Breast Cancer Detection from Histopathological Images using Deep Learning and Transfer Learning

MSc Research Project
Data Analytics

Mansi Chowkkar
Student ID: 18134599

School of Computing
National College of Ireland

Supervisor: Cristina Muntean

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Mansi Chowkkar
Student ID:	x18134599
Programme:	Data Analytics
Year:	2019-2020
Module:	Research Project
Supervisor:	Cristina Muntean
Submission Due Date:	12/12/2019
Project Title:	
Word Count:	6300
Page Count:	23

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:	
Date:	29th January 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 from Histopathological images using Deep Learning and Transfer Learning

Mansi Chowkkar
x18134599

Abstract

Breast Cancer is the most common cancer in women and it's harming women's mental and physical health. Due to complexities present in Breast Cancer images, image processing technique is required in the detection of cancer. Early detection of Breast cancer required new deep learning and transfer learning techniques. In this paper, histopathological images are used as a dataset from Kaggle. Images are processed using histogram normalization techniques. This research project implements the Convolutional Neural Network(CNN) model based on deep learning and DenseNet-121 based on transfer-learning. Transfer learning uses the Imagenet pre-trained model for training. Hyper-parameter tuning is done for increasing accuracy and precision value. Research achieved 90.9 % test accuracy using the CNN model and 88.03 % accuracy by the transfer learning model.

1 Introduction

Breast cancer is the most common cancer among women in the world. As compared to other cancer types, Breast Cancer has the highest mortality rate. 8.8 million deaths happened in 2015 due to breast cancer. This death rate will increase up to 2.1 million per year over the existing rate, according to the world health organization Elelimy and Mohamed (2018). For early detection, there are different imaging modalities such as mammography, ultrasound, MRI, tomography, and biopsy. Mammography is a simple and common technique for early detection of cancer but it cannot detect breast cancer in a dense breast. CT scan method uses radiation which generates genetic mutation. Ultrasound method produces false-positive results. Breast Biopsy is the most precise solution where a sample piece of tissue is taken for the microscopic examination. This study of biological tissues under the microscope is called histopathology as mentioned in Kaushal, Bhat, Koundal and Singla (2019). An expert doctor is required to examine the histopathological images and it is a time-consuming process. The computer-assisted diagnosis(CAD) is used for classifying malignant and benign cancer from images. This study focuses on classifying histopathological images using deep learning and transfer learning techniques. The use of Deep learning will reduce the load from the radiologist and they can more concentrate on only suspicious cases.

As deep learning-based CNN uses more layers, new problems are arising. The gradient is transferred to many layers and information can get lost by the time it reaches to the last layer. This problem can be addressed by using VGG16, ResNet, CNN and DenseNet. Although these models have different network structure and training procedures, they serve the same feature of providing a connecting path from the previous layer to the next layer. DenseNet reduces gradient problem and provides feature reusability. DenseNet reduces the number of parameters used in the network for training the model as mentioned by Huang et al. (2017).

This study will use Histogram normalisation as an image processing technique. Processed images are augmented to avoid biasing and overfitting problems. This research uses CNN with a parameter tuning process to improve the accuracy. This approach increases the true negative number for cancer detection. Transfer learning will be implemented for avoiding the over-fitting problem and training the model on less number of images for specific magnification Sabeena Beevi et al. (2019). Transfer learning is used with the Imagenet pre-trained weight and Sigmoid classification function. Deep learning and fine-tuning method will improve the diagnostic efficiency and provide accurate cancer detection results to the doctors. This behavior will add diagnostic value to the medical domain by saving time.

1.1 Research Question

This research tries to find a solution to the following research question:

How efficiently, accurately and quickly Deep learning and Transfer learning algorithm detect Breast Cancer from Histopathological images?

2 Related Work

Many Deep learning researches on breast cancer detection are done on mammography and ultrasound images. Recent advancement in transfer learning and fine-tuning technology is useful in the medical imaging field for efficient detection of diseases. To detect cancer from image data, various image processing, deep learning and transfer learning methods are discuss and summarised below.

2.1 Deep Learning in Health care

In Machine learning, Deep learning algorithms are rapidly becoming a choice for the medical image analysing technique. Research done by Litjens et al. (2017) summarises image classification techniques using deep learning models. This research study explains various packages in python for CNN models. It compares all the CNN models and prepossessing techniques for medical image data. This study has shown that Deep learning has a good impact on medical image analysis. In the research study Li, Xie and Shen (2019), Atrous Dense Net(ADN) and deep-reverse active learning(DRAL) is applied to pathological image data. DRAL is used to remove mislabelling from the training of cancer image data. To overcome the over-fitting problem RefineNet with six layers of CNN is adopted for the DRAL model in this research. Statistical analysis proved that VGG-16 and ADN models both achieved the same performance. ADN achieved better accuracy as compared to DenseNet-121, ResNet and VGG16. DenseNet-121 has a large model size hence it can perform better on different datasets.

The deep learning approach is used in Vo et al. (2019) for image classification. The deep convolution neural network learning is used in this research for feature extraction from multi-train training image data. This has improved multiple labeled cancer detection accuracy. The use of gradient boosting is also proposed for enhancing the performance of classification. In this study, the Inception-ResNet-v2 model is compared with the Cruz-Roa-CNN model and the Inception-ResNet-v2 model outperformed. To improve the performance of this model, in future use of large image data is suggested to overcome the data limitation problem. Hence deep learning models will provide better accuracy on large histopathological image data. A recent study Azli et al. (2018) classifies cancer from mammography images through a backpropagation neural networks. Research shows that CNN with 100 hidden nodes improves the performance of the classifier than CNN with 3 nodes. CNN with 100 nodes achieved an accuracy of 70%. The architecture of CNN can be modified to improve accuracy with hyper parameter tuning and adding Dense layers. In a study performed by Xiang et al. (2019), CNN is implemented for the classification of histopathological images. Dataset consists of 4 different types of magnifications images. All images with all magnifications are used in the training process of the CNN model and accuracy is calculated. Data augmentation and cross-validation methods are used for improving model accuracy. The softmax classifier is used in the output layer of the CNN. This study achieved 93.2% accuracy. Stochastic gradient descent (SGD) and backpropagation are used here for the parameter updating process. The similar data is being used in this study hence this research can be extended here. To extend Xiang et al. (2019) study different architectures for the CNN model can be used in this research with different parameter tuning.

2.2 Image Processing techniques

Detection of cancerous cells from histopathological images helps in the early detection process in medical diagnosis. In Medical Image processing, Normalisation plays an important role in reducing contrast and intensity. In the Anghel et al. (2019) study, stain normalisation is performed on the image data using the unsupervised stain-vector method. In this study of classification using the neural network, the algorithm is compared before stain normalisation and after stain normalisation. The study proved that Stain normalisation improved the accuracy of the CNN model from 79% to 87%. Hence normalisation has a major impact on the accuracy of CNN model. Colour normalisation and stain separation from the images are difficult processing. In Hidalgo-Gavira et al. (2018) this paper, blind colour deconvolution method based on the Bayesian technique is proposed. This method finds stain concentration from the histopathological images based on the color vector matrix. This method is proved to be faster, easier, and robust to different parameter values. Kaushal, Koundal and Singla (2019) study compares different segmentation techniques to detect cancerous cells from histopathological images. These techniques are evaluated using accuracy, true positive and false positive rate. Segmentation can be performed before applying a classification model to image data for better performance of classification.

Jiang et al. (2019) proposed a deep learning method symmetric residual convolutional neural network (SR-CNN) method for enhancing the images by sharpening the edges of the images. Results are evaluated using real-time patient data by augmenting images under-sample cone-beam computed tomography (CBCT). In Rundo et al. (2019) research study, the MedGa image enhancement process is introduced which uses a bimodal intensity histogram. This method improves image quality and appearance for the region of interest. This method crops the image in the rectangular box to increase the region interest of the image. Enhancing the contrast of the actual region of interest is more important than enhancing the whole contrast of the image in medical image processing. Zero-mean normalisation method is used for pre-processing mammography images for breast cancer classification inLi, Zhuang, Li, Zhao and Ma (2019). This processing enhanced the model robustness and increased the training speed of the model. Image normalisation can also have an impact on accuracy, true negative values which can be evaluated in this research study.

2.3 DenseNet

Deep learning is used by most of the researchers for classifying medical images. In the research Li, Zhuang, Li, Zhao and Ma (2019), mammography images are classified as benign and malignant cancer using the DenseNet-2 model. This model replaces the first convolution layer with Inception-Net which increases the performance of the model. Augmentation is used to overcome the over-fitting problem and data insufficiency problem in this research. Ten fold cross-validation is performed to check which model performed best. The result has shown that DenseNet-2 performed better with 94.55% accuracy than VGGNet, DenseNet, and AlexNet. DenseNet can be evaluated on different datasets to check its performance using some parameter tuning in the network without any modifications. Recent research on CNN shown that Dense Net can provide accurate results with less number of parameters Gottapu and Dagli (2018). In this research DenseNet, architecture is used for brain MRI image segmentation. This method is used for finding abnormalities in the brain. Results are compared with and without compression. This model achieved 95.5 % accuracy. Different dataset (brain segmentation) is tested to make DenseNet model as a benchmark. For benchmarking this model can even be tested on images without segmentation as a classifier.

For classifying pancreatic cancer from CT scan images Li, Reichert, Lin, Tselousov, Braren, Fu, Schmid, Li, Menze and Shi (2019) research study applied Dense Net feature learning model. This method is based on the CAD approach for the early detection and segmentation of pancreatic cancer. For the visualising results for the radiologists, the study implemented the saliency maps analysis tool. DenseNet model achieved 72.8 % accuracy which is higher than the diagnostic accuracy. As compared to other researches this accuracy is lower and hence more research can be done on DenseNet to increase the accuracy. Research Oyama and Yamanaka (2017) implemented the DenseNetSal model for predicting saliency maps which uses VGG16 with 150 convolution layers. DenseNet-161 is used by this research which uses 161 convolution layers and four dense blocks. DenseNet-161 architecture is modified here and it is observed that it performed better than VGG16. Here accuracy of DenseNet-161 increased by using two scales as 0.5 and 1, in future parameters can also be changed to see the impact on accuracy. Huang et al. (2017) compared DenseNet-121, DenseNet-161, DenseNet-201, and DenseNet-BC on different image data and imagenet. A study proved that DenseNet requires less parameter change and less computation time for achieving an accuracy similar to ResNets. As the parameters grow in DenseNet, it shows consistent performance growth.

2.4 Transfer Learning

In the medical image classification and diagnosis, CNN plays an important role and its a growing technique. For training the CNN model we require a large amount of image data but in the medical domain, there is an insufficient amount of image data due to privacy-preserving policies. Hence transfer learning is growing in the medical image diagnosis field rapidly to overcome this problem Vogado et al. (2018). In Talo (2019) study, transfer learning is used for classifying histopathological images. ResNet-50 is used for classification of colored images and DenseNet-161 is used for grayscale image classification and it achieved %95.7 accuracy. These models use pre-trained weights which are trained on ImageNet with 1000 classes of image data. DenseNet-161 uses fewer parameters than CNN and it reduces over-fitting problems. The transfer learning method increases the rate of the training process and accelerates CNN model construction. DenseNet-121 uses even less parameters than DenseNet-161. In the future, this model can be used as a transfer learning method using imagenet as a pretrained weight.

Sabeena Beevi et al. (2019) research study used the transfer learning approach by pre-trained convolutional neural networks for feature extraction. In this study, the output layer of softmax activation is replaced by random forest. In this research automated mitosis detection from the histopathological images is carried out using VGGNet and Caffe model with ImageNet as a pre-trained weight. The use of the transfer learning approach overcome the issue of an insufficient large amount of labeled data. In the research study Vogado et al. (2018), leukaemia diagnosis is made using transferred learning where pre-trained CNN algorithms(VGG, AlexNet)are used for feature extraction. The SVM classifier then classifies these extracted features. In this study robustness of the model is proved using three different data-sets in training. Results are evaluated using confusion matrix, accuracy and kappa index. The transferred learning-based model's performance is calculated in d. Nóbrega et al. (2018) to classifying lung cancer from CT lung images. Transfer learning models are used here to extract the features and then it is provided to the classifier. It is seen that ResNet50 and SVM RBF provided the best results with f1 score 78.83% which is more than the related work on the same dataset. In future models based on transfer, learning can be used to improve accuracy.

2.5 Hyper parameter Tuning

Deep learning has shown progress in image classification in the medical field. Deep learning requires a large number of images for training, validation, and testing. To resolve this problem Oyama and Yamanaka (2017) used a pre-trained CNN model and performed fine-tuning layer by layer using a transfer learning strategy. Accuracy is improved to 94% on contrast-enhanced magnetic resonance images(CE-MRI) data using a cross-validation method. Tao Tan (2018) study implements transfer learning using DenseNet-121 model with fine tuning. Experiment uses numbers of convolutional layers and the number of epochs as a part of the fine tuning. Study obtained 82.0% accuracy overall after fine tuning. Research shown the benefits of sequential fine-tuning in cancer detection. Due to the insufficient amount of image data study could not able to tune a large number of parameters in the model. This can be extended in this research study with comparatively large image data set.

2.6 Summery

This study implements a histopathological image processing technique as we have seen image normalisation increases the accuracy. As mentioned in Xiang et al. (2019), this research will use a basic CNN model built from scratch for classification with parameter tuning. DenseNet has shown great value in medical image classification. A little research is done on DenseNet for histopathological images. Therefore this research paper extending the use of DenseNet based on transfer learning for histopathological data. In DenseNet models, very few research is available on DenseNet-121, this research will extend the transfer learning-based DenseNet-121 model used in Tao Tan (2018). As we have seen parameter tuning increases the accuracy without much modifications and cost to the model. Therefore this research implements parameter-tuning on DenseNet and DenseNet-121 parameters to get the best suitable parameters for histopathological image data.

3 Methodology

This research work applied the CNN model to histopathological images to classify benign and malignant cancer accurately as discussed in 2.6. For implementing the transfer learning method, research uses the DenseNet-121 model as a novelty on the histopathological image data set. The Cross-Industry Standard Process for Data Mining (CRISP-DM) method is used for this research as shown in figure 3. It will be explained in Business understanding, data understanding, data preparation, and, modeling sections below as mentioned inTie et al. (2011).

3.1 Business Understanding

Detection of cancer from mammography images is a very common method but a breast biopsy is a very precise method and its prediction requires more time and expertise. Pathologists and doctors evaluate histopathological images after preprocessing of images. As the preparation of images involves sectioning of images, collection, staining, and fixation step Elelimy and Mohamed (2018). Hence evaluation of biopsy images under the microscope is very time consuming for pathologists and it is highly impacted on the experience of the doctor in this field. As discussed in the related work summary, a deep Learning-based CNN model is used in this research. CNN is used to reduce the time required for evaluating histopathological images. In the medical field, images are insufficient for training the CNN model with a large number of layers. Hence transfer learning-based DenseNet-121 model and parameter-tuning approach are used to overcome this problem.

3.2 Data Understanding

Image data is taken from Kaggle with version 1 [1]. This data is having only two kernels available. Total 7909 images are present with four different magnification as 40X, 100X, 200X, and, 400X. Table 1 explains the image distribution concerning image magnifications and the two classes. Figure 1 shows malignant and benign images of all the magnifications from the dataset. 40X_B indicates a benign image of 40X magnification and 40X_M indicates a 40X magnification image of malignant type. These images are then divided into train, test and valid folders with 80:10:10 proportion as explained in table 1. 5439 images are of type malignant and 2480 images of type benign. All images are three dimensional with each image size 460*700. Considering all these factors about image data, all data preparation steps are discussed in the following section.

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

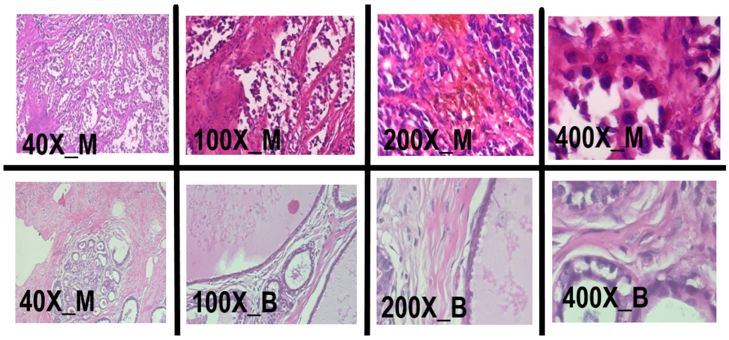

Figure 1: Different Magnification Images

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

Table 1: Image distribution as per the class and magnification

3.3 Data Preparation

For providing images to the CNN and DenseNet-121, images are normalised for more readability. Images then augmented and saved using rotation and batch size parameters. At the time of model execution images are augmented run time using different parameters such as rotation, zoom, scaling, batch size and image size.

3.3.1 Image Normalisation

Medical images require image enhancement to help pathologists for detecting disease and its diagnosis. Images go under an automated deep learning model needs to improve its quality for better results. In Zhuang and Guan (2017), Histogram Equalization method for image enhancement based on mean and variance is performed to increase the contrast and brightness of the image without compromising details of the image data. As mentioned above, in this research the Histogram Equalization method is applied to histopathological image data for increasing contrast and brightness of the image. Histopathological images need to undergo microscopic examination hence they contain some staining process and difficult to read. Image normalization will increase the contrast of the image. Hence it increases the accuracy and performance of the deep learning model. Figure 2 shows the effect of histogram normalisation on one of the benign types of histopathological images.

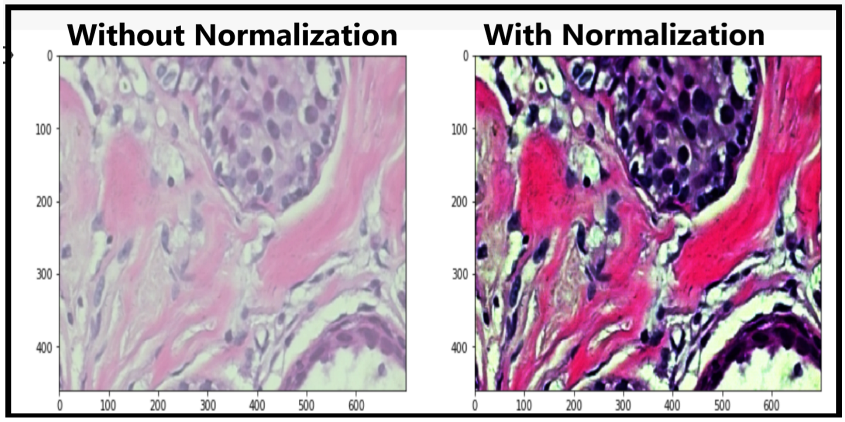

Figure 2: Histogram Normalisation effect

3.3.2 Image Augmentation and Upsampling

After normalising images, images are upsampled before providing it to CNN based DenseNet model and the transfer learning. In the study Tellez et al. (2019), basic augmentation is used with scaling, gaussian blurring, and gaussian noise. Morphological augmentation consists of horizontal and vertical mirroring of images, 90 degree, 180-degree image rotations or scaling of images by 0.5 to 1.0 factor. In this research, benign images are upsampled to match with malignant image count. For augmentation, 90-degrees rotation and batch size of 32 is used. While providing as an input to the classification model run time augmentation is performed by tuning different parameters as explained in detail in the evaluation section. Table 2 explains the number of benign histopathological images with all magnification before and after upsampling. Table 3 explains all the hyperparameters used and their corresponding values tuned for selecting the best suitable parameters for the classifiers. Effect of all values from table 3 is explained in detail in the section 6.

Magnification	Benign images before upsampling	Benign images after upsampling
40X	500	986
100X	515	1017
200X	498	982
400X	470	932
All magnifications	1983	3795

Table 2: Histopathological image data distribution

Parameters in Augmentation	Values used
Image Rotation	60, 90, 180, 270
Width shift range	0.2
Height shift range	0.1
zoom range	0.4, 0.5
target size	224*224, 128*128, 64*64, 32*32
batch size	64, 32,16
shuffle	true
colour mode	RGB
horizontal flip	true

Table 3: Parameters used in Image Augmentation

3.4 Modelling

This section explains the architecture used for the CNN and DenseNet-121 model.

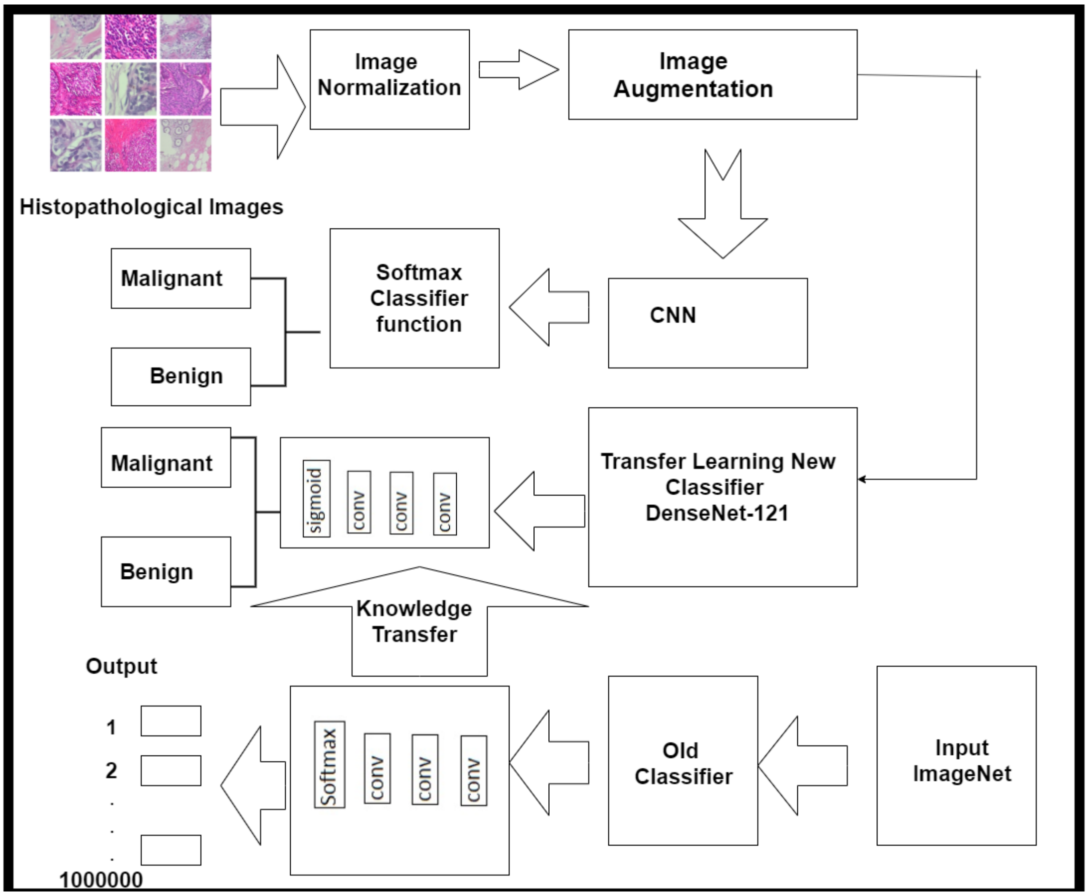

Figure 3: Flow Diagram

3.4.1 CNN model

CNN consists of Transition layers, dense blocks, and convolutional layers. CNN consists of a sequence of neural networks for extracting features. To circulate features at different level it concatenates all features from previous layers. There are mainly five layers present in the architecture of the CNN as mentioned in Sabeena Beevi et al. (2019). To capture the local information from the input, the convolutional layer is used. The filter is used to compute the dot product of its weights and the input image region. The batch normalisation layer controls the output from the previous layer and sends it to the next activation layer. Since normalisation is used in the gradient process, it balances the weights in the network. Rectified Linear Unit (ReLU) is used as an activation function in the CNN. ReLU accelerates the execution of CNN. The complexity of the features increases as CNN becomes deeper. MaxPooling layer is used for shrinking the feature map size. Each fully connected layer neurons connect with the previous layer for calculating class count. The output layer consists of a softmax or sigmoid as a classification function.

3.4.2 DenseNet121 model based on Transfer Learning

Transfer learning requires a small number of images for training the model. The transfer learning approach is used in this research study by using DenseNet-121. The model uses 121 layers and Imagenet as a pre-trained weight as shown in the figure 3. Dense layers are very narrow and classification is based on knowledge of already trained weights as shown in. Therefore transfer learning learns complicated training of images from fewer parameters and less computation time Oyama and Yamanaka (2017). Table 4 explains the architecture of the DenseNet-121 model as explained in Allaouzi et al. (2019). It gives details about output layers and convolutional layers used for each dense block and transition block.

Layers	Output	DenseNet-121 k=12
Convolution	112*112	7 * 7 Conv, stride2
Pooling	56 *56	3*3 max pooling, stride2
Dense Block 1	56*56	$\begin{bmatrix} 1*1conv \\ 3*3conv \end{bmatrix}$ *6
Transition 1	56*56 28*28	1*1 conv 2*2 Avg pool, stride 2
Dense Block 2	28*28	$\begin{bmatrix} 1*1conv \\ 3*3conv \end{bmatrix}$ *12
Transition 2	28*28 14*14	1*1 conv 2*2 Avg pool, stride 2
Dense Block 3	7*7	$\begin{bmatrix} 1*1conv \\ 3*3conv \end{bmatrix}$ *24
Transition 3	14*14 7*7	1*1 conv 2*2 avaerage pool, stride 2
Dense Block 4	7*7	$\begin{bmatrix} 1*1conv \\ 3*3conv \end{bmatrix}$ *16
Classification layer	1*1(Sigmoid)	7*7 global avg pool

Table 4: DenseNet-121 Architecture

4 Design Specification

The following are the specifications used for both the models.

Convolution layer: In this layer the number of filters is present is the same as the growth rate for extracting features Khened et al. (2019). In this experiment growth rate is 12 hence there are 12 filters present in the DenseNet model.

BatchNormalisation: When the output of the previous convolution layer provided to the next layer, BatchNormalisation is used for normalising output which reduces the overfitting problem.

MaxPooling2D: One max-pooling layer is used in this project to reduce the dimensionality of the feature map.

Dropout layer: 0.2 Dropout is set to skip the connection for matching the input channels. Dropout layer helps in reducing overfitting of the model

Dense (Fully Connected): The dense block is placed before the output classifier for aggregating the information which is extracted from previous layers. In this study, three dense blocks are used in the DenseNet model.

Softmax layer: The final output activation function in the dense layer used is a softmax layer which generates the final classification label map.

Loss Function: In medical images, class imbalance occurs between background and region of interest. To overcome this issue here the cross-entropy loss function is used.

Weight: In the transfer learning 'imagenet' pre-trained weight is used for training DenseNet-121.

Optimizer: Adam optimizer is used here with a learning rate of 0.0001.

Sigmoid layer: The final output activation function in the dense layer of DenseNet-121 used is a Sigmoid layer which generates the final classification label map.

5 Implementation

This section provides information on how the two models are implemented to classify breast cancer accurately from histopathological images.

5.1 Setup

CNN and transfer learning models require more time to process the images while training the model. The experiment is carried out on the Google colaboratory with 100 GB drive storage, 12.72 GB RAM and 48.97 GB run-time GPU. Both the models require more time for execution on large image data as they consist of more layers. Runtime usage of GPU accelerates the execution time of CNN and DenseNet-121. For implementing CNN and DenseNet121, python libraries Keras and TensorFlow are used. Python version 3 is used on google colaboratory notebook. Data is uploaded on google drive and accessed by mounting google drive using python API. Image normalisation and up-sampling are done using numpy, PIL and Keras libraries from python. Normalised and Upsampled images are provided to the CNN and DenseNet-121 as shown in the figure 3.

5.2 Data Handling

All the histopathological images are read from google drive and loaded in the data frame. Different data frames according to benign and malignant classes concerning different magnifications are red into different files. Function called get_histogram() is written for histogram normalisation. This function is called from hist0() by providing flatten images and size 256. hist0() function takes the filename as an input with all images which need to be normalised. New normalised images are written into a new folder in a for loop for further use in the deep learning models.

Normalised images are red into data frames and checked for its biasing. It is then converted into .csv file. For getting better performance benign images are upsampled. While training the model on mixed magnification images, first all mixed magnification images are mixed and then upsampled as shown in table 2. ImagedataGenerator API from the Keras library is used for upsampling. Images are rotated by 90 degrees while augmenting and saved in the directory using datagen.flow() function. Run time augmentation is also carried out to provide train images to the model. Parameters in augmentation are changed accordingly in different case studies to study its impact on the accuracy of the model.

5.3 Architecture of the CNN model

For implementing the CNN model, functions are written to create a convolution layers, Dense blocks, transition layers, and Dense network.

1. Convolution layers function is written using 'relu' activation function and 0.2 dropout value. Dropout is changed to 0.3 to check the impact on accuracy but it is seen that 0.2 dropout provides better accuracy than 0.3. Convolution layer takes input as a dense block and filter size. The filter size is the same as that of the growth rate which is set to 12. This layer returns the convolution layer as a function of dropout, normalization, and activation.

2. The dense Block function is written which returns dense block concatenated with the previous input layer. This layer also returns the number of filters in the block.

3. Transition block consists of the Batch normalization layer, 2D convolution layer, 'relu' activation layer and average pooling layer.

4. Finally, a Dense net is created which creates an input layer with input image shape. Input image shape is changed in different case studies to check its impact on accuracy. The output layer is created with the 'softmax' activation function.

5. CNN is created with three dense blocks and five layers. The model has then compiled with 'Adam' optimizer and 'cross-entropy' loss function.

6. model.fit_generator() API is used to run the model with train data and valid data. Epochs are varied in different case studies.

The model is compiled using the model.compile() function with Adam optimizer and 0.0001 learning rate.

5.4 Transfer Learning: DenseNet-121

Transfer learning is used to transfer the knowledge of already trained models to another model for the classification of image data. The DenseNet-121 model is used with the 'Imagenet' pre-trained weight as shown in figure 3. In the output layer, the Sigmoid activation function is used. Model is executed on the number of epochs using model.fit_generator() function. This function takes training images, validation images, steps count and number of epochs as an input. class_weight parameter is used with value 'auto' to improve the true negative values(cancer count). figure 4 shows the code used in DenseNet-121 model. Test accuracy of the model is calculated using model.evaluate_generatot() function.

```
base_model = densenet.DenseNet121(input_shape=(128, 128, 3)
                                  weights='imagenet',
                                  include_top=False,
                                  pooling='avg')

for layer in base_model.layers:
    layer.trainable = True

x = base_model.output
predictions = Dense(2, activation='sigmoid')(x)
model = Sequential()

model = Model(inputs=base_model.input, outputs=predictions)

model_history = model.fit_generator(
    train,
    epochs=10,
    steps_per_epoch=257,
    validation_data=valid,
    validation_steps=24,
    class_weight='auto')
```

Figure 4: DenseNet-121

6 Evaluation

This section discusses the comprehensive analysis of two models and all parameters which are tuned to get the best performance. For this research study, two models are used on the histopathological image dataset from Kaggle. DenseNet-121 is used as a novelty on this data set. For evaluation process loss and accuracy of training and validation data is calculated per epoch for both the models. Accuracy and Loss plot for training and validation is plotted. The test accuracy is calculated as a function of prediction and test data. Confusion metrics for each model is calculated to get a true positive and true negative value. The experiments explained below, discuss the increased accuracy and increased value of true negative as they are important factors for medical diagnosis. Precision (Specificity) is also calculated as a function of true negative and false positive value. As mentioned in Tao Tan (2018), fine-tuning is performed by changing different parameters in the model and the effect on accuracy is observed in the following experiments.

6.1 Experiment 1: Effect of Different magnifications on test Accuracy of the CNN model

All magnification images are provided as training and validation data in the CNN model. In this experiment, test accuracy is calculated for different magnification images and results are listed in table 5. The experiment evaluates accuracy on different image sizes as shown in table 5. Accuracy is also compared against different image rotation (60,90,180) as mentioned by Li, Zhuang, Li, Zhao and Ma (2019). Table 5 shows image size, rotation of the image at the time of augmentation and magnification impact on test accuracy.

Image size	Rotation	Accuracy for 40X	Accuracy for 100X	Accuracy for 200X	Accuracy for 400X	Accuracy for Mix
128*128	60	81.25	76.07	70.79	67.75	71.08
128*128	90	70.62	83.7	85.6	82.4	84.15
128*128	180	77.5	84.6	84.15	81.42	75.59
64*64	60	63.75	83.7	85.14	78.14	72.01
64*64	90	65	86	82.67	80.87	73.2
64*64	**180**	71.25	90.9	87.12	81.42	76.65
32*32	60	58.75	81.3	78.12	74.8	68.96
32*32	90	61.25	87.08	85.04	81.96	72.14
32*32	180	68.12	85.16	84.65	79.23	71.08

Table 5: Test Accuracy of CNN for different Magnifications

In a research study, Kumar et al. (2020) accuracy are compared for different magnifications according to different models. Research shows that accuracy is more when magnification is 100X and 200X since these magnifications capture more area of the region and extract more details. Results in Table 5 proves this fact that 100X and 200X magnification images have more accuracy. The highest accuracy of 90.9% is achieved with 64*64 image size, 180-degree rotation, and 100X test image magnification. As seen from the results, it can be concluded that even 400X magnification gives better accuracy as compared to 40X. When low magnification images are used, to obtain more accuracy large image size is advisable as explained by Table 5. 40X images provide more accuracy when image size is 128*128 as compared to 64*64 and 32*32. Hence it is concluded that higher magnification images provide better results in the image prediction with any size of the image. It is also observed that while augmenting images, 180 and 90 degrees of rotation must be used for better results in CNN as compared to 60 degrees.

6.2 Experiment 2: Effect of class_weight parameter on accuracy and true negative value in CNN

The fit.generator() function takes an input parameter as train data, valid data, number of epochs, steps for train and validation and class_weight. hen data consists of imbalanced classes, class_weight is used to balance a particular class value. In this research, data is already balanced in upsampling. Therefor 'auto' value is used for the class_weight. Effect of class_weight = 'auto' is observed on accuracy and true negative value. The experiment is carried out on 64*64 image size data with all magnification images in training and validation. Rotation is set to 60 degrees for training and validation images. As 60 degree provided less accuracy in experiment 1, in this experiment trying to increase the accuracy of CNN for 60-degree rotation. The test accuracy is calculated on different magnifications with and without class_weight as shown in table 6.

Test Magnification	Accuracy with class_weight = auto	Accuracy without class_weight	True Negative with class_weight	True negative without class_weight	Total malignant images
40X	67.5	63.75	103	64	139
100X	89.9	83.7	104	92	144
200X	86.63	85.14	99	100	139

Table 6: Effect of class_weight on accuracy and the true negative value

As seen from experiment results, weight_class increases the true negative values and accuracy. It is beneficial for the healthcare domain to identify breast cancer cases from images. The precision value is calculated using the following formula. It is observed that for 40X image precision is increased from 45.32% to 74.10%. Accuracy not increased by major difference as compared to precision.

$$Precision(Specificity) = \frac{TN}{TN + FP}$$

6.3 Experiment 3: Comparison of Accuracy of CNN Vs Transfer learning with magnification 100X

As we have seen in experiment one, highest accuracy is achieved when magnification is 100X hence here transfer learning is applied on 100X images. Accuracy from CNN model and DenseNet-121 is compared on 64*64 image size with 100X magnification on different image rotations. For DenseNet-121, Imagenet pre-trained weight is used. Both the models are trained on 12 epochs.

Test Magnification	Rotation	Accuracy of CNN	Accuracy of DenseNet-121
100X	60	65.62	75.5
100X	90	66.8	68.89
100X	180	68.12	83.25
100X	270	75	79.42

Table 7: Effect of Transfer Learning on Accuracy

Table 7 provides the accuracy of CNN and DenseNet-121 model on 2094 training images. Transfer learning increased the accuracy of 100X magnification image data and it is observed that the highest accuracy is obtained when rotation is set to 180 degrees. Transfer learning performed better on fewer images as compared to the CNN model.

6.4 Experiment 4: Effect of different activation functions and weights in transfer learning

To select the best activation function and pre-trained weight in transfer learning this experiment is carried out. 128*128 image size and 100X magnification images are used for training, validation, and testing. Sigmoid and Softmax both activation functions are used and it is observed from results that sigmoid activation gives more accuracy in DenseNet-121. As mentioned in the literature review imagenet is widely used in all research studies as a pre-trained weight, this study compares imagenet and densenet121_weights_tf_dim_ordering_tf_kernels_notop weights for best results. From results explained in table 8 shows that imagenet weight gives more accuracy as it is trained on 1 million different classes.

Rotation	Weight used	Activation function used	Batch size	Accuracy
60	densenet121	softmax	64	41.6
90	densenet121	softmax	64	76.06
180	densenet121	softmax	64	81.33
60	imagenet	softmax	64	82.2
90	imagenet	softmax	64	82.29
180	imagenet	softmax	64	85.64
60	imagenet	sigmoid	64	85.64
90	imagenet	sigmoid	64	88.03
180	imagenet	sigmoid	64	86.12
60	imagenet	sigmoid	32	82.2
90	imagenet	sigmoid	32	82.77
180	imagenet	sigmoid	32	80.38

Table 8: Effect of activation function, batch size and weight on Accuracy

The highest accuracy obtained is 88.03% with 90-degree rotation, imagenet as trained weight, sigmoid output function, and 64 batch size. Table 8 shows that imagenet gives better accuracy than dense121 weight. It is observed from Table 8 that the batch size of 64 is more efficient in DenseNet-121. Softmax and sigmoid function have not much difference in accuracy but sigmoid performed better in DenseNet-121.

6.5 Experiment 5: Parameter tuning in transfer learning to achieve the highest accuracy

As evaluated from experiment 4, here Imagenet as a pretrained weight and Sigmoid as an activation function is used. Effect of different image size and rotations are tested on accuracy in DenseNet-121.

image size	Rotation	Test Accuracy of 100X	Test Accuracy of 200X
224*224	60	64.11	67.32
224*224	90	85.16	75
224*224	180	72.48	84.65
128*128	60	85.64	74
128*128	90	88.03	71.03
128*128	180	86.12	56.43
64*64	60	76.07	76.73
64*64	90	66.98	72.77
64*64	180	62.9	67.81

Table 9: DenseNet-121 Accuracy for different rotation, magnification and image size

It is observed from Table 9 that the highest accuracy for 100X image is achieved when image size is 128*128 and rotation is 90 degrees. The highest accuracy for 200X images is achieved when image size is 224*224 and rotation is 180 degrees. The true positive value achieved is 117 out of 139 images that is 84.17% precise. A true negative value indicates the accurate detection of cancer from images.

6.6 Experiment 6: Effect of Magnification and normalisation on Computation time

To evaluate the performance of CNN and DenseNet-121, computation time for both models is calculated on normalised and non-normalised images. Effect of different magnifications on computation time is also shown in figure 6.

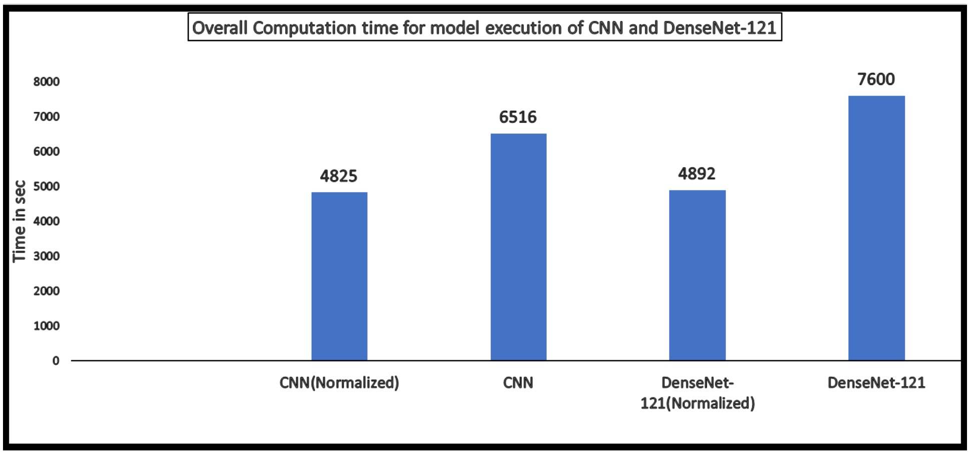

Figure 5: Effect of Normalisation on computation time

Figure 5 shows that Normalised images have less computation time for both models. CNN has less computation time as compared to DenseNet-121. All though DenseNet-121 have 121 layers in the model, its computation time is relatively less. This concludes that transfer learning takes less time for classification as it uses already trained knowledge. Figure 6 shows that magnification has no major impact on computation time.

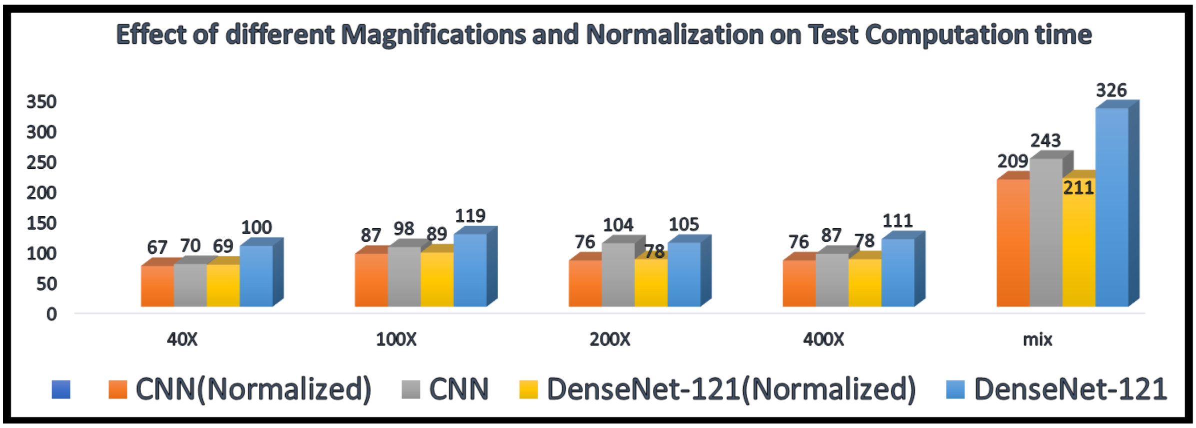

Figure 6: DenseNet-121

6.7 Experiment 7: Effect of image Normalisation on accuracy and computation time

In this experiment CNN and DenseNet-121 models are executed on normalised and without normalised images and accuracy is compared. To run the models, 128*128 image size is used with 180 degrees of rotation and 64 batch size. 10 epochs are executed for both models with a 0.5 zooming factor.

Test Magnification	Classifier	Accuracy without Normalisation	Accuracy with Normalisation
40X	CNN	66.87	71.25
100X	CNN	69.85	70.81
200X	CNN	77.72	68.8
400X	CNN	58.46	68.3
Mix	CNN	65.9	66.9
40X	DenseNet-121	49.3	74.3
100X	DenseNet-121	48	75.11
200X	DenseNet-121	50.99	64.6
400X	DenseNet-121	43.16	62.2
Mix	DenseNet-121	45.88	65.11

Table 10: Effect of image normalisation on accuracy

For 100X magnification images, the accuracy of DenseNet-121 increased from 48 to 75.11. For 400X magnification images, CNN accuracy increased from 58.56 to 68.3. After histogram image normalisation. The true negative value achieved for 200X magnification is 138 out of 139. It gives 99.28% precision for the model.

6.8 Discussion

The main aim of this research study is to increase the accuracy and precision value for cancer detection. Image normalisation preprocessing increased the average accuracy of both the models. Its effect majorly seen in the DenseNet-121. Hyper-parameter tuning increased the accuracy of the CNN model as well as a transfer learning model. However, transfer learning model, DenseNet-121 provided the highest accuracy of 88.03, its accuracy is lower when large image data with mixed magnifications are trained. Figure 7 shows the loss and accuracy of DenseNet-121 on training and validation data on 15 epochs. As seen from the figure, as loss decreases by train data, valid loss also decreases. Similarly, as accuracy goes on increasing by train data, valid data accuracy also increases. Figure 8 shows the configuration metrics for 100X image data with 128*128 size using the CNN model. In this figure 0 indicates benign cases and 1 indicates malignant cases. A true negative value is 119 from 144 malignant cases and it shows 20 true positive value from 65 benign cases. It shows that 30% sensitivity and 82.63% specificity(precision). If the model is trained on a large number of epochs, the model will train more accurately and it may increase the accuracy and sensitivity. Due to limitations on hardware, this experiment is performed on a fewer number of epochs.

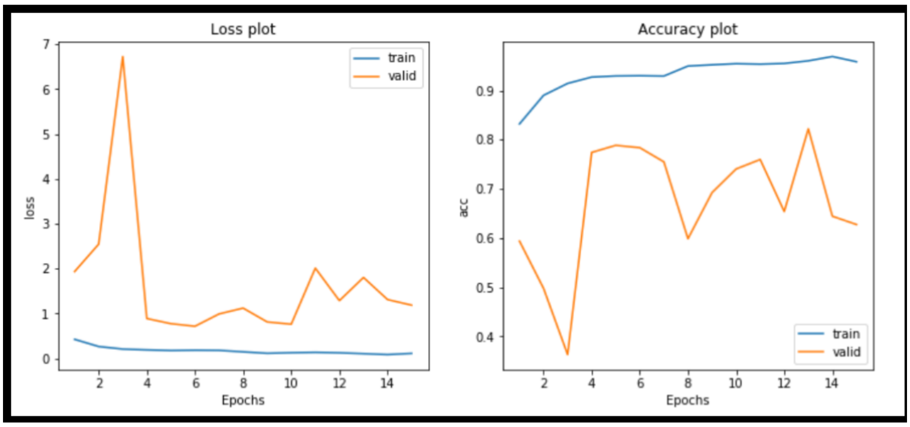

Figure 7: Loss and Accuracy of DenseNet-121 model while training

As compared to state-of-art accuracy of 94.55% reported in Xiang et al. (2019), this CNN model achieved less accuracy of 90.9%. The reason behind this may be that the state-of-art uses Inception-V3 model architecture. CNN architecture of state-of-art is very complex and hence they could not try fine-tuning on more parameters. This research tried fine-tuning on many parameters and it is concluded from results that CNN built here best work with 64*64 image size with dropout value=0.2 and number of transition layers=5. Transfer learning with fine-tuning achieved 88.03% overall accuracy as seen in experiment 5. DenseNet-121 achieved more accuracy as compared to state-of-art Tao Tan (2018). Study compared image size, rotations, weights, activation function and several epochs in the DenseNet-121. Since images were 2165 for 100X magnifications and 2044 for 200X magnification respectively more accuracy is achieved as more parameters are available to fine-tune. DenseNet-121 execution is limited to only two weights for pre-trained weights. CNN trained model could be used as a pre-trained weight in DenseNet-121.

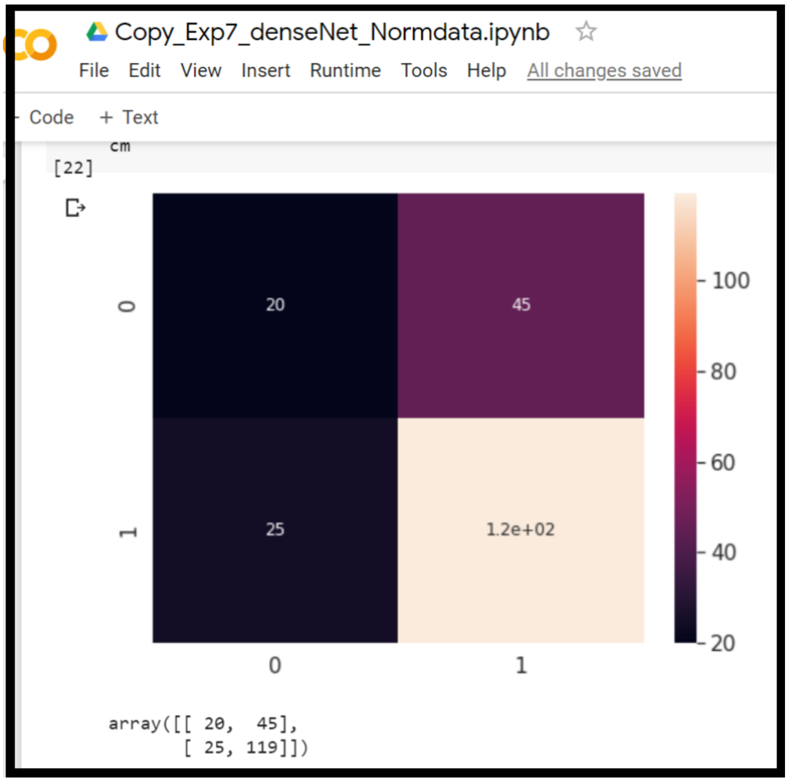

Figure 8: Confusion Metrics

7 Conclusion and Future Work

This research implemented a deep learning model CNN to find breast cancer from histopathological images. The research found cancerous cases 90.9% accurately. The research concludes that 100X and 200X magnification images give more accuracy in the diagnostic study. Hence biopsy results can be improved using 100x and 200X magnification in the clinical field. Histogram normalization method increased the precision value to 99.28% and it reduced the execution time. Hence CNN model can be used on histopathological image data for a quick and accurate finding of the cancer. Histogram normalization, image augmentation and fine-tuning are utilized during the training process to solve the overfitting problem. In the future, Stain-normalization can be applied to histopathological images before classifying. Also, cross-validation can be used in the future for improving the performance of the CNN network Xiang et al. (2019).

Research implemented transfer learning-based DenseNet-121 model for working on individual magnification image data. DenseNet-121 achieved 88.3% accuracy after performing fine-tuning which is greater than state-of-art. It is concluded that Imagenet pretrained weight and Sigmoid output activation function is best suitable for this image data. In transfer learning, 128*128 images with 90-degree rotation outperformed in terms of accuracy and 224*224 image outperformed in terms of cancer detection precision. The use of class_weight parameter in the model execution increased the true negative value in confusion metrics as seen in figure 8. However this experiment did not achieve good true positive value. Research finds out cancerous cases correctly with 99.28% precision but it failed to detect non-cancerous cases correctly. This concludes that our model is best suitable for detecting cancerous cases than detecting non-cancerous cases. In the future to increase the sensitivity of the model more parameters could be tuned. In future, transfer learning-based various models can be tried on this image data. The number of epochs and different image enhancement process can try out for the better performance of the model. To extend this study feature extraction can be performed using transfer-learning. These extracted features can be further classified using the CNN model.

8 Acknowledgment

I would like to express my sincere gratitude to my supervisor Cristina Muntean for mentoring and supporting me in this journey. She guided me in every stage of this tough project which makes it easy and achievable. I would also like to thank my husband and parents for believing and allowing me to achieve my goals.

References

Allaouzi, I., Ahmed, M. B. and Benamrou, B. (2019). An Encoder-Decoder model for visual question answering in the medical domain, *CEUR Workshop Proceedings* **2380**(September 2019): 9–12.

Anghel, A., Stanisavljevic, M., Andani, S., Papandreou, N., Ruschoff, J. H., Wild, P. and Gabrani, M. (2019). A high-performance system for robust stain normalization of whole-slide images in histopathology., *Frontiers in Medicine* .

Azli, E. A., Huddin, A. B., Ibrahim, M. F. and Samad, S. A. (2018). Mass classification in mammograms using neural network, *Proceedings of the 2017 6th International Conference on Electrical Engineering and Informatics: Sustainable Society Through Digital Innovation, ICEEI 2017* **2017-November**: 1–5.

d. Nóbrega, R. V. M., Peixoto, S. A., d. Silva, S. P. P. and Filho, P. P. R. (2018). Lung nodule classification via deep transfer learning in ct lung images, *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 244–249.

Elelimy, E. and Mohamed, A. A. (2018). Towards automatic classification of breast cancer histopathological image, *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pp. 299–306.

Gottapu, R. D. and Dagli, C. H. (2018). Densenet for anatomical brain segmentation., *Procedia Computer Science* **140**(Cyber Physical Systems and Deep Learning Chicago, Illinois November 5-7, 2018): 179 – 185.

Hidalgo-Gavira, N., Mateos, J., Vega, M., Molina, R. and Katsaggelos, A. K. (2018). Blind color deconvolution of histopathological images using a variational bayesian approach, *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 983–987.

Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K. Q. (2017). Densely connected convolutional networks, *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* **2017-January**: 2261–2269.

Jiang, Z., Chen, Y., Zhang, Y., Ge, Y., Yin, F. F. and Ren, L. (2019). Augmentation of CBCT Reconstructed From Under-Sampled Projections Using Deep Learning, *IEEE transactions on medical imaging* **38**(11): 2705–2715.

Kaushal, C., Bhat, S., Koundal, D. and Singla, A. (2019). Recent trends in computer assisted diagnosis (cad) system for breast cancer diagnosis using histopathological images., *IRBM* **40**(4): 211.

Kaushal, C., Koundal, D. and Singla, A. (2019). Comparative analysis of segmentation techniques using histopathological images of breast cancer, *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 261–266.

Khened, M., Kollerathu, V. A. and Krishnamurthi, G. (2019). Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers, *Medical Image Analysis* **51**: 21–45.

Kumar, A., Singh, S. K., Saxena, S., Lakshmanan, K., Sangaiah, A. K., Chauhan, H., Shrivastava, S. and Singh, R. K. (2020). Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer, *Information Sciences* **508**: 405–421.

Li, H., Reichert, M., Lin, K., Tselousov, N., Braren, R., Fu, D., Schmid, R., Li, J., Menze, B. and Shi, K. (2019). Differential diagnosis for pancreatic cysts in ct scans using densely-connected convolutional networks, *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2095–2098.

Li, H., Zhuang, S., Li, D.-a., Zhao, J. and Ma, Y. (2019). Benign and malignant classification of mammogram images based on deep learning., *Biomedical Signal Processing and Control* **51**: 347 – 354.

Li, Y., Xie, X. and Shen, L. (2019). Reverse active learning based atrous DenseNet for pathological image classification, pp. 1–15.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B. and Sánchez, C. I. (2017). A survey on deep learning in medical image analysis., *Medical Image Analysis* **42**: 60 – 88.

Oyama, T. and Yamanaka, T. (2017). Fully convolutional densenet for saliency-map prediction, *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 334–339.

Rundo, L., Tangherloni, A., Nobile, M. S., Militello, C., Besozzi, D., Mauri, G. and Cazzaniga, P. (2019). Medga: A novel evolutionary method for image enhancement in medical imaging systems., *Expert Systems With Applications* **119**: 387 – 399.

Sabeena Beevi, K., Nair, M. S. and Bindu, G. (2019). Automatic mitosis detection in breast histopathology images using convolutional neural network based deep transfer learning., *Biocybernetics and Biomedical Engineering* **39**(1): 214 – 223.

Talo, M. (2019). Automated classification of histopathology images using transfer learning, *Artificial Intelligence in Medicine* **101**: 101743.

Tao Tan, Z. L. (2018). Optimize transfer learning for lung diseases in bronchoscopy using a new concept: Sequential fine-tuning., *IEEE Journal of Translational Engineering in Health and Medicine, Translational Engineering in Health and Medicine, IEEE Journal of, IEEE J. Transl. Eng. Health Med* p. 1.

Tellez, D., Litjens, G., Bándi, P., Bulten, W., Bokhorst, J.-M., Ciompi, F. and van der Laak, J. (2019). Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology., *Medical Image Analysis* **58**.

Tie, J., Jin, J. and Wang, X. (2011). Study on application model of three-tiered architecture, *2011 2nd International Conference on Mechanic Automation and Control Engineering, MACE 2011 - Proceedings* pp. 7715–7718.

Vo, D. M., Nguyen, N.-Q. and Lee, S.-W. (2019). Classification of breast cancer histology images using incremental boosting convolution networks, *Information Sciences* **482**: 123 – 138.

Vogado, L. H., Veras, R. M., Araujo, F. H., Silva, R. R. and Aires, K. R. (2018). Leukemia diagnosis in blood slides using transfer learning in cnns and svm for classification., *Engineering Applications of Artificial Intelligence* **72**: 415 – 422.

Xiang, Z., Ting, Z., Weiyan, F. and Cong, L. (2019). Breast cancer diagnosis from histopathological image based on deep learning, *2019 Chinese Control And Decision Conference (CCDC)*, pp. 4616–4619.

Zhuang, L. and Guan, Y. (2017). Image enhancement via subimage histogram equalization based on mean and variance., *Computational Intelligence and Neuroscience* **2017**.