Breast Tumor Segmentation Using Deep Learning by U-Net Network

A. Parhizkar Tarighat
Payame Noor University, Iran, Rasht
Pr.tarighat@gmail.com

***Abstract*—The methods of image analysis are important for segmentation and detection of breast tumors, where a reliable diagnosis will be supported by portraying crucial morphological. This paper presents the segmentation of a breast tumor using a deep convolutional neural network by U-net. The structure includes a contractile route to capture the background and an a-symmetric expansion path that provides precise localization. It trains end-to-end from a small number of images. This segmentation algorithm is applied to breast tumors within a Region of Interest (ROI) with a High-Power Field (HPF) in biopsy and its corresponding ground-truth, where it has been labeled by pathologists with benign and malignant. The suggested segmentation model produces binary masks as realistic as possible. It supplies an Intersection Over Union (IOU) of 68% and attains an overall accuracy as high as 91%, which shows it performs better than the current works.**

***Index Terms*—Breast Cancer; Deep Learning; Segmentation; U-Net.**

I. INTRODUCTION

Malignant tumor named Breast Cancer (BC) is the most common cancer suffered by women in the world today. Specialist doctors have found that hormonal, environmental and lifestyle factors can influence a person's probability of developing BC. More than five to six percent of BC patients are with gene evolution that depends on the family age. Factors of increasing age, obesity, postmenopausal hormone imbalance are the main causes that lead to BC [1-5].

Early detection of BC has an effective result in reducing the treatment costs. Computer-Aided Diagnosis (CAD) is a good tool to help the radiologist in finding breast tumors and predestine their boundaries. Because of the low signal-to-noise ratio, the breast tumor segmentation is challenging [1-5].

At present, deep learning algorithm has been broadly used for segmentation of organ and tissue. Consisting of the brain, chest, heart, head, neck, abdomen, pelvis, breast, bone, and joint. In segmentation, the patch-based approach is commonly used, because most medical images have high resolution, where images are separated into small patches with a specific scale as the network of neural.

The goal of this study is to validate the U-net architecture, U-net with considering a large area with A Fully Convolutional Residual Neural Network (FC-RNN) [5-7], based on deep learning for a breast tumor within a Region of Interest (ROI) with High-Power Fields (HPF) in biopsy and its corresponding ground-truth in which labeled by a pathologist with benign and malignant.

This paper is structured as follows: the introduction is discussed in Section I. Section II represents the literature review of methods and references. Section III represents the methodology. Results and discussions are explained in Section IV. Finally, Section V concludes the paper.

II. LITERATURE REVIEW

In this section, methods and references are introduced for detection of the breast cancer.

The most famous methods of BC detection are named as; Support Vector Machine (SVM) Classifier, Naïve Bayes Classifier, Decision Tree, Fully Convolutional Network (FCN), R-CNN (Residual Convolutional Neural Networks) Classifier, Bidirectional Recurrent Neural Networks (HA-BiRNN) [1, 4]. These techniques are explained in this part.

SVM Classifier is used to plan input data point into a feature space with high dimension, in which data points are divided into two classes. SVM with linear and nonlinear kernels is utilized to carry out the linear and nonlinear classifications. By utilizing different kernel functions (Gaussian, quadratic, cubic, linear, polynomials) to various histopathology data sets the efficiency of the SVM classifier increased. This technique removes the redundant features (lowest weight feature) in all iterations [8-13].

Naïve Bayesian (BN) networks are highly scalable, requiring some linear parameters in the variables (features/predictors) in a learning problem. Figure 1 illustrates a BN in which each variable has the value of conditional probabilities. In the graph, nodes present arcs and variables show a relation between them. It includes one parents and many children. Child nodes are independent in the background of their parent. These classifiers are utilized to generate probability estimations [14-15].

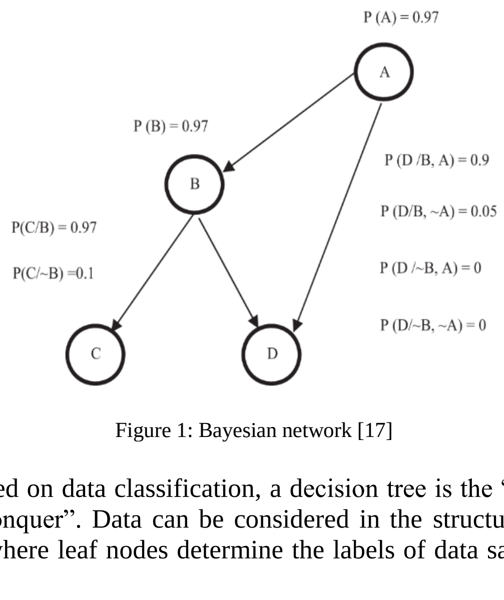


Figure 1: Bayesian network [17]

Based on data classification, a decision tree is the "divide and conquer". Data can be considered in the structure of a tree, where leaf nodes determine the labels of data samples.

Various features are shown by internal nodes. The tree is travelled from root to leaf for the classification of the particular data sample. Finally, the classification results are provided by the leaf node [16].

Convolutional Neural Networks (CNNs) represent objects at various orientations and scales by automatically learning features from the given images. By adding the layer numbers (CNN model depth) more particular specifications can be gained, which have a conclusive role to solve various computer vision schemes. For example, detection of objects, classification, and segmentation. Therefore, different ways have been suggested with deep learning approaches to solve the image segmentation problem [17-18].

Fully Convolutional Network (FCN) architectures are one of the well-known architectures for semantic segmentation, which its base is encoding (convolutional) and decoding (deconvolutional) layers. This is used to transform the image classification into image filtering networks [19]. A new approach of FCN is a hierarchy of decoders and each one is related to each encoder called SegNet. The decoder network does non-linear up-sampling of their input feature maps by utilizing the max-pooling indices taken from the related encoder [20].

R-CNN (Residual Convolutional Neural Networks) classifier is widely utilized in time sequence modeling. It is deep in the time-sequential dimension. Also utilizes the spatial data among the machines (SSVM) and Conditional Random Fields (CRF). The loss function is minimized by graphical models which are built on pixel probabilities supplied by a CNN [21-22].

The limitations of the mentioned technique are as follows: for large datasets, the SVM classifier is not proper and also not impressive on great computer vision applications. Naïve Bayes Classifier in the absence of training data generates bad results. To train the network, RCNN needs more time [22].

III. METHODOLOGY

In this study, U-net a network and training strategy is utilized in which, by the correct usage of data augmentation, the existing marginalized samples are more efficient. The structure includes a contractile way to take the background and a symmetrical expansion way that provides exact localization. By skipping connections between decoding and encoding layers, maintain important information from the input features. By this method, with very few images the network can be trained end-to-end and have a better performance rather than the previous best method (a sliding-window convolutional network) for segmentation of neural structures. The flowchart of proposed work is presented in Figure 2. For segmentation with U-net model, first image processing is done. Then, we have model training and testing process.

A. Dataset

A dataset with well-annotation is a requirement to generate a robust and novel technique for BC detecting. There are not enough available samples and patients' demographic information. In this study, the marginalized dataset is gained from the TCGA[1] (The Cancer Genome Atlas). The number of training data and testing data is 1000 and 500, respectively.

The used pathology images dataset including Whole Slide Images (WSIs) in which the Region of Interest (ROI) is selected, and the section of the High-Power Fields named by the patch is determined. These classified images are benign and malignant tumors. Then, these images are pre-processing and de-noising, after that U-net method applied to them. The suggested method has generated highly efficient and accurate results in comparison to the available methods.

To increase the dataset, augmentation is employed to decrease the problem of limited data size. The selection of the data augmentation approach on the dataset is complex; because augmentation techniques of natural images has not good result for medical images. This method for dataset is explained in the data augmentation section (3.4).

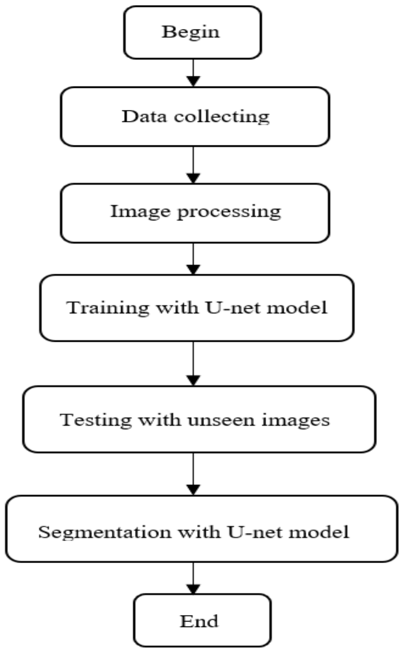


Figure 2: Flowchart of proposed work

B. Architecture of the Network

The architecture of the network is shown in Figure 3. It includes a descending path (the left side of the shape) that has convolution and max-pooling layers and at ascending path there are convolution and up-sampling layers (the right side of the shape). The descending path has a special structure of a convolutional network. In each down-sampling block, two convolutions with a kernel size of 3×3 (padded convolutions) were each tracked by an Exponential Linear Unit (ELU) that tend to converge cost to zero faster and produce more accurate results. Also, has a smooth manner for a negative value. The number of feature channels is doubled at each down sampling step. Therefore, in the down-sampling stage, the size of the max-pooling kernel size is used to divide the size of the input image at each max-pooling layer. Also, there is a 2×2 max pooling operation with stride 2 for down sampling. Also, in this path, the dropout layer with factor of 0.1 is added to avoid over fitting in the network.

The right path includes of an up-sampling feature map traced by a 2×2 convolution ("up-convolution") for decode operation that the number of feature channels are halved and an interpolation with the relating cropped feature map from

[1]https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga

the left (contracting) path for adding data to the image, Then, two 3×3 convolutions, each followed by an ELU activator, were applied to this concatenated image. Therefore, in the up-sampling stage, by convolutions the size of the input image is augmented, where kernel weights are learned during training. In the last layer, a 1×1 convolution layer is utilized to map each 64-component feature vector to the favorable number of classes. Totally, there are 23 convolutional layers.

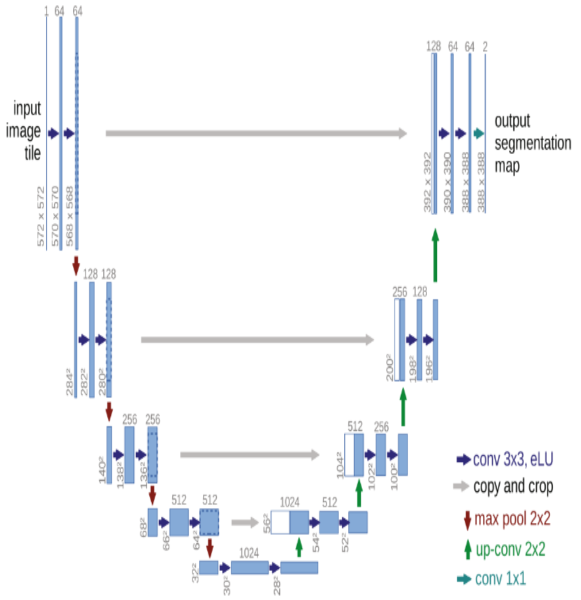


Figure 3: Proposed U-net architecture

The two paths of the U depict the information incorporation that exists between the down-sampling path into the up-sampling path, the convolution layers are copied from the contractile path to the relating expanding part. Thus, complete information gained in the contracting component of the network is utilized at the expanding path. The size of the output and input images is equal.

C. Training

The input images and their related segmentation maps are utilized to train the network with the 'Adam' optimizer with a learning rate = 0.001.

A pixel-wise soft-max over the final feature map mixed with the cross-entropy loss function to calculate the energy function. The soft-max is computed as [5-6]:

$$p_k(x) = e^{a_k(x)} \Big/ \sum_{k'=1}^{K} e^{a_{k'}(x)} \quad (1)$$

where: $a_k(x)$ = Activation in feature channel feature channel k at the pixel position $x \in \Omega$ with $\Omega \subset Z^2$

K = Number of classes

$p_k(x)$ = Approximation of the maximum-function (i.e. $p_k(x)$ is near to 1 for the k with the maximum activation $a_k(x)$, and $p_k(x)$ is near to zero for all other k).

The cross-entropy at each situation is attained as the deviation of $p_{l(x)}(x)$ from one by [5-6]:

$$E = \sum_{x \in \Omega} \omega(x) \log(p_{l(x)}(X)) \quad (2)$$

where: $l : \Omega$ = {1,…,K} is the correct label of every pixel

$\omega : \Omega$ = |R shows a weight map that has an important role in the training

Each ground-truth segmentation with its weight map is pre-processed to recompense the different frequencies of pixels from a specified class in the training data set. The network should learn the small distinct borders that are determined between touching cells.

The distinct border is calculated by morphological operations. The weight map is estimated as:

$$\omega(x) = \omega_c(x) + \omega_0 \cdot \exp\left(-\frac{(d_1(x)+d_2(x))^2}{2\sigma^2}\right) \quad (3)$$

where: $\omega_c : \Omega$ = |R is the weight map to balance the class frequencies

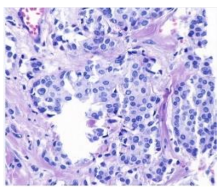
$d_1 : \Omega$ = |R defines the distance to the border of the nearest cell

$d_2 : \Omega$ = |R the distance to the border of the second nearest cell

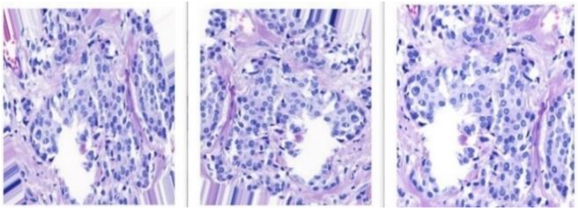
In deep networks with different paths through the network and many convolutional layers, weights with good initialization are very serious. Otherwise, components of the network might present extreme activations, while other sections do not have a contribution. Preferably, the initial weights should have a proper position such that the variance of each feature map is unit, approximately. This network structure (alternative convolution and ELU layers) can be attained by considering the primary weights from a Gaussian distribution in which its standard deviation is $\sqrt{(2/N)}$, which N represents the number of inbound nodes of one neuron.

D. Data Augmentation

Augmentation of data is necessary to train the desired network of immutability and sanity specification, when there are only a few samples of training. In microscopical images we mainly require rotation and shift immutability as well as sanity to deformations and gray value changes. Mainly accidental elastic changing of the teaching examples represent to be the main idea to train a segmentation network with very few annotated images. In Figure 4, three augmentations of the primary image (a), with brightness between 0.8 and 1.2, rotation of 40º, and zoom of 0.2 are done that are labeled by (b), (c), and (d).



(a)



(b) (c) (d)

Figure 4: Three augmentations (b), (c) and (d) of the primary image (a).

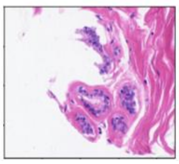
IV. RESULTS AND DISCUSSION

In this part, the simulation results for the suggested U-net method are analyzed. The introduced U-net classification is performed in the python 3.7 and TensorFlow library. The learning rate of the network is tuned at 0.0001. The maximum number of epochs is 100. Segmentation needs a patch-based approach, where small patches with determining size are attained by dividing images as the input of the neural network. The mentioned method can fully use the focused area of the local information.

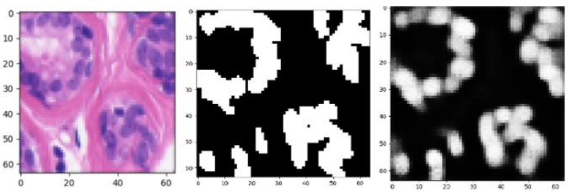
The experimental results are illustrated in Figure 5. It includes input images 1-3 in (a),(e),(i); the patch of an input image 1-3 in (b),(f),(j); the corresponding ground-truth of the input image 1-3 in (c), (g), (k) and the segmented image 1-3 of the proposed method in (d), (h), (l). For model evaluation Intersection Over Union (IOU) is used and calculated as [5].

$$IOU = \frac{target \cap prediction}{target \cup prediction} \tag{4}$$

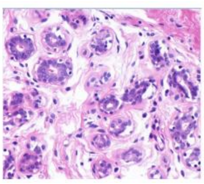
Where, target and prediction are real and predicted class of ground-truth, respectively. The Intersection Over Union of the proposed method is illustrated in Figure 6. The IOU of 68% is achieved after 20 epochs.



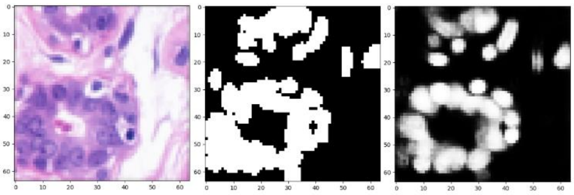
(a)



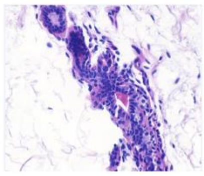
(b) (c) (d)



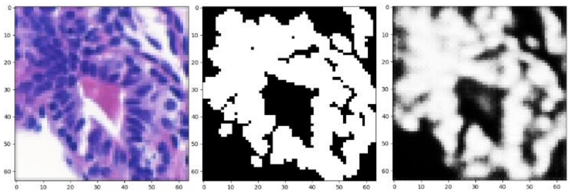
(e)



(f) (g) (h)



(i)



(j) (k) (l)

Figure 5: (a) Input image 1 with, (b) patch of the input image 1, (c) corresponding ground-truth of the input image 1, (d) segmented image of the proposed method; and (e) Input image 2 with, (f) patch of the input image 2, (g) corresponding ground-truth of the input image 2, (h) segmented image of the proposed method; and (i) Input image 3 with, (j) patch of the input image 3, (k) corresponding ground-truth of the input image 3, (l) segmented image of the proposed method.

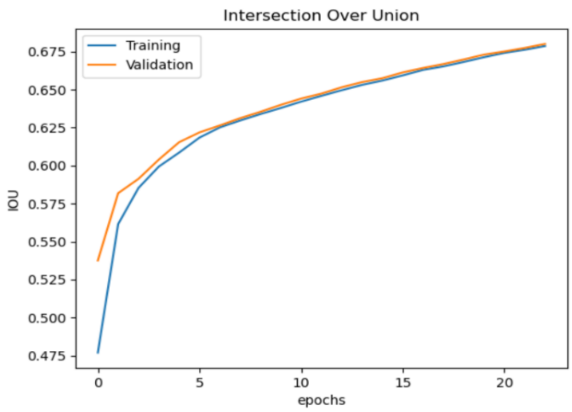


Figure 6: Intersection Over Union (IOU) of the proposed method

Figure 7: The loss of the proposed method

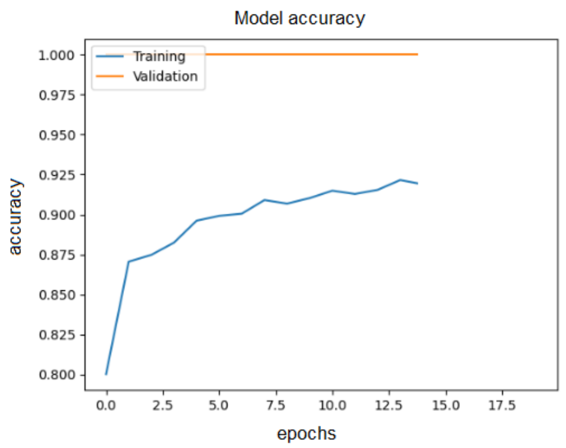


Figure 8: The accuracy of the proposed method

By dimension decreasing the computational complexity of a classifier improve, it causes sacrificing the loss of important features that may help in prediction. The loss of the proposed U-net is illustrated in Figure 7 in which after 20 epochs reaches to lower than 0.2.

The accuracy of the proposed method is shown in Fig 8. This parameter is a metric that generally explain the model performance across all classes. It is calculated as the ratio between the number of correct predictions to the total number of predictions [2].

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}} \quad (5)$$

The results determine that high accuracy was achieved (0.925) after 14 epochs. Fully automatic segmentation by deep learning segmentation with U-net for the breast is possible to perform and has reasonable accuracy rather than segmentation of ground-truth by a Model-based method which is done by a radiologist. These methods for processing need time from minutes to more than two hours, which is in part because of the demand for the post segmentation correction. In addition, testing datasets with validation permits us to approximate whether the improved method of segmentation can be widely used for other datasets. In Table 1, the existing methods are listed. For small datasets, the SVM classifier is proper and also not impressive on great computer vision applications. Naïve Bayes Classifier in the absence of training data generates bad results. U-net with convolutional structure is proper for computer vision and provide good result in comparison to other works.

Table 1
Analysis of the BC Detection Ways

Method	Precision	Accuracy	Recall
HA-BiRNN	80.00	82.00	79.00
SVM	95.00	95.00	93.60
Naïve Bayes	95.60	95.60	93.60
RCNN	91.00	91.00	89.00
U-net	92.00	92.50	90.00

V. Conclusions

This paper presents the segmentation of a breast tumor using a deep convolutional neural network based on the U-net. This architecture attains excellent performance in segmentation of breast tumor patch. Data augmentation is performed by elastic deformation, requiring a small number of annotated images and corresponding training time. By this segmentation method, high percentage of Intersection Over Union (IOU) and accuracy with low losses are attained.

References

[1] A. Reddy Vaka, B. Soni and S. Reddy, “Breast cancer detection by leveraging Machine Learning”, ICT Express 2020, pp.1-5.

[2] Sh. Saxena and M. Gyanchandani, “Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review”, Journal of Medical Imaging and Radiation Sciences xx (2019), pp. 1-12.

[3] P. Wang, Q. Song, Y. Li, Sh. Lv, J. Wang, L. Li, and H. Zhang “Cross-task extreme learning machine for breast cancer image classification with deep convolutional features”, Biomedical Signal Processing and Control J, Vol.57, 2020, pp.1-10.

[4] X. Asri, H. Mousannif, H. Moatassime, Th. Noel, “Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis”, Procedia Computer Science, vol. 83, 2016, pp: 1064 – 1069.

[5] O. Ronneberger, Ph. Fischer, and T-B. Lu, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, Springer International Publishing Switzerland, May 2015, pp. 234–241.

[6] Y. Zhang, J-H. Chen, K-T. Chang, V- Y.Park, M-J. Kim, S. Chan, P. Chang, D. Chow, A. Luk, T. Kwong, M-Y. Su, “Automatic Breast and Fibroglandular Tissue Segmentation in Breast MRI Using Deep Learning by a Fully Convolutional Residual Neural Network U-Net”, *Academic Radiology* J, Vol. 26, 2019, pp.1526-1535.

[7] V- K. Sing, H-A. Rashwana, S. Romania , F. Akram , N. Pandey, Md. M- K. Sarker , A. Saleh , M. Arenas , M. Arquez , D. Puig , J. Torrents-Barrena, “Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network”, *Expert Systems With Applications*, vol. 139 , 2020, pp. 1 – 14.

[8] F. A. Spanhol, L. S. Oliveira, C. Petitjean, & L. Heutte, “A dataset for breast cancer histopathological image classification”, *IEEE Trans Biomed Eng*, vol.63, no. 7, 2016, pp.1455–1462.

[9] J. A. Cruz & D. S. Wishart, “Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2”, 2006, pp. 59–77.

[10] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, & P. Honeine, “Multiple instances learning for histopathological breast cancer image classification”. *Expert Syst Appl 117*, 2019, pp. 103–111.

[11] T. Araujo, G. Aresta, & E. Castro, “Classification of breast cancer histology images using convolutional neural networks” *PLoS One*, vol. 12, no. 6, 2017, pp.1–14.

[12] Y. Li, J. Wu, & Q. Wu, (2019). “Classification of breast cancer histology images using multi-size and discriminative patches based on deep Learning”, *IEEE Access*, 2019, vol. 7, pp. 21400–21408.

[13] A. Osareh, & B. Shadgar, “Machine learning techniques to diagnose breast cancer”, *5th International Symposium on Health Informatics and Bioinformatics*, 2010, pp. 114–120.

[14] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, & D. I. Fotiadis, “Machine learning applications in cancer prognosis and prediction”, *Comput. Struct Biotechnol J*, vol. 13, 2015, pp. 8–17.

[15] S. B. Kotsiantis, (2007), “Supervised machine learning: a review of classification techniques”, *Informatica, 2007, vol.* 31, no. 3, 2007, pp. 249–268.

[16] I. H. Witten, & E. Frank, “Data mining practical machine learning tools and techniques”, (2nd ed.), *Morgan Kaufmann Publishers*, 2005.

[17] Sh. Saxena, M. Tech, & M. Gyanchandani, “Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review”, *Medical Imaging and Radiation Sciences J, 2019, vol.* 51, no. 1, 2019, pp. 182–193.

[18] J. Schmidhuber, “Deep learning in neural networks: *An overview. Neural Networks*, vol. 61, 2015, pp. 85-117.

[19] J. Long, E. Shelhamer, & T. Darrell, “Fully convolutional networks for semantic segmentation”, *In Processing of the IEEE conf on computer vision and pattern recognition (cvpr)*, 2015, pp. 3413-3440.

[20] V. Badrinarayanan, A. Kendall, & R. Cipolla, “Segment: A deep convolutional encoder-decoder architecture for image segmentation”, *IEEE Trans on Pattern Analysis and Machine intelligence*, vol. 39, no. 12, 2017, pp. 2481-2495.

[21] N. Dhungel, G. Carneiro, & A.P. Bradley, “Deep learning and structured prediction for the segmentation of mass in mammograms”, *In Processing of the international conference on medical image*

computing and computer-assisted intervention (MICCAI), 2015, pp. 605-612.

[22] N. Dhungel, G. Carneiro, & A.P. Bradley, "A deep learning approach for the analysis of masses in mammograms with minimal user intervention", *Medical Image Analysis*, vol. 37, 2017, pp. 114-128.