Breast Cancer Prediction Based on the CNN Models

Nuo Chen [1,†], Boyu Han [2,*,†], Zhixin Li [3,†], Haotian Wang [4,†]

[1] Department of Economics, University of Southern California, Los Angeles, United States

[2] Department of Science, McGill University, Montreal, Canada

[3] College of Arts & Science, New York University, New York, United States

[4] Parsons School of Design, The New School, New York, United States

* Corresponding Author Email: boyu.han@mail.mcgill.ca

[†]These authors contributed equally.

Abstract. In modern society, the natural lifespan of an individual increased dramatically benefitting from advanced yet accurate methods of medical treatment. Though many diseases could be treated with a cure, the treatment of cancer has yet to be overcome. Related medical research has proven that the combination of accurate breast cancer diagnoses and treatments at an early stage could prevent the spread of cancer cells as it could increase a person's potential lifespan by a large margin. This research has conducted a comprehensive study on improving the efficiency of autonomous image recognition of breast cancer diagnosis using deep learning models. We use the most advanced CNN baseline models for image recognition, including VGG, ResNet, Efficient, etc. We also select two typical breast cancer datasets and tested the models on them to make our result more convincing. The final enhanced model of ResNet 101 can achieve a recognition rate of 89.98% for the benign and malignant samples.

Keywords: Breast Cancer, Deep Learning, CNN, VGG, ResNet, EfficientNet.

1. Introduction

Cancers are defined as a collection of diseases in which abnormal cells can divide, spread, and cause damage to nearby body tissue cells. It could permanently weaken the immune system and makes the human body more vulnerable to additional potential diseases. The statistics collected from the World Health Organization (WHO) proven that cancer is the second leading cause of death globally and most recently in 2020 it accounts for nearly 10 million deaths, occurring as frequently as one in every six deaths. Cancer could arise in many parts of the human body, while breast cancer is the most prevalent cancer type around the world. As it most commonly occurs among middle-aged women, modern medical treatment can be highly effective for breast cancer especially when the disease is identified at an early stage, which can prevent further spread and growth of malignant tissues, thereby potentially expanding an individual's lifespan.

The initial step of breast cancer diagnosis is usually conducted by a doctor using physical examination, visual analysis of mammography, and ultrasonic images in order to collect premature information regarding the breast' s shape, density, and suspected lesions. The final diagnosis is confirmed by a pathologist using microscopic analysis of the extracted region after surgery. However, there are several difficulties and challenges that remain regarding the traditional method of diagnosing breast cancer. First, hospitals often lack experienced staff to perform accurate diagnosis mammograms [1] as mammogram screening is highly effective to detect early potential signs of breast cancer, resulting in poorly diagnosed cases being reported every year. Additionally, mammograms are still popularly acknowledged by researchers as an exceptionally difficult research field considering the various modalities, densities, and shapes. Both abnormal and normal cells typically appear as high-intensity regions under mammograms, making it challenging for the naked eye to differentiate between abnormal and normal tissues [1] while trying to make the correct diagnosis. Last but not least, the final decision for breast cancer diagnosis remains subjective to different pathologists. In other words, the same patients might receive different diagnosis decisions regards whether they developed breast cancer. While patients could develop cancer without any symptoms, the process of breast cancer diagnosis is laborious and complex as the final diagnosis is also subjective, making it critical for early diagnosis to save lives.

The development and advancement in artificial neural networks ignite new hopes for breast cancer treatment. Over the recent year, computing algorithms implemented in the medical field using deep learning models have the capability of reaching astonishing accuracy.

Convolutional Neural Network (CNN) has been most popularly used for images analysis due to its unique characteristic of hidden convolutional layers within CNN models performing convolutional operations in order to pick out or detect specific patterns and make sense of them, making it vital for autonomous image recognition. As such, it was widely perceived as the most effective candidate for automated breast cancer detection and classification applications.

2. Related Work

Although breast cancer can be detected through physical discomfort or the use of mammograms at an early stage, the efficiency and accuracy of diagnosis are still a thorny issue. Recently, machine learning is a popular tool to aid breast cancer diagnosis. With the aid of machine learning methods including K Nearest Neighbour (KNN), Support Vector Machine (SVM), and Convolutional Neural Network (CNN), detection systems for breast cancer have many effective improvements.

Tsochatzidis et al. increased accuracy in detecting breast cancer using segmentation data integrated into a CNN [2]. They modified each CNN's convolutional layer and added a new loss function in their experiment and demonstrated that performance assessment in diagnosis was conducted using two mass datasets for mammography, DDSM-400 and CBIS-DDSM, given differences in accuracy levels of corresponding ground truth segmentation maps. In [3], Masud et al. examined eight distinct pre-trained models that have been fine-tuned to see how machine learning algorithms classify breast tumors using ultrasound pictures and a shallow customized convolutional neural network that surpasses training models across a range of performance metrics has been proposed. Tahmooresi et al. proposed a hybrid machine learning (ML) model that integrates the Support Vector Machine (SVM), Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), and Decision Tree (DT) techniques for effective breast cancer diagnosis (DT). In the result, with the model of SVM, which is one of the most popular ways used for the diagnosis of cancer, they achieved a maximum accuracy of 99.8%. Under the supervision of structural and statistical data taken from the photographs, Nahid et al. classified a collection of biological breast cancer photos (from the BreakHis dataset). For the purpose of classifying breast cancer images, they suggest CNN, Long-Short-Term Memory (LSTM) or their combination. After extracting features using the novel DNN models, the decision-making stage had been carried out using Softmax and Support Vector Machine (SVM) layers. Their experiment's greatest Accuracy result of 91.00% was reached on the 200x dataset. Similarly, other researchers [4-6] also proposed to utilize CNN in the detection of cancer in their experiments. These connected studies demonstrate how machine learning algorithms can significantly enhance breast cancer diagnosis.

3. Methodology

3.1. Data

3.1.1 Dataset

In this study, we make use of two datasets. The first one is Dataset-BUSI-with-GT from Kaggle, which contains 780 images about breast cancer abnormalities. Both practice and test data are included in this dataset, and each file contains benign, malignant and normal images. Since the number of images in Dataset-BUSI-with-GT is not enough to examine the accuracy of our model, another dataset from Breast Cancer Histopathological Database (BreakHis) is also used in our experiment. According to BreakHis, this dataset contains information of 82 patients, which generates 9,109 microscopic photos at various magnifications of breast tumor tissue (40X, 100X, 200X and 400X). It contains a total of 5,429 malignant and 2,480 benign samples from 8 distinct types of breast tumors.

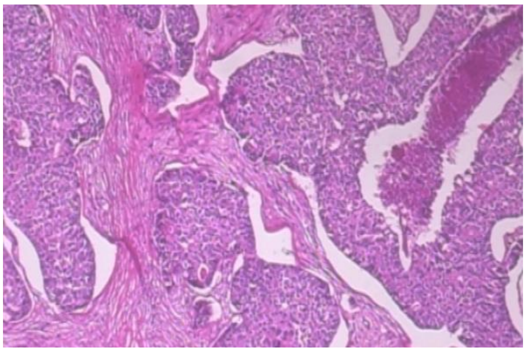

Figure 1. Example picture of BreakHis

Figure 1 shows a breast cancer slide stained with HE and seen at four various magnifications.

3.1.2 Data preprocess

For Dataset-BUSI-with-GT from Kaggle, because the original dataset was classified by benign or malignant tumors rather than by image type, we did the data split by categorizing the original data into two parts, the original x-ray images and their masks. This preprocess allowed us to test our model in the experiment more efficiently. For the dataset from BreakHis, there are eight tumor types in the dataset. Because the data covered by the original dataset is too large, we selectively extracted representative data. We have selected 4 or 5 folders under each tumor classification. On average, each benign tumor has 250 images and each malignant tumor has 350 images.

3.2. Model

3.2.1 Transfer learning using pre-trained VGG

AlexNet achieved outstanding results in the 2012 ImageNet competition, prompting more and more people to pay attention to deep neural networks, creating a series of methods using deep neural networks to solve image problems. So far, rich network models have emerged in this field, and we will introduce several key models in the subsequent content.

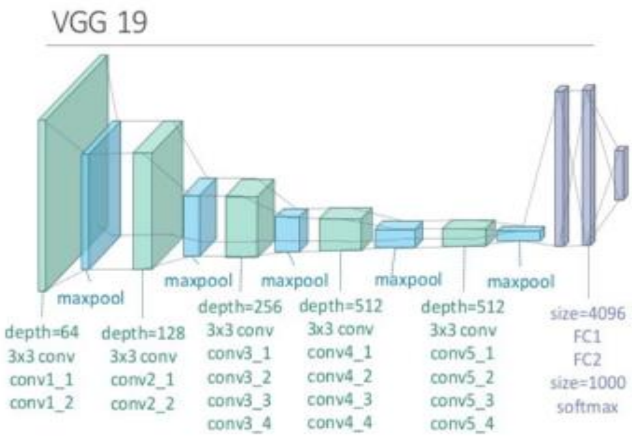

Figure 2. Architecture of VGG19.

AlexNet contains three fully-connected layers at the end and five convolutional layers in the front, using data augmentation, dropout and ReLU activation functions which can suppress overfitting and reduce vanishing gradients. Although AlexNet had a good performance at the time, it did not propose a clear direction for the design of neural networks. This problem is well solved in VGG [7]. The VGG network consists of a combination of several 3×3 convolutional layers and 2×2 pooling layers to do feature extraction, followed by three fully-connected layers with 4096, 4096 and 1000 nodes respectively. VGG uses ReLU as the activation function, Dropout to suppress overfitting, and finally uses the output of softmax as the classification prediction. VGG shows that adopting a small convolution kernel can reduce the number of parameters that need to be trained, thereby improving the training and testing efficiency of the model. For example, using two 3x3 convolutional layers has the same receptive field as using one 5x5 convolutional layer, both being 5, but two 3x3 convolutional layers require fewer parameters to train than one 5x5 convolutional layer. At the same time, the reduction of the convolution kernel means the superposition of more convolution layers, which further deepens the neural network and is conducive to image classification. The success of VGG demonstrates that image feature extraction performance can be improved by increasing the depth of the neural network. In this paper, we choose VGG19 (the number of convolutional layers and fully-connected layers are 16 and 3 respectively) as the baseline model for further experiments. The architecture of VGG19 is shown in Figure 2.

3.2.2 Development of the enhanced CNN model

Inspired by the above models, the industry gradually increases the number of layers of the neural network model and the structure becomes more and more complex to obtain better prediction results. In theory, this idea of deepening the neural network is feasible, but practice shows that with the increase in the number of network layers, the training error tends to increase. The possible reason is that with the continuous increase of network depth, more and more activation functions are introduced, and the data are mapped to a more discrete space. At this time, it is difficult to return the data to the origin (identity map). The problem is so-called degradation which means when deepening the network, it does not perform as well as the shallow network. Using the concept of residual learning, Kaiming He et al. proposed the residual network ResNet [8] to solve the above problems.

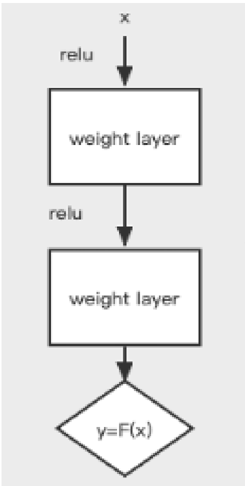

Figure 3. Conventional approach.

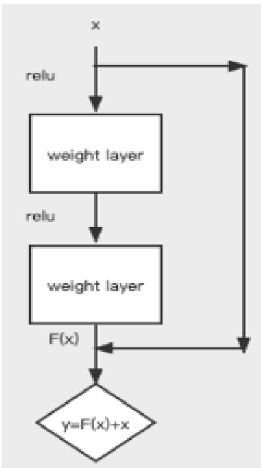

Figure 4. Residual learning.

Figure 3 shows the traditional learning method that maps x to y = F(x) output when adding a network. Figure 4 shows residual learning applied by ResNet with output y=F(x)+x. The straight line (the arc on the right in Figure 4) connecting the input x directly to the output y is called a shortcut. Residual learning does not directly learn the representation of the output feature y, but learns y-x. We only need to set all parameters of F(x) to 0 to learn the representation of the original model. F(x)=y−x is also called the residual term, and y=x is the identity map. If the mapping of x→y is close to the identity map, it is much easier to learn the residual term in Figure 4 than to learn the full mapping form in Figure 3.

ResNet adds a residual element via the shortcut connection, referring to VGG. When down sampling, ResNet directly uses Stride = 2 for convolution and uses the global average pooling layer instead of the fully connected layer. Global average pooling can replace the fully connected layer without affecting its accuracy, and the whole network structure can realize the structure of full convolution. However, the use of a fully connected layer will increase the number of parameters of the entire network, which is easy to cause the phenomenon of overfitting. Because of this shortcut connection, when the network propagates the gradient backward, the gradient does not disappear and the gradient at each layer is updated. In theory, the depth of the network can be increased almost without limit. The training efficiency of ResNet is dramatically be improved by introducing the Depthwise Seperable ResBlock proposed by Haoyu Ren et al. [9]. Even so, deep networks are not necessarily better, since overfitting and deep network problems still often require weeks of training.

A scaling method that combines both speed and accuracy called compound scaling was proposed by Tan, M. et al. in 2019. Tan jumps out of the previous understanding of the scaling model and examines these scaling dimensions from a high level. Tan thought that these three dimensions are mutually influential and explored the best combination of them, and proposed the latest model EfficientNet [10] on this basis.

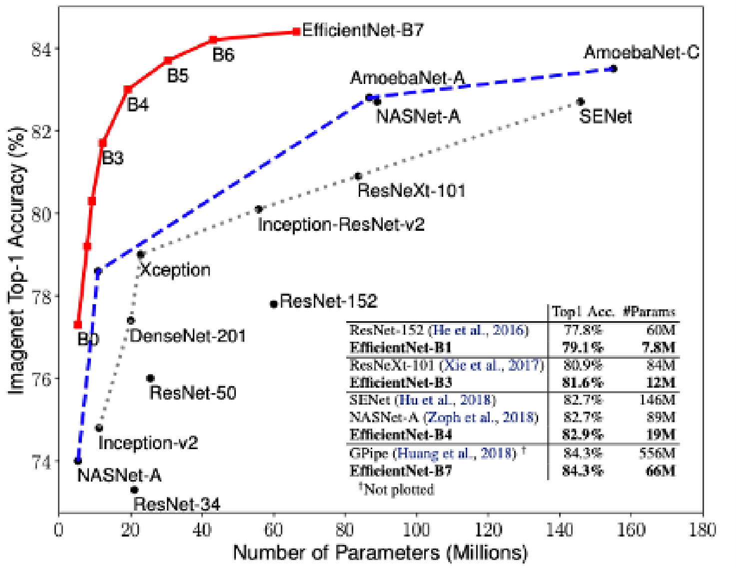

Figure 5. Model Size vs. ImageNet Accuracy

In terms of the EfficientNet-B0, its main building block is mobile inverted bottleneck MBConv, to which also adds squeeze-and-excitation optimization. Figure 5 illustrates the superior performance of EfficientNet. In this paper, because the breast cancer dataset is not very complex, the baselines ResNet50, ResNet101, EfficientNetB0, and EfficientNetB3 with relatively simple structures compared to ResNet152 and EfficientNetB7 are selected for further experiments.

4. Performance evaluation and result analysis

We split the datasets into training set and validation set in order to evaluate the models. The training set's goal is to fit the model's parameters. While the model's hyperparameters are being adjusted, the validation set provide an estimation of the model performance. Initially, we adopted the Dataset-BUSI-with-GT dataset. In addition to the baseline neural network models, we added an Input layer with shape (None, 500, 500, 3) prior to the baseline, a global average pooling 2d layer with shape (None, 1280), and a dropout layer with 0.3 dropout rate.

We used VGG19, ResNet 50, ResNet101, and EfficientNetb0 and compared the performance of those models with different batch sizes(bs). As indicated below in Table 1, the learning rate(lr) is 0.001 and epochs are set to 5. As shown in Table 1, ResNet50 with fine-tuned hyperparameters reached the highest training accuracy of 73.72% and highest validation accuracy which is 78.85%. However, the Dataset-BUSI-with-GT dataset is relatively small with only 780 pictures. Since the results are not satisfying, we introduce another dataset from BreakHis.

Table 1. VGG19, ResNet50, ResNet101, EfficientNetb0 Experiments Results with Dataset-BUSI-with-GTDataset

Model	#Traning_data	#Val_data	lr	bs	Epochs	Train_acc	Val_acc
VGG19	624	156	0.001	16	5	0.5865	0.6923
VGG19	624	156	0.001	32	5	0.6138	0.6474
VGG19	624	156	0.001	64	5	0.5000	0.6346
ResNet50	624	156	0.001	16	5	0.7372	0.7885
ResNet50	624	156	0.001	32	5	0.6699	0.7628
ResNet50	624	156	0.001	64	5	0.6410	0.7038
ResNet101	624	156	0.001	16	5	0.7019	0.7885
ResNet101	624	156	0.001	32	5	0.7147	0.75
ResNet101	624	156	0.001	64	5	0.6522	0.7436
EfficientNetb0	624	156	0.001	16	5	0.6458	0.6859
EfficientNetb0	624	156	0.001	32	5	0.633	0.6282
EfficientNetb0	624	156	0.001	64	5	0.609	0.6282

Like what we have done in the Dataset-BUSI-with-GT dataset, we added an Input layer with shape (None, 460, 700, 3) prior to the baseline, a global average pooling 2d layer with shape (None, 1280), and a dropout layer with 0.2 dropout rate.

The model is further enhanced after using the BreakHis dataset. As indicated down below in Table 2, we set the learning rate(lr) to 0.001 and epochs are set to 5. Table 2 below reveals the simulation results. The highest validation accuracy is 89.98% when bs=16 in the ResNet101 model.

Table 2. VGG19, ResNet50, ResNet101, EfficientNetb0 Experiments Result with BreakHis Dataset

Model	#Traning_data	#Val_data	lr	bs	Epochs	Train_acc	Val_acc
VGG19	1839	459	0.001	16	5	0.7216	0.7603
VGG19	1839	459	0.001	32	5	0.6754	0.7712
ResNet50	1839	459	0.001	16	5	0.8945	0.8845
ResNet50	1839	459	0.001	32	5	0.8608	0.8736
ResNet101	1839	459	0.001	16	5	0.9125	0.8998
ResNet101	1839	459	0.001	32	5	0.8847	0.841
EfficientNetb0	1839	459	0.001	16	5	0.8135	0.8126
EfficientNetb0	1839	459	0.001	32	5	0.7852	0.7625

5. Conclusion

Early detection of breast cancer is important for lowering the disease's fatality rate. The development of deep learning and image processing has helped in detecting and offering timely treatment. The radiologist can now read breast images easier and more accurately with the aid of a model that has been created. A pre-trained VGG, which is capable of capturing richer and more sophisticated features, has been applied to the datasets created from Dataset-BUSI-with-GT and BreakHis. Then, by keeping track of the pre-trained VGG model's learning rate, an improved CNN model was created. The enhanced CNN model attains an identification rate of 89.98% for the benign and malignant samples. Today, there are still limited datasets provided to the researchers. As the additional datasets of breast images taken under the different conditions are available to the research community, our model would be further enhanced, such as using ResNet-RS proposed by Irwan Bello et al. instead of ResNet, and the diagnosis of breast cancer should be improved. Alongside this, the further improvement of technologies would support image processing and classification of benign and malignant examples.

References

[1] M Thilagaraj, N. Arunkumar, Petchinathan Govindan, "Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model",Computational Intelligence and Neuroscience, vol. 2022, Article ID 6785707, 12pages, 2022.

[2] L. Tsochatzidis, P. Koutla, L. Costaridou, and I. Pratikakis, “Integrating segmentation information into CNN for breast cancer diagnosis of mammographic masses,” Computer Methods and Programs in Biomedicine, vol. 200, p. 105913, 2021.

[3] M. Masud, A. E. Eldin Rashed, and M. S. Hossain, “Convolutional neural network-based models for diagnosis of breast cancer,” Neural Computing and Applications, vol. 5, 2020.

[4] M. Malathi, P. Sinthia, F. Farzana, and G. Aloy Anuja Mary, “Breast cancer detection using active contour and classification by deep belief network,” Materials Today: Proceedings, 2021.

[5] M. S. Hossain, G. Muhammad, and N. Guizani, “Explainable AI and mass surveillance system-based healthcare framework to combat COVID-I9 like pandemics,” IEEE Network, vol. 34, no. 4, pp. 126–132, 2020.

[6] D. Singh, S. Singh, M. Sonawane, R. Batham, and P. A. Satpute, “Breast cancer detection using convolution neural network,” International Research Journal of Engineering and Technology, vol. 5, no. 6, pp. 316–318, 2017.

[7] Simonyan, K., & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556, 2014.

[8] He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. arXiv:1512.03385, 2015.

[9] H. Ren, M. El-Khamy, J. Lee. DN-ResNet: Efficient Deep Residual Network for Image Denoising. Asian Conference of Computer Vision 2018.

[10] Tan, M., & Le, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. ICML 2019.