May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

Deep Learning Extraction of Features for Early Detection of Breast Cancer in Histological Images

[1]Deepa B G, [2]S Senthil

[1]Research Scholar, School of Computer Science and Applications, REVA University
Bangalore, India, deepabg03@gmail.com
[2]Professor & Director, School of Computer Science and Applications
REVA University, Bengaluru, India, senthil_udt@rediffmail.com

Article Info
Volume 83
Page Number: 3831-3838
Publication Issue:
May-June 2020

Abstract

The Deep Neural Features Extraction is new method of extracting information from an image, this method focuses on extracting the portion of an image consisting of mitosis in histological images, which need to be segmented and processed for detection of breast cancer in human beings. Thus, we need a method that identifies the region of histology images having mitosis. As these mitosis is a sign of early stages of breast cancer, we proposed an algorithm for detecting the region of images through deep neural features. These histology images are preprocessed for enhancing the region that represents the mitosis, as these mitosis are primary cause of the breast cancer, we incorporated a new method of Deep Neural Features that may be helpful in extracting the required information from an image at different layers. Each layer contributes sufficient information towards detecting the mitosis in histology images, thereby the cancerous tissues are detected and stages of breast cancer is determined, whether the stages of image is benign or malignant. The proposed deep neural feature has yielded an accuracy of 98.38 % over a benchmark dataset.

Keywords: *Filtering, Deep Neural Network, Feature Extraction, Classification, Histology images.*

Article History
Article Received:*19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *12 May 2020*

1. Introduction

The serious issue in women is Breast Cancer, which needs to be detected in early stages of it and processed for further treatment. So that the lives of every women who suffers from breast cancer can be saved. Thus, proposed a new method of detecting the breast cancer in early stages of it and suitable treatment may be given to women. These histology images are considered for detecting the mitosis in images, as these mitosis are the main cause for breast cancer. These breast cancer poses serious problems, if suitable treatments are not taken. In order to detect whether the stages of breast cancer is a beginning stage are in ending stage are determined using the deep neural features extracted from an image. The extracted information from an image is processed through an appropriate classifier. So that the cancerous tissues detected is determined as weather Benign and Malignant.

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

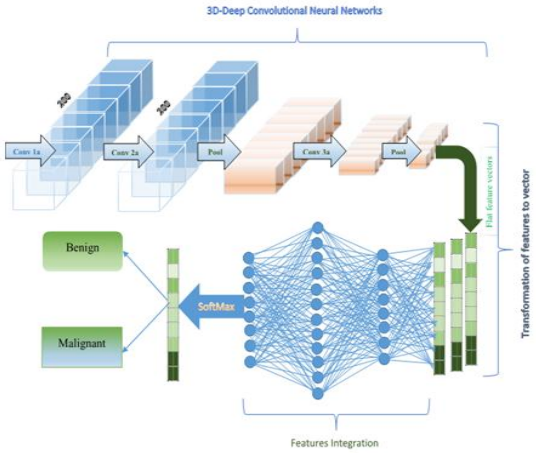


Figure 1: Overall architectural framework of the proposed deep learning Feature Extraction.

The research article has been presented with three (3) significant contributions like:

1. The proposed deep neural features extraction has contributed towards new method features extraction in terms of color and texture based information.
2. The Neural features performs the information retrieval through texture based representation of histology images
3. The Neural features perform the information retrieval through color based representation concatenated with texture based methods.

The overview of the proposed method shall be seen in different sections like section 2 describes the related research papers referred to come-up with new techniques of deep neural features extraction. Section 3 presents the proposed deep neural features extraction for detection breast cancer in early stages of cancer. Section 4 represents the results and analyzes the results of the proposed methods over a benchmark dataset. Section 5 concludes the research article with inadequate contributions made towards detection of cancerous tissues in histology images.

2. Related Work

The deep weighted feature representation learning plays a vibrant role in understanding the features of a dataset consisting of images 40x, 100x, 200x, 400x. The related research articles. Features are extracted using kernels and these kernels are used for global feature extraction learning. The research information mentioned in [30] paves certain ways to extract features of an image. The research articles [24], [26], [28] have given a hint as to how the processing is to be performed initially before detecting and segmenting the nuclei of images.

Deep neural techniques guided by local clustering [7], [20], [21], [22], [23] is also used to extract features of an image with certain methods like grouping of similar items on certain criteria has helped to analyze and understand the features of an image. Computer Aided Diagnosis [2], [3], [6], [8] also plays a suitable role in detecting the nuclei in histological images but also faces certain deficiencies like true positives against ground truth has reduced due to various other parameters such deficiencies have been addressed in our proposed research article while detecting the mitosis in histology images.

3. Proposed Deep Weighted Features Representation Learning

The representation of images is one of the important process of image processing and machine learning, as it is essential for prediction of detected regions of images into different classes. Thus, we proposed a new method of feature extraction Deep Weighted Representation Learning. The process of representation learning is represented pictorially in fig.1, where the features are extracted with kernels. The process of feature representation learning is presented in subsequent sections.

Extraction of Deep Weighted Features

The features of an RGB image is extracted by considering the texture and color based Convolutional neural Networks. The texture includes Local Binary Pattern (LBP) and color based feature extraction is used to extract required information from an image. Further, the extracted information from an image is subjected to classification of information into different classes. The proposed research article has 2 classes like Benign and Malignant.

Training

The detection of certain features in breast images that may pose cevior health issues in women as per [12], [15], and [16]. Thus, we proposed a new method that detects those features from an image and determines the stage of breast cancer whether it is a benign or malignant. We addressed the problem of breast cancer as (1).

$$LBP + Color\ Features \quad (1)$$

$LBP_{1,3}$ [9], [10], [11] is a texture feature descriptor that places a constraint of binarization by applying the threshold values, where the threshold specifies the binarization values that may be used to represent the features of an image. Thus, the Texture based feature representation helps our proposed method in detecting the portion of an image, where exactly the features of benign and malignant information is represented.

The color based feature representation [13], [14] is another strategy that helped us in representing the portion of RGB image consisting of information of either Benign or Malignant. Further, the Color based information representation presents the features in concatenation with texture based feature representation learning. We trained our system to learn the features in images as such all information from an image. LBP based Texture feature representation along with color feature representation is used as a kernel for training a system with deep

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

convolutional neural networks. The deep convolutional features are represented with 3D image representation.

The system is trained with texture and color based information for convolution as a part of training, then pooling task helps in gathering sufficient information for learning the information represented by Texture and Color based information. Further, the concatenation of texture and color based information representation helps in detecting the portion of an image having certain features of benign and malignant. Thus, we concatenated the texture and color based information during convolution and pooling is used to get information from 3Dimensional convolution.

The color based information is represented in the form of RGB Image [1], [3], [4], [5] these color based information representation helps in gathering sufficient information in terms of concatenation. The concatenation us done on all 3Dimensional channels, where the dimension Red, Green and Blue channels are combined together with texture based information during convolution and sufficient information is gathered from an image for imbibing the system with sufficient information. The metrics like Euclidian distance is used to analyse the performance of the proposed method to determine the effectiveness and robustness of proposed method towards detecting the features for early detection of Benign and Malignant.

The texture LBP and color features are extracted in different folds with seed value of 5, and a scaling factor of 4 with 3 CNN encoders together forms the features of size 300 features from an input image. Further, the proposed method utilizes the features descriptors for detection of mitosis in histology images, where the presence of mitosis represents whether the histology image is an indication of breast cancer of benign or malignant. Thus, we proposed a method of extracting deep convolutional features that may be sufficient enough to predict the breast cancer in histology images. The Deep weighted feature representation learning system is trained with certain Local Binary Pattern (LBP) features along with color information together constitutes information form every patch of an image. Initially, we split the given input image into 64x64, where each patch is processed with deep kernels of LBP texture [17], [18] and color based information [19], [21]. These texture and color based information is used as a kernel to the individual patches, these individual patches are convoluted in all dimension (3-Dimensions). The information collected from all 3 dimensions are concatenated to form a color based information from an image, then texture LBP based information is also collected and concatenated with color based information together forms a features vector of every images of training portion of an image dataset.

The probability of training versus testing is considered with different ratios like 80 percent of images are considered for training and the rest 20% of images are considered for testing, case 2 considers 70 percent of images for training and 30 percent of images for testing, similarly case 3 considers 60 percent of images for training and 40 percent of images to testing purpose. These probability [25], [27] based information of training versus testing is considered to check the effectiveness of the proposed Deep weighted feature representation learning over a dataset histological images.

$$LBP_{p,r}(row_r, col_r) = \sum_{p=0}^{-1} S(g_p, g_c)\, 2^p \qquad (2)$$

$$T \approx t\left(LBP_{p,r}(row_r, col_r)\right) \qquad (3)$$

$$Color = \sum_{k=1}^{m,n} I(R_k + G_k + B_k) \qquad (4)$$

The deep weighted feature representation learning makes use of LBP texture based information along with color based information to extract sufficient information in terms of features. Thereby, the trained system with all these features together makes the system robust. The robustness towards predicting the early signs of breast cancer is possible in all datasets of 40x, 100x, 200x, 400x is possible. The result of such robustness shall be identified when the probability of training versus testing is used for predicting the portion of an image having the representation of benign and malignant. The accuracy of the proposed method is a basis for defining the usefulness of proposed method over different datasets. Further, the deep weighted features learning is considered as a state of the art method, as it works on all datasets of 40x, 100x, 200x and 400x. The deep weighted feature representation learning has made significant contribution towards detecting mitosis in histology images very accurately in comparison with other contemporary methods.

Algorithm

The proposed deep weighted feature representation learning detects the mitosis and discriminates the information gathered into different classes as Benign or Malignant.

Input: An histology image

Description: The proposed Deep weighted representation learning detects the mitosis and predicts the early signs of breast cancer and classifies the histology images into Benign or Malignant.

Output: The system predicts the breast cancer is Benign or Malignant.

Step 1: [Initialization]

Input images are trained with features of deep

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

convolution of features

Step 2: [Processing for Features Extraction]

Step 2.1: Solve eq. (2)

Step 2.2: Calculate LBP features

Step 2.3: Solve eq. (3)

Step 2.4: Calculate threshold information

Step 3: [Optimization]

Step 3.1: Concatenate LBP with color information by solving eq. (4).

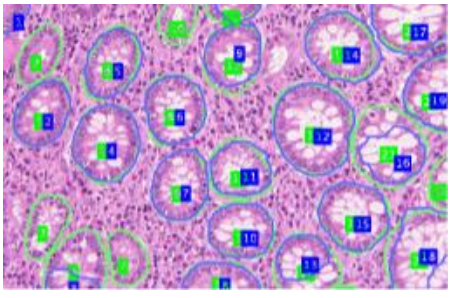
Step 3.2: Use LBP with color features as kernel or mask for Convolution phase of training the system.

Step 4: **[Output]** Predicts the region of mitosis and classifies the histology images into different classes

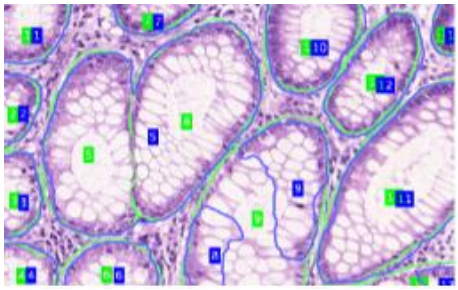
Algorithm 1: Deep weighted feature representation learning for prediction of Breast Cancer and classification of cancerous tissues detected into benign or malignant.

Testing

The testing is a part of assessment of images, as it is dependent on information gathered from training phase. The system is trained with 3-Dimensional features of color and textures, similarly, during the testing phase, the process of features extraction with LBP texture features along with color information is done as a part of kernel or sliding window over every individual patches of an input test image, thereby the proposed method finds the similarity metric learning using Euclidean distance to detect the information represented from training phase. The metric learning helps the proposed method to detect the portion of an image having such similarity in histology images are detected. So that the prevention of breast cancer is possible with appropriate treatment.



(a)



(a) (b)

Figure 3. Predicted results of individual images of a dataset consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

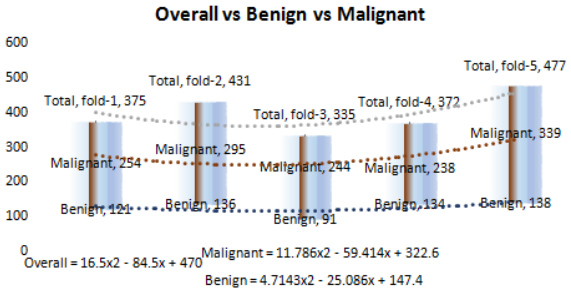


Figure 4: Predicted results of individual images of a dataset 40x consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

The graphical representation of the deep weighted learning is assessed with ground truth results by considering the total number of benign and malignant images across different datasets like 40x, 100x, 200x and 400x.

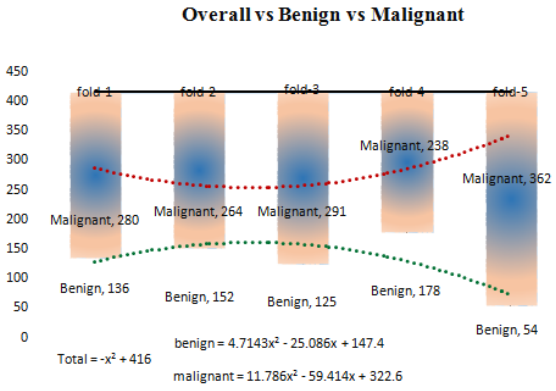


Figure 5: Predicted results of individual images of a dataset 100x consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

The results of deep weighted feature learning with respect to the images of dataset 100x of histology is presented in fig.5. Further, the proposed deep weighted feature learning is much better than other methods on every folds of classification. The fold 2 of the proposed deep weighted feature learning is comparatively less than Rakhlin.et.al [30], and has produced good accuracy of classifying the breast cancer issues comparatively better than all other methods mentioned in fig.8.

Further, the resulting values from a proposed method has been represented with polynomial equation in fig.5 to state that the suggested proposed method has produced improved results than other existing methods and it is an obvious that the proposed method had yielded good results for images of dataset 100x on every folds except fold 2 of Rakhlin.et.al [30].

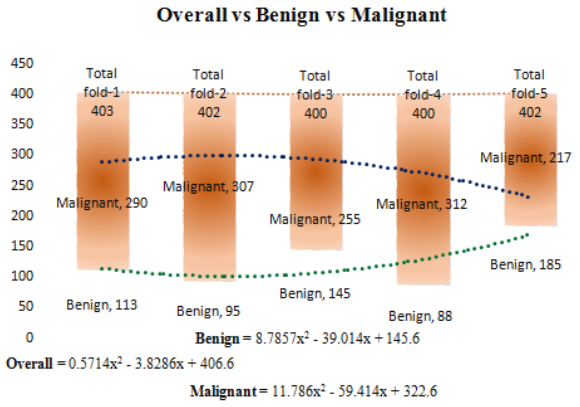


Figure 6: Predicted results of individual images of a dataset 200x consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

The accuracies of proposed deep weighted feature learning shows an evidence that the proposed method produced enhanced results than all other existing methods for classification of benign and malignant issues of women. The ground truth results are used as in a reference to calculate the accuracy of the proposed method with respect to other existing methods over a dataset 200x. Some of the challenges were to be addressed were the problems of over segmentation in existing methods over a dataset 200x of histology images, as there were some regions in images of 200x with nuclei less than the estimated size.

Another challenging issues of existing methods with respect to the dataset of 200x is to detect the size of mitosis is very minute, which needs to be processed with appropriate methodology like texture based processing. Thus, we addressed the problem of understanding the minute areas of images of 200x, which is very much essential for addressing during the segmentation of images as a part of preprocessing. Thus, the proposed method focuses on such issues as a challenging interest.

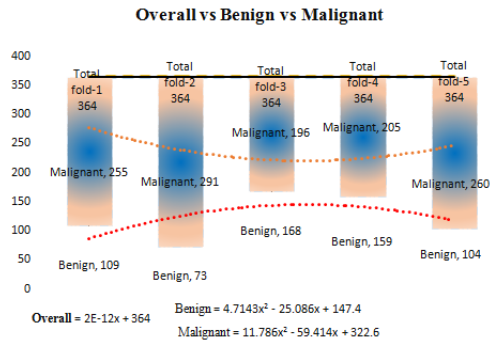


Figure 7: Predicted results of individual images of a dataset 400x consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

Similar to the challenging issues of 200x, the over segmentation was a problem in case of other existing methods. Thus, we addressed the problem of over segmentation of minute regions of images having benign information were addressed with proposed methods as a part of preprocessing. Further, the polynomial equations of proposed methods over images of 400x is an evidence that our methods has out-performed on all folds of the processing.

The predicted results have a out-performed other existing methods in-spite of the fact that the issues are to be addressed with much interest towards identifying the regions of images having benign. Similarly, the malignant regions detected in fig.3 is an evidence that the proposed method had resulted in good accuracy over dataset 400x of histology images.

Comparison of Results

The comparison of proposed deep weighted feature learning with reference to other contemporary methods are in fig.8. Further, the classification accuracies are subjected to different methods used on different datasets of histology images.

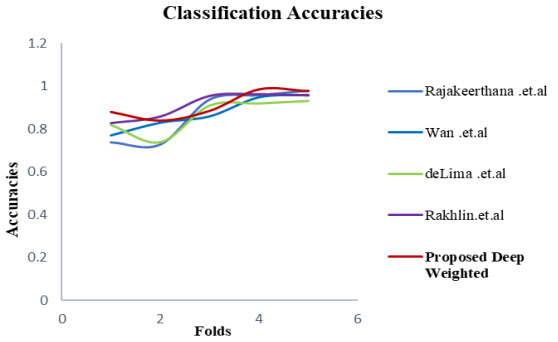


Figure 8: Predicted results of individual images of a dataset 400x consisting of Benign and Malignant. (a) Indicates the prediction of benign and (b) represents the prediction of malignant.

The classification by SVM has yielded good accuracy over individual datasets like 40x, 100x, 200x and

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

400x of [23], [24], [29], [30]. Further, the results of assessment of different contemporary methods with respect to the proposed method is an evidence that the proposed method has yielded good classification accuracies over every datasets, as the deep weighted feature extraction has given good results in comparison with other methods noticed in fig.8.

Tabular Representation of Results

The dataset 40x, 100x, 200x and 400x of histology images were processed with the proposed deep weighted feature learning. The proposed method has yielded good accuracy over other existing methods on every folds of the processing. The proposed method has given good accuracy with SVM classifier except fold 2 of Rakhil.et .al [30].

Table 1: Comparison of proposed Deep Weighted Feature Learning over all datasets of histology having Benign and Malignant information.

Methods	Dataset	Accuracy
RajaKeerthana.et.al [30]	40x, 100x, 200x, 400x	80.23 %
Wan .et.al [31]	40x, 100x, 200x, 400x	82.61 %
deLima .et.al [32]	40x, 100x, 200x, 400x	91.44 %
Rakhlin .et.al [33]	40x, 100x, 200x, 400x	97.26 %
Proposed Deep Weighted Feature Learning	40x, 100x, 200x, 400x	98.38 %

It is evident from the table 1 that the results in accuracy of projected proposed method is comparatively resulted in improved than all other methods. Table.1 is an evidential representation of proposed method with respect to other existing methods. Further, it has been represented as a state of the art

4. Conclusion

The research paper presented a new features extraction method that may be concatenated with LBP texture based information along with color information, so that the robustness of predicting the early signs of breast cancer is made possible. The classification task performs the predicted breast cancer tissues into different classes like Benign or Malignant based on the convolution information processed with mask of LBP texture and color based ideologies. Thus, we state that the proposed deep weighted feature illustration learning is a state of the art technique. Further, the proposed method has made few significant contribution towards identifying the breast cancer at early stages of breast cancer.

References

[1] Agrawal, Rakesh, and Ramakrishnan Srikant. "Fast algorithms for mining association rules." In Proc. 20th international conference. very large data bases, VLDB, vol. 1215, pp. 487-499. 1994.

[2] Garcia-Molina, Hector. Database systems: the complete book. Pearson Education India, 2008.

[3] Wang, X. Sean, Claudio Bettini, Alexander Brodsky, and Sushil Jajodia. "Logical design for temporal databases with multiple granularities." ACM Transactions on Database Systems (TODS) 22, no. 2 (1997): 115-170.

[4] Bruce, Kim B., Luca Cardelli, and Benjamin C. Pierce. "Comparing object encodings." In ISTACS , pp. 415-438. Springer, Berlin, Heidelberg, 1997.

[5] Elston, Christopher W., and Ian O. Ellis. "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long- term follow- up. CW Elston & IO Ellis. Histopathology 1991; 19; 403–410: AUTHOR COMMENTARY." Histopathology 41, no. 3a (2002): 151-151.

[6] Gurcan, Metin N., Laura E. Boucheron, Ali Can, Anant Madabhushi, Nasir M. Rajpoot, and Bulent Yener. "Histopathological image analysis: A review." IEEE reviews in biomedical engineering 2 (2009): 147-171.

[7] Rosen, Paul Peter, ed. Rosen's breast pathology. Lippincott Williams & Wilkins, 2001.

[8] Elmore, Joann G., Gary M. Longton, Patricia A. Carney, Berta M. Geller, Tracy Onega, Anna NA Tosteson, Heidi D. Nelson et al. "Diagnostic concordance among pathologists interpreting breast biopsy specimens." Jama 313, no. 11 (2015): 1122-1132.

[9] Tang, Jinshan, Rangaraj M. Rangayyan, Jun Xu, Issam El Naqa, and Yongyi Yang. "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances." IEEE transactions on information technology in biomedicine 13, no. 2 (2009): 236-251.

[10] Veta, Mitko, Josien PW Pluim, Paul J. Van Diest, and Max A. Viergever. "Breast cancer histopathology image analysis: A review." IEEE Transactions on Biomedical Engineering 61, no. 5 (2014): 1400-1411.

[11] Kowal, Marek, Paweł Filipczuk, Andrzej Obuchowicz, Józef Korbicz, and Roman Monczak. "Computer-aided diagnosis of breast cancer based on fine needle biopsy

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

microscopic images." Computers in biology and medicine 43, no. 10 (2013): 1563-1572.

[12] Filipczuk, Paweł, Thomas Fevens, Adam Krzyżak, and Roman Monczak. "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies." IEEE transactions on MI 32, no. 12 (2013): 2169-2178.

[13] George, Yasmeen Mourice, Hala Helmy Zayed, Mohamed Ismail Roushdy, and Bassant Mohamed Elbagoury. "Remote computer-aided breast cancer detection and diagnosis system based on cytological images." IEEE Systems Journal 8, no. 3 (2013): 949-964.

[14] Belsare, A. D., M. M. Mushrif, M. A. Pangarkar, and N. Meshram. "Classification of breast cancer histopathology images using texture feature analysis." In TENCON 2015-2015 IEEE Region 10 Conference, pp. 1-5. IEEE, 2015.

[15] Brook, Alexander, Ran El-Yaniv, Eran Isler, Ron Kimmel, Ron Meir, and Dori Peleg. Breast cancer diagnosis from biopsy images using generic features and SVMs. No. CS Technion report CS-2008-07. CS Dept, Technion, 2008.

[16] Zhang, Bailing. "Breast cancer diagnosis from biopsy images by serial fusion of Random Subspace ensembles." In 2011 4th International Conference BMEI, vol. 1, pp. 180-186. IEEE, 2011.

[17] Israel Institute of Technology dataset;. Available from: ftp.cs.technion.ac.il/pub/projects/medic-image.

[18] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In ANIPS, pp. 1097-1105. 2012.

[19] Cireşan, Dan C., Alessandro Giusti, Luca M. Gambardella, and Jürgen Schmidhuber. "Mitosis detection in breast cancer histology images with deep neural networks." In International conference on medical image computing and computer-assisted intervention, pp. 411-418. Springer, Berlin, Heidelberg, 2013.

[20] Litjens, Geert, Clara I. Sánchez, Nadya Timofeeva, Meyke Hermsen, Iris Nagtegaal, Iringo Kovacs, Christina Hulsbergen-Van De Kaa, Peter Bult, Bram Van Ginneken, and Jeroen Van Der Laak. "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis." Scientific reports 6 (2016): 26286.

[21] Sirinukunwattana, Korsuk, Shan E. Ahmed Raza, Yee-Wah Tsang, David RJ Snead, Ian A. Cree, and Nasir M. Rajpoot. "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images." IEEE transactions on MI 35, no. 5 (2016): 1196-1206.

[22] Spanhol, Fabio Alexandre, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. "Breast cancer histopathological image classification using convolutional neural networks." In 2016 IJCNN, pp. 2560-2567. IEEE, 2016.

[23] Cruz-Roa, Angel, Ajay Basavanhally, Fabio González, Hannah Gilmore, Michael Feldman, Shridar Ganesan, Natalie Shih, John Tomaszewski, and Anant Madabhushi. "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks." In MI 2014: Digital Pathology, vol. 9041, p. 904103. International Society for Optics and Photonics, 2014.

[24] Macenko, Marc, Marc Niethammer, James S. Marron, David Borland, John T. Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E. Thomas. "A method for normalizing histology slides for quantitative analysis." In 2009 IEEE ISBI: From Nano to Macro, pp. 1107-1110. IEEE, 2009.

[25] Fukushima, Kunihiko, Sei Miyake, and Takayuki Ito. "Neocognitron: A neural network model for a mechanism of visual pattern recognition." IEEE transactions on systems, man, and cybernetics 5 (1983): 826-834.

[26] LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86, no. 11 (1998): 2278-2324.

[27] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." In Proceedings of the 27th ICML-10, pp. 807-814. 2010.

[28] Zeiler, Matthew D. "Adadelta: an adaptive learning rate method." arXiv preprint arXiv:1212.5701 (2012).

[29] Maaten, Laurens van der, and Geoffrey Hinton. "Visualizing data using t-SNE." JMLR-9, no. Nov (2008): 2579-2605.

[30] Nahid, Abdullah-Al, Mohamad Ali Mehrabi, and Yinan Kong. Histopathological breast cancer image classification by deep neural network techniques guided by local clustering." BioMed research international 2018 (2018).

[31] Alexander Rakhlin, Alexey Shvets, Vladimir Iglovikov, and Alexandr A. Kalinin, Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis, biomedical research Feb. 5, 2018.

May-June 2020
ISSN: 0193-4120 Page No. 3831-3838

[32] Teresa Arau ´jo1, Guilherme Aresta, Eduardo Castro, Jose ´ Rouco, Paulo Aguiar, Catarina Eloy, Anto ´nio Polonia, Aure ´lio Campilho1Classification of breast cancer histology images using Convolutional Neural Networks, PLOS one, June 1, 2017

[33] Abdullah-Al, NahidandYinanKong, Involvement of Machine Learning for Breast Cancer Image Classification: A Survey, Hindawi Computational and Mathematical Methods in Medicine, Volume 2017, Article ID 3781951, 31 December 2017