Breast Cancer Prediction

Fatemah Abdelwahed and Youssef Montasser

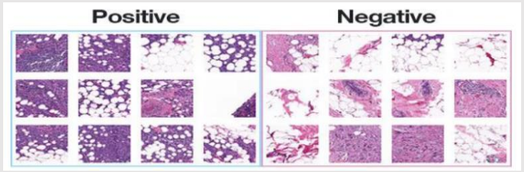
American University in Cairo

Abstract

Recently, there has been a notable emphasis on Deep Learning, particularly in its application within the medical domain. 7.8 million women had been diagnosed with breast cancer within the last five years, establishing it as the most widespread form of cancer globally. In this study, we focus on the use of CNNs to predict Breast Cancer, describe the data employed in constructing a model, present our findings, and discuss potential work for future research.

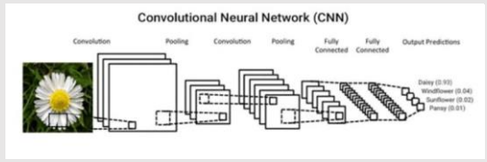
Introduction

We will train a deep learning model to predict Invasive Ductal Carcinoma (IDC) is most common subtype of all breast cancers, in breast histology images. This model will automatically give a prediction after analyzing the images for cancer risk factors.



Methodology

Addressing this challenge necessitates the utilization of a CNN (Convolutional Neural Network) model. Primarily employed for tasks such as image analysis, identification, and classification, CNNs excel in simplifying images for practical analysis. The model's procedure involves passing the image through a convolutional layer to filter and enhance the dataset's features.



We built on a model developed by pyimagesearch. The model initially had an accuracy of 83%.

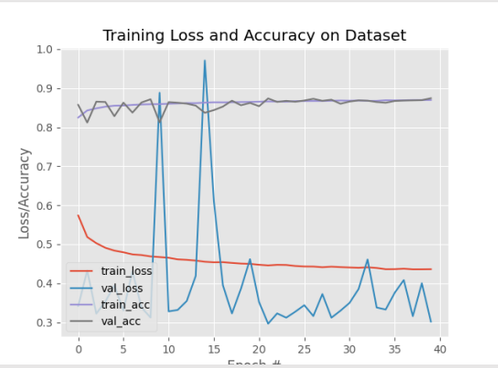
We tried different

- epochs,
- optimizer,
- train/test/validation division,
- augmentation,
- added an attention layer,
- various CNN models

to get different accuracies.

Results

Removing 3 Convolutional layers of 128 filters with their pooling layer had the best overall accuracy for the model at 87%.



Discussion

We employed a substantial dataset of 3 GB for model training, resulting in prolonged processing times. Optimal performance can be achieved by running the model on a local machine equipped with a GPU. Additionally, noteworthy evaluation metrics for the model include sensitivity and specificity.

```
[[61822  9473]
 [ 2910 25538]]
acc: 0.8759
sensitivity: 0.8671
specificity: 0.8977
```

Conclusion

In this study, CNNs were effectively employed to predict Invasive Ductal Carcinoma (IDC), with a notable accuracy improvement to 87% achieved through strategic model adjustments. The removal of three Convolutional layers emerged as a key enhancement, showcasing the model's adaptability to various configurations.

Future Work

Moving forward, the exploration of more sophisticated CNN architectures and the incorporation of additional relevant features could further enhance the predictive capabilities of the model. Validating the model on diverse datasets could contribute to its applicability in real-world clinical settings. Furthermore, extending the research to address potential class imbalances in the dataset would be valuable for comprehensive breast cancer prediction.

References

Rosebrock, A. (2019, February 18). Breast cancer classification with Keras and Deep Learning - PyImageSearch. PyImageSearch. Link

World Health Organization. (n.d.). Breast cancer. World Health Organization. Link.