



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

 9940 572 462

 6381 907 438

 ijircce@gmail.com

 www.ijircce.com

A Survey on Histopathological Image Classification Using Advanced Machine Learning Techniques

Apoorv Dhaygude¹, Abhijeet Kumar¹, Naman Upadhyay¹, Juli Priya¹, Prof. Ankita Kotalwar²

Department of Computer Engineering, Sinhgad College of Engineering, Pune, Maharashtra, India¹

Assistant Professor, Department of Computer Engineering, Sinhgad College of Engineering, Pune, Maharashtra, India²

ABSTRACT: The classification of breast cancer is the subject of interest in the fields of healthcare and bioinformatics, because it is the second main reason of cancer-related deaths in women. Breast cancer is analyzed using a biopsy where tissue is eliminated and studied under microscope. The identification of problem is based on the qualification and experienced of the histopathologist, who will attention for abnormal cells. However, if the histopathologist is not well-trained or experienced, this leads to wrong diagnosis. With the recent proposition in image processing and machine learning domain, there is an interest in experiment to develop a strong pattern recognition-based framework to improve the quality of diagnosis. In this work, we will use the image feature extraction approach and machine learning approach for the classification of breast cancer using histology images into benign and malignant. Using Histopathological image, we can preprocess this image after that apply feature extraction and classify the final result using SVM and Naive Bayes Classification techniques.

KEYWORDS: Histopathological image classification, breast cancer diagnoses, feature extraction, SVM classification, Naive Bayes Classification.

I. INTRODUCTION

A. BACKGROUND

Breast cancer is most common and dangerous intrusive cancer in women and the second main effect of cancer death in women, after lung cancer. The International Agency for Research on Cancer (IARC), which is part of the WHO, the numbers of deaths reasoned by cancer in the year of 2012 only come to around 8.2 million. The number of new cases is expected to growth to about more than 27 million by 2030.

Finding breast cancer quick and getting state of the art cancer treatment are the major plan of action to avoid deaths from breast cancer. In existing, it is a widely-used way to identification of breast cancer by identifying hematoxylin and eosin (H&E) stained histological slide preparations that are checked under a high powered microscope of the changed area of the breast. In medical practice, classification of breast cancer biopsy result into different plans is manually driven by experienced pathologists.

Come out machine learning approaches and enlarging image volume developed automatic system for breast cancer classification possible and can help pathologists to obtain precise identification of problem more efficient.

Breast cancer can be identified using medical images testing using histology and radiology images. The radiology images search will help to find the areas where the difference is located. However, they cannot be used to find or identify whether the area is cancerous. The biopsy process, where a tissue is gives as input and processed under a microscope to see if cancer is present, is the only sure way to find if an area is cancerous. After completing the biopsy, the identification of problem will be based on the qualification of the histopathologist, who will analyze the tissue under a microscope, looking for exceptional or cancerous cells. The histology images allow us to differentiate the cell nuclei types and their flowchart according to specific pattern. Histopathologist particularly examine the consistency of cell shapes and tissue distributions and decided the cancerous regions and malignancy degree. If the histopathologist are not well-trained, this leads to an incorrect identification of problem. Also, there is a lack of intelligent specialists, which maintain the tissue sample on hold for up to two months. There is also the problem of reproducibility, as

histopathology is a subjective science. This is right especially between non-specialized pathologists, where we can get a different identification of problem on the existing same sample. Therefore, there is an insistent demand for computer-assisted identification of the problem.

B. MOTIVATION

Breast cancer can be found or identified using a biopsy where tissue is removed and studied under microscope. The identification of problem is based on the qualification and experience of the histopathologist, who will look for abnormal cells. However, if the histopathologist is not well-trained or experienced, this may lead to wrong identification of problem. The recent proposition in image processing and machine learning domain, there is an interest to try to develop a reliable pattern recognition-based approach to improve the quality of identification of problem.

C. AIM/ OBJECTIVES

- To classify the breast cancer histology images into benign and malignant.
- To work on histopathological image dataset for breast cancer classification
- To develop features-based classification methods

II. RELATED WORK

Breast cancer is a savage disease, executing a huge number of individuals consistently. Creating a robotized dangerous BC recognition framework connected to patient's symbolism can assist managing this issue all the more effectively, making diagnosis more versatile and less inclined to mistakes. DeCaf (or deep) highlights comprise in the middle of arrangement it depends on reusing a formerly trained CNN just as highlight vectors, which is then utilized as contribution for a classifier prepared just for new order assignment. In the more to this, they display an assessment of DeCaf highlights for BC recognition, with a specific end goal to all the more likely see how they contrast with alternate methodologies [1].

This work leads to classify breast cancer histopathology images independent of their magnifications using convolutional neural networks (CNNs). They propose two different architectures like single task CNN is used to predict malignancy and multi-task CNN is used to predict both malignancy and image magnification level concurrently. Evaluations and comparisons with previous results will be carried out on Break His dataset [2].

The reason for this work is to create an insightful deep discovery and finding approach for breast disease in accordance of cytological pictures. Initially, this work exhibits a completely mechanized methodology for cell nuclei recognition and division in breast cytological images. The areas of the cell cores in the cytological picture were identified with roundabout Hough change. The expulsion of false-positive (FP) discoveries (loud circles and platelets) was achieved utilizing Otsu's thresholding procedure and fluffy c-implies grouping strategy. The division of the nuclei limits was proficient with the utilization of the marker-controlled watershed changes. Next, an astute breast malignancy grouping framework was also created [3].

The effectiveness of the treatment of breast cancer depends on its timely detection of it. An early advance in the finding is the cytological picture examination of breast material acquired straightforwardly from the tumor. This work gives in PC supported breast growth recognizable proof of issue in improvement of the examination of cytological pictures of fine needle biopsies to recognize this biopsy as either benevolent or harmful. Rather than focus on the exact division of cell nuclei, the nuclei are found by circles utilizing the roundabout Hough change system. The result circles are then sifted to keep just astounding estimations for additionally think about by a help vector machine which groups in binary identification circles as right or wrong utilizing surface highlights and the level of cores pixels as per a cores veil acquired utilizing Otsu's thresholding system [4].

This work directs some fundamental examinations utilizing the deep learning way to go with arranged breast cancer histopathological pictures from BreakHis, an openly accessible dataset at <http://webinf.ufpr.br/breast/malignancy> database. They propose a strategy in point of the extraction of picture patches for preparing the CNN and the mix of these patches for definite grouping. This strategy leads to permit utilizing the high-quality histopathological pictures from BreakHis as contribution to existing CNN, maintaining a strategic distance from adjustments of the model that can prompt a more unpredictable and computationally exorbitant engineering [5].

Current methodologies mainly depend on handcraft highlight portrayal, for example, shading, surface, and Local Binary Patterns (LBP) in arranging two areas. Contrasted with carefully assembled included based methodologies, which include undertaking subordinate portrayal, DCNN is a conclusion to end highlight extractor that might be straight forwardly gained by the crude pixel force estimation of EP as well as ST tissues in an information driven mold. These abnormal states highlight to add to the development of a directed classifier for separating the two kinds of tissues [6].

The test turns out to be the means by which to cleverly join fix level arrangement results and model the way that not all patches will be differentiative. They propose to prepare a choice combination model to total fix level forecasts given by fix level CNNs, which be the best of our insight has not been appeared previously. They apply this technique to the grouping of glioma and non-little cell lung carcinoma cases into subtypes [7].

Computerized atomic identification is a basic advance for various PC helped pathology related picture examination calculations, for example, for mechanized evaluating of breast disease tissue samples. Be that as it may, computerized core location is muddled through (1) the huge number of nuclei and the measure of high goals digitized pathology pictures, and (2) the inconstancy in estimate, shape and appearance and surface of the individual nuclei. As of late there is enthusiasm for the utilization of "Profound Learning" techniques for order and investigation of enormous picture information [8].

This work present a dataset of 7,909 breast tumor histopathology pictures procured on 82 patients, which is currently openly accessible from <http://web.inf.ufpr.br/vri-breast-cancer-database>. The dataset incorporates both benign and malignant pictures of the cancer. The undertaking related to this dataset is the robotized classification of these pictures in two classes, which would be an important PC helped finding instrument for the clinician without histopathologist. So as to evaluate the trouble of this undertaking, we show some primer outcomes acquired with state-of-the-art image classification systems [9].

There are a few problem which still exist in conventional individual Breast Cancer Diagnosis. To take care of the issues, an individual credit appraisal display in view of help vector order technique is proposed to the technology. Utilizing SPSS Clementine information mining device, the individual credit information is bunching investigation by Support Vector Machine used for binary classification. It is found in detail with the distinctive part capacities and parameters of Support vector machine. Bolster vector machine could be utilized to enhance crafted by medicinal specialists in the consideration of breast cancer growth [10].

Existing Work Disadvantages:

- Previous approaches did not work satisfactorily for a new challenging database of higher-resolution images.
- Due to the low resolution of the previous images existing work not considered textural features
- Time Consuming
- Does not handle overlapped cells.

III. PROPOSED SYSTEM

Classifying breast cancer histopathological images automatically is the most important task in computer assisted pathology analysis. However, extracting informative and non-redundant features for histopathological image classification is difficult and challenging.

In our proposed work using Histopathological image, firstly we will be applying image pre-processing technique to remove the noise of an image. After that we will be applying the feature extraction process. The feature-based approaches consist of the features extraction phase and also then classification phase. This approach mainly focuses on extracting the feature of image and classify them using machine learning classification technique. The extracted features are trained using support vector machines technique and Naive Bayes Classification technique. Finally, we compared the performance using the existing classification methods mentioned.

PROPOSED SYSTEM ARCHITECTURE

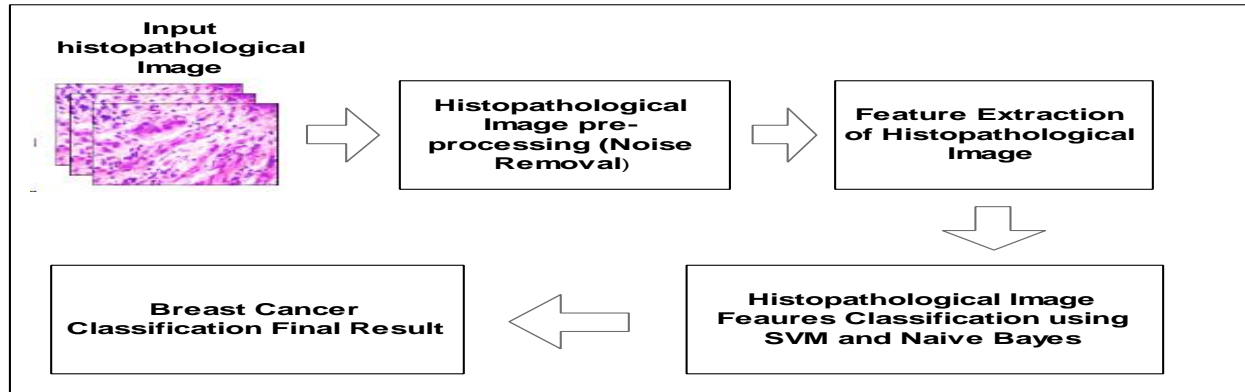


Fig: Proposed System Architecture

Proposed System Advantages:

- Work could be beneficial to obtain fast and precise quantification, reduce observer variability, and increase objectivity.
- Cell nuclei detection using image thresholding and image edge detection.
- We can measure accurate cell features.
- This application can be used by physicians from their homes or any other place.
- This work will be suitable for images with a high degree of noise and blood cells and cell overlapping, as it can successfully detect the cell nuclei.

IV. CONCLUSION

In this work, we work on the histopathological images by using Support Vector Machine (SVM) and Naive Bayes Classification with various configurations for the classification of breast cancer histology images into benign and malignant. The designed SVM topology for binary classification and Naive Bayes Classification worked well on histopathological images features in classification tasks. However, the performance of the SVM classification and Naive Bayes Classification are better compared to the one of the previously existing classification methods. SVM have become state-of-the-art, demonstrating an ability to solve challenging classification tasks for binary classification. This proposed work successfully classifies using breast cancer histology images into benign and malignant.

REFERENCES

- [1]. F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," *2017 IEEE International Conference*
- [2]. N. Bayramoglu, J. Kannala, and J. Heikkila, "Deep learning for magnification independent breast cancer histopathology image classification," *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016.
- [3]. Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," *IEEE Systems Journal*, vol. 8, no. 3, pp. 949–964, Sept 2014.
- [4]. P. Filipczuk, T. Fevens, A. Krzyzak, and R. Monczak, "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," *IEEE Transactions on Medical Imaging*, vol. 32, no. 12, pp. 2169–2178, 2013.
- [5]. F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *International Joint Conference on Neural Networks*, Vancouver, BC, Canada, July 2016, pp. 2560–2567.
- [6]. J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, 2016.



- [7] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," in IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, June 2016, pp. 2424–2433.
- [8] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, and A. Madabhushi, "Stacked sparse autoencoder (SSAE) for nuclei detection of breast cancer histopathology images," IEEE transactions on medical imaging, vol. 35, no. 1, pp. 119–130, 2016.
- [9] F. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455–1462, 2016.
- [10] Shang GaoHongmei Li, "Breast Cancer Diagnosis Based on Support Vector Machine". 2012 International Conference on Uncertainty Reasoning and Knowledge Engineering



INNO  SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details