A COMPARATIVE EXAMINATION OF CONVOLUTIONAL AUTOENCODER AND DENSENET APPLICATIONS FOR BREAST CANCER CLASSIFICATION

Naderan Maryam, Yuri Zaychenko

Abstract: *Breast cancer is one of the most widespread and dangerous cancers among women. It is a disease that requires quick and accurate diagnoses. For this task, convolutional neural networks (CNNs) represent a huge breakthrough in image recognition. Many CNNs however, require large datasets for training which is not always available to researchers. In this paper, the effects of using a small dataset will be compared between our proposed convolutional autoencoder (CA) and DenseNet. It is shown that when using a small dataset there is considerable overfitting when using DenseNet, whereas overfitting does not occur with the proposed CA. Moreover, the training time of the CA was faster than DenseNet, and sensitivity (recall) of the proposed model was 90%.*

Key words*: Breast cancer detection, convolutional autoencoder, DenseNet, image classification*

1. Introduction

Breast cancer is a serious public health risk, affecting a large number of women every year**.** According to the last studies, one in every eight women in the United States will develop breast cancer in her lifetime. In the last year, it is estimated there were 268,600 new cases of invasive breast cancer and 62,930 new cases of non-invasive breast cancer diagnosed in women in the U.S [1]. Breast cancer is one of the largest public health threats requiring early detection to save lives. When it comes to cancer detection, a false negative might leave a patient with a lack of treatment which can have dire consequences for the patient. A false positive will just lead to more testing and analysis, which will eventually lead to the discovery of the false positive. For this reason, recall (sensitivity) is more important than accuracy.

The aim of this study is to develop deep learning methods that could improve the sensitivity and training time of diagnosing breast cancer. In order to reduce the training time of the model, the model should be simplified. However, if the convolutional DenseNet will be used, because the model is more complex, it requires more data to train the model. As a result, the time of training is increased. Alternatively, using a convolutional autoencoder, the number of convolutional layers will be chosen in a way that will simplify the model and reduce the chance of overfitting. Using a small sample size with the convolutional autoencoder, produced a better result than the DenseNet model while reducing the training time.

2. Review of previous works

Authors in [2] proposed a new method they call “end-to-end" method and used the Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM). The authors compared two convolutional neural networks, VGG and Resnet50, with their own method. The accuracy achieved using their proposed model was 84% and 97% for cancer and no-cancer respectively and the sensitivity of their model was 86.1%.

In [3] two modified CNNs were proposed where the average accuracy was 85%. During the experiment, 2420 mammography scans were used for training their proposed models. However, the model was very complex.

Authors in [4] proposed a modified model, where a CNN was used for feature extraction, but the gradient boosted tree model was used for classification. The experiment used 1804 mammography scans, from the Digital Database for Screening Mammography (DDSM). This modified model achieved 85% accuracy at detecting images with masses, with a sensitivity of 85%. The weakness of both [3,4] is that the training time of the model is increased. In this paper, a modified model with a decreased number of parameters is proposed. This model reduces the training time, while maintaining or improving sensitivity.

3. Dataset

The BreakHist dataset was used for this experiment. The dataset includes two classes of tumors Malignant and Benign, which are further organized by tumor type. The dataset is also separated into four magnification zooms 40X, 100X, 200X and 400X. Fig. 1 illustrates some input images that were used for training the model. Figures 1.a - 2.d belong to the benign category and figures 1.e - 1.h belong to the malignant category.

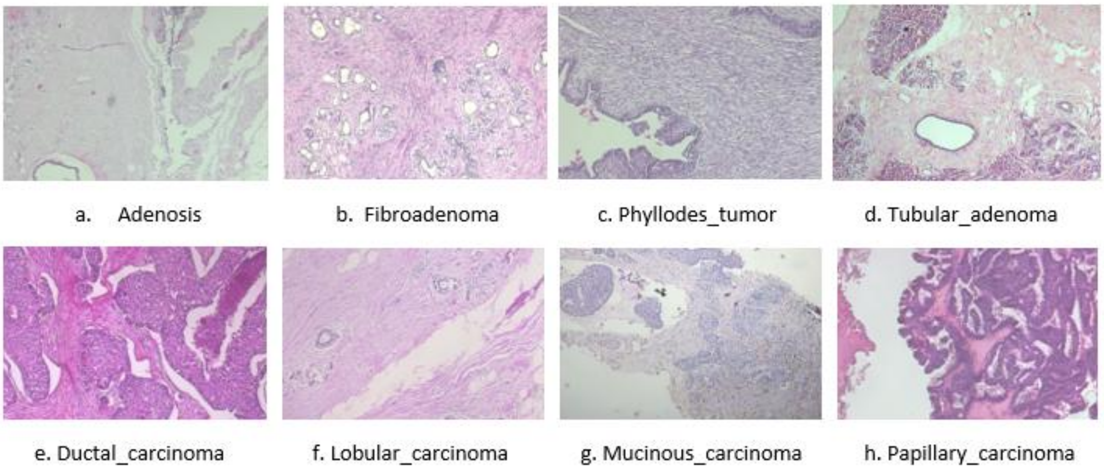

Fig. 1 Sample of dataset.

4. Experimental Investigations and Analysis

In this paper, all experiments were developed using Jupyter Labs, TensorFlow 2 and Python 3. The programs were implemented on a virtual machine with eight Intel CPUs.

Two convolutional networks DenseNet and Autoencoder were proposed in this work. The DenseNet model was both trained from scratch (FS) and fine-tuned (FT). There are some advantages of using the convolutional autoencoder:

1) It is challenging to get access to large datasets with labeled scans. However, since the autoencoder is an unsupervised model it does not require labeled

data, making it a viable solution for image recognition challenges involving a lack of labeled data, or small datasets.

2) The model is simple. Respectively, it has less parameters and as a result, the time of computation and training is drastically reduced.

The DenseNet (FS) model used 24 filters and 54 convolutional blocks. Each block included convolutional, activation, maxpooling and normalization layers. Therefore, the model is very complex and requires a large amount of data for training. Whereas, the convolutional autoencoder is simple with eight convolutional, eight batch normalization and two max-pooling layers. This simple architecture allows the CA to be trained with less data. Figure 2 illustrates the accuracy for the training and validation sets using DenseNet.

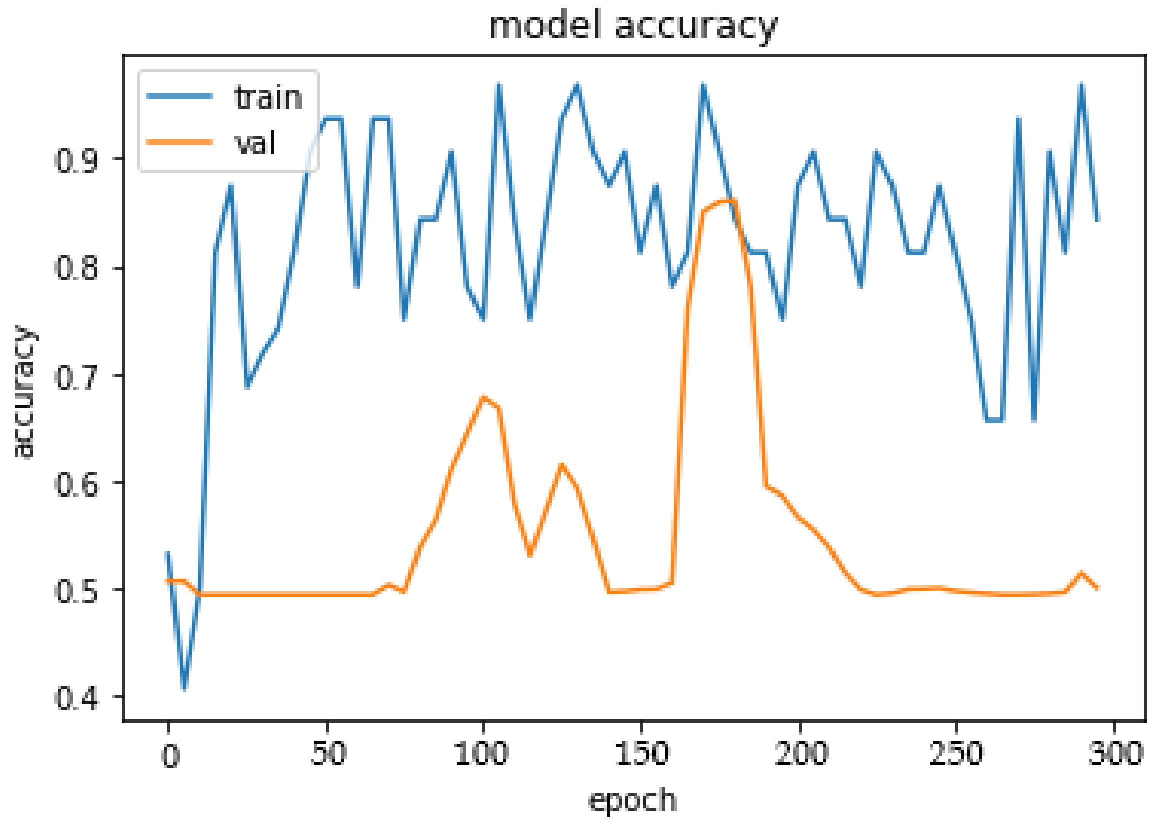

Fig. 2 Accuracy of training and validation using DenseNet

According to figure 2, it can be concluded that the model is overfitting since the accuracy for training data is significantly better than the accuracy for validation data. Using the same dataset, the modified convolutional autoencoder, showed better results than the DenseNet FS model. Figure 3 shows the loss function for the training and the validation set while using the autoencoder model.

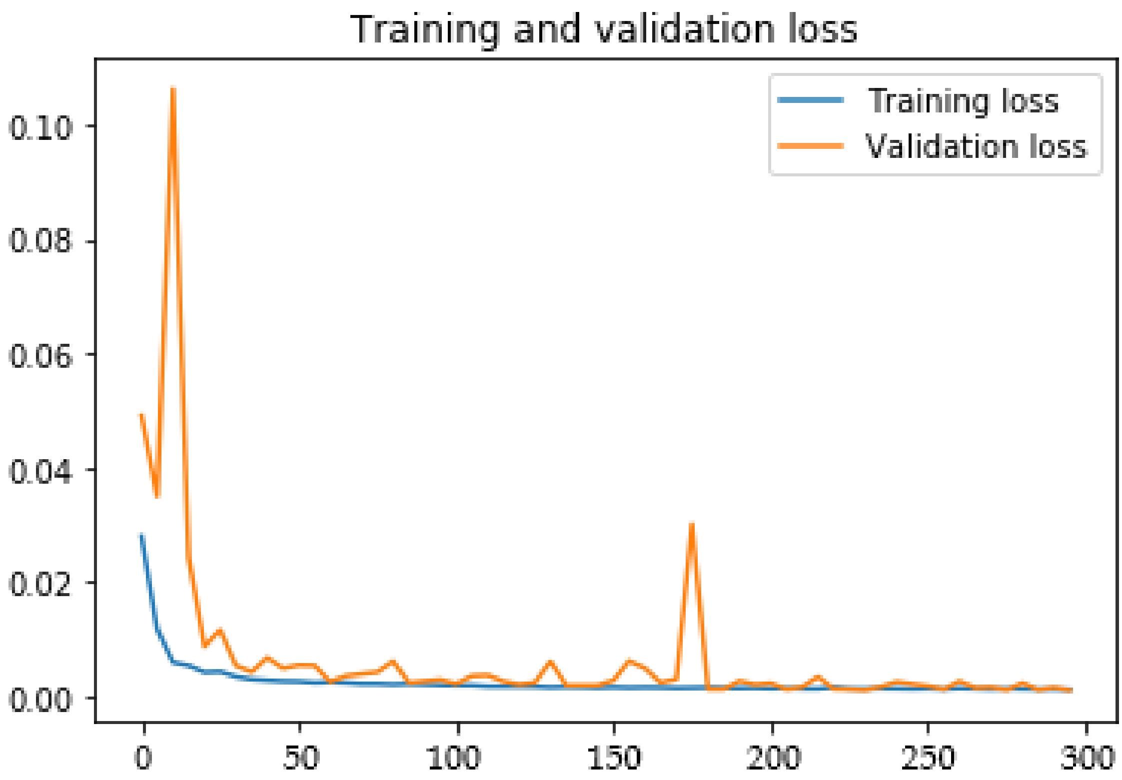

Fig. 3 Loss function of training and validation set using convolutional autoencoder

The loss function presents a measure of mistakes that were made by the network in predicting the output. Figure 3 also shows that after epoch 200, the value of the error function for training and validation data does not change and the loss reaches its minimum value. As a result, the optimal number of epochs in the current task is 200.

The DenseNet model was both trained from scratch and fine-tuned. To fine-tune the model, the output layer (Softmax layer) of pre-trained model was replaced with a new layer recognizing two classes, cancer and no cancer. The rest of the pretrained layers were frozen. The sensitivity of this model as well as the proposed convolutional autoencoder were 90%, however the autoencoder required less data and training time.

Table 1 illustrates the results of the CNNs that were used in this paper.

Table 1 – Comparison of CNNs for Breast Cancer Detection

	Precision	Sensitivity	F1-Score	Accuracy
Convolutional Autoencoder	90.40%	90%	89.50%	90%
DenseNet FT	91%	90%	90.50%	95%
DenseNet FS	70%	65%	62.66%	67.50%

According to table 1, even though the sensitivity of the proposed CA and DenseNet FT are the same, training time in DenseNet FT was longer than CA. The training time of the DenseNet FT was about 3 hours, whereas in CA it was ~1.5 hours.

Conclusion

It is crucial to detect breast cancer at early stages, because if it is left undiscovered patients are at risk of more severe levels of cancer. For this reason, sensitivity is an important measurement in analyzing various CNNs. In our experiments we have shown that a high level of sensitivity (90%) is achieved with the convolutional autoencoder, while seeing other advantages as well. The convolutional autoencoder has less parameters than DenseNet, therefore, the model is less complex and prevents overfitting when using a small dataset. As a result, the training time of the model is dramatically decreased. In contrast, DenseNet requires a big amount of data to train the model and the training time increases considerably.

Bibliography

1 National Breast Cancer Foundation INC "Breast Cancer Facts" www.nationalbreastcancer.org/breast-cancer-facts.

2 Li Shen, Laurie R. Margolies, Joseph H. Rothstein, Eugene Fluder, Russell McBride, Weiva Sieh. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. Scientific Reports. (2019) 9:12495.

3 M. G. Ertosun and D. L. Rubin, "Probabilistic visual search for masses within mammography images using deep learning," *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Washington, DC, 2015, pp. 1310-1315.

4 Kooi T, van Ginneken B, Karssemeijer N, den Heeten A. Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. Med Phys 2017; 44: 1017–27. doi: https://doi.org/10.1002/mp.12110.

5 Zeshan Hussain, Francisco Gimenez, Darvin Yi, Daniel Rubin AMIA Annu Symp Proc. 2017; 2017: 979–984. Published online 2018 Apr 16. PMCID: PMC5977656.

Authors' Information

***Naderan Maryam** - PhD student. Institute for applied system analysis, NTUU "KPI", 03056, Ukraine, Kyiv, Peremogi pr. 37, Corpus 35*

***Yuri Zaychenko** – Professor, doctor of technical sciences,*

Institute for applied system analysis, NTUU "KPI", 03056, Ukraine, Kyiv, Peremogi pr. 37, Corpus 35; e-mail: baskervil@voliacable.com, zaychenkoyuri@ukr.net

Major Fields of Scientific Research: Information systems, Fuzzy logic, Decision making theory