Efficacy of MobileNet Models in Detecting Breast Cancer in Patient Histopathology Images – An Empirical Examination

VenuGopal Balijepally
Oakland University, balijepa@oakland.edu

Uroob Mullick
Oakland University, uroobmullick@oakland.edu

Efficacy of MobileNet Models in Detecting Breast Cancer in Patient Histopathology Images – An Empirical Examination

Emergent Research Forum (ERF)

VenuGopal Balijepally
Oakland University
balijepa@oakland.edu

Uroob Mullick
Oakland University
uroobmullick@oakland.edu

Abstract

Breast cancer, the most common of all the cancers, is treatable when detected early. Histopathology (HP) biopsy images generated from breast tissue samples are commonly used as early screening tools for detecting malignancy. This study has developed and compared the efficacy of two convolutional neural network (CNN) models that are based on MobileNet architecture for automatic detection of Invasive Ductal Carcinoma (IDC), the most common form of breast cancer, in histopathology (HP) images of breast tissues. The best of the two models has shown encouraging classification performance in terms of accuracy, precision and recall. The results attest to the viability of using lighter CNN models such as MobileNet for decision support when screening for breast cancer.

Keywords

Deep Learning Models, Breast Cancer, Histopathology images, CNN, MobileNet.

Introduction

Cancer is a leading cause of death, contributing to nearly 10 million (i.e., nearly one in six) deaths worldwide in 2020. Cancer results from the progressive transformation of healthy cells into precancerous lesions and into malignant tumors. Among the different cancers (i.e., breast, lung, colon and rectum, prostrate, skin and stomach), breast cancer is the most common one accounting for 2.26 million new cases worldwide in 2020 (WHO 2022).

Breast cancer is treatable when detected early. Imaging techniques such as mammograms (breast x-ray images), ultrasound imaging, Computer tomography, magnetic resonance imaging (MRI), and histopathology (HP) biopsy imaging are commonly used as screening tools for early detection/diagnosis (Murtaza et al. 2020).

Histopathology (HP) biopsy imaging involves collecting breast tissue samples and fixing them across glass slides for examining under a microscope. The slides are stained by Hematoxylin and Eosin (HE) dyes that highlight the structure of the tissue sample for visual analysis by expert pathologists. These stained slides are further scanned and digitized into whole slide images (WSIs) so that region of interest (ROI) patches can be extracted from WSIs at different magnification levels to help with the diagnoses (Murtaza et al. 2020).

Manual reading of different HP images to detect malignancy and its type requires a lot of expertise and is a cumbersome and time-consuming task for the pathologists involved. The resulting fatigue could also contribute to misdiagnoses, particularly in cases of initial stages of breast cancer. Given the advances in AI technologies and related machine learning/deep learning algorithms and architectures, computer-aided diagnosis (CAD) systems equipped with pattern recognition tools offer a lot of promise serving as second opinion options for pathologists and radiologists so that they could provide diagnosis with higher levels of confidence (Murtaza et al. 2020).

In this research we develop a convolutional neural network (CNN) model that is based on MobileNet, for automatic detection of malignancy in histopathology (HP) images of breast tissues. We chose MobileNet model in this research as it is a light-weight model with low latency that has shown promising performance in some recent studies (Chorianopoulos et al. 2020). The rest of the paper is organized as follows. First, we provide a brief overview of related literature. Next, we explain the method and present the model and discuss its efficacy in correctly classifying the test dataset HP images after training.

Literature Review

Deep learning, an area falling within the broad domain of machine learning, uses artificial neural networks for facilitating learning to automate classification or feature detection of raw data. It tries to replicate the way the human brain functions in performing unstructured pattern analysis. The deep learning models do not require having prior knowledge of the domain as they can organize and extract different features from the underlying data. Deep learning approaches such as DNN (deep neural networks), RNN (recurrent neural networks), DBN (deep belief networks) and CNN (convolutional neural networks) are being used across industries for applications related to speech recognition, natural language processing, computer vision, image processing, material testing, etc. (Alanazi et al. 2021; Murtaza et al. 2020).

CNN networks are a popular option for feature exploration of images, which is achieved by running layers of convolutions on the image and searching for patterns (Alhussein et al. 2018). The CNN models are designed to have an input layer, multiple hidden layers and an output layer. CNN architecture is considered best suited for complex image classifications as a CNN includes pooling layers and convolutions with a filter or kernel helping to convolve the image pixels. Some previous studies attest to the superiority of CNN models compared to other deep learning models such as MLP in achieving higher accuracy (Desai and Shah 2021). A review study suggests that CNN models are quite popular in breast cancer classification modeling research (Murtaza et al. 2020).

MobileNet is also a convolutional neural network model used for mobile vision and classification of images. However, MobileNet stands out due to its speed and minimal utilization of computer power in aiding transfer learning. It is designed for the model to be functional in mobile applications and other systems with less computational power and GPU while resulting in higher accuracy (Pujara 2020). The model makes use of 1x1 convolution for changing dimensions known as pointwise convolution along with depth-wise convolution. A recent research study used a MobileNet based model for image recognition for pest control in fruits and achieved improvement in the classification accuracy with fewer parameters (Thangaraj et al. 2020).

Within the healthcare domain, automating detection and classification of breast cancer from multi-modal medical images using deep learning approaches has been gaining research attention for some time now. Some previous studies examined using neural network and/or machine learning algorithms for detection of malignancy in mammogram images. For instance, one recent study did a comparison of deep learning (DL) based Convolutional Neural Network (CNN) architectures and machine learning (ML) algorithms for the detection of malignancy in mammogram images (Alanazi et al. 2021) and reported achieving higher accuracy with CNN approaches (87%) compared to ML algorithms (78%). A recent review study provides an excellent summary of the current state of research in this area (Murtaza et al. 2020).

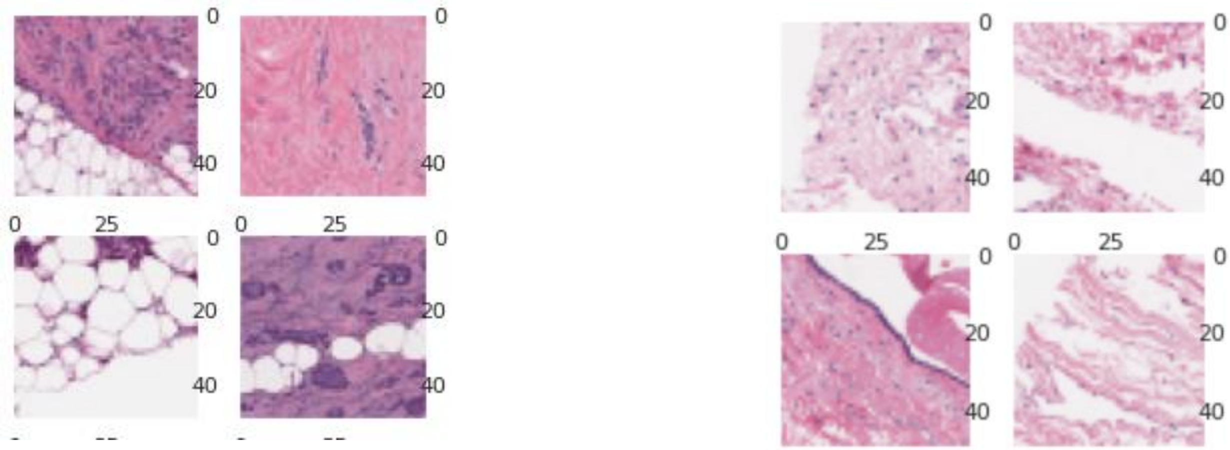
In this study, we used MobileNet based models to help diagnose breast cancer from histopathology images.

Method

We used Kaggle 162 H&E (Janowczyk 2015) dataset for testing the MobileNet model presented in this study. The dataset includes both Invasive Ductal Carcinoma (IDC) positive (malignant) and IDC negative (benign) image patches. The original data has 277,524 digital image patches, each with 50 x 50 pixels in RGB. An example each of positive and negative tissues can be seen in Figure 1. The dataset included 78,786 malignant and 198,738 benign images.

As it is an imbalanced dataset containing a disproportionate number of IDC negative (benign) image patches, we down-sampled it using Python code in Google Colab to generate 6,000 each of IDC negative and positive image patches, which were then used for further analysis. The dataset was randomly divided such that 80% of the data was used for training the model with the rest 20% used for testing.

We applied two MobileNet models, built in Google Colab, for binary classification of the breast cancer HP image dataset. The modeling was done on Intel core i5-10210U CPU @1.60 GHz with 8GB RAM and AMD Radeon M433 TPU through Google Colab Pro. In developing the models, we used Python with the following deep learning frameworks: 'TensorFlow' for training models, 'Keras' for developing and evaluating models, 'NumPy' for working with arrays, and 'pandas' for data manipulation (Costa 2020).



IDC Positive tissue IDC Negative tissue

Figure 1: Examples of IDC positive and negative image patches.

We used two MobileNet models (Models 1 and 2) for training and classifying the dataset. We pre-processed the data by scaling the images differently for the two models–i.e., we rescaled the data to 96 x 96 pixels and 128 x 128 pixels for the MobileNet Model 1 and Model 2, respectively. Model specifications used for the two models are summarized in Table 1.

Item	MobileNet Model 1	MobileNet Model 2
Resized Image size	96 x 96 pixels	128 x 128 pixels
Layers	Input layer (weights="imagenet") Convolutional layer (activation="relu") Pooling layer (GlobalAveragePooling2D) Dense layer (activation="sigmoid")	Input layer (weights="imagenet") Pooling layer (GlobalAveragePooling2D) Dense layer (activation="sigmoid")
Epochs	10	10
Processing Unit used	TPU	TPU

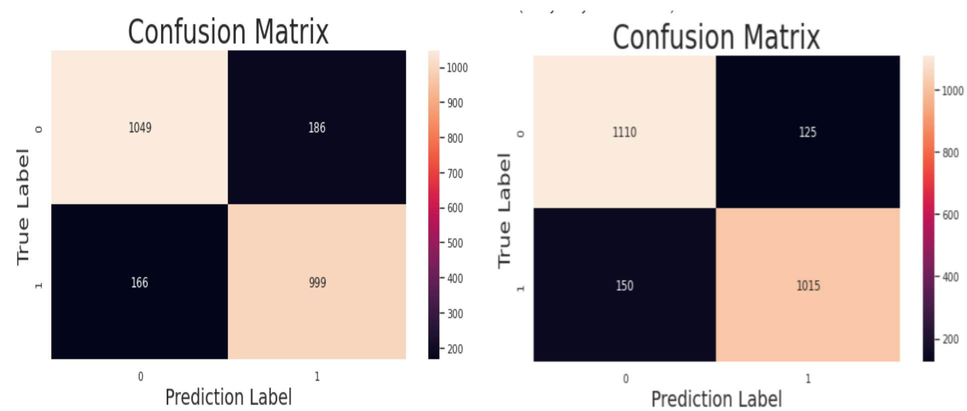
Table 1. Model Specification Comparison

Analysis and Results

We evaluated the efficacy of MobileNet neural network models developed for breast cancer diagnosis using validation accuracy measures, confusion matrices, and classification reports.

Validation accuracy metric reflects how correctly a model classifies the underlying data into the groups of interest on the testing dataset after each iteration/epoch. The loss function is calculated as the sum of errors and is indicative of the behavior of the model after each iteration/epoch with lower loss value indicative of better performance (Kumar 2018). Results suggest that both models had performed at the same validation accuracy level (84%), while Model 2 had performed with a lower loss (0.35 for Model 2 vs. 0.39 for Model 1). Thus, based on the validation accuracy and loss metrics, Model 2 has performed better.

Confusion matrices developed for each of the two models are shown in Figures 2. From the confusion matrix of the MobileNet Model 2 in Fig. 2 we can see that TP (true positive), TN (true negative), FP (false positive) and FN (false negative) values respectively are 42.29%, 46.25%, 5.21%, and 6.25%. These percentages are calculated by dividing the raw count in each quadrant by the total number of observations (i.e., 2400).



Model 1 Model 2

Figure 2: Confusion Matrices of the MobileNet Models

Next, we compared the classification reports (computed from confusion matrices shown above) which are summarized in Table-2. Precision is a probabilistic measure that indicates whether a predicted positive case actually belongs to the positive class. Recall is a probabilistic measure that indicates if an actual positive case is correctly classified with the positive class. The F1 score value is calculated as the geometric mean between precision and recall while the Support value shows the number of samples of the true response that resides in that class (Dabeer et al. 2019).

	MobileNet - Model 1				**MobileNet - Model 2**			
	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
Benign	0.84	0.86	0.85	1235	0.88	0.90	0.89	1235
Malignant	0.85	0.82	0.83	1165	0.89	0.87	0.88	1165
Accuracy			0.84	2400			0.89	2400
Micro Avg	0.84	0.84	0.84	2400	0.89	0.89	0.89	2400
Weighted Avg.	0.84	0.84	0.84	2400	0.89	0.89	0.89	2400

Table 2. Model Performance Comparison

From the results summarized in Table-2, we can see that MobileNet model 2 has performed better in terms of having higher precision and recall for both the benign and malignant classes. The model accuracy is marginally higher than the (88.26%) accuracy score for IDC case detection reported in a previous study (Chorianopoulos et al. 2020) that used a MobileNet based model for similar purposes. Given the high accuracy of this MobileNet model, this could be a candidate for use in decision support systems that pathologists could use for validation or for second opinion to complement their visual judgment when examining HP images for early breast cancer detection.

Conclusion

This study has developed and compared two convolutional neural network models that are based on MobileNet, for automatic detection of malignancy in histopathology (HP) images of breast tissues and demonstrated their efficacy. As MobileNet models offer higher processing speeds with minimal computing resources compared to other deep learning models like CNN, they are highly promising for several healthcare applications. Future studies may test the efficacy of MobileNet models on larger datasets in healthcare and other settings.

REFERENCES

Alanazi, S.A., Kamruzzaman, M., Islam Sarker, M.N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., and Siddiqi, M.H. 2021. "Boosting Breast Cancer Detection Using Convolutional Neural Network," *Journal of Healthcare Engineering* (2021).

Alhussein, M., Muhammad, G., Hossain, M.S., and Amin, S.U. 2018. "Cognitive Iot-Cloud Integration for Smart Healthcare: Case Study for Epileptic Seizure Detection and Monitoring," *Mobile Networks and Applications* (23:6), pp 1624-1635.

Chorianopoulos, A.M., Daramouskas, I., Perikos, I., Grivokostopoulou, F., and Hatzilygeroudis, I. 2020. "Deep Learning Methods in Medical Imaging for the Recognition of Breast Cancer," *2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA*, pp. 1-8.

Costa, C.D. 2020. "Best Python Libraries for Machine Learning and Deep Learning." Retrieved 02/20/2022, from https://towardsdatascience.com/best-python-libraries-for-machine-learning-and-deep-learning-b0bd40c7e8c

Dabeer, S., Khan, M.M., and Islam, S. 2019. "Cancer Diagnosis in Histopathological Image: Cnn Based Approach," *Informatics in Medicine Unlocked* (16), p 9.

Desai, M., and Shah, M. 2021. "An Anatomization on Breast Cancer Detection and Diagnosis Employing Multi-Layer Perceptron Neural Network (Mlp) and Convolutional Neural Network (Cnn)," *Clinical eHealth* (4), pp 1-11.

Janowczyk, A. 2015. "Use Case 6: Invasive Ductal Carcinoma (Idc) Segmentation." Retrieved 02/20/2022, from http://www.andrewjanowczyk.com/use-case-6-invasive-ductal-carcinoma-idc-segmentation/

Kumar, H. 2018. "Technical Fridays. "Loss Vs Accuracy"." Retrieved 03/01/2022, from https://kharshit.github.io/blog/2018/12/07/loss-vs-accuracy

Murtaza, G., Shuib, L., Abdul Wahab, A.W., Mujtaba, G., Mujtaba, G., Nweke, H.F., Al-garadi, M.A., Zulfiqar, F., Raza, G., and Azmi, N.A. 2020. "Deep Learning-Based Breast Cancer Classification through Medical Imaging Modalities: State of the Art and Research Challenges," *Artificial Intelligence Review* (53:3), pp 1655-1720.

Pujara, A. 2020. "Image Classification with Mobilenet." Retrieved 02/20/2022, from https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470

Thangaraj, R., Dinesh, D., Hariharan, S., Rajendar, S., Gokul, D., and Hariskarthi, T. 2020. "Automatic Recognition of Avocado Fruit Diseases Using Modified Deep Convolutional Neural Network," *International Journal of Grid and Distributed Computing* (13:1), pp 1550-1559.

WHO. 2022. "Cancer Fact Sheet." Retrieved 02/28/2022, from https://www.who.int/en/news-room/fact-sheets/detail/cancer