Breast cancer is the most common cancer in women. Recent statistics show that breast cancer affects one of every ten women in Europe and one of every eight in the US. The American Cancer Society estimates 40,920 women will die due to breast cancer in the US in 2018.

Early detection of the breast cancer can decrease considerably the death rate due to this dangerous type of cancer. Mammography is the most widely used for this purpose. However, mammography is still a manual process, quite prone to human error. Significant works have been done in the recent decades regarding mass detection using advanced methods like AI based ones, particularly deep convolutional neural networks (CNN). The main advantage of deep CNN models is generating high-level features compared to the previous hand-built features.

In this project we used a deep learning algorithm for breast cancer tumor classification. The basic idea was taken from a paper that did the classification using the traditional neural network on the same data set. In more details, we worked on a data set involves 322 breast images which are divided in three groups: malignant, benign, and normal. This research has been categorized in two steps: preprocessing and tumor classification using deep learning algorithms. The software we used for this research was MATLAB R2017a.

**Approach**

We have proposed a two-stage algorithm for classification of breast tumors as follows:

**Preprocessing:** The original breast images include some amount of noise in addition to have a low contrast quality. Also, some of the images have been taken from the left breast and some of them from the right breast. For solving the first two challenges we used thresholding and contrast enhancement on the images. For the third problem, we flipped all the left breast images to the right. Finally, we removed the muscle part from each image by use of thresholding again to work with only the breast tissue of the image in the future steps of the project.

**Tumor classification:** After obtaining the preprocessed image, we fed the images to a deep learning algorithm. In more detail, we used AlexNet topology for this task. The AlexNet topology has 25 hidden layer and it is a pretrained algorithm. The output of the algorithm will be the class of the tumor at the breast image which can be benign, malignant or normal.

**Conclusion**

In the previous figure you can see our results (confusion matrix), beside the results of the reference paper. We can say that the result of the deep neural network is good but not completely acceptable yet. Also it can be found that our results are not as good as the reference paper’s result which used traditional neural network. The reasons are:

- In the reference paper, they used exactly the tumor area as input of neural network as well as extracting pre-defined features from this ROI, however our images are not zoomed in and we did not extract any features.
- We used an AlexNet already trained by images with general topics (RGB or grayscale) therefore maybe we are not using the best low-level features.

It can be a good idea to try the network at the future with other alternative network architecture if possible or study the possibility of pre-training the AlexNet with the current dataset, although we would have the small dataset limitation.

**Bibliography**