 22 suspected masses reported by ultrasound, 21 cases of cancer and 1 case of abscess were reported. The mean mass diameter was 29 mm and the mean lymph node diameter was 17.3 mm. Mammographic findings due to dense breast tissue reported 94.21% and 97.10% acuuracy for final classification with two proposed method based of fusion of tree and four CNN classifiers.

Conclusion: Although mammography is the gold standard method of diagnosing breast cancer, but due to its occurrence at a younger age in recent years and the presence of dense breast tissue at this age, complementary ultrasound (supplementary screening method), especially in the lower age group to increase the sensitivity of diagnosis Will be effective..

Keywords: Breast cancer, lesions classification,Convolutional Neural Networks,fusion of classifiers.

1. Introduction

Breast cancer is the most common cancer in women and the second most common cancer overall. Evidence suggests that breast cancer incidence rates increase every year [1]. In fact, one in 8 women in the United States and two in 5 are diagnosed with cancer [2]. The number of cancer cases is projected to increase to 6.2 million people by 2030 [3]. The average annual cost of health care for a new cancer patient in the United States was $21,222 during 2001-2008 [4]. Therefore, early diagnosis is essential for guiding appropriate and timely treatments and a higher chance of survival[5].

Various imaging techniques have been developed to improve the sensitivity and specificity of breast cancer diagnosis [4]. Conventional screening techniques include complicated imaging methods. Meanwhile, there are some limitations to the use of these techniques, such as the pain experienced during mammography, damage caused by long-term exposure to radiation, and high costs [3]. Generally, the sensitivity of mammography depends on the physical characteristics of the tumor, and mammographic sensitivity is rendered a less sensitive tool in younger women with dense breast tissues. Ultrasound is another breast cancer screening tool, which is mostly used in young women to reduce unnecessary radiation exposure. Nonetheless, certain issues must be taken into account in this method, including noise conditions and operators' imaging expertise. In fact, they could cause failure in the discovery of deep breast microtissue[6]. The final technique is thermography. Infrared imaging is a non-contact method, used widely by the medical community in recent years [7]. Given the fact that infrared thermal imaging does not use ionizing radiation or additional invasive techniques, patients undergoing examination by this method feel no discomfort. Owing to its high sensitivity and specificity, thermography has been extensively used as an auxiliary screening tool [8]. Thermography results can be correct 8-10 years before detection by mammography [1]. In addition, infrared thermography shows skin surface temperature as a qualitative and quantitative tool. It is also able to minimize small temperature changes in the anomalies of arteries and reflects angiogenic areas due to breast cancer [9]. Given the higher chemical changes and activities of blood reservoirs in cancer tissues, relative to adjacent areas, cancer breast tissues are detected in thermography due to high temperature[5].

Various image processing and artificial intelligence techniques have been proposed to detect breast cancer by thermography images in the last few decades, which can be more efficient and reduce costs. The clinical interpretation of a breast thermogramis primarily based on the asymmetry analysis of these heat patterns subjectively. The heat patterns are symmetrically distributed in both breasts. Therefore, bidirectional asymmetric analysis has been widely evaluated in previous studies. In fact, it is the classic breast screening method in thermography [10].

There is an asymmetric temperature distribution of heat in two breasts with hot areas due to abnormal growth of blood vessels [1]. While the quantitative analysis of these hot areas provides useful information about breast pathology, its success requires the extraction of some features from the left and right breast to perform asymmetric analysis [11]. Normally, the extracted features are the first moments, such as mean, standard deviation, and the difference between the maximum and minimum temperature of two sides of the breast[12].

Some studies use morphology information, such as the number of veins with higher temperatures, tissue features, the number of hot areas, the geometry of the hot center and the histogram of thermal images [13]. Ultimately, the extracted features are

classified by artificial intelligence algorithms. In this regard, some of the algorithms applied in previous studies include fuzzy classification, Bayesian network, support vector machine (SVM), neural networks and wavelet transform [13-15].

A literature review revealed some limitations to the use of breast thermography, including lack of researchers, especially image processing engineers' access to free databases, the small number of images on the database, and, most important of all, different breast morphologies (various geometrical shapes) in people.

This significance is also related to the asymmetry analysis, which is the main thermography technique. Lack of proper selection and separation of breast area will lead to failure to implement analyses such as the first and second moments by using the asymmetry technique. Another problem is reduced sensitivity and accuracy of result classification in terms of being sick or healthy .

According to what was mentioned and with regard to our knowledge or authors of previous works [8, 24, 25] in the area of breast cancer thermography, it is suggested that breast separation and feature extraction be avoided and diagnosis be made based on the raw thermal image. In this regard, deep learning is a novel technique that can significantly contribute to this area. Various architectures have been proposed for image classification problems in deep networks. In this study, attempts were made to diagnose breast cancer based on thermal images by assessing several deep learning methods.

Table 1. A summary of the analysis of bilateral asymmetry in mammograms based on screening diagnosis methods of thermography

Author	**Method**	**Type of Classifier**	**Accuracy**	**Sensitivity**	**Specificity**	**# Database images**
Schaefer et al. [16]2009	Stochastic tissue features	A classifier based on hybrid fuzzy law	80%	-	-	146
Ng et al. [17]2007	Patient temperature and clinical data	Radial basis function neural networks	80.95%	81.20%	88.20%	82
Mookiah et al. [18]2012	Tissue properties and DWT coefficients	Decision tree	93.30%	86.70%	100%	50
Acharya et al. [13]2012	Tissue features	SVM	88.10%	85.71%	100%	50
Francis et al. [19]2014	Stochastic tissue features	SVM	90.91%	81.82%	100%	22
Araujo et al. [20]2014	Symbolic features Temperature time interval	Distance-based classifier	-	85.70%	86.50%	50
Zadeh et al. [21]2012	The first moments	Self-organizing neural networks	70%	50%	75%	200
Sathish et al. [22]2017	Stochastic tissue features	SVM	90%	87.5%	92.5%	80
Borchartt et al. [12]2011	The first moments stochastic	SVM	85.71%	95.83%	25%	28

	morphology features					
Lashkari et al. [23]2016	Frequency, amplitude, histogram	Adaboost	85.33%	63.33%	90.83%	67
Gogoi et al. [10]2018	Singular values	SVM	98%	98%	98%	145

2. Materials and Proposed Methods

In this study, we used the breast thermography images derived from the BUS database [26], Based on[26] accumulated dataset, contain 2673 images werecollected in three batches in HunanXiangya Hospital over the years, of which 1425 were Benign images and 1248 Malignantmammograms from405 patients. For this paper we use randomly selected.163 images.

As observed, the images are analyzed individually in two pattern categories. Some images are in the form of Figure 1-a, where the chest area is symmetrical and elliptical, which are isolated in previous studies by parabolic separation techniques or methods such as a decision tree. Some other images are in the form of figure 1-b, where the chest area is spherical and can be isolated by methods such as circular Hough transform or contour detection. However, some images (e.g., figures 1-c and 1-d) are not morphologically symmetrical, and the simultaneous use of a single technique on both left and right breast might not yield a suitable result. There are other images (e.g., 1-e), where the breast area cannot be accurately detected. In some images (e.g., 1-F), the breast might have been biopsied or operated on shortly after imaging, which required dressing of the chest area. This leads to lack of responsiveness of asymmetry.

Convolutional Neural Network (CNN) is one of the most important deep learning techniques, through which several layers can be trained in a robust manner [28]. CNN is extremely efficient and has been a dominant method in computer vision tasks. Figure 2 shows an overview of CNN

architecture. In general, a CNN encompasses three main substructures, including convolutional layers, pooling layers, and fully connected layers, each having different functions. There are two stages of feed-forward and backpropagationIn the first .[39]CNN for training in each the points the input image enters the network by multiplying ,stage betweenthe ,Afterwards .input and parameters of each neuron convolutional operations are applied to each layer. In the next step, the network's output is calculated, and the output results of the networks are used to tune the parameters related to network training by estimating the network error rate. To this end, the output of the network is compared to the accurate response (favorable solution) using an error function and the error rate is calculated. In the next step, the backpropagation stage initiates based on the estimated error rate. At this stage, the gradient of each parameter is calculated based on the chain rule, and all parameters are changed based on their effect on the error created in the network. The feed-forward stage starts following tuning all parameters. Ultimately, network training is terminated after repeating an appropriate number of steps.

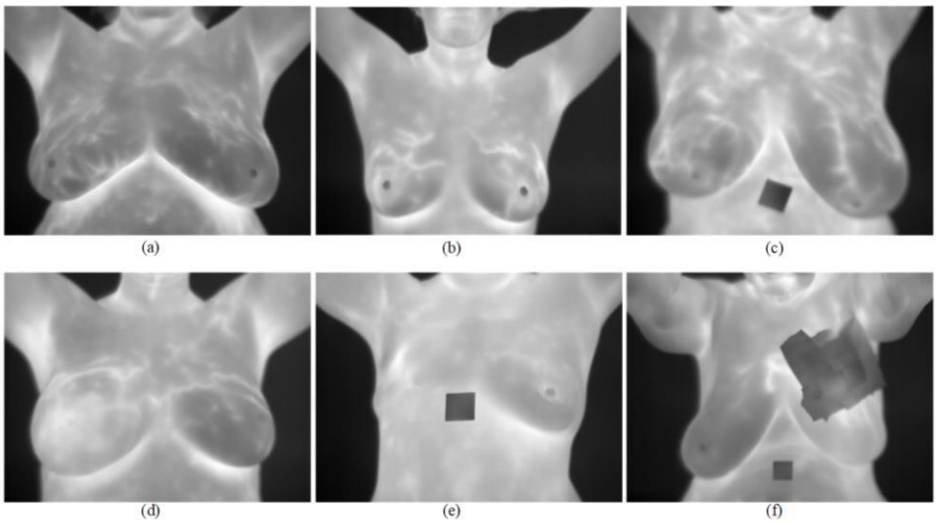

Fig.1 Sample of dataset images

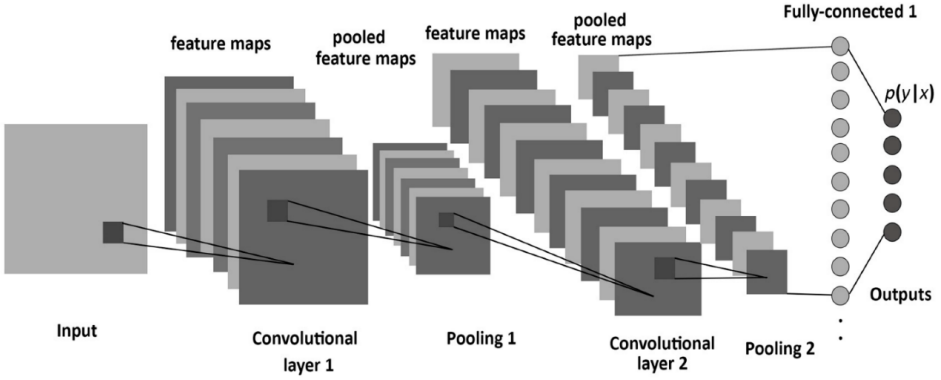

Fig.2 overview of CNN

Now, the following steps must be taken to perform the process based on the mentioned pre-training tasks. In the first stage, the input image is

applied to the network. However, preprocessing must be performed on the images before entering the images into CNNs as input in order to prepare them for processing. In this regard, preparation could involve cutting, changing the size of the image, normalizing the image and zeroing. Therefore, the size of the input image must be changed, if required, to match the dimensions of the input layer of the network. Therefore, the input images must be the standard size of these pre-trained networks in the first stage. Afterwards, the size of the collected color image is changed to 224×224×3. The third dimension is related to the red, green and blue colors. Each entry of the three-dimensional matrix is valued at 0-255, and the minimum and maximum values show the lowest and highest amount of brightness, respectively.

Following the preprocessing of images and adjusting them to the standard size, the first convolutional layer receives the image as input. In the second stage, the images must be convolved by the convolution kernels of the first layer. Each convolution layer comprises three sublayers. In the first sublayer, the pre-processed image is convolved by prespecified convolution kernels. However, these convolution kernels have different amounts and dimensions depending on the type of convolutional network used. In addition, the output dimensions can be the same size as the input image or smaller and larger depending on the image zeroing number. Moreover, several convolution kernels can be used at this stage. Hereon, the convolution operator for data means a local operator that does not change with displacement. The input-to-output mapping for the convolution operator is shown in the equation below:

$$f: RM \times N \times K \rightarrow RM' \times N' \times K'$$

As observed, the convolution operator presented here is mapping from 3D data to another three-dimensional data and is in the set of real numbers. The following equation shows the operation of the convolution operator mathematically:

$$y_{(\iota J \acute{k})'} = \sum_{ijk} y_{(\iota J \acute{k})'} x \quad {}_{i+\iota, J+\acute{J}, k}$$

Where w is a convolutional filter bank and a four-dimensional data. In fact, the fourth dimension is the filter number in the filter bank, and the filter itself is a three-dimensional weight mass. In other words, in any location that the convolutional 3D filter number 1 is placed after sliding over the 3D data, the dot product is found by multiplying the corresponding values in each vector and adding them together. Since only the value of one pixel is attained in each placement position, a 2D feature map is obtained by sliding the filter number 1 over the entire data. Similarly, the next filters are applied, and a 2D feature map is obtained per each filter. The ultimate 3D feature map will be developed by putting the feature maps in a row in the third dimension.

After applying the first convolutional sublayer, the output of the first sublayer is given to the non-linear sublayer. At this stage, a non-linear activation function is applied to the obtained values to achieve higher-level features. In addition, a Rectified Linear Unit (RLU) activation function is applied in all three models as the non-linear function. In general, there is a higher tendency for using RLUs in deep

convolutional networks, compared to other non-linear functions since the mentioned function can be easily solved and does not engage several computational resources. Based on experience, the mentioned function has acceptable accuracy. The following equation defines the RLU:

$$f(x) = x \quad if\ x > 0$$

$$0 \quad else$$

Ultimately, the obtained values are given to the third sublayer to perform the combination operation. In the combination sublayer, a statistical summary of neighboring pixels of the central pixel replaces the value of the central pixel in the merge window. The combination increases the stability of features and reduces sensitivity to unwanted changes. The output data dimensions of this sublayer can be equal to the dimensions of the input data or can have different dimensions depending on the values of zeroing or step parameters. Therefore, another use of this sublayer is maintaining more valuable features and eliminating less significant features in case of a decrease of dimensions in the combination operations. There are different types of combinations, the most popular of which is max pooling, which is described below:

$$y_{ijk} = \max\{y_{i'j'k} : i < i' < i + p, j < j' < j + p\}$$

At this point, an entire convolutional layer is applied to the preprocessed image. This step is repeated a certain number depending on the type of

convolutional network used. After finishing the convolutional layers, one or several fullyconnected layers can be used for the final mapping of the obtained features. While a fullyconnected layer is exactly similar to a convolutional layer, their difference is that no sporadic interactions occur in the former, and just like traditional neural networks, a complete connection is established between this layer and the previous layer. The output of the last layer is a one-dimensional vector, and the number of members of this vector is equal to the number of classification classes. In fact, this layer carries out the classification process. The last fullyconnected layer of all convolutional architecturesis connected to a softmax layer. In fact, a softmax layer performs the classification process in convolutional networks. The layer encompasses a number of neurons that is equal to the number of classes in the classification problem and is used for the final mapping of features and applying the classification [31].

At this point, the output result is used to calculate the network error rate in order to tune its parameters (i.e., network training). To this end, the network's output is compared to a correct solution using an error function, which leads to the calculation of the error rate. The experimental error estimation method is shown below:

$$L(w) = 1/n \sum l(Z_i; f(X_{i;}, w))$$

In the equation above, $l(Z; \hat{Z})$ is a loss function that determines the penalty for incorrect prediction of $\hat{Z}$ instead of Z. In the next stage, backpropagation starts based on the estimated error rate. At this point, the gradient of each parameter is calculated based on the chain rule, and

all parameters change based on their impact on the error made in the network.

$$w^{t+1} = w^t - p_t \frac{\partial f}{}(w^t)$$

The next step (i.e., feed-forward) starts following updating the parameters, and the network training stage terminates after repeating an appropriate number of these stages.

Nevertheless, one of the most important challenges of deep learning and CNNs is their training. Training these networks is very difficult due to the high volume of the layers and their weights and the need for powerful high-speed processors. Therefore, using deep learning networks does not require the training of the entire network, and pre-trained models and algorithms can be applied in this regard. Accordingly, a pre-trained network can be considered instead of designing a deep convolutional neural network from the base, which can be used as an input feature to learn a classification task. During the process, a number of layers that express the generalities of images are frozen based on the existing database so that their weight does not change during training. If there is a low amount of data in the database, only the last layer of the fullyconnected block is trained. Otherwise, a higher number of layers are trained. The latest advancements in the use of CNNs in the field of computer vision have led to the emergence of famous CNN models, including VGG, GoogleNet and ResNet, which can be uploaded. Table 2 shows the characteristics of each layer in the CNN applied in the present study.

Weeused Matlab, which was applied on a high-performance computing

system, which encompasses several computational clusters and integration among them leads to the centralized management of the tasks. Hardware characteristics were: two nodes, 2x NVIDIA® Tesla K80 GPUs, 8x16GB memory, and2x Intel® Xeon® E5-2695 v3 @ 2.30GHzprocessor. It is worth noting that only one specific model was used on an HPC system in all results obtained from various techniques. Table 3 exhibits characteristics related to the neural network adjustment parameters.

3.Results

In this study, three evaluation criteria of accuracy, sensitivity and specificity are used to facilitate the assessment of function of our proposed breast anomaly detection system **:**

The variables of Sen, ACC and SPE are related to sensitivity, accuracy and specificity of the model, whereas the variables of Tp, Tn, Fp and Fn show the number of true positives, true negatives, as well as false positives and false negatives, respectively. Therefore, accuracy will be high if both sensitivity and specificity are high. Out of the 1960 available images, 70% of the data are considered as training data and 30% of the data (588 samples) are considered as a test. Table 2 presents the results related to the sensitivity, accuracy and specificity obtained from the test data **.**

The level of accuracy and area under the ROC Curve will increase when there is an increase in the amount of sensitivity and specificity. Figure 3 includes a diagram related to the amount of accuracy of the validation data used for the three neural network models **.**

In addition, Figure 4 shows the diagram related to the amount of error of the validation data applied for the three neural network models..

Table 2. Evaluation of extracted features using the classifications

Model	TP	TN	FP	FN	ACC	SPE	Sen

GoogleNet	74	426	49	39	85.03%	89.7%	65.48%
Resnet	87	406	36	59	83.8%	91.9%	59.58%
Vgg16	95	405	40	48	85.03%	91.01%	66.43%
Fusion of GoogleNet and Resnet	117	437	18	16	94.21%	96.04%	87.96%
Fusion of GoogleNet and Resnet and Vgg16	126	445	9	8	97.10%	98.01%	94.02%

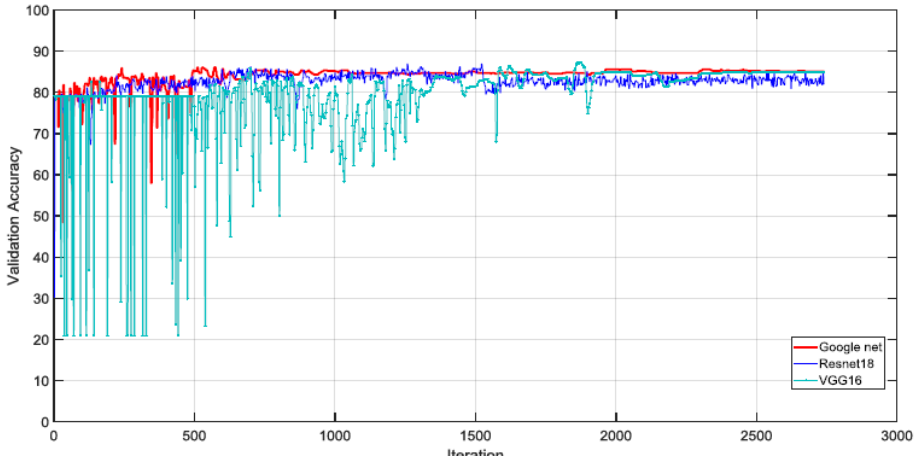

Figure 3. Diagram of accuracy measured at each iteration

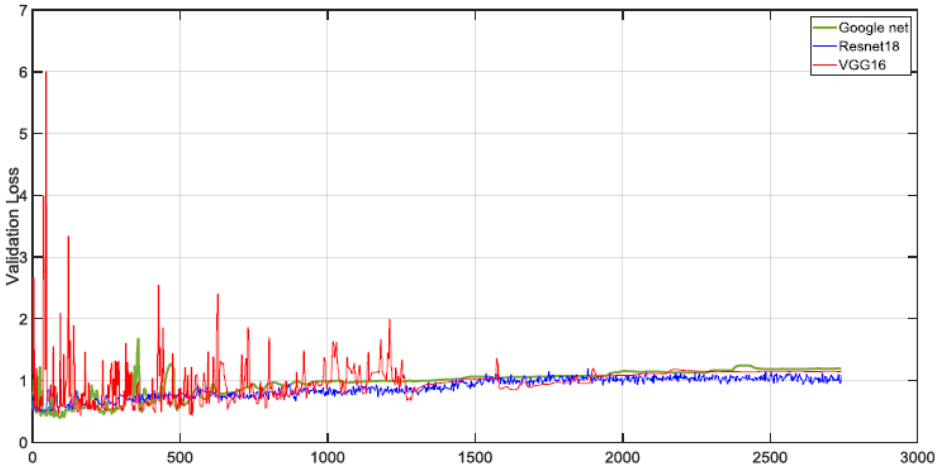

Figure 4. Diagram of error measured at each iteration

Finally figure 5 shows the two proposed method, Accuracy and Error plots.

We propose votingmethod for combining classifiers.

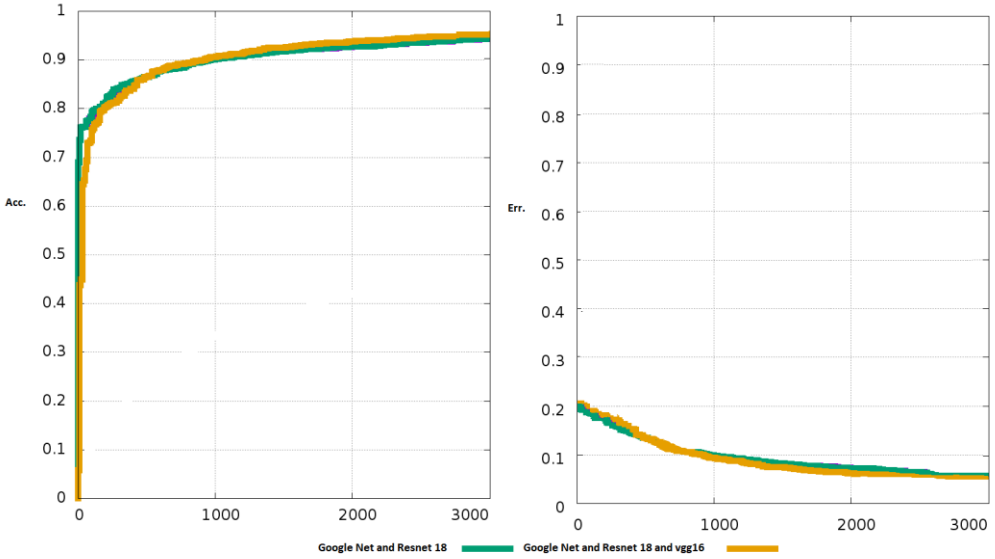

Figure 5. Diagram of accuracy and error measured at each iteration for proposed method

3. Discussion

In the previous section, result of simulations presented. In this section, brief discussion presented.The present study was performed to implement and evaluate a decision support system (DSS) to assist the detection of breast cancer with the use of deep neural networks. The use of deep neural networks in the detection of breast cancer by thermographic images can be discussed in two areas, including the application of deep neural networks in comparison to previous works and analysis of the deep neural network itself .

Most previous works relied on combining stochastic tissue features with a classifier such as SVM, including an article by Acharya and Francis [13, 32]. While the accuracy values of the mentioned research and present study were almost equal, the former used a much lower number of images, compared to the latter [1960 images]. In addition, the accuracy of the proposed model was higher than other articles[16, 17, 21] but similar to studies [12, 20, 23] with a close error rate .

The traditional breast anomaly prediction approach is the two-way analysis of asymmetry, the accuracy of which depends on the proper separation of the left and right breasts from a thermogram. Similar to some studies, Schaefer and Ng proposed the manual extraction of features, which involved a difficult task by the system's developer [16, 17]. Nonetheless, the proposed method detects breast cancer without the need for the feature extraction process in thermal patterns .

The architecture of the proposed method can also be assessed. Some studies have analyzed certain aspects of computational costs (e.g., memory use and inference time), which was similar to a study by Bianco et al., who evaluated the effect of computational costs on detection accuracy.At this point, we compare and analyze the architectures used in the proposed method. VGG16 has five convolutional blocks and one block that is fully connected to four layers. This method is improved by replacing large filters with core size (11 and 5 in the first and second layers, respectively) using size three filters one after the other. However, there are some barriers to the use of the VGG algorithm, including large computational necessities, both in terms of time and memory. On the other hand, the ResNet algorithm is deeper than the VGG algorithm. The

mentioned algorithm also includes feedback, which is due to the error backpropagation education algorithm. The deeper the algorithm becomes and the more increase occurs in the number of layers, the impact of weights and layers might become zero due to the presence of gradient function in the education function by entering deep space. This led to the proposal and use of ResNet algorithms that include feedback [33]. Similar to other neural networks, data placement order during convolutional network training affects the method of network training. Weight tuning is based on calculating the error between the predicted value and the expected value. Evidently, the more network training data there is, the more the model adapts to the training event. If the data consecutively given to the network do not all belong to one class, the network adapts itself during the training in such a way that it can better distinguish the data of these classes. Therefore, the order of data placement is disturbed before entering the convolutional network and data are not entered into a class in a row **.**

Compared to previous studies implemented on this database[34], the proposed method achieves better results. For example in [35] final accuracy is 94.8% and less than our accuracy also in [36] accuracy is 87.3% about 10% less than our method. More than this result we can refer [37] and [38] that achives 86.4% and 90.6% respectively.

Despite all the benefits of deep neural network models, they have some limitations as well. One of the major drawbacks of these methods is the accurate tuning of parameters based on the duration of each program run to achieve favorable results. This forced the authors to run the program several times. Access to a computer hardware system with optimal processing capability is another feature required for such tasks in the field of deep neural networks. Another limitation is the lack of access to large datasets. there is currently only one rich database available to the public. Notably, a deep neural network requires a rich database to achieve good accuracy in such tasks.

4- Conclusion

Breast cancer has been the cause of mortality of many women worldwide. A review of articles revealed that making advancements in the breast cancer detection field could be a great help to this area from

the perspective of computer sciences. While thermography can be an aid in the early detection of breast cancer, it certainly has advantages and disadvantages. Different architectures of CNN and fusion with voatingmethod were used in the present study to classify breast cancer patients using thermography images. The present study aimed to make significant improvements in the accuracy of breast anomaly detection by using deep neural networks and raw thermal images without extracting features. Three pre-trained architectures of Google Net, ResNet18 and VGG16 and fusion of these classifiers with voating method, were used to classify breast cancer in thermal images. According to the results, the GoogleNet architecture was the most table one among the three architectures and yielded excellent results regarding the classification of thermography images. Meanwhile, VGG had a better performance, compared to ResNet. Therefore, CNNs were recognized to be extremely efficient in breast cancer screening. The most important novelty of the present study, compared to previous works in the field of breast thermography, was the lack of effective involvement of different breast morphologies in the results and the lack of manual extraction of features. In addition, the accuracy of the results was due to the proper tuning of the parameters and choosing an appropriate loss function . It is recommended that the classification accuracy be increased by using larger datasets for training and validation. In fact, CNNs will yield better results if there is an increase in the number of training samples. In addition, the deep learning approach will definitely perform well with the use of color datasets, compared to grayscale datasets for static protocols.

Refrences:

1-Hossein Gholizadeh M, GhayoumiZadeh H, FatehiMarj H, Ahmadinejad N. A Study of Deep Convolutional Neural Network for Diagnosing Breast Cancer in Thermographic Images. JundishapurSci Med J 2019; 18(6):615-629.

2-Díaz-Cortés M-A, Ortega-Sánchez N, Hinojosa S, Oliva D, Cuevas E, Rojas R, et al.. Infrared Physics & Technology. 2018;93:346-61.

3-Suganthi S, Ramakrishnan S. Anisotropic diffusion filter based edge enhancement for segmentation of breast thermogram using level sets. Biomedical Signal Processing and Control. 2014;10:128-36.

4-Kandlikar SG, Perez-Raya I, Raghupathi PA, Gonzalez-Hernandez J-L, Dabydeen D, Medeiros L, et al. Infrared imaging technology for breast

cancer detection–Current status, protocols and new directions. International Journal of Heat and Mass Transfer. 2017;108:2303-20.

5-Ng E-K. A review of thermography as promising non-invasive detection modality for breast tumor. International Journal of Thermal Sciences. 2009;48(5):849-59.

6-Kuhl CK, Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. Journal of clinical oncology. 2005;23(33):8469-76.

7-Lahiri B, Bagavathiappan S, Jayakumar T, Philip J. Medical applications of infrared thermography: a review. Infrared Physics & Technology. 2012;55(4):221-35.

8-Ghafarpour A, ZareI ,Zadeh HG, Haddadnia J, Zadeh FJS, Zadeh ZE, et al. A review of the dedicated studies to breast cancer diagnosis by thermal imaging in the fields of medical and artificial intelligence sciences. Biomedical Research. 2016;27(2).

9- Keyserlingk J, Ahlgren P ,Yu E, Belliveau N, Yassa M. Functional infrared imaging of the breast. IEEE Engineering in Medicine and Biology Magazine. 2000;19(3):30-41.

10-Gogoi UR, Bhowmik MK, Bhattacharjee D, Ghosh AK. Singular value based characterization and analysis of thermal patches for early breast abnormality detection. Australasian physical & engineering sciences in medicine. 2018;41(4):861-79.

11-Borchartt TB, Conci A, Lima RC, Resmini R, Sanchez A. Breast thermography from an image processing viewpoint: A survey. Signal Processing. 2013;93(10):2785-803.

12-Borchartt TB, Resmini R, Conci A, Martins A, Silva AC, Diniz EM, et al., editors. Thermal feature analysis to aid on breast disease diagnosis. Proceedings of 21st Brazilian Congress of Mechanical Engineering—COBEM22 ;

2011.

13-Acharya UR, Ng EY-K, Tan J-H, Sree SV. Thermography based breast cancer detection using texture features and support vector machine. Journal of medical systems. 2012;36(3):1503-10.

14-Nicandro C-R, Efrén M-M, MaríaYaneli A-A, Enrique M-D-C-M,

Héctor Gabriel A-M, Nancy P-C, et al. Evaluation of the diagnostic power of thermography in breast cancer using bayesian network classifiers. Computational and mathematical methods in medicine. 2013;2013.

15-Pramanik S, Bhattacharjee D, NasipuriM ,editors. Wavelet based thermogram analysis for breast cancer detection. Advanced Computing and Communication (ISACC), 2015 International Symposium on; 2015: IEEE.

16-Schaefer G, Závišek M, Nakashima T. Thermography based breast cancer analysis using statistical features and fuzzy classification. Pattern Recognition. 2009;42(6):1133-7.

17-Ng E, Kee E. Integrative computer-aided diagnostic with breast thermogram. Journal of Mechanics in Medicine and Biology. 2007;7(01):1-10.

18-Mookiah MRK, Acharya UR, Ng E. Data mining technique for breast cancer detection in thermograms using hybrid feature extraction strategy. Quantitative InfraRed Thermography Journal. 2012;9(2):151-65.

19-Francis SV, Sasikala M, Saranya S. Detection of breast abnormality from thermograms using curvelet transform based feature extraction. Journal of medical systems. 2014;38(4):23.

20-Araújo MC, Lima RC, De Souza RM. Interval symbolic feature extraction for thermography breast cancer detection. Expert Systems with Applications. 2014;14(15):6728-37.

21-Haddadnia J, Hashemian M, Hassanpour K. Diagnosis of breast cancer using a combination of genetic algorithm and artificial neural network in medical infrared thermal imaging. Iranian Journal of Medical Physics. 2012;9(4):265-74.

22-Sathish D, Kamath S, Prasad K, Kadavigere R, Martis RJ. Asymmetry analysis of breast thermograms using automated segmentation and texture features. Signal, Image and Video Processing. 2017;11(4):745-52.

23-Lashkari A, Pak F, Firouzmand M. Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications. Journal of medical signals and sensors. 2016;6(1):12.

24-Zadeh HG, Haddadnia J, Montazeri A. A model for diagnosing breast

cancerous tissue from thermal images using active contour and Lyapunov exponent. Iranian journal of public health. 2016;45(5):657.

25-Ghayoumi Zadeh H, Montazeri A, AbaspurKazerouni I, Haddadnia J. Clustering and screening for breast cancer on thermal images using a combination of SOM and MLP. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. 2017;5(1):68-76.

26-PROENG. Image Processing and Image Analyses Applied to Mastology, 2012,http://visual.ic.uff.br/en/proeng ./

27-Motta L, Conci A, Lima R, Diniz E, Luís S, editors. Automatic segmentation on thermograms in order to aid diagnosis and 2D modeling. Proc of 10th Workshop emInformáticaMédica; 2010.

28-Chambolle A, De Vore RA, Lee N-Y, Lucier BJ. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. IEEE Transactions on Image Processing. 1998;7(3):319-35.

29-Fernández-Ovies FJ, Alférez-Baquero ES, de Andrés-Galiana EJ, Cernea A, Fernández-Muñiz Z, Fernández-Martínez JL, editors .Detection of Breast Cancer Using Infrared Thermography and Deep Neural Networks. International Work-Conference on Bioinformatics and Biomedical Engineering; 2019: Springer.

30-Albelwi S, Mahmood A. A framework for designing the architectures of deep convolutional neural networks. Entropy. 2017;19(6):242.

31-Spanhol FA, Oliveira LS, Petitjean C, Heutte L, editors. Breast cancer histopathological image classification using convolutional neural networks. 2016 international joint conference on neural networks (IJCNN); 2016: IEEE.

32-Wakankar AT, Suresh G, editors. Automatic diagnosis of breast cancer using thermographic color analysis and SVM classifier. The International Symposium on Intelligent Systems Technologies and Applications; 2016: Springer.

33-Szegedy C, Ioffe S, Vanhoucke V, Alemi AA, editors. Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-First AAAI Conference on Artificial Intelligence; 2017.

34- Yousefi B, Castanedo CI, Maldague XP. Low-rank Convex/Sparse Thermal Matrix Approximation for Infrared-based Diagnostic System.

arXiv preprint arXiv:2010.06784. 2020 Oct 14.

35- B. Yousefi, S. Sfarra, C. I. Castanedo, and X. P. Maldague, "Comparative analysis on thermal non-destructive testing imagery applying candid covariance-free incremental principal component thermography (ccipct)," Infrared Physics & Technology, vol. 85, pp. 163–169, 2017.

36- C. Ibarra-Castanedo, J. R. Tarpani, and X. P. Maldague, "Nondestructive testing with thermography," European Journal of Physics, vol. 34, no. 6, p. S91, 2013.

37- B. Yousefi, C. Ibarra-Castanedo, and X. P. Maldague, "Infrared nondestructive testing via semi-nonnegative matrix factorization," in Multidisciplinary Digital Publishing Institute Proceedings, vol. 27, no. 1, 2019, p. 13.

38- B. Yousefi, C. Ibarra-Castanedo, and X. Maldague, "Application of sparse non-negative matrix factorization in infrared non-destructive testing," in Proceedings of the 3rd Asian Conference on Quantitative Infrared Thermography, Tokyo, Japan, 2019, pp. 1–5.