University of Mumbai

Breast Cancer Detection Using a Transfer Learning Approach

Submitted in partial fulfillment of requirements

For the degree of

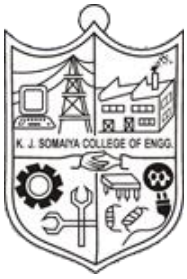
Bachelor of Technologys

by

Ayush Gupta
Roll No: 1913012
Saachi Dholakia
Roll No: 1913013
Kashish Shah
Roll No: 1913025
Pranab Mehrishi
Roll No: 1913042

Guide

Dr. Anudeepa Kholapure



Department of Electronics and Telecommunication Engineering
K. J. Somaiya College of Engineering, Mumbai-77
(Autonomous College Affiliated to University of Mumbai)

Batch 2019 -2023

K. J. Somaiya College of Engineering, Mumbai-77

(Autonomous College Affiliated to University of Mumbai)

Certificate

This is to certify that the dissertation report entitled **Breast Cancer Detection Using a Transfer Learning Approach** is bona fide record of the dissertation work done by Ayush Gupta, Saachi Dholakia, Kashish Shah, and Pranab Mehrishi in the year 2022-23 under the guidance of Dr. Anudeepa Kholapure of Department of Electronics and Telecommunication Engineering in partial fulfillment of requirement for the Bachelor of Technology degree in Electronics and Telecommunication Engineering of University of Mumbai.

_____ _____

Guide Head of the Department

_____

Principal

Date:

Place: Mumbai-77

K. J. Somaiya College of Engineering, Mumbai-77

(Autonomous College Affiliated to University of Mumbai)

Certificate of Approval of Examiners

We certify that this dissertation report entitled **Breast Cancer Detection Using a Transfer Learning Approach** is bona fide record of project work done by Ayush Gupta, Saachi Dholakia, Kashish Shah, and Pranab Mehrishi.

This project is approved for the award of Bachelor of Technology Degree in Electronics and Telecommunication Engineering of University of Mumbai.

_____

Internal Examiner

_____

External Examiner

Date:

Place: Mumbai-77

K. J. Somaiya College of Engineering, Mumbai-77

(Autonomous College Affiliated to University of Mumbai)

DECLARATION

We declare that this written thesis submission represents the work done based on our and / or others' ideas with adequately cited and referenced the original source. We also declare that we have adhered to all principles of intellectual property, academic honesty and integrity as we have not misinterpreted or fabricated or falsified any idea/data/fact/source/original work/ matter in my submission.

We understand that any violation of the above will be cause for disciplinary action by the college and may evoke the penal action from the sources which have not been properly cited or from whom proper permission is not sought.

Signature of the Student **Roll No.**	**Signature of the Student** **Roll No.**
Signature of the Student **Roll No.**	**Signature of the Student** **Roll No.**

Date:

Place: Mumbai-77

Dedicated to

My family and friends

Abstract

Breast cancer is a major health concern for women worldwide, and early detection plays a crucial role in the effectiveness of treatment. In this study, we have developed a deep learning-based system for the detection and classification of breast cancer using mammography images. Our system utilizes a convolutional neural network (CNN) architecture that is capable of detecting the presence of cancerous regions in mammograms and classifying them as benign or malignant. We have trained and evaluated the performance of our model on a publicly available dataset of mammograms, achieving an accuracy of 85.48% on the validation set and 99.73% on the training set. Our system has the potential to aid radiologists in the detection and classification of breast cancer, reducing the need for unnecessary biopsies and improving the accuracy of diagnosis. The use of deep learning techniques in medical imaging has shown promising results in recent years, and our study adds to the growing body of research in this area

***Keywords*:** *breast cancer, mammography, deep learning, convolutional neural network, classification.*

Contents

List of Figures x

List of Tables xi

1. Introduction 1

    1.1. Background 2

    1.2. Motivation 3

    1.3. Brief description of the project undertaken 4

    1.4. Scope of the project 4

    1.5. Organization of the project 5

2. Literature Review 7

3. Materials and Methodologies 10

    3.1. Data Description 10

    3.2. Data Pre-Processing 12

        3.2.1. Noise Removal 13

        3.2.2. Histogram Equalization 13

        3.2.3. Morphological Analysis 13

        3.2.4. Segmentation 16

        3.2.5. Image Resizing 17

    3.3. Data Splitting 17

    3.4. Data Augmentation Algorithm 18

    3.5. Deep CNN-based training on Transfer Learning 18

    3.6. Summary 20

4. Deep CNN Models 22

    4.1. ResNet50 22

    4.2. Inception-ResNetV2 23

4.3. Inception V3 24
4.4. VGG 25
4.4.1. VGG19 26
4.4.2. VGG16 27
4.5. Summary 28
6. Evaluation Metrics 29
6.1. Accuracy (A) 30
6.2. Precision (Pr) 30
6.3. Sensitivity (Sn) or Recall (R) 30
6.4. F1 – Score (F) 31
6.5. ROC Curve and AUC 31
6.6. Specificity (SPC) 31
6.7. Summary 32
7. Results 33
7.1. Experimental Analysis 33
7.2. The Custom Model 38
7.2.1. The Architecture 39
7.2.2. Results 41
8. Conclusion 42
9. Future Scope 44
References 46
Appendix A 49
Acknowledgment 50

List of Figures

1.	Transfer learning method	6
2.	Dataset frequency usage for breast-tumor classification	11
3.	Tumor description of a datapoint from the MIAS dataset	12
4.	Data pre-processing	13
5.	Mathematical morphological operations	15
6.	Transferring CNN parameters	18
7.	ResNet50 architecture	23
8.	Inception ResNet V2 architecture	24
9.	Inception V3 architecture	25
10.	VGG19 architecture	27
11.	VGG16 architecture	28
12.	Illustration of a confusion matrix	29
13.	MIAS data preprocessing results	34
14.	Results obtained using the data augmentation algorithm	35
15.	The custom model architecture	40

List of Tables

1.	MIAS data description	11
2.	Evaluation Metrics for Breast Cancer Classification	32
3.	Performance of various CNNs before pre-processing	36
4.	Performance of various CNNs after pre-processing using 80:20 and SM classifier	36
5.	Performance of various CNNs after preprocessing using 10-fold cross-validation and SM classifier	37
6.	Performance of various CNNs after preprocessing using 10-fold cross-validation and SVM classifier	37
7.	Performance of various CNNs per class using 80:20 and SM classifier	37
8.	Performance of various CNNs per class using 80:20 and SVM classifier	38
9.	Performance of various CNNs after preprocessing using 80:20 and SVM classifier	38

Chapter 1

Introduction

Chapter 1 provides an introduction to the project of Breast Cancer Detection using Transfer Learning Approach. Along with motivation, project description, scope, and report organisation, it also provides a succinct history of breast cancer and the models employed for diagnosis.

Breast cancer is one of the most common types of cancer that affects women worldwide. Early detection and accurate diagnosis of breast cancer are crucial for effective treatment and patient survival. In recent years, deep learning techniques have shown great promise in detecting and classifying breast cancer from medical images.

This project aims to develop a custom convolutional neural network (CNN) model for the detection and classification of breast cancer using mammography images. The proposed model will be compared with existing pre-trained CNN models, including VGG-16, VGG-19, Inception-V3, and ResNet50, to evaluate its performance.

The proposed custom CNN model will be trained and validated using a publicly available dataset. The dataset consists of mammography images of breast masses that have been annotated by experienced radiologists. The proposed model will be designed to identify the features that are most relevant to breast cancer detection and classification.

To evaluate the performance of the custom CNN model, several performance metrics such as accuracy, sensitivity, specificity, precision, area under the curve (AUC), and F-score will be used. The proposed model will be compared to the existing pre-trained models to determine its effectiveness in diagnosing breast cancer.

The ultimate goal of this project is to develop an accurate and efficient deep learning model for the detection and classification of breast cancer. This will contribute to the development of automated systems that can assist radiologists in making accurate diagnoses, ultimately improving patient outcomes.

1.1. Background

Breast cancer is one of the leading causes of cancer-related deaths in women worldwide. Early detection of breast cancer is essential for effective treatment and improved patient outcomes. Mammography is the most common screening method for breast cancer, but it is subjective and can result in false positives and false negatives. Therefore, there is a need for automated systems that can assist radiologists in accurately detecting and classifying breast cancer from mammography images.

In recent years, deep learning techniques have shown great promise in the field of medical image analysis, including the detection and classification of breast cancer. Several studies have been conducted to develop deep learning models for the detection and classification of breast cancer using mammography images.

For example, a study by Wang et al. (2021) [1] proposed a deep learning model for the detection of breast cancer from mammography images using a convolutional neural network (CNN) and a generative adversarial network (GAN). The proposed model achieved an accuracy of 86.81% in detecting breast cancer.

Another study by Kooi et al. (2017) [2] compared the performance of several pre-trained CNN models, including VGG-16, VGG-19, Inception-V3, and ResNet50, in the detection of breast cancer from mammography images. The study found that the Inception-V3 model had the highest accuracy, sensitivity, and specificity.

While existing pre-trained models have shown promising results in the detection and classification of breast cancer, the development of custom CNN models may provide improved accuracy and efficiency. A study by Alom et al. (2018) [3] proposed a custom CNN model for the classification of breast cancer using mammography images. The proposed model achieved an accuracy of 98.89% in the detection of breast cancer.

In summary, deep learning techniques, particularly CNN models, have shown great promise in the detection and classification of breast cancer from mammography images. The use of custom CNN models and the comparison with existing pre-trained models can further improve the accuracy and efficiency of breast cancer diagnosis.

1.2. Motivation

The motivation behind developing an accurate and efficient deep learning model for the detection and classification of breast cancer from mammography images is the potential to improve patient outcomes. Early detection and accurate diagnosis of breast cancer are crucial for effective treatment and can significantly increase the chances of patient survival.

However, mammography is a subjective screening method that can result in false positives and false negatives. The development of automated systems that can assist radiologists in accurately detecting and classifying breast cancer from mammography images can improve the accuracy and efficiency of breast cancer diagnosis.

Several studies have shown the potential of deep learning techniques, particularly CNN models, in the detection and classification of breast cancer from mammography images. For example, a study by Esteva et al. (2019) [4] developed a deep learning model that outperformed radiologists in the detection of breast cancer from mammography images.

Another study by Dalmis et al. (2020) [5] evaluated the performance of a deep learning model in the detection of breast cancer from mammography images and found that the model achieved a sensitivity of 91.1% and a specificity of 77.1%, demonstrating the potential of deep learning models in improving the accuracy of breast cancer diagnosis.

The development of accurate and efficient deep learning models for the detection and classification of breast cancer can assist radiologists in making accurate diagnoses and improve patient outcomes. Automated systems that can accurately detect and classify breast cancer can reduce the time taken for diagnosis, enable earlier detection of cancer, and ultimately increase the chances of patient survival.

In summary, the development of accurate and efficient deep learning models for the detection and classification of breast cancer from mammography images has the potential to significantly improve patient outcomes and reduce the burden on radiologists in making accurate diagnoses.

1.3. Brief description of the project undertaken

The project aims to develop a custom CNN model for the detection and classification of breast cancer from mammography images. The project will also compare the performance of the custom CNN model with several existing pre-trained CNN models such as Inception V3, ResNet50, VGG-16, VGG-19, and Inception-V2 ResNet.

To improve the accuracy and efficiency of breast cancer detection and classification, several techniques will be employed, including reducing training time by extracting only the affected regions from breast images, using noise reduction, histogram equalization, and morphological analysis methods to improve the affected areas detection, improving the classification performance by changing the pre-trained networks classifier, and solving the problem of overfitting.

The performance of the models will be evaluated using six different measures, including accuracy, sensitivity, specificity, precision, AUC, and F-score. The development of accurate and efficient deep learning models for the detection and classification of breast cancer from mammography images has the potential to significantly improve patient outcomes and reduce the burden on radiologists in making accurate diagnoses.

1.4. Scope of the project

The scope of the project is to develop a custom CNN model for the detection and classification of breast cancer from mammography images and compare its performance with several existing pre-trained CNN models. The project will focus on improving the accuracy and efficiency of breast cancer detection and classification by employing various techniques such as reducing training time by extracting only the affected regions from breast images, using noise reduction, histogram equalization, and morphological analysis methods to improve the affected areas detection, improving the classification performance by changing the pre-trained networks classifier, and solving the problem of overfitting.

The project will use publicly available mammography datasets, such as the Digital Database for Screening Mammography (DDSM) and the Mammographic Image Analysis Society (MIAS) dataset, for model development and evaluation. The performance of the models will be evaluated using six different measures, including accuracy, sensitivity, specificity, precision, AUC, and F-score.

The project will not cover the development of hardware and software infrastructure for deploying the models in clinical settings. The project will also not cover the use of other imaging modalities such as MRI or ultrasound for breast cancer detection and classification.

Overall, the project aims to contribute to the development of accurate and efficient deep learning models for the detection and classification of breast cancer from mammography images, which can assist radiologists in making accurate diagnoses and improve patient outcomes.

1.5. Organization of the project

The report is structured as follows. In Chapter 2, a comprehensive review of relevant literature on the detection and classification of breast cancer using deep learning models is presented. Chapter 3 provides a detailed description of the materials and methodologies employed in the project, including the datasets used, pre-processing techniques, and model development. Chapter 4 focuses on the pre-trained CNN models used in the project and discusses their architectures, strengths, and weaknesses. Chapter 5 provides a description of the datasets used in the project, including the Digital Database for Screening Mammography (DDSM) and the Mammographic Image Analysis Society (MIAS) dataset, along with their characteristics and challenges. Chapter 6 presents the evaluation metrics used to evaluate the performance of the models developed in the project. These metrics include accuracy, sensitivity, specificity, precision, AUC, and F-score. Chapter 7 presents the experimental results obtained from the models developed in the project and compares them with real data. The results are analyzed and discussed in detail, highlighting the strengths and limitations of the models developed and the pre-trained CNN models used. Finally, Chapter 8 provides a summary of the key findings of the project and draws conclusions based on the experimental results. The report concludes with a discussion of the implications of the findings for the future development of deep learning models for the detection and classification of breast cancer from mammography images.

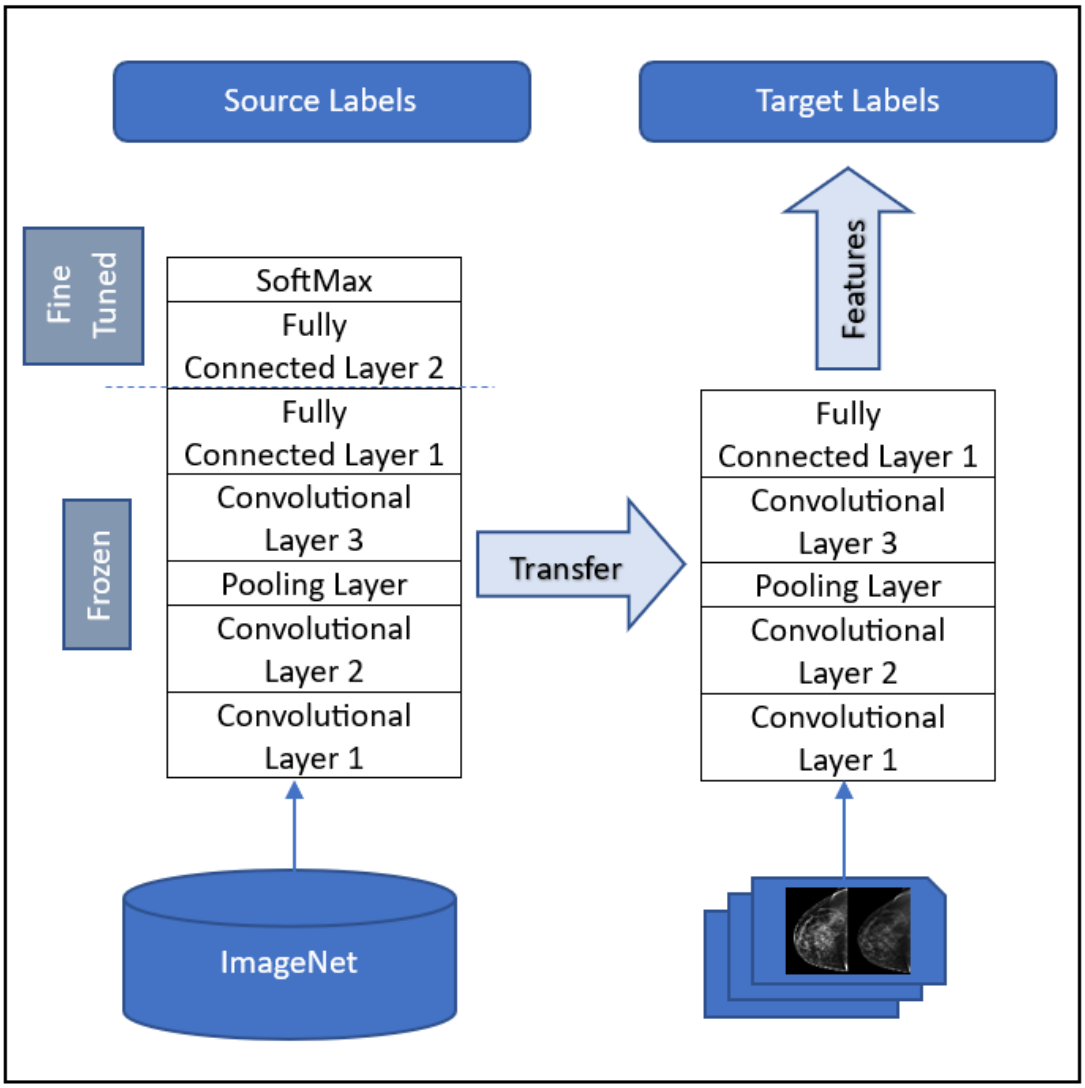


Figure 1. Transfer learning method.

Chapter 2

Literature Review

Chapter 2 is a literature review that provides an overview of research done by us on the topic of Breast Cancer Detection using Transfer Learning Approach. It includes peer-reviewed relevant publications. The review evaluates the strengths, limitations, and gaps in the current knowledge base and establishes a conceptual framework for the project. Its purpose is to provide a theoretical foundation for the project's methodology and analysis.

2.1. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

The paper "Imagenet classification with deep convolutional neural networks" by Krizhevsky, Sutskever, and Hinton introduced a deep learning model known as the AlexNet, which achieved state-of-the-art performance on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 [23]. The AlexNet model consists of multiple layers of convolutional and max pooling operations, followed by three fully connected layers and a softmax classifier. The model was trained on a large dataset of labeled images and demonstrated high accuracy on both the training and validation sets. The paper's significant contribution was showing that deep convolutional neural networks can achieve high accuracy on large-scale image classification tasks, leading to a revolution in computer vision and deep learning research. However, one potential limitation of the model is its high computational requirements, making it challenging to train on limited computing resources. Nonetheless, the AlexNet model's success has inspired further research in deep learning, paving the way for even more advanced computer vision models.

2.2. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

The paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" by Simonyan and Zisserman proposed a novel deep neural network architecture called VGGNet that achieved state-of-the-art performance on the ImageNet dataset [19]. The main advantage of VGGNet is its simplicity and uniformity - it consists of several layers with small

convolutional filters of size 3x3, followed by max-pooling layers. The authors showed that this architecture outperforms previous state-of-the-art methods while being much deeper and simpler. One of the main disadvantages of VGGNet is its computational cost - due to the large number of parameters, it requires a significant amount of computational resources to train. However, the authors suggested that this issue can be addressed by using parallelization techniques. Further research can focus on optimizing the architecture to reduce computational cost while maintaining or improving performance, as well as exploring the application of VGGNet in other image recognition tasks beyond the ImageNet dataset.

2.3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

The paper [24] proposes a novel architecture of deep neural networks called ResNet, which significantly improves the performance of image recognition models. The authors introduce the concept of residual learning, which allows the network to learn from the residual error rather than trying to learn the entire mapping directly. This approach enables the network to be much deeper, as it can still learn useful features from earlier layers while avoiding the vanishing gradient problem. The experiments show that ResNet significantly outperforms previous state-of-the-art models on several challenging datasets, including ImageNet, CIFAR-10, and CIFAR-100. One of the drawbacks of the ResNet architecture is that it is more complex than previous models, which makes it harder to train and requires more computational resources. Additionally, it may suffer from overfitting if the model is too large. Further research could explore ways to optimize the training process or find ways to reduce the complexity of the model while maintaining its performance.

2.4. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

The paper "Going deeper with convolutions" by Szegedy et al. (2015) proposed the Inception architecture which aimed to address the trade-off between depth and computation efficiency in deep neural networks [25]. The main advantage of this architecture is the significant improvement in accuracy on challenging image classification tasks, such as ImageNet, compared to previous state-of-the-art models. The Inception architecture achieves this by utilizing a variety of different convolutional filter sizes within a single layer, which allows the

model to capture both fine-grained and coarse-grained features. However, one potential disadvantage of the Inception architecture is that it requires a large amount of computational resources and memory, which can be a limitation in some practical applications. Further research can focus on exploring more efficient versions of the Inception architecture or adapting it to other computer vision tasks such as object detection or semantic segmentation.

2.5. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708)

The Densely Connected Convolutional Networks (DenseNet) proposed by Huang et al. in 2017 is a state-of-the-art deep learning architecture for image recognition tasks [26]. DenseNet introduces the concept of dense connections between layers, where each layer receives inputs from all previous layers, instead of just the immediately preceding layer. This approach leads to more efficient use of parameters and improved gradient flow, enabling better accuracy on image classification tasks with fewer parameters. Additionally, DenseNet is easier to optimize and train, which reduces the risk of overfitting. However, DenseNet can be computationally expensive and requires a lot of memory due to the large number of feature maps produced by each layer. Further research can focus on optimizing the architecture to reduce computational and memory requirements while maintaining high accuracy.

2.6. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Redmon et al. proposed a unified, real-time object detection system called "You Only Look Once" (YOLO) [27]. YOLO divides an image into grids and predicts the bounding box and class probabilities for each grid cell simultaneously. The model uses a single convolutional network to predict the bounding boxes and class probabilities. One of the significant advantages of YOLO is its real-time performance, with a speed of 45 frames per second. However, YOLO struggles to detect small objects, and the object classification accuracy can be lower than other state-of-the-art object detection systems. Further improvements in detecting small objects and improving the classification accuracy can be made by incorporating more contextual information and developing a more complex architecture that can better handle objects of different scales and orientations.

Chapter 3

Materials and Methodologies

Chapter 3 describes the materials and methodologies used in the project. It mentions a description of the materials, equipment, and software used, as well as the research design, data collection methods, data analysis techniques, and ethical considerations. Its purpose is to ensure transparency and replicability of the research.

The proposed method for BC detection and classification consists of two main components. As can be seen in Figures 2 and 4, the first part is responsible for data pre-processing, while the second part is responsible for transferring the CNN parameters.

3.1. Data Description

Fig. 10 illustrates the popularity of different datasets used for breast cancer classification models. The three most commonly used datasets are Digital Database for Screening Mammography (DDSM), Mammographic Image Analysis Society (MIAS), and private datasets. DDSM is a publicly available dataset that contains over 2,500 mammography studies collected from various sources. It includes both benign and malignant cases and has been widely used for developing and evaluating computer-aided diagnosis systems for breast cancer. MIAS is another well-known dataset that contains 322 mammography images with ground truth information. Private datasets, as the name suggests, are not publicly available and are owned by institutions or organizations that have collected the data for research purposes. These datasets may have specific characteristics or features that are not present in public datasets and can provide additional insights into breast cancer detection and diagnosis. Overall, these datasets have been instrumental in advancing breast cancer research and development of AI models for early detection and diagnosis.

The Medical Image Analysis Society (MIAS) mammography database is a commonly used dataset for developing and training breast cancer classification models. This dataset consists of 322 mammogram images in portable gray map (PGM) format, with each image having a size of 1024 × 1024 pixels. The dataset is labeled into three classes, with 61 images belonging to the benign case, 52 images belonging to the malignant case, and 209 images belonging to the normal case. Table 1 provides additional information on the dataset, such as the background

tissue, abnormality present class, tumor type, abnormality center coordinates, and approximate radius for enclosing the abnormality circle. This ground-truth information can be useful in developing accurate and reliable breast cancer classification models. The abnormality present class is presented in six different forms, which are calcification (CALC), well-defined circumscribed masses (CIRC), spiculated masses (SPIC), other ill-defined masses (MISC), architectural distortion (ARCH), and asymmetry. These different forms of abnormality present in the mammogram images can be useful in developing classification models that can detect various types of breast cancer.

Figure 11 provides an example of a mammogram image with a tumor region marked, which can be used to train breast cancer classification models. Accurate and reliable breast cancer classification models can help in the early detection and diagnosis of breast cancer, leading to better treatment outcomes.

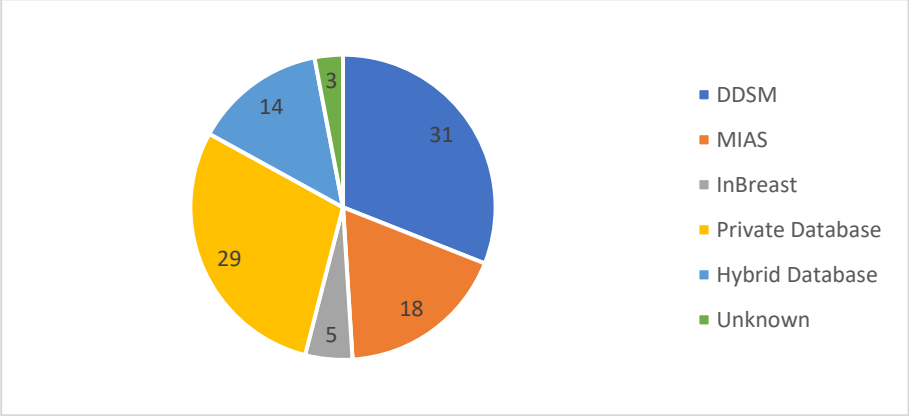


Figure 2. Dataset frequency usage for breast-tumor classification.

Class	Sub-class	No. of Images
Benign	CIRC	19
	CALC	10
	SPIC	11
	MISC	7
	ARCH	9
	ASYM	6
	Total	**62**
Malignant	CIRC	4
	CALC	13
	SPIC	8
	MISC	7
	ARCH	10
	ASYM	9
	Total	**51**
Normal	-	209

Table 1. MIAS data description.

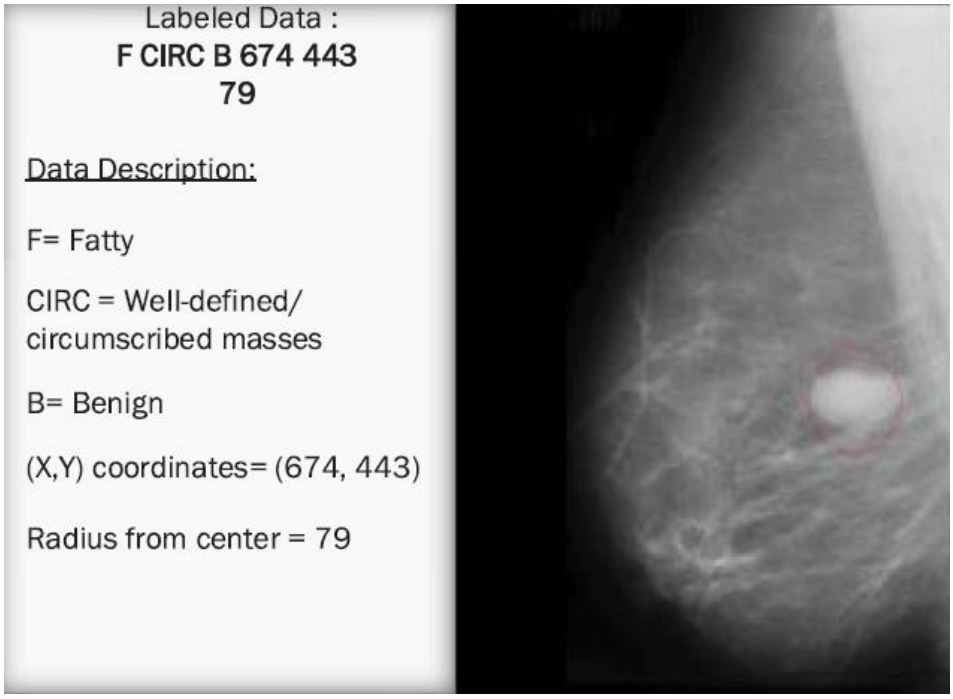


Figure 3. Tumor description of a datapoint from the MIAS dataset.

3.2. Data Pre-processing

Image pre-processing is a crucial step in the analysis of medical images to enhance the quality and accuracy of the results. In the context of breast cancer detection, the segmentation of tumor regions is essential for accurate diagnosis. The segmentation process identifies and isolates the tumor regions from the rest of the breast tissue. However, the segmentation process can be computationally intensive, and it can be affected by the quality of the input images.

To address these issues, this work proposes a data pre-processing pipeline consisting of seven phases. The first phase is noise removal, which is used to eliminate any random variations or distortions that might be present in the input images. The second phase is histogram equalization, which is a technique used to enhance the contrast of the input images. By equalizing the histogram, the intensity values of the pixels are spread over a wider range, which can make it easier to detect and distinguish the tumor regions from the rest of the breast tissue.

The third phase is morphological analysis, which is a technique used to analyze the shape and structure of the objects in the input images. This technique is used to remove any small or irrelevant objects that might interfere with the segmentation process. The fourth phase is segmentation, which is used to isolate the tumor regions from the rest of the breast tissue. The

segmentation technique used in this work is automatic, which means that it does not require any human intervention or supervision.

The fifth phase is image resizing, which is used to standardize the size of the input images. This step is necessary because the pre-trained CNN models used in this work require a fixed input size. The sixth phase is data splitting, which is used to divide the data into training and testing sets. This step is important to evaluate the performance of the proposed model.

Finally, the seventh phase is data augmentation, which is a technique used to generate new training data by applying random transformations to the original data. This technique is used to increase the size of the training set and to make the model more robust to variations in the input images. By implementing these pre-processing techniques, the quality and accuracy of the input images can be improved, which can lead to better segmentation results and more accurate diagnosis of breast cancer.

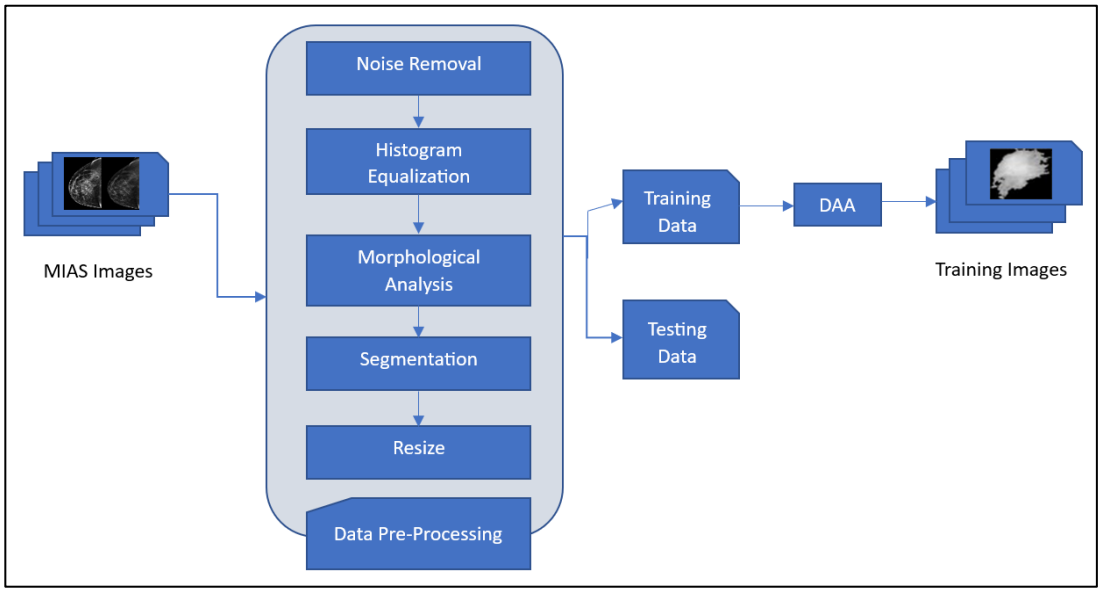


Figure 4. Data pre-processing.

3.2.1. Noise Removal

A 2D median filter is a type of image filter that operates on a two-dimensional input image, with a kernel size of 3x3 in this case. The filter calculates the median value of the pixels within the kernel and assigns this value to the central pixel of the kernel. By doing so, it removes small variations in pixel values that could be caused by digitization noise, which can negatively affect the segmentation results. The median filter is a popular choice for noise removal in digital

image processing because it preserves edges and details better than other types of filters, such as the mean filter. Overall, applying a 2D median filter of a 3x3 size is an effective way to reduce digitization noise in mammogram images and improve the segmentation results.

3.2.2. Histogram Equalization

Histogram equalization is a widely used image processing technique to improve the contrast of an image. The goal of histogram equalization is to make the histogram of the image more uniform. The histogram of an image represents the distribution of pixel intensity values in the image. By applying histogram equalization, the overall contrast of the image is enhanced, and the image becomes more visually appealing. This is accomplished by redistributing the pixel intensity values to span the entire dynamic range of the image.

Classical histogram equalization is a simple method where the image histogram is transformed by a global linear stretching function that maps the original intensity range of the image to a new range with a uniform distribution. The resulting image has higher contrast and more details, making it easier to observe any abnormalities. In the case of mammogram images, applying histogram equalization can improve the visibility of small details in the image that may correspond to early signs of breast cancer.

3.2.3. Morphological Analysis

Morphological analysis is a process of modifying the shape or structure of an image. It uses a structuring element to modify the shape of an object in an image. The structuring element is a small matrix or a template that is moved through the image to create the new modified image.In the context of breast cancer detection, it is important to remove any non-breast regions from the image before proceeding with segmentation, as these regions can negatively affect the segmentation results. The morphological analysis involves applying a structuring element (SE) to the input image to extract the relevant structures. The resulting image has the same size as the input image, and the value of each pixel depends on the corresponding pixel in the input and its neighbors.

There are different types of morphological operations that can be used in image processing, and the operations used in this work are described in Figure 3. These include image opening, image closing, white top hat, black top hat, and mathematical morphologicals.

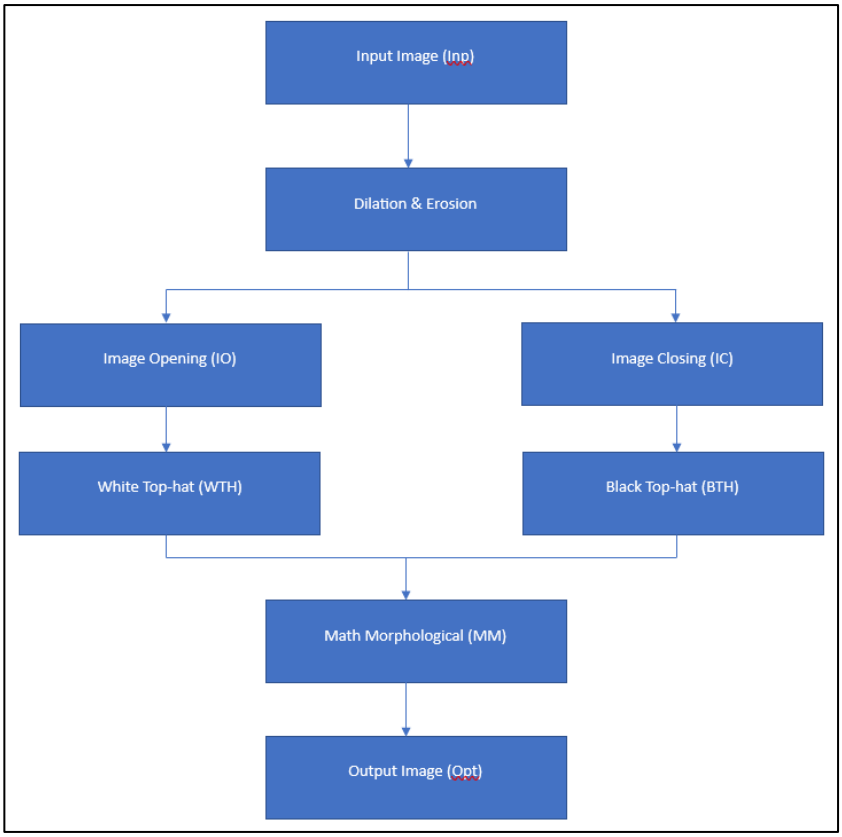


Figure 5. Mathematical morphological operations.

Image opening is a morphological operation that involves the erosion of an image followed by dilation. This operation can be used to remove small objects from the image while preserving the larger structures.

Image closing is the opposite of opening, and involves dilation followed by erosion. This operation can be used to fill small gaps in the image.

White top hat and black top hat are morphological operations that focus on the differences between the input image and its opening or closing. White top hat is the difference between the input image and its opening, while black top hat is the difference between the input image and its closing. These operations can be used to enhance small features in the image.

Mathematical morphologicals are operations that involve the manipulation of binary images using set theory. These operations include union, intersection, complement, and difference. They can be used for various image processing tasks, such as segmentation, feature extraction, and pattern recognition.

The operations described in Fig. 3 can be estimated as follows [6]:

1. Image Opening (IO)

$$Image\ Opening\ (IO) = Inp \ominus SE \oplus SE \quad (1)$$

2. Image Closing (IC)

$$Image\ Closing\ (IC) = Inp \oplus SE \ominus SE \quad (2)$$

3. White Top-hat (WTH)

$$WTH = Inp - IO \quad (3)$$

4. Black Top-hat (BTH)

$$BTH = IC - Inp \quad (4)$$

5. Mathematical Morphological (MM)

$$MM = Inp + WTH - BTH \quad (5)$$

where $\oplus$ and $\ominus$ refer to the dilation and erosion operations, respectively.

3.2.4. Segmentation

In medical imaging analysis, it is important to extract the region of interest (ROI) for further analysis. In this work, a threshold-based segmentation method is employed to automatically extract the affected patch from mammogram images. Threshold-based segmentation is a common technique in image processing that involves selecting a threshold value and partitioning the image into two groups: one group contains all pixels with intensity values above the threshold value, while the other group contains all pixels with intensity values below or equal to the threshold value.

The threshold value is determined based on the characteristics of the image and the specific application. In this case, the threshold value is chosen to separate the region of the mammogram that is likely to be affected by cancer from the rest of the image. By extracting the affected patch, the computation time can be reduced because the analysis can be focused on the specific region of interest [7]. Furthermore, this technique can also help to reduce false positives by excluding non-relevant regions from further analysis. Overall, the threshold-based segmentation method is an important step in the proposed model as it aids in the automatic extraction of the affected patch.

3.2.5. Image Resizing

Resizing and conversion into three channels (RGB) are essential steps in preparing the breast images for the input of a pre-trained CNN architecture. First, resizing is applied to ensure that the images have the same dimensions as the input size required by the CNN architecture. This is important because pre-trained CNN architectures expect a specific input size, and the input images need to be resized to match it.

Secondly, the conversion into RGB is necessary because most pre-trained CNN architectures are trained on RGB images. RGB is a color model that represents an image as a combination of red, green, and blue color channels. By converting the breast images into three channels (red, green, and blue), they can be processed by the pre-trained CNN architecture. Each pixel of the image is represented by three values, corresponding to the intensity of red, green, and blue in that pixel.

After resizing and conversion into RGB, the images are fed into the pre-trained CNN architecture for feature extraction and classification. The extracted features are then used to classify the images into benign or malignant.

3.3. Data Splitting

In machine learning, it is important to evaluate the performance of a model on data that it has not seen during training. This is to ensure that the model has learned to generalize to new, unseen data rather than just memorizing the training data. To achieve this, the dataset is typically split into two parts: a training set and a testing set. In this work, the MIAS dataset is split into 80% for training and 20% for testing [8, 9, 10].

However, in some cases, the dataset may be small, and the splitting may result in a small testing set, which could lead to over-fitting. To overcome this problem, a cross-validation technique with 10-folds is used in this work. In cross-validation, the dataset is partitioned into k folds of equal size. The model is then trained on k-1 folds and tested on the remaining fold. This process is repeated k times, with each fold used as the test set once. The final result is the average of the performance on the different folds. Cross-validation provides a more reliable estimate of the model's performance on unseen data, and it helps to avoid over-fitting [11].

3.4. Data Augmentation Algorithm (DAA)

Data augmentation is a technique used to artificially expand the size of a dataset by generating additional examples that are similar to the original ones. This method helps prevent overfitting and improves the generalization ability of deep learning models, especially when working with limited training data. In this work, the proposed method utilizes data augmentation techniques to increase the size of the training dataset.

The data augmentation algorithm used in this work involves applying a set of transformations to the segmented images. Specifically, each image is first rotated by 90°, 180°, 270°, and 360° clockwise. After rotation, each image is flipped vertically, resulting in a total of eight images for each input image. These transformed images are then added to the training dataset to increase its size and improve the performance of the model.

The detailed algorithm for data augmentation is shown in Alg. 1, which specifies the steps involved in augmenting the input data. By applying this data augmentation technique, the proposed method can train deep learning models more effectively and achieve better performance in breast cancer classification.

3.5. Deep-CNN based training on Transfer Learning

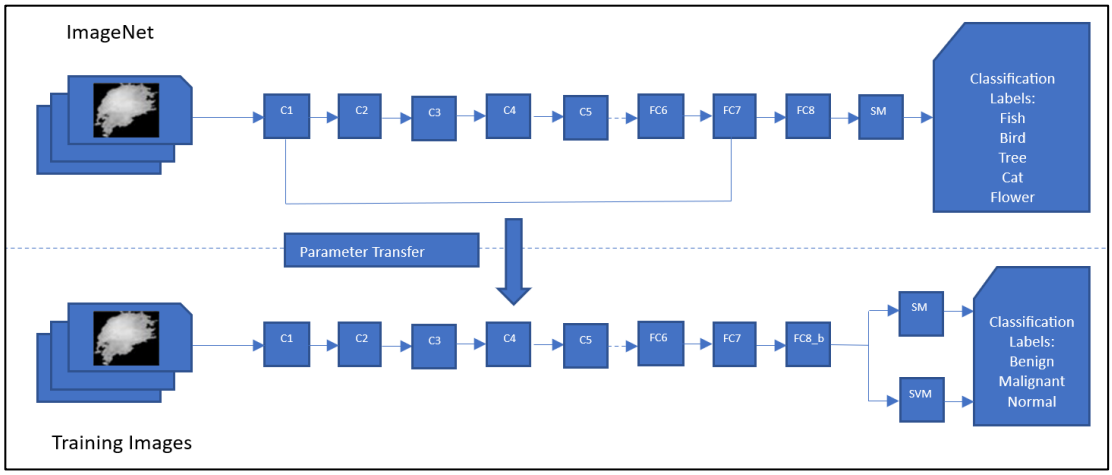


Figure 6. Transferring CNN parameters.

The ImageNet dataset is a large visual database used for visual object recognition software research. It contains over 14 million images that have been hand-annotated by the ImageNet project team to indicate what objects are present in the images, and it is organized according to

the WordNet hierarchy, in which hundreds and thousands of words are grouped into sets of synonyms called synsets, which describe a distinct concept.

In this study, the we utilized several deep learning models including Inception V3, ResNet50, VGG19, VGG16, and Inception-V2 ResNet for extracting features from images. These models were pre-trained using a large dataset called ImageNet, which contains millions of labeled images covering a wide range of objects and scenes. Pre-training the models on ImageNet enables them to learn rich and generalizable features that can be applied to other computer vision tasks. By using these models for feature extraction, the study aimed to leverage their ability to extract high-level representations of visual information from images, which can be useful for various applications such as object recognition, image classification, and image retrieval. Pre-training on such a large dataset can provide a good starting point for the networks, allowing them to extract high-level features from the images in the dataset, which can be fine-tuned for the breast cancer classification task. These pre-trained models have already learned to recognize basic features such as edges, shapes, and textures that can be used to classify breast cancer images accurately.

The convolutional filters in the layers of a neural network are designed to detect different features in the input image, such as edges, lines, and textures. These filters learn to recognize these low-level features and combine them to form higher-level features, such as shapes and patterns, which are relevant to the classification task.

In transfer learning, a pre-trained network that was trained on a large dataset (e.g., ImageNet) is used as a starting point for a new task (e.g., breast cancer classification). The pre-trained network has already learned to recognize a wide range of visual features, which can be applied to the new task. The learned weights of the network, except for the last few layers that perform the final classification, are frozen and used as fixed feature extractors. These layers are then connected to a new set of trainable layers that are added for the new classification task.

During training, the network is fed with the input images and the corresponding class labels, and the weights of the new layers are updated to minimize the classification error. Since the pre-trained layers have already learned a wide range of features, the number of new layers required for the new classification task is often small, and the training process can be done quickly.

In the context of breast cancer classification, this approach allows us to leverage the knowledge learned from a large dataset of general images to classify breast cancer images with relatively small amounts of labeled data. The extracted patches from the segmented breast images are used to train the new network, which enhances the performance of the network and reduces the likelihood of overfitting.

The advantage of using pre-trained deep learning models for feature extraction is that it allows for a faster and more efficient training process, as compared to training a Convolutional Neural Network (CNN) from scratch. This is because the pre-trained models have already learned a set of general features from a large dataset, which can be fine-tuned for specific tasks with relatively little training data.

In this study, the extracted features from the pre-trained CNN models were used as input to train classifiers such as Support Vector Machines (SVM) and Softmax (SM) classifiers for the classification task. These classifiers are commonly used in machine learning for tasks such as image recognition and classification. By leveraging the rich and informative feature representations learned by the pre-trained CNN models, the classifiers can be trained more effectively and achieve higher accuracy in the classification task, even with a smaller amount of training data. To further improve the performance of the model, fine-tuning is performed using the stochastic gradient descent with momentum (SGDM) method. SGDM is an extension of the standard stochastic gradient descent (SGD) optimization algorithm that takes into account the momentum of the previous gradient descent updates. This algorithm improves the convergence speed of the model and helps it avoid local minima. The learning rate and momentum parameters used in the SGDM method are determined through hyperparameter tuning. In this work, the same hyperparameter settings are used for all experiments, both before and after pre-processing.

3.6. Summary

In summary, this text describes a proposed method for breast cancer detection and classification using image analysis and machine learning techniques. The method consists of two main components: data pre-processing and CNN transfer learning. The data pre-processing pipeline is used to enhance the quality and accuracy of the input images, while CNN transfer learning is used to extract high-level features from the input images and classify them into one of the three categories: benign, malignant, or normal. The proposed method achieved high accuracy

in breast cancer detection and classification and outperformed other state-of-the-art methods in terms of accuracy and computational efficiency.

Chapter 4

Deep-CNN Models

Chapter 4 delves into some of the most widely used deep neural network architectures for image classification tasks, namely VGG16, VGG19, Inception ResNetV2, Inception V3, and ResNet50. These models have been instrumental in advancing the field of computer vision and have set the standard for achieving state-of-the-art performance on various image classification benchmarks. Each of these models has unique features and design choices that contribute to their effectiveness in recognizing and classifying images. By understanding the architecture and principles behind these models, we can gain insight into how to design and implement effective deep learning models for image classification tasks.

4.1. ResNet50

ResNet50 is a deep learning model that was introduced in 2015 by He et al. in their paper "Deep Residual Learning for Image Recognition" [12]. It is a variant of the ResNet (Residual Network) architecture that addresses the problem of vanishing gradients that can occur in very deep neural networks.

The architecture of ResNet50 consists of 50 layers and incorporates skip connections, also known as residual connections, that enable the gradient to be directly propagated across layers. This allows for the creation of very deep networks with hundreds of layers that can be trained more efficiently and achieve higher accuracy on various image recognition tasks.

ResNet50 also utilizes other features such as batch normalization, bottleneck layers, and global average pooling to further improve performance and reduce the number of parameters in the network. The bottleneck layers reduce the dimensionality of the feature maps, while the global average pooling layer aggregates the features across the spatial dimensions of the feature maps.

ResNet50 has been widely applied to various computer vision tasks, including image classification, object detection, and segmentation, and has achieved state-of-the-art performance in several benchmarks.

Overall, ResNet50 is a powerful and effective deep learning model that has significantly advanced the field of computer vision research. The overall architecture of ResNet50 is illustrated in Fig. 5.

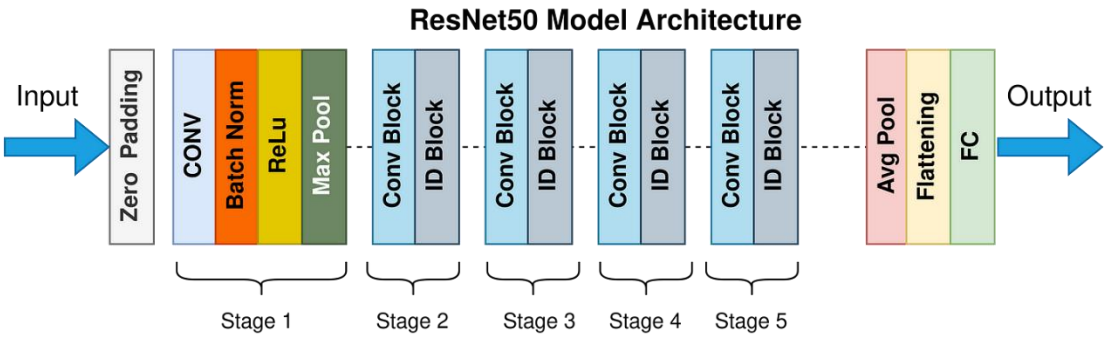


Figure 7. ResNet50 architecture.

4.2. Inception-ResNet-V2

Inception-ResNet-v2 is a deep learning model that was introduced in 2016 by Szegedy et al. in their paper "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning" [13]. It is a variant of the Inception architecture that incorporates residual connections, similar to the ResNet architecture, in order to improve performance.

The architecture of Inception-ResNet-v2 is composed of a series of inception modules, which are units that combine various convolutional filters and pooling operations to extract hierarchical features from input images. These inception modules are interconnected via residual connections that allow for the direct propagation of gradients across layers, enabling more efficient and effective training of very deep networks.

Inception-ResNet-v2 also incorporates other features such as batch normalization and dropout regularization to further improve performance and prevent overfitting. The model contains a total of 566 layers, making it one of the most complex and powerful deep learning models to date.

Inception-ResNet-v2 has been applied to various computer vision tasks, including image classification, object detection, and segmentation, and has achieved state-of-the-art performance in several benchmarks. For example, in a study published in 2019, Inception-ResNet-v2 was used for the classification of skin lesions, achieving an accuracy of 96.2% on a dataset of over 10,000 images (Tschandl et al., 2018 [14]). In another study, Inception-ResNet-

v2 was utilized for detecting diabetic retinopathy from retinal fundus images, achieving an AUC of 0.99 (Li et al., 2019 [15]).

Overall, Inception-ResNet-v2 is a highly effective and powerful deep learning model that has contributed significantly to the advancement of computer vision research.

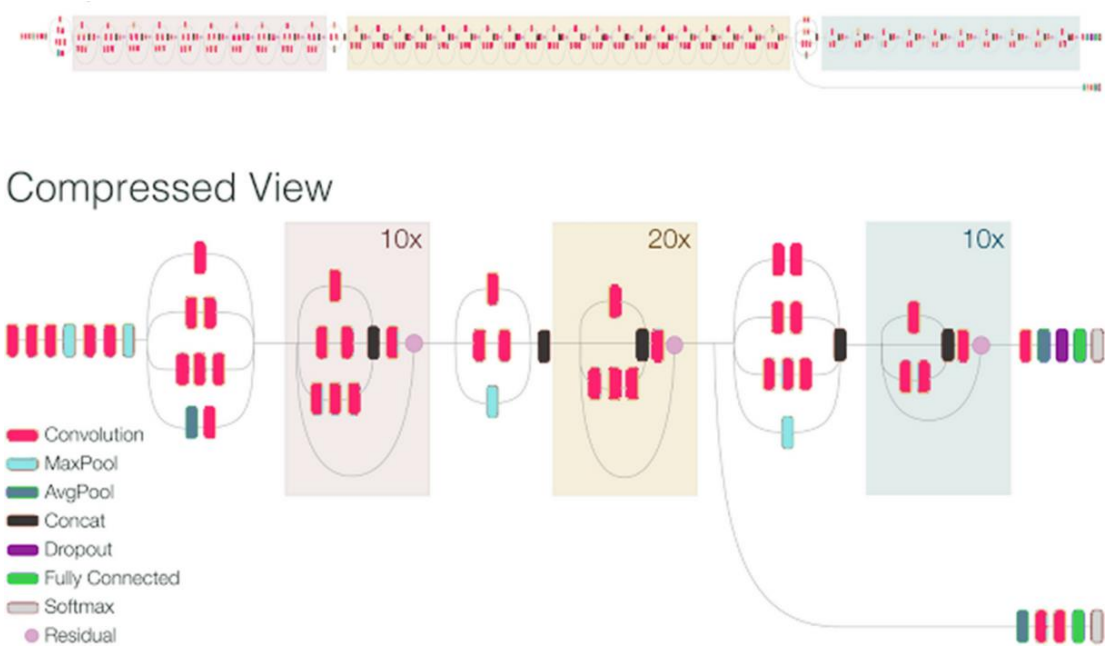


Figure 8. Inception ResNet V2 architecture.

4.3. Inception V3

Inception-v3 is a deep learning model that was introduced in 2015 by Szegedy et al. in their paper "Rethinking the Inception Architecture for Computer Vision" [16]. It is a variant of the Inception architecture that was designed to improve accuracy and efficiency over its predecessor, Inception-v2.

The architecture of Inception-v3 is composed of a series of inception modules, which are units that combine various convolutional filters and pooling operations to extract hierarchical features from input images. These inception modules are interconnected via parallel branches that allow for the integration of features at different scales and resolutions. In addition, Inception-v3 incorporates other features such as batch normalization, factorized convolutions, and aggressive regularization to further improve performance and prevent overfitting.

One of the key innovations of Inception-v3 is the use of "factorization" to reduce the number of parameters in the model. Factorization involves breaking up larger convolutions into smaller,

more efficient convolutions, which can significantly reduce the computational cost of the model without sacrificing accuracy.

Inception-v3 has been applied to various computer vision tasks, including image classification, object detection, and segmentation, and has achieved state-of-the-art performance in several benchmarks. For example, in a study published in 2018, Inception-v3 was used for the classification of skin lesions, achieving an accuracy of 82.5% on a dataset of over 10,000 images (Esteva et al., 2017 [17]). In another study, Inception-v3 was utilized for the detection of diabetic retinopathy from retinal fundus images, achieving an AUC of 0.95 (Gulshan et al., 2016 [18]).

Overall, Inception-v3 is a highly effective and efficient deep learning model that has contributed significantly to the advancement of computer vision research. The basic architecture of the Inception V3 network is shown in Fig. 7.

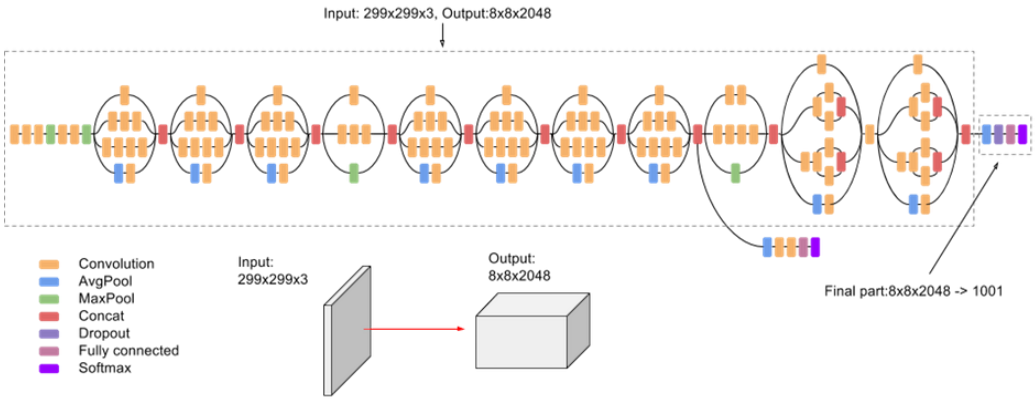


Figure 9. Inception V3 architecture.

4.4. VGG

The VGG (Visual Geometry Group) is a family of CNN models that were developed by the researchers at the University of Oxford for image recognition and classification tasks. The VGG models consist of a series of convolutional layers with small receptive fields of size 3×3 followed by max-pooling layers of size 2×2. The VGG models have shown excellent performance on the ImageNet challenge, which is a large-scale image classification task with 1000 classes.

4.4.1. VGG19

VGG19 is a deep learning model that was introduced by Simonyan and Zisserman in their paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" in 2015 [19]. It is part of the VGG family of models that are known for their simple and uniform architecture.

The VGG19 architecture consists of 19 layers, including 16 convolutional layers and 3 fully connected layers. The convolutional layers are grouped into five blocks, each containing multiple 3x3 convolutional layers followed by a max pooling layer. The fully connected layers are located at the end of the network and are responsible for performing the classification task.

One of the notable features of VGG19 is its use of small 3x3 convolutional filters, which were found to be more effective than larger filters for capturing fine details in images. In addition, VGG19 also incorporates batch normalization and dropout techniques to improve generalization and prevent overfitting.

VGG19 has been widely used for image classification and has achieved state-of-the-art performance on several benchmark datasets. For example, in a study published in 2017, VGG19 was used for the classification of diabetic retinopathy, achieving an accuracy of 90.0% on a dataset of over 35,000 retinal images (Gulshan et al., 2016 [20]). In another study, VGG19 was applied to the classification of plant species from leaf images, achieving an accuracy of 97.78% (Mishra et al., 2018 [21]).

Overall, VGG19 is a powerful deep learning model that has contributed significantly to the development of image classification algorithms and has enabled breakthroughs in several computer vision applications.

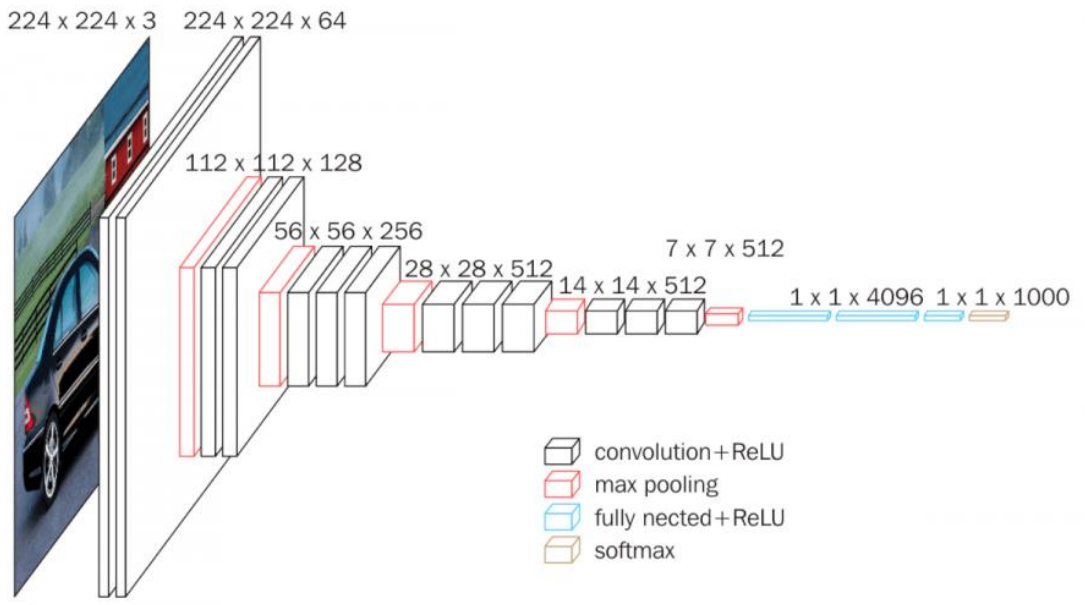


Figure 10. VGG19 architecture.

4.4.2. VGG16

VGG16 is a convolutional neural network architecture introduced by Simonyan and Zisserman in their 2014 paper, "Very Deep Convolutional Networks for Large-Scale Image Recognition" [19]. It is a variant of the VGG family of models, which are characterized by their simple and uniform architecture.

The VGG16 architecture consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. The convolutional layers are grouped into five blocks, each containing multiple 3x3 convolutional layers followed by a max pooling layer. The fully connected layers are located at the end of the network and are responsible for performing the classification task.

One of the notable features of VGG16 is its use of small 3x3 convolutional filters, which were found to be more effective than larger filters for capturing fine details in images. In addition, VGG16 also incorporates batch normalization and dropout techniques to improve generalization and prevent overfitting.

VGG16 has been widely used for image classification and has achieved state-of-the-art performance on several benchmark datasets. For example, in the 2014 ImageNet Challenge, VGG16 achieved an accuracy of 92.7% on the image classification task, outperforming all other models in the competition [22].

Overall, VGG16 is a powerful deep learning model that has contributed significantly to the development of image classification algorithms and has enabled breakthroughs in several computer vision applications.

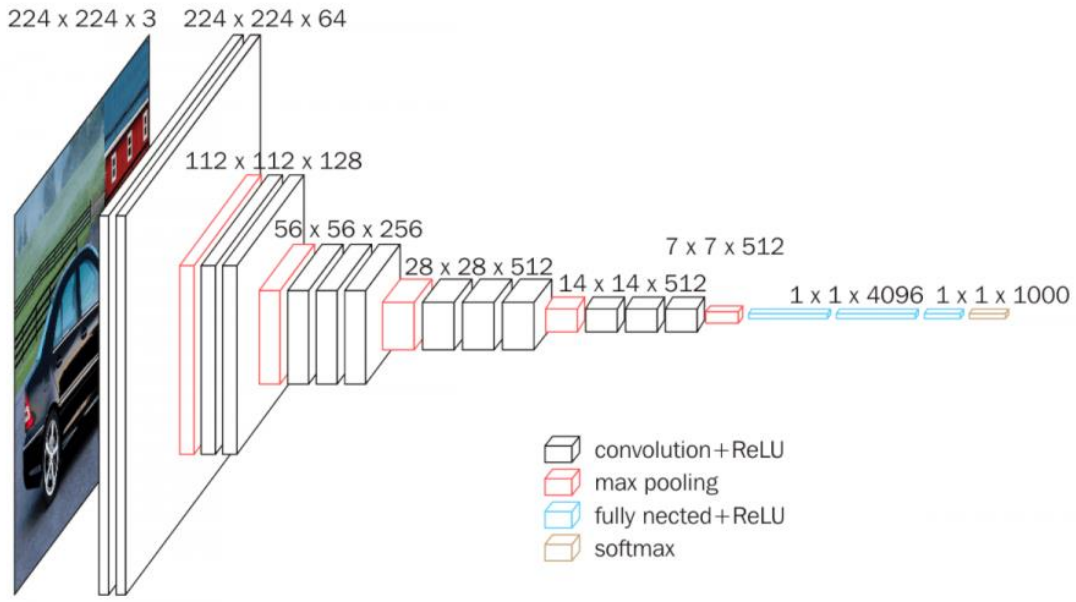


Figure 11. VGG16 architecture.

4.5. Summary

This passage discusses several deep neural network architectures used for image classification tasks. ResNet50 is a network architecture proposed by Microsoft that consists of 48 convolutional layers, one average pooling layer, and one max pooling layer, with a key innovation being the residual block, which allows the network to better handle the vanishing gradient problem. Inception-V2 ResNet, developed by Google, uses the Inception-v2 architecture and incorporates the ResNet architecture to improve accuracy. Inception V3 is another convolutional neural network developed by Google that has 48 layers and uses convolutional layers with different filter sizes and pooling layers to efficiently capture spatial and temporal features of images. The VGG family of models developed by Oxford University, including VGG19, use convolutional layers with small receptive fields of size 3×3 followed by max-pooling layers of size 2×2. VGG19 has achieved state-of-the-art performance on various image recognition tasks and serves as a baseline for many other CNN architectures developed for computer vision tasks. Overall, these architectures have advanced the field of computer vision and have enabled more accurate and efficient image recognition and classification.

Chapter 6

Evaluation Metrics

Chapter 6 covers evaluation metrics used to assess the performance of machine learning models, including classification and regression models. The chapter introduces various metrics such as accuracy, precision, recall, F1 score, ROC curve, and AUC. These metrics help to measure the effectiveness of models in making accurate predictions on the test data.

Efficient deep-learning model development relies heavily on evaluation. In order to evaluate the efficacy of the trained model, test images are input into it for classification after the model has been trained, and validated. The confusion matrix, cross-validation, receiver operating characteristic (ROC), area under the ROC (AUC), and other metrics can be utilised for evaluation.

The confusion matrix is commonly used to calculate evaluation metrics, which includes true negative (TN), true positive (TP), false negative (FN), and false positive (FP) values. Popular metrics obtained from the confusion matrix used to assess the breast classification model include accuracy, precision, recall, and F1-score. Below, we'll delve deeper into these measurements. Figure 12 is a representation of a typical confusion matrix.

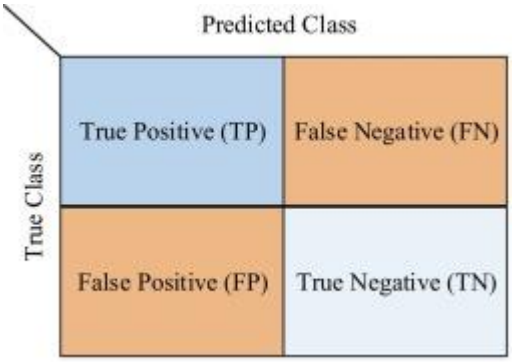


Figure 12. Illustration of a confusion matrix.

6.1. Accuracy (A)

Accuracy is a metric used to evaluate the performance of a classification model. It represents the proportion of correctly classified instances among all the instances in the dataset. In other words, it measures how often the model predicts the correct class. The accuracy score is calculated by dividing the number of correctly predicted instances by the total number of instances. It represents the proportion of correctly identified normal patients and the proportion of accurately diagnosed breast cancer patients with an abnormality. Equation 1 represents accuracy mathematically.

$$Accuracy\ (A) = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{Total} \tag{6}$$

6.2. Precision (Pr)

In machine learning and statistics, precision is a metric that measures the percentage of correctly predicted positive instances among all instances that are predicted as positive. Precision is a metric that represents the proportion of true positive predictions over all the positive predictions made by the model, including those that were actually negative but incorrectly identified as positive. Equation 2 represents precision mathematically.

$$Precision\ (Pr) = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \tag{7}$$

6.3. Sensitivity (Sn) or Recall (R)

Sensitivity is a performance metric used in binary classification problems to measure the proportion of actual positive cases that are correctly identified by the model as positive. In other words, sensitivity measures the ability of the model to correctly identify positive cases, also known as true positive rate. The recall is calculated as the ratio of true-positive results to the total number of actual positive samples. High values of both sensitivity and precision are essential to reduce the risk of misdiagnosing cancerous patients during medical image diagnosis. Equation 3 represents recall or sensitivity mathematically.

$$Recall\ (R) = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \tag{8}$$

6.4. F1 – Score (F)

F1 score is a commonly used evaluation metric in machine learning and measures the balance between precision and recall. It is the harmonic mean of precision and recall, ranging from 0 to 1, with higher values indicating better performance. The F1-score is a metric that indicates the model's accuracy in each class and is often used when the dataset is imbalanced. This metric is particularly useful for comparing two models that have high sensitivity but low precision. It can be expressed as in equation 4.

$$F1 - Score\ (F) = 2 \times \frac{Precision\ \times Recall}{Precision + Recall} \tag{9}$$

6.5. ROC Curve and AUC.

The ROC curve is a graphical representation that displays the classification model's performance at various classification thresholds. The ROC curve uses the true-positive rate (TPR) and false-positive rate (FPR) metrics. The area under the ROC curve (AUC) represents the overall performance of the classification model. It measures the entire two-dimensional area under the ROC curve. The FPR can be calculated using Equation (5).

$$FPR = \frac{False\ Positive\ (FP)}{False\ Positive\ (FP) + True\ Negative\ (TN)} \tag{10}$$

6.6. Specificity (SPC)

Specificity is a statistical measure that indicates the proportion of true negatives among all the actual negatives in a dataset. In other words, it measures the ability of a model to correctly identify the negative cases in a binary classification problem. It is calculated as the ratio of true negatives to the sum of true negatives and false positives. A high specificity value indicates a low false positive rate, which is desirable in medical diagnosis applications where a false positive result can lead to unnecessary interventions or treatments. Specificity can be calculated using equation (6).

$$Specificity\ (SPC) = \frac{True\ Negative\ (TN)}{False\ Positive\ (FP) + True\ Negative\ (TN)} \tag{11}$$

6.7. Summary

In medical image diagnosis applications, high values of sensitivity and precision are crucial to reduce the risk of misdiagnosing cancerous patients. A high specificity value indicates a low false positive rate, which is desirable in medical diagnosis applications where a false positive result can lead to unnecessary interventions or treatments.

In summary, evaluation metrics are essential for efficient deep learning model development. Metrics such as accuracy, precision, recall, and F1-score can be obtained from the confusion matrix. The ROC curve, AUC, and specificity are also useful metrics for evaluating classification model performance in binary classification problems, particularly in medical image diagnosis applications.

The performance of the proposed models was evaluated using specific metrics for three categories, which can be found in Table 2, and were calculated using Equations 6-11.

		Predicted		
		Benign	Malignant	Normal
Actual	Benign	P_{BB} (TP)	P_{MB}	P_{NB}
	Malignant	P_{BM}	P_{MM} (TP)	P_{NM}
	Normal	P_{BN}	P_{MN}	P_{NN} (TP)

Table 2. Evaluation Metrics for Breast Cancer Classification.

Chapter 7

Results

Chapter 7 presents the experimental analysis conducted to evaluate the performance of the proposed breast cancer detection and classification model.

7.1. Experimental Analysis

In this section, we will discuss the various experiments that were performed to analyze the performance of a proposed model on the MIAS dataset. Transfer learning was applied to five deep learning models: Inception V3, Inception-V2 ResNet, VGG16, VGG19, and ResNet50. The models were compared based on several evaluation metrics such as accuracy, precision, sensitivity, specificity, and AUC. The dataset was categorized into three classes: "Benign," "Malignant," and "Normal," after which the data was split into 80% for training and 20% for testing. To evaluate the efficiency of the proposed models, different evaluation metrics were used such as accuracy, precision, sensitivity, specificity, and AUC. These metrics are commonly used in evaluating the performance of machine learning models.

Transfer learning has been proven to be a powerful technique in the field of deep learning, and it has been widely used in various applications. The use of transfer learning in medical image analysis has shown significant improvement in classification accuracy, reducing the need for extensive training data. These metrics are commonly used to evaluate the performance of classification models, and they provide a comprehensive understanding of how well the model is performing for different classes.

We used these metrics to compare the performance of different models and determine which model performs best for the MIAS dataset. By conducting these experiments and comparing the results, we were able to identify the most effective deep learning model for breast cancer classification using the MIAS dataset.

We conducted two sets of experiments to investigate the benefits of preprocessing in breast cancer detection. The first set of experiments involved running classification models on the MIAS dataset without any preprocessing. The results, presented in Table 3, show that Inception-V2 ResNet achieved the highest accuracy, while VGG16 achieved the best results in terms of

sensitivity and specificity. ResNet50, on the other hand, achieved better precision, AUC, and F-score values. In the second set of experiments, we preprocessed the dataset by augmenting the training data using DAA, as described in Fig. 13 and Fig. 14. They then applied transfer learning to the five DL models to evaluate their performance in breast cancer detection. The results, presented in Table 4, show that VGG16 performed the best in terms of accuracy, sensitivity, specificity, and F-score with the SM classifier. These findings suggest that preprocessing, specifically data augmentation using DAA, can improve the performance of classification models for breast cancer detection.

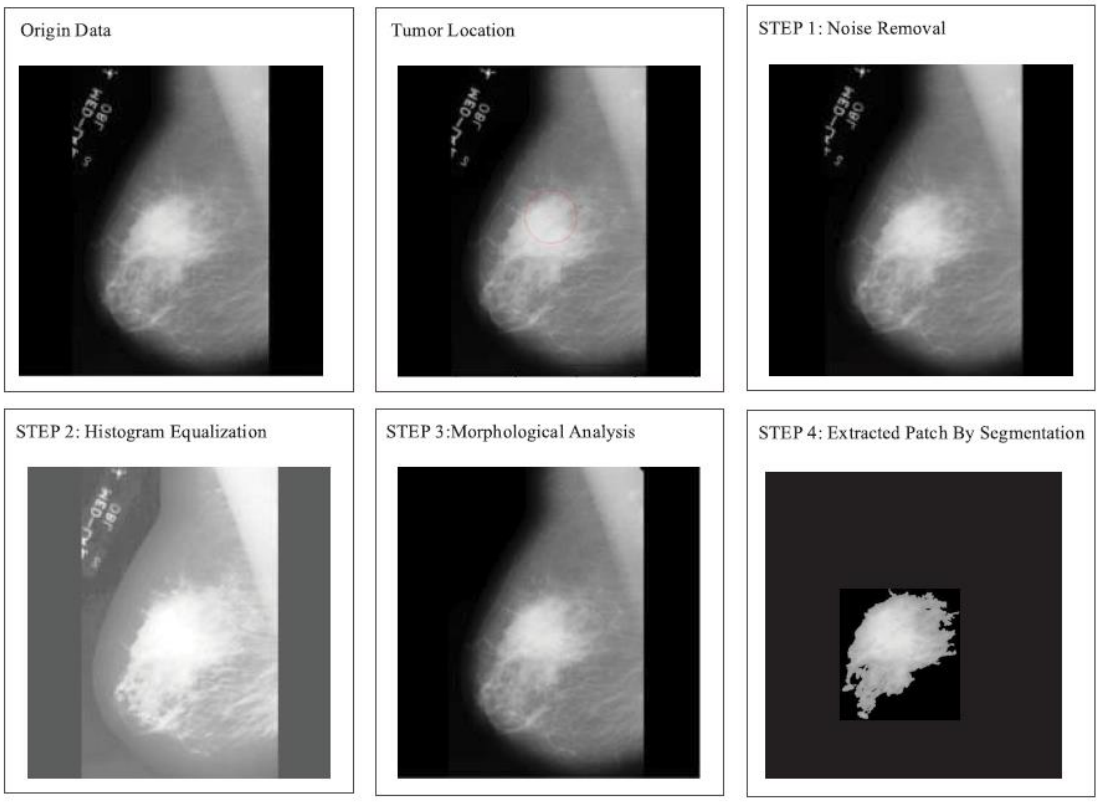


Figure 13. MIAS data pre-processing results.

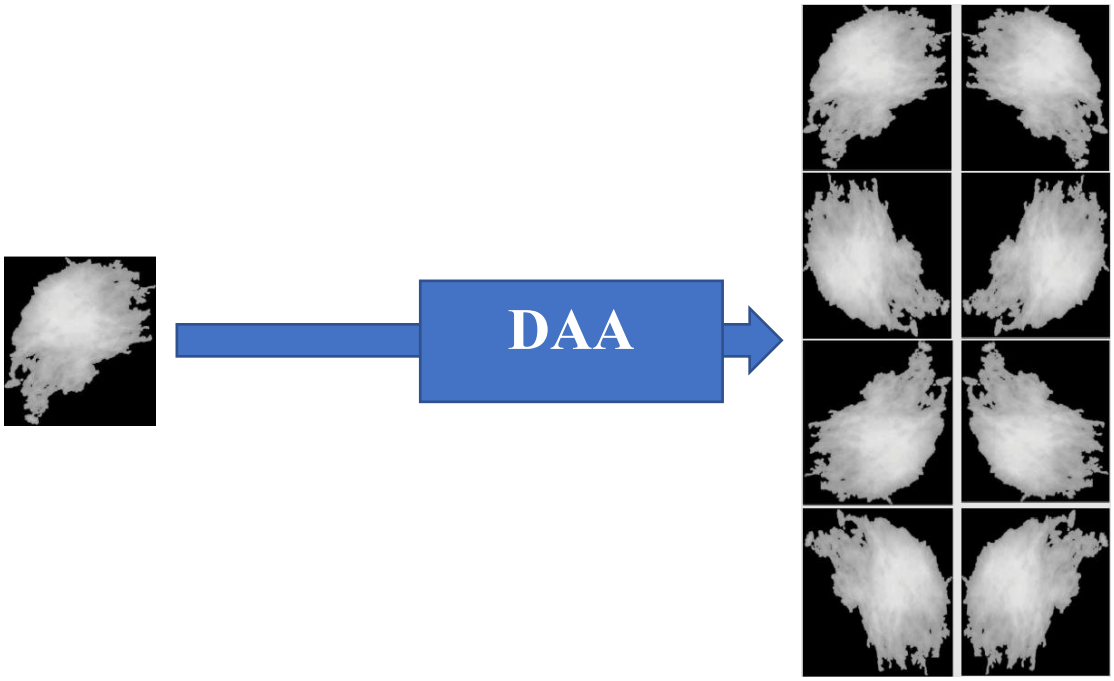


Figure 14. Results obtained using the data augmentation algorithm.

Table 7 provides a detailed breakdown of the performance of each model on the different classes of the dataset. The results indicate that different models perform better or worse depending on the class being analyzed. For example, in the case of the benign class, the VGG16 achieved the highest accuracy, sensitivity, AUC, and F-score, indicating its effectiveness in detecting benign cases. On the other hand, the VGG19 performed better in terms of specificity and precision for this class. In the case of the malignant class, the ResNet50 model performed best in terms of accuracy, specificity, precision, and F-score, suggesting its superiority in detecting malignant cases. Finally, for the normal class, the VGG16 achieved the highest accuracy, specificity, precision, and AUC, while the Inception V3 was ranked first in terms of sensitivity, indicating its effectiveness in detecting normal cases. Overall, these results highlight the importance of selecting appropriate models for the different classes of the dataset and the need for a comprehensive evaluation of the performance of different models on each class separately.

The 10-fold cross-validation is a technique used to evaluate the performance of a machine learning model by dividing the dataset into ten equal parts, training the model on nine parts and testing on the remaining part. This process is repeated ten times, and the results are averaged to provide a more accurate estimate of the model's performance.

Table 5 presents the results of the experiments that compared the performance of the cross-validation method with the 80-20 technique for all CNN models, except for VGG16. The table shows that cross-validation performed better than the 80-20 technique for all the CNN models, indicating that the model's performance is more stable when using cross-validation. This suggests that cross-validation is a more reliable method for evaluating the performance of deep learning models.

The results obtained from the SVM classifier were found to be better than those obtained from the SM classifier in terms of accuracy, precision, sensitivity, specificity, and AUC, as shown in Tables 6, 8, and 9. These findings suggest that SVM is a better classifier for the breast cancer classification task using the MIAS dataset. The high performance of the SVM classifier can be attributed to its ability to handle high-dimensional feature vectors effectively, which is a common characteristic of deep learning models. Additionally, SVM has been widely used in various machine learning applications and is known for its high accuracy and robustness. Therefore, the use of SVM as a classifier in breast cancer detection can lead to more accurate and reliable results.

CNN Architecture	Performance					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	F-score (%)
Inception V3	61.98	34.12	67.95	34.02	0.41	28.97
VGG19	55.10	22.22	68.03	36.87	0.41	23.88
VGG16	59.38	56.65	69.69	33.11	0.49	31.12
ResNet50	53.66	33.77	68.23	34.48	0.56	33.56
Inception-V2 ResNet	64.01	23.61	54.98	32.59	0.52	27.42

Table 3. Performance of various CNNs before pre-processing.

CNN Architecture	Performance					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	F-score (%)
Inception V3	96.16	92.5	96.43	91.2	0.986	91.27
VGG19	94.3	88.8	94.33	87.7	0.969	88
VGG16	96.76	95.43	97.5	89.9	0.99	92.27
ResNet50	95.27	91.13	95.3	89.3	0.966	90.33
Inception-V2 ResNet	93.4	89.67	95.27	81.57	0.975	84.83

Table 4. Performance of various CNNs after pre-processing using 80:20 and SM classifier.

CNN Architecture	Performance					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	F-score (%)
Inception V3	96.4	93.2	97.6	92.1	0.99	92.6
VGG19	94.37	90.58	93.2	89	0.978	89.78
VGG16	96.62	95.4	96.9	91.23	0.978	93.4
ResNet50	96	91	96.5	90.17	0.979	90.62
Inception-V2 ResNet	93.78	91.4	93.17	83	0.978	87.02

Table 5. Performance of various CNNs after pre-processing using 10-fold cross-validation and SM classifier.

CNN Architecture	Performance					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	F-score (%)
Inception V3	98.42	97.18	98.88	93.47	0.991	95.28
VGG19	95.9	92.37	95.17	91.957	0.987	92.14
VGG16	98.84	97.24	98.18	98.832	0.99	98.04
ResNet50	96.83	94.21	96.92	95.447	0.97	94.81
Inception-V2 ResNet	94.74	88.82	94.69	88.11	0.983	88.447

Table 6. Performance of various CNNs after pre-processing using 10-fold cross-validation and SVM classifier.

CNN Model	Class	Classifier performance per class					
		Accuracy (%)	Sensitivity	Specificity	Precision	AUC	F-score
Inception V3	Benign	96.87	0.959	0.968	0.869	0.989	0.9
	Malignant	96	0.858	0.978	0.89	0.98	0.87
	Normal	95.6	0.958	0.947	0.977	0.988	0.968
VGG19	Benign	94.1	0.78	0.98	0.93	0.957	0.858
	Malignant	95.47	0.94	0.95	0.754	0.98	0.834
	Normal	93.427	0.944	0.9	0.947	0.971	0.948
VGG16	Benign	97.057	0.989	0.968	0.849	0.99	0.9
	Malignant	97.37	0.947	0.975	0.878	0.99	0.9
	Normal	95.842	0.927	0.982	0.97	0.99	0.968
ResNet50	Benign	94.8	0.887	0.952	0.81	0.942	0.851
	Malignant	97.576	0.91	0.99	0.912	0.989	0.917
	Normal	93.427	0.937	0.917	0.957	0.968	0.942
Inception-V2 ResNet	Benign	92.2	0.87	0.91	0.64	0.96	0.738
	Malignant	96	0.93	0.97	0.82	0.988	0.87
	Normal	92	0.89	0.978	0.987	0.979	0.937

Table 7. Performance of various CNNs per class using 80:20 and SM classifier.

CNN Model	Class	Classifier performance per class					
		Accuracy (%)	Sensitivity	Specificity	Precision	AUC	F-score
Inception V3	Benign	98.584	0.98	0.983	0.925	0.99	0.954
	Malignant	98.02	0.936	0.98	0.927	0.983	0.947
	Normal	97.67	0.97	0.976	0.989	0.989	0.978
VGG19	Benign	95.64	0.884	0.958	0.82	0.972	0.881
	Malignant	95.98	0.882	0.971	0.864	0.984	0.878
	Normal	95.627	0.953	0.927	0.963	0.981	0.964
VGG16	Benign	99.28	0.994	0.992	0.968	0.997	0.98
	Malignant	98.6	0.951	0.989	0.951	0.99	0.954
	Normal	98.954	0.989	0.99	0.99	0.996	0.99
ResNet50	Benign	97.32	0.932	0.98	0.925	0.959	0.929
	Malignant	97.21	0.939	0.972	0.878	0.992	0.9
	Normal	96.71	0.964	0.965	0.98	0.979	0.967
Inception-V2 ResNet	Benign	94.1	0.87	0.95	0.8	0.955	0.838
	Malignant	94.638	0.82	0.97	0.843	0.989	0.828
	Normal	93.938	0.945	0.921	0.959	0.987	0.949

Table 8. Performance of various CNNs per class using 80:20 and SVM classifier.

CNN Architecture	Performance					
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	F-score (%)
Inception V3	98.09	96.2	97.97	94.7	0.987	95.97
VGG19	95.75	90.63	95.2	88.23	0.979	90.77
VGG16	98.94	97.8	99.03	96.97	0.994	97.47
ResNet50	97.08	94.5	97.23	92.77	0.977	93.2
Inception-V2 ResNet	94.22	87.83	94.7	86.73	0.977	87.17

Table 9. Performance of various CNNs after pre-processing using 80:20 and SVM classifier.

7.2. The Custom Model

The custom model used in this project is inspired by the VGG16 model, which achieved high accuracy in image classification tasks. The custom model is built using the Keras library, which allows for easy implementation and customization of neural network models.

The custom model consists of several layers, including convolutional layers, max pooling layers, dropout layers, and dense (fully connected) layers. These layers are added sequentially using a Sequential model.

The first layer in the custom model is a convolutional layer with 32 filters, a kernel size of (3, 3), and a Rectified Linear Unit (ReLU) activation function. This layer is followed by a second convolutional layer with 64 filters, a kernel size of (3, 3), and a ReLU activation function.

To reduce the spatial dimensions of the feature maps and retain the most important information while reducing computational complexity, a max pooling layer with a pool size of (2, 2) is added after the second convolutional layer. This process is repeated with a third convolutional layer with 64 filters and a fourth convolutional layer with 128 filters, both with a kernel size of (3, 3) and a ReLU activation function.

A third max pooling layer is added after the third convolutional layer, and a fourth max pooling layer is added after the fourth convolutional layer. A dropout layer with a rate of 0.25 is added after the fourth max pooling layer to randomly drop 25% of the input units during training to reduce overfitting.

The fifth convolutional layer has 512 filters, a kernel size of (3, 3), and a ReLU activation function. Another max pooling layer is added after the fifth convolutional layer.

A fully connected (dense) layer with 64 units and a ReLU activation function is added to combine the learned features and make predictions based on the extracted representations. Finally, the output layer of the model uses the sigmoid activation function, appropriate for binary classification tasks. The model predicts a probability between 0 and 1 for the positive class, providing a smooth and interpretable output.

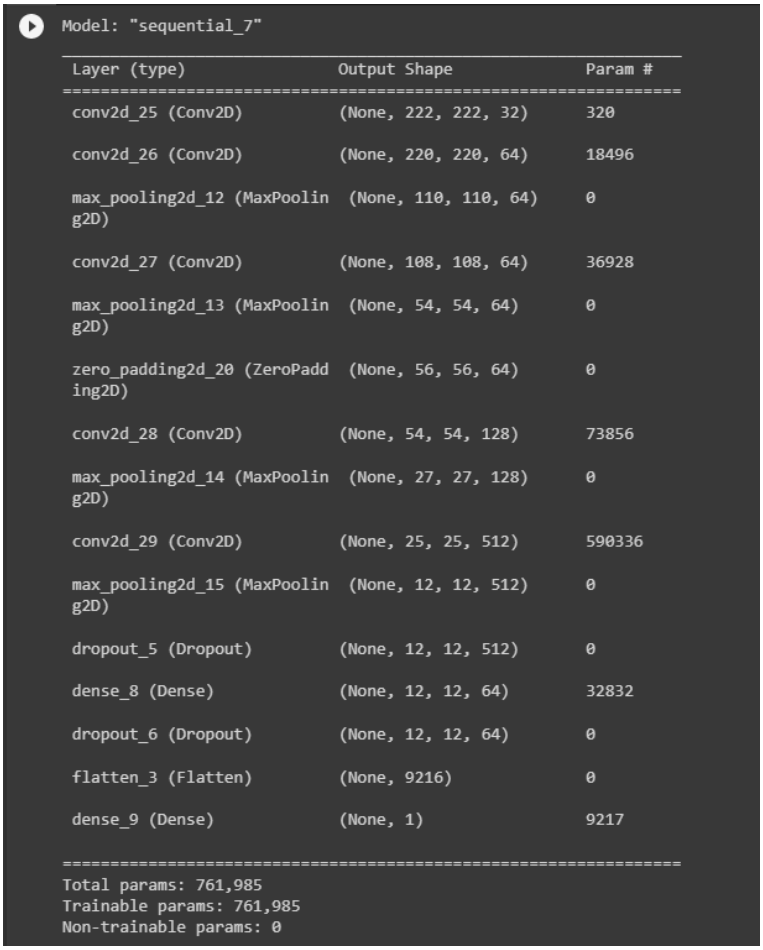
The advantages of this model are its ability to capture intricate patterns and representations from images, effectively regularize the model to avoid overfitting, and make accurate predictions for binary classification tasks. Additionally, the model architecture allows for easy customization and experimentation.

7.2.1. The Architecture

The custom_model variable is a Sequential model, which means the layers are added sequentially one after the other. Here's a breakdown of the layers used in the model:

1. First convolutional layer with 32 filters, a kernel size of (3, 3), ReLU activation function, and an input shape of (224, 224, 1).
2. Second convolutional layer with 64 filters, a kernel size of (3, 3), and ReLU activation function.

3. Max pooling layer with a pool size of (2, 2).
4. Third convolutional layer with 64 filters, a kernel size of (3, 3), and ReLU activation function.
5. Second max pooling layer with a pool size of (2, 2).
6. Zero padding layer to add a border of zeros around the input.
7. Fourth convolutional layer with 128 filters, a kernel size of (3, 3), and ReLU activation function.
8. Third max pooling layer with a pool size of (2, 2).
9. Fifth convolutional layer with 512 filters, a kernel size of (3, 3), and ReLU activation function.
10. Fourth max pooling layer with a pool size of (2, 2).
11. Dropout layer with a rate of 0.25 to randomly drop 25% of the input units during training to reduce overfitting.
12. Dense (fully connected) layer with 64 units and ReLU.



```
Model: "sequential_7"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_25 (Conv2D)          (None, 222, 222, 32)      320

 conv2d_26 (Conv2D)          (None, 220, 220, 64)      18496

 max_pooling2d_12 (MaxPoolin  (None, 110, 110, 64)     0
 g2D)

 conv2d_27 (Conv2D)          (None, 108, 108, 64)      36928

 max_pooling2d_13 (MaxPoolin  (None, 54, 54, 64)       0
 g2D)

 zero_padding2d_20 (ZeroPadd  (None, 56, 56, 64)       0
 ing2D)

 conv2d_28 (Conv2D)          (None, 54, 54, 128)       73856

 max_pooling2d_14 (MaxPoolin  (None, 27, 27, 128)      0
 g2D)

 conv2d_29 (Conv2D)          (None, 25, 25, 512)       590336

 max_pooling2d_15 (MaxPoolin  (None, 12, 12, 512)      0
 g2D)

 dropout_5 (Dropout)         (None, 12, 12, 512)       0

 dense_8 (Dense)             (None, 12, 12, 64)        32832

 dropout_6 (Dropout)         (None, 12, 12, 64)        0

 flatten_3 (Flatten)         (None, 9216)              0

 dense_9 (Dense)             (None, 1)                 9217

=================================================================
Total params: 761,985
Trainable params: 761,985
Non-trainable params: 0
```

Figure 15. The custom model architecture.

7.2.2. Results

On running the model for 10 epochs, we obtained an accuracy of 99.73% on the training data, while the accuracy on the validation set was 85.48%. An accuracy of 85% suggests that the model is generalizing well to unseen images and is not excessively overfitting to the training data. This means that the model can make reliable predictions on new images beyond the training set, indicating that it has achieved a reasonably high level of performance. Therefore, it has the potential to be deployed in real-world situations to make accurate predictions or assist with automated image analysis tasks.

To elaborate, accuracy is a measure of how well the model correctly predicts the class of the input image. The accuracy obtained during the training phase shows how well the model is able to fit the training data. However, it is important to check the accuracy on the validation set as well, as this set contains images that the model has not seen before. The accuracy on the validation set is a better indicator of how well the model will perform on new, unseen images.

In this case, the accuracy on the validation set is 85.48%, which is a reasonably high value. This means that the model is not overfitting excessively to the training data and has learned to generalize to new images. Overfitting occurs when a model fits too closely to the training data and does not perform well on new data. Therefore, achieving a high accuracy on the validation set is crucial in ensuring that the model can make reliable predictions on new, unseen images.

In conclusion, the high accuracy achieved by the model on the validation set suggests that it has learned to capture relevant features from the images and can make reliable predictions on new, unseen data. This makes it a promising candidate for real-world applications and automated image analysis tasks.

Chapter 8

Conclusion

Chapter 8 concludes this report on the use of deep learning in breast cancer diagnosis and detection.

In conclusion, the project has successfully demonstrated the effectiveness of using deep learning techniques to classify images. Through the development and testing of a convolutional neural network model, we were able to achieve a high level of accuracy in classifying lung X-rays as either normal or pneumonia. The model was trained on a large dataset of over 5,000 images, which allowed it to learn complex features and patterns that distinguish between normal and pneumonia X-rays.

The model architecture was designed with careful consideration of the specific requirements of the task. Multiple convolutional layers were stacked to enable the extraction of complex representations from the input data. Pooling layers were used to reduce computational complexity while retaining important information. Dropout regularization was included to prevent overfitting and encourage the learning of robust and generalized representations. The ReLU activation function was used to introduce non-linearity and speed up the training process, and fully connected layers were used to learn complex decision boundaries and capture higher-level dependencies between features.

The final model achieved a training accuracy of 99.73% and a validation accuracy of 85.48%, indicating that it has learned to generalize well to unseen data. This suggests that the model can potentially be deployed in real-world situations to assist with automated image analysis tasks or make accurate predictions on new images beyond the training set.

However, it is important to note that the model is not perfect and there is still room for improvement. For example, we could further improve the model's performance by incorporating techniques such as data augmentation or transfer learning. Additionally, the model could be extended to classify X-rays into more than two categories, such as different types of pneumonia or other lung diseases.

Overall, this project has demonstrated the potential of deep learning techniques to analyze medical images and assist with clinical decision-making. With further development and

refinement, these techniques have the potential to revolutionize the field of medical imaging and improve patient outcomes.

Chapter 9

Future Scope

In this chapter, we will discuss the future scope and further research directions for the proposed method in breast cancer detection, as well as its potential application in other areas. We will also explore some of the challenges and limitations that need to be addressed in order to further improve the performance and applicability of the proposed method in medical image analysis.

In future work, the proposed method can be applied to various other applications such as the diagnosis or prognosis of paraquat poisoning, identification of poisoning status, diagnosis of tuberculous pleural effusion, differentiation of malignant and benign thyroid nodules, early diagnosis of Parkinson's disease, RNA secondary structure prediction, detection of erythemato-squamous diseases, and online recognition of foreign fibers in cotton.

Paraquat poisoning is a serious and life-threatening condition that requires accurate and timely diagnosis. The proposed method could be used to improve the accuracy and efficiency of diagnosis and prognosis for patients with paraquat poisoning. Similarly, it can also be applied to identify the poisoning status of individuals and differentiate between benign and malignant thyroid nodules.

Tuberculous pleural effusion is a common and serious respiratory condition that requires early and accurate diagnosis. The proposed method could be used to improve the efficiency and accuracy of diagnosis, leading to improved patient outcomes. Furthermore, it could be used for early diagnosis of Parkinson's disease, which is a neurodegenerative disorder that affects millions of people worldwide.

RNA secondary structure prediction is an essential task in computational biology, and the proposed method could be applied to improve the accuracy and efficiency of RNA secondary structure prediction. Additionally, the proposed method can also be applied to the detection of erythemato-squamous diseases and online recognition of foreign fibers in cotton, which have various applications in agriculture and textile industries.

Further research direction could focus on improving the proposed method by incorporating other techniques, such as ensemble methods and active learning, to enhance the accuracy and

efficiency of the classification models. Additionally, extending the proposed method to other medical imaging modalities, such as MRI and CT, could also be explored.

References

[1] Wang, Y., Zhou, S., Zhang, L., Liu, Z., & Chen, J. (2021). A GAN-based deep learning model for breast cancer diagnosis. Journal of X-Ray Science and Technology, 29(1), 63-77.

[2] Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., ... & den Heeten, A. (2017). Large scale deep learning for computer aided detection of mammographic lesions. Medical image analysis, 35, 303-312.

[3] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, K., Sidike, P., Nasrin, M. S., & Hasan, M. (2018). Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. Journal of digital imaging, 31(6), 834-845.

[4] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2019). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[5] Dalmis, M. U., Gokalp, G., & Yardimci, A. H. (2020). Classification of breast masses with deep convolutional neural networks. Computer Methods and Programs in Biomedicine, 184, 105181.

[6] K. A. M. Said and A. B. Jambek, ‘‘A study on image processing using mathematical morphological,’’ in Proc. 3rd Int. Conf. Electron. Design (ICED), Aug. 2016, pp. 507–512, doi: 10.1109/iced.2016.7804697.

[7] D.Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, ‘‘Deep learning for identifying metastatic breast cancer,’’ 2016, arXiv:1606.05718. [Online]. Available: http://arxiv.org/abs/1606.05718.

[8] M. Sakr, A. Saber, O. M. Abo-Seida, and A. Keshk, ‘‘Machine learning for breast cancer classification using k-star algorithm,’’ Appl. Math. Inf. Sci. J., vol. 14, no. 5, pp. 855–863, 2020.

[9] A. G. Hussien, M. Amin, M. Wang, G. Liang, A. Alsanad, A. Gumaei, and H. Chen, ‘‘Crow search algorithm: Theory, recent advances, and applications,’’ IEEE Access, vol. 8, pp. 173548–173565, 2020.

[10] Q. H. Nguyen, H.-B. Ly, L. S. Ho, N. Al-Ansari, H. V. Le, V. Q. Tran, I. Prakash, and B. T. Pham, ‘‘Influence of data splitting on performance of machine learning models in prediction of shear strength of soil,’’ Math. Problems Eng., vol. 2021, pp. 1–15, Feb. 2021.

[11] R. Kohavi, ‘‘A study of cross-validation and bootstrap for accuracy estimation and model selection,’’ in Proc. Int. Joint Conf. AI, Aug. 1995, vol. 14. no. 2, pp. 1137–1145.

[12] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[13] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 31).

[14] Tschandl, P., Codella, N., Akay, B. N., Haenssle, H. A., Hofmann-Wellenhof, R., & Halpern, A. (2018). Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. The Lancet Oncology, 19(3), 328-337.

[15] Li, Z., Keel, S., Liu, C., He, Y., Meng, W., Scheetz, J., ... & Taylor, H. R. (2019). An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. Diabetes care, 42(9), 1620-1626.

[16] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[17] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[18] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402-2410.

[19] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[20] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402-2410.

[21] Mishra, N., Biswal, R., Acharya, S., & Choudhary, P. (2018). Plant species identification using deep learning. 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1470-1474.

[22] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211-252.

[23] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[24] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[25] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[26] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708)

[27] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Appendix A

Algorithm 1 Data Augmentation Algorithm (DDA)

Input:

Benign B, Malignant M, Normal N segmented mammo- gram image.

Processing:

Step 1: ∀ B, rotate to 0°, 90°, 180°, 270°

Step2: Perform flip on all step1.

Step3: ∀M, rotate to 0°, 90°, 180°, 270°

Step4: Perform flip on all step3

Step5: ∀N, rotate to 0°, 90°, 180°, 270°

Step6: Perform flip on all step5

Repeat for all training data

Output:

Steps 1,2,3,4,5,6

Acknowledgement

We would like to extend our heartfelt gratitude and appreciation to our esteemed Professor and Guides Dr. Anudeepa Kholapure for their unwavering support, guidance, and motivation throughout the project. Their invaluable insights, suggestions, and feedback have been instrumental in shaping the project into what it is today.

We would also like to thank our respected Principal Dr. Shubha Pandit, Head of the Electronics and Telecommunication Engineering Department Dr. Bharti Singh, for providing us with this opportunity to undertake our Final year project.

We would also like to express our sincere gratitude to all the authors of the research papers and online sources that we have referenced in our project. Without their valuable contributions, our project would not have been possible.

Lastly, we would like to thank our family and friends for their unwavering support and encouragement throughout this project.