International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

A Novel Hybrid Method to Detect and Diagnose Breast Cancer

ANJANA K
Department of Computer Science and Engineering
Bannari Amman Institute of Technology
Sathyamangalam, TamilNadu, India

GANAPATHY S
Department of Electrical and Electronics Engineering
Bannari Amman Institute of Technology
Sathyamangalam, TamilNadu, India

RAMYA D
Department of Artificial Intelligence and Machine Learning
Bannari Amman Institute of Technology
Sathyamangalam, TamilNadu, India

Abstract: A development of abnormal tissue in the breast tissue is called breast cancer. The condition may be associated with serious health concerns. Currently, the most common cancer among women in the world is breast cancer. It is also one of the leading causes of death due to cancer. Its early and accurate diagnosis will alleviate the treatment burden and enhance patient survival. To this end, the paper outlines a new approach for breast cancer diagnosis using histopathological images from the BreakHis dataset of 2,480 benign and 5,429 malignant images. The method has analyzed and interpreted the pictures using the transfer learning model using Convolutional Neural Networks (CNNs), VGG16, and VGG19. Using pre-trained models that VGG19 uses, the suggested strategy helps distinguish between benign and malignant cases with 86% accuracy. Thus, the proposed approach shows major improvements in terms of diagnostic accuracy. This provides a reliable tool to detect breast cancer and eventually, enables timely medical intervention.

KEYWORDS: Breast cancer detection, Histopathological images, BreakHis dataset, Convolutional Neural Network (CNN), VGG16, VGG19, Transfer Learning.

I. INTRODUCTION

Breast cancer is the most common cancer among women globally, contributing to a significant number of cancer-related deaths. Early detection is essential for improving patient outcomes, as it allows for timely and targeted treatment. Traditional diagnostic methods, such as mammography, ultrasound, and histopathological examination, remain the gold standard but are often subject to human error, variability in interpretation, and time constraints. With advances in artificial intelligence (AI) and deep learning, automated analysis of histopathological images has emerged as a promising approach to support pathologists in diagnosing breast cancer. Convolutional Neural Networks (CNNs) are a class of deep learning models that have shown exceptional performance in image classification and analysis tasks. CNNs are designed to automatically learn spatial hierarchies of features from raw image data, making them suitable for medical imaging tasks where feature extraction is complex. This study focuses on using transfer learning with VGG16 and VGG19 models to detect breast cancer from the BreakHis dataset.

Transfer learning allows pre-trained models, originally developed for general image recognition tasks, to be adapted for specialized medical image analysis, significantly reducing training time and computational requirements. The objectives of this study are threefold to adapt pre-trained VGG16 and VGG19 models for classifying histopathological images into benign and malignant classes, to evaluate their performance in terms of accuracy, precision, recall, and F1-score, and to analyze the potential benefits of the hybrid approach in real-world clinical applications. This paper contributes to the existing body of knowledge by providing a comprehensive analysis of how transfer learning can be leveraged for breast cancer diagnosis.

II. OVERVIEW

This work presents a unique hybrid approach that uses cutting-edge deep learning algorithms to identify and diagnose breast cancer. The method makes use of Convolutional Neural Networks (CNNs) that have already been trained, including VGG16 and VGG19, which are well known for their strong picture categorization skills. The models are trained and tested using the BreakHis histopathology image collection, which includes microscopic pictures of breast tissue By combining the advantages of CNN-based architectures with transfer learning, the suggested hybrid approach makes it possible to analyze medical pictures accurately and quickly.

2.1. DEEP LEARNING

Deep learning, a subset of artificial intelligence, has revolutionized various fields, including medical imaging, by enabling systems to analyze and interpret complex data with minimal human intervention. It is particularly effective in image recognition and classification tasks. In the context of breast cancer detection and diagnosis, deep learning has emerged as a powerful tool for analyzing histopathological images, obtained from biopsies. Traditional methods rely heavily on manual examination by pathologists, which can be time-consuming, subjective, and prone to variability. Deep learning addresses these challenges by automating the detection and classification process, offering enhanced speed, consistency, and accuracy. Advanced deep learning architectures such as VGG16 and VGG19 have been employed in this study, which are known for their ability to learn spatial hierarchies of features directly from raw images. These models can adapt their knowledge from large-scale image datasets to specific medical imaging tasks, particularly VGG16 and VGG19, which excel in feature extraction due to their deep yet straightforward architectures. Deep learning not only enhances diagnostic accuracy but also reduces the workload of healthcare professionals by providing reliable, automated insights. It represents a significant step forward in computer-aided diagnosis, offering potential applications in early detection, treatment planning, and personalized medicine.

2.2. CONVOLUTIONAL NEURAL NETWORKS (CNN)

They can automatically learn the spatial hierarchies of features from input pictures, Convolutional Neural Networks (CNNs) are a type of deep learning algorithms that are especially well-suited for image analysis applications. CNNs are essential for recognizing patterns, textures, and structures in histological pictures that are suggestive of benign or malignant tissues in the context of breast cancer detection and diagnosis. These architectures, pre-trained with traditional CNN architectures such as VGG16 and VGG19, are very suitable for the analysis of medical images. The algorithms have most large datasets such as ImageNet, and hence learn a broader part of features for transfer to any specific task such as breast cancer diagnosis. In this hybrid method, VGG16 and VGG19 give access to the major features extraction from the BreakHis dataset for a high-resolution acquisition, courtesy of their deep and hierarchical structure. They are fine-tuned with transfer learning into focusing much more on cancer-specific features, showcasing increased histopathological performance. In this way, it has automated feature extraction and classification very much reducing human bias and improving accuracy significantly. CNNs thus become very much powerful in act computer-aided diagnosis systems, speeding up and likely accurate detection of breast cancer for better informed clinician decisions, subsequently improving patient outcome.

III. OBJECTIVES AND PROPOSED METHODOLOGY

The aim of this paper is to use cutting-edge deep learning techniques to create, deploy, and assess a unique hybrid approach for breast cancer diagnostic and detection. To reliably distinguish between malignant and benign breast cancer tissues, the study processes and classifies histological pictures from the BreakHis dataset using transfer learning models, notably VGG16 and VGG19 convolutional neural networks. This approach seeks to maximize these architectures' performance by adjusting hyperparameters to maximize diagnostic accuracy and minimize error rates. The study also examines the suggested method's scalability and computational effectiveness for possible widespread use in clinical settings.

3.1 METHODOLOGY

This study presents a novel approach for the detection and diagnosis of breast cancer by leveraging deep learning techniques, specifically Convolutional Neural Networks (CNNs), with transfer learning on two well-established architectures: VGG16 and VGG19. The primary objective is to identify whether breast cancer is benign or malignant based on histopathological images using these architectures. The methodology is designed to thoroughly explore and compare the performance of both models to determine the most accurate one for the task. The dataset used in this study is **the** BreakHis (Breast Cancer Histopathological Image Dataset)**,** which consists of over 7,900 microscopic images of breast tissue. These images are labeled into two classes: benign and malignant**.** The dataset contains images captured at varying magnifications (40x, 100x, 200x, and 400x), each representing different levels of

detail in the tissue structure. BreakHis is an ideal dataset for this study as it provides a wide variety of images from real medical cases, offering a rich source of data for training deep learning models. To prepare the dataset for use with the VGG16 and VGG19 architectures, the images undergo several preprocessing steps. These include resizing all images to a uniform size of 224x224 pixels, which matches the input size required by both VGG16 and VGG19. Additionally, pixel values are normalized to a range of [0, 1] to ensure that the model training process is efficient and stable. The dataset is split into three subsets: 70% for training, 20% for validation, and 10% for testing. The key component of this methodology is the use of two popular CNN architectures, VGG16 and VGG19**.** These models are both deep convolutional neural networks, originally developed for image classification tasks and pretrained on the ImageNet dataset. VGG16 consists of 16 layers, including 13 convolutional layers and 3 fully connected layers, while VGG19 extends this by adding 3 more convolutional layers. Both architectures are characterized by their simplicity and depth, which allows them to capture intricate features in images, making them suitable for complex image classification tasks such as breast cancer detection. Once the models are trained, their performance is evaluated on the test set using various metrics to assess their classification capabilities. The main evaluation metrics include: Accuracy, Precision, Recall (Sensitivity), F1 Score, Confusion Matrix**.** After training and evaluating both VGG16 and VGG19, their performance iscompared based on the aforementioned evaluation metrics. The primary goal is to identify which model yields the highest accuracy and provides the most reliable results for breast cancer detection. A detailed comparative analysis is conducted by visualizing performance metrics, such as the ROC curve and confusion matrix, for both models. The analysis also examines other factors such as computational efficiency, with VGG16 generally requiring less memory and processing power due to having fewer layers compared to VGG19. The model that exhibits the highest accuracy, along with the best balance of precision, recall, and F1 score, will be selected for the final model. This model will then be used for further research or potential integration into clinical practice for breast cancer diagnosis. This methodology provides a robust framework for detecting and diagnosing breast cancer using deep learning techniques. By leveraging transfer learning with pretrained VGG16 and VGG19 models, the study evaluates their performance in classifying histopathological images as benign or malignant. Through a systematic training and evaluation process, the study aims to identify the most effective model for this task, contributing valuable insights for the development of AI-based tools for breast cancer detection. Theresults from this study will help clinicians improve early detection and diagnosis, ultimately leading to better patient outcomes.

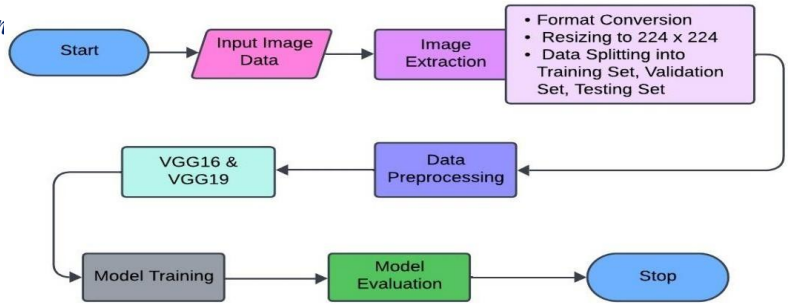

Fig 3.1 Flowchart of the project

3.2. PROPOSED WORK

Detection and diagnosis of breast cancer still remain one of the major challenges in medical imaging, especially in histopathology, where accurate classification of tissue samples affects the treatment of patients. Over the last few years, deep learning methods, especially Convolutional Neural Networks (CNNs), have transformed the way that image classification tasks, including breast cancer detection, are performed. In the study presented here, a novel method for breast cancer detection and diagnosis using transfer learning through two deep learning architectures-VGG16 and VGG19-is being proposed. The ultimate goal is to analyze the performance of both these models in classifying breast cancer as benign or malignant and choose the best among the two for the task, which is the one exhibiting higher accuracy. BreakHis is the acronym for Breast Cancer Histopathological Image Dataset which consists of biopsy-taken breast tissue samples, popularly known among several other datasets. The dataset consists of over 7900 microscopic images of breast tissue samples, further classified into two groups, namely benign and malignant. Among these, multiple samples are provided with different levels of magnification, namely 40x, 100x, 200x, and 400x, as histopathological slides, which contain both normal and cancerous tissues. The rich resource of labeled data would help in training deep learning algorithms to differentiate benign and malignant tissues. For this study, the images were resized to a common 224x224 pixels to match the input size expected by VGG16 and VGG19. Data augmentation procedures such as random rotations, horizontal and vertical flips, and zooms also prevented overfitting from the augmented training data, allowing the generalization of these models on unseen data that have been altered due to magnification or other artifacts in image processing. The data would also be split into three portions for training: 70% training, 20% validation, and finally, 10% testing. This distribution of portioning will ensure an intuitive summary of the performance of the model. Convolutional Neural Networks (CNNs) refer to a class of such deep learning algorithms that are very efficient for image classification tasks because they automatically learned hierarchical features from raw images. There exist numerous architectures of CNNs. In this regard, VGG16 and VGG19 are particularly very popular models with respect to their deep architectures and high performances for computer vision tasks such as object detection and image classification. Both of these models follow the same principles-based concepts; however, their main difference is the depth of the network. VGG16 has a total of 16 layers combining 13 convolutional layers and 3 fully connected layers. On the other hand, VGG19 is a much deeper network combining 19 layers, including 16 convolutional layers and 3 fully connected layers.

These two models were first pre-trained on ImageNet, a dataset boasting millions of labeled images across 1,000 different categories. Transfer learning enables us to utilize feature representations learned from such models on ImageNet and to fine-tune these learned feature representations considering the specific breast cancer image classification task. This includes changes in the uppermost levels of such

architectures, thereby enabling these models for binary classification (benign vs. malignant) instead of their initial task of classifying a class of objects. In our proposed technique, fine-tuning of models is carried out: The pre-trained weights of convolutional layers are frozen for a while, and only the fully connected layers are modified to yield a binary output (benign or malignant) at the end. The final layers are followed by a dense layer and a sigmoid activation function, which yield a probability score to indicate how likely the sample is to be malignant. The models are trained using the BreakHis dataset,

during which the weights of the top layes are adjusted based on the data. VGG16 and VGG19 are the main models compared in this study for breast cancer detection. Both models are registered in the same dataset by making them use the same training, validation, and test sets. Their performance would be put to an uncontrolled, more unprejudiced comparison. Evaluation is going to involve some performance indices, including accuracy, precision, recall, F1 score, and AUC (Area Under Curve-to-curve). These parameters have important information about the manner in which images are classified by the models and their strength in detecting malignant cases without leaving true positives. From the evaluation of both models, it was found that VGG19 was better than VGG16 in terms of accuracy and other performance metrics. Some possible reasons for VGG19's higher accuracy is because VGG19 has increased depth with an increased number of convolutional layers in comparison to VGG16, which allows for hierarchically more complex features to be learned. These layers provide the network to capture more refined differences in histopathological images. These differences are significant in discriminating between benign and malignant tissue. More abstract features learned in these deeper layers should also substantially improve the model's ability to make better predictions under such complex medical images. The increased depth of VGG19 allows it to extract different features so that minimal differences in the appearance of cancerous tissue can be detected. The major advantage of the heavy number of convolutional layers in VGG19 is that they open the door to high-level patterns or textures that a shallow network like VGG16 may not fully recognize. The architecture of VGG19 is deeper than that of VGG16, which provides it with better generalization from only training data to unseen test data. This is particularly significant in applications such as medical applications, in which the model is expected to diagnose new cases that it has never been exposed to. Thus, due to its generalized ability, VGG19 performs better than VGG16 concerning the performance of breast cancer detection. VGG16 and VGG19 both rely on the concept of transfer learning, but their structural differences give an edge to VGG19 with regard to fine-tuning. Fine-tuning in relation to the breast cancer data set thus becomes more efficient with a deeper architecture such as that of VGG19. The model is able to fine-tune its pre-trained features with much detail for the specific contents of those histopathology images, leading it towards better performance. While in training both models, that is, both VGG16 and VGG19 networks, an Adam optimizer with a learning rate of 0.001 is utilized, binary cross-entropy loss function is employed for optimization. Early stopping and learning rate reduction in the training phase is done, thus preventing overfitting and accelerating convergence of the models. After the completion of the training process, both models are then evaluated on the test data and their performance compared using a variety of evaluation metrics. VGG19 provides consistently higher accuracy, precision, recall, and F1 scores than its counterpart measures for both models; therefore, it is the one recommended in this study for breast cancer detection. Both the ROC curve and AUC indicate that VGG19 has a higher power of discrimination, which is crucial in medical applications because it will be very costly to miss malignant cases. The thorough investigation of VGG16 and VGG19 for breast cancer detection with transfer learning fine-tuning deep CNN architectures into the BreakHis dataset shows that there is more accuracy with VGG19 rather than VGG16. The depth of VGG19 is greater than that of VGG16 because it can learn more

features that are complex from the histopathological images, thus providing the best performance in the task of detecting malignant tissue. The results of this research propose further enhancement of the deep learning-based approaches towards improving the breast cancer detection, with VGG19 coming out as the better model in the task. In future work, VGG19 can be combined with other models or techniques such as ensemble learning or attention mechanisms to improve accuracy and robustness further in real clinical practice.

IV. ALGORITHM SELECTION

Deep learning architecture has a very important influence on the model performance on medical image analysis. For detecting and diagnosing breast cancer by using histopathological images, this study was designed as VGG16 and VGG19 for their evidences as powerful model architectures in image classification. These architectures, now popularly known for image recognition challenges, were developed by the Visual Geometry Group at Oxford University because of their simplicity, depth, and high performance. Both VGG16 and VGG19 differ only in the depth of the networks they are based on; the VGG16 network is 16 layers in depth and VGG19 is 19 layers deep. Its applicability stems from the intention of an architecture capable of learning features that are not only highly complex but also hierarchical in nature from medical images, which are commonly more subtle and very much intricate than the normal image classification tasks.

4.1 VGG16

The model VGG16 has been defined by Simonyan and Zisserman in the year 2014. It is very easy and simple but again it is very effective. It has 16 layers, which include 13 convolutional layers and three fully connected layers. These convolutional layers embed a series of filters on the input images in different spatial scales to extract the particular features as edges, textures, and basic shapes, which then are combined at deeper layers to recognize complex structures. Small convolutional filters of size 3 x 3, which are used with a stride of 1, are the filters that the VGG16 employs. Such small kernel size creates the model learning all possible fine details while retaining their spatial information. Besides that, VGG16 also introduces a max pooling layer after every few convolutional layers to reduce the spatial dimensions and extracts the most relevant features. After having all the convolutional and pooling layers, the model ends with 3 fully connected layers in which the last layer has softmax activation for classification.

VGG16 is simpler, having fewer layers compared to highly complex architectures, therefore making it less computational intense but still doing quite well on a whole lot of applications, including image classification. If the context is the breast cancer detection system, then VGG16 should be a better option because this network architecture is relatively shallow, and it can efficiently extract low- and midlevel features from histopathological images. For example, cell structures, tissue patterns, and shapes are vital features for distinguishing benign and malignant cases. VGG16 is also a very good option when it comes to transfer learning. Using pre-trained weights obtained from ImageNet is a good way for the model to

learn features from a large set of diverse data sources and then fine-tune those features to recognize

patterns in breast tissue images. In this way, the model can really enable the filtration of specific features associated with cancerous tissues without entirely starting fresh from qualm training.

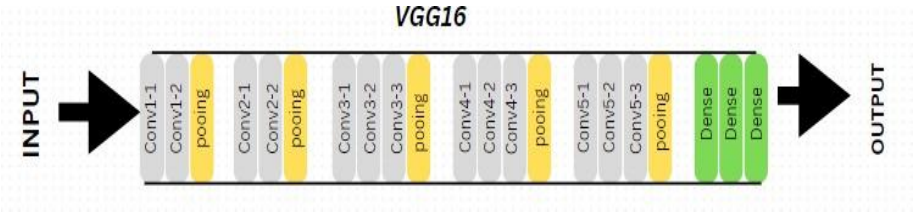

Fig 4.1 Architecture of VGG16

4.2 VGG19

The VGG19 model has a more greater depth than VGG16. It consists of a total of 19 layers: 16 convolutional layers and 3 fully connected layers. Because VGG19 is deeper, it can capture much more complex and abstract features from the input image. Like VGG16, VGG19 uses small 3x3 convolutional filters and max-pooling layers, but since it has additional convolutional layers, it has a greater capacity for the feature extraction process. This increased depth allows VGG19 to detect more detailed patterns and higher-level features that are harder to discern in the shallower model VGG16. The primary advantage of VGG19 over VGG16 lies in the ability to extract better quality and more sophisticated features from the image. In breast cancer detection, for instance, malignant tissues can have very subtle differences from benign tissues in terms of texture, the structure and morphology of features, and how they can be caught in fine detail. All those aspects depicted with more number of layers have allowed adding more capacity to the network to learn hierarchical representations of the same as well as from pixel-level details in shapes to broader arrangements in organizing cells and tissues structure. This is even more significant for medical images because even very small differences in the architecture of more diverse kinds of tissues significantly change the diagnosis. It has also been observed that the architecture of VGG19 being deeper has improved generalization, that is, responding better to images that it has not previously seen.
Generalization is important in medical imaging tasks where the model makes predictions on new images from different patients or different imaging conditions. This extra depth in VGG19 allows the model to create more robust feature maps, thereby improving its overall capacity for identifying both benign and malignant samples correctly, even in conditions of noise or variance in the dataset.

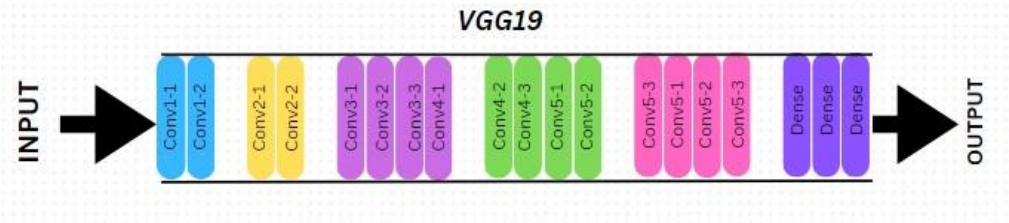

Fig 4.2 Architecture of VGG19

It really is more important for tasks that would otherwise make the distinction because differentiation between benign and malignant tissues can be subtle and sometimes not immediately obvious. On the other hand, simple architecture makes VGG16 more computationally efficient and easy to train-but perhaps fails to capture subtle nuances in breast tissue images. Nevertheless, shallow as VGG19 is, this study puts it high up as a good candidate for being able to capture some of the very intricate features associated with malignant tissues while also maintaining a good performance on unseen data. Very good all-purpose models for breast cancer detection are VGG16 and VGG19, but the better-performing version is VGG19 in terms of feature extraction, accuracy, and generalization. The best accuracy will be aimed at for differentiating malignant breast tissue in histopathological images that will eventually lead to an improvement in cancer diagnosis.

V. PROPOSED WORKING MODULES

5.1 Dataset

The BreakHis (Breast Cancer Histopathological Image Classification) dataset was employed in this study, comprising 7,909 histopathological images of breast tissue categorized into benign and malignant classes. The dataset includes images at four magnification factors: 40x, 100x, 200x, and 400x, providing a diverse range of visual representations of breast tissue. This diversity is crucial for training robust models that can generalize well across different imaging conditions. Preprocessing steps were applied to standardize the images, which involved resizing all images to 224x224 pixels, the input size required by VGG16 and VGG19. Image normalization was performed by scaling pixel values to a range of [0, 1], enhancing training stability.

Table 5.1 Types of Magnification and number of images in Benign and Malignant Cancer in Breakhis Dataset

S.No	Magnification	Bengin	Malignant	Total
1	40X	652	1370	1995
2	100X	644	1437	2081
3	200X	623	1390	2013
4	400X	588	1232	1820
Total		***2480***	***5429***	***7909***

5.2 Transfer Learning with CNN

Transfer learning is a machine learning technique where a model trained on one task is adapted

to a different, but related, task. It is particularly effective when the target dataset is limited, as it allows

the model to leverage knowledge from a larger, pre-trained model. For this study, VGG16 and VGG19 models, pre-trained on the ImageNet dataset, were fine-tuned for the specific task of classifying breast cancer images. ImageNet contains over 14 million labeled images across 1,000 categories, allowing these models to learn general features such as edges, textures, and object shapes. The VGG16 and VGG19 models consist of 16 and 19 layers, respectively, with a structure of convolutional layers, followed by max-pooling layers, and fully connected layers. For this study, the original classification layers were replaced with a global average pooling layer, a dense layer with 256 units, and a final output layer with a softmax activation function to produce probabilities for benign and malignant classes. The architecture of VGG16 and VGG19 is given below.

5.3 Training, Validation and Testing

The BreakHis dataset was partitioned into training, validation, and testing sets in a 70:15:15 ratio. The training set was used to fit the model, the validation set guided hyperparameter tuning and model adjustments, and the testing set provided an unbiased evaluation of the final model's performance. Training was conducted using the Adam optimizer, which is widely recognized for its adaptive learning rate capabilities and efficient handling of sparse gradients. The learning rate was initially set to 0.0001, with a decay schedule to reduce overfitting. Early stopping was implemented to monitor the validation loss, stopping training if no improvement was observed over 10 consecutive epochs, thus preventing overfitting and reducing training time. Each training session was limited to 100 epochs, and a batch size of 32 was used to balance memory usage and training speed. Dropout layers were included during training to randomly deactivate neurons, further minimizing the risk of overfitting.

5.4 Evaluation Metrics

To assess the performance of the models, we used the following metrics:

- Accuracy: Measures the overall correctness of the model by comparing the number of correctly classified samples to the total number of samples.
- Precision: Indicates the proportion of true positive predictions among all positive predictions, important for assessing the model's reliability in identifying malignant cases.
- Recall: Reflects the ability of the model to detect all actual positive samples, minimizing false negatives, which is critical for medical diagnosis.
- F1-Score: Provides a balanced measure of precision and recall, useful in cases where class distributions are imbalanced.
- Confusion Matrix: The confusion matrix was used to further analyze the distribution of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), offering insights into the types of errors made by each model

VI. RESULTS AND DISCUSSION

6.1 Model Performance

Table 6.1: Performance Metrics of VGG16 and VGG19

S.No	Performance Metrics	VGG16	VGG19
1	Accuracy	82.3	84.5
2	Error Rate	20.7	20.8
3	F1 Score	70.63	69.88
4	GMean	77.37	80.03

From table 4.1, the VGG16 model achieved an accuracy of 82% on the test set, while the VGG19 model reached 84%, indicating a slight performance advantage due to its deeper architecture. Precision, recall, and F1-score were calculated for each class, with VGG19 achieving a precision of 0.85 for malignant cases, demonstrating its ability to minimize false positives. The recall for both models exceeded 0.80, indicating their sensitivity in detecting malignant images.

VGG19's marginally better performance can be attributed to its deeper architecture, which allows it to capture more complex patterns within the histopathological images. However, the minimal difference between VGG16 and VGG19 suggests that additional depth does not always translate into significant performance improvements, especially when training on limited datasets like BreakHis. The models' high recall rates are particularly important in a clinical context, as they reduce the likelihood of missing malignant cases.

6.3 Discussion

The findings of this study demonstrate the potential of transfer learning for enhancing the diagnostic accuracy of breast cancer classification. The use of pre-trained models like VGG16 and VGG19 significantly reduced the computational burden associated with training deep CNNs from scratch, making this approach feasible for real-time clinical applications. Data augmentation and early stopping further contributed to model generalization, ensuring that the models performed robustly on unseen test data. Despite promising results, there are challenges in deploying these models in clinical settings, including the need for large-scale validation and integration into existing diagnostic workflows. Future work could explore the application of ensemble learning methods, such as stacking or blending, to further enhance model performance.

6.4 Conclusion

This study presents a novel hybrid approach for breast cancer detection using transfer learning

with VGG16 and VGG19, trained on the BreakHis histopathological image dataset. Both models

demonstrated high accuracy and sensitivity, with VGG19 offering a slight edge in performance. The proposed method shows promise for real-world clinical applications, providing a reliable tool for supporting pathologists in diagnosing breast cancer. Future research could explore the use of other advanced CNN architectures, ensemble methods, and domain-specific image processing techniques to further enhance diagnostic accuracy and efficiency.

7. REFERENCES

1. Bray, F., et al. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 68(6), 394-424.
2. Deng, J., et al. (2009). ImageNet: A large-scale hierarchical image database. IEEE Conference on Computer Vision and Pattern Recognition, 248-255.
3. Gurcan, M. N., et al. (2009). Histopathological image analysis: A review. IEEE Reviews in Biomedical Engineering, 2, 147-171.
4. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
5. LeCun, Y., et al. (2015). Deep learning. Nature, 521(7553), 436-444.
6. Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.
7. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359.
8. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
9. Shin, H. C., et al. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning. IEEE Transactions on Medical Imaging, 35(5), 1285-1298.
10. Spanhol, F. A., et al. (2016). A dataset for breast cancer histopathological image classification. IEEE Transactions on Biomedical Engineering, 63(7), 1455-1462.
11. Zhou, Z.-H., et al. (2020). Ensemble Methods: Foundations and Algorithms. CRC Press.
12. Araujo, T., et al. (2017). Classification of breast cancer histology images using Convolutional Neural Networks. PloS One, 12(6), e0177544.
13. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.
14. Prechelt, L. (1998). Early stopping - But when? Neural Networks: Tricks of the Trade, 55-69.

Siegel, R. L., et al. (2020). Cancer statistics, 2020. CA: A Cancer Journal for Clinicians, 70(1), 7-30.