Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus Querétaro

Breast cancer diagnosis from mammography images using Deep Learning

María de los Ángeles Contreras Anaya

19/08/2022

Summary

Breast cancer is the most common cancer among women worldwide, and it is one of the leading causes of death across women of ages 35 to 54. It has been proven that breast cancer screening with mammography images reduces about a 30% the mortality rate associated to the disease, sadly studies have shown that its usage leads to high false-positive rates and thus endangers the patient's wellbeing. Approximately 70% of biopsies made from the result of mammography images interpretations are unnecessary, and that is why the usage of Deep Learning (DL) to diagnose breast cancer has been increasing in popularity. Studies using Deep Learning techniques to diagnose breast cancer show sensitivity between 72% and 85% and a specificity between 86% and 88% which remains to be improved and hopefully bring us a step closer to using this DL models to aid breast cancer diagnosis in the practice. To improve the results found on the state-of-the-art, a Convolutional Neural Network (CNN) model will be designed and implemented using the CCD-CESM dataset that is a collection of standard and Contrast Enhanced Spectral Mammography (CESM) images.

Abbreviations

AI	Artificial Intelligence
DL	Deep Learning
DNN	Deep Neural Network
CNN	Convolutional Neural Network
CADx	Computer-Assisted Diagnosis
CADe	Computer-Assisted Detection
FDA	Food and Drug Administration
CESM	Contrast Enhanced Spectral Mammography

I. Introduction and background

Every thirteen minutes, a woman dies from breast cancer [1]. Breast cancer is the most common cancer among women worldwide and the second leading cause of death across women of ages between 35 and 54. Approximately, 287,850 cases of invasive breast cancer will be diagnosed in women in 2022 [1]. Breast cancer is a disease in which cells in the breast mutate and grow in an uncontrolled way, thus creating a mass of tissue called a tumor. Early diagnosis reduces the mortality of breast cancer patients, that is why screening with mammography images was introduced.

Mammography is the most cost-effective imaging modality to date and is proven to reduce mortality in approximately 30% [2]. Nonetheless, some lesions are often ignored because of various reasons, whether it is dense breast tissue present in about half of the screened breasts, false positioning at time of screening or interpretation error [3]. Mammography, involves the acquisition of a single two-dimensional image of the breast. These images can end up having tissue superposition that covers up the presence of malignant lesions, reducing sensitivity or mimicking lesions, which reduces specificity. These limitations of mammography images gave rise to the usage of computer assisted diagnosis (CADx) tools.

CADx tools have been used in screening mammography since they were approved by the FDA in 1988, but have failed to improve drastically the performance of screening because they use programmed-in features [4]. The algorithms of these tools search for specific features that humans identify as representative of suspicious lesions, thus they may be biased or error-prone [2]. In an attempt to eliminate human intervention and improve both the sensitivity and the specificity of diagnosis in the early stages of the disease, the usage of Artificial Intelligence (AI) more specifically Deep Learning algorithms has been increasingly growing.

II. Problem statement and justification

AI-assisted software based on Deep Learning models for breast cancer detection and diagnosis have shown an increase in the specificity and sensitivity, surpassing the performance of CADx tools, however this metrics can certainly be improved with the design and implementation of a better Deep Neural Network (DNN) model or with the usage of Contrast-Enhanced Spectral Mammography (CESM) images. This type of images are specialized mammograms that use iodinated intravenous contrast in conjunction with a standard mammogram, thus improving the visibility of some lesions.

According to some studies, CESM images have a better diagnostic accuracy as compared with standard mammography images [5] which is why my proposed solution to achieve a higher sensitivity and specificity in breast cancer diagnosis is to design and implement a Convolutional Neural Network but with the CDD-CESM dataset. The dataset is a collection of low-energy images with their corresponding subtracted CESM images, gathered over the period from January 2019 to February 2021 from the Radiology Department of the National Cancer Institute [6].

To evaluate my proposed solution, the metrics of recall, precision and F1 score will be used in conjunction with the confusion matrix. These evaluation metrics were selected considering we are dealing with a health-related topic. Recall will be used since our principal goal is to diagnose the majority of malignant cancers, but I will also be looking at precision since this is the current problem of existing CADx and CADe tools, given that there is a negative impact on the wellbeing of patients and the incurred costs of unnecessary biopsies if results are not precise. Because of the importance of both the recall and precision in this topic, I decided to also use F1 score, since it is the weighted average of the above-mentioned metrics. Additionally, a confusion matrix will be obtained to perform a deeper analysis of my model.

Since there are no other studies where CSEM images are used, instead of standard mammograms, this study is the first step to see the impact of this specialized images on the new generation of image interpretation methods. Furthermore, a greater performance is expected since studies show CESM images by themselves are far more accurate than standard mammograms, thus increasing the effectivity and hence it will bring us one step closer to seeing the full potential and real-world impact of these methods.

III. State of the Art

Software aided detection and diagnosis of breast cancer is not a new topic, CADe and CADx tools have been used in the medical field since 1998 nonetheless researchers have employed different Deep Learning methods instead of these traditional methods because of the rate of false-positive cases that it has and thus the impact to the wellbeing of diverse patients.

A. J. T. Wanders et al. [8] proposed a new model for the detection of interval cancer, which combined the usage of AI, specifically Neural Networks and breast density (BD) assessments in mammography images. The combined assessment was significantly better than either method carried out alone, achieving a sensitivity of 50.9% at a 90% specificity in comparison to the BD 22.4% or the AI system 37.5% specificity. These results showed improvement in current CADe or CADx systems, but there is an improvement opportunity if instead of using a Machine Learning algorithms we opt for using a DL one which is far more powerful.

H. J. Kim et al. [4] proposed the usage of a Deep Learning AI-based diagnosis supporting software for beast cancer detection and diagnosis. The software was developed using a deep CNN. The results showed that the software was able to diagnose correctly 31.3% of occult cancers in mammography images, thus reducing the rate of mortality of women with this disease and discovered the effectiveness of AI in detecting cancer on dense breasts. Although an improvement in diagnosis rate was mentioned, there was a lack of data of the specificity and sensitivity metrics of the AI-software, thus impeding us to make further comparisons.

I. Kizildag Yirgin et al. [3] evaluated the performance of AI algorithms in a simulated screening setting and its effectiveness in detecting missed and interval cancers. The proposed model used the Lunit INSIGHT MMG, the same AI- based software from the previous study, to diagnose breast cancer from 211 mammograms. Detection rates of radiologists in the screening program were compared with the AI system results, which showed that cancer detection rates went up from 67.3% to 72.7% as compared with radiologists. Furthermore, a combination between radiologists and the software was evaluated obtaining an 83.6% detection rate, concluding that AI may assist professionals in cancer screening programs, with the best result of sensitivity being 72.8% and specificity of 88.3%.

IV. Objectives

The main objective of the present study is to diagnose breast cancer using Artificial Intelligence (AI) specifically by implementing a Deep Learning algorithm. This objective will be accomplished by completing the secondary objectives described on image 1.

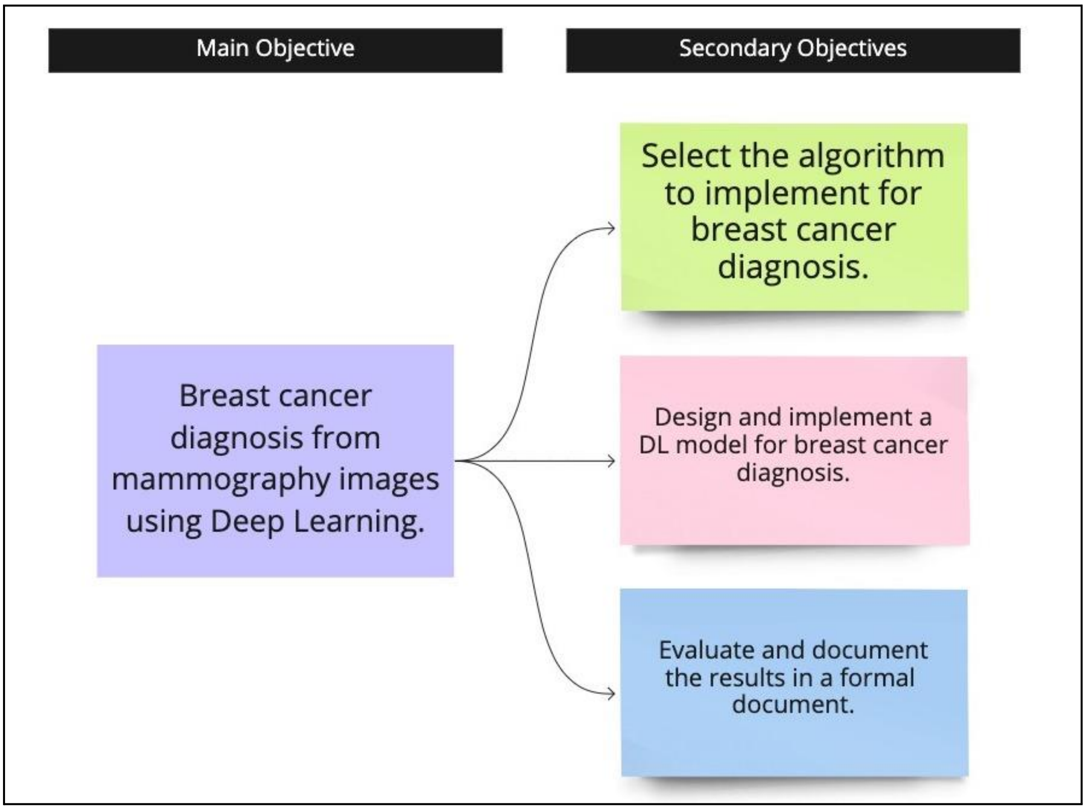

Image 1. Definition of the main objective of this study and its mapping to the secondary objectives.

V. Expected contributions and deliverables

The expected contributions of this research are:

- A Deep Learning model with greater sensitivity and specificity than the AI-software or models in the state of the art.
- A baseline for future research on the impact of the usage of CSEM images combined with Deep Learning models for the diagnosis of breast cancer.
- A preprocessed dataset with the new image modality for breast cancer screening, CSEM images.
- A Deep Learning model available to all public that will enable future researchers to tune it and try to improve the performance of it.

These contributions are mapped directly to the deliverables that will be the outcome of this investigation:

- Research proposal
- Preprocessed dataset
- Code for the implementation of the CNN algorithm for breast cancer diagnosis.
- Undergraduate dissertation

Finally, these contributions will bring us one step closer to using deep learning models in practice to diagnose breast cancer and therefore enable the detection of the disease in early stages which consequently will decrease the cases of mammography missed cancers, interval cancers and unnecessary biopsies. By decreasing the mentioned cases, mortality rate will also decrease, given that cancer caught in early stages has a 5-year relative survival rate of 99%.

VI. Methodology and work plan

To achieve the proposed solution the defined objectives were mapped against the tasks and final deliverables, this mapping is shown in image 2. The image also shows the sequence in which activities will be performed.

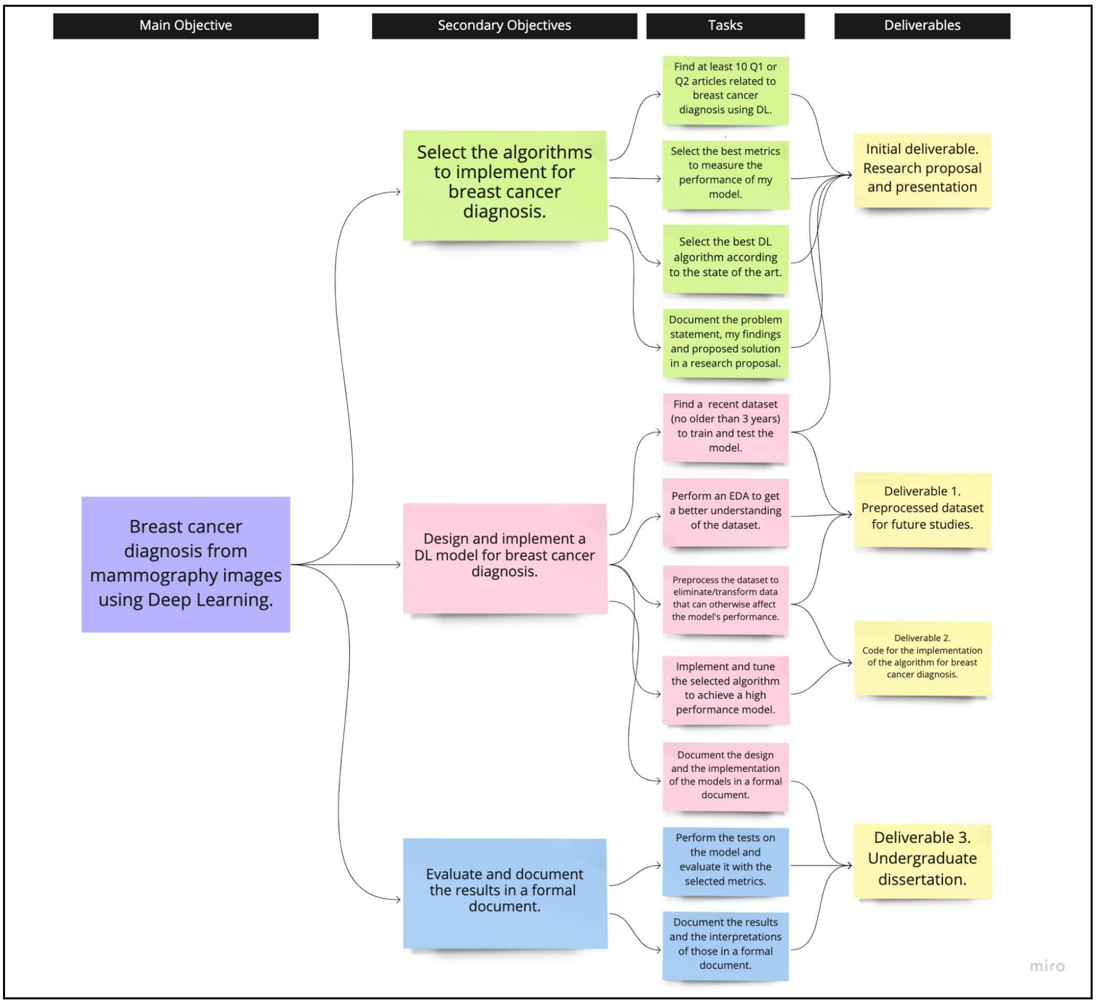

Image 2. Relationship mapping between the defined objectives, tasks to be performed and final deliverables.

All the code generated for the proposed solution will be developed using the high-level programming language Python. A detailed description of the frameworks and libraries that will be used is presented below:

- Pandas for data analysis and manipulation
- Seaborn library for data visualization
- Scikit-image library for image processing

- Scikit-learn library for model’s evaluation
- Keras for the CNN design and implementation running on top of the open-source framework for Deep Learning TensorFlow

Finally, image 3 shows all activities in the given timeframe, this work plan was made to ensure all objectives will be met before the final deadline.

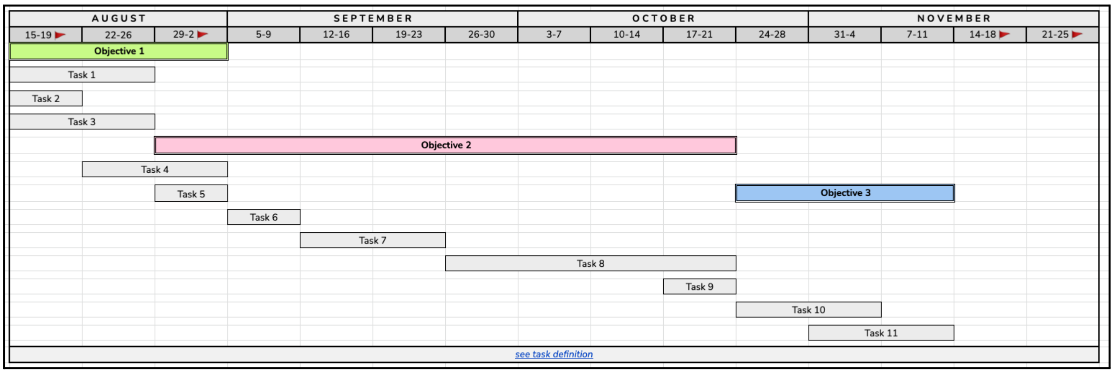

Image 3. Sequence and time-frame that will be dedicated to complete each task.

VII. References

[1]
"Breast Cancer Statistics | Facts & Figures | NBCC," *National Breast Cancer Coalition*, Jan. 14, 2022. https://www.stopbreastcancer.org/information-center/facts-figures/ (accessed Aug. 20, 2022).

[2]
I. Sechopoulos, J. Teuwen, and R. Mann, "Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art," *Seminars in Cancer Biology*, Jun. 2020, doi: 10.1016/j.semcancer.2020.06.002.

[3]
I. Kizildag Yirgin *et al.*, "Diagnostic Performance of AI for Cancers Registered in A Mammography Screening Program: A Retrospective Analysis," *Technology in Cancer Research & Treatment*, vol. 21, p. 153303382210751, Jan. 2022, doi: 10.1177/15330338221075172.

[4]
H. Li *et al.*, "Application of deep learning in the detection of breast lesions with four different breast densities," *Cancer Medicine*, Jun. 2021, doi: 10.1002/cam4.4042.

[5]
M. Helal, M. F. Abu Samra, M. A. Ibraheem, A. Salama, E. E. Hassan, and N. E.-H. Hassan, "Accuracy of CESM versus conventional mammography and ultrasound in evaluation of BI-RADS 3 and 4 breast lesions with pathological correlation," *The Egyptian Journal of Radiology and Nuclear Medicine*, vol. 48, no. 3, pp. 741–750, Sep. 2017, doi: 10.1016/j.ejrnm.2017.03.004.

[6]
R. Khaled *et al.*, "Categorized contrast enhanced mammography dataset for diagnostic and artificial intelligence research," *Scientific Data*, vol. 9, no. 1, p. 122, Mar. 2022, doi: 10.1038/s41597-022-01238-0.

[7]
H. J. Kim *et al.*, "Mammographically occult breast cancers detected with AI-based diagnosis supporting software: clinical and histopathologic characteristics," *Insights into Imaging*, vol. 13, no. 1, Mar. 2022, doi: 10.1186/s13244-022-01183-x.