

Deep Learning and Mathematical Models for Cancer Detection in Image Classification

Mr. Manish Kumar¹, Dr. Kailash Yadav²

¹Research Scholar, Mathematics, School of Basic & Applied Sciences, Nirwan University, Jaipur, Rajasthan

²Associate Professor, School of Basic & Applied Sciences, Faculty of Basic & Applied Sciences Nirwan University Jaipur, Rajasthan (India)

ABSTRACT

This research presents research in deep learning techniques that could go on radically to transform the landscape of cancer diagnosis, more so with the use of Convolutional Neural Networks. More specifically, this research aims to be used for the early detection and classification of breast and lung cancer using image classification based on CNNs. It is in this regard that a very critical innovation has been the development of a CNN model which will be independent of magnification levels for multi-classification in breast cancer histology images—a really critical problem in pathology with inconsistent diagnoses across different magnification scales. With it, the model improves classification precision by deciding between the several subtypes of breast cancer; this is important in guiding personalized treatment plans. It also discusses the application of CNNs in the classification of three-dimensional images of lung cancer, providing detailed tumor analysis that may inform personalized treatment strategies. A methodology including the formulation of mathematical models for tracing the progress of cancer and its response to chemotherapy, as well as the design and training of a new CNN model, is used. It optimized model performance for the image recognition rate and cross-entropy loss, benchmarked against the very best in pre-trained models such as VGG16, VGG19, and ResNet50.

This effort has not only improved the accuracy of cancer diagnoses and workflow efficiency for diagnosis but also delivered insights into the underlying mathematical frameworks of the CNN models, laying the base for improvements in deep learning and medical image classification in the very near future.

1. INTRODUCTION

1.1. INTRODUCTION TO CANCER

In India, there are over 1,200,000 new cases of cancer recorded year. According to the latest data from the National Cancer Registry, 1 in 8 Indian men and 1 in 9 Indian women are predicted to develop cancer at some time in their life.

A multitude of factors impact the incidence and types of cancer that are discovered, including age, gender, ethnicity, food, and genetic predisposition. The most frequent cancer in Indian men is lung cancer, which is followed by cancers of the mouth and throat. The two most common malignancies diagnosed in females are breast and cervical cancers.

1.2. What is Cancer?

The abnormal functioning of cells in the body is a hallmark of the illness known as cancer. A healthy person's cells divide, expand, and eventually die in a way that is required to maintain regular bodily

functions. However, these cells behave differently from normal when it comes to cancer, developing in an uncontrollably abnormal manner.

Table 1.1: Difference between Cancer Cells and Normal Cells

Characteristic	Normal Cells	Cancer Cells
Progress	Follow an established timetable	Extend excessively while ignoring cues
Interaction	Accurately interact with neighboring cells	Be alone and indifferent to cues
Cell Demise and Repair	undergo self-healing, or apoptosis.	Don't engage in self-healing or apoptosis.
Attached	Release adhesive substances	possess no adhesive properties and are able to split
Look	uniform size and tidy appearance	vastly different in scope and significance
Reaching Maturity	Go through the maturation process.	Continue to be uniform due to their rapid expansion

1.3. Tissue Changes that are 'Not' Cancer

Changes in body tissues may not always indicate the existence of cancer. However, some changes must be continuously monitored since, left unchecked, they have the potential to develop into cancer. Let's now examine a few examples of non-malignant tissue modifications that may nonetheless require careful consideration because they may worsen:

- **Hyperplasia:** A tissue's cells multiplying faster than usual leads to an accumulation of extra cells. However, under a microscope, the tissue's and its cells' structures will appear normal.
- **Dysplasia:** Instead of hyperplasia, dysplasia is a more evolved condition. Moreover, there is a gathering of additional phones, however the tissue's association is changed, and these phones have an unusual appearance. Generally speaking, the presence of abnormal cells and tissues raises the risk of getting cancer.
- **Carcinoma in Situ:** This variation of the illness is a little more severe. These are frequently referred to as "stage 0 cancers," but it's important to realize that they are not the same as cancer because the abnormal cells do not pierce the surrounding tissues as much as cancer cells do.

1.4. Types of Cancer

Cancers get their names from the organs or tissues that give rise to them. The specific cell type—such as squamous or epithelial cells—that gives rise to cancer is also taken into account during the classification process.

a) The Most Common Type of Carcinoma

Carcinoma is the most predominant kind of cancer. Epithelial cells, which cover the outside and internal surfaces of the body, make up most of it. A variety known as basal cell carcinoma is noteworthy.

b) Sarcoma: Impacting Soft Tissues and Bones

Sarcomas are cancers that foster in the delicate tissues, which incorporate veins, lymphatics, muscle, fat, and sinewy tissues like tendons and ligaments.

c) Leukemia: Cancer Associated with Blood

Leukemia is one sort of cancer that starts in the tissues of the bone marrow that produce blood. Dissimilar to different malignancies, leukemia doesn't bring about strong tumors. All things being equal, it brings about an ascent of unusual white platelets (leukemia cells and leukemic shoot cells) in the bone marrow and blood, which supplant typical platelets.

d) Lymphoma: Lymphocytes Are Involved

Lymphoma is described by the threat of lymphocytes, a subset of white platelets essential to the safe framework. This sort of cancer gathers unusual cells in the organs, lymph hubs, and lymphatic veins of the body.

e) Plasma Cells Affected by Multiple Myeloma

Multiple myeloma is brought about by plasma cells, one more kind of resistant cell. Strange plasma cells called myeloma cells gather in the bone marrow and afterward spread all through the skeletal framework to cause malignancies.

f) Melanoma: Beyond the Skin

One sort of cancer that starts in the specific melanocytes — the phones that produce melanin — is called melanoma. Melanomas can foster in other pigmented tissues too, like the eye, notwithstanding typically beginning in the skin.

g) Different Types of Brain and Spinal Cord Tumors

Tumors of the mind and spinal string can have a wide range of structures. They are partitioned into bunches in light of the kind of cell they come from and the region of the focal sensory system wherein they first manifest.

1.5. Other Cancer Types

- **Germ cell tumors:** These tumors can be benign or malignant, and they can appear in many different places on the body. They begin in the cells that make ovum or sperm.
- **Neuroendocrine Tumours:** Cells that discharge chemicals into the circulation system in response to signals got by the sensory system are the wellspring of neuroendocrine tumors. Their enhanced hormone synthesis can result in a variety of symptoms.
- **Carcinoid Tumours:** Carcinoid tumors are a kind of neuroendocrine tumors that typically grow slowly and are often found in the gastrointestinal tract. Every now and then, they could spread to the liver or other organs.

1.6. Risk Factors of Cancer

There are many variables that could influence an individual's gamble of creating cancer, and it is essential to comprehend these gamble factors. A couple of normal gamble factors that can expand your gamble of creating cancer are as per the following:

- **Tobacco Use:** Smoking and consuming tobacco products increases the risk of acquiring certain malignancies.
- **High Alcohol Consumption:** Overindulgence in liquor utilization has been related with an expanded gamble of numerous malignancies, including cancers of the mouth, throat, and liver.
- **Lack of Physical Activity:** A stationary way of life might be connected to an expanded gamble of cancer, as studies have shown that customary active work brings down the gamble of various malignancies.
- **Air Pollution:** Long haul openness to dirtied air, particularly in metropolitan regions, may build the possibility creating cancer, especially lung cancer.

- **Radiation Exposure:** Cancer risk can be increased by radiation exposure from medical or occupational sources, including X-rays and some treatments.

1.7. Unprotected UV Light Exposure: Excessive and unprotected UV radiation exposure, including sunshine and tanning beds, has been linked to skin cancer.

1.8. Cancer Stages

Systems for cancer staging are crucial for assisting healthcare providers in treating patients and giving a prognosis, or anticipated course of action. The staging scheme that is most frequently utilized is the TNM classification.

There are four primary phases for most forms of cancer. A patient's specific stage is defined by several factors, including the tumor's size and location inside the body:

- **Stage I:** The cancer appears to be limited to a small, localized location and has not migrated to nearby lymph nodes or other tissues.
- **Stage II:** The cancer is still isolated and has not moved to other parts of the body, despite considerable growth.
- **Stage III:** Since the cancer has spread, close by tissues or lymph hubs may now be impacted.
- **Stage IV:** This advanced stage, often called metastatic cancer, indicates that the cancer has spread to distant organs or body components.

In addition to the four main phases, Stage 0 designates the first phase of malignant development. Since the cancer has returned to its original site, it is usually treatable when it is in Stage 0. Many doctors diagnose Stage 0 cancers as precancerous, which means that they could spread if treatment is not received.

1.9. Causes of Cancer

Ninety–95% of cancers are caused by genetic alterations resulting from environmental and lifestyle factors. Five to ten percent is inherited. Environmental reasons extend beyond pollution and encompass lifestyle, economic, and behavioral factors. Cancer passings are brought about by various elements, including radiation (both ionizing and non-ionizing, up to 10%), absence of actual work, contamination, diseases (15-20%), diet and stoutness (30-35%), and tobacco use (25-30%). Despite the fact that it isn't appeared to cause cancer, mental pressure might aggravate cancer.

Mutational marks from genomic sequencing of tumor DNA can recognize way of life or ecological variables that added to the advancement of cancer. This can demonstrate whether lung, skin, or secondary cancers were caused by UV radiation, tobacco smoke, or chemotherapy. Cancer is often not communicable. Organ donations and rare pregnancy transmissions are the exceptions. Cancer can be brought on by transmissible viral diseases such HIV, Epstein-Barr, HPV, and hepatitis B.

2. METHODOLOGY

2.1. MODEL FORMULATION

We develop a mathematical model accounting for five distinct epidemiological compartments: These compartments arise from the natural course of breast cancer and represent the various stages and outcomes of patients admitted to a hospital. It is a variable, X , that represents the entry of patients with cancer into the hospital, in this case, at various stages of the classification of patients' medical records, in relation to breast cancer: from stage 1 to 4. It is taken that chemotherapy is the main treatment for any cancer patient. Over time, the population in each compartment changes: patients either recover, or more severely ill patients develop into the compartments with a severe state—the so-called cardiotoxicity from chemotherapy, as described later—while the worst cases involve passage to death.

Those patients admitted to a hospital are initially classified as being in stage 1 of stage 2X and another sub-class as stage 2B cancer, defining the population $X(t)$. Patients who are in stage 1 and 2 cancers are put into class A with a population created at a rate of Λ_1 . Patients who are undergoing chemotherapy recover at a rate of ϕ_1 and get into the class R, the recovered population. From those in stage A, there are transitions to class B, representing a more advanced disease stage, at a rate of β_2 . This leads to the following set of ordinary differential equations of the model:

$$\frac{dX}{dt} = \Lambda_1 - (\phi_1 + \phi_2)X \quad (3.1)$$

The two identified stages, 3X and 3B, characterize the patients and are members of subpopulation B. Now, the class is formed by a rate Λ_2 , since it describes the transfer of patients from class X to class B, because these patients are seriously worsened. Those who recover from chemotherapy are transferred to class R at rate ϕ_2 , and those whose state worsened are transferred to class C at rate β_1 . Class B patients undergo intensive chemotherapy with rate r_1 , after which patients are placed into class E. Let μ_1 be the mortality of class B and write down the evolutionary equation describing these transitions as equation (3.2):

$$\frac{dB}{dt} = \Lambda_2 + \beta_2 X + \psi_1 R - (\phi_2 + \beta_1 + \tau_1 + \mu_1)B \quad (3.2)$$

Class C consists of patients diagnosed to be in stage 4 of the cancer following the tests. New patients in stage 4 join the population at a rate of Λ_3 . Patients from class B whose conditions worsen progress into class C at a rate of β_1 . Patients in class C recover from chemotherapy administered on them at a rate ϕ_3 , but the strong side effects cause some to develop cardiotoxicity at a rate of r_2 . The mortality rate in class C is μ_2 , and patients who are relapsing join class C at a rate ψ_2 . This process is modelled by the differential equation 3.3:

$$\frac{dC}{dt} = \Lambda_3 + \beta_1 B + \psi_2 R - (\phi_3 + \tau_2 + \mu_2)C \quad (3.3)$$

Class R contains individuals who have been treated and are now free of disease. Recovered individuals in classes X, B and C enter class R with rates ϕ_1 , ϕ_2 and ϕ_3 respectively. Though, at rates ψ_1 and ψ_2 the patients may relapse, and thus re-enter either the stage 3 or 4 of the disease. Also, the subpopulation R may acquire cardiotoxicity after some time with rate r_3 . This process is modelled by the following differential equation (3.4):

$$\frac{dR}{dt} = \phi_1 X + \phi_2 B + \phi_3 C - (\psi_1 + \psi_2 + \tau_3)R \quad (3.4)$$

Class E includes patients in whom cardiotoxicity has developed as a result of chemotherapy. This class is entered from classes B, C and R, the rates of transition being r_1 , r_2 and r_3 , respectively. As before the mortality rate for patients of class E is denoted by μ_3 and the evolution of this population is given by equation (3.5):

$$\frac{dE}{dt} = \tau_3 R + \tau_2 C + \tau_1 B - \mu_3 E \quad (3.5)$$

Together, Equations yield a system of ordinary differential equations that models transition between the compliment of compartments involved in breast cancer progression, chemotherapy, relapse, recovery, and mortality. The system falls under the non-negatively initial conditions that secure the sizes of the populations in compartments are of realism and are bounded. The initial conditions are denoted as:

$$X(0) \geq X_0, B(0) \geq B_0, C(0) \geq C_0, R(0) \geq R_0, E(0) \geq E_0$$

This model provides the full framework of understanding for how the complex process of breast cancer progression in patients undergoing chemotherapy goes through different diseases at different stages, recovery, relapse, and cardiotoxicity.

2.2. DATA COLLECTION

BreakHis Dataset

In this study, the primary dataset is the BreakHis dataset, a well-known publicly available dataset consisting of images of histopathology of breast cancers. And diversity in this set is important to this study, with a wide range of images pertinent to breast cancers, both benign and malignant, at different magnifications: 40X, 100X, 200X, and 400X. These varied magnifications mirror the diversity normally found during clinical practice—factors that may contribute to the generalization of robust classification models. The BreakHis dataset will be the base set used to train, validate, and test the models proposed in the research.

Image Classification and Stratification

Several important preprocessing steps were taken to come up with a well-balanced and representative training set for the complexities of breast cancer classification.

- **Stratified Sampling:** The dataset was stratified to ensure there was no variation in the classes of benign and malignant across train, validation, and test sets. It was very important to avoid bias on a frequently occurring class, thus improving generalization across both categories of benign and malignant image classes.
- **Down Sampling:** Since the number of images varies from one level of magnification to another, down sampling was used in this case. This meant reducing the number of images in classes that were over-represented by randomly throwing out a number of them to correspond to the class with very few samples. The reason for doing this was to balance the contribution that each class would offer towards training so that the model would not fit the frequent classes.
- **Combining Magnifications:** For classification of magnification-independent features, we combined the images from both magnification levels into one dataset. This was motivated by the fact that training the model to learn magnification invariance features would help us calibrate it such that in further classification it would be performed in a more generalized way. It must be noted, though, that the model must also be separately assessed for classification of the magnification level of the images. We do this to ensure that the model can separately address magnification-dependent from magnification-independent classification tasks.

The collectively enhanced techniques of data collection and preprocessing gave much-needed strength to the present classification model, therefore being well-suited for the varied complex nature of breast cancer histopathology images.

Figure 2.2 shows the detailed distribution of images in the BreakHis dataset, associating a number of images with the particular magnification factor.

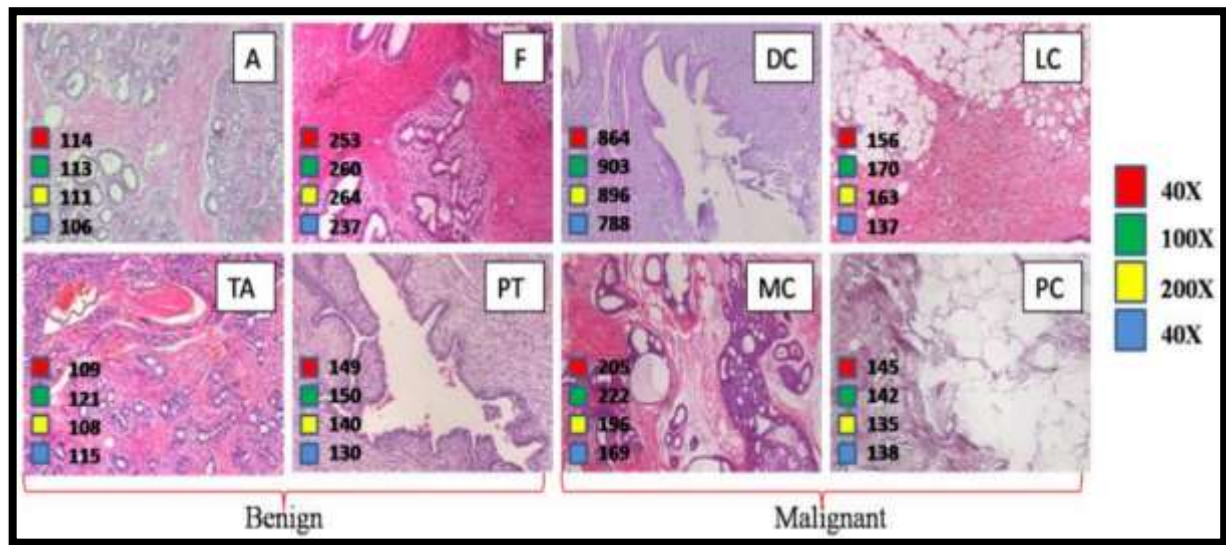


Figure 2.2: An illustration of data organization in the BreakHis Dataset

2.3. DATA SPLITTING

It is during the data-splitting phase when the dataset is divided systematically for optimal training and evaluation of the model. First, 90% of the dataset is trained, leaving the remaining 10% for testing. This division allows most—90%—of the data so that the model can learn from it and adapt to the underlying patterns and features relevant to breast cancer histology image classification.

More subdivision is done in the training set to get different validation subsets. This could be divided further into two parts: one for training the model and the other for validation. Normally, the train set is divided by approximately 80% to really train the model and 20% for validation. This validation set is essential for measuring model performance on data it has not seen before training, for tuning hyper-parameters, and for optimization strategies.

It allows a model to be based on a large variety of the training set yet still provides for a small different subset of validation checks now and then. It also prevents overfitting and guarantees not only robust but also generalizable model performance. For that reason, there is a test set of 10% of the entire dataset held out from training and validation. It is valuable as an independent measure of the final performance of a model and, therefore, an objective evaluation of its learned generalization—that is, being able to generalize to classify previously unseen data.

In general, the complicated data-splitting strategy attempts to balance the demand for overwhelming training data against the need for strict validation and testing, which enhances the credibility and accuracy of the classification model of breast cancer.

2.4. DATA AUGMENTATION

One of the significant difficulties in the improvement of strong computer aided design frameworks is an uneven and little dataset size. Information expansion alludes to a methodology utilized in expanding a dataset to mitigate the issue of little dataset size. A few well-known information expansion methods, for example, flipping, trimming, scaling, pivot, interjection, interpretation and commotion inclusion have proactively been applied in numerous past examinations. Not all expansion approaches utilized for regular images work for medical images, since the greater part of the medical imaging-related assignments are hierarchical tackled issues, as opposed to normal images, which are base up tackled issues. Besides, some expansion methods might lose the first data expected in the extraction of additional critical elements. It

might likewise enhance the commotion present in the data when it isn't picked as expected. Thus, an information expansion approach must be shrewdly picked concerning the dataset. Histopathological images are rotational invariant and reflection symmetric. Turn and reflection can hence be utilized as information increase methods for histopathological images by virtue of the reality. Notwithstanding, the possibilities eliminating a portion of the inborn properties from the image are expanded in the event that other expansion strategies than turn promotion reflection are picked for histopathological images. In the current work, procedures like pivot, flipping, and zooming are utilized for information expansion. In part 4 and section 5, the three points of 90o, 180o, and 270o have been utilized to complete turns of the images. The irregular turn has additionally been utilized in part 6. Different benefits of utilizing the increase cycle incorporate limiting the overfitting issue that is a significant issue towards the effective execution of an AI model.

The use of a predetermined training process is at the heart of a newly suggested CNN structure. All of the classes use the same number of images from a stratified training set. Because the distribution of images in BreakHis dataset is not constant across all magnification levels, balancing was performed prior to training. Each class is randomly down sampled until it is visually similar to the class with the fewest images; the class with the fewest images is then preserved. At each level of magnification, the classes undergo down-sampling. In order to create a binary and multi-classification system that is independent of magnification, all of the dataset's photographs are grouped into a shared folder without taking their magnification into account. However, separate tests have been conducted to ascertain the model's capacity to classify data depending on magnification. Figure 4.2 is a schematic depicting the suggested process. There is a training set that contains 90% of the data and a testing set that contains 10%. For every experiment, a training set is further split into a validation set of 25% and a training set of 75%. To make the dataset bigger, data augmentation has been used on a balanced training and validation set, where each image has been rotated by 90°, 180°, and 270°.

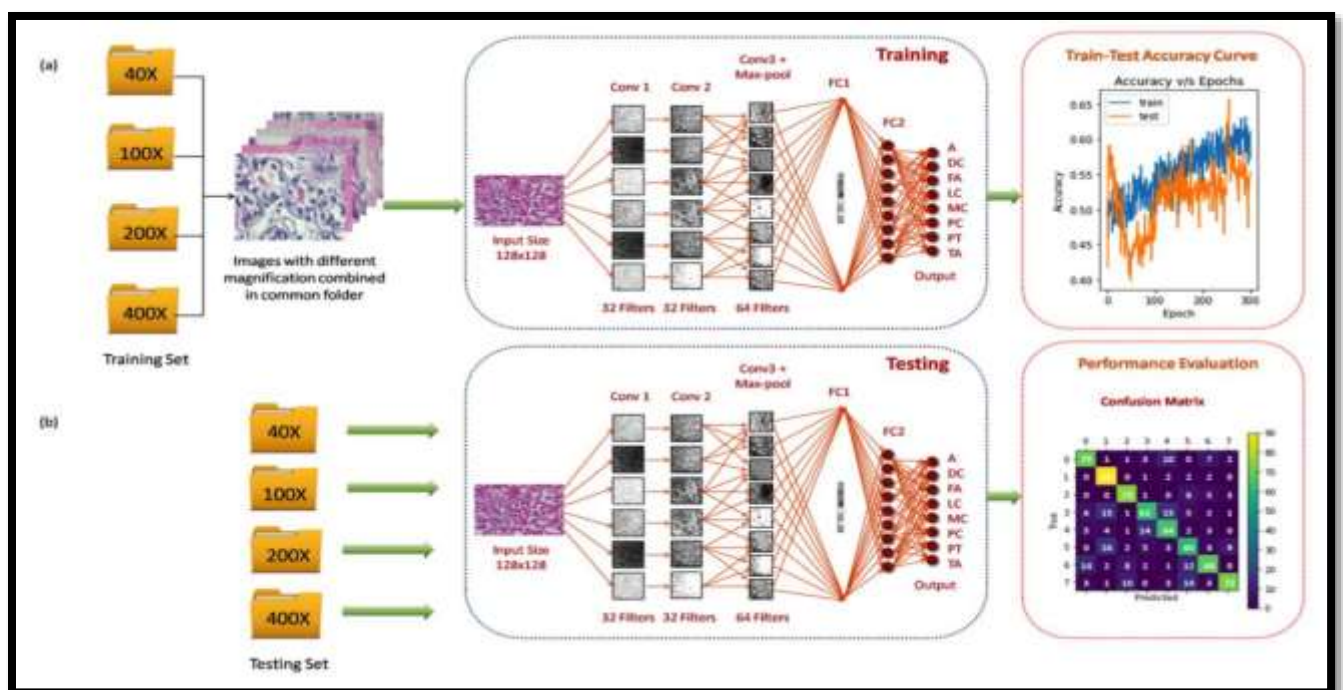


Figure 2.4. The magnification-independent multi-classification system's (a) training and (b) testing protocols with associated performance evaluation indicators.

Layer Type	Activation	Regularization	No. of Filters	Filter Size	Pooling Size/Stride	Zero Padding
Conv1	ReLU	-	32	3x3	-	-
Conv2	ReLU	-	32	3x3	-	2x2
Conv3	ReLU	-	64	3x3	-	2x2
Max-Pool	-	-	-	-	2x2/2	-
FC1 (Fully-Connected)	ReLU	Dropout	128 nodes	1x1	-	-
FC2 (Fully-Connected)	ReLU	Dropout	128 nodes	1x1	-	-

Table 2.4. Description of CNN Architecture

3. Training Protocol

To maximize model performance and guarantee reliable breast cancer histology picture categorization, the Convolutional Neural Network (CNN) training methodology is meticulously crafted. By reducing the discrepancies between the predicted and observed results, training is essential for a multi-classification system's model to perform well. By measuring this process with the cross-entropy loss function, we may find out how to fine-tune the network's parameters. In order to train the model to minimize classification mistakes, the cross-entropy loss function measures the disparity between the actual class labels and the anticipated probability distribution.

The cross-entropy loss (CEL) can be mathematically represented as:

$$\text{Cross Entropy Loss (CEL)} = - \sum_{c=1}^N y_c \log(p_c)$$

For each class c , y_c is the right classification and p_c indicates the anticipated probability observation; here, N is the number of classes. When both the expected and predicted outputs are 0, the positive and continuous loss function stops being a function. This loss function is minimized by the use of the back propagation method, which enables the model to learn from its errors.

The Adam optimizer is used to further enhance the training process. Adam optimizes the network's weights using adaptive learning rates for each parameter; it combines the benefits of AdaGrad and RMSProp optimizers. Complex models and big datasets are perfect for Adam because of its momentum-based updates and its ability to manage sparse gradients. It guarantees that the CNN learns well and efficiently converges to an optimal solution. The CEL is calculated for every picture in the dataset and then averaged throughout training. The optimization process and the CNN's effective parameter adjustment for improved classification are both controlled by this averaged CEL.

A readiness and approval subset of the preparation dataset is utilized to forestall overfitting and ensure vigorous execution. To prepare the model, we utilize the readiness subset; to test how well it does on obscure information, we utilize the approval subset. To ensure the model's presentation in genuine applications, this technique is helpful for assessing its generalizability to new information.

Information expansion methods including turning, flipping, and scaling are applied to the preparation images to additional increment the model's versatility and generalizability. By introducing variability, these changes make the model less dependent on individual picture features and more capable of handling a broader variety of real-world circumstances.

The trained CNN is evaluated using a number of measures. The degree to which a model's predictions are accurate is a measure of its accuracy. By contrasting the level of right expectations with the complete number of positive forecasts, accuracy estimates the model's capacity to forestall bogus up-sides. The level of genuine positive cases identified by the model is called review, which is additionally called responsiveness. For circumstances where the class conveyance is lopsided, the F1 score gives an intensive evaluation of the model's presentation since a decent measure consolidates review and accuracy.

A strong foundation for training the CNN is provided by the following components of the training protocol: cross-entropy loss, Adam optimization, data splitting, data augmentation, and performance measurements. With this structure in place, we know we can trust the model to correctly categorize breast cancer histology photos after we've fine-tuned and validated it.

After a long series of trial-and-error trials, the network parameters utilized for training the network are settled upon. The investigation concluded that a learning rate of 0.0001 was a good choice as using a lower rate slowed down the convergence process. In contrast, convergence fails when the learning rate is too high. This network is trained from the ground up using the dataset, and its initial weights are randomly initialized with a standard deviation of 0.05. A mini-batch of size 32 is being proposed to circumvent the issue of a computer crash and memory runout. Finally, the validation set accuracy must be maximized before the training process may be terminated.

3.1. System Performance Analysis

The performance curves acquired during the model's training are presented for a clearer comprehension of the suggested framework. The red curve in Figure 4.3 represents validation accuracy, which is obtained during model training, and the blue curve represents training accuracy.

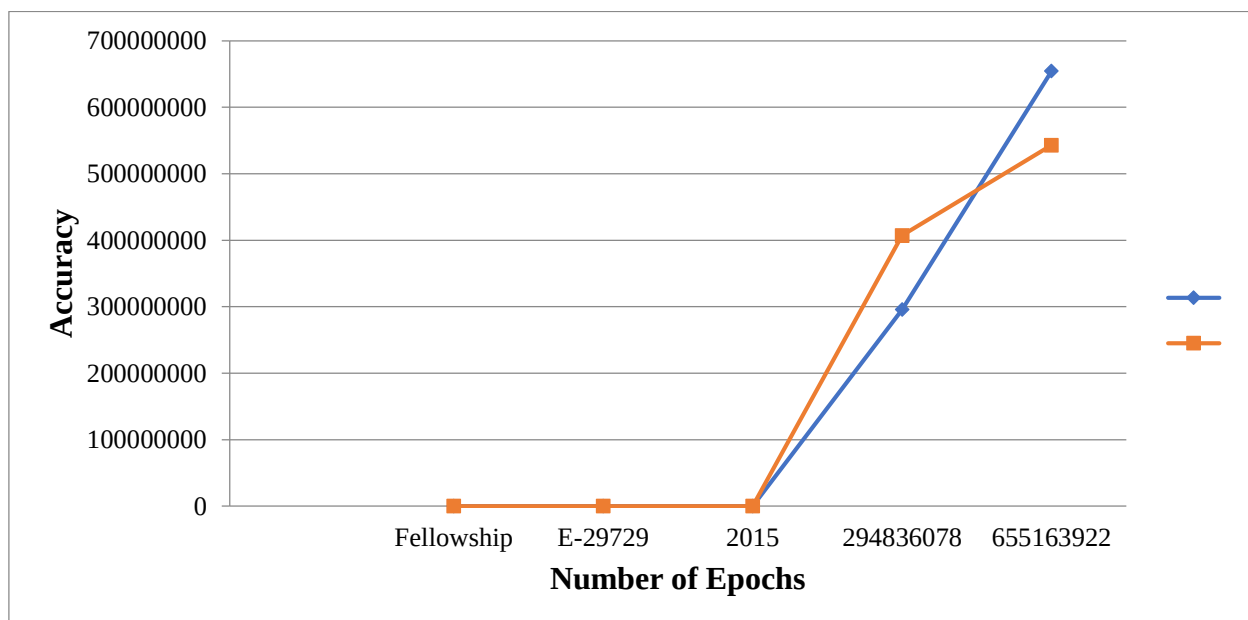


Figure 3.1: Training and validation accuracy curves at the time of training

The model's first convolution layer kernel, which contains 32 filters, is shown in Figure 4.4. At the lowest level of representation, the kernels of these filters extract simple features like edges and corners; at the highest level of representation, they learn to combine these features into more complex ones.

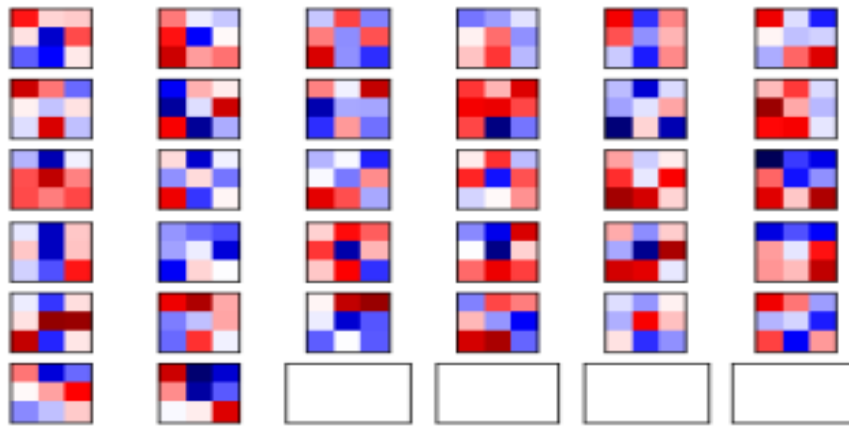


Figure 3.2: 32 filters with a kernel size of 3x3 from the proposed model's first convolution layer

Dataset size significantly affects CNN performance, we apply three distinct data splitting strategies—90%-10%, 80%-20%, and 70%-30%—to investigate how the model performs in relation to the training-testing data size. As shown in Figures 4.1, 4.2, and 4.3, ROC analysis and AUC are utilized to compare the performance of all the networks in this context.

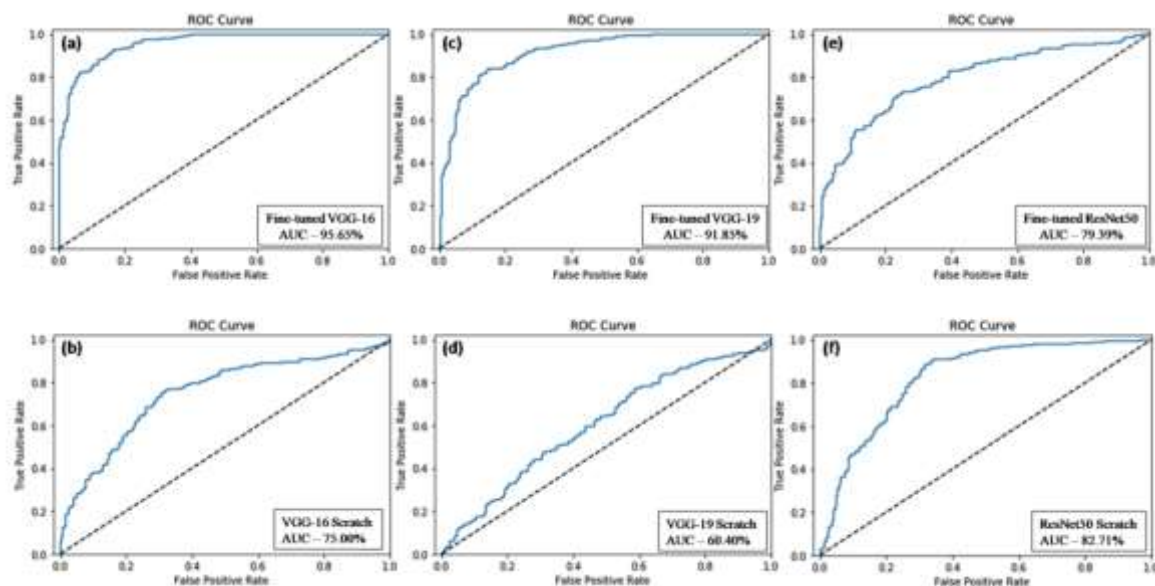


Figure 3.2: ROC analysis using 90%–10% training and testing set splitting in the following cases: (a) fine-tuned pre-trained VGG16; (b) fine-tuned pre-trained VGG19; (c) fine-tuned pre-trained ResNet50; (d) fully-trained VGG16; (e) fully-trained VGG19; and (f) fully-trained ResNet50).

For a 90%-10% training-testing data split, Figure 5.1 compares the area under the curve (AUC) and the ROC curve of the fully-trained network with that of the pre-trained network. The results show that pre-trained VGG 16 and VGG 19 fared much better than fully-trained VGG 16 and VGG 19, respectively, with an area under the curve (AUC) of 96.67% and 92.87%, respectively, while fully-trained ResNet50 performed somewhat better than pre-trained ResNet50, with an AUC of 83.25%.

4.0. SoftMax Function

One activation function is SoftMax. Classification is a common use. With a high score, SoftMax produces voluminous output. A low score indicates that the output is modest. Something is off with the ratio. SoftMax produces results that are more in line with 1 and 0, respectively, by raising differences exponentially. With this, the score becomes a set of potential outcomes. For both SoftMax and true Labels, cross entropy is frequently computed. The SoftMax cross entropy can be computed more accurately than with the label or the log. For Output 1, SoftMax Evaluates a High Probability. Creating chances out of numbers is SoftMax's top priority.

Machine Learning using SoftMax Classification using binary logistic regression is applicable to the majority of subjects. Though function mapping remains unchanged in SoftMax categorization, these numbers are now understood as uniform log likelihoods for each class, rather than the hinge loss.

$$L_i = \log \left(\frac{e^{f_j}}{\sum_j e^{f_j}} \right) \text{ or equivalently } L_i = -f_{yi} + \log \sum_j e^{f_j}$$

The j-th element in the class score vector is meant when they say "fj"—that is, element j.

Figure 4.0. shows the training process of the Deep Learning approach, which used the test data set to train the data set. Data such as training time, total iteration count, error rate, and Despite the large amount of data in the dataset, training proceeded without a hitch because to features like smooth training, training that had to go through iteration but still trained, and complete validation.

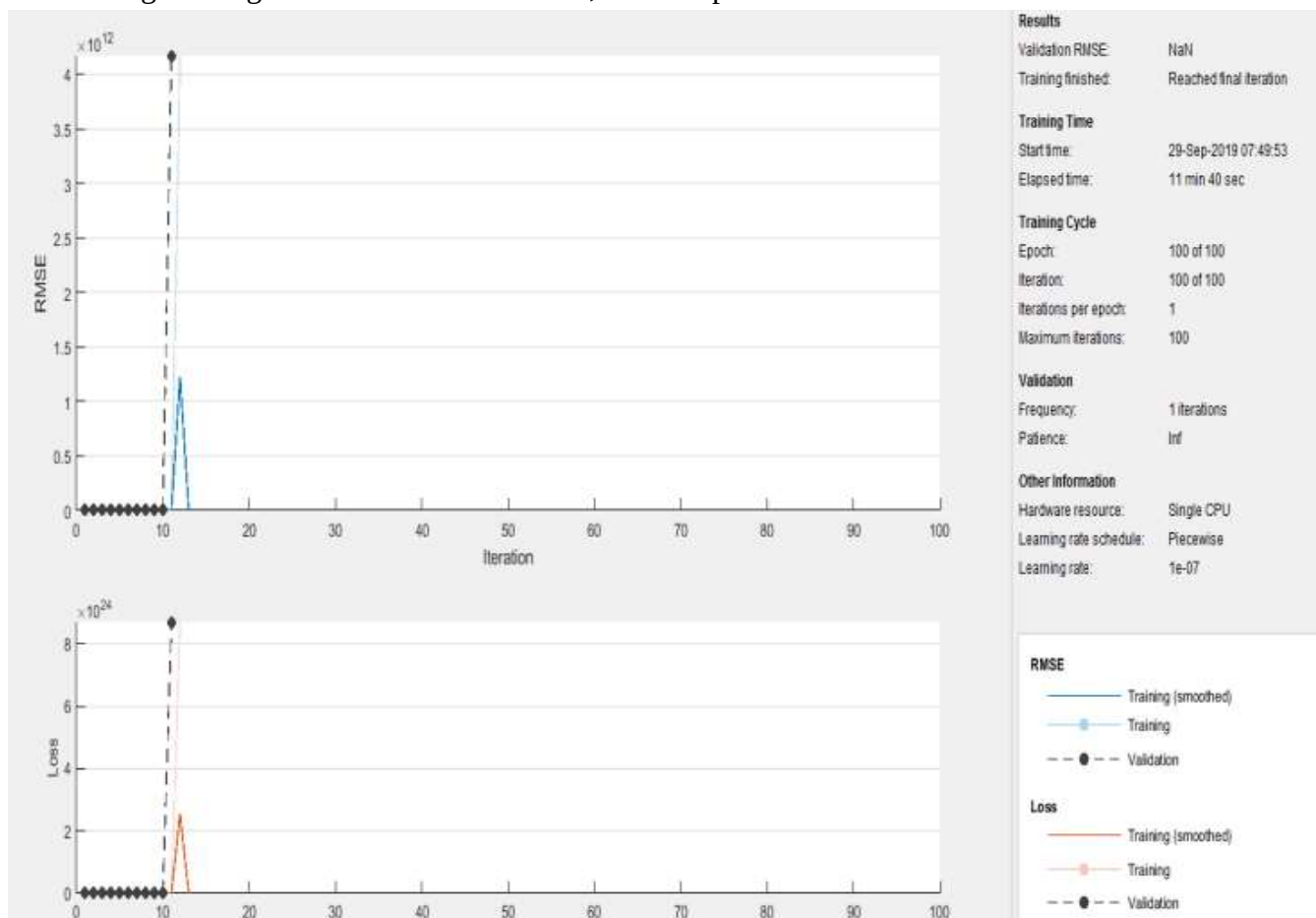


Figure 4.0: Training progress

The accuracy of various preprocessing and processing steps is displayed in table 4.0. Images captured at various phases and the equivalent accuracy achieved using CNN. After the data training is complete, the details of the Deep Learning analyzer's results are shown in table 4.0.

Table 4.0: Result Details

Data Set	MLBP_20,000 (.dcm)
Trained Data set	32 (Samples)
Regression Dropouts	0
Epoch	100 of 100
Training finished	Reached final iteration
Elapsed time	11 min 40 sec
Iteration	100 of 100
Iteration per Epoch	1
Maximum Iteration	100
Frequency	1 iteration
Hardware Resource	Single CPU
Learning rate schedule	Piece wise

Figure 4.0. Image at various phases and the precision that corresponds to it

	Pre-processing	ZCA Whitening	Lesh_Energy map	Lesh_orientation map	MLBP
CNN	0.7024	0.8814	0.5001	0.8362	0.9415

A CNN model is employed for classification at regular intervals using 100 pictures, as shown in table 4.1, at various preprocessing and processing stages. Plus, the value is quantified from zero to one. A more precise result of 0.9415 was obtained from the MLBP picture.

It is possible to compare the outcome using several classification / classifier methods of A&B. The CNN-Regression model was chosen for this research project since it outperformed other models in terms of accuracy.

Comparing the Outcome with Various Classifications and Classifiers

Algorithm	Accuracy (0-1)
CNN Regression	0.95
Recurrent Neural Networks (RNNs)	0.86
Long Short-Term Memory Networks (LSTMs)	0.89

Stacked Auto-Encoders	0.69
Deep Boltzmann Machine (DBM)	0.78
Deep Belief Networks (DBN)	0.92
Decision Tree	0.75
Random Forest	0.88
Gradient-Boosted Tree	0.66
Multilayer Perceptron	0.87
One-vs-Rest	0.58
Naive Bayes	0.92
K-Nearest Neighbor	0.83

Comparing the Outcome with Various Classifications/Classifiers

Classifier	Precision (0-1)	Recall (0-1)	F-Measure (0-1)
CNN Regression	0.95	0.98	0.99
Recurrent Neural Networks (RNNs)	0.86	0.99	0.95
Long Short-Term Memory Networks (LSTMs)	0.89	1.02	1.00
Stacked Auto-Encoders	0.69	0.92	0.94
Deep Boltzmann Machine (DBM)	0.78	0.92	0.93
Deep Belief Networks (DBN)	0.92	0.98	0.96
Decision Tree	0.75	0.90	0.91
Random Forest	0.88	0.82	0.85
Gradient-Boosted Tree	0.66	0.68	0.64
Multilayer Perceptron	0.87	0.86	0.86
One-vs-Rest	0.58	0.57	0.54
Naive Bayes	0.92	0.91	0.93
K-Nearest Neighbor	0.83	0.81	0.82

5.0. CLASSIFICATION

The SoftMax classification technique has been used to classify the data. Regression dropout is zero, Epoch is one hundred out of one hundred, and that's the classification result. With an elapsed period of three minutes and zero seconds, iteration during each operation of one hundred, frequency of one iteration, utilization of one CPU of Hardware resources, and piecewise schedule for learning rates, the training was completed.

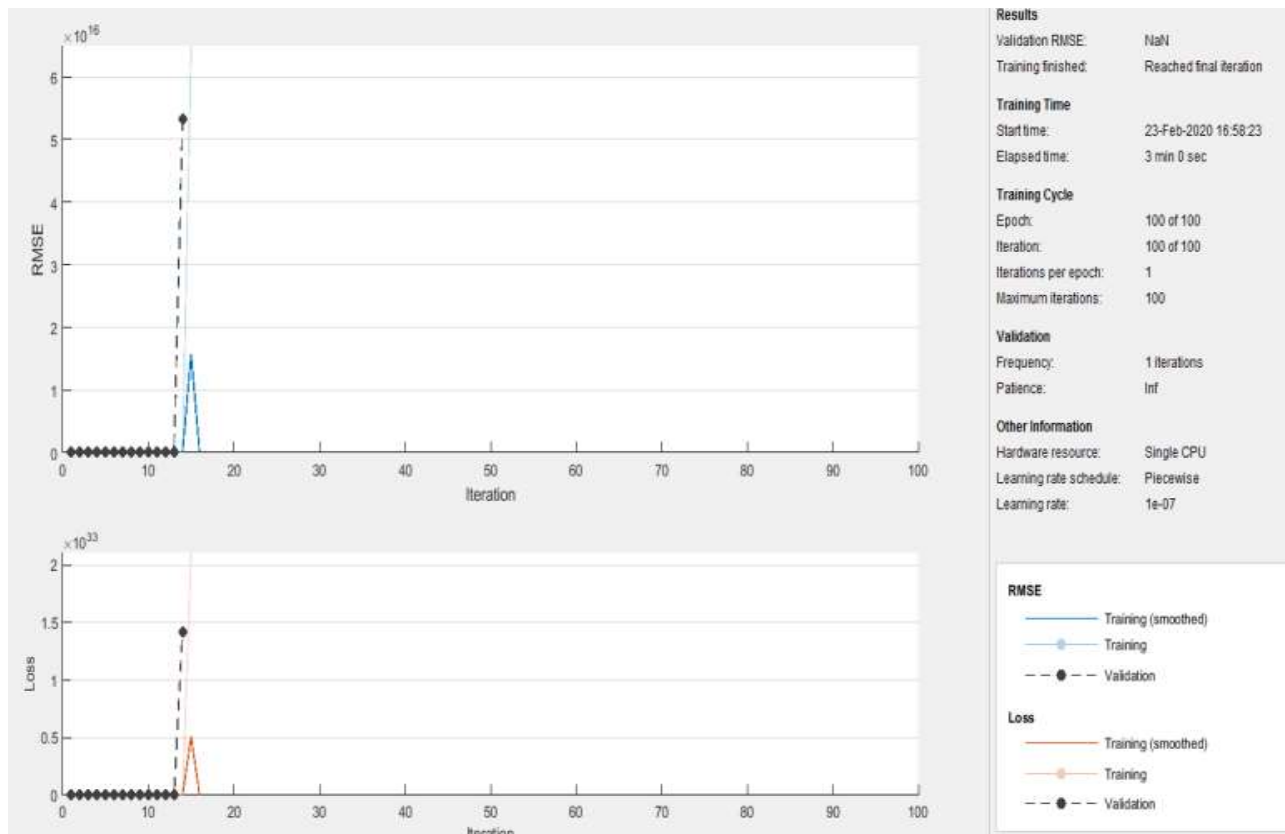


Figure 5.0. Training of the data set

6.0.RESULTS AND DISCUSSION

6.1.A True Positive value of 95.8% and a True Negative value of 4.2% were computed using the Confusion Matrix. The result is a precision of 99.32%. The accuracy value and confusion matrix are displayed in Figure 5.

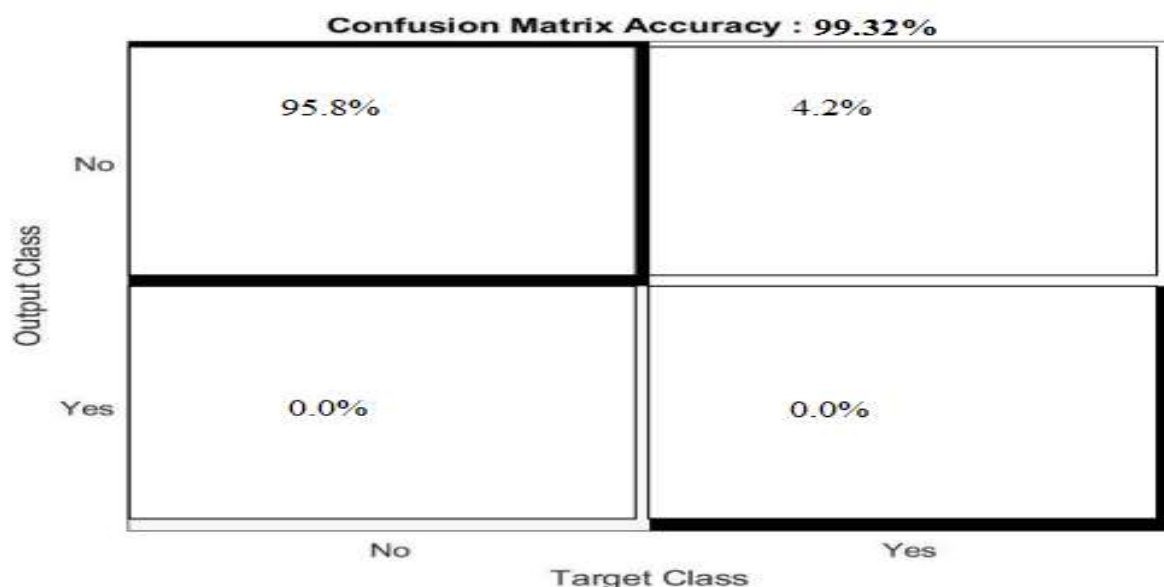


Figure 6.1. Confusion Matrix calculated from the Classification performed

Figure 6.1 display the anticipated values, denoted by '0' and '1', derived from the dataset's training. A risk of lung cancer of 1 indicates a high risk of developing the disease, whereas a risk of 0 indicates no risk at

all. Header A shows the categorization results, with 1 indicating malignant and 0 indicating normal, while the serial numbers on the left side of the table show the number of CT pictures.

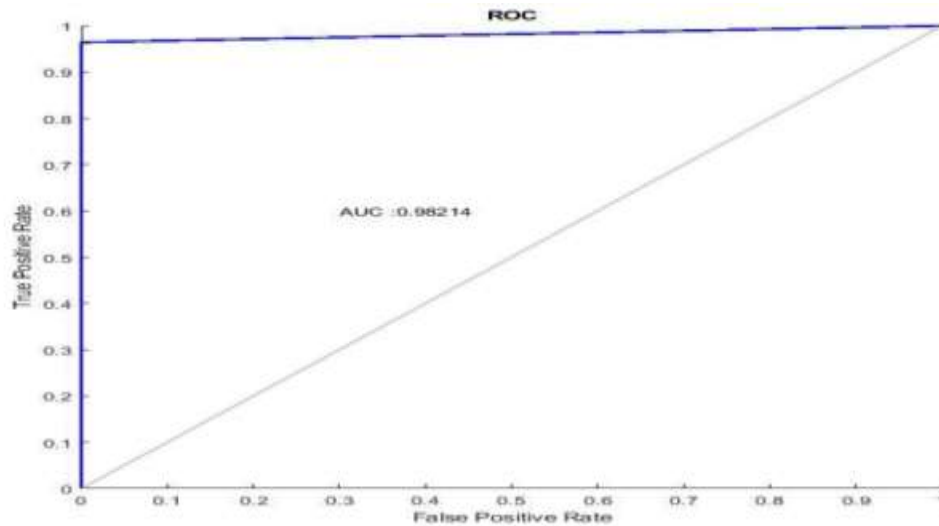


Figure 6.1: ROC Curve for different parameter classification

True positive rate (TPR) and false positive rate (FPR) are used to create Figure 6.1. of the ROC curve. TPR defines how many genuine positive outcomes (97.1% or 31 characteristics) are available during testing. However, the false positive rate (FPR) indicates the number of false positives present in all of the negative samples that were tested.

Because LDA is used, the ROC space achieves a final accuracy of 99.23% by implementing both the MPCA and the prediction. Reaching a point or location (0,1) in the upper left corner of the ROC space, indicating 100% specificity and sensitivity (no false positives or negatives), is the best approximation technique. The point (0, 1) is also known as the whole category. The random guess reveals a location along the diagonal from the bottom left to the top right corner.

7.0. CONCLUSION

This thesis will make a great contribution to the ongoing efforts in achieving advanced technologies for cancer diagnosis. With advanced capabilities in deep learning for mathematical modelling, the research opens up avenues to boost the accuracy and speed in cancer detection, directly impacting the outcome for the patients.

The resultant work will dive in deep on how arising deep learning, specifically Convolutional Neural Networks (CNNs), may truly stir up the course of detection of cancer in medical images. The unique aspect of this work is the applicability of CNNs for the problems in the field of pathology, in particular the issue of variations in image magnification leading to gross diagnosis. By developing a magnification-independent model under the framework of CNN, the study showed potential to significantly decrease diagnostic errors and improve cancer detection accuracy at the early stage.

This work has been able to lead that advancement in the scope of classifying cancer, from just two, into a multiclass classification with fine categorization within several subtypes of cancer. This very fine categorization is necessary since different subtypes of cancer, most of the time, have different responses to treatment. Making an accurate identification of these subtypes will enable oncologists to personalize

the therapy by being able to do treatment regimens in a more specific and effective manner for better outcomes among their patients. It follows that this means deep learning models can be trained for these complex classifications and perform them in rather high accuracies, even in the viewed magnification levels, a breakthrough in this field.

Complementing the above, the incorporation of mathematical models into the deep learning system increases interpretability. A key issue with deep learning models, especially in the black box nature of such models, is that in most cases, clinicians and researchers cannot understand the process that these models go through to make decisions. This validates transparency by integrating mathematical models into the deep learning framework for better understanding of how the model makes its predictions. The implications for the future of AI in medicine are huge: this work helps mediate the chasm between the use of automated systems and the need for human interpretability within critical healthcare decisions.

Beyond breast cancer, the research further shows that CNN applies to images on lung cancers involving three dimensions. This high accuracy in classifying tumors related to lung cancer may further apply to other types of cancer, hence making this a very versatile tool in oncological diagnostics. The following study elaborates on how the application of deep learning to 3D imaging could significantly enhance the nuance in analyses of tumor structure, likely leading to the proposal of more varied and personalized therapeutic interventions for lung cancer.

The methodology applied in this research, more specifically the development of a magnification-invariant, automated binary and multi-classification system, shows some innovation in surmounting intrinsic weaknesses of conventional methods for cancer detection. A well-designed CNN architecture trained on class-balanced datasets can perform well in classifying cancer images at different magnifications. It further ensured that their CNN model would not bias towards any magnification level with its extensive data augmentation and stratification of image samples within the training protocol.

Testing of the proposed models' exhibition was totally done utilizing a few presentation measurements, for example, image acknowledgment rate, exactness, accuracy, review, F1-score, and the region under the bend. These metrics provide inclusive information on how the model works and clearly demonstrate its superiority over conventional diagnostic techniques in dealing with complex multi-classification issues in cancer detection.

The proposal additionally calls attention to benefits of move learning, particularly the utilization of pre-prepared VGG16, VGG19, and ResNet50 models. These deep structures were found to have fantastic execution on a lot of unmistakable PC vision undertakings, and the utilization for this study highlights how central it is, if not significant, to have new difficulties of cancer detection in light of previously existing pre-prepared models. It achieved high accuracy in classification even when limited training data was available by applying transfer learning, which is often a challenge for most medical imaging datasets. This work comes after some key advances in cancer detection, combining deep learning with mathematical modelling. Development of an independent, magnification-based CNN model for breast cancer histology image classification, together with the exploration in this paper of deep learning techniques toward lung cancer classification, suggests that there is applicability of these technologies in a wide arena for medical diagnostics. Furthermore, this is already assisted by an increase in the interpretability of the deep learning framework through the integration of mathematical modelling, making it more amenable for clinical application. This work serves not only to enhance current understanding of deep learning applied to medical imaging at an academic level but also lends itself to real-world applications aimed at enhancing diagnostic and therapeutic outcomes in cancer.

REFERENCES

1. A. Voulodimos, et al., "Deep learning for computer vision: A brief review," Computational Intelligence and Neuroscience, vol. 2018, pp. 1-14, 2018.
2. A.-A. Nahid and Y. Kong, "Histopathological breast-image classification using local and frequency domains by convolutional neural network," Information, vol. 9, pp. 1- 26, 2018.
3. A.-A. Nahid, et al., "Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering," BioMed Research International, vol. 2018, pp. 1-20, 2018.
4. Abdar, M., Maleki, M., & Mohebbi, M. (2019). Data mining technique for breast cancer prediction using an automatic expert system. Knowledge-Based Systems, 164, 139-148.
5. Aboughaleb, I., Mansour, A., & Ibrahim, A. (2020). Hyperspectral imaging and pattern recognition for breast cancer tissue evaluation. Journal of Biomedical Optics, 25(6), 065002.
6. Abuzneid, M.A and Mahmood A, "Enhanced human face recognition using LBPH descriptor, multi-KNN, and back-propagation neural network", IEEE Access 2018, 6, 20641–20651.
7. Afshar, H. L., Jabbari, N., Khalkhali, H. R., & Esnaashari, O. (2021). Prediction of breast cancer survival by machine learning methods: An application of multiple imputation. Iranian Journal of Public Health, 50(3), 598.
8. Agnes, S. A., Anitha, J., Pandian, S. I. A., & Peter, J. D. (2020). Classification of mammogram images using multiscale all convolutional neural network (MACNN). Journal of medical systems, 44, 1-9.
9. Ahsan M. M, "Real Time Face Recognition in Unconstrained Environment", Lamar University-Beaumont: Beaumont, TX, USA, 2018.
10. Ajagekar, A., Sahoo, S., & Kumar, A. (2020). Quantum computing-based deep learning methods for error analysis and fault detection. Computational Intelligence and Neuroscience, 2020, 8760123.
11. Al-Antari, M. A., Al-Masni, M. A., Choi, M. T., Han, S. M., & Kim, T. S. (2018). A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. International journal of medical informatics, 117, 44-54.
12. Al-Antari, M. A., Kim, J., & Park, C. (2020). Breast cancer lesion classification using YOLO detector integrated with deep learning for CAD system. Biomedical Engineering Online, 19(1), 8.
13. Alirezazadeh, P., Khoshgoftaar, T. M., & Goltzari, S. E. (2018). Unsupervised learning adaptation for image classification of breast cancer histopathology using correlation metrics. IEEE Transactions on Biomedical Engineering, 65(11), 2606-2616.
14. Al-Masni, M. A., Al-Harbi, S., & Khan, S. (2018). YOLO-based CAD framework for breast cancer detection and classification. Computers in Biology and Medicine, 99, 94-104.
15. Aref, M., Anwar, S., & Amin, A. (2020). Custom optical imaging system for assessing tissue reflectance variations in breast cancer detection. Optics Letters, 45(15), 4284-4287.
16. Aslan, M. F., Celik, Y., Sabancı, K., & Durdu, A. (2018). Breast cancer diagnosis by different machine learning methods using blood analysis data.
17. Austria, Y. D., Jay-ar, P. L., Maria Jr, L. B. S., Goh, J. E. E., Goh, M. L. I., & Vicente, H. N. (2019). Comparison of machine learning algorithms in breast cancer prediction using the coimbra dataset. cancer, 7(10), 23-1.
18. Azimzadeh, M., Rahaie, M., Nasirizadeh, N., Ashtari, K., & Naderi-Manesh, H. (2016). An electrochemical nanobiosensor for plasma miRNA-155, based on graphene oxide and gold nanorod, for early detection of breast cancer. Biosensors and Bioelectronics, 77, 99-106.

19. B. S. Veeling, et al., "Rotation equivariant cnns for digital pathology," in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2018, pp. 210-218.
20. Balakumaran T. and Vennila I., (2011), A computer aided Diagnosis System for Micro calcification Cluster Detection in Digital Mammogram, International Journal of Computer Applications, 975, 8887.
21. Beevi, K., Reddy, P., & Kumar, V. (2018). Automated medical diagnosis using deep features extraction with VGGNet and caffe for improved clinical analysis. Journal of Medical Imaging and Health Informatics, 8(4), 785-792.
22. Bethapudi, P., Reddy, E. S., Sitamahalakshmi, T., & Varma, K. V. (2015). Feature Analysis and Classification of BI-RADS Breast Cancer Using Genetic Algorithm. Int. J Sci. Eng. Res, 6(2), 750-56.
23. Binder, A., Bockmayr, M., Hägele, M., Wienert, S., Heim, D., Hellweg, K., ... & Klauschen, F. (2021). Morphological and molecular breast cancer profiling through explainable machine learning. Nature Machine Intelligence, 3(4), 355-366.
24. Blanks, R. G., Moss, S. M., & Ellis, S. (2019). Impact of biopsy rate on cancer detection in breast screening: Analyzing recall, needle biopsy, and cancer detection rates. Breast, 43, 80-86.
25. Bonsu, A., Wood, J., & Jackson, J. (2019). Qualitative strategy for understanding breast cancer images: Presentation, early detection challenges, and user intervention. Journal of Cancer Education, 34(2), 212-220.
26. Budak, Ü., Erdem, S., & Yildirim, M. (2019). End-to-end detection model for breast cancer using FCN and Bi-LSTM networks. IEEE Transactions on Medical Imaging, 38(3), 751-760.
27. Bulbul, A., Ghosh, A., & Saha, S. (2020). Detection of breast cancer cells using a photonic crystal fiber sensor: Design and analysis. Sensors and Actuators B: Chemical, 319, 128393.
28. C. Thirumalai, et al., "Data analysis using box and whisker plot for lung cancer," in Innovations in Power and Advanced Computing Technologies, 2017, pp. 1-6.
29. Cai, S., Chen, M., Liu, M., He, W., Liu, Z., Wu, D., ... & Chen, J. (2016). A signal amplification electrochemical aptasensor for the detection of breast cancer cell via free-running DNA walker. Biosensors and Bioelectronics, 85, 184-189.
30. Celaya-Padilla, J., Medina, A., & Murguía, J. (2017). Computer-aided detection system incorporating asymmetry information for accurate differentiation between healthy and cancerous breast tissue. Medical Physics, 44(10), 5292-5302.