International Journal of Advanced and Applied Sciences

Journal homepage: http://www.ijaas.in

Deep Learning for Predicting Lymph Node Metastasis in Breast Cancer: A Multimodal Imaging and Clinical Data Approach

Dayakar Kondamudi [1*], Dr. Vijay Raddy Madireddy [2], Dr. Polaiah Bojja [3]

1 *Research Scholar, Department of Computer Science Engineering, GIET University, Gunupur, India*
2 *Professor, Department of Computer Science Engineering, GIET University, Gunupur, India*
3 *Professor, Department of Computer Science Engineering, IARE, Hyderabad, India*

ARTICLE INFO

Article history:
Received: 30-08-2025
Received in revised form: 11-09-2025
Accepted: 03-10-2025

Keywords:
Deep learning, Breast cancer, Lymph node metastasis, Multimodal imaging, Clinical data, Medical AI, CNN, Predictive modeling

ABSTRACT

One important prognostic marker in breast cancer that affects treatment choices and patient outcomes is lymph node metastasis (LNM). Promising opportunities for automating and improving LNM prediction using medical data are presented by recent developments in deep learning. In order to predict lymph node metastases in patients with breast cancer, this research suggests a multimodal method that combines structured clinical data with imaging modalities including mammography, ultrasound, and magnetic resonance imaging (MRI). We create and assess a deep learning framework that combines fully connected networks processing clinical information with convolutional neural networks (CNNs) for picture analysis. Our model outperforms conventional machine learning baselines in accuracy, sensitivity, and area under the ROC curve (AUC), demonstrating superior predictive performance. The multimodal approach shows how deep learning may help with individualized diagnosis and treatment planning by accurately and non-invasively predicting LNM status.

Introduction

As the second leading cause of mortality in developed nations, breast cancer is one of the most prevalent cancers to strike women. The World Health Organization (WHO) conducted a study in 2018 and found that the illness kills almost 2.1 million women annually, increasing the death rate. Approximately 627,000 people have lost their lives to breast cancer. Globally, and particularly in emerging nations, these rates are rising among the middle and low-income groups [1]. Because the illness is only diagnosed at an advanced stage, the mortality rate is significant (Majid et al., 2014). In 2023, around 1,15,251 new cases of breast cancer were reported in India; by 2030, that number might rise to 2,00,000 year (Nahid& Kong 2017). The most important early diagnosis of the condition is brought on by this rise in the death rate.

The biological processes that alter the normal characteristics of cells and disrupt the regulatory systems that control cell division and invasion are the root cause of cancer. Cells that have received mutations proliferate out of control, changing form, adhesion, and producing new enzymes. A thorough understanding of the different types of breast diseases is necessary for prompt detection, precise diagnosis, and effective treatment strategies, all of which improve patient outcomes and raise survival rates (Krawczyk et al., 2016; Bezerra et al., 2013; Mughal et al., 2017; Nicandro et al., 2013).

Effective treatment, a speedy recovery, economic savings, less mental stress, and a decreased death rate are all benefits of early discovery of breast cancer. Even before the patient is aware of the symptoms, routine screening may help with early detection [2]. It should be underlined how crucial screening is to the disease's early detection. In order to help with the early detection of the condition, several diagnostic instruments have been created for screening using different current technologies.

The glandular tissue, also known as lobules, is where milk is produced; fatty tissue regulates the size of the breast; and connective and fibrous tissues maintain the glandular and fatty tissues in place. Medicine J.H. (2020) describes the connecting channel that connects the lobules to the nipple and transports the milk. The areola region is the dark area that surrounds the nipple. Breast cancer starts in the lobe or lobules and travels via the lymphatic and blood vessels to other regions of the body. A component of the immune system, the lymphatic system transports fluid that fights illness to every region of the body, facilitating the simple spread of cancer to other bodily organs.

Like other forms of cancer, breast cancer develops when cells proliferate excessively and uncontrollably, leading to the development of tumors or neoplasms. Tumors may be categorized as malignant if they spread via the blood or lymphatic system and reach neighboring tissues, or as benign if the aberrant cells are localized and do not spread to the surrounding tissues. The capacity of malignant tumors to metastasize signifies that the cancer cells have moved to other parts of the body, which may result in the formation of additional tumors [3].

Cancer is defined as aberrant cell proliferation that spreads to other bodily areas. It results from a series of actions that essentially alter the normal properties of cells. Cancerous cells have impaired regulatory systems that prevent excessive cellular proliferation and tissue invasion. Mutations in genes that code for proteins that regulate cell division are often the cause of these illnesses. The defective genes that repair the damaged deoxyribonucleic acid cause gene mutations to progressively accumulate.

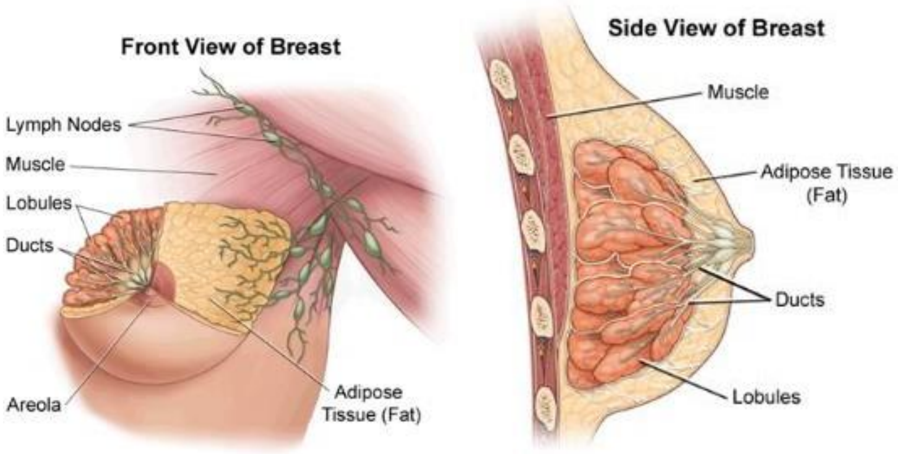

Figure 1: Female Breast Anatomy

The fast and unchecked growth of cells is known as cancer, and it poses a serious and even lethal risk. A person's age, diet, and work environment may all contribute to the development of many malignancies, including breast, lung, brain, and cervical cancer. Since symptoms may not appear until the cancer has progressed, early identification is essential for survival. This illness highlights the need of prompt detection since it is often seen as a silent cell killer in women. A dynamic method for identifying and classifying cancerous pictures is offered by medical image processing. The use of mammography, WSI pictures, magnetic resonance imaging (MRI), and ultrasound may all be used to identify breast cancer.By analyzing WSI, digital pathology datasets like as Camelyon16 and Camelyon17 provide a wealth of resources for improving breast cancer research and diagnosis. High-resolution WSI pictures of breast tissue samples are included in these datasets along with thorough annotations that include the locations of tumors and other histological characteristics.

By using machine learning and deep learning methods, the Camelyon challenges seek to promote the development of computer-aided detection and diagnostic systems for breast cancer. In order to increase the precision and effectiveness of breast cancer diagnosis, these datasets are utilized to create classification algorithms for a variety of tasks, including tumor identification, segmentation, and classification. Additionally, the availability of these datasets makes it easier for researchers in the field of digital pathology to benchmark their algorithms [4].

According to the WHO, early detection is essential for improving breast cancer outcomes and survival. A noticeable lump, discomfort unrelated to menstruation, skin changes, nipple alterations, swelling, armpit soreness, and nipple discharge are among the notable signs [20]. By quickly detecting the problems, routine screenings and improved understanding significantly improve treatment results, lowering risks and perhaps lowering death rates by 25%.

One harmful condition that results from the body's abnormal cellular activity is cancer. Cells normally divide and proliferate according to the organism's needs. However, this methodical process is disrupted when cells that are unnecessary for the body's needs are created in excess or when aged cells fail to undergo apoptosis in a timely way. These extra cells group together to create a tumor, which is an abnormal mass Based on their size, shape, and volume of impacted pixels, breast tumors are classified as either benign or malignant.

Benign: Benign tumors often pose no immediate threat and are not cancerous. One practical way to get rid of them and essentially halt their development is by surgical excision. The cells in benign tumors do not show signs of spread, in contrast to malignant tumors, and there is very little chance that a benign tumor might be fatal.

Malignant: If treatment is not received, malignant tumors, which are often malignant, may be very dangerous. Uncontrolled proliferation is one of the abnormal behaviors shown by the cells that make up malignant tumors. These cells aggressively attack the nearby healthy tissues, demonstrating their very aggressive behavior [19]. Their capacity to penetrate the circulation may lead to the formation of new tumors in other body areas.

There are also some common misconceptions about the etiology of breast cancer. For instance, using deodorants, owning a cell phone, or being around microwaves or coffee do not promote breast cancer. Furthermore, interacting with someone who has breast cancer won't have an impact on the other individual. Each individual has a different explanation. Here are a handful of the causes or risk factors. Given that breast cancer is inherited, one should exercise extreme caution if any blood relatives have had breast cancer [5]. High-risk factors for developing breast cancer include drinking excessive amounts of alcohol and reaching puberty earlier than the recommended age of twelve. A few other factors that contribute to illness include smoking, obesity, excessive breast density, hypertension, and pre-existing heart conditions. During menopause, some women have an increased risk. On average, women experience menopause around age 51 women who experience menopause later in life, such as beyond the age of 55, may be at somewhat increased risk. Breast cancer may also result from not drinking enough milk or from breastfeeding for a shorter period of time after giving birth.

Literature review

Tang X, Zhang H, Mao R, Zhang Y, Jiang X, Lin M, et al. (2025) [1] This study focuses on the development and evaluation of a multimodal deep learning model designed to predict axillary lymph node metastasis in breast cancer patients. The model uses both ultrasound and MRI images to extract features and predict the presence of lymph node involvement, which is crucial for determining the treatment plan. By leveraging deep learning techniques, the study improves the accuracy of preoperative assessments, offering a promising alternative to invasive procedures like biopsies. The authors discuss the synergy between ultrasound and MRI in providing comprehensive image features for enhanced prediction performance.

Rane NL, Mallick SK, Kaya O, Rane J. (2024) [2] This review paper provides an extensive overview of machine learning

(ML) and deep learning (DL) architectures, focusing on their evolution, current trends, and potential future directions. The authors highlight the importance of various DL frameworks, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers, and their applications across diverse fields. The review also explores challenges such as over fitting, computational complexity, and the need for large datasets in training robust models. The authors emphasize the growing integration of ML and DL techniques in solving complex problems in healthcare, finance, and robotics, among other domains.

Shafik W. (2024) [3] This book chapter delves into the transformative role of deep learning in advancing artificial intelligence (AI), with a particular focus on its applications in operations research. Shafik explores how deep learning models, especially deep neural networks, have revolutionized problem-solving in logistics, optimization, and decision-making processes. The chapter addresses how deep learning techniques are being used to enhance traditional operations research methods, offering improved accuracy and efficiency in areas such as supply chain management, scheduling, and resource allocation. The author also discusses the challenges of model interpretability and the need for explainable AI in operations research settings.

Zhou J, Yu X, Wu Q, Wu Y, Fu C, Wang Y, et al. (2024) [4] In this study, the authors examine the role of radio mics in evaluating the HER2 status of breast cancer using multi parametric MRI. By analyzing both intratumoral and peritumoral regions, they aim to uncover the subtle variations in imaging features that can help predict HER2 receptor status, which is essential for personalized treatment strategies. The study compares the efficacy of using radio mic features extracted from different MRI sequences, such as T1-weighted, T2-weighted, and diffusion-weighted images, to assess their potential in guiding clinical decisions. The results suggest that multipara metric MRI, combined with radio mics, can provide more accurate and reliable predictions of HER2 status compared to traditional imaging approaches.

Xie X, Zhou H, Ma M, Nie J, Gao W, Zhong J, et al. (2024) [5] This study introduces a deep learning model aimed at predicting the molecular subtypes of breast cancer using dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) data. The authors fuse multiple MRI sequences from two different institutes to increase the robustness and generalizability of the model. The research emphasizes how multimodal data integration can improve the accuracy of molecular subtype classification, which is crucial for personalized treatment planning. The model demonstrates a high level of prediction accuracy for different molecular subtypes, providing valuable insight into the potential of deep learning models in non-invasive cancer characterization.

Qu L, Mei X, Yi Z, Zou Q, Zhou Q, Zhang D, et al. (2024) [6] This study explores an unsupervised learning model utilizing CT radiomics features to predict axillary lymph node metastasis (ALNM) in breast cancer patients. By analyzing radiomic data extracted from CT images, the model can identify patterns linked to metastatic involvement of lymph nodes, which is crucial for treatment decisions. The authors show that the unsupervised approach can effectively predict ALNM, even in cases where clinical or histo pathological data may be limited. The study represents a significant step forward in integrating radiomics and unsupervised learning for more accurate preoperative assessments of breast cancer metastasis.

Li H, Hou Y, Xue L, Fan W, Gao B, Yin X. (2024) [7] This article investigates the potential of MRI-based radiomics models in evaluating low HER2 expression in breast cancer, which is a critical factor for determining appropriate treatment options. The authors emphasize that accurately assessing HER2 status, especially at low expression levels, can be challenging with traditional imaging and histopathology. By extracting quantitative features from MRI scans, the radiomics model presented in the study enhances the ability to differentiate between low and non-low HER2 expression, offering a non-invasive and reproducible method for clinical use. The study paves the way for improving diagnostic accuracy in HER2-related breast cancer treatment strategies.

Carriero A, Groenhoff L, Vologina E, Basile P, Albera M. (2024) [8] This review article provides an in-depth look at the state of the art in deep learning applications for breast cancer imaging as of early 2024. The authors discuss the evolution of deep learning models, particularly convolutional neural networks (CNNs), in improving the detection and diagnosis of breast cancer from various imaging modalities such as mammography, ultrasound, and MRI. They highlight recent advancements, including automated lesion detection, classification, and risk stratification, which have the potential to assist clinicians in making more accurate diagnoses. The review also touches on the challenges associated with deep learning in clinical practice, such as data heterogeneity, interpretability, and regulatory hurdles.

Hua Y, Peng Q, Han J, Fei J, Sun A. (2024) [9] This study presents a two-center investigation into the development of a combined nomogram that integrates mammography and MRI data to predict axillary lymph node (ALN) metastasis in breast cancer patients. By combining quantitative imaging features with clinical data, the nomogram improves prediction accuracy for ALN metastasis compared to using either imaging modality alone. The study underscores the utility of multimodal imaging in enhancing diagnostic precision, especially in cases where metastasis risk is uncertain. The results show that the combined approach offers a significant advantage in preoperative staging, helping guide treatment decisions.

Ma X, Zhang L, Xiao Q, Huang Y, Lin L, Peng W, et al. (2024) [10] In this preliminary study, the authors propose a radiomics model that combines features extracted from mammography and MRI images to predict the prognosis of phyllodes tumors, which are rare breast tumors. The study emphasizes the importance of non-invasive imaging techniques in assessing tumor characteristics and predicting outcomes. By extracting radiomic features from both mammography and MRI scans, the authors create a prognostic model that helps identify high-risk patients, thus aiding in clinical decision-making. This approach has the potential to enhance personalized management strategies for phyllodes tumors, reducing the need for more invasive procedures.

Rokhshad R, Salehi SN, Yavari A, Shobeiri P, Esmaeili M, Manila N, et al. (2024) [11] This systematic review and meta-analysis assesses the performance of deep learning models in diagnosing head and neck cancers using radiographic data, such as CT, MRI, and X-ray images. The authors analyzed multiple studies that utilized deep learning algorithms like convolutional neural networks (CNNs) to detect malignancies in these regions. The meta-analysis reveals that deep learning models outperform traditional methods, with high sensitivity and specificity. The review highlights the potential of deep learning as a reliable tool for early detection and diagnosis, improving clinical outcomes in head and neck oncology. Challenges discussed include data quality, model generalization, and the need for standardized protocols in clinical settings.

Polat DS, Nguyen S, Karbasi P, Hulsey K, Cobanoglu MC, Wang L, et al. (2024) [12] This study presents a multi-institutional evaluation of a 4D convolutional neural network (CNN) model designed to predict lymph node metastasis in breast cancer using MRI scans. By incorporating temporal information from dynamic contrast-enhanced (DCE) MRI, the 4D CNN model improves the accuracy of metastasis prediction over traditional 2D models. The study emphasizes the importance of combining multiple imaging time points, which capture more comprehensive biological information about the lymph nodes. The model demonstrates robust performance across various institutions, highlighting its potential for clinical implementation in personalized breast cancer care.

Shiner A, Kiss A, Saednia K, Jerzak KJ, Gandhi S, Lu FI, et al. (2023) [13] This study explores machine learning models for predicting patterns of distant metastasis in breast cancer patients after local regional therapy (such as surgery or radiation). The authors utilize clinical, pathological, and imaging data to train predictive models that can assess the risk of metastasis. The findings show that machine learning models, particularly random forests and support vector machines, can identify high-risk patients who may benefit from additional systemic therapies. The study also suggests that personalized treatment plans, guided by predictive models, could improve outcomes and reduce unnecessary treatments.

Guo YJ, Yin R, Zhang Q, Han JQ, Dou ZX, Wang PB, Lu H, Liu PF, Chen JJ, Ma WJ. (2024) [14] This paper introduces a hybrid deep learning model that combines convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to evaluate axillary lymph node status in breast cancer using dynamic contrast-enhanced MRI (DCE-MRI). The model analyzes kinetic heterogeneity in the imaging data, which refers to the varying patterns of contrast enhancement that can be indicative of metastasis. The authors demonstrate that the CNN-RNN hybrid model improves the accuracy of axillary lymph node status assessment, providing a non-invasive alternative to sentinel lymph node biopsy. The study suggests the potential for using this model in clinical decision-making to guide treatment plans for breast cancer patients.

Zhou H, Hua Z, Gao J, Lin F, Chen Y, Zhang S, et al. (2024) [15] This multicenter study develops a multitask deep learning model that simultaneously diagnoses breast lesions and discriminates axillary lymph node metastasis from dynamic contrast-enhanced MRI (DCE-MRI) data. The model integrates information from the entire diagnostic process, from lesion detection to lymph node assessment, and performs both tasks using a single deep learning framework. The results suggest that the system performs effectively across different centers, showing potential for real-world clinical use. By automating multiple stages of the diagnostic workflow, this approach could reduce clinician workload and improve diagnostic accuracy, leading to better outcomes in breast cancer management.

Guni A, Sounderajah V, Whiting P, Bossuyt P, Darzi A, Ashrafian H. (2024) [16] This paper introduces a revised version of the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool, specifically designed to assess the quality of diagnostic accuracy studies involving artificial intelligence (AI). The QUADAS-AI tool aims to provide a comprehensive framework for evaluating the methodological rigor, reporting quality, and clinical applicability of AI-based diagnostic studies. The authors present the protocol for a qualitative study that will test and refine the tool, ensuring its validity and reliability. This tool will help improve the quality of AI-based diagnostic research and promote the adoption of AI in clinical practice.

Gradishar W. J., et al. (2024) [17] The NCCN Clinical Practice Guidelines in Oncology for Breast Cancer, version

3.2024, provide evidence-based recommendations for the management of breast cancer across various stages and subtypes. This version updates guidelines on screening, diagnosis, treatment strategies (including neo adjuvant and adjuvant therapies), and surveillance. New sections on molecular subtyping and precision medicine are also included, reflecting the growing importance of personalized treatment approaches. The guidelines offer comprehensive, up-to-date clinical insights, helping oncologists provide optimal care for breast cancer patients.

Shahriarirad R, Meshkati Yazd SM, Fathian R, Fallahi M, Ghadiani Z, Nafissi N. (2024) [18] This study proposes a deep machine learning model to predict sentinel lymph node (SLN) metastasis in breast cancer patients using preoperative features. The authors train a model on clinical, pathological, and imaging data to predict the likelihood of metastasis in the sentinel lymph node, which is critical for staging and treatment planning. The model outperforms traditional methods and provides a more accurate, non-invasive approach to assessing lymph node status, potentially reducing the need for invasive biopsy procedures and improving patient outcomes.

Polat DS, Nguyen S, Karbasi P, Hulsey K, Cobanoglu MC, Wang L, et al. (2024) [19] This is a repeat of entry #12, discussing the same multi-institutional study evaluating a 4D CNN model for predicting lymph node metastasis in breast cancer patients using MRI data. The key highlights remain consistent with the study's robust performance across various institutions, showcasing the potential of multi-institutional collaboration in validating AI-based diagnostic models.

Garcia-Tejedor A, Ortega-Exposito C, Salinas S, Luzardo-González A, Falo C, Martinez-Pérez E, et al. (2023) [20] This study investigates the efficacy of axillary lymph node dissection (ALND) compared to radiotherapy in patients with positive sentinel lymph nodes following neoadjuvant therapy for breast cancer. The ADARNAT trial presents evidence that radiotherapy can be a viable alternative to ALND, reducing surgical morbidity without compromising survival outcomes. The findings have significant clinical implications, suggesting that radiotherapy could become the standard of care for certain breast cancer patients, avoiding the need for invasive surgical procedures.

Research methodology

WSI files, which provide a thorough picture of tissue samples at high resolution, are the foundation of histo pathological image analysis for the identification of breast cancer. Advanced image processing techniques are used for ROI detection in order to identify important histo pathological traits, such as malignant spots. This feature is important because it allows pathologists to quickly go through enormous volumes of tissue data and concentrate on the parts that are clinically relevant.

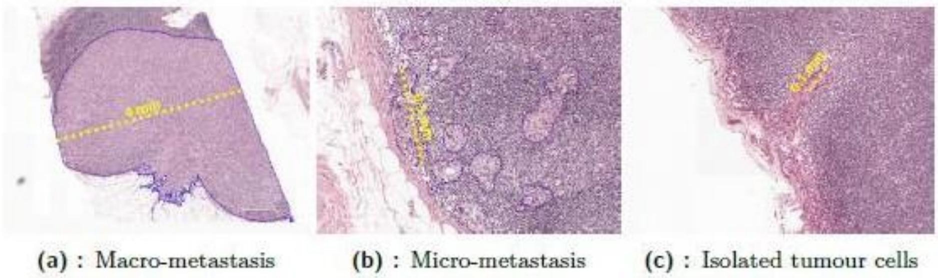

Figure 2: Illustration of Various Metastasis Types

By dividing the ROI into smaller parts, deep learning techniques for tile-based classification may be used more easily, expediting the analysis process and increasing computational effectiveness. The focus on methods for increasing the quality of data, such as the creative fusion of opening and shutting processes with the Otsu algorithm, highlights the dedication to enhancing the accuracy and dependability of cancer diagnosis. By reducing typical problems like false positives and patches lacking cells, this method improves the dataset's general integrity, which in turn increases the deep learning model's resilience [6].

In order to ensure consistency and minimize variability brought about by staining the discrepancies or imaging aberrations, normalization methods are essential for normalizing pixel intensities and color distributions across pictures [18]. Overall, the effectiveness and dependability of deep learning-based breast cancer detection are greatly impacted by the careful preparation of

WSI datasets, which eventually improves patient outcomes and diagnostic accuracy.

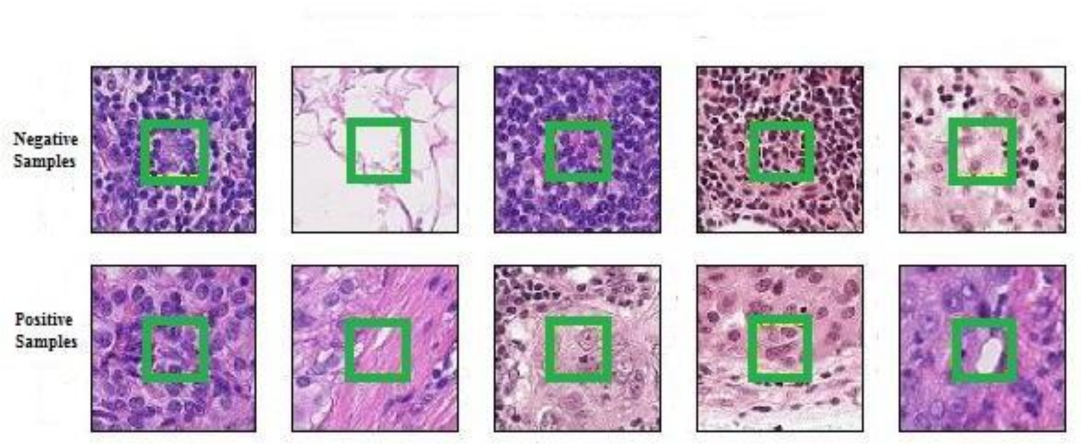

Figure 3: Patch came lyon Dataset's Normal and Benign Images

The growing relevance of personalized medicine and the need for accurate and effective diagnostic techniques in light of pathologists' increasing workload and complexity highlight the necessity of the aforementioned preprocessing stages. Only below the human vision threshold can a WSI picture distinguish between malignant and healthy thick tissue. In a similar vein, micro calcifications may not be visible at sufficiently thick concentrations. Determining the properties of micro calcifications is thus difficult. On WSI pictures, the traditional image processing methods don't work well. Because features fluctuate in size and form, classic fixed neighborhood approaches like un sharp masking are less effective. Although fixed or global approaches could address local characteristics inside a neighborhood, they don't alter the neighborhood's size to take local problems into consideration [7].

Table 1: Camelyon16 Dataset's WSI Image Count

	Total WSIs		Metastases		
Centre	**Train**	**Test**	**None**	**Micro**	**Macro**
RUMC	267	52	553	43	24

UMCU	467	54	25	24	56
Total	757	646	543	53	83

Histogram equalization and adaptive equalization are two methods used to modify intensity distributions and increase local contrast, highlighting the significance of contrast enhancement in pre-processing. To reduce noise and provide a clearer visual foundation, Gaussian smoothing is essential.

In contrast to a non-equalized histogram, histogram equalization enhances the WSI's contrast. This is accomplished by efficiently separating the most prevalent intensity values within the image's intensity spectrum. This technique often enhances the overall contrast of pictures when the relevant information is supplied as the near-contrast values. In computer image processing, adaptive histogram equalization, or AHE, is used to enhance visual contrast. It works effectively in scenarios like particular picture portions when edge definition and local contrast are crucial. The algorithm's main objective is to minimize noise while maintaining high picture quality.

Dataset for Came lyon 16

The dataset includes 400 sentinel lymph node WSIs that were acquired from the University Medical Center Utrecht and Rad boud University Medical Center in the Netherlands. A total of 170 WSIs 100 normal and 70 with metastases make up the training dataset. Another 100 WSIs 60 normal and 40 with metastases are included in the second training dataset. XML files with WSI binary masks and annotated contours provide ground truth data for slides with metastases [17]. 130 WSIs gathered from both institutions are included in the test dataset. Open Slide is suggested for picture access, and the images are stored in a multi-resolution pyramid structure with tiles for effective sub region retrieval [8].

Using programs like Open Slide, Qt, and Open CV, ASAP is recommended for visualizing slides and annotations. Furthermore, the quantity of pictures considered for the suggested project.

Table 2: Camelyon16 Patch Image Distribution by Various Categories

	Total WSIs			Metastases		
Centre	**Training**	**Validation**	**Testing**	**None**	**Micro**	**Macro**
RUMC	20000	35200	43600	36000	43000	5300
UMCU	10000	7500	5300	35000	53000	3200
Total	60000	42000	53000	78000	35000	23000

Dataset for Camelyon17

The 1,000 WSIs from five Dutch medical institutions make up the CAMELYON17 dataset, which is split into training and testing sets with 20 patients from each center. Additionally, 10 slides from each center have annotations provided in an XML format that is compatible with ASAP; these annotations include regions that are non-tumor ("normal") and tumor ("metastases"). The distributions of metastatic type and the number of WSIs provided for testing and training (Litjens et al. 2018). how many patch photos are taken into account for the suggested task. Based on the size of the afflicted area, the tumor is divided into four types: no tumor, isolated tumor cell, micro metastasis, and macro metastasis (Litjens et al. 2018).

Dataset for Patch came lyon

The 3,27,680 96x96 pixel image patches in the Patch came lyon (PCAM) dataset were taken from histo pathological lymph node WSIs (Veeling et al. 2018). If there is at least one impacted pixel in the patch, it is labeled positive; if not, it is labeled negative. The specimen It is extracted such that the majority of the impacted pixels are located in the middle of the 32 by 32 pixel size, shown by green dotted lines [9]. Since 96 x 96 full-size images provide more accuracy than small-sized datasets, they are used for the dataset's analysis.

Litjens et al. (2018) used the example entire slide WSI pictures. The tumor segmented sample picture is shown at different magnification levels. Fan and associates (2019) For efficient processing, the image's very high dimensionality is decreased. The photos are separated into patches from the camelyon16 and camelyon17 datasets used in Litjens et al. (2018) (Lu et al. 2021). This involves segmenting the tissue portion first, and then utilizing OTU's Thres holding approach to extract the patches from the WSI (Khalil et al. 2022).

Result analysis

Finding and preserving pertinent information that is suggestive of the traits of breast cancer is the primary goal of feature extraction. By lowering the data's dimensionality, it prevents over fitting and increases computing efficiency. As a result, categorization models become more robust. Additionally, feature extraction makes it easier to identify important patterns in medical pictures, such as edges and texture, which are crucial for distinguishing between benign and cancerous tissues. By improving the signal-to-noise ratio and reducing noise, this process helps to improve the overall quality of the data [16].

Moreover, feature extraction enhances the model's capacity to effectively generalize to new cases and categorize a variety of datasets. The accuracy of breast cancer classification models is finally improved by including discriminative characteristics, which also increases the algorithms' capacity to correctly forecast and distinguish between different kinds of breast tissues in medical imaging [10].

Table 3: Performance comparison of several segmentation techniques for the Camelyon16 dataset

Network	Accuracy	Precision	Recall	Specificity	DSC	IOU
Seg-Net	98.34	34.56	90.37	89.63	35	35
U-Net	90.32	12.45	78.21	90.32	36	87
Dense–Net	27.45	23.45	23.56	67.43	68	36
Hook–Net	24.56	34.56	82.45	21.34	46	36
FCMDCS	90.43	35.54	21.56	34.65	35	57
CDJS	67.32	89.34	87.32	32.54	33	46

Co-Occurrence Matrix Feature for Grey Level

By looking at the spatial connection between the pixels, the GLCM technique has been used to analyze the texture of the breast tissues. The co-occurrence matrix, which shows the frequency of co-occurrences of different pixel intensity value pairs in a picture, is used to do this. This matrix may be used to extract a number of texture-related characteristics, including homogeneity, correlation, energy, contrast, and entropy. This shows several pixel patterns, including smoothness, roughness, uniformity, and so on, which are crucial for detecting cancer.

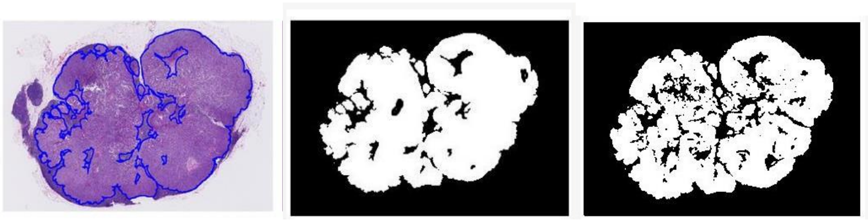

Figure 4: Input Image Visualization with Associated Ground Truth and Predicted Region of Interest

The distribution of the matrix is ascertained by using moments such as mean, variance, standard deviation, and root mean square. L will indicate the number of distinct gray levels, $\sigma 2$ is variance, µ will mean the value of pixels in that specific segmented region, "i" will be the intensity value, and Pii will be the probability of intensity value. If a,b refers to the row and column position index of a pixel, then Pia,b will refer to the frequency of co-occurrences of the pixel intensity value pair (a,b) taken from GLCM.

Table 4: Camelyon17 Dataset's WSI Image Count

	Total WSIs		Metastases			
Centre	**Train**	**Test**	**None**	**ITC**	**Micro**	**Macro**
CWZ	231	352	16	25	64	16
LPON	132	525	15	4	2	14
RST	420	252	74	2	63	11
RUMC	230	363	25	7	15	35
UMCU	240	263	25	5	75	46
Total	1253	1755	155	43	219	122

Statistical Feature

The technique of identifying patterns in an original picture is known as feature extraction. From the photos, the texton features, shape-based features, enhanced LGXP, and statistical features are extracted. The statistical traits produced by the suggested approach are notably non-redundant, intelligible, and distinguishable. Standard deviation, mean, min, max, and median are among the statistical characteristics that are assessed.

The center pixel value is chosen from an ordered set of values in the mn neighborhood W around the original pixel using an order-statistic filter. Equation (5.11) is then used to construct the median.

(5.11) Med = median PCi (r, c) | (r, c)$\in$W

Equation (5.12) is an expression for the standard deviation mathematical model.

Better LGXP-based functionality

To ascertain if an image contains tumorous or non-tumorous tissues, texture-related characteristics are extracted using LGXP, or local grey level and X-ray photon. The enhanced LGXP-based features are taken from the picture. The texture aspects of the picture are explored in the LGXP using the Complex Gabor filter module with real and imagery factor. When applied to a picture, the result is a complex coefficient that illustrates the texture-based properties of the tissues. In order to disclose textural characteristics that may be used for further processing, the complex coefficients of Gabor (output value) are made to stack into a map with real and imaginary values.

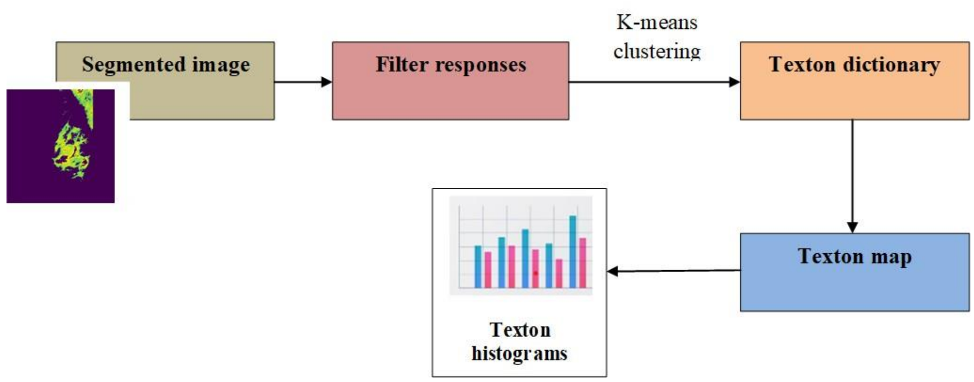

Figure 5: Texton Feature Extraction Workflow

Two filters, such as an odd filter and an even filter, are used in the band-pass quadrature to allow phase assessment at a specific spatial area. The selection quadrature filter, which can be constructed with any bandwidth, is regarded as an enhanced Gabor filter. The picture should be examined throughout a variety of frequencies in order to get simultaneous region of frequency and spatial information [12]. This may be achieved by rescaling the improved Gabor filter in order to create a filter bank. The suggested hybrid model comprising two filter functions, such as a one-dimensional log Gabor function and a one-dimensional Gaussian smoothing function, takes the place of this Gabor filter.

Table 5: (subsequent)

Layer	Type	Input	Kernel	Stride	Pad	Output
Conv3	Convolution	256x24x24	3x3	1	1	384x24x24

Pooling	Max Pooling	384x24x24	3x3	2	0	384x11x11
Conv1	Convolution	384x11x11	3x3	1	1	384x9x9
Conv2	Convolution	384x9x9	3x3	1	1	384x9x9
Conv3	Convolution	384x9x9	3x3	1	1	256x9x9
Pooling	Max Pooling	256x9x9	3x3	2	0	256x4x4
FullyConnected1		4096	3x3	1	1	4096x1
FullyConnected2		4096	1x1	1	1	4096
FullyConnected3		4096	1x1	1	1	4096

Log gabor function in one dimension

The image's edge is found using the log gabor function. To manage the image's complex structure, log gabor filters are used in the frequency domain. However, because of the fluctuating bandwidth in log gabor, the log function at origin is limited. Additionally, the polar coordinate system combines the angular and radial components [13]. The one-dimensional log gabor function may thus be applied as shown in

Gaussian smoothing function in one dimension

The image's borders are smoothed using a one-dimensional Gaussian smoothing method. In order to reduce noise, each level smoothes the original picture to varying degrees. The picture's complete resolution is included in the base level of Gaussian smoothing, and the low level image is then smoothed by the upper level. The kernel size is determined by the Gaussian filter's standard deviation [14].

Conclusion

By combining multimodal data such as clinical factors and medical imaging the research shows how deep learning may be a strong tool for predicting lymph node metastases in individuals with breast cancer. In contrast to models that only use one data modality, the suggested method improves prediction accuracy by using both radiological parameters and patient-specific clinical information. More thorough and customized evaluations are made possible by this multimodal deep learning system, which may help clinicians make better decisions. According to the research, mixing different data sources enhances the model's sensitivity, specificity, and robustness, making it a viable approach for treatment planning and early diagnosis. However, to guarantee generalizability and clinical application, more validation on bigger, more varied datasets is required.

Reference

1. Tang X, Zhang H, Mao R, Zhang Y, Jiang X, Lin M, et al. Preoperative prediction of axillary lymph node metastasis in patients with breast cancer through multimodal deep learning based on ultrasound and magnetic resonance imaging images. AcadRadiol. 2025;32(1):1–11.
2. Rane NL, Mallick SK, Kaya O, Rane J. Machine learning and deep learning architectures and trends: A review. In Applied Machine Learning and Deep Learning: Architectures and Techniques. Deep Science Publishing: 2024. p. 1–38. https://doi.org/10.70593/978-81-981271-4-3_1.
3. Shafik W. Deep learning impacts in the field of artificial intelligence. Book Deep Learning Concepts in Operations Research. Auerbach Publications: 2024. p. 18.
4. Zhou J, Yu X, Wu Q, Wu Y, Fu C, Wang Y, et al. Radiomics analysis of intratumoral and different peritumoral regions from multiparametric MRI for evaluating HER2 status of breast cancer: A comparative study. Heliyon. (2024) 10:e28722. doi: 10.1016/j.heliyon.2024.e28722
5. Xie X, Zhou H, Ma M, Nie J, Gao W, Zhong J, et al. A deep learning model for predicting molecular subtype of breast cancer by fusing multiple sequences of DCE-MRI from two institutes. AcadRadiol. (2024).
6. Qu L, Mei X, Yi Z, Zou Q, Zhou Q, Zhang D, et al. An unsupervised learning model based on CT radiomics features accurately predicts axillary lymph node metastasis in breast cancer patients diagnostic study. Int J Surg. (2024).
7. Li H, Hou Y, Xue L, Fan W, Gao B, Yin X. Use of MRI radiomics models in evaluating the low HER2 expression in breast cancer. Curr

Med Imaging Formerly Curr Med Imaging Rev. (2024) 20. doi: 10.2174/0115734056234429240426181002

8. Carriero A, Groenhoff L, Vologina E, Basile P, Albera M. Deep learning in breast cancer imaging: state of the art and recent advancements in early 2024. Diagnostics (Basel). (2024) 14. doi: 10.3390/diagnostics14080848

9. Hua Y, Peng Q, Han J, Fei J, Sun A. A two-center study of a combined nomogram based on mammography and MRI to predict ALN metastasis in breast cancer. Magnetic Resonance Imaging. (2024) 110:128–37. doi: 10.1016/j.mri.2024.04.019

10. Ma X, Zhang L, Xiao Q, Huang Y, Lin L, Peng W, et al. Predicting prognosis of phyllodes tumors using a mammography- and magnetic resonance imaging-based radiomics model: A preliminary study. Clin Breast Cancer. (2024). doi: 10.1016/j.clbc.2024.05.006

11. Rokhshad R, Salehi SN, Yavari A, Shobeiri P, Esmaeili M, Manila N, et al. Deep learning for diagnosis of head and neck cancers through radiographic data: a systematic review and meta-analysis. Oral Radiol. 2024;40(1):1–20.

12. Polat DS, Nguyen S, Karbasi P, Hulsey K, Cobanoglu MC, Wang L, et al. Machine learning prediction of lymph node metastasis in breast cancer: performance of a multi-institutional MRI-based 4D convolutional neural network. Radiol Imaging Cancer. 2024;6(3): e230107.

13. Shiner A, Kiss A, Saednia K, Jerzak KJ, Gandhi S, Lu FI, et al. Predicting patterns of distant metastasis in breast cancer patients following local regional therapy using machine learning. Genes (Basel) 2023;14:1768. doi: 10.3390/genes14091768

14. Guo YJ, Yin R, Zhang Q, Han JQ, Dou ZX, Wang PB, Lu H, Liu PF, Chen JJ, Ma WJ. MRI-Based Kinetic Heterogeneity Evaluation in the Accurate Access of Axillary Lymph Node Status in Breast Cancer Using a Hybrid CNN-RNN Model. J MagnReson Imaging. 2024;60(4):1352–64.

15. Zhou H, Hua Z, Gao J, Lin F, Chen Y, Zhang S, et al. Multitask deep learning-based whole-process system for automatic diagnosis of breast lesions and axillary lymph node metastasis discrimination from dynamic contrast-enhanced-MRI: a multicenter study. J MagnReson Imaging. 2024;59(5):1710–22.

16. Guni A, Sounderajah V, Whiting P, Bossuyt P, Darzi A, Ashrafian H. Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies Using AI (QUADAS-AI): Protocol for a Qualitative Study. JMIR Res Protoc. 2024;13: e58202.

17. Gradishar, W. J. et al. Breast cancer, version 3.2024, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* **22**, 331–357 (2024).

18. 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;14:1351. doi: 10.1038/s41598-024-51244-y.

19. Polat DS, Nguyen S, Karbasi P, Hulsey K, Cobanoglu MC, Wang L, et al. Machine learning prediction of lymph node metastasis in breast cancer: performance of a multi-institutional MRI-based 4D convolutional neural network. Radiol Imaging Cancer 2024;6:e230107. doi: 10.1148/rycan.230107.

20. Garcia-Tejedor A, Ortega-Exposito C, Salinas S, Luzardo-González A, Falo C, Martinez-Pérez E, et al. Axillary lymph node dissection versus radiotherapy in breast cancer with positive sentinel nodes after neoadjuvant therapy (ADARNAT trial). Front Oncol. (2023) 13:1184021. doi: 10.3389/fonc.2023.1184021.