Breast Cancer Detection using Deep Learning Techniques

MSc Research Project
Data Analytics

Melwin Francis Rodrigues
Student ID: X18183000

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	MELWIN FRANCIS RODRIGUES		
Student ID:	X18183000		
Programme:	Data Analytics	**Year:**	2020
Module:	MSc Research Project		
Supervisor:	Dr. Catherine Mulwa		
Submission Due Date:	17/08/2020		
Project Title:	Breast Cancer Detection Using Deep Learning Techniques		
Page Count:	20		
Word Count:	5418		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: MELWIN FRANCIS RODRIGUES

Date: 17/08/2020

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Breast Cancer Detection using Deep Learning Techniques

Melwin Francis Rodrigues
X18183000

Abstract

Breast cancer is a major health problem in women especially for the women who are above the age of 50. Early detection can help in saving many lives. Cancer begins when cells grow out of control to form a mass called a tumour. A tumour can be cancerous or benign. The most common symptoms of breast cancer are formation of lump. The detection of cancerous cells can prevent the loss of lives and help the women to take corrective actions before the condition gets worse. This was the motivation for this research in this research the data is divided into two classes i.e malignant and benign. The data is again sub divided into magnifications from 40 x, 100x, 200x, 400x. The dataset was pre-processed by using zoom, horizontal flip and rescale. Classification methods such as Dense Net 121, Inception V3 and CNN was used. The best classification results were given by Dense Net 121 with 95% accuracy. The accuracy and validation accuracy obtained was the better than most of the previous research results. This research will benefit the women and the doctors to detect the cancer cells in the preliminary stage.

1 Introduction

Breast cancer is most common form of cancer among women in Ireland. Most of the breast cancer is above 50 years of age for women. It is very difficult for the women and their families to go through such a difficult time. It is curable cancer if detected on earlier stages. Women should check their breast on regular intervals of time and get examined by the doctors. The most common symptoms are breast lumps. The change in breast size and shape is also another symptom. Age plays an important factor in breast cancer. As the condition is more common in women who are above the age of 50. Women who use contraceptive pill have high tendency to develop breast cancer. Lifestyle factors also play an important role. Alcohol, obesity etc. play a vital role in breast cancer development. Breast cancer can be prevented by having a healthy diet, exercising, maintaining a balanced weight and intake of low amount of alcohol and fats. Psychological problems like depression, anxiety also have negative impact (Dona, et al., 2018). Breast cancer screening uses two tissues to determine whether there is a breast cancer or no. The two tissues are benign and malignant. Histopathological images are very detailed and very useful in providing useful insights regarding breast tumor (Sara, et al., 2019). The detection through histopathological images can help in prior detecting the breast cancer and by using the machine learning techniques such as CNN and transfer learning algorithms the accuracy of detection can be enhanced (S, et al., 2019). The architecture proposed is a comparative analysis between Dense Net 121, CNN and inception V3 to classify the breast cancer data. It will increase the efficiency as preprocessing is done on it by using flipping, zooming and transforming the image dataset into proper readable form. Which will take less amount of time to process the data. This model would be very useful and efficient for the radiologist on histopathological images. This method will help in detecting the breast cancer in less time with better accuracy.

1.1 Research Question

Detection of breast cancer can help in saving many lives. Early detection can help the doctors and patients to identify the problem and can help the patients to correctly identify the cancerous cells. Early detection can help patients to change their routine and lifestyle which will help in reducing the cancerous cells.

RQ: *"To what extent can multiple machine learning algorithms be used to classify breast cancer in histopathological images?"*

1.2 Objectives and Contributions

Objectives 1:

Obj1.1 Involves investigation and review of literature of breast cancer research projects from 2011-2020.

Obj1.2 Merging the dataset programmatically and creating a proper readable dataset for the machine learning models.

Obj1.3 Creating Meta Data from the Datasets.

Objective 2: Implementation, Evaluation and Results of classification models.

Obj2.1 Implementation, Evaluation and Results of DenseNet121

Obj2.2 Implementation, Evaluation and Results of Inception V3

Obj2.3 Implementation, Evaluation and Results of CNN

2 Related Work

2.1 Convolutional Neural Network used for classification

In (P, et al., 2019) the dataset was taken from BreakHis which had 7909 images out of which 2440 were benign and 5429 were malignant. The method used was CNN with accuracy of 73.68%. In future by optimizing and using preprocessing methods the accuracy could be increased. (B, et al., 2019) used dataset from Kaggle which was publicly available it had 277,524 patches. Convolutional Neural Network was used Adam optimization. The AUC value obtained was 0.935 which is very good when it comes to 27753 images.

2.2 Approaches used from Transfer Learning

This (Sara, et al., 2019) paper used was from breakhis which consisted of 7909 images. The dataset is divided into benign and malignant. There were 4 magnification from 40, 100, 200 and 400. The other dataset used is from BACH challenge. VGG19, mobileNet, DenseNet were used for ensembled network. The accuracy was more than the traditional methods.
(N, et al., 2019) applied image preprocessing, classification and evaluation to the image dataset which was taken from IRMA database Germany. In Image preprocessing image resize and image conversion were performed. The methods used are VGG16 and ResNet50 of which VGG16 gave accuracy of 94%. In (A, et al., 2019) the dataset was taken from Kaggle competition of breakhis. Which consisted of 7909 images of 82 patients. The method used

was AlexNet which was fine tuned for better performance. The accuracy obtained was between 93.8 to 95.7% for classification.

2.3 Most frequently used Methods SVM, Decision trees, Random forest, Naïve Bayes, KNN

In (Dona, et al., 2018) survey was conducted which did comparison of multiple classification and clustering algorithms. The output proved that classification algorithms were better than clustering algorithms. In multiple classification used SVM and C5.0 performed better than other classification algorithms. (Rashmi, et al., 2015) used Naïve Bayes classification and prediction algorithm. The dataset used were from Wisconsin University. The attributes of the database are as follows: Clump Thickness, Cell Size, Cell shape, Class (benign, malignant) etc. The attributes values were between 1 to 10 which are the layers of penetration. There were 444 benign and 239 malignant instances. Both the naïve based classification and prediction algorithms showed similar success rate. In (H, et al., 2017) the dataset was taken from 82 breast cancer patients from sri kuppuswamy Naidu hospital Coimbatore. The method used is Bayesian linear discriminant analysis (BLDA). Basically, BLDA is extension of fishers LDA. The only difference is regularization is used in BLDA to avoid overfitting. The accuracy obtained was 91.66%. In (M, et al., 2018) the dataset used is donated to University of California. There are total of 11 attributes in dataset. The methods used were Naïve bayes and KNN of which KNN gave slightly better result 97.51% than Naïve based which was quite nearby 96.19%. If the dataset would have been large than KNN would have taken more computational time than Naïve Bayes. In (A, et al., 2011) the dataset is taken from WBDC which consist of 569 instances. Out of 569 instances 357 were benign and 212 malignant. Dimension reduction by ICA and SVM are used. The performance was measured by confusion matrix. The results showed SVM with quadratic kernel yield highest accuracy of 94.41%. (S, et al., 2018) used Random forest, KNN and Naïve Bayes to the dataset which was taken from Wisconsin Diagnosis. Comparison between the algorithms were done. Evaluation metrics used were Accuracy, Recall, Precision and F1 score. The results showed that KNN performed better and was the most effective while comparing the other algorithms.

In (S, et al., 2017) the dataset was taken from Wisconsin Breast cancer UCI repository. The main objective is to classify cancer is benign or malignant. The methods used were linear regression, Decision tree and Random forest for prediction of the type of treatment. The accuracy obtained was 84.14% for linear regression, 88.14% for random forest. The accuracy could have been increased by preprocessing. (B & Akbugday, 2019) used the dataset from UCI machine learning repository. The methods used are Naïve bayes, KNN and SVM for classification. The accuracy of SVM was better than the rest of the methods. The entire methods were implemented on WEKA in future python or R can be used for more accurate results.

In (E, et al., 2019) the dataset was taken from Wisconsin Breast cancer which consisted of 699 instances. The methods used were Artificial Neural Network and SVM. The results showed that SVM performed better than ANN with 96.99% accuracy. The entire process was implemented from WEKA tool. In (P, et al., 2019) the dataset was taken from Wisconsin database. Total number of images were 698 out of which 457 were benign and 241 were malignant. Comparison was made between deep learning, Random forest, SVM, Vote and Random Forest. Deep learning proved to be successful with high performance. In (Anon., 2019) the dataset was taken from University of California Irvine. It contained 116 samples. The methods used were decision tree, SVM, muti layer perceptron, K- nearest neighbors, logistic regression, and random forest. The accuracy was the highest for KNN which was 87.5%.

In (M. I. H. Showrov, 2019). The models implemented were Linear SVM, RBF Neural Network out of linear SVM gave better accuracy of 96.72%. (A, et al., 2018) the dataset was taken from Wisconsin diagnostic breast cancer. It had 357 benign and 212 malignant images. KNN, Naïve Bayes, logistic regression and SVM were used out of which KNN gave best results.

2.4 Approaches used from Neural Networks

In (S, et al., 2019) paper the dataset used was from Cancer Genome Atlas it had 27,397 image patches. This experiment showed logistic regression gave better result in AUC. The ensembled based active learning increase the performance to a greater extent and can be imported in HistomicsML framework. In (Jasmir, et al., 2018) the dataset was taken from multiple sources like Medical Center University, Institute of oncology and Ljublijana. The dataset had 286 instances with 10 attributes. The preprocessing was done by labelling data and cleaning and imputation of the missing values. Multilayer Perceptron was used as a classifier. The evaluation was done by 10-fold cross validation. The final output was 96.5% accuracy. In (M, et al., 2018) the dataset used was taken from Wisconsin Diagnostic it contained 699 records. Extreme learning algorithm was used with hidden neurons. This method was compared to other systems and performed better than tradition method.
In (S, et al., 2017) the dataset was taken from Wisconsin Brest cancer. The accuracy of J48 classifier was the best with 75.54% accuracy. In (H, et al., 2016) the dataset was taken from Mathworks. The dataset had malignant and benign images. The classification was done by using back propagation. The accuracy was satisfactory. (S, et al., 2019) the dataset used was from 2018 challenge on breast cancer recognition. Pretrained model such as xception and DCNN was used of which Xception performed better than DCNN with 92.50%.

Table 1: Comparison of all Methods

Authors	Methods Used	Dataset Used	Results
(P, et al., 2019)	CNN	BreakHis	73.68%
(B, et al., 2019)	CNN	Kaggle	AUC 0.935
(Sara, et al., 2019)	VGG19, mobileNet, DenseNet	Breakhis	Better than traditional methods
(N, et al., 2019)	VGG16 and ResNet50	IRMA	94%
(A, et al., 2019)	AlexNet	Breakhis	Between 93.8 to 95.7%
(Dona, et al., 2018)	SVM and C5.0	Wisconsin University	Better than other classification algorithms
(Rashmi, et al., 2015)	Naïve Bayes	Wisconsin University	Better than other classification methods
(H, et al., 2017)	Bayesian linear discriminant analysis	Sri kuppuswamy Naidu hospital Coimbatore	91.66%
(M, et al., 2018)	Naïve Bayes and KNN	University of California	97.51%
(A, et al., 2011)	Dimension reduction by ICA and SVM	WBDC	94.41%
(S, et al., 2018)	Random forest, KNN	Wisconsin Diagnosis	The results showed

	and Naïve Bayes		that KNN performed better
(S, et al., 2017)	Linear Regression, Decision Tree and Random Forest	Wisconsin Breast cancer UCI repository	88.14%
(B & Akbugday, 2019)	Naïve bayes, KNN and SVM	UCI machine learning repository	Accuracy of SVM was better than the rest of the methods
(E, et al., 2019)	Artificial Neural Network and SVM	Wisconsin Breast cancer	96.99%
(P, et al., 2019)	Random forest, SVM, Vote and Random Forest	Wisconsin database	Deep learning proved to be successful with high performance
(Anon., 2019)	Decision tree, SVM, muti layer perceptron, K-nearest neighbors, logistic regression, and random forest	University of California Irvine	87.5%
(M. I. H. Showrov, 2019)	Linear SVM, RBF Neural Network out of linear SVM	Wisconsin dataset	linear SVM gave better accuracy of 96.72%
(A, et al., 2018)	KNN, Naïve Bayes, logistic regression and SVM	Wisconsin dataset	KNN gave best results
(S, et al., 2019)	logistic regression	Cancer Genome Atlas	logistic regression gave better result in AUC
(Jasmir, et al., 2018)	Multilayer Perceptron	Medical Center University, Institute of oncology	96.5%

3 Research Methodology

Introduction:

In this research we will be using KDD as it suits the models which will be applied on the dataset. This research will bring value in the medical field as it will be useful in faster detection of breast cancer in women and save many lives. There will be two-tier structure i.e User tier and business logic tier.

3.1 Breast Cancer Detection Using Knowledge Discovery Methodology

The below Figure1 is a modified KDD consists of the following stages. (1) Data collection of images is done which is in .png format. Total of 7909 images are selected. (2) Data is moved programmatically and combined with malignant and benign in one folder named (Combined). Meta data is created based on the image dataset programmatically. (3) Pre-processing is implemented to remove noise or any distortion in the images. (4) Classification models are applied such as CNN, DenseNet121, Inception V3. (5) Models are evaluated and compared based on accuracy, specificity, sensitivity.

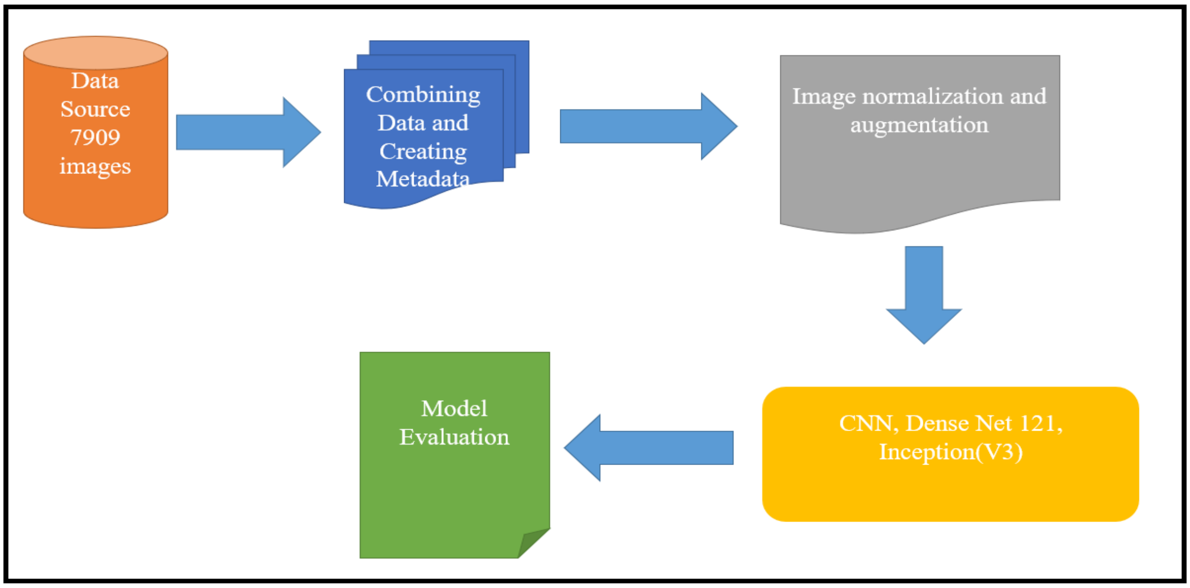

Figure 1: Modified KDD

Project Understanding and Data Gathering

First task of any project is project understanding. To actually understand what the project is about and understand the requirements of the project. The need to do this project. The major task was to understand the dataset and clean it in order to pass it when building the models.
The dataset was taken from kaggle[1] with only two kernerls. The dataset is originally of the breakhis competition. The Figure1 descibes the dataset which has 7909 images of 4 magnifications from 40x, 100x, 200x, 400x. Table 1 gives the depth count and magnifications of the dataset. The dataset was already divided into multiple folders of magnifications and also had sub folders of malign and benign. The 40X_M indicates the magnifiaction of image which is 40X and M stands for malignant type. In the similar way 200X_B indicates 200X magnifaction and Bdenotes the benign type.

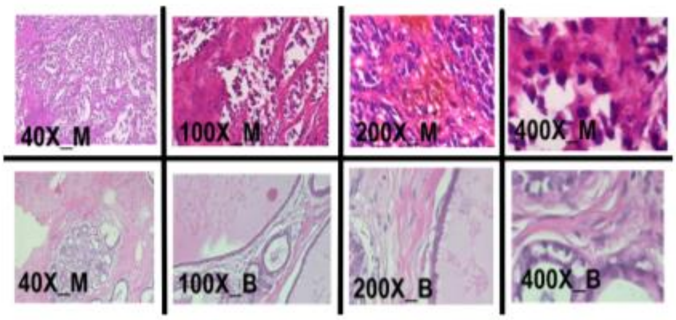

Figure 1: Different Magnifications

[1] https://www.kaggle.com/kritika397/breast-cancer-dataset-from-breakhis

Table 2: Distribution of Images based on Magnifications

Magnification	Benign	Malignant	Total
40X	train: 500 valid: 62 test: 63	train: 1176 valid: 92 test: 97	1990
100X	train: 515 valid: 64 test: 65	train: 1149 valid: 144 test: 144	2081
200X	train: 498 valid: 62 test: 63	train: 1112 valid: 139 test: 139	2013
400X	train: 470 valid: 59 test: 59	train: 985 valid: 123 test: 124	1820
Total	2480	5429	7909

Data Preparation

The data was cleaned after aquaring it from kaggle. As the data needed to be cleaned and into a standard form which machine learning algorithms understand. Data was prepared for modelling and anlysis. The meta data was prepared based on the image dataset programatically. The below Figure3 shows the names and labels columns. The names column has the names of the images from all the dataset and the labels column has the cancerous cell i.e M for Malignant and B for Benign image. The data is split into 60% Train, 20% Test and 20% Validation.

1	Names	Labels
2	SOB_M_DC-14-14015-100-003.png	M
3	SOB_M_DC-14-14926-200-012.png	M
4	SOB_M_DC-14-13993-200-002.png	M
5	SOB_M_DC-14-17901-200-004.png	M
6	SOB_M_DC-14-20636-100-001.png	M
7	SOB_M_PC-14-19440-200-014.png	M
8	SOB_B_F-14-14134-100-006.png	B
9	SOB_B_F-14-14134E-100-011.png	B
10	SOB_M_MC-14-16456-100-056.png	M
11	SOB_B_F-14-23060CD-400-009.png	B

Figure 3: Meta Data

Data Preprocessing

The data had only images in it. It didn't had meta data. Meta data had to be created by using computation logic using python in jupyter note book. The meta data was created with columns named the image name and labels. Every image in the dataset had unique names with codes in their names. The images were defined by the unique number, manification and type of cancer i.e malignant or benign. Meta data was important for the modelling. The images were of different magnifications. Hence rescale of images were done, shear range and zoom range and horizontal flip was implemented to make the images into stadard form which the model would accept. The columns in meta data were of image names and the type of image i.e malign or belign signified by M and B.
The data is divided into Train, Valid and Test folders. In Train Folder there are 4745 images, the valid and test both has 1582 images. The data after rescaling, zoom and horizontal flip was in a standard form but still had to check wheather there was data leakage or no.
Data leakage could have caused the model to give baised output. As the training would have been improper. Separate folders were created on google colab. As it was convenient for model training.
The dataset after splitting were rescaled, zoom range and horizontal flip was implemented in order the image should be standardized with respect to the model which is shown in Figure4. For the preprocessing part image data generator was used. Image data generator is used as it takes the original data and transforms the data into the new transformed data.

```
from keras.preprocessing.image import ImageDataGenerator
train_datagen = ImageDataGenerator(
        rescale=1./255,
        shear_range=0.2,
        zoom_range=0.2,
        horizontal_flip=True)
test_datagen = ImageDataGenerator(rescale=1./255)
```

Figure 4: Image Pre-processing

Modelling

Modelling is used to train the model and predict the values adjusting to business requirements and then validating and testing them. Breast cancer data from breakhis is used by many researchers to predict the data correctly. Many algorthims have been applied on this data set. Various algorithms like Dense Net 121, Inception V3 and CNN can be applied on this data set. Dense Net 121 outperformed in previous researchs hence this version of dense net was used which is trained on Image datasets.

3.2 Design Specification

This project followed Two- Tier Architecture shown in Figure5. It consists of Client Layer and Business layer which is also called data layer. The client can be Doctors or radiologists who will feed the image datasets. Metadata is prepared based on the image datasets and divided into train, test and validation CSV files. Data cleaning and pre-processing is done on the image datasets. Also, data leakage is checked as it will affect in the modelling during classification. The cleaning and pre-processing were done by using python on google colab. After the data is ready it is sent to the models for classification for training and testing. Three classification models where applied CNN, Inception V3 and Dense Net 121. The output obtained was evaluated using accuracy, sensitivity and specificity. The final evaluation output was given to client for implementation.

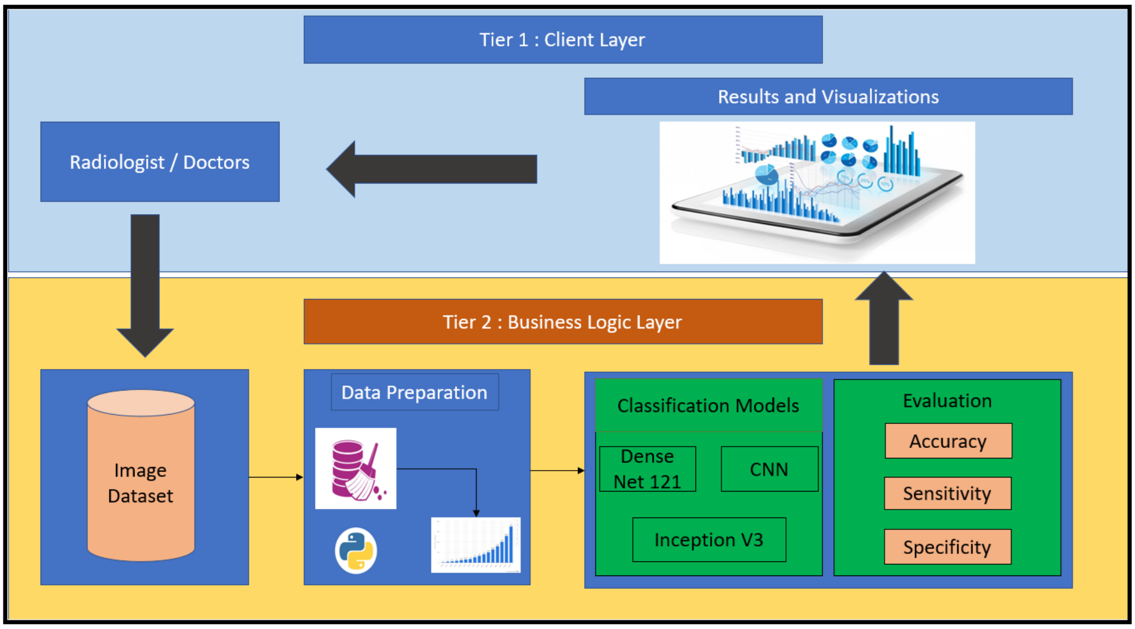

Figure 5: Process Flow

4 Implementation, Evaluation and Results of Classification Models.

Implementing the classification models were an important task. Various techniques were used for classification of breast images. 4.1 DenseNet121, 4.2 Inception V3, 4.3 CNN.
The results are evaluated based on evaluation metrics.

Accuracy is the percentage of correctly classified instances, which is given by the formula below.
Sensitivity is the actual positives which are correctly identified. It was calculated by the below formula.
Specificity is the rate of positive instances were correct. It was calculated by using the formula.

Accuracy = (TP + TN) / (TP + TN + FP + FN)
Recall = TP (TP + FN)
Precision = TP (TP + FP)
where, TP = True Positive, sentiments that are positive and are classified as 1
TN = True Negative, which are negative and are classified as 0
FP = False Positive, which are negative but are classified as 1
FN = False Negative, which are positive but are classified as 0

4.1 Implementation of Dense Net 121

As Dense net was already used on this data set as suggested by the previous research.
Dense Net have many advantages such as vanishing gradient problem is resolved they strengthen feature propagation. Dense Net 121 was used because it was already trained in Image Net and gave better results than Normal Dense Net. The 121 denoted the depth of ImageNet Models. 121 can be computed as: - 5+(6+12+24+16) *2 = 121 where 5 is conv, pooling + 3 transition layers+ classification layer. We multiply by 2 because each dense block has 2 layers. The dense Net is used as each layer receives the additional inputs from the previous layers and then passes its own features to sub sequent layers. Dense Net 121 is used as it is trained on images. In Dense Net 121 we have given 2 classes as there were 2 classes names malign and benign. The top layer was set as False as we wanted our pre trained model to learn from the dataset itself. Average pooling layer is set, and the pool size is (3,3). There are no strides given hence it will default set as pool size. The activation function used is sigmoid. The reason for using sigmoid is it lies between 0 and 1. Since our dataset consist of probability between malign and benign i.e 0 and 1. The below Figure6 shows the code of dense net 121 model.

```
def build_model(base, layer_units, num_classes):
    for layer in base.layers:
        layer.trainable = False

    x = base.output
    x = Flatten()(x)
    for num_units in layer_units:
        x = Dense(num_units, activation='relu')(x)
    predictions = Dense(num_classes, activation='sigmoid')(x)
    model = Model(inputs=base.input, outputs=predictions)
    return model
```

```
from keras.optimizers import Adam, SGD
from keras.layers import Activation, Dense, Flatten
from keras.models import Model, load_model

adam = Adam(lr=0.0001)
model = build_model(dn121, [1024], 1)
model.compile(adam, loss='binary_crossentropy', metrics=['accuracy'])
```

Figure 6: Building of Dense Net 121 model

Dense Net 121 gave 95% accuracy when compared to CNN and inception its validation accuracy was also maximum. The accuracy of Dense Net 121 as the number of epochs is increased the accuracy is also increased. The final accuracy was 95% also shown in Figure7 with 10 epochs.

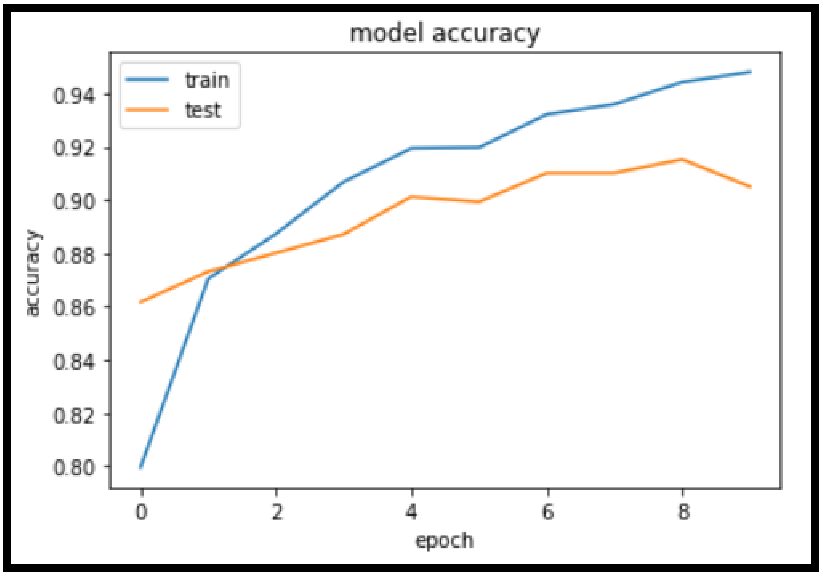

Figure 7: Model accuracy of Dense Net 121

The sensitivity score is 0.68 which is the true positive rate. That means that 68% of benign data is correctly identified whereas Specificity is the True Negative rate which predicts the actual negatives that is malign data which was 58% (0.58).
The Figure8 shows the overall loss and accuracy of train and test dataset.

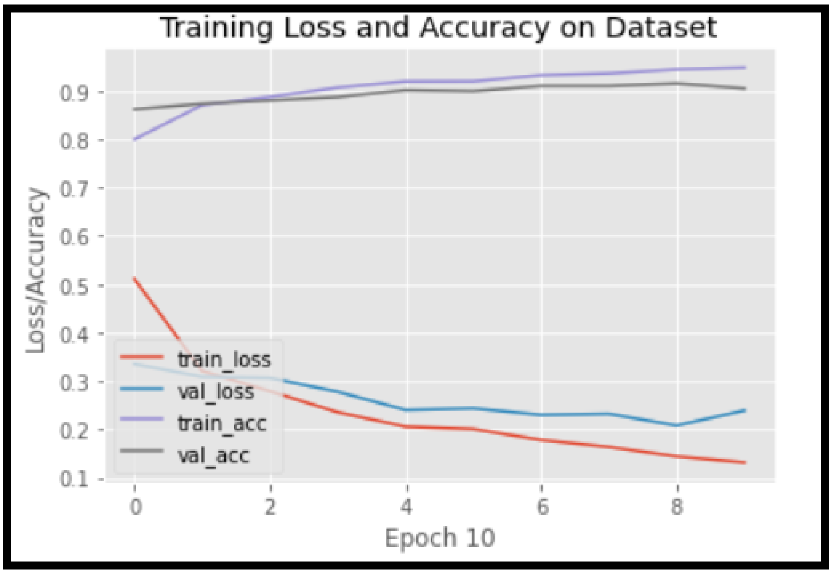

Figure 8: Overall Loss and Accuracy.

4.2 Implementation of InceptionV3

Inception V3 was used because of RMSProp Optimizer. Has factorized 7*7 convolutions, Batch auxiliary Classifiers and has label smoothing. The previous models of inception used attained greater accuracy on image datasets. The model is made of symmetric and asymmetric building blocks, with convolution average pooling and max pooling.
This model was trained over millions of datasets and over 1000 classes.

The below Figure9 is Pre trained layer which is set to false because we wanted to train the model based on our dataset.

```
[3]  for layer in pre_trained_model.layers:
       layer.trainable=False
```

Figure 9: Pre-Trained layer is set to false.

The optimizers used was Adam to update network weights. The loss function used is (binary_crossentropy) because there are 2 classes in the Dataset i.e malign and benign. In the first layer relu is applied because it should consider all the positive values and in the last layer sigmoid is used because of the 2 classes as it was binary i.e 0 and 1 (malign and benign). The Figure10 show the models code which was used in building it.

```
from tensorflow.keras.optimizers import RMSprop
from tensorflow.keras.optimizers import Adam, SGD
from tensorflow.keras import layers
from tensorflow.keras import Model
x = pre_trained_model.output

x = layers.Flatten()(x)
x = layers.Dense(1024,activation = 'relu')(x)
x = layers.Dropout(0.4)(x)
x = layers.Dense(1, activation='sigmoid')(x)
model = Model(pre_trained_model.input,x)
model.compile(optimizer = Adam(lr=0.0001), loss = 'binary_crossentropy', metrics = ['acc'])
```

Figure 10: Building of Inception V3 model

The below graph Figure11 shows the accuracy of test and train dataset of inception V3. The number of epochs used was 85 which gave better results.

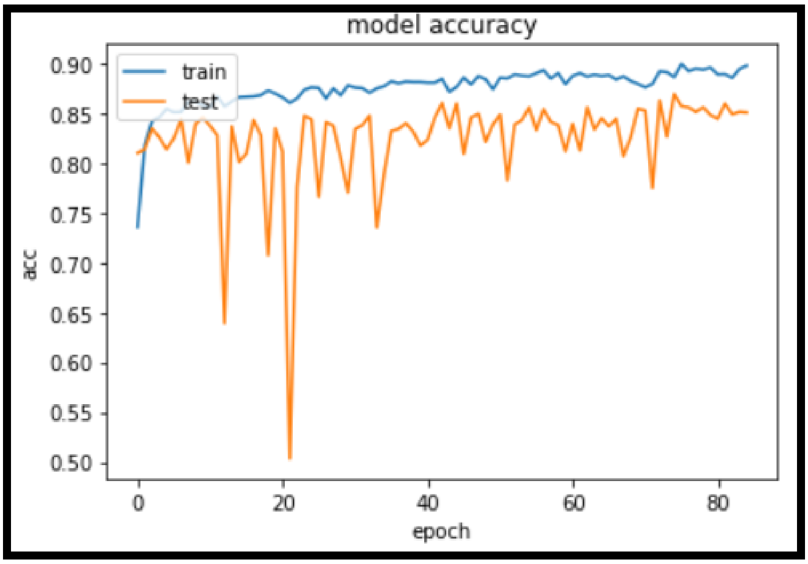

Figure 11: The model accuracy of Train and Test dataset

The below Figure12 shows the validation accuracy, loss, Train accuracy and loss. As the fig shows that there was significant loss in first 20 epochs but later it was almost linear after 40 epochs.

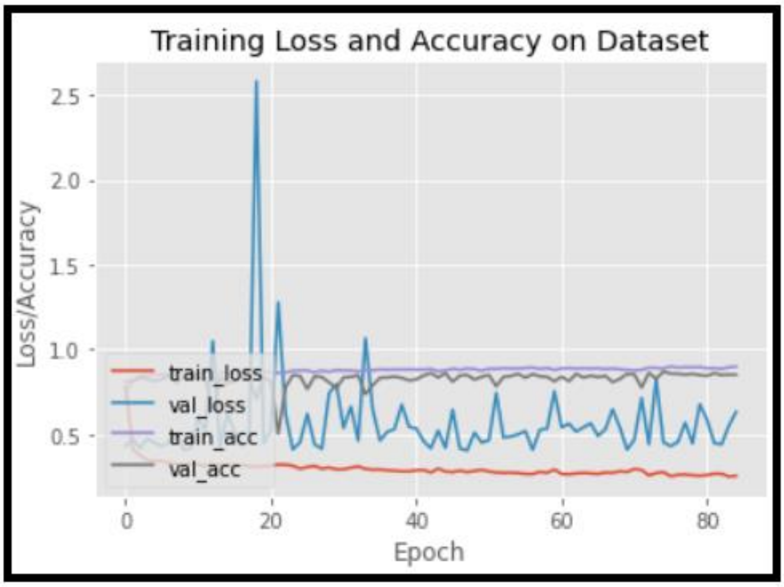

Figure 12: Overall loss and accuracy of dataset

4.3 Implementation of CNN

CNN perform very well on image datasets and are very popular. CNN with better implementation using pre-processing can give better results. Hence CNN was used. Therefore, the results were better for this dataset as the dataset was transformed by using zoom, rescale etc. In CNN model the activation function used is relu in the first layer and sigmoid in the last layer. The optimizer used was Adam as it is a combination between RMSprop and stochastic gradient descent. Binary crossentropy was used as there were 2 classes in the dataset. The Figure13 shows how the CNN model was built.

```
model = Sequential()
model.add(Conv2D(32, (5, 5), activation='relu', input_shape=(250,250,3)))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(64, (5, 5), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(1000, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(500, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(250, activation='relu'))
model.add(Dense(1, activation='sigmoid'))
```

```
model.compile(loss='binary_crossentropy',
              optimizer = Adam(lr=0.0001),
              metrics=['accuracy'])
```

Figure 13: Model code for CNN

The accuracy of CNN was 87.74%. Figure14 is the graph showing the test and train accuracy result with the number of epochs. The number of epochs was 20.

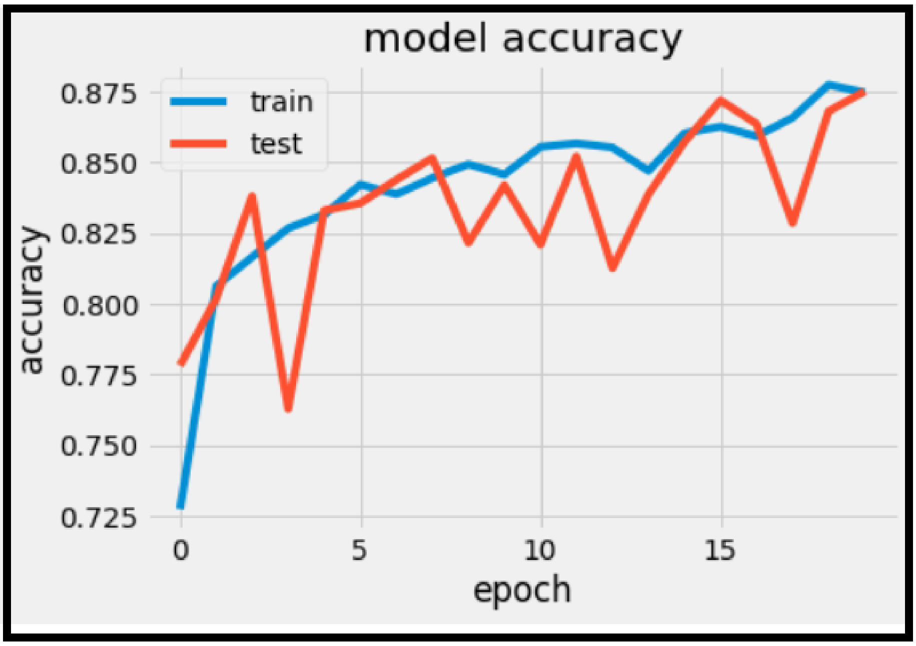

Figure 14: Model accuracy of test and train dataset.

The below Figure15 shows the overall model accuracy and loss of train and test dataset. The sensitivity obtained is 0.65 and specificity was 0.56.

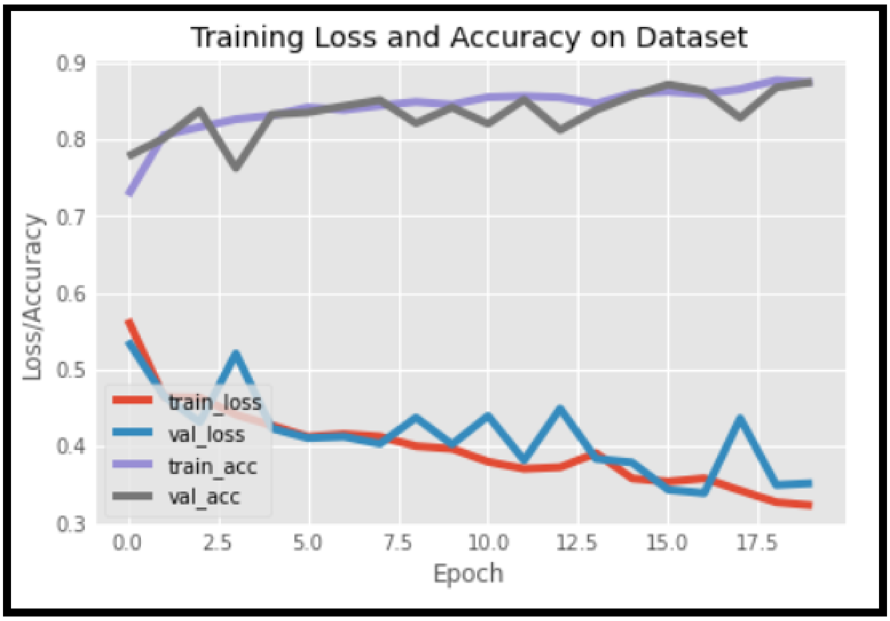

Figure 15: Shows the model loss and accuracy of the model

4.4 Evaluation and Results

The below Table3 clearly indicates that dense Net 121 performed better than Inception V3 and CNN with the accuracy of 95%. The data was imbalanced but still the sensitivity which provide the true positive values was 68% that means it classified 68% of data correctly and specificity which was 58% correctly identified the 58% negative data.

Table 3: Comparison of all 3 implemented models

Model Name	Accuracy	Sensitivity	Specificity
Dense Net 121	95%	0.68	0.58
Inception V3	89.79%	0.64	0.56
CNN	87.74%	0.65	0.56

Comparison of Previous Results

Table 4: Comparison between past research

Author	Models Used	Accuracy
(Sara, et al., 2019)	VGG19, mobileNet, DenseNet	Better than traditional methods
(N, et al., 2019)	VGG16 and ResNet50	94%
(S, et al., 2019)	Xception, DCNN	92%

When compared to the above Table4 in the past research the accuracy obtained was comparatively less from this research which is 95% as desired. As in this research we have used multiple pre-processing techniques and created meta data for better results.

5 Discussion

As from the above table Dense Net 121 outperformed both CNN and inception V3 on the image dataset. Inception V3 came near to dense Net with accuracy of 89.79%. The sensitivity and specificity were 0.68 and 0.58. The above objectives have been implemented successfully and the models with metadata and preprocessing have achieved better results. From the above it is safe to say Breast cancer for histological image best works with Dense Net 121 and Inception V3 models. In this research I have gained immense knowledge in how to use python and google colab using GPU to build machine learning models. Also, how to manipulate the image data and obtain a better result by using proper preprocessing technique.

6 Conclusion and Future Work

The objects of the project were successfully met. The research question stating “To what extent can machine learning algorithms give better results for histopathological image. All three algorithms performed well but Dense Net121 performed better than the other two. The performance could be improved further as there was data imbalance between the classes in the dataset. In Future work the imbalance of the data can be removed by using different preprocessing techniques and the performance metrics can be increased. As seen in the above table it is proven that Dense Net 121 gave better results for histopathological image. In future scope ensembled method can be used with using multiple machine learning algorithms.

Acknowledgment

The implementation of research project was possible due to the help and guidance of Dr. Catherine Mulwa. She played a key role and motivated me to complete this project. I would also like to pay my gratitude to my family members for financial support and moral support needed through my Masters. Special thanks to my colleague as well.

References

A, Mert, N., Kilic, A. & Akan, 2011. *Breast cancer classification by using support vector machines with reduced dimension.* s.l., s.n., pp. 37-40.

B. N. Narayanan, V. Krishnaraja and R. Ali, "Convolutional Neural Network for Classification of Histopathology Images for Breast Cancer Detection," *2019 IEEE National Aerospace and Electronics Conference (NAECON)*, Dayton, OH, USA, 2019, pp. 291-295, doi: 10.1109/NAECON46414.2019.9058279.

D. S. Jacob, R. Viswan, V. Manju, L. PadmaSuresh and S. Raj, "A Survey on Breast Cancer Prediction Using Data MiningTechniques," *2018 Conference on Emerging Devices and Smart Systems (ICEDSS)*, Tiruchengode, 2018, pp. 256-258, doi: 10.1109/ICEDSS.2018.8544268.

H, et al., 2016. *Neural Network architecture for breast cancer detection and classification.* Beirut, s.n., pp. 37-41.

H. Rajaguru and S. Kumar Prabhakar, "Bayesian linear discriminant analysis for breast cancer classification," *2017 2nd International Conference on Communication and*

Electronics Systems (ICCES), Coimbatore, 2017, pp. 266-269, doi: 10.1109/CESYS.2017.8321279.

M. Amrane, S. Oukid, I. Gagaoua and T. Ensarİ, "Breast cancer classification using machine learning," *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, Istanbul, 2018, pp. 1-4, doi: 10.1109/EBBT.2018.8391453.

M. NEMISSI, H. SALAH and H. SERIDI, "Breast cancer diagnosis using an enhanced Extreme Learning Machine based-Neural Network," *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*, Guelma, Algeria, 2018, pp. 1-4, doi: 10.1109/SIVA.2018.8661149.

N. S. Ismail and C. Sovuthy, "Breast Cancer Detection Based on Deep Learning Technique," *2019 International UNIMAS STEM 12th Engineering Conference (EnCon)*, Kuching, Malaysia, 2019, pp. 89-92, doi: 10.1109/EnCon.2019.8861256.

Jasmir *et al*., "Breast Cancer Classification Using Deep Learning," *2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*, Pangkal Pinang, 2018, pp. 237-242, doi: 10.1109/ICECOS.2018.8605180.

P. T. Nguyen, T. T. Nguyen, N. C. Nguyen and T. T. Le, "Multiclass Breast Cancer Classification Using Convolutional Neural Network," *2019 International Symposium on Electrical and Electronics Engineering (ISEE)*, Ho Chi Minh, Vietnam, 2019, pp. 130-134, doi: 10.1109/ISEE2.2019.8920916.

P. Mekha and N. Teeyasuksaet, "Deep Learning Algorithms for Predicting Breast Cancer Based on Tumor Cells," *2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON)*, Nan, Thailand, 2019, pp. 343-346, doi: 10.1109/ECTI-NCON.2019.8692297.

Rashmi G D, A. Lekha and N. Bawane, "Analysis of efficiency of classification and prediction algorithms (Naïve Bayes) for Breast Cancer dataset," *2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, Mandya, 2015, pp. 108-113, doi: 10.1109/ERECT.2015.7498997.

S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider and R. Deters, "Breast Cancer Diagnosis with Transfer Learning and Global Pooling," *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Korea (South), 2019, pp. 519-524, doi: 10.1109/ICTC46691.2019.8939878.

Sara, et al., 2019. *Classification of histopathological biopsy images using ensemble of deep learning networks.* s.l., s.n., pp. 92-99.

S, et al., 2019. *Breast Cancer Diagnosis with Transfer Learning and Global Pooling.* s.l., s.n., pp. 519-524.

S. Lee, M. Amgad, M. Masoud, R. Subramanian, D. Gutman and L. Cooper, "An Ensemble-based Active Learning for Breast Cancer Classification," *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, San Diego, CA, USA, 2019, pp. 2549-2553, doi: 10.1109/BIBM47256.2019.8983317.

S. Murugan, B. M. Kumar and S. Amudha, "Classification and Prediction of Breast Cancer using Linear Regression, Decision Tree and Random Forest," *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, Mysore, 2017, pp. 763-766, doi: 10.1109/CTCEEC.2017.8455058.

S. Sathya, S. Joshi and S. Padmavathi, "Classification of breast cancer dataset by different classification algorithms," *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, 2017, pp. 1-4, doi: 10.1109/ICACCS.2017.8014573.

S. Sharma, A. Aggarwal and T. Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms," *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Belgaum, India, 2018, pp. 114-118, doi: 10.1109/CTEMS.2018.8769187.