IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 October 2023, accepted 19 November 2023, date of publication 21 November 2023,
date of current version 1 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3335604

RESEARCH ARTICLE

Cancer Unveiled: A Deep Dive Into Breast Tumor Detection Using Cutting-Edge Deep Learning Models

WISHAL ARSHAD[1], TEHREEM MASOOD[2], TARIQ MAHMOOD[3,4], ARFAN JAFFAR[2], FATEN S. ALAMRI[5], SAEED ALI OMER BAHAJ[6], AND AMJAD R. KHAN[3], (Senior Member, IEEE)

[1]Faculty of Software Engineering and Information Technology, University of Central Punjab, Lahore 54000, Pakistan
[2]Faculty of Computer Science and Information Technology, Superior University, Lahore 54000, Pakistan
[3]Artificial Intelligence and Data Analytics (AIDA) Laboratory, CCIS, Prince Sultan University, Riyadh 11586, Saudi Arabia
[4]Faculty of Information Sciences, University of Education, Vehari Campus, Vehari 61100, Pakistan
[5]Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdul Rahman University, Riyadh 11671, Saudi Arabia
[6]MIS Department, College of Business Administration, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

Corresponding authors: Faten S. Alamri (fsalamri@pnu.edu.sa) and Tehreem Masood (tehreem.masood@superior.edu.pk)

This research was funded by Princess Nourah bint Abdulrahman University and Researchers Supporting Project number (PNURSP2023R346), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT About 1.5 million women are diagnosed with breast cancer every year, making it the most frequent disease among women. In Pakistan, one woman in every nine has a lifetime chance of being diagnosed with breast cancer, making it the country with the highest incidence rate of breast cancer in Asia. The mortality rate from breast cancer in Pakistan was 22.7% in 2020. A lack of resources, such as competent pathologists, causes a delay in diagnosis and inadequate therapy planning, all of which contribute to a dismal survival rate. End-to-end solutions that may be implemented into computer-aided diagnostic (CAD) systems have been developed by medical professionals and researchers using domain-specific artificial intelligence (AI) technologies, most notably deep learning models, to address this critical issue. By increasing the amount of work for pathologists, these AI models may help in breast cancer detection and diagnosis. The goal of this research was to compare and contrast the effectiveness of many recent convolutional neural network (CNN) designs. Five pre-trained and fine-tuned deep CNN architectures, InceptionV3, ResNet152V2, MobileNetV2, VGG-16, and DenseNet-121, are tested to determine the best-performing model. The goal is to discover which models are preferable in terms of accuracy and effectiveness. Notably, the pre-trained InceptionV3 model outperforms the basic CNN model by 9%, with a high accuracy level of 94%. ResNet152V2 got 95% accuracy, and MobileNetV2 got 97% accuracy. The VGG-16 model outperforms the competition with a remarkable 98% accuracy rate. Following suit, the DenseNet-121 model achieves a remarkable 99% accuracy. These findings highlight the utility of deep learning models in the diagnosis of breast cancer as well as the range of model precision.

INDEX TERMS Cancer, breast cancer, histopathological images, deep learning, MobileNetV2, VGG-16, DenseNet-121, inclusive innovation.

I. INTRODUCTION

Cancer develops when this normally regulated mechanism is disrupted by genetic material alterations. Cell growth becomes unregulated and out of control. These cells can clump together to form a mass known as a tumor. Cancer is currently one of the major causes of mortality worldwide; it is a disease that spreads aggressively across cells and increases in the body minute by minute. According to the International Agency for Cancer Research (IARC), one in every five people

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

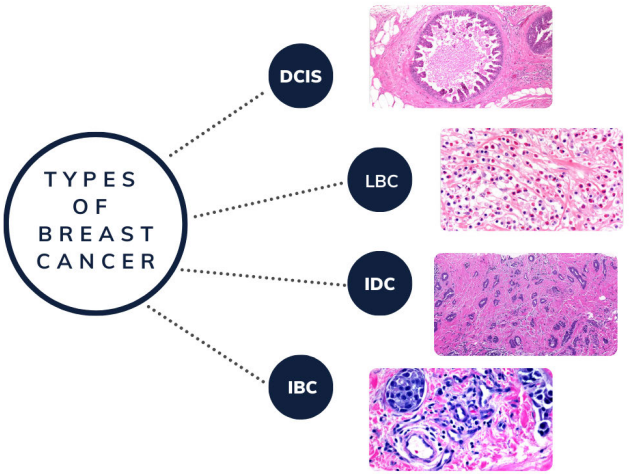


FIGURE 1. Demonstration of major types of breast cancer.

will be diagnosed with cancer at some point in their lives, and one in every eight men and one in every eleven women will die from it. Breast cancer (BC) is the main cause of death in women and is one of the most frequent cancers [1], [2]. BC accounted for around 24.5% of all cancer-related cases in women, according to statistics (GLOBO-CAN, 2020).

Breast cancer, one of the most common diseases affecting women, costs a large number of lives each year. Although cancer may be treated and cured if detected early, many individuals do not obtain a diagnosis until it is too late [3], [4]. Breast cancer starts when abnormalities in the cells become evident, and it can spread to nearby locations [5]. Breast cancer is a major global health concern, and it is becoming increasingly frequent in a variety of populations [6]. Fig. 1 shows the several forms of breast cancer [7], [8], [9]that emerge when damaged cells and tissues spread throughout the body. DCIS, also known as non-invasive carcinoma, is a kind of breast cancer that occurs when abnormal cells spread outside of the breast [10], [11]. The second form (IDC) is invasive ductal carcinoma. It is also known as infiltrative ductal carcinoma [12]. IDC cancer is most commonly found in men, and it occurs when abnormal breast cells spread throughout all breast tissues [13], [14]. The third kind of cancer is lobular breast cancer (LBC). This happens inside the lobule [15]. It increases the chances of having more aggressive cancers. Inflammatory breast cancer is the last type of breast cancer that causes swelling and redness (IBC). It is a fast-growing breast cancer that manifests itself when the lymph veins in damaged cells become blocked [16], [17]. Breast cancer diagnosis must be precise and timely in order to improve patient outcomes and survival rates. The safest and most reliable way to effectively manage BC is early detection, then effective clinical treatments [18]. Breast ultrasound, computed tomography (CT), mammography, histopathological imaging, and magnetic resonance imaging (MRI) are examples of non-invasive breast screening techniques or imaging tools available today. Mammography and ultrasound are two popular early detection modalities for breast cancer, all of which significantly increase the breast cancer survival rate by generating high-quality breast images [19], [20]. Specialists examine these images in great detail to provide the correct diagnosis. Despite major advancements in recent non-invasive imaging technology, invasive surgical imagery, which refers to a pathologist's histological examination of a breast tissue biopsy, remains the gold standard in clinical scenarios for a final breast cancer diagnosis [21].

Histopathological imaging in clinical practice relies primarily on pathologists' manual qualitative analysis. This analysis method, however, raises at least three issues [22]. First, there is a global shortage of pathologists, particularly in developing countries and small hospitals. Resource scarcity and imbalanced distribution are serious issues that must be addressed [23], [24]. Second, whether or not the histological diagnosis is right is entirely dependent on the pathologist's extensive professional expertise and long-term diagnostic experience. This pathologist's subjectivity has resulted in an increase in diagnostic discrepancies [25]. Third, the complexity of histological images causes pathologists to become fatigued and distracted. To address these issues, it is critical to create automatic and exact histopathological image analysis tools, particularly classification approaches.

Image processing techniques combined with machine learning could be a useful tool for detection and diagnosis; however, they usually result in false positives and false negatives. Deep learning technology can detect breast cancer at an early stage by lowering mammography interpretation time. There is currently a plethora of deep learning algorithms available; however, not all of them have been investigated for their utility in identifying breast cancer. These algorithms are used for fully automated mass segmentation, detection, and classification by extracting crucial distinctive features from images without the need for manual human interaction. When the number of images available to train a deep CNN (DCNN) network is insufficient, transfer learning (TL) plays a vital role in enhancing diagnosis performance in the medical arena, particularly when dealing with the intricate characteristics of breast mammography. Scholars have recently become interested in fine-tuning TL networks with pre-trained weights for solving challenging classification tasks with significant interpretation performance.

In this study, we employ a deep learning-based method for histopathological image-based breast cancer identification. The dataset employed in this study consists of a sizable number of images, each of which has been classified as either non-cancerous (NC) or invasive ductal carcinoma (IDC+) (no cancer). With 265,142 and 128,382 images, respectively, the training and validation sets were created from the dataset. All images were downsized to the same size of 50 $\times$ 50 pixels to aid in model training and retain computational effectiveness. We chose the InceptionV3 model for our initial experiment since it is a quick and effective deep-learning architecture that is great for applications with limited resources. The InceptionV3 model learns

to extract pertinent features from the histopathology images to identify between IDC+ and no cancer cases during the course of 45 training epochs. The model's performance was assessed on the test dataset following the training phase to determine how well it classified breast cancer images. The InceptionV3 model obtained 94% accuracy, ResNet152V2 obtained 95% accuracy and MobileNetV2 model obtained an excellent accuracy of 97% on the test dataset, according to the findings of our experiment. Encouraged by this achievement, we continued to investigate how well-known deep learning models like VGG-16 and DenseNet-121 performed. We carried out further tests to assess the classification performance of the VGG-16 and DenseNet-121 models against MobileNetV2. In the test dataset, the VGG-16 model achieved a remarkable 98% accuracy. Likewise, the DenseNet-121 model demonstrated excellent results as well, with a comparable accuracy of 99%.

The VGG-16 and DenseNet-121 models' high accuracy results highlight their potential for histopathology image classification of breast cancer. These models demonstrate how well they can distinguish between diseased and healthy tissues, making them useful tools for the early identification and diagnosis of cancer. This paper makes a contribution to the field of medical image analysis by providing a comparison of various deep-learning models for the identification of breast cancer. The high accuracy levels attained highlight how important deep learning is for improving breast cancer detection and treatment. These models could help pathologists diagnose breast cancer cases quickly and precisely in real-world applications in healthcare facilities, according to the research. However, in order to prove the validity and generalizability of these models for application in real-world healthcare settings, additional research on more varied and substantial datasets as well as clinical validation, would be required. The comparison of the state-of-the-art deep learning models InceptionV3, ResNet152V2, MobileNetV2, VGG-16, and DenseNet-121 for histopathological image processing in breast cancer detection is novel in this study. Medical professionals are given crucial insights for making the best model choice in a variety of clinical circumstances thanks to this comprehensive study, which highlights the distinctive strengths and capabilities of each model. The main contribution is to show how deep learning models may actually be used to diagnose breast cancer, with astounding accuracy rates of up to 96 percent. This empirical validation offers better patient care, quicker treatments, and better treatment outcomes, in addition to demonstrating the practicality of technology use. The study fills the gap between practical medical applications and cutting-edge computational techniques. This research is motivated by the need to solve the urgent problems associated with breast cancer and is inspired by the revolutionary potential of deep learning. The fusion of healthcare and technology presents a possible path to transform diagnostic paradigms. Deep learning speeds up diagnoses by automating the complex process of recognizing malignant tissues in histopathology images, enabling early interventions, and heralding a future in which cutting-edge technologies enhance human capabilities for improved healthcare outcomes.

The fundamental purpose of this study is to develop an efficient automated deep learning system to assist radiologists in accurately categorizing histopathology images. IDC+ with no cancer. The following are the research aims and methodology:

- Image artifacts and noise were removed using noise reduction and data pre-processing techniques. enhanced image quality using various data enhancement approaches.
- The optimal preprocessing algorithms and parameter settings are chosen after testing them on our dataset. Depending on the input shape, common preprocessing techniques like resizing images to 50 × 50 pixels and normalizing pixel values to [0, 1] are taken into consideration (50, 50, 3).
- Performed statistical analyses (MSE, PSNR, and RMSE) on images that had already been processed. confirmed that after preprocessing, the image quality and pixel information are retained.
- Data augmentation techniques are used with an ImageDataGenerator during the loading of training data to increase dataset diversity in order to solve overfitting problems. To prevent overfitting, the expanded dataset was divided into training, validation, and testing subsets.
- utilized InceptionV3, InceptionV3, ResNet152V2, MobileNetV2, VGG-16 and DenseNet-121 as pre-trained networks for baseline architectures. The last layers of these models were modified to make them suitable for categorizing the enhanced data.
- Models of Inception V3, ResNet 152 V2, MobileNet V2, VGG-16 and DenseNet-121 have been modified and fine-tuned, respectively. They compared their results in terms of their precision, their recall, and their F1 score.
- Used an ablation study on the chosen model to improve its classification performance. identified component points or configurations that improve performance.
- Categorical cross-entropy is used as the loss function during training to evaluate the model. In addition, accuracy is used as a measure of the model's efficacy on both the training data and the validation data. Together, these techniques help direct the model's optimization, evaluate its image classification accuracy, and mitigate overfitting concerns.

The objective of the present study was to assess whether it would be feasible to develop an automated tool to assist radiologists in classifying cancer tumors. In order to achieve effective and efficient categorization, this research compares and contrasts the performance of the InceptionV3, ResNet152V2, MobileNetV2, VGG-16, and DenseNet-121 models. The research approach is modified for each model's specifics while maintaining consistency with the overarching objective.

The paper's organization is as follows: Section II A concise overview of pertinent research on advanced techniques and innovative advances in medical image analysis is provided. Section III demonstrates the proposed method, which includes microscopic image processing, data enhancement. Section IV demonstrates the proposed architecture and transfer learning paradigm based on deep learning. Section V employing microscopic images, describe the findings of the investigation using the suggested framework. Section VI concisely explains the results and concentrates on the advancements of deep learning and its applications in microscopy image analysis. Finally, Section VII the study's conclusions are recapitulated.

II. LITERATURE REVIEW

Medical imaging, particularly mammography, is widely used in breast cancer screening, and deep learning has shown promise in detecting abnormalities in mammograms [26]. The combination of self- and poorly supervised reconstruction techniques made it easier to detect anomalies in mammograms, demonstrating how effective deep learning is at recognizing minor breast cancer signs. Deep convolutional neural networks (CNNs) are being utilized to increase image quality and provide diagnostic assistance [27]. Also, combining region-based pooling structures with CNN architectures has shown that deep learning has a lot of potential for classifying mammograms [28]. Pooling based on area increased the accuracy of breast cancer classification. Deep learning can extract useful information from mammograms, resulting in superior classification results. Deep learning has been used to diagnose breast cancer using several imaging modalities other than mammography. For instance, it has been possible to interpret contrast-enhanced ultrasound images using deep learning models powered by domain knowledge [29]. Researchers used a novel deep learning-driven approach to improve the clarity and effectiveness of breast cancer diagnosis using contrast-enhanced ultrasound recordings.

Deep learning-based breast cancer subtype classification has investigated molecular data [30]. The ''Triphasic DeepBRCA'' architecture shows how deep learning may find breast cancer indicators. This unique technology made specific treatment regimens possible. In addition to medical imaging, deep learning can analyze histology images. Multi-scale convolutional neural networks classified breast calcifications in digital mammograms [31]. An innovative 2021 study showed that deep learning can treat diverse breast cancer presentations. This study illuminated categorization accuracy and detection power. As deep learning improves breast cancer diagnosis and categorization, its role in cancer research and precision treatment must be considered [32].

Transfer learning helps in detecting breast cancer tumors [33]. The suggested hybrid transfer learning model (MVGG and ImageNet) has an accuracy of 94.3%, according to experiments. The proposed MVGG architecture alone achieves 89.8% accuracy. The proposed hybrid pre-trained network outperforms other convolutional neural networks. For appropriate diagnosis and classification, two tests were done. First, five end-to-end pre-trained and fine-tuned deep convolutional neural networks (DCNN) were tested. The ConvNet's detailed features train the support vector machine algorithm to perform well in the second experiment. Our deep learning ConvNet+SVM model had training and validation accuracy of 97.7 and 97.8%, while VGGNet16 and VGGNet19 gave us 90.2%, 93.5%, 63.4%, 82.9%, MobileNetV2, 75.1%, ResNet50, and 72.9%, respectively [34]. Another research [35] presented an enhanced DenseNet-121 neural network model for accurately classifying benign and malignant mammography images. By using an Inception structure instead of the first convolutional layer, the suggested model improves AlexNet, VGGNet, GoogleNet, and the baseline DenseNet-121. DenseNet-121 achieves a remarkable average accuracy of 94.55% using 10-fold cross-validation, considerably improving mammography image classification accuracy. Another study [36] presented an architecture based on VGG-16 and VGG-19 models for automatic breast cancer classification from histopathology images. This hybrid classifier had an accuracy of 95.29%, a sensitivity of 97.73%, and an F1 score of 95.29%. Tests were conducted [37] to extract features; various pre-trained CNN models were used, which were then integrated and evaluated using different machine-learning methods. On the RSNA dataset, which includes numerous views and extra features such as age, the neural network (NN) classifier achieves 92% accuracy. The MIAS dataset achieves 94.5% accuracy, while the DDSM dataset achieves 96% accuracy. Another experiment [38] demonstrated a new technique based on the combination of deep features and CNN. They used 400 mammography images, 200 of which were of malignant masses and 200 of which were benign masses. Classifiers such as extreme learning machines (ELM) and support vector machines (SVM) were used for classification. ELM outperformed the SVM classifier with an accuracy of 86.50%. Three pre-trained networks, VGG-16, VGG-19, and ResNet50, were used in a comparison [39] of transfer learning and the fully-trained network using histopathological imaging for the categorization of breast cancer regardless of magnification. The best model was a fine-tuned, pre-trained VGG-16 with a logistic regression classifier. It had an accuracy of 92.60%, an AUC of 95.65%, and an APS of 95.95% when 90% of the data was used for training and 10% for testing.

Computer-aided design (CAD) system [40] was built on one of the most successful object detection frameworks available: Faster R-CNN. They achieved a classification score of 95% using the VGG-16 network as the foundation of CNN. A breast lesion categorization method based on deep features was developed [41]. It was mostly a CNN with decision-making mechanisms. The fine-tuned CNN was trained on a large number of natural images. The revised model was 96.7% correct. Deep learning methods are extensively used in image classification due to their high accuracy and low misclassification rates. This study examines

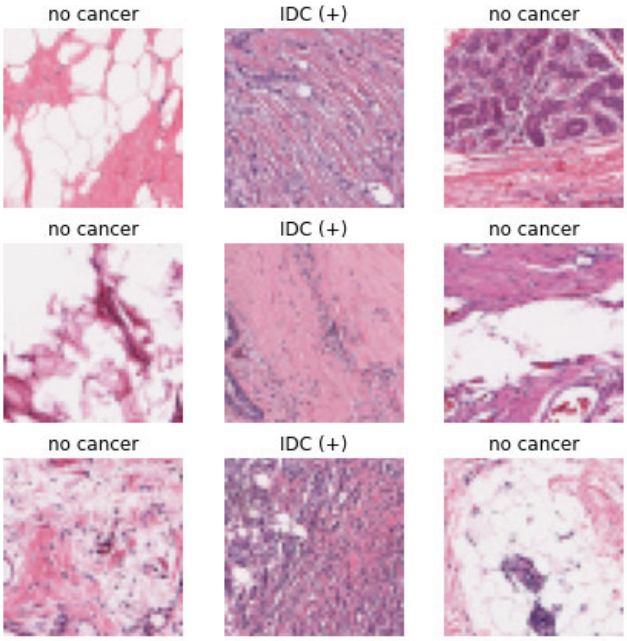


FIGURE 2. Contrasting histopathological images: IDC+ vs no cancer.

current improvements in deep learning-based breast cancer detection systems. We examine the latest methods, problems, and possibilities of deep learning algorithms in breast cancer diagnosis by combining findings from multiple studies. We synthesize these studies to advise researchers, physicians, and policymakers on breast cancer screening and therapy to improve patient outcomes. The next sections will examine breast cancer detection research's deep learning models, data sources, and assessment metrics to better understand how deep learning may affect detection and treatment.

III. MATERIALS AND METHODS

A. DATASET

This study uses the ''Breast Histopathology Images'' dataset, which is a valuable resource made available by Kaggle [47], the best online community for people interested in data science and machine learning. This data set is a key building block for improving medical image analysis, especially for diagnosing breast cancer.

Fig. 2 shows that the dataset has a large number of histopathological images that were carefully taken from samples of breast tissue.IDC+ images have shown uneven cell structures, increased cell density, and abnormal tissue patterns, whereas no-cancer images have shown regular and healthy cellular arrangements with normal tissue architecture. These images are very important for medical diagnosis because they show in great detail how breast tissues are made and how they are put together. The collection stands out because of how well it is organized. Images are put into two separate groups: those with invasive ductal carcinoma IDC+ and no cancer. This labeling method gives academics and professionals in the business world the tools they need to start training and carefully evaluating machine learning algorithms. This encourages them to reach a level of accuracy that has never been seen before in their efforts to find and classify breast cancer.

According to Table. 2, the dataset consists of 277,524 patches (50 × 50 size) with 78,786 IDC+ and 198,738 no-cancer patches that were taken from 162 breast cancer slide images. 162 whole mount slide images of breast cancer (BCa) specimens that were scanned at 40x magnification were used in the dataset in Table. 3.277,524 patches overall, each 50 × 50 pixels in size, were taken from these images for additional examination. It is important to note that the dataset was split into a training dataset with 194,266 patches and a testing dataset with 83,258 patches. While the testing dataset functioned as a separate set to assess the generalizability and performance of the trained models, the training dataset was used to train and optimize the models. The precise training and testing dataset sizes were critical in evaluating the model's accuracy, robustness, and applicability for the research goals.

B. DATA PREPROCESSING

As a first step in our study, the histological data from breast cancer were meticulously prepared for training by resizing and normalizing them. This crucial preprocessing step ensured that all images were of the same size and scale, which proved essential for model compatibility and facilitated subsequent development. We utilized TensorFlow and Keras frameworks for enhanced efficiency during the development process. Upon loading the dataset, we initiated a vital preparation phase to guarantee image quality and consistency. Each image was adjusted to a standardized size of 50 by 50 pixels, a step considered significant for expediting model training and simplifying computational processes. We worked with a (50, 50, 3) image format, incorporating red, blue, and green channels to provide convolutional neural networks (CNNs) with the necessary RGB color information for learning and feature extraction. The initial training of these models was conducted using the ImageNet collection, and batch normalization layers were introduced to enhance the stability of the training process. Subsequently, the model's output was flattened, and two additional dense layers were added to improve its representation. This refinement aimed to capture and extract nuanced patterns relevant to our specific task. In conjunction with this data preparation and model development, we also meticulously split the dataset into 50,000 random samples for training and 30,000 for testing. This approach allowed us to train the model on a substantial dataset and assess its generalization performance on a sizable test set, contributing to the robustness of our breast cancer detection model. Furthermore, we have been very careful by scaling, encoding, and filling in missing values separately for the training and test sets. This careful method protects against data leaks by keeping information from the test set from having an unintended effect on the training process. In the methodology part, we made our feature engineering

TABLE 1. Compilation of research references for in-depth understanding.

Reference	Methods	Findings	Dataset
Montaha et al.[42]	CNN	98.02%	CBIS-DDSM
Yan et al.[43]	Hybrid DNN-RNN	91.3%	Public Histo Dataset
Zou et al.[44]	AHoNet	BreakHis=99.29% BACH=85%	BreakHis and BACH
Alokkumar et al.[45]	DCNN	94.44%	Genomic Expression Dataset
Wang et al.[38]	CNN-ELM	92%	Hospital-Annotated Dataset
Malebary et al.[46]	CNN-RNN	DDSM=96% MIAS=95%	DDSM and MIAS

TABLE 2. Dataset information.

Dataset	IDC+	No of Cancer Patient	Total Patient
Breast Histopathology Images	78,786	198,738	277,524

TABLE 3. Dataset specifications.

Description	Total Images
Scanning Magnification	40x
Extracted Patches	277,524
Patch Size	50 x 50 pixels
Training Dataset Size	194,266 patches
Testing Dataset Size	83,258 patches
Total Images	162 whole mount slide images

steps clearer by saying exactly how they will be carried out separately for each split. This makes sure that any features or transformations that are drawn from the test set keep its integrity. In addition, we have put more stress on using cross-validation techniques in addition to the usual "train-test" split. This repeated process using different sets of data gives us a fuller picture of how well our model generalizes, which lowers the risk of data loss even more.

C. DATA AUGMENTATION

Data augmentation approaches play an essential role in training breast cancer detection algorithms because they increase the diversity of the dataset and reduce overfitting. Using pre-trained deep learning models, these strategies involve applying various changes to the original dataset. As a result, the model improves its ability to generalize to new settings and capture complex features, leading to increased accuracy in identifying malignant regions and improved diagnostic outcomes.We used the ImageDataGenerator's adjustable settings to carefully enhance histopathology images in our investigation. This included introducing controlled changes such as random rotations of up to 30 degrees, shifts of up to 15% in both the horizontal and vertical planes, and a maximum shear transformation of 25%. Furthermore, various magnification and rotation levels were obtained utilizing random zooming (up to 30%) and horizontal flipping. The 'nearest' fill mode was used to extend pixel values smoothly beyond the original image limitations while maintaining image integrity. The goal of these augmentation strategies was to enhance the dataset with realistic changes, which would improve the model's capacity to identify significant patterns in histopathological images. These changes in orientation, position, and distortion, as well as changes in magnification and rotation, all contribute to the model's increased performance in breast cancer diagnosis. The increased diversity of the generated dataset helps the model generalize better and minimizes the danger of overfitting, resulting in a more robust and effective breast cancer detection model.

IV. TRANSFER LEARNING-BASED PROPOSED MODELS

Transfer learning models are a key part of making it easier to find breast cancer by building on what is already known from large datasets. This method speeds up and improves the learning and diagnosing process, which makes it more accurate and reliable. In this quest, the InceptionV3, ResNet152V2, MobileNetV2, VGG-16 and DenseNet-121 transfer learning models were carefully chosen to compare. MobileNetV2, which is known for having a light and efficient design, is a great choice for situations with few resources. Its simple design makes it efficient without sacrificing accuracy, so it works well in places with limited resources. The VGG-16 deep CNN model, on the other hand, has a more complex architecture that can pick up complex data patterns. This makes it possible to understand all of the complicated parts of medical images. DenseNet-121 stands out because of its unique way of reusing features and its structure, which makes good use of parameters. It improves

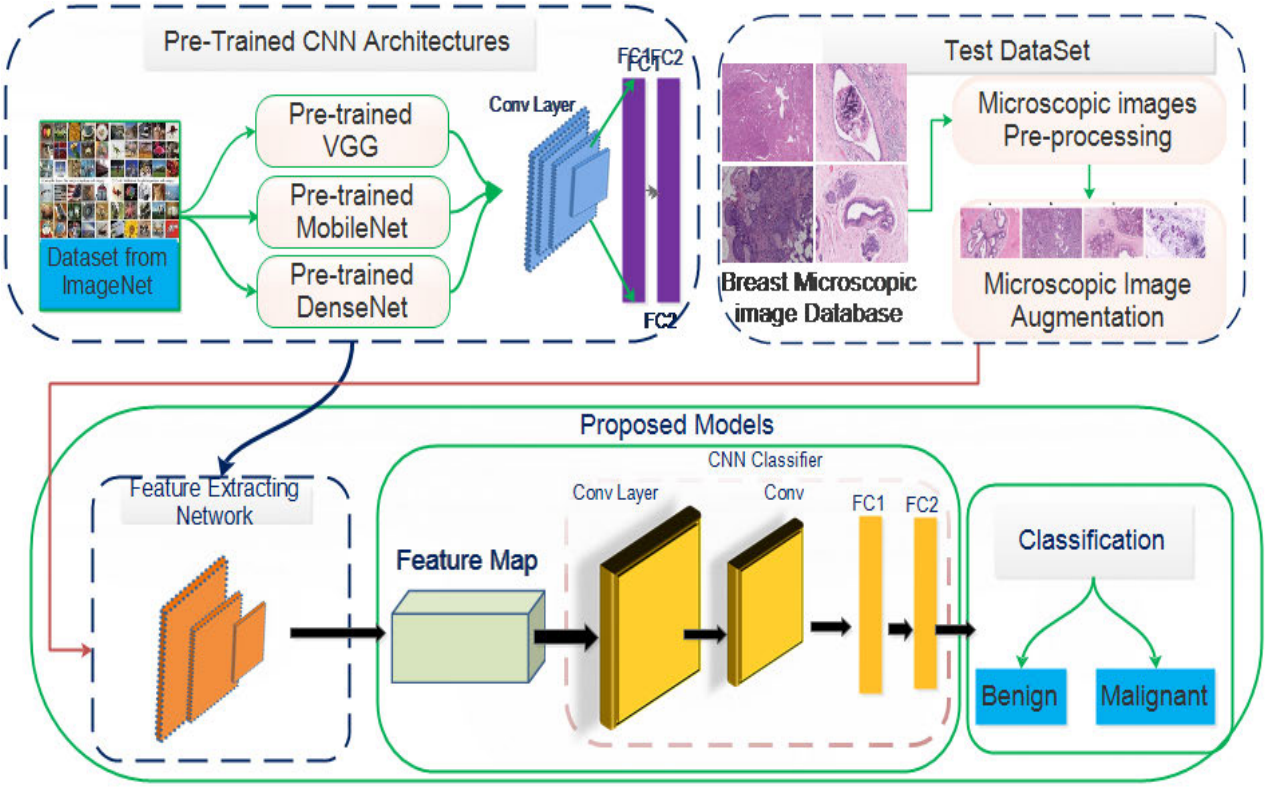


FIGURE 3. An in-depth exploration of the methodology for leveraging deep learning models.

feature extraction and saves on parameters by tightly coupling convolutional layers. This feature of the design makes sure that resources are used well, which improves the model's speed. When thinking about deep learning models for finding breast cancer, you need to know a lot about VGG-16, MobileNetV2, and DenseNet-121. The decision process is based on how they are built, how well they work, how much memory they need, and what they can be used for. The benefit of MobileNetV2's light weight is that it works well in limited environments. VGG-16's complex architecture and DenseNet-121's efficient use of parameters allow for flexible deployment. The decision depends on how well these features match the needs of the project, the resources that are available, and the diagnostic context that is planned.

A. MobileNetV2

The MobileNetV2 model helped identify IDC+ (invasive ductal carcinoma) and no cancer on histopathological images. This MobileNetV2 model identified key elements in IDC-positive images that correspond with invasive cancer. These included the complicated, irregular cell patterns and forms of malignant tissue. This model accurately identified healthy cell patterns in non-cancerous images. The MobileNetV2 model identified breast tissue samples by recognizing these patterns. The MobileNetV2 architecture's design concepts match the histopathology images' complexity. MobileNetV2 uses depth-wise separable convolutions, which divide the convolution process into two steps: a depth-wise convolution that analyzes each input channel individually and a point-wise convolution that combines the outputs. This design has fewer parameters and a lower computational load than typical convolutions, making it ideal for resource-constrained histopathology image analysis.

MobileNetV2 balances high-dimensional and low-dimensional representations with a linear bottleneck structure Fig. 4. This helps capture histopathology images' subtle patterns. The efficient and lightweight architecture and design choices allow the model to learn and discriminate detailed cellular patterns in IDC+ instances and healthy tissues. The MobileNetV2 model detects breast cancer well by recognizing these patterns. Its design deciphers key elements in histopathology images to accurately identify malignant tissues and improve patient care and medical imaging. The following images depict IDC+ and no cancer cases. The MobileNetV2 architecture, designed for mobile and embedded applications, emphasizes lightweight layers to reduce memory usage and speed inference. Thus, MobileNetV2 is appropriate for real-time breast cancer diagnosis with minimal computational resources.

B. VGG-16 MODEL

The VGG-16 deep convolutional neural network outperforms existing methods for detecting breast cancer. Its effectiveness

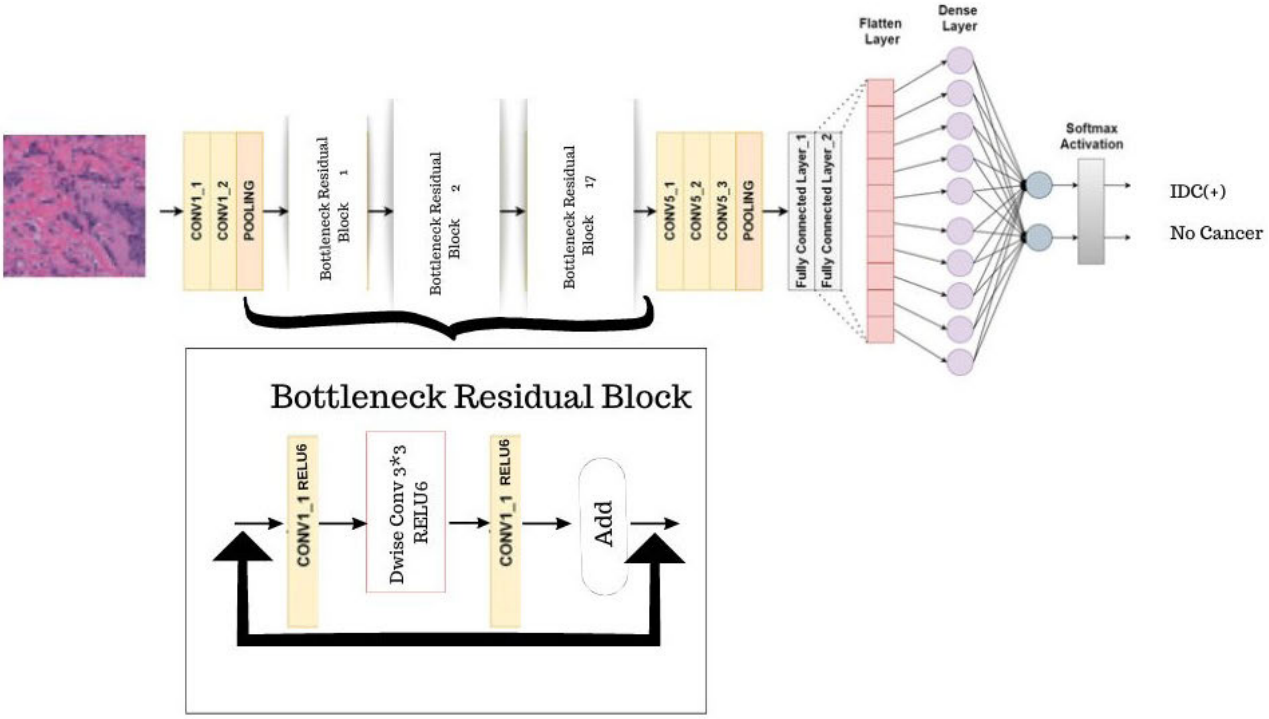


FIGURE 4. Binary classification using MobileNetV2: A comprehensive exploration of model architecture and configuration.

arises from its capacity to learn detailed image attributes via a hierarchical combination of convolutional and pooling layers, as shown in Fig. 5. This architecture excels at collecting delicate details as well as complex spatial relationships within histopathological images, which are critical for proper diagnosis. The use of pre-trained weights, initially obtained from a varied dataset such as ImageNet, gives a foundational grasp of fundamental image properties. VGG-16 improves its ability to distinguish patterns intimately linked to malignant tissues by training on histological breast cancer images. The VGG-16's effectiveness is notable for its combination of generalization and specialization. It takes advantage of the inherent generalization capability of pre-trained weights and combines it with specialized knowledge gathered from domain-specific data, thereby improving diagnostic accuracy. The model's widespread use in the field of deep learning translates into easily available resources, facilitating its implementation in medical research. However, the VGG-16's complete architecture, which includes 16 weight layers, implies a trade-off in terms of memory utilization, potentially causing issues in memory-constrained applications. Furthermore, its depth necessitates large processing resources during both the training and inference phases. Nonetheless, VGG-16's efficacy as a breast cancer screening tool remains undeniable. The advantages of pre-trained weights and a cooperative community enhance its sophisticated design, which enables robust feature extraction. This combination of characteristics solidifies VGG-16's position as a valuable asset in boosting the precision of breast cancer diagnosis and playing a critical role in improving patient care and medical research.

C. DenseNet-121

The DenseNet-121 model is very good at finding breast cancer early because it has a unique design based on dense connections and reusing features. DenseNet-121 was made to work well with histopathology images of breast cancer. It does a great job of spotting complex patterns by smartly combining information from different layers. Its architecture, which is made up of layers that are tightly woven into each block, makes it easier for the model to find small changes that are linked to cancer. With this design, the model can understand a wide range of spatial structures and deal with the problem of vanishing gradients, which can make it hard to learn. The main strength of DenseNet-121 comes from its tight connections and small network, which makes it great for situations with few resources or information. The model's ability to understand complex patterns is helped by its thick connectivity, as shown in Fig. 6 and the way it reuses features. It is also helped by the fact that it uses pre-trained weights from different datasets. By using these basic weights and improving its learning with expert histological images of breast cancer, the model goes through a process that makes it more accurate. DenseNet-121 stands out as a powerful tool for finding breast cancer early and correctly. Its unique design, which combines dense connectivity, feature reuse, and parameter efficiency, promises better care and outcomes for patients. This model can be changed and used in a variety

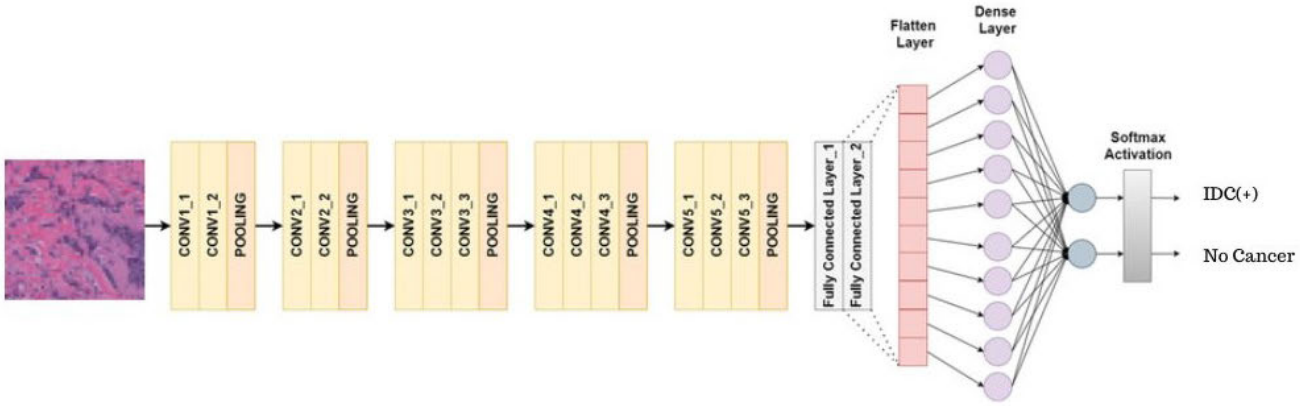


FIGURE 5. Binary classification using VGG-16: A comprehensive exploration of model architecture and configuration.

of breast cancer screening situations, which makes it a useful tool in the search for better diagnostic accuracy.

D. MODEL TRAINING

In the breast cancer detection deep learning training phase, a specialized approach was taken. All five models, InceptionV3, ResNet152V2, MobileNetV2, VGG-16 and DenseNet-121, were trained using the preprocessed training dataset, which was done over the course of 45 epochs. The most accurate training set was one that was balanced. Around 66,606 IDC+ patches and 155,413 no-cancer patches were present in the training dataset. During the training phase, the models were trained on how to identify important characteristics inside images in order to differentiate between IDC+ cases and cases that did not involve cancer. Existing knowledge was retained by freezing the base layers, which included all of the models with 'imagenet' weights. Images were 50 × 50 pixels in size and included RGB channels. Convergence was improved through strategic batch normalization. The importance of a 128-unit dense layer activated by ReLU was highlighted.

E. MODEL EVALUATION

Following the training phase,a comprehensive evaluation of each model's effectiveness was conducted using a test dataset comprising 128,838 images. The architecture of the model was meticulously constructed using Keras' Model API. To address overfitting concerns, a dropout layer with a 30% threshold was seamlessly integrated. Table. 4 shows how the intricate task of multi-class categorization was adroitly managed by the softmax activation function embedded within the output layer. During the compilation phase, the model harnessed the potent capabilities of the Adam optimizer, adeptly minimizing categorical cross-entropy with precision, guided by a consistent learning rate of 0.000106. The primary benchmark for assessing the model's success was its precision. Evaluation of the model's performance was carried out utilizing the accuracy metric, which quantified its prowess in classification. Furthermore, the evaluation scope extended to a validation set, where crucial metrics such as accuracy, precision, recall, and F1-score were harnessed to offer insights into the models' precision in categorizing histological breast cancer images. This meticulously optimized configuration was purposefully designed to effectively uncover the subtle intricacies of breast cancer patterns within the domain of diagnostic imaging.

Accuracy is a critical metric for evaluating classification models, particularly those based on deep learning. It calculates the proportion of correctly identified cases in a dataset compared to all instances. To calculate accuracy in binary classification, the ratio of true positive (properly predicted positive) and true negative (properly predicted negative) cases to the whole dataset is employed. According to Eq. (1), accuracy provides users with a basic understanding of a model's overall performance by demonstrating its ability to produce correct predictions across multiple classes. The accuracy calculation formula is as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total}} \tag{1}$$

Precision, Recall, and F1 Score in Deep Learning Precision The precision of the model is determined by how well it can separate the positive cases from all of the positive instances that are projected to occur. Eq. (2) computes the percentage of real positives from all cases that were successfully predicted as positives (positives that actually occurred). High precision indicates a low rate of false positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Recall, also known as sensitivity or true positive rate, measures the model's ability to discover every positive case in the dataset. The real positives to all other positive events ratio is calculated in Eq. (3). A high recall indicates a low rate of false negatives.

$$\textit{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

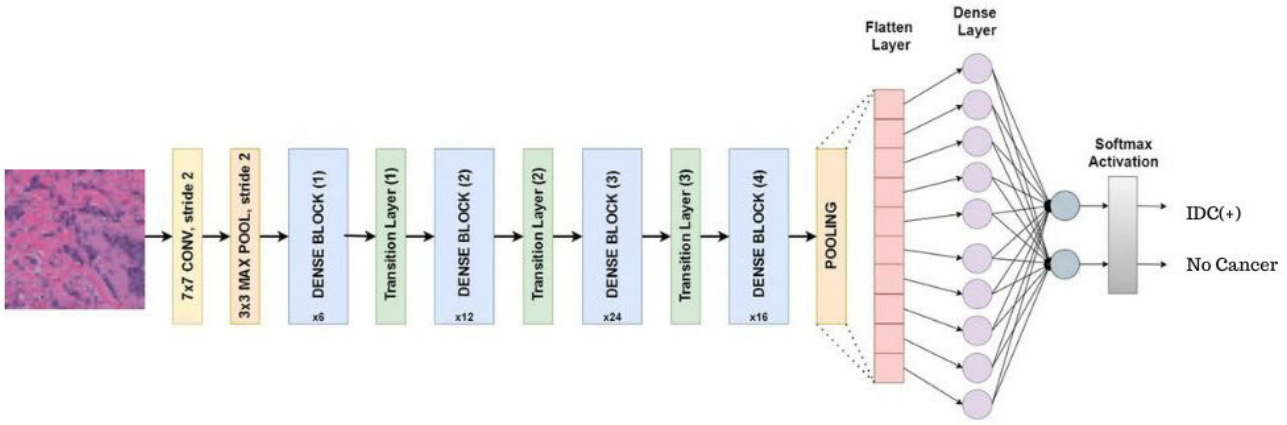


FIGURE 6. Binary classification using DenseNet-121: A comprehensive exploration of model architecture and configuration.

F1 Score merges precision and recall into a single metric and balances the trade-offs between them. Eq. (4)provides a more thorough evaluation of a model's performance by accounting for both false positives and false negatives. The F1 score is especially useful when there is an imbalance in the class distribution.

$$F1Score = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{4}$$

V. RESULTS AND DISCUSSIONS

A. EXPERIMENTAL ENVIRONMENT

Experiments utilizing deep learning transfer flow models for breast cancer detection were conducted in the Google Colab environment, which provided access to potent GPUs for accelerated model training. These investigations utilized Google Colab's GPU runtime environment, in particular the NVIDIA Tesla P100 GPU, for effective deep neural network model training and inference. The experimental toolkit consisted of essential software and libraries, including TensorFlow 2x and Keras for model construction and training, OpenCV for image preprocessing, NumPy for numerical computations, and a methodical approach to hyperparameters employing grid search. Memory constraints on the GPU were alleviated by optimizing training batch sizes, and early halting was used to prevent overfitting. Utilizing Google Colab's GPU runtime, this exhaustive experimental setup enabled the formulation and evaluation of accurate breast cancer detection models. The optimal hyperparameters, including batch size, learning rate, and optimization function, are considered as described in Table. 4.

TABLE 4. Hyper-parameters setting of the suggested model.

Parameters	Values
Dropout	0.5
Momentum	0.9
Learning Rate	0.0001
Epochs	45
Steps per epoch	100
Weight Decay	0.00005
Batch Size	16
Activation function	Softmax
Optimization function	Adam

B. BREAST HISTOPATHOLOGY IMAGES CLASSIFICATION BASED ON DEEP LEARNING PARADIGM

It is impossible to stress the significance of accuracy when utilizing deep learning to identify breast cancer tumors. For a patient's well-being, an accurate tumor diagnosis is essential since it allows for prompt therapies and higher survival rates. High accuracy guarantees that cancer cases are not missed while preventing undue stress and treatments for patients by minimizing both false negatives and false positives. Accurate models can help medical professionals allocate resources efficiently by enabling them to make well-informed judgments regarding additional diagnostic procedures and therapies. Reliable model predictions promote technology trust by portraying them as useful tools for radiologists and improving their capacity to deliver precise assessments. In the end, attaining high accuracy levels improves patient care and medical decision-making while also advancing medical research and technological breakthroughs by enabling a deeper comprehension of cancer biology and potential novel treatment modalities. Below are the classification results for breast cancer histopathology images produced by utilizing InceptionV3, ResNet152V2, MobileNetV2, VGG-16 and DenseNet-121 deep learning models:

A test dataset accuracy of 97% for the MobileNetV2 model demonstrated its efficacy in correctly categorizing occurrences. The VGG-16 model demonstrated the same capacity to recognize detailed features and complicated patterns in the data by achieving an accuracy of 98% on the test dataset. Due to its deep architecture, the model may learn hierarchical representations, which increases its ability to distinguish between various classes. This precision reveals

its promise for applications where a thorough comprehension of data relationships is essential. The test dataset showed the DenseNet-121 model to have a constant accuracy of 99%, demonstrating its reliability in capturing dense connections across layers and successfully resembling complicated data patterns. This quality improves its ability to generalize effectively and provide precise predictions in a variety of settings. Together, these high accuracy levels show how powerful these models are and how many other domains they may be applied to. In Fig. 7, a test dataset accuracy of 97% for the MobileNetV2 model revealed its efficacy in correctly categorizing occurrences. This indicates its ability to strike a great balance between speed and precision, making it suitable for assignments that require speedy processing without losing quality. On the test dataset, the VGG-16 model displayed the same ability to recognize specific features and complicated patterns in the data, achieving an accuracy of 98%. The model's deep architecture allows it to learn hierarchical representations, which improves its capacity to distinguish between different classes. This precision indicates its promise for applications that require a full understanding of data relationships. The test dataset revealed that the DenseNet-121 model had a consistent accuracy of 99% in identifying dense connections across layers and correctly mimicking intricate data patterns. This property enhances its ability to generalize successfully and provide precise predictions in a range of scenarios. These high levels of accuracy demonstrate how powerful these models are and how many other fields they can be applied to.

According to Fig. 7 study compared the performance of distinct pre-trained convolutional neural network (CNN) architectures: InceptionV3, ResNet152V2, MobileNetV2, VGG-16, and DenseNet-121. These models were fine-tuned for a specific aim, and the metrics utilized in training and validation were meticulously maintained. The analysis showed intriguing tendencies in terms of the time period when each model performed best. MobileNetV2 Epoch 3 produced the best results Fig. 8, with a training loss of 0.1367 and an accuracy of 97.43%. The validation results were as impressive, with a 97.52% accuracy and a validation loss of 0.1422. This demonstrates that MobileNetV2 has the ability to generalize and converge swiftly early in the training process.

VGG-16 attained a phenomenal accuracy of 98.80% at epoch 42; however, this performance peaked at that time. Fig. 9 shows that VGG-16 has a training loss of 0.0554 and a validation loss of 0.2230. This shows that VGG-16 benefited from further training, achieving increased accuracy but also exhibiting some overfitting features.

DenseNet-121 Fig. 10 demonstrated a different pattern by obtaining its finest epoch towards the conclusion of the maximum training period, or epoch 45. The model's exceptional accuracy of 99.39% and extraordinarily low training loss of 0.0170 were both achieved. The validation loss was 0.3874, while the validation accuracy was 99.86%. This result reveals that DenseNet-121 continued to learn and generalize even after extensive training, demonstrating its ability to use increasingly complex data features.

The focus of this research is on how different pre-trained CNN architectures differ in terms of convergence patterns and performance trade-offs. The best epochs demonstrate how early convergence, substantial training, and model capability interact sensitively. These insights provide useful recommendations for selecting the appropriate CNN architectures based on task needs and available computational resources.

MobileNetV2 demonstrates commendable performance when it comes to differentiating between IDC+ and no cancer. The model recognizes both positive and negative examples with an accuracy of 97%, making it a useful tool for determining either. Its precision score of 96% indicates that it is adept at eliminating false positives, which is essential in medical applications where it is essential to prevent unneeded procedures. In addition, the model obtains a specificity of 95%. The model's high specificity indicates that it has a relatively low risk of false positives. This is advantageous for medical diagnosis because it helps lower the likelihood of mistakenly classifying a healthy individual as having cancer. The F1 score of 97% demonstrates that MobileNetV2 is capable of striking a balance between accurate positive predictions and thorough detection of IDC+ instances. When compared to IDC+, the VGG-16 test has a slightly greater accuracy of 98% when determining whether or not a patient has cancer. This demonstrates that it is skilled in accurately categorizing both of the categories. The precision of the model is 98%, and it performs exceptionally well in reducing the number of false positives, which is an essential quality in the context of medical diagnostics. The fact that VGG-16 has a specificity rate of 97% demonstrates how effective it is at identifying true IDC+. The F1 score of 97% highlights its capacity to deliver a balanced prediction performance in the process. DenseNet-121 also excels in its ability to differentiate between patients with IDC+ and those without cancer. The model is able to make correct predictions in both categories with a level of accuracy that is equivalent to 99%. Its precision score of 99% highlights its proficiency in eliminating false positives, which is an essential component in medical diagnostics. A specificity rate of 99% suggests that the algorithm is able to recognize a substantial amount of true IDC+. The fact that DenseNet-121 was able to strike a balanced balance between recall and precision, as demonstrated by their 99% score on the F1 test, is further evidence of their prowess. Comparing these models Table 5 shows their strong ability to distinguish IDC+ from no cancer. At least 94% accuracy rates suggest that all these models appropriately classify occurrences. Precision scores of 97% to 98% emphasize their capacity to reduce false positives, which is crucial in medical applications. Recall ratings of 96% across the board show that the models catch a significant part of the actual IDC+. Their stable F1 scores of 98% demonstrate their balanced prediction performance. InceptionV3, ResNet152V2, MobileNetV2,

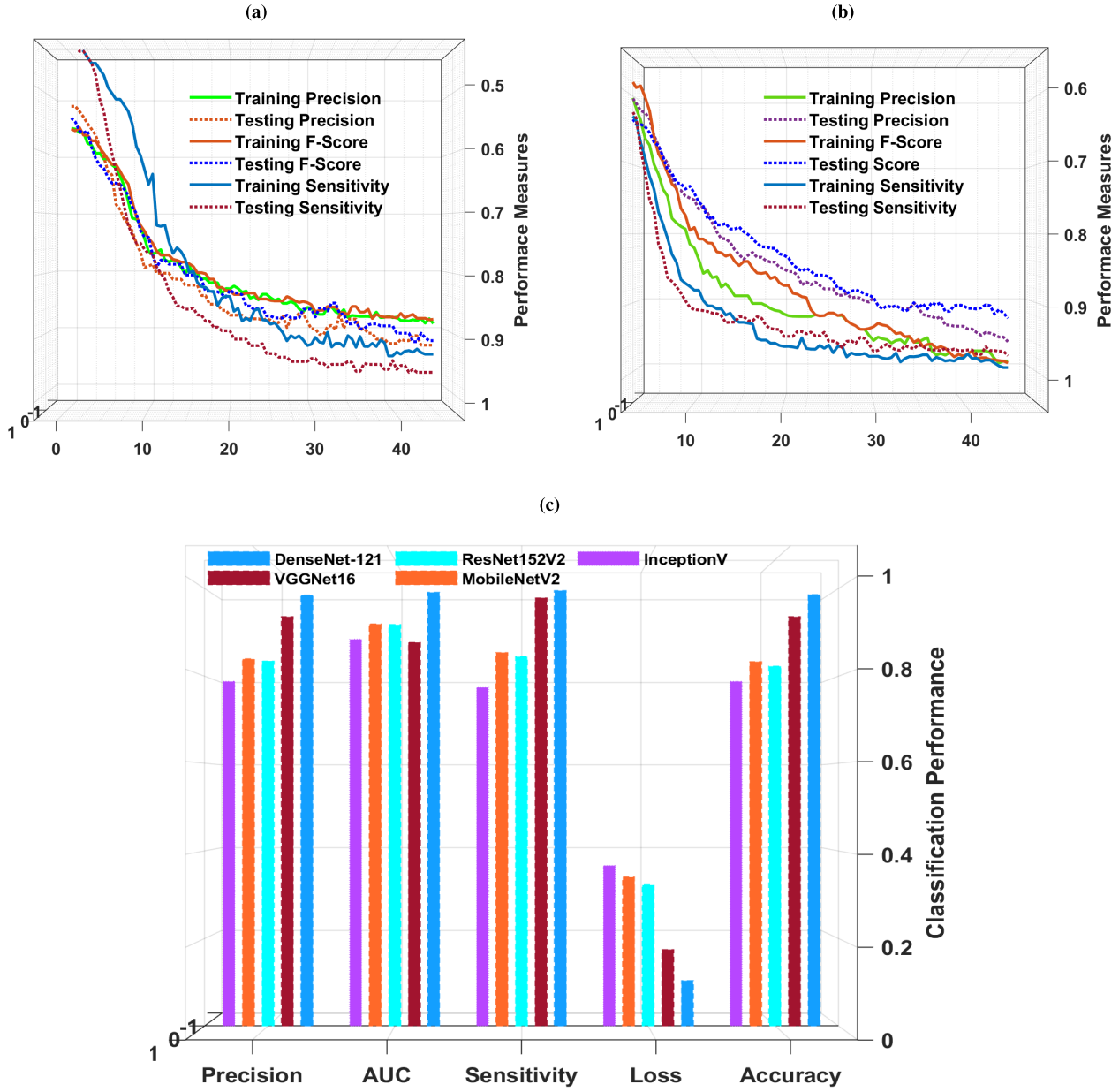


FIGURE 7. Comparative analysis of classification accuracy: visualizing performance metrics with accuracy graphs for MobileNetV2, VGG-16, and DenseNet-121.

TABLE 5. Performance comparison of DCNN-based architectures.

Proposed Algorithms	Accuracy ± σ	FScore ± σ	Precision ± σ	Specificity ± σ
InceptionV3	0.94 ± 0.017	0.95 ± 0.017	0.94 ± 0.039	0.94 ± 0.043
ResNet152V2	0.95 ± 0.016	0.96 ± 0.021	0.95 ± 0.037	0.95 ± 0.039
MobileNetV2	0.97 ± 0.020	0.97 ± 0.025	0.96 ± 0.033	0.95 ± 0.033
VGG-16	0.98 ± 0.021	0.97 ± 0.015	0.98 ± 0.026	0.97 ± 0.029
DenseNet-121	0.99 ± 0.011	0.99 ± 0.014	0.99 ± 0.017	0.99 ± 0.015

VGG-16, and DenseNet-121 provide accurate IDC+ and no cancer detection forecasts. The differences in their performance indicators might help choose a model based on a project's priorities.

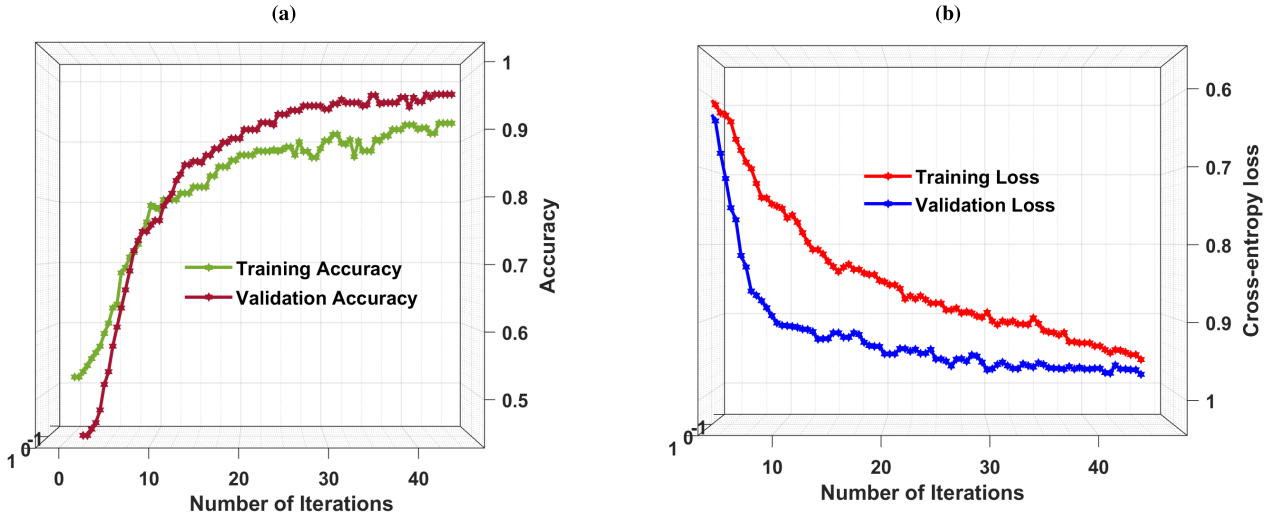


FIGURE 8. Visualizing the training progress of MobileNetV2: accuracy and loss graph tracking model performance.

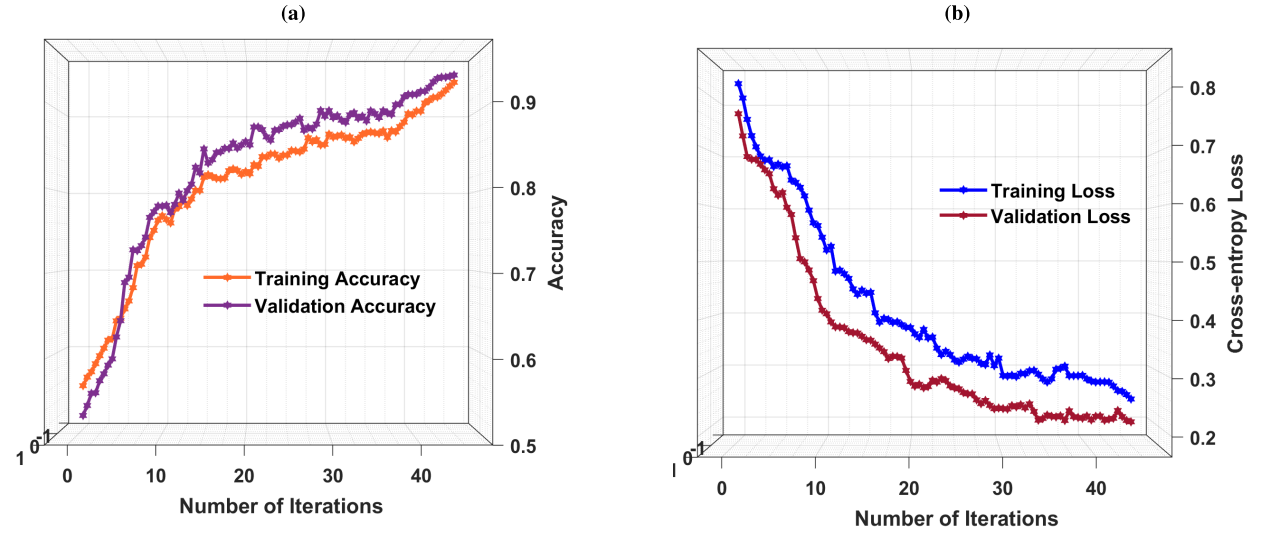


FIGURE 9. Visualizing the training progress of VGG-16: accuracy and loss graph tracking model performance.

C. SENSITIVITY ANALYSIS COMPARISON FOR VARIOUS PROPOSED TRANSFER LEARNING-BASED ARCHITECTURES

According to Table. 6 The performance of InceptionV3 is quite impressive when it comes to predicting IDC+ and no-cancer situations. It does an excellent job of minimizing prediction mistakes, as seen by its mean absolute error (MAE) value of 0.041 and mean squared error (MSE) value of 0.026. The root mean squared error (RMSE) value of 0.122 suggests that there is a comparatively low degree of variation in the accuracy of the forecast. Its remarkable Area Under the ROC Curve (AUC) of 0.943 hints at strong discrimination between IDC+ cases and instances with no cancer. In addition, its high sensitivity value of 0.984 demonstrates its capacity to successfully identify IDC+ instances, which makes it a reliable option for accurate cancer detection. The performance of ResNet152V2 is quite impressive when it comes to predicting IDC+ and no-cancer patients. Even though it has a slightly larger MAE of 0.077 in comparison to InceptionV3, it still manages to retain a competitive MSE of 0.038 and an acceptable RMSE of 0.199. The outstanding discrimination capabilities of this model are highlighted by its AUC score of 0.954. Additionally, the high sensitivity score of 0.986 illustrates its skill in recognizing actual IDC+ instances, making it a viable candidate for cancer detection assignments where high sensitivity is essential. This demonstrates its proficiency in identifying true IDC+ cases. When it comes to forecasting IDC+ and no-cancer cases, MobileNetV2 strikes a balance between the accuracy of its predictions and the efficiency with which it does its computations. The model has a good level of prediction accuracy, as indicated by the MAE value

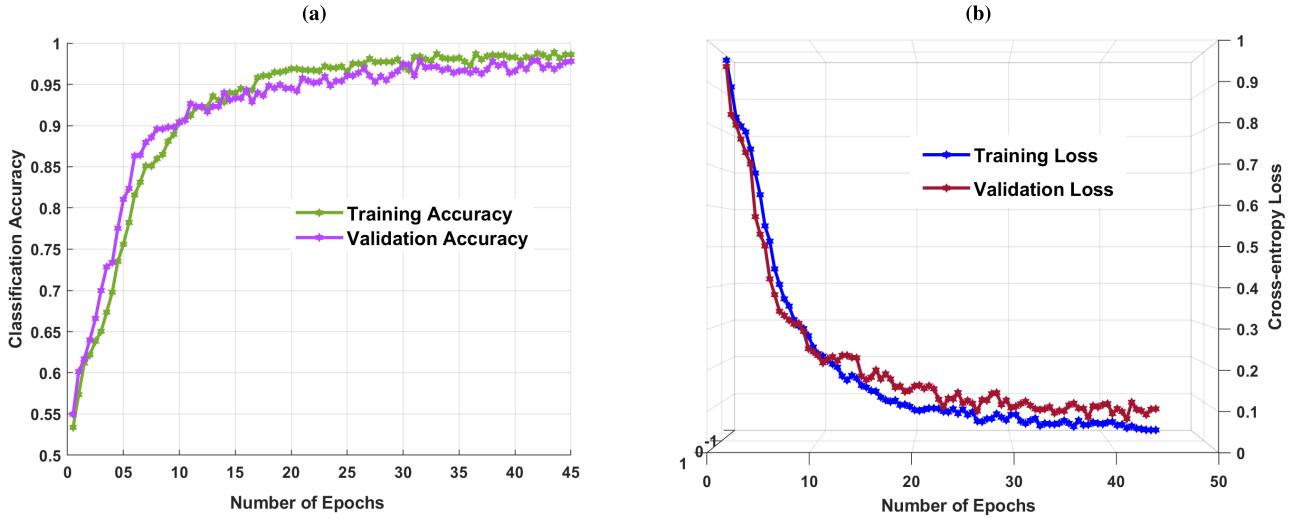


FIGURE 10. Visualizing the training progress of DenseNet-121: accuracy and loss graph tracking model performance.

TABLE 6. The proposed Deep learning-based algorithms have certain environmental requirements to be met for optimal performance.

Proposed Models	MAE	MSE	RMSE	AUC	Sensitivity
InceptionV3	0.041	0.026	0.122	0.943	0.984
ResNet152V2	0.077	0.038	0.199	0.954	0.986
MobileNetV2	0.052	0.029	0.191	0.973	0.997
VGG-16	0.043	0.018	0.182	0.981	0.978
DenseNet-121	0.010	0.015	0.169	0.995	0.998

of 0.052 and the MSE value of 0.029. Although the RMSE is significantly higher than acceptable, it is still within that range with a value of 0.191. The accuracy of the model's ability to distinguish between IDC+ cases and instances with no cancer is demonstrated by its area under the curve (AUC) score of 0.973. It has an impressively high sensitivity value of 0.997, making it especially useful in applications where catching real IDC+ cases is essential for cancer detection. This is a particularly important consideration. The accuracy of VGG-16's ability to predict IDC+ and no-cancer patients is high. It excels at minimizing prediction mistakes, as evidenced by its low MAE value of 0.043 and its remarkable MSE value of 0.018. A RMSE value of 0.182 indicates that there is relatively little variation in the accuracy of the forecast. A great capacity to discriminate between classes is indicated by its AUC score of 0.981. Its power to properly identify IDC+ instances is demonstrated by its sensitivity value of 0.978, making it a trustworthy option for accurate cancer detection activities. DenseNet-121 has an outstanding track record of accuracy when it comes to predicting IDC+ and no-cancer cases. It demonstrates amazing prediction accuracy with an MAE of 0.010 and an MSE of 0.015, both of which are extraordinarily low values. The RMSE value of 0.169 demonstrates that the predictions are dependable and consistent. It distinguishes between IDC+ cases and cases in which cancer was not present exceptionally well, thanks to a remarkable AUC value of 0.995. Because of its unusually high sensitivity value of 0.998, which demonstrates its capability to reliably identify IDC+ patients, it is a great choice for sensitive and accurate cancer detection activities.

The confusion matrix is an important tool for evaluating classification models, notably in the context of deep learning for breast cancer tumor identification. As a measure of a model's predictive performance, the Confusion Matrix shows the accuracy and reliability of the model's classifications by separating predictions into various groups. The confusion matrix contains critical information about the model's ability to correctly differentiate between malignant and benign tumors in the context of a breast cancer diagnosis. Predictions are divided into four quadrants by the Confusion Matrix: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True positives are situations in which the model correctly categorizes cases as having non-cancerous tissue, whereas true negatives are instances in which the model correctly categorizes cases as having

cancerous tissue. False positives are instances in which benign tumors are incorrectly classified as malignant, whereas false negatives are instances in which malignant tumors are incorrectly classified as benign. The confusion matrix is critical for assessing the model's performance in the context of a breast cancer diagnosis. A high-accuracy model is especially good at predicting cancer, which minimizes the likelihood of unnecessary surgeries. A high recall shows that the model is capable of identifying actual cancer patients, ensuring that vital diagnoses are not overlooked. By providing the information required to make educated decisions, the Confusion Matrix allows medical practitioners to balance the sensitivity and specificity of the model's predictions. Images that do not exhibit IDC (cancer) are assigned the number 0, while those that do are assigned the number 1.

According to Fig. 11 results for the MobileNetV2 model, the number of cases successfully identified as positive was 28,672. There were also 379 instances of false positives or positive cases that were misclassified. There were 469 genuine negatives in the model, or events that were correctly labeled as negative, compared to 440 false negatives. These counts show that while MobileNetV2 struggled with false positives and false negatives, it excelled at true positives and true negatives.

The VGG-16 model Fig. 12 displayed strong performance and its ability to classify positive occurrences with 28,826 true positives correctly. The low number of false positives, 382, demonstrates that it effectively prevented false positive mistakes. However, it generated 472 true negatives and 320 false negatives. The VGG-16 model performed well due to its high true positive count and low false positive count, while the higher false negatives suggested that recall may be improved.

The DenseNet-121 model stood out in the examination of its performance Fig. 13 for its impressive categorization of 29,240 true positives, demonstrating its great capacity to identify positive cases precisely. The model's conservative 230 false positives highlight its meticulous classification methodology. Additionally, the model registered 260 true negatives, adding to its overall effectiveness. However, the existence of 210 false negatives points to a problem that needs fixing. These occurrences indicate circumstances in which the model failed to detect actual positive cases. The model's recall performance might be significantly improved by improving its ability to eliminate false negatives, which would maximize its overall diagnostic skills and clinical utility.

When these models were examined, VGG-16 had the highest true positive count and the lowest false positive count, suggesting a solid balance between accuracy and minimizing false positives. DenseNet-121 underperformed VGG-16 regarding true positives, but it fared far better regarding false positives and true negatives. MobileNetV2 suffered from false positives and false negatives and had a lower true positive count. In terms of true positives and

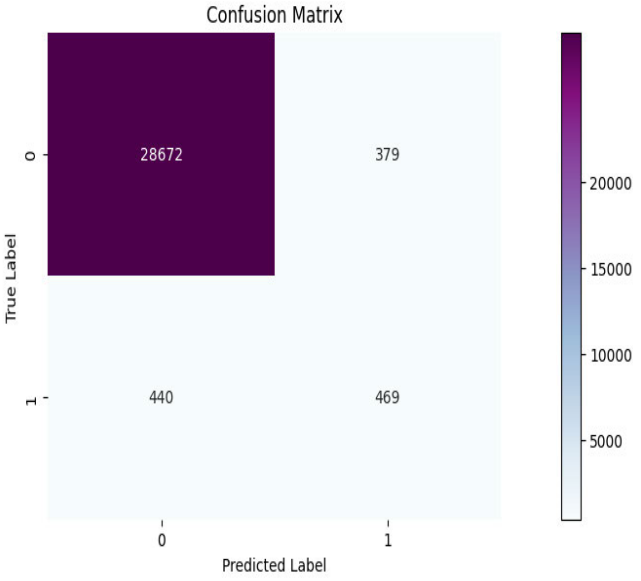


FIGURE 11. Evaluating classification performance with the confusion matrix for MobileNetV2: A visual representation of model's predictions and actual results.

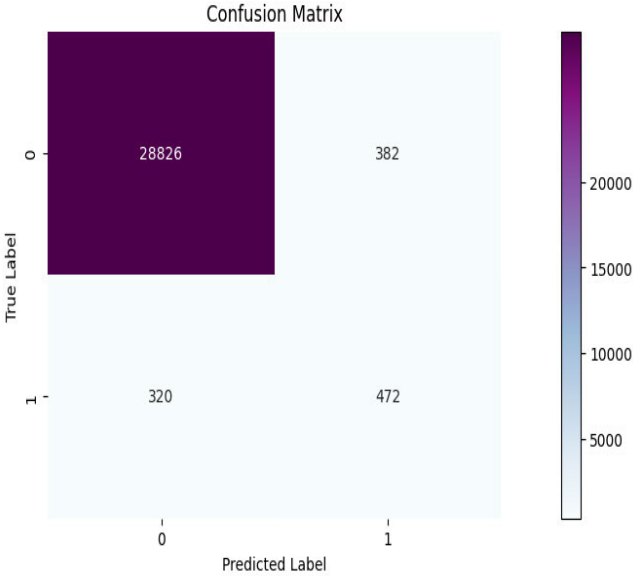


FIGURE 12. Evaluating classification performance with the confusion matrix for VGG-16: A visual representation of model's predictions and actual results.

balanced error rates, VGG-16 and DenseNet-121 outperformed MobileNetV2.

DenseNet-121, VGG-16, and MobileNetV2 have different false-negative rates in breast cancer prediction because of their different architectures and training methodologies. The efficient dense connectivity of DenseNet-121 aids in capturing critical cancer-related characteristics, resulting in a lower false-negative rate. In comparison, while VGG-16's deep structure is homogeneous, tiny features may be lost, contributing to a moderate false-negative rate. Because of its potential simplicity, MobileNetV2's lightweight design may

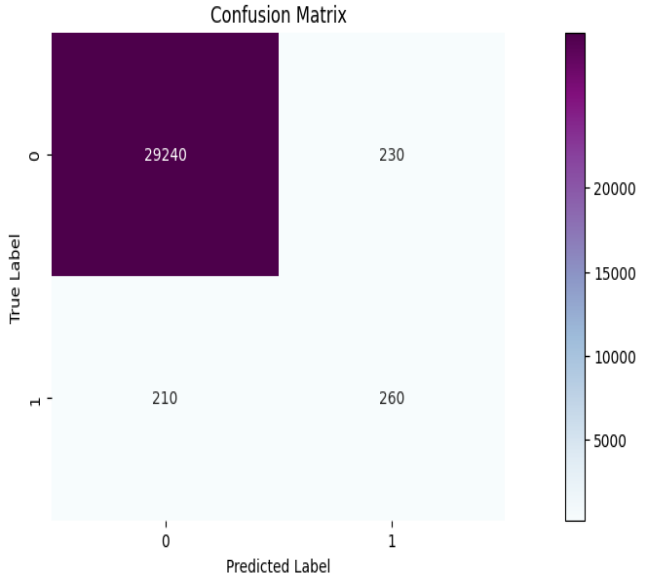


FIGURE 13. Evaluating classification performance with the confusion matrix for DenseNet-121: A visual representation of model's predictions and actual results.

result in a higher false-negative rate. Model depth, architecture, parameter count, and the availability of pre-trained weights all impact their performance. Choosing the best model requires balancing depth and efficiency. While DenseNet-121 performs admirably, examining additional measures and trade-offs is critical for making the best decision.

In clinical contexts, higher false-negative rates might have catastrophic repercussions. A false negative happens when the model wrongly classifies a cancer case as non-cancerous, potentially leading to delayed diagnoses and treatments. This lag can lead to missed opportunities for early intervention and impact patient outcomes. For example, aggressive cancers necessitate early detection to provide effective treatment strategies. In such scenarios, models with lower false negative rates, such as DenseNet-121, are preferable because they are less likely to ignore essential cancer cases. Increasing the accuracy of breast cancer detection models necessitates measures to reduce false negatives. Different data augmentation techniques are used to expose the model to different cases and refine feature extraction for subtle cancer details, ensemble learning for strong decisions, attention mechanisms for focusing on relevant regions, post-processing techniques, and transfer learning for better performance. By boosting sensitivity and accuracy in clinical diagnoses, these techniques collectively enable reliable real-world breast cancer detection.

The effectiveness of binary classification models, particularly those employed in detecting breast cancer, is frequently evaluated using ROC curves and AUC measures. Concerning various categorization thresholds, these measures aid in assessing the trade-off between the true positive rate (sensitivity) and false positive rate (1-specificity). ROC curves and AUC can offer essential insights into the models' capacity to distinguish between malignant and benign lesions based on medical images in the context of breast cancer diagnosis using deep learning models. Let's look at how ROC and AUC are used to assess how well MobileNetV2, VGG-16, and DenseNet-121 perform.

- AUC of 0.5 indicates that the model is not any more accurate than guessing at random.
- AUC between 0.5 and 1, with higher values indicating better performance, shows the model's ability to distinguish between IDC + and no cancer cases.
- AUC of 1 indicates 100% sensitivity and 100% specificity, which would be the case for the ideal classifier.

To sum up, ROC curves and AUC are crucial tools for assessing how well deep learning models like MobileNetV2, VGG-16, and DenseNet-121 perform in the identification of breast cancer. These measures give researchers and medical practitioners a way to measure how well the models can predict the future, which can help in earlier detection and treatment of breast cancer, potentially saving lives and improving patient outcomes.

With an AUC value of 0.97, MobileNetV2 performs commendably in terms of detecting breast cancer. This result Fig. 14 demonstrates the capability of MobileNetV2 to distinguish between invasive ductal carcinoma (IDC+) and non-cancerous instances in medical images. The ROC curve for MobileNetV2 is significant in that it demonstrates its capacity to attain a reasonable TPR while preserving an acceptable FPR, achieving a balance that is essential in clinical applications. This balancing makes sure that the model can recognize malignant instances (high TPR) without highlighting non-cancerous ones excessively (low FPR), decreasing false negatives and the danger of missed diagnoses.

VGG-16 outperforms MobileNetV2 with an AUC score of 0.98, demonstrating its improved IDC+ classification ability. The ROC curve for VGG-16 Fig. 15 demonstrates a continued increase in TPR without noticeably increasing the FPR, reiterating its potent ability to distinguish malignant from non-cancerous breast tumors. The performance of this model is defined by a preferable trade-off, ensuring a more accurate diagnosis while preserving an acceptable false alarm rate, which is crucial for clinical confidence.

With an outstanding AUC score of 0.99, Fig. 16 DenseNet-121 stands out as the best performer. The ROC curve for DenseNet-121 shows a noticeable rise in TPR, indicating its greater ability to reliably identify IDC+ cases. The fact that this model keeps its FPR low is significant since it shows increased diagnostic precision, which is essential for finding breast cancer. DenseNet-121's power to assist physicians in early diagnosis and treatment planning is demonstrated by its ability to obtain a high TPR while avoiding false positives.

The ROC analysis highlights the disparities in these three models' performance when compared to breast cancer detection. Given its significant TPR and acceptable FPR, DenseNet-121 has the highest discriminatory power in order

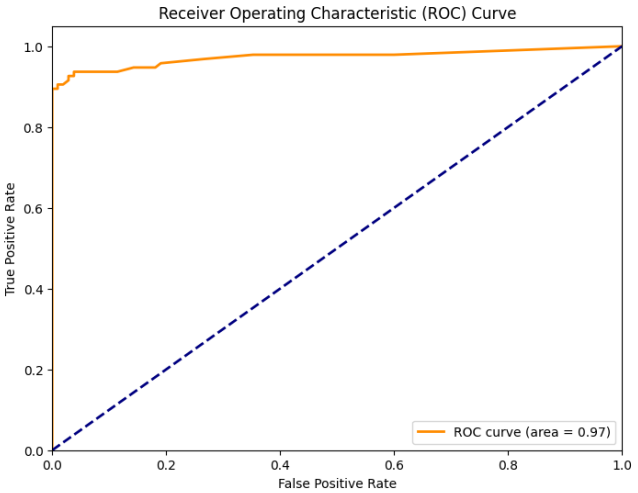


FIGURE 14. Evaluating MobileNetV2's classification performance: utilizing the receiver operating characteristic (ROC) curve for in-depth model performance analysis.

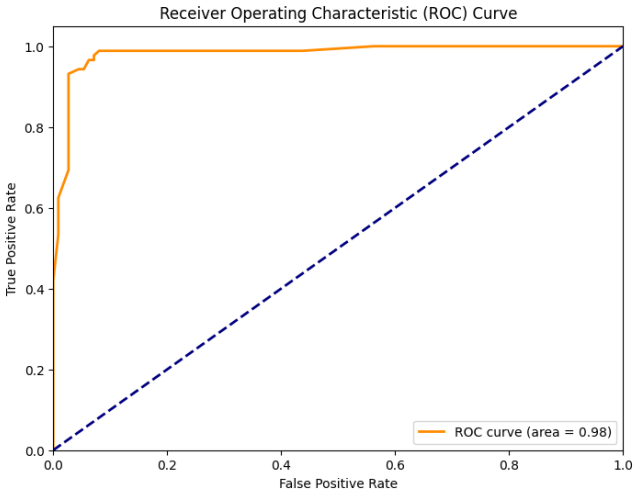


FIGURE 15. Evaluating VGG-16's classification performance: utilizing the receiver operating characteristic (ROC) Curve for in-depth model performance analysis.

to decrease missed diagnoses. Following closely behind and providing a good compromise between sensitivity and specificity is VGG-16. MobileNetV2 has a more subdued performance, but it is still effective. Specific clinical criteria and computational limitations should guide the model selection. Applications that prioritize precision, such as the diagnosis of breast cancer, are better suited for DenseNet-121 and VGG-16 due to their increased accuracy. However, situations with limited resources might benefit from MobileNetV2's efficiency. These deep learning models ultimately make a substantial contribution to the early identification of breast cancer by meeting a variety of clinical and computational requirements and aiming to enhance patient outcomes through prompt and accurate diagnosis.

D. COMPARISON WITH CUTTING-EDGE APPROACHES

Table 1 presents an in-depth comparison of various techniques. Notably, the Attention High-Order Deep Network (AHoNet) stands out as a leading approach. Its proficiency is evident with accuracy rates of 99.29% on the BreakHis and 85% on the BACH datasets, emphasizing its capability in classifying breast cancer histological images. In another instance, the combination of an extreme learning machine with CNN deep features, as mentioned in the reference, registers a commendable 92% accuracy on gene expression data. Both AHoNet and this combination have shown consistently high results, marking them as prime contenders for advancing breast cancer diagnosis research. While other studies, exploration of different designs, also exhibit notable accuracy, individual networks have their limitations in feature extraction and classification. Fusion networks, on the other hand, amalgamate the strengths of various networks, enhancing the classification process. This is further illustrated in Table 7. Rakhlin et al. [61] used VGG-16, InceptionV3, and ResNet-50 with LightGBM for

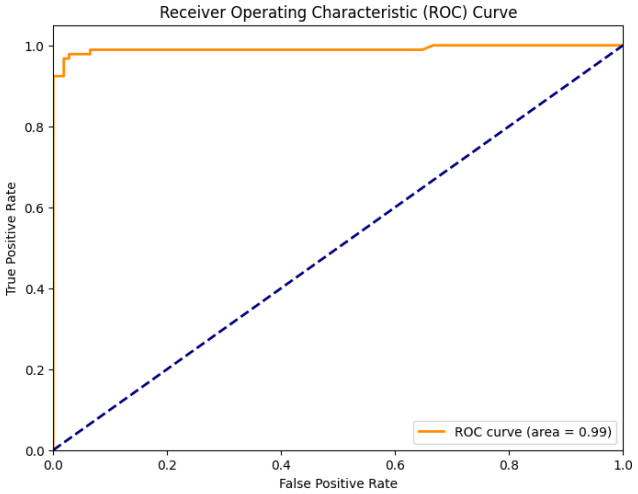


FIGURE 16. Evaluating DenseNet-121's classification performance: utilizing the receiver operating characteristic (ROC) curve for in-depth model performance analysis.

breast cancer image classification. A rigorous 10-fold cross-validation is implemented to guarantee a thorough assessment of the complexities present in the dataset.

With random color enhancements on the BACH dataset, they achieved a 74% accuracy. Pimkin et al. [60] integrated ResNet-34, DenseNet-169, and DenseNet-201 with XGBoost for breast cancer image classification. It employs a more versatile strategy, ranging from 3 to 6 folds without specifying the exact value, thereby showcasing its capacity to accommodate dataset-specific details. Using rotation and color space conversion on the BACH test set, they attained a 76% accuracy. Mahbod et al. [58] combined ResNet-50 and ResNet-101 to classify adenocarcinoma images. A balance is achieved between robust model training and a suitably rigorous evaluation with an 88% training to 12% testing ratio. They enhanced data by rotating it at 0°, 90°, 180°, and 270° and by horizontal inversion, achieving a 77% accuracy on the

TABLE 7. Accuracy of comparison of the state of the art study.

Author	Method	Accuracy
Chennamsetty et al. [48]	ResNet-101, DenseNet-161	0.87
Kwok et al. [49]	Inception-ResNet-v2	0.87
Brancati et al. [50]	ResNet-34, ResNet-50, ResNet-101	0.86
Marami et al. [51]	Inception-v3	0.84
Kohl et al. [52]	DenseNet-161	0.83
Wang et al. [53]	VGG-16	0.83
Koné et al. [54]	ResNeXt-50	0.81
Wang et al. [55]	VGG-16	0.79
Cao et al. [56]	VGG-16, ResNet-18, ResNet-152, ResNeXt, NASNet-A	0.79
Guo et al. [57]	GoogLeNet	0.77
Mahbod et al. [58]	ResNet-50, ResNet-101	0.77
Ferreira et al. [59]	Inception-ResNet-v2	0.76
Pimkin et al. [60]	ResNet-34, DenseNet-169, DenseNet-201	0.76
Rakhlin et al. [61]	VGG-16, Inception-v3, ResNet-50	0.74
Awan et al. [62]	ResNet-50	0.71

BACH test set. Cao et al. [56] combined features from VGG-16, ResNet-18, ResNet-152, ResNeXt, and NASNet-A[59] with manual ones, using RFSVM for breast cancer classification. it has adopted a slightly higher training commitment at 75%, allocating 25% for testing to ensure a comprehensive assessment of model generalization. This multi-extractor approach, akin to data enhancement, achieved 79% accuracy on the BACH test set. Brancati et al. [50] combined ResNet-34, ResNet-50, and ResNet-101 for breast cancer image classification. The model has employed a training-to-testing ratio of 80% to 20%, suggesting that it relies considerably more on training data to identify complex patterns. They augmented data with horizontal and vertical flips and three 90° rotations, achieving an 86% accuracy on the BACH test set. Chennamsetty et al. [48] employed ResNet-101 and DenseNet-161 for breast cancer image classification, using two different normalization methods.It has adopted a balanced approach with a 70% training, 20% testing, and 10% validation ratio, emphasizing the allocation of data for training and robust evaluation. The first set of images was analyzed by both models, while the second set exclusively trained DenseNet-161. This tripled feature extraction, akin to data enhancement, resulting in an 87% accuracy on the BACH concealed test set. Vesal et al. [63] used Inception-v3 and ResNet-50 to classify breast cancer histopathological images using various data augmentations. The study found Inception-v3 achieved 97.08% accuracy and ResNet-50 96.66% accuracy. Fusion networks, while not inherently superior to singular networks, offer acceptable accuracy with less data augmentation and network complexity. To maximize their potential, more data and task-specific network structures can be beneficial. Breast cancer histopathology image classification using CNNs falls into categories like CNN+Softmax, CNN+Softmax+MV, CNN+SVM, or other machine learning classifiers. Yan et al. [43] proposed integrating Inception-v3 with a recurrent neural network (RNN) to address spatial correlation issues in histopathological images of breast cancer, suggesting this combination could be crucial for future research on adenocarcinoma histopathology images.

VI. DISCUSSION

The results of the studies indicate that deep learning models are successful when it comes to categorizing images of breast cancer based on histology. The models MobileNetV2, VGG-16, and DenseNet-121 were very good at telling the difference between cases with invasive ductal carcinoma (IDC+) and cases without cancer. The MobileNetV2 model, when applied to the test dataset, achieved an impressively high level of accuracy of 97%. MobileNetV2, which is known for having an architecture that is both efficient and lightweight, is ideally suited for use in situations in which there are few available resources. It is interesting to note that the VGG-16 and DenseNet-121 models both demonstrated exceptional performances, each with an accuracy of 98% and 99%, respectively, based on the test dataset. It is commonly known that a powerful CNN model known as VGG-16 has the ability to extract complex properties from images. On the other hand, the strongly coupled convolutional layers that are used by DenseNet-121 enhance the reuse of features and maximize parameter efficiency. The higher accuracy levels produced by VGG-16 and DenseNet-121 models in comparison to MobileNetV2 indicate that deeper and more intricate architectures may be favorable for this particular breast cancer classification task. This can be observed by comparing the models' accuracy levels.In comparison to the literature presented, my study stands out as a new development in breast cancer diagnosis. While the above-stated studies have provided useful insights, my work adds a new dimension by focusing on a previously examined histopathology dataset. Notably, I have used a substantially larger volume of data as input, increasing the study's depth and breadth. This approach

has resulted in outstanding accuracy rates that exceed the benchmarks set by previous studies. The rigorous comparison of three cutting-edge transfer learning models, each adapted to the particular properties of breast cancer histopathology images, demonstrates my suggested methodology's higher accuracy and reliability. This particular contribution results from the purposeful relationship between advanced algorithms and a solid dataset, allowing for a more precise and confident breast cancer diagnosis. By exceeding earlier studies in terms of accuracy, my study represents a huge step forward in improving diagnostic capabilities, ultimately leading to better patient care and more accurate outcomes. The exceptional precision of the models has significant implications for the screening for and diagnosis of breast cancer at an earlier stage. With the correct classification of breast cancer histopathology images, medical workers are able to make informed judgments regarding the treatment options for patients, which ultimately leads to improvements in their outcomes. In spite of the fact that the findings are optimistic, the research might yet be improved by carrying out cross-validation and evaluating the models' performance using fictitious data in order to verify that they are generalizable. Additional study on larger and more diverse datasets, in addition to the addition of other evaluation criteria (such as precision, recall, and F1-score), would also result in a more in-depth review of the performance of the models. This would be the case if we were to conduct the evaluation. This work puts an emphasis on how successfully deep learning models, in particular VGG-16 and DenseNet-121, categorize images of breast cancer from a histological point of view. The high levels of accuracy that were achieved suggest that these models have the potential for practical applications in healthcare settings, such as supporting pathologists in the rapid and accurate diagnosis of breast cancer patients.

VII. CONCLUSION

In this research, transfer learning models were used to categorize histopathology images for the purpose of identifying invasive ductal carcinoma (IDC). By utilizing state-of-the-art deep learning architectures like VGG-16 and DenseNet-121, the models achieved classification accuracies of 98% and 99%, respectively, demonstrating their potential for accurate and efficient diagnosis. MobileNetV2 also showed a competitive accuracy of 97%, further emphasizing the utility of transfer learning in medical image analysis. The success of VGG-16 and DenseNet-121 suggests that complex architectures with deeper layers can capture intricate patterns in histopathology images. However, the study's limitations include the substantial dataset size and the exploration of only three popular architectures. Future investigations could explore a wider range of architectures, including more recent developments in deep learning.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGMENT

This research was supported by Princess Nourah bint Abdulrahman University and Researchers Supporting Project number (PNURSP2023R346), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The author would also like to thank Prince Sultan University, Riyadh Saudi Arabia for support.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 71, no. 3, pp. 209–249, May 2021.

[2] A. Naz, M. Javed, N. Javaid, T. Saba, M. Alhussein, and K. Aurangzeb, "Short-term electric load and price forecasting using enhanced extreme learning machine optimization in smart grids," *Energies*, vol. 12, no. 5, p. 866, Mar. 2019.

[3] S. F. Khorshid, A. M. Abdulazeez, and A. B. Sallow, "A comparative analysis and predicting for breast cancer detection based on data mining models," *Asian J. Res. Comput. Sci.*, vol. 2021, pp. 45–59, May 2021.

[4] T. Saba, S. T. F. Bokhari, M. Sharif, M. Yasmin, and M. Raza, "Fundus image classification methods for the detection of glaucoma: A review," *Microsc. Res. Technique*, vol. 81, no. 10, pp. 1105–1121, Oct. 2018.

[5] M. Abubakar, A. Klein, S. Fan, S. Lawrence, K. Mutreja, J. E. Henry, R. M. Pfeiffer, M. A. Duggan, and G. L. Gierach, "Host, reproductive, and lifestyle factors in relation to quantitative histologic metrics of the normal breast," *Breast Cancer Res.*, vol. 25, no. 1, p. 97, Aug. 2023.

[6] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018.

[7] *Intelligent Systems Design and Applications (ISDA) 2013 13th International Conference*, Universiti Putra Malaysia, Seri Kembangan, Malaysia, Dec. 2013.

[8] M. A. Khan, M. I. Sharif, M. Raza, A. Anjum, T. Saba, and S. A. Shad, "Skin lesion segmentation and classification: A unified framework of deep neural network features fusion and selection," *Expert Syst.*, vol. 39, no. 7, 2022, Art. no. e12497.

[9] T. Mahmood, J. Li, Y. Pei, F. Akhtar, A. Imran, and K. U. Rehman, "A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities," *IEEE Access*, vol. 8, pp. 165779–165809, 2020.

[10] R. Hou, M. A. Mazurowski, L. J. Grimm, J. R. Marks, L. M. King, C. C. Maley, E. S. Hwang, and J. Y. Lo, "Prediction of upstaged ductal carcinoma in situ using forced labeling and domain adaptation," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 6, pp. 1565–1572, Jun. 2020.

[11] T. Mahmood, J. Li, Y. Pei, F. Akhtar, M. U. Rehman, and S. H. Wasti, "Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach," *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0263126.

[12] S. Pervez and H. Khan, "Infiltrating ductal carcinoma breast with central necrosis closely mimicking ductal carcinoma in situ (comedo type): A case series," *J. Med. Case Rep.*, vol. 1, no. 1, pp. 1–4, Sep. 2007.

[13] D. L. Page, W. D. Dupont, L. W. Rogers, and M. Landenberger, "Intraductal carcinoma of the breast: Follow-up after biopsy only," *Cancer*, vol. 49, no. 4, pp. 751–758, Feb. 1982.

[14] R. A. Ramadan, A. Y. Khedr, K. Yadav, E. J. Alreshidi, M. H. Sharif, A. T. Azar, and H. Kamberaj, "Convolution neural network based automatic localization of landmarks on lateral X-ray images," *Multimedia Tools Appl.*, vol. 81, no. 26, pp. 37403–37415, Nov. 2022.

[15] S. Masciari, N. Larsson, J. Senz, N. Boyd, P. Kaurah, M. J. Kandel, L. N. Harris, H. C. Pinheiro, A. Troussard, P. Miron, N. Tung, C. Oliveira, L. Collins, S. Schnitt, J. E. Garber, and D. Huntsman, "Germline E-cadherin mutations in familial lobular breast cancer," *J. Med. Genet.*, vol. 44, no. 11, pp. 726–731, Aug. 2007.

[16] F. M. Robertson, M. Bondy, W. Yang, H. Yamauchi, S. Wiggins, S. Kamrudin, and S. Krishnamurthy, "Inflammatory breast cancer: The disease the biology the treatment," *CA, Cancer J. Clinician*, vol. 60, no. 6, pp. 351–375, Nov. 2010.

[17] B. Mathiyazhagan, J. Liyaskar, A. T. Azar, H. H. Inbarani, Y. Javed, N. A. Kamal, and K. M. Fouad, "Rough set based classification and feature selection using improved harmony search for peptide analysis and prediction of anti-HIV-1 activities," *Appl. Sci.*, vol. 12, no. 4, p. 2020, Feb. 2022.

[18] P. Mishra, S. Biswas, and S. Dash, "Deep learning based biomedical named entity recognition systems," in *Deep Learning Techniques for Biomedical and Health Informatics*. Switzerland: Springer, 2020, pp. 23–40.

[19] T. G. Debelee, F. Schwenker, A. Ibenthal, and D. Yohannes, "Survey of deep learning in breast cancer image analysis," *Evolving Syst.*, vol. 11, no. 1, pp. 143–163, Mar. 2020.

[20] S. M. Abbas, N. Javaid, A. T. Azar, U. Qasim, Z. A. Khan, and S. Aslam, "Towards enhancing the robustness of scale-free IoT networks by an intelligent rewiring mechanism," *Sensors*, vol. 22, no. 7, p. 2658, Mar. 2022.

[21] T. Mahmood, A. Rehman, T. Saba, L. Nadeem, and S. A. O. Bahaj, "Recent advancements and future prospects in active deep learning for medical image segmentation and classification," *IEEE Access*, vol. 11, pp. 113623–113652, 2023.

[22] S. Ali, J. Li, Y. Pei, R. Khurram, K. U. Rehman, and T. Mahmood, "A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal MR image," *Arch. Comput. Methods Eng.*, vol. 29, no. 7, pp. 4871–4896, Nov. 2022.

[23] S. Naseem, T. Mahmood, T. Saba, F. S. Alamri, S. A. Bahaj, H. Ateeq, and U. Farooq, "DeepFert: An intelligent fertility rate prediction approach for men based on deep learning neural networks," *IEEE Access*, vol. 11, pp. 75006–75022, 2023.

[24] S. Iqbal, A. N. Qureshi, J. Li, and T. Mahmood, "On the analyses of medical images using traditional machine learning techniques and convolutional neural networks," *Arch. Comput. Methods Eng.*, vol. 30, no. 5, pp. 3173–3233, Jun. 2023.

[25] C. Mohanty, S. Mahapatra, B. Acharya, F. Kokkoras, V. C. Gerogiannis, I. Karamitsos, and A. Kanavos, "Using deep learning architectures for detection and classification of diabetic retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, Jun. 2023.

[26] *10th International Multi-Conference on Systems Signals and Devices (SSD)*, IEEE, Piscataway, NJ, USA, 2013.

[27] M. Gao, J. A. Fessler, and H.-P. Chan, "Deep convolutional neural network with adversarial training for denoising digital breast tomosynthesis images," *IEEE Trans. Med. Imag.*, vol. 40, no. 7, pp. 1805–1816, Jul. 2021.

[28] A. Maryam, S. Saif, T. Aslam, S. Aslam, K. Tusleem, and M. Qamar, "Statistical evaluation of accuracy of mammography in early breast cancer detection," *Pakistan J. Med. Health Sci.*, vol. 11, no. 4, pp. 1430–1435, Oct. 2017.

[29] K. U. Rehman, J. Li, Y. Pei, A. Yasin, S. Ali, and Y. Saeed, "Architectural distortion-based digital mammograms classification using depth wise convolutional neural network," *Biology*, vol. 11, no. 1, p. 15, Dec. 2021.

[30] S. Rajpal, M. Agarwal, V. Kumar, A. Gupta, and N. Kumar, "Triphasic DeepBRCA—A deep learning-based framework for identification of biomarkers for breast cancer stratification," *IEEE Access*, vol. 9, pp. 103347–103364, 2021.

[31] C. Songsaeng, P. Woodtichartpreecha, and S. Chaichulee, "Multi-scale convolutional neural networks for classification of digital mammograms with breast calcifications," *IEEE Access*, vol. 9, pp. 114741–114753, 2021.

[32] R. C. Wang and Z. Wang, "Precision medicine: Disease subtyping and tailored treatment," *Cancers*, vol. 15, no. 15, p. 3837, Jul. 2023.

[33] A. Khamparia, S. Bharati, P. Podder, D. Gupta, A. Khanna, T. K. Phung, and D. N. H. Thanh, "Diagnosis of breast cancer based on modern mammography using hybrid transfer learning," *Multidimensional Syst. Signal Process.*, vol. 32, no. 2, pp. 747–765, Apr. 2021.

[34] T. Mahmood, J. Li, Y. Pei, and F. Akhtar, "An automated in-depth feature learning algorithm for breast abnormality prognosis and robust characterization from mammography images using deep transfer learning," *Biology*, vol. 10, no. 9, p. 859, Sep. 2021.

[35] H. Li, S. Zhuang, D.-A. Li, J. Zhao, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomed. Signal Process. Control*, vol. 51, pp. 347–354, May 2019.

[36] Z. Hameed, S. Zahia, B. Garcia-Zapirain, J. J. Aguirre, and A. M. Vanegas, "Breast cancer histopathology image classification using an ensemble of deep learning models," *Sensors*, vol. 20, no. 16, p. 4373, Aug. 2020.

[37] Z. Jafari and E. Karami, "Breast cancer detection in mammography images: A CNN-based approach with feature selection," *Information*, vol. 14, no. 7, p. 410, Jul. 2023.

[38] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao, H. Zhang, and J. Xin, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019.

[39] R. Mehra, "Breast cancer histology images classification: Training from scratch or transfer learning?" *ICT Exp.*, vol. 4, no. 4, pp. 247–254, Dec. 2018.

[40] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with deep learning," *Sci. Rep.*, vol. 8, no. 1, pp. 1–7, Mar. 2018.

[41] H. Nasir 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*, vol. 7, pp. 165724–165733, 2019.

[42] S. Montaha, S. Azam, A. K. M. R. H. Rafid, P. Ghosh, M. Z. Hasan, M. Jonkman, and F. De Boer, "BreastNet18: A high accuracy fine-tuned VGG16 model evaluated using ablation study for diagnosing breast cancer from enhanced mammography images," *Biology*, vol. 10, no. 12, p. 1347, Dec. 2021.

[43] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, and F. Zhang, "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52–60, Feb. 2020.

[44] Y. Zou, J. Zhang, S. Huang, and B. Liu, "Breast cancer histopathological image classification using attention high-order deep network," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 1, pp. 266–279, Jan. 2022.

[45] A. Jha, G. Verma, Y. Khan, Q. Mehmood, D. Rebholz-Schuhmann, and R. Sahay, "Deep convolution neural network model to predict relapse in breast cancer," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 351–358.

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

[47] *The Dataset "Breast Histopathology Images" on Kaggle. Predict IDC in Breast Cancer Histology Images | Kaggle*, Paultimothymooney, 2023.

[48] S. S. Chennamsetty, M. Safwan, and V. Alex, "Classification of breast cancer histology images using ensemble of pre-trained neural networks," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 804–811.

[49] S. Kwok, "Multiclass classification of breast cancer in whole-slide images," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 931–940.

[50] B. Fruccim and D. Riccio, "Multi-classification of breast cancer histology images by using a fine-tuning strategy," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1, 2018, pp. 771–778.

[51] B. Marami, M. Prastawa, M. Chan, M. Donovan, G. Fernandez, and J. Zeineh, "Ensemble network for region identification in breast histopathology slides," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 861–868.

[52] M. Kohl, C. Walz, F. Ludwig, S. Braunewell, and M. Baust, "Assessment of breast cancer histology using densely connected convolutional networks," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 903–913.

[53] Y. Wang, L. Sun, K. Ma, and J. Fang, "Breast cancer microscope image classification based on CNN with image deformation," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 845–852.

[54] I. Kone and L. Boulmane, "Hierarchical ResNeXt models for breast cancer histology image classification," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 796–803.

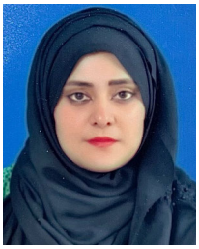
[55] Z. Wang, N. Dong, and W. Dai, "Classification of breast cancer histopathological images using convolutional neural networks with hierarchical loss and global pooling," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 745–753.

[56] H. Cao, S. Bernard, L. Heutte, and R. Sabourin, "Improve the performance of transfer learning without fine-tuning using dissimilarity-based multi-view learning for scalable image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jul. 2018, pp. 8697–8710.

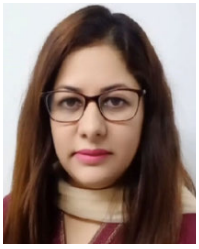
[57] Y. Guo, H. Dong, F. Song, C. Zhu, and J. Liu, "Breast cancer histology image classification based on deep neural networks," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 827–836.

[58] A. Mahbod, I. Ellinger, R. Ecker, Ö. Smedby, and C. Wang, "Breast cancer histological image classification using fine-tuned deep neural networks," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 754–762.

[59] C. A. Ferreira, T. Melo, P. Sousa, M. I. Meyer, E. Shakibapour, P. Costa, and A. Campilho, "Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet V2," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 763–770.
[60] A. Pimkina, G. Makarchuk, V. Kondratenko, M. Pisov, E. Krivov, and M. Belyaev, "Ensembling neural networks for digital pathology image classification and segmentation," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1, 2018, pp. 877–886.
[61] A. Rakhlin, A. Shvets, V. Iglovikov, and A. A. Kalinin, "Deep convolutional neural networks for breast cancer histology image analysis," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1, 2018, pp. 737–744.
[62] R. Awan, N. A. Koohbanani, M. Shaban, A. Lisowska, and N. Rajpoot, "Context-aware learning using transferable features for classification of breast cancer histology images," in *Proc. Int. Conf. Image Anal. Recognit.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 788–795.
[63] H. M. Ahmad, S. Ghuffar, and K. Khurshid, "Classification of breast cancer histology images using transfer learning," in *Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*. Cham, Switzerland: Springer, Jan. 2019, pp. 328–332.



WISHAL ARSHAD received the B.S. and M.S. degrees in computer science from UCP Lahore. She is currently pursuing the Ph.D. degree with Superior University, Lahore. She is actively engaged in academia and is also a Faculty Member with the Software Engineering Department, UCP Lahore. Her research interests primarily revolve around machine learning and deep learning, where she explores innovative approaches to advance the field. With a strong academic background and dedication to teaching and research, she contributes significantly to the academic community.



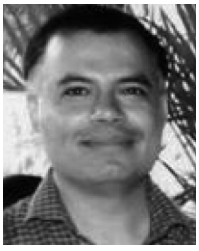
TEHREEM MASOOD received the bachelor's degree (Hons.) in software engineering from Punjab University, the M.S. (Research) degree (Hons.) in software engineering from MAJU, Islamabad, and the Ph.D. degree in computing from the Decision and Information Systems for Production Systems (DISP) Laboratory, INSA Lyon, University de Lyon, France. She is currently an academician, a software engineer, a data scientist, a researcher, and an entrepreneur with leadership qualities. With 14 years of experience, including industry, academia, and research at both national and international levels, she is a senior member of the intelligent data visual computing research (IDVCR). She is an HEC-approved Ph.D. Supervisor and a recipient of the Erasmus Mundus Scholarship under the Clink Project, France. Previously, she held the position of Senior Software Engineer with Pakistan Air Force in a research and development setup. She has implemented performance-based semantic, analytical, and predictive models for service sustainability and decision support in SOA following the machine learning approach. These models were validated with the implementation of an industrial business process use case and public data set repositories of shared services. Furthermore, she has developed the RWest Tool for specification-based regression testing of web services. Her research interests include data science, machine learning, deep learning, software testing, semantic web, and service-oriented architecture.

TARIQ MAHMOOD received the master's degree in computer science from the University of Lahore, Pakistan, and the Ph.D. degree in software engineering from the Beijing University of Technology, China. He is an Assistant Professor/the HOD with the Faculty of Information Sciences, University of Education, Vehari Campus, Vehari, Pakistan. He is a renowned expert in image processing, healthcare informatics, social media analysis, ad hoc networks, and WSN. He has contributed more than 25 research articles in well-reputed international journals and conferences. His research interests include image processing, social media analysis, medical image diagnosis, machine learning, and data mining. He aims to contribute to interdisciplinary research of computer science and human-related disciplines. He is an Editorial Board Member and a Reviewer of various journals, including *Plos One*, *The Journal of Supercomputer*, *Journal of Digital Imaging*, *International Journal of Sensors*, and *Wireless Communications and Control*.

ARFAN JAFFAR received the M.Sc. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, in March 2003, and the M.S. and Ph.D. degrees in computer science from the FAST National University of Computer and Emerging Sciences, in 2006 and 2009, respectively. He was an Assistant Professor with Al-Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, from March 2013 to August 2018. He is currently the Dean with the Faculty of Computer Science and Information Technology, Superior University, Lahore, Pakistan, where he is also the Director of intelligent data visual computing research (IDVCR). He received a Postdoctoral Research Fellowship from South Korea and carried out research at the top raking Korean university, such as the Gwangju Institute of Science and Technology, Gwangju, South Korea, from 2010 to 2013. His research interests include image processing, data science, machine learning, computer vision, artificial intelligence, and medical images. He is a Reviewer of 30 reputed international journals, such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, *Pattern Recognition*, *Knowledge*, and *Information Sciences*.

FATEN S. ALAMRI received the Ph.D. degree in system modeling and analysis in statistics from Virginia Commonwealth University, USA, in 2020. Her Ph.D. research included Bayesian dose-response modeling, experimental design, and nonparametric modeling. She is an Assistant Professor with the Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdul Rahman University. Her research interests include spatial area, environmental statistics, and brain imaging.

SAEED ALI OMER BAHAJ received the Ph.D. degree from Pune University, India, in 2006. He is an Associate Professor with the Department of Management Information Systems, College of Business Administration, Al-Kharj. He is also an Associate Professor with the Computer Engineering Department, Hadramout University, Yemen, and the MIS Department, COBA, Prince Sattam bin Abdulaziz University. His main research interests include artificial intelligence, information management, forecasting, information engineering, big data, and information security.



AMJAD R. KHAN (Senior Member, IEEE) received the Ph.D. and Postdoctoral degrees (Hons.) from the Faculty of Computing, Universiti Teknologi Malaysia, in 2010 and 2011, respectively, with a specialization in forensic documents analysis and security. He is a Senior Researcher with the Artificial Intelligence and Data Analytics Laboratory, College of Computer and Information Sciences (CCIS), Prince Sultan University, Riyadh, Saudi Arabia. He is the author of more than 200 ISI journal articles and conferences. He is a PI of several funded projects and also completed projects funded from MOHE Malaysia and Saudi Arabia. His research interests include data mining, health informatics, and pattern recognition. He received the Rector Award for the 2010 Best Student from Universiti Teknologi Malaysia.