A Deep Image Segmentation Approach to Breast Keypoint Detection

Tiago Gonçalves [1,2]
tiago.f.goncalves@inesctec.pt
Wilson Silva [1,2]
wilson.j.silva@inesctec.pt
Jaime S. Cardoso [1,2]
jaime.cardoso@inesctec.pt

[1] Faculdade de Engenharia
Universidade do Porto
Porto, Portugal
[2] INESC TEC
Porto, Portugal

Abstract

The main aim of breast cancer conservative treatment is the optimisation of the aesthetic outcome and women's quality of life, without jeopardising local cancer control and overall survival. Recently, a deep learning algorithm, used in conjunction with a shortest-path algorithm that models images as graphs, has been proposed and achieved state-of-the-art results. However, it is both time-consuming and computationally complex. In this work, we studied a novel algorithm based on the interaction of deep image segmentation and deep keypoint detection models capable of improving both performance and execution-time on the breast keypoint detection task.

1 Introduction

Breast cancer ranks as the most frequent cancer among women [1, 5]. Despite being a highly mutable and rapidly evolving disease, it is estimated that most breast cancers are curable if properly detected and treated [8]. Under this paradigm, it is possible to surgically treat most cancers with a breast cancer conservative treatment (BCCT), which does not require the removal of the entire breast, as in mastectomy [12]. Currently, to perform the aesthetic assessment of BCCT, the majority of the extracted features are related to asymmetry measurements [4]. To facilitate the extraction of such features, it is fundamental to mark breast keypoints. The advent of machine learning and deep learning opened the possibility to design novel algorithms based on deep neural networks (DNN) which may be fully end-to-end (*i.e.*, receive an image and output the aesthetic assessment score). Until the publication of this work [6], the state-of-the-art algorithm for keypoint detection was a hybrid method based on a DNN and on traditional computer vision methods, which made it computationally heavy and slow. This work presents the development of a novel breast keypoint detection algorithm that addresses the efficiency problem while maintaining or improving the accuracy. A study of algorithm performance based on execution time has also been conducted, since, at the long term, the intention is to deploy such algorithms into a web-based application that could be accessed by a diversity of devices and operative systems.

2 Deep Keypoint Detection Algorithm

Silva *et al.* proposed a method [10] that uses a deep neural network (DNN) for the keypoint detection task, opening the possibility to follow an integrated learning approach. Following the ideas of Cao *et al.* [3] and Belagiannis *et al.* [2], Silva *et al.* proposed an architecture that first learns how to regress *heatmaps* (in which the ground-truth was obtained by applying a Gaussian kernel to the keypoints) and, after iterative refinement of this heatmap regression, it can predict keypoint localisation. The heatmap regression is performed with the U-Net model [9]. Then, to do the keypoint regression, the original images are multiplied by the refined heatmaps (to improve the initial fuzzy localization of keypoints) and are fed to a keypoint regression module composed of three blocks: top layers of the VGG16 [11] (*i.e.*, only the convolutional layers), four additional convolutional layers and three dense layers. The entire model is trained, using the iterative refinement of the regression of heatmaps as a regularisation term of the loss function. The endpoints and the nipples are obtained with this deep learning algorithm, whereas the contour is refined with the shortest-path algorithm, which models images as graphs, based on gradients (see Figure 1). However, this procedure is very time-consuming when compared with the inference process of a deep learning model only. Also, if one intends to integrate such algorithms into a web application, it is of utmost importance that performance measurements (*e.g.*, loading time, execution time) are taken into account when testing and designing novel methods.

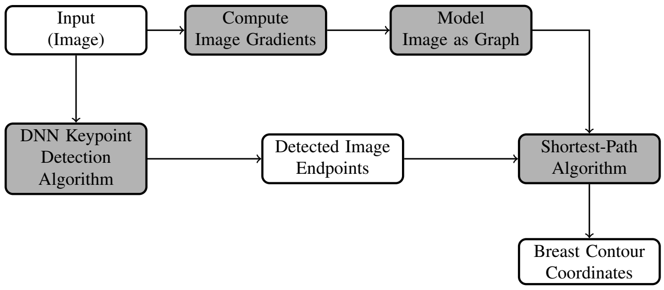

Figure 1: Hybrid keypoint detection algorithm, from [10].

3 Deep Image Segmentation for Breast Keypoint Detection

The intuition behind this approach is that it is easier to detect breast contours if one is capable to detect breasts first [7]. This can be seen as a problem of semantic segmentation, where both breasts are considered the foreground and the rest is considered the background of the image. The main hypothesis is that if it is possible to perform the segmentation of both breasts with high precision, one could proceed to an algorithm of contour detection and then accurately extract the keypoints related to the breast contours. With segmentation, the goal is to learn a single solution (*i.e.*, one image corresponds to one mask). This is important because, if the DNN is capable of predicting the correct mask, the set of points of the detected contour will contain a subset of points that belong to the real breast contour. On the other hand, with keypoint regression, there is a higher degree of variability, where the DNN can predict points that belong to the real breast contour and points that do not, negatively influencing the algorithm performance. Furthermore, when compared with the hybrid keypoint detection algorithm, one expects that this approach will bring improvements in terms of results and performance, *i.e.*, it would be faster than the hybrid keypoint detection algorithm.

4 Implementation and Results

The available dataset (221 images) has 37 ground-truth keypoints (4 endpoints, 30 points along the breast contours, 2 nipples and the sternal notch) resulting in a total of 74 coordinates per image. All experiences were done taking into account 5-fold cross-validation split into train and test sets. First, we trained a deep image segmentation model that could achieve good results in breast segmentation. To train this model, it was necessary to generate ground-truth breast masks, which were obtained with the support of the ground-truth keypoints and images.

Taking into account previous results with other models [7], for this experiment, we decided to use the U-Net++ [13] architecture as the deep segmentation model. The U-Net++ model was trained and used to generate segmentation masks. From these masks, contours were extracted. This first step outputs a variable number of contour keypoints, some of which are not desired, since they do not belong to what is considered the breast contour. As a post-processing step, we project the Silva *et al.* DNN predicted keypoints onto to the mask contours through the minimization of the Euclidean Distance between the mask contour keypoint and the predicted keypoint (see Figure 2). At the end of this processing step, the final set of breast keypoints is obtained (see Figure 3). We also studied the

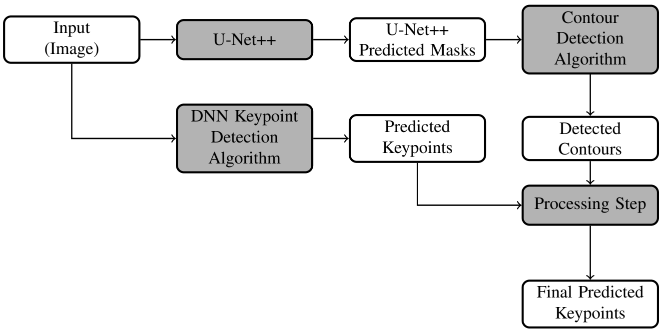

Figure 2: Scheme of the deep image segmentation algorithm for breast keypoint detection.

performance of both ours and Silva *et al.* algorithms to assess which one would fit better into a web-version of BCCT.core, capable of real-time interaction with deep learning models. To perform this study, the execution time of each algorithm on CPU (Intel® Core™ i7-2600 CPU @ 3.40GHz $\times$ 8) was measured on the test set of each cross-validation fold. Table 1 presents the average error distance (measured in pixels) and the average execution time (measured in seconds) of each model inference on the test set. It can be seen that our proposed method surpasses both the DNN and hybrid keypoint detection algorithms from Silva *et al.* in the endpoints and breast contours detection tasks, which were the state-of-the-art breast keypoint detection algorithms. Moreover, this novel algorithm achieves lower values of standard deviation and maximum error, which suggests it is even more robust when compared with the other two. Regarding the study of performance, it can be understood that the DNN keypoint detection algorithm achieves better execution time, however, it has the highest error for the breast contour. Our method presents the best balance between time-efficiency and accuracy, as it is the most accurate model, and has a time efficiency comparable to the most time-efficient method.

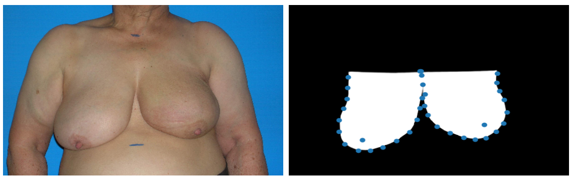

Figure 3: Example of the results obtained with the proposed deep image segmentation method. The first image is the input photograph and the second image is the U-Net++ predicted mask with the detected breast keypoints (after the processing step).

5 Conclusions and Future Work Recommendations

In this work we presented a novel algorithm based on the interaction of two deep learning models (a segmentation model, and a keypoint detection model) that can surpass the state-of-the-art algorithms present in the literature [6]. Furthermore, a comparative study regarding algorithms performance was done to assess which one would fit better a web-based application for the aesthetic assessment of BCCT. Our proposed model revealed itself as the best in terms of keypoint prediction, while being very competitive in terms of executing time. As future work, the next step will be to improve results on nipples detection task and to modify this novel segmentation-based keypoint detection algorithm by integrating all the tasks of its pipeline in a unique DNN with a combined loss function. The integration and full deployment of this algorithm in a web-application are also planned.

Table 1: Average error distance for endpoints, breast contours and nipples, measured in pixels and average execution time of the models' inferences. Best results are highlighted in bold. Note: STD stands for standard deviation and Max stands for maximum error.

Model	Endpoints			Breast Contours			Nipples			Execution Time (s)
	Mean	STD	Max	Mean	STD	Max	Mean	STD	Max	
DNN keypoint detection algorithm	40	**33**	**182**	21	8	72	**70**	**39**	**218**	**150**
Hybrid keypoint detection algorithm	40	**33**	**182**	13	14	104	**70**	**39**	**218**	1704
Our keypoint ketection algorithm	**38**	34	195	**11**	**5**	**34**	**70**	**39**	**218**	280

Acknowledgements

The project "TAMI - Transparent Artificial Medical Intelligence" (NORTE-01-0247-FEDER-045905) leading to this work is co-financed by ERDF - European Regional Fund through the Operational Program for Competitiveness and Internationalisation - COMPETE 2020, the North Portugal Regional Operational Program - NORTE 2020 and by the Portuguese Foundation for Science and Technology - FCT under the CMU - Portugal International Partnership and the PhD grants "SFRH/BD/139468/2018" and "SFRH/BD/06434/2020".

References

[1] *World cancer report 2014*. ISBN 978-92-832-0443-5.

[2] Vasileios Belagiannis and Andrew Zisserman. Recurrent Human Pose Estimation. *arXiv:1605.02914 [cs]*, May 2016. arXiv: 1605.02914.

[3] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime Multi-Person 2d Pose Estimation using Part Affinity Fields. *arXiv:1611.08050 [cs]*, November 2016. arXiv: 1611.08050.

[4] Jaime S. Cardoso and Maria J. Cardoso. Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. *Artificial Intelligence in Medicine*, 40(2):115–126, June 2007. ISSN 09333657. doi: 10.1016/j.artmed.2007.02.007.

[5] Sharon H Giordano, Aman U Buzdar, and Gabriel N Hortobagyi. Breast Cancer in Men. page 11.

[6] Tiago Gonçalves, Wilson Silva, Maria J Cardoso, and Jaime S Cardoso. A novel approach to keypoint detection for the aesthetic evaluation of breast cancer surgery outcomes. *Health and Technology*.

[7] Tiago Gonçalves, Wilson Silva, and Jaime Cardoso. Deep Aesthetic Assessment of Breast Cancer Surgery Outcomes. Springer International Publishing, Cham, 2020. doi: 10.1007/978-3-030-31635-8_236.

[8] Helder P. Oliveira, Jaime S. Cardoso, Andre Magalhaes, and Maria J. Cardoso. Methods for the Aesthetic Evaluation of Breast Cancer Conservation Treatment: A Technological Review. *Current Medical Imaging Reviews*, 9(1):32–46, April 2013. ISSN 15734056. doi: 10.2174/1573405611309010006.

[9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]*, May 2015. arXiv: 1505.04597.

[10] Wilson Silva, Eduardo Castro, Maria J. Cardoso, Florian Fitzal, and Jaime S. Cardoso. Deep Keypoint Detection for the Aesthetic Evaluation of Breast Cancer Surgery Outcomes. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI'19)*, 2019.

[11] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv:1409.1556 [cs]*, September 2014. URL `http://arxiv.org/abs/1409.1556`. arXiv: 1409.1556.

[12] Williams Street. Cancer Facts & Figures 2018. page 76, 2018.

[13] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-030-00889-5_1.