International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Survey on Breast Cancer analysis

Shruthi G.K[1], Dr. Pushpa Ravikumar[2]

[1]Research Scholar, Dept. of CS & E, Adichunchanagiri Institute of Technology, Chikkamagaluru, Karnataka, India
[2]Professor and Head, Dept. of CS & E, Adichunchanagiri Institute of Technology, Chikkamagaluru, Karnataka, India

Abstract
Breast cancer constitutes 12% of all newly diagnosed cases globally, making it the most predominant type of cancer as a whole. Histopathology images, among all the medical image modalities, maintain the cancer's path and provide richer phenotypic relevant information. The patient's immune system is progressively predictable as a critical feature in defining the suitable treatment. Tumor-infiltrating lymphocytes (TILs), immune cells found within tumors, are emerging as key biomarkers in breast cancer. The detection of lymph node metastasis affects the management of patients with primary breast cancer significantly in terms of staging, treatment, and prognosis.Tumoour infiltrating lymphocytes score and lymph node metastasis identification plays a very important role in staging of breast cancer.This review aims to provide a comprehensive overview on how deep learning-based method is used to compute the Tumour-Infiltrating Lymphocytes (TILs) score and lymph node metastasis from breast cancer histopathological images.

Keywords: Breast cancer, T-lymphocytes, lymph node metastasis, Image Processing, CNN, Deep Learning.

I. INTRODUCTION

Medical science report states that breast cancer is the furthermost common type of cancer amongst women today. The majority of the time, tissue is taken out and examined under a microscope as part of the biopsy procedure to identify it. Untrained histopathologists run the risk of making the incorrect diagnosis. Better diagnosis is made easier with the help of automatic a [1]. Breast cancer is the most common type of cancer among humans, with 7.5% of women receiving a diagnosis in the last five years [2]. Effective treatment and improved patient outcomes depend on early detection and precise diagnosis. Histopathology, which is regarded as the standard for diagnosing breast cancer, examines tissue samples under a microscope to find anomalies that could be cancerous. On the other hand, manual histopathological image interpretation is laborious, subject to disagreements between observers, and dependent on pathologist skill. As a result, computational approaches—in particular, deep learning-driven image investigation—offer viable ways to improve the accuracy and effectiveness of diagnosis [3].

Histopathologists visually examine the pictures to identify patterns in the distribution of cancerous tissue and cell shapes and to determine the level of malignancy. Despite advances in diagnostic imaging technologies, pathologists still rely on pictorial review of histological models for accurate diagnosis, grading, and staging. It can be difficult to manually diagnose cancer from histopathological images

because human limitations may result in detection errors. Histopathological image analysis can be made more accurate and efficient by introducing computer-vision-aided diagnosis systems.

There are currently automated methods based on deep learning that can differentiate between images of benign and malignant cancer [4]. We can explain data collection, visualization, long-term storage, and synchronous browsing with the use of big data technology in the medical field. There are no longer geographical or temporal limitations on the processing of gathered pathological resources. As a result, digital pathology is now frequently utilized in pathology-related fields [5].

In the case of cancer, where the clinician's ability to prescribe the appropriate treatment for patients depends greatly on accurate diagnosis, molecular alteration, and quantification of druggable targets, pathologists must furnish dependable and repeatable data on quantitative tissue biomarkers The International Immuno-Oncology Biomarker Working Group (www.tilsinbreastcancer.org), also known as the TIL-WG, sought to learn from the experiences of other cancer biomarkers, such as estrogen and progesterone receptor quantification, HER2 IHC, histological grade, and more recently, Ki67, by guaranteeing that the TIL quantification method could be standardized across pathologists in different laboratories worldwide. [6].

Because TILs offer a glimpse into the tumor environment and are among the best illustrations of the connection between innate defenses and carcinogenesis, they are predictive prognostic markers in BC[7].The first step in the process of using deep learning to compute the Tumor-Infiltrating Lymphocytes (TILs) score in breast histopathological images is to gather a labeled dataset of histopathology images, preferably with annotations for lymphocytes and tumor regions, from public sources or with pathologists' manual assistance.PanopTils Dataset is used in the suggested method. For pathologists, examining Whole-slide images (WSIs) of lymph node sections requires a lot of time and more attention. When the cancer spreads to the lymph nodes, the prognosis procedure for pathologists becomes difficult and time consuming, and also tends to be misinterpreted. The burden of diagnosis procedure of metastasis in lymph node images can be reduced with computer-aided diagnosis systems.The dataset used for this purpose is PCAM(patch Cameleon)[8].

II. ROLE OF DEEP LEARNING IN BREAST CANCER ANALYSIS

The function of DL in the recognition and categorization of breast cancer Tumor detection and classification comes first [9]. Cruz-Roa et al. created a CNN to categorize if invasive ductal carcinoma is present in a breast cancer whole section image (WSI) plaque. The model attained a pixel-level F1 score of 76% after being trained on 300 slides and verified on 100 slides. [10]. On the other hand, Han et al. used the BreAKHis dataset to train a classifier that had an accuracy of 93.2% in differentiating between all different 8 types of malignant and benign breast cancers[11]. An object identification DL model called Faster Region Based Convolutional Neural Network (FRCNN) is employed. It is a substantially faster and more accurate variant of the R-CNN series, which also includes Fast R-CNN [12]. Yap et al. used the FRCNN model to diagnose breast cancer masses, and their outcomes were outstanding. By using FRCNN an average recall rate of 0.9236 is obtained,0.9408 precision rate is obtained, F1-score of 0.9321 is gained, and false alarm rate of 0.0621 is obtained on several sets of trained data. The FRCNN model performed even better when RGB images were used; recall rate was 0.9572, precision rate was 0.9020, F1-score was .9288, and false alarm rate was down to 0.1111 [13]. These findings show that the FRCNN model for diagnosing breast cancer has a good diagnostic accuracy and a low false positive rate. Agarwal et al. also presented a Faster Region-Based Convolutional Neural Network (Faster-RCNN) model for Full-Field

Digital Mammography (FFDM) breast cancer mass detection that same year. Using public databases, their model produced an accuracy of 0.93 and a false positive rate (FPI) of 0.78. Their model also performed exceptionally well on the INbreast dataset, obtaining an FPI of 1.69 and an accuracy of 0.99 for the analysis of malignant breast cancer masses, and an FPI of 1.0 and an accuracy of 0.85 for the analysis of benign masses[14]. DL model RetinaNet is intended for object detection. This model's central concept is the addition of a loss function known as Focal Loss, which improves detection performance by addressing the problem of class unevenness between positive and negative samples [15]. In one investigation, breast lesions were identified in ultrafast DCE-MRI sequences using a modified 3D RetinaNet model. The model improves the detection of tiny lesions by utilizing both geographical and temporal information. The model obtained a detection rate of 0.90, a sensitivity of 0.95, and a benign lesion identification rate of 0.81 in a dataset of 572 lesions from ultrafast MRI scans [16]. The RetinaNet model was also used in another study to detect cancer in mammography images. With an AUC of 0.93 in the identification of benign and malignant masses, this model showed exceptional performance across a number of datasets, standing out in particular for its ability to recognize masses, architectural deformation, and microcalcifications—all signs of breast cancer [17]. A deep learning (DL) model known as You Only Look Once (YOLO) was introduced by Joseph Redmon et al. in 2016 [18] for real-time object detection. Unlike traditional object recognition techniques, YOLO transforms the detection process into a single regression task, enabling efficient end-to-end object detection. Al-Masni et al. were the first to propose a Computer-Aided Diagnosis (CAD) system utilizing regional DL strategies, incorporating a convolutional neural network (CNN) called YOLO, specifically for breast cancer diagnosis. The methodology consists of four key phases: mammogram pre-processing, feature extraction through deep convolutional networks, confident detection of masses, and mass classification using fully connected neural networks (FC-NNs). To train and evaluate the system, they utilized an additional 2,400 mammograms alongside the original 600 mammograms from the Digital Database for Screening Mammography (DDSM). The system achieved a mass localization accuracy of 99.7% and differentiated between benign and malignant lesions with an accuracy of 97%, based on the findings from fivefold cross-validation tests [19]. Su et al. introduced a dual model for efficient mass detection and segmentation simultaneously, combining YOLO with a LOGO (Local-Global) architecture. When evaluated on two different mammography datasets (INBreast and CBIS-DDSM), this model outperformed previous studies significantly. The mass segmentation F1-score was 74.5%, the intersection over union (IoU) reached 64.0%, and the true positive rate for mass detection on the CBIS-DDSM dataset was 95.7%, with an average precision of 65.0%. [20]

III. ROLE OF DEEP LEARNING IN TILS SCORE ASSESSMENT AND LYMPH NODE METASTASIS IDENTIFICATION

The Faster R-CNN object detector is employed for the automatic identification and counting of lymphocytes in immunohistochemistry images of breast cancer, which is crucial for determining the Immunoscore. The tested feature extractors—ResNet-50, VGG-16, and ResNet-101—showed that ResNet-101 delivered the highest detection accuracy, while VGG-16 offered the quickest detection speed. The study using ResNet-101 achieved 98% accuracy in classifying bounding boxes and outperformed earlier models in terms of efficiency. It excelled in all performance metrics, whereas VGG-16 demonstrated faster detection times [21].

The use of immune-checkpoint inhibitors (ICIs) has greatly enhanced patient outcomes in breast cancer (BC), though optimizing patient selection based on tumor biomarkers continues to be a challenge.

Incorporating tumor-infiltrating lymphocytes (TILs) into breast cancer management enhances prognostic evaluations and treatment decisions, particularly for triple-negative and HER2-positive subtypes [22]. A specialized Convolutional Neural Network (CNN) known as BCF-Lym-Detector was developed to tackle the challenge of identifying tumor-infiltrating lymphocytes (TILs) in cancer histology images, which is crucial for evaluating cancer prognosis and treatment strategies. The varied morphology of TILs makes their detection more difficult. [23]. The objective is to develop a two-step deep learning approach for TILs detection. First, a dataset is established by labeling and annotating TILs in the stroma on 63 pathology slides following IIOBWG criteria. Next, various machine learning models were trained to identify the stroma, achieving 98% accuracy with U-Net. Subsequently, TILs were identified using a Mask R-CNN model [24]

This study introduces a two-step pipeline for the detection of lymphocytes and Ki-67 on histopathology slides utilizing commonly employed deep learning models [25]. This paper suggests a deep learning-based method that leverages ResNet architectures and a 1cycle learning rate policy to effectively identify cancer metastases in lymph node images with high precision. The proposed model enhances depth while preserving accuracy. It demonstrates quicker convergence and is easy to optimize. "Shortcut connections" in the ResNet architecture facilitate layer or time skips between layers [26]. With rapid advancements in digital histopathology, convolutional neural network (CNN) modeling has been successfully utilized for the detection and classification of images related to skin cancer, breast cancer, and other tumors. It was shown that the accuracy of lymph node quantification was 97.13%. The Xception and DenseNet-121 models combined exhibit a negative predictive value (NPV) of 97.99% and a positive predictive value (PPV) of 93.53% [27].An end-to-end deep learning network known as TabNet was introduced in this study to predict sentinel lymph node (SLN) involvement in patients with breast cancer. The TabNet encoder comprises several stages where batch normalization is applied to the raw features. In comparison to the logistic regression model, which achieved an accuracy of 70%, precision of 73%, specificity of 65%, and sensitivity of 79% on the dataset, the TabNet model averaged 75% accuracy and an area under the curve (AUC) of 0.74 on the dataset [28].

IV. PROPOSED METHODOLOGY

The work flow of the proposed system is as shown in the figure1.The process begins with the collection of datasets for identifying TILs score and lymph node metastasis in breast Histopathological images. Data is collected from TIGER and Patch Cameleon Dataset which is particularly meant Dataset for finding out TILs score and Lymph node metastasis in Breast cancer Histopathological images.

Next step is pre-processing of the data to make it suitable for feature extraction. The feature which are relevant to the topic are extracted. The next step is biomarker Quantification. Followed by that step classification and prediction is done. Analysis of the classification result is done followed by interpretation. For analysis accuracy, Sensitivity, Specificity, F1score is considered as evaluation parameters. In order to analyse the Breast cancer TILs and Lymph node metastasis plays a very important role as Biomarkers. With the continuous development of AI and deep DL technologies, the future application of these technologies in breast cancer diagnosis and treatment is promising. First, AI and DL are expected to significantly improve the diagnostic performance of breast cancer and its metastasis. Histopathological characteristics are microscopic findings observed after analyzing tissue samples (biopsies) that have been stained to enhance cellular structures. These features are crucial for assessing abnormalities in cells and tissues. Important aspects include cellular morphology, such as the shape, size,

and appearance of individual cells (e.g., pleomorphism, nuclear atypia), and tissue architecture, which refers to the organization or disorganization of tissue layers (e.g., glandular structures in adenocarcinomas). The presence of mitotic figures, representing dividing cells, and necrosis, signifying dead or degenerating tissue, are vital in determining disease progression. Additionally, the inflammatory response, including immune cell infiltration, abnormal growth patterns like dysplasia or hyperplasia, and specific staining techniques (e.g., hematoxylin and eosin, immunohistochemistry), help detect particular proteins, cells, or pathogens for diagnostic and therapeutic purposes.

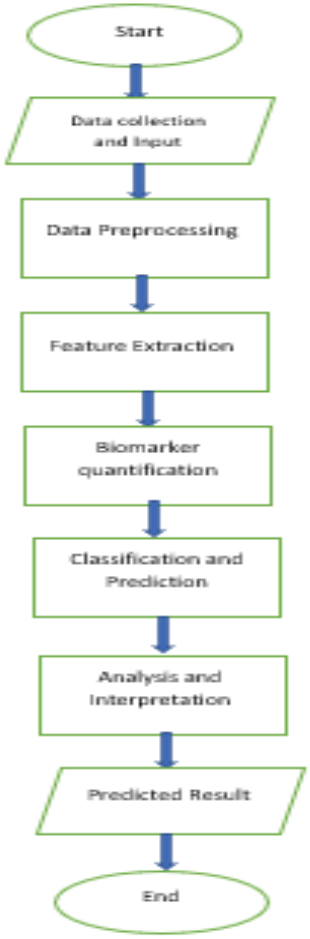

Figure1: Work flow of the Proposed System

VI REFERENCES

1. Karan Gupta and Nidhi Chawla,"Analysis of Histopathological Images for Prediction of Breast Cancer Using Traditional Classifiers with Pre-Trained CNN",Procedia Computer Science,Volume 167,2020,Pages 878-889,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.03.427.
2. Priya C V L, V G B, B R V and Ramachandran S. Deep learning approaches for breast cancer detection in histopathology images: A review. Cancer Biomark. 2024;40(1):1-25. doi: 10.3233/CBM-230251. PMID: 38517775; PMCID: PMC11191493.
3. Balasubramanian AA, Al-Heejawi SMA, Singh A, Breggia A, Ahmad B, Christman R, Ryan ST and Amal S. Ensemble Deep Learning-Based Image Classification for Breast Cancer Subtype and Invasiveness Diagnosis from Whole Slide Image Histopathology. Cancers (Basel). 2024 Jun 14;16(12):2222. doi: 10.3390/cancers16122222. PMID: 38927927; PMCID: PMC11201924.
4. Faisal Bin Ashraf, S.M. Maksudul Alam, Shahriar M. Sakib,Enhancing breast cancer classification via

histopathological image analysis: Leveraging self-supervised contrastive learning and transfer learning,Heliyon,Volume10,Issue2,2024,e24094,ISSN2405-8440,https://doi.org/10.1016/j.heliyon.2024.e24094.

5. Zhao Y, Zhang J, Hu D, Qu H, Tian Y, Cui X. Application of Deep Learning in Histopathology Images of Breast Cancer: A Review. Micromachines (Basel). 2022 Dec 11;13(12):2197. doi: 10.3390/mi13122197. PMID: 36557496; PMCID: PMC9781697.
6. S. Loi, S. Michiels, S. Adams, S. Loibl, J. Budczies, C. Denkert, R. Salgado,The journey of tumor-infiltrating lymphocytes as a biomarker in breast cancer: clinical utility in an era of checkpoint inhibition,Annals of Oncology,Volume 32, Issue 10,2021,Pages 1236-1244,ISSN 0923-7534,https://doi.org/10.1016/j.annonc.2021.07.007.
7. Fiorin, A., López Pablo, C., Lejeune, M. *et al.* Enhancing AI Research for Breast Cancer: A Comprehensive Review of Tumor-Infiltrating Lymphocyte Datasets. *J Digit Imaging. Inform. med.* (2024). https://doi.org/10.1007/s10278-024-01043-8
8. Ertan Bütün, Murat Uçan, Mehmet Kaya,Automatic detection of cancer metastasis in lymph node using deep learning,Biomedical Signal Processing and Control,Volume 82,2023,104564,ISSN1746 8094,https://doi.org/10.1016/j.bspc.2022.104564.
9. Duggento A, Conti A, Mauriello A, et al. Deep computational pathology in breast cancer. Semin Cancer Biol. 2021;72:226–37.
10. Cruz-Roa A, Gilmore H, Basavanhally A, et al. High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: application to invasive breast cancer detection. PLoS ONE. 2018;13:e0196828.
11. Han Z, Wei B, Zheng Y, et al. Breast Cancer multi-classification from histopathological images with structured Deep Learning Model. Sci Rep. 2017;7:4172.
12. Ren S, He K, Girshick R, Sun, JJAinips. Faster r-cnn: towards real-time object detection with region proposal networks. 2015; 28.
13. Yap MH, Goyal M, Osman F, et al. Breast ultrasound region of interest detection and lesion localisation. Artif Intell Med. 2020;107:101880.
14. Agarwal R, Díaz O, Yap MH, et al. Deep learning for mass detection in full field Digital mammograms. Comput Biol Med. 2020;121:103774.
15. Lin T-Y, Goyal P, Girshick R et al. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision. 2017; 2980–2988.
16. Ayatollahi F, Shokouhi SB, Mann RM, Teuwen J. Automatic breast lesion detection in ultrafast DCE-MRI using deep learning. Med Phys. 2021;48:5897–907.
17. Ueda D, Yamamoto A, Onoda N, et al. Development and validation of a deep learning model for detection of breast cancers in mammography from multiinstitutional datasets. PLoS ONE. 2022;17:e0265751.
18. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, realtime object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; 779–788.
19. Al-Masni MA, Al-Antari MA, Park JM, et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLObased CAD system. Comput Methods Programs Biomed. 2018;157:85–94.

20. Su Y, Liu Q, Xie W, Hu P. YOLO-LOGO: a transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms. Comput Methods Programs Biomed. 2022;221:106903
21. Keren Evangeline, I., Glory Precious, J., Pazhanivel, N. et al. Automatic Detection and Counting of Lymphocytes from Immunohistochemistry Cancer Images Using Deep Learning. J. Med. Biol. Eng. 40, 735–747 (2020). https://doi.org/10.1007/s40846-020-00545-4
22. El Bairi K, Haynes HR, et al,; International Immuno-Oncology Biomarker Working Group. The tale of TILs in breast cancer: A report from The International Immuno-Oncology Biomarker Working Group. NPJ Breast Cancer. 2021 Dec 1;7(1):150. doi: 10.1038/s41523-021-00346-1. PMID: 34853355; PMCID: PMC8636568.
23. Rauf Z, Khan AR, Sohail A, Alquhayz H, Gwak J, Khan A. Lymphocyte detection for cancer analysis using a novel fusion block based channel boosted CNN. Sci Rep. 2023 Aug 28;13(1):14047. doi: 10.1038/s41598-023-40581-z. PMID: 37640739; PMCID: PMC10462751.
24. Yosofvand M, Khan SY, Dhakal R, Nejat A, Moustaid-Moussa N, Rahman RL, Moussa H. Automated Detection and Scoring of Tumor-Infiltrating Lymphocytes in Breast Cancer Histopathology Slides. Cancers (Basel). 2023 Jul 15;15(14):3635. doi: 10.3390/cancers15143635. PMID: 37509295; PMCID: PMC10377197.
25. Karol, M., Tabakov, M., Markowska-Kaczmar, U. et al. Deep learning for cancer cell detection: do we need dedicated models?. Artif Intell Rev 57, 53 (2024). https://doi.org/10.1007/s10462-023-10699-7
26. Ertan Bütün, Murat Uçan, Mehmet Kaya,Automatic detection of cancer metastasis in lymph node using deep learning,Biomedical Signal Processing and Control,Volume 82,2023,104564,ISSN1746-8094,https://doi.org/10.1016/j.bspc.2022.104564.
27. Hu Y, Su F, Dong K, Wang X, Zhao X, Jiang Y, Li J, Ji J, Sun Y. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. Gastric Cancer. 2021 Jul;24(4):868-877. doi: 10.1007/s10120-021-01158-9. Epub 2021 Jan 23. PMID: 33484355.
28. Shahriarirad R, Meshkati Yazd SM, Fathian R, Fallahi M, Ghadiani Z, Nafissi N. Prediction of sentinel lymph node metastasis in breast cancer patients based on preoperative features: a deep machine learning approach. Sci Rep. 2024 Jan 16;14(1):1351. doi: 10.1038/s41598-024-51244-y. PMID: 38228684; PMCID: PMC10791698.