An Efficient Transfer Learning Based Cross Model Classification (TLBCM) Technique for Breast Cancer Prediction using Cyber-Physical System

Sudha Prathyusha Jakkaladiki [Corresp., 1], **Filip Maly** [1]

[1] Department of Informatics and Quant., University of Hradec Králové, Hradec Kralove, Hradec Kralove, Czech Republic

Corresponding Author: Sudha Prathyusha Jakkaladiki
Email address: sudha.jakkaladiki@uhk.cz

Breast cancer has been the most life-threatening disease in women in the last few decades. The maximum death among women is due to breast cancer because of less awareness and a minimum number of medical facilities to detect the disease in the early stages. In the recent era, the situation has changed with the help of many technological advancements and medical equipment to see breast cancer development. The machine learning technique supports vector machines (SVM), logistic regression, and random forests have been used to analyze the images of cancer cells on different data sets. Although the particular technique has performed better on the smaller data set, accuracy still needs to catch up in most of the data, which needs to be fairer to apply in the real-time medical environment. In the proposed research, state-of-the-art deep learning techniques, such as Transfer Learning, Based Cross Model classification (TLBCM), Convolution Neural Network (CNN) and transfer learning, Residual network (Resnet), and Densenet proposed for efficient prediction of breast cancer with the minimized error rating. The convolution neural network and transfer learning are the most prominent techniques for predicting the main features in the data set. The sensitive data is protected using a Cyber-Physical System (CPS) while using the images virtually over the network. CPS act as a virtual connection between human and networks. While the data is transferred in the network, it must monitor using CPS. The Resnet changes the data on many layers without compromising the minimum error rate. The Densenet conciliates the problem of vanishing gradient issues. The experiment is carried out on the data sets Breast Cancer Wisconsin (Diagnostic) and Breast Cancer Histopathological Dataset (BreakHis). The convolution neural network and the transfer learning have achieved a validation accuracy of 98.3%. The results of these proposed methods show the highest classification rate between the benign and the malignant data. The proposed method improves the efficiency and speed of classification, which is more convenient for discovering breast cancer in earlier stages than the previously proposed methodologies.

An Efficient Transfer Learning Based Cross Model Classification (TLBCM) Technique for Breast Cancer Prediction using Cyber-Physical Systems

Sudha Prathyusha Jakkaladiki[1] **and Filip maly**[1]

[1]**Department of Informatics and Quant. methods, University of Hradec kralove , Czech Republic**

Corresponding author:
Sudha Prathyusha Jakkaladiki[1]

Email address: Sudha.jakkaladiki@uhk.cz

ABSTRACT

Breast cancer has been the most life-threatening disease in women in the last few decades. The maximum death among women is due to breast cancer because of less awareness and a minimum number of medical facilities to detect the disease in the early stages. In the recent era, the situation has changed with the help of many technological advancements and medical equipment used to detect breast cancer development. The machine learning technique supports vector machines (SVM), logistic regression, and random forests have been used to analyze the images of cancer cells on different data sets. Although the particular technique has performed better on the smaller data set, accuracy still needs to catch up in most of the data, which is not fair enough to apply in the real-time medical environment. In the proposed research, state-of-the-art deep learning techniques, such as Transfer Learning, Based Cross Model classification (TLBCM), Convolution Neural Network (CNN) and transfer learning, Residual network (Resnet), and Densenet proposed for efficient prediction of breast cancer with the minimized error rating. The convolution neural network and transfer learning are the most prominent techniques for predicting the main features in the data set. The sensitive data is protected using a Cyber-Physical System (CPS) while using the images virtually over the network. CPS act as a virtual connection between human and networks. While the data is transferred in the network, it must monitor using CPS. The Resnet changes the data on many layers without compromising the minimum error rate. The Densenet conciliates the problem of vanishing gradient issues. The experiment is carried out on the data sets Breast Cancer Wisconsin (Diagnostic) and Breast Cancer Histopathological Dataset (BreakHis). The convolution neural network and the transfer learning have achieved a validation accuracy of 98.3

INTRODUCTION

In the modern world, although there are much medical equipment and advancements in medicine, nearly two million women are affected by the cruelest disease, breast cancer. Cancer occurs in the human body when there is a cell growth called mutations. Few prominent techniques and methods are available to predict breast cancer in its early stages in middle-aged women. Mutation divides the cell and grows unconditionally in a chaotic way. Moreover, the progress of this mutation will be abnormal. Consequently, it ends up in the formation of tumor cells in the human body. Breast cancer occurs particularly in a female body when there is a malignant called a cancerous cell. As it grows continuously, it may spread to other parts of the human body and could automatically result in death. Early detection of the disease will help the women to survive and decrease the death rate. The classification of mammogram images to detect breast cancer is based on techniques such as fuzzy systems, machine learning, and deep learning. Image processing converts the data obtained from medical devices such as mammograms. Also, they are converted into a digital form of processing. During the conversion of images to digital format, useful information will be extracted from the images for analyzing the data. Specifically, the pixels are the

information about a particular value in a precise location. The image processing is performed in the following steps.

1. The image is obtained from medical devices such as optical scanners or X-rays.
2. Enhancement of the images, including data compression to extract the features.
3. The result will be obtained by analyzing the quality and the classification.

Classification of mammogram images is done by using supervised or unsupervised machine learning algorithms. These algorithms are based on the shape of the cancer cell, which is gentle, malignant, or abnormal.

The computer-aided design detection (CAD) on mammogram images is used to detect the occurrence of breast cancer (33; 25). Although the performance of the techniques is better, it has many disadvantages. Specifically, handcrafted feature extraction is the most tedious process. Compared to the other machine learning technologies, the traditional method involves a lot of handcrafted features. It is very difficult and has no generalized procedure. The solution for the existing traditional handcraft-based feature extraction is the convolutional neural network. (23; 14).

According to previous research, the convolutional neural network performs better than the traditional method for getting features from the images. The classic convolutional neural network alexnet has won the challenge in imagenet, where it contains 1000 classes of colored images, and it has got 83.6

In this proposed research work, the cross-modal deep learning approach is used to predict breast cancer from mammographic images. Preprocessing of images and classification is a part of the model. The images are preprocessed for enhancement and extraction of important information from the images. The image is resized for clarity so as to fit into the model. The model includes Convolutional Neural Network (CNN), transfer learning, Resnet, and Densenet.

The key contribution of our research work is as follows:

1. The features are extracted using the convolution neural network along with the transfer learning technique.
2. The weight and biases of the pre-trained model are fine-tuned automatically to analyze the features of the Mammogram images
3. The Resnet and Densenet models are employed to compare the accuracy
4. The learning rate change and data augmentation apply to avoid the over-fitting of the model.

The remaining parts of the paper are arranged in the following order. In section 2, the literature reviews on convolutional neural network transfer learning and augmentation on the mammographic image dataset are analyzed. In section 3, the dataset and the experimental setup are discussed. In section 4, the results of the proposed model are discussed. In section 5, the conclusion is given.

RELATED WORK

Deep learning is a concept that has been established to effectively extract the pertinent information from the raw images and use it for the classification process in order to overcome the limitations of classic machine learning approaches. (19),(3).

The convolutional neural network can be applied in three major ways; 1. pre-trained model of CNN. 2. training the CNN from the scratch. 3. fine-tuning of CNN (12; 24), has shown numerous machine learning algorithms that are employed for detecting breast cancer on mammographic images. In the study, the databases that are most commonly used are Digital Database for Screening Mammography (DDSM) and MIAS, as proposed by (PUB et al.). M. Heath at. al. used a 10-fold cross-validation technique to test the model on the database. S. Khan (17), Y.-D. Zhang (39) used the automatic feature extraction method with CPS instead of the handcraft feature extraction technique such as fractional Fourier transform and Gabor filter along with the classifiers such as Support Vector Machine (SVM).

(10) has proposed many neural networks and example methods for the classification of breast cancer. (29) had focused on convolutional neural networks. It produced excellent results when compared to the

other neural networks in terms of feature extraction and classification in mammogram images. (13) had proved in his study that the convolutional neural network was superior to the other traditional machine learning techniques for extracting the features along with transfer learning using CPS. Deep learning techniques are implemented in the medical field for obtaining better outcomes, along with tough challenges in the input data compared to the other fields. Transfer learning is used by different physicians and technicians for analyzing various medical images, which gives better results in diagnosis (34; 22)].

(13), had achieved 88 (1; 26) enhanced the study accuracy of breast cancer by utilizing the method of Double Shot Transfer Learning (DSTL) with the help of pre-trained networks. Unlike the other models which use the smallest medical data, the proposed model employs a larger data set as a target data set for fine-tuning. The weight and the bias are fine-tuned from the pre-trained model. The fine-tuning of parameters is achieved with the help of the target dataset.

An approach for Deep Learning (DL) has been proposed by Dongdong Sun et al. dubbed D-SVM for human breast cancer prediction prognosis. The program has successfully discovered hierarchy and created an abstract representation from raw input data combined approach of traditional classification (37). In order to identify the architectural distortion from digital mammography, (30) have examined various deep-learning techniques. The primary objective of their research was to find the abnormalities that are the signs of advanced diseases like masses and micro-calcification. Only 12

(31) proposed a transfer learning approach with the help of three kinds of networks; 1. VGG-f, 2. Caffe, and 3. VGG-m. The output was tested in two cases while fine-tuning the model. In the first case, image normalization was applied to find the abnormal letters in mammogram images. In the second case, the model of no image normalization was applied for testing. In the proposed model, both support vector machines and convolution neural networks are employed. Features are extracted using the convolution neural network, and the classification is handled by the support vector machine. The extracted features by the convolution neural network are passed to the support vector machine for the classification of mammogram images for identifying the possibilities of breast cancer.

It is aimed to develop an end-to-end deep learning framework for multi-label breast lesion detection, which is motivated by the success of Convolutional Neural Networks (CNNs) in single-label mammography classification (15),(2),(21). In order to create an autonomous multi-labeling framework that can aid the radiologist in providing his patients with a thorough report and more accurate diagnoses, we aim to make use of the highly expressive convolutional neural network architecture (CNN) (20),(27),(9),(38),(35).Numerous microwave imaging techniques, including microwave tomography and radar-based imaging, were researched. Some imaging diagnostic techniques were described by Lu et al. [42-46] for the diagnosis of breast cancer. It looked into how computer vision and machine learning could detect breast cancer. On the basis of mammographic images, the effectiveness of various approaches was examined.

(16) proposed a new feature extraction technique, Speed-Up Robust Features (SURFs), which is k-means convergence on a mini dataset. An additional layer is introduced in the classification process. 70% of the data from the dataset is used for training and 30% of the data is used for testing models. Classification support vector machines and deep neural networks are employed. The result got from this model is sure that this deep neural network performs better than the decision tree model. The accuracy is improved by employing the convolution neural network model on many datasets, which is equal to 87.5%.

Deep learning techniques built on Convolution Neuronal Networks (CNN) have recently seen considerable success in the field of biomedical image analysis. Examples include the recognition of mitotic cells in microscopic pictures (28), the detection of tumors (6), the segmentation of neural membranes (7), the detection and classification of immune cells (5), and the quantification of mass in mammograms (11),(4),(8).

PROPOSED TLBCM ARCHITECTURE

The best cancer detection method is shown in figure 1. The deep learning approaches convolution neural networks (CNN), Resnet, and Densenet are used to extract the salient features from the mammogram image dataset. The dataset is processed to identify and classify the images based on benign and malignant.

Preprocessing

In this section, the pre-processing of data is described. Data augmentation is applied to the data set to increase the number of data to avoid over-fitting during training and testing of the model. The main

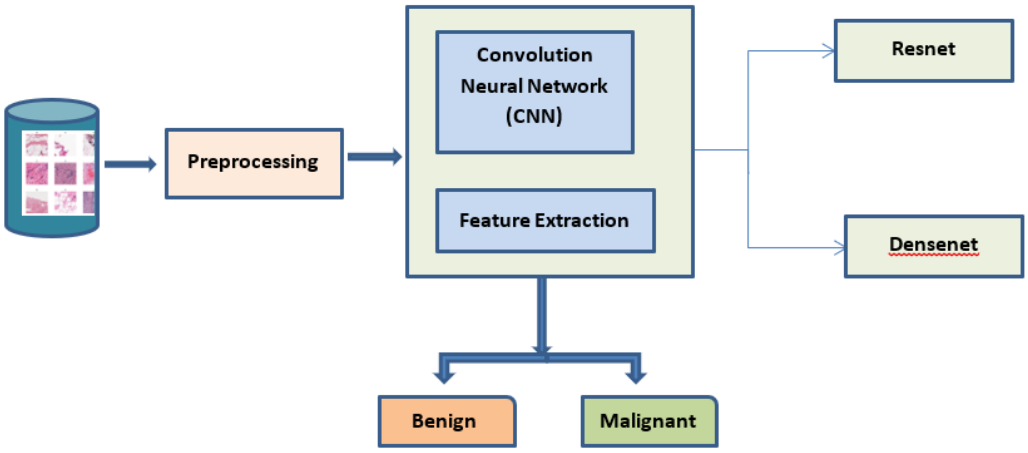


Figure 1. Proposed System

issue with the medical dataset is the minimum number of samples. This dataset can be enlarged with the help of data augmentation and scaling. The missing data cells are replaced with zero. The benign and the malignant data are balanced by introducing a random noise in the data. The 241 malignant data are doubled as 482. This is achieved by the array of random numbers. The standard deviation sigma value is 0.1. The feature values are scaled using the formula

$$\frac{v_f - v_{min}}{v_{max} - v_{min}} \tag{1}$$

v_f is the feature value. v_{min} is the minimum value, v_{max} is the maximum value

The square root transformation is applied to the resultant dataset.

Convolution Neural Network with transfer learning

Transfer learning is the concept of taking the already learned features from other works to apply to the current problem without the need to start from the beginning. Normally, transfer learning is built on the convolution neural network (CNN) on a well-known dataset. The convolution neural network reduces the input and analyzes the features to differentiate it from the other images. It contains several layers that include a convolution layer, max pooling layer, fully connected layer, and batch normalization. The convolution layer in the CNN extracts the features from the given input mammogram images. In this layer, the images are convolved with a kernel or filter. The kernel is multiplied by the match patch. The filter size is matched with the input size and width, and height of the filter according to the network deployment. After the convolution, the input is transformed through subsampling, which can be max pooling, min pooling, or average pooling. The pooling filter is chosen as an odd number. It is responsible for dimensionality reduction, which results in minimizing the overfitting issues. The max-pooling layer in the model is used to minimize the sample in the proposed model. This is used to reduce the complexity of the model when the data is transferred from one layer to another layer. It is used to initiate invariance. Fully connected players are on the top of the model that is connected to each other.

In the proposed method, Resnet 50 is used, which is inherited from the base model Resnet. Figure 2 shows the architecture of Resnet 50. It comprises nearly 50 layers with the residual block. Further, it reduces the computation process and the complexity of the model. Resnet 50 comprises convolution layers, normalization layers, and residual blocks. There are 16 residua block modules in between the pooling layers. The filter size used here is 3×3. This pointer is used to perform the spatial convolution operation for the classification of benign and malignant images in the dataset. The input value x is mapped to the features of output$O(x)$.

$$O(x) = F(x) + x \tag{2}$$

The error from the output is $F(x)$. If the value of $F(x)$ is 0, then the output feature is exactly the same as the target output. Otherwise, it deviates from the target and the weight needs to be adjusted in the

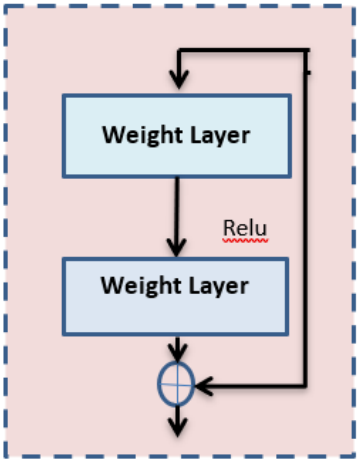


Figure 2. Resnet Architecture

Table 1. CNN training architecture integrated with Transfer learning

Layer	Particulars
Input	RGB image
Convolution layer 1	Conv_3-32 + ReLU
Max pooling Layer 1	MaxPool_2
Convolution layer 2	Conv_3-32 + ReLU
Max pooling Layer 2	MaxPool_2
Convolution layer 3	Conv_3-64+ ReLU
Max pooling Layer 3	MaxPool_2
Fully Connected Layer	FC_64 + ReLU(with Dropout = 0.5)
Output	(Softmax)

hidden layers.

$$y = F\left(x, \{V_j\}\right) + V_j x \tag{3}$$

V_j is the parameter value which implies the weight to be adjusted in the input shape.

Densenet is the inherited idea of Resnet. The Densenet does not contain any constraints for the number of convolution layers and the width of the layers. The parameters are reduced in the Densenet; hence it avoids redundant features with the help of the feature-reused method. Densenet is used to avoid the vanishing gradient problem. The dense net concatenates the feature map from the different layers to reuse the learned features from the previous layers. Let $D_i$is the output of the i^{th}layer.

$$D_i = H(D_{i-1}) \tag{4}$$

Then the dense connection of the layers is termed as,

$$D_i = H(D_0, D_1, D_2 \ldots\ldots, D_{i-1}) \tag{5}$$

The entire architecture of TLBCM architecture is shown in figure 3 The transfer learning is done with the help of Resnet 50 architecture, which is used as the pre-trained model to reduce the time complexity and space complexity. Densenet uses the features extracted from the previous layer for the training. The dropout value of the model is 0.5.

Table 1 Shows the CNN architecture, which is integrated with the pre-trained model. The convolution layer and the fully connected layers are associated with the activation function (Rectified Linear Unit) ReLU. The output layer is associated with the Softmax activation function, which is used to distinguish the malignant and benign data.

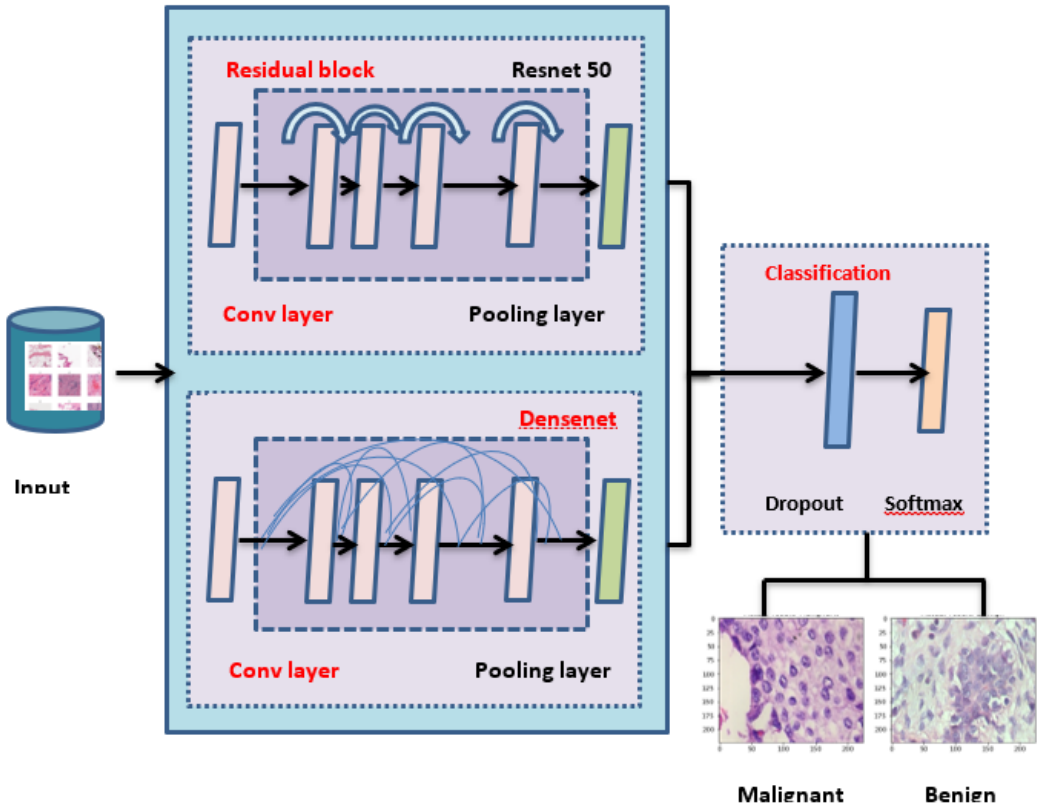


Figure 3. TLBCM architecture

EXPERIMENTAL SETUP

The experiment is carried out on the Python 3.6 environment, Keras as a framework and Tensor Flow as a backend. Two datasets are employed for the proposed research work; Breast Cancer Wisconsin (Diagnostic) and Breast Cancer Histopathological Database (BreakHis).

Dataset

The experiment is carried out on the Breast Cancer Wisconsin (Diagnostic) Data Set, a multivariate dataset that comprises 569 instances with 32 attributes. The dataset comprises 357 malignant and 212 benign data. Figure 4 Shows the data distribution of the Wisconsin database. The dataset comprises 357 malignant and 212 benign data.

Figure 5shows the dataset features' mean value. The mean values are calculated in terms of benign and malignant. It shows the linear pattern between the area, perimeter, and radius.

The dataset comprises ten real-valued features for each cell nucleus. Table 2 Shows the real-valued features and their mean values for sample data. It includes radius, texture, area, perimeter, compactness, smoothness, concave points, concavity, fractal dimensions, and symmetry.

Figure 6 shows the sample of benign and malignant cell structures. The malignant cancer cells are the breast tumors and the benign cell structures are the non-cancer and normal cell structures. The benign are the lumps that are not identified as cancer cells.

Another dataset used in the proposed paper is Breast Cancer Histopathological Database (BreakHis). The dataset comprises microscopic images from 82 different people. Table 3 shows the dataset counts based on the magnification. The magnifying factors are 40X, 100X, 200X, and 400X. The total number of benign in the dataset is 2480. The total number of malignant data is 5429.

The figure 7 shows the data distribution. A large number of data is identified in the 100X magnification. The number of data is 1437. The least data identified in the magnification is 400X. The number of data in this field is 1232.

Figure 6 (36) shows the different magnification samples of malignant cancer. The magnificent factor

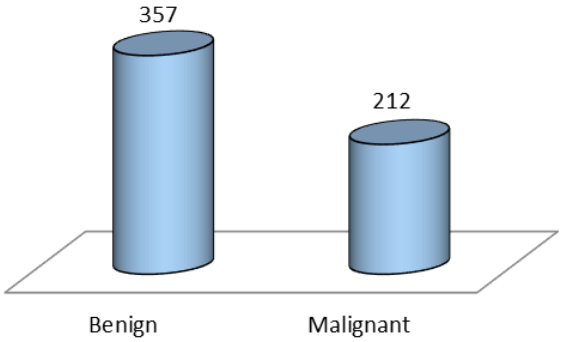


Figure 4. Wisconsin [Diagnostic] data distribution

Table 2. Feature names of sample data

Parameter	Mean Value of Sample 1	Mean Value of Sample 2
mean radius	17.99	20.57
mean texture	10.38	17.77
mean perimeter	122.8	132.9
mean area	1001	1326
mean smoothness	0.1184	0.08474
mean compactness	0.2776	0.07864
mean concavity	0.3001	0.0869
mean concave points	0.1471	0.07017
mean symmetry	0.2419	0.1812
mean fractal dimension	0.07871	0.05667
worst texture	17.33	23.41
worst perimeter	184.6	158.8
worst area	2019	1956
worst smoothness	0.1622	0.1238
worst compactness	0.6656	0.1866
worst concavity	0.7119	0.2416
worst concave points	0.2654	0.186
worst symmetry	0.4601	0.275
worst fractal dimension	0.1189	0.08902

Magnification	Benign	Malignant	Total
40X	652	1,370	1,995
100X	644	1,437	2,081
200X	623	1,390	2,013
400X	588	1,232	1,820
Total of Images	2,480	5,429	7,909

Table 3. Breast Cancer Histopathological Database (BreakHis) dataset counts based on magnification

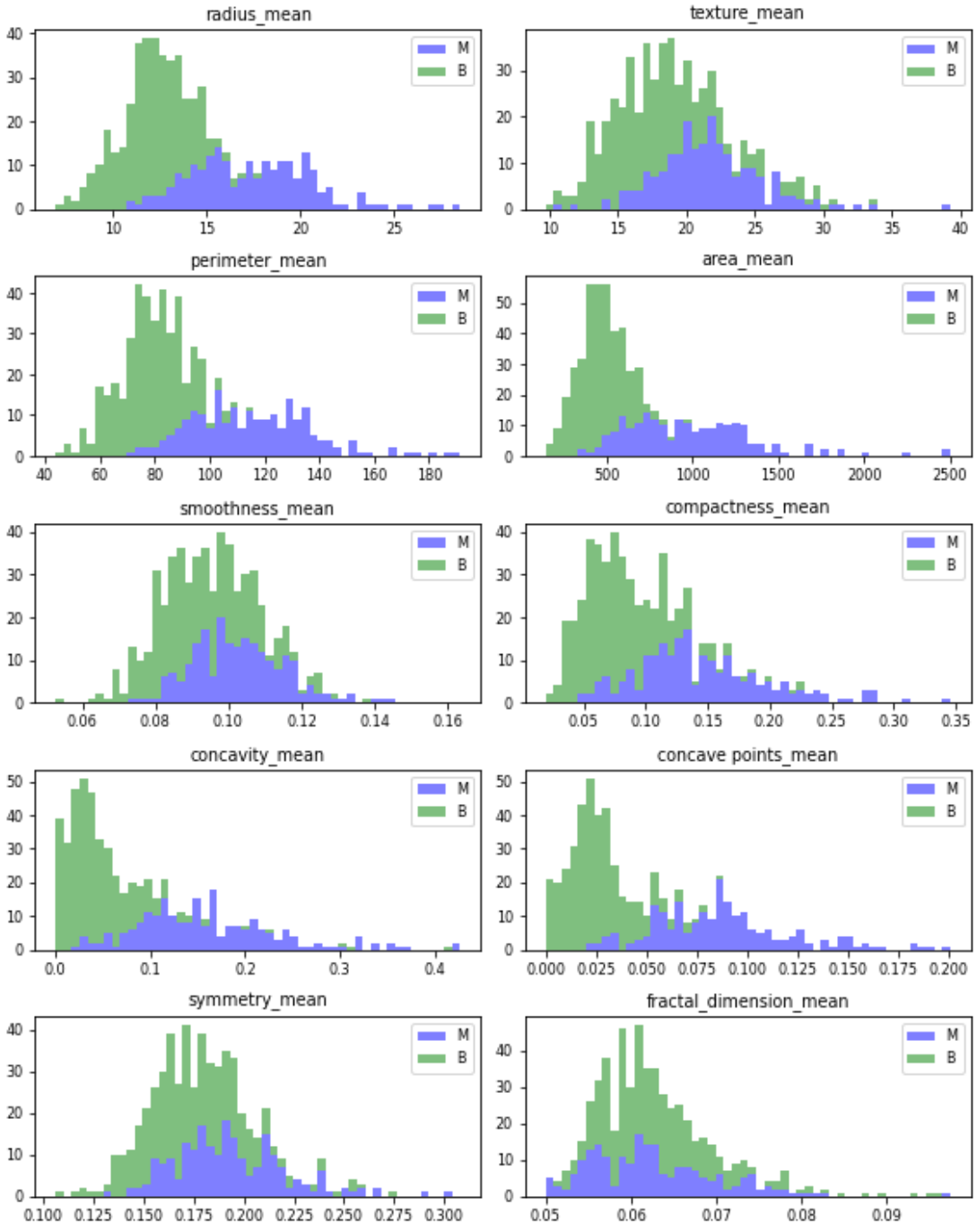


Figure 5. Mean values of Wisconsin's real-valued features

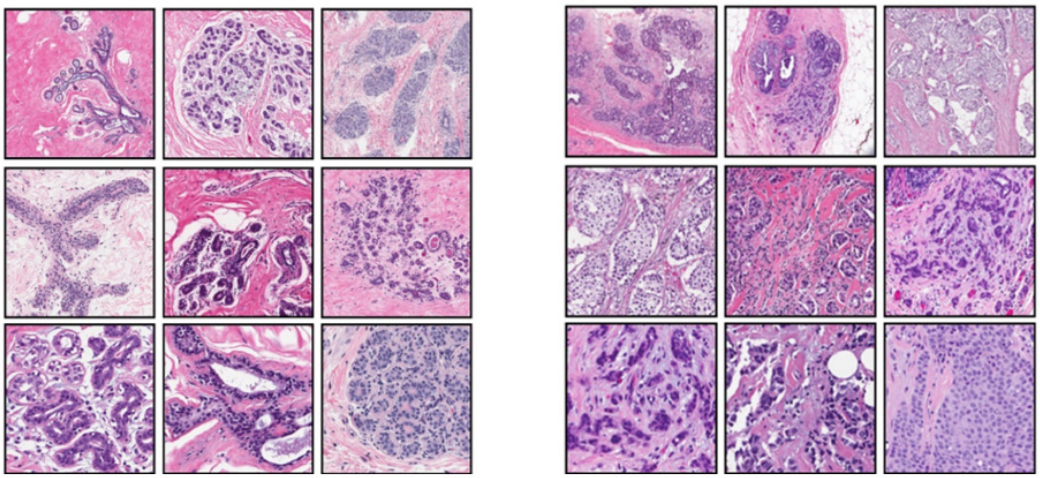


Figure 6. Benign and malignant data sample

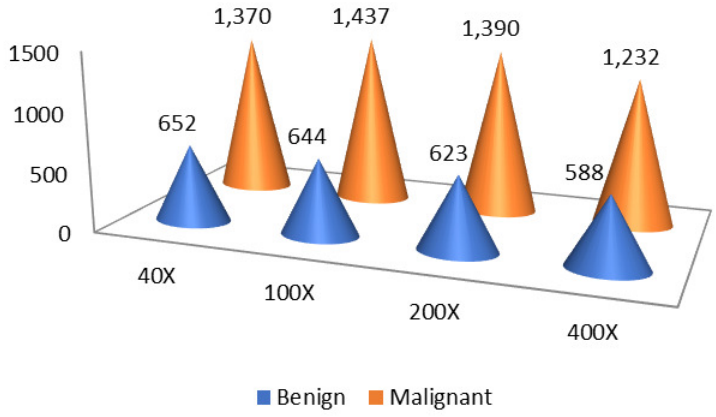
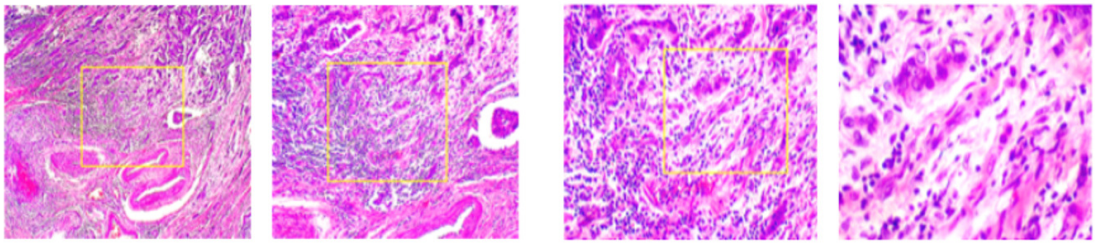


Figure 7. BreakHis data distribution



(a) 40X (b) 100X (c) 200X (d) 400X.

Figure 8. Different Magnification of Malignant Cancer

Table 4. TLBCM model construction Model: $Sequential_1$

Layer (type)	Output Shape	Param #
densenet201 (Model)	(None, 7, 7, 1920)	18321984
obal_average_pooling2d_1	(None, 1920)	0
dropout_1 (Dropout)	(None, 1920)	0
batch_normalization_1 (Batch)	0	0
(None, 1920)	7680	0
dense_1 (Dense)	(None, 2)	3842

400X shows the clear and magnified view of cancer. The enormous size of tissue is removed from the patient under the anesthetic and the tissue is analyzed to capture this magnified view of the cancer cells. The benign cells are always relatively slowed-growing and unharmed. Malignant cancer is fast-growing and has the possibility of spreading to the nearest cells and tissues.

Total params: 18,333,506

Trainable params: 18,100,610

Non-trainable params: 232,896

Table 4 shows the parameter values from the model construction for the Densenet. The total parameters from the input are 18,333,506, and the trainable parameter for the model is 18,100,610.

Hyperparameter Tuning

Adam is the most widely used optimizer for the medical-related data set to train and test. Compared to SGD, Adam produces good accuracy in recognizing the output. When the data set comprises a minimum number of medical data, over-fitting will be a severe issue during model testing and training. The initial learning rate is taken as 0.0001. For the upcoming epoch, the learning rate is reduced to 0.1 to improve the accuracy of the recognition. The model is trained for the 20 epochs. The dropout layer is added to the model to inactivate some neurons during the training process. The dropout layer is added to the Resnet, and then the Densenet model is with a rate of 0.2. For the fully connected layer, the dropout value is 0.5.

RESULT AND DISCUSSION

In this experimental setup, 80

Figure 7 (18) shows the sample predicted and actual results.

The model's performance is assessed using the accuracy recall, precision, F1 score, and accuracy, which are the vital data collection and measurement factors. Both precision and accuracy predict the measurement against the actual output. Accuracy is the measurement of the expected value, and precision measures the reproducible value even if it deviates from the target output.

$$Accuracy = \frac{Total\ number\ of\ correct\ prediction}{total\ number\ of\ prediction} \tag{6}$$

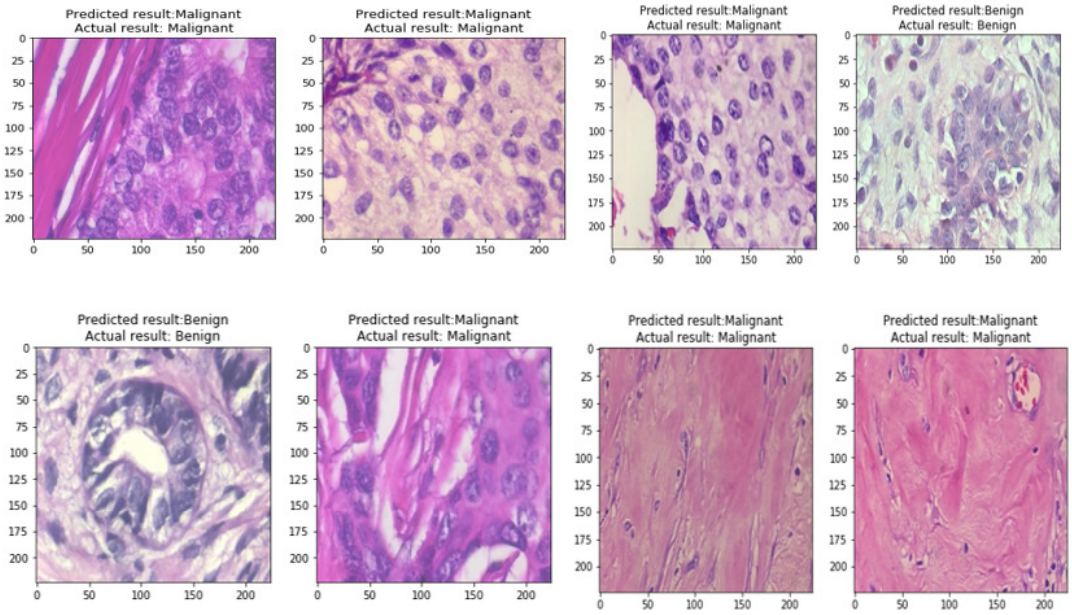


Figure 9. Predicted vs. Actual result

$$Precision = \frac{True\ positive}{All\ positive} \tag{7}$$

The measurement of True positive recalls and the F1 score is the mean value of precision and recall.

$$F1\ Score = \frac{2 * (Recall * Precision)}{(Recall + Precision)} \tag{8}$$

Where ,

$$Recall = \frac{correctly\ predicted\ positive\ value}{Total\ positive\ value} \tag{9}$$

Figure 10 Shows the Receiver Operating Characteristic curve (ROC) to measure the performance of the proposed classification. The final area under the ROC curve (AUC) is predicted as 0.692. Moreover, the proposed model has a positive and a negative rate. The true positive rate is calculated using the formula.

$$T_r = \frac{T_p}{T_p + F_n} \tag{10}$$

The false positive rate is calculated as:

$$F_r = \frac{F_p}{F_p + T_n} \tag{11}$$

T_p is True Positive and $T_n$is True negative. $F_p$is False Positive and $F_n$is False negative. The AUC Is the aggregate measurement of the performance across the all epoch.

Model Evaluation

Figure **??** The confusion matrix is used to analyze the performance of the proposed model. It is a two × two matrix that indicates the target classes. There are two target classes, namely benign and malignant. The predicted values are represented as columns, and the target classes are defined as rows. From the confusion matrix, it is observed that only a few data samples are misclassified as malignant instead of

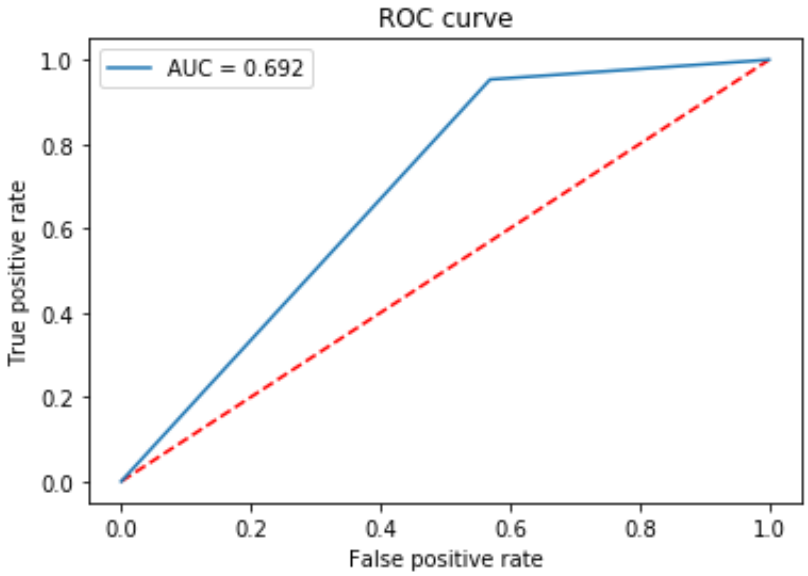


Figure 10. ROC curve

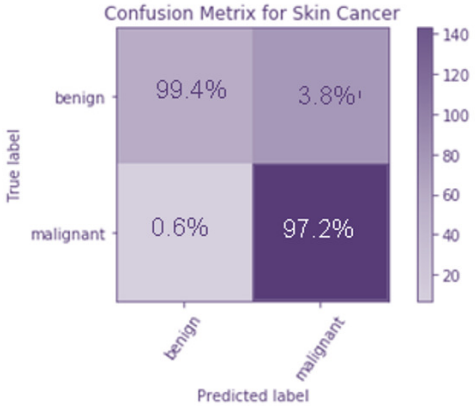


Figure 11. Confusion matrix of TLBCM model

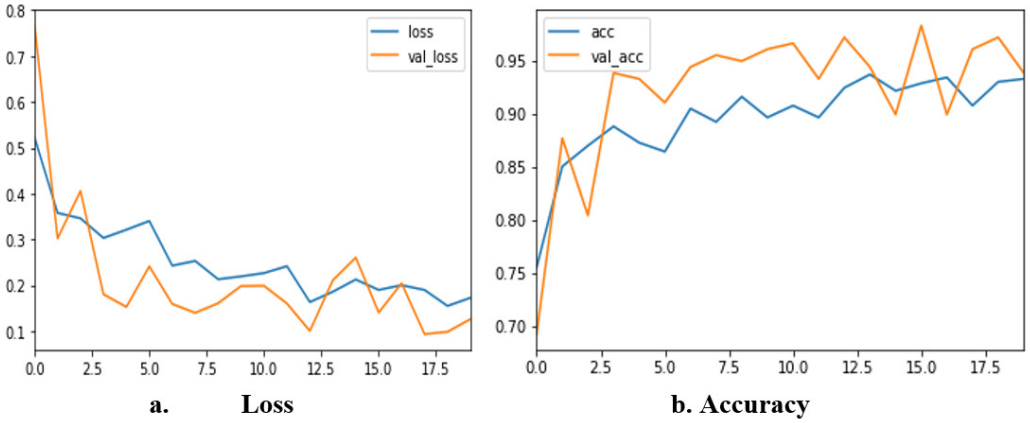


Figure 12. Accuracy and Loss of TLBCM model

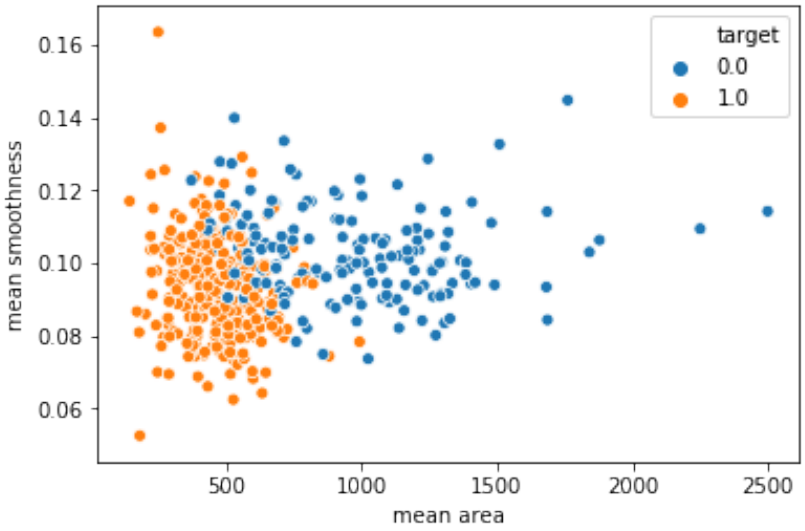


Figure 13. Mean area vs Mean smoothness

benign. Figure 9 shows the confusion matrix of the TLBCM model, which evaluates the value predicted against the target output. 97.2

Figure 12 shows loss and accuracy for all iterations of training and testing. The loss value changes continuously from the beginning of the training to the end. At the end of the training, the value of the loss is nearly 0.1. The accuracy of the training starts from 70% and reaches the highest of 98.3%.

The mean area against the smoothness of the cancer cell can be predicted from Figure **??**. The values produced after the model improvisation. The values correctly predicted in the orange color values as 1 and the wrongly predicted data was indicated as 0 in the blue color.

The figure 13 shows the mean area against the smoothness of the cancer cell. The values produced after the model improvisation. The values correctly predicted in the orange color values as 1 and the wrongly predicted data was indicated as 0 in the blue color.

Comparative Analysis

The proposed model is compared with other recent models implemented by other researchers. Compared to the other studies, two types of datasets are used instead of focusing on one dataset, which increases the learning capacity of the model. The additional added advantage of the proposed work is the pre-trained classifier. In addition, the model has a fully connected layer that could be integrated into the pre-trained

Model	Accuracy	Precision	Recall	F1 score
TLBCM	98.3%	0.65	0.95	0.77
CNN+LSTM	82.5%	0.62	0.85	0.70
ResNet + LSTM	90.1%	0.59	0.92	0.75
SVM	87%	0.95	0.95	0.95
Decision Tree	82%	0.90	0.90	0.91

Table 5. Proposed Model Comparison with previous research methods

model as a Convolution Neural Network (CNN).

The table 5 Shows the accuracy value of the proposed model against the previous models. The model CNN+LSTM achieves only 82.5

CONCLUSION AND DISCUSSION

Many researchers use the machine learning algorithms KNN, SVM, and Decision trees to detect breast cancer from different data sets. This image feature data is monitored continuously using CPS in the network. Although these methods perform well, the handcrafted extracted features make the model complex, and cancer detection is also delayed. This paper proposes the method TLBCM, which combines Resnet, Dense net, and CNN network to effectively classify breast cancer to predict whether it is malignant or benign. The Resnet and Densenet are used as transfer learning for identifying the dataset's features. The experiment has demonstrated the benefits of transfer learning for the job of interest, even when it is between unrelated tasks. To begin the learning process, pre-trained weights should be used as initialization. The loaded weights must then be gradually fine-tuned to train the new datasets to the network. Backpropagation on the layers with a low learning rate is resumed to accomplish this. The experimental results show that TLBCM with CPS performs better than the existing method. The dataset Breast Cancer Wisconsin (Diagnostic) and Breast Cancer Histopathological Database (BreakHis) are used in the proposed paper. The proposed work has achieved 98.3Several tips are offered to improve the creation and training of a neural network model based on the proposed experiments. It is demonstrated that neural networks outperform state-of-the-art techniques for huge training set sizes. Similarly, it is shown that transfer learning is feasible for a short training set and significantly enhances the model's performance. The effects of the critical hyper-parameters are assessed, and the performance of neural networks against the current best practices is also evaluated. Additionally, transfer learning strategies' efficacy is examined in various experimental configurations. The security of the health or patient health data is protected sensitively using CPS. It results in approaching the earlier therapy and reducing the death rate of women due to breast cancer. Future work can be extended to localize the breast tissues with many datasets. Future studies will concentrate on finding the strategies to use the unique characteristics of the multi-labeled problem to develop a joint image-label embedding that characterizes both the semantics label dependency and the relevance of the image-label relationship. To construct an effective, reliable, and potent computer-aided diagnosis system for early breast cancer diagnosis, it is also better to include imaging modalities other than mammography in the learning process. As a result, it will benefit from additional rich representations that already exist.

CONFLICT OF INTEREST

There are no competing interests were present.

ACKNOWLEDGEMENT

The study was supported by the SPEV project, run by the faculty of informatics and management, at the University of Hradec Kralove, Czech Republic.

REFERENCES

[1] Alkhaleefah, M., Ma, S.-C., Chang, Y.-L., Huang, B., Chittem, P. K., and Achhannagari, V. P. (2020). Double-shot transfer learning for breast cancer classification from x-ray images. *Applied Sciences*, 10(11):3999.

[2] Andreadis, I. I., Spyrou, G. M., and Nikita, K. S. (2014). scheme for mammography empowered with topological information from clustered microcalcifications' atlases. *IEEE Journal of Biomedical and Health Informatics*, 19(1):166–173.

[3] Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828.

[4] Chen, T. and Chefd'Hotel, C. (2014). Deep learning based automatic immune cell detection for immunohistochemistry images. In *International workshop on machine learning in medical imaging*, pages 17–24. Springer.

[5] Ciresan, D., Giusti, A., Gambardella, L., and Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. *Advances in neural information processing systems*, 25.

[6] Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., and Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *Medical Imaging 2014: Digital Pathology*, volume 9041, page 904103. SPIE.

[7] Cruz-Roa, A. A., Arevalo Ovalle, J. E., Madabhushi, A., and González Osorio, F. A. (2013). A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In *International conference on medical image computing and computer-assisted intervention*, pages 403–410. Springer.

[8] Dhungel, N., Carneiro, G., and Bradley, A. P. (2015). Deep learning and structured prediction for the segmentation of mass in mammograms. In *International Conference on Medical image computing and computer-assisted intervention*, pages 605–612. Springer.

[9] Ekici, S. and Jawzal, H. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. *Medical hypotheses*, 137:109542.

[10] ElOuassif, B., Idri, A., Hosni, M., and Abran, A. (2021). Classification techniques in breast cancer diagnosis: a systematic literature review. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 9(1):50–77.

[11] Esteva, A., Kuprel, B., and Thrun, S. (2015). Deep networks for early stage skin disease and skin cancer classification. project report.

[12] Ganesan, K., Acharya, U. R., Chua, C. K., Min, L. C., Abraham, K. T., and Ng, K.-H. (2012). Computer-aided breast cancer detection using mammograms: a review. *IEEE Reviews in biomedical engineering*, 6:77–98.

[13] Hussain, Z., Gimenez, F., Yi, D., and Rubin, D. (2017). Differential data augmentation techniques for medical imaging classification tasks. In *AMIA annual symposium proceedings*, volume 2017, page 979. American Medical Informatics Association.

[14] Jamieson, A. R., Drukker, K., and Giger, M. L. (2012). Breast image feature learning with adaptive deconvolutional networks. In *Medical Imaging 2012: Computer-Aided Diagnosis*, volume 8315, pages 64–76. SPIE.

[15] Karahaliou, A. N., Boniatis, I. S., Skiadopoulos, S. G., Sakellaropoulos, F. N., Arikidis, N. S., Likaki, E. A., Panayiotakis, G. S., and Costaridou, L. I. (2008). Breast cancer diagnosis: analyzing texture of tissue surrounding microcalcifications. *IEEE transactions on information technology in biomedicine*, 12(6):731–738.

[16] Kaur, P., Singh, G., and Kaur, P. (2019). Intellectual detection and validation of automated mammogram breast cancer images by multi-class svm using deep learning classification. *Informatics in Medicine Unlocked*, 16:100151.

[17] Khan, S., Hussain, M., Aboalsamh, H., and Bebis, G. (2017). A comparison of different gabor feature extraction approaches for mass classification in mammography. *Multimedia Tools and Applications*, 76(1):33–57.

[18] Kwon, A., Chae, I. H., You, E., Kim, S. H., Ahn, S.-y., Lee, O.-J., Park, Z.-Y., Rhee, S., Huh, Y. H., and Song, W. K. (2020). Extra domain a-containing fibronectin expression in spin90-deficient fibroblasts mediates cancer-stroma interaction and promotes breast cancer progression. *Journal of Cellular Physiology*, 235(5):4494–4507.

[19] LeCun, Y., Bengio, Y., Hinton, G., et al. (2015). Deep learning. nature, 521 (7553), 436-444. *Google Scholar Google Scholar Cross Ref Cross Ref*.

[20] LeCun, Y., Kavukcuoglu, K., and Farabet, C. (2010). Convolutional networks and applications in

vision. In *Proceedings of 2010 IEEE international symposium on circuits and systems*, pages 253–256. IEEE.

[21] Li, H., Meng, X., Wang, T., Tang, Y., and Yin, Y. (2017). Breast masses in mammography classification with local contour features. *Biomedical engineering online*, 16(1):1–12.

[22] 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.

[23] Lo, S.-C. B., Chan, H.-P., Lin, J.-S., Li, H., Freedman, M. T., and Mun, S. K. (1995). Artificial convolution neural network for medical image pattern recognition. *Neural networks*, 8(7-8):1201–1214.

[24] Maheswaran, S., Sathesh, S., Gayathri, M., Bhaarathei, E., and Kavin, D. (2020a). Design and development of chemical free green embedded weeder for row based crops. *Journal of Green Engineering*, 10(5):2103–2120.

[25] Maheswaran, S., Sathesh, S., Priyadharshini, P., and Vivek, B. (2017). Identification of artificially ripened fruits using smart phones. In *2017 international conference on intelligent computing and control (I2C2)*, pages 1–6. IEEE.

[26] Maheswaran, S., Vivek, B., Sivaranjani, P., Sathesh, S., and Pon Vignesh, K. (2020b). Development of machine learning based grain classification and sorting with machine vision approach for eco-friendly environment. *Journal of Green Engineering*, 10(3):526–543.

[27] Mahrooghy, M., Ashraf, A. B., Daye, D., McDonald, E. S., Rosen, M., Mies, C., Feldman, M., and Kontos, D. (2015). Pharmacokinetic tumor heterogeneity as a prognostic biomarker for classifying breast cancer recurrence risk. *IEEE Transactions on Biomedical Engineering*, 62(6):1585–1594.

[28] Malon, C. D. and Cosatto, E. (2013). Classification of mitotic figures with convolutional neural networks and seeded blob features. *Journal of pathology informatics*, 4(1):9.

[29] Nassif, A. B., Talib, M. A., Nasir, Q., Afadar, Y., and Elgendy, O. (2022). Breast cancer detection using artificial intelligence techniques: A systematic literature review. *Artificial Intelligence in Medicine*, page 102276.

[30] Oyelade, O. N. and Ezugwu, A. E.-S. (2020). A state-of-the-art survey on deep learning methods for detection of architectural distortion from digital mammography. *IEEE Access*, 8:148644–148676.

[31] Perre, A. C., Alexandre, L. A., and Freire, L. C. (2018). Lesion classification in mammograms using convolutional neural networks and transfer learning. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*.

[PUB et al.] PUB, M. H., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, P. The digital database for screening mammography. In *Proceedings of the Fifth International Workshop on Digital Mammography*, pages 212–218.

[33] Rao, V. M., Levin, D. C., Parker, L., Cavanaugh, B., Frangos, A. J., and Sunshine, J. H. (2010). How widely is computer-aided detection used in screening and diagnostic mammography? *Journal of the American College of Radiology*, 7(10):802–805.

[34] Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221.

[35] Shu, X., Zhang, L., Wang, Z., Lv, Q., and Yi, Z. (2020). Deep neural networks with region-based pooling structures for mammographic image classification. *IEEE transactions on medical imaging*, 39(6):2246–2255.

[36] Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering*, 63(7):1455–1462.

[37] Sun, D., Wang, M., Feng, H., and Li, A. (2017). Prognosis prediction of human breast cancer by integrating deep neural network and support vector machine: supervised feature extraction and classification for breast cancer prognosis prediction. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE.

[38] Wang, Z., Zhang, L., Shu, X., Lv, Q., and Yi, Z. (2020). An end-to-end mammogram diagnosis: A new multi-instance and multiscale method based on single-image feature. *IEEE Transactions on Cognitive and Developmental Systems*, 13(3):535–545.

[39] Zhang, Y.-D., Wang, S.-H., Liu, G., and Yang, J. (2016). Computer-aided diagnosis of abnormal breasts in mammogram images by weighted-type fractional fourier transform. *Advances in Mechanical Engineering*, 8(2):1687814016634243.