Convolutional Neural Networks for Static and Dynamic Breast Infrared Imaging Classification

Matheus F O Baffa
and Lucas G Lattari
Department of Computer Science
Federal Institute of Southeast of Minas Gerais
Rio Pomba, MG – Brazil
Email: mfreitas826@gmail.com,
lucas.lattari@ifsudestemg.edu.br

***Abstract*—Breast cancer is the most frequent type of cancer among women. Since early diagnosis provides a better prognosis, different techniques have been developed by researchers all over the world. Several studies proved the efficiency of infrared image as a breast cancer screening technique. This paper proposes a methodology for analyzing infrared thermography of breast, considering distinct protocols, in order to classify patients images as healthy or non-healthy due to anomalies such as cancer. The major contribution of this work is to provide accurate classification using Convolutional Neural Networks, which were not exploited in previous works. Many methods relies on hand-crafted features and traditional classificators, such as Support Vector Machines. We obtained competitive results compared to other works and we design an appropriate modelling which takes advantage of this type of deep learning architecture. Our proposal obtained 98% of accuracy for static protocol and 95% for dynamic protocol.**

I. Introduction

Cancer is a disease which can spread quickly and impact the major organs and functions of the body. According to the World Health Organization – WHO [1], cancer is an uncontrollable growth or reproduction of abnormal cells, beyond their usual boundaries. Metastasis is one of the hallmarks of cancer, which the tumor cells spread to other parts of the body damaging organs and tissues, leading patients to death.

Several glands and tissues compose the breasts. Lobules are responsible to produce milk which is transported to the nipples by ducts, a breast structure similar to a pipe. Breasts are also composed by fat, lymph nodes and blood vessels in a region called stroma.

Breast cancer occurs when the tumor is located in one of the breast structure, mentioned before. According to The VisualMD Website [2], about 80-85% of breast cancer cells occurs in the ducts, 10-15% of the cases appears in lobules cells, and the other 5-10% occurs in the stroma region.

Breast cancer is the second most common type of cancer in the world [1]. In Brazil, the National Institute of Cancer [3] estimated that in 2018, 28% of new registered cases of the disease, are breast cancer. It is also responsible for the death of an average of 14,000 people.

Despite its high occurrence, when diagnosed early, the patient may obtain 95% chance of cure [4]. This is the major reason for many researchers to keep developing new methodologies to diagnose breast cancer earlier. Also, these technologies assist doctors to make a more accurate diagnose and also, let them able to indicate preventive treatments for patients at risk to develop the disease.

There are several ways to diagnose breast cancer. The breast self-exam is the easiest option since the patients can do it by themselves. More detailed exams rely on images to visualize and detect any abnormal pattern inside the breast that indicates anomalies.

Mammography is the most used screening technique in order to identify patients at initial stages of breast diseases. When this exam presents any abnormal results, doctors may conduct their patients to a more detailed screening exam, considering the BI–RADS score. However, if breasts have high density or protesis, it requires a different screening exam, such as Ultrasound or Magnetic Ressonance Imaging – MRI, which is more expensive than other methods. The combination of different techniques is important, since most of them are complementary [5].

Considering the most popular breast diagnosis exams, infrared imaging provide important information about the presence of abnormalities on the breast. Based on the asymmetry of the temperature distribution along the sagital axis, it is possible to investigate the presence of abnormalities on the breast.

Also, infrared imaging introduces several advantages. They do not use ionizing radiation, venous access or other invasive procedures. It is painless and has no contact with patients skin and it has low cost when compared with traditional exams, such as mammography, ultrasound and others. Section II presents more details about infrared imaging in the context of breast cancer detection.

Computer aided detection or diagnosis (CAD) systems may assist medical doctors on cancer diagnostics. This task requires from the doctor a specific study and experience in order to provide an accurate diagnosis based on image. Also, due to some limitation each screening technique has, a computer-aided detection system may work as a peer based diagnosis method.

These systems consist of several stages. Initially, acquisition and storage of infrared images are necessary. Several proposals

of acquisition protocols of breast thermal images includes static and dynamic approaches. The static protocol does not take into account data related to time. Normally it consists of a single image with thermal maps. The dynamic protocol, in contrast, encompasses several images captured during an interval of time.

After the definition of the protocol, it is necessary to use image processing techniques for segmentation of the region of interest (ROI), in order to extract the relevant area processed for later stages. The extraction of features in ROI is necessary for further use in a machine learning classifier, which detect patterns with anomalies (positive samples) and without anomalies (negative samples).

Considering the increasing demand for automatic methods for breast cancer detection in infrared imaging, we analyze different methodologies to automatically classify infrared images into distinct subsets: healthy or abnormal. The contributions of this paper are stated:

- It was developed a supervised breast cancer detection methodology based on Convolutional Neural Networks (CNNs), which was never used for classifying infrared breast images.
- We propose 4 strategies to determine how the dynamic protocol fits better in a CNN model. Also, we describe an appropriate CNN architecture design for the described problem.
- We evaluate the performance of two distinct protocols (static and dynamic) of breast images acquisition, considering cross-validation. That avoid the effects of overfitting.
- We discuss the results quantitatively for seven accuracy measures and time (in seconds) of network training, considering CPU and GPU architectures.

This paper is organized as follows. The Section II describes how infrared imaging works and how they are used to classify breast thermal images as healthy or non-healthy. Also, it analyzes how accurate infrared image methods detect the disease. The section III summarizes the state-of-the-art of breast infrared imaging classification and compares our approach to them. Finally, in the Section IV we describe the proposed methodology considering new approaches for introducing the breast thermal dataset with distinct protocols in a deep learning framework. Finally, Section V summarizes our experiments and quantitatively results. The conclusion and future works are depicted in Section VI.

II. The Role of Infrared Image Detecting Breast Cancer

Infrared imaging used on clinical applications is a procedure used to detect, record and produce an image representation of the infrared radiation, invisible for humans eyes [6]. The infrared radiation is part of the electromagnetic spectrum which is not visible but can be represented as a temperature.

According to Gore and Xu [7], it is possible to observe, for a long period of time, a certain symmetry in the thermal temperature distribution of the body. The presence of abnormal cells, such as a tumor, may affect this temperature distribution.

Three main reasons for this thermal asymmetry can be mentioned, such as the angiogenesis process, the high metabolic activity of the cancer cells and the vasodilatation caused by nitrix oxide release [6], [8].

It is possible to capture the breast temperature using simple and non-invasive methods, such as thermal cameras. This type of camera is sensible to infrared radiation, acquiring the information and converting into one of the visible bands of the electromagnetic spectrum.

The Database for Mastology Research [9] has two different protocols of acquiring images, the static and the dynamic protocol. The static records only one image per patient, in different angles, while the dynamic records a set of twenty images per patient, over a given period, after passing through a cooling step.

III. Related Papers

Classification of medical images is a complex task that requires some expertise to analyze and detect abnormalities. Computers can master this task with a good dataset and the right machine learning algorithm. Is this section, we are going to review some of the methods available in the literature and compare them to the proposed method of this paper.

In the work of Lessa and Marengoni [10], it was developed a method based on Artificial Neural Networks (ANN) to classify breast thermal images into abnormal and normal samples. Similar to our approach, Lessa and Marengoni used infrared images from DMR [9] using the static aquisition protocol. In their experiments, it was selected 47 patients and 94 breast images, in which 48 came from healthy patients and 46 had some abnormality. They reported 87% in sensitivity, 83% in specificity and 85% in accuracy.

Their method consists of 4 steps. First, the data is preprocessed, such that the image is converted to grayscale and the background is removed; Then, they segment the breast thermograms using the Canny edge detector with a threshold for the lower part of each breast; After that, they compute the histogram of each thermogram, in order to understand the relationship between the temperature of healthy and sick samples using statistical analysis; and finally, they train ANNs to realize the classification based on this feature.

Differently from Lessa and Marengoni, Silva et al. [11] worked only with the dynamic aquisition protocol. He developed a hybrid approach to classify samples into normal or "with abnormalties", using time series. His approach achieved 100% in accuracy.

Their work extracted the temperature matrix from the patients exam. Next, it is developed a step of image processing, which the Region of Interest (ROI) is segmented and the image is registered. Then, the ROI is divided into small squares which the maximum temperature is computed over the patients thermograms, generating the time series. After that, K-Means algorithm is applied in order to construct k groups. Clustering validation index is used to evaluate the results obtained by

K-Means. Finally, the values obtained in the previous step is used as features and submitted to classification. Over 30 classification algorithms were used in his experiment, reaching 100% of accuracy with K-Star and Bayes Net.

Gaber et al. [12] developed a CAD system classifier. His approach has two stages, the automatic segmentation and the classification. First, the segmentation has three basic steps, the pre-, segmentation and post-segmentation. Basically these stages prepare the data, segment and enhance the result of the segmentation. After that, 30 features based on gabor coeficients, statistics and GLCM, were extracted from the dataset and transferred to the classifier. At the end, the Support Vector Machine (SVM) has been used to classify the samples. The authors reported a total of 92.06% of accuracy and 87.50% of precision.

Sathish et al. [13] also developed a breast thermogram image classifier. In their approach, a automatic segmentation of the ROI is developed and used to extract two features, the histogram and the grayscale co-occurrence matrix-based texture. According to the authors, using statistical tests, these two features are more significant in order to detect breast cancer. Their approach uses SVM RBF to classify the samples into sick and healthy patients. They reported an accuracy of 90% and 87.5% of sensitivity.

A similar work was developed by Borchartt et al. [14]. Their work uses the DMR static images for its experiments, with a sample of 51 images, which 37 came from healthy patient and 14 have some abnormality. The methodology proposed by the authors include a segmentation step which can be a manual or an automatic algorithm to segment the ROI. After that, the authors used statistical measures, histogram, Higuchis fractal dimension and three geostatistics methods as features to train the SVM algorithm. And finally, the Genetic Algorithm (GA) is used in order to optimize the results of the classification. The authors reported a 88% of accuracy and 79% of specificity.

Another type of work have been conducted by Gogoi et al. [15] in which it was proposed a investigation of the importance of selecting the right features to improve the classification accuracy. According to the authors, a set of 24 features have been used in order to select the best discriminative features for breast cancer detection. Using a random selected images from different databases (DMR and DBT-TU-JU), the authors used statistical and texture features with several classifiers in order to find the algorithm with the best accuracy. With ANN and Linear SVM, the authors reached 87.5% of accuracy and 80% of sensitivity.

IV. Materials and Methods

A. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a specialized kind of neural network for processing data that has a known, grid-like topology [16]. A typical layer of a CNN consists of three stages. In the first stage, several convolution operations are performed in parallel to produce a set of linear activations. These convolutions compute feature maps of the previous layer with a set of filters.

Later, each linear activation is computed through a nonlinear activation function, such as reLU (rectified linear activation function). Normally reLU is used because it leads to a faster training stage in deep nets. This stage is called the detector stage.

Finally, it uses a pooling function to modify the output of the further layer, reducing the size of feature maps. A pooling layer with pooling size of 2 x 2 reduces a patch with 2 x 2 pixels to one pixel. That results in faster convergence and better generalization.

These stages are depicted in Figure 1. In this context, the convolutional net is viewed as a small number of complex layers, with each layer having many stages.

The training of a CNN ressembles a feedforward neural network. Each sample (image) is forwarded through the layers until a loss function is computed. Then, the loss is back-propagated into the net, changing the weights in accordance to gradient descent methods [18]. This procedure is looped until a pre-defined number of training iterations (epochs) is reached.

B. Training a CNN for Breast Thermal Dataset using Static and Dynamic Protocols

The infrared images used in this work belongs to the *Database for Mastology Research with Infrared Image* - DMR [9]. This dataset comprises static and dynamic protocols. Protocols for thermal imaging can be classified based on the behavior of the body, related to heat transfer.

In static acquisitions, the body of the patient must achieve thermal balance in a controlled environment; dynamic protocols are used to inspect the skin temperature recovery caused by thermal stress after cooling the patient by eletric fan.

The static dataset of DMR is composed of 177 images of healthy patients and 42 pictures of patients with cancer. Since balance of image classes is necessary to use CNN architecture, images of cancerous set is augmented until both classes have the same amount of images. These images are chosen randomly among the abnormal set.

For the dynamic protocol, 95 healthy patients has 20 images each, totalizing 1900 figures. On the non-healthy cases, 42 patients compose 20 images (840 pictures in total). As in static approach, abnormal set of images are augmented randomly, until the healthy and non-healthy sets has the same number of samples (images).

For the static approach, CNN classification is performed using each image of each patient, considering cross-validation. However, since we use CNN, it is not necessary to define features to be selected into the breast thermal dataset. The CNN is responsible to determine which patterns and features separates among the healthy and non-healthy sets. That is a fundamental difference between our paper and traditional works such as [10], [11], [19].

For the dynamic protocol, several strategies were proposed in this work to verify which approach works better for CNN classification of breast thermal images. The first strategy

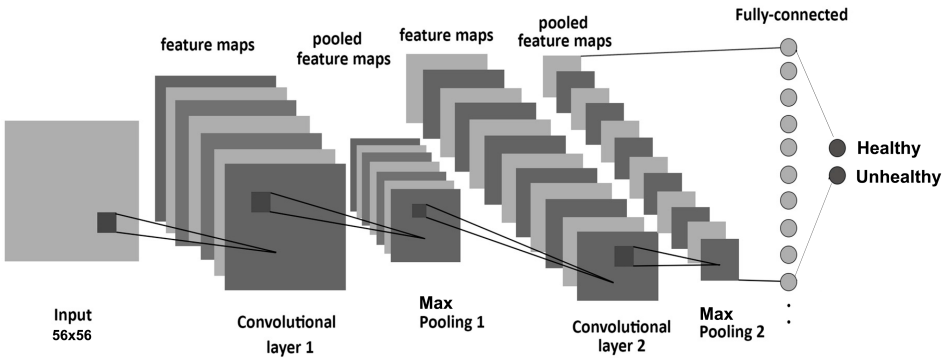

Fig. 1. The structure of a typical Convolutional Neural Network adapted to our proposed architecture. Source: [17], adapted.

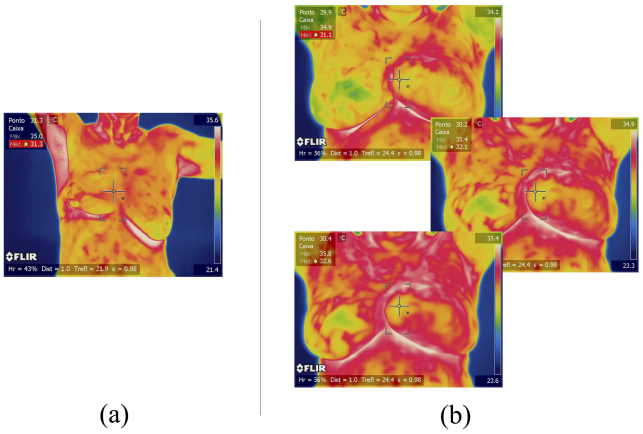

Fig. 2. Example of the breast thermal database. In (a) is represented the static image acquisition protocol, which each patient has only one image and in (b), it is presented the dynamic protocol, which each patient has a set of 20 images obtained during a certain interval of time.

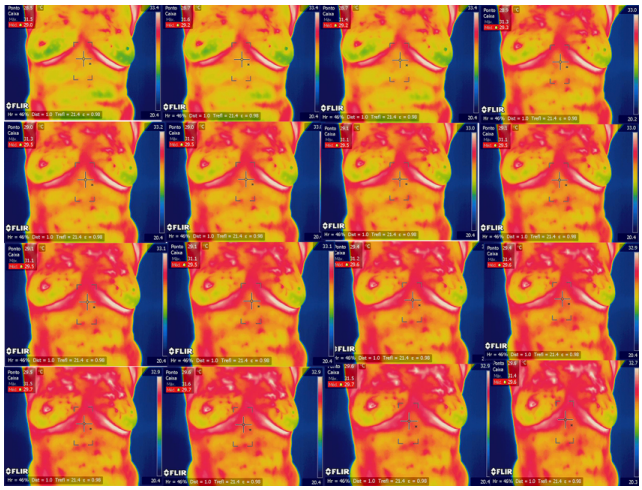

Fig. 3. Illustration of several images that composes a single patient thermal images in dynamic protocol of DMR database. Each set of images composes a dataset for CNN classification.

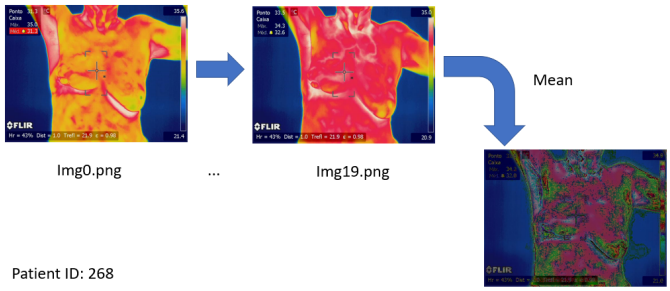

Fig. 4. Example of the second strategy, where all image colors are summed and their mean is computed. The final generated image is used in the network training.

assigned all data (20 images) in a single array, as if they were part of a single image. This strategy is depicted in Figure 3.

The second strategy takes all 20 images of a patient set and compute their mean. In this case, each patient generates a single image which is used to train the network model. This approach is stated in Figure 4.

The third strategy encompasses the first and the last image of the set and computes their mean. Consequently, only two images with a significant difference in their temperature are considered. The average image is used to train the network. This approach is depicted in Figure 5.

Finally, the last strategy uses image subtraction to produce the picture used in the neural network. The last image of the set is subtracted by the first one. The result is a transformed image similar to the presented in Figure 6.

Independent of the strategy adopted for static or dynamic protocols, both sets are divided into train and test classes. The first one is used by the CNN to fit the model. The second one is used to provide an unbiased evaluation of a final model fit on the training dataset.

V. Experiments and Results

A. Database Methodology

The DMR database is a public breast thermogram database [9]. It contains the breast infrared imaging of 287 patients which has healthy and sick labels. For acquisition of thermograms, they used a FLIR SC-620 Thermal Camera. Each patient image has spatial resolution of 640x480 pixels, and

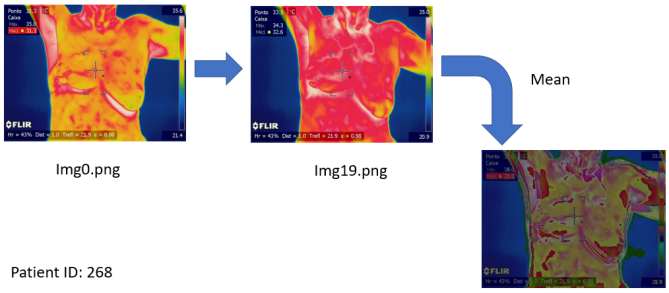

Fig. 5. Example of the third strategy, where only the first and the last image of the patient set is summed and their mean is computed. The final average image is used in the network training.

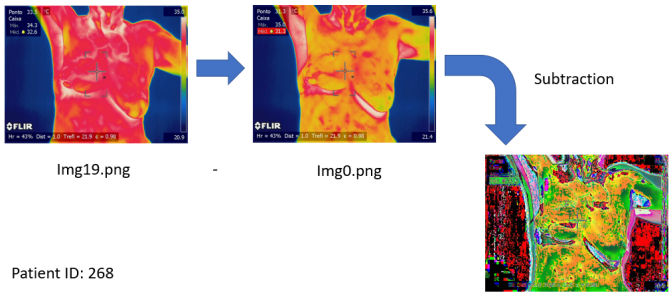

Fig. 6. Example of the fourth strategy, where the last image is subtracted by the first one. The generated image is used in the network training.

grayscale or colored images that represents their heat temperature.

For the static approach, we partitioned the original dataset, composed of 300 images into two sets, in a proportion of 10% of test images and 90% of training images. These values were obtained empirically. Consequently, our methodology is composed of 10 experiments, such that 15 images belong to the training set and 15 figures are assigned to the test set.

Considering the dynamic methodology, we used data of 137 patients, which 95 patients were considered healthy and 42 were assigned to the non-healthy set. The proportion of train and test sets is 88%-12%, respectively. Analogously to the static approach, these values were chosen in an empirical fashion. Once exist a disproportion between the healthy and non-healthy sets, we got a higher degree of positive set during the cross-validation stage. Consequently, in this case a total of 6 experiments were evaluated for each strategy.

The dynamic protocol required these strategies to represent each patient, with 20 imagens, in only one picture, so the proposed architecture for CNN could fit for both protocols. In total, we used 190 images in which 95 was normal, or healthy, samples and 42 abnormal, or unhealthy, samples. Since the data is unbalanced, we used data augmentation to balance and increase the number of unhealthy samples to 95.

B. Network Design and Computer Architecture

In order to classify breast images as healthy or malignant, we designed a Convolutional Neural Network for static and dynamic protocols. This type of network become state-of-the-art on image classification problems [20].

We design our CNN architecture with learning rate of 0.001 and dropout rate of 0.75. We also used the Adam Optimizer to optimize our deep learning model. Also, in the proposed architecture, it is used two textit Convolutional Layers with the size of 5x5 and 32 outputs, followed by two *Max Pooling Layers* with 5x5 in size and a threshold of 3. The output layer is a fully connected layer and classify the data into two classes, called here healthy and unhealthy.

For the construction of each test and train sets, we resize all images to 56 x 56 pixels, which is the resolution used tipically in CNN applications. We also use cross-validation to guarantee that every image is evaluated in test and train sets. Also, it minimizes generalization problems caused by overfitting.

To finish the training network, it is defined a halting criterion. It considers the error function of the method. First we evaluate the difference between the error function of the previous iterations minus the current error function is greater than 1000, or, $(prevloss - currloss) > 1000.0$. In that case, if this condition is true for more than 10 iterations, then the training is stopped. After evaluation of several experiments, we consider that a good halting criterion.

Since we are working with an unbalanced database, which has more healthy samples than unhealthy ones, we used data augmentation on botch dynamic and static approach. We used operations, such as crop and duplicate to set the unhealthy samples with the same amount as the other class.

About the computer architecture, our experiments were evaluated in a notebook with Processor AMD Ryzen 7 1700 3.2Ghz with 8GM of RAM DDR4. We also tested part of our results in a GPU nVidia GeForce GTX 1060 6GB. Our software consists of Cuda Toolkit 8.0, cuDNN 5.1, Python 3.6.3 and Tensorflow 1.4.0.

C. Comparison with Other Works

Conceive a breast abnormality classification system is difficult, since several feature sets and classifier models are available. Choose the most efficient to deal with breast thermograms is a challenging task. Based on that, our choice of CNN deal with the feature selection problem and the classification as well.

Most works relies on combination of statistical and texture features with a SVM (Support Vector Machine) classifier. This approach requires hand-crafted features extraction, which involves a hard work by the system developer. On the contrary, CNNs deal with automatic feature selection in a robustness fashion.

Since our work it the first to deal with CNN for automatic classification of breast thermograms as healthy or malignal, we compared our results with articles that used the same breast database as ours ([10]–[15]). We summarize their and our results in Table I.

Based on Table I, our work performed better on static protocol with more images than the compared works. That is

expected, since CNN outperforms traditional image classifiers [20].

For the dynamic protocol, Silva et al [11] achieves a perfect score. However, this accuracy suggests that their work has overfitting. Differently, our work had experiments with mean accuracy of 95%, for colored images even considering cross-validation of the entire image set.

D. Detailed Results

All results of our proposal are fulfilled in Tables II, III, IV, V, VI. Mean values for static and dynamic protocols are highlighted.

Evaluation results show that our deep learning approach using the colored image dataset presented good performance in comparison with grayscale dataset for static protocol (Table II). Furthermore, our results outperformed other works that deals with the same dataset in every accuracy measure. That is expected, since much more information about temperature is stored in the color set, and CNN features capture patterns in a more efficient manner than handcrafted feature selection.

Considering the dynamic protocol, the first and the third strategies performed better when considering the accuracy (Tables III and V). The first strategy comprises all data of each patient, giving much more information to the deep network. Consequently, more redundant data is used to identify different patterns of cancerous cells.

The third strategy performs very well. Only the mean between the first and the last image are considered. Based on our results, it is possible to reduce the amount of data necessary to produce an effective model to classify anomalies on breast thermal images. Also, all strategies generalizes well, avoiding overfitting.

VI. Conclusion

One of the most common malignancies among women is breast cancer. Further exploration of techniques for early detection of cancer is necessary. Based on that, we developed a Convolutional Neural Network optimized for breast infrared imaging. The results of this study indicated that CNNs obtained competitive results for both protocols: static and dynamic. Without overfitting, it is possible to assume that our proposal can be easily generalized for much more data, obtaining similar or even more accurate results.

In future works we expect to construct a complete Computed-Aided Diagnosis system, realizing segmentation and classification in the same methodology. We also expect that our previous segmentation works in consonance with CNN-classificator can yield good results.

References

[1] W. H. Organization. (2017) Cancer. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs297/en/

[2] TheVisualMD. (2017) When things go wrong. [Online]. Available: https://www.thevisualmd.com/health_centers/cancer/breast_cancer/when_things_go_wrong

[3] I. N. do Cancer. (2017) Mama. [Online]. Available: http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home+/mama/cancer_mama

[4] H. do Cancer de Barretos. (2015) Informao: Saiba quais so os tipos de cncer mais comuns no brasil. [Online]. Available: https://www.hcancerbarretos.com.br/

[5] T. B. Borchartt, A. Conci, R. C. Lima, R. Resmini, and A. Sanchez, "Breast thermography from an image processing viewpoint: A survey," *Signal Processing*, vol. 93, no. 10, pp. 2785 – 2803, 2013, signal and Image Processing Techniques for Detection of Breast Diseases. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168412002794

[6] W. C. e. a. Amalu, "Infrared imaging of the breast - an overview," *Medical Devices and Systems: The Biomedical Engineering Handbook*, pp. 25–1, 2006.

[7] J. Gore and L. Xu, "Thermal imaging for biological and medical diagnostics," *Biomedical Photonics Handbook*, pp. 446–457, 2003.

[8] J. Scicinski, B. Oronsky, S. Ning, S. Knox, D. Peehl, M. M. Kim, P. Langecker, and G. Fanger, "No to cancer: The complex and multifaceted role of nitric oxide and the epigenetic nitric oxide donor, rrx-001," *Redox biology*, vol. 6, pp. 1–8, 2015.

[9] L. Silva, D. C. M. Saade, G. Sequeiros Olivera, A. Silva, A. Paiva, R. Bravo, and A. Conci, "A new database for breast research with infrared image," *Journal of Medical Imaging and Health Informatics*, vol. 4, pp. 92–100, 03 2014.

[10] V. Lessa and M. Marengoni, "Applying artificial neural network for the classification of breast cancer using infrared thermographic images," *Computer Vision and Graphics*, vol. 9972, pp. 429–438, 09 2016.

[11] L. F. Silva, A. A. S. Santos, R. S. Bravo, A. C. Silva, D. C. Muchaluat-Saade, and A. Conci, "Hybrid analysis for indicating patients with breast cancer using temperature time series," *Computer Methods and Programs in Biomedicine*, vol. 130, pp. 142 – 153, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260715300249

[12] T. Gaber, G. Ismail, A. Anter, M. Soliman, M. Ali, N. Semary, A. E. Hassanien, and V. Snasel, "Thermogram Breast Cancer Prediction Approach based on Neutrosophic Sets and Fuzzy C-Means Algorithm," Dec. 2013. [Online]. Available: https://doi.org/10.5281/zenodo.34913

[13] D. Sathish, S. Kamath, K. Prasad, R. Kadavigere, and R. Martis, "Asymmetry analysis of breast thermograms using automated segmentation and texture features," *Signal, Image and Video Processing*, vol. 11, no. 4, pp. 745–752, 5 2017.

[14] T. B. Borchartt, R. Resmini, L. S. Motta, E. W. G. Clua, A. Conci, M. J. A. Viana, L. C. Santos, R. C. F. Lima, and A. Sanchez, "Combining approaches for early diagnosis of breast diseases using thermal imaging," *Int. J. Innov. Comput. Appl.*, vol. 4, no. 3/4, pp. 163–183, Oct. 2012. [Online]. Available: http://dx.doi.org/10.1504/IJICA.2012.050054

[15] U. R. Gogoi, M. K. Bhowmik, A. K. Ghosh, D. Bhattacharjee, and G. Majumdar, "Discriminative feature selection for breast abnormality detection and accurate classification of thermograms," April 2017, pp. 39–44.

[16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[17] S. Albelwi and A. Mahmood, "A framework for designing the architecture of deep convolutional neural networks," *Entropy*, vol. 19, no. 6, 2017.

[18] L. F. Rodrigues, M. C. Naldi, and J. F. Mari, "Exploiting convolutional neural networks and preprocessing techniques for hep-2 cell classification in immunofluorescence images." IEEE Computer Society, 2017, pp. 170–177.

[19] R. Rastghalam and H. Pourghassem, "Breast cancer detection using mrf-based probable texture feature and decision-level fusion-based classification using hmm on thermography images," *Pattern Recogn.*, vol. 51, no. C, pp. 176–186, Mar. 2016. [Online]. Available: https://doi.org/10.1016/j.patcog.2015.09.009

[20] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, "Recent advances in convolutional neural networks," *Pattern Recogn.*, vol. 77, no. C, pp. 354–377, May 2018. [Online]. Available: https://doi.org/10.1016/j.patcog.2017.10.013

TABLE I
SUMMARY OF SIMILAR WORKS ON BREAST THERMOGRAM CLASSIFICATION USING DMR DATABASE.

Authors	Features / Classifiers	Number of Images	Acquisition Protocol	Accuracy
Gaber et al [12]	Gabor Coefficients / SVM RBF	63 (29 healthy and 34 malignant)	Static	88.41%
Sathish et al [13]	Statistical Texture Features / SVM RBF	80 (40 normal and 40 abnormal)	Static	90%
Lessa and Marengoni [10]	Statistical Features / Artificial Neural Networks	94 (48 normal and 46 abnormal)	Static	85%
Borchartt et al [14]	Statistical Features / Genetic Algorithm	51 (14 abnormal and 37 normal)	Static	88.2%
Gogoi et al [15]	Statistical and Texture Features / SVM Linear and Artificial Neural Networks	80 (45 normal and 35 abnormal)	Static	87.50%
Silva et al [11]	Several using Feature Selection / Bayes Nets	80 patients with 20 images each (40 healthy and 40 non-healthy)	Dynamic	100%
Our Work	CNN Features / Convolutional Neural Networks	300 images (126 abnormal and 174 normal)	Static	98% (color) and 95% (grayscale)
Our Work	CNN Features / Convolutional Neural Networks	137 patients with 20 images each (95 healthy and 95 non-healthy)	Dynamic	95% (color) and 92% (grayscale)

TABLE II
OUR RESULTS USING COLORED IMAGES AND GRAYSCALE IMAGES ON THE STATIC PROTOCOL.

Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time (GPU)	Time (GPU)
0	1	1	1	1	1	1	1	28.67s	28.67s
1	0.96	0.93	1	1	0.93	1	0.96	53.35s	24.68s
2	0.93	0.87	1	1	0.86	1	0.93	88.04s	63.36s
3	0.96	0.93	1	1	0.93	1	0.96	121.08s	57.72s
4	1	1	1	1	1	1	1	136.71s	78.99s
5	1	1	1	1	1	1	1	152.13s	73.14s
6	1	1	1	1	1	1	1	184.69s	111.55s
7	1	1	1	1	1	1	1	208.72s	97.17s
8	0.96	0.93	1	1	0.93	1	0.96	219.96s	122.79s
9	1	1	1	1	1	1	1	239.04s	116.25s
Mean	0.98	0.97	1	1	0.97	1	0.98	-	77.43s
Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time (CPU)	Time (CPU)
0	0.98	1	0.94	0.93	1	1	0.93	618s	618s
1	0.92	0.92	0.82	0.8	0.93	0.8	0.86	1614s	996s
2	0.94	1	1	1	1	1	1	2329s	715s
3	0.92	0.87	0.87	0.87	0.87	0.87	0.87	3260s	931s
4	0.96	0.93	0.94	0.95	0.96	0.97	0.98	5476s	561s
5	1	1	1	1	1	1	1	4915s	996s
6	0.96	0.93	0.94	0.95	0.96	0.97	0.98	5476s	561s
7	0.94	0.87	0.93	0.93	0.86	0.93	0.9	6216s	740s
8	0.94	0.93	0.94	0.95	0.96	0.97	0.98	6818s	602s
9	0.92	0.93	0.87	0.87	0.93	0.87	0.9	7616s	799s
Mean	0.95	0.94	0.92	0.92	0.95	0.94	0.94	-	761.7s

TABLE III
OUR RESULTS USING COLORED IMAGES ON THE FIRST STRATEGY OF DYNAMIC PROTOCOL.

Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	0.96	1	0.93	0.92	1	0.92	0.96	3223.80s	3223.80s
1	0.92	0.87	1	1	0.85	1	0.93	10255.68s	7031.88s
2	1	1	1	1	1	1	1	13173.04s	2917.26s
3	0.96	1	0.93	0.92	1	0.92	0.96	16543.41s	3370.37s
4	0.86	1	0.77	0.71	1	0.71	0.83	21112.49s	4569.08s
5	1	1	1	1	1	1	1	24007.71s	2895.22s
Mean	0.95	0.97	0.93	0.92	0.97	0.97	0.94	-	3515.05s

TABLE IV

OUR RESULTS USING COLORED IMAGES AND GRAYSCALE IMAGES ON THE SECOND STRATEGY OF DYNAMIC PROTOCOL.

Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	1	1	1	1	1	1	1	252.32s	252.32s
1	1	1	1	1	1	1	1	478.18s	225.86s
2	0.8	0.84	0.76	0.73	0.86	0.73	0.78	762.75s	536.89s
3	0.93	1	0.88	0.86	1	0.86	0.92	928.09s	391.2s
4	0.9	0.92	0.87	0.86	0.93	0.86	0.89	1180.92s	789.72s
5	0.83	0.81	0.85	0.86	0.8	0.86	0.83	1342.36s	552.64s
Mean	0.91	0.93	0.89	0.89	0.93	0.89	0.90	-	458.11s
Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	1	1	1	1	1	1	1	155.33s	155.33s
1	0.83	0.81	0.85	0.86	0.8	0.86	0.83	282.7s	127.37s
2	0.96	0.93	1	1	0.93	1	0.96	479.68s	353.31s
3	0.83	0.81	0.85	0.86	0.8	0.86	0.83	605.86s	253.55s
4	0.90	0.92	0.87	0.86	0.83	0.86	0.89	753s	499.45s
5	0.90	0.83	1	1	1	1	0.9	870s	370.55s
Mean	0.90	0.88	0.93	0.93	0.89	0.93	0.90	-	293.09s

TABLE V

OUR RESULTS USING COLORED IMAGES AND GRAYSCALE IMAGES ON THE THIRD STRATEGY OF DYNAMIC PROTOCOL.

Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	1	1	1	1	1	1	1	190.39s	190.39s
1	1	1	1	1	1	1	1	318.23s	127.84s
2	0.83	0.81	0.85	0.86	0.8	0.86	0.83	523.24s	395.4s
3	0.93	0.88	1	1	0.86	1	0.93	630.12s	234.72s
4	0.96	0.93	1	1	0.93	1	0.96	789.04s	554.32s
5	0.96	0.93	1	1	0.93	1	0.96	919.68s	365.36s
Mean	0.95	0.93	0.98	0.98	0.92	0.98	0.95	-	311.34s
Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	0.93	0.88	1	1	0.86	1	0.93	91.13s	91.13s
1	0.93	0.88	1	1	0.86	1	0.93	195.19s	104.06s
2	0.96	0.93	1	1	0.93	1	0.96	324.69s	220.63s
3	0.90	0.92	0.87	0.86	0.93	0.86	0.89	437.63s	217s
4	0.96	0.93	1	1	0.93	1	0.96	527.87s	310.87s
5	0.86	0.78	1	1	0.73	1	0.88	605.53s	294.66s
Mean	0.92	0.89	0.98	0.98	0.87	0.98	0.93	-	206.39s

TABLE VI

OUR RESULTS USING COLORED IMAGES AND GRAYSCALE IMAGES ON THE FOURTH STRATEGY OF DYNAMIC PROTOCOL.

Experiment	Accuracy	Sensibility	Specificity	Preditivity Pos	Preditivity Neg	Precision	F1	Accumuled Time	Time
0	1	1	1	1	1	1	1	155.94s	155.94s
1	0.83	0.81	0.85	0.86	0.8	0.86	0.83	246.41s	90.47s
2	0.93	1	0.88	0.86	1	0.86	0.92	410.67s	320.2s
3	0.96	0.93	1	1	0.93	1	0.96	549.62s	229.42s
4	0.93	0.88	1	1	0.86	1	0.93	711.74s	482.32s
5	0.90	0.83	1	1	0.8	1	0.9	807s	324.68s
Mean	0.93	0.91	0.96	0.95	0.90	0.95	0.92	-	267.17s
0	0.96	0.93	1	1	0.93	1	0.96	214.39s	214.39s
1	0.83	0.81	0.85	0.86	0.8	0.86	0.83	332.46s	118.07s
2	0.83	0.75	1	1	0.66	1	0.85	423.23s	305.16s
3	0.90	0.92	0.87	0.86	0.93	0.86	0.89	561.53s	226.37s
4	0.83	0.81	0.85	0.86	0.80	0.86	0.83	652.43s	396.06s
5	0.93	0.88	1	1	0.86	1	0.93	756.07s	360.01s
Mean	0.88	0.85	0.93	0.93	0.83	0.93	0.88	-	275.01s