Breast Cancer Image Classification using Transfer Learning and Convolutional Neural Network

Tripti Sharma
IT Department,
Maharaja Surajmal Institute of Technology, New Delhi, India
E-mail: tripti_sharma@msit.in

Rajit Nair*
CSE Department,
Vellore Institute of Technology, Bhopal, India
E-mail: rajitnitbpl@gmail.com

S. Gomathi
Training and Development,
Atna Technologies India Pvt. Ltd., Coimbatore, India
E-mail: gms@atnatechnologies.com

Corresponding author: Rajit Nair

Abstract
Breast cancer is the world's second most frequent cancer among women. In 2012, new cancer cases made up 12% of all new cases, with female malignancies accounting for 25% of all cancer diagnoses. These cells can be detected by an x-ray or a bump on the body. To be cancerous, a tumour must have cells that have invaded or spread to other areas of the body. Develop an algorithm that can determine whether or not a patient has breast cancer based on biopsy photos. To protect human life, the algorithm must be extremely precise. In this study, a database of breast cancer photos is used for analysis, and the categorization is done using a deep learning approach. The deep learning model is applied by implementing a Convolutional Neural Network with transfer learning. The accuracy has achieved more than 96%, which is better than other states of the art algorithms.
Keywords- Densenet; pre-processing; biopsy; maxpooling; random forest; fully connected.

1. Introduction

The medical imaging community is becoming increasingly interested in applying machine learning, particularly deep learning, to increase cancer screening accuracy (Zhe et al., 2020). Breast cancer is the second-leading cause of cancer mortality among women in the United States, and screening mammograms have been shown to reduce mortality (Yao and Liu, 1999). Unfortunately, a high rate of false positives and false negatives exist with routine screening mammography, despite its advantages. Digital screening mammography in the United States has an average sensitivity of 86.9 percent and an average specificity of 88.9 percent (Nair et al., 2021). There have been CAD software4 in use since the 1990s to assist radiologists in improving the accuracy of screening mammograms, which are used to screen for breast cancer. A decade after

they were first released, evidence showed that early commercial CAD systems had not led to considerable improvements in performance. Several research groups are improving deep learning technology to help radiologists with mammography screening and accuracy (Panta et al., 2020). According to the study, deep learning-based CAD systems perform as well as radiologists in stand-alone mode and even increase radiologists' performance in support mode.

It is challenging to detect subclinical breast cancer on screening mammograms since the tumours only occupy a small percentage of the image. When it comes to full-field digital mammography (FFDM), the normal image size is 4000 x 3000 pixels, whereas the probable tumour area of interest (ROI) is typically only a few hundred square pixels (Khushi et al., 2020). The classification of annotated lesions has thus been the focus of several. Despite the requirement for manually labeled ROI detection, an automated software solution must process the complete mammography to provide extra information beyond the diagnosed lesions and improve clinical interpretations. Traditional methods such as the region-based convolutional neural network (R-CNN), its variations, and others would be simple to employ if ROI annotations were widely available in mammography datasets (He et al., 2017).

Annotations of ROI may not be possible in large mammography databases due to this limitation on approaches that rely on them. Some of the public mammography databases are not fully annotated. Annotated mammograms were used to train neural networks in a few researches, and the results were mixed (Vaishnav et al., 2021). Finally, it's impossible to know if these networks were able to locate the clinically significant lesions and base their predictions on the matching parts of the mammograms. A large training dataset is necessary for deep learning to be most effective. Both fully annotated datasets and larger datasets that are just labeled with cancer status are needed to improve the accuracy of breast cancer classification systems (Abdel-Zaher and Eldeib, 2016).

Pre-training to train a classifier in the absence of ideal broad and complete training datasets is a promising method (Soni and Singh, 2021). This pre-training technique was used by Hinton et al. to determine the weight parameters of their DBN, which has three hidden layers, and then to fine-tune it for classification purposes afterward. Researchers found that applying pre-training for handwritten digit identification increased both speed and accuracy (Gomathi et al., 2021). Deep learning models can be fine-tuned after training on a large database like ImageNet (Krizhevsky et al., 2012). Weight parameters for primitive characteristics like edges, corners, and textures are already established in the model and can be used for a new assignment. This can save a lot of time in the modeling process while improving results (Kumar and Dhiman, 2021).

An "end to end" technique is proposed in this paper, in which a model for identifying local picture patches is pre-trained using an annotated dataset with ROI information (Soni et al., 2021). These weight values can then be used to build a complete image classifier, which can be fine-tuned using datasets without ROI annotations to improve accuracy. Patch and full image classifiers were created using a huge public FFDM database, then moved to a smaller database with hundreds of images (Asbeutah et al., 2020). We used a variety of network configurations to find the most effective patch and full picture classifiers. The pipeline for building a complete image classifier and the advantages and disadvantages of various training procedures is discussed here.

The structure of the paper follows these guidelines: There is an introduction in Section 1, followed by a literature review in Section 2. Section 3 provides an explanation of the proposed

process, and Section 4 shows the resulting data. Finally, the conclusion is stated, followed by the citations.

2. Literature Review

For microscopic biopsy image categorization, (Zhang et al., 2013) developed a new cascade random subspace ensemble technique with rejection choices in 2012. This classification approach uses two random subspace classifier ensembles. K binary classification difficulties (K = 3) are solved using support vector machines in the first ensemble. Rejected samples from the first multi-layer perceptron ensemble focus on the second one. There were 361 photographs of normal tissue, including 119 images of healthy tissue, 102 images of cancer in situ, and 140 images of lobular and/or invasive cancer. It was decided that 20% of each class's photographs would be utilized for testing, and the remainder would be used for instruction. It achieved 99.25 percent classification accuracy and 97.65 percent classification dependability with only a 1.94 percent rejection rate. Four clustering methods were employed to conduct nuclei segmentation on 500 pictures from 50 breast cancer patients in 2013 (Singh et al., 2021).

These images were classified into benign and malignant tumours using three distinct methodologies. For each of the 500 photos, there were 10 photographs for each benign and 25 images for each cancerous instance. Using a 50-fold cross-validation technique, they could attain classification accuracy of 96 to 100%. Traditional machine learning approaches such as KNN (K-next neighbour with K = 5), NB (naive Bayes classifier using kernel density estimate) and DT (decision tree) were utilized to create classifiers for the 25 nucleus features of the biopsies. Tested on 737 microscopic pictures from fine needle biopsies taken from 67 individuals, 25 of which were benign and 42 malignant. Up to 98.51 percent has been reported as the best success rate.

Breast cancer diagnosis can be made utilizing nuclear segmentation of cytological images (Singh et al., 2021). Multilayer perceptron's using the back propagation technique, PNNs, LVQs, and SVMs were all used in the classification process. George et al. list each model's parameters. Using 10-fold cross-validation, the classification accuracy is 76% to 94% using only 92 photos, 45 of which are benign tumours and 47 of which are malignant tumours. Using the Wisconsin Breast Cancer dataset, 699 instances, researchers conducted a performance comparison of four machine learning algorithms, including SVM, DT, NB, and KNN, in 2016. (including 458 benign and 241 malignant cases). According to the experiments, SVM had the maximum accuracy of 97.13 percent using a 10-fold cross-validation.

Researchers have demonstrated a DenseNet-based breast cancer classification algorithm. Results from the BreaKHis experiment were 95.4 percent correct. BreaKHis was then used to perform diagnosis of benign and malignant tumours, as well as multi-class classification of various subtypes of histological images of breast cancer, using the pre-trained model of ResNet. A 96.4–98.7 percent accuracy rate was achieved for binary and multi-class categorization. Although BreaKHis has 7,909 histological pictures from 82 individuals, they are insufficient for utilising deep learning methods. Since the BreaKHis dataset only contains 7,909 histopathological images, we hypothesized that deep learning and transfer learning might be used together to detect breast cancer.

In the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC), the Inception V3 network achieved 78.0–93.9 percent accuracy on top-1 and top-5 criteria, respectively. A network called Inception ResNet V2 achieved the same results. Deep learning is the best option for

detecting and segmenting breast cancer. Mammogram images can be classified and categorized using deep convolutional neural networks without having to train the complete image set for detection in mammography. Using VGGNet architecture, the researchers achieved an accuracy of 92.53 per cent. Cropped photos were used to train the network, then evaluated on the original images. Grad-CAM is then applied. Preoperative tumour reduction is the goal of neoadjuvant chemotherapy (NAC). It is possible to reduce NAC toxicity and it's time for effective treatment by predicting NAC response. Grad-CAM heatmaps for post-hoc analysis were constructed using DCE-MRI images with a CNN. A range of.69 to.88 is stated as the accuracy.

3. Proposed Methodology

Using a breast cancer picture dataset, a CNN-based model is applied to the dataset for supervised classification in this study. Firstly, pre-processing technique will be used to process the dataset and later on classification will be done. Given below Figure 1 has shown the flowchart of the proposed work.
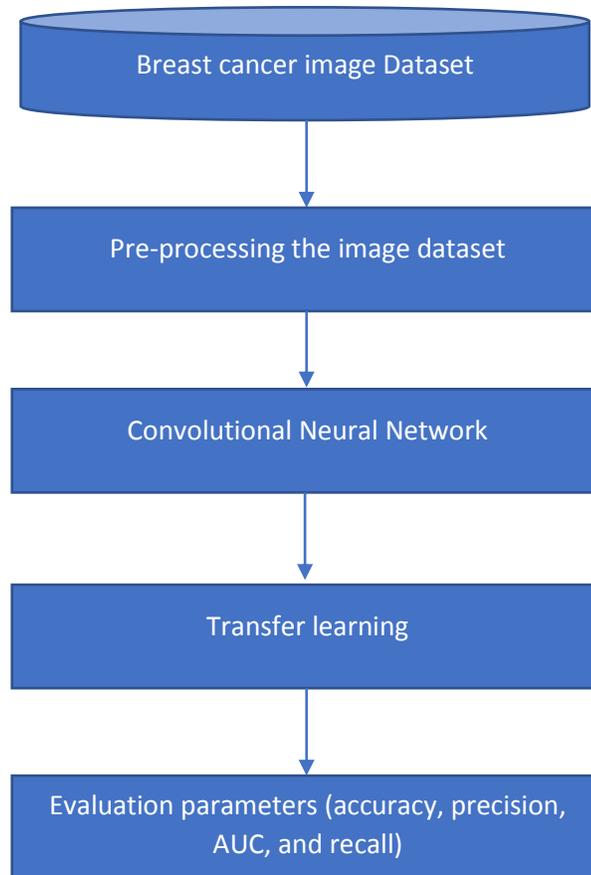

Figure 1: Flowchart of the proposed model.

3.1. Datasets

The Breast Cancer Histopathological Image Classification was created using 9,109 microscopic photos of breast tumour tissue taken from 82 different women. There are 2,480 normal samples in the database and 5,429 malignant samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG format).

3.2. Experimentation and Technology

(i) Using biopsy images, develop an algorithm that can tell if a patient has breast cancer or not. To protect human life, the algorithm has to be exceedingly precise (Gupta et al., 2022). There are 1000 photographs in each category in the training folder, but only 250 in the validation folder. Given below Figure 2 has shown the breast cancer images.
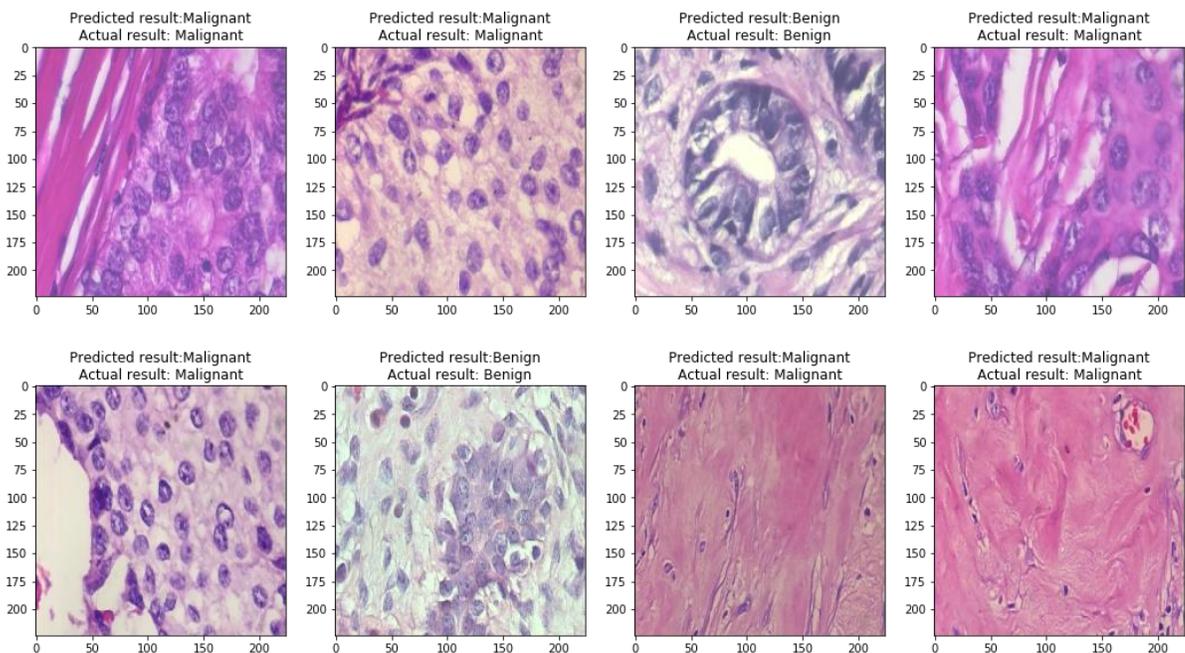

Figure 2: Cancer image dataset snapshot.

(ii) Convolution: Feature maps are needed for this layer's purpose. We often begin with a small number of filters for low-level feature detection. However, detecting high-level characteristics becomes more accessible as we move deeper into CNN. To generate a feature map, feature detection uses a filter of a specific size to scan the input.

(iii) Pooling of Co-evolutionary Operation: Because of the spatial variance provided by this layer, the system can recognize objects even when their appearance changes. For example, to get [16x16x12] for pooling size=16, a down sampling operation is applied to the spatial dimensions (width and height) (2, 2). Figure 3 has represented the architecture of convolutional layers used in this implemented work
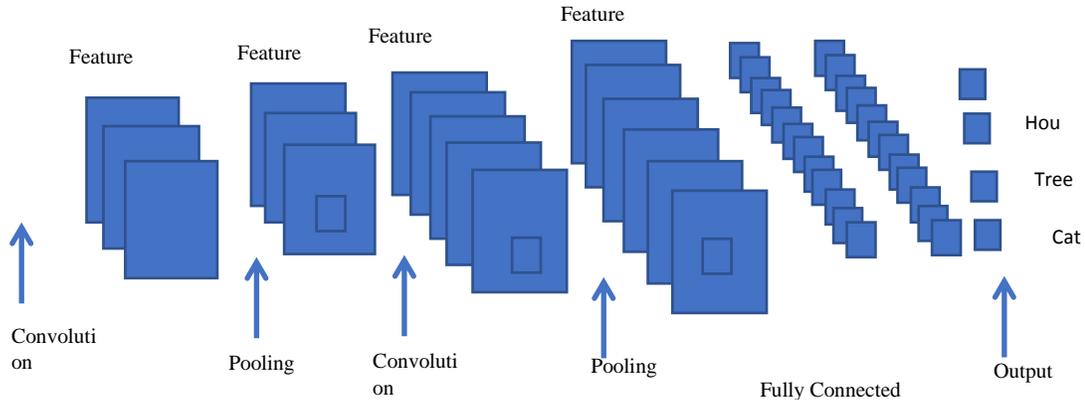
Figure 3: Convolutional architecture.

(iv) Fully Connected Pooling Operation: Every node in the current layer is connected to every node in the next layer in a completely linked layer. As in conventional Neural Networks, all neurons in the previous layer have complete connections to all activations in the current layer. Figure 4 has shown the fully connected pooling layer.
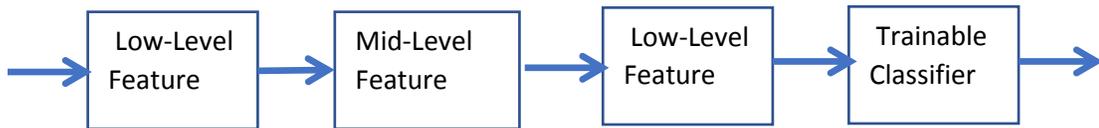

Figure 4: Fully connected pooling layer.

3.3. Classification of Images by CNN

Image classification can be summarized in this manner: A training dataset of N photos, each labeled with two possible classes, serves as our input. Using this training set, a classifier may then be trained to recognize each of these classes. Classifier quality can be assessed by testing it on new images that the classifier has never seen before. These photographs will be compared to their anticipated labels, and we'll see if they match up.

3.4. Transfer Learning

DenseNet model is used as a transfer learning approach to train the model. Pre-trained weights are incorporated with 0.0001 was chosen as the learning rate. The global average pooling layer followed by 50 percent dropouts is used to avoid over-fitting. With softmax as the activation function, batch normalization is integrated and a dense layer with two neurons for two output classes, benign and malignant. The optimizer is Adam, and the loss function is binary-cross-entropy.

4. Experimental Results

Python is the programming language of choice for this project. In this study, a transfer learning approach is used to create a convolutional neural network. The model's accuracy, precision, recall, and AUC score are evaluated using these metrics. According on the results provided in Table 1, the proposed convolutional neural network outperforms existing algorithms in comparison.

Batch sizes ranged from 16 to 32. Batch size is one of the most important hyperparameters in deep learning. The parallelism of GPUs will be leveraged to speed up the training of the suggested models in big batches. It is widely accepted that large batch sizes lead to inaccurate generalizations. However, adopting a batch size equal to the full dataset assures that the objective function's global optimality is reached. However, this comes at the cost of a slower convergence to the ideal solution.

In contrast, smaller batch sizes have been shown to yield better results more quickly. This can be explained by the fact that smaller batch sizes allow the model to start learning before it has seen all of the data. Smaller batches have the drawback that the model's convergence to the global optimum isn't always guaranteed. There are numerous studies that have shown that a small batch size is better than a large batch size when it comes to speeding up the training dynamics.

Data augmentation, for example, can be used to increase the size of the training set, which is advantageous. For example, using training examples allows the network to see data points from a wider range while still being representative. Finally, a data generator to automate the data transfer from our folders to Keras. For this, Keras provides handy python generator routines.

Table 1: Experimental Results.

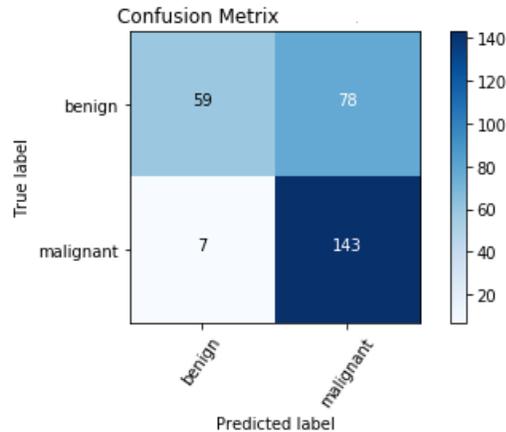| Model	Accuracy	Precision	Recall	ROC-AUC
Naïve Bayes	81.1	0.83	0.86	87
Support Vector Machine	87.2	0.86	0.87	88
Random forest	85.6	0.84	0.83	86.4
Convolutional Neural Network	91.8	0.86	0.92	88.4
Convolutional Neural Network with transfer learning	98.4	0.87	0.95	89.56

Figure 5: Confusion matrix of the implemented model.

Figure 5 has shown the confusion matrix of the implemented model through the model can be easily analyzed.

5. Conclusion

In this paper, the analysis of breast cancer is done through deep learning algorithm that is Convolutional Neural Network with transfer learning. Previously most of the cancer analysis is done through machine learning algorithms and the have achieved satisfactory results. But in this work the analysis is done on the medical images and the deep learning has performed better than other states of the art algorithms. The model has achieved an accuracy of 98.4%. In the future, some more algorithms based on convolutional neural network such as capsule network will be applied on the medical images, it might happen that these types of capsule networks can generate more accurate results than the proposed one.

References

Asbeutah, A. M., AlMajran, A. A., Brindhaban, A., & Asbeutah, S. A. (2020). Comparison of radiation doses between diagnostic full-field digital mammography (FFDM) and digital breast tomosynthesis (DBT): a clinical study. *Journal of Medical Radiation Sciences*, *67*(3), 185-192.

Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Systems with Applications*, *46*, 139-144.

Gomathi, S., Soni, M., Dhiman, G., Govindaraj, R., & Kumar, P. (2021). A survey on applications and security issues of blockchain technology in business sectors. *Materials Today: Proceedings*.

Gupta, V. K., Shukla, S. K., & Rawat, R. S. (2022). Crime tracking system and people's safety in India using machine learning approaches. *International Journal of Modern Research*, *2*(1), 1-7.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).

Khushi, M., Naseem, U., Du, J., Khan, A., & Poon, S. K. (2020, November). Data mining encode data predicts a significant role of sina3 in human liver cancer. In *International Conference on Neural Information Processing* (pp. 15-25). Springer, Cham.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

Kumar, R., & Dhiman, G. (2021). A comparative study of fuzzy optimization through fuzzy number. *International Journal of Modern Research*, *1*(1), 1-14.

Nair, R., Vishwakarma, S., Soni, M., Patel, T., & Joshi, S. (2021). Detection of COVID-19 cases through X-ray images using hybrid deep neural network. *World Journal of Engineering*.

Panta, A., Khushi, M., Naseem, U., Kennedy, P., & Catchpoole, D. (2020, November). Classification of neuroblastoma histopathological images using machine learning. In *International Conference on Neural Information Processing* (pp. 3-14). Springer, Cham.

Singh, P. D., Kaur, R., Singh, K. D., Dhiman, G., & Soni, M. (2021). Fog-centric IoT based smart healthcare support service for monitoring and controlling an epidemic of Swine Flu virus. *Informatics in Medicine Unlocked*, *26*, 100636.

Soni, M., Dhiman, G., Rajput, B. S., Patel, R., & Tejra, N. K. (2021). Energy-effective and secure data transfer scheme for mobile nodes in smart city applications. *Wireless Personal Communications*, 1-21.

Soni, M., & Singh, D. K. (2021). LAKA: lightweight authentication and key agreement protocol for internet of things based wireless body area network. *Wireless Personal Communications*, 1-18.

Vaishnav, P. K., Sharma, S., & Sharma, P. (2021). Analytical review analysis for screening COVID-19 disease. *International Journal of Modern Research*, *1*(1), 22-29.

Yao, X., & Liu, Y. (1999, July). Neural networks for breast cancer diagnosis. In *Proceedings of the 1999 congress on evolutionary computation-CEC99 (Cat. No. 99TH8406)* (Vol. 3, pp. 1760-1767). IEEE.

Zhang, Y., Zhang, B., Coenen, F., & Lu, W. (2013). Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Machine vision and applications*, *24*(7), 1405-1420.

Zhu, W., Xie, L., Han, J., & Guo, X. (2020). The application of deep learning in cancer prognosis prediction. *Cancers*, *12*(3), 603.