Online ISSN: 2682-2628
Print ISSN: 2682-261X

IJCBR

International Journal of Cancer and Biomedical Research

https://jcbr.journals.ekb.eg

Editor-in-chief
Prof. Mohamed Labib Salem, PhD

Recognition of ultrasound images of breast cancer based on discrete sinc transform and deep learning

Mohamed Moustafa Gad

EKB Egyptian Knowledge Bank
بنك المعرفة المصري

PUBLISHED BY
EACR EGYPTIAN ASSOCIAN FOR CANCER RESEARCH
Since 2014

INTERNATIONAL JOURNAL OF CANCER AND BIOMEDICAL RESEARCH (IJCBR)

RESEARCH ARTICLE

Recognition of ultrasound images of breast cancer based on discrete sinc transform and deep learning

Mohamed Moustafa Gad

Biomedical Engineering Department, Medical Research Institute, Alexandria University, Alexandria, Egypt

ABSTRACT

Background: : The most common cause of cancer-related death among women all over the world is breast cancer. Altough ultrasound emaging I useful in diagnosis of the patients, itsaccuracy needs more improvement. Aim: This paper presents a new automated system to precisely classify images of ultrasound of breast cancer to help radiologists for better diagnosis. Materils and Methods: Deep learning neural network algorithm was used in the classification step. Features of images are extracted using discrete wavelet transform. After that, discrete sinc transform (DSNT) was obtained. Then gray-level co-occurrence matrix was used. After that mean was calculated. Several discrete transforms are applied as discrete cosine transform (DCT), discrete tan transform (DTT), discrete sine transform (DST) and discrete sinc transform (DSNT). DSNT was chosen because it has the largest accuracy rate after the classification step. Results: The obtained accuracy percent is 99%. The specificity rate is 98%. The sensitivity rate is 100%. F-measure rate is 99.0%. F-score is 0.99. Conclusion: Our study point pout that future studies are needed to devlope a new feature extraction method to achieve higher accuracy rate.

Keywords: Cosine transform, Discrete cosine transform, Sinc transform, Sine transform, Tan transform

Editor-in-Chief: Prof. M.L. Salem, PhD - Article DOI: 10.21608/jcbr.2021.65002.1185

ARTICLE INFO

Article history
Received: April 24, 2021
Revised: July 21, 2021
Accepted: October 12, 2021

Correspondence to
Dr. Mohamed Moustafa Gad,
Biomedical Engineering Dept., MRI,
Alexandria University, Egypt
Tel.: +201005090782
Email: drmazmi@alexu.edu.eg

INTRODUCTION

The most common cause of death among women in the world is due to breast cancer. In 2020, the number of detected women with breast cancer was 2.3 million and number of deaths are 685,000 globally. Early detection is preferred to reduce the death rate according to WHO. Many researches are conducted about classification of breast cancer using computer techniques. Rodrigues et al. in 2006 used computer techniques to classify breast cancer images. Attia et al. in 2012 used optical image analysis for diagnosing breast cancer. Raha et al. in 2017 obtained 96.5% accuracy percent of classification of breast mass of ultrasound images. Shastri et al. in 2018 classified images of breast cancer using mammography applying descriptors of texture exploiting. Berbar in 2018 extracted features using hybrid methods for classifying breast masses. Adel et al. in 2019 obtained 94.12% when classifying ultrasound images of breast cancer. Durga and her colleagues in 2019 obtained 98% of accuracy percent using deep learning. In 2019, Chen et al. applied new image analysis method to obtain features of mammographic images for diagnosing breast cancer. The aim of the study is to classify ultrasound images of breast to know whether this patient is benign or malignant by using computer techniques.

METHODS

Wavelet Transform

Signal is decomposed into group of band pass filters as a function of time by wavelet transform. Wavelet transform is used for compressing data and reducing noise from data. The wavelet transform has faster speed of computation than Fourier transform. It can also represent the signal in multidimensional domains in time domain and frequency domain. It is suitable for biomedical signals because they are non-stationary signals.

Discrete Transforms

Discrete cosine transform (DCT) is used for data reduction as shown in equation (1). Discrete tan transform (DTT), discrete Sine transform (DST), DSNT are deduced from DCT as shown in the next equations (2-6)respectively. They can be used for applications of data reduction also.

$$y(k) = w(k)\sum_{n=1}^{N} x(n)\cos(\frac{\pi}{2N(2n-1)(k-1)}) \quad (1)$$

$$y(k) = w(k)\sum_{n=1}^{N} x(n)\tan(\frac{\pi}{2N(2n-1)(k-1)}) \quad (2)$$

$$y(k) = w(k)\sum_{n=1}^{N} x(n)\sin(\frac{\pi}{2N(2n-1)(k-1)}) \quad (3)$$

$$y(k) = w(k)\sum_{n=1}^{N} x(n)\mathrm{sinc}(\frac{\pi}{2N(2n-1)(k-1)}) \quad (4)$$

$$\mathrm{sinc}(t) = \begin{cases} \frac{\sin(\pi t)}{\pi t}, t \neq 0 \\ 1, t = 0 \end{cases} \quad (5)$$

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}}, k = 1 \\ \sqrt{\frac{2}{N}}, 2 \leq k \leq N \end{cases} \quad (6)$$

N is the length of x. The size of x and y are the same. MATLAB vectors run from 1 to N instead of from 0 to N–1. So, the series is indexed from n=1 and k=1 instead of the usual n=0 and k=0. All these discrete transforms are used as shown below in analysis and extracting features from ECG signals (Jain in 1989 and Pennebaker in 1993).

Co-occurrence Matrix

The statistical method of gray-level co-occurrence matrix (GLCM) is used for examining texture that describes the spatial relationship between pixels, also known as the gray-level spatial dependence matrix. A co-occurrence matrix can also be known as co-occurrence distribution. The gray-scale image is used to produce GLCM. GLCM can be used to calculate how often a pixel with gray-level (grayscale intensity or Tone) value i occurs either horizontally, vertically, or diagonally to adjacent pixels with the value j (Mokji and Abu Bakar in 2007).

$$C_{m,n,\phi} = \sum_{x}\sum_{y} P\{I(x,y) = m \,\&\, I(x \pm d\phi_0, y \mp d\phi_1) = n\} \quad (7)$$

Where d is relative distance between pixel pair with relative orientation ϕ (0,45,90,135). I is image to be considered, m represents the gray level of pixel(x,y) where n represents the gray level of pixels ($x \pm d\phi_0$, $y \pm d\phi_1$) with L level of gray tones where 0≤ x≤M-1, 0≤y≤N-1 and 0≤ m, n≤L-1. $C_{m,n}$is GLCM. P{.}=1 if the argument is true and otherwise, P{.}=0.

Classification Method

Deep learning is used here as classification method. It is a type of machine learning techniques (Faust et al. in 2018). Here, learning uses a network to learn and extract features from each hidden layer of the selected neurons. Number of hidden layers in the structure of Conventional Neural Network (CNN) is large so it is called deep learning. CNN overcomes the disadvantages of Neural Network (NN). CNN is considered feed forward NN. It is composed from 3 layers pooling, convolution, and fully-connected layers. The Recurrent Neural Network (RNN) is architecture of deep learning. RNN is feed-forward recurrent network that means the network makes a routine task with the output which depends on the previous computations. The memory unit which operates this task. The Long Short-Term Memory (LSTM) network is type of RNN. The LSTM consists from memory block with three gates. These gates are the input, output, and forget gate. They are used to control forgetting or storing information from the network. This process is repeated for every input. The new information is decided to be stored and updated in the cell state by input gate. Output gate decides what information is used according to the cell state. Forget gate removes redundant information from the cell state. LSTM layer considers the sequence of time in the forward direction. The bidirectional LSTM layer BILSTM can consider time sequence of time in forward and backward direction.

Feature Extraction

The ultrasound images of breast are collected from 250 patients (Rodrigues in 2017). 150 images of patients are selected for training. 100 images of patients are selected for testing. As shown in Figures 1,2 ultrasound images of breast of cancer and benign patient. Features of ultrasound images are extracted (Figure 3).

Figure 1. Cancer of breast ultrasound

Figure 2. Benign image of breast patient

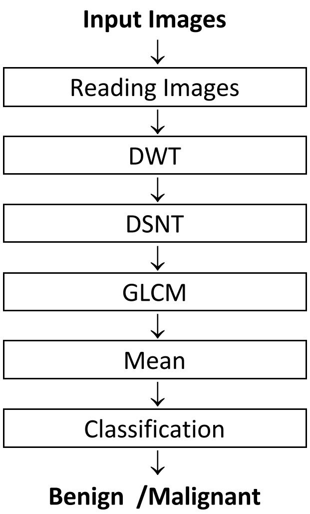

Figure 3. Feature Extraction

First ultrasound images are read by MATLAB then discrete wavelet transform is calculated. After that discrete transforms are extracted such as DCT or DST or DSNT or DTT. Then GLCM is calculated. After that mean is calculated.

Classification is achieved by deep learning (Figure 4). Number of input layers equals 2. Number of hidden units is 100. Number of output layers is 2. The gradient threshold equals 1. The maximum number of epochs is 100. The minimum batch size is chosen to be 150 to decrease the amount of padding.

RESULTS AND DISCUSSION

As shown in Table 1, there are comparisons between several transforms. DSNT is chosen because it achieved the highest accuracy. The obtained accuracy of training process is 96.7% as shown in Figure 5. The obtained accuracy of testing process is 99% (Figure 6). The area under the curve is shown in Figure 7. It is near 1. The specificity rate is 98%. The sensitivity rate is 100%. F-measure rate is 99.0%. F-score is 0.99.

CONCLUSION

The most common cause of death of women in all over the world is due to breast cancer. Here, this paper presented new automated system by computer algorithms for classification of ultrasound images of breast cancer to help radiologists. Deep learning is used as a classifier. Features of images are extracted by applying discrete wavelet transform then DSNT. After that then GLCM is obtained then mean is calculated. The obtained accuracy percent is 99%. The specificity rate is 98%. The sensitivity rate is 100%. F-measure rate is 99.0%. F-score ratio is 0.99. In future, a new method of extraction will be introduced to obtain higher accuracy percent of classification of ultrasound images of breast cancer.

CONFLICTOF INTEREST

All authors declare no conflicts of interest.

FUNDING

No fund was received for this work.

Table 1. Comparison of Calssifiction rateof Several dicrete Transforms

	Accuracy%	Sensitivity%	Specifity%	Precision%	Recall%	F_Measure%	F-score
DCT	88	100	76	80	100	89	0.89
DST	90	100	80	83	100	90	0.9
DTT	89	100	78	81	100	90	0.9
DSNT	99	100	98	98	100	99	0.99

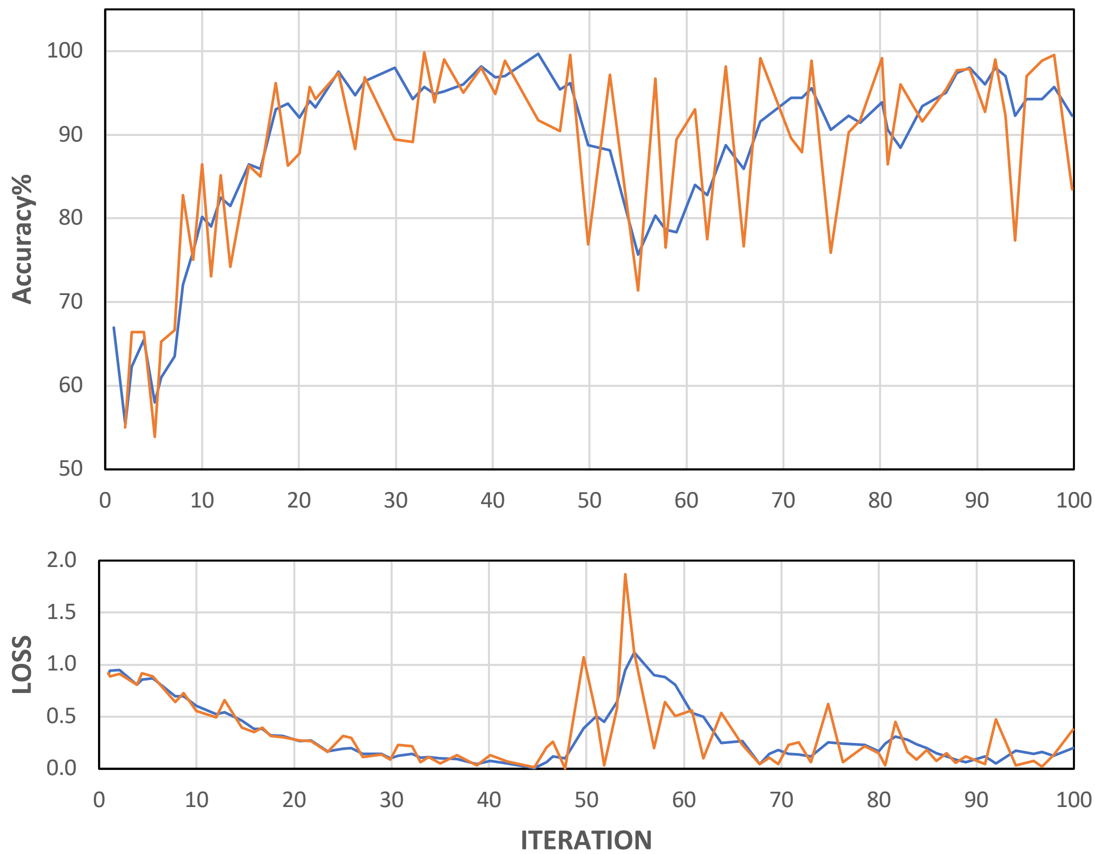

Figure 4. Training Dataset using deep learning

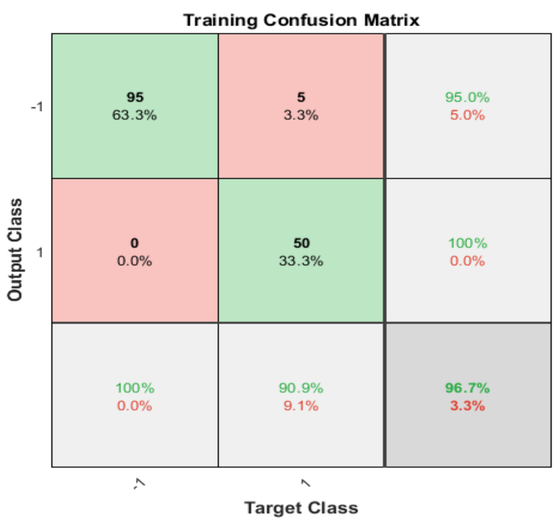

Figure 5. Training confusion matrix

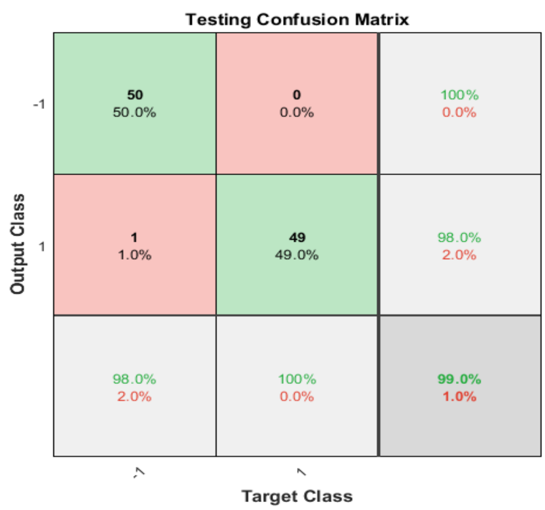

Figure 6. Testing Confusion matrix

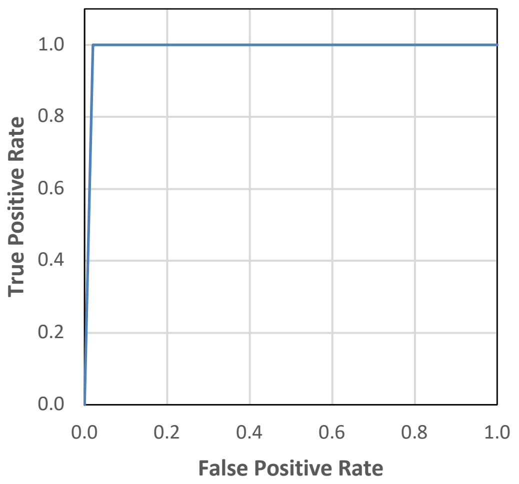

Figure 7. Area Under The Curve

REFERENCES

AttiaSJ ,Blackledge JM , Abood ZM and Agool IR (2012). Diagnosis of Breast Cancer by Optical Image Analysis, ISSC 2012, NUI Maynooth, June 28–29, 1-4.

Raha P, Menon RV, Chakrabarti I (2017). Fully Automated Computer Aided Diagnosis System for Classification of Breast Mass from Ultrasound Images, IEEE international conference of Wireless Communications, Signal Processing and Networking WISPNET, 48-51.

Shastri AA, Tamrakar D, Ahuja K (2018). Density-wise two stage mammogram classification using texture exploiting descriptors, Expert Systems With Applications Vol.99, 71–82.

Berbar MA (2018). Hybrid methods for feature extraction for breast masses classification, Egyptian Informatics Journal, 19, 63–73.

Adel M ,Kotb A, Farag O, Darwees MS, Mostafa H (2019). Breast Cancer Diagnosis Using Image Processing and Machine Learning for Elastography Images,2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), 1-4.

DurgaKameswari SS, Vijayakumar V (2019). A CNN Based Breast Tumor Classifier Using Mendeley BUS Dataset", International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8:6, 1340-1343.

Chen X, Zargari A , Hollingsworth AB ,Liu H , Zheng B, Qiu Y (2019). Applying a new quantitative image analysis scheme based on global mammographic features to assist diagnosis of breast cancer, Computer Methods and Programs in Biomedicine (179), 1-4.

Jain AK (1989). Fundamentals of Digital Image Processing,Englewood Cliffs, NJ: Prentice-Hall.

Pennebaker WB, Mitchell JL. (1993). JPEG Still Image Data Compression Standard, New York, NY: Van Nostrand Reinhold. Chapter 4.

Mokji MM, Abu Bakar SAR (2007).Gray Level Co-Occurrence Matrix Computation Based OnHaar Wavelet,Computer Graphics, Imaging and Visualization (CGIV 2007),

Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR, (2018). Deep learning for healthcare applications based on physiological signals: A review , Computer Methods and Programs in Biomedicine, 161, 1-13, ISSN 0169-2607.

Rodrigues, Paulo S (2017). Breast Ultrasound Image, MendeleyData, http://dx.doi.org/10.17632/wmy84gzngw

Rodrigues PS, Giraldi GA, Chang RF, Sur JS (2006). Non-extensive entropy for CAD systems of breast cancer images, The 19th Brazilian symposium on computer graphics and image processing, 121-128.