2019 5th International Conference for Convergence in Technology (I2CT)
Pune, India. Mar 29-31, 2019

Comparative Study of Breast Cancer Detection Techniques

Royston D'mello, Vishwa Pardeshi, Hasti Mehta, Dr. Sudhir Dhage
Department of Computer Engineering
Sardar Patel Institute of Technology
Mumbai
{royston.dmello, vishwa.pardeshi, hasti.mehta, sudhir_dhage}@spit.ac.in

***Abstract*—Diagnosis of breast cancer always comes with some level of uncertainty. Depending on the method of diagnosis, various issues can create uncertainty in a patient's diagnosis. An undiagnosed cancer and a falsely diagnosed one, both are bad for the patient in their own way. To prevent these kind of situations we can use machine learning techniques to aid the doctors in providing a more confident diagnosis. Several techniques have been used so far, which we summarise in this paper.**

***Index Terms*—Breast Cancer, Machine Learning, Diagnosis**

I. Introduction

Cancer refers to the group of diseases which are characterized by an abnormal and uncontrolled cell growth which leads to the formation of tumors. These tumors can be benign or cancerous. A cancerous tumor is malignant, it has the potential to spread or invade other parts of the body. Out of the 100 types of cancer that affects human, breast cancer is the most commonly diagnosed in women worldwide. The promotion of regular breast screening has been responsible for the overall decline in the incidences of breast cancer. Despite this, it is one of the leading cause of deaths in women, as it can only be treated and not prevented. Hence, the decrease in the death rates due to breast cancer have been rightly attributed to early detection and diagnosis which are the key to rehabilitation and recovery.

A. Motivation

The risk factors associated with breast cancer is not limited to genetics and includes lifestyle factors such as alcohol consumption. Some risk factors cannot be influenced or modified. Hence, early detection of breast cancer is the best defence as many significant risk factors such as age, gender are beyond ones control.The cancer incidences in older women are higher as longer lives gives more opportunities for certain breast cancer gene mutations to take place which might lead to formation of malignant tumors.

B. Existing Detection Techniques

Detection of breast cancer is done by using physical examination, scans etc. Once, a change in the shape or size or any abnormality is observed in the breasts, a physical examination is conducted, followed by X-ray of the breast. The image produced is called a mammogram. These mammograms are then studied for signs of breast cancer by trained radiologist. Once detected by the radiologist, a needle test or biopsy is conducted to confirm the presence of cancerous tissues in cell. This however is an invasive method. An alternative is to eliminate the human error and inefficiency by introducing machine learning algorithms to detect the presence of malign tumors in the breast by observing images or other form of data about the scans. This has to overcome the problem of false positives and false negatives by improved classification of masses and micro-calcification.

C. Organisation of the paper

In this paper, various machine learning techniques for detection of Breast Cancer using extensive datasets have been studied. The paper is organised as follows: Section II explains the various types of predictions - Stage Classification, Binary Classification, Segmentation. Section III categorizes the type of input that a breast cancer detection algorithm can utilize and the characteristics of each of the input sets. The method of acquiring the input sets have been discussed briefly. The two major categories discussed are image scans and numerical data gathered from a variety of techniques such as biopsy, imaging, MRI etc. Section IV discussed the various datasets that are generated as a result of the above gathered inputs from various sources. The defining features of each dataset is stated. Section V details the experimental setup that will be required for the accurate prediction of breast cancer in a patient. The experimental setup includes the hardware, software and instrument required for a particular type of prediction as listed in section II. Section VI elaborates on the dataset preprocessing required to select, extract and transform the different features from the input dataset to distinguish malignant and benign tumors and detect breast cancer accurately. Section VII lists the performance metrics for each type of prediction listed in section II. Section VIII enlists the various machine learning algorithms that can identify and extract defining characteristics from the extensive datasets to make precise predictions regarding the detection and classification of the malignant subjects, if any. Section IX summarizes the result of the previously discussed algorithms. Section X talks about the scope of future improvements. Section XI gives the conclusion of the study.

II. Types of Prediction

Breast cancer can be classified depending on range of factors such as invasiveness, location etc. The various machine learning algorithms which will be discussed extract the most distinguishing feature from the available dataset to indicate the presence of malignant subject using three different kinds of prediction. The three types are described as follows (In the methodology we only focus on Binary Classification):

978-1-5386-8075-9/19/$31.00 ©2019 IEEE

A. Stage Classification

The stages are described based on the presence of the tumor, its size and metastasis in the axillary lymph nodes. Classification of breast cancer into different stages characterized by the extent of the development or spread of the cancerous tissue helps the doctor make informed medical decisions about the course of medical response to adapt in a particular case. Once classified, history of previous similar cases can be referred to make quick, efficient decisions regarding the suitable treatment. The treatment can be aggressive or less aggressive, depending on the stage of breast cancer. Hence, the classifier algorithms eliminate error due to human inexperience and lack of expertise by defining appropriate performance metrics and evaluating the accuracy of the results by using the standard datasets.

The existing stage classification for breast cancer is based on the TNM classification. This staging system is based on the size of the tumor, the spread to distant sites (metastasis) and nearby lymph nodes, the estrogen and progesterone receptor status, Her2 status, and grade of the cancer. These physical and cellular characteristics of the examined tumor is each assigned a number. A suitable machine learning technique that exploits these numerical inputs predicts the likelihood of malignancy.

From the multitude of machine learning approaches available for detection and classification of breast cancer, Naive Bayesian Classifier and k- nearest neighbor are classifier algorithms. The Naive Bayesian Classifier is efficient in determining classification models when clubbed with other classifiers. It handles data which influence each other to model decisions. Based on probability and statistics, it is used to predict chemical and biological properties on the datasets subject. A challenging problem that arises is the false positive and false negatives. K- nearest Neighbour measures the performance of false positive rates.

B. Binary Classification

The objective of a mammogram of breasts is to capture the presence of any abnormal growth by exposing the breasts to a low level of ionizing radiation. By visual observation of the size, shape and location of the tumors, lesions or cyst, the nature of the subject - malignant or benign can be determined. This classification of tumors into either malignant or benign form the basis of binary classification.

The digitized mammogram which are obtained from standard datasets are feed to machine learning algorithms. While these methods are more efficient in accurate diagnosis and detection of cancerous tumors than a radiologist, limitations of mammograms in screening dense breast tissues leads to cases of false positives and false negatives. A good classifier model should have a low false positives and false negatives rates. This can be achieved by careful identification of region of interests. The classifying algorithms digital input is exhaustive and redundant information needs to be eliminated. Thus, identifying features required for accurate classification of the tumors is vital to the success of the classifier algorithm. These identifying features can be the statistical image texture, contrast, dissimilarity etc[1]. For accurate classification apart from the physical and cellular characteristics of the tumor, patients history and previous diagnostics, if any play a significant role in the characterization.

Once the features that encapsulate the difference between malignant and benign masses are chosen, a classifier model is applied. The training efficiency of such models depend heavily on the extracted features from the digitized pictures and performance metrics.

C. Segmentation

One of the most challenging problem of digital mammography is to find a robust and accurate breast profile segmentation technique. In all machine learning algorithms the preprocessing of the data is an essential step. It should focus on extraction of regions of interest (ROI) and improve the overall efficiency of the detection and classification algorithms. This concentrated ROI narrow the are of search thereby eliminating background information that interferes with the visibility of the abnormalities in the breast tissues.

Once preprocessing removes the digitization noise and pectoral muscles, the result is an accentuated breast profile. This breast profile is then subjected to mammogram segmentation. The segmentation involves categorization into distinct regions. Each region, such as the nipple, the breast border etc affect the search space and adequacy. There are various approaches proposed for segmentation. The focus of segmentation can be on using breast contours, non-breast region, thresholding, active contours and gradients.

The thresholding approach was one of the earliest segmentation approach.The shortcoming of thresholding is related to the misclassification due to the overlap between the background and breast region. To overcome the above limitation, different thresholds for different regions can be defined to isolate the ROIs and detect the edges. Breast contours are then developed and suitable algorithm is applied to meet the stated performance metrics.

III. Types of Inputs

Detection and diagnosis of Breast Cancer is done by employing a combination of approaches which include physical examination, imaging techniques, biopsy etc. Classification of breast cancer requires labeled data to develop and train models to make accurate predictions. This data can be acquired from the humongous databases maintained by the hospitals and clinics containing the records of patients with their symptoms, diagnosis and progress of the disease[2]. The machine learning algorithm make use of this historic data to make medical inferences. The following are the two major categories of input which is used by machine learning algorithms.

A. Numerical Input

In this type of input, physical characteristics of the subject i.e tumor, lesions etc are captured in form of numerical values and often arranged in an easily readable form such as an array.

The characteristics often captured in these types of datasets are physical properties of the cyst and cells. The mentioned cellular features can only be observed from a digitized image of an FNA other invasive procedure of a breast mass[3]. These datasets contain benign as well as malignant samples. These datasets, hence, can be used for binary and stage classification of breast cancer.

The Wisconsin Breast Cancer Database is an example of this input which is multivariate and mostly commonly used for classification.

B. Image Input

There are generally two types of image inputs used. Mammograms and Histopathology images are used. Mammograms are created using X-rays and Histopathology images are created using FNA.

IV. Datasets

Although some of the research uses proprietary datasets, most of the surveyed literature uses publicly available datasets. The datasets are categorized according to the input types defined above.

A. Numerical Datasets

1) Wisconsin Breast Cancer Database (WBCD): It contains elements having various scalar observations. It contains samples that are benign as well as malignant. The nine features it consists of are as follows; clump thickness ,uniformity of cell size ,uniformity of cell shape ,marginal adhesion ,single epithelial cell size ,bare nuclei ,bland chromatin ,normal nuclei, and mitoses . They are represented as an integer between 1 and 10. These features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass.

B. Image Datasets

The image datasets are described in the table I

Dataset	Reference	Type	Number of samples	Benign (%)	Malignant (%)
WBCD	[4]	FNA	699	65.52	34.47
DDMS	[5]	Mammogram	899	25.06(100 out of 399)	21(105 out of 500)
MIAS	[6]	Mammogram	322	58.69	41.3
DREAM	[7]	Mammogram	500	36.25(29 out of 80)	63.75(51 out of 80)
BreakHis	[8]	Histopathology	9109	27.22	59.6

TABLE I: Summary of datasets surveyed for this paper

V. Data Collection

For Wisconsin Breast Cancer Dataset, FNA is done to remove cells from a cyst using a very thin needle. The properties of the cyst are given values from 1 to 10. 1 is closest to bening and 10 is closest to malignant. Mammography: This requires an X-ray machine. Minimum 2 pictures of each breast have to be taken from 4 different viewing angles. For DDMS (Digital Database for Screening Mammography), the mammograms collected are decompressed and then filtered and re-annotated by mammographers. These images are then saved in DICOM format. For BreakHis dataset, the samples are created using excisional biopsy. It removes large part of a tissue and is generally performed in a hospital under general anesthesia.

VI. Pre-processing

A. Feature Selection

In [9] Akay et al. use the F-score [10] of the WBCD features to determine the most discriminant features. They rank the features using F-scores and make 9 subsets by iteratively adding one more lower scored feature into the subset.

Karabatak et al. [11] use an algorithm called AR-1 for feature selection. It eliminates features based on rules it finds that denote the discrimination power of the features.

In [12] authors use two datasets and employ different feature selection techniques for the two. For the first dataset containing fine needle aspirate data, the authors propose using an (signal-to-noise ratio) SNR filter to rank the features and then use the 3 highest ranked features. For the second dataset which is a prognosis dataset using gene data, the authors use a wrapper based on Sequential Forward Selection which optimises an objective function in a hill-climbing fashion.

In [13] the authors have used mammography images from the DREAM challenge. The images in this dataset have labels on them, so the authors use binary image thresholding to remove those labels. The authors then use histogram equalisation to normalise the intensity of the image to the range of 0-255. The contrast of the mammograms are then enhanced. To enhance the mammograms further, nonsubsampling contourlet transform is used on the images.

B. Feature Extraction

[14] use SIFT, morphological entropy based, EFDS and texture of the image as features and try them with different kernels of an SVM.

In [15] the authors use the MIAS dataset, regions of interest are extracted based on the centre and radius of the cyst given in the dataset. Statistical texture analysis is used to create 15 1st and 2nd order features.

C. Feature transformation

In [3] 1024x1024 images are first resized to 224x224. Noise reduction is done by using binarization, masking to find ROI.

[16] Uses ICA for dimensionality reduction. ICA uses higher order statistics to get Independent Components which condense more information than PCA. The Wisconsin Dataset is reduced to 2 dimensions.

[17] uses a non-public mammogram dataset. The authors first perform cropping to remove the margins from the images. A Wiener filter is then used to remove the blurriness from the images. They then use Discrete Wavelet Transform (DWT) to transform the images from time to frequency domain.

VII. Performance Metrics

1) Accuracy: Accuracy is defined as the total number of correct predictions divided by the total number of predictions. Accuracy is not a good metric for skewed datasets.
2) Sensitivity: This measures false negatives. In terms of breast cancer detection it is the percentage of actually malignant cases among those predicted as benign by the system.
3) Specificity: This measures true positives. In terms of breast cancer detection it is the percentage of actually benign cases among those predicted as benign by the system.
4) Area under ROC: This helps us to measure the performance at various threshold settings. ROC is a probability curve while AUC represents measure of separability.
5) Matthews correlation coefficient: It is used to measure performance of binary classifiers. It goes from -1 to 1.
6) Positive Predictive Value (PPV) and Negative Predictive Value (NPV): The percentage of malignant and benign cases predicted correctly respectively.
7) False Positive Rate (FPR): It is the percentage of cases predicted by the algorithm as malignant which were actually benign.

VIII. Methodology

Depending on the kind of features and required prediction, different kinds of deep and shallow learning techniques have been applied.

A. Shallow Learning

Also known as conventional machine learning, it is best used when the number of features is less, something which is typical to numerical datasets. Some of the most common conventional machine learning algorithms are described as follows:

1) Support Vector Machines (SVM): A support vector machine separates the classes of a dataset using a hyperplane and its objective tries to maximize the distance between the hyperplane and the samples on each side of it. It offers in-built support for regularisation. Kernel functions are used when the decision boundary between the classes is not very simple i.e. the classes are not very linearly separable. [9] have used a linear kernel and RBF kernel and polynomial kernel svm, on the Wisconsin dataset and use a grid search to optimise the hyper-parameters. A linear kernel is as good as not using a kernel. It works in the same feature space as the input features.
2) K-Nearest Neighbours (KNN): A data sample is classified based on its distance from the kth closest sample in the dataset. The hyper-parameter k controls the bias and variance. As k decreases the variance increases and as k increases the bias. [18] use k = 3 on the Breast Cancer Dataset.
3) Naive Bayesian Classifier: This is a classifier based on Bayes theorem and models decisions on conditionally independent variables.
4) Random Forest Classifier: A decision tree models decisions as nodes which decide which path and input sample will follow based on some of its features. The leaves give the final decision. But a decision tree tends to overfit, hence an ensemble of decision trees with a random set of features and random set of training samples for each tree is created. The ensemble is believed to be better than an single tree since they cover each others mistakes.
5) Relevance Vector Machines: It gives probabilistic predictions and does not a regularisation hyper-parameter. It uses kernel functions like an SVM but requires lower computational costs.
6) Apriori Algorithm: It is an association rule based algorithm. Association rules models the patterns of co-occurrence of feature values in a dataset.
7) Artificial Neural Networks (ANN): An artificial neural network is a universal function approximator, which is modelled after the human brain. It performs computations in layers. The more number of layers the more number of computations. It is basically used to transform a lesser linearly separable feature space into a more separable one. Although, this is something that is closest to Deep Learning, the ANNs used in [19] are shallow as they have only 1 hidden layer. A shallow network was used as the features used were from WBCD. [12] have used a different architecture called PNN
8) Extreme Learning Machine: This is a Neural Network except that its weights are estimated using analytic methods and not backpropagation. Hence it is more computationally feasible than a conventional Neural Network.

B. Deep Learning

Neural Networks are at the center of deep learning. Their ability to transform the input feature space into a more usable and decisive one make them very favourable for a lot of problem statements. However due to their high parametric nature they work best on problems with a big feature size. Hence they are typically used on image datasets. What makes neural network research so intricate is getting the right architecture.

A further improvement on the performance of a neural networks was made by using Convolutional Layers. These layers perform convolutions of filters over the images and and transform the feature space into a highly complex one. These filters are optimised using back-propagation. [3] use 6 convolutional layers along with 4 average pooling layers on the MIAS dataset while [13] use only 3 convolutional and 3 max pooling layers. [20] use a model called the DenseNet model which does not use standard convolutional and pooling layers. Instead it uses Dense blocks which themselves contain multiple convolutional layers and non-linearities. And instead of pooling layers it uses transition layers which themselves

contain convolutional layers with kernel size 1 and average pooling layers to reduce the size.

Transfer Learning is a technique where the weights are not randomized and in fact are initialised from a pretrained network that has been trained on another dataset. This is done because it is believed that some feature transformation techniques are transferable from one dataset to another. The depth to which the network filters are well transferred depends on the similarity of the two datasets. If the two datasets are of similar domains then almost the entire network can be reused. But if they are very different then the lower layers are optimum as they usually contain edge and colour detectors which can be used on most images. [20] initialises the DenseNet described above with layers from a network trained on ImageNet and optimises the weights for the domain of histopathological images using the BreakHis dataset. [21] uses the InceptionV3 network as the base and re-trains it using the BreakHis dataset while [22] use an AlexNet network and re-train it on the DDSM dataset in combination with data augmentation.

The U-net architecture is a very popular choice for segmentation problems. [23] have used a modified version of u-net and trained on mammogram dataset which they created and manually annotated themselves.

IX. Results

We compare the results of all the papers examined in this survey and group the results according to the prediction type and dataset type and present the best summarize the best results from each paper in the tables II and III

Reference	Dataset	Preprocessing	Methodology	Metric	Value
[2]	WBCD	None	Random Forest	Area under ROC	99.9
[9]	WBCD	F-score sorting	SVM (RBF kernel)	Accuracy	99.51
[11]	WBCD	ART feature elimination	Neural Networks	Accuracy	97.4
[16]	WBCD	ICA with 2 components	SVM quadratic	Accuracy, Sensitivity	94,40, 97.77
[18]	BCD	None	KNN	Accuracy	97.51
[19]	WBCD	None	ANN	Accuracy	98
[24]	WBCD	None	LS-SVM	Sensitivity, Specificity, Accuracy	97.87, 97.77, 97.08
[25]	WBCD	None	SVM	Accuracy	97.13
[26]	WBCD	Selected only 4 features	Relevance Vector Machine	Accuracy	97

TABLE II: Results of Numerical Input Methods

Reference	Dataset	Preprocessing	Methodology	Metric	Value
[13]	DREAM challenge	Binary thresholding + Histogram equalisation + contrast + NSCT	CNN	Accuracy	100
[14]	DDMS	Texture	Bayes	Sensitivity, Specificity, PPV, NPV, TA, FPR, AUC	1,
		EFDs	Bayes		1,
		EFDs	SVM Polynomial		1,
		Morphology	SVM rbf		1,
		Entropy	SVM rbf		0,
		EFDs	SVM gaussian		1
[15]	mini-MIAS	Statistical Texture analysis	Extreme Learning Machine	Accuracy, Sensitivity, Specificity	91, 90, 98,
[17]	Proprietary	Cropping, filtering, DWT	Neural Network	Accuracy	93,7
[20]	BreakHis	None	DenseNet with fine tuning	Accuracy	95.4
[21]	BreakHis	None	InceptionV3 with fine tuning	Accuracy	Benign: 0.83, Malignant: 0.89

TABLE III: Results of Image Input Methods

X. Future Works and Conclusion

In this paper we focus on only breast cancer detection which is a binary classification type of prediction. We described the history of the field briefly followed by the various ways in which breast cancer can be detected. We then described the various datasets used in the field according to the type of input. Most of the datasets used mammograms, which others used numerical data and histopathology of data. The data collection mechanisms were described for all of the datasets, so that more data can be collected for future work. Pre-processing of many different kinds has shown to be very effective in increasing the performance metrics of the classifiers. We then summarised the results of all the papers under consideration showing their best possible result along with the algorithm and preprocessing used. For future work, more focus should be given on Deep Learning methods as they are shown to be very effective in image processing tasks. Hybrid prediction i.e. predicting both the segmentation mask and the type of the tumor should be explored. As segmenting the tumor and then predicting the type can help in better classification of the tumor. Using segmentation, the physical characteristics of the cysts can also be predicted and used as a feature for classification. Stage classification has not be explored much due to the fact that it gives a lower accuracy due to many overlapping cases. Efforts can be made to disambiguate the overlapping cases and find better stage classification for better and immediate treatment.

References

[1] G. Vani, R. Savitha and N. Sundararajan : Classification of Abnormalities in Digitized Mammograms using Extreme Learning Machine, 2010 11th Int. Conf. Control, Automation, Robotics and Vision Singapore.

[2] Bazazeh, D., & Shubair, R.M. (2016). Comparative study of machine learning algorithms for breast cancer detection and diagnosis. 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), 1-4.

[3] S. Charan, M. J. Khan and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, 2018, pp. 1-5. doi: 10.1109/ICOMET.2018.8346384

[4] W.N. Street, W.H. Wolberg and O.L. Mangasarian. Nuclear feature extraction for breast tumor diagnosis. IS&T/SPIE 1993 International Symposium on Electronic Imaging: Science and Technology, volume 1905, pages 861-870, San Jose, CA, 1993.

[5] The Digital Database for Screening Mammography, Michael Heath, Kevin Bowyer, Daniel Kopans, Richard Moore and W. Philip Kegelmeyer, in Proceedings of the Fifth International Workshop on Digital Mammography, M.J. Yaffe, ed., 212-218, Medical Physics Publishing, 2001.

[6] Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., et al. (2015). Mammographic Image Analysis Society (MIAS) database v1.21

[7] http://sagebionetworks.org/research-projects/digital-mammography-dream-challenge/

[8] Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L., A Dataset for Breast Cancer Histopathological Image Classification, IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, 2016.

[9] Akay, M.F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst. Appl., 36, 3240-3247.

[10] Chen YW., Lin CJ. (2006) Combining SVMs with Various Feature Selection Strategies. In: Guyon I., Nikravesh M., Gunn S., Zadeh L.A. (eds) Feature Extraction. Studies in Fuzziness and Soft Computing, vol 207. Springer, Berlin, Heidelberg

[11] Murat Karabatak and M. Cevdet Ince. 2009. An expert system for detection of breast cancer based on association rules and neural network. Expert Syst. Appl. 36, 2 (March 2009), 3465-3469. DOI=http://dx.doi.org/10.1016/j.eswa.2008.02.064

[12] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," 2010 5th International Symposium on Health Informatics and Bioinformatics, Antalya, 2010, pp. 114-120. doi: 10.1109/HIBIT.2010.5478895

[13] Natalia Pirouzbakht, Jose Mej a: Algorithm for the Detection of Breast Cancer in Digital Mammograms Using Deep Learning

[14] Hussain, L., Aziz, W., Saeed, S., Rathore, S., & Rafique, M. (2018). Automated Breast Cancer Detection Using Machine Learning Techniques by Extracting Different Feature Extracting Strategies. 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 327-331.

[15] G. Vani, R. Savitha and N. Sundararajan, "Classification of abnormalities in digitized mammograms using Extreme Learning Machine," 2010 11th International Conference on Control Automation Robotics & Vision, Singapore, 2010, pp. 2114-2117. doi: 10.1109/ICARCV.2010.5707794

[16] A. Mert, N. Kilic and A. Akan, "Breast cancer classification by using support vector machines with reduced dimension," Proceedings ELMAR-2011, Zadar, 2011, pp. 37-40.

[17] M. R. Al-Hadidi, A. Alarabeyyat and M. Alhanahnah, "Breast Cancer Detection Using K-Nearest Neighbor Machine Learning Algorithm," 2016 9th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, United Kingdom, 2017, pp. 35-39. doi:10.1109/DeSE.2016.8

[18] M. Amrane, S. Oukid, I. Gagaoua and T. Ensar, "Breast cancer classification using machine learning," 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), Istanbul, 2018, pp. 1-4. doi: 10.1109/EBBT.2018.8391453

[19] Abdel-Ilah, Layla & ahinbegovi, Hana. (2017). Using machine learning tool in classification of breast cancer. 3-8. 10.1007/978-981-10-4166-2_1.

[20] Majid Nawaz, Adel A. Sewissy and Taysir Hassan A. Soliman, Multi-Class Breast Cancer Classification using Deep Learning Convolutional Neural Network International Journal of Advanced Computer Science and Applications(IJACSA), 9(6), 2018. http://dx.doi.org/10.14569/IJACSA.2018.090645

[21] Chang, J., Yu, J., Han, T., Chang, H. J., & Park, E. (2017). A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer. In 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services, Healthcom 2017 (Vol. 2017-December, pp. 1-4). Institute of Electrical and Electronics Engineers Inc.. DOI: 10.1109/HealthCom.2017.8210843

[22] Lévy, Daniel & Jain, Arzav. (2016). Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks.

[23] T. D. Moor, A. Rodriguez-Ruiz, R. Mann, A. G. Mrida, and J. Teuwen, Automated lesion detection and segmentation in digital mammography using a u-net deep learning network, 14th International Workshop on Breast Imaging (IWBI 2018), Jun. 2018.

[24] Kemal Polat, Salih Gne, Breast cancer diagnosis using least square support vector machine, Digital Signal Processing, Volume 17, Issue 4, 2007, Pages 694-701, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2006.10.008.

[25] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, Thomas Noel, Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis, Procedia Computer Science, Volume 83, 2016, Pages 1064-1069, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.04.224.

[26] B. M. Gayathri and C. P. Sumathi, "Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-5. doi: 10.1109/ICCIC.2016.7919576