Segmentation Of Medical Images

Randeep Kaur Saini
Information Technology
SIES Graduate School of Technology
Navi Mumbai, India

Stefy S Raj
Information Technology
SIES Graduate School of Technology
Navi Mumbai, India

Anshu Kumari Yadav
Information Technology
SIES Graduate School of Technology
Navi Mumbai, India

Prof. Stuti Ahuja
Information Technology
SIES Graduate School of Technology,
Navi Mumbai, India.

Abstract*—Breast cancer is the most prevailing type of cancer in women and men in andaround the world and mostly in India. Breast cancer is growing as one ofthe most chronic diseases of this age, because ofinadequate of advancementin medical field pathologists perform qualitative analysis of tissues throughvisual inspection. This could produce subjective results and is not suitablefor large scale processing. The prediction of the accurate percentage of canceron H and E image is not convincing able. H and E -Hematoxylin and Eosinis a conventional, highly seasoned method that allows the localization ofan antigen within a cell or a tissue with high resolution. H and E imagesare used for disease diagnose, biological researchers and drug development.The methodology used is the neural network (NN) to segment H and Eimages to acquire the quantitative measurements of the stained nucleus forthe determination of over-expressed HER2 protein.

I. INTRODUCTION

Breast cancer is the most prevailing type of cancer inwomen and men inand around the world and more in India. Breast cancer is growing as oneof the most chronic diseases in this age because of the lack of advancementin the medical field, In the recent trends in image processing, it shows howthe medical field can be engineered with the latest technologies. Few problemsdue to the manual assessment of histology images, such as inter- andintra-observer variability, inability to assess subtle visual features, and thetime taken to examine whole slides, can be avoided with the use of computationalpathology with the improvements in computer vision techniques andhardware, For the sake of conceptualizing under the microscope, differentstructures of the tissue are colored with different stains. then, a detailedstudy on these stained tissues has been for the detection of lesions or tumorsdone by a pathologist.
Quantitative examination can be used to supportpathologists' decision about the closeness or thenonappearance of a disease,besides to help in infection development evaluation. With the use ofnew advancements in image processingtechniques in computer vision severalmethods have been used for the detection of breast cancer accurately. Thisproject will give the most accepted computer-aided technique for the analysisof breast cancer from a histopathology image.

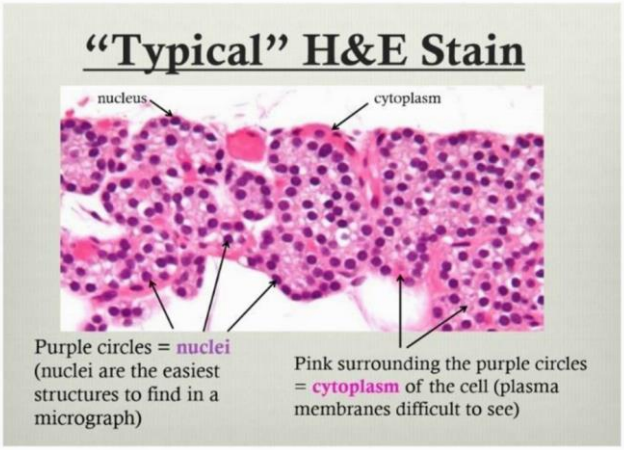

Figure 1: H&E Image

II. NEED OF PROJECT

Segmentation of medical image is done so that the accurate percent of protein can be quantified from an H and E image so that it will help in predicting the stage of cancer with accurate measures. At present no such system is developed inpathology field that can quantify the accurate percentage of proteins in a H and E image so in this project we are going to overcome the present system challenges.
Limitation of present visual system are:
- Difficulty in counting positive (or negative) cells
- Overlapped stained nuclei
- Variability of immunostaining presence of other irrelevant objects
- Takes time to manually count these cells
- Wide intra and inter observer results because of subjectivity.

III. LITRATURE SURVEY

- Breast Cancer in IHC -Nucleus:

(a) Di Cataldo, Santa, et al. "Automated segmentation of tissue images for computerized IHC analysis." *Computer methods and programs in biomedicine* 100.1 (2010): 1-15. This paper[1] presents two automated methods for the segmentation of immunohistochemical tissue images that overcome the limitations of the human intervention for assessment and the use of the existing computerized techniques. The first independent method, based on unsupervised color clustering, recognizes the target cancerous areas automatically in the specimen and disregards the stroma; the second method, based on color separation and morphological processing, detects automated segmentation of the nuclear membranes of the cancerous cells. Extensive experimental results on real tissue images demonstrate the accuracy of our techniques compared to manual segmentation; additional experiments show that our

techniques are more performed in immunohistochemical images than the popular approaches based on active contours or supervised learning. The proposed procedure can be exploited for any applications that require tissues and cellexploration and to perform reliable and standardized measures of the activity of specific proteins involved in multi-factorial genetic pathologies.

(b) Azimi, Vahid, Young Hwan Chang, Guillaume Thibault, Jaclyn Smith, Takahiro Tsujikawa, Benjamin Kukull, Bradden Jensen, Christopher Corless, Adam Margolin, and Joe W. Gray. "Breast cancer histopathology image analysis pipeline for tumor purity estimation." In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 1137-1140. IEEE, 2017. In this paper[7] we develop a quantitative image analysis pipelinewhich includes annotation, classification, and segmentation. The method was introduced to provide a systematic comparison betweenpathologists' TP and image-based TP estimations. We envisionthat this framework will allow us to do a better understandingof TP estimation based on quantitative image analysis.

- Breast Cancer In IHC Membrane

(a) Kumar, Neeraj, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. "A dataset and a technique for generalized nuclear segmentation for computational pathology." *IEEE transactions on medical imaging* 36, no. 7 (2017): 1550-1560. In this paper[8] Nuclear segmentation in digital microscopic tissue images can enable the extraction of high-quality features for nuclear morphometrics and other analysis in computational pathology. Conventional image processing techniques such as watershed segmentation and Otsu thresholding do not work effectively on challenging cases, such as crowded nuclei and chromatinsparse. In contrast, segmentation using machine learning can generalize across variousnuclearappearances. However,trainingalgorithmsof machine learning require datasets of images in whicha huge number of nuclei have been annotated. Publicly accessibleand annotated datasets, along with widely agreedupon metricsto compare techniques, have catalyzed tremendous innovationand progress on other image classification problems, specifically inobject recognition.Inspired by their success, we introduce a hugedataset of HE stained tissue images that can be accessed publiclywith more than 21,000 pink stained annotated nuclear boundaries,whose quality was validated by a medical doctor. Our dataset istaken from multiple hospitals that include various nuclear appearancesfrom several patients, organs and disease states, techniquestrained on it are likely to generalize well and work right on otherHE stained images. We also propose a new metric to evaluate thesegmentation of nucleus results that penalizes object and pixellevelerrors in a uniform manner. Unlike previous metrics whereonly one type of error is penalized. We also propose a deep learningbased on segmentation technique that lays special emphasison identifying the nuclear boundaries, including those betweenthe touching or overlapping nuclei, and works well on a diverseset of test images.

(b) Pezoa, Raquel, Rodrigo Rojas, Luis Salinas, Luis Pizarro, J. Reyes, and R.Gonzales. "Segmentation of IHCstained Breast Tissues Images using SVM."Centerof TechnologicalInnovation in HPC (CTIHPC) Universidad TecnicaFederico.2015. The present study[3] is focused on the processing of breast cancer images, specifically immunohistochemistry (IHC) stainedimages. IHC is a technique which allows the detection of specific proteins in tissue sections. This technique is fundamentalfor patient treatment and the prognosis of certain cancers. Thepresent study is focused on a specific IHC technique, that allowsthe detection of HER2 protein over-expression and is displayedby membrane staining known as HercepTest. As a rule, pathologistsperform qualitative analysis of tissues through visual inspection,this produces subjective results and which is not suitable for large scale processing. This paper describes the use of SupportVector Machine (SVM) to segment IHC images in order to obtain quantitative measurements of the stained membranes for the determination of over-expressed HER2 protein.

(c) Pezoa, R., Maureira-Fredes, C., Arce, P., H•artel, S., Salinas, L. and Torres.Segmentation of HER2 protein overexpression in immunohistochemicallystained breast cancer images using Support Vector Machines." In J. Phys.Conf. Ser. Vol. 762, p. 012050,2016. Breast cancer is one of the most common cancers in womenworldwide. Patient therapy is done after the analysis of immunohistochemically(IHC) stained tissue sections of the patient. Theanalysis of HER2 overexpression by immunohistochemistry helpsto determine when patients are suitable for HER2-targeted treatment.Computational HER2 overexpression analysis is still a challengingtask and an open problem principally because of the subjectivityof the specialists to assess the samples and variabilityof immunohistochemistry tissue samples. In addition, the immunohistochemistryprocess can produce diverse artifacts thatmakes it difficult for the HER2 overexpression assessment. Inthis paper[5] we study the segmentation of HER2 overexpressionin IHC stained breast cancer tissue images using a support vectormachine (SVM) classifier. We assess the SVM performanceusing diverse color and texture pixel-level features including theRGB, CMYK, HSV, CIE L*a*b* color spaces, color deconvolutionlter; andHaralickfeatures:We measure classification performance for three datasets containinga total of 1750 H and E images that were previously labelled by a pathologist.

IV. METHODOLOGY

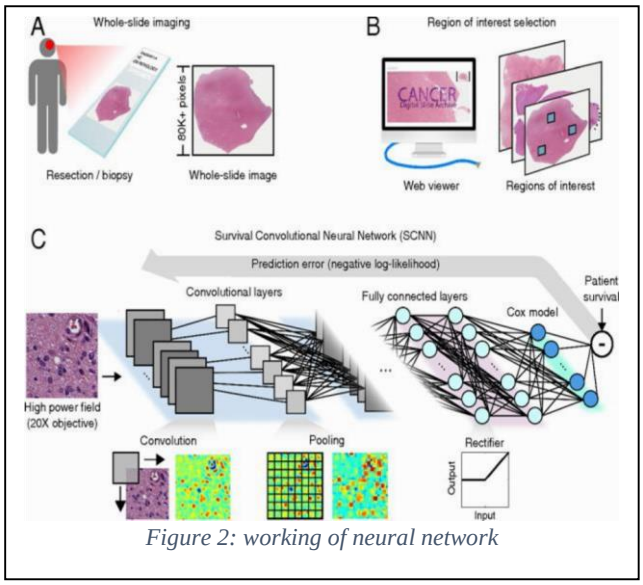

Figure 2: working of neural network

1. In this paper neural networks technique is used initialization is the first step we do in neural networks this is done because before we use weights we have to initialize them, because we don't have values for the weights yet, we use random values between 0 and 1. In python we use

   random. Seed to generate random numbers.

2. The second step is to define and format network input and target data.

3. The third step is to divide the data into three sets training data, test data and validation.

4. The fourth step is to create feed forward, back propagation three network layers and this is done to train our data in a better way so that it gives us a better accuracy

5. The fifth step is to train the network and for training neural network has to undergo many epochs or iteration and this is done in order to reach the optimal weights and biases that will give us the desired output and accurate prediction for the given input.

6. The sixth step is to make the total error zero if the total error is less than the final target error we can simulate the network and end the training and if the total error is not zero we need to do the process gain and make it zero or less than the final target error.

7. Last and final step is to simulate the network after all the above steps are satisfied and end the process.

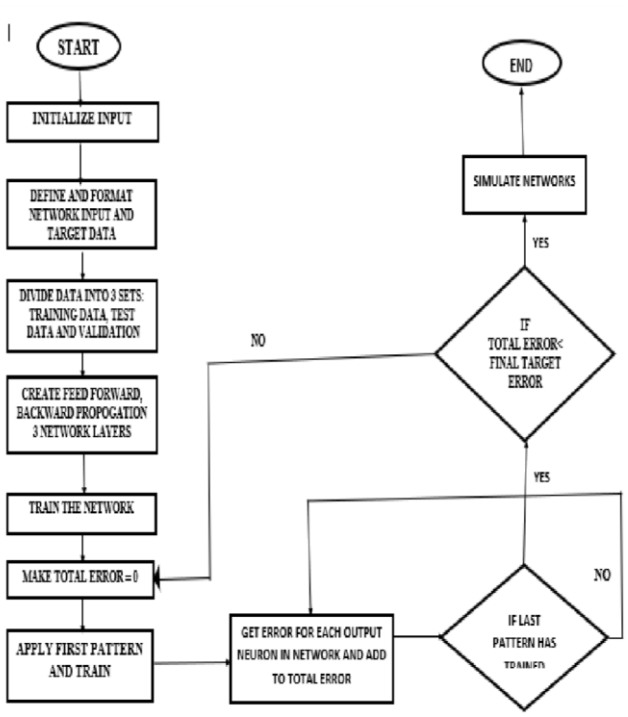

Figure 3:Flow Chart Of Program

V. RESULTS & DISCUSSION

Accuracy Percentage:

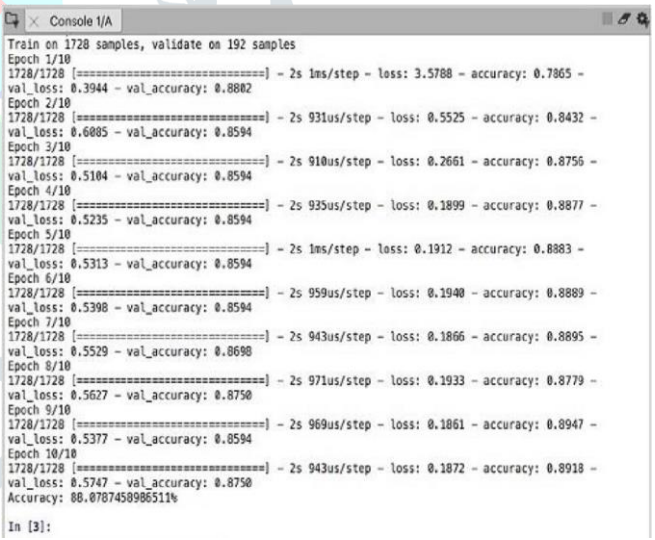

Figure 4: Results

The algorithm used to train the model is neural networks and the modelwill detect the malignant and benign H and E stain tissues images of breastcancer and these images are of [50,50] which is the resized form of the original image and [50.50] are given as an input for better computation and result ofthe model, 1200 breast cancer tissue images are used as dataset on which the neural network will work. The model uses two activation functions, first is 'relu'($y = max(0, x)$) and the second is 'softmax' (**(σ)**: $g(z) = 1 / (1 + e^{-z})$) to increase the accuracy and the model is trainedwith 30 epochs and three layers of networks to retain a better accuracy.Withthe use of these activation function it gave a very good percentage of accuracy of 80.07 percentage.

Activation Functions used here are:

i. **SOFTMAX:**

- The softmax function squashes the outputs of each unit to be between 0 and 1, just like a sigmoid function. But it also divides each output such that the total sum of the outputs is equal to 1 (check it on the figure above).
- The output of the softmax function is equivalent to a categorical probability distribution, it tells you the probability that any of the classes are true.
- Mathematically the softmax function is shown below, where z is a vector of the inputs to the output layer (if you have 10 output units, then there are 10 elements in z). And again, j indexes the output units, so j = 1, 2, ..., K.

((σ): *g(z) = 1 / (1 + e^{-z}))*

ii. **RELU:**

A rectified linear unit has output 0 if the input is less than 0, and raw output otherwise. That is, if the input is greater than 0, the output is equal to the input. ReLUs' machinery is more like a real neuron in your body. **(*y = max(0, x)*)**

ReLU activations are the simplest non-linear activation function you can use, obviously. When you get the input is positive, the derivative is just 1, so there isn't the squeezing effect you meet on back propagated errors from the sigmoid function. ReLUs result in much faster training for large networks. Most frameworks like Tensor Flow and TFLearn make it simple to use ReLUs on the the hidden layers, so you won't need to implement them yourself.

ACKNOWLEDGMENT

We extend our sincere thanks to Dr. Sanjay N Doke for your support and guidance regarding medical image dataset throughout the project.

FUTURE SCOPE

Currently, it is possible to detect the cancer cells in a given picture in future it can be used for finding other skin diseases by providing images to the system.

CONCLUSION:

Thus with respect to the above lying scope decided for the project, we have successfully implemented 'segmentation of medical images' which is a prototype which can be used in hospital and pathology to detect whether the H and E stained breast cancer tissue is malignant or benign and this is achieved by using neural networks algorithm .

REFERENCES

[1] Di Cataldo, Santa, Elisa Ficarra, Andrea Acquaviva, and Enrico Macii."Automated segmentation of tissue images for computerized IHC analysis."Computer methods and programs in biomedicine 100, no. 1,115(2010).

[2] Irshad, Humayun, Antoine Veillard, Ludovic Roux, and Daniel Racoceanu."Methods for nuclei detection, segmentation,and classification indigital histopathology: a review|current status and future potential." IEEEreviews in biomedical engineering7,97-114,2014

[3] Pezoa, Raquel, Rodrigo Rojas, Luis Salinas, Luis Pizarro, J. Reyes, and R.Gonzales. "Segmentation of IHC-stained Breast Tissues Images using SVM."Centerof TechnologicalInnovation in HPC (CTIHPC) Universidad TecnicaFederico.2015.

[4] Apou, Gr_egory, Nadine S. Schaadt, Beno_tNaegel, Germain Forestier, Ralf Sch•onmeyer, Friedrich Feuerhake, C_edricWemmert, and Anne Grote. "Detectionof lobular structures in normal breast tissue." Computers in biology and medicine 74, 91-102,2016.

[5] Pezoa, R., Maureira-Fredes, C., Arce, P., H•artel, S., Salinas, L. and Torres.Segmentation of HER2 protein overexpression in immunohistochemicallystained breast cancer images using Support Vector Machines." In J. Phys.Conf. Ser. Vol. 762, p. 012050,2016

[6] Aswathy, M. A., and M. Jagannath. "Detection of breast cancer on digital histopathology images: Present status and future possibilities." Informatics in Medicine Unlocked 8,74-79,2017.

[7] Azimi, Vahid, Young Hwan Chang, Guillaume Thibault, Jaclyn Smith,TakahiroTsujikawa, Benjamin Kukull, Bradden Jensen, Christopher Corless, Adam Margolin, and Joe W. Gray. "Breast cancer histopathology image analysis pipeline for tumor purity estimation." In 2017 IEEE 14th InternationalSymposium on Biomedical Imaging (ISBI 2017), pp. 1137-1140. IEEE,2017.

[8] Kumar, Neeraj, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. "A dataset and a technique for generalized nuclear segmentation for computational pathology." IEEE transactions on medical imaging 36, no. 7,pp 15501560,(2017).

[9] Mouelhi, Aymen, Hana Rmili, Jaouher Ben Ali, Mounir Sayadi, RaoudhaDoghri, and Karima Mrad. "Fast unsupervised nuclear segmentation and classification scheme for automatic allred cancer scoring in immunohistochemical breast tissue images." Computer methods and programs in biomed medicine 165,37-5,2018