Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

Escuela Nacional de Ingeniería y Ciencias

Programa de Graduados

Maestría en Ciencias en Intelligent Systems

Propuesta de Tesis

Early Detection and Diagnosis of Breast Cancer Lesions using Deep Convolutional Networks in Digital Mammograms

por

Erick Michael Cobos Tandazo
1184587

Monterrey, N.L., April 7 de 2015

Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

Escuela Nacional de Ingeniería y Ciencias

Programa de Graduados

Los miembros del comité de tesis recomendamos que la presente propuesta de Erick Michael Cobos Tandazo sea aceptada para desarrollar el proyecto de tesis como requisito parcial para obtener el grado académico de **Master in Science**, especialidad en:

Intelligent Systems

Comité de Tesis:

Dr. Hugo Terashima Marín
Asesor Principal

Por definir
Sinodal

Por definir
Sinodal

Dr. Ramón Brena Pinero
Director del Programa de Maestría en
XXXX

April 7 de 2015

Contents

Abstract

Breast cancer is one of the most common and deadly cancers in woman around the world and the best tools used today for early diagnosis are screening mammograms. Mammograms are x-ray pictures of the breast used by radiologists to identify micro-calcifications and breast masses, signs of early breast cancer development. Traditional computer systems use handmade features and complex image techniques to detect these lesions in mammographic images. In this work, we plan to use convolutional networks, a recent development in computer vision, which can automatically learn the important features for the task given enough training data. Convolutional networks have been used in some studies for breast cancer detection but we hope to introduce newer features and carefully tune the architecture to produce improved results. In addition, this will be the first approximation to use deep learning techniques as part of an ongoing project in the institution which aims to develop a computer-aided diagnosis system for breast cancer. This thesis proposal is presented for approval to obtain the degree of Master of Science in Intelligent Systems.

1 Introduction

Yet to write

2 Problem Definition

Breast cancer is the most commonly diagnosed cancer in woman and its death rates are among the highest of any cancer. It is estimated that about 1 in 8 U.S. women will be diagnosed with breast cancer at some point in their lifetime [Howlader et al., 2014]. Early detection is key in reducing the number of deaths from breast cancer; detection in its earlier stage (*in situ*) increases the survival rate to virtually 100% [Howlader et al., 2014].

With current technology, a high quality mammogram is "the most efective way to detect breast cancer early" [National Cancer Institute, 2014]. Mammograms are X-ray images of each breast used by radiologists to search for early signs of cancer such as tumors or microcalcifications. About 85% of breast cancers can be detected with a screening mammogram [Breast Cancer Surveillance Consortium, 2013]. This high sensitivity is the product of the careful examination of the mammograms by experienced radiologists. A computer-aided diagnosis tool (CAD) could automatically detect and diagnose these abnormalities saving the time and training needed by expert radiologists and avoiding any human error. Computer based approaches could also be used by radiologists as a help during the screening proccess or as a second informed opinion on a diagnostic.

CAD systems are based on image and classification techniques coming from Artificial Intelligence and Machine Learning. Traditional CAD tools for breast cancer diagnosis are composed of three steps: feature extraction, feature selection and classification. In the feature extraction phase, the system uses filters and image transformations to preprocess the mammogram and find geometric patterns which are used to produce a set of features for the image; expert knowledge is sometimes used in this phase. Feature selection or regularization is used to focus only on the important features for the classification task. Once a vector of features is obtained for each image, an standard binary classifier can be used to perform the final detection or diagnosis. These techniques have been used for many years and are

standard in the industry [1].

Despite its widespread use and efficiency, systems based on traditional computer vision techniques have various limitations that should be addressed to further improve its performance:

- There is no standard way of preprocessing mammograms. Some filters are commonly used but their performances can vary.
- It uses handcrafted features. The features extracted from the image are chosen beforehand (maybe designed with the help of experts) and special filters and image techniques are used to extract them.
- It normally uses a small patch of the mammogram and makes a prediction on that patch but it does not consider the entire mammogram neither to make a prediction on the patient or to account for correlation between patches.
- To produce good results it requires knowledge in various fields such as radiology, oncology, image processing, computer vision, machine learning, etc.
- It is composed of many sequential steps. At each stage, there are many techniques from which the researcher can choose and many parameters which have to be estimated. This represents a cost in time and results as it is improbable that the optimal selection of techniques and parameters is achieved.
- As it is a complex system with different subsystems involved many other issues can arise such as non desired or unknown dependencies between subsystems, difficulty to localize errors, maintainability, etc.
- The techniques currently used are complex but the improvements achieved are not substantial. Much work is needed to make only incremental improvements and it is hard to know to which part of the system dedicate more resources.

This project will center around using Convolutional Networks, a recent development in computer vision, to tackle some of these limitations, especifically automate preprocessing and feature extraction, use entire mammogram images and simplify the system pipeline by using a convolutional network as a replacement for many steps traditionally performed in succesion.

3 Objectives

The main goal of this work is to succesfully apply convolutional networks in digital mammograms to detect and diagnose breast cancer lessions, microcalcifications and breast masses, and to compare our results to those obtained by other groups working in convolutional networks for breast cancer diagnosis.

Particularly, there are various subgoals which we expect to achieve as the project advances:

- Develop a working pipeline for processing the mammographic images from our database and training a convolutional network. Essentially, this tool could also be used for other image classification tasks.

[1] See [Hernandez, 2014] for an example of a CAD system developed in this institution.

- Kickstart the research on deep learning in the institution.
- Use a simple convolutional network to perform detection and diagnosis and study these initial results to guide further research.
- Show the viability of convolutional networks for breast cancer diagnosis.
- Use convolutional networks on an entire mammogram instead of only on small patches.
- Improve the performance of convolutional networks reported on the literature.
- Generate results that could result in a conference or journal article.
- Propose new ideas and methods for future research in the topic.

Initial exploratory research has not yet been performed and some of these particular objectives may be modified as the project progresses. Furthermore, some new research avenues could be taken if they seem promising, for instance, using convolutional networks with digital tomosynthesis images (3-dimensional X-ray images of the breast).

4 Hypotheses

Although a considerable amount of work on breast cancer detection and diagnosis has been done in the institution, this project will be the first approximation to using convolutional networks for efficiently detecting and diagnosing breast cancer. Convolutional networks are widely used for object recognition tasks and have shown very good results [Russakovsky et al., 2014, Taigman et al., 2014, Dieleman et al., 2015]. They have a big research community and have become one of the preferred methods to perform image classification tasks.

Due to the exploratory nature of this work we are not truly certain of the results that will be obtained. Nevertheless, we have a well established idea of what we expect to obtain. We will apply some of these newly developed techniques expecting to produce similar or better results than those obtained using more traditional computer vision techniques. We believe that implementing convolutional networks for mamographic images will not be very difficult as it has already been done (see Section 5.7). We do not think that a simple convolutional network will suffice to obtain acceptable results; we will need to use a more refined convolutional network with well fitted parameters.

4.1 Research Questions

Some of the questions which will be answered in this work are:

- Can we improve the results reported by other groups using convolutional networks? Is training a convolutional network on mammographic images better than computing numeric features from the mammograms and training a simple classifier?
- Is deep learning feasible with the resources we have? Is the data and computational power we posses enough? Is there any advantage to use GPU acceleration?

- Can we simplify the pipeline for breast cancer diagnosis? Can preprocessing be replaced by more layers on the same convolutional network? Could we use an entire mammogram for diagnosis instead of only small patches or could we automatically join results for small patches to generate results on the entire mammogram?
- What are the best parameters for our convolutional networks (number of layers, number of units, kernel sizes, regularization, activation functions, etc)? Is there a big improvement on refining the network and tuning parameters?
- What are the advantages of using a deep versus a shallow convolutional network?
- Could we use a convolutional network trained on a different database (such as the ImageNet database) to obtain features for mammographic images and use these features for classification?
- Are convolutional networks a good option for future research?

5 Background

We offer an introduction to some of the essential concepts needed to understand the rest of this document. We start by discussing breast cancer and mammograms in Section 5.1, we describe the specifics of the mammography database used for this thesis in Section 5.2, we offer some basic concepts about classification and evaluation measures in Section 5.3, in Sections 5.4 and 5.5 we give a short introduction into Artificial Neural Networks and Convolutional Neural Networks and finally we offer an overview of how convolutional networks have been used for breast cancer diagnosis in Sections 5.6 and 5.7.

5.1 Breast Cancer

Cancer is an umbrella term to refer to a group of diseases caused by abnormal cell growth in different parts of the body. The accumulation of extra cells usually forms a mass of tissue called a *tumor*. Tumors can be benign or malignant: *benign tumors* are noncancerous, lack the ability to invade surrounding tissue and will not regrow if removed from the body; malignant or *cancerous tumors* are harmful, can invade nearby organs and tissues (*invasive cancer*), can spread to other parts of the body (*metastasis*) and will sometimes regrow when removed [National Cancer Institute, 2012].

Breast cancer is the cancer that forms in tissues of the breast. The two most common types of breast cancer are *ductal carcinoma* and *lobular carcinoma*; these cancers begin in the breast ducts and lobules, respectively (see Fig. 1). Breast cancer *incidence rate*, the number of new cases in a specified population during a year, is the highest of any cancer among American women. Its *mortality rate*, the number of deaths during a year, is also one of the highest of any cancer [Howlader et al., 2014].

The *cancer stage* depends on the size of the tumor and whether the cancer cells have spread to neighboring tissue or other parts of the body. It is expressed as a Roman numeral ranging from 0 through IV; stage I cancer is considered *early-stage breast cancer* and breast cancer at stage IV is considered *advanced*. Stage 0 describes non-invasive breast cancers, also known as *carcinoma in situ*. Stage I, II and III describe invasive breast cancer, i.e., cancer that has invaded normal surrounding breast tissue. Stage IV is used to describe metastatic cancer, i.e., breast cancer has spread beyond nearby tissue to other organs of the body.

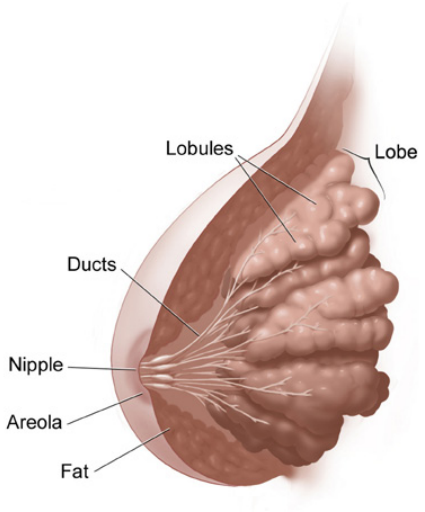

Figure 1: Anatomy of the female breast. Image courtesy of NCI.

5.1.1 Mammograms

A *mammogram* is an x-ray image of the breast. *Screening mammograms* (normally composed of two mammograms of each breast) are used to check for breast cancer signs on women who have not shown symptoms of the disease. If an abnormality is found, a *diagnostic mammogram* is ordered, these are detailed x-ray pictures of the suspicious region [National Cancer Institute, 2014]. A standard mammogram is shown in Fig. 2.

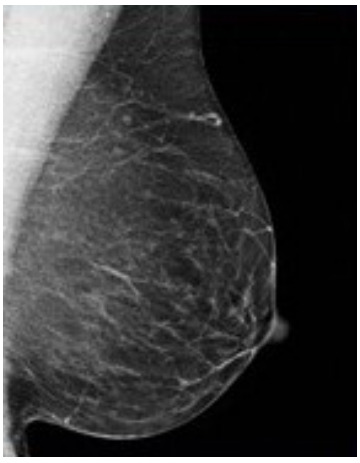

Figure 2: A standard mammogram.

Having a screening mammogram in a regular basis is the most effective method for detecting early breast cancer; around 85% of breast cancers can be detected in a screening mammogram [Breast Cancer Surveillance Consortium, 2013]. Nevertheless, screening mammograms have many limitations: a high false positive rate, overtreatment in Stage 0 cancer, false negative results for women with high breast density, radiation exposure and physical and psychological discomfort [National Cancer Institute, 2014].

Mammograms are read by expert radiologists. The radiologist looks primarily for micro-

calcifications and breast masses. *Microcalcifications* are tiny deposits of calcium in the breast tissue which can be a sign of early breast cancer if found in clusters with irregular layout and shapes. *Breast masses* or breast lumps are possibly a variety of things: fluid-filled cysts, fatty tissues, fibric tissues, noncancerous or cancerous tumors, among others. A mass can be a sign of breast cancer if it has poorly defined shape and margins. See Fig. 3 for an example of possible signs of breast cancer. Radiologists will also consider the breast density of the patient when reading a mammogram given that high breast density is linked to a higher risk of breast cancer and it also difficults the interpretation of the mammogram [American Cancer Society, 2014].

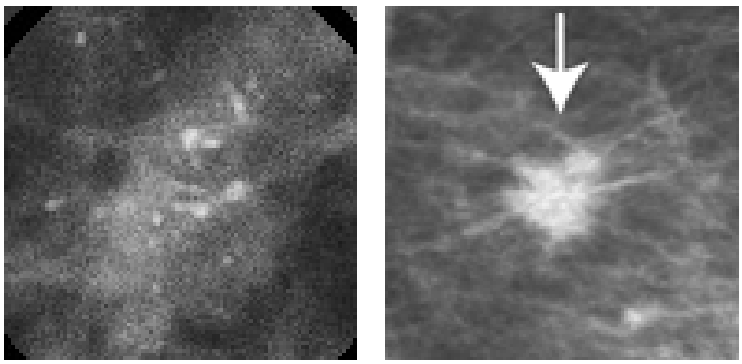

Figure 3: Signs of possible breast cancer in a mammogram. Left: A cluster of microcalcifications in an irregular layout. Right: A poorly defined breast mass.

Conventional mammography uses film to record x-ray images of the breast. *Digital mammography* on the other hand uses digital receptors to convert the x-rays into electric signals and stores the image electronically. Digital mammograms offer a clearer picture of the breast and can be digitally manipulated and shared between health care providers. Its effectiveness to identify breast cancer over film mammograms, however, is still debated [Kerlikowske et al., 2011, Pisano et al., 2008, Skaane et al., 2007]. Digital mammography is steadily becoming the standard for breast cancer screening, Fig. 2 is, in fact, a digital mammogram.

Digital tomosynthesis, also called three-dimensional mammography, is a new technology that essentially produces 3-dimensional x-ray images of the breast and is expected to improve the efficacy of regular 2-d mammograms. Studies comparing the two techniques have not yet been published, though [National Cancer Institute, 2014].

In this thesis we will center on using mammograms, either digital or manually digitized from film, to detect microcalcifications and masses and predict the likelihood of breast cancer on the patient.

Much of this section was written using information from the National Cancer Institute. We recommend to visit its website (`www.cancer.gov`) for more information.

5.2 Classification

Yet to write

5.3 Artificial Neural Networks

5.4 Convolutional Networks

5.5 Convolutional Networks applied to Breast Cancer

5.6 Related Work

In this section we offer a summary of some of the first work on using convolutional networks for breast cancer diagnosis as well as some of the articles that have influenced this thesis.

[Lo et al., 1995] were the first group to use convolutional networks for breast cancer detection. They used a CNN with two hidden layers to detect microcalcifications. A high sensitivity image processing technique was used to obtain a set of 2104 patches (16 by 16 pixels) of all potential disease areas from 68 digital mammograms; of these, 265 were true microcalcifications and 1821 were "false subtle microcalcifications". Prior to training the CNN, a wavelet high-pass filtering technique was used to remove the background of these images. Each image was flipped over (left-right) and 4 rotations for each the original and flipped images were used for training (0°, 90°, 180° and 270°). The CNN was composed of one input unit (16×16), 12 units in the first hidden layer (12×12), 12 units in the second hidden layer(8×8) and two output nodes(one for YES and one for NOT). The input size (16), number of hidden layers (2) and kernel size (5×5) was obtained via cross validation, altough not many other options were explored: they tried input sizes of 8, 16 or 32, one or two hidden layers and kernel sizes of 2, 3, 5 or 13. The CNN reached 0.87 average AUC when identifying individual microcalcifications and 0.97 AUC for clustered microcalcifications. Only a minimum of three calcifications was considered a detection. Sensitivity and specificity test results were not reported. This article proved that simple convolutional networks can be efficiently used for medical image pattern recognition.

6 Methodology

In order to achieve the proposed objectives and test our hypotheses we will need to carry out various tasks. We list them here in the order in which we plan to execute them:

- Literature review: A thorough review of the published work using the databases and resources available in the institution. By the end of this task, a complete theoretical background should be obtained and written. This will also help refine the scope of the project and the experiments to be conducted.
- Software review: Once a clear idea of what are the possible experiments to be executed, we will need to find appropiate software to perform them. Software for database managing, preprocessing and implementation of different neural networks should be either located or developed.
- Database preprocessing: We will ready the database images for the experiments; these implies joining different databases, obtaining the required features, preprocessing the images, assigning labels, etc.
- Assesing image preprocessing: We will train a standard convolutional network with fixed parameters on mammograms with three different preprocessings: no preprocessing, image enhancement using median or gaussian filters and wavelet filtered images.

Furthermore, we will train a deeper convolutional network on nonpreprocessed images. We want to answer three research questions: which is the best preprocessing for convolutional networks, is using the best filter significantly better than using nonpreprocessed images and can a convolutional network automatically preprocess the images?

- Exploratory experiments: We will train standard convolutional networks in two different inputs: small image patches obtained from mammograms and whole mammogram images. We will also train a linear classifier, probably rectified linear units, on the features obtained from a convolutional network trained on the ImageNet database, i.e., we will use an already trained convolutional network instead of one trained specifically in mammograms. Here we will use the image preprocessing technique that showed better results in the previous step. We want to answer two research questions: Can a convolutional network trained on whole mammograms perform as well as one trained on small patches and can we use an already trained convolutional network to classify mammograms?

- Model selection: Using the insights from previous sections and the current literature on convolutional networks, we will select a network architecture along with novel features, preprocessing, training and regularization procedures. We aspire to find the best convolutional network configuration for mammogram classification.

- Further experiments: We will train the chosen convolutional network on our mammographic database. We will perform crossvalidation to adjust the most important network parameters and use regularization to avoid possible overfitting. We want to answer two research questions: is the performance of the convolutional network considerably improved by parameter tuning and, more importantly, is this a good performance?.

- Gathering results: Produce results on the test set and elaborate figures and tables. This could be obtained directly from software output or from further program executions.

- Reporting results: Write the thesis and any article or technical guide which may result from this work. Both this and the previous step will be performed along the execution of the project, hopefully benefiting from the supervisors' feedback.

Finally, we would like to note that this is an idealized workflow and some changes may occur due to time limitations or lack of resources. In the unlikely case that the work is finished before the project deadline, we will either reiterate on model selection, experiments, result gathering and reporting or look into digital tomosynthesis, network ensembles or evolving convolutional networks.

7 Work Plan

We present here the expected work plan for this master's thesis. A description of the activities can be found in Section 6

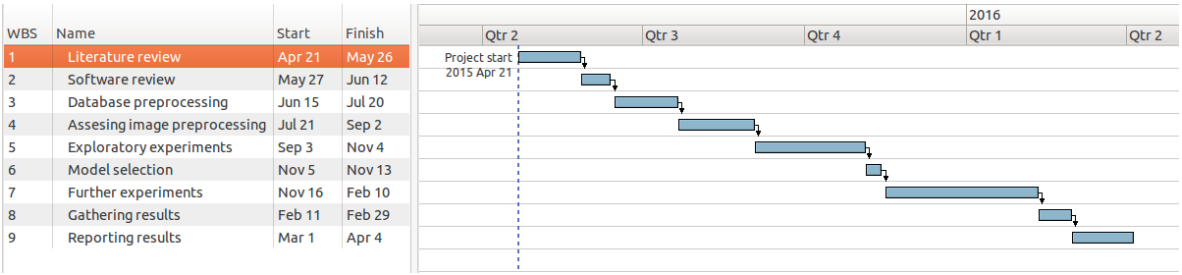

Figure 4: Thesis work plan.

References

[American Cancer Society, 2014] American Cancer Society (2014). Mammograms and other breast imaging tests. electronic, Atlanta, GA. Available online on `cancer.org/acs/groups/cid/documents/webcontent/003178-pdf.pdf`.

[Breast Cancer Surveillance Consortium, 2013] Breast Cancer Surveillance Consortium (2013). Performance measures for 1,838,372 screening mammography examinations from 2004 to 2008 by age–based on BCSC data through 2009. electronic. Available on `breastscreening.cancer.gov/statistics/performance/screening/2009/perf_age.html`.

[Dieleman et al., 2015] Dieleman, S., Willett, K. W., and Dambre, J. (2015). Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*.

[Hernandez, 2014] Hernandez, J. L. (2014). Selección de características y clasificacíon de masas por medio de redes neuronales, máquinas de vector de soporte, análisis discriminante y regresíon logística en mamografías digitales. Master's thesis, Tecnológico de Monterrey, Monterrey, Mexico.

[Howlader et al., 2014] Howlader, N., Noone, A. M., Krapcho, M. F., Garshell, J., Miller, D. A., Altekruse, S. F., Kosary, C. L., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A. B., Lewis, D. R., Chen, H. S., Feuer, E. J., and Cronin, K. A. (2014). SEER cancer statistics review, 1975-2011. review, National Cancer Institute, Bethesda, MD. Available on `seer.cancer.gov/csr/1975_2011/`.

[Kerlikowske et al., 2011] Kerlikowske, K., Hubbard, R. A., Miglioretti, D. L., Geller, B. M., Yankaskas, B. C., Lehman, C. D., Taplin, S. H., and Sickles, E. A. (2011). Comparative effectiveness of digital versus film-screen mammography in community practice in the united states: A cohort study. *Annals of Internal Medicine*, 155(8):493–502.

[Lo et al., 1995] Lo, S.-C. B., Chan, H.-P., Lin, J.-S., Li, H., Freedman, M. T., and Mun, S. K. (1995). Artificial convolution neural network for medical image pattern recognition. *Neural Networks*, 8(7–8):1201 – 1214. Automatic Target Recognition.

[National Cancer Institute, 2012] National Cancer Institute (2012). What you need to know about breast cancer. electronic, Bethesda, MD. Available online on `cancer.gov/publications/patient-education/WYNTK_breast.pdf`.

[National Cancer Institute, 2014] National Cancer Institute (2014). Mammograms. electronic. Available on `cancer.gov/cancertopics/types/breast/mammograms-fact-sheet`.

[Pisano et al., 2008] Pisano, E. D., Hendrick, R. E., Yaffe, M. J., Baum, J. K., Acharyya, S., Cormack, J. B., Hanna, L. A., Conant, E. F., Fajardo, L. L., Bassett, L. W., D'Orsi, C. J., Jong, R. A., Rebner, M., Tosteson, A. N. A., and Gatsonis, C. A. (2008). Diagnostic accuracy of digital versus film mammography: Exploratory analysis of selected population subgroups in DMIST. *Radiology*, 246(2):376–383. PMID: 18227537.

[Russakovsky et al., 2014] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2014). ImageNet Large Scale Visual Recognition Challenge. *CoRR*, abs/1409.0575. Available online on `arxiv.org/abs/1409.0575`.

[Skaane et al., 2007] Skaane, P., Hofvind, S., and Skjennald, A. (2007). Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: Follow-up and final results of oslo II study. *Radiology*, 244(3):708–717. PMID: 17709826.

[Taigman et al., 2014] Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In *Computer Vision and Pattern Recognition (CVPR)*, Columbus, Ohio.