JOURNAL OF
EMERGING INVESTIGATORS

Article

Predicting the Instance of Breast Cancer within Patients using a Convolutional Neural Network

Sanjay Adhikesaven[1], Aradhya Kapoor[1] , Ali Shadab Khowaja[1], Vivian Li[1], Karthik Sabhanayakam[1], Larry McMahan[1]

[1] Department of Computer Science & Engineering, Aspiring Scholars Directed Research Program, Fremont, CA

SUMMARY

Breast cancer is a widespread disease that affects millions of people across the world. This makes early detection and diagnosis critical for heightened chance of survival. Although there are many ways for breast cancer to be detected, many patients are diagnosed late in their disease progression. In this paper, we present a robust and automated way of early breast cancer detection within patients using machine learning methods. Our method uses medical images of possible breast cancer and classifies them as either non-cancerous or cancerous using a convolutional neural network (CNN). Our results showed that our CNN model had a mean accuracy of 97.25% when averaged using the final 50 percent of samples. These results show that machine learning can be used as an effective way to clinically diagnose breast cancer due to the high accuracy.

INTRODUCTION

In 2020 alone, 2.3 million women were diagnosed with breast cancer, and 685,000 died from breast cancer (1). The survival rate after 5 years varies 66 percent in India and 40 percent in South Africa (1). Thus, it is critical to detect breast cancer within patients (both male and female) early in order to increase survival rates in many nations. When breast cancer is detected early within the localized stage (before it spreads to distant sites), the survival rate is 99% (2). Despite having multiple methods to detect breast cancer within patients, many diagnoses are delayed. Within Libyan women, for example, there is a median diagnosis delay of 7.5 months, meaning that the time between the first symptoms to the diagnosis is 7.5 months (3). The cause of this delay is both patient delay, which is when the patient delays seeking medical attention, and system delay, which is when the healthcare system delays procedures such as appointments, testing, or results that ultimately delay the diagnosis (3). It is crucial for the medical system to decrease this system delay, as a diagnosis delay of three months or more can lead to larger tumors, high instances of late clinical stage, and metastatic disease (3).

Presently, there are four main methods being used to detect breast cancer: breast ultrasound, diagnostic mammogram, breast magnetic resonance imaging (MRI), and biopsy. A breast ultrasound creates sonograms of areas inside the breast, a diagnostic mammogram is a detailed X-ray of the breast, a breast MRI is a scan that creates pictures of inside the breast, and a biopsy removes tissue or fluid from the breast to do more testing (4). However, in all of these circumstances, a doctor has to manually look at the images to determine whether someone has breast cancer. After diagnosing someone with breast cancer, doctors determine the stage of the cancer, which is determined by where the cancerous cells have spread within the body (4). We aim to reduce this diagnosis delay by using a convolutional neural network (CNN) in order to classify breast cancer screening data. This model would have the benefit of being both robust and automated. A CNN is a deep learning algorithm that consists of multiple layers and is often used for image classification (8). The input image is fed into the CNN, and the layers are used to learn different features within the image to aid in classification. After it is trained using many input images, it is then tested using unseen data, which is used to report the accuracy of the algorithm.

First, we aim to answer whether it is possible to use a convolutional neural network to accurately detect breast cancer. Additionally, we aim to answer to what extent convolutional neural networks can predict the instance of breast cancer. We hypothesize that CNNs can be used to detect breast cancer with relatively high accuracy. To calculate our results, we tested our CNN using 50 MRI images and calculated the average accuracy using the final 50 percent of samples. These results showed that the CNN model had a mean accuracy of 97.25%. Thus, these results showed that a CNN is a robust method of breast cancer diagnosis and has potential to be used in a clinical setting.

Table 1. Distribution of the split of the 250 data images. Within each set, the images were split evenly between both the normal (non-cancerous) and cancerous groups.

	Training	Validation	Testing	Row Total
Cancerous	75	25	25	125
Normal	75	25	25	125
Column Total	150	50	50	125

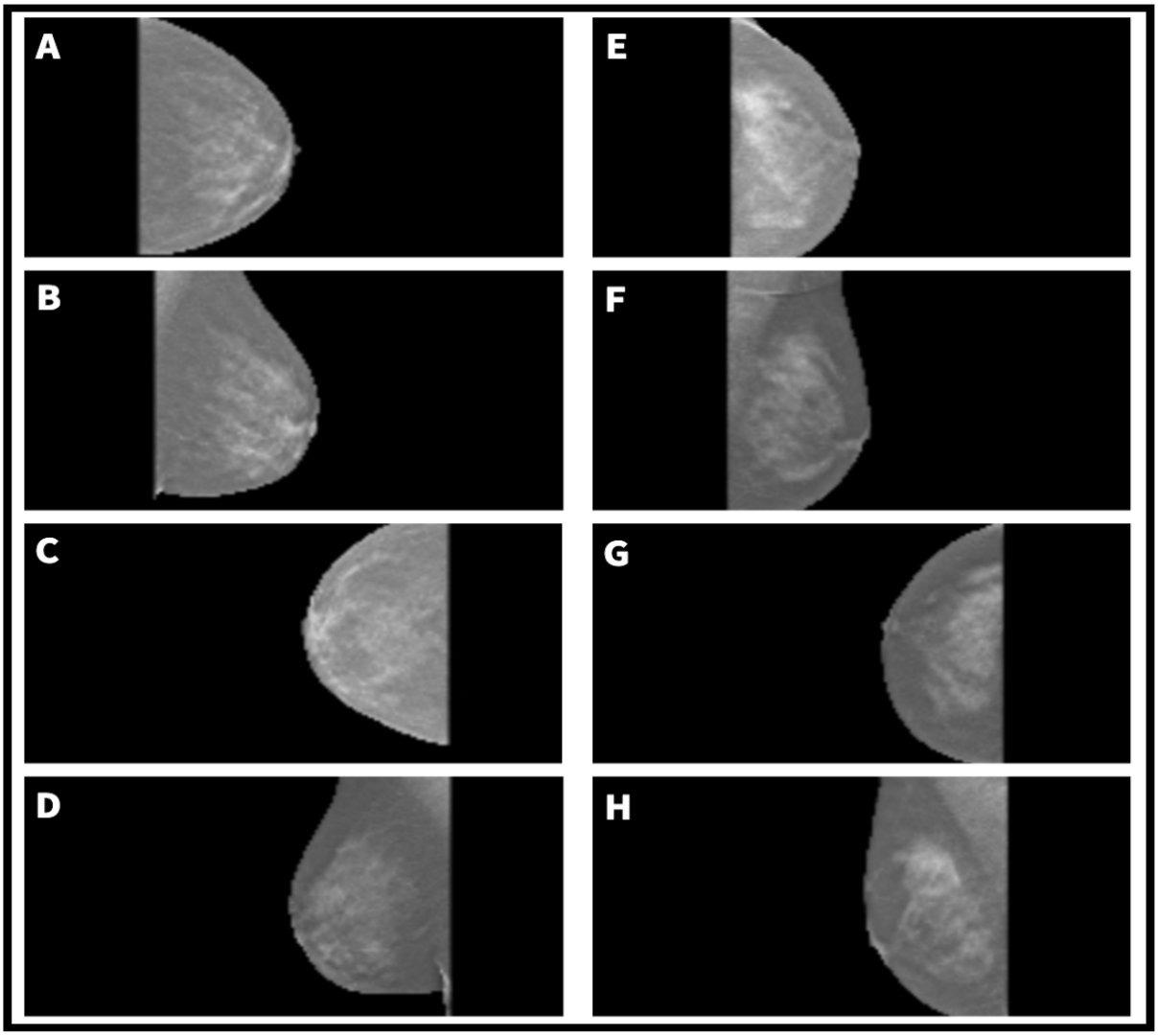

Figure 1. Comparison of cancerous and non-cancerous sample images. Images A–D are from a cancerous sample, and images E–H are from a non-cancerous sample. Images A and E are from the LCC view, B and F are from the LMLO view, C and G are from the RCC view, and D and H are from the RMLO view. Images A, B, E, and F show the left angle of the breast, and images C, D, G, and H show the right angle of the breast.

RESULTS

Our proposed system consisted of MRI scans being inputted into a CNN classifier (**Figure 1**). We used the Breast Cancer Screening - Digital Breast Tomosynthesis data provided by Duke University and used a CNN to classify the images. We tested our CNN using a total of 40 epochs, which is the number of times the dataset went through the CNN to be tested for accuracy.

The model started off with an accuracy of 33%, but it progressed on to 44%, then 56%, and 67%, and 89% within 10 epochs. Here, the accuracy refers to the amount of correctly classified images divided by the total number of images. On the 10th epoch, the model obtained its first 100% accuracy. During the next 20 epochs, the model achieved a rating of 78% accuracy 4 times, 89% accuracy 9 times, and yielded 100% accuracy on 7 instances. Next, during an additional 10 epochs, the model yielded 100% accuracy 9 of those 10 times. Finally, after having the model trained and tested, we calculated the accuracy over the final 20 epochs and found that the mean accuracy of the model was 97.25%.

DISCUSSION

Our results showed that the mean accuracy of our CNN model was 97.25% when averaged using the final 50 percent of samples (20 epochs). We were able to improve prior accuracy metrics significantly using our model architecture. In comparison to other work, our CNN performs at a higher accuracy than the state of the art, but other studies did not use the same dataset we used. For example, a study conducted by Zhang *et al.* introduced a CNN that had a 95.24% accuracy

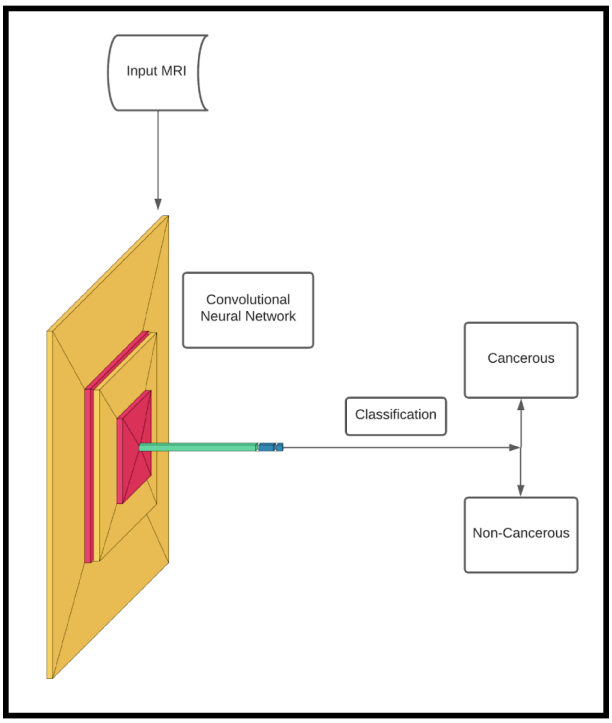

Figure 2. Architecture of MRI input into our CNN. Each MRI is input, and the CNN then decides whether a given MRI should be classified as cancerous or non-cancerous. This decision-making process is shown by the two options for the classification. The CNN architecture is illustrated by the diagram; the layer sizes represent the input dimensions into each layer, and the width of

when classifying benign (non-cancerous) versus malignant (cancerous) mammograms (5). Another model proposed by Alanzi et al. exhibited an accuracy of 87%, meaning our CNN is more accurate (6). Our model accounts for the left and the right view of the breast and employs three rounds of the convolutional and max pooling layer in comparison to past models described above only using two. This unique data set and architecture could account for the improved accuracy.

However, there are some limitations to our study. One such limitation would be the sample size of our study. We used 250 images, and using more training examples could improve the model's accuracy. In future work, we aim to mitigate this issue using data augmentation in order to generate more data. This would provide more data with which we could train and test the model. We also aim to expand our model beyond MRI images to other types of breast cancer screenings, like mammograms or ultrasound images. Presently, the model is only trained on MRI images that were within the data set we used, which is also a limitation. This could make the model biased based on the type of patients that the dataset contained. More research must be done to train a CNN using more views and types of scans the CNN can classify, making it more useful in healthcare and making the model more applicable for a general purpose.

In the future, our model can be used to aid in clinical diagnosis. This model can automate the breast cancer diagnosis process, and the model's high accuracy shows its effectiveness. This can aid in early detection of breast cancer by decreasing the time it takes to diagnose breast cancer, as our model requires MRI images to perform the classification. Since human input is not needed, our model can aid in prompt breast cancer diagnosis. This can be impactful in areas where diagnosis is severely delayed, and through aiding prompt diagnosis, the model can help increase the breast cancer survival rate.

MATERIALS AND METHODS

We used a convolutional neural network (CNN) for this study. The most prominent layers used are convolutional layers and pooling layers. Each of these layers is made up of numerous nodes, each of which serves as a regression model. Multiple convolution filters are used in the convolutional layer to extract distinct features from the input layer and build a feature map using convolutional operations. Filter size, dimensions, stride, the number of pixels the filter moves after each operation, and padding (adding zeros around the edge of the input picture to extend its dimensions), are all hyperparameters for convolution filters. A pooling layer is usually added following the convolutional layer to lower the spatial size of the convolved features and reduce overfitting. Max pooling, which takes the maximum value of each zone, and average pooling, which averages the values of each region, are the two most prevalent forms of pooling (9).

In our study, our CNN was built using three Convolutional layers, three Pooling layers, one Flatten layer, and two Dense layers. The first convolutional layer was the Conv2D layer, which kept 16 filters each sized 3 x 3. This convolution layer used the rectified linear activation function (ReLU), a piecewise linear function, and required the images to have an input size of 300 x 300 pixels. Since the images were RGB standard color models, the input was three dimensional. The next layer was the first pooling layer, MaxPool2D. This layer reduced the spatial dimensions of the output volume and had a pooling size of (2, 2). These two layers were followed by four very similar convolutional and pooling layers. Layer three was another similar Conv2D layer with 32 filters. Layer four, similar to Layer two, reduced the spatial dimensions of the output volume from layer three. Layer five was another Conv2D layer with an increased number of 64 filters, and Layer six was another MaxPool2D layer. Layers one and two were very similar to layers three and four and layers 5 and 6. However, the number of filters increased within the convolutional layers. We increased the number of filters within the CNN as layers progressed to increase the number of abstractions that the neural network could extract from image data as the data processes through the layers. This further helped the model learn more features from the input. The dense layers classified the images based on how the model was trained. Our first dense layer used the ReLU and the second one used

a sigmoid activation function. These layers are used to add non-linearity in a machine learning model. Finally, the model was compiled, and results outputted.

This study used Duke University's Breast Cancer Screening - Digital Breast Tomosynthesis data (7). We used this data set because it had a large sample size and had high-resolution images that were usable for our study. There were four possible views for each patient in the dataset: the left craniocaudal (LCC) view, the left mediolateral oblique (LMLO) view, the right craniocaudal (RCC) view, and the right mediolateral oblique (RMLO) view (**Figure 1**). From this data set, we randomly selected 125 non-cancerous MRIs and 125 cancerous MRIs, making the total data used 250 images. This sample size would allow for a large enough testing and training dataset. We then divided this data into three categories: training, validation, and testing. Using a 60/20/20 split, 150 of the images were randomly assigned to the training group, 50 of the images were randomly assigned to the validation group, and 50 of the images were randomly assigned to the testing group (**Table 1**). The training data was used to originally train the model, the validation data was used to tune the hyperparameters of the CNN, and the testing data was used to evaluate the accuracy of the CNN.

To create our model, we used Python 3.10.5 with the TensorFlow, NumPy, and Matplotlib libraries. We generated the three data sets using the TensorFlow ImageDataGenerator function, and the data was normalized such that every value in the matrix representing the images was between 0 and 1. After having initialized both the training and validation data sets, we used the TensorFlow train.flow_from_directory function, which took in the training/validation dataset path, the target size of the images (300 x 300), the batch size of the images (3), and the class mode (binary). We chose the binary class mode because our CNN classifies each image using only two categories: normal (non-cancerous) and cancerous. We then created the CNN model and compiled it. Since we used a classification model, we passed in a loss of "binary_crossentropy". Our optimizer was "RMSprop" (library: tensorflow) with a learning rate of 0.001, though the optimizer "adam" could be used as well. To test the model, we used the model.fit() function, which was given the training data, validation data, and the number of epochs. The number of epochs determines the number of times the dataset will go through the whole model and be tested for accuracy. In our case, the epochs were 40. The proposed system is represented in a simplified diagram (**Figure 2**).

ACKNOWLEDGMENTS

We would like to thank ASDRP for providing us with the resources and guidance required for this research

Received: March 20, 2022
Accepted: August 19, 2022
Published: September 19, 2022

REFERENCES

1. "Breast Cancer." World Health Organization, \www.who.int/news-room/fact-sheets/detail/breast-cancer. Accessed 22 February 2022.
2. "Breast Cancer Early Detection." National Breast Cancer Foundation,www.nationalbreastcancer.org/early-detection-of-breast-cancer/. Accessed 22 February 2022.
3. Caplan, L. "Delay in Breast Cancer: Implications for Stage at Diagnosis and Survival." Front. Public Health, vol. 2, no. 87, 2014, doi:10.3389/fpubh.2014.00087.
4. "How Is Breast Cancer Diagnosed? | CDC." Centers for Disease Control and Prevention, www.cdc.gov/cancer/breast/basic_info/diagnosis.htm. Accessed 22 February 2022.
5. Zhang C, Zhao J, Niu J, Li D (2020) New convolutional neural network model for screening and diagnosis of mammograms. PLOS ONE 15(8): e0237674. doi: 10.1371/journal.pone.0237674
6. Alanazi SA, Kamruzzaman MM, Islam Sarker MN, Alruwaili M, Alhwaiti Y, Alshammari N, Siddiqi MH. Boosting Breast Cancer Detection Using Convolutional Neural Network. J Health Eng. 2021 Apr 3;2021:5528622. doi: 10.1155/2021/5528622.
7. Buda M, Saha A, Walsh R, *et al.* A Data Set and Deep Learning Algorithm for the Detection of Masses and Architectural Distortions in Digital Breast Tomosynthesis Images. JAMA Netw Open. 2021;4(8):e2119100. doi:10.1001/jamanetworkopen.2021.19100
8. Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018). doi: 10.1007/s13244-018-0639-9
9. Gholamalinezhad, Hossein, and Hossein Khosravi. "Pooling methods in deep neural networks, a review." arXiv preprint arXiv:2009.07485 (2020).