FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

BREAST CANCER DETECTION USING MICROSCOPIC HISTOPATHOLOGICAL IMAGES: A TRANSFER LEARNING APPROACH.

Bilyaminu MUHAMMAD

Master's Thesis

DEPARTMENT OF SOFTWARE ENGINEERING

Division of Technology Faculty

JUNE 2022

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

BREAST CANCER DETECTION USING MICROSCOPIC HISTOPATHOLOGICAL IMAGES: A TRANSFER LEARNING APPROACH.

Author
Bilyaminu MUHAMMAD

Supervisor
Associate. Prof. Dr. Fatih ÖZKAYNAK

JUNE 2022
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

Title:	Breast Cancer Detection using Microscopic Histopathological Images: A Transfer Learning Approach.
Author:	Bilyaminu MUHAMMAD
Submission Date:	27 May 2022
Defense Date:	30 June 2022

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was unanimously approved after the defense exam made open to the academic audience.

Signature

Supervisor:	Associate. Prof. Dr. Fatih ÖZKAYNAK Firat University, Faculty of Technology	Approved
Chair:	Assist. Prof. Dr. Harun BİNGÖL Malatya Turgut Özal University, Faculty of Engineering and Natural Sciences	Approved
Member:	Assist. Prof. Dr. Bihter DAŞ Fırat University, Faculty of Technology	Approved

This thesis was approved by the Administrative Board of the Graduate School on

..... / ..... / 20 .....

Signature

Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled “Breast Cancer Detection using Microscopic Histopathological Images: A Transfer Learning Approach.” in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

30 June 2022

Bilyaminu MUHAMMAD

PREFACE

Breast cancer disease is an uncontrolled growth of cells that gradually spreads to other parts of the body and has become one of the world's leading causes of death among women. Many studies have been published in the literature on breast cancer diagnosis using microscopic biopsy and histopathological images, including feature extractions. However, extraction of features from histological images is a critical and challenging part of the computer-aided detection of breast cancer using histopathology images. Pathologists apply visual examination to breast cancer images to find the affected region of interest. This procedure is slow and time-consuming. Furthermore, the procedure depends on the pathologist's vast experience, which requires scanning image tissue under different magnification levels of the microscope to find clinical assessment guides that produce correct diagnoses. Thus, this process is subjective to certain factors, such as fatigue, that influence the outcomes of the medical practitioner's precise image interpretations. Recently, many fields have benefited from deep learning applications, including medical images and feature extractions, especially histological ones. However, insufficient medical health datasets to train the deep learning model from scratch affected its accuracy. Furthermore, data dependency is one of the challenging issues in deep learning because it requires a large quantity of training data to comprehend and uncover the hidden patterns in the data.

This thesis presents a novel deep feature extraction engineering for subtypes of breast cancer diagnosis-based transfer learning approach to overcome the stated limitations in deep learning because transfer learning utilizes the knowledge of already pre-trained convolutional neural networks trained on millions of image classification tasks, thereby saving additional computational power and resources required to train deep learning models from scratch. Based on the obtained results in this research thesis, the proposed model outperforms some of the studies in the literature and is proving to be very robust at handling deep feature extraction from raw histological images of breast cancer prior to applying any image processing techniques. Thus, it is becoming beneficial to obtain fast and accurate results in detecting breast cancer subtypes in their early stages based on the feature extraction engineering approach. Additionally, the model can serve as a second opinion for medical specialists, mainly assisting inexperienced practitioners, lowering workload, and improving the objectivity of final diagnosis. This thesis is written to meet the Graduate School of Natural and Applied Sciences requirements of the Software Engineering master's program at Firat University. The research study was carried out from December 2021 to May 2022.

I want to thank and acknowledge Fabio Spanhol for providing the dataset (BreaKhis) used in this research and making it available to the public as benchmark data for researchers in the area of a breast cancer diagnosis. My appreciation goes to my former supervisor for his immense contributions, Prof. Dr. Asaf Varol, who started supervising this research in its initial stages until he retired from the school. I have been lucky to be assigned to a new supervisor, Associate Prof. Dr. Fatih ÖZKAYNAK, who cared and was passionate about my work and responded promptly to my questions and queries. I appreciate Associate Prof. Dr. Türker Tuncer for his immense contribution and words of encouragement to carry out this work with perfection at every stage. My sponsor, the National Information Technology Development Agency (NITDA), is duly acknowledged for believing in me to see this through. My appreciation also goes to Dr. Mohammed Bawa Gummi. He happened to be a brother, mentor, and my guidance for his words of encouragement, morally and financially, at every stage of my educational career. Finally, I dedicated this thesis to my late parents.

Bilyaminu MUHAMMAD
ELAZIG, 2022

CONTENTS

ABSTRACT

Breast Cancer Detection using Microscopic Histopathological Images: A Transfer Learning Approach.

Bilyaminu MUHAMMAD

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Software Engineering

June 2022, Page: xii + 54

The extraction of features from histological images is a challenging part of the computer-aided detection of breast cancer. This work presents a novel deep feature extraction engineering for subtypes of Breast detection-based transfer learning approach using the BreaKhis dataset. This approach consists of five phases: feature extraction, concatenation, transformation, selection, and classification. In the first phase, nineteen pre-trained convolutional neural networks were utilized as feature extractors to extract features from the input images. Support vector machines were used in this phase to compute the loss value function. In the classification phase, it was used as a classifier. According to the feature extraction results, the two networks attained the highest accuracies on the dataset, outperforming other networks. The two networks considered were chosen and concatenated to produce the DRNet model, a combination of ResNet50 and DenseNet201 pre-trained networks. The extracted features were decomposed into five sub-hand low-level features using a multilevel discrete wavelet transform in the transformation phase. Iterative neighborhood component analysis was utilized to choose the minimum number of features required in the classification phase. Cubic support vector machines and k-nearest neighbors were used as classifiers in the final phase. k-nearest neighbor produced an average classification rate accuracy of 98.74%, 97.7%, 97.3%, and 98% on the 40×, 100×, 200×, and 400× levels of magnification, respectively, where support vector machines produced an average classification rate accuracy of 98.61%, 98.04%, 97.68%, and 97.71% on the 40×, 100×, 200×, and 400× levels of magnification, respectively, using the one-gainst-all and 10-fold cross-validation method.

Keywords: Feature extraction, Histological images, Pretrained networks, Transfer learning.

ÖZET

Mikroskobik Histopatolojik Görüntülerle Meme Kanseri Tespiti: Bir Transfer Öğrenme Yaklaşımı.

Bilyaminu MUHAMMAD

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Haziran 2022, Sayfa: xii + 54

Histolojik görüntülerden özelliklerin çıkarılması, bilgisayar destekli meme kanseri tespitinin zorlu bir parçasıdır. Bu çalışma, BreaKhis veri kümesini kullanarak Meme algılama tabanlı aktarım öğrenme yaklaşımının alt türleri için yeni bir derin özellik çıkarma mühendisliği sunmaktadır. Bu yaklaşım beş aşamadan oluşur: özellik çıkarma, birleştirme, dönüştürme, seçme ve sınıflandırma. İlk aşamada, girdi görüntülerinden öznitelik çıkarmak için on dokuz önceden eğitilmiş evrişimsel sinir ağı öznitelik çıkarıcı olarak kullanılmıştır. Bu aşamada kayıp değeri fonksiyonunu hesaplamak için destek vektör makineleri kullanıldı. Sınıflandırma aşamasında sınıflandırıcı olarak kullanılmıştır. Öznitelik çıkarma sonuçlarına göre, iki ağ, diğer ağlardan daha iyi performans göstererek, veri seti üzerinde en yüksek doğruluğa ulaşmıştır. Göz önünde bulundurulan iki ağ, ResNet50 ve DenseNet201 önceden eğitilmiş ağların bir kombinasyonu olan DRNet modelini üretmek için seçilmiş ve birleştirilmiştir. Çıkarılan öznitelikler, dönüşüm aşamasında çok düzeyli ayrık dalgacık dönüşümü kullanılarak beş alt-el düşük düzey özniteliğe ayrıştırılmıştır. Sınıflandırma aşamasında gereken minimum öznitelik sayısını seçmek için yinelemeli komşuluk bileşeni analizi kullanılmıştır. Son aşamada sınıflandırıcı olarak kübik destek vektör makineleri ve k-en yakın komşular kullanılmıştır. k-en yakın komşu 40×, 100×, 200× ve 400× büyütme seviyelerinde sırasıyla %98,74, %97,7, %97,3 ve %98'lik bir ortalama sınıflandırma oranı doğruluğu üretti; burada destek vektör makineleri bir ortalama üretti 40×, 100×, 200× ve 400× büyütme seviyelerinde sırasıyla %98,61, %98,04, %97,68 ve %97,71 sınıflandırma oranı doğruluğu, tek elde tüm ve 10 kat çapraz doğrulama kullanılarak yöntem.

Anahtar Kelimeler: Özellik çıkarma, Histolojik görüntüler, Önceden eğitilmiş ağlar, Aktarım öğrenimi.

LIST OF FIGURES

Page

LIST OF TABLES

Page

ABBREVIATIONS

A	: Adenosis
AI	: Artificial Intelligence
ANNs	: Artificial Neural Networks
BC	: Breast Cancer
BT	: Benign Tumor
CADS	: Computer- Aided Diagnosis System
CNNs	: Convolutional Neural Networks
CT	: Computer Tomography
DC	: Ductal Carcinoma
DL	: Deep Learning
DRNet	: A Combined Networks of ResNet50 and DenseNet201
DWT	: Discrete Wavelet Transform
F	: Fibroadenoma
FN	: False Negative
FP	: False Positive
HLF	: High Level Feature
INCA	: Iterative Neighborhood Component Analysis
k-NN	: k-Nearest Neighbor
LC	: Lobular Carcinoma
LLF	: Low-Level Feature
LRD	: Learning Rate Decay
LSTM	: Long-Short Term Memory
MC	: Mucinous Carcinoma
MDWT	: Multilevel Discrete Wavelet Transform
ML	: Machine Learning
MRI	: Magnetic Resonance Imaging
MT	: Malignant Tumor
NN	: Neural Network
PCA	: Principal Component Analysis
PT	: Phyllodes Tumor
RNNs	: Recurrent Neural Networks
SVM	: Support Vector Machine(s)
TA	: Tubular Adenoma
TL	: Transfer Learning
TN	: True Negative
TP	: True Positive
VGG	: Visual Geometry Group

1. INTRODUCTION

Breast cancer (BC) is a prevalent type of cancer disease worldwide. BC disease is an uncontrolled growth of cells that gradually spreads to other parts of the body and has become one of the world's leading causes of death among women. BC cancer can also occur among men but with less severe effects [1]. Invasive ductal carcinoma and lobular carcinoma are the most common types of BC among women[2]. Reported incidence by World Health Organization (WHO) in 2021 shows 2.3 million women population diagnosed with BC, and 685 thousand deaths were recorded across the globe [2]. Factors such as aging, excessive use of alcohol, exposure to radiation, Etc., spread the disease's incidence[2].

A microscopic examination of the cells by computerized imaging tests such as ultrasound, MIR, X-ray images, positron emission tomography (PET), electrical impedance-based imaging, computer tomography (CT), optical imaging, and mammography images are the most commonly used imaging capturing techniques to identify suspected inflammatory in the breast by the medical practitioners. These tests' imaging assists them in determining whether cancer is feasible by showing development or mass. Usually, a final decision is made through various patient laboratory trials [3].

Another method used by Scientists is called a biopsy[4]. A surgeon carries out a biopsy. The Scientist takes a sample of the tissue that is under study. Cancer cells differ from cells in healthy conditions. They are in a distinct order, differ in size, and are poorly organized. However, a medical image like CT or MRI cannot display all patterns and data on cancer or cancer subtypes.

In order to make the diagnosis of cancer more effective and quicker, this method needs to be automated. These processes are essential in the early detection of the disease, which will improve the survival rates of the patients because an early cancer diagnosis can help save a patient's life.

Furthermore, an earlier diagnosis may stop the disease from further spreading to other body parts. This early diagnosis and prognosis have facilitated these techniques to explore the Large quantities of cancer information gathered, accessible to the medical research community. Biomedical teams are getting better results with the application and substantial improvement in the precise application of ML approaches for cancer, recurrence, and survival prediction throughout the years, ranging from 15% to 20% [5].

However, the precise prediction of the disease outcome at an early stage is one of the most challenging tasks for physicians[6]. Consequently, ML and DL techniques among medical researchers are becoming more common and essential. These techniques can identify patterns and relations from complex datasets while effectively predicting future cancer outcomes. Most research was carried out using different approaches on a real dataset. However, privacy and ethics made most of the dataset not readily available for other researchers using a different approach.

1.1. The motivation of the research study

The procedure for breast image examination is slow and time-consuming. It requires scanning image tissue under different magnification levels of the microscope to find clinical assessment guides that produce correct diagnoses. In addition, the procedure depends on the pathologist's vast experience. Furthermore, it is subjective to certain factors that influence the outcomes of the precise result of image interpretations by medical practitioners. As a result, each incorrect examination induced by false negative (FN) or false positive (FP) results in severe drug resistance and expensive hardship for patients. Hence, early detection of BC may save additional women's lives and resources. Therefore, diagnosing it earlier with greater precision is critical since it may allow for appropriate medication or chemotherapy for a diagnosed patient. The primary motivation of this research project is to develop and automate some of the functions in the diagnosis workflow with better accuracy. Thus, becoming beneficial to obtain fast and precise quantification to decrease spectator variability.

1.2. Aims and objectives of the research

This study aims to demonstrate and present a novel feature extraction technique for identifying breast cancer using pre-trained networks as feature extractors. The research study will employ these networks to extract the most significant features from the Breakhis raw image database and identify them as benign or malignant histopathology images. The following are the research objectives: To develop an efficient and robust breast cancer classification model capable of utilizing microscopic histopathology images.

i. To conduct an exploration of the breakhis dataset.
ii. To utilize pre-trained neural networks as feature extractors
iii. To determine the two most relevant feature extractors from a raw image database.
iv. To develop Deep models that magnification-dependent
v. To address multiclass classification problems using the breakhis breast cancer image database.
vi. To assess the model's correctness.
vii. To determine if the magnification factor affects how medical professionals interpret images.

1.3. Statement of the problems

Pathologists apply visual examination on breast cancer image database. However, the procedure is slow and time-consuming. Furthermore, the procedure depends on the pathologist's

vast experience. The analysis requires scanning image tissue under different magnification levels of the microscope to find clinical assessment guides that produce correct diagnoses.

Thus, it is subjective to certain factors that influence the outcomes of the precise result of image interpretations by medical practitioners. Recently, deep learning (DL) has garnered increased interest from researchers worldwide and has been effectively used in various fields, including the analysis of medical images.

However, insufficient medical health datasets to train the DL model from scratch affect its accuracy. Its algorithms attempt to learn from higher-level features of a vast dataset, making DL far beyond machine learning in its conventional form. It can extract data features automatically using supervised or unsupervised feature learning techniques and hierarchical feature extraction. Unfortunately, data dependency is one of the most challenging issues in deep learning.

Additionally, deep learning requires a large quantity of training data to comprehend and uncover hidden patterns in the dataset.

Transfer learning is increasingly being utilized to address the issue of inadequate data. As a result, the present work will employ a transfer learning method to detect and classify BC histopathological images from the Breakhis database as either Benign or Malignant. The method has become popular in image analysis. Thus, the proposed research study will demonstrate the potential of transfer learning to improve performance on classification issues involving BC histopathology images.

1.4. Hypothesis

As stated in the problem statement, the proposed research will establish the following primary hypothesis:

"Can the Transfer learning method improve BC diagnosis using histopathological images from the Breakhis dataset?"

1.5. Contributions

The main contribution of this research study are stated below:

i. To demonstrate and provide a novel feature extraction technique for breast cancer diagnosis utilizing pre-trained image networks and pick the two highest-performing feature extractors from the Breakhis image database's raw images.

ii. The research will advance our understanding of precision medicine. The study aims to develop a predictive model for classifying the two most prevalent breast cancer (Benign and Malignant).

iii. Additionally, the project will develop a hybrid model for distinguishing various subtypes of benign and malignant cancers using the breakhis image database.

This hybrid model identifies the pathological characteristics of breast cancer and allows for the accurate differentiation of distinct subtypes of breast tumors. The model can serve as a second opinion for medical specialists, mainly assisting inexperienced practitioners, lowering workload, and improving the objectivity of final diagnosis. Additionally, the developed model may be used for additional histopathological classification problems involving images similar to those analyzed.

1.6. Overview of the research

Organization of this thesis is explained below.

- **Chapter one**, In the current, the relevance of the research to aid BC diagnosis among the female population has been discussed. This initial Chapter presents diagnosis through histopathological analysis by the pathologists, whose visual examinations are time-consuming and subject to the misinterpretation of the results. Finally, the motivation to work with this problem and the contributions of the research study was well presented in this Chapter.
- **Chapter two**: Summarizes the relevant literature of the research. This background research study is essential for correctly understanding the research approach and the achieved results in this research area. It also provides a proper comprehension of the computational techniques employed. Results achieved by the different methods applied to the Breakhis dataset have been analyzed and well presented.
- **Chapter three**: comprises a background about Artificial Intelligence, neural networks. It also explains deep learning, transfers learning concepts, discusses the convolutional neural network, and how transfer learning can be accomplished.
- **Chapter four**: Summarizes the planned study's materials, methodology, and model. Additionally, it serves as the Chapter for development and implementation.
- **Chapter five**: Presents results of the study, its analysis, and conclusions.

2. LITERATURE REVIEW

Recently, many studies have been published in the literature on breast cancer diagnosis using microscopic biopsy and histopathological images. Most existing studies classify the two basic types of BC (i.e., Benign and Malignant) using computer-assisted diagnosis tools (CADTs). Krawczyk and Schaefer [7] proposed another approach for breast cancer detection. Their strategy was to acquire image attributes that described the asymmetry of healthy and unhealthy breasts. They extract fundamental statistical features, histogram features, image occurrences, and numerous texture features for feature extraction. They merge multiple classifiers in the pattern classification section, utilizing a neural network fusion technique and genetic algorithms. This approach was validated using 146 static infrared breast images, 29 of which were unhealthy and 117 of which were healthy. They reported an accuracy of 90.03 %, a sensitivity of 80.35 %, and a specificity of 90.15 %. Spanhol et al. [8] introduced the Breakhis dataset, which is used to diagnose breast cancer using a histological image. Since its release, this dataset has garnered significant interest from academics looking to build accurate algorithms for classifying breast cancer histopathology images autonomously. Spanhol et al. [9] presented a CNN model trained on the Breakhis image database to classify breast cancer histopathology images (DeCAF). In order to train a new classifier, feature vectors were employed as input. These features were a promising alternative to the rapid development of high-accuracy BC recognition in the experimental assessment. The suggested model's accuracy was 86.3% at a magnification level of 200x and 84.6% at a magnification level of 40x. Breast histopathology images stained with Hematoxylin and Eosin (HE) are frequently used to diagnose cancer. However, potential differences in slide preparation conditions and stain color in HE images may increase the likelihood of misdiagnosis. As a result, the experimental work of Kausar et al.[10] suggested a new deep learning-based technique for breast histopathology images stained with Hematoxylin and Eosin (HE). The pipeline of the intended technique was the color normalization method and feature extraction strategy that are utilized to pre-process HE histological images. Additionally, deep feature computation was evaluated using deep convolutional neural networks (CNNs), and classification was carried out using a variety of deep classifiers. The performance of the proposed model approach was evaluated using the publicly available BreakHis dataset. The results indicate that the proposed approach has significant promise for cancer type classifications. Brancati et al. [11] investigated deep learning approaches for analyzing histology images of breast and lymphoma stained with hematoxylin and eosin (H&E). Their experimental work developed a deep learning technique into two case scenarios: first, detecting invasive ductal carcinoma in breast histology images, and second, categorizing lymphoma subtypes. Both use cases were handled in this study by using a residual convolutional neural network as part of a convolutional autoencoder network (i.e., FusionNet). Their results were compared to those

generated by employing other deep neural networks and evaluated using public datasets of digital histology images (UNet and ResNet). The results of this experiment work yielded an accuracy of 87.52%. The experimental study conducted by Deniz et al.[12] proposed a transfer learning-based histopathological image classification for breast cancer diagnosis. In this study, two types of pre-trained CNN models were considered. AlexNet is utilized for feature extraction, while Vgg16 was used for fine-tuning. The author further classifies the acquired features using the SVM classifier. This experimental work yielded 84.87% accuracy on 40x, 89.21% accuracy on 200x, 88.65% accuracy on 300x, and 86.75% accuracy on 400x. They concluded that transfer learning performed much better than feature extraction. Das et al. [13]suggested a CNN-based method. They examined images from a random number of locations in the tissue segment at various magnifications, without the need for viewing correspondence across different magnifications, as opposed to previous approaches. A majority voting-based technique was adopted for slide-level diagnosis. The posterior estimates across random multi-views at many magnifications were filtered to obtain a slide-level diagnosis. This study examined their experimental performance using a patient-level, 5-folded cross-validation approach with 58 malignant and 24 benign incidences of breast cancer, respectively. The model's accuracy was determined using metric criteria, and a 94.67% accuracy level was achieved. Chan and Tunszynski [14] used an SVM classifier to automatically categorize benign and malignant cancer types, including subtypes, using the Breakhis image database. Their results for multiclass classification at a magnification level of 40x resulted in an accuracy of 0.556%. Murtaza et al.[15] proposed an ensemble tree-based multiclassification of breast tumors using histopathology images via deep learning. The authors evaluated a deep learning model (DL) to extract discriminative features for multiclassification issues with less processing capacity to obtain more accurate results. The research was conducted using images from the Breakhis image database. They further upgraded the misclassification model by incorporating a misclassification algorithm. By utilizing fewer computer resources, an accuracy of 87.5% was attained for four subtypes of breast tumors. The experimental findings demonstrate that the suggested model outperforms previously published state-of-the-art baseline research for four subtypes of breast tumors. Bayramoglu et al. [16] presented a DL model for BC diagnosis using histology images. The classification was performed using CNNs. An accuracy of 83.33% and 83.25% for patient and CNN levels was achieved. According to Saber et al. [17], a novel deep CNN model based on transfer-learning (TL) was built to help detect and diagnose the BC suspicious region using two methodologies, namely 80-20 and cross-validation. A problem-specific model of DL architecture is used in this study. Demir [19] introduced a novel convolutional-LSTM model for classification the Breakhis image database to detect cancer subtypes automatically. MWSA (marker-controlled watershed segmentation algorithm) was applied as a processing method. A fine-tuned SVM classifier with Bayesian optimization yielded the best results for binary and multi-classification

tasks in his study. Zhang et al. [20] proposed a novel breast cancer high-order deep network (AHoNet) capable of detecting and classifying additional deep features in breast cancer pathological images by embedding an attention mechanism and a high-order statistical representation into a residual convolutional network simultaneously. The strategy consisted of two phases. Initially, AHoNet was used as an efficient channel attention module capable of reducing non-dimensionality. Then, local cross-channel interaction is used in the second phase to get conspicuous deep features of breast cancer pathological images at the local level. Then, using matrix power normalization, their second-order covariance statistics are computed further, resulting in a more robust global feature presentation of breast cancer images. On the public BreakHis and BACH breast cancer image datasets, the AHoNet model was assessed. The experimental results indicate that AHoNet achieves optimal patient-level classification accuracy of 99.29% and 85%, respectively, using the BreakHis and BACH image databases. Kataoka et al. [21] Created feature assessments of deep CNNs for object identification and detection with the well-known Alexnet and Vggnet. The authors carried out feature representations by utilizing the convolutional layer next to fully connected layers. In addition, feature concatenation and transformation were carried out using basic tuning and PCA on a publicly available benchmark image database called Caltech and Daimlerpedestrian. Based on the results achieved in this work, the authors concluded a significant improvement in image classification tasks when feature concatenation of convolutional layers is performed. In the work of Talaat et al. [22] proposed an improved hybrid image classification model using DL. They employed CNN to extract features and a swarm-based technique to identify the most relevant features. Yang et al. [23] created a diagnosis method based on support vector machines (SVMs) that primarily comprise three steps. In the first stage, PCA was used to eliminate redundant data and extract representative patterns from the original data. This procedure reduces the dimension of the feature space. They then use the differential evolution algorithm to find the optimal parameter values for the SVM in the subsequent stage. In the last stage, the SVM classifier is trained to distinguish between incoming tumors. The researchers assessed the classifier's performance objectively and thoroughly using classification accuracy, sensitivity, and specificity metrics. Compared to K-nearest neighbor, random forest, bagging naive bayes, decision tree, and other classification algorithms, the suggested method outperforms all when evaluated using the University of California's Wisconsin Diagnostic Breast Cancer (WDBC) dataset with five-fold cross-validation. Zhang et al.[24] developed an unsupervised feature learning framework for identifying distinct properties from gene expression profiles by merging a PCA technique with an autoencoder neural network. The AdaBoost method is used to build an ensemble classification system (PCA-AE-Ada), which was created to use the collected characteristics to predict clinical outcomes in breast cancer.

Additionally, the authors built a second classifier using the identical classifier learning approach (PCA-Ada) to act as a baseline for the proposed method, with the only variation being the training inputs. Their experiments indicated that the suggested strategy outperforms others using deep learning techniques. Yang et al.[25] developed a CNN architecture consisting of a convolutional layer, a tiny SE-ResNet module, and a fully-connected layer. The suggested tiny SE-ResNet module is an enhanced hybrid of the residual module and the Squeeze-and-Excitation block, achieving similar performance. On the Breakhis image database, they utilize their algorithm to classify breast cancer histology images into benign and malignant categories and eight subcategories. According to their findings, the suggested model attained an accuracy ranging from 98.87% to 99.34% for binary classification and accuracy ranging from 90.66% to 93.81% for multiclass classification. The use of histopathological images can be a highly successful approach to diagnosing breast cancer. The extraction of information from these images, on the other hand, is a complex process. As a result, Carvalho et al. [26] suggested a technique that utilized phylogenetic variety index values to categorize images to develop an algorithm to detect histopathological breast images into four categories: invasive carcinoma, in situ cancer (cancer in the breast that has not spread), normal tissue, and benign lesion. In this work, classifiers such as XGBoost, random forest, multilayer perceptron, and support vector machine were used to differentiate between the types of breast cancer investigated. They also performed content-based image retrieval to check the classification results. They proposed a rating for groups of images that had not been labeled, among other things. The results acquired were quite robust. They proved to help create a CADx system to assist experts in important health centers, which is now under development. Zewdie et al. [27] suggested a multiclass classification system for breast cancer categories, subcategories, and grades based on deep learning techniques. Histological images collected under four different magnification levels (40x, 100x, 200x, and 400x) were used in the training and validation of the algorithm. The created method can categorize breast cancer into binary (benign and malignant) and multiclass categories (sub-types). For tumors categorized as invasive ductal carcinomas, cancer grade was determined. Their findings revealed that the suggested technique attained an accuracy of 96.75 % for two-class classification tasks, 96.7 % accuracy for benign sub-types, and 95.78 % accuracy for malignant sub-types. In comparison, 93.86% was achieved for the grade identification task. The experimental study of [28] proposed a breast cancer image multi-classification using random patch aggregation and a depth-wise convolution-based deep-net model. This study utilized pre-trained VGG16 for Breakhis classification with 40X magnification. This study obtained a classification accuracy of 89.6% in multiclass classification. Most of the studies in the literature focused more on enhancing accuracy either using ML or DL-based techniques. The research project presents a novel feature extraction technique for breast cancer diagnosis, using a publicly available image database (Breakhis) suitable for multiple classification problems of breast cancer subtypes.

The Breakhis image database is a handy tool for researchers to set up future benchmarking and possible evaluations. This research project employed pre-trained image networks as feature extractors to select the most relevant features to ease BC diagnosis tasks. In this process, the two most performing feature extractors have been selected and concatenated to form a novel hybrid model capable of distinguishing Benign and Malignant subtypes of BC from this imge database. The developed hybrid model identifies breast cancer's pathological characteristics/features that allow for the accurate differentiation of distinct subtypes of breast tumors. The model can serve as a second opinion for medical specialists, mainly assisting inexperienced practitioners, lowering workload, and improving the objectivity of final diagnosis. Additionally, the developed model may be used for additional histopathological classification problems involving images similar to those analyzed in the dataset.

3. DEEP LEARNING

DL, also known as deep neural learning or deep neural network, is a category of machine learning techniques that employs numerous layers to extract higher-level (HL) characteristics from raw data. Lower layers (LL) in image processing, for example, may recognize edges. In contrast, higher layers may identify human-relevant entities like numerals, characters, or faces [29]. DL is simply a neural network with three (3) or more layers. A neural network is a structure like a human brain. It consists of artificial neurons, known as nodes, stacked next to each other in three layers, comprising the input layer, the hidden layer, and the output layer, as depicted in Figure 3.1. These layers seek to replicate the function of the human brain, allowing it to learn from enormous volumes of data. Whereas a neural network with such a single layer may still produce reasonable predictions, more hidden layers can assist in optimizing and tuning for accuracy. DL is at the heart of much artificial intelligence (AI) products and services that improve automation by executing analytical and physical activities without human interaction. DL is also at the heart of everyday products and services such as digital assistants, voice-enabled TV remotes, credit card fraud detection, and emerging AI technologies such as self-driving cars.

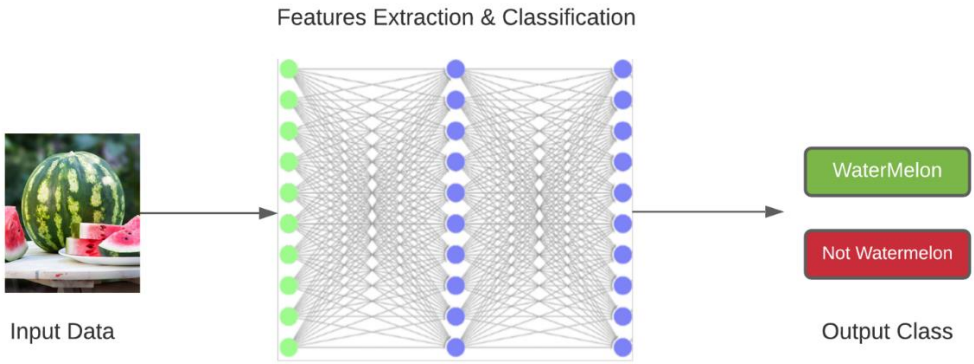

Figure 3.1. Example of a Deep learning model.

Figure 3.1 shows a typical DL architectural model consisting of the input layer, hidden layer, and output layer, respectively.

DL removes some of the pre-processing data tasks generally associated with machine learning. These algorithms can absorb and interpret unlabeled data such as text and images and automate feature extraction, which reduces the need for a human specialist, making them more potent than traditional ML algorithms. DL can automatically learn, extract, and distinguish each hidden feature or pattern from the images, unlike ML, which requires a human expert to establish this hierarchy of features manually. The deep learning algorithm then changes and fits itself for

accuracy via gradient descent and backpropagation, allowing it to generate more precise predictions about a fresh image of an object. All of these technologies share something in common, which is to help in a better understanding of data used for classification, prediction, and analytics. However, each of these techniques takes a very different approach to achieve the desired goal. Understanding the difference between them will provide a clearer picture of their harmony for better results.

3.1.1. Deep learning methods

To build powerful deep learning models, a variety of approaches can be utilized. Learning rate decay (LRD), transfer learning (TL), training model from scratch, and dropout are some of the approaches used in this process.

3.1.2. Learning rate

The learning rate decay, also known as adaptable learning rates, is a way of adjusting the learning rate to improve performance while decreasing training time. A hyperparameter component defines a system or establishes operating conditions for it preparatory to the learning process that governs how much change the model undergoes in response to an estimated error each time the model weights are modified. Overly rapid learning rates may result in inconsistent training processes or the learning of a poor set of weights, depending on the situation. Learning rates that are too low might result in a lengthy training procedure that tends to become stale. Techniques to lower the learning rate over time are the simplest and most common modifications to the learning rate that may be made throughout training.

3.1.3. Transfer learning

Humans have the inherent ability to transfer knowledge across different tasks. For example, what we acquire as knowledge while learning about one task, we utilize it in the same way to solve related tasks. Thus, the more related tasks, the easier it is for us to transfer or cross-utilize our knowledge. Thus, it brought about the idea of transfer learning.

Transfer learning is the knowledge of the already trained machine learning model and applied to another but related problem. It is a way of utilizing what has been learned to improve generalization in another task. It is a method of transforming the weight that a network has learned at ‘’ task A’’ to a new ‘’ task B’.

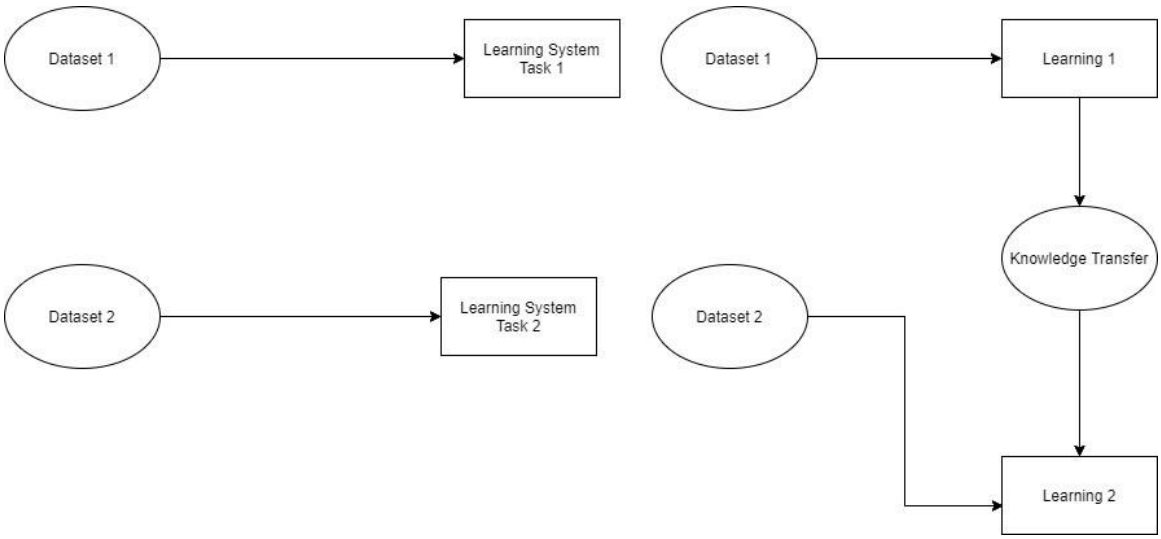

Figure 3.2. Traditional ML and TL Processes

Figure.3.2, as demonstrated shows traditional ML and TL processes.

Conventional machine learning has been traditionally designed to work in isolation because the algorithms are trained to solve specific tasks. Thus, the model must be built from scratch once the feature-space distributions change, making it difficult for knowledge to be retained or accumulated. Therefore, transfer learning is overcoming the isolated learning paradigm and utilizing knowledge acquired for one task to solve related ones. In transfer learning, the learning of new tasks relies on the previously learned tasks, and its learning process is faster, more accurate, and needs less training data.

3.1.4. Training from scratch

Unlike TL utilizing the knowledge of an already pre-trained network, making it possible to retrain the model with less computer power and little amount of data, training from scratch requires an application developer first to compile an extensive labeled data collection and then construct a network architecture that can learn the features and model in order for this strategy to operate. In particular, this strategy is beneficial for both new applications and those with many different output categories to sort through. On the other hand, this strategy is generally less prevalent since it requires an excessive quantity of data, which causes training to take days or even weeks, depending on the dataset.

3.1.5. Dropout

In order to overcome the problem of overfitting in networks with a large number of parameters, this strategy involves randomly deleting units and their connections from the neural

network while it is still being trained. In supervised learning applications such as speech recognition, document categorization, and computational biology, it has been demonstrated that the dropout strategy may increase the performance of neural networks.

3.1.6. Types of Deep Learning Algorithms

This study explored the most popular deep learning algorithms as follows:

1. Convolutional Neural Network
2. Long-Short Term Memory Networks
3. Recurrent Neural Network

3.1.7. Convolutional Neural Networks

As the name implies, a convolutional neural network (also known as ConvNet or CNN) is a network design for deep learning that learns directly from the input data, removing the need for human extracting features. Generally speaking, convolutional neural networks are a sort of neural network that is most commonly used to solve image processing challenges. Several layers make up this structure. Projects involving natural language processing can benefit from applying convolutional neural networks. Due to their importance in deep learning and artificial intelligence today, they are beneficial in these rapidly expanding fields. They are pretty effective for image classification, feature extraction, and processing. When different input types are received, the input layer accepts them. The hidden layers use these inputs to compute results. After that, the results of the analyses and extractions are sent through the output layer. CNN receives an image as its input and processes it using several weights and biases, extracts features, and outputs the process results. It tends to minimize the dimension of the images through a kernel operation, which makes it easier for feature extractions, unlike generic dense neural networks. CNNs have proven to be very effective in image classification and recognition, making them the essential ML and computer vision tools.

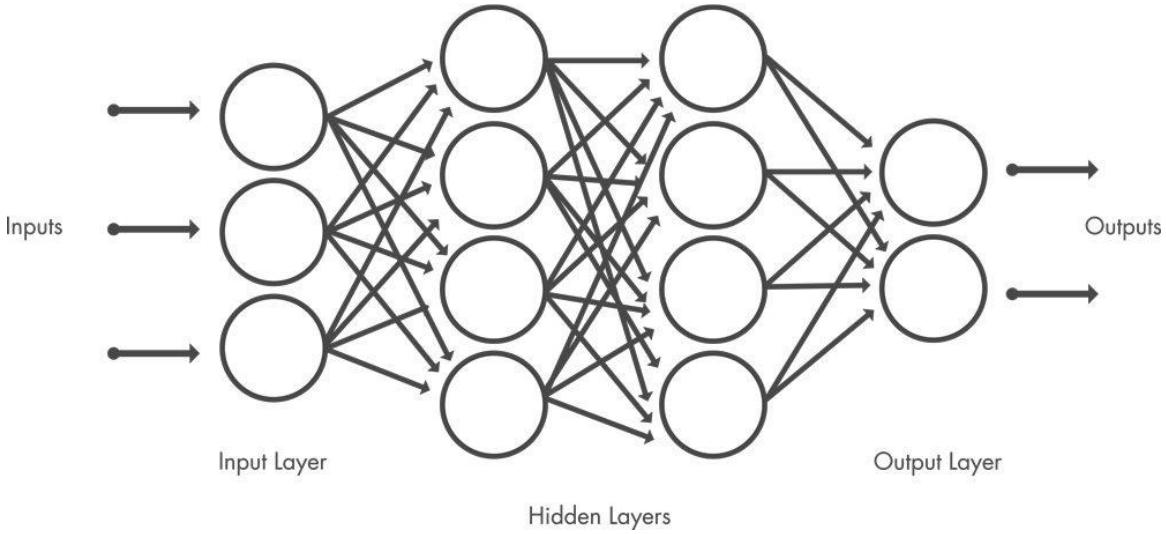

Figure 3.3. A typical CNN layer [30].

As shown in Figure 3.3, a neural network component comparable to CNN consists of an input layer, numerous hidden layers in the space in between them, and an output layer.

3.1.8. Architectural overview of CNN

CNNs are a kind of neural network that is most commonly used to solve image processing challenges. They are composed of three primary sorts of layers, which are as follows:

- Convolutional layer: The convolutional layer is the core block of CNN and works by playing a filter on an array of image pixels. This creates what is called convolved feature map.
    a) Flattening layer: The pooled feature map must first be collected before proceeding to the next stage, flattening the map. In order to do this, the whole pooled feature map matrix must first be converted into a single column before being supplied to the NN for processing.
    b) Kernel layer: is a feature extraction filter that extracts features from images. It is also a form of a matrix that moves over the input data based on the stride value by performing the dot product with the sub-region of the input data and then getting the output as the matrix of dot products.
    c) Stride: As the filter is moved across the image, left to right, top to bottom, with non-pixel, on horizontal movements, a one-pixel row change is made, and on vertical movements, a one-pixel row change is made. The stride is the amount of movement between filter applications to the input image, and it is nearly always symmetrical in height and width dimensions.

d) Padding: Padding is the best approach. The number of pixels required for the convolutional kernel to process the edge pixels is added onto the outside, copying the pixels from the edge of the image, effectively fixing border problems.

- Pooling layer: The pooling layer plays the role of down sampling the sample size by decreasing the number of parameters the network needs to process. The output of this is called pooled feature map. There are two ways by which this process can achieve:

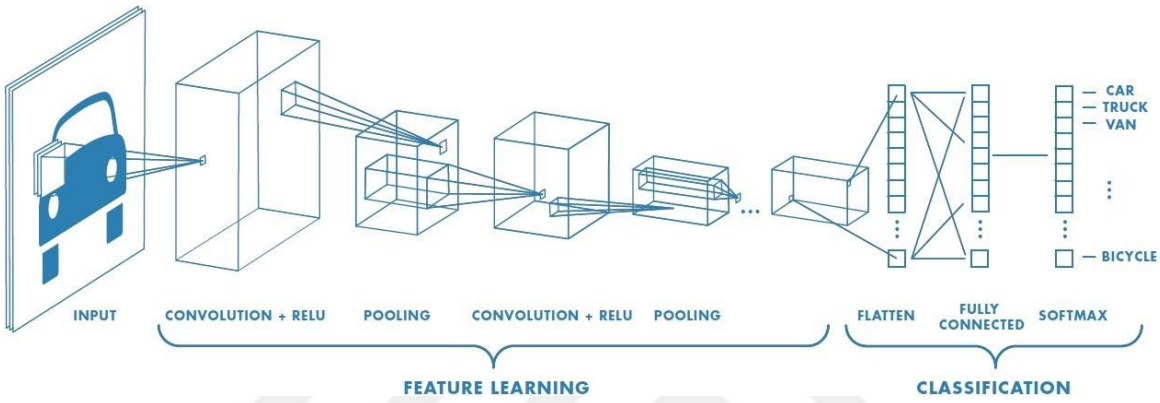

Figure 3.4. A typical CNN Architecture [30].

As illustrated in Figure 3.4 shows a typical CNN Architecture commonly used for image processing challenges.

e) Maximum pooling takes the maximum input of a particular convolved feature map.

f) Average pooling: This takes the average feature map values or summarizes the presence of a feature. The concept of pooling is to reduce image dimensionality, provided all the essential parameters are reserved.

- Relu: Relu acts as an activation function by ensuring non-linearity as the data moves through the network. The dimensionality that must be preserved in the data input into each layer would be lost without Relu.
- Fully connected layer: This layer allows classification tasks to be carried out on the dataset based on the extracted features obtained from the previous layers and their corresponding filters.

3.1.9. Types of Convolutional Neural Network

Pre-trained CNNs come in a variety of architectures. However, these architectures use the same universal framework of applying convolutional layers to the input data to build feature maps. Here are some of the most often used CNN architectures for feature extraction in various deep learning tasks, spanning simple applications to complex challenges. AlexNet, GoogLeNet, LeNet, and ResNet are some of the CNN architectures.

3.1.10. LeNet

Yann LeCun et al. [31] created the LeNet framework for convolutional neural networks in 1998. It has been in use ever since. "LeNet" generally refers to LeNet-5, a primary convolutional neural network with no hidden layers. They are a feedforward neural network in which artificial neurons can respond to a subset of the surrounding cells within their coverage range. Convolutional neural networks perform well in large-scale image processing because their artificial neurons can respond to a subset of the surrounding cells within their coverage range. This CNNs is made up of three (3) convolutional layers, two (2) pooling layers, one (1) fully connected layer, and one (1) output layer, making a total of seven (7) layers depth. Furthermore, activation functions are added between the NN layers in order to solve the problem of not being able to perform nonlinear operations. This type of activation function essentially serves as a trigger to the neuron, which is accomplished using a function known as mapping, similar to other types of the CNN approach. In addition, it enhances the neural network's ability to deal with nonlinear challenges that may develop in the future. LeNet accepts images with a resolution of 32 × 32 pixels.

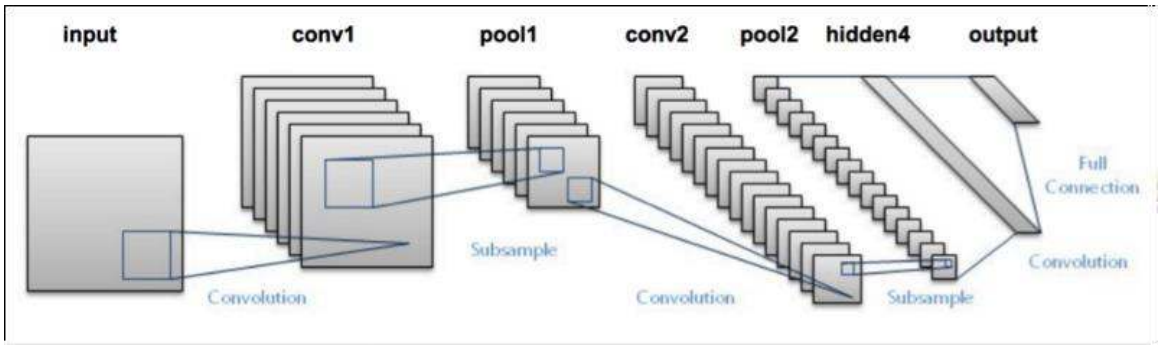

Figure 3.5. Examples of LeNet-5 Architecture, a convolutional neural network [31].

Figure 3.5 shows the general architecture of LeNe-5, a kind of convolutional neural network with no hidden layer.

Table 3.1. provided detailed information about the model architecture of the LeNet-5 convolutional neural network

Table 3.1 Architectural information of CNN-LeNet-5

Layer	Layer Name	Layer Type	Output Size	Feature Map	Kernel Size/Stride
1	**Input**	**Image**	32 x 32	1	
2	'Conv1'	Convolution	28 x 28	6	5 x 5 / 1
3	'Avg Pooling'	Average Pooling	14 x 14	6	2 x 2 / 2
4	'Conv2'	Convolution	10 x 10	16	5 x 5 / 1
5	'Avg Pooling'	Average Pooling	5 x 5	16	2 x 2 / 2
6	'Conv3'	Convolution	1 x 1	120	5 x 5 / 1
7	FC		84		
8	FC	Output	10		

3.1.11. GoogleLeNet

GoogLeNet is a CNN-based network that has been pre-trained and is 22 layers deep, with an image input size of (224 ×224 × 3). Also used to train this network was the ImageNet dataset, which classifies images into 1000 output categories, including monitor, cup, pencil, and various other items and animals. Among the most frequently utilized networks is the GoogLeNet (Inception Network). It has learned several characteristic rendering strategies for a range of images. It is known as the Inception architecture. This network is based on discovering the ideal local sparse structure in a convolutional network, which may be calculated by adding dense layer components to the convolutional network. The name GoogLeNet was chosen as a tribute to Yann LeCuns, a pioneer in constructing the LeNet-5 pre-trained network and was instrumental in its evolution. A group of Google researchers created the Inception network in 2014, and it is still in use today. This network won first place in the 2014 classification and detection competition in the ImageNet challenge. The Inception model is composed of a cardinal unit known as an Inception cell, the smallest unit in the model. This unit carries out a succession of convolutions at different scales. In this model, (1 × 1) convolutional filter reduces the distance between the input channels from top to bottom. As a result, for each cell on the input, the network learns a succession of (1 × 1), (3 × 3), and (5 × 5) filters, which are then used to extract features with varying weights from the input parameters, one for each cell on the input, as shown in Table 3.2, represent the detail information.

Table 3.2. Architectural information of CNN-GoogleLeNet (Inception)

Layer	Layer Type	Layer Information	Layer Depth
1	**Input**	224 x 224 x 3	1
2	Convolution	112 x 112 x 64 with patch size 7 x 7 and stride 2	1
3	Max Pool	56 x 56 x 64 with patch size 3 x 3 and stride 2	0
4	Convolution	56 x 56 x 192 with patch size 3 x 3 and stride 1	2
5	Max Pool	28 x 28 x 192 with patch size 3 x 3 and stride 2	0
6	Inception (3a)	28 x 28 x 256-layer output size	2
7	Inception (3b)	28 x 28 x 480-layer output size	2
8	Max Pool	14 x 14 x 480 with patch size 3 x 3 and stride 2	0
9	Inception (4a)	14 x 14 x 512-layer output size	2
10	Inception (4b)	14 x 14 x 512-layer output size	2
11	Inception (4c)	14 x 14 x 512-layer output size	2
12	Inception (4d)	14 x 14 x 528-layer output size	2
13	Inception (4e)	14 x 14 x 832-layer output size	2
14	Max Pool	7 x 7 x 832 with patch size 3 x 3 and stride 2	0
15	Inception (5a)	7 x 7 x 832-layer output size	2
16	Inception (5b)	7 x 7 x 1024-layer output size	2
17	Average Pool	1 x 1 x 1024 with patch size 7 x 7 and stride 1	0
18	FC	1 x 1 x 1024	0
19	Dropout	1 x 1 x 1024	0
20	Linear	1 x 1 x 1024	1
21	SoftMax	1 x 1 x 1024	0
22	Output		

3.1.12. AlexNet

AlexNet is a pre-trained CNNs that has gained a rich feature representation to classify various images ranging from various applications. The network was developed by [32]. The basic structure of AlexNet is identical to that of the LeNet-5, although this model is much larger. AlexNet's rise to prominence began when it won first place in the 2012 ImageNet competition; this performance persuaded the AI community to embrace deep learning techniques in computer vision problems. The AlexNet input layer receives an input with a size of (227 x 227 x 3). Table 3.3 depicts a typical AlexNet network architecture.

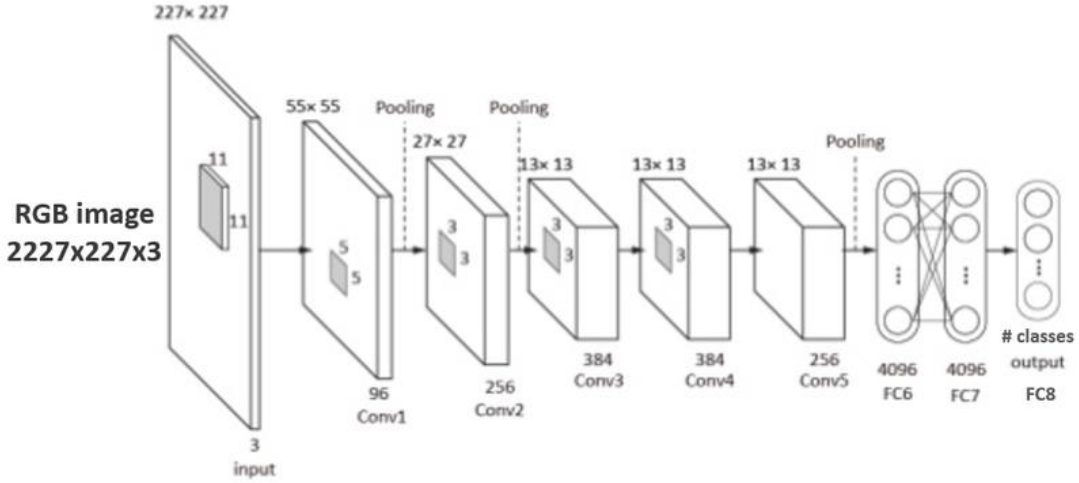

Figure 3.6 Architecture of AlexNet-CNNs [33]

As shows in Figure 3.6, AlexNet has eight weighted layers; the first five are convolutional layers. The final three are fully connected.

Table 3.3. Architectural information of AlexNet-CNNs.

Layer	Layer type	Layer Name	Layer Information
Input	**Image**	**'input'**	**227 x 227 x 3 output size**
1	Convolution	'conv1'	55 x 55 x 96 output size, 11 x 11 kernel size, stride 4, ReLU activation functions
	Max Pooling	'pool1'	27 x 27 x 96 output size, 3 x 3 kernel size, stride 2, ReLU activation functions and padding [0 0 0 0]
2	Convolution	'conv2'	27 x 27 x 256 output size, 5 x 5 kernel size, stride 1, ReLU activation functions and padding [2 2 2 2]
	Max Pooling	'pool2'	13 x 13 x 256 output size, 3 x 3 kernel size, stride 2, ReLU activation functions and padding [0 0 0 0]
3	Convolution	'conv3'	13 x 13 x 384 output size, 3 x 3 kernel size, stride 1, ReLU activation functions and padding [1 1 1 1]
4	Convolution	'conv4'	13 x 13 x 384 output size, 3 x 3 kernel size, stride 1, ReLU activation functions and padding [1 1 1 1]
5	Convolution	'conv5'	13 x 13 x 256 output size, 3 x 3 kernel size, stride 1, ReLU activation functions and padding [1 1 1 1]
	Max Pooling	'pool5'	6 x 6 x 256 output size, 3 x 3 kernel size, stride 2, ReLU activation functions
6	Fully Connected Layer	'fc'	9216 output size, ReLU activation functions
7	Fully Connected Layer	'fc'	4096 output size, ReLU activation functions
8	Fully Connected Layer	'fc'	4096 output size, ReLU activation functions
Output	Classification Layer	'output'	SoftMax activation functions

3.1.13. VGG

Visual Geometry Group (VGG) is a deep CNN with various levels of layering and shapes. An example is VGG16 and VGG19, each with 16 and 19 convolutional layers. In the context of

large-scale image recognition, the authors in [34] investigated the effect of the depth of the convolutional network on its accuracy. In particular, they present a thorough evaluation of networks with increasing depth using an architecture with a small (3x3) convolution filter, demonstrating that increasing the depth to 16–19 weighted layers significantly improves over the prior state-of-the-art setups. As a result of these findings submitted an entry to the ImageNet Challenge 2014, where their team was awarded first and second place in the localization and classification tracks, respectively. The VGG architecture serves as the foundation for groundbreaking algorithms in object recognition. Furthermore, it is one of the most widely used image recognition architectures today.

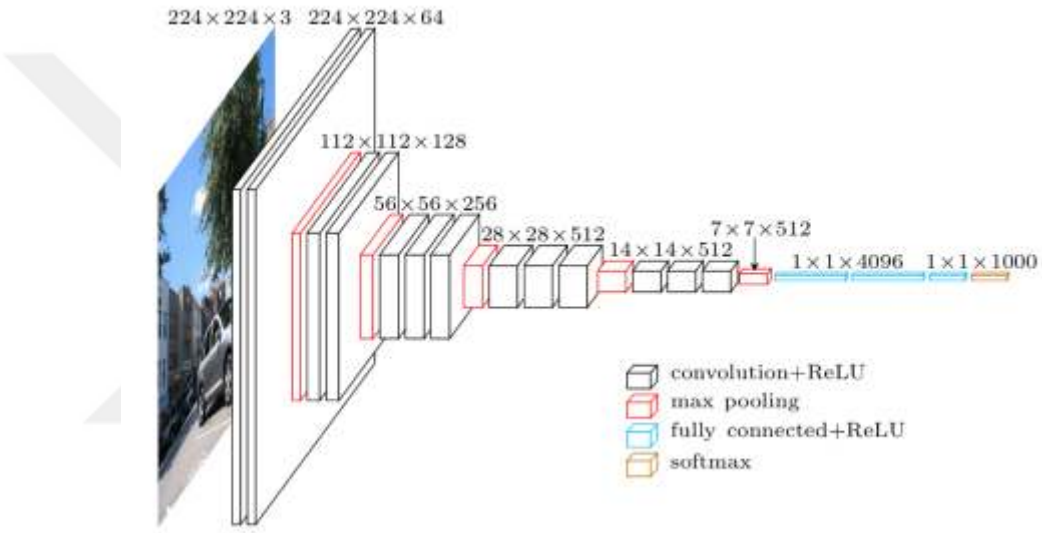

Figure 3.7 The architecture of VGG-CCNs [35].

The input parameter for this CNN-based VGG network is a (224 x 224 x 3) pixel image, as represented in Figure 3.7, which depicts the VGG network's overall architecture.

3.1.14. Resnet

ResNet-18 is a deep convolutional neural network (DCNN) with 18 layers of depth created by Kaiming He et al.[36] The network has been pre-trained to categorize images into 1000 object categories, including keyboards, mouse, pencils, and various animal species. As a result of its training, the network has amassed a rich set of feature representations for various images. Images having a resolution of 224 × 224 pixels are acceptable on the network. The residual CNN-based network approach adds alternate routes as a link on every two layers, similar to the VGG network approach's architectural representation. Other types of these residual networks include Resnet50 and Resnet101.

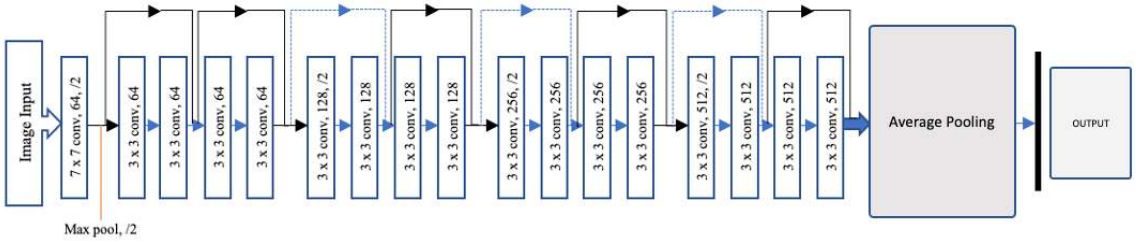

Figure 3.8 Architecture of ResNet-18.

As illustrated in Figure 3.8, it is an architecture of the residual network-18.

3.2. Recurrent Neural Network

RNNs are a form of ANN (Artificial Neural Network) that employs sequential and time-series data to learn from the past to enhance network performance on present and future inputs. These DL algorithms are often used in temporal and ordinal issues, including NLP, speech recognition, and language translation. The presence of loops and the hidden state distinguishes RNNs from other DL algorithms, making it possible for the network to store previous information in its hidden state and operate on sequences because of the looping structure. Typical feedforward NNs are only suitable for data points unrelated to one another.

However, suppose the data is presented in a sequence in which one data point is dependent on the preceding data point. In that case, it is necessary to change the neural network to account for the dependencies between the data points. As a solution to this problem, RNNs are provided with a sort of memory that allows them to retain information about or maintain the state of prior inputs while generating the subsequent output in the sequence. As a result of these characteristics, RNNs are suitable for solving a wide range of problems involving sequential data at different lengths.

During each time step, the hidden state of the cell acts on the inputs to produce the output, and the hidden state is transmitted to the next step in the process.

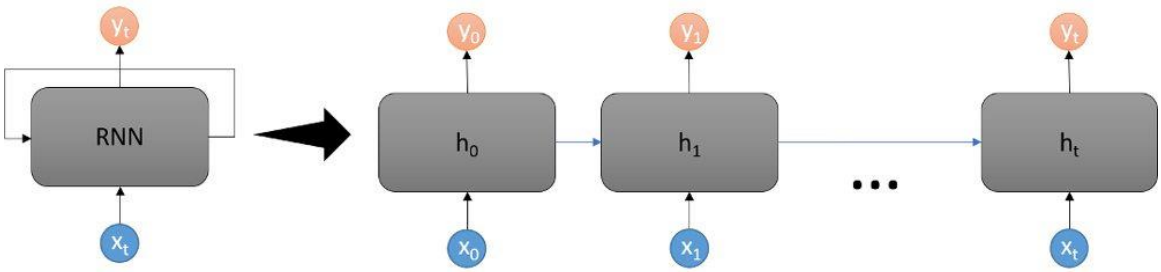

Figure 3.9. The basic architecture of a single cell RNN [37].

Figure 3.9. It shows a single cell of an RNN being unrolled to demonstrate how information passes through the network in the context of a data sequence.

The network consists of hidden state weights and other weights for the input. A neural network's learning process involves determining the weights that should be applied to both the inputs and the hidden state. The result is based on the current input. The hidden state is based on the previous inputs when the method is implemented.

3.2.1. Types of RNNs

There are different types of RNNs depending on the network architecture, as shown in Figure 3.10, Figure 3.11, and Figure 3.12, respectively.

- One–to–one (1-1).

Figure 3.10. A simple Diagram of 1-1 RNNs architecture.

As shown in Figure 3.11, there is only one pair (xt, yt), and it is used in traditional RNNs.

- One-to-many (1-M)

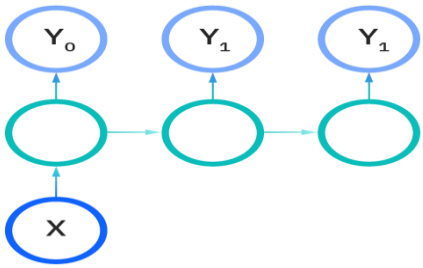

Figure 3.11. A simple Diagram of 1-M RNNs architecture.

For example, a single input can yield several outputs in a 1-M network. One-to-many networks are used in the composition of music.

- Many–to–many (M-M)

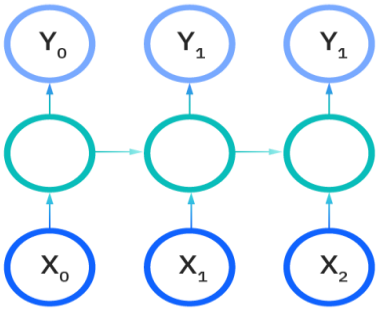

Figure 3.12. A simple Diagram of 1-M RNNs architecture.

There are several alternatives for many to many. An example of this is seen in the above figure, in which many inputs result in many outputs. Machine translation makes use of many-to-many networks.

Table 3.4. Advantages and disadvantages of RNNs.

Advantages	**Disadvantages**
It can handle sequencing data.	It requires high computational power.
It can recall and store historical information.	In order to make choices, the network does not consider future inputs.
It can deal with inputs at different lengths.	In a situation where a gradient is employed for the purpose of computing the weight of the network, the network may fail to update its weights due to the gradient vanishing problem.

Table 3.4 provides a summary of the advantages and disadvantages of recurrent neural networks.

3.3. Long Short-Term Memory

LSTM is a type of artificial RNNs architecture [38] that is used in the field of DL. LSTM, unlike traditional feedforward NNs, contains feedback connections that can analyze the whole data sequences (speech and video) and the individual data elements such as images. Because of its ability to retain previous inputs, LSTM, like other forms of recurrent networks, is commonly used in time-series prediction issues. LSTM networks are a machine learning algorithm that explains how to deal with sequence and time-series data for classification and regression. Using this network to categorize sequence data is a realistic example of LSTM use. Because of their high efficiency and effectiveness, LSTM networks are efficient and successful in classifying, processing, and predicting sequential data. As a result, these types of RNNs are considered among the most powerful, if not the most powerful, classifiers available in the AI and deep learning communities.

Four components make up a general LSTM unit, which is as follows: cell A memory cell is a component that stores values for a random period. Its components include an input gate, an output gate, and a forget gate. These components' primary function is to govern the data and information flowing into and out of the cell. The LSTM network has specific, unique units commonly referred to as memory blocks. These particular units are generally located in the recurrent hidden layer of the LSTM RNN-based network and are used to store data. In order to store the temporary state of the LSTM, these memory blocks are constructed from memory cells connected by self-connecting connections.

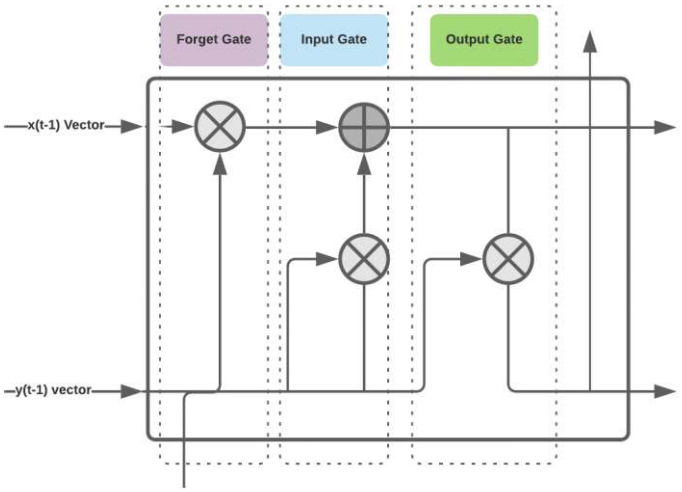

Figure 3.13. The architecture of the LSTM cell.

A diagram representing LSTM network architecture used in the field of DL is shown in Figure 3.13.

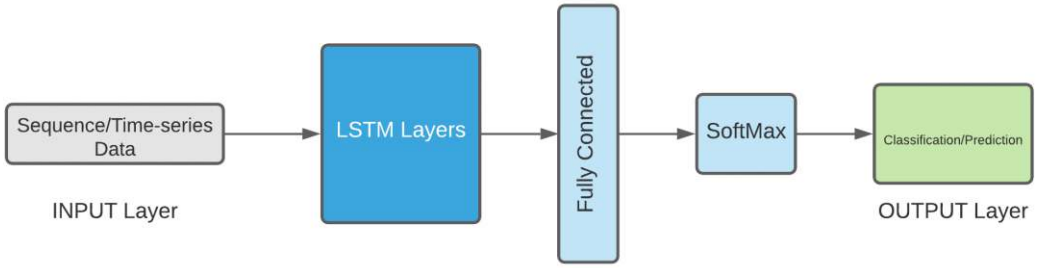

Figure 3.14. LSTM network architecture for a classification problem

A diagram representing a basic RNN-based LSTM network architecture that addresses a classification problem is shown in Figure 3.14.

4. MATERIAL AND METHOD

4.1. Dataset

In this research, the Breakhis image database was used. It is a breast cancer image database first introduced by Spanhol et al. [8], available free on https://web.inf.ufpr.br/vri/breast-cancer-database/. The datasets consist of 7909 instances of histopathology images, collected from 82 patients at P & D Lab Brazil, composed of benign and malignant images, each of 700×460×3 size. The image database consists of 2480 non-cancerous images (benign) and 5429 cancerous images (malignant) obtained under different microscope magnification levels, such as 40×, 100×, 200×, and 400×, as shown in Figure 4.3. The image samples were created from biopsy slides of breast tissue and stained with H&E (hematoxylin and eosin) under different magnification levels. Each of the cancerous and non-cancerous breast cancer types is further categorized into subtypes by the following: Adenosis, represented by a letter (A), Fibroadenoma, represented by a letter (F), Tubular adenoma, represented by letters (TA), Phyllodes tumor, represented by letters (PT) for non-cancerous cancer, while Ductal carcinoma, represented by letters (DC), Lobular carcinoma, represented by letters (LC), Mucinous carcinoma (MC), and Papillary carcinoma (PC), for cancerous types, as depicted in Figure 4.2. All the images in the database have been encoded into RGB channels of 8-bit each, in PNG format. Table 4.1 represents information about the distribution of images in the Breakhis image database.

Table 4.1 Distribution of breast cancer images in the Breakhis image database based on the level of magnifications.

Classes	Sub-classes	40×	100×	200×	400×	Total
Benign	Adenosis (A)	114	113	111	106	444
	Fibroadenoma (F)	253	260	264	237	1014
	Phyllodes (Tumor)	109	121	108	115	453
	Tubular (Tumor)	149	150	140	130	569
Malignant	Ductal carcinoma (DC)	864	903	896	788	3451
	Lobular carcinoma (LC)	156	170	163	137	626
	Mucinous carcinoma (MC)	205	222	196	169	792
	Papillary carcinoma (PC)	145	142	135	138	560
Total	8 Sub-classes	1995	2081	2013	1820	7909

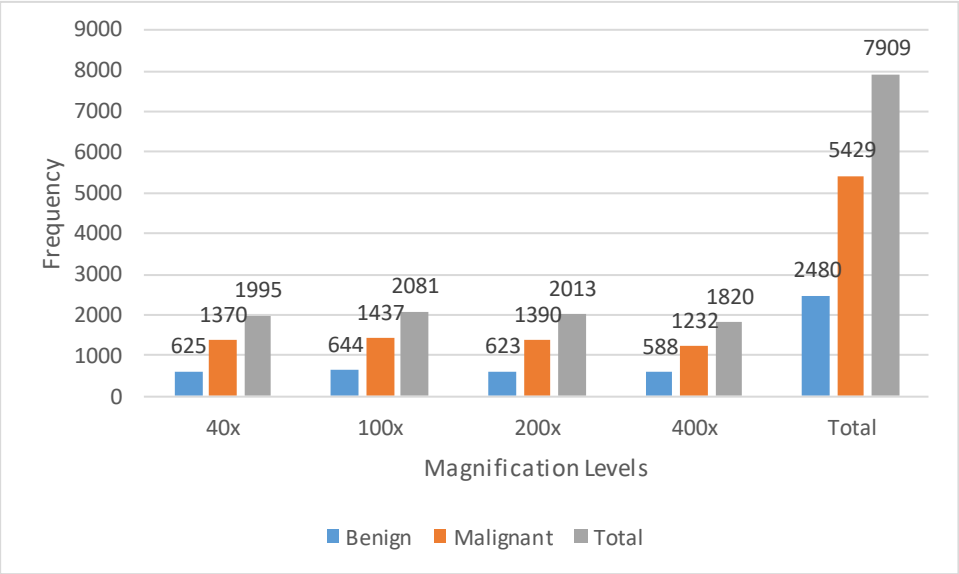

Figure 4.1. Analysis of the benign and malignant in the Breakhis image database.

Figure 4.1, shows the analysis of the number of sample images in the dataset collected under different level of magnifications.

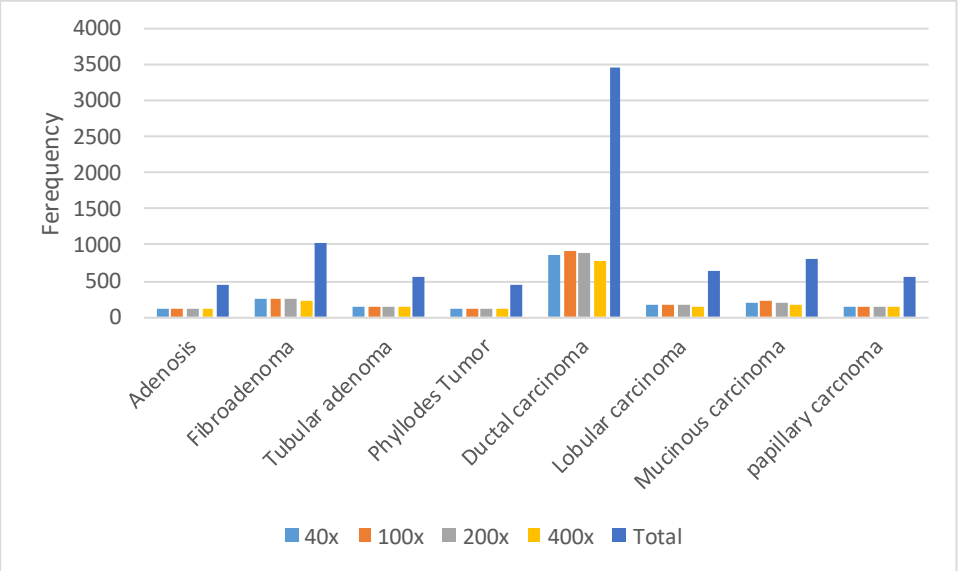

Figure 4.2. Analysis of the benign and malignant subtypes in the BreaKhis image database.

Figure 4.2 shows the analysis of the number of breast cancer subtypes sample images in the dataset collected under different level of magnifications.

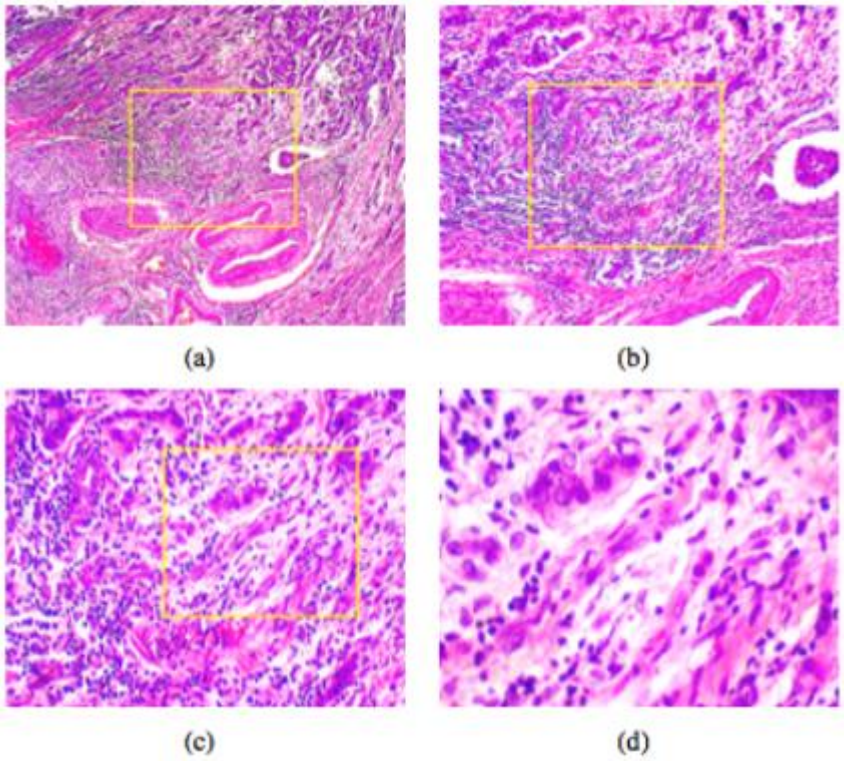

Figure 4.3. Samples of the malignant tumor under different magnification levels from the BreaKhis image database[39].

4.2. Deep feature extraction and selection

4.2.1. Feature extraction using pre-trained CNNs

Features extraction is an important phase in processing an image. As a result, CNNs are the most commonly used and most influential networks in computer vision. It shows how an image is changed into a set of features. The extraction of features from histological images is a critical and challenging part of the computer-aided detection of breast cancer using histopathology images. Several machine learning approaches for extracting features from histopathology images have been developed and implemented in recent years. It is suggested in [40] to use convolutional neural networks to extract features from histopathological images using a new nucleus-guided feature extraction approach. The nuclei are first recognized from images and then utilized to train a convolutional neural network with three hierarchy structures that have been built explicitly for this task. Image features such as the nuclei pattern and spatial distribution are retrieved using the trained network to analyze the image. Evaluation of the suggested characteristics is carried out by a classification experiment on a histological image database of breast lesions.

The experimental findings demonstrate that the extracted features accurately depict histopathological images. The proposed framework achieves a superior classification performance for breast lesions than the state-of-the-art approaches evaluated in the comparison. Transfer

learning-based feature extraction from raw image databases is the first phase of the proposed model framework, as shown in Fig 4.4. The raw images from the public benchmark datasets are first pre-processed by resizing them into a uniform size that the networks or algorithms will accept as input. The supplied image contains excessive information that is not required for categorization. As a result, after pre-processing the image, the first step is to simplify it by extracting critical information and discarding non-essential data. This process simplifies and accelerates the process of identifying images based on their features. In this process, the proposed DRNet model considers nineteen (19) pre-trained networks such as ResNet50 [36], ResNet101 [36], ResNet18 [36], Vgg16 [41], Vgg19 [41], AlexNet [41], DarkNet19 [42], DarkNet53 [42], Mobilenetv2 [43], Nasnetlarge [44], Nasnetmobile [44], Xception [45], InceptionResNetTv2 [46], EfficientnetB0 [47], Shufflenet [48], Inceptionv3 [49], Densenet201 [50], Googglenet [51], and Squeeznet [52], denoted by N1...N19, were utilized as feature extractors. The raw input images were fed to feature extraction algorithms. The convolutional neural network will then learn the features of the images in the convolutional layers of the pre-trained networks. The extraction algorithms produced the final feature vectors that contained all the lists of features F1...F19. The learned features will then be passed to the fully connected layer of the network for classification. Then, the best two performing feature vectors were selected based on the calculated SVM's loss values. The extracted features can then be input into an ML model like k-NN or SVM for classification. Table 4.2 shows the number of extracted features based on the magnification levels.

Table 4.2. Deep feature extractions based on magnification levels.

Magnification levels	Used classifiers	Number of extracted features
40×	KNN and SVM	369
100×	KNN and SVM	392
200×	KNN and SVM	643
400×	KNN and SVM	346
Total		1750

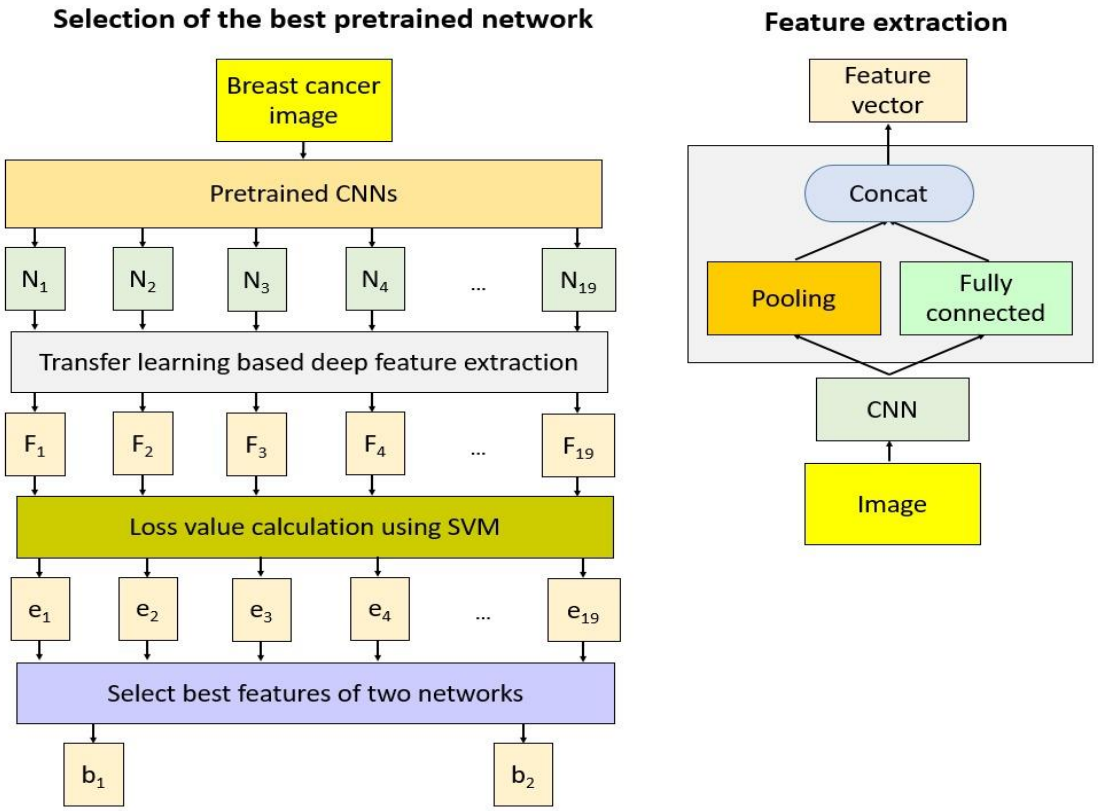

Figure 4.4. A framework of the pre-trained CNNs based feature extraction.

Figure 4.4 shows a framework of the used pre-trained convolutional neural networks used to extract features from the row image data.

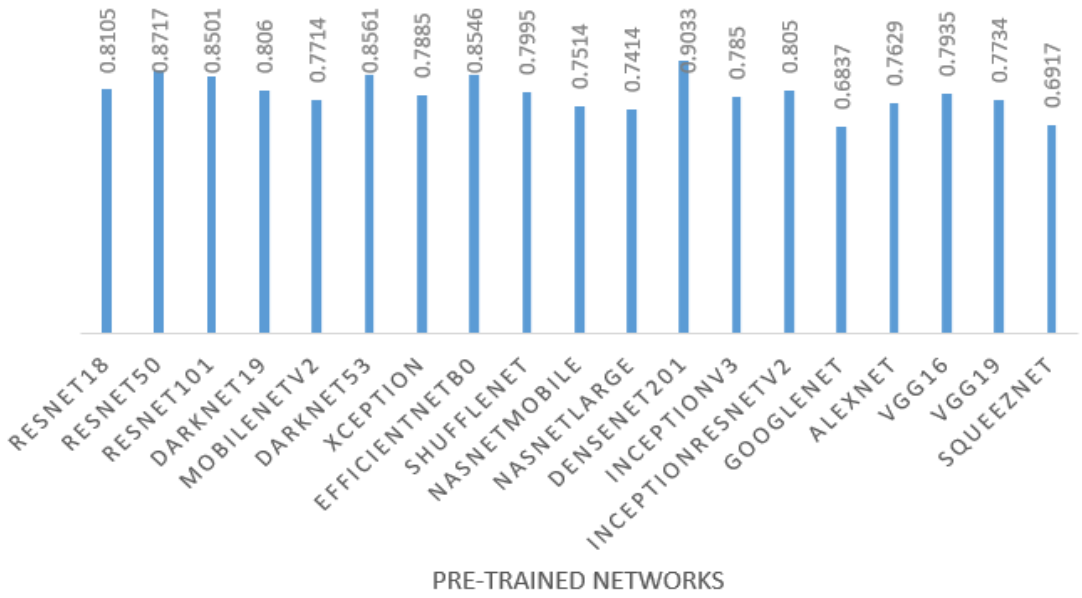

Figure 4.5. Performance of the used pre-trained CNNs used for feature extractions on Breakhis dataset.

Figure 4.5 shows the performance accuracy of these networks used to extract features.

4.2.2. Algorithm of the proposed model framework-based feature extraction.

4.2.3. Steps for the pre-trained CNNs based feature extraction.

The following are the four (4) steps for the model framework:

Step 1: Read input images from the Breakhis image database.

Step 2: Employ 19 pre-trained CNNs for feature extraction.

Step 3: Calculate the misclassification rates of the generated features of each CNN using the SVM Classifier.

Step 4: Choose the best two performing CNNs for feature extraction.

4.2.4. Feature transformation engineering

Among the rising study areas in the biomedical domain is image processing, which is proliferating among researchers. In image processing, various strategies effectively extract relevant information for analytical purposes. However, they are also efficient in computing time and memory storage. One sort of image processing technology that falls under this category is transformation. Examples of transformation techniques are Hilbert, Fourier, Radon, and wavelet transform [53]. DWT has recently gained popularity for analyzing, de-noising, and compressing signals and images among researchers [54]. This work utilizes a multilevel discrete wavelet transform to decompose the extracted images in the Breakhis image database into five (5) low-level sub-hand images, as depicted in Figure 3.5. According to the model framework results in Figure 4.5, among the 19 pre-trained CNNs feature extractors, ResNet50 and DenseNet201 outperformed other CNNs employed in this study. The best two performing pre-trained CNNs in consideration were then combined (concatenated) to form the proposed hybrid DRNet model, a combination of ResNet50 and DenseNet201, as depicted in Figure 4.6.

4.2.5. Feature selection engineering

The challenge with microscopic datasets is high dimensionality, resulting in poor generalization and a long execution time, necessitating feature selection [55]. The most exciting concept in selecting features adopted in this study is eliminating the extracted features that are no longer relevant in the classification task, thereby saving additional resources. There are various feature selection algorithms in the literature, including PSO, PCA, and relief-based. This study utilized iterative neighborhood component analysis (INCA) to choose the minimum number of features as a feature selector. It was first presented in the study of Tuncer et al. [56] as an upgraded

version of PCA to overcome feature redundancy and the failure to generate the optimal number of features in the PCA working mechanism. INCA-based feature selection is illustrated in Figure 4.6.

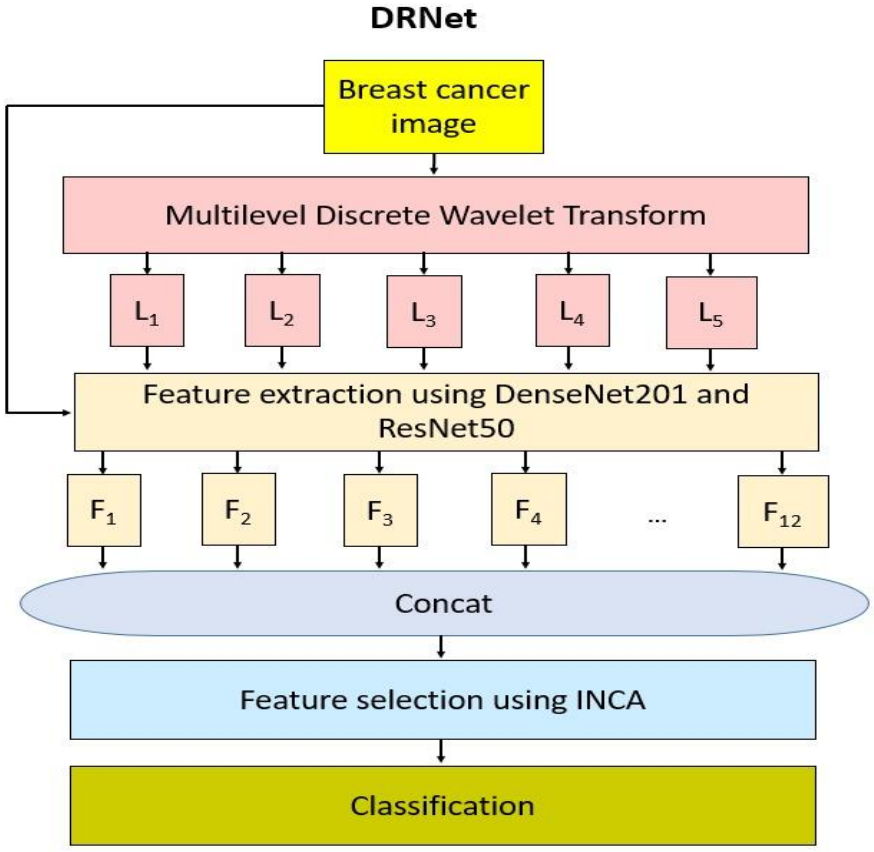

Figure 4.6. The proposed DRNet model for deep feature extractions of benign and malignant BC subtypes.

4.2.6. Classification techniques

Classification of BC subtypes is one of the most challenging tasks due to similarity and the high number of features among the various subtypes of histopathological BC images, making it challenging to have an accurate class prediction for different cancer subtypes. According to Sotiriou et al. [57], different cancer subtypes respond to different treatments. The authors further suggested that it is crucial to classify the cancer disease so that patients can receive more precise and decent treatment while also saving medical costs associated with an unneeded diagnosis. Support vector machines, naive bayes, and k-nearest neighbor are the most frequently utilized approaches in cancer classification problems [55]. In this study, Cubic-SVM and KNN classifiers were adopted to classify the extracted features into different subtypes of benign and malignant BC, respectively.

4.2.7. K-NN classifier

K-nearest neighbor, or k-NN, is one of the types of ML supervised algorithms mostly applicable to regression and classification tasks, as depicted in Figure 4.7, especially classification involving binary problems. For a given known value of K, say, in this case, the classification uses only one nearest neighbor to conclude. For an unknown value of K, say, in this case, the classification uses only the n-nearest neighbor to conclude. When there is more than one neighbor, the one closest is used to categorize the element of interest. The distance between the data and all other clusters is measured to classify a given dataset. There are many distance equations to measure the distance between the data points. Euclidean is the most commonly used to measure distances.

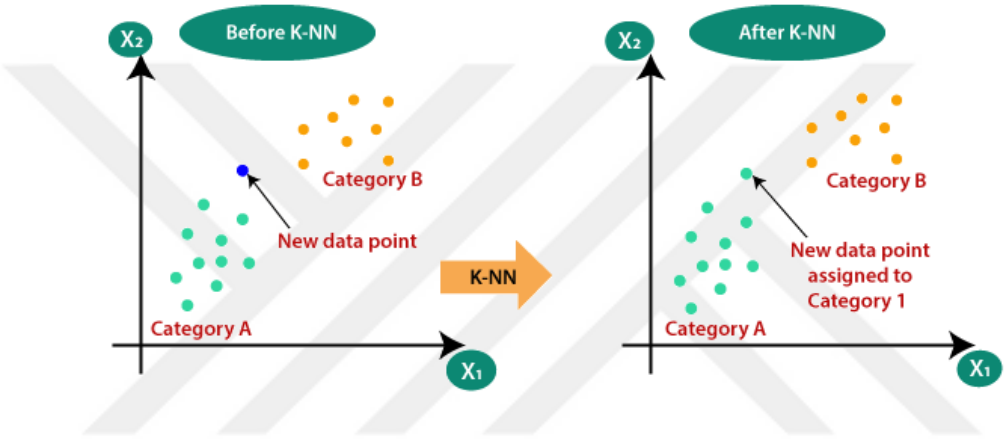

Figure 4.7. An example of KNN algorithm [58].

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{4.1}$$

The Euclidean distance can be measured using equation (4.1) for the two sets of features in the data. For the three sets of features in the data, the distance can be measured as defined by equation (4.2).

$$\sqrt{\sum_{\substack{i=1 \\ j=1}}^{3} (x_i - x_j)^2} \tag{4.2}$$

Where $x_j^1 = feature\ of\ element\ of\ interest$

After calculating, k is chosen, which defends on the closest distances and logical thinking. The most common feature of the nearest neighbor is used to classify it (x^1).

4.2.8. Support vector machine

SVM is one of the most robust tools for multiclass support vector classification and regression problems. Each object to be classified is represented as an object in space -1 , and the coordinates of these points are called features. The classification task in SVM is done by drawing a hyperplane that separates all points from one category of data on one side and all points from the second category on the other side, as can be seen in Figure 4.8 representations of linearly and non-linearly separable data. There could be multiple hyperplanes in a given classification problem in this case. Finding the best hyperplanes that separate the two categories and maximizing the distance between points in each category is called margin. The point that falls exactly on the margin is called the support vector. However, the support vector has a limit in a situation whereby some data points cannot be separated by hyperplanes, necessitating transformation. For example, using a kernel can enable transformation, which will allow a hyperplane to be drawn among the categories. Other methods include hard and soft margins SVM. SVM is also known as the maximum margin classifier. Its job is to find the plane with the maximum separation between samples.

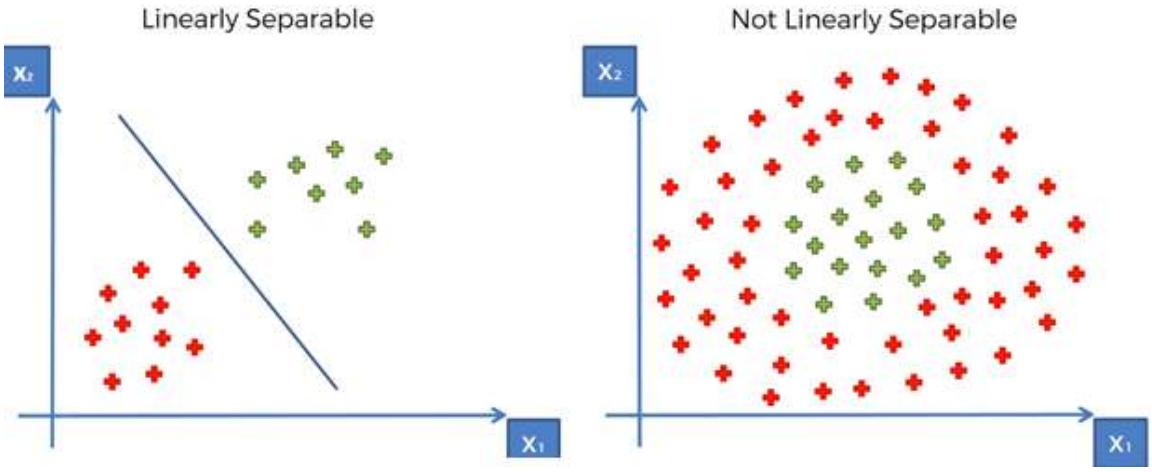

Figure 4.8. Example of linearly and non-linearly seperable data [59].

The margin linear equation of SVM classifier function for a hyperplane can be represented by the following defined equation:

$$\bar{w} \,.\, \bar{x} + b = 0 \tag{4.3}$$

Where $\bar{w}$ is the moving vector, $\bar{x}$ is the feature, and b is the bias. The primary aim of linear hyperplane is to correctly categorize samples according to the classes. For positive sample, the prediction can be written as:

$$\bar{w} \,.\, \bar{x} + b \geq 0 \tag{4.4}$$

And for negative prediction, the equation can be rewritten as:

$$\bar{w} \,.\, \bar{x} + b \leq 0 \qquad (4.5)$$

However, for non-linearly separable data, it's a situation whereby some data points cannot be separated by hyperplanes, necessitating transformation, through kernel and polynomial kernel methods. Kernel method can be used to replace $K(\bar{x_\iota} \,.\, \bar{x_j})$ in the lagrange defined in equation

$$L = -\sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i \cdot y_j \cdot \bar{x}_i \, \bar{x}_j \qquad (4.6)$$

For two features, the kernel can be rewritten as:

$$K(\bar{x_\iota} \,.\, \bar{x_J}) = (x_1^2 + x_2^2) \qquad (4.7)$$

Such that the Lagrange defined in equation (4.6) becomes:

$$L = -\sum_i \alpha - \frac{1}{2}\sum_i \sum_j \alpha_{i\alpha_j} \cdot y_i y_j \cdot (x_1^2 + x_2^2) \qquad (4.8)$$

The equation defined in (4.8) enables hyperplanes to be drawn in non-linearly separable data. Other techniques are the polynomial kernel and Gaussian kernel. According to [60], the author suggested that SVM is a robust classifier for handling multiclass classification problems, especially BC, as adopted in this study.

.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

5.1. Experimental setup

This section of the research study describes the configuration of the model that has been presented. The suggested DRNet-based feature extraction of breast cancer images from the Breakhis dataset model is coded in the MATLAB 2021b programming environment. In addition, deep neural networks (CNN) were incorporated into the MATLAB 2021b version. The convolutional neural networks (CNNs) used in this study were pre-trained using the ImageNet dataset. Currently, the pre-trained network output layers still use 1000 labels for the classes, so this will cause a mismatching problem as the information is passed without changing the value of the default network output classes. Moreover, the breakhis dataset's original image size is 700 × 460 ×3 dimensions. There is variation in the size of the input images among the pre-trained networks, ranging from 224 × 224 to 229 × 229. However, to address mismatching problems while passing images to the networks, all the input images from the Breakhis dataset were resized to 224 × 224 dimensions for uniformity and accepted as input images. The default number of output layer classes in the pre-trained was replaced with the new eight (8) classes of output layer adapted to the dataset. Other parameters, such as the learning rate, were employed using their default values. The features were extracted using these CNNs. A personal computer with a simple configuration is required. There is no requirement to utilize GPU-supported hardware to run the program. The configuration utilized for the experiment is given in Table 5.1.

Table 5.1. The specifications of the computer that was utilized to experiment on the dataset

Hardware/Software Configuration	Attributes
CPU	Intel @i9 9th generation
RAM	48 GB
HDD	512 GB
Operating system	Windows 10.1 Pro
Programming environment	MATLAB-Based (2021b)

5.1.1. Performance measure evaluations

The performance of the proposed DRNet model is evaluated in terms of accuracy, precision, sensitivity (recall), and F1 score metrics as defined by equations (5.1) to (5.4), based on the confusion matrix obtained for the classification tasks of BC. The confusion matrix acquired will help to measure the overall model performance for the four (4) levels of magnification (40×, 100×, 200×,400×), on the Breakhis dataset. The primary goal is to present novel feature engineering

techniques to extract the most relevant features from these benign and malignant breast cancer subtype images and classify them accordingly, using k-NN and SVM classifiers based on the image magnification levels. The resulting confusion matrix describes the various combinations of actual and predicted class values.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5.1)$$

$$Precision = \frac{TP}{TP + FP} \quad (5.2)$$

$$Recall = \frac{TP}{TP + FN} \quad (5.3)$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5.4)$$

The terms "true positive" (TP) and "true negative" (TN) values utilized in this study represented the number of malignant histopathological instances in the breakhis image database, correctly classified as malignant (cancerous). At the same time, TN refers to the number of benign (non-cancerous) correctly identified by the model as benign. The terms "false positive" (FP) and "false negative" (FN) were used to describe the number of benign and malignant histopathological instances that were wrongly classified as benign and malignant by the model, respectively.

Table 5.2 shows that the proposed DRNet model-based feature extractions attained an average classification rate accuracy of 98.74%, precision of 94.5%, Recall 94.9%, and F1 Score 94.6%, respectively, at a 40× magnification level. Its corresponding confusion matrix is depicted in Figure 5.1. It was used to classify subtypes of benign (A, F, PT, TA) and malignant (DC, LC, MC, PC) images from the Breakhis image database using a k-NN classifier.

Table 5.2. DRNet model evaluation results for 40× magnification level using k-NN classifier on Breakhis.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	115	114	99.85	99.0	98.0	99.0
2	864	864	95.89	95.0	95.0	95.0
3	257	253	99.40	98.0	97.0	98.0
4	154	156	96.59	78.0	79.0	78.0
5	207	205	99.40	98.0	97.0	97.0
6	145	145	99.50	97.0	97.0	97.0
7	101	109	99.50	92.0	99.0	95.0
8	152	149	99.75	99.0	97.0	98.0
Average	-	-	98.74	94.5	94.9	94.6

The number of actual cases in a class is represented by n(truth). In contrast, the number of categorized cases in a class is represented by n(classified). INCA was used to select 369 features

for eight (8) subtypes of benign and malignant breast tumors from the Breakhis image database, then categorized using k-NN. For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the highest classification accuracy.

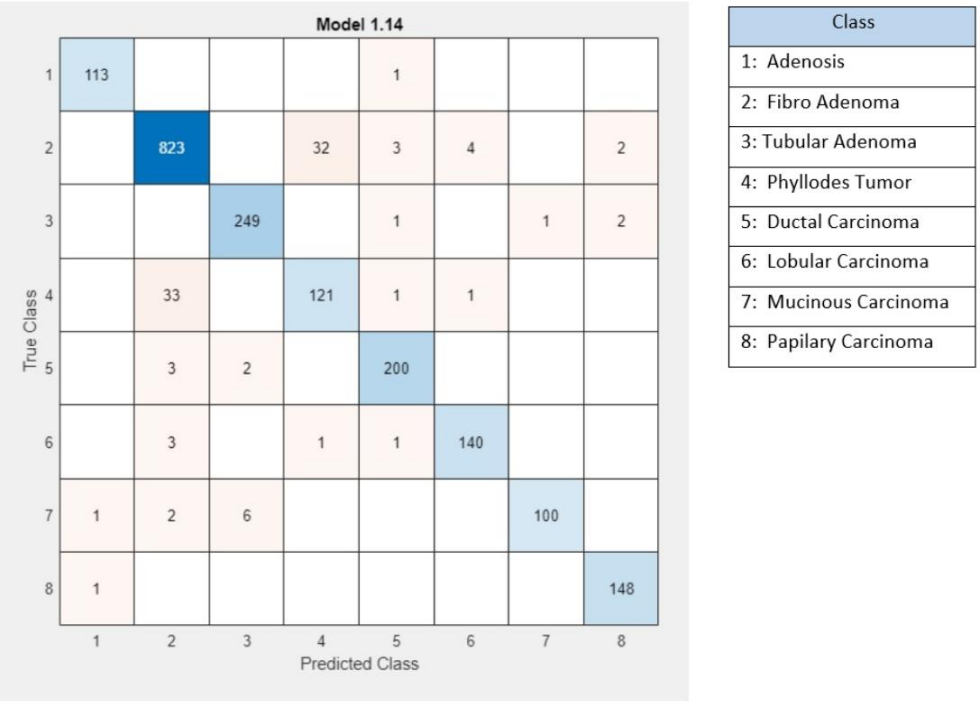

Figure 5.1. Confusion matrix of 40× magnification level acquired using k-NN classifier.

Table 5.3 shows that the proposed DRNet model-based deep feature extractions produced an average classification rate accuracy of 97.71%, a precision of 89.63%, a recall of 89.88%, and an F1 score of 89.75%, respectively, at a 100× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.2, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using the k-NN classifier.

Table 5.3. DRNet model evaluation results for 100× magnification level using k-NN classifier on Breakhis.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	115	113	99.42	96.0	94.0	95.0
2	259	260	98.03	92.0	92.0	92.0
3	123	121	98.56	88.0	87.0	88.0
4	149	150	99.38	95.0	96.0	96.0
5	910	905	93.90	93.0	93.0	93.0
6	173	170	95.44	73.0	72.0	72.0
7	216	222	97.98	89.0	92.0	90.0
8	138	142	98.94	91.0	93.0	92.0
Average	-	-	97.71	89.63	89.88	89.75

INCA was used to select 392 features for eight (8) subtypes of benign and malignant breast tumors from the Breakhis image database, then categorized using k-NN. For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the maximum classification accuracy.

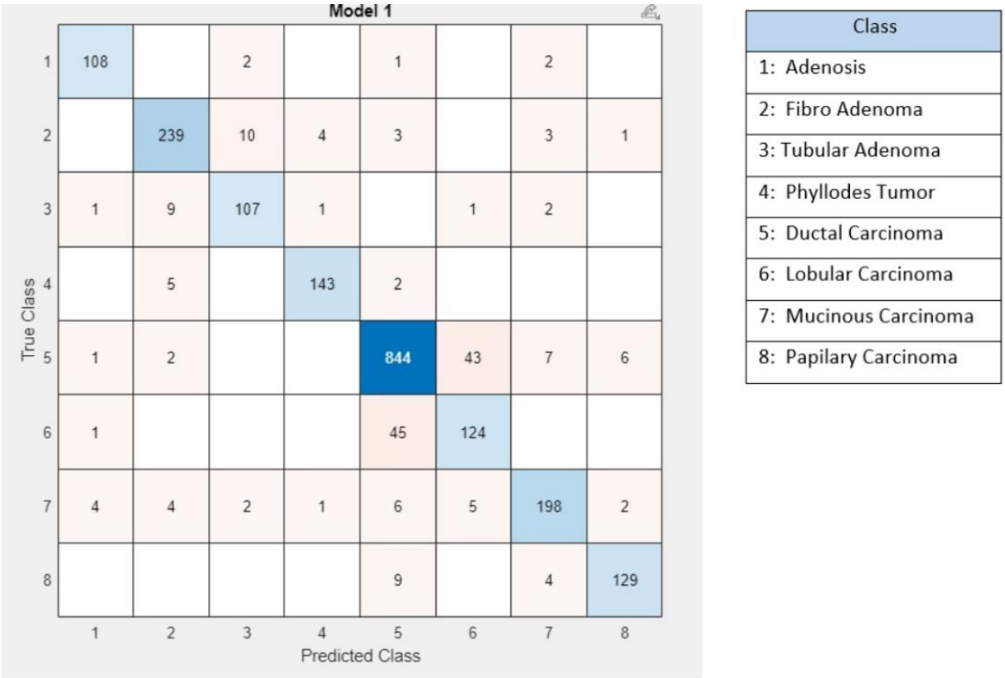

Figure 5.2. Confusion matrix of 100× magnification level acquired using k-NN classifier.

Table 5.4 shows that the proposed DRNet model-based deep feature extractions attained an average classification rate accuracy of 97.30%, a precision of 87.38%, a recall of 87.63%, and an F1 score of 87.64%, respectively, at a 200× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.3, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using a k-NN classifier.

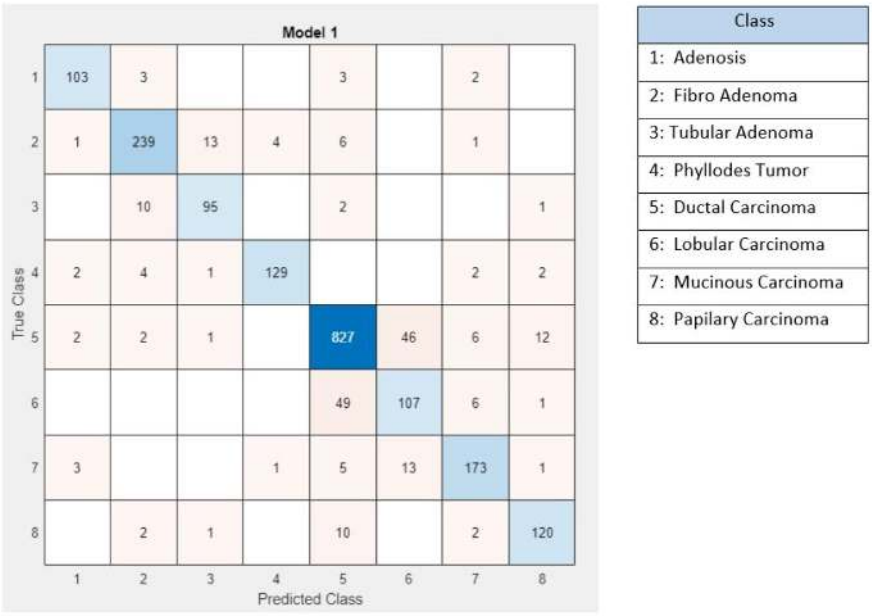

Figure 5.3. Confusion matrix of 200× magnification level acquired using k-NN classifier.

In this process, INCA was used to select 643 features for eight (8) subtypes of benign and malignant breast tumors from the Breakhis image database, then categorized using k-NN. For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the maximum classification accuracy.

Table 5.4. DRNet model evaluation results for 200× magnification level using k-NN classifier on Breakhis.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	111	111	99.21	93.0	93.0	93.0
2	260	264	97.71	91.0	92.0	91.0
3	111	108	98.56	88.0	86.0	87.0
4	134	140	99.21	92.0	96.0	94.0
5	902	896	92.85	92.0	92.0	92.0
6	166	163	94.29	66.0	64.0	65.0
7	192	196	97.91	88.0	90.0	89.0
8	137	135	98.41	89.0	88.0	88.0
Average	-	-	97.30	87.38	87.63	87.64

Table 5.5 shows that the proposed DRNet model-based deep feature extractions produced an average classification accuracy rate of 98.40%, a precision of 89.5%, a recall of 89.25%, and an F1 score of 89.5%, respectively, at a 400× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.4, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using a k-NN classifier.

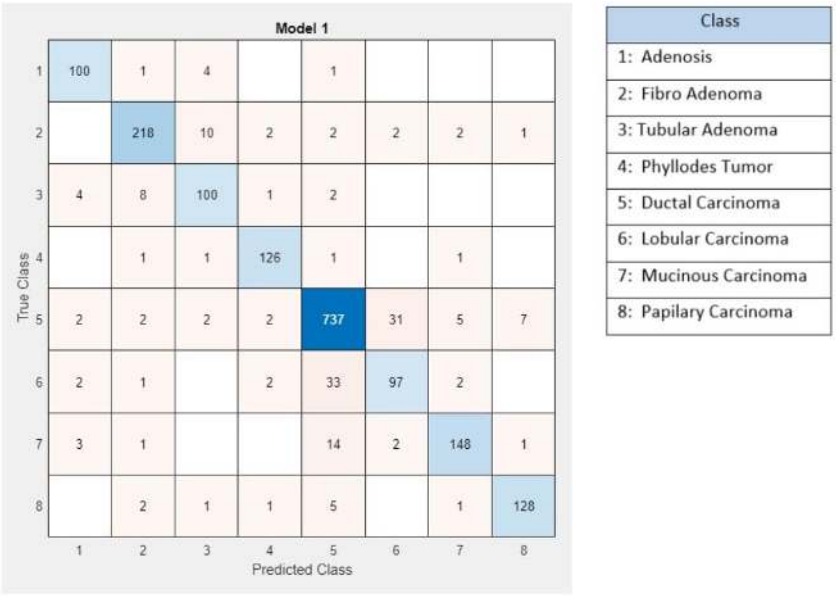

Figure 5.4. Confusion matrix of 400× magnification level acquired using k-NN classifier.

INCA was also utilized to select 346 features from the Breakhis image database for eight (8) subtypes of benign and malignant breast cancers, then classified using KNN. In order to get the maximum classification accuracy, 1-all and 10-fold cross-validation were used.

Table 5.5. DRNet model evaluation results for 400× magnification level using k-NN classifier on Breakhis.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	111	106	99.07	94.0	90.0	92.0
2	234	237	98.08	92.0	93.0	93.0
3	118	115	98.19	87.0	85.0	86.0
4	134	130	99.34	97.0	94.0	95.0
5	795	788	99.01	94.0	93.0	93.0
6	132	137	95.88	71.0	73.0	72.0
7	159	169	98.24	88.0	93.0	90.0
8	137	138	98.96	93.0	93.0	93.0
Average	-	-	98.40	89.5	89.25	89.25

Table 5.6 shows that the proposed DRNet model-based deep feature extractions yield an average classification accuracy rate of 98.61%, a precision of 93.4%, a recall of 94.75%, and an F1 score of 94.0, respectively, at a 40× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.5, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using an SVM classifier.

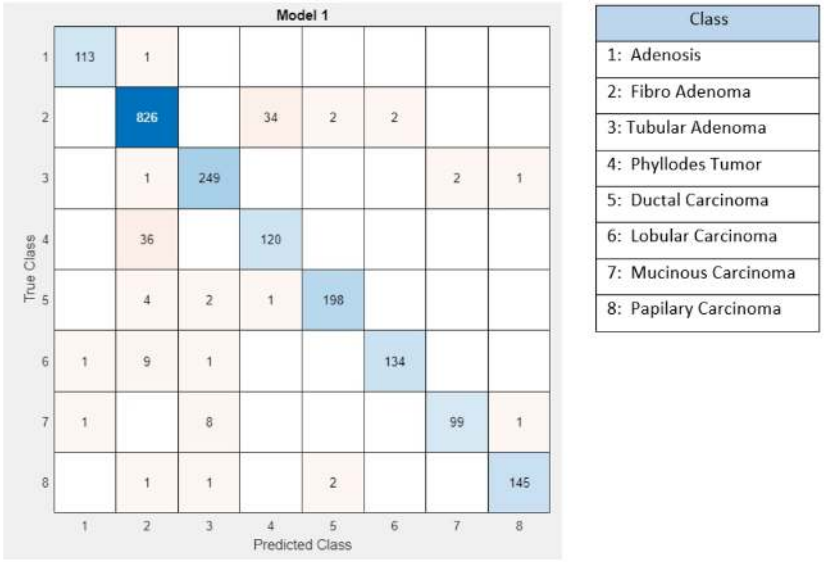

Figure 5.5. Confusion matrix of 40× magnification level acquired using SVM classifier.

In this process, INCA was used to select 369 low-level features for eight (8) subtypes of benign and malignant breast tumors from the Breakhis image database, then categorized using KNN. For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the maximum accuracy.

Table 5.6. DRNet model evaluation results for 40× magnification level using the Cubic-SVM classifier.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	115	114	99.85	99.0	98.0	99.0
2	878	864	95.49	96.0	94.0	95.0
3	261	253	99.20	98.0	95.0	97.0
4	155	156	96.44	77.0	77.0	77.0
5	202	205	99.45	97.0	98.0	97.0
6	136	145	99.35	92.0	99.0	95.0
7	101	109	99.40	91.0	98.0	94.0
8	147	149	99.70	97.0	99.0	98.0
Average	-	-	98.61	93.4	94.75	94.0

Table 5.7 shows that the proposed DRNet model-based deep feature extractions attained an average classification accuracy rate of 98.07%, a precision of 91.0%, a recall of 91.75%, and an F1 score of 91.25%, respectively, at a 100× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.6, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using an SVM classifier.

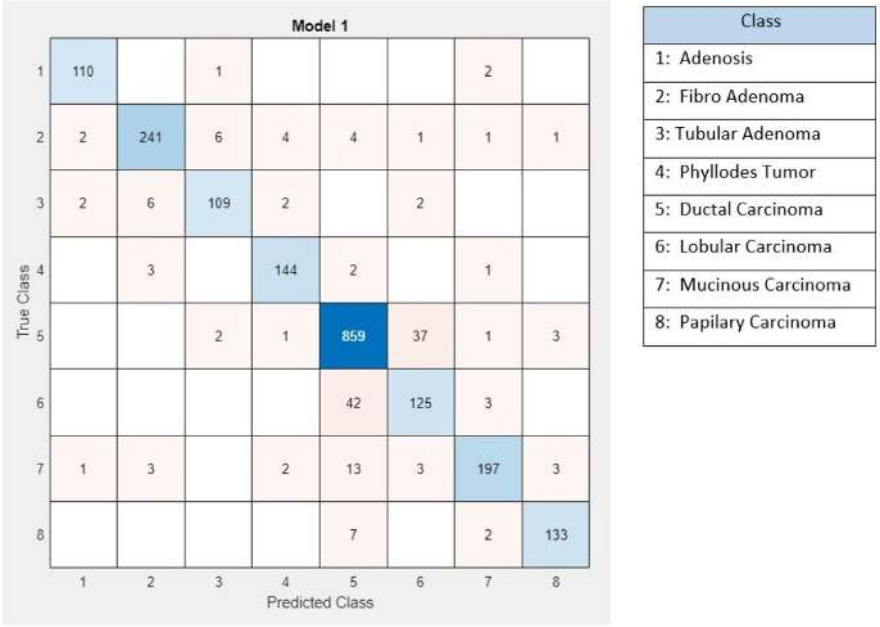

Figure 5.6. Confusion matrix of 100× magnification level acquired using SVM classifier.

For the classification accuracy, 1-all and 5-fold cross-validation were adopted to obtain the maximum accuracy. INCA was also utilized to select 392 low-level features from the Breakhis image database for eight (8) subtypes of benign and malignant breast cancers, then classified using an SVM classifier.

Table 5.7. DRNet model evaluation results for 100× magnification level using the Cubic-SVM classifier.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	115	113	99.62	97.0	96.0	96.0
2	253	260	98.51	93.0	95.0	94.0
3	118	121	98.99	90.0	92.0	91.0
4	153	150	99.28	96.0	94.0	95.0
5	927	903	94.62	95.0	93.0	94.0
6	168	170	95.77	74.0	74.0	74.0
7	207	222	98.32	89.0	95.0	92.0
8	140	142	99.23	94.0	95.0	94.0
Average	-	-	98.04	91.0	91.75	91.25

Table 5.8 shows that the proposed DRNet model-based deep feature extractions produced an average classification accuracy rate of 97.68%, a precision of 88.63%, a recall of 90.0%, and an F1 score of 89.38%, respectively, at a 100× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.7, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using an SVM classifier.

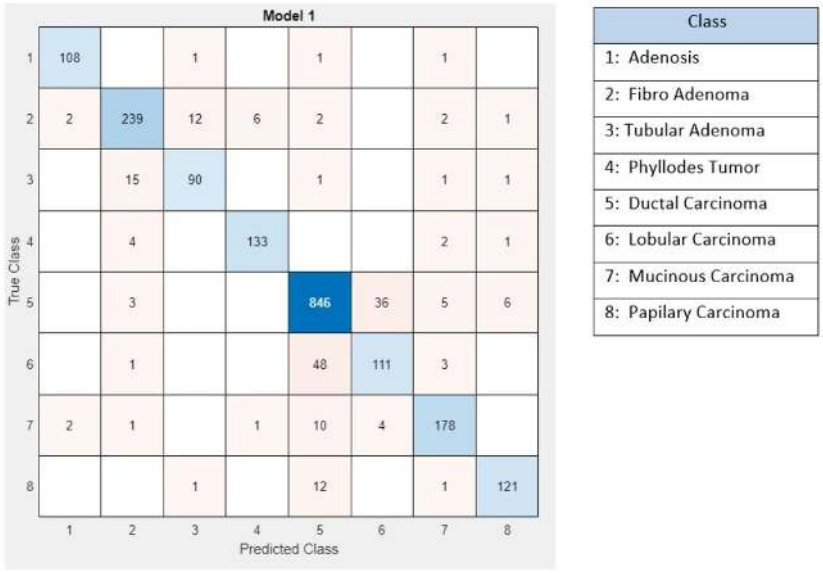

Figure 5.7. Confusion matrix of 200× magnification level acquired using SVM classifier.

For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the maximum accuracy. INCA was also utilized to select 643 low-level features from the Breakhis image database for eight (8) subtypes of benign and malignant breast cancers, then classified using an SVM classifier.

Table 5.8. DRNet model evaluation results for 200× magnification level using the Cubic-SVM classifier.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	112	111	99.65	97.0	96.0	97.0
2	263	264	97.57	91.0	91.0	91.0
3	104	108	98.41	83.0	87.0	85.0
4	140	140	99.30	95.0	95.0	95.0
5	920	896	93.84	94.0	92.0	93.0
6	151	163	95.43	68.0	74.0	71.0
7	193	196	98.36	91.0	92.0	92.0
8	130	135	98.86	90.0	93.0	91.0
Average	-	-	97.68	88.63	90.0	89.38

Table 5.9 shows that the proposed DRNet model-based deep feature extractions yield an average classification accuracy rate of 99.18%, a precision of 97.71%, a recall of 88.5%, and an F1 score of 89.25%, respectively, at a 400× magnification level. Its obtained corresponding confusion matrix is depicted in Figure 5.8, as it was used to classify eight (8) subtypes of benign and malignant images from the Breakhis image database using an SVM classifier.

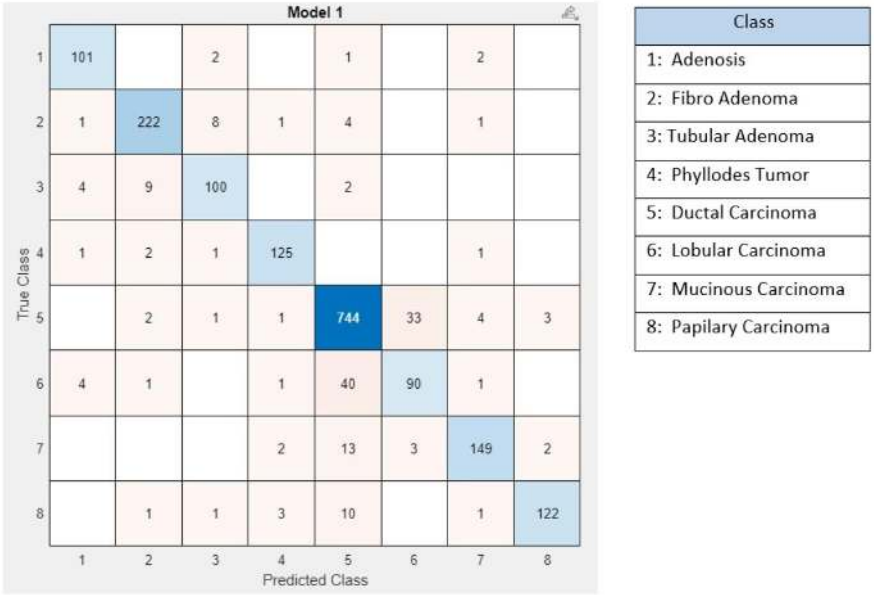

Figure 5.8. Confusion matrix of 400× magnification level acquired using SVM classifier.

For the classification accuracy, 1-all and 10-fold cross-validation were adopted to obtain the maximum accuracy. INCA was also utilized to select 346 low-level features from the Breakhis image database for eight (8) subtypes of benign and malignant breast cancers, then classified using an SVM classifier. Tables 5.2 to 5.5 shows the different performance measure evaluation results obtained for the proposed DRNet model using a k-NN classifier at different magnification levels. Tables 5.6 to 5.9 presents the different performance measure evaluation results obtained using an SVM classifier. For the 40× magnification, the DRNet_k-NN model attains an average classification accuracy rate of 98.74% in identifying different subtypes of benign and malignant BC histopathological images in the breakhis dataset based on the extracted features. In comparison, the DRNet_SVM model attains 98.61% for the 40× magnification. Thus, k-NN proved to perform slightly better than SVM at classifying each of the subtypes of benign and malignant based on the extracted features.

Table **5.9.** DRNet model evaluation results for 400× magnification level using the Cubic-SVM classifier.

Class	n(truth)	n(classified)	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	111	106	99.18	95.0	91.0	93.0
2	237	237	98.35	94.0	94.0	94.0
3	113	115	98.46	87.0	88.0	88.0
4	133	130	99.29	96.0	94.0	95.0
5	814	788	93.74	94.0	91.0	93.0
6	126	137	95.44	66.0	71.0	68.0
7	159	169	98.35	88.0	94.0	91.0
8	127	138	98.85	88.0	96.0	92.0
Average	-	-	97.71	88.5	89.9	89.25

For the 100× magnification, DRNet_SVM attains an average classification accuracy rate of 98.32% compared to the 97.71% attained by DRNet_k-NN. Thus, this shows that SVM slightly performs better than k-NN at classifying each of the subtypes of benign and malignant based on the extracted features.

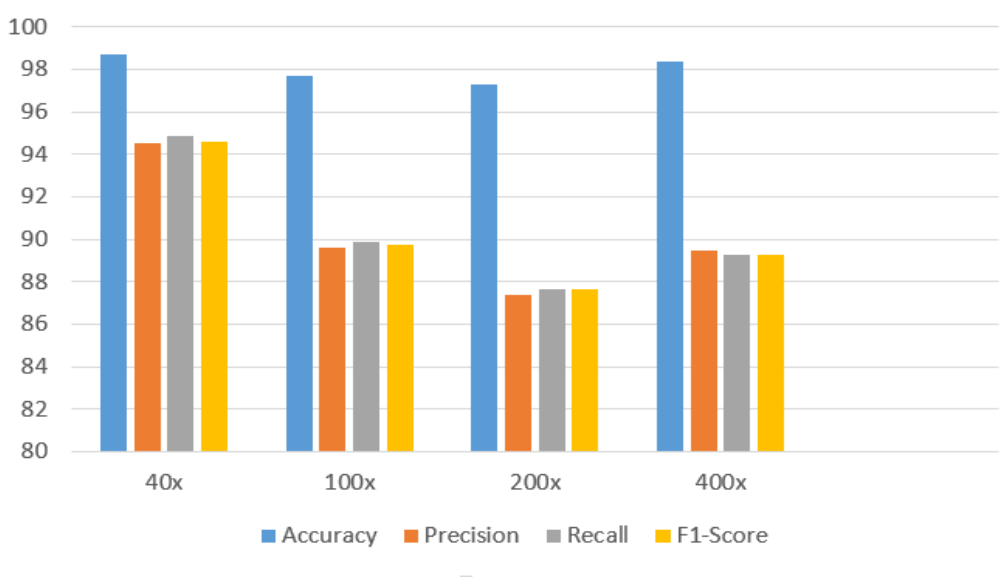

Figure 5.9. A graphical representation of the performance measure evaluation results of the k-NN classifier.

Figure 5.9 shows the graphical representation of the results based on the performance measure evaluations defined in equations (5.1) to (5.4), using k-NN classifier.

For the 200× magnification, DRNet_SVM produced a classification accuracy rate of 97.68% when compared with the 97.30% accuracy attained by DRNet_k-NN. The results show that DRNet_SVM is proving to be more effective at 100× and 200× magnification levels, respectively, in classifying subtypes of benign and malignant-based feature extractions. As for the 400× magnification, DRNet_k-NN produced an average classification accuracy rate of 98.40%, whereas DRNet_SVM had 97.71%. Thus, DRNet_k-NN is proving to be more effective at 40× and 400×

magnification levels, respectively, in classifying the subtypes of benign and malignant BC histopathological images in the dataset based on the extracted features. At the same time, all of the four (4) performance measures adopted in this study achieved higher classification rates based on the image magnification levels.

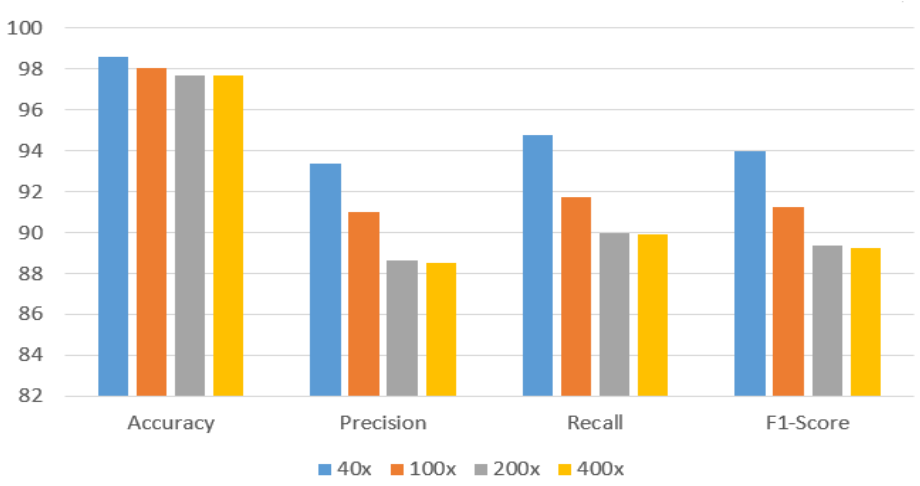

Figure 5.10. A graphical representation of the performance measure evaluation results of the SVM classifier.

Figure 5.11 shows the graphical representation of the results obtained based on the performance measure evaluations defined in equations (5.1) to (5.4), using SVM classifier.

Table 5.10. Comparison of the proposed DRNet model with other state-of-the-art studies.

References	Dataset	Model	Accuracy (%)
Spanhol et al. [2]	Breakhis	Benign and Malignant classification	83±0.125
Spanhol et al. [3]	Breakhis 40× 100× 200× 400×	Deep feature extraction using Decaf, a variant of AlexNet	 84.6±2.9 84.8±4.2 84.2±1.7 81.6±3.7
Deniz et al. [6]	Breakhis a. 40× b. 100× c. 200× d. 400×	Fine-tuned AlexNet	90.96±1.5 90.58±1.96 91.37±1.72 91.30±0.74
Bayramoglu et al. [5]	Breakhis a. 40× b. 100× c. 200× d. 400×	Deep learning model for magnification independent	83.08±2.08 83.17±3.51 84.63±2.72 82.10±4.42
Murtaza et al. [4]	Breakhis	A tree-based multiclassifcation	83.33
Proposed DRNet model	Breakhis 40× 100× 200× 400×	A combined ResNet50 and DenseNet201 with SVM classifier	98.61 98.04 97.68 97.71

Table 5.10 shows the comparison of the proposed DRNet model with its prior state-of-the-art counterpart studies explored in the literature. Based on the obtained results presented in this study, the DRNet model offers the best level of accuracy in classification in all performance measure evaluations under all magnification levels. Thus, it is proving to be very robust at extracting features of basic histopathology images from the breakhis image database compared to some of the studies explored in the literature.

6. CONCLUSIONS

This thesis work presented a novel deep feature extraction DRNet hybrid model to classify different categories of BC using the Breakhis dataset. It also demonstrated the potential of the transfer learning approach for feature extraction by utilizing the knowledge of an already pre-trained convolutional neural network with minimum resources and less computing power to extract the most relevant features needed to carry out the classification of BC subtypes. Data dependency is one of the challenging issues in DL because it requires a large quantity of training data to comprehend and uncover the hidden patterns in the data. Also, insufficient medical health datasets to train the deep learning model affected its accuracy. This thesis presents a novel deep feature extraction engineering for subtypes of breast cancer diagnosis-based transfer learning approach to overcome the stated limitations in deep learning because transfer learning utilizes the knowledge of already pre-trained convolutional neural networks trained on millions of image classification tasks, thereby saving additional computational power and resources required to train deep learning models from scratch. The proposed deep feature extraction engineering for subtypes of BC diagnosis-based transfer learning model approach consists of five phases: the feature extraction phase, concatenation, transformation, selection, and classification. In the first phase, nineteen (19) pre-trained convolutional neural networks were utilized as feature extractors to extract relevant features from the input images. SVM was used in this phase to compute the loss value function, and in the classification phase, it was used as a classifier. According to the feature extraction results, two networks attained the highest accuracies on the dataset, outperforming other networks. The two networks in consideration were chosen and concatenated to produce the DRNet model (a combination of ResNet50 and DenseNet201). In the transformation phase, the extracted features were transformed and decomposed into five sub-hand low-level features using a multilevel discrete wavelet transform (MDWT). Iterative neighborhood component analysis (INCA), a variant of principal component analysis (PCA), was applied as a feature selector in the classification phase to determine the needed minimum number of characteristics. In the last phase, SVM and KNN were used to classify the eight subtypes of breast cancer present in the breaKhis dataset. Five (5) performance indicators were used to evaluate the proposed DRNet model: accuracy, precision, recall, and F1-Score. The evaluation was based on magnification levels. Using the k-NN classifier, 1-all, and 10-fold cross-validation, the average classification rate accuracy of 98.74%, 97.7%, 97.3%, and 98 % at 40x, 100x, 200x, and 400x levels of magnification, respectively, were produced. Using the SVM classifier, 1-all, and 10-fold cross-validation, an average classification rate accuracy of 98.61%, 98.04%, 97.68%, and 97.71% was obtained for 40x, 100x, 200x, and 400x levels of magnification, respectively. According to the results obtained, the proposed model outperforms some of the studies pointed out in the literature. The approach is robust at handling

deep feature extraction from raw histological images of breast cancer before applying any image processing techniques. Thus, the proposed model enables medical experts to obtain fast and accurate results and to know the treatments needed for different breast cancer subtypes in their early stages, thereby saving additional women's lives and resources.

RECOMMENDATIONS

The developed deep feature extraction DRNet model for BC subtypes of benign and malignant can serve as a second opinion for medical specialists, mainly assisting inexperienced practitioners, lowering workload, and improving the objectivity of final diagnosis. The developed model is recommended for additional histopathological image classification problems involving skin cancer, prostate cancer and brain tumors similar to those analyzed in the breakhis image dataset used in this study.

REFERENCES

[1] Breast cancer, (n.d.). https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed November 19, 2021).

[2] Breast cancer, (n.d.). https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed November 20, 2021).

[3] Breast cancer - Diagnosis and treatment - Mayo Clinic, (n.d.). https://www.mayoclinic.org/diseases-conditions/breast-cancer/diagnosis-treatment/drc-20352475 (accessed March 20, 2022).

[4] Biopsy | Cancer.Net, (n.d.). https://www.cancer.net/navigating-cancer-care/diagnosing-cancer/tests-and-procedures/biopsy (accessed March 20, 2022).

[5] J.A. Cruz, D.S. Wishart, Applications of Machine Learning in Cancer Prediction and Prognosis, Cancer Inform. 2 (2006) 59. https://doi.org/10.1177/117693510600200030.

[6] K. Kourou, T.P. Exarchos, K.P. Exarchos, M. V. Karamouzis, D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Comput. Struct. Biotechnol. J. 13 (2015) 8–17. https://doi.org/10.1016/J.CSBJ.2014.11.005.

[7] B. Krawczyk, G. Schaefer, Breast thermogram analysis using classifier ensembles and image symmetry features, IEEE Syst. J. 8 (2014) 921–928. https://doi.org/10.1109/JSYST.2013.2283135.

[8] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A Dataset for Breast Cancer Histopathological Image Classification, IEEE Trans. Biomed. Eng. 63 (2016) 1455–1462. https://doi.org/10.1109/TBME.2015.2496264.

[9] F.A. Spanhol, P.R. Cavalin, L.S. Oliveira, C. Petitjean, L. Heutte, Deep features for breast cancer histopathological image classification, 2017 IEEE Int. Conf. Syst. Man, Cybern. SMC 2017. 2017-January (2017) 1868–1873. https://doi.org/10.1109/SMC.2017.8122889.

[10] T. Kausar, M.A. Ashraf, A. Kausar, I. Riaz, Convolution Neural Network based Approach for Breast Cancer Type Classification, Proc. 18th Int. Bhurban Conf. Appl. Sci. Technol. IBCAST 2021. (2021) 407–413. https://doi.org/10.1109/IBCAST51254.2021.9393249.

[11] N. Brancati, G. De Pietro, M. Frucci, D. Riccio, A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images, IEEE Access. 7 (2019) 44709–44720. https://doi.org/10.1109/ACCESS.2019.2908724.

[12] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, Ü. Budak, Transfer learning based histopathologic image classification for breast cancer detection, Heal. Inf Sci Syst. 6 (2018) 18. https://doi.org/10.1007/s13755-018-0057-x.

[13] K. Das, S.P.K. Karri, A. Guha Roy, J. Chatterjee, D. Sheet, Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification, Proc. - Int. Symp. Biomed. Imaging. (2017) 1024–1027. https://doi.org/10.1109/ISBI.2017.7950690.

[14] A. Chan, J.A. Tuszynski, Automatic prediction of tumour malignancy in breast cancer with fractal dimension, R. Soc. Open Sci. 3 (2016). https://doi.org/10.1098/RSOS.160558.

[15] G. Murtaza, A.W. Abdul Wahab, G. Raza, L. Shuib, A tree-based multiclassification of breast tumor histopathology images through deep learning, Comput. Med. Imaging Graph. 89 (2021) 101870. https://doi.org/10.1016/j.compmedimag.2021.101870.

[16] N. Bayramoglu, J. Kannala, J. Heikkila, Deep learning for magnification independent breast cancer histopathology image classification, Proc. - Int. Conf. Pattern Recognit. 0 (2016) 2440–2445. https://doi.org/10.1109/ICPR.2016.7900002.

[18] A. Saber, M. Sakr, O. Abo-Seida, A.K.-I. Access, undefined 2021, A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique, Ieeexplore.Ieee.Org. (n.d.). https://ieeexplore.ieee.org/abstract/document/9427477/ (accessed March 21, 2022).

[19] F. Demir, DeepBreastNet: A novel and robust approach for automated breast cancer detection from histopathological images, Biocybern. Biomed. Eng. 41 (2021) 1123–1139. https://doi.org/10.1016/j.bbe.2021.07.004.

[20] Y. Zou, J. Zhang, S. Huang, B. Liu, Breast cancer histopathological image classification using attention high-order deep network, Int. J. Imaging Syst. Technol. (2021) 1–14. https://doi.org/10.1002/ima.22628.

[21] H. Kataoka, K. Iwata, Y. Satoh, Feature Evaluation of Deep Convolutional Neural Networks for Object Recognition and Detection, (2015). https://doi.org/10.48550/arxiv.1509.07627.

[22] A.T. Sahlol, D. Yousri, A.A. Ewees, M.A.A. Al-qaness, R. Damasevicius, M.A. Elaziz, COVID-19 image classification using deep features and fractional-order marine predators algorithm, Sci. Reports 2020 101. 10 (2020) 1–15. https://doi.org/10.1038/s41598-020-71294-2.

[23] L. Yang, Z. Xu, Feature extraction by PCA and diagnosis of breast tumors using SVM with DE-based parameter tuning, Int. J. Mach. Learn. Cybern. 10 (2019) 591–601. https://doi.org/10.1007/S13042-017-0741-1.

[24] D. Zhang, L. Zou, X. Zhou, F.H.-I. Access, undefined 2018, Integrating feature selection and feature extraction methods with deep learning to predict clinical outcome of breast cancer Thesis papers, Ieeexplore.Ieee.Org. (n.d.). https://ieeexplore.ieee.org/abstract/document/8361862/ (accessed March 21, 2022).

[25] Y. Jiang, L. Chen, H. Zhang, X. Xiao, Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module, PLoS One. 14 (2019) e0214587. https://doi.org/10.1371/JOURNAL.PONE.0214587.

[26] E.D. Carvalho, A.O.C. Filho, R.R.V. Silva, F.H.D. Araújo, J.O.B. Diniz, A.C. Silva, A.C. Paiva, M. Gattass, Breast cancer diagnosis from histopathological images using textural features and CBIR, Artif. Intell. Med. 105 (2020) 101845. https://doi.org/10.1016/j.artmed.2020.101845.

[27] E.T. Zewdie, A.W. Tessema, G.L. Simegn, Classification of breast cancer types, sub-types and grade from histopathological images using deep learning technique, Health Technol. (Berl). 11 (2021) 1277–1290. https://doi.org/10.1007/S12553-021-00592-0/TABLES/9.

[28] View of Breast Cancer Image Multi-Classification Using Random Patch Aggregation and Depth-Wise Convolution based Deep-Net Model, (n.d.). https://online-journals.org/index.php/i-joe/article/view/18513/8559 (accessed March 23, 2022).

[29] J. Ahmad, H. Farman, Z. Jan, Deep Learning Methods and Applications, SpringerBriefs Comput. Sci. (2019) 31–42. https://doi.org/10.1007/978-981-13-3459-7_3.

[30] What is a Convolutional Neural Network? - MATLAB & Simulink, (n.d.). https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html (accessed April 18, 2022).

[31] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE. 86 (1998) 2278–2324. https://doi.org/10.1109/5.726791.

[32] T.F. Gonzalez, Handbook of approximation algorithms and metaheuristics, Handb. Approx. Algorithms Metaheuristics. (2007) 1–1432. https://doi.org/10.1201/9781420010749.

[33] AlexNet architecture. Includes 5 convolutional layers and 3... | Download Scientific Diagram, (n.d.). https://www.researchgate.net/figure/AlexNet-architecture-Includes-5-convolutional-layers-and-3-fullyconnected-layers_fig3_322592079 (accessed April 19, 2022).

[34] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. (2014). https://doi.org/10.48550/arxiv.1409.1556.

[35] VGG Very Deep Convolutional Networks (VGGNet) - What you need to know - viso.ai, (n.d.). https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/ (accessed April 19, 2022).

[36] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-December (2016) 770–778. https://doi.org/10.1109/CVPR.2016.90.

[37] 1637052992200.jpg (1000×241), (n.d.). https://www.mathworks.com/discovery/rnn/_jcr_ content /main Parsys/ image.adapt.full.medium.jpg/1637052992200.jpg (accessed April 21, 2022).

[38] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, Neural Comput. 9 (1997) 1735–1780. https://doi.org/10.1162/NECO.1997.9.8.1735.

[39] Screen-Shot-2017-06-20-at-09.58.29.png (422×384), (n.d.). https://web.inf.ufpr.br/vri/wp-content/uploads/sites/7/2017/06/Screen-Shot-2017-06-20-at-09.58.29.png (accessed March 25, 2022).

[40] Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, Y. Zhao, Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification, Pattern Recognit. 71 (2017) 14–25. https://doi.org/10.1016/J.PATCOG.2017.05.010.

[41] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L. Fei-Fei, ImageNet Large Scale Visual Recognition Challenge, Int. J. Comput. Vis. 115 (2015) 211–252. https://doi.org/10.1007/S11263-015-0816-Y/FIGURES/16.

[42] Darknet: Open Source Neural Networks in C, (n.d.). https://pjreddie.com/darknet/ (accessed April 4, 2022).

[43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.C. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2018) 4510–4520. https://doi.org/10.1109/CVPR.2018.00474.

[44] B. Zoph, V. Vasudevan, J. Shlens, Q. V. Le, Learning Transferable Architectures for Scalable Image Recognition, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2017) 8697–8710. https://doi.org/10.48550/arxiv.1707.07012.

[45] F. Chollet, Xception: Deep Learning with Depthwise Separable Convolutions, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017. 2017-January (2016) 1800–1807. https://doi.org/10.48550/arxiv.1610.02357.

[46] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, (n.d.). www.aaai.org (accessed April 4, 2022).

[47] M. Tan, Q. V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 36th Int. Conf. Mach. Learn. ICML 2019. 2019-June (2019) 10691–10700. https://doi.org/10.48550/arxiv.1905.11946.

[48] X. Zhang, X. Zhou, M. Lin, J. Sun, ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2017) 6848–6856. https://doi.org/10.48550/arxiv.1707.01083.

[49] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the Inception Architecture for Computer Vision, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem (2016) 2818–2826. https://doi.org/10.1109/CVPR.2016.308.

[50] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely Connected Convolutional Networks, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017. 2017-January (2016) 2261–2269. https://doi.org/10.48550/arxiv.1608.06993.

[51] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 07-12-June-2015 (2015) 1–9. https://doi.org/10.1109/CVPR.2015.7298594.

[52] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size, (2016). https://doi.org/10.48550/arxiv.1602.07360.

[53] S. Thakral, P. Manhas, Image Processing by Using Different Types of Discrete Wavelet Transform, Commun. Comput. Inf. Sci. 955 (2019) 499–507. https://doi.org/10.1007/978-981-13-3140-4_45.

[54] N.I. Chervyakov, P.A. Lyakhov, N.N. Nagornov, Quantization Noise of Multilevel Discrete Wavelet Transform Filters in Image Processing, Optoelectron. Instrum. Data Process. 2018 546. 54 (2019) 608–616. https://doi.org/10.3103/S8756699018060092.

[55] I. Journal, O.F. Engineering, D. Classification, T. In, B. Cancer, G. Data, Review : a Combine Approach of Feature Selection and, 6 (2017) 170–177.

[56] T. Tuncer, S. Dogan, F. Ozyurt, S.B. Belhaouari, H. Bensmail, Novel multi center and threshold ternary pattern based method for disease detection method using voice, IEEE Access. 8 (2020) 84532–84540. https://doi.org/10.1109/ACCESS.2020.2992641.

[57] C. Sotiriou, P. Wirapati, S. Loi, A. Harris, S. Fox, J. Smeds, H. Nordgren, P. Farmer, V. Praz, B. Haibe-Kains, C. Desmedt, D. Larsimont, F. Cardoso, H. Peterse, D. Nuyten, M. Buyse, M.J. Van de Vijver, J. Bergh, M. Piccart, M. Delorenzi, Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis, J. Natl. Cancer Inst. 98 (2006) 262–272. https://doi.org/10.1093/JNCI/DJJ052.

[58] k-nearest-neighbor-algorithm-for-machine-learning2.png (600×300), (n.d.). https://static.javatpoint.com/tutorial/machine-learning/images/k-nearest-neighbor-algorithm-for-machine-learning2.png (accessed May 5, 2022).

[59] 1*yLQxzD1GR9Wt2-kcAMVARg.png (620×256), (n.d.). https://miro.medium.com/max/620/1*yLQxzD1GR9Wt2-kcAMVARg.png (accessed May 5, 2022).

[60] F.F. Chamasemani, Y.P. Singh, Multi-class Support Vector Machine (SVM) classifiers - An application in hypothyroid detection and classification, Proc. - 2011 6th Int. Conf. Bio-Inspired Comput. Theor. Appl. BIC-TA 2011. (2011) 351–356. https://doi.org/10.1109/BIC-TA.2011.51.