A Comparative study between the Statistical Approach and the Deep Learning Approach in Classifying Histopathological Images

A PROJECT REPORT

Submitted by

BL.EN.U4CSE19091 Nagulapati Kavya Reddy

BL.EN.U4CSE19110 Preethi Reddy Mudireddy

BL.EN.U4CSE19140 Vattikuti Shravya

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

AMRITA SCHOOL OF COMPUTING, BANGALORE

AMRITA VISHWA VIDYAPEETHAM

BANGALORE 560 035

DECEMBER – 2022

AMRITA VISHWA VIDYAPEETHAM

AMRITA SCHOOL OF COMPUTING, BANGALORE, 560035

BONAFIDE CERTIFICATE

This is to certify that the project report entitled **"A Comparative study between the Statistical Approach and the Deep Learning Approach in Classifying Histopathological Images"** submitted by

BL.EN.U4CSE19091	Nagulapati Kavya Reddy
BL.EN.U4CSE19110	Preethi Reddy Mudireddy
BL.EN.U4CSE19140	Vattikuti Shravya

in partial fulfillment of the requirements as part of **Bachelor of Technology** in "**COMPUTER SCIENCE AND ENGINEERING"** is a bonafide record of the work carried out under my guidance and supervision at Amrita School of Computing, Bangalore.

Ms. Radha D	Dr. Sriram Devanathan
Assistant Professor	Professor and Chairperson,
Dept. of CSE	Dept. of CSE

This project report was evaluated by us on 31/12/2022

EXAMINER1 EXAMINER2

ACKNOWLEDGEMENTS

The satisfaction that accompanies successful completion of any task would be incomplete without mention of people who made it possible, and whose constant encouragement and guidance have been a source of inspiration throughout the course of this project work.

We offer the sincere pranams at the lotus feet of **"AMMA", MATA AMRITANANDAMAYI DEVI** who showered her blessing upon us throughout the course of this project work.

We owe the gratitude to **Prof. Manoj P.**, Director, Amrita School of Engineering, Bangalore.

We thank **Dr. Sriram Devanathan,** Principal and Chairperson-CSE, Amrita School of Engineering, Bangalore for his support and inspiration.

It is a great pleasure to express the gratitude and indebtedness to the project guide **Ms. Radha D,** Assistant Professor, Department of Computer Science and Engineering, Amrita School of Computing, Bangalore for her/his valuable guidance, encouragement, moral support, and affection throughout the project work.

We would like to express the gratitude to project panel members for their suggestions, encouragement, and moral support during the process of project work and all faculty members for their academic support. Finally, we are forever grateful to the parents, who have loved, supported and encouraged us in all the endeavors.

ABSTRACT

One of the most prevalent cancer tumors in women, Breast cancer has quite a significant effect on a female's physical and intellectual well-being and potentially poses a life-threatening risk. As it can be exceedingly challenging to determine the true cause of breast cancer, early diagnosis is crucial to lowering the disease's fatality rate. The likelihood of survival increases by up to 8% with early cancer identification. Radiologists examine breast shots obtained from X-rays, Magnetic resonance imaging or mammogram scans to find anomalies. Even experienced radiologists, however, struggle to recognize characteristics like micro-calcifications, lumps, and masses, which results in significant false positive and false negative rates. Using various methodologies like machine learning, deep learning, statistics, etc., this task becomes less burdening. This commitment seeks to compare the efficiency of the statistical approach and deep learning approach when given a task to classify histopathological images into two categories, benign and malignant.

Various models such as SVM, Decision tree, and random forest were employed after preprocessing the dataset to help classify the histopathological images under the statistical approach. Deep learning models such as AlexNet, ResNet, DenseNet, MobileNet, VGG16, and ShuffleNet are used to achieve the same goal. The likelihood of the patient making a quicker recovery can be increased with early identification. This objective drives us to develop a model that can quickly and accurately classify histopathological images.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER – 1
INTRODUCTION

Many researchers have lately mentioned that the mammary cancer related deaths among women have risen. The disease's tardy detection and challenging therapy are the primary contributors to the poor survival rate. Therefore, It is crucial to diagnose breast cancer sooner to ensure adequate treatment and lessen the likelihood that carcinoma may spread to other body cells.

Cancer develops when normal cells undergo genetic alteration, spreads throughout the body, and becomes fatal if detected late enough to be treated. invasive and non-invasive breast cancer are the two subtypes.

The first is considered to be malignant, noxious and capable of traveling to other organs The latter doesn't hurt the body and doesn't spread to other organs.

To study the development of cancer in organs, histopathology is a process in which a specialist/pathologist performs a microscopic examination and thorough analysis of a biopsy sample to examine the growth of cancer in organs.

Digital histopathology on human tissue samples has been made possible by the use of computerized image assessment and techniques for machine learning. Over the past decade, the development of digital pathology was from the use of Complete tissue samples that are now routinely digitally scanned using microscopes and cameras. Computer-aided design (CAD) systems have been integrated into the everyday operations of pathology laboratories because of factors including large gains in processing power, less costly storage devices, and major breakthroughs in photogrammetric techniques in recent years. These medical technologies were created to aid in illness detection, prognosis, and diagnosis, working in tandem with the pathologist's experienced human judgment. This improved CAD system's automated analysis of biological data can assist the pathologist in making a prompt or early identification.

In order to help in the early diagnosis of breast cancer and lower illness mortality, Computer aided software solutions are still needed to lessen the workload on pathologists by isolating and filtering noticeably benign sections.

CHAPTER - 2

LITERATURE REVIEW

2.1.

Main author FaezehSadat Shahidi along with Salwani Mohd and others proposed "**Breast Cancer Classification using Deep Learning Approaches and Histopathology Image: A comparison study**" in 2020.

Methodology

Histopathological images from 5 different publicly available datasets were combined to form the dataset. To increase the model's functionality, data is augmented. The different data augmentation techniques include, rotation, rescaling, random cropping, color shifting, flipping etc. The images are all resized and the dataset is checked for the balancing problem. If the data is imbalanced, sampling is done to remove the biasing. As a part of the pre-processing, the images are all normalized in the contrast of the image so as to be able to clearly demarcate the region of interest. This data is then fed to various deep learning models and performance of each is compared.

Observations

It was established that, out of all the estimated models, the SENet-154 network that was trained beforehand had the best score of 99.87%. DualPathNet-131 earned the second spot in the accuracy rankings with a score of 98.73%.

Future work

Future work in the paper was to improve the accuracies currently obtained and also add more models to the study.

2.2.

Hanan Aljuaid, Nazik Altuki, Najah Alsubaise, Lucia Cavallaro, Antonio Liotta proposed "**Computer aided diagnosis for breast cancer classification using deep neural networks and transfer learning**" in 2022.

Methodology

BreakHis dataset was used for the implementation of this project. The 7909 cumulative specimens that are currently accessible were gathered from 82 patients. The distinct amplification factors, specifically 40, 100, 200, and 400, are contained in the dataset. It had 2480 benign and 5429 malignant entries. The first step was to remove noise. Two different filters were used for this purpose, gaussian filter and median filter. These filters retained the main features of the images. Data augmentation was implemented to increase the accuracy as well. This also helped overcome the overfitting problem. All images were resized to the required size of the Deal learning model. Of these images 65% were sent for training while the rest 35% were sent for testing.

Observations

The accuracy of InceptionV3, NetResNet, and ShuffleNet were 97.66%, 99.7%, and 96.44%, respectively., respectively in the binary class classification and 96.07%, 97.81%, and 95.79% respectively for the multi class classification.

Future work

Future study will try to expand the capability of the model and boost the accuracy of breast cancer diagnosis by using more reliable datasets or patient-based picture datasets.

2.3.

Jin Huang, Liye Mei, Mengping Long, Yiqiang Liu, Wei Sun, Xiaoxiao Li proposed "**BM-Net: CNN-Based MobileNet-V3 and Bilinear Structure for Breast Cancer Detection in Whole Slide Images**" in 2022.

Methodology

n this paper, WSI-whole image slides (tissues under the microscope) are taken as input and are classified if the nature of the tissue is normal or it is affected by carcinogenic nodes. Since the WSI would be too large to take as an input, an option that can be done is to compress the image, but this leads to loss of important features. To solve this issue, the WSI is divided into patches, and each patch is given a designation before the neural network is taught how to recognize the target characteristics. Once each patch is segmented to find the region of interest and the patches that fall under the background are deleted, they are stitched back together.

Observations

These techniques have proved to work well with the binary classification task but faced challenges to identify 4 classes. In the paper, they have used mobienet-v3 and the bilinear model - net (BM-Net) to extract features and train. This reduced the computational time and space required as these are lightweight deep learning models. BM-Net had better performance ratio compared to MobileNet but one model outperforms the other by a very small margin that is essentially negligible.

2.4.

Sobia Shakeel along with Gulistan Raja proposed, in their paper published in 2021 about the classification of cancerous tissues using DCNN

Methodology

The novel Computer aided method put out in this article uses specialized DCNN to identify and categorize breast cancer into either malignant and benign variants. First, a neighborhood segmentation approach is used to extract the area of focus. This technique is then improved with counterbalanced adaptive histogram equalization. Later, to learn properties from mammograms, a tailored deep convolutional neural network is applied. Breast masses are classified as benign or malignant using a distance-based hyperplane separator.

Observations

88.7% accuracy is achieved by the model. System specific, responsiveness, precision, F1-measure, and AUC are also recorded as 0.93, 0.841, 0.917, 0.877, and 0.885, respectively.

2.5.

Feyza YILMAZ, Ahmet DEMİR, Onur KÖSE Proposed, "**Comparison of Two Different Deep Learning Architectures on Breast Cancer**" in 2019.

Methodology

In this investigation, the deep learning architectures DN-201 and Xception are used. On the breast cancer dataset, the performance of these two distinct deep learning algorithms is examined. The collection includes photos of both aiding and hostile tumors. There are 5713 testing photos and 20748 learning images.

Observations

When trained on the utilized dataset, the DenseNet-201 technique yields a score of 92.24%, while the Xception method achieves an F-1 average accuracy of 92.41%.

2.6.

Majid Nawaz along with others proposed "**Multi-Class Breast Cancer Classification using Deep Learning Convolutional Neural Network**" in 2018.

Methodology

The suggested method seeks to identify tumors as malignancy, but at the same time also forecast the subtype of tumors such as Lesion, Lobular carcinoma, and so on. Each of these classes have various class specific features which will be extracted from each image. This method is done using deep learning models.

Observations

The DenseNet CNN model exhibited good processing capabilities with 95.4% accuracy throughout the multi-class carcinoma classification test utilizing histopathology pictures from the BreakHis dataset.

Future work

The performance achieved is aimed to be enhanced if offered more information drawn from a larger dataset.

2.7.

Xiaofei Zhang and his team proposed "**Whole mammogram image classification with convolutional neural networks**" in 2018.

Methodology

They constructed and validated many CNN strategies for whole-mammography picture classification in this work. All pictures were taken at the University of Kentucky's Department of Radiology. There are 3,018 unfavorable and 272 favorable mammary pictures in the collection. To improve the performance of the classifiers, techniques such as data enhancement and manipulation along with transfer learning are used with CNN models. In augmentation, the original and reflected images were rotated 270, 90, and 180 degrees. An original image was therefore expanded to eight photos. Thus, every image in the supplemented dataset was then scaled up to 832 x 832 pixels, the grade set with the intention of preserving tumor digital signals. These images are fed to different deep learning models. They tested 7 different CNN architectures.

Observations

The conclusion arrived by this research was that the combination of various data augmentation techniques performed on the dataset works better to train and model and also helps achieve better accuracies.

2.8.

Ghada Hamedand his team of seven proposed "**Deep Learning in Breast Cancer Detection and Classification**" in 2017.

Methodology

Five various datasets containing mammary images were selected and combined to perform this analysis. A comparative analysis for dataset preparation for recent work in breast cancer detection and classification was prepared. They are studied in terms of several aspects such as the dataset they utilized, whether or not they performed processing beforehand on this data, whether or not data manipulation and amplification was done, and the mammography images they were trained upon.

Observations

According to the findings of this study, the most recent and widely used suggested and referred models for mass detection and classification are AlexNet, GoogleNet. A newer model, YOLO, and RetinaNet were included as well. YOLO and RetinaNet are the most recent and most accurate breast detection and categorization researchers.

SUMMARY

In all the papers mentioned above, A pattern where data augmentation has helped improve the performance of the models and also helped overcome the overfitting problem has been noticed. It has been found that AlexNet and ResNet seem to work the best with mammogram image classification with accuracies spanning between 95% for multiclass classification to 99.7% for binary classification.

CHAPTER – 3
SYSTEM SPECIFICATIONS

3.1 Software requirements

Front-end requirements:

- Google Colaboratory

  Its primary uses were for running Python programs and importing different libraries, like openCV, tensorflow, numpy, etc.

Back-end requirements:

- Google Drive

  It was primarily utilized to store files securely and to make the dataset accessible across a variety of devices.

3.2 Hardware requirements

a. Processor: i5, 10^{th} gen, GPU

b. Graphics: 8GB

3.3 Packages used

This project was built using a variety of Python packages. Among them, the potentially significant ones are,

- OpenCV

  The term "open cv" refers to an open-source computer vision library. This library's primary goal is to offer support for applications involving computer vision. It is also employed in image processing.

- Keras

  Keras is a Google API that enables us to develop and employ multiple deep learning models.
- Scikit-learn

  It is a free and open-sourced analytical library that aids in the implementation of various machine learning algorithms and models.
- Skimage

  Skimage is a free and freely available image processing library. It supplies us with many methods, such as regionprops(), which allows us to retrieve image features.

CHAPTER - 4
SYSTEM DESIGN

This chapter describes the project's process. The fundamental purpose of the study is to discriminate between benign and malignant histopathological images. Two approaches are used to do this: a statistical method and a deep learning-based technique. Fig 4.1 displays the project's workflow as a flowchart.

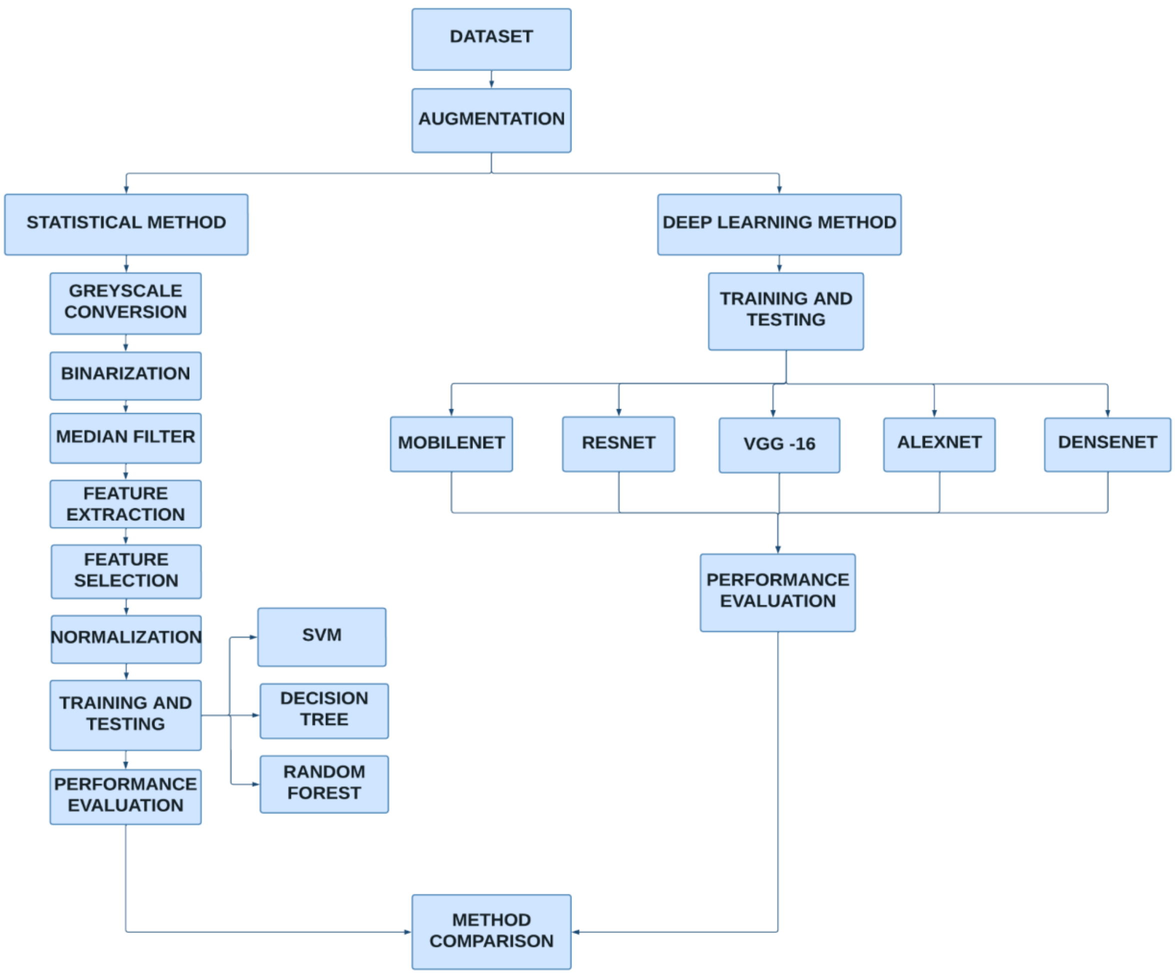


Fig 4.1 *Workflow diagram of the project.*

The objective is to evaluate the performance of each methodology and draw a comparison between them. Fig 4.1 depicts the various modules that the project undergoes in the statistical method shown on the left and in the deep learning-based approach shown on the right of Fig 4.1. Getting a collection of histological pictures that included both normal and malignant tissue was the initial step. Then, in order to increase the quantity of the data and improve the training of the deep learning models, image augmentation was performed. The new augmented dataset was used as the primary dataset for both the approaches.

The statistical technique employed a rule-based formula in which the connection between distinct picture elements, also known as features, was used as a metric to categorize images as benign or malignant.

The entire categorization process is a black box in the deep learning technique. The data is loaded into deep learning models, which provide an output.

CHAPTER – 5
SYSTEM IMPLEMENTATION

Based on the project workflow, this chapter describes the many modules that were implemented in the project. The first stage is to collect a database, which is then augmented. On this data, a statistical and deep learning technique is implemented.

In the statistical approach the various modules implemented are,

- Grayscale conversion
- Binarization
- Median Filter
- Feature extraction
- Feature selection
- Normalization
- Machine learning model implementation

In the deep learning approach, we feed the data to various deep learning models such as,

- AlexNet
- DenseNet
- ResNet-50
- VGG-16
- MobileNet

Deep learning models are tested using their basic code, pre-trained models, and pre-trained models using an SVM classifier.

5.1 Dataset Acquisition

The dataset includes histopathological images captured from cancer patients with benign and malignant tumors. There are three folders in the dataset: train, test, and valid. Each of these folders contain two separate classes of images: benign and malignant. Images are added from the valid folder to the training images in order to increase the number of images in the training set.

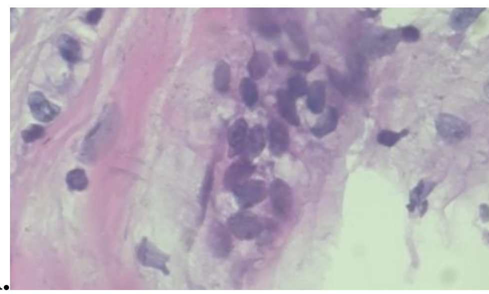
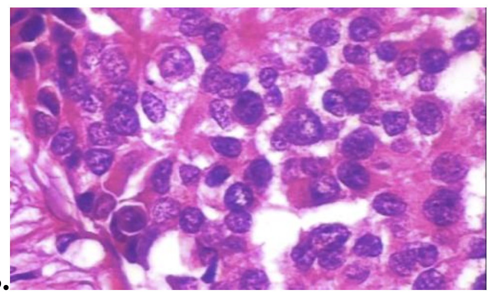
A.  **B.** 

Fig 5.1.1 *Benign and Malignant images*

A sample picture of a benign cancerous tissue is shown in Fig 5.1.1 A, and a tissue attacked by malignant cancer is shown in Fig 5.1.1 B.

Below in Table 5.1.1 are the statistics for the dataset. It shows that there are a total of 386 images under benign and 422 under malignant for training.

Table 5.1.1 *Dataset Description*

FOLDER	BENIGN	MALIGNANT
TRAIN	386	422
TEST	64	65
VALID	97	92

5.2 Data Augmentation

Data augmentation is a technique used to lessen the issue of overfitting a model. Overfitting is the term used to describe a situation when a model has excellent training accuracy but produces inaccurate predictions during testing. By exposing the model to numerous iterations of the picture, the issue may be solved.

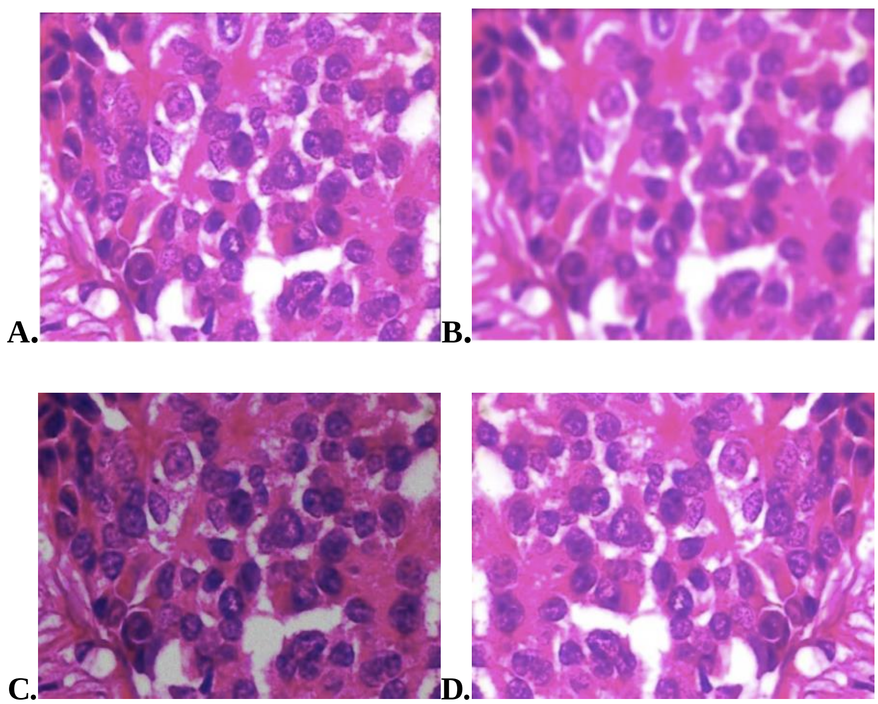


Fig 5.2.1 *Original image and its augmented images*

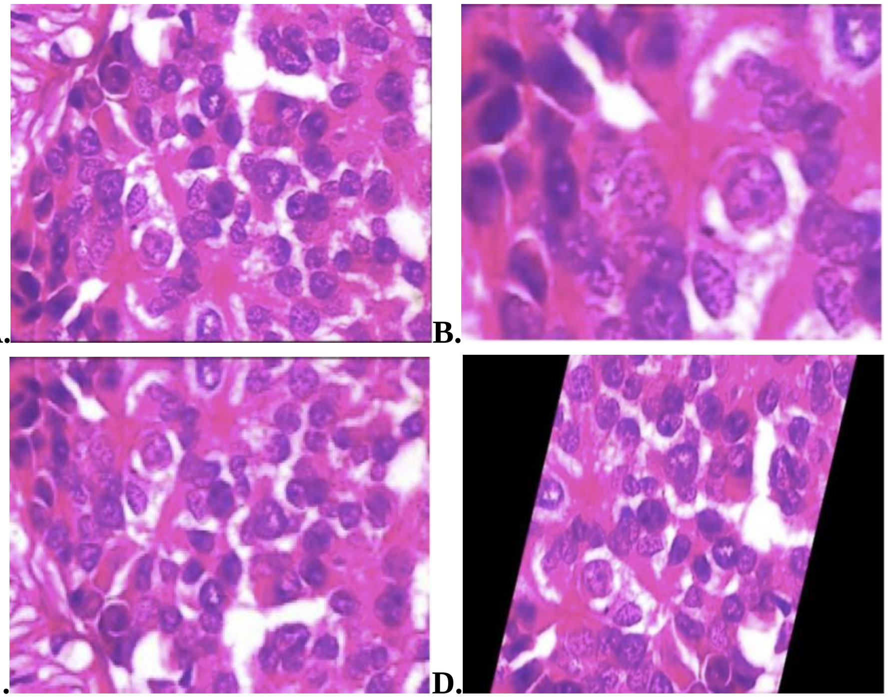
A. B. C. D.


Fig 5.2.2 *Augmented images of the original image*

The original image is shown in Fig. 5.2.1 A, and several augmentation techniques, such as blurring of the original image is shown in Fig 5.2.1 B, the horizontal flip augmentation in Fig 5.2.1 C, the addition of gaussian noise to the image in Fig 5.2.1 D, the resized image in Fig 5.2.2 A, the rotation of the image by 120 degrees in Fig 5.2.2 B, the vertically flipped image in Fig 5.2.2 C, and the cropped original image in Fig 5.2.2 D. For the project, it was decided to include three different kinds of noise: gaussian, speckle, and salt & pepper. Additionally, data augmentation methods including rescaling, rotation, translation, shearing, horizontal and vertical flip, etc were employed.

5.3 Statistical Approach

In this approach the relationship between various picture elements based on which the features of that particular image are found is deduced. These features are used to train models that will help us in the task of prediction. It is known as the statistical approach as it follows a rule-based approach to find the relation between pixels and uses mathematical formulae to analyze them.

5.3.1 GrayScale conversion

The first step in the statistical method is to convert an image from the RGB scale to a grayscale image. The values of the pixels in a grayscale image 0 to 255, where 0 indicates black dots and 255 represents white dots. The colors denote the intensity of the brightness in the original image. This is done so as to reduce the complexity of computation.

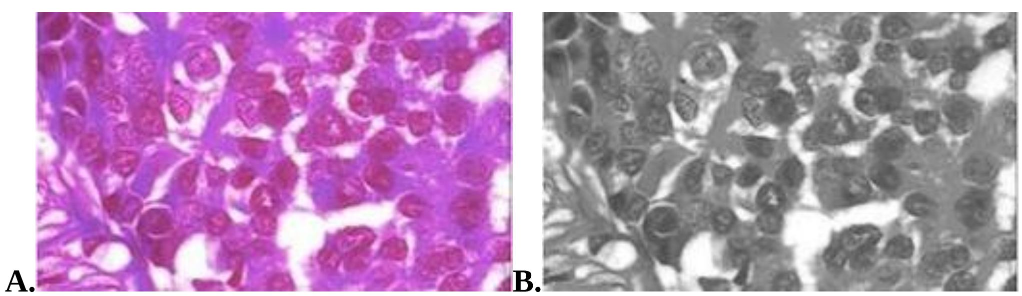


Fig 5.3.1.1 *Original image and its grayscale image*

Fig 5.3.1.1 A. shows an original image in the RGB channels and Fig 5.3.1.1 B shows an image converted to grayscale.

5.3.2 Binarization

The next step done is to binarize the image. This step is crucial to remove the presence of any background noise and gives us the foreground. The foreground is the region of interest. Binarization is the process where each pixel value is converted to either 1 depicting white or 0 depicting black. The region of interest or the foreground is usually converted into white.

A. B.

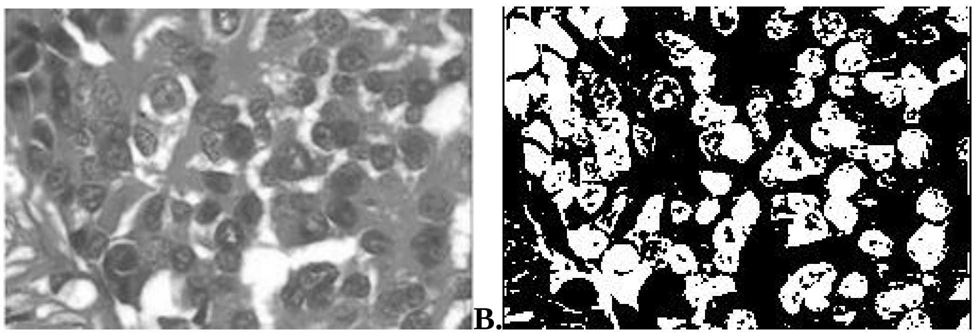


Fig 5.3.2.1 *Image before and after binarization.*

Fig 5.3.2.1 A shows the grayscale image and Fig 5.3.2.1 B shows the image when it undergoes binarization.

5.3.3 Median Filter

Median filter is a noise removal filter which removes unwanted noise from an image. It removes data points that do not contribute much to the prediction of the class. These points only add on to the complexity. In the median filter a particular threshold value is fixed. If a particular region falls below this threshold, it is treated as noise and is removed. The rest of the data is kept.

A. B.

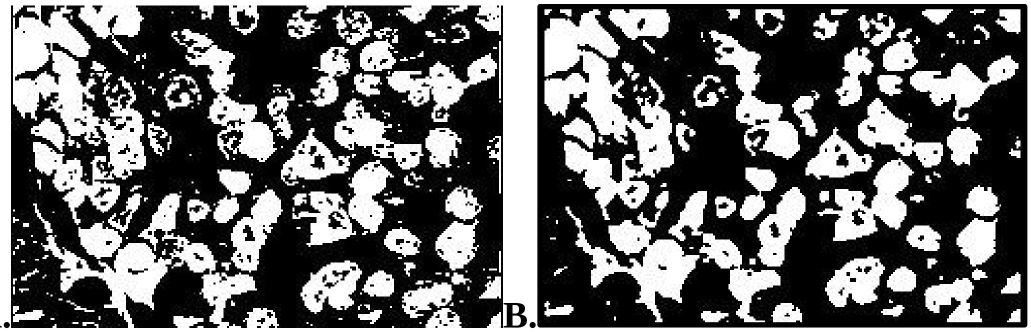


Fig 5.3.3.1 *Images before and after binarization*

Fig 5.3.3.1 A shows an image with noise and Fig 5.3.3.1 B is the same image after the removal of noise using the median filter.

5.3.4 Feature Extraction

In this step, features are extracted from the region of interest. Built-in Python methods called regionprops are utilized to get the features. The skimage package includes this built-in function. A region properties table, which extracts information like area, center of mass, orientation, bounding box, etc., is used to carry out the activity. These features can be found in every region of the image. Each region's properties are referred to as a connected component. The image is made up of several regions, thus for each feature the mean of the connected components is determined. This serves as the dataset that all models are trained on.

The features included are listed below.

- Solidity: percentage of the convex hull's pixels that are also present.
- extent: Pixels in the territory as a percentage of all the pixels in the bounding box.
- Eccentricity: determined by dividing the ellipse's major axis total

- length by the distance in between the focuses.
- Convex Area: ConvexImage's pixel count.
- Area: Count of actual pixels in the area.
- Orientation: the angle formed between the region's second moments and the region's main axes of the elliptical.
- Bbox: acronym for "BoundingBox." Position and dimensions of the region's smallest box.

5.3.5 Feature Selection

Feature selection is performed so as to use features that contribute largely to the prediction of the classes of the dataset. features that are redundant and un-insightful. Feature selection aids in escaping the dimensionality or overfitting curse. For the project forward feature selection is employed.

5.3.6 Normalization

Sometimes, few features play a more significant role in classification due to their highly varying values. This causes a bias in the prediction. to ensure that each feature is given equal importance, normalization is performed. Normalization brings each feature to the same scale. There are many normalization methods that can be employed such as min max normalization, z- scale normalization, decimal normalization etc. In the project decimal normalization technique was implemented to the dataset.

5.3.7 Machine learning model

Various machine learning models were trained on the dataset to categorize the images into either slow paced and fast paced cancer.

- **SVM**

  A machine learning approach that requires data labels may be used for both categorization and regression works. SVM plots all the training points on a graph and finds hyperplanes in the graph that can classify these data points into classes. The margin is the separation between the hyperplane and the closest data point in either collection. The hyperplane chosen has the largest feasible buffer between itself and the nearest point in the training set. The new datapoint is categorized in terms of this.

- **Decision Tree**

  A supervised learning approach called a decision tree is employed for classification. Entropy is a concept that decision trees employ to categorize elements into different groups. Depending on the information gain that was obtained from the entropy, it divides a certain group of objects in the dataset. The objects in this group are shown as nodes in a tree. A decision tree's leaf nodes contain data points that belong to a single class.

- **Random Forest**

  Machine learning with oversight methods Random forests, for example, are commonly in use in designation and regression situations. As the name implies, it combines multiple decision trees. It builds decision trees on innumerable samples and utilizes their overwhelming vote for identification and the norm of the findings for regression.

5.4 Deep Learning based Approach

In the deep learning approach, various hidden layer-based models are employed to classify the specimens into benign and malignant. The whole process of the feature extraction and training is done in the various layers of each model. It acts as a black box where the input is fed and the output arrives. For the project, models that are pre-trained on publicly available datasets were used. These models are much faster compared to a model that is executed from scratch.

5.4.1 Deep learning models

- **AlexNet**

  AlexNet is an eight-layer convolutional neural network. consists of five convolution layers, which each uses Relu activation with the deviation of the outputting layer, and uses a mix of pooling the maximum value layer, three fully densely meshed layers, and these layers. The learning process was enhanced using relu's activation function. dropout layers, which prevented the model from fitting too well. The images that make up this model's input are 227 x 227 x 3 pixels.

During development, the characteristics of the filtration systems that comprise a convolutional layer must be learned. After converging with the image, each filter generates an activation map.

The phrase "max pooling" refers to a grouping procedure that extracts the biggest unit from the convolutional feature area defined by the lens. It reduces the image's size.

- **ResNet**

Residual network is referred to as ResNet. It is a particular AlexNet architectural variant. It resolves the vanishing gradient problem, which is the primary issue brought on by increasing the number of hidden layers. The problem with vanishing gradients is that as the number of layers rises, the gradient zeroes out. Skip connections, which connect the activations of a layer to subsequent layers by bypassing certain levels in between, are a feature of ResNet that address this issue. This network accepts a 224 x 224 x 3 input and employs a 34-layer simple network.

- **DenseNet**

The Dense Convolutional Network architecture strives to improve the degree of networks using deep learning while simultaneously boosting training efficiency by exploiting smaller linkages between stages. Each level of the Compact convolutional neural network is tied to all of the deeper layers.

This is done to optimize knowledge transfer between network stacks.

Contrary to Resnets, it concatenates the features rather than combining them by sum.

DenseNet provides two critical building blocks in addition to the core convolutional and merging layers. They are termed as Bridging Layers and Dense Blocks.

The first convolution block contains 64 filters with a stride of 2 and a 64x7 filter size. A MaxPooling layer with 3x3 maximum pooling and a stride of 2 follows it.

- **MobileNet**

Depthwise Separable convolution is used in the considerably quicker and smaller CNN architecture known as MobileNet. These models are often deemed to be extremely useful for implementation on telephonic and dedicated devices due to their modest size.

MobileNet only requires 16–18 MB of disc space, but a VGG16 model can need up to 500 MB. Because of this, it is perfect for loading on mobile devices.

Therefore, they can be executed on mobile / embedded devices without powerful graphical processing units by reducing the number of computations and making them less power-hungry through the layers developed in the MobileNet architecture.

The key distinction here between MobileNet architecture and a traditional CNN is that the earlier employs a batches normative and Rectified linear function after each 3x3 convolutions, whereas MobileNet distributes the convolution into a 1x1 element - wise convolution and a 3x3 depth - wise separable convolution.

- **VGG-16**

VGG16 is a top object identification and segmentation method. It is a popular approach for picture classification that is straightforward to use with transfer learning. VGG16 has sixteen loaded layers, thirteen convolutions, five Max - pooling, three High - density layers, and a total of 21 layers, albeit only sixteen load layers are present. The main distinguishing characteristic of VGG16 is that it prioritized convolutions of a 3x3 filter with step 1 over a great number of hyper-parameters and used the same pad and maxpool layer of a 2x2 filter with step 2 on a regular basis. Because 3 x 3 convolutions were consistently used throughout the network, it was highly straightforward, elegant, and simple to utilize. In the field of computer vision (CV), significant effort has been made over many decades to enhance this skill. One of the important innovations that paved the path for a number of subsequent improvements in this area is VGG16.

5.5 Performance evaluation

To obtain the classification effectiveness of the models, accuracy, precision, recall, and F1-Score are evaluated and examined.

CHAPTER - 6

SYSTEM TESTING

In this chapter, the few snapshots of the code implementation and the testing samples conducted on a few images are displayed.

CODE IMPLEMENTATION SNIPPETS

```
# Pre trained ResNet-50
base_model = ResNet50(weights='imagenet')
model = Model(inputs=base_model.input, outputs=base_model.get_layer('avg_pool').output)

image_size = 224
img_paths = df.image.tolist()
features_array = np.zeros((4117,2048))

for i, img_path in enumerate(img_paths):
    img = image.load_img(img_path, target_size=(image_size, image_size))
    x = image.img_to_array(img)
    x = np.expand_dims(x, axis=0)
    x = preprocess_input(x)
    features = model.predict(x)
    features = features.reshape(2048,)
    features_array[i,:] = features
```

Fig 6.1 *The pretrained ResNet-50 model implementation*

```
import time
start = time.time()

from sklearn.svm import SVC
clf = SVC()
clf.fit(Xtrain, ytrain)
preds = clf.predict(Xtest)

print('Running time: %.4f seconds' % (time.time()-start))
```

Fig 6.2 *Classification by a SVM classifier implementation*

Fig 6.1 is a code implementation of the pre-trained ResNet - 50 model which was used to extract features from the image. The features are extracted from the last average pooling layer of the ResNet-50 architecture.

Fig 6.2 is the code implementation screenshot of the SVM model used to classify the features obtained from the pretrained ResNet-50 model.

Fig 6.3 *User Interface*

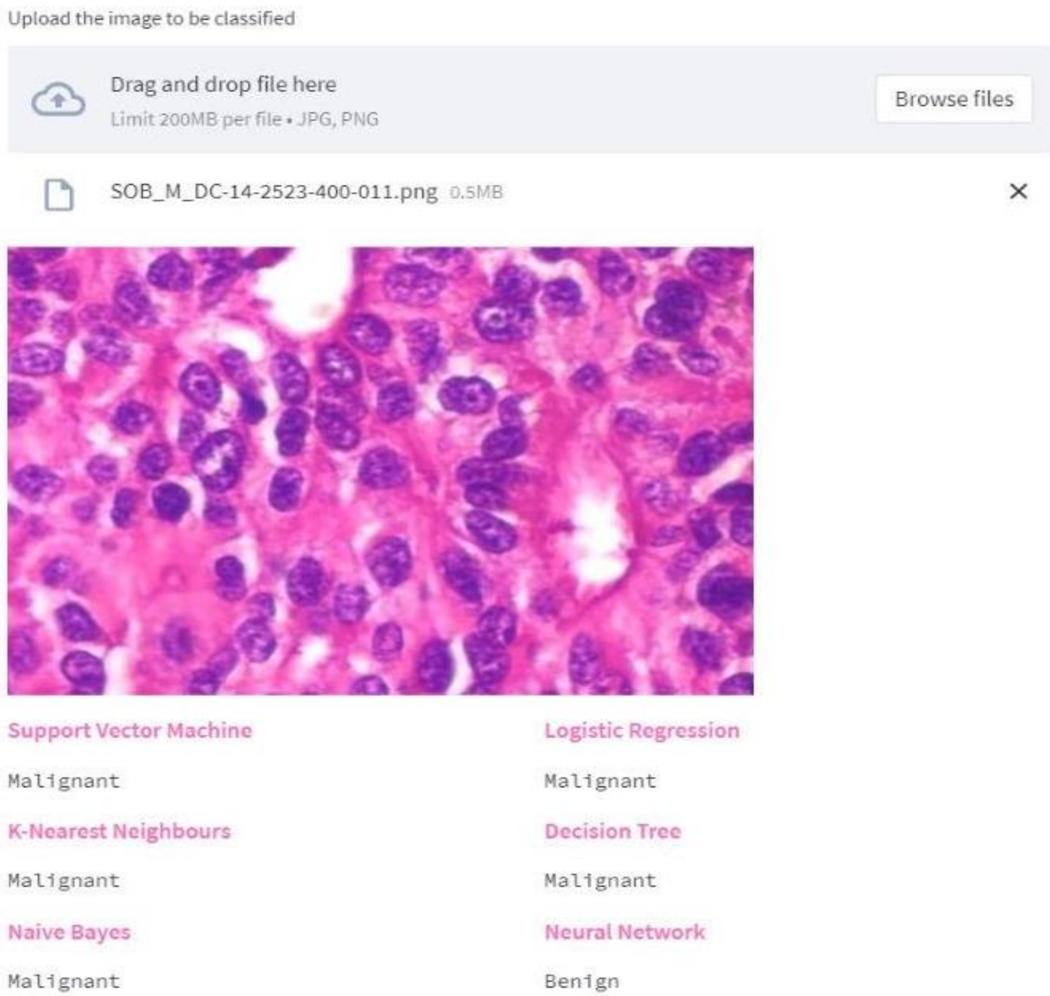


Fig 6.4 *The UI correctly classifies the malignant image*

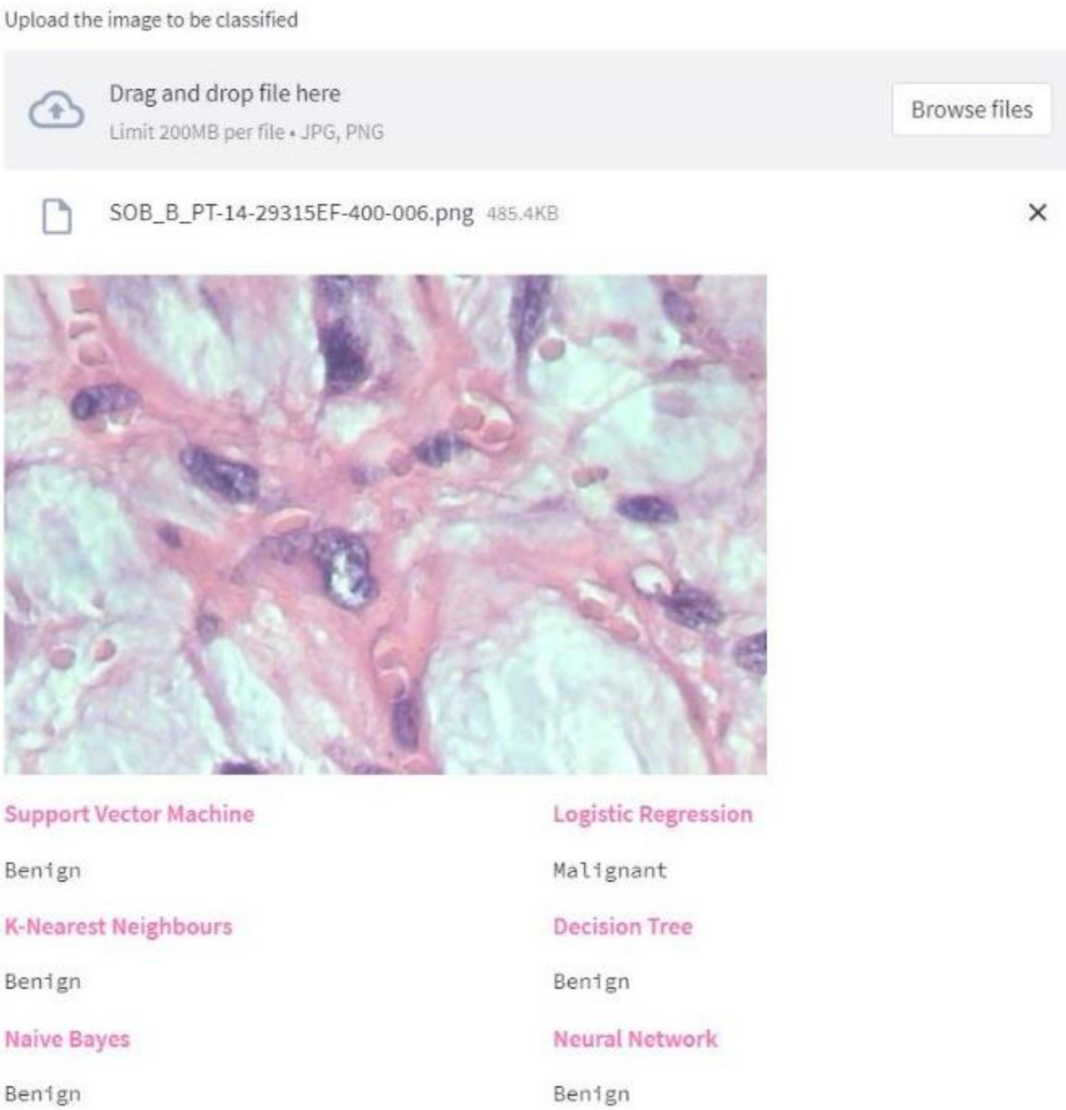


Fig 6.5 *The UI correctly classifies the benign image*

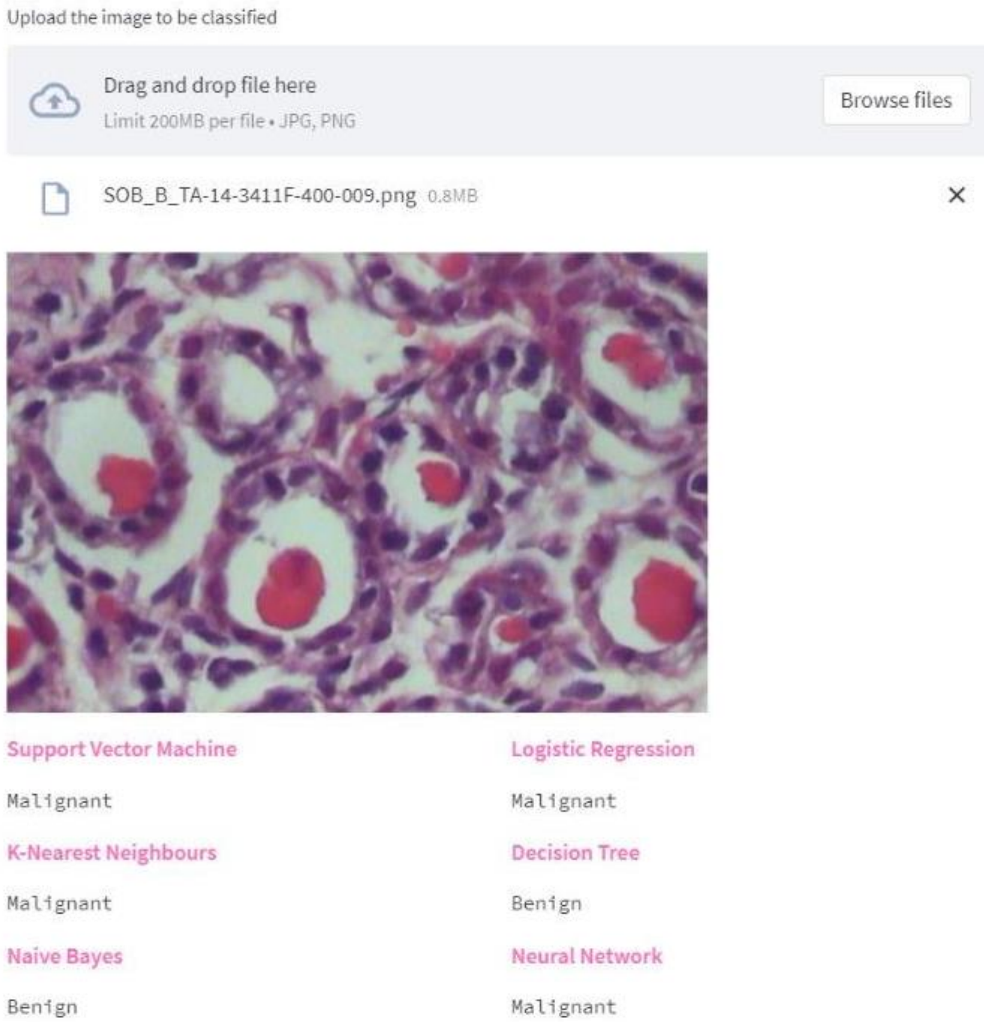


Fig 6.6 *The UI incorrectly classifies the benign image*

.

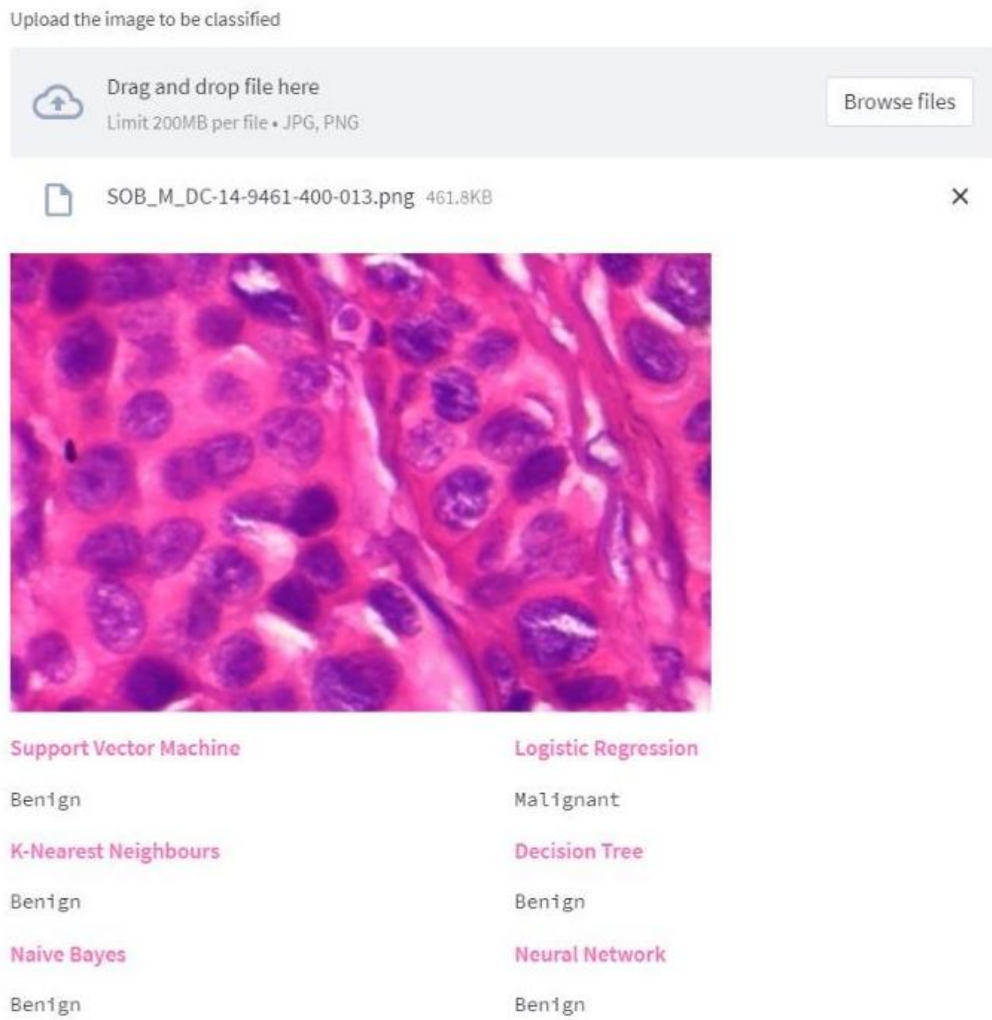


Fig 6.7 *The UI incorrectly classifies the malignant image*

Fig 6.3 is an UI interface which enables a user to upload an image into the database. It feeds the image as an input to various classifiers which assign the image a label. The UI then using a voting mechanism selects the class which has the greatest number of predictions and declares it as the final class. The project testing has a few good predictions as shown in Fig 6.4 where a malignant image is correctly classified into the malignant class

And Fig 6.5 where a benign image is correctly classified into the benign class. There are a few incorrect or erroneous predictions. Fig 6.6 shows a sample case where a benign image is incorrectly classified into the malignant class and in Fig 6.7, the malignant image is incorrectly classified into the benign class.

CHAPTER – 7
RESULTS AND ANALYSIS

In this chapter, the results and observations after implementing the statistical and the deep learning approach are detailed. Learning accuracy, testing accuracy, preciseness, recall, and F1-Score are few metrics used to analyze the results.

Below are the graphs for the few of the best performing models for the dataset.

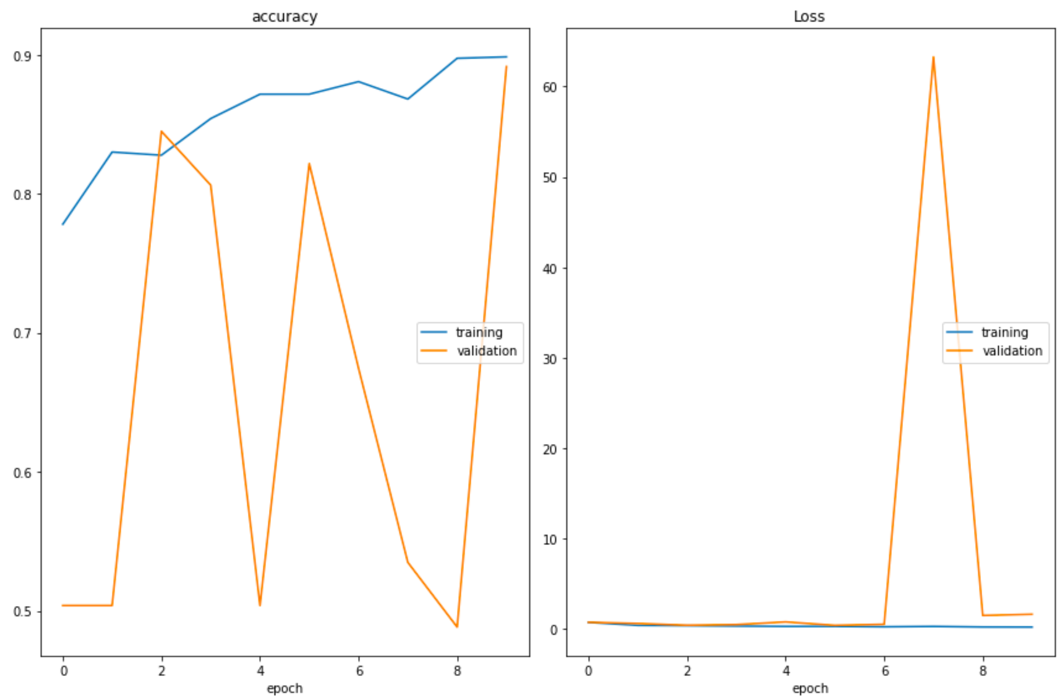


Fig 7.1 *ResNet-50 liveloss plot.*

Fig 7.1 is a plot that is obtained by running the pretrained ResNet-50 model. It shows the varying loss and accuracy for the model in each epoch of the execution. The liveloss plot for ResNet-50 model shows that at the end of the 10th epoch, the training accuracy obtained is 89% and is close to the validation accuracy.

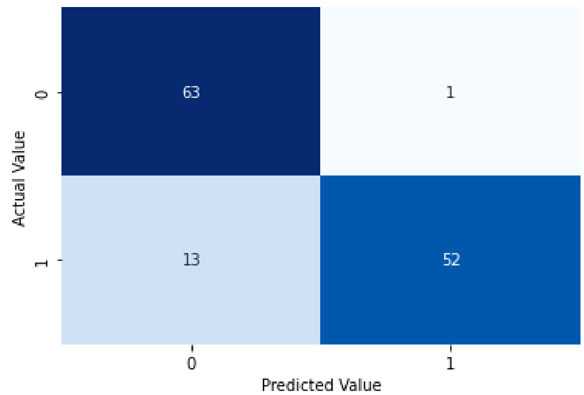


Fig 7.2 *ResNet-50 confusion matrix.*

Fig 7.2 is the confusion matrix obtained by the ResNet-50 model for the classification of histopathological images into benign and malignant. It shows that 115 images were classified correctly and 14 images were classified incorrectly. It shows that there has only been 1 false positive prediction and also relatively less false negative predictions. This initiated that the working of the model was efficient.

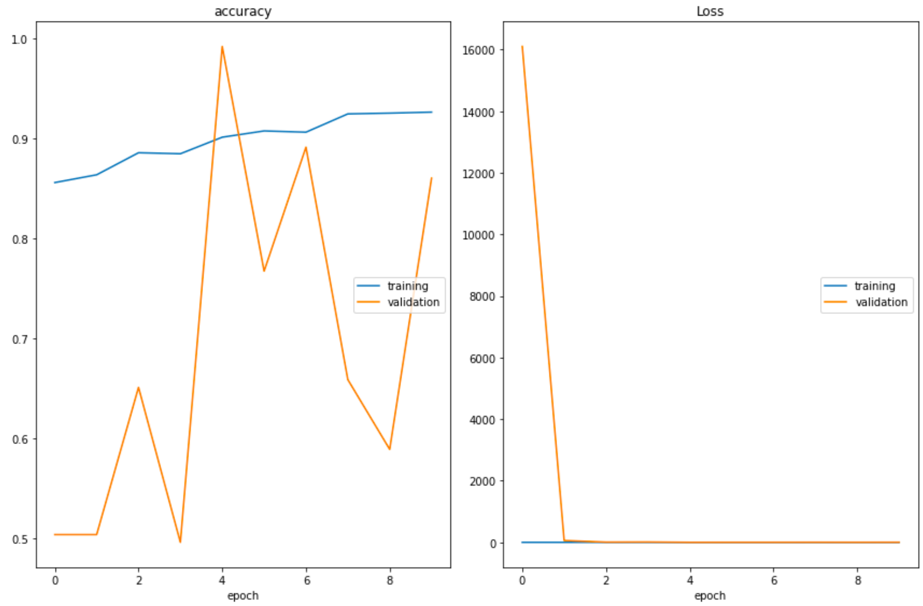


Fig 7.3 *DenseNet liveloss plot.*

Fig 7.3 is a plot that is obtained by running the pretrained DenseNet model. It shows the varying loss and accuracy for the model in each epoch of the execution. The plot indicates that the training accuracy for the model is fluctuating and at the end of the 10 epochs, it arrives at a value of 85% accuracy.

```
Text(33.0, 0.5, 'Actual Value')
```

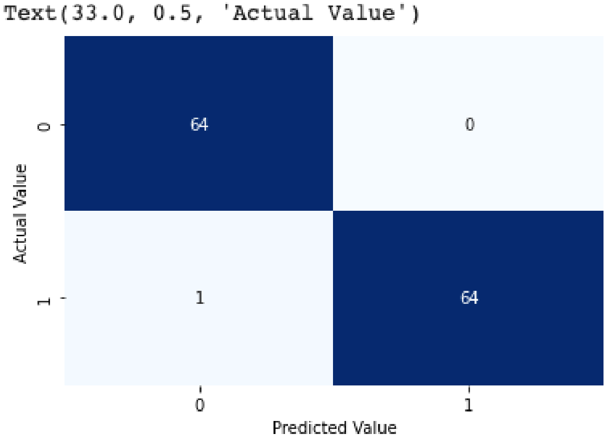


Fig 7.4 *DenseNet Confusion Matrix..*

Fig 7.4 is the confusion matrix obtained by the DenseNet model for the classification of histopathological images into benign and malignant. It shows that 128 images were classified correctly and 1 image was classified incorrectly. The matrix details the almost perfect prediction of the data points by the DenseNet model. It shows that there was only one negative prediction.

Statistical approach

Table 7.1 *Statistical approach results*

CATEGORIZE R	LEARNING ACCURACY	TESTING ACCURACY	PRECISE NESS	SENSIT IVITY	F - MEASURE
SVM	54.26%	53%	54%	53%	50%
Decision Tree	59.68%	62.79%	63%	63%	63%
Random Forest	76.3%	61.2%	63%	61%	60%

The values of learning accuracy, testing accuracy, preciseness, sensitivity and F1-Measure obtained by implementing SVM, Decision Tree and Random Forest classifiers are detailed in Table 7.1

Deep Learning Based Approach

- USING BASIC MODELS

Table 7.2 *Deep learning-based approach results using basic models*

CATEGORIZER	LEARNING ACCURACY	TESTING ACCURACY	PRECISENESS	SENSITIVITY	F - MEASURE
AlexNet	81.1%	91.47%	93%	92%	91%
ResNet-50	89.84%	89.15%	91%	89%	89.98%
DenseNet	92.65%	99.22%	99%	99%	99%
VGG-16	73%	70%	72%	67	69.4%
MobileNet	85.68%	70.54%	81%	71%	75.67%

The values of learning accuracy, testing accuracy, preciseness, sensitivity and F1-Measure obtained by implementing AlexNet, ResNet-50, DenseNet, VGG-16 and MobileNet models using the basic models are detailed in Table 7.2

- USING PRE- TRAINED MODELS

Table 7.3 *Layer traversal learning based approach results using pre trained models*

CATEGORIZER	LEARNING ACCURACY	TESTING ACCURACY	PRECISE NESS	SENSITI VITY	F - MEASURE
AlexNet	83.2%	80%	95%	95%	95%
ResNet-50	73.8%	79.8%	86%	80%	82.89%
DenseNet	87.6%	88.9%	88%	88%	88%
VGG-16	78%	75.9%	80%	76%	78%
MobileNet	72.75%	67%	78%	71%	69%

The values of learning accuracy, testing accuracy, preciseness, sensitivity and F1-Measure obtained by implementing AlexNet, ResNet-50, DenseNet, VGG-16 and MobileNet models using the pre-trained models are detailed in Table 7.3

- USING DEEP-LEARNING MODELS WITH SVM

Table 7.4 *Deep learning-based approach results using pre trained models using SVM*

CATEGORIZER	LEARNING ACCURACY	TESTING ACCURACY	PRECISENESS	SENSITIVITY	F - MEASURE
AlexNet	83.5%	82.97%	80%	78%	78.9%
ResNet-50	88.6%	86%	86%	84%	85%
DenseNet	82%	80%	79%	78%	78%
VGG-16	87.7%	85.3%	85%	84%	84%
MobileNet	75.4%	73.5%	77%	70%	73.3%

The values of learning accuracy, testing accuracy, preciseness, sensitivity and F1-Measure obtained by implementing AlexNet, ResNet-50, DenseNet, VGG-16 and MobileNet models using the pre-trained models to extract the features and SVM classifier to classify them are detailed in Table 7.4

After utilizing machine learning techniques to complete the classification, it was found as shown in table 7.1 that the model efficiency is not satisfactory. SVM provides an accuracy of 54.26%, decision tree algorithm provides a 59.68% accuracy, and the random forest approach provides a 76.3% accuracy. According to the results, the Random Forest classifier would be the best option if a machine learning classifier were to be used to categorize the photographs.

However, in the field of medical image processing, there cannot be inaccurate predictions; as a result, deep learning-based classifiers to increase the prediction rate of the model are applied. These deep learning models have been used in three different ways.

- Basic models

  Basic models are deep learning algorithms that are created from the ground up and trained solely on the dataset. It was discovered as shown in table 7.2 that DenseNet had the greatest performance using fundamental learning models, with an accuracy of approximately 99%. However, this strategy takes a lot of time, making it challenging to use in day-to-day operations. Furthermore, a lot of computing power was needed.

- Pre-trained models

  pre-trained models were employed to solve the aforementioned issue. Pre-trained models are those that have already undergone training using a publicly accessible dataset, most often the ImageNet corpus. This helps the model run more quickly and effectively. Later, the models were trained using the data. As a result, it was discovered as shown in table 7.3 that DenseNet had the highest performance but was less accurate than its fundamental model. The benefit is that the time required has greatly improved, though.

- Pre-trained models with SVM

  In this instance, the feature extraction is left up to the deep learning classifiers, which subsequently pass the input to an SVM classifier for classification. It is shown in table 7.4 that the best performance came from compiling ResNet-50 in this way.

CHAPTER – 8
CONCLUSION AND FUTURE SCOPE

As a result of the research and analysis, it has been concluded that random forest outperformed other machine learning models. If the choice was to be made to decide which deep learning models to use, DenseNet would be a superior choice. A ResNet-50 model with an SVM classifier would be used if the methods are to be mixed.

The current effort examines the classification of benign and malignant cancer in histopathological images. Examining cancerous tissues can be done in a variety of ways. Mammograms, computerized tomographic pictures, and magnetic device images are all variations of x-ray images to identify cancer. In the future, the hope is to broaden the research to conduct categorization on all sorts of images. Classifying photos into the numerous subgroups of cancer cells present would be an additional enhancement. This would be a more thorough investigation into the type and occurrence of cancer in women.

REFERENCES

[1] Shahidi, F. *et al.* (2020) "Breast Cancer Classification using Deep Learning Approaches and histopathology image: A comparison study," *IEEE Access*, 8, pp. 187531–187552. Available at: https://doi.org/10.1109/access.2020.3029881.

[2] Aljuaid, H. *et al.* (2022) "Computer-aided diagnosis for breast cancer classification using Deep Neural Networks and transfer learning," *Computer Methods and Programs in Biomedicine*, 223, p. 106951. Available at: https://doi.org/10.1016/j.cmpb.2022.106951.

[3] Huang, J. *et al.* (2022) "BM-net: CNN-based MobileNet-V3 and bilinear structure for breast cancer detection in whole slide images," *Bioengineering*, 9(6), p. 261. Available at: https://doi.org/10.3390/bioengineering9060261.

[4] Shakeel, S. and Raja, G. (2021) "Classification of breast cancer from mammogram images using Deep Convolution Neural Networks," *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)* [Preprint]. Available at: https://doi.org/10.1109/ibcast51254.2021.9393191.

[5] Yilmaz, F., Kose, O. and Demir, A. (2019) "Comparison of two different deep learning architectures on breast cancer," *2019 Medical Technologies Congress (TIPTEKNO)* [Preprint]. Available at: https://doi.org/10.1109/tiptekno47231.2019.8972042.

[6] Nawaz, M., A., A. and Hassan, T. (2018) "Multi-class breast cancer classification using deep learning Convolutional Neural Network," *International Jthenal of Advanced Computer Science and Applications*, 9(6). Available at: https://doi.org/10.14569/ijacsa.2018.090645.

[7] Ribli, D. *et al.* (2018) "Detecting and classifying lesions in mammograms with deep learning," *Scientific Reports*, 8(1). Available at:

https://doi.org/10.1038/s41598-018-22437-z.

[8]Zhang, X. *et al.* (2017) "Whole mammogram image classification with Convolutional Neural Networks," *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* [Preprint]. Available at: https://doi.org/10.1109/bibm.2017.8217738.