Deep LearningTechniques For Breast Cancer Risk Prediction Using Ibm Cloud

CONTENTS

1. INTRODUCTION

1.1Overview:

Breast cancer is one of the main causes of cancer death worldwide. Computer-aided diagnosis systems showed the potentialfor improving diagnostic accuracy. But early detection and prevention can significantly reduce the chances of death. It is important to detect breast cancer as early as possible. The goal is to classify images into two classifications of malignant and benign. As early diagnostics significantly increases the chances of correct treatment and survival. In this application, we are helpingthe doctors and patients to classify the Type of Tumour for the specificimage given with the help of Neural Networks.

1.2Purpose:

Breast cancer is the most commonmalignancy among women,accounting for nearly1 in 3 cancers diagnosed among women in the United States, and it is the second leading cause of cancer death among women. Breast Cancer occurs as a results of abnormal growth of cells in the breast tissue, commonly referred to as a Tumor. A tumor does not mean cancer - tumors can be benign (not cancerous), pre-malignant (pre-cancerous), or malignant (cancerous). Tests such as MRI, mammogram, ultrasound and biopsyare commonly used to diagnosebreast cancer performed. By classifying the imagewe can speedup the process of diagnosis.

2. LITERATURE SURVEY

2.1 Existing problem:

S.NO	AUTHOR	DATASET USED	TOOL USED	TECHNIQUE USED	ADVANTAGES	ACCURACY	ERROR RATE
1	Wang et al.	Electronic health records	WEKA	Logistic regression	5-year survivability prediction using logistic regression	96.4 %	0.33
2	Keles, M. Kaya	Wisconsin Diagnostic Breast Cancer dataset	Python	SVM vs KNN, decision trees and Naives bayes	The marginal distance between thedecision hyperplane and the instances that are closest to boundary is maximized	up to 96.91%	0.33
3	Kavithaa et. at	CancerSociety	MATLAB	Ensemble method with Logistic and NeuralNetwork	Multiple Learners are combined givinghigher accuracy.	96.3 %	
4	Shravya atal.	UCI repository	Spyder	SVM	Hyperplane separates two classes which helps in higheraccuracy.	92.7%	
5	Medjahedet. al	Wisconsin breast cancer dataset	WEKA	Decision Trees	Helps in splitting	96.1 %	

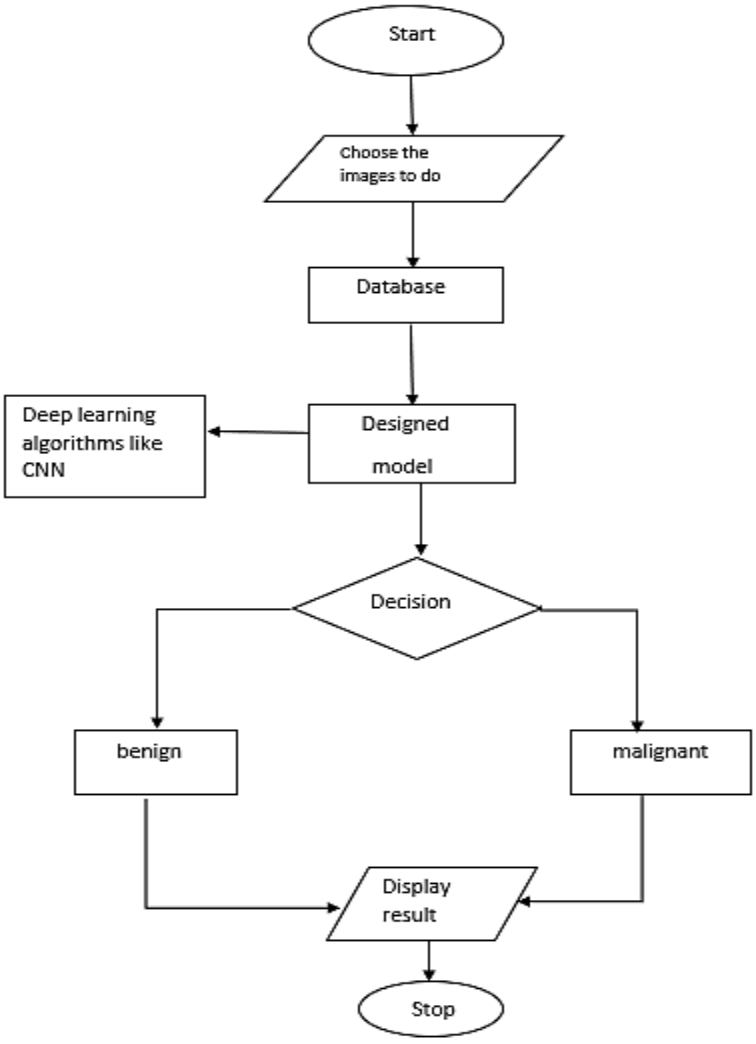
2.2 Proposed solution:

Deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC) and apply sigmoid function to classify an object with probabilistic values between 0 and 1. The model has been developed for two conditions, first when the dataset includes numerical values of images which

are scaled to classifyan image as benign or malignant .The model uses CNN architecture to predict the model accuracy.The second condition is when the model is trained using training set consisiting of images which are benign and malignant. The model's accuracy is determined by deployingit on test set and predicting the accuracy and loss.

3.THEORITICAL ANALYSIS

3.1Block diagram:



3.2 Hardware / Software designing:

Software Requirements:

- **Anaconda Navigator** : Anaconda Navigator is a free and open-source distribution of thePython and R programming languages for data science and machine learning related applications
- **Tensor flow**: TensorFlow is an end-to-end open-source platform for machine learning.
- **Keras**: Kerasleverages various optimization techniques to make high-level neuralnetworkAPI easier and more performant.
- **Flask**: Web framework used for buildingWeb applications.

- **Python libraries:**

- Numpy
- Pandas
- Matplotlib
- opencv-python

- **Jupyter Notebook**: It is a powerfulway to iterate and writeon your Pythoncode fordataanalysis .

Hardware Requirements:

Processor: Intel® Core™ i3-2350M CPU @ 2.30GHzInstalled memory (RAM):4.00GB

System Type: 64-bit Operating System

4.EXPERIMENTAL INVESTIGATIONS

while working on the solutionwe investigated on the what is AL and what is ML and how to buildmodels using them and how to do image processing. And mainly we had studied about the CNN because our solution mainly need this so we worked on these aspects.

Artificial Intelligence: Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems enabling it to even mimic human behaviour. Its applications lie in fields of Computer Vision, Natural Language Processing, Robotics, Speech Recognition, etc.

Basic Operation of Neural Networks: Neural Networks (NN) form the base of deep learning, a subfield of machine learningwhere the algorithms are inspired by the structure of the human brain. NN take in data, train themselves to recognize the patterns in this data and then predict the outputs for a new set of similar data. NN are made up of layers of neurons. These neurons are the core processing units of the network.

Transfer Learning: A major assumption in many machine learning and data mining algorithms is that the trainingand future data must be in the same featurespace and have the same distribution. However, in many real-world applications, this assumption may not hold. For example, we sometimeshave a classification task in one domainof interest, but we only have sufficient training data in another domain of interest, where the latter data may be in a different feature space or follow a different data distribution. In such cases, knowledge transfer,if done successfully, would greatly improvethe performance of learning by avoiding much expensive data labelling efforts.In recent years,transfer learning has emerged as a new learning framework to address this problem.

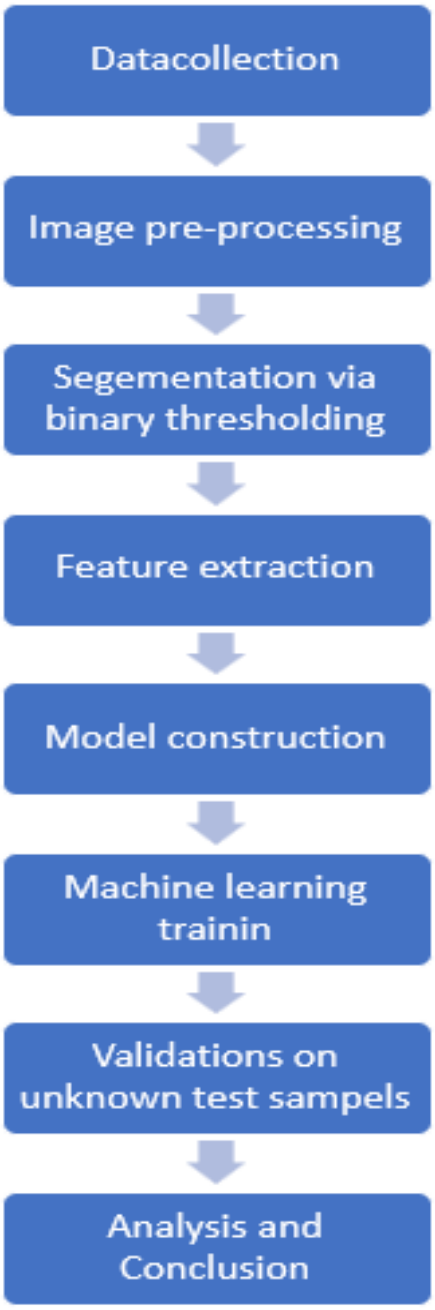
Convolutional Neural Network: Classifier models can be basically divided into two categories respectively whichare generative modelsbased on hand-crafted features and discriminativemodels based on traditional learningsuch as support vector

machine(SVM), Random Forest(RF) and Convolutional Neural Network (CNN). One difficulty with methods based on hand-crafted features is that they often require the computation of a large number of features in order to be accurate when used with many traditional machine learning techniques.This can make them slow to compute and expensive memory-wise. More efficient techniques employ lower numbers of features, using dimensionality reduction like PCA (Principle Component Analysis) or feature selection methods, but the reduction in the number of features is often at the cost of reduced accuracy. Brain tumor segmentation employ discriminative models because unlike generative modelling approaches, these approaches exploitlittle prior knowledge on the brain‘sanatomy and instead rely mostly on the extraction of [a large number of] low level image features, directly modelling the relationship between these featuresand the label of a given voxel.

Activation Function: Sigmoid function rangesfrom 0 to 1 and is used to predictprobability as an output in case of binary classification while Softmax functionis used for multi-class classification. tanh function ranges from -1 to 1 and is considered better than sigmoid in binary classification using feed forward algorithm. ReLU (Rectified Linear Unit) ranges from 0 to infinity and Leaky ReLU (better version of ReLU) ranges- from -infinity to +infinity. ReLU stands for Rectified LinearUnit for a non-linear operation.

The output is $f(x) = \max(0,x)$.ReLU‘s purposeis to introduce non-linearity in our ConvNet.Since, the real world data would want our ConvNet to learn would be non-negative linear values. There are other nonlinear functions such as tanh or sigmoidthat can also be used instead of ReLU. Most of the data scientists use ReLU since performance wise ReLU is better than the other two. Stride is the number of pixels that would move over theinput matrix one at a time.Sometimes filter does not fit perfectly fit the input image. We have two options: either pad the picture with zeros (zero-padding) so that it fits or drop the part of the image where the filter did not fit. This is called valid padding which keeps only valid part of the image

5.FLOW CHART


Datacollection
Image pre-processing
Segementation via binary thresholding
Feature extraction
Model construction
Machine learning trainin
Validations on unknown test sampels
Analysis and Conclusion

6.RESULT

The suggested CNN-based breast cancer screening technique has shown satisfactory results in preliminary testing.There were two different training and testing techniques used. Initially, the data was divided into two categories: benign and malignant tumors.

```
Epoch 1/10
5/5 [==============================] - ETA: 0s - loss: 0.4682 - accuracy: 0.2857WARNING:tensorflow:Your input ran out of data;
interrupting training. Make sure that your dataset or generator can generate at least `steps_per_epoch * epochs` batches (in th
is case, 10 batches). You may need to use the repeat() function when building your dataset.
5/5 [==============================] - 4s 798ms/step - loss: 0.4682 - accuracy: 0.2857 - val_loss: 216.9595 - val_accuracy: 0.3
333
Epoch 2/10
5/5 [==============================] - 2s 357ms/step - loss: 0.4198 - accuracy: 0.2755
Epoch 3/10
5/5 [==============================] - 2s 380ms/step - loss: 0.4315 - accuracy: 0.2818
Epoch 4/10
5/5 [==============================] - 2s 375ms/step - loss: 0.4231 - accuracy: 0.2653
Epoch 5/10
5/5 [==============================] - 2s 411ms/step - loss: 0.4175 - accuracy: 0.2727
Epoch 6/10
5/5 [==============================] - 2s 337ms/step - loss: 0.3976 - accuracy: 0.2755
Epoch 7/10
5/5 [==============================] - 2s 370ms/step - loss: 0.3692 - accuracy: 0.2551
Epoch 8/10
5/5 [==============================] - 2s 394ms/step - loss: 0.3673 - accuracy: 0.2818
Epoch 9/10
5/5 [==============================] - 2s 446ms/step - loss: 0.4311 - accuracy: 0.2857
Epoch 10/10
5/5 [==============================] - 2s 410ms/step - loss: 0.6421 - accuracy: 0.2857
```

7.ADVANTAGES & DISADVANTAGES

Advantages:

- The main advantage of **CNNs** compared to a traditional neural network is that theyautomatically **detect important features** without any human supervision. For example, given any pictures of cats and dogs, it can learn the key features for each class by itself. Besides, the number of parameters learned during training

reduces significantly since all the features have been extracted already.

Disadvantages:

- Classification of Images with differentPositions
- Adversarial examples
- Coordinate Frame
- Otherminor disadvantages like performance

8. APPLICATIONS

- The main application of this model is to preprocess the images to a machine-readableformat. It is well trained so that it will predict the correct data.
- Breast cancer risk prediction modelsin clinical practice

9. CONCLUSION

We created aweb app which predicts type of breast tumouras benign and malignant tumors .

10. FUTURE SCOPE

The analysis of the results signifies that the integration of multidimensional data along with different classification, feature selection and dimensionality reduction techniques can provide auspicious tools for inference in this domain. Further research in this field should be carried out for the better performance of the classification techniques so that it can predict on more variables. We are intending how to parametrize our classification techniques hence to achieve high accuracy. We are looking into many datasets and how further deep Learning algorithms can be used to characterize Breast Cancer. We want to reduce the error rates with maximum accuracy.

11.BIBILOGRAPHY

[1] Wang, D. Zhang and Y. H. Huang “Breast Cancer Prediction Using Machine Learning”

[2] B. Akbugday, "Classification of Breast Cancer Data Using Machine Learning Algorithms," Technologies Congress (TIPTEKNO), Izmir, Turkey, 2019, pp. 1-4.

[3] Keles, M. Kaya, "Breast Cancer Prediction and Detection Using Data Mining Classification A Comparative Study." Tehnicki Vjesnik - Technical Gazette, vol. 26, no. 1, 2019, p. 149+.

[4] V. Chaurasia and S. Pal, “Data Mining Techniques: To Predict and Resolve Breast Cancer Survivability”,

[5]] Delen, D.; Walker, G.; Kadam, A. Predicting breast cancer survivability: A comparison of three data mining methods

[6] R. K. Kavitha1, D. D. Rangasamy, “Breast Cancer Survivability Using Adaptive Voting Ensemble

Machine Learning Algorithm Adaboost and CART Algorithm” Volume 3, Special Issue 1, February 2014

[7] P. Sinthia, R. Devi, S. Gayathri and R. Sivasankari, “Breast Cancer detection using PCPCET and ADEWNN”, CIEEE’ 17, p.63-6

Appendix

CNN code:

```
from keras.preprocessing.image import ImageDataGenerator
 from keras.models import Sequential
 from keras.layers import Dense
 from keras.layers import Convolution2D
 from keras.layers import MaxPooling2D
 from keras.layers import Flatten
```

```
 import tensorflow as tf

history = model.fit(training_set.repeat(),
   steps_per_epoch=int(8000/batch_size),
   epochs=25,
   validation_data=test_set.repeat(),
   validation_steps=int(2000/batch_size))

train_datagen = ImageDataGenerator(rescale = 1./255,shear_range = 0.2,zoom_range =
0.2,horizontal_flip = True)
 test_datagen = ImageDataGenerator(rescale = 1)

x_train = train_datagen.flow_from_directory(r'D:\mlprj\Breast-Cancer-Risk-Prediction-
main\breastcancerdataset\train',target_size = (64,64),batch_size = 32, class_mode = 'binary')
 x_test =  test_datagen.flow_from_directory(r'D:\mlprj\Breast-Cancer-Risk-Prediction-
main\breastcancerdataset\test',target_size = (64,64),batch_size = 32, class_mode = 'binary')

print(x_train.class_indices)

Model = Sequential()
 Model.add(Convolution2D(32,(3,3),input_shape = (64,64,3),activation = 'relu'))

Model.add(MaxPooling2D(pool_size = (2,2)))

Model.add(Flatten())

Model.add(Dense(40 ,kernel_initializer = 'uniform',activation = 'relu'))

Model.add(Dense(1,activation = 'softmax',kernel_initializer ='uniform'))

Model.compile(loss = 'binary_crossentropy',optimizer = "adam",metrics = ["accuracy"])

Model.fit_generator(x_train, steps_per_epoch =5,epochs = 10,validation_data =
x_test,validation_steps = 10)

model.save("breastcancer.h5")
```