**Classification of Breast Cancer with Deep Learning**

**Michael Oroyo**

### **Abstract**

Breast cancer is one of the diseases that represent a large number of mortalities in the world. Data mining classifications techniques will be used to effectively classify cancer data to facilitate decision-making. The objective of this paper is to compare the performance of different machine learning algorithms in the diagnosis of breast cancer, to define exactly if this type of cancer is a benign or malignant tumor.

I have selected the best algorithms for both classification and regression to check which algorithms work best in detecting breast cancer.

### **Introduction**

Breast cancer is a disease where cancer cells form in the tissue of the breast of the woman and can propagate to the other organs of the body. It represents the second cause of death for women after lung cancer [1]. Early diagnosis can reduce the breast cancer mortality rate by 40% [2]. Data mining and machines learning algorithms have become interesting tools in the field of health. Because they can process and analyze massive data to extract useful information in decision-making. So they will be effective solutions to predict and diagnose breast cancer also to classify it into its two categories either benign or malignant tumor.

The project will be performed using CNN on IDC regular datasets to determine whether an input of an image of a breast cancer dataset is benign or malignant.

**1.1** **Breast Cancer (Overview)**

Cancer begins in cells, the building blocks that make up all tissues and organs of the body, including the breast. Normal cells in the breast and other parts of the body grow and divide to form new cells as they are needed. When normal cells grow old or get damaged, they die, and new cells take their place. Sometimes, this process goes wrong. New cells form when the body doesn‘t need them, and old or damaged cells don‘t die as they should. The buildup of extra cells often forms a mass of tissue called a lump, growth, or tumor. Tumors in the breast can be benign (not cancer) or malignant (cancer):

Benign tumors:

Are usually not harmful.

Rarely invade the tissues around them

Don‘t spread to other parts of the body

Can be removed and usually don‘t grow back

Malignant tumors:

May be a threat to life

Can invade nearby organs and tissues (such as the chest wall)

Can spread to other parts of the body

Often can be removed but sometimes grow back

1.2 **Risk Factors**

Although risk factors don‘t tell everything. Many risk factors may increase chances of having breast cancer; it is not yet known just how some of these risk factors cause cells to become cancer (American Cancer society, 2002).

* Gender: Breast cancer is about 100 times more common in women than in men.
* Age: The chance of getting breast cancer goes up as a woman gets older.
* Genetic risk factors: Inherited changes (mutations) in certain genes like BRCA1 and BRCA2 can increase the risk.
* Family history: Breast cancer risk is higher among women whose close blood relatives have this disease.
* Personal history of breast cancer: A woman with cancer in one breast has a greater chance of getting a new cancer in the other breast or in another part of the same breast. Race: Overall, white women are slightly more likely to get breast cancer than African-American women. Asian, Hispanic, and Native-American women have a lower risk of getting and dying from breast cancer.
* Dense breast tissue: Dense breast tissue means there is more gland tissue and less fatty tissue. Women with dense breast tissue have a higher risk of breast cancer.
* Certain benign (not cancer) breast problems: Women who have certain benign breast changes may have an increased risk of breast cancer. Some of these are more closely linked to breast cancer risk than others.
* Lobular carcinoma in situ: In this condition, cells that look like cancer cells are in the milk-making glands (lobules), but do not grow through the wall of the lobules and cannot spread to other parts of the body. It is not a true cancer or pre-cancer, but having LCIS increases a woman's risk of getting cancer in either breast later.
* Menstrual periods: Women who began having periods early (before age 12) or who went through the change of life (menopause) after the age of 55 have a slightly increased risk of breast cancer.
* Breast radiation early in life: Women who have had radiation treatment to the chest area (as treatment for another cancer) as a child or young adult have a greatly increased risk of breast cancer. The risk from chest radiation is highest if the radiation were given during the teens, when the breasts were still developing.
* Treatment with DES: Women who were given the drug DES (diethylstilbestrol) during pregnancy have a slightly increased risk of getting breast cancer
* Not having children or having them later in life: Women who have not had children, or who had their first child after age 30, have a slightly higher risk of breast cancer. Being pregnant many times or pregnant when younger reduces breast cancer risk.
* Certain kinds of birth control: Studies have found that women who are using birth control pills or an injectable form of birth control have a slightly greater risk of breast cancer than women who have never used them.
* Using hormone therapy after menopause: Taking estrogen and progesterone after menopause increases the risk of getting breast cancer.

**2 Related Works**

In previous works, Spanhol, et. al [2] used a CNN architecture inspired by AlexNet [1] to classify H&E breast tissue biopsy samples in benign and malignant tumors, using multiple magnifications and two patch extraction methods and others, [2]– [3][4] have applied to digital pathology image analysis to breast cancer.

Also, the research is supported by Analysis of Histopathology Images, using Pre-trained CNN [5], [6] that utilizes diverse pre-trained models to arrange various sorts of harmful tissues. Whereas In this paper [7], they propose a neural conditional random field framework to detect cancerous tissue in the WSI. The framework uses feature extracted from CNN and considers the spatial correlation between neighboring patches through a CNN, which provides a score of 0.8096 for FROCs. Different work also carried out on the same data set as [8] author had proposed a method of integration based on random forest dissimilarity to merge different feature groups together. Which gives 87.10% accuracy.

3 **Method**

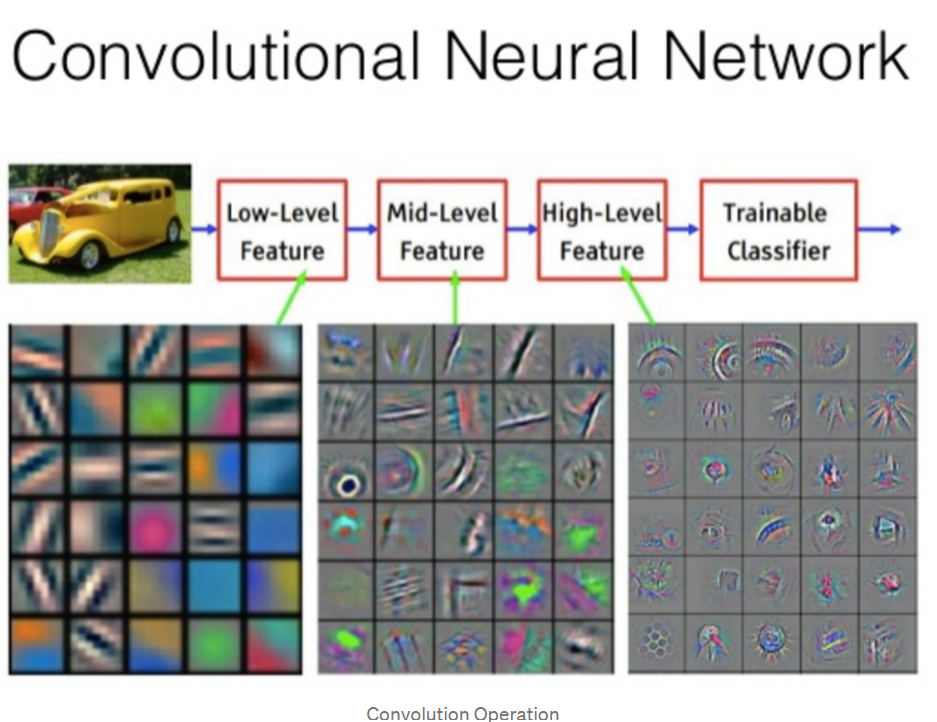
**CNN Architecture**

**3.1 Input**

A Matrix of pixel values in the shape of [WIDTH, HEIGHT, CHANNELS]. Let’s assume that our input is [32x32x3].

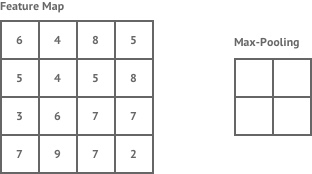
**3.2 Convolution**

The purpose of this layer is to receive a feature map. Usually, we start with a low number of filters for low-level feature detection. The deeper we go into CNN, the more filters we use to detect high-level features. Feature detection is based on ‘scanning’ the input with the filter of a given size and applying matrix computations in order to derive a feature map.



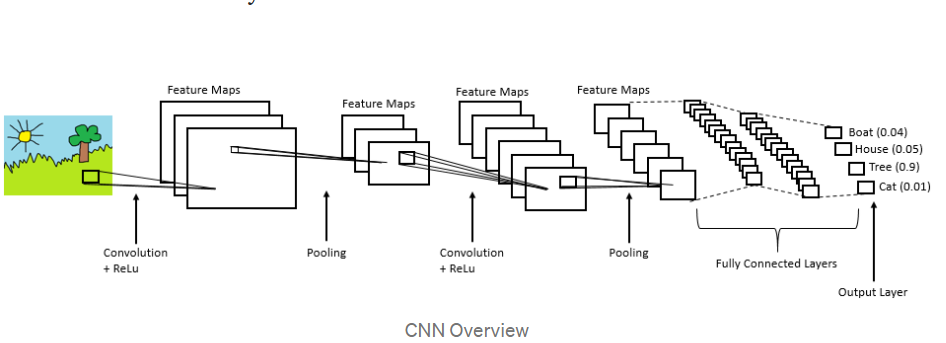
**3.3 Pooling**

The goal of this layer is to provide spatial variance, which simply means that the system will be capable of recognizing an object even when its appearance varies in some way. Pooling layer will perform a downsampling operation along the spatial dimensions (width, height), resulting in output such as [16x16x12] for pooling\_size= (2, 2).



**3.4 Fully Connected**

In a fully connected layer, we flatten the output of the last convolution layer and connect every node of the current layer with the other nodes of the next layer. Neurons in a fully connected layer have full connections to all activations in the previous layer, as seen in regular Neural Networks and work in a similar way.



**3.5 Image Classification**

The complete image classification pipeline can be formalized as follows:

1. Our input is a training dataset that consists of N images, each labeled with one of 2 different classes.
2. Then, we use this training set to train a classifier to learn what every one of the classes looks like.
3. In the end, we evaluate the quality of the classifier by asking it to predict labels for a new set of images that it has never seen before. We will then compare the true labels of these images to the ones predicted by the classifier.

**3.6 Data Augmentation**

Data augmentation is commonly used in computer vision. In vision, we can almost certainly flip, rotate, or mirror an image without risk of changing the original label. The practice of data augmentation is an effective way to increase the size of the training set. Augmenting the training examples allow the network to see more diversified, but still representative data points during training.

Then a data generator was created to get the data from our folders and into Keras in an automated way. Keras provides convenient python generator functions for this purpose.

**3.7 Building the model**

The network we’ll build will be a CNN (Convolutional Neural Network) and call it CancerNet. This is an updated model from the one presented in the presentation. This network performs the following operations:

Use 3×3 CONV filters

Stack these filters on top of each other

Perform max-pooling

Use depthwise separable convolution (more efficient, takes up less memory)

We use the Sequential API to build CancerNet and SeparableConv2D to implement depthwise convolutions. The class CancerNet has a static method build that takes four parameters- width and height of the image, its depth (the number of color channels in each image), and the number of classes the network will predict between, which, for us, is 2 (0 and 1)

In this method, we initialize model and shape. When using channels\_first, we update the shape and the channel dimension.

Now, we’ll define three DEPTHWISE\_CONV => RELU => POOL layers; each with a higher stacking and a greater number of filters. The softmax classifier outputs prediction percentages for each class. In the end, we return the model.

4 **Experimental Results**

**4.1 Datasets**

The datasets that were used in the experiment have been updated from the one proposed in the research proposal. The updated dataset is a Breast Cancer Histopathological Image Classification (BreakHis) composed of 9,109 microscopic images of breast tumor tissue collected from 82 patients using different magnifying factors (40X, 100X, 200X, and 400X). To date, it contains 2,480 benign and 5,429 malignant samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG format). This database has been built in collaboration with the P&D Laboratory – Pathological Anatomy and Cytopathology, Parana, Brazil (http://www.prevencaoediagnose.com.br). We believe that researchers will find this database a useful tool since it makes future benchmarking and evaluation possible.

The dataset BreaKHis is divided into two main groups: benign tumors and malignant tumors. Both breast tumors benign and malignant can be sorted into different types based on the way the tumoral cells look under the microscope. Various types/subtypes of breast tumors can have different prognoses and treatment implications. The dataset currently contains four histological distinct types of benign breast tumors: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA); and four malignant tumors (breast cancer): carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and papillary carcinoma (PC).

The dataset was was initialized into the path to the input dataset (datasets/original), that for the new directory (datasets/base), and the paths for the training, validation, and testing directories using the base path. We also declare that 80% of the entire dataset will be used for training, and of that, 10% will be used for validation.

This will split our dataset into training, validation, and testing sets in the ratio mentioned above- 80% for training (of that, 10% for validation) and 20% for testing. With the ImageDataGenerator from Keras, we will extract batches of images to avoid making space for the entire dataset in memory at once.

Then build a list of original paths to the images, then shuffle the list. Followed by calculating an index by multiplying the length of this list by 0.8 so we can slice this list to get sublists for the training and testing datasets. Next, we further calculate an index saving 10% of the list for the training dataset for validation and keeping the rest for training itself.

Now, datasets is a list with tuples for information about the training, validation, and testing sets. These hold the paths and the base path for each. For each setType, path, and base path in this list, we’ll print, say, ‘Building testing set’. If the base path does not exist, we’ll create the directory. And for each path in originalPaths, we’ll extract the filename and the class label. We’ll build the path to the label directory(0 or 1)- if it doesn’t exist yet, we’ll explicitly create this directory. Now, we’ll build the path to the resulting image and copy the image here- where it belongs.

**4.2 Performance Analysis**

The most common metric for evaluating model performance is accuracy. However, when only 2% of your dataset is of one class (malignant) and 98% some other class (benign), misclassification scores don’t really make sense. You can be 98% accurate and still catch none of the malignant cases which could make a terrible classifier.

**4.3** **Precision, Recall and F1-Score**

For a better look at misclassification, we often use the following metric to get a better idea of true positives (TP), true negatives (TN), false positive (FP) and false negative (FN).

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

Recall is the ratio of correctly predicted positive observations to all the observations in actual class.

F1-Score is the weighted average of Precision and Recall.

F = (2 \* (Precision \* Recall))/ (Recall + Precision)

The higher the F1-Score, the better the model. For all three metrics, 0 is the worst while 1 is the best.

**4.4 Confusion Matrix**

Confusion Matrix is a very important metric when analyzing misclassification. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class. The diagonals represent the classes that have been correctly classified. This helps as we not only know which classes are being misclassified but also what they are being misclassified as.

5  **Results**

The following below is the result of the deep learning program used in the classification of the datasets.

precision recall f1-score support

0 0.89 0.89 0.89 177

1 0.82 0.82 0.82 104

accuracy 0.86 281

macro avg 0.85 0.85 0.85 281

weighted avg 0.86 0.86 0.86 281

**6 Conclusion**

In summary, we have demonstrated that given data, deep learning methods such as convolutional neural networks can make an accurate diagnosis of breast cancer. But there is a need to pay special attention in making a consistent model that is free from errors because the datasets are not just numbers but a representation of a human condition where if a mistake is made, that may mean the end of the patient which represent that data. Future works will be to reinforces consistent accuracy and a model that is free from overfitting and noisy datasets.

**References**

[1] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural

networks", *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[2] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image on.,

pp. classification using convolutional neural networks", Neural Networks (IJCNN) 2016 International

Joint Conference 2560-2567, 2016.

[3] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A

Comprehensive tutorial with selected use cases", *Journal of natholovv informatics*, vol. 7, 2016.

[4] .Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, et al., "Classification

of breast cancer histology

images using convolutional neural networks", *PloS one*, vol. 12, no. 6, pp. e0177544, 2017.

[5] Lu Liangqun and Bernie Daigle, "Prognostic Analysis of Histopathological Images Using Pre-Trained

Convolutional Neural Networks", *bioRxiv*, 2019.

[6] Bejnordi Babak Ehteshami et al., "Diagnostic assessment of deep learning algorithms for detection of

Lymph node metastases in women with breast cancer", *Jama*, vol. 318, no. 22, pp. 2199-2210, 2017.

[7] Li Yi and Wei Ping, Cancer metastasis detection with neural conditional random field, 2018.

[8] Cao Hongliu et al., "Improve the performance of transfer learning without fine-tuning using *and*

*recognition*, dissimilarity-based multi-view learning for breast cancer histology images", *International*

*conference image analysis* 2018.

[9] S. Angara, M. Robinson and P. Guillén-Rondon, "Convolutional Neural Networks for Breast

Cancer Histopathological Image Classification," 2018 4th International Conference on Big

Data and Information Analytics (BigDIA), 2018, pp. 1-6, doi: 10.1109/BigDIA.2018.8632800.

[10] R. Roslidar, K. Saddami, F. Arnia, M. Syukri and K. Munadi, "A Study of Fine-Tuning CNN Models

Based on Thermal Imaging for Breast Cancer Classification," 2019 IEEE International Conference

On Cybernetics and Computational Intelligence (CyberneticsCom), 2019, pp. 77-81, doi:

10.1109/CYBERNETICSCOM.2019.8875661.

[11] M. F. Jamaluddin et al., "Cell Classification in ER-Stained Whole Slide Breast Cancer Images Using

Convolutional Neural Network," 2018 40th Annual International Conference of the IEEE

Engineering In Medicine and Biology Society (EMBC), 2018, pp. 632-635, doi:

10.1109/EMBC.2018.8512386.

[12] N. Ojha and A. Kumar, "A Comparison Based Breast Cancer High Microscopy Image Classification

Using Pre- trained Models," 2020 IEEE Students Conference on Engineering & Systems (SCES), 2020,

pp. 1-6, doi: 10.1109/SCES50439.2020.9236708.