NON-INVASIVE CANCER DETECTION FROM BIOMEDICAL IMAGES USING EFFICIENT MACHINE LEARNING APPROACH

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Electronics and Communication Engineering

By

EELANDULA KUMARASWAMY

Registration Number: 41900670

Supervised By

Dr. Sumit Kumar (UID: 24786)

School of Electronics & Electrical Engineering (Associate Professor)

Lovely Professional University

LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2024

DECLARATION

I, hereby declared that the presented work in the thesis entitled "**Non-Invasive Cancer Detection from Biomedical Images Using Efficient Machine Learning Approach**" in fulfilment of degree of **Doctor of Philosophy (Ph. D.)** is outcome of research work carried out by me under the supervision **Dr. Sumit Kumar**, working as Associate Professor, in the School of Electronics and Electrical Engineering of Lovely Professional University, Punjab, India. In keeping with general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of other investigator. This work has not been submitted in part or full to any other University or Institute for the award of any degree.

(Signature of Scholar)

Name of the scholar: EELANDULA KUMARASWAMY

Registration No.: 41900670

Department/school: Electronics and Communication Engineering

Lovely Professional University,

Punjab, India.

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled "**Non-Invasive Cancer Detection from Biomedical Images Using Efficient Machine Learning Approach**" submitted in fulfillment of the requirement for the reward of degree of **Doctor of Philosophy (Ph.D.)** in the Department of Electronics and Communication Engineering, is a research work carried out by **EELANDULA KUMARASWAMY**, Registration No. **41900670,** is bonafide record of his original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.

(Signature of Supervisor)

Name of supervisor: Dr. Sumit Kumar

Designation: Associate Professor

Department/school: School of Electronics and Electrical Engineering

University: Lovely Professional University

ABSTRACT

Medical image processing is a non-invasive diagnosis tool for identifying abnormalities and determining disease stages and progression. Imaging modalities help pathologists to allow visual representation of the function of organs of a body for clinical analysis. Biopsy, a microscopic study conducted on blood samples to visualize abnormalities and generate digital histopathological images, entails a rigorous process. This process begins with collecting a blood sample on a glass slide, followed by stain normalization and color augmentation steps. Histopathological images consist of highly informative data and a variety of approaches used to analyze disease details from it. In the decision-making process, manually studying the visual patterns in histopathological images by medical professionals is complicated, limited precision, difficulty in identifying the accurate boundary of the region of interest, time-consuming, and human interpretation is subjective.

With emerging technology, the CAD system is widely used to analyze histopathological images. Even though, cancer statistics monotonically increasing over the past decade. In this connection, an essential need is to focus continuously and update the medical studies along with the new approaches in automated CAD systems to diagnose the disease, increase the life span of cancer patients, and reduce the mortality rate. Accurate identification and classification of the histopathological images is a very challenging task due to the variety in the tissue's patterns and challenges in automated CAD systems, such as scarcity of data, model under-fitting, over-fitting, computational complexity of the model, etc. In addition, automated systems mainly depend upon the type of data used to train the model, feature extraction, and classifier approaches opted.

In this work, a non-invasive cancer diagnosis approach using histopathological images has been proposed. The model has been trained and tested on histopathological images of the newly available dataset, called Databiox. CNN as a potential feature extractor, and various ML classifiers to categorize the grades. Later, magnification-independent and dependent classification of grades performed with different pre-trained CNNs using a transfer learning approach, achieved significant improvement in the classification accuracy. Finally, an ensemble CNN has been developed, achieved the highest classification accuracy with data augmentation, and the obtained results are also compared with the existing state-of-the-art techniques.

ACKNOWLEDGEMENT

"No one who achieves success does so without acknowledging the help of others"
-Alfred North Whitehead

During the virtuous journey of this thesis work, I have come across many people who have supported and motivated me throughout my research work and thesis preparation.

I would like to express my deep and sincere gratitude to my supervisor, Dr. Sumit Kumar, Associate Professor, School of Electronics and Electrical Engineering, Lovely Professional University, Punjab, for allowing me to do research and providing valuable support, and motivation throughout the research work. It was a great privilege and honor to work under his guidance. I extend my thanks for his consistent guidance, support, and help in the research, publications, and writing of my thesis.

I am grateful to the Dean and the members of the Centre for Research Degree Programme (CRDP) for being part of my entire research progress review panel and providing valuable suggestions during this period. I extend my thanks to all anonymous reviewers of my research papers submitted to various international journals and conferences, whose feedback enabled significant improvements in the work presented here. I am indebted to the Honourable Chancellor, Dr. Ashok Mittal, Lovely Professional University for providing all the unwavering support throughout the execution of this work. I express my heartfelt gratitude to the Head and staff of the School of Electronics and Electrical Engineering for their cooperation and support.

Additionally, I would like to express my gratitude to Dr. Shallu Sharma, a Scientist at the National Brain Research Centre, Manesar, Haryana, India, for her valuable advice. No words of thanks are sufficient to convey my deepest gratitude and sincerest love to my parents, whose unwavering faith in me kept me strong in all circumstances. I am also thankful to my wife, Sridevi, for her wholehearted, endless support and patience.

I extend my gratitude to everyone who has supported me in completing the research work, whether directly or indirectly. Finally, I bow my head before the Almighty God to shower his mercy on me for every moment.

EELANDULA KUMARASWAMY

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
AUC	Area Under the ROC Curve
CAD	Computer-Aided Diagnosis
CNN	Convolution Neural Networks
DCIS	Ductal Carcinomas in Situ
DCNN	Deep Convolutional Neural Networks
DT	Decision Tree
FN	False Negatives
FP	False Positives
GLCM	Gray Level Co-Occurrence Matrix
GPU	Graphics Processing Unit
H&E	Hematoxylin and Eosin
HOG	Histogram of Oriented Gradient
IARC	International Agency for Research on Cancer
IDC	Invasive Ductal Carcinomas
KNN	K-Nearest Neighbour
LDA	Linear Discriminant Analysis
LR	Logistic Regression
MRI	Magnetic Resonance Imaging
NB	Naive Bays
PCA	Principal Component Analysis
PET	Positron Emission Tomography
ResNet	Residual Network
RF	Random Forest
ROC	Receiver Operating Characteristic
SVM	Support Vector Machines
TN	True Negatives
TP	True Positives
VGG	Visual Geometry Group
WHO	World Health Organization

LIST OF TABLES

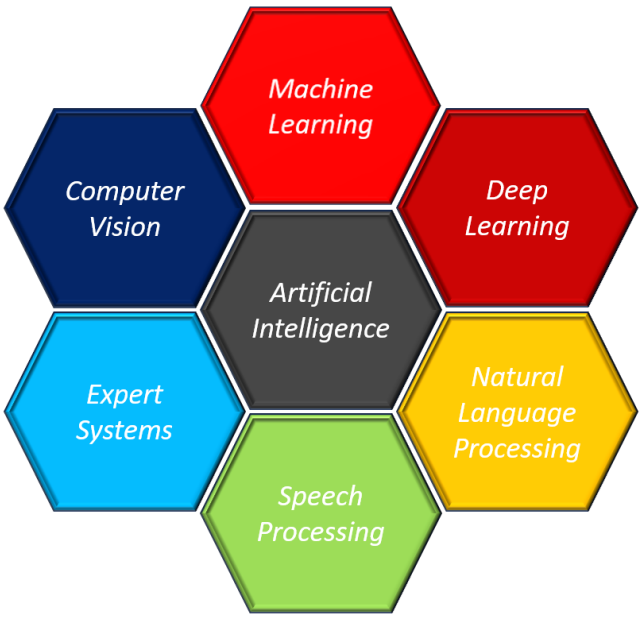
LIST OF FIGURES

1

Introduction

"Artificial Intelligence and humans will solve society's biggest challenges by working together"

– Fei Fei Li



In this chapter, an introduction to cancer particularly breast cancer, and related statistics across the world were presented. In the detection of cancer, image classification approaches especially, histopathological image classification using conventional machine learning, and deep learning techniques in the categorization of breast cancer were discussed. The feature extraction approaches such as conventional feature extractors, and convolutional neural networks, emphasized cancer detection applications. Additionally, the progression of convolutional neural networks from LeNet is also discussed. A detailed description of the transfer learning approach and the designed model's performance evaluation metrics were considered.

1.1 Introduction to Cancer

A disease is a particular irregular disorder that affects an organism's or a cell's portion or entire function differently, but that might not be traced instantly to any exposure to outside injury. However, health conditions are regarded as diseases when they are linked to particular signs or symptoms. A disease could have multiple sources like infections, organ functional disorders, or internal injuries. Hence, irregularities in internal immune cells are the source of several diseases. Cancer is a complex disease marked by the uncontrolled and abnormal growth of cells, forming masses or tumors [1]. In general, the immune system within the body regulates uncontrolled cell division; however, due to changes in a specific gene, some abnormal cells expand, and cover, but are unable to notice. The abnormal cells multiply uncontrollably and spread across the body, which further forms a mass of tissues, called tumors. There are about 100 diverse forms of cancer that have been found, each with its causes and symptoms. The names of the cancers are given based on the area of the body from where they originate to identify one from another. The tissues can form anywhere in the body, and the type of cancer is determined by the position of the tissue. There have been found A to Z forms of cancer in human beings, including breast cancer, lung cancer, liver cancer, and so on. It may be classified as malignant or benign depending on the severity, spread, or invasion of tissues to further parts of the body. As the mass of tissue grows and does not invade nearby tissue, called benign. If the tumor spreads and forms as distinct tissues, called malignant. Metastasis is the process of breaking the cancer tissues from the primary tumor and traveling to other locations of the body, forming a secondary tumor [2].

The four main forms of cancer are carcinomas, sarcomas, leukemias, and lymphomas. Carcinomas, which grow in the skin or the exterior layer of cells lining the inner organs, glands, and body cavities, contribute to around 90% of all human cancers [3]. Carcinomas tend to occur in the breast, colon, lung, prostate, and skin tissues [4]. Sarcomas, which are less frequent than carcinomas, are caused by the alteration of cells in connective tissues such as bones, cartilage, muscles, or fats. The leukemias kind of cancer does not produce solid tumors, which causes an excessive number of undeveloped white blood cells to be generated from the bone marrow. Lymphomas are cancers of the lymphatic system, a portion of the immune system that filters the blood

and tissues and is made up of lymph, lymph arteries, and lymph nodes.

As per the Global Cancer Observatory [5], the total estimated cancer incidents are 1,23,56,779 cases across ten different countries and 62,51,255 deaths in 2020, shown in Figure 1.1. Nearly one-third of the new cases (74%) and more than one-third of cancer deaths (78%) in 2020 occurred together in China, USA, India, and Japan. Out of these four countries, only the USA has a lesser percentage of deaths as compared to cancer incidents, because of the availability of a diagnosis system, early-stage treatment, and also, awareness among the people in these countries. Delay in cancer diagnosis and treatment is the major reason for the higher mortality rates, and updated diagnosis systems are not available in most of the countries.

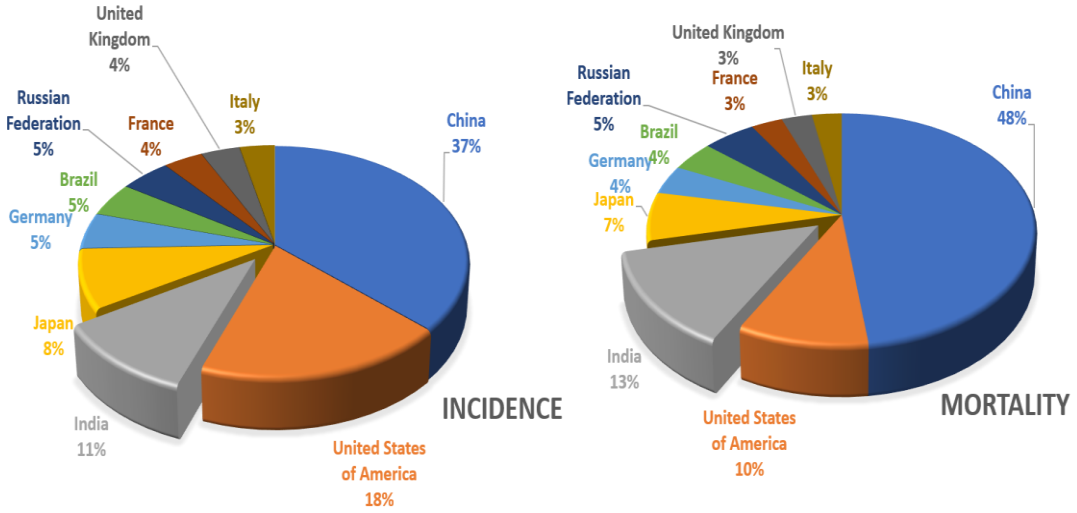


Figure 1.1 Global cancer incident cases and deaths across ten different countries.

It is necessary to update, develop efficient methods for cancer diagnosis, and provide medical care to reduce the death rate from the cancer. Among all diagnosis techniques, images are more helpful tool for visualizing everything, including the internal and external organs of the human body. Many different fields, including health, agriculture, the military, security, and others, have benefited from improvements in image processing techniques. To obtain the necessary information for a given task from image data, variety of image processing techniques turn the image into a digital form and modify a few pixel values. Additionally, computers can extract the most pertinent data from digital images using machine learning, a branch of artificial intelligence, used in decision-making processes [6]. Computers learn from the data supplied to them directly

without the intervention of human beings. Day by day, most of the activities of humans are automated by systems with artificial intelligence. For example, biometrics and iris are used for security authentication, virtual assistance in most problem-solving situations, pattern recognition, face recognition, speech recognition, fraud detection, and many more. Nowadays, most researchers, scientists, and investigators work continuously in this field to develop innovative algorithms and modify existing approaches with advancements in computer technology. Aiding the machine learning approaches in the analysis of medical images helps radiologists and doctors to diagnose cancer quickly, and effectively [7, 8].

1.2 Breast Cancer

Breast cancer is the foremost cancer incidence type across the world. The different breast components, normal and abnormal tissues in the duct are shown in Figure 1.2.

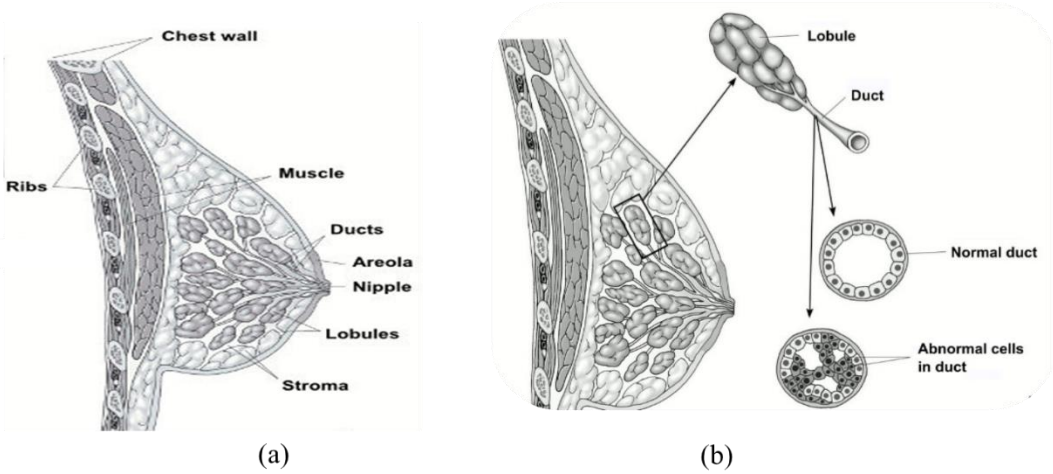


Figure 1.2 Illustration of (a) different breast components, (b) normal and abnormal cells in ducts. (Source:https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html)

The most common cancers are listed here in ascending order according to the number of cases accounted in the year 2020, as per the World Health Organization (WHO) and International Agency for Research on Cancer (IARC). Breast cancer of 11.7 % (22,61,419 cases), lung cancer of 11.4 % (22,06,771 cases), Colorectum of 10% (19,31,590 cases), prostate of 7.3% (14,14, 259 cases), the stomach of 5.6% (10,89,103 cases), liver of 4.7% (9,05,677 cases), Oesophagus and Cervix uteri cancers of 3.1% (each of 6,04,100 cases) remaining cancers of 42.9% (82,75,743 cases) out of total 192,92,789 cancer incident cases, worldwide, in 2020. The various cancer types versus

number of cancer cases (incidence and mortality) bar graph plot is shown in Figure 1.3. (Source: https://gco.iarc.fr/today/data/factsheets/cancers/20-Breast-fact-sheet.pdf).

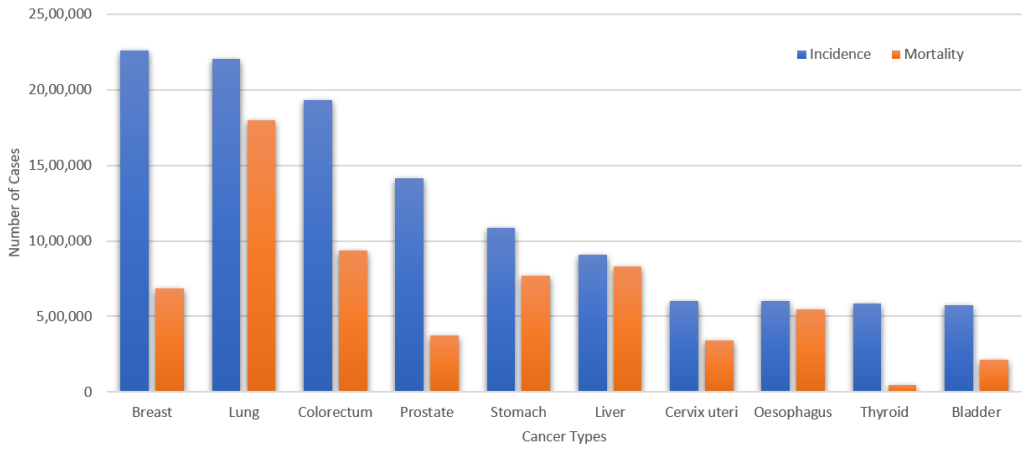


Figure 1.3 Number of cancer incidence and mortality cases of different cancer types in the year 2020 according to Global Cancer Observatory.

Detection of cancer at an initial stage diminishes the death rate, it is possible with the help of advanced screening techniques. Non-invasive screening techniques such as X-ray, ultrasound scan, mammography, MRI, breast-specific gamma imaging, positron emission tomography (PET), and thermal imaging are the most adopted imaging techniques [9, 10]. A small amount of tissue from any abnormal portion of the body part used to investigate under the microscope to confirm whether the tissues are cancerous or normal, called Biopsy. All the imaging-based screening methods are the prime and provide detailed information to diagnose the disease. Despite of variety of image-based screening methods available, the final decision of cancer diagnosed by analysis of cells and tissues by pathologist. The biopsy is the only method to get detailed information about the cells and tissues. Biopsy is an invasive screening technique, in which histopathological images generated using Hematoxylin and eosin (H&E) staining process on tissue sections viewed through microscopy [11]. In the biopsy, the collected tissue samples are fixed over a glass slide. The sample tissue is sliced into sections or slides, H&E staining process helps in extracting the important information about the shape and size of the cells from the tissue sample slides. The histology screening approach is used to investigate a variety of cancer types. Digital histopathological images are widely used to investigate breast cancer diagnosis tools with efficient machine learning techniques to reduce the global cancer mortality rate.

1.3 Motivation

As per recent cancer statistics, the number of cancer incidences and mortalities rises globally. Invasive type of biopsy modality image screening allows detailed analysis of tissue patterns. Pathologists visually inspect the histopathological images to differentiate between cancer and normal tissues and identify the stage of cancer. As the number of incidences increases, a proportional rate of pathologists is not available, and the burden on limited pathologists to get the quality analysis. Also, visual inspection of tissue patterns is difficult and time-consuming process, whereas the diagnosis of these images by a medical professional is subjective, and the chance of occurrence of error is depending on the skills of a medical professional. A minute error may lead to a delay in the detection of cancer and might cause serious impacts on patient's health condition.

The digitization of tissue samples and histopathological images enables us to use digital computational approaches to analyze the pathological images using machine learning, and image processing techniques [12]. The technological advancements in digital histology tend to find small details in the histopathological images, which are difficult to observe with the human eye. Advanced algorithms also derived revolutions to assist the detection and development of the CAD [13]. Nowadays, various factors influence the execution of the CAD system. Pathologists considered the analysis of computer-aided system reports in the determination process. CAD with machine learning algorithms helps in this part, most of the research has implemented CAD-based approaches in the health sector to diagnose cancer. However, absence of CAD, humans take a long time to analyze plenty of medical images with no precision, even a specialist cannot guarantee precision. CAD with machine learning overcomes the stated problems and added advantages to pathologists to detect and identify accurately in the decision-making process. In this context, the development of a robust and reliable automated cancer disease analysis and detection system is essential. An efficient and effective CAD system is a challenging task to develop and has potential to reduce the mortality rate.

1.4 Breast Cancer Classification

Here, we discussed a generalized system for breast cancer tissue image classification shown in Figure 1.4. Any automated system with artificial intelligence-based algorithm is used to diagnose and categorize the tissue classes. The classification system consists

of the following steps.

1.4.1 Data Selection and Collection

Implementation of any diagnosis system and its performance depends on the type of data. How the selected data is going to be collected is also an important task.

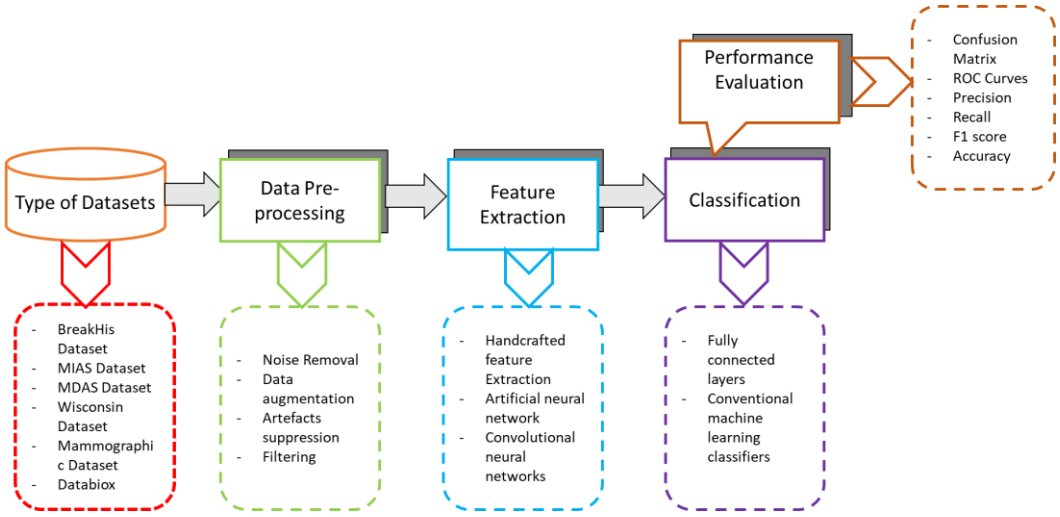


Figure 1.4 Flow chart of breast cancer classification system

The first step in the development of any automated diagnosis system is the type and size of the dataset. Data selection and collection is deciding the type of system to design, that is, if the dataset is in CSV file format the algorithms are to sort the data and identify the features, and decide whether it is relevant to benign or malignant. The histopathological image data is the image file format dataset, which contains a huge amount of information that needs to be extracted effectively to classify it into a particular class.

- **Databiox Dataset**

For this work, Histopathological image dataset called Databiox is used. The dataset available publicly at https://databiox.com/ has been selected and downloaded. The Databiox dataset consists of images collected from 124 IDC (invasive ductal carcinomas) diagnosed patients between the years 2014 and 2019 for grade classification. The IDC breast cancer dataset is a new database that contains histopathological images at 4x, 10x, 20x, and 40x magnification factors [14]. A total of 922 images, 259 of grade 1, 366 of grade 2, and 297 of grade 3; collected from 37, 43, and 44 patients, respectively. All images are RGB color images collected through microscopic imaging stained with H&E and saved in JPEG-type format with a

resolution of 1276 x 956 and 2100 x 1574 pixels. All the specimens in this dataset are labelled based on modified Bloom Richardson histologic grading and organized accordingly. An illustration of image samples of grade 1, grade 2, and grade 3 from the Databiox dataset are depicted in Figure 1.5.

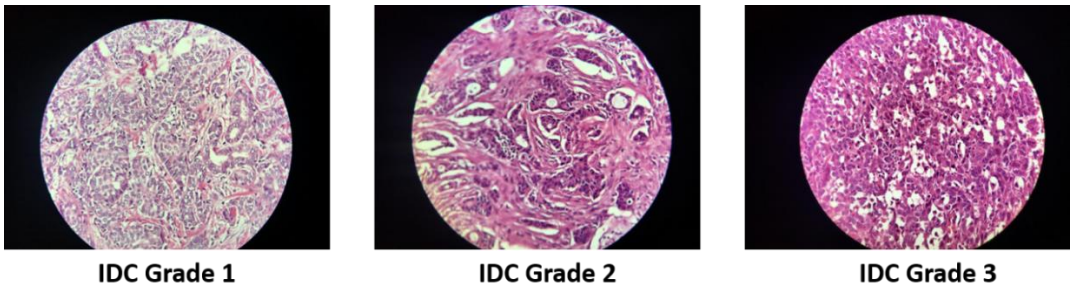


Figure 1.5 An illustration of image samples of different grades namely, IDC Grade 1, IDC Grade 2, and IDC Grade 3 from the Databiox dataset.

The tubular formation, mitotic count, and nuclear pleomorphism are the three important factors based on which the IDC grading is performed. Each factor has a range of scores from 1 to 3, and scores of all three factors are added. If the added score lies in the range 3 to 5 assigned as grade 1, 6 to 7 as grade 2, and 8 to 9 assigned as grade 3, respectively. The analysis of histopathological images depends on various factors such as fixation, staining, tissue sample preparation, and mounting. Inter and intra-observer variability in the manual analysis is also one of the major issues in analyzing the histopathological images and makes it challenging to classify the images of different grades accurately. Thus, it becomes essential to develop a robust and accurate grade classification algorithm for precise diagnosis and prognosis of breast cancer [15-17].

Table 1.1 The distribution of image samples in the Databiox dataset at four different levels of magnification. The IDC grades also referred to as grade class 0 (Grade I), grade class 1 (Grade II) and grade class 2 (Grade III).

IDC grades	**Number of Patients**	**Magnification levels**				**Total number of images in grade**
		4x	**10x**	**20x**	**40x**	
Grade I	37	45	40	43	131	259
Grade II	43	59	64	63	180	366
Grade III	44	56	49	49	143	297
Total	124	160	153	155	454	922

In Table 1.1, the dataset's image distribution is tabulated. The images taken from the dataset are RGB images in the 3024x4032x3 pixel JPEG format. The file names in this dataset are in the format i.e., 01_BC_IDC_9054_4x_1. The description for this format is demonstrated in Table 1.2

Table 1.2 Databiox dataset file structure details

Sl. No	Name in the part of the file structure	Sample
1	Sample number	01
2	Cancer type	BC_IDC
3	Pathology archive number	9054
4	Magnification level	4x
5	Number of specified magnification level	1

1.4.2 Image Pre-processing

Most of the time, the available raw data cannot be used directly in the training of the model; it must first undergo some pre-processing. A variety of image pre-processing methods are used to convert obtained raw data into the most appropriate form for developing a model because, in most cases, raw data may contain image artifacts and inconsistencies. The type of pre-processing techniques may have a significant impact on how well the model performs. To improve the performance of the diagnosis system, redundant information and artifacts in the obtained raw data has to be removed. Some of the pre-processing operations include scaling, blur removal, noise reduction, artifact removal, and geometric alterations.

The classification problems might be executed directly on the histopathology images or after performing segmentation on histopathological images. Some methods partition the region of interest (ROI) into segments, and extracting the features from the ROI. The majority of methods, however, simply extract the features directly from the provided images.

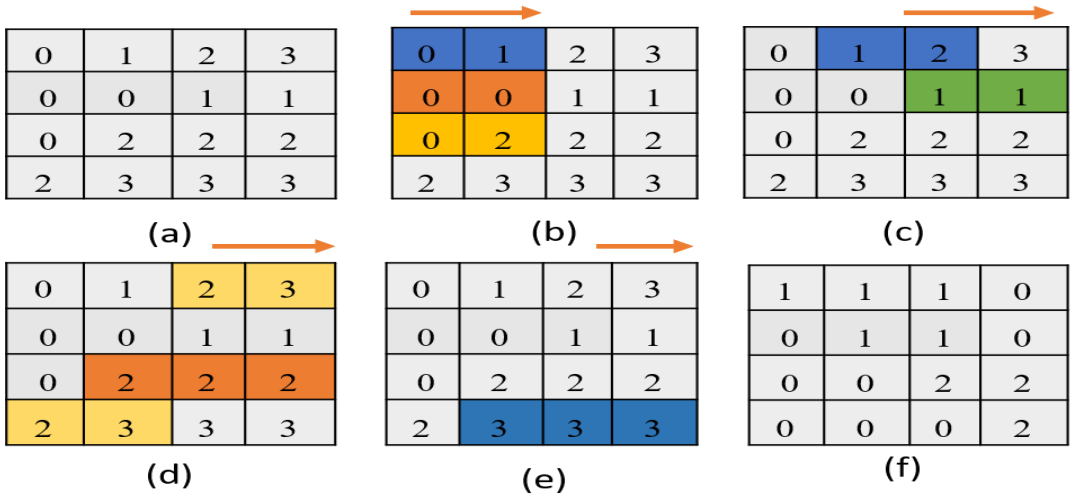
1.4.3 Feature Extraction

Since raw data typically has a high degree of dimension, using it for classification tasks can be challenging. Feature space with low dimensionality leads to more relevance to particular class in the classification tasks. Therefore, selecting important features is a

crucial step. The feature extraction is initiated to select informative values derived from the image and eliminate redundant data. Wherein the input data is too large to process, it needs to be converted into a reduced set of features (feature vectors) [18]. Whereas feature extraction is a function to create a new set of features from the original set of huge features. The extraction of significant features is a vital and thought-provoking approach in the computer-aided diagnosis of breast cancer from histopathological images. The feature is the most significant attribute of the information and is used in the process of identification of abnormal tissues, and diagnosing any type of disease. Therefore, efficient extraction of features is the prime role in the computational task involved in image processing applications. The significant features are precisely extracted from the histopathological image and represented in a reduced set of attributes, called feature vectors. For histopathological image feature extraction, the most important features, color, shape, and texture are considered. As per the separation level, colour is a pixel-level descriptor, whereas shape and texture are global-level descriptors. Each of these descriptors is significantly utilized to identify some special characteristics and distribution to differentiate the cancerous and normal tissues of the histopathological images. However, to evaluate the diagnosis system performance the extracted features are the essential thing that decides its efficiency.

- **Color Histogram:** Generally, a histopathological image staining procedure is performed to differentiate the different attributes of the microscopic tissue sample such as nuclei, stroma, or cytoplasm. For that reason, color is one of the important features of a histopathological image, which needs to be considered while identification of cancerous cells. Hematoxylin dyes cell nuclei with blue, whereas eosin stains cytoplasm, connective tissue, and other tissues with red or pink. The stained procedure is used to enhance color, contrast, and other features under the microscope. Color histogram is the approach generally used to identify the color distribution in histopathological images. Color histogram plot describes the distribution of colors and number of pixels having that color [19]. The histogram determines the probability of occurrence of each color pixel in the image.
- **Haralick Texture:** The texture is another important feature, which brings information about the surface and presence of objects in the image. Based on the degree of randomness texture property either random or stochastic basis.

Stochastic texture patterns appeared mostly in histopathological images and to assess the texture feature analysis, it is required to compute statistical attributes. Gray level co-occurrence matrix (GLCM) is used to compute the statistical attributes in the spatial domain [20]. The commonly used GLCM method was proposed by Haralick in 1973. This method is used to determine the characterization of texture using various statistical features, such as correlation, sum average, variance, angular second moment, etc. A GLCM is a matrix with the same number of rows and columns as the number of gray levels in the image. The GLCM frame work has been shown in Figure 1.6.



Reference pixel	Neighbours pixel value			
	0	1	2	3
0	(0,0)	(0,1)	(0,2)	(0,3)
1	(1,0)	(1,1)	(1,2)	(1,3)
2	(2,0)	(2,1)	(2,2)	(2,3)
3	(3,0)	(3,1)	(3,2)	(3,3)

Figure 1.6 GLCM framework, (a) 5x5 image, (b) to (e) computation of gray level co-occurrence along the horizontal direction with the combination of pixel values shown in (g), and (f) is the calculated GLCM matrix along the horizontal direction of the input (a).

- **Hu moment:** The shape of cells is another main feature to determine the characteristics of the cells, either malignant or not. Irregular development of the cells, random shape, and size of the cells are the morphological characteristics of

cancer cells. The Hu-moment is widely used to determine the shape or match the shape of normal tissue among the abnormal. Various descriptors such as geometric moments, rational moments, complex moments and orthogonal moments are commonly used to characterize the patterns [21]. Moment variants are widely used to recover information from the image. The significant property of these moment invariants is independent to shape, position, and alignment of image content. However, in the classification of histopathological images, color distribution, texture and the shape of the biological structure are the major features.

The distinctive varieties of feature extraction approaches are proposed by researchers in the past. However, these methods are broadly categorized into two types: statistical and learning-based methods. Some feature descriptors include scale-invariant feature transform, histogram of oriented gradient, graph-based features, mixed features such as mixing scale invariant feature transform, color histogram, and local binary pattern [22].

1.4.4 Classification

Extracted features generally feed to one or more classifiers to classify into one of the categories. Generalized strategies in each stage and a flowchart of the classification of the breast cancer system are depicted in Figure 1.4. Classification of breast cancer is performed by using various types of approaches, including machine learning techniques, ensemble learning techniques, and deep learning techniques as depicted in Figure 1.7.

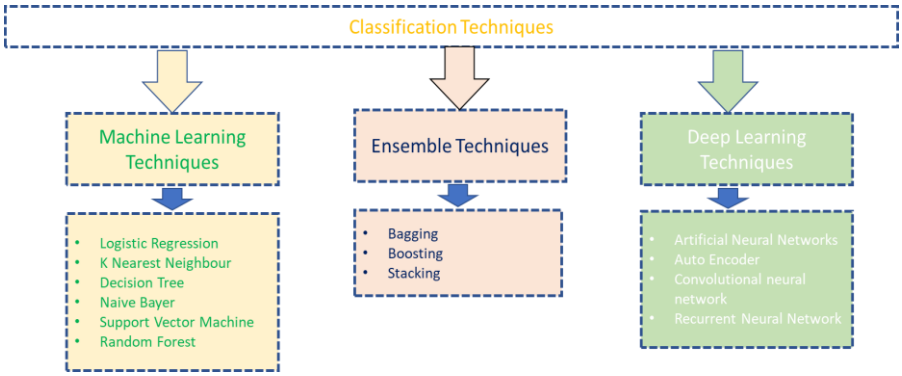


Figure 1.7 Classification techniques for breast cancer

1.5 Machine Learning Techniques

The field of machine learning encompasses a broad categorization, including supervised, unsupervised, semi-supervised, and reinforcement learning algorithms.

Among these categories, few algorithms were most popular for classification tasks are listed in the following section. The breast cancer classification system developed with conventional feature extraction approaches and classifier by machine learning algorithms shown in Figure 1.8.

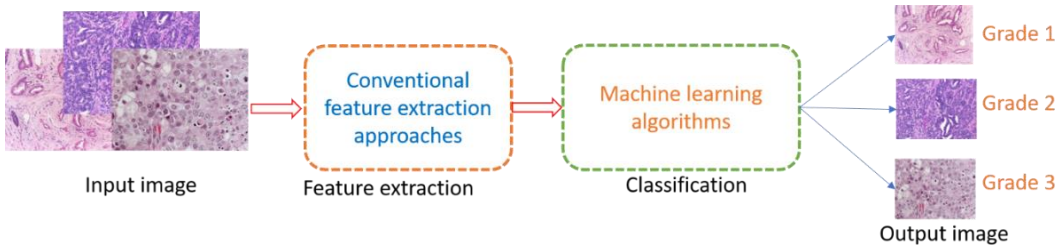


Figure 1.8 Machine learning strategy to classify breast cancer

The machine learning algorithms used to classify extracted features correctly to related classes or grades as follows.

- **Logistic Regression (LR**): It is employed for addressing both regression and classification problems, offering continuous outcomes for regression tasks and categorical outcomes for discrete class classification. The LR functions on sigmoid function or logistic function. It is used to determine the particular class based on the probability of that class and dependent variables. For instance, predicting whether the incoming email is spam or not spam. The Logit function is used in the classification problems, referred to as logistic regression [23].
- **K Nearest Neighbours (KNN):** It is a supervised algorithm used for the recognition and classification of patterns. It works based on the distance between the target, and its neighbors' count categorized into a particular class [24]. The distance function used to calculate the similarity between different classes in this algorithm is Euclidian, Minkowski, Manhattan and Hamming distance. Wherein Hamming distance function is used for categorical target problems and the rest of the functions for continuous target problems.
- **Decision Tree (DT):** DT is an algorithm used for classification tasks and regression problems. A tree is formed based on all-possible outcomes by comparing with threshold, branching methodology. The classification and regression tree is the subset of the decision tree approach [25]. The algorithm is used in various applications such as banks, finance, remote sensing and medical field, etc.

- **Naive Bayer (NB):** This algorithm suits the large training dataset and produces fruitful results using the Bayesian approach [26]. In a noisy environment, this approach could provide good accuracy. NB is a low-bias and high-variance classifier, applications of this algorithm include sentiment analysis, email spam filtering, document categorization and news article classifications.
- **Support Vector Machine (SVM):** SVM is an algorithm utilized for classification and regression problems. In this, the training data is classified by a hyperplane. It works well for nonlinear and complex problem tasks. In SVM support vectors are formed and based on these support vectors, the algorithm work [27]. The SVM algorithm provides better accuracy with linear, sigmoid, and radial basis function kernels.
- **Random Forest (RF):** It is the most frequently employed supervised learning algorithm for addressing both categorization and regression problems [28]. Generally, the RF algorithm utilized in the prediction of categorical class labels by fine-tuning a pre-trained network. This algorithm generates a bunch of decision trees with a random subset of data and bagging the decision trees called the bagging approach. This algorithm is used in various applications such as the automobile industry, banks, speech recognition, image and text classification.
- **Linear Discriminant Analysis (LDA):** It is a linear transformation procedure used for dimensionality reduction and categorization of two or more classes. The LDA and principal component analysis (PCA) are closely associated with the data reduction approach but PCA is unsupervised whereas LDA is multi class classification supervised algorithm [29].

1.6 Ensemble Techniques

The ensemble is a technique used to combine homogeneous and heterogeneous algorithms to configure a new method for classification tasks [30]. Homogeneous algorithms combined to configure the approach called bagging and boosting. Heterogeneous algorithms combined two or more base methods to configure the approach called stacking.

- **Bagging:** The name implies that it bags various individual models into a single model. These individual models were trained separately, in parallel and combined to perform the classification tasks. Bagging is used to reduce the

variance of the prediction model.

- **Boosting:** A classification model formed by combining the models sequentially to strengthen the weak learners and boost their performance. Step-by-step implementation of the weak models is trained individually and combined to boost its performance.
- **Stacking:** Combining weak models that were implemented based on different algorithms but using the same dataset. The name implies that heterogeneous algorithms merge to form a new model in the classification task.

1.7 Deep Learning Techniques

The deep learning approach, a subset of machine learning, serves to surpass the limitations of conventional machine learning algorithms, particularly in addressing classification problems. A deep form of a convolutional layers or layers of neural network is called deep learning. Deep networks are incorporated with sequence of convolutional layers, pooling layers, and fully connected layers in the architecture as shown in Figure 1.9. Deep learning models are designed deeper to enhance and extract the in-depth features of the images. The network can recognize and classify the patterns into different categories. The SoftMax function used in the last layer to categorize the particular class based on the probability distribution of all considered classes.

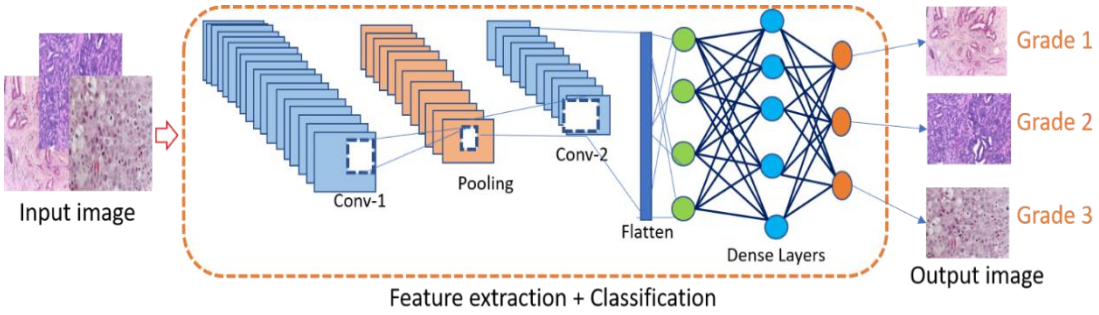


Figure 1.9 Deep learning strategy to classify breast cancer

1.7.1 Artificial Neural Networks (ANN)

With advances in computer technology, ANN is the most frequently used technique for data classification problems [31]. Networks are incorporated with fully connected neurons as an input layer, hidden layers, and output layers; In ANN, each layer consists fully connected artificial neurons. These techniques are popular in the classification of patterns in the histopathological dataset. Algorithms are used to simplify complex problems with parallel processing, distributed memory, and collective solutions.

1.7.2 Auto Encoder

The name implies encoding the input data and decoding it back to get the original [32]. The basic idea behind the encoder is to encode the important features from a huge number of data samples. The trained model ignores irrelevant and noisy information.

1.7.3 Convolutional Neural Network (CNN) Architecture

The convolution of image data with the kernels will extract features during the feature extraction phase. In the deep learning approach, CNN is the potential feature extractor and decides the efficiency of the model [33]. The various convolutional kernels are available for extracting features, used to reduce the redundancy in features, and minimizes complex calculations. It is based on the type of filters, and number of layers, more precise features extracted from the input image. CNN layers are divided into convolutional layers, and pooling layers and activation functions generally come after these layers. The final stage might be a fully connected layer or any conventional classifiers. It could classify the categorical data. All of these layers are combined, and named CNN.

- **Convolutional layer:** It is the prime layer in the construction of CNN, and performs filtering operations with the kernel to extract features and the local region of the input. The number of filters used represents the depth of CNN, filter slide movement over the image is called stride. Zeros are padded around the image matrix to ensure it maintains its original size as the input image matrix. Thus depth, stride and zero padding are the hyperparameters of the CNN, which controls the dimensionality of the output image.
- **Pooling layer:** Pooling serves as a nonlinear transformation that consolidates the output into a singular value, providing a summarized representation. In deep CNN, after each convolution layer, a pooling layer is inserted to reduce the redundancies and computational complexities. Generally, this single value is attained from the statistical analysis of neighbouring outputs of convolution operation, which makes the feature descriptions added robust and invariant to insignificant changes in the input data. The pooling operation gradually decreased the dimensionality of the input, accordingly reducing the need of memory for storing the parameters and advancing the computational efficiency. Inherently, pooling layers control the overfitting in deep networks.

- **Fully connected layer:** A full connection of all the neurons in the present and former layers, called fully connected layer. As per the problem, the final output neurons selected and fully connected. The output of each neuron depends on the associated weight and output of the preceding neuron. Compute the final stage neuron output score to the category in each case of classification.

The CNNs produce a significant performance for image processing applications. Local features connectivity and parameter sharing make CNNs computationally efficient by minimizing the number of training parameters. The stride value, padding value, size of the filter and number of channels are the influence factors to produce the output of CNN layers. Gradient descent (GD) algorithms are loss optimization functions used to diminishes the training loss while training the networks. One of the types of gradient descent is stochastic gradient descent (SGD), which diminish the training loss by changing the weights. In mini-batch gradient descent, the training data samples partitioned into small batches, referred to as mini-batch GD. In which, sum the gradient among the various batches leads to improve the model training speed.

CNNs are now the preferred approach for tackling challenging image classification, recognition problems, and other significant disciplines. They are also the method of preference for extracting features. Feature extraction is the most difficult task in classification problems, significantly influencing the overall performance of the model. Required to extract the features of the images by using global attributes like shape, color, and texture are important, and need to be extracted properly. Effective feature extractors for given problems are essential. Various CNN models are used to extract features by using different convolutional layers with multiple kernel sizes. The following are important factors that need to be considered while implementing any CNN such as the number of convolutional layers, size of the kernels in convolutional layers, type of the pooling layers used and followed, and type of activation functions used. CNN layer formation in each case may vary, but the operation principle of CNN remains the same. CNNs are the foundation for deep learning-based computation in the extraction of image features, which play a major role in classification tasks. There are diverse variants of CNNs are available. The development of CNNs over the years in chronological order is discussed here.

- **LeNet**

LeNet is the first convolutional neural network used to classify two-dimensional data such as binary and grayscale images. In the year 1998, at that time LeNet was implemented, with eight convolutional layers. To train the model there is no GPU and a low CPU available, it is not possible to train the model with high-resolution images, hence only grayscale images are used to train the model. The model is employed for the recognition of handwritten digits and the detection of digits on cheques [34].

- **Alex Net**

The typical Alex Net consists of five convolutional layers with the size of the kernel 11x11, 5x5, and 3x3, followed by three fully connected layers. It is an eight layered CNN, designed by the Super Vision group in 2012. The second network is Alex Net, which has a lot of similarities between Alex Net and LeNet, one significant difference or development in Alex Net is the depth. With more filters per layer and a greater number of convolution layers, it is substantially deeper. That is why it is called a deep neural network. In contrast to older models, the dropout function ensures that overfitting is prevented [35]. Some neurons with defined probability will be removed randomly in the dropout process by keeping all other input and output layers and neurons unaltered, and model parameters will also update accordingly to the learning method of the network. It will continue in the next iteration that randomly removes some neurons until the end of the training process or the number of epochs.

- **VGG Net**

The Visual Geometry Group (VGG) from Oxford University is the mastermind behind the development of this network; hence it is named VGG and its architecture shown in Figure 1.10. VGG16 has 16 different layers, these are the layers got have tunable parameters, and other layers like max pooling, do not contain any tunable parameters [36].

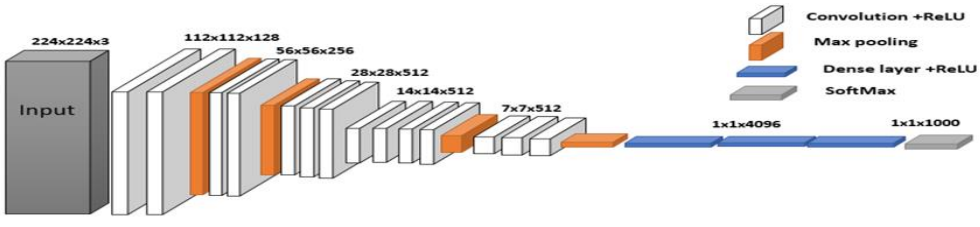


Figure 1.10 VGG architecture

Rectangular boxes indicate the convolutional layers along with rectified linear units (ReLU). Overall, there are 16 layers, out of which 13 convolutional layers and 3 fully connected layers. Hence, it is called VGG 16. The final layer is the SoftMax which is 1000 classes of the image Net dataset. Input image of size 224x224x3, which is a color image, CNN accepts color images. Applied input image passing through a sequence of convolutional layers, to preserve the size and resolution of convolution output, row and column padding used. Convolution operation has a relatively small receptive field of 3x3 and stride of 1 (Feature map). There are 13 convolution layers and 5 max pooling layers of window size 2x2, with stride 2. There are 3 fully connected (FC) layers, out of which, 2 FC layers have 4096 channels each with ReLU as an activation function and the last FC layer has 1000 channels, SoftMax layer with 1000 channels, one for each class of images in ImageNet database. All variable size kernels as in Alex Net can be realized using multiple kernels such as 3x3, 5x5, 11x11, etc. This realization is in terms of the size of the receptive field covered by the kernels. The architectures are Google Net and other is Residual Net, how these two networks address the issue of the vanishing gradient problem will look while studying these networks [37].

- **Inception module**

The networks in the network (NIN) strategy, called inception module, computes 5x5, 3x3, and 1x1 convolutions within the same network and covers a huge area, by preserving tiny details and resolutions of the images [38]. By using various kernel sizes in parallel from the lowest kernel 1x1 to a larger one 5x5, and the use of a 1x1 kernel reduces the number of computations significantly. Every inception module acts as a multi-level feature extractor. There are 9 such inception modules in Google Net, and Google Net tackles the vanishing gradient problem using an auxiliary classifier [39].

- **Residual Network (ResNet)**

ResNet uses a shortcut connection called a skip construction or identity construction that omits one or more levels [40]. The performance should not be affected by layering as opposed to its shallow counterpart. The weight layer learns F(X)=H(X)-X. By stacking identity mappings, the resultant deep network should give at least the same performance as its shallow counterpart. A deeper network should not give higher training errors than a shallow network. During learning the gradient can flow to any earlier network through shortcut connections alleviating the vanishing gradient

problem. Figure 1.11 illustrates the skip connection used in ResNet architecture.

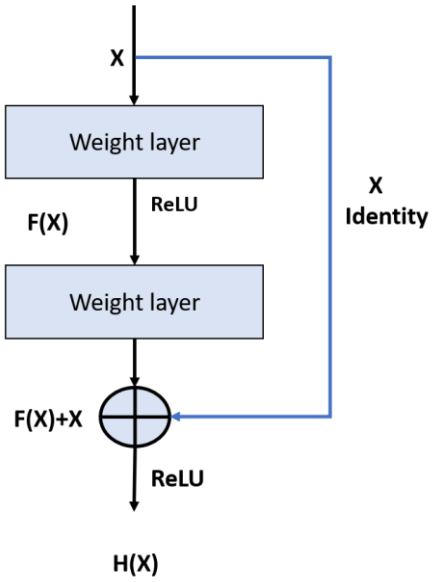


Figure 1.11 Skip connection used in ResNet architecture

1.8 Transfer Learning Approach

CNNs naturally need huge amounts of data and high computational graphics processing units (GPU) for quick training. But in reality, there are many fields where enormous data is not obtainable. For instance, the collection of annotated large dataset of medical data is a very tough task due to various privacy and other issues involved while collecting data from the patients. A huge amount of medical dataset involves, incomplete samples owed to which the task of CNNs training from scratch turned out to be a long process. To overcome these issues, the idea of a transfer learning approach introduced [41]. Transfer learning is a design methodology, that a model trained from a well annotated huge amount of data adopted for another new task, a shortage in training data. In a CNN, diverse layers follow to extract the lower-level features, mid-level features and problem related features by initial, mid and final layers respectively. In order to use the pre-trained network for another task, end layers of the network need to train on the new dataset in transfer learning approach. This method offers the ability to design a compacted machine learning model under little amount of training dataset since the employed model is previously trained. Beside that it supports in dropping a huge computational complex problem otherwise it may yield longer time to train the same model from scratch for a complex task. The transfer learning approach in deep networks are applied in three different ways to utilize the pre-trained networks as: i) baseline

model or end to end learning, ii) feature extractor, and iii) fine-tuning of the pre-trained network. In transfer learning approaches, these strategies available to implement a model and it helps in the following points to reduces the computational complexity, hardware, and software required to train the model from scratch, training time, and GPU used for accessing huge of data.

The first approach is the baseline model, in which a pre-trained network is used to categorize problems by training from scratch with arbitrarily initialized weights. In this strategy, the constructed pre-trained model used and alter it to the task specific model trained with respect to the new dataset. The final FC is changed with a new one which contains the number of nodes corresponding to the targeted application [42].

In the second approach, a pre-trained network is used as feature extractor and a fully connected network with activation functions or a conventional machine learning classifier used in the last layers of the network [43]. Generally, lower layers of the pre-trained network extract the most common features applicable for all types of tasks and higher layers of the network are task-specific.

Finally, the representation learning capability of CNNs is the key reason, because this capability allows the initial layers of CNN to capture fundamental lower-level features that are universally applicable to a wide range of computer vision problems [44]. However, the higher layers specific to problem related features from input image. Since fine-tuning of some high-level layers on another problem but relevant, sometimes acceptable for significant performance. The fine tuning a pre-trained model has two strategies. One is, layer-wise fine-tuning, second, partial training of the network. In the first strategy, layer-wise fine-tuning is initialized with the training of the layer and followed by other layers. It is an operative approach to test and compute the efficient depth of the network for fine-tuning. In the second strategy, partial training of the network performed, that is weights of initial layers are reserved and only higher layers are experienced for training on the new task related data.

1.9 Data Augmentation

It is the process used to increase the size of the dataset to solve the problem of limited dataset related tasks [45]. However, limited and unbalanced data size is one of the primary challenges in the development of automated systems. Data augmentation approaches such as flipping, cropping, interpolation, translation and rotation are the

operations performed on input data to grow the data size. The augmentation techniques applied to regular images may not be suitable for medical images due to the distinct problem-solving approaches required in the context of medical image analysis.

In addition to that, some of the augmentation approaches lead to the loss of information and others increase the noise of the images. Histopathological images have the properties such as rotational invariant and reflection symmetric. Therefore, rotation and reflection can be suitable data augmentation approaches for histopathological images. Another advantage of the augmentation approach is that it minimizes the overfitting problem as well.

1.10 Performance Evaluation Metrics

In the field of automated computing systems, the effectiveness of any machine learning based model is measured based on performance evaluation metrics. The choice of these metrics depends on the type of the problem statement. The present work deals with the classification problem so confusion matrix, accuracy, recall, receiver operating characteristics (ROC), F1 score, precision, and area under the ROC curve (AUC) are employed as the performance evaluation metrics. The performance of the particular classification model is assessed by using suitable metrics. The following are the defined metrics along with mathematical expressions.

- **Confusion Matrix**

  The most intuitive and simplest metric in classification problems to calculate the accuracy and precision of the model is the confusion matrix as shown in Figure 1.12. It is appropriate for binary classification [46].

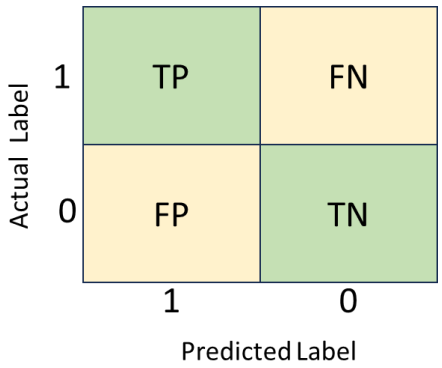


Figure 1.12 Confusion matrix

The confusion matrix by itself is not a performance metric but allows calculating some valued performance measures based on the value of true positive (TP), false

positive (FP), true negative (TN), and false negative (FN). Study a problem of binary classification, where the person with cancer is denoted as 1 and the person without cancer as 0. Wherein, the actual labels are represented in rows and predicted labels in columns. The terms TP, FP, FN and TN are represented as follows.

TP: if a person has cancer (1) and is predicted as cancer (1), termed as true positive.

FP: if a person without cancer (0) and is predicted as cancer (1), then is referred to as false positive.

FN: It is fair opposite to FP, where the person with cancer (1) in actual is predicted as without cancer (0).

TN: If a person has no cancer (0) and is predicted as without cancer (0), then is referred as TN.

Some key performance metrics derived from the confusion matrix include precision, recall, F1 score, and accuracy, as outlined below.

- **Precision, also called positive predictive value**

Precision defined as the ratio of true positives among the total positives retrieved. Mathematically expressed as

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

For a classification task, precision is about one of the particular class label predictions stated that count of only true positives divided by the total count of both true and false positive items categorized.

- **Recall, also called sensitivity or true positive rate**

Recall is defined as the ratio of true positives among the total retrieved. Mathematically expressed as

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

For a classification task, recall is about the count of true positive class by the total count of items that belong to that class. Both precision and recall are based on true positives. However, precision and recall are inversely related. To improve one at the cost of decreasing the other. For example, in the cancer tissue elimination process - higher recall improves the likelihoods of eliminating healthy cells and probabilities of removing all cancer cells. Higher precision declines the removing probability of

healthy cells and eliminating probability of cancer cells.

- **F1 score**

Basically, precision and recall metrics are inversely related. The performance of precision and recall are together measured by using new metrics called F-score.

The F1-score can be described as the weighted harmonic mean of both recall and precision.

Mathematically expressed as

$$F1\ Score = \ 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

Another metric used to compute the performance of the classifier is the true-negative rate. Specificity, another name of the true negative rate.

Mathematically expressed as

$$True\ negative\ rate(TNR) = \frac{True\ negative}{True\ positive + True\ negative}$$

- **Accuracy**

For only a balanced dataset, accuracy metric is used to assess the performance of the classifier. It can be a confusing measure for unbalanced data sets. Mathematically expressed as

$$Accuracy = \frac{True\ positive + True\ negative}{True\ positive + True\ negative + False\ positive + False\ negative}$$

- **Balanced Accuracy**

The mathematical average of true-positive rate and true negative rate, called balanced accuracy. It can assess the performance of a model trained on either a balanced or imbalanced dataset, contributing to the regularization of the model's overall performance.

Mathematically expressed as

$$Balanced\ accuracy = \frac{True\ positive\ rate + \ True\ negative\ rate}{2}$$

- **ROC and AUC**

Receiver operating characteristic (ROC) is a graph-based tool used to assess the performance of classification model. In 1941, military radar receiver operating characteristics were evaluated using these metrics, referred as ROC [47]. It is plotted

between the true positive rate and the false positive rate at various threshold points. As it compared two operating characteristics TPR and FPR, hence also called a relative ROC [48]. The diagonal line in the plot divides ROC equally and the points above the diagonal curve give good classification whereas the below points represent bad results. The AUC determines the accuracy of the model [49]. For instance, an ROC curve with an AUC of 1 represents a perfect model, while an AUC of 0.5 denotes an inefficient model and AUC defines the ability of model in predicting the correct class for a given input. ROC is another most normally used tool to evaluate the performance of a classification model [50].

1.11 Thesis Organization

This thesis is organized into seven chapters as follows

Chapter 1 delivers an introduction to breast cancer and its various aspects. The necessity of implementation of breast cancer diagnosis approaches based on cancer statistics has been highlighted. The potential of conventional machine learning and deep learning approaches are discussed. The distinctive steps in the classification system including data selection, preprocessing, feature extraction, and classification are also elaborated. To elucidate the developed model, a variety performance evaluation metrics are also described for comparison analysis with the existing state-of-the-art approaches. Beside the transfer learning approach, the importance of an early-stage diagnosis of breast cancer in the CAD system using histopathological images have been presented. In addition, performance evaluation metrics are also discussed in this chapter.

Chapter 2 discusses the different cancer classification techniques proposed by academicians, researchers and scientists in the field of biomedical image processing for the development of efficient CAD systems. The chapter also elaborates the comparative analysis of existing state-of-the-art techniques for histopathological image classification through recent contributions to automatize the CAD systems for improving the diagnosis. Apart from that, we have outlined challenges in the analysis of histopathological images, research objectives, and methodologies of the proposed work.

Chapter 3 provides a comparative study of CNN based approaches used for grade classification. Automated classification system developed using a pre-trained model as feature generator and fully connected layer with SoftMax function to categorize the IDC grades. The classification of segmented image and entire histopathological image dataset are performed by employing segmentation approach and eight different pre-trained neural networks, respectively to assess the effectiveness of CAD system in differentiating IDC BC grade classes under different magnifications are discussed.

Chapter 4 determines the potential of CNN as an effective feature extractor, and the last stage in the CNN replaced with conventional machine learning RF algorithm. The proposed model performance elaborated with train-test data split ratio of 90-10, 80-20 using five different CNNs viz. VGG16, Inception_V3, ResNet152_V2, DenseNet102 and NASNetMobile with RF classifier to perform magnification independent histopathological image classification of IDC breast cancer grades.

Chapter 5 provides the performance analysis of the transfer learning-based CNN model for IDC breast cancer grade classification using six different pre-trained models including VGG16, VGG19, ResNet50, DenseNet201, NASNetLarge and MobileNet_V2 as feature extractors and in last stage six different machine learning classifiers RF, LDA, LR, KNN, DT, and SVM are deployed for this study. Further, the potential of transfer learning approach is analyzed for magnification-independent and dependent grade classification. In this regard, the chapter outlines the executions of proposed CNNs along with ML classifiers for achieving efficient classification under different magnification factors.

Chapter 6 determines the performance of different CNNs combined by an ensemble algorithm for classification purpose. At the initial stage, data augmentation is performed to increase the size of the of databiox dataset. Here, the performance of an individual CNN network base models evaluated with a fully connected layer with SoftMax activation function. This study has been performed in a comparative manner among the different data samples, and the best accuracy-produced CNN models in order to form an ‘Ensemble’ model. Here, individual CNNs are trained with different number of epochs as it leads to better accuracy. The results revealed that the proposed ensemble model with data augmentation provides the highest classification accuracy when

compared with existing state-of-the-art techniques for the ultimate goal of enhanced IDC breast cancer grade classification in Databiox dataset.

Chapter 7 summarize the work presented in the thesis and also outlined the future scope for the proposed work.

2

Literature Review

"Study the Past if You Would Define the Future"
—Confucius

In this chapter, the literature review related to breast cancer classification has been discussed. The recent advancements that have been produced by academicians, researchers and scientists in the field of medical image processing applications, particularly cancer diagnosis is presented. This study thoroughly examines traditional machine learning and deep learning methods aimed at addressing challenges in computer vision, with a specific focus on their application in biomedical contexts are discussed in detail to identify the research gaps. In this view of preceding, the present chapter is divided into two sections where the literature based on conventional machine learning, and deep learning including transfer learning techniques are explained in a systematized manner. The results attained by some state-of-the-art techniques in cancer classification have been brief in table for an easy comparison and evaluation.

2.1 Introduction

The work that has been done by researchers and scientists in the process of diagnosis of cancer with biomedical images using machine learning, deep learning algorithms, and transfer learning approaches are studied. The existing techniques and results attained in each approach are considered to identify the research gaps. The extensive review of the literature determines the developments and enhancements in CAD systems. In this connection, first considering the diverse techniques available in the field of research is the factor in making innovations, and new precise discoveries, especially in the design process of computer aided diagnosis systems. Therefore, relevant studies in the identification and diagnosis of cancer have been discussed. In the process of classification of breast cancer, feature extraction approaches are one of the keys to making better performance of the automated CAD system. Wherein conventional machine learning and deep learning are the major techniques that are employed with a variety of feature extraction-based classification approaches. Finally, summarized the better performance produced techniques in the classification of cancer using histopathological images.

2.2 Literature Based on Conventional Machine Learning

One of the important steps in the classification system is feature extraction because the extracted features influence the performance of the system. In this context, conventional feature extraction approaches used along with the machine learning algorithms are discussed here. The performance of automated system depends on the type of dataset used to train and test a model. In addition, feature extraction is a challenging task as the histopathological images consist of complexity in pattern and resolution. It leads to difficulty in classifying between normal and cancer tissues. Therefore, the researchers proposed a variety of feature extraction approaches to meet the maximum efficiency of the automated system. High dimensionality is the other factor that influences the performance of the model. The use of a large number of features leads to overfitting the model, meanwhile lesser number of features leads to underfitting the models. In this connection, feature extraction and selection of a set of feature vectors or feature maps are important in the process of the automated diagnosis system.

Some of the recent studies proposed by researchers on feature extraction including GRLM [51], LBP [52, 53], HOG [54], SIFT [55], SURF [56], completed LBP [57], LPQ [58] are employed to diminish the computational burden and time for feature-based classification systems in different biomedical applications.

L. Hussain et al. [22] discussed a paper on the most frequently diagnosed breast cancer among women and the second most common cause of deaths. It has been mentioned that breast cancer is not enough identified by radiologists because of the complex existence of micro unclassified masses of tissues. In this conjunction, CAD systems have been developed that allow radiologists to efficiently diagnose anomalies in breast tissues. In this work, they suggested a reliable machine learning classification approach, that includes a DT and SVM to distinguish cancer mammography from non-cancerous mammograms. The elliptic descriptor, texture, and mammography entropy are some recommended features for improving the classification with the aid of the above-mentioned classifiers.

R. Fakoor et al.[59] described the usage of automated CAD systems, namely machine learning, to advance and improve medical research and diagnosis. They described how feature space extraction from microarray data may be utilized for cancer discovery and cancer type classification. The major advantage of the proposed strategy over the remaining methods for malignancy identification is the ability to apply data from various varieties of cancer to dynamically shape features that support advance the prediction and identification of particular types of cancer. In this paper, the strategy is broadened to include the identification and categorization of cancerous forms using microarray data. In this area, they show that output of this method is better than that of previous approaches, offering a more systematic and generic cancer detection and diagnostic approach.

J. F. Mccarthy et al. [60] exclusively discussed the use of high-dimensional data visualization and data mining for cancer detection at the molecular level. Such an interpretation of molecular genetics may have a huge effect on the clinical environment, how to identify, diagnose, and treat cancer cases. Additionally, the medical oncologist faced that clinically valuable information from the vast amount of existing molecular data available locally is the major challenge.

N. B. Bahadure et al. [61] discussed a paper on image processing for the identification and extraction of MRI-based brain tumours using biologically influenced BWT and SVM. The primary problem, however, is the segmentation, identification, and extraction of MR images of diseased lump mass, which is a slow and laborious task carried out by clinical professionals or radiotherapists. That is the reason why the most interest is focused on resolving these limitations with the help of computer aided methods based on the barkeley wavelet transform (BWT). BWT increased the efficiency and decreased the difficulty in the segmentation process involved in the medical images to identify the type of brain tumor. With the help of experimental results conducted on the different images in the process of identification of brain tumours the proposed method is a quick one and the accuracy is more as compared to the manual identification performed by the clinical experts or radiologists. Several factors were evaluated to verify the performance of the recommended work including average, mean square error, peak signal to noise ratio, sensitivity, and specificity.

M. De Bruijne [62] addressed machine learning methods for medical image processing, ranging from identification to diagnosis. There is still some ambiguity when using training methods as a "black box" for the purposes of diagnosis and risk analysis. An adaptable learning system can work randomly in a high dimensional functional space and can be difficult to detect. It is interesting to trust that they may be effective when measurable biological change is linked to a particular illness or disease, it helps a good separability between various classes of diseases. Diagnostic decisions are mostly based on the training data in the learning model, but not on signs of illness whereas the signs of illness indicate the status of the disease. The effect of particular infection is higher in one type on gender as compared to other, it is complicated to design a learning algorithm to determine the disease and prediction of risk with that disease. Age, gender, other controlling variables are introduced carefully in the predictor to design a good predictor.

J. Ker et al. [63] introduced and explored machine learning techniques for cancer detection. With the aid of a health information system, the effectiveness of a certain machine learning procedure that is based on image processing tasks in the field of investigative imaging is improved. A review on the use of CNNs in the domain of health

care and medical image processing with the use of machine learning is performed. The advantages of machine learning in the context of medical data are that they can be found algorithmically without the need for extensive hierarchical connections within the data. It covers prime areas of research and applications in machine learning for localization, identification, registration, segmentation, and classification of medical images.

A. Depeursinge et al. [64] conferred a work on three dimensional biomedical dense textures in the form of advanced description is the key to achieve large scale and high extent screening of imaging data. With higher resolution, three-dimensional data is more readily available in testing and clinical settings which enables numerous image processing algorithm trends in the context of biomedical texture analysis. It has been shown for larger structures that only non-separable multi-scale directional convolutionary approaches enable unbiased three-dimensional biomedical textures to be modelled. In the future, these models are predictable to better exploit the wealth of three-dimensional texture scrutiny of biomedical images with improve in the resolution isotropic image protocols in medical practice and science.

K. Rawal and G. Sethi [65] discussed study on medical image processing approaches used to detect the abdominal diseases. By interpreting the medical history of many intestinal organs, including the kidney, liver, pancreas, gall bladder, and grumpiness, variety of abnormalities has been identified. Usually, tumours, stones, cysts, are the types of abdominal disorders and that are detected by their grayscale characteristics. They have also explored diagnostic techniques for identifying various diseases, including liver tumours and breast cancer, and their effects while using CT scan imaging.

D. Komura and S. Ishikawa [66] illustrated the enormous demand for CAD methodologies that use machine learning to diagnose a variety of diseases utilizing a limitless collection of accumulated digital pathological images. Usually, digital pathological images are used to address several disorders. This paper looked at digital pathological image analysis utilizing machine learning techniques, however there is still a huge amount of space for improving the datasets and machine learning models. Most problems are likely to be resolved if a large quantity of Whole Slide Image (WSI) data is available. To attain huge amount of data, one of the methods is to collect WSIs from different organizations to jointly interpret them along with identical tags or

representations and make it public, so that more digital pathological images to be analysed. To resolve this a lot of feature selection methods have been implemented to compute pathological images for solving high dimensionality problems.

T. Jagadesh and R. J. Rani [67] explored how image processing is utilized by radiologists or clinicians to identify cancers. Usually, noise decays the clarity of medical images and distorting them. Edge destruction, structural data, distorting thresholds, etc. are all factors in this degradation. Denoising would seem to be a very important process, and can be removed without altering important image features. Principal component analysis (PCA) is an assembly technique used to categorize new data that is less detailed than the original image while retaining a significant amount of research information. The wavelet PCA output is extended and enhanced by integrating the relationship of intra and inter-scale constraints between the wavelet coefficients. In their work on ultrasound images, a multilevel description is performed by using wavelet transform and its inverse form. Vector feature is extracted for all pixel values with help of wavelet transform. In PCA method, a soft thresholding method and enhanced soft thresholding are used to eliminate the noise.

S. Sharma et al. [68] proposed a study that lung cancer is the fatal category of cancer and reported excess mortality than colon, prostate, and breast cancers all together. In this article, the radiomic approach proposed, which is not only the cost-effective method it also provides a non-invasive provision to detect the cancers. The additional feature of this method is that it is easy to execute regardless of the cancer type, with help of various types of databases. The conversant lung cancer research is possible with the help of proposed radiomic approach. Therefore, in clinical practice, radiomics has a great effect, and is a strong alternative to histopathological methodology, and offers many possibilities for improving decision support in cancer detection. Low-cost identification is the importance of this method since this technique can recognize the any type of cancer. Although, due to the need for multiple trials to create data sets of enough size and effective statistical description, reproducibility, and sharing of big data are key challenges. Therefore, certain strategies need to be formulated in a manner that can solve such challenges.

Z. Othman and S. Saleem [69] presented a paper with the goal of assisting medical professionals in their diagnosis-related decisions. Thus, the doctor has an influence on the crucial final decision. From this work, it can be deduced in a number of ways that the obvious palm image can only be captured using the superb image acquisition method that has been selected. Over the last decades, the computer aided design with help of biomedical imaging systems the critical decision becomes easy. This plays a key role while taking clinical decisions in the range of biomedical imaging investigation and diagnosis. This article suggests a method for healthcare choice facility to identify sicknesses based on the hue of the hand using palmistry.

S. Sharma and R. Mehra [70] discussed that a well-adjusted training dataset is the main condition for optimizing a convolutional deep neural network. The prime problem is the severe alteration and deficiency that creates it challenging to train the system from the raw or scratched information that influences medical findings. In addition to the advantage of faster preparation, it is detected that the transfer learning methodology has a huge impact to compete with insufficient datasets.

S. Sharma and R. Mehra [71] discussed a study on current difficulties for machine learning applications where pooling is a component of the deep neural network. The system computing amount and training period, which are equally important when developing a model are considerably reduced by pooling approaches. The advantages and disadvantages of the suggested pooling approaches are explained and addressed in this study, which validates several pooling strategies based on the accessibility of several datasets.

A. Alarabeyyat and M. Alhanahnah [72] discussed a paper on breast cancer detection using machine learning technique. Breast cancer is quite common around the world among women. Intelligent to notice cancer during its early stages, indeed, improve the chance to save lives. The radiologists are able to predict whether or not the mammogram images have cancer, but some fifteen percent of them could be missing. In the proposed article, a new method of accurate breast cancer detection discussed is discussed. The initial level of this method involves image segmentation to set up the mammography image for collecting features, while another method employed machine learning approach using LR and back propagation neural networks.

L. Zhang et al. [73] proposed a novel approach for prostate cancer identification and lesion partition in MRI images with GrowCut and Zernik functionality through integrated machine learning approaches. To achieve this, they combine a variety of machine learning approaches including KNN, MLP, and SVM in the assembly learning framework in conjunction with GrowCut algorithms. It has been observed that integration of multiple ML techniques improved the system performance by around 20 percent and also strengthened the other evaluation metrics such as recall, accuracy, and error of the system.

2.3 Literature Based on Deep Learning

Deep learning is the subset of machine learning, where the features are extracted by deeper feature extraction approaches. In deep learning extraction of features carried in different manner as compared to traditional machine learning. Sophisticated hierarchical approaches used to capture the features from the data in contrast to the traditional hand-crafted feature extraction approaches used in the machine learning classification system. Moreover, the training of deep networks is a time-consuming process, and required high graphical processing units to handle huge amount of data; however, it improves the system performance significantly. In this section, the implementation of deep learning approaches is discussed which includes artificial neural networks, and convolutional neural networks for efficient classification, while the transfer learning approaches are also discussed to further reduce the computational burden.

C. Cao et al. [74] discussed a survey on variety of deep learning algorithms for biomedical applications pertaining to feature extraction, and classification. Wherein interpreting the medical images and identification of discriminative features is the key task to perform classification, segmentation of region of interest and detection of variety of tissue abnormalities.

X. Zhou et al. [75] outlines the artificial neural network techniques particularly used for Breast Histopathological Image Analysis (BHIA) including breast cancer tissue segmentation and classification. In this context, ANNs interpret histopathological images by using two approaches as classical and convolutional networks as deep. Furthermore, the deep neural network summarized into the different groups based on

the applied datasets in the models. The most commonly used models in this study are the classical MLP and probabilistic neural network for performing feature extraction and classification of breast cancer histopathology images through morphological and textural features. This helps in enabling the detection of breast cancer at early stage and make better therapy possible.

A. M. Kist and M. Döllinger [76] proposed a tiny U-Net architecture for semantic image segmentation which helps to reduce the computational burden involved in the analysis of biomedical datasets. With 95% of segmentation performance of the baseline model, they reduced 99.8% of computational burden. However, semantic segmentation is used in encoder decoder-based U-Net architecture and provide high performance, which requires high end CPUs and GPUs that lead to high computational cost.

S. S. Koshy et al. [77] presented a survey on deep learning approaches employed for histopathological images. In this study, outlined the state-of-the-art techniques used and the importance of high computational GPU requirement when the deeper neural network layers trained with variety of datasets. In addition, addressed the computational complexities of huge of data and overfitting problems against the conventional machine learning feature extractors and classifiers.

H. Wieslander et al. [78] presented a paper on cancer detection using deep CNN networks by detecting cellular changes due to malignancy. Cervical cancer detection automated system developed by using two datasets pertaining to oral cell and cervical cell information. When used only the oral cell dataset for the diagnosis of patient, then the most of them looked normal, but all cells taken from patient were malignant. In this connection, assessed to check the malignancy of the patient with two different VGG and ResNet network architectures trained with two different datasets. The performance of the model evaluated and concluded that ResNet performs better than the VGG.

Y. Liu et al. [79] discussed the difficult work involves designing a process to offer real-world decision support in health information as a result of faith level and level of exactness. In the proposed deep learning model, the performance of the disease diagnosis evaluated along with the second diagnosis process with medical text data of human physiological parameters including urine parameters, hematologic parameters and biochemical reports. Different data quantification techniques also effect the performance of the proposed model, for intelligent model diagnosis. The development

of the accurate and strong secondary diagnosis system with help of medical data helps a lot and reduces the workload of medical experts. Future research concentrates on the auxiliary diagnosis system for detection of various types of diseases in human beings.

X. Ren et al. [80] presented a paper on task decomposition and synchronization procedures for biomedical image semantic segmentation. Semantic segmentation is crucial for the investigation of biomedical images to identify and detect the abnormalities in the tissue. In this article, they propose that the single task be divided into three smaller tasks, the first of which is the pixel-by-pixel separation of the image, the second of which is the class mark prediction of the matters inside the image, and the third of which is the object classification to which the item belongs. These primarily three subtasks demonstrate that the task-task framework ensemble facilitates collaboration even though they are trained to optimize diverse intuitive levels of their loss functions. Additionally, proposed a new sync-regularization method to compensate the discrepancy between classification performance tasks and the output of pixel-level segmentation.

S. Pang et al. [81] proposed a novel CNN for the biomedical images and universal non-biomedical image classifiers to identify more compressed biomedical image characteristics and detect the small difference among similar images of various forms of diseases. In this approach, combining shallow layer features and deeper layer features of a CNN model to discovered that though shallow layers mainly carry extraction of local features and deeper layers that carry high-level semantic information. It helps to discriminate and detect the diverse diseases.

S. Sharma and R. Mehra [82] presented a deep CNN model for the categorization of breast cancer tissues and BreakHis dataset utilized to train the model. The developed system consists the pooling layer in the ending of the convolution network only, which reduced the loss of information. The model performance increased while using the data augmentation procedure. In addition, the proposed model achieved robust performance with characteristic feature of the model is that independency among the different magnification levels of the dataset. Using cross-entropy and updating weights by Adam optimizer, it minimizes the loss and attains an accuracy of 85.3 percent on average.

M. Toğaçar et al. [83] proposed a novel CNN-based model called BrainMRNet and used for the identification and detection of brain tumours. The importance of this model is that it consists attention module to select the important areas of the image. The important areas of the images transferred to convolutional layers to extract the features, and these features are saved in the form of array structure by using hyper column technique in the last layer. The proposed model achieved better performance as compared to the Google Net, Alex Net, VGG16 by selecting efficient features from the array structure.

S. Makaju et al. [84] discussed the limitations and drawbacks in the detection of cancer with the state-of-the-art techniques and proposed a model for lung cancer detection using CT scan images. The important goal of this study is to evaluate the best approach with biomedical image analysis using an artificial intelligence method. The proposed model consists watershed approach for segmentation task and SVM machine learning classifier for the classification task and achieved an accuracy 92% better than the existing techniques. Additionally, suggested that to enhance the performance of the proposed technique by a suitable pre-processing approach.

A. Osmanović et al. [85] proposed a feedforward backpropagation neural network with single hidden layer and TANSIG transfer function with 20 neurons. The model performance evaluated using the trial-and-error method by updating three parameters: number of layers (either 1,2,3,4, or 5 layers), number of neurons (either 10,20,30,40, and 50), and transfer functions (training algorithms TANSIG, Logsig, and pure line). The breast cancer classification accuracy of 98.9%, and 99% achieved with train and test datasets, respectively. For the implementation of this model, UCI machine learning respiratory, breast cancer data set (Wisconsin) with 699 samples are used.

M. Z. Alom et al. [86] discussed the classification model performance by using dual diverse datasets including the BreakHis dataset with 7909 samples and the breast cancer classification challenge 2015 dataset. Additionally, the samples in the dataset are increased using the data augmentation process including rotating, cropping, shearing, as well as translation. The proposed model design strategy follows that Inception_V3, Residual Net and Recurrent convolutional neural network and named as an Inception Recurrent Residual Convolutional Neural Network (IRRCNN). The

proposed model trained for 50 epochs to classify binary as well as multi-class breast cancer identification.

S. Sharma et al. [87] proposed a CNN model for magnification independent classification using BreakHis dataset. The proposed CNN architecture followed a sophisticated training procedure to learn the discernible features in contrast to the different magnification factors of histopathological images. In this context, data augmentation, pooling strategy, and optimization techniques were used for observing cancer types from magnified histopathological images. The proposed system achieved an accuracy of 80.47% with 40x magnification of histopathological images, can improve the accuracy by Adam, AdaGradient, and gradient descent optimizers are investigated, and achieved the best results with a learning rate of 0.0001.

S. Sharma and R. Mehra [88] discuss a comparative study on histopathological breast cancer multi classification by using pre-trained CNN as a baseline model and feature extractor with machine learning classifier in contrast to handcrafted feature extraction based conventional classification approach. The proposed automated system utilized VGG19, ResNet50, and VGG16 models as baseline models and feature generators to determine the potential of transfer learning, and achieved the best accuracy with VGG16 fine-tuned deep neural network with linear SVM classifier among the other employed approaches.

D. A. Ragab et al. [89] discussed a deep CNN based framework for breast cancer categorization. The proposed CAD method classifies breast cancer lesions in mammograms by using deep learning techniques for feature extraction, and categorization. In this context, four different forms of experiments are planned (1) pre-trained deep CNN as a baseline model for classification task. (2) pre-trained deep CNN as a feature generator, and SVM for classification task. (3) pre-trained deep CNN and, deep feature fusions stages combined for improved feature extraction, SVM classifier for categorization task. (4) pre-trained deep CNN and, deep feature fusions stages are used for extracting features, and PCA reduce the redundancies in features and select important features for further processing, SVM classifier for classification task. The deep feature-based approaches produced better results in contrast to other methods.

G. Murtaza et al. [90] addressed several issues and challenges while developing automated CAD techniques for predicting progression in breast cancer using several medical imaging modalities. In this review paper, the issues and challenges addressed are types of datasets, size of the dataset, types of imaging modalities, pre-processing steps, and categories, types of deep convolution neural networks are the various parameters used to measure the performance of the model. To reduce inconsistencies in breast cancer images, some studies used pre-processing techniques includes augmentation, and image normalization.

S. Boumaraf et al. [91] discussed a transfer learning strategy for automatic CAD system to perform binary and multi-class, magnification independent and dependent classification breast cancer. In the proposed deep model, ResNet-18 network as a feature generator which has been trained using ImageNet dataset and refined on histopathological images using block-wise fine-tuning strategy so that the last blocks are trained to target histopathological image dataset. The proposed model produced better results, in addition, avoids the model overfitting problem.

S. Saxena et al. [92] addressed the issues of class imbalance in the classification of breast cancer automated CAD system with help of histopathology. The proposed hybrid machine learning model, consists ResNet50 for extracting features and the kernelized weighted learning, more weight is allotted to the minority class in contrast a lower weight is allotted to the majority class. BreakHis and BisQue datasets are used to train and test the proposed model, and the obtained results are significantly improved with the state-of-the-art technique for both minority and majority class cases.

Y.D. Zhang et al. [93] introduced a novel hybrid deep learning approach for the breast cancer lesion identification by using a graph convolution network (GCN) and CNN. Initially, 8-layer CNN developed with two improvements including dropout (DO) and batch normalization (BN) techniques, later employed rank-based stochastic pooling (RSP) in place of max-pooling. The mix of CNN, BN, DO, and RSP called BDR-CNN. In addition, a two-layer GCN mixed with BDR-CNN to form a hybrid model. The proposed hybrid model achieved an accuracy of 96.10±1.60% with the breast mini- (mammographic image analysis society) MIAS dataset.

A. Saber et al. [94] discussed breast cancer classification using a transfer learning approach with help of the MIAS database. In the proposed work, six pre-trained CNNs

including Inception-V2, VGG16, VGG19, ResNet 50, Inception -V3, and Resnet models used in the experimental study. Different assessment metrics are utilized to measure the performance of the model with 80-20 data splitting and 10-fold cross-validation strategy. The proposed models attained an accuracy of 98.96%.

I. Hirra et al. [95] proposed a patch based Deep Belief Network (DBN) model used to classify breast cancer. The proposed model extracts the features from the patch images and logistic regression classifier used in the categorization stage. The model attained accuracy of 86% when trained and tested with the whole slide histopathological images.

X. Y. Liew et al. [96] presented a deep learning based multiclass histopathological image classification. In this work, DensNet201 as a feature extractor and eXtreme boosting algorithm as a classifier. The proposed model performance assessed and achieved accuracy of 97% with 5-fold cross-validation for both binary and multiclass problems.

S. Singh et al. [97] proposed a functional link artificial neural network (FLANN) to classify breast cancer of Wisconsin diagnostic breast cancer, as WDBC dataset and Wisconsin breast cancer dataset, with 569 and 699 sample images, respectively. The proposed model avoids the overfitting problem and achieved maximum accuracy of 99.41%.

S. Sharma and S. Kumar [98] discussed the potential of the Xception deep learning model for the effective feature extraction and SVM classifier with different 'radial basis function' kernel used to classify breast cancer images. The proposed model utilized BreakHis dataset utilized with magnification levels includes 40x, 100x, 200x, and 400x. In addition, the performance compared with the RF classifier with 6000 decision trees. However, Xception with SVM classifier achieved the best and consistent results at magnification levels 40x, 100x ,200x and 400x with 96.25%, 96.25%, 95.74% and 94.11% accuracies, respectively.

F. A. Spanhol et al. [99] introduced a histopathological image dataset called, BreakHis and make it available to research community. The produced dataset consists of 7909 images taken from 82 patients, which consists of benign and malignancy cancer types and each type further classified into four subtypes. In addition, automated CAD

system developed to classify the cancer types of the proposed dataset and achieved an accuracy in the range of 80% to 85%.

2.4 Literature Based on the Databiox Dataset

The following are the literature survey papers related to Databiox dataset.

H. Bolhasani et al. [14] in the year 2020 published a report on a grade of histopathological images from a dataset called Databiox. The dataset includes 922 images collected from 124 patients of distinct grades of with different magnification levels 4x, 10x, 20x and 40x.

E. Kumaraswamy et al. [100] reported to classify IDC breast cancer grades using a transfer learning approach. In order to categorize the target IDC breast cancer grades of the Databiox dataset, multiple pre-trained networks that are utilized as feature extractors and final stage in the pre-trained networks substituted with various machine learning RF classifier. The performance evaluation metrics AUC used to measure the area under the ROC curve for the following grades: DensNet201- 98% for grade class 0, DensNet201-75% for grade class 1, and Inception V3 - 69% for grade class 2.

P. H. Zavareh et al. [101] presented a new BCNet model to classify histopathological images of the IDC breast cancer grade classification. The proposed method uses a VGG16 pre-trained model and output section is comprised of a Global average pooling layer with 512 nodes, and two fully connected layers with 1024 and 3 nodes respectively with SoftMax function in the classification layer. Finally, the proposed BCNet attained a validation accuracy of 88% and an evaluation accuracy of 72%.

S. Talpur et al. [102] proposed sequential Convolutional Neural Network Two-Dimensional (CNN2D) model to classify IDC breast cancer grades. The proposed sequential CNN yields an accuracy of 92.81%. and individual grade accuracies are 92.00% for G-1, 90.41% for G-2, and 96.55% for G3. The sequence of convolutional layers used to extract the features of histology images and fully connected layers used to detect IDC breast cancer grades.

R. Sujatha et al. [103] discussed a deep learning-based transfer learning approach to classify IDC breast cancer grade classification. In the proposed work, five different pre-trained networks including VGG16, VGG19, InceptionReNetV2, DenseNet121, and DenseNet201 utilized and last layer neurons adjusted to the target problem. The

DenseNet121 model with SoftMax function with 50 epochs produces the best accuracy of 92.64% as compared to the existing techniques.

E. Kumaraswamy et al. [104] presented an ensemble of convolutional neural networks used to classify the IDC breast cancer grades of databiox dataset. The proposed approach utilized the individual CNN models and evaluated the performance of the models. Later, the best performance produced models combined by using ensemble approach to get the best results from the individual models and achieved the highest accuracy of 94% classification accuracy compared to the state-of-the-art methods.

V. Mudeng et al. [105] proposed a novel method for classification of IDC grades by using domain and histopathologic transformation based multistage transfer learning approach. For this study, different pre-trained networks considered as feature generator, which is initially trained with ImageNet dataset. Using transfer learning approach, the pre-trained model trained and tested for BreakHis dataset so that the model transformed from natural ImageNet domain to histopathologic domain. Later, the same model used for categorization of IDC grades by deploying transfer learning technique. In this context, multistage transfer learning using domain and histopathologic transformation work produced significant result at magnification level.

The results and comparison among the recent state-of-the-art techniques deployed for histopathological image classification are tabulated in Table 2.1. Apart from that, a comparison has been made for IDC breast cancer grade categorization of histology images in Databiox dataset and tabulated in Table 2.2.

Table 2.1 Comparison among classification techniques using histopathological images

Author (Year) [Ref.]	**Dataset**	**Feature Extractor**	**Classifier**	**Accuracy (%)**
F. A. Spanhol et al. (2016) [106]	BreakHis	CNN	Fully-connected layer with SoftMax activation	90.0 ± 6.7
R. Yan et al. (2020) [107]	Bioimaging2015 dataset	Google's Inception-V3	Fine tuning	91.3
M. Gour et al. (2022) [108]	BreakHis	Residual learning-based 152-layered CNN	Fully connected layer	84.3
Y. Sun et al. (2018) [109]	BreakHis	CNN	SVM	83.3
M. Z. Alom et al. (2019) [86]	BreakHis and BC classification challenge 2015	Inception Recurrent Residual CNN	Dense layer	97.5 ± 0.8
S. Khan et al. (2019) [110]	Dataset collected from LRH hospital Peshawar, Pakistan and BreakHis	Transfer learning	Fully-connected layer with SoftMax activation	97.5
G. Murtaza et al. (2020) [111]	BreakHis	Deep learning	KNN classifier	95.4
I. Hirra et al. (2021) [95]	Histopathology images	Deep CNN	Logistic regression	86.0
Y. Zou et al. (2022) [112]	BreakHis	ResNet + channel attention + Pooling	Dense layer	85.0
N. Ahmad et al. (2022) [113]	BreakHis	Efficient-Net	SVM	93.6 to 96.9
S. Diao et al. (2023) [114]	BCSS2021 dataset	Inception_v3 + Inception Reduction Module block	Fully-connected layer with SoftMax	93.7 ± 0.7
K. Datta Gupta et al. (2023) [115]	The Malignant Lymphoma dataset	Reduced Fire Net	Dense layer	96.8

Table 2.2 Comparison of the state-of-the-art techniques of IDC BC grade classification of the histopathological images of the Databiox dataset

Author (Year) [Ref.]	**Approach**	**Classifier**	**Magnification Factors**	**Accuracy (%)**
E. Kumaraswamy et al. (2021) [100]	Transfer learning approach - VGG16, Inception_V3, ResNet152V2, DenseNet201, and NASNetMobile	RF classifier	---	72.0
P. H. Zavareh et al. (2021) [101]	VGG16	Dense layer	---	72.0
S. Talpur et al. (2022) [102]	A sequential CNN	Dense layer	---	92.8
R. Sujatha et al. (2022) [103]	Transfer learning approaches	Dense layer	---	92.6
E. Kumaraswamy et al. (2023) [104]	Ensemble model (EfficientNetV2L+ ResNet152V2 + DensNet201)	Dense layer	---	94.0
E. Kumaraswamy et al. (2023) [116]	Xception model	FC with SoftMax function	4x 10x 20x 40x	71.0 91.0 91.0 91.0
E. Kumaraswamy et al. (2023) [117]	Transfer learning	ML classifiers	--- 4x 10x 20x 40x	75.0 79.0 100.0 100.0 91.0
V. Mudeng et al. (2023) [105]	Multistage Transfer learning	FC with SoftMax function	--- 4x 10x 20x 40x	89.6± 0.8 97.6 ±1.0 95.6 ±1.6 95.8 ±1.9 90.6 ±1.3

2.5 Challenges in the Analysis of Histopathological Images

The following research gaps are identified during the literature review,

1) Many theoretical and practical problems need to be directed to unlock the full potential of machine learning techniques including how to develop efficient models

on available small data, how to pre-processing the data, how and when to make the best use of the various features of an image like structure, shape, and information of medical images in the design, how to interpret outcomes, and how to utilize these results in the field of medicine.

2) Due to data high dimensionality, there is redundancy in features and sometimes irrelevant to the task under consideration. These irrelevant features might affect the performance of the model in an adverse manner, called overfitting. Another view is that machine learning techniques are data-hungry, it is desirable to train a model with enough amount of data, to eliminate the underfitting problem. The success of any classification algorithm particularly depends on its feature extraction ability.
3) There is a problem in representing the connection between feature variables and target variables. The classifier estimates to learn that relationship. So, it is quite necessary to design a robust classifier that can efficiently classify the cancer types and able to differentiate between the benign and malignant types. Although various existing machine learning models' performance is limited to type of dataset and it is measured by some evaluation parameters like accuracy and precision. There is a scope for the improvement of these parameters by proposing a new machine learning model.
4) The early and accurate detection of a disease is completely depending on the way of diagnosis, a computer-aided technique used, and the type of machine learning model adopted.

2.6 Research Objectives

Research objectives are

1. To segment the region of interest from biomedical images.
2. To extract the set of feature vectors from the segmented region of interest.
3. To develop an efficient machine learning-based model for biomedical image classification.
4. To evaluate the performance of the developed model and its comparison with the existing state-of-the-art models.

2.7 Research Methodology

The step-wise procedure is followed to fulfil the research objectives discussed below and shown in Figure 2.1.

1. **Determination of issues and challenges in the detection of cancer:** This process enhances the understanding of techniques and methodologies that have been used earlier in the domain of image data classification. The challenges which have been identified while solving these problems are scarcity of data, under-fitting of the network, over-fitting of the network, and the computational complexity of the network, etc.
2. **Data collection and pre-processing:** Most of the datasets are incomplete and generally not reliable for evaluation purposes. Collection of various datasets and pre-processing tasks are performed on raw data to make it suitable for the machine learning model implementation.

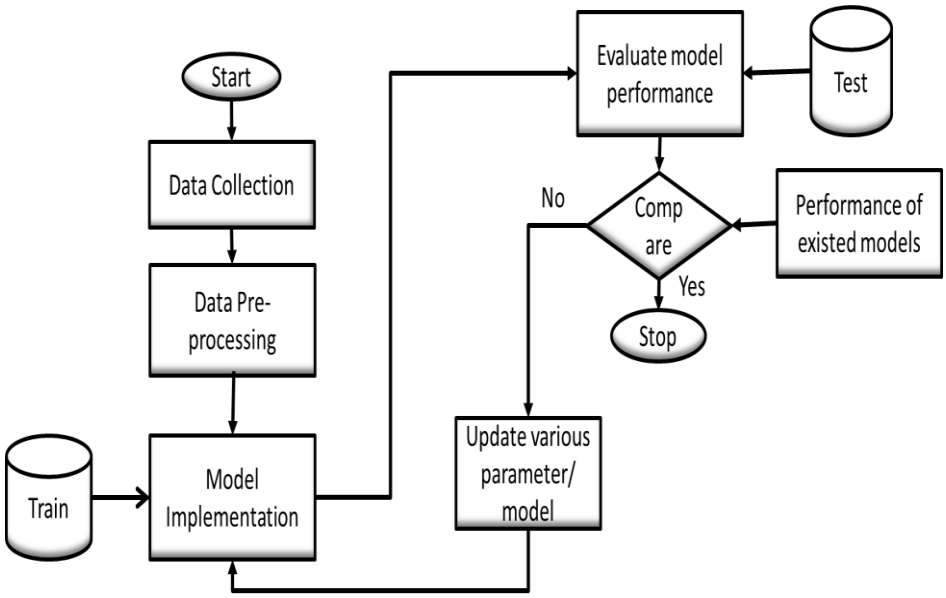


Figure 2.1 Flow chart illustrating the proposed work to address the research gaps

3. **Development of machine learning-based classifier:** A machine learning method will be implemented to categorize the cancer based on the efficient extraction of features from the biomedical image. The proposed model will be trained and tested using the Python framework (TensorFlow, Keras) and Python libraries (NumPy, matplotlib, sci-kit learn, etc.).

4. **Training and testing:** Training and testing are the two important factors in validating the machine learning model. This can be obtained by partitioning the collected data into two sets in which the first set utilized for training and the second set will be utilized to test the model.
5. **Performance evaluation and data validation:** The performance of the proposed system will be assessed through the parameters like precision, accuracy, recall, and AUC. A comparison between the proposed system and existing models will also be performed to determine the state-of-the-art technique.

2.8 Conclusion

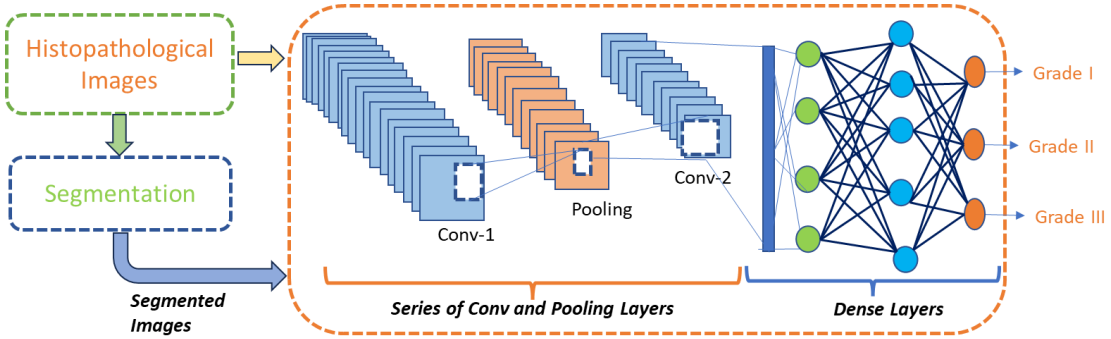
The literature survey explores both conventional machine learning and deep learning approaches, including transfer learning, for the analysis of histopathological images. Additionally, we have conducted a comprehensive review of the literature on a specific dataset named Databiox. This summary captures the diverse methods used in classifying histopathological images. From this literature, we have identified the challenges encountered in analysing these images, especially in the staging of cancer. To tackle these challenges, research objectives have been framed, and to fulfil these objectives, a broad research methodology has been outlined for the execution of the proposed research work.

3

Revealing Insights in Breast Cancer Grade Classification: A Comparative Study of CNN-Based Approaches with and without Segmentation

"Creativity is just connecting things"

– Steve Jobs



In this work paper, we have employed pre-trained neural networks to assess their effectiveness in breast cancer grading across different magnifications. we leverage the capabilities of eight state-of-the-art, pre-trained deep CNN models, namely MobileNet, NASNetMobile, Inceptionv3, InceptionResnetv2, VGG16, VGG19, ResNet50, and Xception, for the purpose of grade classification in breast cancer. Among them, a fine-tuned pre-trained Xception model excelled, achieving consistent accuracy: 71% at 4X and an impressive 91% at 10X, 20X, and 40X magnification levels without segmentation. Challenges like over-segmentation, under-segmentation, and color range selection were identified due to staining and tissue variations.

3.1 Introduction

Breast cancer remains a critical global health challenge, necessitating accurate and timely diagnostic methods for effective treatment [118]. In recent years, the fusion of cutting-edge machine learning techniques, such as CNNs with the nuanced application of segmentation has ushered in a new era of breast cancer classification using histopathological images[119, 120]. This convergence offers the potential to not only enhance diagnostic accuracy but also unravel previously hidden insights within these complex tissue samples. Histopathological analysis, traditionally conducted by skilled pathologists, provides a detailed understanding of tissue structures. However, the integration of machine learning, particularly CNNs, has demonstrated the capacity to revolutionize this process. By automatically learning intricate patterns and features from histopathological images, CNNs present an opportunity for objective, standardized, and scalable diagnostic solutions.

This study focuses on harnessing the groundbreaking potential of merging digital histopathology images with advanced segmentation techniques, alongside state-of-the-art machine learning and deep learning methods, to enhance breast cancer diagnosis. Segmentation, the process of isolating regions of interest within an image, acts as a vital precursor to classification for many computer vision applications. It enhances the accuracy and relevance of feature extraction, enabling CNNs to focus on specific components, enriching the interpretability and applicability of the classification process. Machine learning algorithms, such as SVM, RF, and DT, have proved their prowess in classifying histopathological images [119]. However, the recent surge in deep learning, particularly CNNs, has pushed the boundaries of image analysis, enabling automated feature extraction and hierarchical representation learning.

The need for this study is underscored by several factors. First, breast cancer grade identification remains a complex and challenging task, demanding precise recognition and characterization of diverse tissue components. Second, the integration of segmentation techniques and advanced machine learning and deep learning models may lead to more accurate and efficient cancer grade detection, aiding pathologists in clinical decision-making. Third, the vast amount of digital histopathological data available provides an opportunity to harness the power of data-driven algorithms for improved diagnosis. While significant strides have been made in this field, there are

notable research gaps and areas of novelty. Existing literature showcases the effectiveness of machine learning and deep learning techniques in diagnosing breast cancer through the analysis of histopathological images. However, limited attention has been given to the seamless integration of segmentation as a preprocessing step, which could significantly enhance feature extraction and classification. Furthermore, the challenge of training robust models with smaller annotated datasets remains an active area of research.

This study aims to bridge these gaps by presenting a comprehensive analysis of the impact of segmentation on breast cancer diagnosis. Additionally, it explores inventive approaches to enhance the accuracy and robustness of deep learning models in making predictions. Moreover, it embarks on a comparative exploration of breast cancer classification in histopathological images. We explore the power of CNN-based approaches, assessing their performance both with and without segmentation enhancements. By unravelling the impact of segmentation on CNN-based classification, we aim to uncover a deeper understanding of its potential benefits and limitations. Through this comprehensive investigation, we intend to provide valuable insights into the synergy of CNNs, segmentation, and histopathological images for improved breast cancer classification. Ultimately, this research contributes to a broader understanding of how machine learning can reshape the landscape of medical image analysis, opening new avenues for precision medicine and enhanced patient care.

3.2 Methodology

In this study, we leverage the capabilities of eight state-of-the-art, pre-trained deep CNN models, namely MobileNet [121], NASNetMobile [122], Inceptionv3 [123], InceptionResnetv2 [124], VGG16, VGG19, ResNet50, and Xception [125], for the purpose of grade classification in breast cancer. The selection of these models is driven by their exceptional track record in accurately classifying breast cancer across various publicly available datasets, including BreakHis [99], BACH [126], and BRACS2021 [127]. In this research, we employ a unique approach wherein all the chosen pre-trained networks serve as feature generators. We extract feature maps from various layers preceding the final fully-connected layer in each network. These feature maps are then utilized as feature vectors for training a novel classifier tailored specifically for breast cancer grade classification. To make the ultimate decision, we employ a SoftMax layer

to compute the final probabilities for the three grade classes. It is important to note that our methodology deviates from that used by previous studies, as we refrain from performing any segmentation prior to applying the pre-trained CNN models. Instead, our novel approach focuses on investigating the influence of color segmentation schemes on the performance of these pre-trained networks in the context of breast cancer grade classification

3.2.1. Segmentation Scheme

Image segmentation, a fundamental computer vision technique, plays a crucial role in breaking down an image into distinct regions based on specific attributes, such as color, texture, or intensity [128]. One of the most prevalent methods for image segmentation involves converting the image from the Red, Green, Blue (RGB) color space to the Hue, Saturation, Value (HSV) color space and then generating a mask to identify particular regions of interest. In present research, the following approach has been adopted for image segmentation.

- **RGB to HSV Conversion:** RGB serves as a widely used color representation where each pixel is characterized by its intensity in the three-color channels. In contrast, HSV provides an alternative color representation that separates color information from brightness information, making it well-suited for certain image processing tasks. In this context, H represents the color itself, allowing for the isolation of specific colors or color ranges. S represents the intensity of the color, distinguishing between vivid and pale colors, while V signifies the brightness of the color.
- **Mask Creation:** After converting the image to HSV, a mask is created to segment specific regions or objects of interest. Common techniques for mask creation include thresholding, color range selection, region growing, blob detection, and edge detection. In this case, the color range selection technique is employed to generate masks for segmenting the region of interest. The primary reason behind selecting the color range selection method is its suitability for histopathological images. Firstly, it offers flexibility and adaptability to different staining techniques (e.g., H&E staining, immunohistochemistry (IHC), or other staining methods) and tissue types. The color range thresholds can be fine-tuned to match the staining's specific characteristics. Secondly, by setting

thresholds in the color space, regions of interest can be separated from the background and other irrelevant structures in the image. This aids in the identification and segmentation of specific tissue types, cells, or abnormalities within the histopathological sample. Thirdly, this technique allows for a level of manual control and adjustment by modifying the color range parameters. Furthermore, this approach ensures consistent results when dealing with images from similar staining protocols and sources, a critical aspect in research and clinical settings.

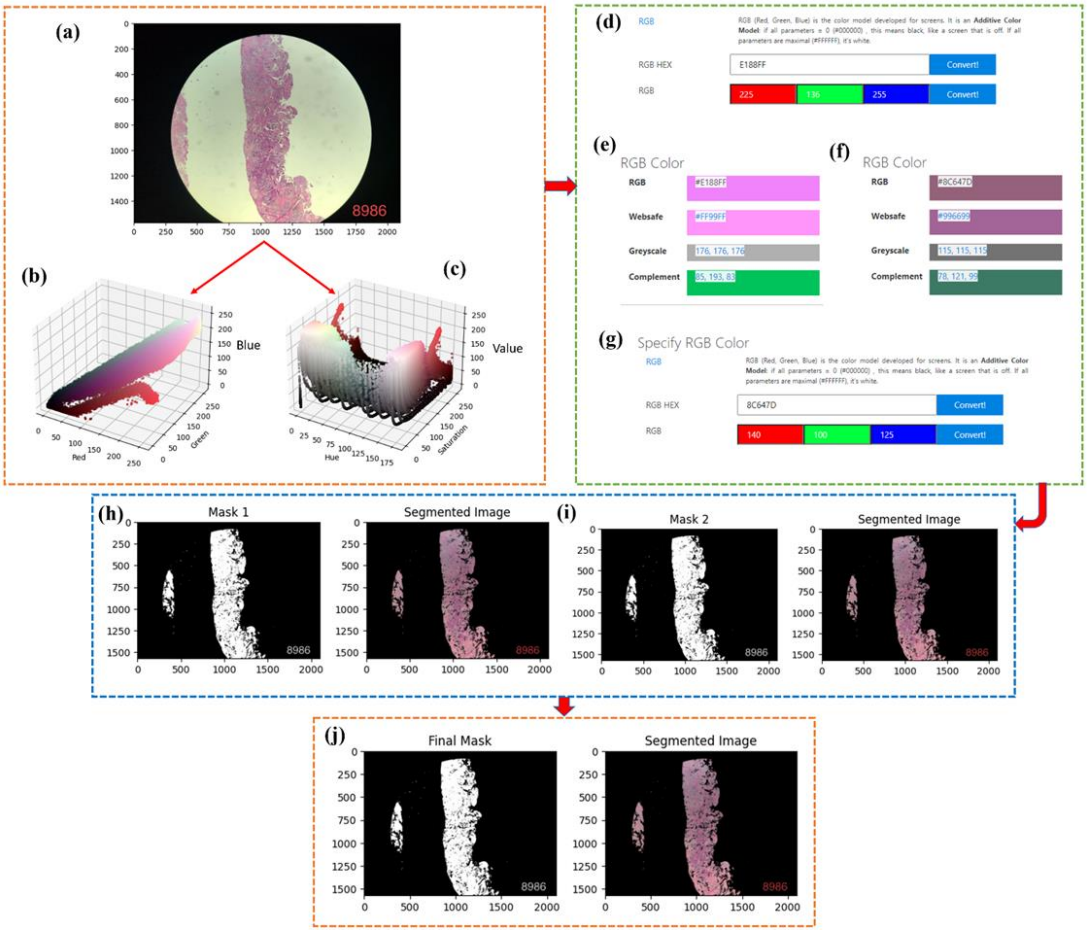


Figure 3.1. Segmentation Scheme follows: (a) histopathological image, (b) and (c) 3D color space of RGB and HSV, (d) RGB HEX value to intensity level, (e) and (f) RGB and its web safe, complement of two colors produced grayscale ranges, (g) Specify the RGB intensity values, (h) mask 1 and segmented image, (i) mask 2 and segmented image, (j) final mask and segmented image.

To illustrate this segmentation approach, Figure 3.1 depicts the process, where original digital histopathological images are initially converted from RGB to HSV space.

Subsequently, to segment the region of interest from the background, two masks are generated. Since the Databiox dataset is based on the H&E staining protocol, color shades in the range of purple, pink, and blue are of particular interest. For this reason, RGB color ranges are selected using a web application available at https://toolstud.io/color/rgb.php to define lower and upper bounds for mask 1 and mask 2. The resulting mask1 and mask2, generated through the HSV-based segmentation scheme, are employed to isolate and extract tissue from the background in the original histopathological image. These segmented regions are then used to train machine learning models for cancer grade detection in previously unseen histopathological images.

3.2.2. Data Augmentation

The Databiox dataset consists of 160, 153, 155, and 454 image samples for each grade at 4x, 10x, 20x, and 40x magnification factors shown in Figure 3.2, which leads to a significant challenge due to its small size and imbalance. This imbalance is particularly problematic when developing a robust model for cancer grade classification. To address this issue, we employ data augmentation techniques on the original dataset, effectively expanding our training data [129, 130].

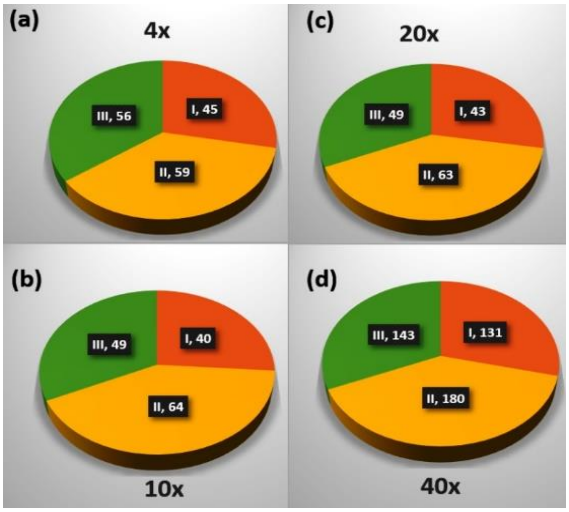


Figure 3.2 Distribution of invasive ductal carcinoma (IDC) samples in databiox dataset for three grades (I, II, and III) at four levels of magnification (a) 4x, (b) 10x, (c) 20x and (d) 40x.

We make use of the 'Image Data Generator' tool from the Keras library for image data augmentation [133], configuring specific parameters to enhance diversity:

Rotation_range: We set this to 40% to allow rotations between -40° and +40°. This aids the model in adapting to various object orientations.

Width_shift_range: With a 20% range, we introduce horizontal shifts, helping the model better handle objects that aren't perfectly centered.

Height_shift_range: Similarly, a 20% range is used for vertical shifts in the image.

Shear_range: We set this to 20%, introducing distortions into the image.

Zoom_range: A 20% range enables random zooming in or out.

Horizontal_flip: Enabled (set to True) for random horizontal flipping of images, accommodating different object orientations.

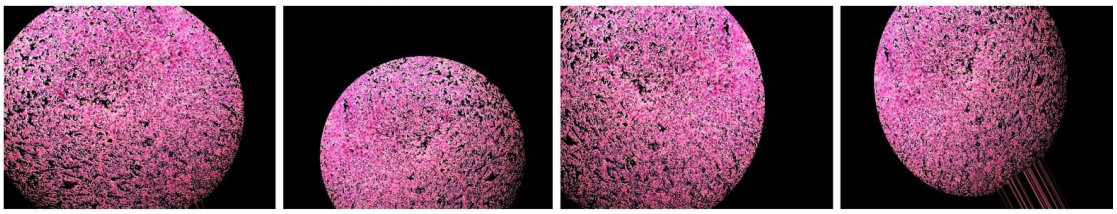


Figure 3.3. Outcome of data augmentation technique after implementing rotation, width shift, height shift, shearing, zoom in, zoom out and horizontal flipping on a segmented image.

Some sample images generated by data augmentation process are shown in Figure 3.3. Following this process, we generate twenty augmented images for each original input. To establish a well-balanced dataset, we remove randomly selected images from this augmented pool. As a result, we now have a newly augmented training dataset that maintains equilibrium, containing an equal number of samples for each grade. This balance plays a pivotal role in ensuring a robust and coherent basis for model training and evaluation. The updated augmented and balanced dataset consists of 225, 100, 104, and 304 image samples for each grade at 4x, 10x, 20x, and 40x magnification factors, respectively. In parallel, the original dataset also undergoes augmentation and balancing, maintaining an equivalent number of samples for each grade.

3.3 Results and Discussion

In order to evaluate the classification performance of both the pre-trained network on segmented images and the entire dataset, the complete dataset was divided into training and testing sets with a ratio of 90% for training and 10% for testing. In our research, we employ an image segmentation process to identify and extract regions of interest within histopathological images.

Table 3.1. Performance analysis for breast cancer grade classification on segmented images using pre-trained CNNs.

*WS – With Segmentation, **WOS – Without Segmentation

Classifier	Mag	Precision		Recall		F_1 Score		Accuracy		AUC					
										WS			WOS		
		WS*	WOS**	WS	WOS	WS	WOS	WS	WOS	I	II	III	I	II	III
VGG 16	4x	10	9	33	30	15	13	30	25	72	64	56	55	50	60
	10x	7	7	33	33	11	11	20	20	60	61	56	72	64	56
	20x	10	35	30	48	15	39	28	47	49	49	60	82	67	71
	40x	45	43	40	36	28	21	36	32	60	60	74	74	68	65
VGG 19	4x	19	9	33	33	15	15	28	28	57	45	59	56	62	59
	10x	7	7	33	33	11	11	20	20	59	58	58	73	46	86
	20x	10	10	33	33	16	16	31	31	57	49	60	77	56	74
	40x	9	9	33	33	15	15	28	28	61	61	76	69	72	52
ResNet50	4x	58	45	56	45	54	44	54	44	77	67	62	55	61	57
	10x	53	57	49	58	47	53	50	53	64	58	51	72	52	78
	20x	42	49	40	50	40	49	41	50	51	48	62	81	73	73
	40x	61	48	57	48	56	48	57	49	71	66	75	77	70	63
InceptionV3	4x	67	65	66	64	66	64	66	63	89	78	84	89	73	79
	10x	75	93	74	94	74	93	80	93	100	94	87	99	99	99
	20x	89	85	87	84	87	84	88	84	100	98	100	100	88	93
	40x	89	83	89	82	89	83	89	83	98	98	95	96	93	94
InceptionResNetV2	4x	56	57	56	58	56	58	56	57	79	70	78	88	73	79
	10x	85	77	84	80	84	76	87	77	79	70	78	94	92	90
	20x	89	82	88	81	87	81	88	81	100	94	80	98	90	94
	40x	82	72	82	71	82	71	82	72	94	88	94	96	88	88
MobileNet	4x	67	73	63	72	64	73	63	72	95	76	81	91	80	89
	10x	87	82	85	87	86	83	90	83	100	93	83	96	97	99
	20x	94	94	94	94	94	94	94	94	100	99	100	97	100	100
	40x	89	82	89	82	89	82	89	83	98	97	96	96	92	91
NASNetMobile	4x	70	61	70	60	69	60	69	60	80	79	86	81	80	80
	10x	76	74	77	76	76	74	80	77	99	89	80	84	96	88
	20x	88	91	88	91	87	90	88	91	98	95	97	95	97	100
	40x	85	69	82	69	82	69	83	70	97	95	93	91	85	91
Xception	4x	79	72	76	70	77	70	**76**	71	94	84	94	**88**	**86**	**93**
	10x	82	87	81	87	81	86	83	**91**	99	95	90	**99**	**98**	**97**
	20x	87	91	84	90	84	90	84	**91**	98	93	100	**100**	**95**	**94**
	40x	87	91	87	91	87	91	87	**91**	98	94	97	**99**	**99**	**98**

This aids in cancer grade detection, where we utilize pre-trained convolutional neural networks fine-tuned using segmented images. To validate our approach, we assess the performance of models trained on segmented images in comparison to those trained on the entire images. These experiments are conducted on a balanced dataset obtained through our discussed data augmentation protocol.

We evaluate the results of these experiments in terms of precision, recall, and accuracy for each class individually. To facilitate comparisons, we calculate an average of results across the three classes for each experiment. Additionally, we analyze the classification performance using ROC analysis and determine the AUC for result validation. These significant findings are presented in Table 3.1 for easy reference.

The comparison of the proposed model's performance is illustrated in Figures 3.4, 3.5, 3.6 and 3.7. In these Figures, we compare the confusion matrix (CM), ROC curve of the pre-trained models on a balanced and augmented Databiox dataset without segmentation at 4x, 10x, 20x and 40x level of magnification. Notably, the Xception model outperforms other models with an AUC of 99%, 98%, and 97% at 10x, and AUC of 100%, 95%, and 94% at 20x, and AUC of 99%, 99%, and 98% at 40x for Grade I, II, and III, respectively. This outperformance is observed for whole images, while at 4x magnification, the Xception model trained over segmented samples surpasses the other models. It achieves an accuracy of 71%, 91%, 91%, and 91% at 4x, 10x, 20x, and 40x levels of magnification, demonstrating consistent performance for whole images as opposed to segmented ones.

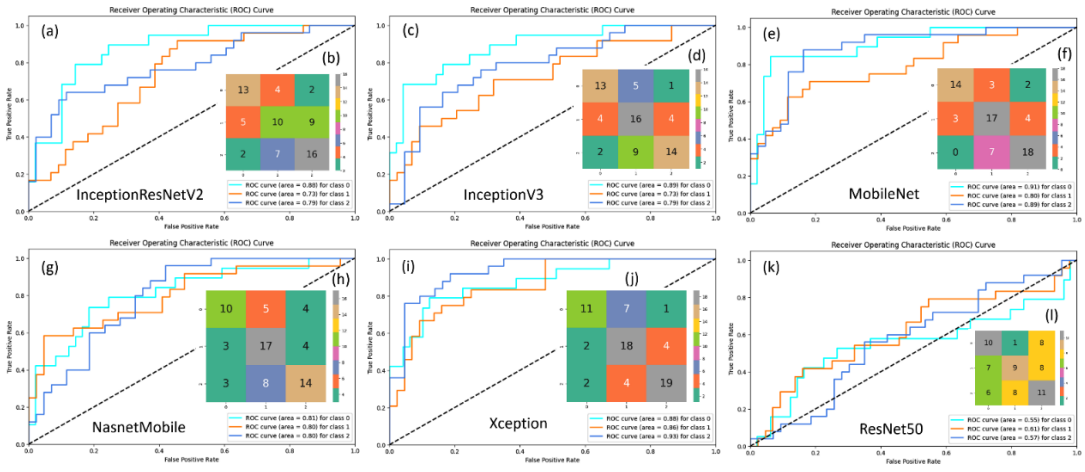


Figure 3.4 ROC and confusion matrix analysis for breast cancer grade classification without segmentation at 4x magnification level (a, b) InceptionResNetV2, (c, d),

InceptionV3, (e, f) MobileNet, (g, h) NASNetMobile, (i, j) Xception, and (k, l) ResNet50.

From the findings in Table 3.1, it is revealed that the considered models exhibit different behaviours in classifying breast cancer grades at four different levels of magnification (4x, 10x, 20x, and 40x). VGG16 and VGG19 models show poor performance over the segmented Databiox dataset, and this trend continues for the unsegmented but augmented dataset. This suggests that these models underfit the Databiox dataset due to their inability to learn from its complexity.

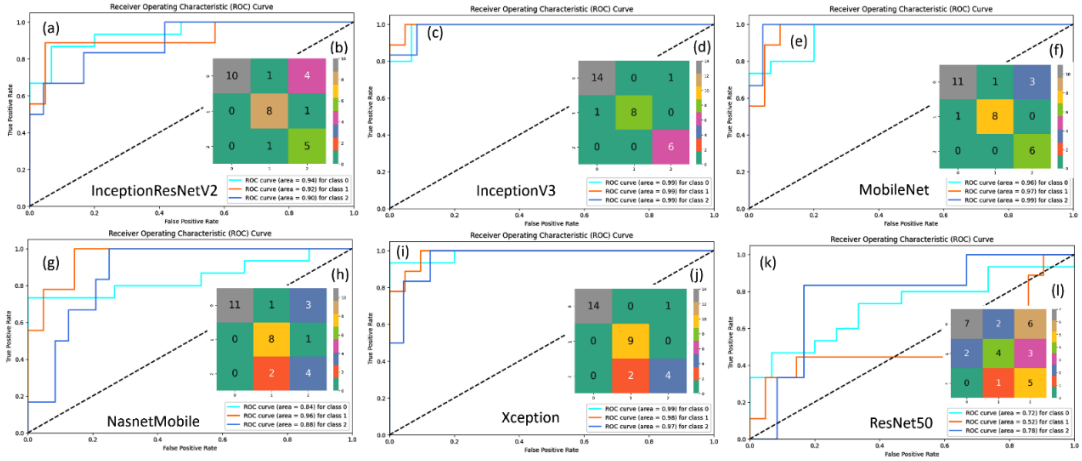


Figure 3.5 ROC and confusion matrix analysis for breast cancer grade classification without segmentation at 10x magnification level (a, b) InceptionResNetV2, (c, d) InceptionV3, (e, f) MobileNet, (g, h) NASNetMobile, (i, j) Xception, and (k, l) ResNet50.

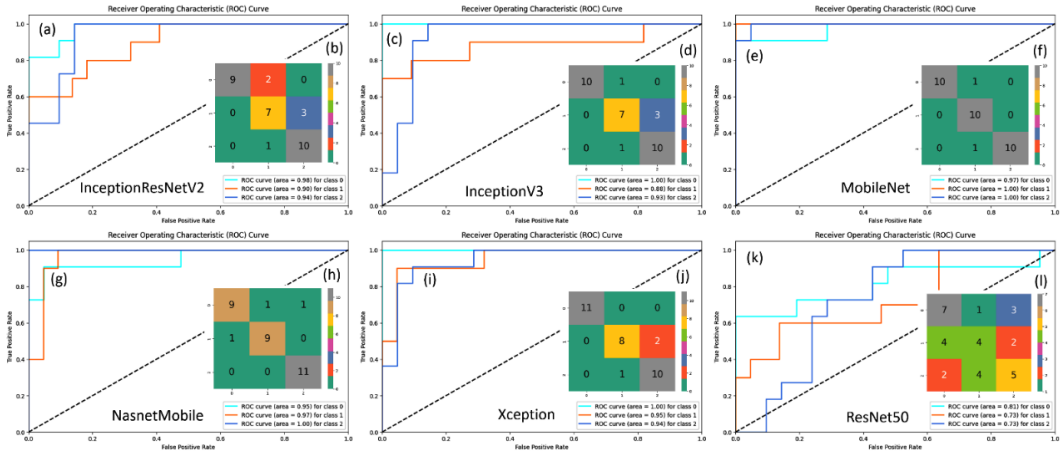


Figure 3.6 ROC and confusion matrix analysis for breast cancer grade classification without segmentation at 20x magnification level (a, b) InceptionResNetV2, (c, d),

InceptionV3, (e, f) MobileNet, (g, h) NASNetMobile, (i, j) Xception, and (k, l) ResNet50.

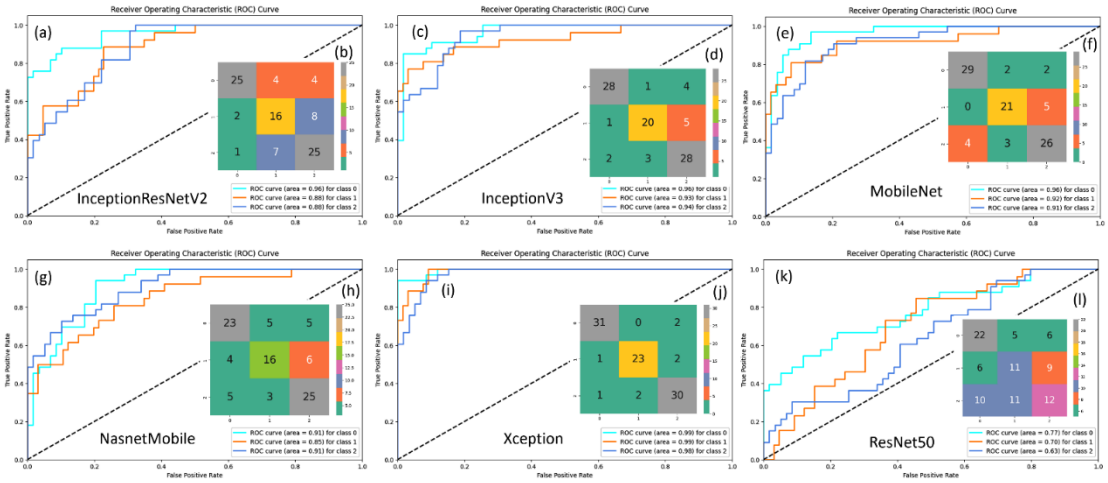


Figure 3.7 ROC and confusion matrix analysis for breast cancer grade classification without segmentation at 40x magnification level (a, b) InceptionResNetV2, (c, d), InceptionV3, (e, f) MobileNet, (g, h) NASNetMobile, (i, j) Xception, and (k, l) ResNet50.

Although ResNet50's performance is unsatisfactory, it is better than VGG16 and VGG19. Its suboptimal performance may stem from overfitting, wherein the network's capacity surpasses the size of the dataset. On the other hand, InceptionresnetV2, InceptionV3, NASNetMobile, and MobileNet models provide comparable and acceptable performance over both segmented and whole images. However, it's crucial to note that the segmentation results did not meet our initial expectations in terms of accuracy and reliability. Several issues and limitations were encountered during the segmentation process, including over-segmentation, under-segmentation, and challenges in color range selection. Over-segmentation led to an excessive number of segmented regions, making it difficult to distinguish relevant and non-relevant structures. In Figure 3.8(a)-(b), we can observe a case from a Grade I class at the 4x level, illustrating an instance of over-segmentation. Notably, this leads to the inclusion of the slide's number, 9057, as part of the regions of interest. However, in Figure 3.8(c)-(d), we encounter a different scenario, characterized by under-segmentation, where not all the relevant regions of interest are effectively captured. This deficiency is evident in the image, as certain regions remain unsegmented following the implementation of the proposed segmentation protocol. The choice of color range selection proved

challenging because of variations in staining procedures and tissue types, leading to inconsistent segmentation results. While the color range thresholds were adjusted to match specific staining techniques and tissue types, there were cases where it was difficult to find a one-size-fits-all solution. These challenges significantly impacted the precision and usability of the segmented data for Databiox dataset classification at different magnification levels. In Figure 3.8(e)-(f), we can observe another case from a Grade I class at the 4x level, illustrating an instance of inconsistent segmentation due to choose of color range selection which does not fit for that particular image. Furthermore, the segmentation method struggled to adapt to variations in staining protocols, tissue types, and image quality when applied to unseen histopathological images.

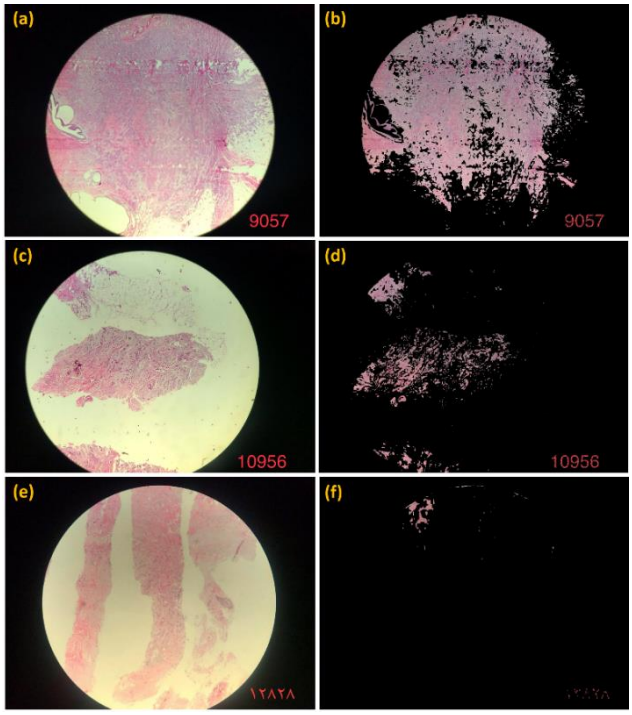


Figure 3.8 Segmentation issues shown for Grade I image samples at 4x level of magnification, (a)-(b) Over segmentation, (c)-(d) Under segmentation, (e)-(f) Choice of color range selection.

In light of these limitations, it is clear that further refinement of the segmentation approach is necessary to achieve more robust and reliable results. Future research should explore advanced segmentation techniques, leverage advanced deep learning

methods, and optimize color range selection strategies to address these issues. This will lead to improved outcomes in cancer grade detection and histopathological analysis.

- **State-of-the-art Comparison**

In this section, performance comparison of our proposed method is conducted with the existing state-of-the-art findings using the Databiox dataset (refer to Table 3.2). Kumaraswamy et al. [104] employed ensemble model of three pre-trained neural networks without considering magnification level and achieved a remarkable performance with an accuracy of 94%. Zavareh et al. [101] applied a transfer learning strategy, employing a pre-trained VGG16 CNN without fine-tuning, and attained a validation accuracy of 88%. Notably, there is limited prior research available for the Databiox dataset, and no research work is reported on Databiox with segmented images, making it challenging to perform extensive comparisons with our results at this point.

Table 3.2. Performance comparison of our proposed method is conducted with the current state-of-the-art findings using the Databiox dataset

Method	Approach	Acc.	Pre.	Recall	F1-Score	AUC
E. Kumaraswamy et al. [104]	Ensemble Model	94%	97%	93%	95%	96%
P. H. Zavareh et al. [101]	VGG-16	88%	--	--	--	--
Present Work	Xception (4x)	71%	72%	70%	70%	93%
	Xception (10x)	91%	87%	87%	86%	97%
	Xception (20x)	91%	91%	90%	90%	94%
	Xception (40x)	91%	86%	91%	91%	98%

3.4 Conclusion

This study explores the profound impact of segmentation techniques on the classification performance of pretrained neural networks deployed to the grading of breast cancer. To validate our findings, we compare model performance using the Databiox dataset, both in its segmented, augmented, and balanced form and the original full images. Key insights derived from this investigation include:

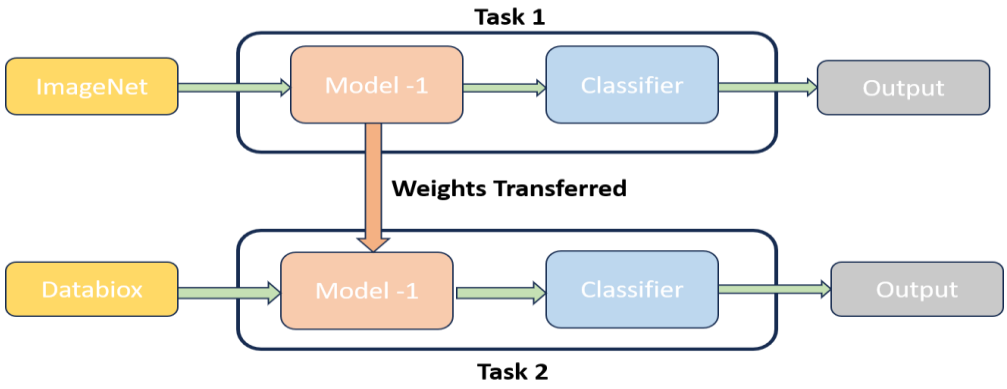
- Fine-tuned pre-trained Xception models consistently demonstrated outstanding performance, achieving accuracy rates of 71%, 91%, 91%, and 91% at 4x, 10x, 20x, and 40x magnifications, respectively, when processing the entire images without any segmentation.
- The challenges of over-segmentation, under-segmentation, and color range selection loom large in the field of segmentation, mainly because of variations in staining protocols and tissue types. This underscores the complexity involved in accurate image segmentation for diverse scenarios.
- Classifying images at 4x magnification proved to be a significant challenge, as all considered models displayed lower accuracy. This suggests that distinguishing subtle features at the smallest magnification level remains a notable obstacle.
- Misclassification of most class 1 samples occurred due to their intermediate position between class 0 and class 2. As a result, some clinical characteristics from adjacent classes contaminated class 1, posing a challenge for the classifier. Employing a soft decision boundary could potentially resolve this issue.

Nonetheless, it is crucial to acknowledge that color range selection has its limitations, particularly when confronted with staining variations and varying image qualities. In cases involving complex staining patterns or overlapping colors, defining suitable color ranges can be a formidable task. In such instances, more advanced segmentation techniques, such as machine learning-based methods or multi-channel image analysis, may prove necessary to attain more accurate and reliable results.

4

Invasive Ductal Carcinoma Grade Classification in Histopathological Images Using Transfer Learning Approach

"The measure of intelligence is the ability to change"
-Albert Einstein



In this chapter, breast cancer grade classification implemented by means of the transfer learning method. Herein, deeper pre-trained models, ResNet152V2, DenseNet201, VGG16, NASNetMobile and Inception_V3 used for the feature extraction tasks and RF classifier in the last stage for the classification task. The transfer learning approach we introduced demonstrated exceptional results, achieving AUC scores of 98% and 83% for the DensNet201 and VGG16 models, respectively, in grade 1 classification. For grade 2 classification, the DensNet201 model scored 75% AUC, and the NASNetMobile model scored 72% AUC. In grade 3 classification, the Inception_V3 model achieved a 65% AUC, while the NASNetMobile model secured a 69% AUC. In addition, the performance of the model elaborated with train-test data splitting of 80%-20% and 90%-10%. The proposed model attained the highest performance in contrast to the existing state-of-the-art techniques.

4.1 Introduction

Breast cancer stands as the most prevalent form of cancer among women globally. It is difficult to identify an actual cause of breast cancer as there are plenty of factors such as unbalanced nutrition, women's lifestyle, lack of proper physical activity, premature or late marriages, breastfeeding which may cause the formation of tissue masses. Some of these causative factors cannot be regulated like menstruation at young age and menopause at an older age and inherited changes in genes. As per the World Health Organization (WHO), breast cancer ranks as the second leading cause of cancer-related deaths worldwide. As per 2021 cancer statistics, 18.1 million new cases were recorded and 9.5 million cancer deaths were accounted for around the world [131]. Digitized histopathological images are supposed to be the most reliable source of information to detect breast cancer at an early stage where stained tissue slides are scanned to produce images and analyzed by the pathologists. Breast cancer is generally classified into two broad categories: In-situ (benign) and Invasive (malignant). Invasive is the type of cancer that invades other body parts along with breast tissues and forms new masses of tissues while the in-situ type of cancer is confined to breast tissues only. Invasive breast cancer is further divided into two types i.e., invasive ductal carcinoma (IDC) and invasive lobular carcinomas (ILC). The cancerous cells grow outside the ducts in the case of IDC. However, grows in lobules for ILC. [https://www.breastcancer.org]. Therefore, detecting cancer early and accurately is crucial to decrease mortality rates by enabling optimal treatment during the initial stages.

Manual analysis of biomedical images is very difficult especially for the cancer types which share similar expressions. The probability of making errors in manual diagnosis of cancer type is very high and tedious too for a radiologist. The problem of manual analysis can be overcome by utilizing robust and advanced CAD systems that will help in providing accurate diagnosis of cancer [132]. Therefore, it is highly required to develop a smart, high-end CAD system for cancer disease.

In this context, deep neural networks (DNN) are drawing special attention from researchers for designing CAD systems due to their automatic representation learning ability[133]. The availability of high-end configured computers with increased computational power makes the use of DNN more popular among the science community. Among the variety of DNNs, a CNN is the most common and applicable

variant of DNNs [134, 135]. Their wide-ranging applications in computer vision stem from their capacity for weight sharing and local feature connectivity. These two attributes allow the CNN to behave like local feature extractors for extracting the same patterns throughout the image with fewer trainable parameters and low computational power. The CNN requires a huge amount of training data which further requires high-performance GPU for faster training of the CNN. In the medical domain, the availability of adequate data for CNN training is a big question and also difficult to accumulate.

In this context, the transfer learning technique could be a feasible option to obtain the significant performance of the model which can avoid the problem of over-fitting, encountered in the absence of big data for training purposes. In the transfer learning technique, the weights of the pre-trained network (trained on the extensive ImageNet dataset) are transposed to a new network for performing a related task [136]. This technique unfolds in three distinct applications: baseline model, feature extractor, and fine-tuning [137]. In the baseline model application, the architecture of the pre-existing model is harnessed and initiated from scratch, with modifications limited to adjusting the number of neurons in a dense layer to align with the target classes. The transfer learning approaches illustrated in Figure 4.1.

In this work, the following contributions have been made in terms of experimentation:

- Magnification independent classification of grade for IDC has been performed where the data is obtained from Databiox dataset (http://databiox.com).
- The impact of the data-splitting ratio on the performance of the classifier is studied by using two different training-testing data splitting ratios. The motivation for the experimental work stems from Shallu et al. research contribution in which they classified histopathological images of breast cancer obtained from the BreakHis dataset using three different pre-trained networks namely, VGG16, VGG19 and ResNet50.
- The main objective of this work is to determine the existing state-of-the-art architectures including ResNet152V2, VGG16, NASNetMobile, DenseNet201, and Inception_V3 in the classification of IDC grade by employing transfer learning approach of network's training.

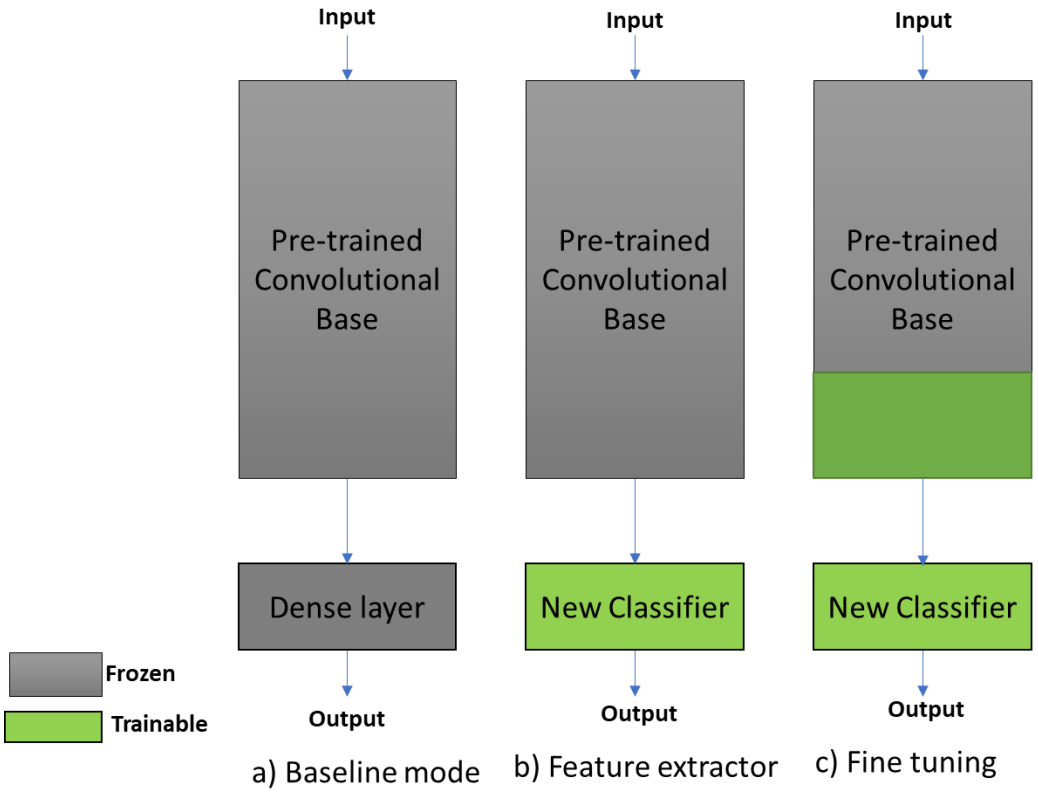


Figure 4.1 Transfer learning approaches a) baseline model b) feature extractor c) fine tuning of layers.

4.2 Methodology

In this work, the prevailing state-of-the-art deep CNN models: ResNet152V2, VGG16, NASNetMobile, DenseNet201, and Inception_V3 are employed as feature extractors [138]. All layers in pre-trained networks remain unchanged, with the exception of the penultimate layer, which is substituted with a conventional classifier. Here, RF is employed as the classifier where the feature vector obtained from the penultimate layer is embedded into the classifier for the outcome. Since the task to classify IDC breast cancer grade based on histopathological images are very complex and tedious so potential architectures have been considered to solve this task. The ResNet152V2, VGG16, NASNetMobile, DenseNet201, and Inception_V3 all are popular and in-depth architectures that have proved their potential in the classification of the ILSVRC-2015 challenge.

4.2.1 Data Augmentation

It is a technique employed to expand the dataset size, addressing challenges posed by limited and imbalanced datasets. A variety of data augmentation approaches like

rotation, shearing, transformation, interpolation, zooming, flipping, cropping, and translation, are accessible as a pre-processing step to increase the size of the dataset [139]. Some popular data augmentation techniques such as cropping, scaling, flipping, rotation, and translation have been applied in many previous studies [140]. All these techniques do not apply to histopathology images. In this context, we have applied only resizing, rotation, zooming, and flipping to avoid the loss of information. The random rotation has been used in the augmentation process instead of fixed rotation. Moreover, the size of each image in the databiox dataset is fixed to 224x224 by applying to resize the technique of data augmentation.

4.2.2 Magnification Independent Classification

The magnification factor plays a key role in the analysis of histopathological images [141]. The process of magnification is meant to enlarge the size of the region of interest (ROI). A magnification factor can be of any resolution such as 4x, 10x, 20x, 40x so that the ROI can be viewed comfortably by the radiologists. The analysis of histopathological images is difficult at low magnification factors because of complex patterns. Since acquiring images at divergent magnification factors introduces the heterogeneity in the background. Therefore, it is essential to develop an automated CAD system that can learn discriminative features from histopathological microscopy images of different magnifications to make a distinguishable diagnosis.

4.3 Results and Discussion

In this study, we have showcased the effectiveness of pre-trained CNN models in the classification of IDC breast cancer grades through histopathological images where the experiments are performed on local machines using TensorFlow and Keras libraries. The whole dataset is split into training and testing data to assess the classification performance of the considered pre-trained network in classifying the IDC grade histopathological images. Dividing the complete dataset into train-test data with a ratio of 90%-10%, 80%-20%, and 70%-30% is the general practice in the analysis of deep CNN network performance. However, we have considered only two training-testing data splitting ratios i.e., a 90%-10%, and 80%-20% to evaluate the classification performance of the proposed models. Precision, Recall, F1 score, accuracy, and ROC are the performance evaluation parameters, computed for all the considered pre-trained networks for a coherent comparison. Area under the curve (AUC) the ROC curve is also

computed to validate the performance of the classifiers.

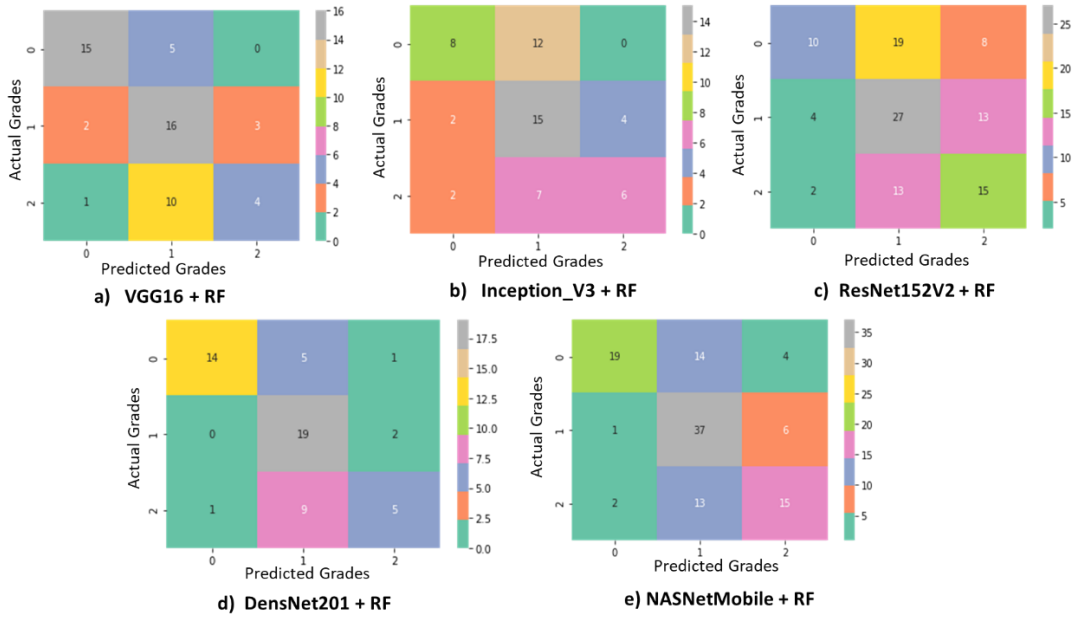


Figure 4.2 Confusion matrix for IDC breast cancer classification using a) VGG16+RF b) Inception_V3+RF c) ResNet152_V2+RF d) DensNet201+RF e) NASNetMobile+RF transfer learning models.

ResNet152V2, VGG16, NASNetMobile, DenseNet201, and Inception_V3, all are pre-trained on the ImageNet dataset and used as feature extractors along with RF classifier to classify IDC grade into Grade I (class 0), II (class 1), and III (class 2). Further, to determine the performance of the model with the change in training-testing data splitting ratio, two different training-testing data splitting ratios (90%-10% and 80%-20%) are used because the size of the dataset has a considerable impact on the performance of the CNNs. The proposed approach is used for magnification independent classification of three IDC grade class classification of breast cancer. In this context, the confusion matrix for all the considered models with a splitting ratio of 90%-10% is illustrated in Fig. 4.2. Moreover, the performance of the classifiers is compared on the basis of ROCs and AUC, depicted in Fig. 4.3. It has been observed from the comparative analysis that the DenseNet201 significantly outperformed all other pre-trained networks in terms of accuracy. ResNet152_V2 network achieved unsatisfactory performance over the databiox dataset and represented their inability in transferring the gained knowledge on the new task of classification. Unfreezing more layers in the network can be a way to improve the efficiency of the networks as it

reduces the capacity of the network and will further reduce the problem of overfitting.

Furthermore, it has been observed from the ROC curve analysis that DensNet201 produced AUC of 98% and VGG16 produced AUC of 83% surpass in the classification of Grade-I. DensNet201, AUC of 75% and NASNetMobile, AUC of 72% produced in identification or classification of Grade-II as compared to the other pre-trained models. However, the Inception_V3, AUC of 65% and NASNetMobile, AUC of 69% achieved significantly for Grade-III classification. Nearly the same trend has been observed for the remaining data splitting ratio of 80%-20%. The results obtained from the proposed transfer learning models are tabularized in Table 4.1. It is further observed that when the training data is 90% of the whole dataset, the two networks (VGG16 and DenseNet201) performed well, while, NASNetMobile, Inception V3, and ResNet152V2 performed well for the training–testing data partitioning of 80-20%. NASNetMobile, Inception V3, and ResNet152V2 deviate from the norm because these are not equally sensitive to the images of Grade I, II, and III during the 80-20% splits.

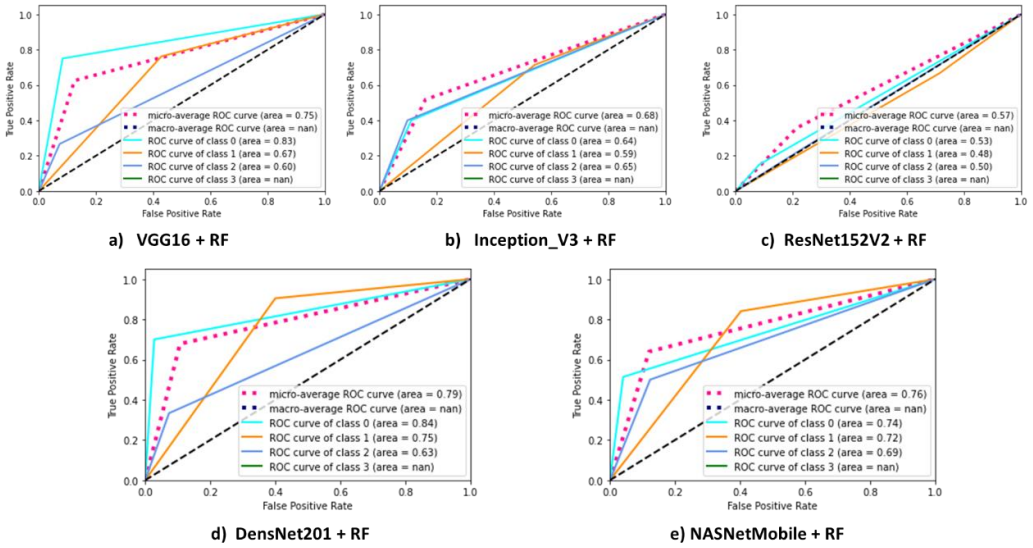


Figure 4.3 ROC performance analysis of IDC BC classification for a) VGG16+RF b) Inception_V3+RF c) ResNet152_V2+RF d) DensNet201+RF e) NASNetMobile+RF transfer learning models.

Table 4.1 Performance analysis of fine-tuned pretrained networks in the classification of histopathological images (VGG16, Inception_V3, ResNet152v2, DenseNet201 and NASNetMobile)

Classifier	Splitting of train & test	IDC Grade	Precision	Recall	F1 score	Accuracy	ROC
VGG16+RF	**80-20**	I	0.71	0.54	0.62	0.58	0.72
		II	0.53	0.61	0.57	0.57	0.63
		III	0.50	0.53	0.52	0.57	0.67
	90-10	I	0.83	0.75	0.79	0.64	0.83
		II	0.52	0.76	0.62	0.62	0.67
		III	0.57	0.27	0.36	0.61	0.60
Inception_v3+RF	**80-20**	I	0.81	0.35	0.49	0.59	0.66
		II	0.50	0.73	0.59	0.53	0.62
		III	0.45	0.47	0.46	0.52	0.63
	90-10	I	0.67	0.40	0.50	0.56	0.64
		II	0.44	0.71	0.55	0.52	0.59
		III	0.60	0.40	0.48	0.51	0.65
ResNet152V2 +RF	**80-20**	I	0.62	0.27	0.38	0.50	0.59
		II	0.46	0.61	0.52	0.47	0.57
		III	0.42	0.50	0.45	0.46	0.62
	90-10	I	0.50	0.15	0.23	0.39	0.53
		II	0.36	0.67	0.47	0.36	0.48
		III	0.27	0.20	0.23	0.32	0.50
DenseNet201+ RF	**80-20**	I	0.86	0.51	0.64	0.68	0.74
		II	0.58	0.84	0.69	0.64	0.72
		III	0.60	0.50	0.55	0.63	0.69
	90-10	I	0.93	0.70	0.80	0.72	0.84
		II	0.58	0.90	0.70	0.68	0.75
		III	0.62	0.33	0.43	0.67	0.63
NASNetMobile+ RF	**80-20**	I	0.67	0.32	0.44	0.54	0.62
		II	0.48	0.66	0.56	0.50	0.60
		III	0.45	0.50	0.48	0.50	0.64
	90-10	I	0.62	0.40	0.48	0.47	0.63
		II	0.41	0.57	0.48	0.45	0.54
		III	0.36	0.33	0.34	0.45	0.56

4.4 Conclusion

The present work demonstrates the efficient utilization of the transfer learning approach to extract the significant features from the pre-trained networks ResNet152V2, VGG16, NASNetMobile, DenseNet201, and Inception_V3 for IDC grade classification with fine-tuning strategies. The pre-trained networks have been utilized as feature extractors and the extracted feature vectors are used for identifying the particular class using a RF classifier.

A few of the conclusions of this research include:

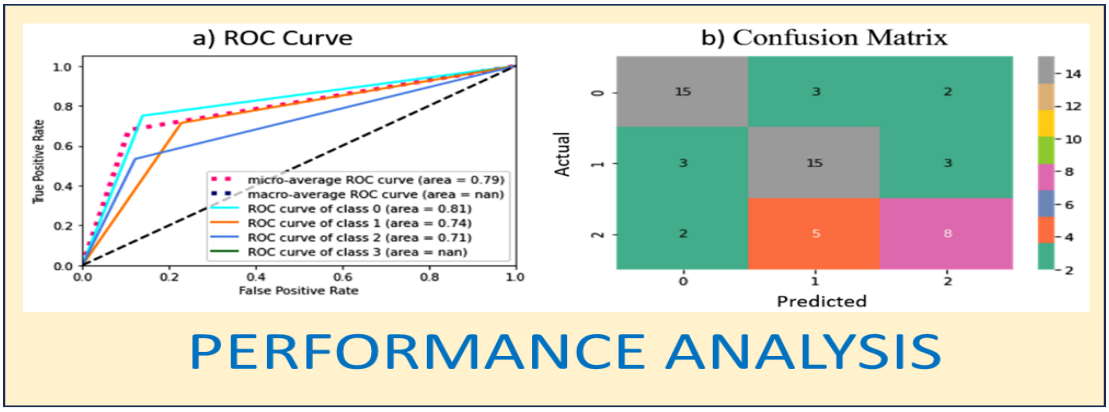
- Fine-tuned random forest classifier network performs well in contrast to fully trained network with dense layer.
- Fine-tuning of pre-trained networks is simple instead of implementing a new deep network trained from scratch with new weights.
- Fine-tuned pre-trained networks are robust to the size of the dataset splitting, that is even though the size of the training data is decreased, the performance of the network does not deteriorate.
- The performance of the network will be greatly reduced if the network is biased towards the data belonging to a particular grade.
- The network's capacity plays a crucial role in shaping its performance, potentially giving rise to challenges like overfitting and underfitting.

As a result, the optimal capacity of a network significantly depends on the application to be handled that could be improved by considering several factors such as type and size of data set, data augmentation techniques, and layer-wise fine-tuning approaches as a future aspect of this work to further enhance the performance of the adapted model.

5

Transfer Learning-Based CNN Model for Efficient Invasive Ductal Carcinoma Breast Cancer Grade Classification using Databiox Dataset

"Making good decisions is a crucial skill at every level"
- Peter Drucker



In this chapter, we have implemented a transfer learning-based histopathological image analysis method to determine the type of IDC breast cancer grade, wherein the deep convolution neural networks (CNN) are utilized. The proposed method performance is evaluated for magnification-independent and magnification-dependent classification to ascertain the complexity of the problem. The results demonstrate that the complexity of the classification task gets reduced when we consider the image of each magnification factor separately. The VGG19 + SVC network with linear SVM provides the highest accuracy of 79% at 4x, MobileNetV2 + KNN provides an accuracy of 100% at 20x, and NASNetLarge + SVC yields the highest accuracy of 100% and 91% at 10x and 40x, respectively. However, the ResNet50 + LR provides the highest accuracy of 75% for magnification-independent classification that ensures the complexity of this classification problem but still surpasses the existing state-of-the-art technique for Databiox dataset classification.

5.1 Introduction

Breast cancer is the most commonly diagnosed cancer worldwide. Several types of chemicals, pesticides, food additives, consumer products, and drinking water contaminants, are the primary sources of breast cancer which induce more secretion of hormones i.e., estrogen or progesterone [142]. Massive production of these hormones is one of the primary causes of breast cancer. Global Cancer Observatory [143], datasheet states that 1,92,92,789 number of people are living with the cancer across the globe, wherein 13,24,413 new cancer cases and 8,51,678 deaths are registered in India. Wherein 1,78, 361 women were newly detected with breast cancer in India alone for the year 2020. Moreover, breast cancer accounted for 26.3% of all newly detected cancers in women. American Cancer Society [144] also estimated 19,58,310 new cancer cases and 6,09,820 cancer deaths in 2023 in the U.S. The estimated number of breast cancer incidence and mortality cases from 2020 to 2040 year depicted in Figure 5.1. Histopathology is a powerful study to visualize and analyze diseases of the tissues which provides greater contrast between normal and abnormal tissues after staining for diagnostic purposes [145]. Advancements in the area of image processing techniques help in accelerating the diagnosis of cancer tissues as time is the most crucial factor in cancer treatment.

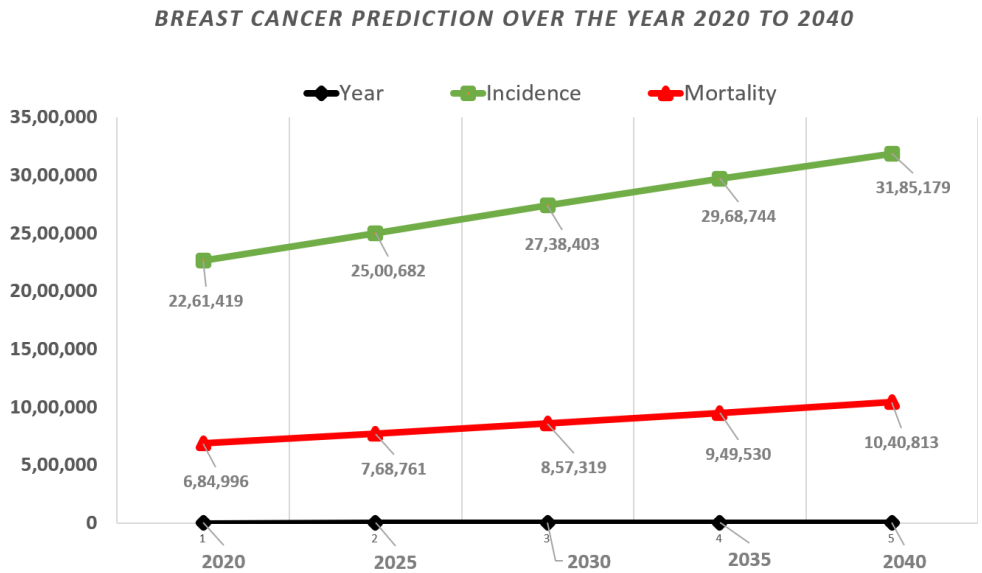


Figure 5.1 Estimated number of cancer incident cases and mortality cases over the year 2020 to 2040 across the worldwide. (Source: Global Cancer Observatory. Available online: https://gco.iarc.fr/today/online-analysis-multi-bars (accessed on 15 May 2023).

The transfer learning approach is a modern technique in the deep learning domain wherein the weights of a CNN model trained on large sized datasets (e.g., ImageNet Dataset) are transferred to perform new but relevant tasks [146]. Transfer learning approach can be implemented in three manners using the pre-existing architecture wherein the feature extractor and fine-tuning approach is implemented using the pre-trained model. While the pre-existing CNN models are learned from the scratch and used as a baseline model in the third approach of transfer learning. In fine-tuning, the last layer of the network is fine-tuned to a new specific task. The proposed model attained the best performance when fine-tuning is applied. Shallu et al. [44] also implemented the transfer learning approach for the classification of breast cancer. They applied layer-wise fine-tunning approach using pre-trained Alex Net model to determine the most appropriate layers of pre-trained model for fine-tuning. They observed that the level of fine-tunning contrasts with the level of image magnification factor. Moreover, they also verified the ability of pre-trained Xception model as feature extractor in the classification of breast cancer. In this scenario, the Xception model paired with an SVM classifier utilizing the 'radial basis function' kernel has consistently demonstrated the best performance across all four levels of magnification [98].

In this work, we have implemented the transfer learning approach using pre-existing network along with conventional classifier for detecting IDC grades. In this context, VGG16, VGG19, ResNet50, DenseNet201, NASNetLarge and MobileNetV2 pre-trained models are considered. The major reason behind the application of transfer learning approach is the scarcity of the dataset which may provide inaccurate results when end to end learning is applied on the small dataset. The major contribution of this work is outlined as follows:

- Six CNN models have been compared for their performance in predicting breast cancer grading using Databiox dataset. In best of our knowledge, this is the very first study where the transfer learning approach as feature extractor is implemented and compared for IDC breast cancer grade classification.
- Magnification independent classification is performed wherein the images at four different levels of magnifications i.e., 4x, 10x, 20x, and 40x are considered collectively for final classification.
- Impact of magnification level is also demonstrated by performing magnification

dependent classification wherein the images acquired at different levels of magnification are considered separately.

- Eventually, a coherent comparison is performed between magnification dependent and magnification independent classification approach to determine the most efficient approach for grade classification.

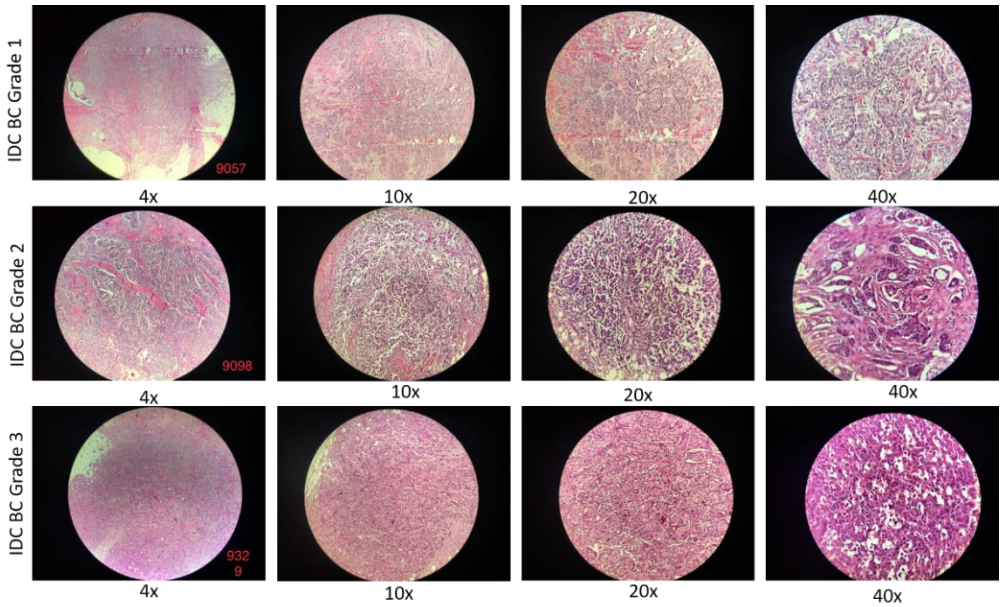


Figure 5.2 Invasive Ductal Carcinomas (IDC) breast cancer grades (grade 1, grade 2 and grade 3) with magnification levels 4x, 10x, 20x and 40x.

5.2 Methodology

A brief description of the pre-existing networks is provided that have been utilized to classify the breast cancer histopathological images. Inter and intra-observer variability in the manual analysis is one of the major issues in analyzing the histopathological images and makes it challenging to classify the images of different grades accurately. Thus, it becomes essential to develop a robust and accurate grade classification algorithm for precise diagnosis and prognosis of breast cancer at different magnification levels. The Databiox dataset selected for execution of this study and various magnification factors 4x, 10x, 20x and 40x of histopathological images shown in Figure 5.2. In this work, various pre-trained networks deployed including VGG16, VGG19, ResNet50, DenseNet201, NASNetLarge, and MobileNetV2 for feature extraction. However, the preceding layer of each model is substituted with different types of conventional classifiers namely, LR, LDA, KNN, DT, RF, and SVM to determine the most optimal and efficient model for classifying the three grades of IDC breast cancer.

The proposed fine-tuning of the pre-trained model for IDC breast cancer grade categorization is depicted in Figure 5.3.

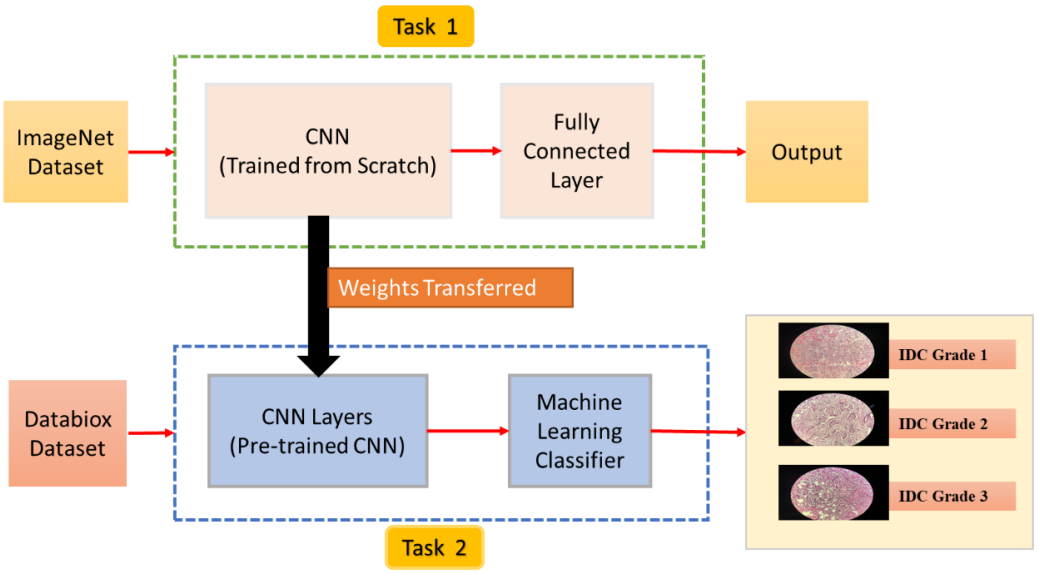


Figure 5.3 A schematic diagram depicting transfer learning approach utilized for IDC breast cancer grade classification.

5.2.1 VGG16 and VGG19

These two networks are proposed and developed by the Visual Geometry Group (VGG) from Oxford University. VGG16 is the first runner-up in the visual recognition challenge in 2014. The VGG16 and VGG19 architectures comprise a sequence of convolutional layers, pooling layers, fully connected layers, and a SoftMax layer, which will be discussed in the following sections.

- The input image passed through a stack of convolutional layers, where each convolutional filter has a small receptive field of size 3x3 with a stride value of 1. Zero paddings are also performed to maintain the spatial resolution of each feature map obtained after performing convolution.
- The VGG16, contains 13 convolution layers to extract the features and five max-pooling layers with a non-overlapping window of size 2x2, stride value 2, and tailed by three fully connected (FC) layers. Wherein two FCs consist 4096 nodes and ReLU as an activation function. However, the last FC layer has 1000 nodes corresponding to 1000 categories. The final layer is the soft-max layer which provides the final output. The major difference in the architecture of VGG16 and

VGG19 is the presence of 16 convolutional layers in VGG19 instead of 13 convolutional layers as in VGG16.

5.2.2 ResNet50

Residual Network (ResNet) is incorporated with several residual blocks that provide the shortest route among various layers. ResNet was introduced by Kaiming He, Xiangyu Zhang in 2015. In ResNet architecture, the shortest path connections enable hundreds of additional layers while accomplishing higher performance [147]. ResNet50 is implemented with 50 deep layers which are composed of 49 convolution layers followed with max-pooling layers and one FC layer. It has been utilized to categorize 1000 classes of the ImageNet database.

5.2.3 DenseNet201

In this CNN, the convolutional layer extracts features from the input image and obtains higher-level features with deeper layers [148]. As we know that the increasing number of convolutional layers cause the problem of gradient vanishing. Within the DenseNet architecture, the challenge of gradient vanishing is addressed through element-wise addition. Each layer receives additional inputs from all preceding layers and shares its own feature maps with all subsequent layers. This dense interconnection, where each layer interacts with all preceding layers results a closely compacted network, termed as Dense Net. Within the DenseNet architecture, the challenge of gradient vanishing is addressed through element-wise addition. In this framework, every layer acquires extra inputs from all previous layers and conveys its own feature maps to all succeeding layers.

5.2.4 NASNetLarge and MobileNetV2

The blocks in the Neural Architecture Search (NAS) Net architecture are predefined by reinforcement learning search method. Normal blocks and reduction blocks are collectively used in the architecture [149]. The convolutional blocks that return a feature map of the same dimension, are known as normal blocks. While the convolutional blocks that return a feature map with reduced height and width, are termed as reduction blocks. MobileNetV2 is used for embedded applications, designed with smaller parameters to reduce computational complexities [150]. This network consists of three layers for two types of residual blocks, i.e., one with stride of 1 and another with stride of 2. The initial layer is 1x1 convolutional layer with ReLU activation, subsequent layer

is depth-wise convolution and third layer is also 1x1 convolutional layer but without any non-linear activation function.

5.3 Results and Discussion

The proposed methodology is applied to the Databiox dataset to evaluate the results. Furthermore, the results are discussed thoroughly to determine the most efficient model for grade classification in both the cases, magnification dependent and independent classification. In the case of magnification independent classification, the available dataset is divided into training, validation, and testing data with a ratio of 60%, 20% and 20%, respectively. According to this training-testing data splitting scheme, 552 images, 370 images and 111 images are utilized for model's training, validation, and testing, respectively. The histopathological image data used for the training consists of 155, 219, and 178 images of grades 1, 2, and 3, respectively. The training-testing data splitting scheme for the Databiox dataset is presented in Figure 5.4. In feature extractor approach of transfer learning, the weights of pre-trained networks are utilized to train the same network on the new dataset that is not adequate to train the network from scratch.

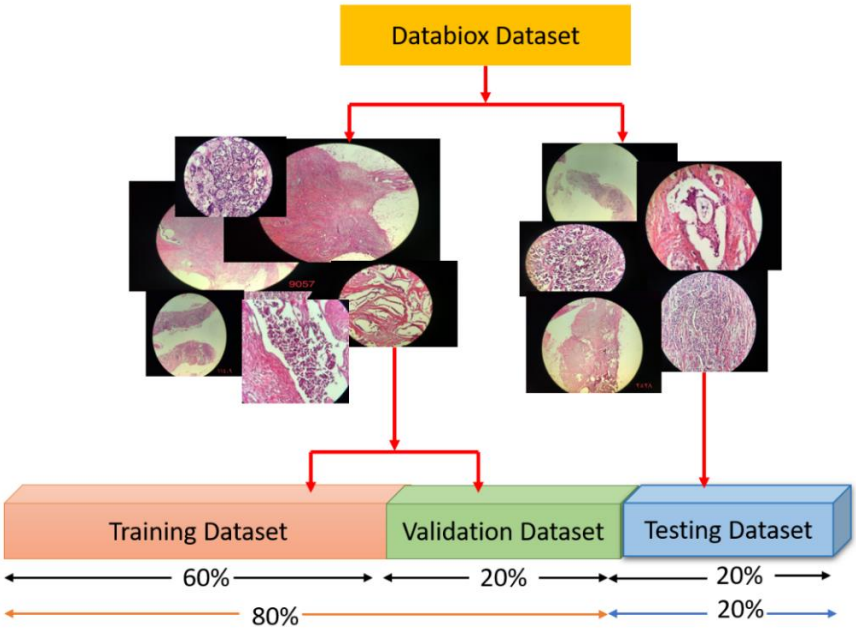


Figure 5.4 Databiox dataset distribution in training, validation, and testing with a ratio of 60%, 20%, and 20%, respectively.

The ImageNet dataset comprises 1,281,167 training images, 50,000 validation images and 100,000 test images corresponding to 1000 object classes. In the experimental

study, VGG16, VGG19, ResNet50 DenseNet201, NASNetLarge, and MobileNetV2, CNN models are trained using the ImageNet database. All layers (except the last layer) of the model are frozen, and learned weights are utilized from the pre-trained models for IDC grade classification. The final layer of the network is substituted with a conventional machine learning classifier for the final decision.

5.3.1 Classification Performance

A Confusion matrix is an authentic metric to evaluate model's performance that also used to compute other performance evaluation parameters such as recall, precision, specificity, accuracy, and AUC [156]. ROC Curve is another excellent method for evaluating the performance of a classification model, wherein, the AUC represents the degree of separability among the classes. Higher AUC ensures the great capability of the model in distinguishing among classes. A distinctive example for confusion matrix obtained for the IDC grade classification i.e., grade class 0, grade class 1, and grade class 2 (also, known as IDC grade 1, IDC grade 2 and IDC grade 3 respectively) are discussed and presented in Table S.1 (Supplementary Information).

The results obtained from the combination of VGG16 and six conventional classifiers (RF, LDA, LR, KNN, DT, and SVC) are tabularized in Table S.2(Supplementary Information). It has been observed from the ROC curve analysis that the VGG16 + RF classifier achieved the maximum AUC for all the three grades i.e., 70%, 63%, and 57% for grade class 0, 1, and 2, respectively, as shown in Figure S.1(Supplementary Information). Despite the highest accuracy, the sensitivity or recall is very less for grade class 0 and grade class 2. It confirms that the model is not much capable in differentiating between grade class 0 and grade class 2. The confusion matrix in Figure S.2 (Supplementary Information) also illustrates that the most of the samples from grade class 0 and grade class 2 are miss-classified as grade class 1 because of extreme confusion among all the three grades. Furthermore, a significant drop has been observed in the performance of the classification models namely, VGG16 + LR, VGG16 + DT, VGG16 + KNN, VGG16 + LDA, and VGG16 + SVC in respective manner. Since the linear kernel is used for SVC so the poor performance of VGG16 +SVC demonstrate that the features extracted from VGG16 are not linearly separable. Further, the VGG19 pre-trained model's performance along with the considered conventional machine learning algorithms are tested for the grade classification. The

results obtained from the proposed models are tabularized in Table S.3 (Supplementary Information). It has been noticed that the RF classifier provides the best performance in the classification of the grade class using features extracted by the pre-trained VGG19 network. The AUC obtained through VGG19 + RF model for grade class 0, 1, and 2 are 71%, 59%, and 67%, respectively, as shown in Figure S.3 (Supplementary Information). Although the RF classifier provided the maximum accuracy but there is still confusion among the grades due to similar clinical manifestations and leads to the mis-classification of data. This is well evident from the confusion matrix, depicted in Figure S.4 (Supplementary Information). However, a substantial decline has been observed in the performance of the rest classifiers namely, VGG19 + DT, VGG19 + LDA, VGG19 + LR, VGG19 + SVC, and VGG19 + KNN, respectively.

A different trend has been observed for the ResNet50 pre-trained model. Table S.4(Supplementary Information) demonstrate that the features extracted using the pre-trained ResNet50 are classified efficiently by the LR classifier. The maximum AUC obtained by the ResNet50 + LR classifier is 81%, 74%, and 71% for the grade class 0, 1 and 2, respectively (Figure S.5 (d)) (Supplementary Information). The confusion matrix in Figure S.6 (Supplementary Information) demonstrates the ability and superiority of ResNet50+LR classifier over other proposed classifiers. The ResNet50+LR classification model classified the test data more accurately and precisely. While a significant drop has been observed in the performance of other classifiers i.e., RF, SVC, LDA, KNN, DT. Skip connections in the ResNet network architecture is the major rationale behind their best performance as they alleviate the issue of vanishing gradient and make it considerably easier for the layers to learn identity functions. Eventually, they more efficiently train the deeper networks than previously possible.

It has been observed from Table S.5 (Supplementary Information) that the DenseNet201+SVC classifier achieved the highest performance with an AUC of 74%, 71%, and 78% for grade class 0, grade class 1, and grade class 2, respectively. DenseNet201 + RF classifier also provided an equivalent performance with an AUC of 71%, 70%, and 70% for grade class 0, grade class 1 and grade class 2, respectively, shown in Figure S.7 (e) and (f) (Supplementary Information).

Although the DenseNet was introduced to provide higher accuracy with fewer

parameters than ResNet, DenseNet201 achieved a lower performance than ResNet50 for cancer grade classification problems. This may happen due to excessive connection in the DenseNet architecture which increases the computational efficiency and memory efficiency but also makes the network more prone to overfitting. In DenseNet, every layer takes input from all preceding layers, allowing the classifier to leverage features across various levels of complexity.

Table 5.1 Performance of the models for magnification independent classification

Pre-trained network	RF		LDA		LR		KNN		DT		SVC	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
VGG16	54	69	47	58	50	61	51	60	50	61	43	**57**
VGG19	61	70	52	63	51	60	43	57	52	64	50	61
ResNet50	68	77	64	75	**75**	**79**	52	64	40	54	67	76
DenseNet201	64	73	63	72	61	70	57	68	50	61	65	74
NASNetLarge	49	68	63	72	53	67	57	68	47	58	54	69
MobileNetV2	58	69	**61**	**70**	57	68	53	67	44	56	49	68

Consequently, it provides more smooth decision boundaries but we are utilizing a pre-trained network (trained on the ImageNet dataset). As a matter of this fact, it is not able to provide generalized features which further leads to the problem of overfitting. Hence, the model is providing lower performance than ResNet50. Despite, the deep architecture of NASNetLarge and MobileNetV2 network, these two networks also not provided the significant results over ResNet50, VGG16 and VGG19. It has been observed from the Table S.6 and S.7 (Supplementary Information) that the NASNetLarge and MobileNetV2 network achieved a classification performance equivalent to VGG16 and VGG19, respectively. It is confirmed from the performance evaluation metrics shown in Figure S.9, S.10, S.11, and S.12 (Supplementary Information) that both the networks are not able to utilize their higher capacity. Therefore, these networks need more regularization to improve the performance of the proposed classifiers. Eventually, the ResNet50 + LR classifier achieved the highest performance for magnification independent data classification as compared to all other considered pretrained networks

combinations, shown in Table 5.1. The ROC curves and confusion matrix of ResNet50+LR classifier as shown in Figure 5.5. In the case of magnification dependent classification, we have followed the same training protocol as for magnification independent classification. According to which the dataset of each magnification factor (4x, 10x, 20x, and 40x) is divided into training, validation, and testing data with a ratio of 60%, 20%, and 20%, respectively. Further, all the six pre-trained classifiers along with six different conventional classifiers are implemented to determine the impact of the magnification factor on the performance of the proposed classification models.

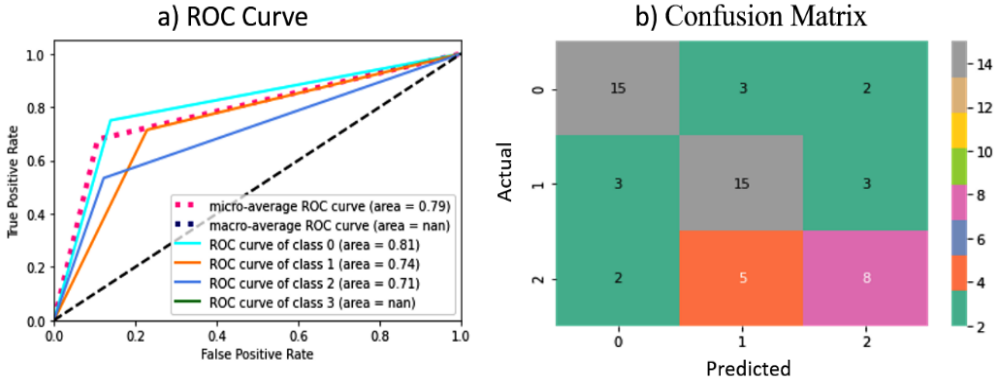


Figure 5.5 The ROC curve and confusion matrix for the best performing model i.e., ResNet50 + LR for magnification independent grade classification.

The results obtained from the proposed models are tabularized in Table 5.2 in terms of accuracy and AUC. At 4x magnification, the VGG19 + SVC and VGG19 + LR surpassed all the models with accuracy and AUC of 79% and 84%, respectively. At 10x, NASNetLarge along with LR, DT, and SVC, while at 20x, MobileNetV2 along with KNN achieved the maximum performance (i.e., 100% accuracy and 100% AUC) for the grade classification, respectively. Thus, both the networks utilized their higher capacity for magnification-dependent classification. At 40x magnification, NASNetLarge along with DT and SVC surpassed other models with accuracy and AUC of 91% and 93%, respectively. Figures 5.6 and 5.7 illustrated the best performance produced models ROC curve analysis and confusion matrix at magnification levels 4x, 10x, 20x and 40x of histopathological images.

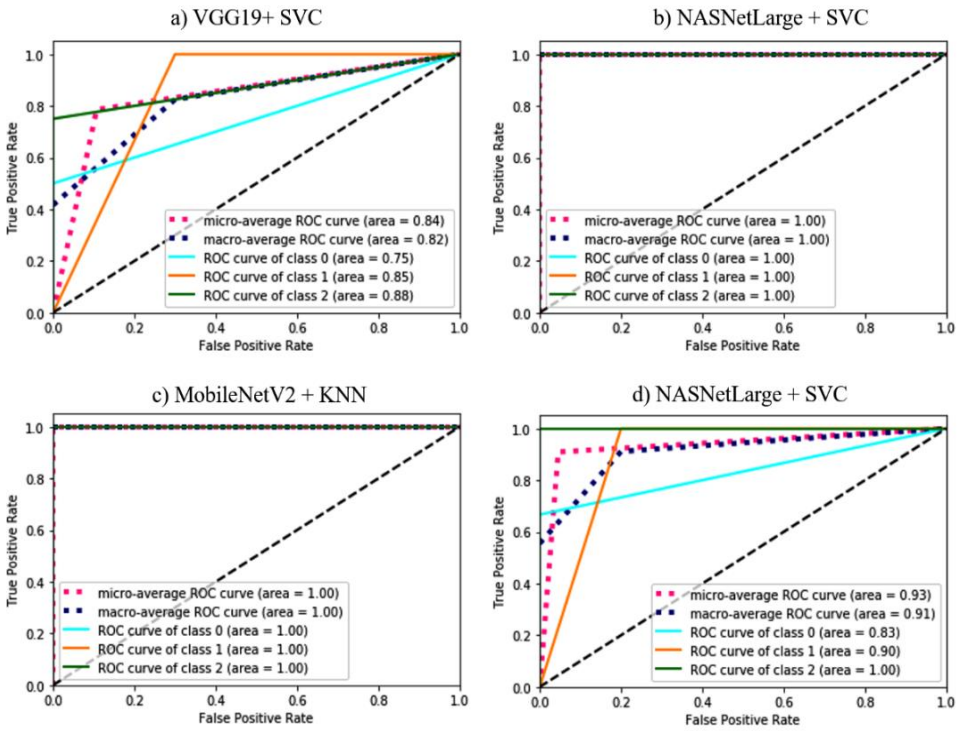


Figure 5.6 The ROC curve analysis of the best performing model, a) VGG19 + SVC at 4x, b) NASNetLarge + SVC at 10x, c) MobileNetV2 + KNN at 20x, and d) NASNetLarge + SVC at 40x for magnification dependent grade classification.

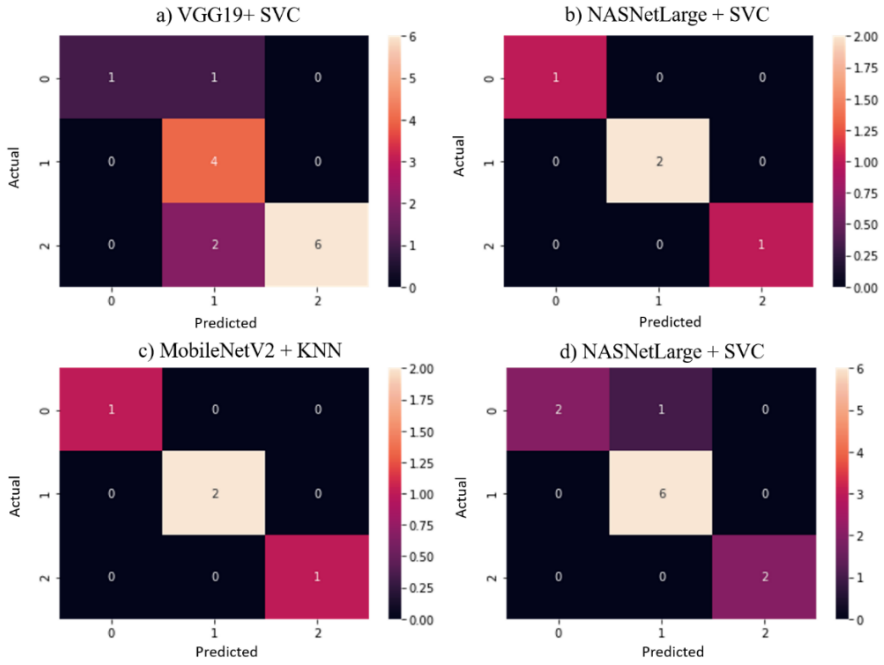


Figure 5.7 The confusion matrix of the best performing model, a) VGG19 + SVC at 4x, b) NASNetLarge + SVC at 10x, c) MobileNetV2 + KNN at 20x, and d) NASNetLarge + SVC at 40x for magnification dependent grade classification.

Table 5.2 Performance of the models for magnification dependent classification.

Pre-trained network	Classifier	4X		10X		20X		40X	
		Acc.	AUC	Acc.	AUC	Acc.	AUC	Acc.	AUC
VGG16	**RF**	64	73	50	62	75	81	64	73
	LDA	57	68	50	62	50	62	82	86
	LR	50	62	50	62	75	81	82	86
	KNN	64	62	50	62	50	62	55	66
	DT	43	57	75	81	50	62	64	73
	SVC	54	65	75	81	75	81	64	73
VGG19	**RF**	64	73	75	81	75	81	73	80
	LDA	64	73	64	73	75	81	73	80
	LR	79	84	64	73	75	81	73	80
	KNN	64	73	43	57	75	81	55	66
	DT	64	73	50	62	75	81	45	59
	SVC	79	84	50	62	75	81	73	80
ResNet50	**RF**	64	73	29	46	58	70	49	63
	LDA	64	73	64	73	52	72	51	63
	LR	64	73	64	73	48	66	49	68
	KNN	29	46	50	62	48	60	45	62
	DT	36	52	50	62	42	60	43	59
	SVC	64	73	29	46	61	79	49	56
DenseNet201	**RF**	56	67	61	71	68	76	60	70
	LDA	53	65	55	66	58	69	57	68
	LR	56	67	58	69	65	73	67	75
	KNN	50	62	71	78	45	59	57	68
	DT	31	48	55	66	45	59	51	63
	SVC	50	62	61	71	65	73	67	75
NASNetLarge	**RF**	44	58	75	81	75	81	73	80
	LDA	47	60	75	81	75	81	64	73
	LR	53	65	100	100	75	81	82	86
	KNN	41	55	75	81	50	62	82	86
	DT	38	53	100	100	50	62	91	93
	SVC	50	62	100	100	75	81	91	93
MobileNetV2	**RF**	56	67	61	71	75	81	64	73
	LDA	41	55	45	59	75	81	73	80
	LR	69	77	68	76	75	81	91	93
	KNN	47	60	55	66	100	100	55	66
	DT	38	53	42	56	50	62	73	80
	SVC	62	72	55	66	75	81	91	93

It has been observed that NASNetLarge achieved a coherent performance at 10x and 40x magnification factor but not capable to provide significant performance at 4x and 20x magnification factor. Moreover, magnification factor has great impact on the performance of the classification model. This happen because the field of view and light intensity decreases as magnification increases. The features extracted at higher magnification are more distinctive but at much higher magnification the field of view is limited and also the intensity of light decreases due to which extraction of distinctive features become complex. Therefore, an appropriate trade-off between the field of view and resolution can only provide significant features and highest performance.

5.3.2 State-of-the-art Comparison

This study demonstrates the potential of transfer learning approach deploying a pre-trained network as feature extractor for breast cancer grade classification. Wherein, the performances of proposed models are evaluated on the Databiox dataset for magnification dependent and independent classification. In case of magnification independent categorization, the ResNet50 + LR achieved considerable performance because of skip connections in the ResNet50 architecture (75% accuracy and 79% AUC). Despite the significant performance of ResNet50 + LR this model did not perform up to the mark.

Table 5.3 Performance comparison among the proposed model and the existing state-of-the-art classifiers for magnification independent and magnification dependent grade classification

Study	**Approach**	**Magnification Independent**	**Magnification Dependent**			
			(4X)	(10X)	(20X)	(40X)
P. H. Zavareh et al. [101]	VGG-16	**72%**	----	----	----	----
Present Work	ResNet50 + LR	75%	----	----	----	----
	VGG19 + SVC	----	79%	----	----	----
	NASNetLarge + SVC	----	----	**100%**	----	**91%**
	MobileNetV2 + KNN	----	----	----	**100%**	----

This all happened due to the consideration of images of each magnification factors together in model training. Consequently, it became complicated for the model to extract discriminative but equivalent features from the images of different

magnification factors. In conjunction to confirm the effect of image different magnification levels on the performance of the model, we implemented the similar protocol for the magnification dependent classification of breast cancer grades. Wherein, the images of same magnification are considered separately. It has been analyzed from the results that the performance of the models shoots up drastically when applied on images of each magnification separately. VGG19, and MobileNetV2 achieved highest accuracy of 79% and 100%, respectively, for 4x and 20x magnification factor. While NASNetLarge provided outstanding performance of 100% and 91%, respectively, for 10x and 40x magnification factor as shown in Table 5.3.

P. H. Zavareh al. [101] also performed magnification independent grade classification using transfer learning strategy without fine-tuning by considering VGG16 convolutional neural network and achieved a validation accuracy of 88% and evaluation accuracy of 72%. We have surpassed the performance of the model proposed by P. H. Zavareh et al. for magnification independent classification. Since this is the only study that has been reported on the Databiox dataset till data so we are not capable to perform much comparison on our result at this stage.

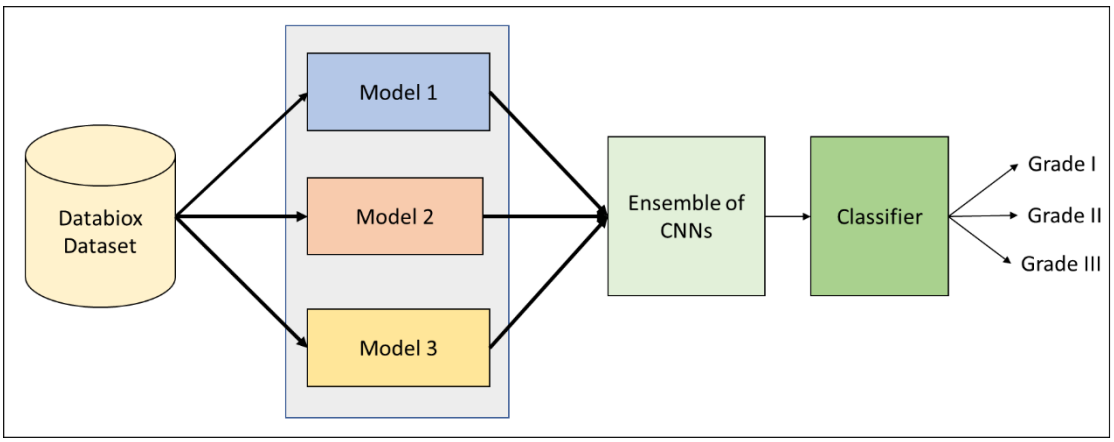
5.4 Conclusion

In this study, the potential of the transfer learning approach is analyzed for magnification independent and magnification dependent grade classification. It is observed through this study that the proposed model provides a remarkable performance in magnification dependent grade classification as compared to magnification independent classification. Among different combinations of classifiers, VGG19 + SVC, and MobileNetV2 + KNN provide the best result for 4x, and 20x magnification factors. While NASNetLarge achieves the highest classification accuracy for 10x and 20x magnification factors. The major rationale behind the insignificant performance for magnification independent classification was the complication in extracting discernible features. Because the extraction of discriminative but equivalent features from the images of different magnification factor becomes complicated in the case of magnification independent classification. Therefore, it is necessary to develop a new technique of data pre-processing that can remove this complication from the images of different magnifications. Moreover, the implementation of advanced data augmentation techniques and different methods of weight initialization can enhance the performance.

6

An Invasive Ductal Carcinomas Breast Cancer Grade Classification Using an Ensemble of Convolutional Neural Networks

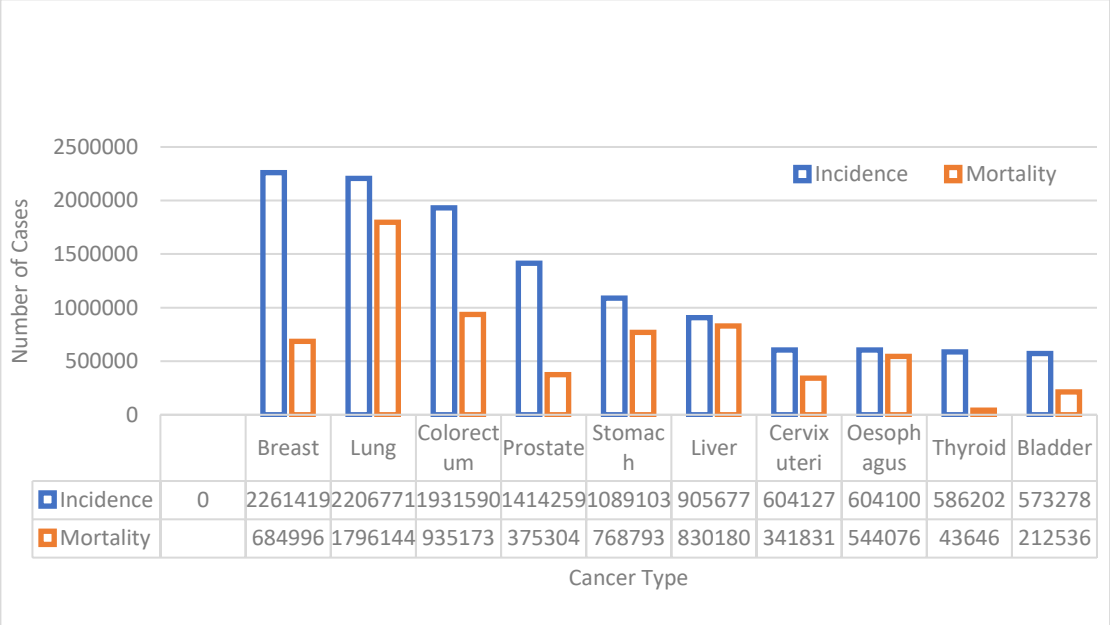
"The key to artificial intelligence has always been the representation"
—Jeff Hawkins



In this chapter, the potential of the pre-trained Convolutional Neural Networks (CNNs) (i.e., EfficientNetV2L, ResNet152V2, DenseNet201) as a single and an ensemble of CNNs has been thoroughly explored. Data augmentation has been used to avoid the issue of data scarcity and data imbalances. The performance of the best model has been compared for three different dataset samples of Databiox (i.e., 1200, 1400, and 1600 images) to determine the implications of data augmentation. Further, the effects of the number of epochs have also analysed to ensure the coherency of the most optimal model. The experimental result analysis revealed that the proposed ensemble model outperformed the existing state-of-the-art techniques related to classifying IDC breast cancer grades of the Databiox dataset. The proposed ensemble model of CNNs achieved 94% classification accuracy and attained a significant area under the ROC curves for grades 1, 2, and 3, i.e., 96%, 94%, and 96%, respectively.

6.1 Introduction

Breast Cancer is a prevalent type of cancer and has become the fifth leading cause of cancer-related deaths among women in recent years [151]. The most common signs and symptoms of BC include heaviness, stiffness, pain, redness or swelling in the breast, as well as abnormalities like shrinking, blood discharge and nipple erosion. Generally, cancer tissues originate either in ducts or lobules of the breast and form a lump, which is referred to as a tumour. As per the position and nature of the mass of cancer tissues, the tumour is categorized into two types: Ductal Carcinoma In Situ (DCIS) and Invasive Ductal carcinoma (IDC). Invasive cancer spreads to other parts of the body, while in-situ cancer does not invade other body parts [152]. Invasive cancer is life-threatening, and 80% of breast cancer deaths are caused by it. Breast cancer has been observed in younger age groups as well, and the incidence has been increasing rapidly. In India, a significant proportion of the younger population in their thirties and forties are affected by breast cancer. According to the International Association of Cancer Registries (IACR) and Global Initiative for Cancer Registry Development, cancer incidence cases and mortality across the world in both sexes and all age groups are depicted in Figure 6.1. In the year 2022, the estimated number of new cases and deaths in both females and males due to breast cancer was 22,61,419 and 6,84,996, respectively [153].



		Breast	Lung	Colorectum	Prostate	Stomach	Liver	Cervix uteri	Oesophagus	Thyroid	Bladder
Incidence	0	2261419	2206771	1931590	1414259	1089103	905677	604127	604100	586202	573278
Mortality		684996	1796144	935173	375304	768793	830180	341831	544076	43646	212536

Figure 6.1 Estimated number of incident cancer cases and deaths worldwide due to various types of cancers in the year 2022.

The diagnosis of breast cancer involves various screening methods such as thermography, mammography, ultrasound scan, positron emission, magnetic resonance imaging, and breast-specific imaging [154]. Microscopy biopsy or surgical incision is another gold standard method used to diagnose the location of the lump and size of the tumour. However, the interpretation and analysis of BC type based on tissue textures, colour, and shape differentiation are challenging due to similar clinical manifestations. Researchers are developing computer-aided diagnosis (CAD) systems based on artificial intelligence techniques to overcome such challenges [155]. CAD systems increase diagnostic accuracy by reducing human error which may further assist radiologists in proper diagnosis. Among the various artificial intelligence-based techniques, Deep neural networks are widely used in developing CAD systems due to their automatic feature extraction, deeper network layers, and representation learning ability. Besides, advances in computer vision and the increased availability of computational power have contributed to the popularity of deep learning techniques. A vast amount of training data, more time, and high computational GPUs are required to train deep neural networks. However, a transfer learning approach can be used to overcome the requirement of large training data. Shallu Sharma and Rajesh Mehra, perform histopathological image classification using both transfer learning and training from scratch techniques [44]. They evaluated the performance of ResNet50, Xception, and Densenet121 pre-trained models as baseline models for feature extraction and achieved significant results.

In the line with this, an ensemble of pre-trained CNNs has been developed which includes EfficientNetV2L [156], ResNet152V2 [157], and DenseNet201[158] to extract potential features to classify IDC breast cancer histopathological grades of a new dataset called databiox [14]. The goal of this paper is to achieve better performance as compared to the state-of-the-art methods for breast IDC grade classification using histopathological images. The experimental work achieved the best results as compared to existing methods on the databiox dataset, as best of the author's knowledge concern.

Contribution

- In this work, an ensemble model comprised of three pre-trained CNNs has been designed to make grading predictions on the Databiox dataset, which consists of histopathological images of IDC-diagnosed patients for grade classification.

Different pre-trained base models have been analyzed for their performance individually and in combination to determine the most optimal and coherent solution for breast cancer grade classification.

- In Databiox, the dataset is imbalanced and the distribution of images among the different grades and the total number of images in each grade is insufficient to train a CNN. It may lead to the problem of biasing towards one particular class with more images. To overcome this issue, data augmentation techniques has been employed to balance the dataset. The performance of the best model has been compared for three different balanced data samples of Databiox (i.e., 1200, 1400, and 1600 images) to ensure the limit of data augmentation.
- Additionally, the implications of the number of epochs have also been demonstrated through this experimental work. The performance of the models has been observed for four different number of epochs which further determines the robustness and coherency of the proposed ensemble model. The performance of the proposed models has been analyzed by utilizing evaluation parameters namely, precision, recall, f1 score, accuracy, ROC curve, and the area under the ROC curve (AUC).

6.2 Methodology

In this work, CNNs have been used as a potential feature extractor and a fully connected neural network as a classifier. The performance of CNNs is enhanced to extract the features effectively, and the extracted features are mapped to corresponding categories using the dense layers of CNNs. Here, histopathological images are analyzed to classify the IDC breast cancer grades with help of individual CNNs and an ensemble of CNNs. Classification in different grades of IDC breast cancer in histopathological images is a challenging task due to the undifferentiable variants in tissue images. To address this challenge, a robust and accurate grade classification ensemble model needs to be developed.

6.2.1 Data pre-processing and augmentation

In Databiox, the dataset is imbalanced and the distribution of images among the different grades and the total number of images in each grade is insufficient to train a CNN. Training CNNs with imbalanced data leads to a biased classification. To overcome this issue, the raw data of 922 images were transformed into a more balanced

dataset of 1200, 1400, and 1600 images using data augmentation techniques [159]. This not only avoids the bias problem for each class but also reduces the data scarcity issues.

To expand the dataset, an augmenter was used in the machine learning library, which is essentially a collection of options for common image-editing tasks, including rotating, shearing and cropping. It is always important to select appropriate data augmentation techniques to avoid any discriminating features or details. A random number generator was used to select values between -5 to 10, forming an augmentation pipeline. This resulted in the creation of a distinct image for each image sent through the pipeline. The original 922 histopathological images from Databiox were increased to 1200, 1400, and 1600 images for this experimental work by rotating with probability= 0.7, maximum left rotation = 10, maximum right rotation = 10, and zooming with probability= 0.3, minimum factor = 1.1, maximum factor = 1.6. After the augmentation process, the produced dataset was split into training and validation with a ratio of 80% and 20% respectively. During training, the validation set was used to validate the performance of the model as well as to control the overfitting problem.

A batch normalization procedure was applied to normalize the selected features [160]. Normalization is a commonly used process in machine learning that aims to scale down the features to make model training less sensitive to the scale of the data. This process helps the model to perform better and maintain training stability. It is an essential step in CAD system designing, as it normalizes the data samples to obtain more consistent data for further processing. Therefore, batch normalization is performed in the pre-processing stage, which changes all sample attributes to a single scale in the range of 0 to 1.

6.2.2 Convolutional Neural Networks (CNNs)

CNNs have indeed become a popular choice for deep learning-based CAD expert systems, owing to their ability to learn spatial hierarchies of features from raw input data. Adding more layers to a CNN can increase its capacity to learn complex features from the data, potentially leading to higher accuracy [161]. Currently, various types of CNNs are available shown in Figure 6.2 and used as efficient feature extractors, and pre-trained CNNs are used to adopt a transfer learning approach when fewer computational resources are available.

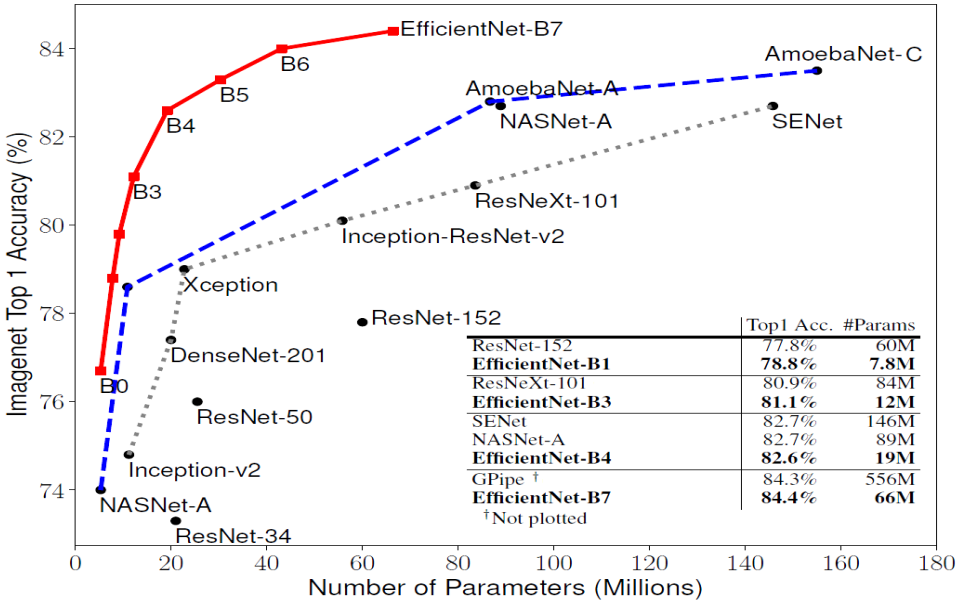


Figure 6.2 Comparative analysis of Top-1 accuracy and model size of different CNN models on ImageNet dataset where EfficientNet outperforms other CNNs significantly [162].

Generally, the transfer learning approach makes the CNN model easy to use, particularly for task-specific classification [163]. Convolutional layers, pooling layers, and fully connected layers are the common types of layers in a CNN model, while different types of CNNs based on the arrangement of different layers, like depth-wise separable convolution layers and skip connection between convolution blocks, have also been proposed. One of the popular CNN architectures is ResNet, which uses a stack of linear separable depth-wise convolutional layers with residual connections. One of the biggest problems of any deep learning network is vanishing and exploding gradients, which restricts us from going much deeper into the network. The core idea of ResNet is based on skip connections, which allow taking activation from one layer and feeding it to a future layer, even if it is much deeper. ResNet has achieved significant success in many computer vision tasks, including the ImageNet competition, where the Residual Network with 152 layers variant (ResNet152) won in 2015.

EfficientNet is a smaller and faster training system launched by Google AI that aims to provide a more efficient approach, as suggested by its name. EfficientNet uses an efficient compound scaling mechanism, gradually increasing model attributes such as depth, width, and resolution, instead of arbitrarily scaling these parameters in CNN design. The recently introduced EfficientNetV2 improves upon EfficientNet in terms of

training time and parameter effectiveness. DenseNet is another CNN architecture that requires fewer parameters than an equivalent traditional CNN, as it does not learn redundant feature maps. Unlike ResNet, DenseNet concatenates the output feature maps of the layer with the incoming feature maps, instead of summing them up.

6.2.3 Ensemble of CNNs

This study aims to compare the performance of individual CNNs with an ensemble of CNNs for the classification of IDC BC grades. The proposed model uses an ensemble of CNN models to achieve better performance. Initially, the CNN model was trained using the ImageNet dataset. Later, the pre-trained CNN model's first few layers were frozen, and the last layer's output neurons were adjusted based on the number of classes in the target problem, and an ensemble of CNNs as shown in Figure 6.3. Individual CNN model performance was observed, and the best-performing individual models were saved for use in the ensemble architecture. Later, the best accuracy-providing models were combined to create the ensemble.

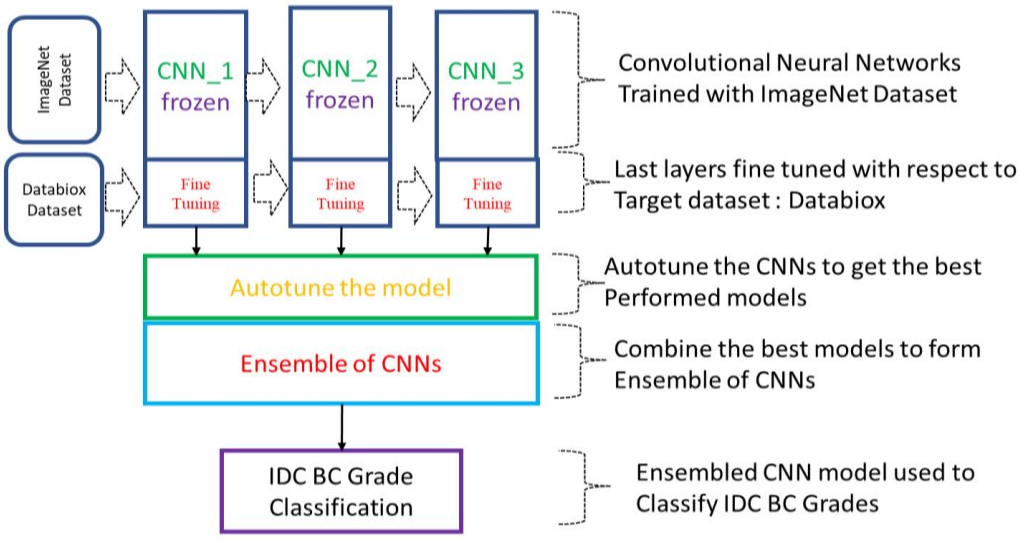


Figure 6.3 Ensemble of CNNs model to classify IDC breast cancer grades

6.3 Results and discussion

In this work, the IDC breast cancer grade classification task was performed on the Databiox dataset after applying pre-processing, and data augmentation techniques. Three different augmented data samples were produced and used, consisting of 1200, 1400, and 1600 images from the original dataset of 922 images. For instance, in the case of 1200 histological images, 960 were used for training and 240 for validation.

Similarly, for the 1400 histopathological images, 1120 were used for training and 280 for validation. Finally, for the 1600 images, 1280 were used for training and 320 for validation purposes. The data were split into an 80%-20% ratio, and the Google Co-lab platform was used to perform the experimental work. To evaluate the model's performance, various metrics such as confusion matrix, precision, accuracy, the ROC curve, recall, F1 score, and AUC were calculated using a python workflow.

Tables 6.1, 6.2, and 6.3 represent the experimental results produced by individual models (via EfficientNetV2L, ResNet152V2, and DenseNet201), and the proposed ensemble of CNN models. The performance of all the models is evaluated for three different data sample sizes (1200, 1400 and 1600) at four different numbers of epochs (5, 10, 15 and 20). From Table 6.1, it has been observed that the performance of the models improved when the number of epochs increased from 5 to 20. It shows that the models get trained well with a higher number of epochs. The performance of the individual model was not consistent; however, the proposed ensemble model consistently achieved the best performance from all the models in comparison. Specifically, the proposed ensemble model achieved a maximum accuracy of 85%, a recall of 86%, an F1 score of 84% and a precision of 87% for 20 epochs with a sample size of 1200. These results are significantly higher than the 72% accuracy obtained by P. H. Zavareh et al. on the same dataset [101].

Table 6.1 Classification report of CNN models (Data augmentation performed and obtained 1200 images, in which 960 training images, and 240 validation images).

Number of epochs = 5					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.78	0.74	0.76	0.71
	Grade 1	0.72	0.62	0.67	
	Grade 2	0.66	0.79	0.72	
Model 2: ResNet152V2	Grade 0	0.35	0.82	0.49	0.45
	Grade 1	0.66	0.24	0.36	
	Grade 2	0.55	0.40	0.46	
Model 3: DenseNet201	Grade 0	0.87	0.33	0.48	0.47
	Grade 1	0.73	0.12	0.20	
	Grade 2	0.41	0.96	0.57	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.78	0.77	0.78	0.75
	Grade 1	0.67	0.82	0.74	
	Grade 2	0.85	0.65	0.73	
Number of epochs = 10					

CNN model	Grade	Precision	Recall	F1-Score	Accuracy
Model 1: EfficientNetV2L	Grade 0	0.70	0.87	0.77	0.76
	Grade 1	0.78	0.78	0.78	
	Grade 2	0.80	0.67	0.73	
Model 2: ResNet152V2	Grade 0	0.66	0.61	0.63	0.61
	Grade 1	0.65	0.53	0.58	
	Grade 2	0.55	0.69	0.61	
Model 3: DenseNet201	Grade 0	0.56	0.75	0.64	0.69
	Grade 1	0.77	0.57	0.66	
	Grade 2	0.75	0.78	0.76	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.78	0.80	0.79	0.78
	Grade 1	0.74	0.84	0.79	
	Grade 2	0.84	0.69	0.76	
Number of epochs = 15					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.00	0.00	0.00	0.37
	Grade 1	0.38	0.83	0.52	
	Grade 2	0.34	0.16	0.22	
Model 2: ResNet152V2	Grade 0	0.28	1.00	0.43	0.28
	Grade 1	0.00	0.00	0.00	
	Grade 2	0.00	0.00	0.00	
Model 3: DenseNet201	Grade 0	0.00	0.00	0.00	0.34
	Grade 1	0.00	0.00	0.00	
	Grade 2	0.34	0.99	0.51	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.82	0.76	0.79	0.78
	Grade 1	0.76	0.80	0.78	
	Grade 2	0.78	0.78	0.78	
Number of epochs = 20					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.74	0.83	0.79	0.82
	Grade 1	0.83	0.83	0.83	
	Grade 2	0.87	0.79	0.83	
Model 2: ResNet152V2	Grade 0	0.64	0.55	0.59	0.62
	Grade 1	0.66	0.72	0.69	
	Grade 2	0.55	0.56	0.55	
Model 3: DenseNet201	Grade 0	0.66	0.67	0.66	0.69
	Grade 1	0.66	0.80	0.72	
	Grade 2	0.83	0.56	0.67	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	**Grade 0**	**0.82**	**0.88**	**0.85**	**0.85**
	Grade 1	**0.85**	**0.87**	**0.86**	
	Grade 2	**0.89**	**0.80**	**0.84**	

Table 6.2 Classification report of CNN models (Data augmentation performed and obtained 1400 images, in which 1120 training images, and 280 validation images).

Number of epochs = 5					
CNN model	**/Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.64	0.94	0.76	0.79
	Grade 1	0.82	0.73	0.78	
	Grade 2	0.95	0.74	0.83	
Model 2: ResNet152V2	Grade 0	0.38	0.72	0.49	0.45
	Grade 1	0.51	0.38	0.43	
	Grade 2	0.52	0.33	0.40	
Model 3: DenseNet201	Grade 0	0.48	0.51	0.49	0.37
	Grade 1	0.26	0.11	0.15	
	Grade 2	0.35	0.56	0.43	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.89	0.80	0.84	0.81
	Grade 1	0.76	0.83	0.80	
	Grade 2	0.80	0.78	0.79	
Number of epochs = 10					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.84	0.93	0.88	0.85
	Grade 1	0.88	0.80	0.84	
	Grade 2	0.84	0.86	0.85	
Model 2: ResNet152V2	Grade 0	0.94	0.21	0.34	0.53
	Grade 1	0.45	0.81	0.58	
	Grade 2	0.65	0.45	0.53	
Model 3: DenseNet201	Grade 0	0.54	0.77	0.64	0.62
	Grade 1	0.85	0.36	0.50	
	Grade 2	0.61	0.81	0.70	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.87	0.92	0.89	0.88
	Grade 1	0.86	0.84	0.85	
	Grade 2	0.90	0.88	0.89	
Number of epochs = 15					
CNN model	**/Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.96	0.93	0.94	0.88
	Grade 1	0.84	0.87	0.86	
	Grade 2	0.86	0.84	0.85	
Model 2: ResNet152V2	Grade 0	0.59	0.70	0.64	0.59
	Grade 1	0.67	0.45	0.54	
	Grade 2	0.53	0.65	0.59	
Model 3: DenseNet201	Grade 0	0.70	0.65	0.67	0.72
	Grade 1	0.65	0.84	0.74	
	Grade 2	0.86	0.63	0.73	

Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.87	0.94	0.91	0.91
	Grade 1	0.91	0.88	0.89	
	Grade 2	0.95	0.92	0.93	
Number of epochs = 20					
CNN model	**/Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.29	0.39	0.34	0.32
	Grade 1	0.35	0.24	0.28	
	Grade 2	0.33	0.36	0.34	
Model 2: ResNet152V2	Grade 0	0.00	0.00	0.00	0.36
	Grade 1	0.00	0.00	0.00	
	Grade 2	0.36	1.00	0.53	
Model 3: DenseNet201	Grade 0	0.23	0.58	0.33	0.27
	Grade 1	0.32	0.31	0.32	
	Grade 2	0.00	0.00	0.00	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	**0.97**	**0.93**	**0.95**	**0.94**
	Grade 1	**0.90**	**0.95**	**0.92**	
	Grade 2	**0.97**	**0.93**	**0.95**	

Based on the results available in Table 6.2, it is clear that the individual model did not exhibit consistent performance but the ensemble of CNNs produced stable results. As the number of epochs increased from 5 to 20, the accuracy of the ensemble model also increased from 88% to 94%, with a sample size of 1400. It is evident from the results that the representation learning ability of the model gets enhanced with the large dataset. However, the sensitivity of all the models except the ensemble model gets influenced towards one particular class when the sample size gets increased from 1200 to 1400 and eventually reduced their performance. These results significantly outperformed those achieved by S. Talpur et al. [102] and R. Sujatha et al. [103], who reported an accuracy of 92.81% and 92.64%, respectively.

Table 6.3 Classification report of CNN models (Data augmentation performed and obtained 1600 images, 1280 training images, and 320 validation images).

Number of epochs = 5					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.59	0.90	0.71	0.70
	Grade 1	0.92	0.50	0.64	
	Grade 2	0.68	0.80	0.73	
Model 2: ResNet152V2	Grade 0	0.31	0.90	0.47	0.37
	Grade 1	0.49	0.29	0.36	
	Grade 2	0.55	0.06	0.10	

Model 3: DenseNet201	IDC Grade 0	0.38	0.82	0.52	0.44
	Grade 1	0.45	0.10	0.16	
	Grade 2	0.53	0.60	0.56	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.78	0.72	0.75	0.79
	Grade 1	0.78	0.83	0.81	
	Grade 2	0.82	0.80	0.81	
Number of epochs = 10					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.27	0.03	0.06	0.32
	Grade 1	0.39	0.09	0.15	
	Grade 2	0.32	0.85	0.46	
Model 2: ResNet152V2	Grade 0	0.00	0.00	0.00	0.40
	Grade 1	0.40	1.00	0.57	
	Grade 2	0.00	0.00	0.00	
Model 3: DenseNet201	Grade 0	0.00	0.00	0.00	0.38
	Grade 1	0.39	0.93	0.55	
	Grade 2	0.22	0.02	0.04	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.79	0.89	0.83	0.86
	Grade 1	0.92	0.84	0.87	
	Grade 2	0.86	0.86	0.86	
Number of epochs = 15					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.36	0.04	0.07	0.34
	Grade 1	0.35	0.91	0.51	
	Grade 2	0.21	0.05	0.09	
Model 2: ResNet152V2	Grade 0	0.32	1.00	0.48	0.32
	Grade 1	0.00	0.00.	0.00	
	Grade 2	0.00	0.00	0.00	
Model 3: DenseNet201	Grade 0	0.26	0.56	0.36	0.25
	Grade 1	0.26	0.22	0.24	
	Grade 2	0.00	0.00	0.00	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	0.92	0.81	0.86	0.86
	Grade 1	0.81	0.94	0.87	
	Grade 2	0.88	0.83	0.85	
Number of epochs = 20					
CNN model	**Grade**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Model 1: EfficientNetV2L	Grade 0	0.24	0.64	0.35	0.24
	Grade 1	0.45	0.04	0.07	
	Grade 2	0.19	0.15	0.16	
Model 2: ResNet152V2	Grade 0	0.23	0.14	0.17	0.39
	Grade 1	0.42	0.88	0.57	

	Grade 2	0.00	0.00	0.00	
Model 3: DenseNet201	Grade 0	0.43	0.28	0.34	0.47
	Grade 1	0.48	0.67	0.56	
	Grade 2	0.48	0.39	0.43	
Ensemble of CNN model (Model 1+ Model 2 + Model 3)	Grade 0	**0.86**	**0.85**	**0.86**	**0.87**
	Grade 1	**0.85**	**0.89**	**0.87**	
	Grade 2	**0.89**	**0.84**	**0.87**	

Further, a similar trend is also observed in Table 6.3, where the performance of the individual model is not significant and inconsistent with 1600 samples. On the other hand, the proposed ensemble model is still outperforming in comparison. The performance of the proposed ensemble model increased from 79% to 87% in terms of accuracy with the increment in the number of epochs but decreased with the increment in the sample size from 1400 to 1600. The overfitting of the model with increased sample size is the major rationale behind its degraded performance. To get more insight into the models' performance, confusion matrices and ROC curve analysis have also been performed. Figure 6.4 (a, b and c) illustrates the confusion matrix and ROC curve for the ensemble model over the sample size of 1200, 1400 and 1600, respectively.

It is analyzed from the confusion matrices in Figure 6.4 (a, b and c) that the samples from grade 0 and grade 2 are wrongly predicted as of grade 1 in most of the cases. It is because the grade 1 stage lies in between the stage of grade 0 and grade 2, therefore they share some clinical expressions with both the stages, grade 0 and grade 2 and this leads to a state of confusion for the model. Despite this, the proposed ensemble model is overall the best one with a sample size of 1400 with an AUC of 0.96, 0.94 and 0.96 for grades 0, 1 and 2, respectively (refer, to Table 6.4).

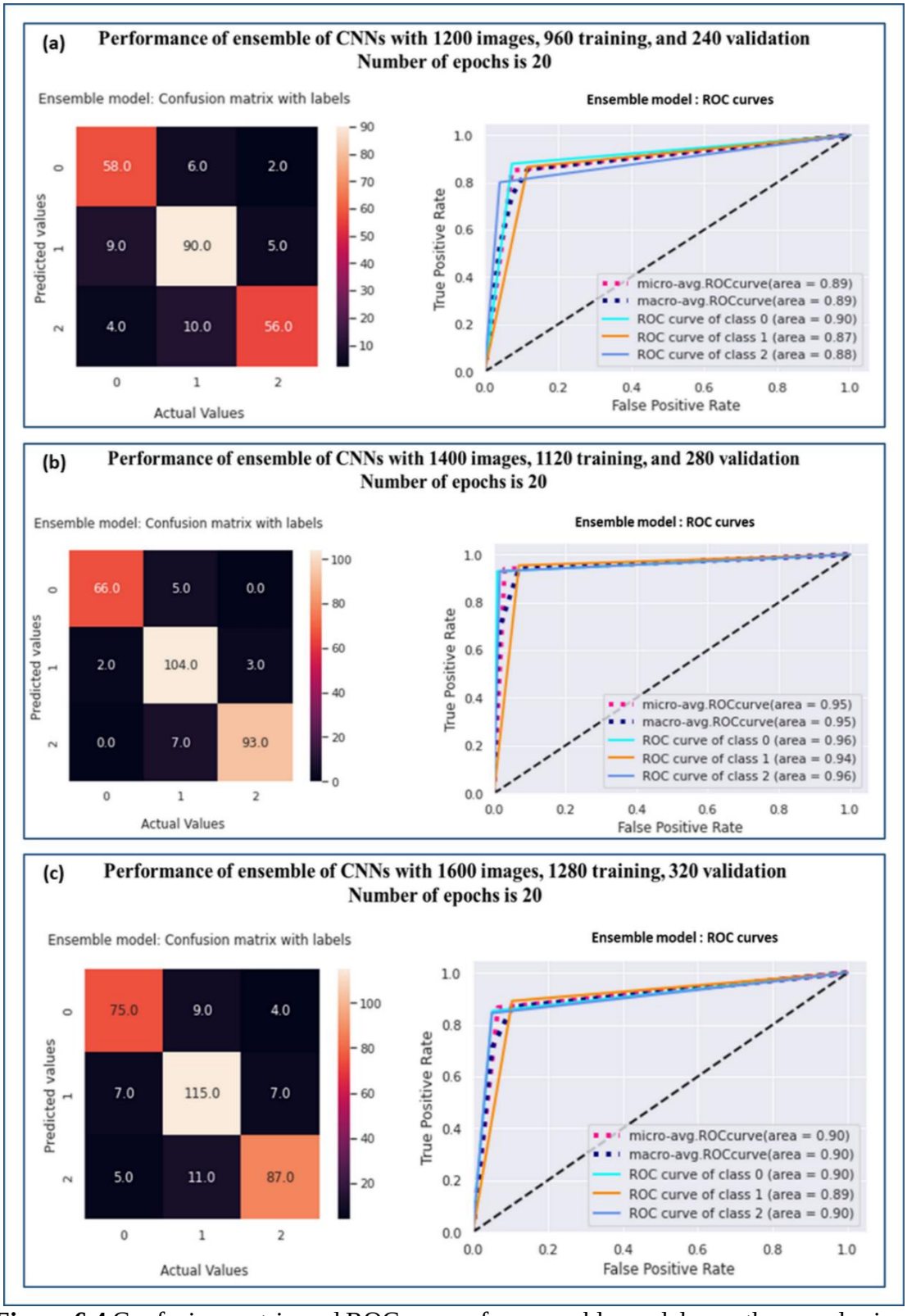


Figure 6.4 Confusion matrix and ROC curves for ensemble model over the sample size (a) 1200, (b) 1400, and (c) 1600.

Table 6.4 Comparative analysis of the proposed ensemble model performance for different sample sizes at 20 epochs

Sample Size	Grade	Precision	Recall	F1-Score	AUC	Accuracy
1200	Grade 0	0.82	0.88	0.85	0.90	0.85
	Grade 1	0.85	0.87	0.86	0.87	
	Grade 2	0.89	0.80	0.84	0.88	
1400	Grade 0	0.97	0.93	0.95	0.96	**0.94**
	Grade 1	0.90	0.95	0.92	0.94	
	Grade 2	0.97	0.93	0.95	0.96	
1600	Grade 0	0.86	0.85	0.86	0.90	0.87
	Grade 1	0.85	0.89	0.87	0.89	
	Grade 2	0.89	0.84	0.87	0.90	

- **Time Complexity**

Time complexity is an important parameter that determines the time taken by an algorithm to execute each instruction of code. It has been observed from the analysis that the computation time of the proposed model is positively correlated to the number of epochs and the sample sizes, as shown in Table 6.5. The time complexity increases with the increment in the number of samples and epochs.

Table 6.5 Time complexity of the proposed ensemble model for different sample sizes and the number of epochs

Samples	Epochs	Time for Data augmentation (In minutes)	Time for training and validation (In minutes)	Total time (In minutes)
1200	5	21	27	48
	10	21	33	54
	15	21	46	67
	20	21	50	71
1400	5	24	30	54
	10	24	39	63
	15	24	48	72
	20	24	53	77
1600	5	28	34	62
	10	28	42	70
	15	28	49	77
	20	28	58	86

6.3.1 Comparison of the State-of-the-art Techniques

This experimental work reveals the potential of data augmentation and an ensemble of CNNs for IDC BC grade classification.

Table 6.6. Performance comparison of the proposed ensemble model with the existing state-of-the-art techniques on the Databiox dataset for the classification of IDC-BC grade images

Reference	Year	Approach	Performance metric
P. H. Zavareh et al. [101]	2021	Transfer learning approach (VGG16 used as feature extractor)	Accuracy of 72%
E. Kumaraswamy et al. [100]	2021	Transfer learning approach pre-trained CNNs: DensNet201 and NASNetMobile used as feature extractors).	Accuracy of 72%. AUC for Grade 1, and Grade 2 is 98% and 75%, respectively with DensNet201 AUC for Grade 3 is 69% with NASNetMobile
R. Sujatha et al. [103]	2022	Transfer learning approaches	DenseNet121 produced the highest accuracy of 92.64%,
S. Talpur et al. [102]	2022	A sequential convolutional neural network is utilized	Accuracy of 92.81%
Present Work	2023	**Proposed Ensemble Model** (EfficientNetV2L+ ResNet152V2 + DensNet201)	**Accuracy of 94%**

The performance of the developed model was assessed on the augmented dataset of Databiox with 1200, 1400, and 1600 histopathological images. In the case of 1200 data size, the proposed ensemble model achieved considerable performance with an accuracy of 85%, AUC of 90% (grade 0), 87% (grade 1), and 88% (grade 2). Despite this significant performance, the model did not reach the existing state-of-the-art technique. This all happened due to the scarcity of data to get the improved performance of the system, we increased the data size to 1200, 1400, and 1600 by applying data augmentation techniques and implemented the same protocol. It has been observed that the performance of the ensemble model drastically increased with 1400 data size in comparison to the individual model. A remarkable result of 94% accuracy is achieved by the proposed ensemble model. P. H. Zavareh et al. achieved an accuracy of 72% for

grade classification using a transfer learning approach, R. Sujatha et al. achieved an accuracy of 92.64% by employing a transfer learning approach with DenseNet121, and S. Talpur et al. achieved an accuracy of 92.81% with the sequential CNNs. The proposed ensemble model surpassed the existing state-of-the-art techniques which have been reported on the Databiox dataset to date. A comparative analysis of the existing state-of-the-art and the proposed ensemble model is illustrated in Table 6.6.

6.4 Conclusion

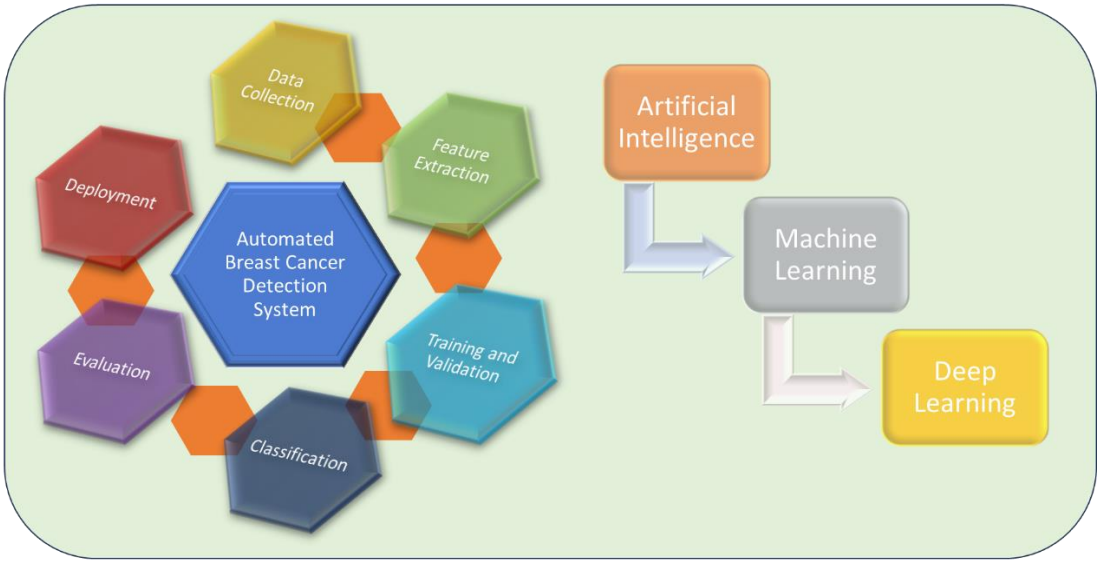
The present work proposes an ensemble model that improves the classification of IDC-BC grade and investigates the impact of sample size and several training epochs. By implementing various data augmentation approaches, the sample size is increased and analyzed for three sample sizes:1200, 1400, and 1600. The obtained results reveal that there is always a trade-off between the model's performance and sample size. Training the model on a largely augmented dataset may introduce the error of generalization, which ultimately leads to an overfitting problem. Therefore, the dataset should be sufficiently large through data augmentation techniques to ensure the optimal performance of the model. Furthermore, our study concludes that the ensemble of CNNs consistently achieved the most stable and robust performance in contrast to the individual models.

7

Conclusion and Future Scope

"The past, like the future, is indefinite and exists only as a spectrum of possibilities"

-Stephen Hawking



This chapter summarizes the prominent features of efficient machine learning approaches employed for the implementation of automated classification system for breast cancer grade classification through different ways described in brief. Additionally, the chapter highlights the future aspects of the presented work.

7.1 Conclusion

The primary goal of this work is to implement an efficient machine learning approach for the classification of histopathological grades for an early-stage diagnosis of breast cancer. The classification accuracy is the main concern and particularly targeted to improve the accuracy of the model throughout this experimentation. Recent data from the Global Cancer Observatory highlights a persistent increase in both the incidence and mortality rates of breast cancer. This alarming trend is linked to various challenges within the diagnosis system that need urgent attention. In this connection, the pathologists prompt us to develop an efficient machine learning algorithm for an automated CAD system for breast cancer classification. Therefore, conventional feature extractors, machine learning, and deep learning approaches are used to analyze the histopathological images. Additionally, deep learning strategies has been intellectualized by the mean of advanced transfer learning approaches for improving the breast cancer classification system.

In this context, we have deployed a deep learning approach to perform the classification of IDC grades using CNN as a feature generator, and SoftMax function in the last layer of classifier. This investigation explores the impact of segmentation techniques on the classification accuracy. The experimental results revealed that the entire images produced significantly better results in contrast to the segmented images when incorporated with CNN. This due to the fact that segmented images have lesser discernible features to distinguish among different grades. However, the transfer learning approaches provide better classification accuracies with the entire histopathological images by deploying the Xception as a baseline model and produced accuracies of 71%, 91%, 91%, and 91% at 4x, 10x, 20x and 40x, respectively. On the other hand, another approach of where deeper pre-trained model is utilized for feature extraction task and RF classifier in the last stage. From the experimental results it has been noticed that a fine-tuned CNN with RF classifier produced significantly better results in contrast to a fully trained network. In addition, the fine-tuned pretrained network robust to the size of the dataset splitting, wherein the size of the training data decreased, the performance of the network not decreased.

Afterward, the experiment work implemented for magnification-independent and dependent classification to ascertain the complexity of the problem. The results

demonstrate that the complexity of the classification task gets reduced when we consider the image of each magnification factor separately. The VGG19 with linear SVM has the accuracy of 79% at 4x, MobileNetV2 with KNN produced 100% at 20x, and NASNetLarge with SVM yields the highest accuracy of 100% and 91% at 10x and 40x, respectively. However, the ResNet50 network with LR provides 75% accuracy for magnification-independent classification. It is observed through this work that a remarkable performance in magnification-dependent grade classification is achieved as compared to magnification-independent. The major rationale behind the insignificant performance for magnification-independent classification was the complication in extracting discernible features. This complexity arises due to the challenge of extracting distinctive yet comparable features from images at various magnification levels in the context of magnification-independent classification. Therefore, it is necessary to develop a new technique that can remove this complication from the images of different magnifications.

In the above context, an ensemble of advanced CNNs is deployed to enhance the model performance through data augmentation techniques. The experimental results of the ensemble of CNNs method outperformed in contrast to individual CNNs and provides the highest classification accuracy of 94%. Additionally, it has been investigated that the impact of sample size and the number of training epochs influences model performance, and there is always a trade-off between the model's performance and sample size. When the model trained on a largely augmented dataset introduced the error of generalization, it ultimately led to the overfitting problem. Therefore, the dataset should be sufficiently large through data augmentation techniques to ensure the optimal model performance. Furthermore, our study concludes that the ensemble of CNNs consistently achieved the most stable and robust performance in contrast to the individual models.

7.2 Future Prospects

According to the results of the current experimental work, the machine learning techniques can be effectively applied to the development of an automated system that performs IDC breast cancer grade categorization, and diagnosis. Though the current work highlights the significance of machine learning techniques in the identification of cancer grades, there nevertheless exists scope for advancement of the models. Further

developments including novel and creative algorithms, and innovations will certainly make it feasible in the near future to create models that are more effective and efficient. In addition, the optimal capacity of a network significantly depends on the various factors including data type, size of the dataset, number of data samples of each class, type of pre-processing, data augmentation, and distinct parameters of the CNN network such as number of convolutional layers, pooling techniques, activation functions, layer wise fine-tuning, methods of weight initialization, training, and testing data size, number of iterations, and optimizer, can influence the performance.

Appendix-A

Supplementary Information

"New information makes new and fresh ideas possible"
- Zig Ziglar

In this appendix, we have summarized the supplementary material of chapter 5 that may be useful to the readers in order to validate the obtained results.

Table S.1 A confusion matrix of IDC grade classes- 0, 1 and 2

Confusion Matrix		***Predicted Class***		
		Grade Class 0	**Grade Class 1**	**Grade Class 2**
Actual Class	**Grade Class 0**	GC_{00}	GC_{01}	GC_{02}
	Grade Class 1	GC_{10}	GC_{11}	GC_{12}
	Grade Class 2	GC_{20}	GC_{21}	GC_{22}

For Grade Class 0:

$T_P = GC_{00}$

F_N =sum of the corresponding row cells except T_P, i.e., $GC_{01}+ GC_{02}$

F_P =sum of the corresponding column cells except T_P, i.e., $GC_{10} + GC_{20}$

T_N =sum of the remaining cells, i.e., $GC_{11} + GC_{21} + GC_{12} + GC_{22}$

For Grade Class 1:

$T_P = GC_{11}$

F_N =sum of the corresponding row cells except T_P, i.e., $GC_{10}+ GC_{12}$

F_P =sum of the corresponding column cells except T_P, i.e., $GC_{01} + GC_{21}$

T_N =sum of the remaining, i.e., $GC_{00} + GC_{20} + GC_{02} + GC_{22}$

For Grade Class 2:

$T_P = GC_{22}$

F_N =sum of the corresponding row cells except T_P, i.e., $GC_{20}+ GC_{21}$

F_P =sum of the corresponding column cells except T_P, i.e., $GC_{02} + GC_{12}$

T_N =sum of the remaining, i.e., $GC_{00} + GC_{10} + GC_{01} + GC_{11}$

Table S.2 Performance analysis of pre-trained VGG16 model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-Trained Network Model	**Classifier**	**Class**	**Precision**	**Recall**	**F1 score**	**AUC**	**AP**
VGG16	**RF**	0	**0.82**	**0.45**	**0.58**	**0.70**	**0.40**
		1	**0.47**	**0.81**	**0.60**	**0.63**	
		2	**0.44**	**0.27**	**0.33**	**0.57**	
	LDA	0	0.45	0.50	0.48	0.58	0.30
		1	0.44	0.19	0.27	0.52	
		2	0.28	0.47	0.35	0.51	
	LR	0	0.70	0.35	0.47	0.63	0.32

		1	0.35	0.67	0.46	0.46	
		2	0.33	0.13	0.19	0.52	
	KNN	0	0.48	0.50	0.49	0.60	
		1	0.36	0.43	0.39	0.49	0.31
		2	0.30	0.20	0.24	0.51	
	DT	0	0.50	0.40	0.44	0.59	
		1	0.35	0.43	0.38	0.47	0.32
		2	0.43	0.40	0.41	0.60	
	SVC	0	0.00	0.00	0.00	0.50	
		1	0.36	0.95	0.53	0.48	0.29
		2	0.00	0.00	0.00	0.49	

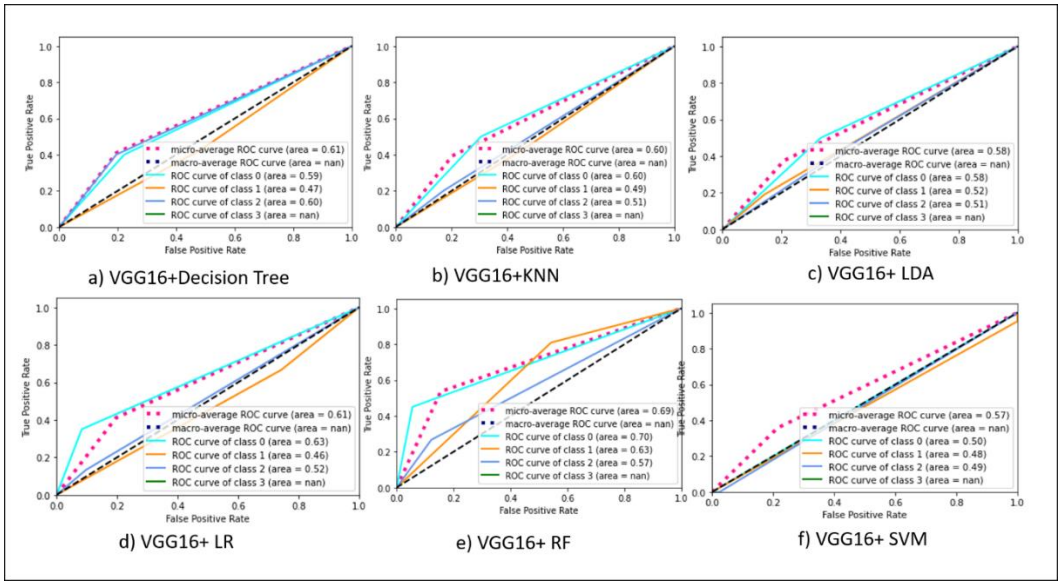


Figure S.1 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained VGG16 model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

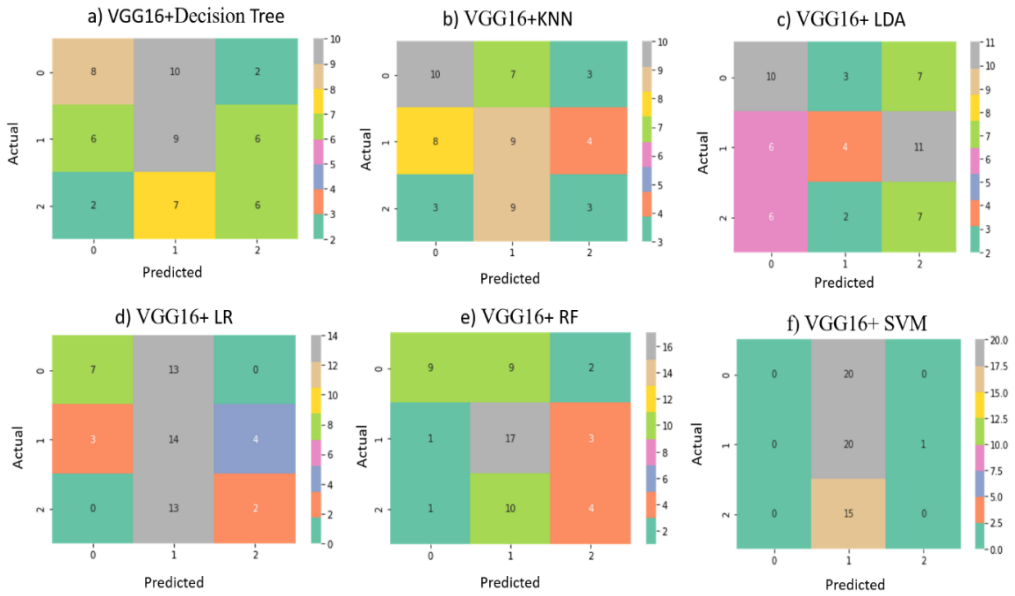


Figure S.2 Confusion matrix of fine-tuned pre-trained VGG16 model with various classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

Table S.3 Performance analysis of pre-trained VGG19 model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-trained network model	Classifier	Class	Precision	Recall	F1 score	AUC	AP
VGG19	**RF**	0	**0.77**	**0.50**	**0.61**	**0.71**	**0.42**
		1	**0.45**	**0.67**	**0.54**	**0.59**	
		2	**0.58**	**0.47**	**0.52**	**0.67**	
	LDA	0	0.50	0.45	0.47	0.60	0.34
		1	0.60	0.43	0.50	0.63	
		2	0.30	0.47	0.37	0.54	
	LR	0	0.40	0.10	0.16	0.51	0.31
		1	0.36	0.71	0.48	0.47	
		2	0.56	0.33	0.42	0.62	
	KNN	0	0.32	0.30	0.31	0.47	0.29
		1	0.38	0.52	0.44	0.50	
		2	0.38	0.20	0.26	0.54	
	DT	0	0.55	0.55	0.55	0.65	0.35
		1	0.45	0.43	0.44	0.56	
		2	0.38	0.40	0.39	0.58	

	SVC	0	0.28	0.22	0.10	0.50	0.32
		1	0.38	0.95	0.55	0.52	
		2	0.75	0.20	0.32	0.59	

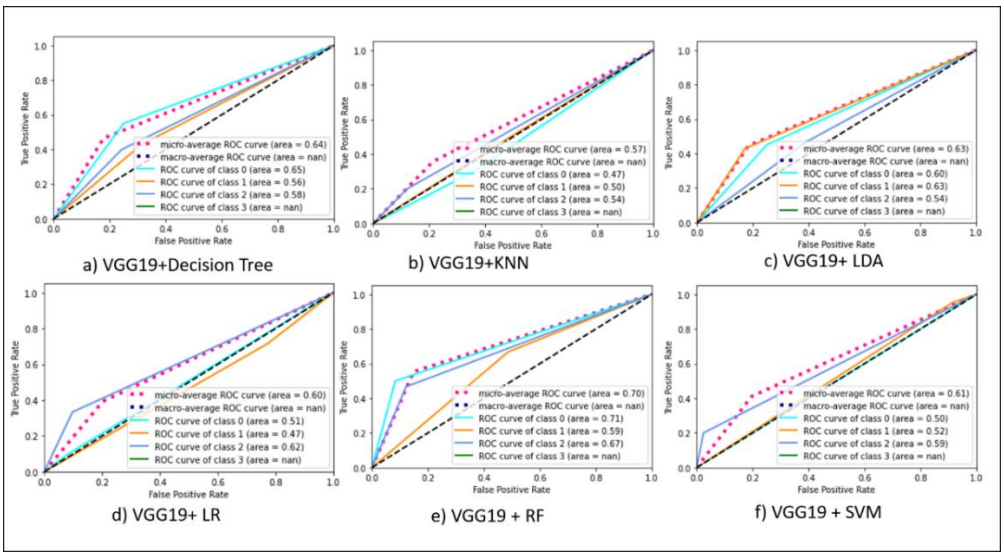


Figure S.3 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained VGG19 model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

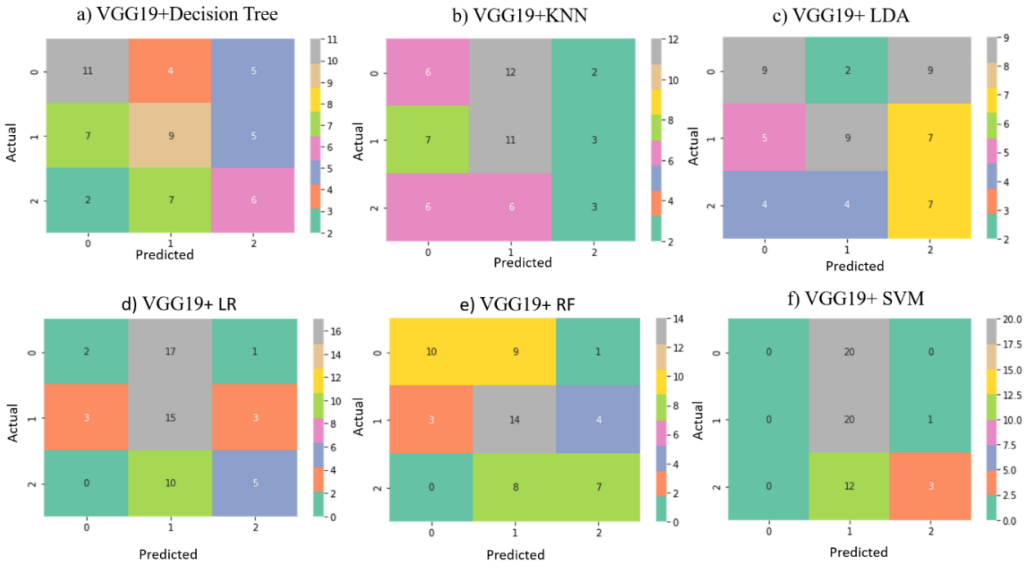


Figure S.4 Confusion matrix of fine-tuned pre-trained VGG19 model with different classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

Table S.4 Performance analysis of pre-trained ResNet50 model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-trained network model	Classifier	Class	Precision	Recall	F1 score	AUC	AP
ResNet50	RF	0	0.88	0.75	0.81	0.85	0.52
		1	0.56	0.86	0.68	0.73	
		2	0.57	0.27	0.36	0.60	
	LDA	0	0.76	0.65	0.70	0.77	0.48
		1	0.60	0.71	0.65	0.71	
		2	0.50	0.47	0.48	0.65	
	LR	0	**0.75**	**0.75**	**0.75**	**0.81**	**0.54**
		1	**0.65**	**0.71**	**0.68**	**0.74**	
		2	**0.62**	**0.53**	**0.57**	**0.71**	
	KNN	0	0.59	0.50	0.54	0.65	0.35
		1	0.43	0.57	0.49	0.56	
		2	0.36	0.27	0.31	0.55	
	DT	0	0.35	0.30	0.32	0.50	0.27
		1	0.29	0.24	0.26	0.45	
		2	0.27	0.40	0.32	0.50	
	SVC	0	0.68	0.75	0.71	0.78	0.50
		1	0.67	0.67	0.67	0.64	
		2	0.54	0.54	0.47	0.64	

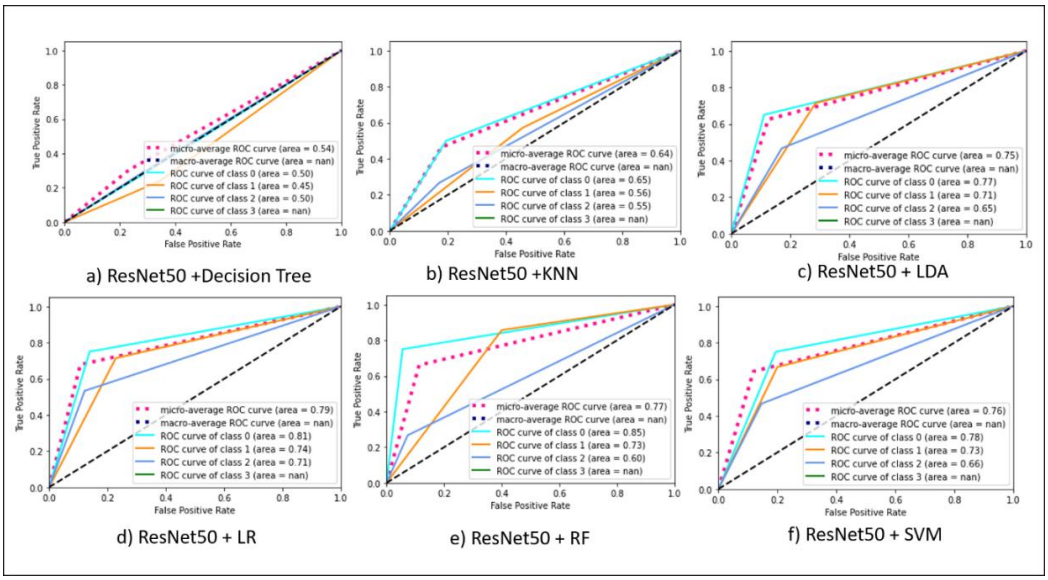


Figure S.5 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained ResNet50 model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

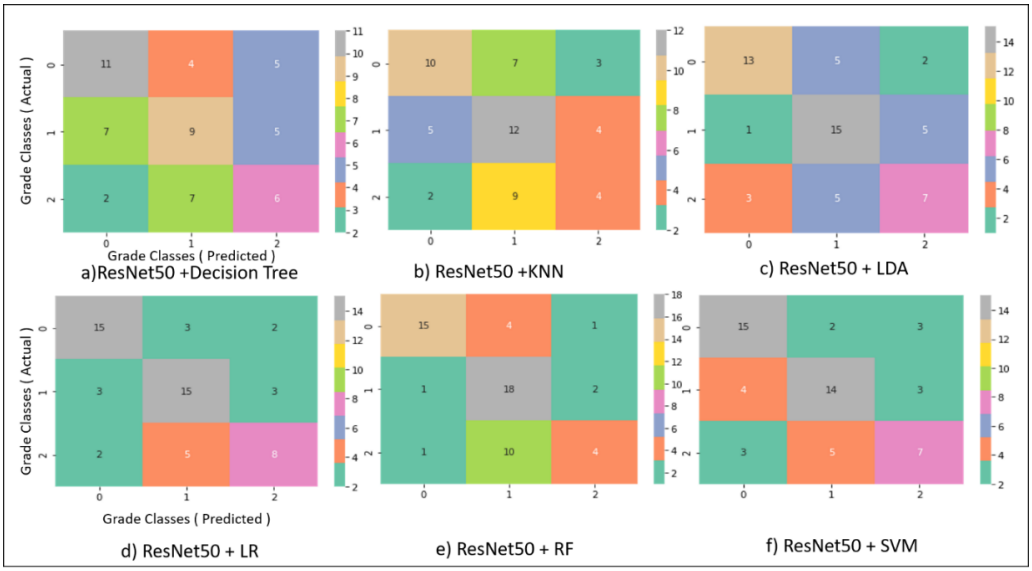


Figure S.6 Confusion matrix of fine-tuned pre-trained ResNet50 model with various classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers

Table S.5 Performance analysis of pre-trained DenseNet201 model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-Trained Network Model	Classifier	Category (Grade /Class)	Precision	Recall	F1 score	AUC	AP
DenseNet201	**RF**	0	0.76	0.47	0.58	0.71	0.50
		1	0.58	0.84	0.69	0.70	
		2	0.70	0.51	0.59	0.70	
	LDA	0	0.60	0.53	0.56	0.70	0.49
		1	0.61	0.66	0.63	0.68	
		2	0.69	0.67	0.68	0.76	
	LR	0	0.54	0.55	0.55	0.70	0.49
		1	0.65	0.58	0.62	0.68	
		2	0.60	0.67	0.64	0.73	
	KNN	0	0.62	0.50	0.55	0.69	0.44
		1	0.53	0.78	0.64	0.66	
		2	0.67	0.38	0.49	0.65	
	DT	0	0.42	0.60	0.50	0.66	0.35
		1	0.49	0.48	0.49	0.56	
		2	0.50	0.36	0.42	0.59	
	SVC	0	0.59	0.62	0.60	**0.74**	**0.51**
		1	0.68	0.62	0.65	**0.71**	
		2	0.68	0.72	0.70	**0.78**	

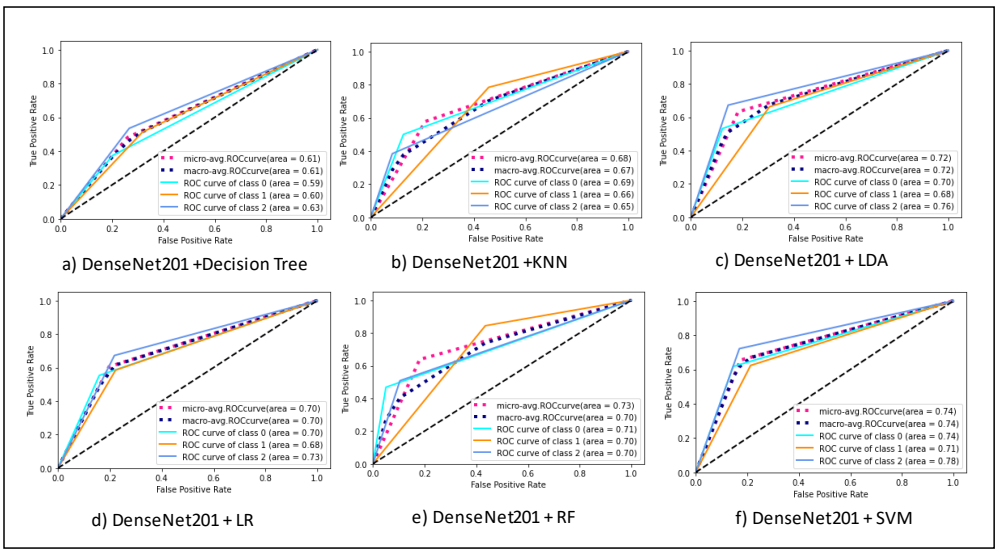


Figure S.7 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained DenseNet201 model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

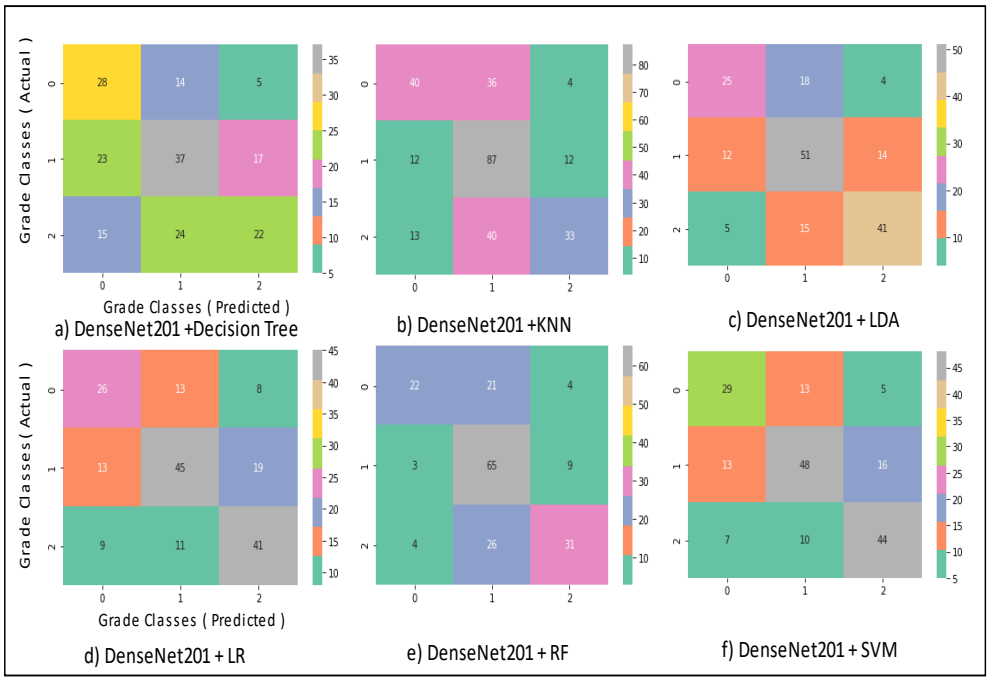


Figure S.8 Confusion matrix of fine-tuned pre-trained DenseNet201 model with various classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

Table S.6 Performance analysis of pre-trained NASNetLarge model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-Trained Network Model	Classifier	Class	Precision	Recall	F1 score	AUC	AP
NASNetLarge	RF	0	**0.76**	**0.47**	**0.58**	**0.65**	**0.44**
		1	**0.58**	**0.84**	**0.69**	**0.65**	
		2	**0.70**	**0.51**	**0.59**	**0.69**	
	LDA	0	0.62	0.38	0.47	0.65	0.41
		1	0.53	0.70	0.60	0.63	
		2	0.54	0.48	0.50	0.64	
	LR	0	0.50	0.47	0.48	0.65	0.42
		1	0.60	0.62	0.61	0.66	
		2	0.54	0.54	0.54	0.66	
	KNN	0	0.34	0.57	0.43	0.60	0.34
		1	0.54	0.58	0.56	0.61	
		2	0.57	0.20	0.29	0.56	
	DT	0	0.41	0.38	0.40	0.60	0.34
		1	0.49	0.45	0.47	0.56	
		2	0.42	0.48	0.45	0.58	
	SVC	0	**0.53**	**0.57**	**0.55**	**0.70**	**0.44**
		1	**0.62**	**0.65**	**0.64**	**0.69**	
		2	**0.57**	**0.51**	**0.54**	**0.66**	

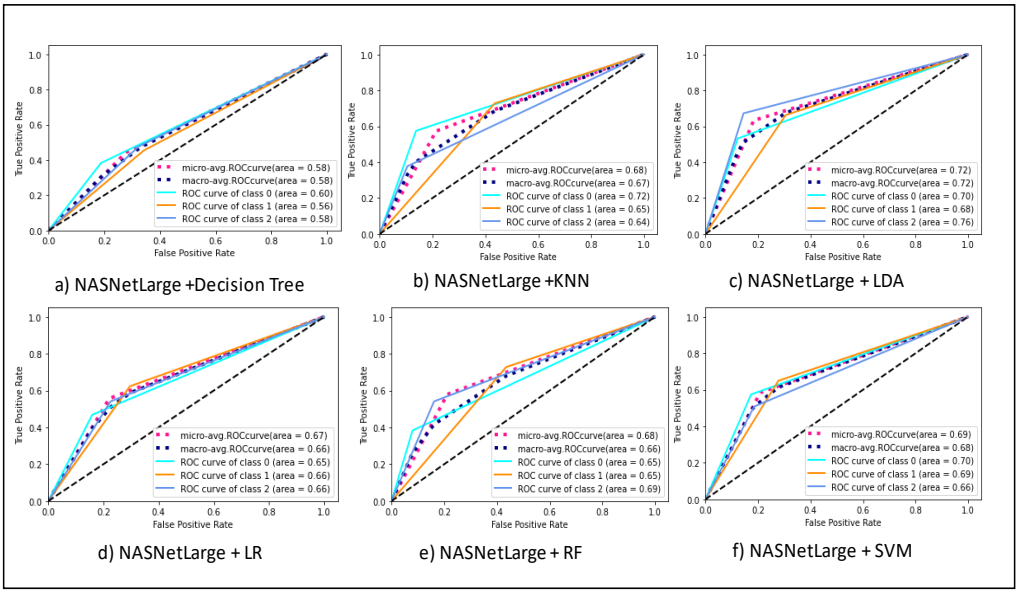


Figure S.9 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained NASNetLarge model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

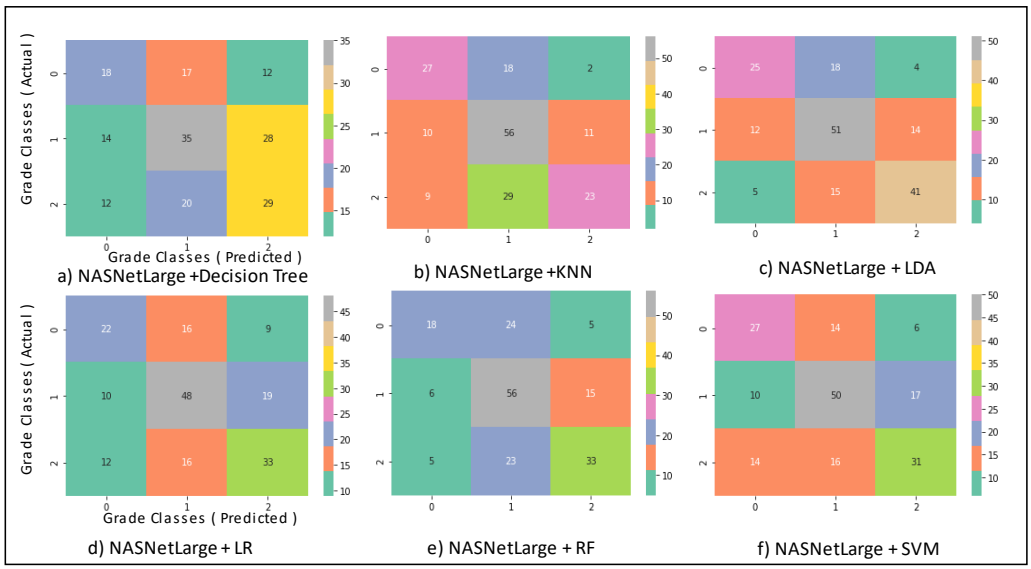


Figure S.10 Confusion matrix of fine-tuned pre-trained NASNetLarge model with various classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

Table S.7 Performance analysis of pre-trained MobileNetV2 model as feature extractor along with various conventional classifiers for breast cancer IDC grade classification.

Pre-Trained Network Model	Classifier	Class	Precision	Recall	F1 score	AUC	AP
MobileNetV2	RF	0	0.67	0.34	0.45	0.64	0.45
		1	0.55	0.79	0.65	0.67	
		2	0.63	0.52	0.57	0.69	
	LDA	0	0.60	0.53	0.56	0.65	0.45
		1	0.61	0.66	0.63	0.68	
		2	0.69	0.67	0.68	0.72	
	LR	0	0.50	0.43	0.46	0.64	0.43
		1	0.59	0.66	0.62	0.66	
		2	0.59	0.56	0.57	0.68	
	KNN	0	0.56	0.38	0.46	0.64	0.42
		1	0.52	0.70	0.60	0.62	
		2	0.63	0.51	0.56	0.68	
	DT	0	0.29	0.32	0.30	0.53	0.32
		1	0.48	0.51	0.49	0.56	
		2	0.44	0.38	0.41	0.57	
	SVC	0	0.51	0.47	0.49	0.74	0.44
		1	0.59	0.66	0.62	0.71	
		2	0.60	0.54	0.57	0.78	

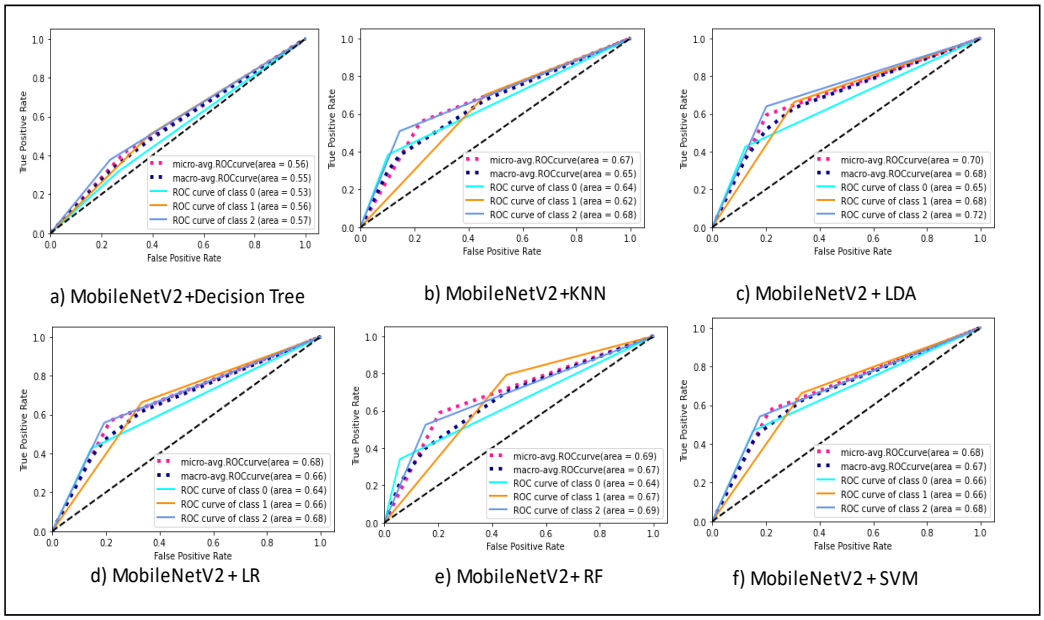


Figure S.11 ROC analysis for IDC breast cancer grade classification by using fine-tuning of pre-trained MobileNetV2 model with a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

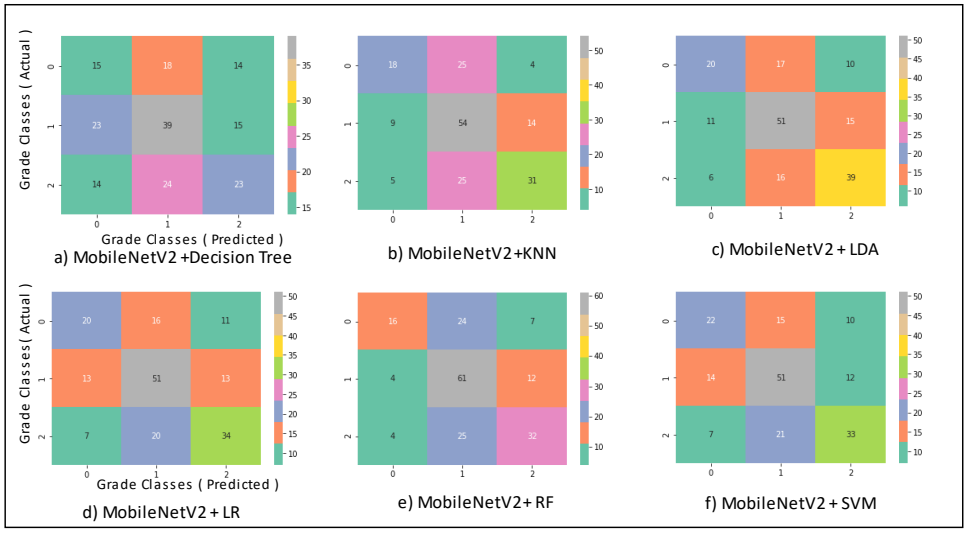


Figure S.12 Confusion matrix of fine-tuned pre-trained MobileNetV2 model with various classifiers such as a) DT, b) KNN, c) LDA, d) LR, e) RF, f) SVM classifiers.

Table S.8 Performance comparison among the best performing pre-trained models i.e., VGG16, VGG19, ResNet50, DenseNet201, NASNetLarge and MobileNetV2 networks for the IDC grade classification.

Pre-Trained Network	**Classifier**	**(Grade /Class)**	**Precision**	**Recall**	**F1 score**	**AUC**	**AP**
VGG16	**RF**	0	0.82	0.45	0.58	0.70	0.40
		1	0.47	0.81	0.60	0.63	
		2	0.44	0.27	0.33	0.57	
VGG19	**RF**	0	0.77	0.50	0.61	0.71	0.42
		1	0.45	0.67	0.54	0.59	
		2	0.58	0.47	0.52	0.67	
ResNet50	**RF**	0	0.88	0.75	0.81	0.85	**0.52**
		1	0.56	0.86	0.68	0.73	
		2	0.57	0.27	0.36	0.60	
	LR	**0**	**0.75**	**0.75**	**0.75**	**0.81**	**0.54**
		1	**0.65**	**0.71**	**0.68**	**0.74**	
		2	**0.62**	**0.53**	**0.57**	**0.71**	
DenseNet201	RF	0	0.76	0.47	0.58	0.71	**0.50**
		1	0.58	0.84	0.69	0.70	
		2	0.70	0.51	0.59	0.70	
	SVC	0	0.59	0.62	0.60	0.74	**0.51**
		1	0.68	0.62	0.65	0.71	
		2	0.68	0.72	0.70	0.78	
NASNetLarge	RF	0	0.76	0.47	0.58	0.65	0.44
		1	0.58	0.84	0.69	0.65	
		2	0.70	0.51	0.59	0.69	
	SVC	0	0.53	0.57	0.55	0.70	0.44
		1	0.62	0.65	0.64	0.69	
		2	0.57	0.51	0.54	0.66	
MobileNetV2	RF	0	0.67	0.34	0.45	0.64	0.45
		1	0.55	0.79	0.65	0.67	
		2	0.63	0.52	0.57	0.69	
	LDA	0	0.60	0.53	0.56	0.65	0.45
		1	0.61	0.66	0.63	0.68	
		2	0.69	0.67	0.68	0.72	

References

[1] C. A. Almeida and S. A. Barry, *Cancer: basic science and clinical aspects*. John Wiley & Sons, 2011.

[2] T. N. Seyfried and L. C. Huysentruyt, "On the origin of cancer metastasis," *Critical Reviews™ in Oncogenesis,* vol. 18, no. 1-2, 2013.

[3] J. E. Visvader, "Cells of origin in cancer," *Nature,* vol. 469, no. 7330, pp. 314-322, 2011.

[4] Y.-S. Sun *et al.*, "Risk factors and preventions of breast cancer," *International journal of biological sciences,* vol. 13, no. 11, p. 1387, 2017.

[5] "Global cancer observatory." https://gco.iarc.fr/today/online-analysis-multi-bars (accessed May 25, 2023).

[6] P. Hamet and J. Tremblay, "Artificial intelligence in medicine," *Metabolism,* vol. 69, pp. S36-S40, 2017.

[7] M. Ataollahi, J. Sharifi, M. Paknahad, and A. Paknahad, "Breast cancer and associated factors: a review," *Journal of medicine and life,* vol. 8, no. Spec Iss 4, p. 6, 2015.

[8] Z. Chen *et al.*, "Invasive lobular carcinoma of the breast: a special histological type compared with invasive ductal carcinoma," *PloS one,* vol. 12, no. 9, p. e0182397, 2017.

[9] A. Karellas and S. Vedantham, "Breast cancer imaging: a perspective for the next decade," *Medical physics,* vol. 35, no. 11, pp. 4878-4897, 2008.

[10] S. V. Sree, E. Y.-K. Ng, R. U. Acharya, and O. Faust, "Breast imaging: a survey," *World journal of clinical oncology,* vol. 2, no. 4, p. 171, 2011.

[11] H. A. Alturkistani, F. M. Tashkandi, and Z. M. Mohammedsaleh, "Histological stains: a literature review and case study," *Global journal of health science,* vol. 8, no. 3, p. 72, 2016.

[12] J. F. Mccarthy *et al.*, "Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management," *Annals of the New York Academy of Sciences,* vol. 1020, no. 1, pp. 239-262, 2004.

[13] J. Yanase and E. Triantaphyllou, "A systematic survey of computer-aided diagnosis in medicine: Past and present developments," *Expert Systems with Applications,* vol. 138, p. 112821, 2019.

[14] H. Bolhasani, E. Amjadi, M. Tabatabaeian, and S. J. Jassbi, "A histopathological image dataset for grading breast invasive ductal carcinomas," *Informatics in Medicine Unlocked,* vol. 19, p. 100341, 2020.

[15] M. De Bruijne, "Machine learning approaches in medical image analysis: From detection to diagnosis," vol. 33, ed: Elsevier, 2016, pp. 94-97.

[16] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *Ieee Access,* vol. 6, pp. 9375-9389, 2017.

[17] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical image analysis,* vol. 42, pp. 60-88, 2017.

[18] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," *Journal of Applied Science and Technology Trends,* vol. 1, no. 2, pp. 56-70, 2020.

[19] P. Liu, J.-M. Guo, K. Chamnongthai, and H. Prasetyo, "Fusion of color histogram and LBP-based features for texture image retrieval and classification," *Information Sciences,* vol. 390, pp. 95-111, 2017.

[20] T. Löfstedt, P. Brynolfsson, T. Asklund, T. Nyholm, and A. Garpebring, "Gray-level invariant Haralick texture features," *PloS one,* vol. 14, no. 2, p. e0212110, 2019.

[21] A. Hmimid, M. Sayyouri, and H. Qjidaa, "Fast computation of separable two-dimensional discrete invariant moments for image classification," *Pattern Recognition,* vol. 48, no. 2, pp. 509-521, 2015.

[22] L. Hussain, W. Aziz, S. Saeed, S. Rathore, and M. Rafique, "Automated breast cancer detection using machine learning techniques by extracting different feature extracting strategies," in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, 2018: IEEE, pp. 327-331.

[23] E. Christodoulou, J. Ma, G. S. Collins, E. W. Steyerberg, J. Y. Verbakel, and B. Van Calster, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models," *Journal of clinical epidemiology,* vol. 110, pp. 12-22, 2019.

[24] L.-Y. Hu, M.-W. Huang, S.-W. Ke, and C.-F. Tsai, "The distance function effect on k-nearest neighbor classification for medical datasets," *SpringerPlus,* vol. 5, no. 1, pp. 1-9, 2016.

[25] Y.-Y. Song and L. Ying, "Decision tree methods: applications for classification and prediction," *Shanghai archives of psychiatry,* vol. 27, no. 2, p. 130, 2015.

[26] F.-J. Yang, "An implementation of naive bayes classifier," in *2018 International conference on computational science and computational intelligence (CSCI),* 2018: IEEE, pp. 301-306.

[27] M. Awad, R. Khanna, M. Awad, and R. Khanna, "Support vector machines for classification," *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers,* pp. 39-66, 2015.

[28] D. Denisko and M. M. Hoffman, "Classification and interaction in random forests," *Proceedings of the National Academy of Sciences,* vol. 115, no. 8, pp. 1690-1692, 2018.

[29] A. Sharma and K. K. Paliwal, "Linear discriminant analysis for the small sample size problem: an overview," *International Journal of Machine Learning and Cybernetics,* vol. 6, pp. 443-454, 2015.

[30] M. Hosni, I. Abnane, A. Idri, J. M. C. de Gea, and J. L. F. Alemán, "Reviewing ensemble classification methods in breast cancer," *Computer methods and programs in biomedicine,* vol. 177, pp. 89-112, 2019.

[31] J. Vrábel, E. Képeš, P. Pořízka, and J. Kaiser, "Artificial neural networks for classification," *Chemometrics and Numerical Methods in LIBS,* pp. 213-240, 2022.

[32] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook,* pp. 353-374, 2023.

[33] J. Qin, W. Pan, X. Xiang, Y. Tan, and G. Hou, "A biological image classification method based on improved CNN," *Ecological Informatics,* vol. 58, p. 101093, 2020.

[34] M. R. Islam and A. Matin, "Detection of COVID 19 from CT image by the novel LeNet-5 CNN architecture," in *2020 23rd International Conference on Computer and Information Technology (ICCIT),* 2020: IEEE, pp. 1-5.

[35] M. Z. Alom *et al.*, "The history began from alexnet: A comprehensive survey on deep learning approaches," *arXiv preprint arXiv:1803.01164*, 2018.

[36] S. Tammina, "Transfer learning using vgg-16 with deep convolutional neural network for classifying images," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, pp. 143-150, 2019.

[37] J. Raja, P. Shanmugam, and R. Pitchai, "An automated early detection of glaucoma using support vector machine based visual geometry group 19 (VGG-19) convolutional neural network," *Wireless Personal Communications*, vol. 118, pp. 523-534, 2021.

[38] P. Ballester and R. Araujo, "On the performance of GoogLeNet and AlexNet applied to sketches," in *Proceedings of the AAAI conference on artificial intelligence*, 2016, vol. 30, no. 1.

[39] W. Shi, F. Jiang, and D. Zhao, "Single image super-resolution with dilated convolution based multi-scale information learning inception module," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017: IEEE, pp. 977-981.

[40] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: Generalizing residual architectures," *arXiv preprint arXiv:1603.08029*, 2016.

[41] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, 2020.

[42] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345-1359, 2009.

[43] Z. N. K. Swati *et al.*, "Brain tumor classification for MR images using transfer learning and fine-tuning," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 34-46, 2019.

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

[45] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1-48, 2019.

[46] A. Tharwat, "Classification assessment methods," *Applied computing and informatics*, vol. 17, no. 1, pp. 168-192, 2020.

[47] S. Yang and G. Berdine, "The receiver operating characteristic (ROC) curve," *The Southwest Respiratory and Critical Care Chronicles,* vol. 5, no. 19, pp. 34-36, 2017.

[48] J. N. Mandrekar, "Receiver operating characteristic curve in diagnostic test assessment," *Journal of Thoracic Oncology,* vol. 5, no. 9, pp. 1315-1316, 2010.

[49] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation," *Caspian journal of internal medicine,* vol. 4, no. 2, p. 627, 2013.

[50] A. Pandya, S. Sy, S. Cho, S. Alam, M. C. Weinstein, and T. A. Gaziano, "Validation of a cardiovascular disease policy microsimulation model using both survival and receiver operating characteristic curves," *Medical Decision Making,* vol. 37, no. 7, pp. 802-814, 2017.

[51] D. J. Marceau, P. J. Howarth, J.-M. M. Dubois, and D. J. Gratton, "Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 28, no. 4, pp. 513-519, 1990.

[52] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 24, no. 7, pp. 971-987, 2002.

[53] S. Liao, M. W. Law, and A. C. Chung, "Dominant local binary patterns for texture classification," *IEEE transactions on image processing,* vol. 18, no. 5, pp. 1107-1118, 2009.

[54] T. J. Alhindi, S. Kalra, K. H. Ng, A. Afrin, and H. R. Tizhoosh, "Comparing LBP, HOG and deep features for classification of histopathology images," in *2018 international joint conference on neural networks (IJCNN)*, 2018: IEEE, pp. 1-7.

[55] Y. Wang, Z. Li, L. Wang, and M. Wang, "A Scale Invariant Feature Transform Based Method," *J. Inf. Hiding Multim. Signal Process.,* vol. 4, no. 2, pp. 73-89, 2013.

[56] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer vision and image understanding,* vol. 110, no. 3, pp. 346-359, 2008.

[57] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE transactions on image processing,* vol. 19, no. 6, pp. 1657-1663, 2010.

[58] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing: 3rd International Conference, ICISP 2008. Cherbourg-Octeville, France, July 1-3, 2008. Proceedings 3*, 2008: Springer, pp. 236-243.

[59] R. Fakoor, F. Ladhak, A. Nazi, and M. Huber, "Using deep learning to enhance cancer diagnosis and classification," in *Proceedings of the international conference on machine learning*, 2013, vol. 28: ACM New York, NY, USA, pp. 3937-3949.

[60] J. F. Mccarthy *et al.*, "Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management," *Annals of the New York Academy of Sciences,* vol. 1020, no. 1, pp. 239-262, 2004.

[61] N. B. Bahadure, A. K. Ray, and H. P. Thethi, "Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM," *International journal of biomedical imaging,* vol. 2017, 2017.

[62] M. De Bruijne, "Machine learning approaches in medical image analysis: From detection to diagnosis," vol. 33, ed: Elsevier, 2016, pp. 94-97.

[63] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *Ieee Access,* vol. 6, pp. 9375-9389, 2017.

[64] A. Depeursinge, A. Foncubierta-Rodriguez, D. Van De Ville, and H. Müller, "Three-dimensional solid texture analysis in biomedical imaging: review and opportunities," *Medical image analysis,* vol. 18, no. 1, pp. 176-196, 2014.

[65] K. Rawal and G. Sethi, "Medical Image Processing in Detection of Abdomen Diseases," *Advancement of Machine Intelligence in Interactive Medical Image Analysis,* pp. 153-166, 2020.

[66] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Computational and structural biotechnology journal*, vol. 16, pp. 34-42, 2018.

[67] T. Jagadesh and R. J. Rani, "A novel speckle noise reduction in biomedical images using PCA and wavelet transform," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016: IEEE, pp. 1335-1340.

[68] Shallu, P. Nanglia, S. Kumar, and A. K. Luhach, "Detection and analysis of lung cancer using radiomic approach," *Smart Computational Strategies: Theoretical and Practical Aspects*, pp. 13-24, 2019.

[69] Z. Othman and S. Saleem, "Diseases diagnosis using medical palmistry fuzzy model," in *MATEC Web of Conferences*, 2016, vol. 76: EDP Sciences, p. 04021.

[70] S. Sharma and R. Mehra, "Effect of layer-wise fine-tuning in magnification-dependent classification of breast cancer histopathological image," *The Visual Computer*, vol. 36, no. 9, pp. 1755-1769, 2020.

[71] S. Sharma and R. Mehra, "Implications of pooling strategies in convolutional neural networks: A deep insight," *Foundations of Computing and Decision Sciences*, vol. 44, no. 3, pp. 303-330, 2019.

[72] A. Alarabeyyat and M. Alhanahnah, "Breast cancer detection using k-nearest neighbor machine learning algorithm," in *2016 9th International Conference on Developments in eSystems Engineering (DeSE)*, 2016: IEEE, pp. 35-39.

[73] L. Zhang, L. Li, M. Tang, Y. Huan, X. Zhang, and X. Zhe, "A new approach to diagnosing prostate cancer through magnetic resonance imaging," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 897-904, 2021.

[74] C. Cao *et al.*, "Deep learning and its applications in biomedicine," *Genomics, proteomics & bioinformatics*, vol. 16, no. 1, pp. 17-32, 2018.

[75] X. Zhou *et al.*, "A comprehensive review for breast histopathology image analysis using classical and deep neural networks," *IEEE Access*, vol. 8, pp. 90931-90956, 2020.

[76] A. M. Kist and M. Döllinger, "Efficient biomedical image segmentation on EdgeTPUs at point of care," *IEEE Access*, vol. 8, pp. 139356-139366, 2020.

[77] S. S. Koshy, L. J. Anbarasi, M. Jawahar, and V. Ravi, "Breast cancer image analysis using deep learning techniques–a survey," *Health and Technology,* vol. 12, no. 6, pp. 1133-1155, 2022.

[78] H. Wieslander *et al.*, "Deep convolutional neural networks for detecting cellular changes due to malignancy," in *Proceedings of the IEEE international conference on computer vision workshops*, 2017, pp. 82-89.

[79] Y. Liu *et al.*, "Detecting diseases by human-physiological-parameter-based deep learning," *IEEE Access,* vol. 7, pp. 22002-22010, 2019.

[80] X. Ren *et al.*, "Task decomposition and synchronization for semantic biomedical image segmentation," *IEEE Transactions on Image Processing,* vol. 29, pp. 7497-7510, 2020.

[81] S. Pang, A. Du, M. A. Orgun, and Z. Yu, "A novel fused convolutional neural network for biomedical image classification," *Medical & biological engineering & computing,* vol. 57, pp. 107-121, 2019.

[82] Shallu and R. Mehra, "Automatic magnification independent classification of breast cancer tissue in histological images using deep convolutional neural network," in *Advanced Informatics for Computing Research: Second International Conference, ICAICR 2018, Shimla, India, July 14–15, 2018, Revised Selected Papers, Part I 2*, 2019: Springer, pp. 772-781.

[83] M. Toğaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," *Medical hypotheses,* vol. 134, p. 109531, 2020.

[84] S. Makaju, P. Prasad, A. Alsadoon, A. Singh, and A. Elchouemi, "Lung cancer detection using CT scan images," *Procedia Computer Science,* vol. 125, pp. 107-114, 2018.

[85] A. Osmanović, S. Halilović, L. A. Ilah, A. Fojnica, and Z. Gromilić, "Machine learning techniques for classification of breast cancer," in *World Congress on Medical Physics and Biomedical Engineering 2018: June 3-8, 2018, Prague, Czech Republic (Vol. 1),* 2019: Springer, pp. 197-200.

[86] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent

residual convolutional neural network," *Journal of digital imaging,* vol. 32, pp. 605-617, 2019.

[87] S. Sharma, R. Mehra, and S. Kumar, "Optimised CNN in conjunction with efficient pooling strategy for the multi-classification of breast cancer," *IET Image Processing,* vol. 15, no. 4, pp. 936-946, 2021.

[88] S. Sharma and R. Mehra, "Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight," *Journal of digital imaging,* vol. 33, pp. 632-654, 2020.

[89] D. A. Ragab, O. Attallah, M. Sharkas, J. Ren, and S. Marshall, "A framework for breast cancer classification using multi-DCNNs," *Computers in Biology and Medicine,* vol. 131, p. 104245, 2021.

[90] G. Murtaza *et al.,* "Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges," *Artificial Intelligence Review,* vol. 53, pp. 1655-1720, 2020.

[91] S. Boumaraf, X. Liu, Z. Zheng, X. Ma, and C. Ferkous, "A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images," *Biomedical Signal Processing and Control,* vol. 63, p. 102192, 2021.

[92] S. Saxena, S. Shukla, and M. Gyanchandani, "Breast cancer histopathology image classification using kernelized weighted extreme learning machine," *International Journal of Imaging Systems and Technology,* vol. 31, no. 1, pp. 168-179, 2021.

[93] Y.-D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, "Improved breast cancer classification through combining graph convolutional network and convolutional neural network," *Information Processing & Management,* vol. 58, no. 2, p. 102439, 2021.

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

[95] I. Hirra *et al.,* "Breast cancer classification from histopathological images using patch-based deep learning modeling," *IEEE Access,* vol. 9, pp. 24273-24287, 2021.

[96] X. Y. Liew, N. Hameed, and J. Clos, "An investigation of XGBoost-based algorithm for breast cancer classification," *Machine Learning with Applications,* vol. 6, p. 100154, 2021.

[97] S. Singh *et al.*, "Feature importance score-based functional link artificial neural networks for breast cancer classification," *BioMed Research International,* vol. 2022, 2022.

[98] S. Sharma and S. Kumar, "The Xception model: A potential feature extractor in breast cancer histology images classification," *ICT Express,* vol. 8, no. 1, pp. 101-108, 2022.

[99] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *Ieee transactions on biomedical engineering,* vol. 63, no. 7, pp. 1455-1462, 2015.

[100] E. Kumaraswamy, S. Sharma, and S. Kumar, "Invasive Ductal Carcinoma Grade Classification in Histopathological Images using Transfer Learning Approach," in *2021 IEEE Bombay Section Signature Conference (IBSSC)*, 2021: IEEE, pp. 1-6.

[101] P. H. Zavareh, A. Safayari, and H. Bolhasani, "BCNet: A deep convolutional neural network for breast cancer grading," *arXiv preprint arXiv:2107.05037,* 2021.

[102] S. Talpur, M. Rashid, S. J. Khan, and S. A. Syed, "Automatic Detection System to Identify Invasive Ductal Carcinoma by Predicting Bloom Richardson Grading from Histopathological Images," *Journal of Independent Studies and Research Computing,* vol. 20, no. 1, pp. 45-55, 2022.

[103] R. Sujatha, J. M. Chatterjee, A. Angelopoulou, E. Kapetanios, P. N. Srinivasu, and D. J. Hemanth, "A transfer learning-based system for grading breast invasive ductal carcinoma," *IET Image Processing,* vol. 17, no. 7, pp. 1979-1990, 2023.

[104] E. Kumaraswamy, S. Kumar, and M. Sharma, "An Invasive Ductal Carcinomas Breast Cancer Grade Classification Using an Ensemble of Convolutional Neural Networks," *Diagnostics,* vol. 13, no. 11, p. 1977, 2023.

[105] V. Mudeng, M. N. Farid, G. Ayana, and S.-w. Choe, "Domain and Histopathology Adaptations–Based Classification for Malignancy Grading System," *The American Journal of Pathology*, 2023.

[106] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 international joint conference on neural networks (IJCNN)*, 2016: IEEE, pp. 2560-2567.

[107] R. Yan *et al.*, "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52-60, 2020.

[108] M. Gour, S. Jain, and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image classification," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 621-635, 2020.

[109] Y. Sun *et al.*, "Detection of breast tumour tissue regions in histopathological images using convolutional neural networks," in *2018 IEEE international conference on image processing, applications and systems (IPAS)*, 2018: IEEE, pp. 98-103.

[110] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, vol. 125, pp. 1-6, 2019.

[111] G. Murtaza, L. Shuib, G. Mujtaba, and G. Raza, "Breast cancer multi-classification through deep neural network and hierarchical classification approach," *Multimedia Tools and Applications*, vol. 79, pp. 15481-15511, 2020.

[112] Y. Zou, J. Zhang, S. Huang, and B. Liu, "Breast cancer histopathological image classification using attention high-order deep network," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 266-279, 2022.

[113] N. Ahmad, S. Asghar, and S. A. Gillani, "Transfer learning-assisted multi-resolution breast cancer histopathological images classification," *The Visual Computer*, vol. 38, no. 8, pp. 2751-2770, 2022.

[114] S. Diao *et al.*, "Deep Multi-Magnification Similarity Learning for Histopathological Image Classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1535-1545, 2023.

[115] K. Datta Gupta, D. K. Sharma, S. Ahmed, H. Gupta, D. Gupta, and C.-H. Hsu, "A novel lightweight deep learning-based histopathological image classification model for IoMT," *Neural Processing Letters*, vol. 55, no. 1, pp. 205-228, 2023.

[116] E. Kumaraswamy, S. Kumar, and S. Sharma, "Revealing Insights in Breast Cancer Grade Classification: A Comparative Study of CNN-Based Approaches with and without Segmentation," *Evolutionary Intelligence Journal*, Reseach paper - Submitted and Under review, 2023.

[117] E. Kumaraswamy, S. Kumar, and S. Sharma, "Transfer Learning-Based CNN Model for Efficient Invasive Ductal Carcinoma Breast Cancer Grade Classification using Databiox Dataset," *Multimedia Tools and Applications Journal*, Reseach paper - Submitted and Under review, 2023.

[118] C. H. Barrios, "Global challenges in breast cancer detection and treatment," *The Breast*, vol. 62, pp. S3-S6, 2022.

[119] A. Cruz-Roa *et al.*, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," in *Medical Imaging 2014: Digital Pathology*, 2014, vol. 9041: SPIE, p. 904103.

[120] M. Puttagunta and S. Ravi, "Medical image analysis based on deep learning approach," *Multimedia tools and applications*, vol. 80, pp. 24365-24398, 2021.

[121] H.-Y. Chen and C.-Y. Su, "An enhanced hybrid MobileNet," in *2018 9th International Conference on Awareness Science and Technology (iCAST)*, 2018: IEEE, pp. 308-312.

[122] I. Naskinova, "Transfer learning with NASNet-Mobile for Pneumonia X-ray classification," *Asian-European Journal of Mathematics*, vol. 16, no. 01, p. 2250240, 2023.

[123] J. Huang, W. Gong, and H. Chen, "Menfish classification based on Inception_V3 convolutional neural network," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 677, no. 5: IOP Publishing, p. 052099.

[124] A. DEMİR and F. YILMAZ, "Inception-ResNet-v2 with LeakyReLU and averagepooling for more reliable and accurate classification of chest X-ray images," in *2020 Medical Technologies Congress (TIPTEKNO)*, 2020: IEEE, pp. 1-4.

[125] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251-1258.

[126] G. Aresta *et al.*, "Bach: Grand challenge on breast cancer histology images," *Medical image analysis*, vol. 56, pp. 122-139, 2019.

[127] N. Brancati *et al.*, "Bracs: A dataset for breast carcinoma subtyping in h&e histology images," *Database*, vol. 2022, p. baac093, 2022.

[128] D. D. Patil and S. G. Deore, "Medical image segmentation: a review," *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 1, pp. 22-27, 2013.

[129] Z. Hussain, F. Gimenez, D. Yi, and D. Rubin, "Differential data augmentation techniques for medical imaging classification tasks," in *AMIA annual symposium proceedings*, 2017, vol. 2017: American Medical Informatics Association, p. 979.

[130] Á. Casado-García *et al.*, "CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks," *BMC bioinformatics*, vol. 20, pp. 1-14, 2019.

[131] "Cancer " https://www.who.int/news-room/fact-sheets/detail/cancer (accessed 25 March 2022, 2022).

[132] G. Aresta *et al.*, "Bach: Grand challenge on breast cancer histology images," *Medical image analysis*, vol. 56, pp. 122-139, 2019.

[133] P. Nanglia, S. Kumar, A. N. Mahajan, P. Singh, and D. Rathee, "A hybrid algorithm for lung cancer classification using SVM and Neural Networks," *ICT Express*, vol. 7, no. 3, pp. 335-341, 2021.

[134] S. Kumar and S. Sharma, "Sub-classification of invasive and non-invasive cancer from magnification independent histopathological images using hybrid neural networks," *Evolutionary Intelligence*, vol. 15, no. 3, pp. 1531-1543, 2022.

[135] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational intelligence and neuroscience*, vol. 2018, 2018.

[136] P. Nanglia, A. N. Mahajan, D. S. Rathee, and S. Kumar, "Lung cancer classification using feed forward back propagation neural network for CT images," *International Journal of Medical Engineering and Informatics,* vol. 12, no. 5, pp. 447-456, 2020.

[137] P. Nanglia, S. Kumar, D. Rathi, and P. Singh, "Comparative investigation of different feature extraction techniques for lung cancer detection system," in *Advanced Informatics for Computing Research: Second International Conference, ICAICR 2018, Shimla, India, July 14–15, 2018, Revised Selected Papers, Part I 2*, 2019: Springer, pp. 296-307.

[138] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik,* vol. 29, no. 2, pp. 86-101, 2019.

[139] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621,* 2017.

[140] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*, 2018: IEEE, pp. 117-122.

[141] N. Bayramoglu, J. Kannala, and J. Heikkilä, "Deep learning for magnification independent breast cancer histopathology image classification," in *2016 23rd International conference on pattern recognition (ICPR)*, 2016: IEEE, pp. 2440-2445.

[142] K. Z. Guyton and M. K. Schubauer-Berigan, "Invited Perspective: Prioritizing Chemical Testing and Evaluation Using Validated in Vitro Assays Relevant to Key Characteristics," *Environmental Health Perspectives,* vol. 129, no. 7, p. 071303, 2021.

[143] "Global Cancer Observatory." https://gco.iarc.fr/today/online-analysis-multi-bars (accessed 15 May, 2023).

[144] "American Cancer Society." https://cancerstatisticscenter.cancer.org/#!/ (accessed 15 May 2023).

[145] H. A. Alturkistani, F. M. Tashkandi, and Z. M. Mohammedsaleh, "Histological stains: a literature review and case study," *Global journal of health science,* vol. 8, no. 3, p. 72, 2016.

[146] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering,* vol. 19, pp. 221-248, 2017.

[147] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[148] Y. Zhu and S. Newsam, "Densenet for dense flow," in *2017 IEEE international conference on image processing (ICIP)*, 2017: IEEE, pp. 790-794.

[149] Y. Zhang, "Lung segmentation with NASNet-Large-Decoder Net," *arXiv preprint arXiv:2303.10315,* 2023.

[150] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.

[151] "Global Cancer Observatory." https://gco.iarc.fr/today/online-analysis-multi-bars (accessed 25 March, 2023).

[152] A. Sharma and C. Ives, "Management of breast cancer: Basic principles," *Surgery (Oxford),* vol. 40, no. 2, pp. 113-120, 2022.

[153] "Breast Cancer India." https://www.breastcancerindia.net/statistics/trends.html (accessed 25 March 2023).

[154] R. C. Miles, S.-H. Chou, C. Vijapura, and A. Patel, "Breast Cancer Screening in Women With Dense Breasts: Current Status and Future Directions for Appropriate Risk Stratification and Imaging Utilization," *Journal of Breast Imaging,* vol. 4, no. 6, pp. 559-567, 2022.

[155] T. M. Allweis, N. Hermann, R. Berenstein-Molho, and M. Guindy, "Personalized screening for breast cancer: rationale, present practices, and future directions," *Annals of surgical oncology,* vol. 28, pp. 4306-4317, 2021.

[156] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," in *International conference on machine learning*, 2021: PMLR, pp. 10096-10106.

[157] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer," *Procedia Computer Science,* vol. 179, pp. 423-431, 2021.

[158] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer

learning," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 15, pp. 5682-5689, 2021.

[159] C. Lei, B. Hu, D. Wang, S. Zhang, and Z. Chen, "A preliminary study on data augmentation of deep learning for image classification," in *Proceedings of the 11th Asia-Pacific Symposium on Internetware*, 2019, pp. 1-6.

[160] C. Garbin, X. Zhu, and O. Marques, "Dropout vs. batch normalization: an empirical study of their impact to deep learning," *Multimedia Tools and Applications*, vol. 79, pp. 12777-12815, 2020.

[161] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *2017 international conference on communication and signal processing (ICCSP)*, 2017: IEEE, pp. 0588-0592.

[162] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, 2019: PMLR, pp. 6105-6114.

[163] H. Jang, S. M. Plis, V. D. Calhoun, and J.-H. Lee, "Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: Evaluation using sensorimotor tasks," *NeuroImage*, vol. 145, pp. 314-328, 2017.

List of Publications

A. International Journals

1. **Eelandula Kumaraswamy**, Sumit Kumar, and Manoj Sharma. "An Invasive Ductal Carcinomas Breast Cancer Grade Classification Using an Ensemble of Convolutional Neural Networks." **Diagnostics Journal**, Volume 13, Issue 11. Published: 5 Jun 2023. Impact factor 3.6. **SCIE and Scopus Indexed. (** https://doi.org/10.3390/diagnostics13111977)
2. **Eelandula Kumaraswamy**, Sumit Kumar, and Shallu Sharma. "Revealing Insights in Breast Cancer Grade Classification: A Comparative Study of CNN-Based Approaches with and without Segmentation." **Evolutionary Intelligence Journal**, (2023). (Communicated & Under Review) (Scopus and SCI Indexed).
3. **Eelandula Kumaraswamy**, Sumit Kumar, and Shallu Sharma. " Harnessing Transfer Learning through CNN Models on the Databiox Dataset for Improved Grading." **Multimedia Tools and Applications Journal,** (2023). (Communicated & Under Review) (Scopus and SCI Indexed).

B. International Conferences

1. **Eelandula Kumaraswamy**, Shallu Sharma, and Sumit Kumar. "Invasive Ductal Carcinoma Grade Classification in Histopathological Images using Transfer Learning Approach." 2021 **IEEE Bombay Section Signature Conference (IBSSC).** IEEE, 2021. **Scopus Indexed.** **(** https://ieeexplore.ieee.org/document/9673156)
2. **Eelandula Kumaraswamy**, Shallu Sharma, and Sumit Kumar. "A review on cancer detection strategies with help of biomedical images using machine learning techniques." **AIP Conference Proceedings.** Vol. 2418. No. 1. AIP Publishing, 2022. **Scopus Indexed. (** https://doi.org/10.1063/5.0081936)
3. **Eelandula Kumaraswamy**, Shallu Sharma, and Sumit Kumar. "Performance analysis of feature extraction and deep learning approaches on whole and segmented histopathological images for cancer grade classification." **IEEE International Conference on Interdisciplinary**

Approaches in Technology and Management for Social Innovation (ITAMSI-2024). IEEE, 2024. **Scopus Indexed**.
(https://ieeexplore.ieee.org/document/10502483)

C. Other Publications

1. **Eelandula Kumaraswamy**, Shallu Sharma, and Sumit Kumar. "Retrospective study of convolutional neural network for medical image analysis and a deep insight through histopathological dataset." **Computational Intelligence: Select Proceedings of InCITe 2022.** Singapore: **Springer Nature Singapore,** 2023. 47-58. **Scopus Indexed. (**https://link.springer.com/chapter/10.1007/978-981-19-7346-8_5)