As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology
Volume 9, Issue 10, pages 53-61.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol9issue10.html
DOI: https://doi.org/10.46501/IJMTST0910007

Check for updates

Detection of IDC in Breast Histopathology Images

S. Suhasini[1] | Dr. M. Sumender Roy, Professor[2]

[1] Department of CSE, Lenora Engineering College, Ramapachodavaram
sabarapusuhasini1993@gmail.com
[2] Professor, Department of CSE, Lenora Engineering College, Ramapachodavaram
sumender1@yahoo.com

To Cite this Article
S. Suhasini and Dr. M. Sumender Roy. Detection of IDC in Breast Histopathology Images, International Journal for Modern Trends in Science and Technology, 2023, 9(10), pages. 53-61. https://doi.org/10.46501/IJMTST0910007

Article Info
Received: 17 October 2023; Accepted: 28 October 2023; Published: 30 October 2023.

ABSTRACT

Detection of Invasive Ductal Carcinoma (IDC) is a difficult, time taking and mandatory job. Moreover precise detection of IDC is important to give proper medical treatment to the patients. In this medical situation, Deep Learning (DL) approach is much reliable. The evolution of Deep Learning is taken off in the past few years due to rise in more and more data, better hardware and it supports with robust computational modelling algorithms to train a framework. Based on the advantage of DL, we introduce a framework to automatically detect IDC in Breast Histopathology Images (BHI) of Breast Cancer (BC) either as a malign or benign. This ResNet is provided with labelled BHI data to extract common features and trained to classify the future images thus able to detect the presence of IDC in breast tissue lesion image. It takes less effort, gives better accuracy within less time. Thus our framework helps to speed up a pathologist's work and provides diagnosis support.

Keywords—Breast Histopathology Images, Breast tissue lesion, Computational modelling algorithms, Deep Learning, Invasive Ductal Carcinoma

1. INTRODUCTION

Invasive Ductal Carcinoma is an attention requiring worldwide problem in women. It is a widely occurring type of Breast Cancer which leads with approximately 80% among diagnosed cases [1, 2]. The top most cancer threats overall in women are lung cancer and breast cancer respectively [3]. Many countries like Belgium, France, Lebanon have been suffering with breast cancer and millions of new cases were registered in 2018 [4]. It is also registering many cases in Indian cities like Mumbai, Delhi and Bangalore and also in rural areas [5]. Almost 60% breast cancer Indian patients were not diagnosed until the disease was in its late stage due to the fatal lack of awareness on it [6]. IDC in U.S. women is as common as 1 over 8. So early detection and treatment is crucial, in order to reduce cancer mortality [7]-[9].

Invasive ductal carcinoma alias infiltrating ductal carcinoma is caused because of the abnormal growth of the cells in the milk ducts which breaks the walls of them and starts to spread to the surrounding tissues or lesions of the breast. The suspected breast lesion is digitized as histopathology images for the automatic analysis and detection.

In early machine learning algorithms doing the feature extraction, feature segmentation and classification tasks are very tedious and error-prone as they require handcrafting features to do them. Later, machine learning algorithms are evolved as deep learning to give better results. Deep learning, nothing but a subset of machine learning, is data-intensive which uses layers of algorithms on high dimensional input data, uses suitable architectures to yield better accuracy with simplicity [10].

In this paper, we presented an approach to detect the invasive ductal carcinoma in histopathology images using deep learning ResNet model for feature extraction.

2. RELATED WORK

To detect the IDC in the breast lesion, a pathologist does the biopsy on the suspected sample and histopathological analysis is performed on it [7]. The manual process takes much long time and cumbersome, sometimes accuracy is not as expected. Here medical image analysis techniques are very helpful [11].

The previous research on the detection of breast cancer by the researchers has been much helpful to evolve further. Medical imaging development is led to visualize the different organs of the body and its activities. This helps in diagnosis of Breast histopathological changes [12]. Earlier the system diagnosis tools are more helpful in detecting carcinoma to the pathologists work and can also be treated as second opinion to improve the accuracy in short time [13]. Oberai et al. [14], stated that using Breast ultrasound elastography imaging technique, the images are analyzed in such a way that first need to determine the displacements in the tissue, thereafter need to determine physical laws of machines and spatial distribution of mechanical properties and then the appropriate features are identified and quantified to classify the tissue is malign or not. The classical machine learning algorithms like support vector machine [15], random forest algorithms [16], logistic regression [17] etc, has strong position in BC detection [18]. Further, a trained Convolutional Neural Network model learns the filters and also requires much lower pre-processing [19] to find and figure out the abnormalities in the breast lesion image patches [14], [20], [21]. Angel Cruz-Roa et al. [21] stated that their deep learning CNN model trained with 113 slide images for training and 49 images for testing to identify IDC got an accuracy of 84.23%.

Almost in each field like image and video recognition, recommendation systems, natural language processing the deep learning has been showing its potential and greatly in the analysis of medical images. Many Medical imaging produced images go under diagnosis require finding the abnormalities in them, quantify measurement and changes over time. Deep learning algorithms have their potential stand in giving accurate diagnosis and interpretation in automated medical image analysis tools [22], [23].

3. DEEP LEARNING

Deep learning has progressed quite fast over the past few years in almost every domain when high configured computers with the ability of quick data processing, GPU (graphics processing units) were available to perform high calculations within short time. Nowadays, deep neural networks are handling huge complex training data like images and sound and producing efficient results with amazing speed [24].

Deep learning architectures include convolutional neural networks, recurrent neural networks, residual networks, etc. In practice, deep layers allow for smooth activation functions to provide learned hyper-planes which find the underlying complex interactions and regions without having to see an exponentially large number of training samples.

A. *Convolutional Neural Network (CNN or ConvNet)*

CNN is a feed-forward neural network overcomes fully connected neural network drawbacks with an ease of training and generalization as shown in Figure-1. It is much better than the networks with full connectivity between adjacent layers. It contains various different components as layers, activation functions and regularization functions which are described below [23], [25]-[27].

CNN Layers: Convolutional Neural Network contains number of various layers in between input layer and output layer as hidden layers and those hidden layers consist of Convolutional Layers, Pooling Layers, Fully Connected Layers and Softmax or Logistic Layer.

Input Layer: An image is a matrix of pixel values. This layer contains image data as input according to our proposed model. Image data is represented by three dimensional matrices. So the image matrix is flattened as a single column vector and fed into the input layer. If the input image is n by n and if there are "m" training examples then dimension of input will be ($n \times n$, m).

Convolutional Layers: These layers are responsible for learning low level features (such as edges, corners etc.), middle level features (the features with the valuable information of an image by connecting two or more low level features such as colour, gradient orientation with visual content of image) and high level features (such as the features built on top of previous analyzed features to detect objects and larger shapes in the image) using relevant kernels or filters($f \times f$ matrices that slid through the image during forward pass) to produce feature maps. The output of these layers is often forwarded to other convolutional layers. The number of convolutional layers in the network depends on the network requirements. The number of filters that are used to produce multiple feature maps is called as depth of feature maps. Mathematically, the feature value at location (i, j) in the feature map of layer is calculated by $\times$ + where x is input patch value at layer at location (i, j).

Pooling Layer: This layer decreases the computational burden through decreasing the resolution or spatial size of the convolved features by reducing the number of connections between the convolutional layers. To achieve this, appropriate pooling methods [28] (max pooling, average pooling, mixed pooling, etc.,) are used based on network requirement.

Fully Connected Layer: This layer exists before the Softmax layer and the number of these layers varies depending on requirement. As the name indicates in this layer every neuron is connected to every neuron in the previous layer to achieve linearity in the networks.

Softmax or Logistic Layer: This layer is a linear layer with classifier helps to determine the final output of the CNN depending on highest probability value. The number of neurons in this layer is depends on number of classifications that have in the neural network requirement. The Softmax classifier is most appropriate for multi classification and logistic for binary.

Output Layer: This layer contains the label which is in the form of one-hot encoded.

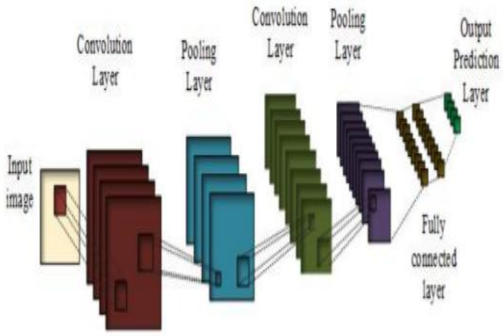

Figure-1: Model of Convolutional neural networks.

B. Activation Functions

The activation functions work on each neuron and find the output in neural network. Each activation function has its own output range to normalize. The values of some activation functions are 0 or any positive integer or the values ranges between -1 and 1 or 0 and 1. The output of these functions decides to make the neuron dead or alive, based on whether the neuron's input (which is from a single or set of previous neurons) is relevant for the model's prediction.

A list of few activation functions are logistic function (sigmoid curve), hyperbolic tangent function (tanh), Maxout, ELU, rectified linear units (ReLU) and LeakyReLU [29], [30]. The most successful activation function of deep learning networks is a nonlinear activation function called Rectified Linear unit (ReLu). It replaced linear activation functions due to the advantages in accelerating the convergence speed and to solve the exploding or vanishing gradient descent problem [31]. ReLU ensures faster computations compared to other functions because it does activate positive neurons only but not all neurons at the same time [24] and thus it is well to rapid dimension reduction [23], [10]. As ReLu function output either zero or positive integer, it converts negative values to zero and sustains the positive values.

C. Backpropagation

The backpropagation is a strong algorithm that makes the network learns the features. It uses of errors in training deep learning models to measure how fast an activation function is changing. It calculates the errors and gradient of cost function. This cost function is computed in terms of weights and bias. If the cost of a particular output neuron is low means it does not depending on that neuron. If the cost is high then update the weights and biases. Weight in the final layer will learn slowly if the output neuron is either low or highly activated or the input neuron has low activation value. This is a situation called saturation where the output neuron stopped or very slowly learning. This algorithm adjusts the weights and biases by performing backward passes through each layer until the error becomes reduced. Backpropagation works faster than classical deep networks to learn the features [23], [26]. Stochastic Gradient descent is an algorithm which uses the approximation gradient to perform fast learning [32].

D. Dropout Regularization

Dropout Regularization controls overfitting. It removes subsets of features randomly based on the probability values in each iteration of training procedure. It proves that computations are reduced because of dropping, thus generalizing well and making feature independent and effective. Mathematically it is represented as, if p indicates the probability to not to drop the feature then the probability (1- p) of a feature indicates the probability to drop it in the network. When a feature is dropped then its activation turned to 0. Each feature has its own probability to dropout. The default optimal value of p is 0.5 [23], [33].

E. Batch Normalization

Batch normalization is a technique for improving optimization. While in dropout regularization, it produces some noise into neural network. Then batch normalization makes the network learn to deal with noise and to generalize well. Batch Normalization improves the training or learning speed and especially when dealing with large dataset. Both normalization and dropout can be used at the same time [10].

F. Residual Network (ResNet)

The Residual Neural Network [34] is a popular deep Neural Network. It differs with the regular deep neural network with the key point 'shortcut connections'. These shortcut connections are those which are parallel to the normal convolutional layers. The basic residual building block is depicted in Figure-2.

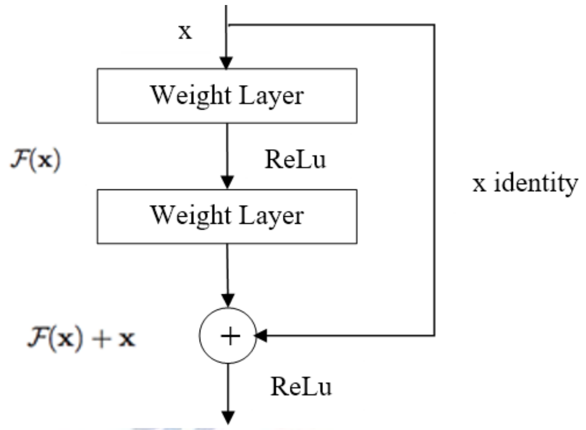

Figure-2: A block diagram of Residual building block.

In ResNet, the residual blocks are connected in a series along with skip connections that are from a residual block to adjacent residual block. Here the shortcut connections do identity mapping that is just to add the output from the previous layer to the layer ahead. If the input size and output size for a basic block is same then skip connection do just the identity mapping else it uses padding or max pooling to adjust the size.

The process of a ResNet 50 is shown as block diagram in Figure-3. Contrary to convolution layers, these shortcut connections are always alive and the gradients can easily backpropagate through them, which results in a faster training. ResNet has reduced complexity and expensive design problems when the network becomes deeper to learn more complex features. The ResNet reduces these problems by using the proper optimization function and normalized initialization of the network while training very deep neural networks [35]. Skipping makes the network simple because of few layers participates in training, but when depth increases in the plain nets it can cause over fitting error. This obviously increases the learning speed and solves the gradient problem either it is vanishing or exploding. The residual blocks reduce the extra training data to recover because they reduce exploring of more feature space.

For the ResNet 50 model, each residual block follows same order of convolution with same dimension. To save every layer computational time, do stride convolution by rising the depth of channel as 3×3 so that feature mapping regularly down sampled. In this model network in network filter 1×1 reduces the computations and preserves the channel depth.

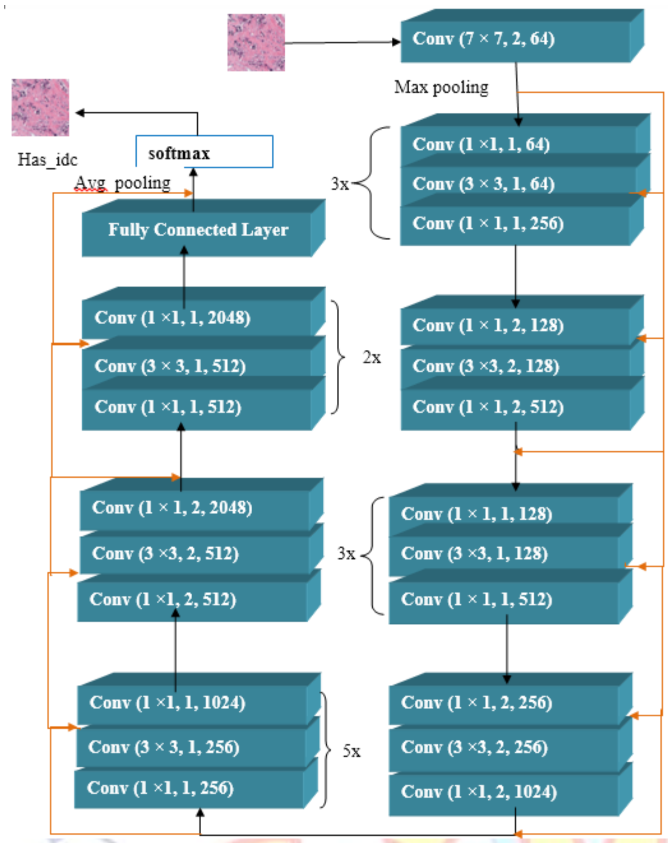

Figure-3: A model of Pre-trained ResNet 50 architecture where brown lines representing identical mapping, group of three consecutive convolutional layers represents a residual block, 2x, 3x, 5x represent 2, 3, 5 times that the corresponding residual block consecutively present in network and Conv (f × f, s, n) represents f × f filter size, s stride size and n number of filters used in the convolution.

4. PROPOSED METHODOLOGY

The block diagram of the proposed deep learning model is described in a sequence starting from acquisition of raw images to classification of the image as output. The design of this proposed methodology is depicted in Figure-4. The entire methodology is mainly divided into six steps: collect breast histopathology (BH) images, Image pre-processing, Image visualization, creation of ResNet Model, model training and finally model evaluation.

Step-1: Collect the images from Breast Histopathology (BH) images source site. This data plays a key role to model our Neural Network as it is the base of our task. Here collecting the data into dataset and splitting the dataset into train dataset and test dataset are performed to yield better performance.

Step 2: Pre-processing the BH images- In this stage, generally the data augmentation techniques such as padding, cropping, horizontal flipping, etc., are performed. In our project vertical flipping, zooming, rotating and lighting are applied to significantly increase the diversity of image data available for training the model without collecting new data on runtime to generalize. Augmentation is required because the medical images are small in size in general. By doing this our model can achieve high results.

Step 3: Images Visualization-In this stage, training data after pre-processing are plotted to know the patterns in it.

Step 4: Set up of ResNet Model-In this stage ResNet 50 architecture is used to create our model. This ResNet gives better accuracy in detecting IDC where 1% accuracy is also a big matter in medical field.

Step 5: Training ResNet Model for detection of IDC- In this stage the features that went through the model are processed to be optimized through updating the weights and those analyzed features are converted in predictions for the malign or benign. Proposed model uses transfer learning to achieve good results. After creating ResNet 50 model, it has frozen initial layers. Initially, we train the weights of the last fully connected layers of this ResNet model with acquired train dataset. Default settings used RGB (Red Green Blue) colour filters. We run the model with the default setting and explore the accuracy. Later to improve the accuracy, repeated the fine tune using call backs to modify the training as required with weight initialization, dropout regularization and stochastic gradient descent optimization. Generally in transfer learning, tweaking initial layers are performed with low learning rate. With an observation of learning rate of the model, we used slice function with the learning rate between 10e-6 to 10e-4 for different layers in the network. This slice function actually trains the initial layers with lowest learning rate and increases it for later hidden layers. To do this explicitly need to unfreeze the model.

Step 6: Evaluation of ResNet Model- After training the model, the model is evaluated using the accuracy metric after every epoch to find if it is learnt as supposed by testing with some unseen validation images. After a number of evaluations and modified training, proposed model has achieved 91.38% accuracy.

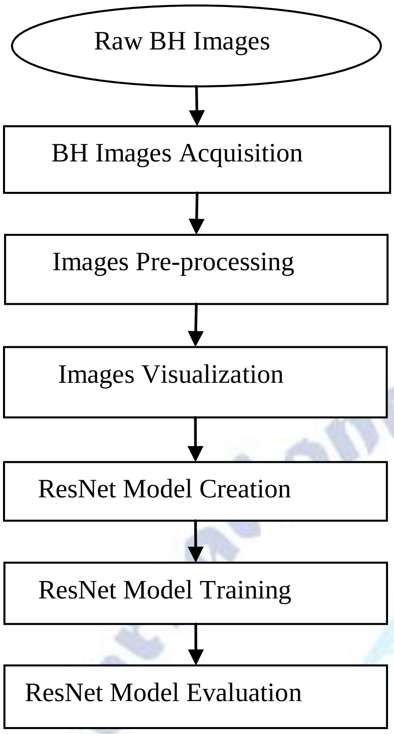

Figure-4: Block diagram of Proposed Methodology

5. EXPERIMENTAL RESULTS

Data Set

The proposed model acquires data from public Kaggle domain [36]. The original dataset in Kaggle, created on December 2017, was scanned at 40x, means that an approximate resolution of 0.25 μm per pixel. It contains 162 whole mount slide images of Breast Cancer specimens. From that 277524 patches were extracted with a size of 50×50 where 198738 images are IDC negative and 78786 are IDC positive labelled. The dataset clearly shows that there are double negative samples than the positive samples. Each image in the dataset has a specific file format.

For each image, a file name was assigned with uniform format that contains patient ID, x-axis and y-axis of where this patch was cropped from respectively and the class of the patch. For example a filename 10278_idx5_x1051_y1251_class1.png in the dataset indicates 10278_idx5 is patient ID, 1051 is the x coordinate and 1251 is y coordinate of where this patch was cropped from and 1 is the class of this image where 1 indicates has-IDC(malign) and 0 indicates no-IDC(benign). This dataset consists of 543 series.

We create a data frame to extract the labels from image files. Table-1 displays the first ten images information in the data frame. In that image_id is identity number of the patient and target is the actual label of the image based on which the images are classified. These labels are picked from their corresponding image_id.

6. RESULTS AND DISCUSSIONS

The proposed model is built on using pre-trained ResNet 50 and trained it with large breast histopathology image dataset to give better performance in predicting the IDC by minimizing the loss function with stochastic gradient descent optimizer. After training the model, it is quite able to learn better results for all the weights and the bias from labelled images. The proposed pre-trained model uses PyTorch library and Fastai framework to be trained and which provide wide support for neural network implementations. These libraries are open source libraries developed using Python programming language to enable the model run on GPU (Graphical Processing Unit).

To normalize whole data, perform augmentation technique on training dataset through different ways of processing. That enhances the size and quality of training dataset. Such that, this deep learning model learns better. In this model the whole Image dataset was divided into two datasets for training as train dataset and for validation as test dataset and for the accuracy comparison we tested on three sets of ratios as 90:10, 75:25 and 50:50. For example 90:10 indicates 90% images for train dataset and 10% for test dataset from whole image dataset. The performance of the proposed pretrained model is evaluated on test data. Figure-5 shows randomly picked few augmented image patches by the model on executing the command data.show_batch (rows= 3, figsize= (7, 8)). In this figure the image label has_idc means malign and no_idc means benign.

Table-1: Creation of a Dataframe from source Dataset [36]

	image_id	patient_id	target
0	9267_idx5_x1901_y901_class1.png	9267	1
1	13022_idx5_x2951_y1301_class0.png	13022	0
2	15634_idx5_x701_y1051_class1.png	15634	1
3	13613_idx5_x2051_y2401_class1.png	13613	1
4	13021_idx5_x1701_y701_class1.png	13021	1
5	8956_idx5_x1801_y351_class1.png	8956	1
6	12891_idx5_x2851_y1401_class0.png	12891	0
7	10308_idx5_x1251_y751_class1.png	10308	1
8	10262_idx5_x2351_y1201_class1.png	10262	1
9	14157_idx5_x1001_y901_class0.png	14157	0

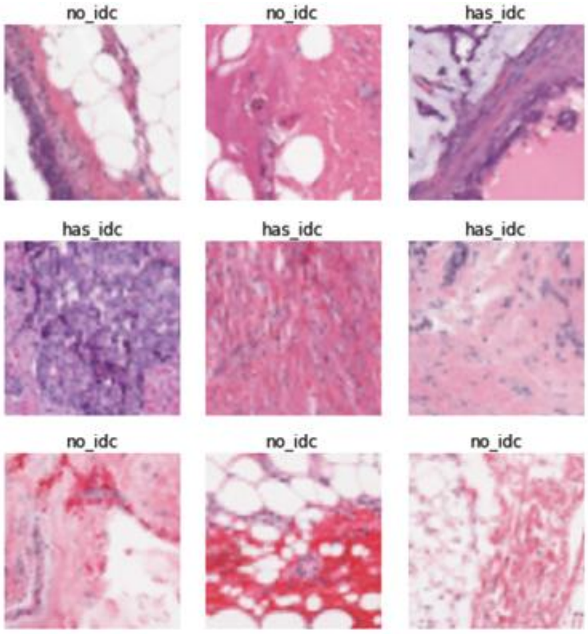

Figure-5: Example of Augmented image patches for training set after transformation (vertical flip, lighting, rotation, zoom).

To compare it's accuracy with other model, we trained ResNet 34 and ResNet 50 on the same train dataset. These two models differ with their residual block layers deep. ResNet 34 contains 2 layers deep whereas ResNet 50 contains 3 layers deep in the residual block. When compared these models accuracies, it showed that ResNet 50 always giving better accuracy than ResNet 34. The results of ResNet 34 and ResNet 50's fine tuning cycle with 4 epochs where the dataset split in the ratio of 90 and 10 for train dataset and test dataset respectively are shown in the below Table-2 and Table-3 respectively. These tables give the training loss, validation loss, error rate and time values for each epoch.

Epoch is one forward and one backward pass through the entire training dataset only once. Loss indicates the model's prediction is not good as the number indicates and its range is from 0 to 1. Training error shows the loss on train dataset and valid loss shows the loss on test dataset through the trained network. Error rate is a predefined metric in Fastai library gives the error rate by calculating (1-accuracy), where accuracy is also predefined metric computes accuracy with targets when input tensor is batchsize × number of classes. From Table 3 and Table 4 for each epoch ResNet 50 showing lower error rate than ResNet 34 from the starting.

After unfreezing the starting layers, train again the whole model using fit ( ) with a range of two learning rates where the learning rate started nudging to optimize the training. The results are shown in the below Table-4 and Table-5 respectively.

Any model during training must consider the overfitting problem as too many epochs can cause overfitting the train dataset. Overfitting means the model gives good performance on train or seen dataset but gives poor results on test or unseen dataset. By observing the above table it is clear that the performance on a validation dataset started to degrade. So to avoid overfitting problem prefer early stopping technique. It is clear that ResNet 34 achieved approximately 90% accuracy whereas ResNet 50 got 91% accuracy. Here the time consumed by ResNet 34 is less than ResNet 50 to predict but in medical field accuracy is more important.

In order to further evaluate the effectiveness of our proposed method, ResNet 50 and ResNet 34 models were created and trained with three different ratios of training and testing datasets and the results were analyzed. The results were depicted to compare the classification accuracy results using ResNet-50 and ResNet-34 models as shown in Table-6. So this proposed model with 90:10 data split achieved accuracy of 91% with ResNet 50 and when compared with Angel crozs model it achieved 7% more accuracy.

Table-2: ResNet 34 values with fine tuning cycle of 4 epochs

epoch	train_loss	valid_loss	error_rate	time
0	0.327232	0.319004	0.134091	12:48
1	0.305335	0.278070	0.111118	12:36
2	0.294650	0.264734	0.105914	12:29
3	0.301546	0.262571	0.105280	13:05

Table-3: ResNet 50 values with fine tuning cycle of 4 epochs

epoch	train_loss	valid_loss	error_rate	time
0	0.327580	0.298450	0.123937	18:46
1	0.286503	0.260109	0.103566	18:40
2	0.273711	0.247817	0.099695	18:40
3	0.271879	0.244322	0.098299	18:39

Table-4: ResNet 34 values after fine-tuning in second cycle with 4 epochs

epoch	train_loss	valid_loss	error_rate	time
0	0.284724	0.255420	0.101028	14:02
1	0.262002	0.243880	0.098680	13:59
2	0.242914	0.227625	0.091953	13:57
3	0.240702	0.222806	0.091255	13:44

Table-5: ResNet 50 values after fine-tuning in second with 4 epochs

epoch	train_loss	valid_loss	error_rate	time
0	0.256882	0.249013	0.098045	26:02
1	0.253358	0.233641	0.093921	25:59
2	0.235600	0.216973	0.084528	26:00
3	0.219059	0.215806	0.083957	26:01

Table-6: Classification accuracy rates of IDC using ResNet-50 and ResNet-34 models with three different ratios of training and testing datasets.

Category	Data (%)	no-idc/ has-idc	No. of Samples	ResNet-50	ResNet-34
				Accuracy Rate (%)	
Training	90%	Total No. of Samples	141814	91%	90%
		no-idc	70907		
		has-idc	70907		
Testing	10%	Total No. of Samples	15758		
		no-idc	7879		
		has-idc	7879		
Training	75%	Total No. of Samples	118180	88%	86%
		no-idc	59090		
		has-idc	59090		
Testing	25%	Total No. of Samples	39392		
		no-idc	19696		
		has-idc	19696		
Training	50%	Total No. of Samples	78786	79%	79%
		no-idc	39393		
		has-idc	39393		
Testing	50%	Total No. of Samples	78786		
		no-idc	39393		
		has-idc	39393		

7. CONCLUSIONS

Nowadays, public awareness about the Breast cancer has been improved. So earlier diagnosis and prognosis help to receive better treatment and thus can increase the survival rate of the patients. 80% of the breast cancers are IDC. Our ResNet 50 framework helps the pathologist to make better diagnosis of the patients' cases to detect IDC with 91% accuracy and in less time.

Acknowledgment

The authors would like to express their gratitude to Sri K.V.V.Satyanarayanan Raju, Founder & Chairman and Sri K.Sasi Kiran Varma, Vice Chairman, Chaitanya group of Institutions for providing necessary infrastructure to carry out the research work in GIET, Rajahmundry.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1]. Invasive Ductal Carcinoma (IDC), Johns Hopkins medicine center, Available: https://www.hopkinsmedicine.org/breast_center/breast_cancers_other_conditions/invasive_ductal_carcinoma.html

[2]. BREASTCANCER.ORG2, Invasive Ductal Carcinoma (IDC), Last modified on March 2019, https://www.breastcancer.org/symptoms/types/idc

[3]. Carol E. DeSantis et al, "Breast cancer statistics, 2019", ACS Journals, 02 October 2019, https://acsjournals.onlinelibrary.wiley.com/doi/full/10.3322/caac.21583

[4]. BREAST CANCER STATISTICS, 2018, https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics

[5]. BREAST CANCER INDIA, 2018, http://www.breastcancerindia.net/statistics/trends.html

[6]. Kerean Watts, "Breast Cancer Awareness Month: Lessons for India", 2019, https://www.healthissuesindia.com/2019/10/02/breast-cancer-awareness-month-lessons-for-india/

[7]. American Cancer Society," How Common Is Breast Cancer?", September 18, 2019, https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html

[8]. SanjeevN.Jain ,Bhagyashri G. Patil, "Cancer Cells Detection Using Digital Image Processing Methods", International Journal of Latest Research in Science and Technology, Vol. 3(4):45-49 , March 2014

[9]. Edmndo C Mauad et al, "Adherence to cervical and breast cancer programs is crucial to improving screening performance",Volume 9 Issue 3, 22 September 2009.

[10]. Nicholas G. Polson, " Deep Learning", stat.ML, Aug 2018

[11]. Peikari, Mohammad, Mehrdad J. Gangeh, Judit Zubovits, Gina Clarke, and Anne L. Martel. "Triaging diagnostically relevant regions from pathology whole slides of breast cancer: A texture based approach." 2019.

[12]. Yousif Mohamed Y. Abdallah, "History of Medical Imaging", Volume 5(2), page:275-278, 2017

[13]. Dundar MM et al, " Computerized classification of intraductal breast lesions using histopathological images", IEEE Trans Biomed Eng., 58(7):1977-84, 2011 Feb 4

[14]. Oberoi et al," Magentic resonance imaging of breast masses: Comparison with mammography", Indian Journal of Radiol Imaging, Volume : 15, Page : 381-387, Year : 2005
[15]. Ireaneus Anna Rejani et al, "Early Detection of Breast Cancer using SVM Classifier Technique", International Journal on Computer Science and Engineering Vol.1(3), 127-130, 2009
[16]. Yi-juTsing et al.,"Predicting breast cancer metastasis by using serum biomarkers and clinic pathological data with machine learning technologies", Article in International Journal of Medical Informatics, May 2019
[17]. Turgay Ayer et al, "Comparison of Logistic Regression and Artificial Neural Network Models in Breast Cancer Risk Estimation1", Radiographics, 30(1):13-22 ,November 2009
[18]. Bikesh Kumar Singh ,"Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm", Article in Biocybernetics and Biomedical Engineering 39(2):393-409 • 2019
[19]. Sumit Saha, "A Comprehensive Guide to Convolutional Neural Networks", 2018
[20]. Savelli et al, "A multi-context CNN ensemble for small lesion detection", Volume 103, 2019
[21]. Angel Cruz-Roa et al, "Automatic detection of invasive ductal carcinoma in whole slide images with Convolutional Neural Networks", Proc. of SPIE Vol. 9041, 2014
[22]. Muhammad Imran Razzak et al," Deep Learning for Medical Image Processing: Overview, Challenges and Future", Classification in BioApps pp 323-350, 2017
[23]. Tasneem Gorach," Deep Convolutional Neural Networks- A Review", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 07, July-2018
[24]. Keith D. Foote," A Brief History of Deep Learning", February 7, 2017 .Yann LeCun et al, "Deep learning", 436 , NATURE , VOL 521 ,28 MAY 2015
[25]. JiuxiangGu et al," Recent Advances in Convolutional Neural Networks", last revised 19 Oct 2017
[26]. Wei hu et al, " Deep Convolutional Neural Networks for Hyperspectral Image Classification", Volume 2015 ,Article ID 258619, 30 Jul 2015
[27]. Nwankpa et al, "Activation functions: Comparison of trends in practice and research for deep learning" ,cs.LG, 8 Nov 2018
[28]. KhushbooMunir et al, "Cancer Diagnosis Using Deep Learning: A Bibliographic Review"Cancers (Basel),11(9). Aug 23 2019
[29]. Bing Xu et al," Empirical Evaluation of Rectified Activations in Convolutional Network", last revised 27 Nov 2015
[30]. Yu-Dong Zhang et al, " Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling", Journal of Computational Science, Volume 27, Pages 57-68, , July 2018
[31]. Gregory Morse et al, "Simple Evolutionary Optimization Can Rival Stochastic Gradient Descent in Neural Networks", Proceedings of the Genetic and Evolutionary Computation Conference 2016, Pages 477–484, July 2016
[32]. Stefan Wager et al,"Dropout Training as Adaptive Regularization", Neural Information Processing Systems 2013,
[33]. Mohammad SadeghEbrahimi et al," Study of Residual Networks for Image Recognition", [v1], 21 Apr 2018, Kaiming He et al," Deep Residual Learning for Image Recognition",2015
[34]. Breast Histopathology images, 2017 https://www.kaggle.com/paultimothymooney/breast-histopathology-images