IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

++++++Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

High Performance Breast Cancer Diagnosis from Mammograms Using Mixture of Experts with EfficientNet Features (MoEffNet)

HOSAMELDIN O. A. AHMED [1], and ASOKE K. NANDI[1], Fellow, IEEE

[1]Department of Electronic and Electrical Engineering, Brunel University London, Uxbridge UB8 3PH, U.K.

Corresponding author: ASOKE K. NANDI (e-mail: asoke.nandi@ brunel.ac.uk).

"This work was supported in part by Brunel University London research funding scheme."

ABSTRACT As breast cancer is a leading cause of death for women globally, there is a critical need for better diagnostic tools. To address this challenge, we propose MoEffNet, a cutting-edge framework that offers high-performance breast cancer diagnosis. MoEffNet is characterised by its innovative hybrid integration of EfficientNet and Mixture of Experts (MoEs), two powerful techniques developed to enhance accuracy and efficiency. EfficientNet, known for its robust feature extraction capabilities, utilises compound scaling and depth-wise separable convolutions to capture image features across multiple levels of abstraction. This is combined with MoEs framework, which employs specialised expert networks to analyse distinct aspects of mammograms. MoEffNet analyses features at various levels: low-level for basic patterns, mid-level for detailed analyses, and high-level for complex contents. Features extracted from various EfficientNet model stages are assigned to specialised experts to optimise diagnostic precision. A dynamic gating mechanism (EffiGate) is introduced to ensure that the most relevant experts contribute to each diagnostic decision, by dynamically adjusting their influence based on input data characteristics. This approach ensures that the most effective experts are utilised for each case, resulting in superior accuracy. The scalability of MoEffNet is highlighted by its ability to adapt to various computational constraints and accuracy requirements, using EfficientNet's architecture, which ranges from B0 to B7 models. We have validated MoEffNet's effectiveness on three mammographic datasets (MIAS, CBIS-DDSM, and INbreast) achieving outstanding results (AUC > 0.99 across all datasets), outperforming existing methods. Particularly, EfficientNet B1 and B2 models with three or four experts achieved the highest accuracy, demonstrating MoEffNet's potential as a robust diagnostic tool for early breast cancer detection. Through its innovative hybrid model, robust feature extraction, dynamic gating, and specialised expert networks, MoEffNet sets a new benchmark in automated mammogram analysis, offering a powerful tool for more accurate and reliable breast cancer diagnosis.

INDEX TERMS Breast cancer diagnosis, Computer-aided diagnosis, Deep learning, Machine learning, Multi-view analysis, Mammography.

I. INTRODUCTION

Breast cancer represents a major global public health burden, with an estimated 2.3 million new cases diagnosed leading to 685,000 deaths in 2020 alone. This number remained significant in 2022, with approximately 670,000 deaths attributed to the disease [1, 2]. As the most common cancer affecting women, it necessitates continuous efforts to raise awareness, improve early detection strategies, which is vital for improved treatment outcomes [3], and develop effective treatment modalities and comprehensive care for patients worldwide. Breast cancer staging groups the disease based on the degree of its spread, which helps doctors decide the best course of treatment. The earliest stage, stage 0 (in situ cancer), involves abnormal cells confined within milk ducts or lobules. Stages I-IV incorporate invasive cancers that have spread into surrounding tissue. Stage I features small tumours with possible lymph node involvement, while stages progress with increasing tumour size, more lymph node involvement, or even chest wall/skin invasion. In stage IV (metastatic), cancer has spread to distant organs. Generally, a lower stage indicates a better prognosis, but other factors like cancer cell biology also influence treatment decisions [4 - 6]. For example, in Canada, the chance of surviving for 5 years (net survival) is very high (nearly 99.8%) for women diagnosed with stage I breast cancer. This survival rate decreases significantly for more advanced stages of breast cancer, with a 5-year net survival rate of 91.9% for stage II, 74.0% for stage III, and 23.2% for stage IV [7]. Also, a recent study conducted by the American Cancer

Society revealed significant variations in a woman's likelihood of surviving for five years (net survival) following a breast cancer diagnosis, depending on the cancer's stage at diagnosis. Patients who were diagnosed with stage I breast cancer from 2012 to 2018 had an outstanding 5-year survival rate of more than 99%. However, as the cancer progresses through stages, the survival rate significantly decreases. For stage II, the 5-year survival rate is 93%, for stage III it is 75%, and for stage IV it drops to 29% [8]. Supporting this, a study conducted in the Netherlands analysing breast cancer trends from 1989 to 2017 confirms the global importance of early detection. This study shows that while initial breast cancer diagnoses rose, recent years have shown a decline. Treatment advancements are evident, with less invasive surgeries and a rise in systemic therapies. Most importantly, survival rates have improved significantly across all stages, and overall breast cancer mortality rates have declined regardless of age [9]. This highlights the critical role early detection plays in improving a woman's chances of successful treatment and long-term survival.

II. REVIEW OF SELECTED METHODS

Advancements in diagnostic medical imaging have produced a diverse toolbox for breast cancer detection, including mammograms, breast thermography, magnetic resonance imaging, ultrasound, positron emission tomography, histopathology, and computed tomography [10-16]. Understanding these techniques empowers informed decision-making regarding screening and early detection strategies. Mammography is an essential component of early detection for breast cancer and is known to significantly improve patient outcomes in terms of treatment success and survival rates. It excels at identifying abnormalities before they become noticeable or cause symptoms. Numerous studies have shown that digital mammography is highly effective in accurately diagnosing breast cancer [17 - 27]. This makes it a fundamental tool for both screening and detailed evaluation of abnormal findings. Besides detecting breast cancer, mammography plays an important role in treatment planning by providing detailed information about tumour size and location, which is essential for surgery and other therapies. Furthermore, it has been shown that digital mammography offers comparable accuracy to screen-film mammography (SFM) in detecting breast cancer, but it has a significant advantage in identifying tumours, particularly in women with dense breast tissue [28]. Additionally, despite limited healthcare access, mammography remains a valuable tool for early breast cancer detection in low and middle-income countries [29]. These studies set mammography's critical role in breast cancer diagnosis, with its value across settings and its continuous evolution for better patient outcomes.

The main benefit of mammogram screening lies in its capability to reduce breast cancer mortality. It achieves this by detecting the disease at an early stage, allowing for earlier intervention, and potentially improving patient prognosis [30]. While breast cancer screening saves lives, it has limitations such as high costs, lengthy procedures, false positives, and human error. Furthermore, existing imaging methods have certain limitations [31]. For instance, up to 35% of breast cancers can be missed during mammography screening, often due to factors such as dense breast tissue or overlying breast tissue. This can result in interval cancers that are detected between regular screenings [32]. Hence, advanced technologies are necessary to improve accuracy, efficiency, and eventually, patient experience. Advancements in AI and automatic diagnostic systems are transforming the field of breast cancer screening. These systems act as powerful assistants to physicians, employing advanced learning and analysis to improve accuracy and efficiency. By detecting subtle patterns that may be missed by the human eye, they can reduce the number of false positives and provide more precise initial assessments. In the end, this supports better decision-making for patient care [33].

Computer-aided diagnostic (CAD) systems powered by Artificial Intelligence (AI) are being developed to detect breast cancer. Studies show that using these systems can lead to a significant increase of 7.62% in detection rates with minimal impact on recall rates, which increased by only 0.93%. By using advanced imaging analysis, these systems help to improve the accuracy and efficiency of diagnosis and could potentially result in reduced mortality and morbidity rates associated with breast cancer [34 – 36]. Traditional machine learning methods have proved to be an essential tool in advancing the field of breast cancer detection through the development of CAD systems using mammograms. These systems assist radiologists by automatically identifying and segmenting suspicious regions, including masses and calcifications [37]. For example, early breast cancer detection in mammograms utilises various segmentation and classification techniques. These techniques range from analysing the entire image (non-segmentation) to segmenting the breast tissue based on its distance from the skin. More precise segmentations can be obtained using advanced methods such as Fuzzy C-Means, Fractal Analysis, and Statistical Analysis. After segmenting the tissue, its shape and texture are analysed to extract features. Finally, a Bayesian framework combining k-Nearest Neighbours and C4.5 decision trees is used to classify tissue. This study shows that Breast segmentation using internal information improves cancer detection. The results demonstrated that the fuzzy C-Means technique significantly enhances breast cancer detection, achieving an accuracy of 82% compared to the 62% accuracy of non-segmentation methods, highlighting the importance of this initial step in improving breast cancer detection [38].

Furthermore, a technique using the Contourlet transform, a combination of Laplacian Pyramidal and Steerable

Gaussian Filters, was proposed to detect architectural distortions in mammograms. This approach aimed to analyse textural features through an Artificial Neural Network (ANN) for classification. However, despite its innovative design, the system generated a high number of false positives (1255 out of 1502 regions flagged), emphasizing the need for further development to improve accuracy [39]. Moreover, a breast cancer identification system utilises image processing techniques and neural networks to distinguish between benign and malignant tumours in mammograms. The system works in stages: first, image processing techniques including grayscale conversion, intensity adjustment, and filtering enhance tumour visibility. Then, segmentation methods incorporating thresholding and morphological operations isolate the tumour region. Next, feature extraction analyses the segmented area using the Gray-Level Co-occurrence Matrix (GLCM) to capture texture and by directly measuring properties like asymmetry and roundness to capture shape. Finally, a neural network classifier, trained on these extracted features, distinguishes between benign and malignant tumours with an impressive 92% identification rate, indicating the effectiveness of this combined image processing and neural network approach [40]. Nevertheless, these techniques rely heavily on hand-crafted features, such as morphological, topological, and textural descriptors. This dependence makes them challenging to develop and highly sensitive to the quality of the selected features.

While these conventional techniques were valuable, especially when dealing with limited annotated data, there has been a significant idea towards deep learning. Deep learning offers a powerful alternative by automating feature extraction directly from raw mammogram images. This eliminates manual feature engineering, allowing the model to identify and learn effective features automatically. This shift shows a major advancement in enhancing breast cancer detection through mammography, possibly leading to improved performance metrics such as accuracy, sensitivity, and specificity. For instance, a novel approach for breast cancer detection in mammograms using convolutional neural networks (CNNs) for feature extraction was introduced in [32]. In this approach, the most informative features are selected from multiple pre-trained CNN models and used to train various machine learning algorithms including neural networks (NN) and support vector machines (SVMs). The experimental results of this method achieved remarkable accuracy (92-96%) across three datasets (RSNA, MIAS, and DDSM), highlighting its effectiveness and potential to outperform other methods in breast cancer detection [41]. In [42], a deep-learning methodology for breast cancer detection in mammograms is presented. This approach segments the breast tissue using a modified U-Net model and then classifies the isolated region as benign or malignant using several CNN architectures such as InceptionV3. This approach employs data augmentation and transfer learning techniques to address the challenge of limited data used in this study. Additionally, both CC and MLO views were used to improve the accuracy. This approach achieved notable results, particularly on the DDSM dataset, with 99% accuracy and under 1.2 seconds processing time [42]. In [43], a deep learning method employing CNNs is used to analyse mammogram and tomosynthesis images for breast cancer detection. Over 3,000 images with confirmed pathology results are utilised to develop CNN models. The validation results showed promising accuracy, suggesting CNNs' potential for automatic breast cancer detection in mammograms and tomosynthesis [43].

Moreover, a Breast Mass Classification (BMC) system, combining deep learning and ensemble learning has been introduced for breast mass classification in mammograms. This system integrates k-means clustering, Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), Random Forest, and Boosting techniques. By segmenting mammograms and extracting features through LSTM, CNNs, and pre-trained CNN models, the system achieves high accuracy (over 95%) and strong generalisability across datasets [44]. In [45], a deep learning model employing transfer learning is proposed for breast cancer detection and classification. Pre-trained CNNs such as Inception V3 are improved to analyse mammograms. Data preprocessing, segmentation, and augmentation techniques prepare the images. The model achieves excellent performance with an overall accuracy exceeding 98%, demonstrating its effectiveness in identifying breast cancer from mammograms [45]. In [46], a Fully Connected Layer First CNN (FCLF-CNN) technique is proposed to address the limitations of traditional CNNs on structured data. It places fully connected layers before convolutions, acting as encoders to transform raw data into localized representations. This structure significantly improves classification performance, as shown by its remarkable accuracy (over 98%) and sensitivity/specificity on two breast cancer datasets, namely, the Wisconsin Diagnostic Breast Cancer (WDBC) database and the Wisconsin Breast Cancer Database (WBCD), outperforming both multi-layer perceptron (MLP) and pure CNNs [46].

Additionally, a deep learning framework applying a dual-path CNN is proposed for breast mass segmentation and diagnosis in mammography simultaneously. This dual approach employs the Locality Preserving Learner (LPL) to extract image features for classification, while the Conditional Graph Learner (CGL) focuses on pixel-wise relationships for segmentation. By combining these learned features, the method achieves superior performance in both tasks, surpassing other methods on benchmark datasets with segmentation accuracies of 92.27% (DDSM) and 93.69% (INbreast), alongside strong classification results

[47]. In [48], researchers propose a Shallow-Deep CNN (SD-CNN) to improve breast cancer diagnosis. This dual-CNN approach utilises a shallow CNN to create virtual images from standard mammograms and a deep CNN to analyse these alongside the originals for better feature extraction. By combining information from both standard and enhanced imaging, the SD-CNN achieves an accuracy of 89% and an AUC of 91%, significantly outperforming the analysis of mammograms alone (accuracy: 85%, AUC: 84%). This suggests the SD-CNN effectively uses additional imaging data to improve diagnostic accuracy.

Recently, researchers introduced a new method using Deep Convolutional Generative Adversarial Networks (DCGANs) that aims to create synthetic mammograms identical to real ones to improve breast cancer detection. Researchers evaluated the technique by having radiologists assess image authenticity, revealing a significant gap between real and synthetic images. This highlights the need for further development in DCGANs to ensure synthetic mammograms can be reliably used for medical diagnosis [49]. In [50], a new approach employing radiomics and deep learning techniques for breast cancer diagnosis with multiparametric mammography is presented. This methodology employed adaptive filtering and data augmentation for robust model training. A novel Chaotic Leader Selective Filler Swarm Optimization (cLSFSO) extracts textural features to locate suspicious lesions, while modified deep learning models (VGGNet and SE-ResNet152) with transfer learning classify normal from concerning regions. Hybrid models incorporating CNNs, Long Short-Term Memory (LSTM) networks, and SVMs further enhance diagnosis and grading. Finally, Grad-CAM techniques highlight crucial areas within mammograms, improving interpretability and accuracy. These advancements achieved a sensitivity result of 99% and an AUC of 99% [50].

Deep learning algorithms have also allowed for advances in the use of multi-view mammography in breast cancer diagnosis, offering a novel approach to diagnosis. Multi-view learning utilises multiple views of the same data, where each view represents a different subset of features or a different representation of the data. The main idea is that combining these views can result in a broader insight and better performance than using a single view alone. A comprehensive overview of the current literature on multi-view information fusion in mammograms is provided in reference [51], which provides an in-depth discussion on the application of multi-view information fusion (MVIF) within the framework of computer-aided diagnosis (CAD) for breast cancer. It elaborates on how screening mammography, which provides two views of each breast (MLO and CC), benefits from MVIF to enhance diagnostic accuracy. The paper categorises MVIF methods into detection, classification, and content-based mammogram retrieval (CBMR), with each category further subdivided into various approaches based on how they utilize the information from different views. The review highlights the advantages of combining these views to reduce false positives and improve detection and classification rates, ultimately aiding radiologists in decision-making [51]. In [52], the Anatomy-Aware Graph Convolutional Network (AGN) method, was designed for mammography mass detection with multi-view reasoning capabilities. The AGN process involves three key steps. First, the AGN employs a Bipartite Graph Convolutional Network (BGN) to model the relationships of ipsilateral views. Second, an Inception Graph Convolutional Network (IGN) captures the structural similarities of bilateral views, aiding in lesion detection. Finally, AGN distributes multi-view information across its network, enhancing feature analysis for multi-view reasoning. Benchmark tests confirm AGN's considerable improvements in detection capabilities [52].

The researchers in [53] developed a Multi-View Feature Fusion (MVFF) based Computer-Aided Diagnosis (CADx) system using deep learning to classify mammograms. This system processes four different mammogram views through Convolutional Neural Networks (CNNs) to extract features. These features are then combined into a single predictive layer to enhance classification accuracy. The system is trained on augmented data from public datasets like CBIS-DDSM and mini-MIAS, showing improved performance over single-view systems in detecting normal vs. abnormal, mass vs. calcification, and malignant vs. benign classifications. This multi-view approach uses complementary information, achieving superior performance (AUC of 93.2% for mass/calcification, 84% for malignant/benign, and 93% for normal/abnormal) compared to single-view systems, highlighting its potential for improving mammogram classification accuracy. Also, an automated deep learning-based analysis of unregistered multi-view mammograms is introduced to assess breast cancer risk. It uses deep learning models, initially trained on large, non-medical image datasets through transfer learning and subsequently fine-tuned on mammogram data. This method analyses MLO and CC views and their lesion segmentation maps holistically, rather than focusing on individual lesions. The approach utilises defined and automatically generated segmentation maps, enhancing flexibility and effectiveness [54]. A breast cancer diagnosis technique using an EfficientNet-based convolutional network trained end-to-end on two-view mammograms incorporates three stages of transfer learning. It starts with a patch classifier developed from a model trained on natural images, progresses to a single-view whole-image classifier, and culminates in a two-view classifier integrating both mammographic views. This technique achieves high diagnostic effectiveness accuracy, with an AUC of up to 93.44%, proving its effectiveness in breast cancer detection.

TABLE I
SUMMARY OF MIXTURE OF EXPERTS (MOES) APPLICATIONS AND THEIR BENEFITS ACROSS VARIOUS MEDICAL CONTEXTS

Ref	Aspect	Medical Application	Data Type	Detail	Benefits
[57]	Data Integration	Cancer Diagnosis	Clinical and Genetic Markers	Integrates clinical factors with gene markers for cancer diagnosis, handling diverse data types.	Enhances diagnostic accuracy by utilising richer data analyses.
[58]	Repeated Measures	Rectal Cancer Monitoring	MRI Data	Adapts MoEs for MRI data in rectal cancer monitoring via perfusion MRI.	Allows dynamic modeling of disease progression, adjusting to changes over time.
[59]	Improved Diagnostic Accuracy	Breast Cancer	Clinical Records	Uses MoEs in breast cancer diagnosis, achieving higher accuracy rates compared to standalone models.	Increases the reliability of diagnostic outcomes, supporting better clinical decisions.
[60]	Prognostic Modeling	Cancer Treatment Response (understanding how different proteins and genetic markers affect survival outcomes among breast cancer patients)	survival times and protein expression levels from breast cancer patients	A Bayesian approach is used to enforce sparsity on regression coefficients, pinpointing key factors for each sub-group. The mixture model allows modeling of data from multiple, unknown sub-groups with distinct survival characteristics.	The study's benefits include improved sub-group identification, precise feature selection, potential for personalized medicine, and a robust model capable of handling complex survival data through a Bayesian mixture-of-experts framework.
[61]	Identify hidden subgroups and assess liver stiffness in patients, potentially indicating liver cirrhosis.	Identifying liver cirrhosis by non-invasively measuring liver stiffness using transient elastography (Fibroscan).	liver stiffness measurements from 228 HBsAg-positive patients, were collected using the Fibroscan system over six months.	MoEs to process input data (liver stiffness measurements), using the expectation-maximization algorithm for effective training and classification of patient data into subgroups indicative of liver health statuses.	improved diagnostic accuracy for liver cirrhosis, providing a non-invasive and rapid method for medical professionals to assess liver fibrosis and cirrhosis, which is crucial for timely and effective patient management.
[62]	Classify 3D optical coherence tomography (OCT) images	Automatic diagnosis of retinal diseases	Optical coherence tomography (OCT) images	Wavelet-based Convolutional Mixture of Experts (WCME) model combines wavelet-based feature extraction with CNNs in a MoEs framework to classify OCT images using high-level features and spatial-frequency information, achieving high precision and accuracy without needing manual retinal layer segmentation.	The WCME model significantly enhances the diagnosis of macular diseases.
[63]	A decision tree-based ensemble model using a mixture of discriminative experts optimized with the Epistocracy algorithm for classifying COVID-19 infections	Diagnosis of COVID-19.	chest X-ray imaging	EpistoNet uses two mixtures of discriminative experts, trained on similar X-ray image clusters and optimized with the Epistocracy algorithm, enhancing accuracy and adapting the model architecture and hyperparameters effectively.	Increased Diagnostic Accuracy.

Together, these studies show that a multi-view approach can involve multiple models, each trained on different views or a single model that combines several views into its design. In this context, ensemble methods such as co-training, where two models are trained at the same time on different viewpoints and teach each other are common.

Another aspect of multi-view analysis highlighting the power of ensemble and composite approaches in supervised machine learning is the Mixture of Experts (MoEs) technique. This technique consists of multiple separate networks, each specialised in managing a subset of the full training dataset [56]. Several studies show that MoEs models play a key role in advancing personalised medicine and medical diagnostics by efficiently processing various data types like genetic markers, clinical data, and images. These models enhance diagnostic precision and treatment results in areas such as oncology and neurology. MoEs support real-time disease monitoring and customised treatment modifications, significantly improving individual patient care [57 – 63]. Table I details their wide-ranging applications and benefits across different medical fields and data types.

Advanced techniques such as ensemble models contribute to the robustness and accuracy of disease diagnosis systems. One of the potentials of these advanced techniques is to assist clinical professionals by offering a promising tool for mammography interpretation, which can represent a reliable second opinion and support complex diagnostic decisions. Looking forward, advancements in the field could be driven by developing further innovations, incorporating larger, more diverse datasets, and introducing real-time adaptive learning models. These innovations could continuously refine diagnostic capabilities and adapt to new challenges in breast cancer detection. In this vein, our research explores the application of MoEs for breast cancer diagnosis in mammograms.

MoEs use a divide-and-conquer strategy, employing a collection of expert networks. Each network within the MoEs ensemble specialises in handling specific aspects of the data. This allows the model to become more complex and adaptable without a corresponding increase in computational demands during inference. Only a subset of the experts is activated for a given input, promoting efficient resource allocation. While there are currently limited direct references to MoEs being used for breast cancer diagnosis through mammograms, the inherent flexibility and capacity for detail-specific processing make MoEs a promising solution for such complicated tasks. Our research aims to address this gap by investigating the potential of MoEs in this domain. We propose to integrate MoEs with EfficientNet features, known for their state-of-the-art performance in image classification. EfficientNet applies depth-wise separable convolutions, which contribute to model efficiency and facilitate compound scaling [64].

By integrating the adaptability of MoEs with the robust capabilities of EfficientNet, we aim to develop a novel method for diagnosing breast cancer in mammograms. This innovative hybrid approach, termed MoEffNet, distinguishes itself by efficiently extracting and processing features across multiple levels of abstraction. MoEffNet analyses low-level features for basic patterns, mid-level features for detailed analyses, and high-level features for complex content, assigning these features to specialised expert networks to optimise diagnostic precision. Additionally, MoEffNet incorporates a dynamic gating mechanism, named EffiGate, which evaluates the relevance of each expert network based on specific mammographic data characteristics, thereby enhancing diagnostic accuracy. These attributes allow MoEffNet to achieve high-performance breast diagnosis accuracy, as validated across diverse mammogram datasets. Our key contributions to the field include:

1. Investigating the application of MoEs for breast cancer classification in mammograms: This research explores the potential of MoEs for this specific task, potentially paving the way for further exploration within the medical imaging domain.
2. Integrating MoEs with EfficientNet features: This integration uses the adaptability of MoEs and the robust feature extraction capabilities of EfficientNet, taking advantage of the strengths of both methodologies.
3. Optimised MoEs with EfficientNet features for mammogram diagnosis: Our work examines the impact of hyperparameters (number of experts, network architecture) on the MoEs performance when combined with EfficientNet feature extraction.
4. Achieving Superior Diagnostic Accuracy Across Multiple Datasets: MoEffNet achieves high accuracy in breast cancer diagnosis, validated through rigorous testing on three distinct mammographic datasets (MIAS, CBIS-DDSM, and INbreast). The model demonstrates AUC values of 99.2% for MIAS, 99.5% for CBIS-DDSM, and 99.7% for INbreast, significantly outperforming existing methods and establishing a new technique in automated mammogram analysis.

The remainder of the paper proceeds as follows: Section III is devoted to descriptions of the proposed method. Section IV is dedicated to a description of the performed experimental study. Finally, section V draws some conclusions from this study.

III. PROPOSED METHOD

This section details the MoEffNet methodology, an innovative approach for the diagnosing of breast cancer. As illustrated in Figure 1, MoEffNet integrates a CNN

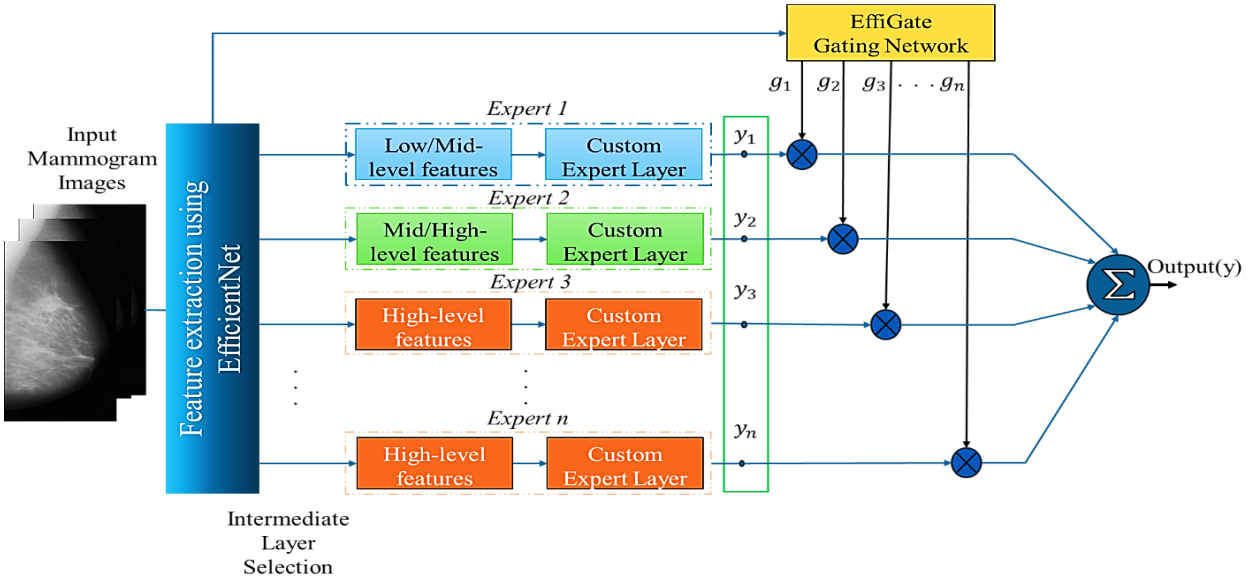


FIGURE 1. Overview of the Proposed MoEffNet Architecture: Integrating EfficientNet for Multi-Level Feature Extraction and Mixture of Experts (MoEs) for Adaptive Breast Cancer Classification.

architecture with a dynamic gating mechanism. This combination is designed to enhance the accuracy and efficiency of breast cancer detection from mammographic images. This developed methodology adopts a hybrid strategy, employing EfficientNet for its robust capability to extract features across multiple levels of abstraction. It is integrated with a Mixture of Experts (MoEs) framework, wherein each expert is adapted to process distinct subsets of features extracted from one or more intermediate layers of EfficientNet to capture different scales and complexities of the image data. The decision-making process is guided by an advanced gating network, named "EffiGate." This network adjusts how much influence each expert has, effectively combining their insights to produce a precise and consistent diagnostic result. This ensures that the final output is accurate and fits the specific examined features. The design of this system is aimed at optimising diagnostic accuracy by utilising the extensive depth of feature extraction provided by EfficientNet, coupled with the specialised analytical capabilities of the multiple expert networks. The methodology ensures an advanced, accurate, and reliable diagnostic process that utilises deep and expert computational insights. The following subsections provide a detailed description of the methods and techniques employed in the proposed approach MoEffNet.

A. ADVANCED FEATURE EXTRACTION USING EFFICIENTNET

As depicted in Figure 1, MoEffNet utilises the EfficientNet architecture for its powerful ability to extract features at various levels of abstraction. The EfficientNet series encompasses a collection of CNN models designed for enhanced accuracy and efficiency. These models were developed through a broad analysis of model scaling, focusing on three key dimensions: the depth, width, and resolution of the networks [64]. The key novelty of the EfficientNet architecture lies in its use of a multiple coefficient to scale the network's width, depth, and resolution uniformly and systematically, rather than adjusting them independently. This multiple scaling technique enhances effectiveness by preserving a balanced proportion among all dimensions, which is essential for achieving higher accuracy without sacrificing the computational cost. EfficientNets are derived from a baseline model crafted through neural architecture search, which optimises for both accuracy and computational efficiency. This baseline model is scaled up, resulting in a series of models from EfficientNet B0 to EfficientNet B7 each offering different levels of accuracy and efficiency. These models consistently outperform previous ConvNet architectures like ResNet and MobileNet, achieving state-of-the-art accuracy on benchmarks such as ImageNet and other datasets. Notably, EfficientNet B7 reaches a top 1 accuracy of 84.4% on ImageNet, while being considerably smaller and faster than other top-performing networks like GPipe.

EfficientNet models are formed in blocks, each capable of capturing features at varying levels of abstraction. The fundamental component of EfficientNet-B0, the mobile inverted bottleneck MBConv layer, is illustrated in Figure 2. All eight models of the EfficientNet series (B0 - B7) incorporate these common blocks, though each model introduces slight variations and increasing complexities in their architectural designs [65, 66].

Given its exceptional balance between high accuracy and computational efficiency, EfficientNet was selected as the feature extraction backbone of our proposed MoEffNet. Its proven capability to outperform previous ConvNet architectures with fewer parameters and reduced

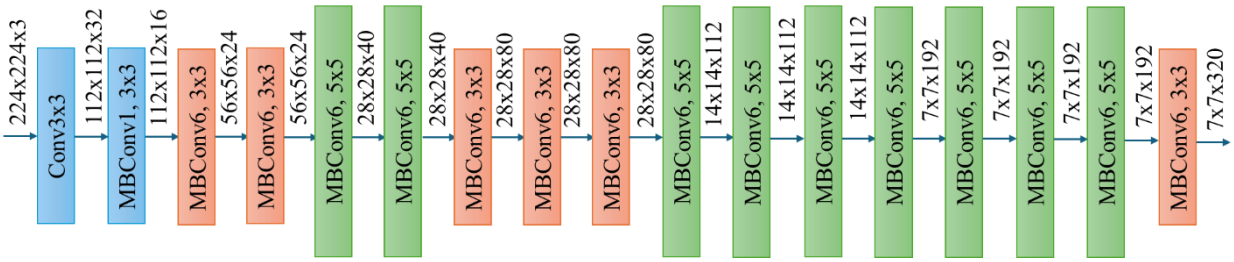


FIGURE 2. Detailed Architecture of EfficientNet-B0: Depicting Key Blocks and Layers Used for Feature Extraction in MoEffNet [65, 66].

computational load makes it an ideal choice for achieving state-of-the-art results without suffering high computational costs. The scalable design of EfficientNet, ranging from B0 to B7 models, also allows for adapted formations to match the specific needs of our research framework.

B. INTERMEDIATE LAYER SELECTION

As described above, EfficientNet is a scalable deep-learning model structured as a series of sequential blocks. Each block is designed to process image features with progressively greater complexity and abstraction. This hierarchical structure is particularly effective in medical imaging applications, such as mammograms, where distinguishing between benign and malignant features is crucial for precise diagnoses. MoEffNet applies a Mixture of Experts (MoEs) model, where each expert is designed to process specific subsets of features extracted from one or more intermediate layers of EfficientNet. The features extracted from EfficientNet models, as CNN-based pre-trained architectures, can be categorised as follows:

a) Low-level features: Typically captured in the initial layers, these are essential for identifying simple patterns and textures. In breast cancer diagnosis using mammograms, we believe that the initial layers of an EfficientNet are crucial for detecting low-level features. These features, which include fundamental elements such as edges, lines, and simple textures, are essential in analysing mammograms. They enable the identification of basic outlines and contours of breast tissues, which are critical for accurate diagnosis.

b) Mid-level features: Captured in middle layers, these features focus on shapes and specific parts of the input. These layers may identify more complex shapes and specific regions within the input image, such as masses or calcifications in mammograms. They serve as middle features, connecting the basic textural elements detected by the lower-level layers with the more abstract features recognised in the deeper layers.

c) High-level features: Extracted from the deeper layers, representing complex content such as objects. Generally, they may illustrate entire objects or complex configurations within the images, providing a comprehensive perspective that could be necessary for determining the final diagnostic decisions.

In the MoEffNet architecture, the integration of features extracted from the EfficientNet model plays a key role in enhancing efficiency and performance. Within the MoEffNet system, features are gathered from various stages - early, middle, and late layers - of a single EfficientNet model. This approach ensures a comprehensive representation across a diverse set of feature maps, which are critical for acquiring different levels of image complexity. Once extracted, these features are allocated to experts in the MoEffNet ensemble, which includes varying numbers of experts, typically ranging from two to four. Each expert is explicitly adapted to handle features based on their complexity and abstraction level. For example, experts, such as Expert 1, typically process low-level features to identify basic patterns. Mid-level features are assigned to Experts 1 and 2, to conduct more detailed analyses. High-level features are directed to other experts (such as experts 3 and 4) capable of interpreting more complex content, including entire objects.

The structured distribution of tasks within the MoEffNet system allows each expert to focus on definite types of features, enhancing the system's adaptability and capability to generalise. Each expert focuses on different levels of image complexity, enabling the ensemble to effectively process various image types. This targeted approach enhances both the accuracy and robustness of the model, making it exceptionally effective in diagnosing breast cancer through mammographic imaging.

C. MIXTURE OF EXPERTS (MoEs) MODEL

MoEffNet employs the MoEs model, where each expert is proposed to process given subsets of features extracted from one or more intermediate layers of EfficientNet. Figure 3 presents a graphical representation of the MoEs model. As shown in Figure 3 the MoEs is a hierarchical machine learning architecture consisting of multiple expert networks and a central gating network. This structure can be imagined as a tree where the expert networks (labelled

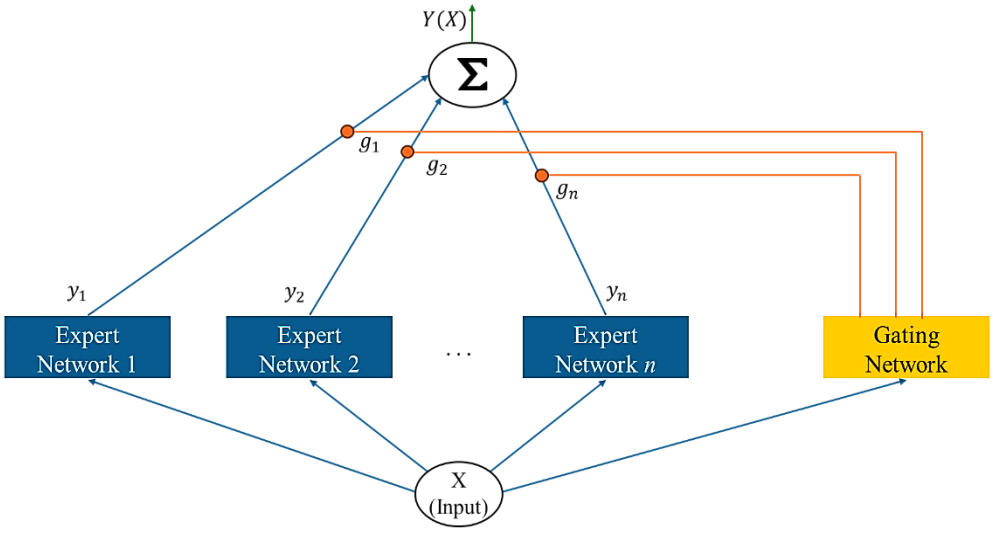


FIGURE 3. Graphical Representation of the Mixture of Experts (MoEs) Model: Showing the Integration of Expert Networks and the Gating Mechanism for Enhanced Diagnostic Accuracy [47].

Expert Network 1, Expert Network 2, through Expert Network n), placed at the leaves of the tree structure, process input vectors individually. Each expert network produces an output vector (y_i), where $i = 1, 2, \ldots, n$ denotes the number of experts, in response to every input vector (x). The output is a probability distribution over the number of classes. This distribution reflects the likelihood of each class being the correct classification for the given input. The output of each expert is thus a vector of probabilities where each component of the vector (corresponding to a class) represents the confidence level of the input belonging to that class.

The gating network highlighted in yellow, plays an essential role in integrating the outputs from the expert networks. Upon receiving an input vector x, the gating network generates scalar outputs (g_i) that form a partition of unity across the input space, fundamentally distributing the influence among the expert networks based on the characteristics of the input. The gating network computes and assigns linear combination coefficients, acting as probabilistic weights for the outputs of the expert networks. These weights determine the relative contribution of each expert network's output to the final decision. Accordingly, the final output of the architecture is a convex weighted sum of all the output vectors from the expert networks.

Suppose that there are n expert networks in the MoEs architecture. The ith expert network produces its output $y_i(x)$ as a generalised linear function of the input x such that.

$$y_i(x) = f(W_i x) \tag{1}$$

Here W_i is a weight matrix and $f(.)$ is usually considered the logistic function or the identity function. The gating network operates as a generalised function, where its $i-th$ output g_i is determined by applying a multinomial logit, or softmax function, to an intermediate variable ξ_i [56, 67 – 68].

$$g_i(x, v_i) = \frac{e^{\xi_i}}{\sum_{k=1}^{n} e^{\xi_k}} \tag{2}$$

Here $\xi_i = v_i^T x$ and v_i represents a weight vector. The overall output $Y(X)$ can be represented as follows:

$$Y(X) = \sum_{i=1}^{n} g(x, v_i) y_i(x) \tag{3}$$

$$Y(X) = \sum_{i=1}^{n} g_i(X) y_i(X) \tag{4}$$

This structure allows the gating network to output a set of probabilities that sum to one, making it suitable for conducting classification tasks where decisions are distributed across multiple categories.

$$\sum g_i(X) = 1 \tag{5}$$

D. CUSTOM EXPERT LAYER

The MoEffNet algorithm utilises the power of multiple expert networks, each designed to focus on different feature levels of the input data, thus enhancing the model's ability to handle complex and varied datasets efficiently. This capability is developed through the custom expert layer within our proposed method. The custom expert layer is fundamental to defining and adapting the layers for each expert within the ensemble. Each expert network is constructed using a sequence of neural network layers designed to improve performance for the kind of data it focuses on. These layers include dense layers with ReLU activation to introduce non-linearity and enable learning of complex patterns, followed by dropout layers that randomly deactivate a portion of neurons (specified by a dropout rate, of 30%) during training [69]. This prevents the network from becoming overly dependent on any single or small group of neurons, thereby reducing the risk of overfitting, and enhancing the model's ability to generalise

to new, unseen data. Additionally, L2 regularisation is applied to the weights of the dense layers, penalising large values which helps keep the model simpler and further guards against overfitting. The outputs from these experts are then collected using the gating mechanism's weights to produce a final, weighted output.

E. EFFIGATE GATING NETWORK

The decision-making process in MoEffNet is guided by an advanced gating network, known as EffiGate. This mechanism dynamically evaluates the contribution of each expert network within the ensemble, utilising features extracted from mammogram input images via an EfficientNet model. By doing so, EffiGate ensures that the most relevant and effective experts are selected, based on the input data's unique characteristics. EffiGate contains two main components that define its functionality within MoEffNet. The first component includes the input features, sourced from an EfficientNet model known for its effectiveness in handling image data, particularly in medical imaging contexts like mammograms. These features, which are high-level abstractions of the input data, are rich in detail and critical for subsequent processing.

The second component is the dense layer, which is configured as a fully connected neural network layer with units equal to the number of expert networks in the ensemble. Each unit in this dense layer outputs a score indicating the significance of each expert's network relative to the current input features. These scores are then processed through a SoftMax activation function, which converts them into normalised gating weights. These weights represent probabilities that measure the confidence or expected efficiency of each expert's contribution to the final decision-making process. The SoftMax ensures that these gating weights are non-negative and sum to one, making them interpretable as the likelihood of each expert's relevance to the specific input. This gating mechanism acts dynamically, continuously adjusting the weights of the various expert models based on the input data processed by EfficientNet, thus adjusting MoEffNet's response to diverse diagnostic scenarios.

IV. EXPERIMENTAL STUDY

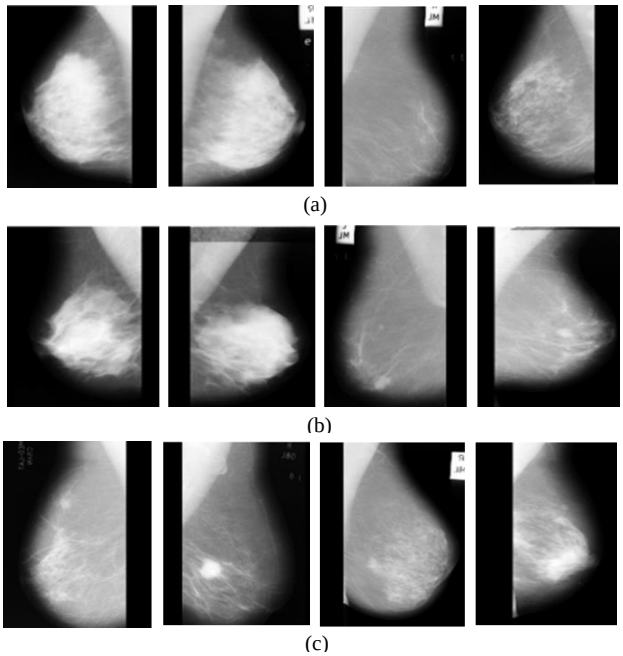
This section presents the validation of the proposed method, MoEffNet, utilising three distinct datasets: MIAS (Mammographic Image Analysis Society database), CBIS-DDSM (Curated Breast Imaging Subset of the Digital Database for Screening Mammography), and INbreast. Each dataset was chosen to highlight different characteristics and capabilities of our methodology under varying conditions. Detailed descriptions of the datasets and the experimental setup are provided below:

A. DATASETS

Three diverse datasets, each with unique strengths, form the foundation of our study. These datasets encompass a wide range of mammographic cases with high-quality annotations, serving as the critical training and testing ground for our novel breast cancer diagnosis method MoEffNet.

a) Mammographic Image Analysis Society (MIAS) Database:

The MIAS database is one of the oldest and most widely utilized mammographic databases in breast cancer research. Developed by a consortium of UK-based research groups, MIAS provides a platform for the evaluation of computer-aided diagnosis (CAD) systems. The database includes 322 images across 161 cases, each annotated with details of the lesions present, including location and type. This dataset comprises 68 benign, 151 malignant, and 203 normal images, each with a resolution of 1024 x 1024 pixels in PGM format. The database exclusively contains MLO views from both the left and right breasts. The mammographic images in this collection were digitised using a high-precision scanning micro densitometer, achieving a resolution of $50\mu m \; x \; 50\mu m$ with each pixel represented at an 8-bit depth. For enhanced usability, the original MIAS database images have been downscaled to a resolution of $200\mu m$ per pixel, as documented in references [70, 71]. This database has been instrumental in developing and validating algorithms for detecting and diagnosing breast cancer, ensuring consistency



(a) (b) (c)

FIGURE 4. **Representative Samples from the MIAS Mammogram Dataset: Illustrating (a) Normal, (b) Benign, and (c) Malignant Cases Used for Training and Evaluation.**

across different studies. Figure 4 presents examples from the MIAS dataset.

b) Curated Breast Imaging Subset of DDSM (CBIS-DDSM)

The CBIS-DDSM is a refined and standardized version of the Digital Database for Screening Mammography (DDSM) [10, 72]. It includes full-field digital mammographic images formatted in DICOM, offering decompressed images, data selection by trained mammographers, updated mass segmentation, and formatted similarly to modern computer vision data sets. This dataset includes 1,644 cases divided into four main categories: Benign Calcification, Benign Mass, Malignant Calcification, and Malignant Mass, with 753 cases of calcifications and 891 instances of masses. Our analysis focused exclusively on mass cases, comprising a benign training set of 355 cases, a benign testing set of 117 cases, a malignant training set of 336 cases, and a malignant testing set of 83 cases. Cases containing calcifications have been delayed for future investigation to allow a focused study on masses in the current research phase. The CBIS-DDSM is critical for our research due to its high-quality images and diverse case presentations, enabling the development of robust CAD algorithms. Figure 5 shows examples from the CBIS-DDSM mammogram dataset.

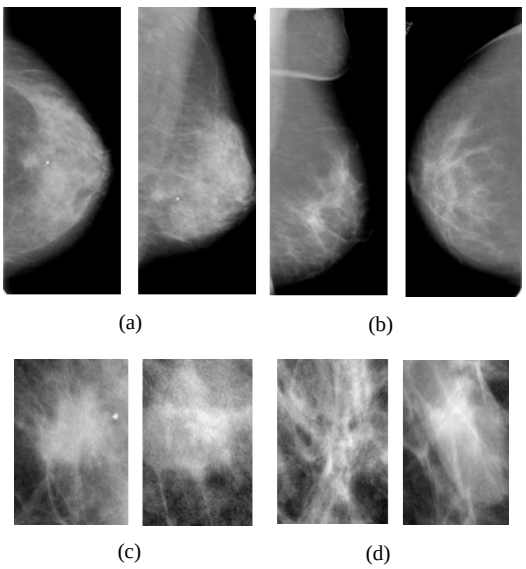


FIGURE 5. Sample Images from the CBIS-DDSM Dataset: Comparison of Full and Cropped Mammograms for Malignant (a, c) and Benign (b, d) Masses.

c) INbreast Database

The INbreast dataset is a more recent addition to the available mammographic databases and is noted for its high-resolution full-field digital mammography images [73]. Compiled at the Centro Hospitalar de S. João in Porto, Portugal, the dataset includes 410 images from 115 cases. It features a variety of mammographic findings such as masses, calcifications, and

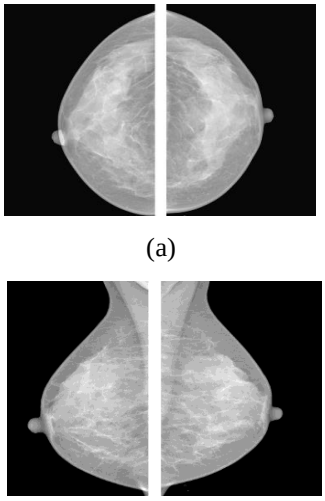


FIGURE 6. Illustrative Images from the INbreast Dataset: (a) Craniocaudal (CC) and (b) Mediolateral Oblique (MLO) Views of Both Breasts Demonstrating the Diversity of Mammographic Presentations in the Dataset.

architectural distortions, all annotated with calcifications, asymmetries, and distortions. Specialists accurately outline each lesion, and the annotations are provided in XML format, enhancing the dataset's utility for precise algorithm development. This level of detail supports advanced algorithmic development and validation, particularly in the accurate detection and classification of subtle mammographic features. The INbreast dataset's comprehensive and detailed annotations make it an invaluable resource for enhancing the diagnostic accuracy of CAD systems. Figure 6 presents illustrative images from the INbreast dataset.

B. DATA PROCESSING

In the pre-processing stage of our study, we employed various data augmentation techniques to enhance the robustness and generalizability of our deep learning model for breast cancer classification. Given the sensitive variations in mammogram appearance, which can arise from patient positioning and acquisition angles, we introduce controlled variability in the training data to enable the model to recognise features consistently across different images. Specifically, all images are resized to a uniform resolution of 224x224 pixels to ensure that the Convolutional Neural Network (CNN) processes standardized inputs. This standardization allows the network to effectively apply its filters and kernels during feature extraction. To resize the original mammogram images, we employed TensorFlow's image resizing capabilities. The process involved the following key steps:

(a) Aspect Ratio Preservation: We ensured the original proportions of the mammogram were maintained during the resizing process to prevent any distortion of the image content.

(b) Interpolation Method: Bilinear interpolation was chosen as the resizing method. This approach calculates the new pixel values by taking a weighted average of the four nearest pixels in the original image, resulting in smooth transitions and preserving the image quality, which is crucial for retaining the critical features in medical images.

(c) Padding to Match Target Size: After resizing the images while maintaining their aspect ratio, the resulting images might not exactly match the target size of 224x224 pixels. To address this, we ensured the final output image was exactly 224x224 pixels by symmetrically adding padding if the resized image was smaller than the target size or cropping if it exceeded the target size. Zero-padding was used, adding black pixels to the borders to avoid modifying the original image content.

Figure 7 illustrates the mammogram image resizing process employed to ensure compatibility with the EfficientNet architecture while preserving critical diagnostic details.

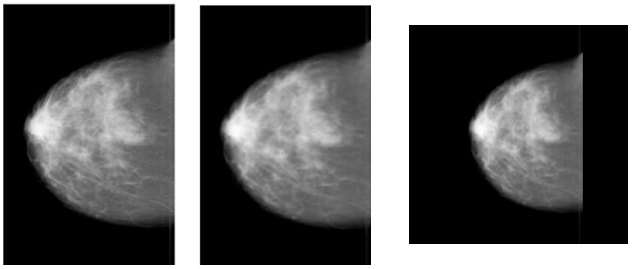


FIGURE 7. Resizing Process for Mammogram Images: (a) Original, (b) Aspect Ratio Preserved, and (c) Final Resized 224x224 Image.

To enhance the robustness and applicability of the datasets employed in this study, we have implemented a series of geometric transformations that simulate the realistic variations often observed during the acquisition of mammograms. These variations include rotations at subtle angles of -10, 0, and 10 degrees to emulate the slight misalignments that can occur during patient positioning. Additionally, we applied translations with pixel offsets of (-11, 0, 11) along the horizontal and vertical axes, reflecting the potential shifts in imaging due to patient movement or equipment handling. Furthermore, we adjusted the scale of the images by factors of 0.9, 1.0, and 1.1, which accounts for the natural fluctuations in image size due to variations in the distance between the imaging device and the breast tissue. Additional augmentations include horizontal flipping to account for laterality differences in lesions and minor adjustments to contrast and brightness (alpha=1.1 and beta=10 respectively) to aid the model in learning features less sensitive to brightness variations. These transformations were applied only to the training sets to enhance the diversity of the training data, improving the model's ability to generalise. The testing sets were intentionally left unaltered to provide a consistent and reliable metric for evaluating model performance. We tested and refined each transformation through multiple trials, achieving high-performance settings. This careful testing and validation process ensures our datasets are comprehensive and effective for training models that produce reliable diagnostic results. Through these methods, we expand the training dataset, providing a broader range of patterns for the CNN to learn from, which is crucial for improving performance on unseen mammograms and addressing the challenge of limited data availability in medical imaging.

C. ASSIGNING FEATURES TO EXPERTS

In this study, we utilised the MoEffNet architecture to extract and assign features from an EfficientNet model to enhance performance across varied complexities of image data. EfficientNet models are structured with a consistent number of building blocks across different variants, but the assembly and information of the blocks can vary slightly depending on the version. Generally, the EfficientNet architecture consists of 7 blocks, followed by a top layer. Each of these blocks can have varying numbers of individual layers and might use different scaling parameters in terms of width, depth, and resolution in the different versions (B0 – B7). The EfficientNet architecture series (B0 to B7) employs a systematic scaling approach using a compound method to adjust network depth, width, and resolution based on a set of fixed scaling coefficients. This scaling is directed by the formula $d = \alpha^{\phi}, w = \beta^{\phi}, r = \gamma^{\phi}$, where d is the depth, w denotes width, r represents resolution, and ϕ is a compound coefficient that increases progressively across the models from B0 to B7. Each subsequent model in the series increases in complexity and capability, with more layers.

We evaluated four EfficientNet models including EfficientNet B0, B1, B2, and B4 in our experiments. Our approach involves extracting features from multiple layers of the EfficientNet—early, middle, and late—which capture low, mid, and high-level features respectively. These are then systematically assigned to the experts within our MoEffNet ensemble. Depending on the experiment, our ensemble included configurations with two, three, and four experts, each designed to handle different types of features based on their complexity. This method allows for detailed and adaptive handling of image features, promoting understanding of image complexities among the experts. Table II summarises how features from an EfficientNet model are assigned to different experts in the MoEffNet system based on the number of experts used. For visualisation purposes, figure 8 displays the first 16 feature maps extracted from Blocks 3, 5, and 7 of the EfficientNet B0 model. These feature maps are arranged in a 4x4 grid and visualised using the 'jet' colour map to enhance contrast and detail. The images used for this visualisation are samples of mammogram images from the CBIS-DDSM dataset, specifically selected to represent cases with malignant and benign conditions.

TABLE II
DISTRIBUTION OF EFFICIENTNET FEATURES ACROSS MOEFFNET EXPERTS: DETAILING LOW, MID, AND HIGH-LEVEL FEATURE ALLOCATION

Number of Experts	Expert1	Expert2	Expert3	Expert4
Two Experts	Low and Mid-level features (Block 3, 5)	Mid and High-level features (Blocks 5 and 7)	-	-
Three Experts	Low-level features (Block 3)	Mid-level features (Block 5)	High-level features (Block 7)	-
Four Experts	Low-level features (Block 3)	Mid-level features (Block 5)	High-level features (Block 7)	High-level features (Block 7)

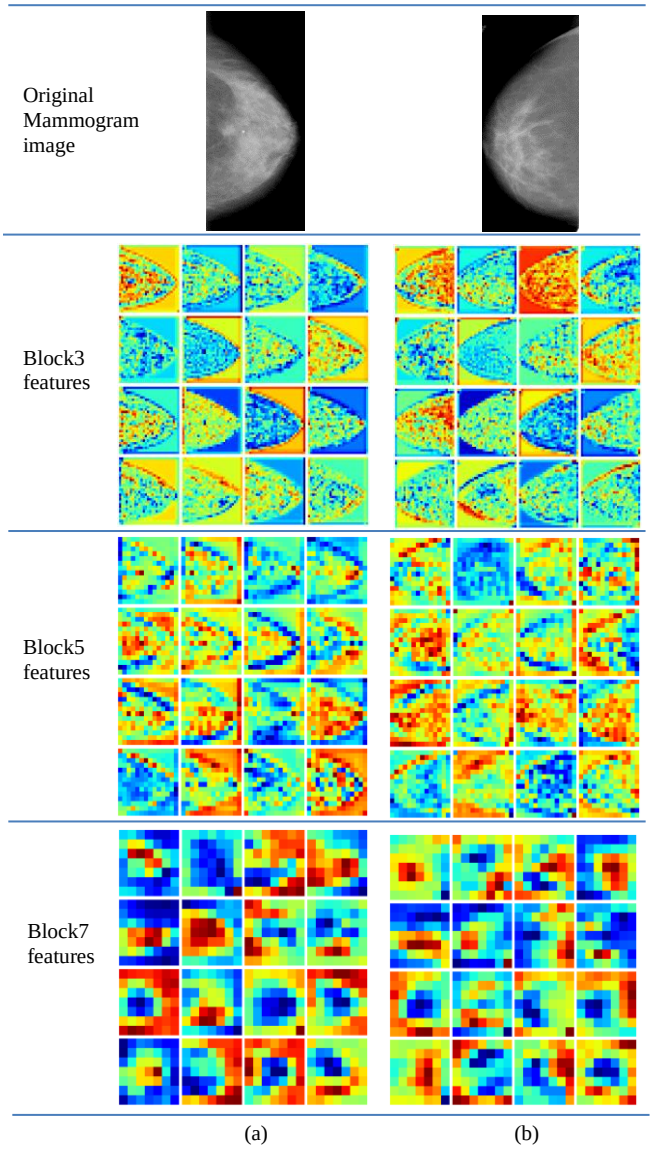


FIGURE 8. Visualisation of Multi-Level Feature Maps Extracted by EfficientNetB0: Detailed Views from Blocks 3, 5, and 7 for (a) Malignant and (b) Benign Mammographic Conditions.

D. MODEL TRAINING AND EVALUATION CONFIGURATION

In the training configuration of the machine learning models in MoEffNet, the Stochastic Gradient Descent (SGD) optimiser is used with an initial learning rate of 0.001, chosen for its straightforward approach to navigating the optimisation landscape. Training involves batches of 32 to optimise computation speed and memory utilisation over up to 100 epochs, allowing the network sufficient learning time. Key callbacks include Early Stopping, which halts training if there's no improvement in validation loss after 10 epochs to prevent overfitting, and ReduceLROnPlateau, which decreases the learning rate if no progress is seen after 3 epochs, helping the model navigate potential local minima more effectively [74]. Furthermore, the model's reliability and consistency are evaluated over 20 separate experiment runs, using accuracy, specificity, precision, recall, F1-score, area under the curve (AUC) of receiver operating characteristics (ROC), and Distance from the Ideal Position (DIP) to provide a detailed assessment of its classification capabilities across different training and testing cycles. Table III displays the hyperparameter settings and the options selected for training across the various scenarios in our experiments.

TABLE III
DETAILED CONFIGURATION OF TRAINING AND EVALUATION PARAMETERS FOR MOEFFNET ACROSS DIFFERENT DATASETS

	Hyperparameter/Dataset	Value
Classes	MIAS	Normal/Benign/Malignant
	CBIS-DDSM	Benign/Malignant
	INbreast	Benign/Malignant
Training Configuration	Optimizer	SGD
	Learning Rate	0.001
Batch Size		32
Epochs		100
Callbacks	Early Stopping Patience	10
	Restore Best Weights	True
	ReduceLROnPlateau Factor	0.2
	ReduceLROnPlateau Patience	3
Evaluation Metrics	Runs for Evaluation	20
	Metrics	Accuracy, Precision, Recall, Specificity, F1-Score, AUC, and DIP

E. EVALUATION METRICS

As described above, we employed several validation metrics to examine the efficacy of our method, including accuracy, specificity, precision, recall, F1-score, distance from the ideal position (DIP). and the area under the curve (AUC) of the receiver operating characteristics (ROC). To understand these metrics, it is essential first to define the components they are calculated from: TP (True Positives), TN (True Negatives), FP

(False Positives), and FN (False Negatives), which are described as follows:

- TP: True positives are the correctly predicted positive cases, which means that the actual class of the data point was positive, and the predicted class is also positive.
- FN: False negatives are the cases where the actual class is positive, but the predicted class is negative.
- TN: True negatives are the correctly predicted negative cases, which means that the actual class of the data point was negative, and the predicted class is also negative.
- FP: False positives are the cases where the actual class is negative, but the predicted class is positive.

a) **Accuracy:** Accuracy measures the overall accuracy of the model by computing the ratio of correct predictions (both TP and TN) to the total number of predictions made such that,

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

b) **Specificity:** Specificity measures the proportion of actual negatives that are correctly identified, which indicates the model's ability to identify negative outcomes. It can be computed using the following equation:

$$Specificity = \frac{TN}{TN+FP} \quad (7)$$

c) **Precision:** Precision represents the accuracy of positive predictions, indicating the ratio of positive identifications that were correct such that,

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

d) **Recall or Sensitivity:** Recall measures the percentage of real positives that are correctly identified, highlighting the model's ability to observe all relevant cases such that

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

e) **F1-score:** The F1-score is a harmonic mean of precision and recall that can be calculated using the following equation:

$$F1-score = 2\,x\ \frac{Precision\ x\ Recall}{Precision+Recall} \quad (10)$$

f) **Area under the curve (AUC):** The AUC is the area under the Receiver Operating Characteristics (ROC) curve, which plots the true positive rate (Recall) against the false positive rate. The AUC measures the entire two-dimensional area underneath the ROC curve from (0,0) to (1,1). A higher AUC indicates a better performing model, capable of correctly classifying positive and negative cases with high probability.

g) **Distance from the Ideal Position (DIP):** The DIP is a performance measure used to evaluate an algorithm's quality when multiple independent and bounded metrics are considered. It can be computed by measuring the Euclidean distance from the ideal value, which is 1 for each metric for all metrics, normalising this by the number of metrics, and then transforming this distance into a score that ranges between 0 to 1. The formula for the DIP is:

$$DIP = 1 - \frac{\sqrt{\sum_{i=1}^{N}(1-m_i)^2}}{\sqrt{N}} \quad (11)$$

Here m_i is the value of the *i-th* metric, and N is the number of metrics. A higher DIP value indicates better performance, with 1 being the best possible score. DIP has been proven to be superior to F1 at the higher performance end [75].

F. RESULTS

We conducted several experiments to verify the validity of MoEffNet for breast cancer diagnosis using the aforementioned preprocessing, feature selection, and model training and evaluation configurations. These experiments utilised three mammogram image datasets: MIAS, CBIS-DDSM, and INbreast. The results for each dataset are detailed in the following subsections.

1. MIAS dataset

The MIAS dataset used in our study was originally composed of 338 images designated for the training set, with 10% reserved for validation, and 84 images designated for the testing set. After the validation split, the total number of training images was 304. The details of the augmentation process are summarised in Table IV.

TABLE IV
MIAS DATASET COMPOSITION AND AUGMENTATION DETAILS.

Technique	Parameters	Combinations	No. of Images Before	No. of Images After
Rotation	-10°, 0°, 10°	3	304	912
Translation	(-11, 0, 11) pixels (x and y)	9	912	8,208
Scaling	0.9, 1.0, 1.1	3	8,208	24,624
Horizontal Flipping	Flipped, Not Flipped	2	24,624	49,248
Contrast/ Brightness	Uniform (Alpha = 1.1, Beta = 10)	1	49,248	49,248

Before investigating the detailed results from applying MoEffNet in the MIAS dataset, it is important to understand the training progress of our model. Figure 9 shows an example of the training and validation accuracy over 30 epochs from one of our experiments. The blue line represents training accuracy, and the orange line represents validation accuracy. Both accuracies increase rapidly at first, with training accuracy steadying close to 100% and validation accuracy just below it, indicating effective learning and generalisation.

Table V presents the testing results of our investigation into the performance of MoEffNet with various EfficientNet

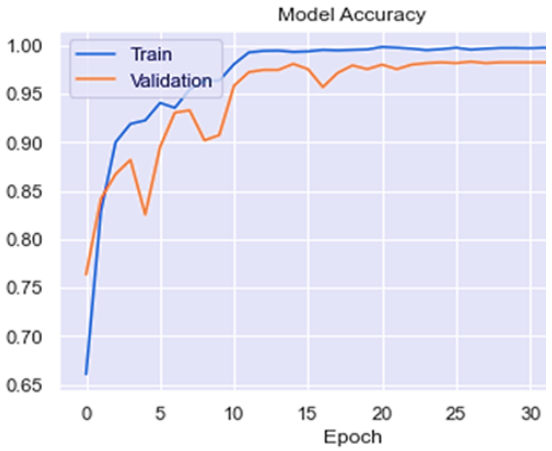


FIGURE 9. Training and Validation Accuracy Trends Over 30 Epochs: Demonstrating the Convergence and Generalisation Capabilities of the MoEffNet Model.

models and different numbers of experts for breast cancer diagnosis using the MIAS dataset. The results provide valuable insights into the relationship between model complexity, the number of experts, and diagnostic accuracy. For EfficientNet B0, accuracy peaks at 99.4% with two experts using low to high-level features, but additional experts slightly reduce accuracy. EfficientNet B1 shows a similar pattern, achieving 99.2% with two experts, but accuracy drops with three and slightly improves with four experts. EfficientNet B2 peaks at 98.8% with three experts focusing on different feature levels, while adding a fourth expert offers reduced returns. EfficientNet B4 demonstrates continuous improvement, achieving the highest accuracy of 99.2% with four experts, benefiting from a combined input across all feature levels. These findings highlight that while all models benefit from expert involvement, the optimal number of experts varies. EfficientNet B0 and B1 perform best with two experts, EfficientNet B2 with three, and EfficientNet B4 with four, highlighting the importance of both model complexity and expert input in enhancing diagnostic accuracy. Furthermore, the low standard deviations across most models suggest that the classification accuracies are stable and reliable, with EfficientNet B2 and B4 showing slight increases in variability as more experts are added but achieving higher accuracy overall.

TABLE V
DETAILED CLASSIFICATION ACCURACY AND STANDARD DEVIATION OF MOEFFNET ACROSS DIFFERENT EFFICIENTNET MODELS AND EXPERT CONFIGURATIONS ON THE MIAS DATASET

	Number of experts		
Pre-trained model	**2**	**3**	4
EfficientNet B0	99.4± 0.05	98.8± 0.04	98.0± 0.07
EfficientNet B1	99. 2± 0.07	98.0± 0.06	98.6± 0.04
EfficientNet B2	98. 6± 0.03	98.8± 0.05	98.2± 0.07
EfficientNet B4	98.1 ± 0.09	99.1 ± 0.08	99.2 ± 0.08

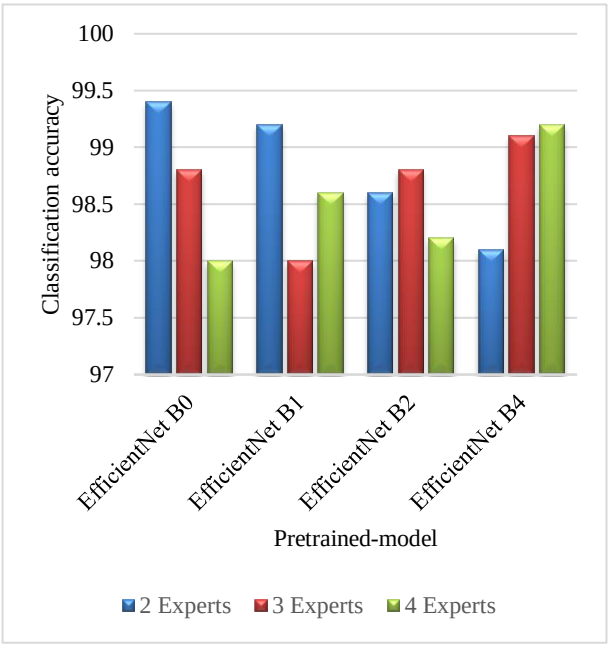


FIGURE 10. Summary of Classification Results Using MoEffNet on the MIAS Dataset: Comparison Across Different Numbers of Experts.

Figure 10 provides a clearer visualisation of the results in Table IV. The figure shows the results with two, three, and four experts involved. EfficientNet B0 and EfficientNet B1 achieve peak performance with two experts, while EfficientNet B2 performs best with three experts. EfficientNet B4 demonstrates the highest accuracy with four experts. This trend suggests that while simpler models like EfficientNet B0 and B1 benefit most from the involvement of fewer experts, the more complex EfficientNet B4 model gains significant improvements from the input of more experts.

Figure 11 presents the performance metrics (Precision, Recall, Specificity, F1 score, and DIP) for MoEffNet using various EfficientNet models (B0, B1, B2, and B4) with two experts. EfficientNet B0 and B1 demonstrate the highest precision and recall, achieving values close to 99.5% and 99.0%, respectively, indicating their effectiveness in correctly identifying positive cases and capturing the most actual positives. They also maintain high specificity around 99.5%, suggesting strong performance in correctly identifying negative cases, and achieve the highest F1 scores around 99.0%, indicating a balanced performance. EfficientNet B2 and B4, while performing well with precision, recall, specificity, and F1 scores around 98.5% and 98.0%, are slightly lower across these metrics. This suggests that for MoEffNet with two experts, EfficientNet B0 and B1 offer the best performance in terms of diagnostic accuracy and reliability.

Figure 12 displays the Receiver Operating Characteristic (ROC) curve for MoEffNet with two experts using EfficientNet B0. The ROC curve illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across various threshold settings.

The curve is plotted with the true positive rate on the y-axis and the false positive rate on the x-axis. The orange line represents the performance of the model, while the diagonal blue dashed line represents the performance of a random classifier.

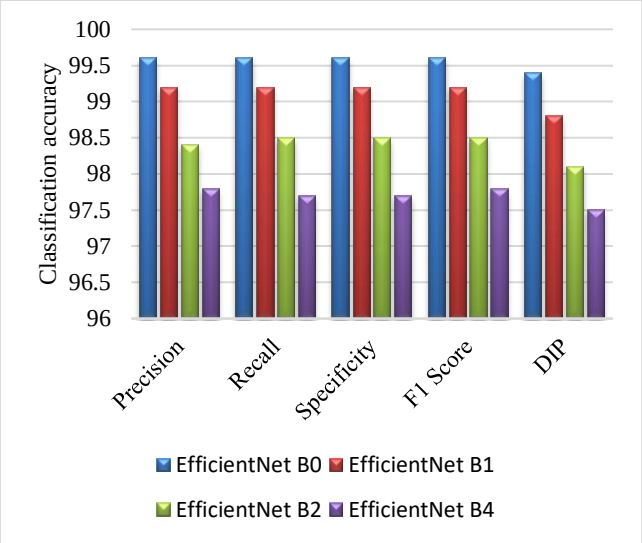


FIGURE 11. Comparison of Key Performance Metrics for MoEffNet: Evaluating Precision, Recall, Specificity, F1 Score, and DIP Across Different EfficientNet Models with Two Experts.

The Area Under the Curve (AUC) is 0.992, which indicates an excellent level of discrimination by the model. An AUC of 0.992 means that the model has a 99.2% chance of correctly distinguishing between positive and negative cases. This high AUC value demonstrates that MoEffNet with EfficientNet B0 and two experts performs exceptionally well in identifying true positives while minimising false positives, thereby confirming its effectiveness in breast cancer diagnosis using the MIAS dataset.

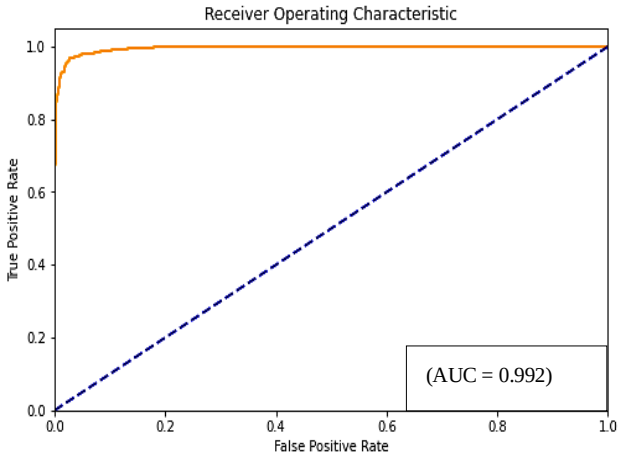


FIGURE 12. Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC) for MoEffNet: Assessing Diagnostic Accuracy with Two Experts Using EfficientNetB0.

2. CBIS-DDSM dataset

In our study, the CBIS-DDSM dataset was used, comprising a benign training set of 355 cases, a benign testing set of 117 cases, a malignant training set of 336 cases, and a malignant testing set of 83 cases. Each case includes both CC and MLO views, doubling the number of images. Thus, the dataset contained a total of 1,782 images: 710 benign training images (with 10% reserved for validation), 234 benign testing images, 672 malignant training images (with 10% reserved for validation), and 166 malignant testing images, leading to a total of 1,382 training images before augmentation. The details of the augmentation process are summarised in the following Table VI.

TABLE VI
CBIS-DDSM DATASET COMPOSITION AND AUGMENTATION DETAILS

Technique	Parameters	Combinations	No. of Images Before	No. of Images After
Rotation	-10°, 0°, 10°	3	1,382	4,146
Translation	(-11, 0, 11) pixels (x and y)	9	4,146	37,314
Scaling	0.9, 1.0, 1.1	3	37,314	111,942
Horizontal Flipping	Flipped, Not Flipped	2	111,942	223,884
Contrast/ Brightness	Uniform (Alpha = 1.1, Beta = 10)	1	223,884	223,884

Table VII presents the testing classification accuracy and standard deviation for validating MoEffNet using various EfficientNet models (B0, B1, B2, and B4) with different numbers of experts on the CBIS-DDSM dataset. EfficientNet B0 demonstrates the best performance with three experts, achieving 99.4% accuracy with a minimal standard deviation of 0.01%, indicating high stability. With two experts, it shows 99.1% accuracy with a standard deviation of 0.02%, and with four experts, the accuracy drops to 98.7% with a standard deviation of 0.04%, indicating increased inconsistency with more experts. EfficientNet B1 performs consistently well, maintaining an accuracy of 99.4% across three and four experts with standard deviations of 0.03% and 0.04%, respectively and 99.3% with two experts, demonstrating consistent performance with slightly increasing variability. EfficientNet B2 achieves the highest accuracy of 99.6% with both three and four experts, showing excellent performance and stability. With two experts, it achieves 99.4% accuracy. EfficientNet B4 shows good performance with two experts at 99.4% accuracy but exhibits a slight decrease in accuracy and increased variability with three experts at 99.3% with a standard deviation of 0.05% and four experts at 99.2% with a standard deviation of 0.06%. Taking together, these results suggest that EfficientNet B2, particularly with three or four experts, offers the best performance and stability for breast cancer diagnosis using the CBIS-DDSM dataset, validating the effectiveness of the MoEffNet method with optimal expert involvement.

TABLE VII
COMPREHENSIVE ACCURACY AND STANDARD DEVIATION ANALYSIS OF MOEFFNET USING EFFICIENTNET VARIANTS AND EXPERT CONFIGURATIONS ON THE CBIS-DDSM DATASET

	Number of experts		
Pre-trained model	**2**	**3**	**4**
EfficientNet B0	99.1± 0.02	99.4± 0.01	98.7± 0.04
EfficientNet B1	99. 3± 0.02	99.4± 0.03	99.4± 0.04
EfficientNet B2	99. 4± 0.03	99.6± 0.04	99.6± 0.02
EfficientNet B4	99.4 ± 0.05	99.3 ± 0.05	99.2 ± 0.06

Figure 13 presents the performance metrics (Precision, Recall, Specificity, F1-score, and DIP) for MoEffNet using three experts and various EfficientNet models (B0, B1, B2, and B4) on the CBIS-DDSM dataset. EfficientNet B2 consistently outperforms the other models across all metrics, achieving approximately 99.6% in Precision, Recall, Specificity, and F1-score, indicating superior performance. EfficientNet B0 and B1 also perform very well, with metrics around 99.3% to 99.4%, demonstrating high effectiveness and reliability. EfficientNet B4, while still showing strong performance with metrics around 99.2%, is slightly lower compared to the other models. These results validate the effectiveness of MoEffNet, particularly with EfficientNet B2, for breast cancer diagnosis using the CBIS-DDSM dataset.

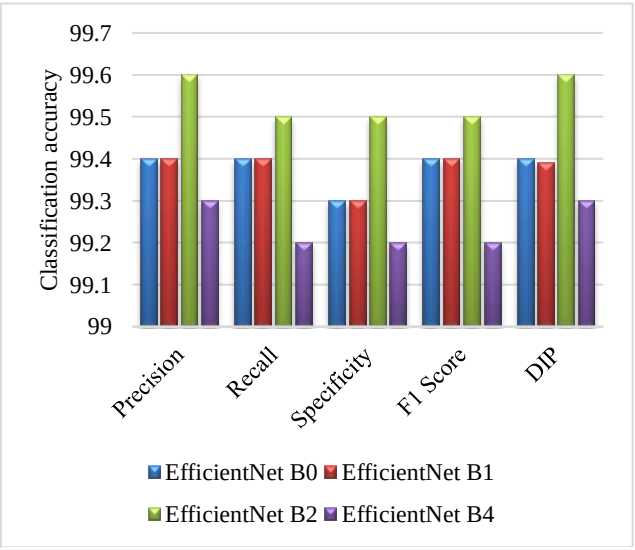


FIGURE 13. Detailed Performance Analysis of MoEffNet: Metrics Comparison Across EfficientNet Variants with Three Experts for the CBIS-DDSM Dataset.

Figure 14 shows the ROC curve for MoEffNet with three experts using EfficientNet B2. It demonstrates excellent performance with an AUC of 0.995, showing that the model is highly effective in distinguishing between positive and negative classes. An AUC of 0.995, close to the ideal value of 1.0, indicates that the model has excellent discriminatory power, performing significantly better than random guessing

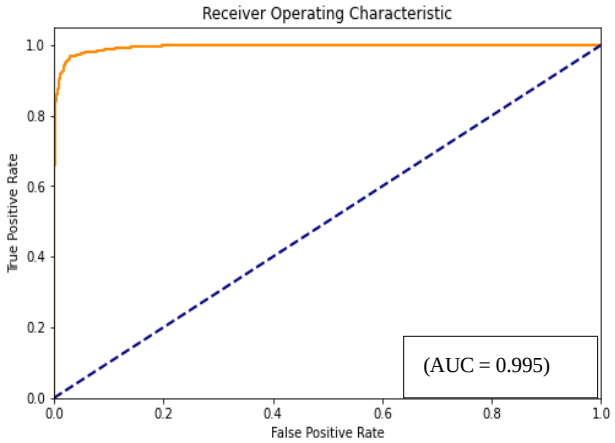


FIGURE 14. ROC Curve and AUC Analysis for MoEffNet with Three Experts: Demonstrating High Diagnostic Precision Using EfficientNetB2 on the CBIS-DDSM Dataset.

(AUC of 0.5). This high AUC value suggests that the MoEffNet model is highly accurate in its predictions, making it particularly useful for applications where precise classification is crucial, such as medical diagnostics.

3. INbreast dataset

The INbreast dataset used comprises a total of 410 mammogram images, with 287 normal cases and 123 abnormal cases. For our study, we divided the dataset into training, validation, and testing sets. The training set contains 328 images, of which 10% (33 images) were reserved for validation purposes. The testing set consists of 82 images. The details of the augmentation process are summarised in the following Table VIII.

TABLE VIII
DETAILS OF GEOMETRIC TRANSFORMATIONS APPLIED TO THE INBREAST DATASET

Technique	Parameters	Combinations	No. of Images Before	No. of Images After
Rotation	-10°, 0°, 10°	3	328	984
Translation	(-11, 0, 11) pixels (x and y)	9	984	8,856
Scaling	0.9, 1.0, 1.1	3	8,856	26,568
Horizontal Flipping	Flipped, Not Flipped	2	26,568	53,136
Contrast/ Brightness	Uniform (Alpha = 1.1, Beta = 10)	1	53,136	53,136

Table IX shows the testing classification accuracy and standard deviation for validating MoEffNet using various EfficientNet models (B0, B1, B2, and B4) with different numbers of experts on the INbreast dataset. According to the table, EfficientNet B0 shows a steady improvement in accuracy, from 99.1% with 2 experts to 99.5% with 4 experts. EfficientNet B1 also demonstrates an upward trend, achieving the highest accuracy of 99.8% with 4 experts. EfficientNet B2 demonstrates very high accuracy across all configurations, with a slight edge at 3 and 4 experts, both achieving around 99.8%. In contrast, EfficientNet B4's performance decreases

as more experts are added, from 99.4% with 2 experts to 99.1% with 4 experts, suggesting potential overfitting or increased complexity not benefiting this variant. Overall, all configurations display high accuracy, confirming MoEffNet's efficacy for this task. However, EfficientNet B2 with 3 or 4 experts stands out as the optimal choice for the highest accuracy, whereas the reducing returns with EfficientNet B4 indicate the need for careful tuning of the number of experts to avoid performance drops.

TABLE IX
PERFORMANCE EVALUATION: CLASSIFICATION ACCURACY AND STANDARD DEVIATION OF MOEFFNET ACROSS EFFICIENTNET MODELS AND EXPERT CONFIGURATIONS ON THE INBREAST DATASET

	Number of experts		
Pre-trained model	**2**	3	4
EfficientNet B0	99.1± 0.02	99.3± 0.04	99.5± 0.07
EfficientNet B1	99.5± 0.04	99.7± 0.04	99.8± 0.06
EfficientNet B2	99.7± 0.03	99.8± 0.06	99.8± 0.08
EfficientNet B4	99.4 ± 0.05	99.3 ± 0.05	99.1 ± 0.06

Figure 15 depicts the performance of the MoEffNet model on the INbreast mammogram dataset using various EfficientNet models (B0, B1, B2, B4) evaluated on Precision, Recall, Specificity, F1-score, and DIP with four experts. As demonstrated in the figure, EfficientNet B0 shows the lowest performance across all metrics, with lower precision and recall, leading to the lowest F1 score. EfficientNet B1 stands out with the highest precision, recall, and F1-score, indicating fewer false positives and effective detection of true positives. EfficientNet B2 also performs exceptionally well, closely following B1 in precision and recall while achieving the highest specificity. EfficientNet B4, although competitive,

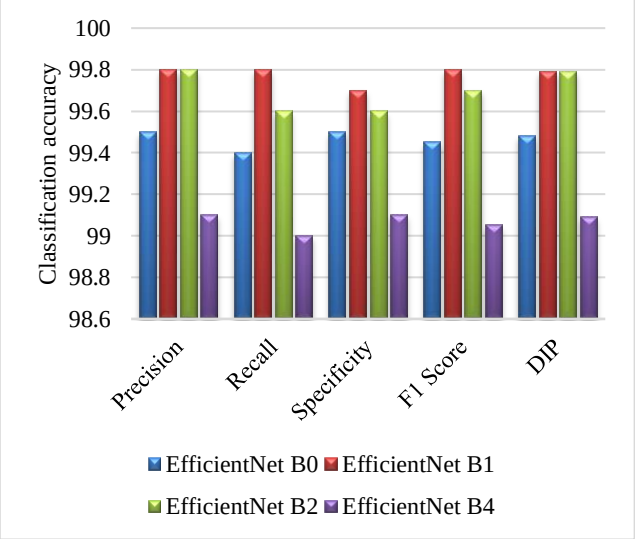


FIGURE 15. Comprehensive Performance Metrics for MoEffNet Across Different EfficientNet Models with Four Experts: Evaluation of Precision, Recall, Specificity, F1-Score, and DIP on the INbreast Dataset.

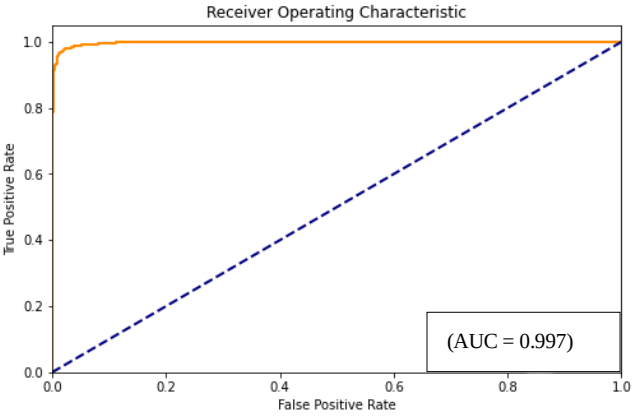


FIGURE 16. Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC) Analysis for MoEffNet: Evaluating the Diagnostic Performance with Four Experts Using EfficientNet B1.

does not surpass B1 and B2, with its metrics being lower but still better than B0. Overall, EfficientNet B1 and B2 are the best performers, making them the optimal choices for the MoEffNet model on this dataset, while B4 and B0 show relatively lower effectiveness. Figure 16 presents the ROC curve for MoEffNet with four experts using EfficientNet B1. It shows excellent performance with an AUC of 0.997, indicating that the model is highly efficient in classifying between positive and negative classes.

In summary, these results show that the proposed MoEffNet model, when applied to three distinct datasets—MIAS, CBIS-DDSM, and INbreast—demonstrates remarkable performance in breast cancer diagnosis using various EfficientNet models (B0, B1, B2, B4) and configurations with multiple experts. For the MIAS dataset, EfficientNet B0 and B1 achieve peak accuracies with two experts at 99.4% and 99.2%, respectively, while EfficientNet B2 performs best with three experts at 98.8%, and EfficientNet B4 reaches the highest accuracy of 99.2% with four experts, suggesting that simpler models benefit from fewer experts while more complex models benefit from more. On the CBIS-DDSM dataset, EfficientNet B2 with three and four experts shows the best performance with an accuracy of 99.6%, indicating its superiority in handling diverse and high-quality mammographic images. EfficientNet B0 and B1 also perform well, with slight variability in accuracy and standard deviation across different numbers of experts. For the INbreast dataset, EfficientNet B1 and B2 consistently achieve the highest accuracy, reaching up to 99.8% with three and four experts, while EfficientNet B4 shows a decrease in performance as more experts are added. Overall, EfficientNet B1 and B2 stand out across all datasets, particularly with three or four experts, making them the optimal choices for high accuracy and stability in breast cancer detection using MoEffNet.

G. DISCUSSION

The experimental evaluation of the MoEffNet model across the three well-known mammographic datasets, MIAS, CBIS-

DDSM, and INbreast, demonstrates the robustness and adaptability of the model in handling varying degrees of complexity and image quality. These datasets, each with its unique characteristics, provided a comprehensive testing ground to assess the efficacy of MoEffNet, particularly when integrated with different EfficientNet variants (B0, B1, B2, B4).

The MIAS dataset, known for its lower resolution and less complex imaging, presented a moderate challenge. Despite these challenges, MoEffNet achieved consistently high accuracy across all EfficientNet variants, with the simpler models, EfficientNet B0 and B1, performing optimally when paired with two experts. This indicates that for datasets with lower complexity, a simpler model combined with a limited number of experts can efficiently capture and analyse the necessary features, leading to accurate diagnostic outcomes. The high performance on the MIAS dataset suggests that MoEffNet is well-suited for application in environments where computational resources are limited or where the imaging data is less complex.

The CBIS-DDSM dataset, which includes high-resolution digital mammograms, allowed us to evaluate MoEffNet's performance on more complex cases, particularly in the detection of masses. The results were particularly impressive with the EfficientNet B2 model, especially when configured with three or four experts. The model achieved the highest accuracy (99.6%) with minimal variance, highlighting the importance of using a more complex model with an optimal number of experts to handle the complex details and variability present in this dataset. The consistent performance across different configurations further validates the model's adaptability to complex diagnostic tasks, making it a reliable tool for mass detection in digital mammography.

The INbreast dataset, known for its high-resolution images and detailed annotations, provided a rigorous test for MoEffNet. The EfficientNet B1 and B2 models showed superior performance, particularly when four experts were utilized. This highlights the necessity of combining a more complex model with a higher number of experts when dealing with high-resolution images that contain fine-grained details, as found in the INbreast dataset. The results suggest that MoEffNet can effectively use the detailed information available in such high-quality datasets to improve diagnostic accuracy.

Across all datasets, the number of experts was a critical factor influencing performance metrics such as accuracy, precision, recall, F1-score and DIP. Simpler models like EfficientNet B0 and B1 reached peak performance with two experts, indicating that fewer experts are sufficient to extract and analyse features effectively from less complex datasets. However, for more complex models like EfficientNet B2 and B4, increasing the number of experts to four generally led to improved performance, especially on the more challenging CBIS-DDSM and INbreast datasets. This finding emphasises the importance of making the model's complexity and the number of experts to the specific characteristics of the dataset being analysed. It is also important to note that adding too many experts, particularly in simpler models, can lead to diminishing returns or even a slight decrease in accuracy. This suggests that while multiple experts can enhance the model's ability to capture diverse features, there is an optimal number of experts beyond which the benefits start to plateau or decline.

The consistent high performance of MoEffNet across all three datasets, particularly when using the EfficientNet B2 variant, underscores the robustness of the model. Its ability to adapt to different dataset characteristics—ranging from low-resolution images in MIAS to high-resolution, complex images in INbreast—demonstrates its potential as an adaptable tool for breast cancer diagnosis. The low standard deviations observed in performance metrics across multiple experimental runs further validate the robustness of MoEffNet, indicating that the model's predictions are stable and reproducible, regardless of the specific dataset or model configuration used. This robustness is crucial in clinical settings, where consistency and reliability are paramount for effective diagnosis. Moreover, the results also reveal the exceptional discriminatory power of MoEffNet, as indicated by the high Area Under the Curve (AUC) values observed in the ROC analyses. For instance, the EfficientNet B2 model, particularly when configured with three experts, consistently achieved AUC values close to 1.0, highlighting its ability to accurately distinguish between benign and malignant cases. This level of performance is critical for clinical applications, where the ability to correctly identify positive cases while minimising false positives and negatives can significantly impact patient outcomes.

In short, the experimental results confirm that MoEffNet, particularly when combined with the EfficientNet B2 model and an optimal number of experts, offers a robust and adaptable solution for breast cancer diagnosis across diverse mammographic datasets. The model's ability to maintain high performance across datasets with varying complexity and resolution demonstrates its potential as a powerful tool in clinical diagnostics, capable of improving the accuracy and reliability of breast cancer detection in real-world settings.

H. COMPARISON OF RESULTS

To further evaluate the effectiveness of our proposed method, Table X, the comparison with some recently published results [42, 45, 55, 76-83] with the same datasets used in this study. The first left column presents the datasets used while the second column shows the reference number. Columns three, four, and five display validation metrics used in the comparison, namely, accuracy, sensitivity, specificity, and AUC respectively. In [76] a technique uses a 2D-Fourier-Bessel decomposition method to extract texture features from mammogram images, which are then enhanced using a linear regression-based feature space for improved classification of benign and malignant masses. In [77] a method involves segmenting regions of interest (ROIs) in mammograms with a

modified K-means algorithm, then applying the bidimensional empirical mode decomposition (BEMD) algorithm to derive multiple layers (BIMFs) from these ROIs, extracting texture features from these layers, and finally classifying the features using a support vector machine (SVM) classifier to automatically assess breast cancer. In [78] a methodology includes applying a two-dimensional discrete wavelet transform (2D-DWT) to extract texture features from mammogram regions of interest (ROIs), followed by feature selection using grey-level co-occurrence matrix (GLCM), and classification using a back-propagation neural network (BPNN) to differentiate between normal, benign, and malignant breast tissues.

In [45] a technique utilises a deep-learning model based on transfer learning, where features from mammogram images are extracted using pre-trained CNN architectures like VGG-16, ResNet50, and Inception V3, and then fine-tuned to classify breast cancer. In [79] a method integrates multi-feature fusions for breast mass classification by extracting complementary features (SIFT, GIST, HOG, LBP, ResNet, DenseNet, and VGG), mining cross-modal pathological semantics, and applying dynamic weight computation for mid-level fusion, followed by ensemble learning with voting strategies for final classification. In [80] a technique involves using two automated methods for breast tumour classification. The first method employs region-growing segmentation with thresholds determined by a trained artificial neural network (ANN). The second method uses cellular neural network (CNN) segmentation with parameters optimized by a genetic algorithm (GA), followed by feature extraction and classification using ANN and other classifiers. In [81] a method encompasses integrating pre-trained CNN models (such as EfficientNet) with ensemble learning using majority and soft voting strategies to classify mammogram images. In [55] an approach applying a three-stage transfer learning process using EfficientNet for breast cancer diagnosis in two-view mammography. It trains sequentially on natural images, mammogram patches, and whole mammogram views, achieving high accuracy using complementary information from both views. Reference [42] presents a method using a modified U-Net model for segmenting mammogram images, followed by classification using pre-trained CNN models (InceptionV3, DenseNet121, ResNet50, VGG16, and MobileNetV2) with transfer learning and data augmentation to enhance performance. In [82] a technique involves training a deep learning model for breast cancer diagnosis using discriminative fine-tuning, which assigns different learning rates to each layer of the deep CNN, and mixed-precision training to reduce computational demands. Data augmentation is also employed to enhance the model's performance on a small dataset, achieving rapid convergence and high accuracy. Reference [83] introduces a method using deep Convolutional Neural Networks (CNNs) with transfer learning and fine-tuning strategies to classify mammogram images, achieving high accuracy by leveraging pre-trained models like VGG16, ResNet50, and Inception v3. This approach enhances the model's ability to differentiate between benign and malignant breast lesions by optimising the networks with large datasets.

TABLE X
COMPARATIVE ANALYSIS OF MOEFFNET'S PERFORMANCE ON MULTIPLE DATASETS AGAINST STATE-OF-THE-ART DIAGNOSTIC MODELS

		Validation Metrics			
Dataset	**Ref**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**AUC**
MIAS	[76]	96.2	96.02	98.48	0.96
	[77]	98.04	98.12	98.31	0.9817
	[78]	94.2	100	90.0	0.95
	[45]	98.96	97.83	99.13	0.995
	MoEffNet (proposed method)	**99.4**	**99.2**	**99.2**	**0.992**
CBIS-DDSM	[76]	99.06	98.48	99.74	0.99
	[77]	98.62	98.60	98.65	0.9818
	[79]	90.91	82.96	98.38	0.983
	[80]	96.47	96.87	95.94	-
	[81]	96.05	-	-	-
	[55]	92.98	85.13	85.13	93.44
	[42]	98.87	98.98	-	09888
	MoEffNet (proposed method)	**99.6**	**99.5**	**99.5**	**0.995**
INbreast	[77]	98.26	97.60	98.21	0.9823
	[82]	99.8	-	-	-
	[83]	95.5	-	-	0.97
	MoEffNet (proposed method)	**99.8**	**99.8**	**99.7**	**0.997**

It is apparent from Table VII that the comparison of various methods for classifying mammograms across the MIAS, CBIS-DDSM, and INbreast datasets demonstrates the superior performance of the proposed method, MoEffNet. In the MIAS dataset, MoEffNet achieves the highest accuracy of 99.4%, significantly higher than other methods, with a sensitivity of 99.2%, specificity of 99.2%, and an AUC of 0.992, indicating its excellent ability to identify both positive and negative cases. Remarkably, the technique introduced in [45] achieves competitive results with an accuracy of 98.96% and an AUC of 0.995. Similarly, in the CBIS-DDSM dataset, MoEffNet outperforms other methods with an accuracy of 99.6%, sensitivity of 99.5%, specificity of 99.5%, and an AUC of 0.995, demonstrating its robustness and high precision in detecting breast cancer. Competitive methods include [76],

with an accuracy of 99.06% and an AUC of 0.99, and [42], with an accuracy of 98.87% and an AUC of 0.9888. In the INbreast dataset, MoEffNet achieves the highest recorded accuracy of 99.8%, along with a sensitivity of 99.8%, specificity of 99.7%, and an AUC of 0.997, surpassing all other methods. However, the technique introduced in [82] achieves competitive accuracy at 99.8%. These results highlight MoEffNet's consistent and outstanding performance across different datasets, making it an effective method for breast cancer diagnosis in mammograms compared to existing techniques.

V. CONCLUSION

In this study, we have introduced MoEffNet, an advanced integration of EfficientNet and MoEs, coupled with its scalability, dynamic gating mechanism, and expert network specialisation for high-performance breast cancer diagnosis. Using EfficientNet's advanced feature extraction and MoEs' adaptive specialization, MoEffNet processes features at multiple levels. The EffiGate mechanism further refines precision by dynamically weighting each expert network based on input characteristics. Our broad validation of three mammographic datasets, MIAS, CBIS-DDSM, and INbreast, demonstrated MoEffNet's outstanding performance. The proposed model achieved high diagnostic accuracy, with AUC values of 0.992 for MIAS, 0.995 for CBIS-DDSM, and 0.997 for INbreast, outperforming existing methods. These results highlight the effectiveness of integrating MoEs with EfficientNet, showing that EfficientNet B1 and B2 models, particularly with three or four experts, offer the highest accuracy across all datasets.

In conclusion, MoEffNet with its characteristics has proven high performance in early breast cancer diagnosis using mammographic images. These not only differentiate MoEffNet from existing methods but also establish it as a new benchmark in automated mammogram analysis. Future work will focus on expanding the diversity of datasets and MoEs, and on further optimising adaptive learning models to enhance MoEffNet's diagnostic capabilities even further.

REFERENCES

[1] Breast cancer: World Health Organisation (WHO) webpage [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer, Accessed on: April. 4, 2024

[2] Cancer: World Health Organisation (WHO) webpage [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer, Accessed on: April. 4, 2024

[3] D. R. Youlden, S. M. Cramb, N. A. Dunn, J. M. Müller, C. M. Pyke, and P. D. Baade, ''The descriptive epidemiology of female breast cancer: An international comparison of screening, incidence, survival and mortality,'' Cancer Epidemiol., vol. 36, no. 3, 2012, pp. 237–248.

[4] M. B. Amin et al., "The eighth edition ajcc cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging," CA: a cancer journal for clinicians, vol. 67, no. 2, 2017, pp. 93–99.

[5] G. N. Sharma, R. Dave, J. Sanadya, P. Sharma, and K. Sharma, "Various types and management of breast cancer: An overview," J. Adv. Pharm. Technol. Res., vol. 1, no. 2, 2010, pp. 109–126.

[6] M. Akram, M. Iqbal, M. Daniyal and A. U. Khan, "Awareness and current knowledge of breast cancer." Biological research 50, 2017, pp. 1-23.

[7] L.F. Ellison and N. Saint-Jacques "Five-year cancer survival by stage at diagnosis in Canada. Health reports", 34(1), 2023, pp.3-15.

[8] A.N. Giaquinto, H. Sung, K.D. Miller, J.L. Kramer, L.A. Newman, A. Minihan, A. Jemal, and R.L. Siegel, "Breast cancer statistics, 2022. CA": a cancer journal for clinicians, 72(6), 2022, pp.524-541.

[9] D. J. van Der Meer, I. Kramer, M. C. van Maaren, P. J. van Diest, S. C Linn, J.H. Maduro, ... & A. C. Voogd, "Comprehensive trends in incidence, treatment, survival and mortality of first primary invasive breast cancer stratified by age, stage and receptor subtype in the Netherlands between 1989 and 2017." International journal of cancer, 148(9), 2021, pp.2289-2303.

[10] M. Heath, K. Bowyer, D. Kopans, P. Kegelmeyer, R. Moore, K. Chang, S. Munishkumaran, "Current status of the digital database for screening mammography." Digit. Mammogr.: Nijmegen 1998, pp. 457–460.

[11] R. Lavayssière, A.E. Cabée, J.E. Filmont, "Positron emission tomography (PET) and breast cancer in clinical practice." Eur. J. Radiol. 1 69(1), 2009, pp. 50–58.

[12] M. Moghbel, S. Mashohor, "A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection." Artif. Intell. Rev. 39, 2013, pp. 305–313.

[13] M. Veta, J.P. Pluim, P.J Van Diest, M.A Viergever, "Breast cancer histopathology image analysis: A review." IEEE Trans. Biomed. Eng. 30 61(5), 2014, pp. 1400–1411.

[14] C.H Lee, D.D Dershaw, D. Kopans, P. Evans, B. Monsees, D. Monticciolo, R.J Brenner, L. Bassett, W. Berg, S. Feig, E. Hendrick, "Breast cancer screening with imaging: Recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer." J. Am. Coll. Radiol. 1 7(1), 2010, pp. 18–27.

[15] C.H. Chang, J.L Sibala, S.L Fritz, S.J Dwyer, A.W Templeton, F. Lin, W.R Jewell, "Computed tomography in detection and diagnosis of breast cancer." Cancer 46(4 Suppl), 1980, pp. 939–946.

[16] C.K. Kuhl, S. Schrading, K. Strobel, H.H. Schild, R.D Hilgers, H.B Bieling, "Abbreviated breast magnetic resonance imaging (MRI): First postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI." J. Clin. Oncol. 1 32(22), 2014, pp. 2304–2310.

[17] M. Zeeshan, B. Salam, Q.S.B Khalid, S. Alam, and R. Sayani. " Diagnostic accuracy of digital mammography in the detection of breast cancer". *Cureus*, *10*(4), 2018, pp. 1-10.

[18] W.A. Berg, L. Gutierrez, M.S NessAiver, W.B. Carter, M. Bhargavan, R.S Lewis, and O.B Ioffe. "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer". *radiology*, *233*(3), 2004, pp.830-849.

[19] H O A Ahmed and A K Nandi, "Colour clustering and deep transfer learning techniques for breast cancer detection using mammography images", In: Strumiłło, P., Klepaczko, A., Strzelecki, M., Bociąga, D. (eds) The Latest Developments and Challenges in Biomedical Engineering. PCBEE 2023. Lecture Notes in Networks and Systems, vol 746. Springer, Cham. 2023, pp. 105 – 119, DOI: 10.1007/978-3-031-38430-1_9.

[20] R M Rangayyan, T M Nguyen, F J Ayres, and A K Nandi, "Effect of pixel resolution on texture features of breast masses in mammograms", Journal of Digital Imaging, vol. 23, no. 5, 2010, pp. 547-553.

[21] A R Dominguez and A K Nandi, "Development of tolerant features for characterization of masses in mammograms", Computers in Biology and Medicine, vol. 39, no. 8, 2009, pp. 678-688.

[22] A Rojas and A K Nandi, "Toward breast cancer diagnosis based on automated segmentation of masses in mammograms", Pattern Recognition, vol. 42, no. 6, 2009, pp. 1138-1148.

[23] A Rojas and A K Nandi, "Detection of masses in mammograms via statistically-based enhancement, multilevel-thresholding segmentation, and region selection", Computerized Medical Imaging and Graphics, vol. 32, no. 4, 2008, pp. 304-315.

[24] T Mu, A K Nandi, and R M Rangayyan, "Analysis of breast tumors in mammograms using the pairwise Rayleigh quotient classifier", Journal of Electronic Imaging, vol. 16, no. 4, 2007, doi: 043004:1-11.

[25] A Rojas and A K Nandi, "Improved dynamic-programming-based algorithms for segmentation of masses in mammograms", Medical Physics, vol. 34, no. 11, 2007, pp. 4256-4269.

[26] R J Nandi, A K Nandi, R M Rangayyan and D Scutt, "Classification of breast masses in mammograms using genetic programming and feature selection", Medical and Biological Engineering and Computing, vol. 44, no. 8, 2006, pp. 683-694.

[27] T Mu, A K Nandi, and R M Rangayyan, "Classification of breast masses via nonlinear transformation of features based on a kernel matrix", Medical and Biological Engineering and Computing, vol. 45, no. 8, 2007, pp. 769-780.

[28] Y. Faridah. "Digital versus screen film mammography: a clinical comparison." *Biomedical imaging and intervention journal*, *4*(4), 2008. pp. 1-6.

[29] P. Singh, B. Shahi, S. Paudyal, P. Sayami, and P.R. Neupane. "Risk Factors of Breast Cancer among Patients from Lower Middle Income Country; A Case-Control Study from Nepal", MAR Oncology & Hematology 4: 01, 2024, pp. 1-18.

[30] M.G. Marmot, D.G. Altman, D.A Cameron, J.A Dewar, S.G Thompson, M. and Wilcox, " The benefits and harms of breast cancer screening: an independent review." *British journal of cancer*, *108*(11),2013, pp.2205-2240.

[31] M. Madani, M.M. Behzadi, S. Nabavi, "The role of deep learning in advancing breast cancer detection using different imaging modalities: A systematic review". Cancers 29 14(21), 2022, pp. 5334.

[32] K. Freeman, J. Geppert, C. Stinton, D. Todkill, S. Johnson, A. Clarke, S. Taylor-Phillips. "Use of artificial intelligence for image analys is in breast cancer screening programmes: Systematic review of test accuracy." Bmj, 2021, pp. 374.

[33] A. Gangwal, R.K. and Gautam, "Artificial Intelligence-Driven Decisions in Breast Cancer Diagnosis." *Drug and Therapy Development for Triple Negative Breast Cancer*, 2023, pp.131-151.

[34] N.I. Yassin, S. Omran, E.M. El Houby, H. Allam. "Machine learning techniques for breast cancer computer aided diagnosis using diff erent image modalities: A systematic review." Comput. Methods Programs Biomed.156, 2018, pp. 25–45.

[35] E.J Topol, "High-performance medicine: the convergence of human and artificial intelligence." Nat. Med. 25(1), 2019, pp. 44–56.

[36] J. Tang, R.M Rangayyan, J. Xu, I. El Naqa, Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances." IEEE Trans. Inf Technol. Biomed. 13(2), 2009, pp. 236–251.

[37] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor, and R. Zwiggelaar, "Deep learning in mammography and breast histology, an overview and future trends." *Medical image analysis*, *47*, 2018, pp.45-67.

[38] A. Oliver, J. Freixenet, R. Martí, and R. Zwiggelaar, "A comparison of breast tissue classification techniques." In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II 9*, Springer Berlin Heidelberg., *2006*, pp. 872-879.

[39] S. Anand and R.A.V. Rathna, "Architectural distortion detection in mammogram using contourlet transform and texture features." *International Journal of Computer Applications*, *74*(5), 2013, pp.12-19.

[40] A. Helwan and R. Abiyev, "October. Shape and texture features for the identification of breast cancer." In *Proceedings of the world congress on engineering and computer science* , Vol. 2, 2016, pp. 19-21.

[41] Z. Jafari and E. Karami. "Breast cancer detection in mammography images: A CNN-based approach with feature selection." *Information*, *14*(7), 2023, p.410.

[42] W.M. Salama and M.H Aly. "Deep learning in mammography images segmentation and classification: Automated CNN approach." *Alexandria Engineering Journal*, *60*(5), 2021, pp.4701-4709.

[43] X. Zhang, Y. Zhang, E.Y. Han, N. Jacobs, Q Han, X. Wang, and J. Liu, " Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks". *IEEE transactions on nanobioscience*, *17*(3), 2018, pp.237-242.

[44] S.J. Malebary and A. Hashmi, " Automated breast mass classification system using deep learning and ensemble learning in digital mammogram." *IEEE Access*, *9*, 2021, pp.55312-55328.

[45] A. Saber, M. Sakr, O.M. Abo-Seida, A. Keshk, and H. Chen. " A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique." *IEEE Access*, *9*, 2021, pp.71194-71209.

[46] K. Liu, G. Kang, N. Zhang, and B. Hou, " Breast cancer classification based on fully-connected layer first convolutional neural networks." *IEEE Access*, *6*, 2018, pp.23722-23732.

[47] H. Li, D. Chen, W.H. Nailon, M.E Davies, and D.L Laurenson. "Dual convolutional neural networks for breast mass segmentation and diagnosis in mammography." *IEEE Transactions on Medical Imaging*, *41*(1), 2021, pp.3-13.

[48] F. Gao, T. Wu, J. Li, B. Zheng, L. Ruan, D. Shang, and B. Patel, "SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis." *Computerized Medical Imaging and Graphics*, *70*, 2018, pp.53-62.

[49] D. Shah, M.A.U. Khan, M. Abrar, F. Amin, B.F Alkhamees, and H. AlSalman, "Enhancing the Quality and Authenticity of Synthetic Mammogram Images for Improved Breast Cancer Detection." *IEEE Access*, *2024*, pp. 12189- 12198.

[50] T. Mahmood, T. Saba, A. Rehman, and F.S.Alamri, F.S., "Harnessing the power of radiomics and deep learning for improved breast cancer diagnosis with multiparametric breast mammography. *Expert Systems with Applications*, 2024, p.123747.

[51] A. Jouirou, A. Baâzaoui, W. Barhoumi. "Multi-view information fusion in mammograms: A comprehensive overview.", Information Fusion. 2019, 1;52, pp. 308-21.

[52] Y. Liu, F. Zhang, C. Chen, S. Wang, Y Wang, Y Yu. "Act like a radiologist: towards reliable multi-view correspondence reasoning for mammogram mass detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 1;44(10), pp. 5947-61.

[53] H.N. Khan, A.R. Shahid, B. Raza, A.H. Dar, and H. Alquhayz, "Multi-view feature fusion based four views model for mammogram classification using convolutional neural network." *IEEE Access*, *7*, 2019, pp.165724-165733.

[54] G. Carneiro, J. Nascimento, and A.P. Bradley. "Automated analysis of unregistered multi-view mammograms with deep learning". *IEEE Transactions on medical imaging*, *36*(11), 2017, pp.2355-2365.

[55] D.G. Petrini, C. Shimizu, RA. Roela, GV. Valente, MA. Folgueira, HY. Kim. "Breast cancer diagnosis in two-view mammography using end-to-end trained efficientnet-based convolutional network", *IEEE Access*, 21;10, 2022, pp. 77723-31.

[56] Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE. Adaptive mixtures of local experts. Neural computation. 1991 Mar;3(1):79-87.

[57] K. A. Lê Cao, E. Meugnier, and G.J. McLachlan,. "Integrative mixture of experts to combine clinical factors and gene markers." *Bioinformatics*, *26*(9), 2010, pp.1192-1198.

[58] S. Myoung. "Modified Mixture of Experts for the Diagnosis of Perfusion Magnetic Resonance Imaging Measures in Locally Rectal Cancer Patients. Healthcare Informatics Research". *Healthcare Informatics Research*, 19(2), 2013, pp. 130.

[59] E.D. Übeyli, "A mixture of experts network structure for breast cancer diagnosis." *Journal of medical systems*, *29*(5), 2005, pp.569-579.

[60] S. Raman, T.J. Fuchs, P.J. Wild, E. Dahl, J.M Buhmann, and V. Roth, "Infinite mixture-of-experts model for sparse survival regression with application to breast cancer." *BMC bioinformatics*, *11*, 2010, pp.1-10.

[61] S. Myoung, J.H. Chang, and K. Song. "A mixture of experts model for the diagnosis of liver cirrhosis by measuring the liver stiffness." *Healthcare Informatics Research*, *18*(1), 2012, p.29.

[62] R. Rasti, A. Mehridehnavi, H. Rabbani, and F. Hajizadeh. "Wavelet-based convolutional mixture of experts model: An application to automatic diagnosis of abnormal macula in retinal optical coherence tomography images." In *2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP)*, *IEEE*, 2017, pp. 192-196.

[63] S.Z. Mousavi Mojab, S. Shams, F. Fotouhi, and H. Soltanian-Zadeh. "EpistoNet: an ensemble of Epistocracy-optimized mixture of experts for detecting COVID-19 on chest X-ray images." *Scientific Reports*, *11*(1), 2021, p.21564.

[64] M, Tan, and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks." In *International conference on machine learning*. PMLR, 2019, pp. 6105-6114.
[65] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen. "Mobilenetv2: Inverted residuals and linear bottlenecks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.
[66] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, " Classification of remote sensing images using EfficientNet-B3 CNN model with attention." *IEEE access*, *9*, 2021, pp.14078-14094.
[67] P. McCullagh. Generalized linear models. Routledge; 2019.
[68] S. Myoung, J. H. Chang, and K. Song, "A mixture of experts model for the diagnosis of liver cirrhosis by measuring the liver stiffness." *Healthcare Informatics Research*, *18*(1), 2012, p.29.
[69] G. Wang, G. B. Giannakis, and J. Chen, "Learning ReLU networks on linearly separable data: Algorithm, optimality, and generalization." *IEEE Transactions on Signal Processing*, *67*(9), 2019, pp.2357-2370.
[70] J. Suckling. "The mammographic images analysis society digital mammogram database." In *Exerpta Medica. International Congress Series*, Vol. 1069, 1994, pp. 375-378.
[71] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, & J. Savage. "Mammographic Image Analysis Society (MIAS) database v1.21." Apollo - University of Cambridge Repository, 2015. https://doi.org/10.17863/CAM.105113
[72] R.S. Lee, F. Gimenez, A. Hoogi, K.K. Miyake, M. Gorovoy, and D.L. Rubin. "A curated mammography data set for use in computer-aided detection and diagnosis research." *Scientific data*, *4*(1), 2017, pp.1-9.
[73] I.C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M.J. Cardoso, and J.S. Cardoso, "Inbreast: toward a full-field digital mammographic database." *Academic radiology*, *19*(2), 2012, pp.236-248.
[74] A. Thakur, M. Gupta, D.K. Sinha, K.K. Mishra, V.K. Venkatesan, and S. Guluwadi, "Transformative Breast Cancer Diagnosis using CNNs with Optimized ReduceLROnPlateau and Early Stopping Enhancements." *International Journal of Computational Intelligence Systems*, *17*(1), 2024, p.14.
[75] A. K. Nandi, "From multiple independent metrics to single performance measure based on objective function." *IEEE Access*, *11*, 2023, pp.3899-3913.
[76] P.K. Chaudhary and R.B. Pachori, "Differentiation of Benign and Malignant Masses in Mammogram Using 2D-Fourier-Bessel Intrinsic Band Functions and Improved Feature Space." *IEEE Transactions on Artificial Intelligence*, 2024, pp.1 – 10
[77] A. Elmoufidi. "Deep multiple instance learning for automatic breast cancer assessment using digital mammography." *IEEE transactions on instrumentation and measurement*, *71*, 2022, pp.1-13.
[78] S. Beura, B. Majhi, and R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer." *Neurocomputing*, *154*, 2015, pp.1-14.
[79] H. Zhang, R. Wu, T. Yuan, Z. Jiang, S. Huang, J. Wu, J. Hua, Z. Niu, and D. Ji. "DE-Ada*: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions." *Information Sciences*, *539*, 2020, pp.461-486.
[80] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian. "Benign and malignant breast tumors classification based on region growing and CNN segmentation." *Expert Systems with Applications*, *42*(3), 2015, pp.990-1002.
[81] F. Azour, and A. Boukerche. "An efficient transfer and ensemble learning based computer aided breast abnormality diagnosis system." *IEEE Access*, *11*, 2022, pp.21199-21209.
[82] A.P. Adedigba, S.A. Adeshina, and A.M. Aibinu. "Performance evaluation of deep learning models on mammogram classification using small dataset." *Bioengineering*, *9*(4), 2022, p.161.
[83] H. Chougrad, H. Zouaki, and O. Alheyane. "Deep convolutional neural networks for breast cancer screening." *Computer methods and programs in biomedicine*, *157*, 2018, pp.19-30.

HOSAMELDIN AHMED received the PhD degree in Electronic and Computer Engineering from the University of Brunel London. Dr. Ahmed is a distinguished researcher and author specialising in signal and image processing, machine learning, and condition monitoring. Currently, he is focusing on advancing breast cancer diagnosis by applying deep transfer learning and image processing techniques. His recent work involves developing innovative methods for detecting breast cancer using mammography images, demonstrating his commitment to improving healthcare outcomes through technology. In addition to his healthcare-focused research, he has collaborated extensively with Prof. Asoke Nandi, contributing to numerous publications on various aspects of machine condition monitoring. He co-authored the book "*Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines*" (IEEE - John Wiley & Sons, 2020) which explores advanced methods for monitoring rotating machinery health using compressive sampling and machine learning. Furthermore, he has published widely on topics such as bearing fault diagnosis, intelligent fault diagnosis frameworks for modular multilevel converters, and internet addiction disorder detection using machine learning. His interdisciplinary research has also contributed to cultural heritage preservation through digital image inpainting and 3D visual interaction techniques. The H-index of his publications is 11and the i10-index is 13 (Google Scholar).

PROFESSOR NANDI received the PhD degree in Physics from the University of Cambridge (Trinity College). He held academic positions in several universities, including Oxford, Imperial College London, Strathclyde, and Liverpool as well as Finland Distinguished Professorship. In 2013 he moved to Brunel University London.

In 1983 Professor Nandi co-discovered the three fundamental particles known as W^+, W^- and Z^0, providing the evidence for the unification of the electromagnetic and weak forces, for which the Nobel Prize for Physics in 1984 was awarded to two of his team leaders for their decisive contributions. His current research interests lie in signal processing and machine learning, with applications to machine health monitoring, functional magnetic resonance data, gene expression data, communications, and biomedical data. He made fundamental theoretical and algorithmic contributions to many aspects of signal processing and machine learning. He has much expertise in "Big Data". Professor Nandi has authored over 650 technical publications, including 300 journal papers as well as six books, entitled *Image Segmentation: Principles, Techniques, and Applications* (Wiley, 2022), *Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines* (Wiley, 2020), *Automatic Modulation Classification: Principles, Algorithms and Applications* (Wiley, 2015), *Integrative Cluster Analysis in Bioinformatics* (Wiley, 2015), *Blind Estimation Using Higher-Order Statistics* (Springer, 1999), and *Automatic Modulation Recognition of Communications Signals* (Springer, 1996). The H-index of his publications is 91 (Google Scholar) and ERDOS number is 2.

Professor Nandi is a Fellow of the Royal Academy of Engineering and a Fellow of six other institutions including the IEEE. In 2023, he has been honoured by the Academia Europaea and the Academia Scientiarum et Artium Europaea. He has received many awards, including the IEEE Heinrich Hertz Award in 2012, the Glory of Bengal Award for his outstanding achievements in scientific research in 2010, the Water Arbitration Prize of the Institution of Mechanical Engineers in 1999, and the Mountbatten Premium of the Institution of Electrical Engineers in 1998. Professor Nandi is an IEEE Distinguished Lecturer (EMBS, 2018-2019).