CNNS On Breast Cancer Diagnosis

Alaa Ashraf
artificial intelligence
Cairo University
Cairo, Egypt
alaaashra709@gmail.com

Kareem Ahmed
artificial intelligence
Cairo University
Cairo, Egypt
karimx791@gmail.com

***Abstract*—Breast cancer is still a big health concern, affecting 13 out of 100 women worldwide. It is also one of the main cancer deaths causes worldwide. In this paper, we will be discussing how we can use CNNS to improve breast cancer diagnosis.**

Keywords—CNNs, Breast Cancer, optimizations

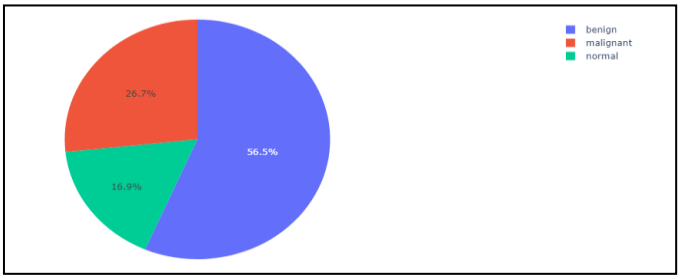
INTRODUCTION

Detecting breast cancer has developed over the years. We started with human-based detection, then moved on to use some machine learning algorithms, where researchers used SVMS, Naïve Bayes classifiers, and RCNN classifiers. We even took it further and are now using deep learning models like convolutional neural networks (CNN). CNN is a deep learning algorithm used in image processing, and researchers have found it very useful and promising in many diagnoses. And That's why we will be using it to determine whether a tumor is benign, malignant, or normal in nature. The breast cancer diagnosis flow is as follows: a breast exam, a mammogram, a breast ultrasound, then removing a sample of breast cells for testing (biopsy), and moving on, among some other additional steps. Our work will take place in the biopsy stage, where we will make the detection CNN-based instead of human-based.

DATASET

We will be using the data collected at baseline, which includes ultrasound images among 600 women in ages between 25 and 75 years old. The images are categorized into three classes which are benign, malignant, and normal. The drawbacks of this dataset are that it does not have that much of number of samples (only 1578). and, it has an unbalanced class distribution which might cause overfitting, but we will discuss how to solve these issues later in the paper.

Fig. 1. The class distribution



RELATED WORKS

A. *Diagnostic methods*

There are many methods used for the diagnosis of breast cancer but the most used one, and the one that has been extensively studied for breast cancer detection is Magnetic Resonance Imaging (MRI), It is good, but it suffers from the false positive rate which is not highly dangerous, so machine learning will be used by mammography radiology that takes images of breast as an x-ray image.

B. *Machine Learning & Deep Learning in Brest Cancer Detection*

There are many developed machine learning algorithms for breast cancer detection, and we will show their results.

Method	**accuracy range**
K-Nearest Neighbor [1]	83 to 86 19.28
Pre-Trained Networks [2]	80 to 89 4.74
Feature Extracted Using CNN [3]	83 to 90 4.28
Deep Convolution [4]	91.54

C. *Challenges & limitations*

Datasets in the context of breast cancer detection are always imbalanced as the benign class is the majority class, and malignant is the minority class.

PROPOSED MODEL

Applying pre-trained ResNet50 to extract features from the images then applying dimensionality reduction using PCA and feature selection and then the images will go through the ML classifier which is Random Forest.

OPTIMIZATIONS

We applied many optimization techniques which led to accuracy improvements and helped the model to converge.

A. *Data augmentation*

As we mentioned earlier in the paper our data suffers from unbalanced class distribution, so we need to handle that. There are some tricks you can do maybe using generative adversarial networks or diffusion models. However, using these advanced methods carries a higher

risk of generating inaccurate images compared to traditional augmentation techniques like shearing and rotating. The inaccurate generation of images may lead to incorrect training and potentially harm the model's performance. And that's why we will be sticking with the traditional augmentation techniques.

B. *PCA*

PCA is one of the most famous dimensionality reduction techniques. It is quite solid and will accelerate the model convergence.

C. *Grid Search*

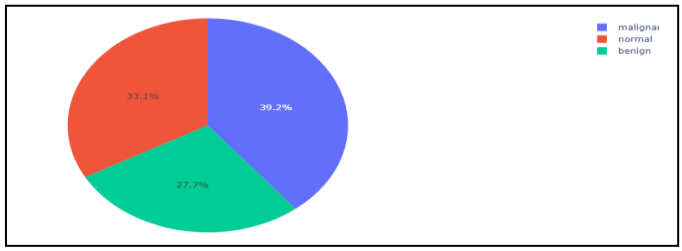
For finding which hyperparameter combinations are best for a machine learning model, we could use grid search which tries them all. It entails creating a hyperparameter grid then trains the model using every possible combination and evaluates results. However, Grid Search can be computationally expensive and may overlook ideal values if the grid is too coarse.

EXPERIMINTAL RESULTS

A. *Data preproccing results*

We split the data to train, validation, and test with ratios 0.7, 0,15, 0.15, then we applied the data augmentation to balance the unbalanced data.

Fig. 2. The class distribution after data augmentation



B. *Classifing on the Data befor using Grid Search*

	precision	recall	f1-score	support
0.0	0.95	0.49	0.65	41
1.0	0.72	0.96	0.81	135
2.0	0.85	0.44	0.58	64
accuracy			0.74	240
macro avg	0.83	0.63	0.68	240
weighted avg	0.78	0.74	0.72	240

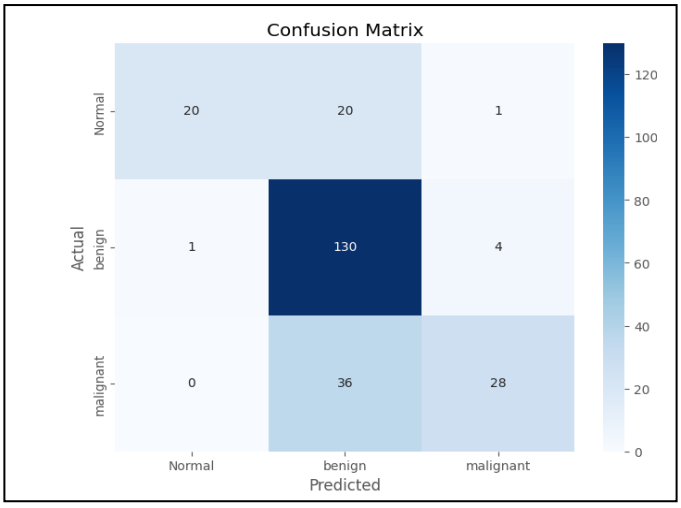


Fig. 3. Confusion matrix before Grid Search

C. *Classifing on the Data after using Grid Search*

	precision	recall	f1-score	support
0.0	0.95	0.51	0.67	41
1.0	0.72	0.96	0.82	135
2.0	0.8	0.53	0.66	64
accuracy			0.77	240
macro avg	0.85	0.67	0.72	240
weighted avg	0.80	0.77	0.75	240

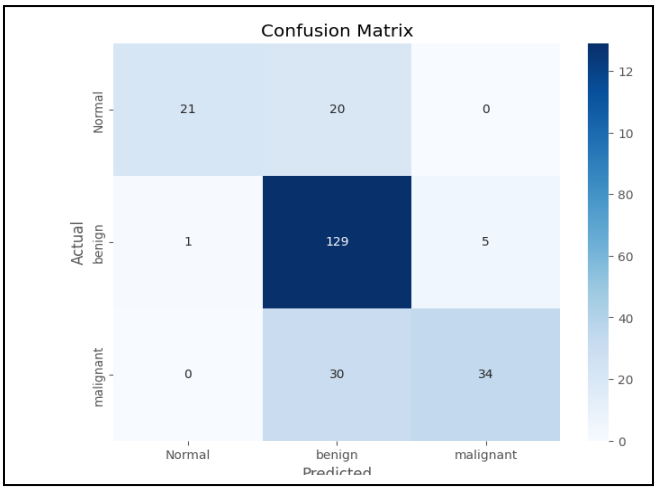


Fig. 4. Confusion matrix after Grid Search

***conclusion*—CNNs seems to be very promising when it comes to Breast Cancer Diagnosis it's quite important to look into improving it so that we don't have to suffer from the human error anymore when it comes to those dangerous disease.**

REFERENCES

[1] S. Dabeer and S. Islam "Cancer diagnosis in histopathological image: CNN based approach" 2019.

[2] IslamS.A. Adeshina, A.P. Adedigba, A.A. Adeniyi, A.M. Aibinu “Breast cancer histopathology image classification with deep convolutional neural networks” April 2018

[3] Song Y, Zou JJ, Chang H, Cai W. Adapting Fisher vectors for histopathology image classification. In: 14th IEEE international symposiumon biomedical imaging, ISBI 2017, melbourne, Australia, april 18-21, 2017; 2017. p. 600–3

[4] Samah AA, Fauzi MFA, Mansor S. Classification of benign and malignant tumors in histopathology images. In: 2017 IEEE international conference on signal and image processing applications, ICSIPA 2017, kuching, Malaysia, september 12-14, 2017; 2017. p. 102–6

[5] A. Adeshina P.adedigba and A.Adeniyi “Breast Cancer Histopathology Image Classification with Deep Convolutional Neural Networks” 2018.