Atlas based breast registration and segmentation in the Mediolateral Oblique and Craniocaudal Views

Athira K S[a*], Janaki Peruvamba Dharmarajan[b], Vijaykumar D K[c] and Nagesh Subbanna[d]
[a*, d]Center for Wireless Networks & Applications (WNA), Amrita Vishwa Vidyapeetham, Amritapuri, India
[b]Department of Radiology, [c]Department of Breast and Gynaec - oncology, Amrita Institute of Medical Sciences and Research Center, Kochi, India

athirakalladayil@gmail.com, janaki19727@aims.amrita.edu, dkvijaykumar@aims.amrita.edu, snag1978@yahoo.com

***Abstract*— In the clinical evaluation of the risk of breast cancer, breast density obtained from analyzing mammograms is of high importance. Higher breast density increases the likelihood of getting breast cancer and also the likelihood of radiologists missing small lesions. Consequently, we analyze breast mammograms to determine which mammograms need greater radiological review. In this paper, we come up with a new atlas-based segmentation technique to register and segment the breast region in mammograms. Atlas-based segmentation is a technique that seeks to segment the pertinent anatomy in medical images. This method is suitable for segmentation of images, whenever there are ambiguous relationships between regions and pixel intensities. This paper evaluates how an atlas-based method can be used to segment the breast region and breast tissue from a mammogram. In our dataset, we have exactly 50 CC images and 15 MLO images. We show the segmented images and, using the dice similarity coefficient to find the performance of the segmentation, show that the accuracy of the segmentation using atlas-based techniques are at 0.87.**

Keywords*— Atlas; Registration; Segmentation; Advanced Normalization Tool (ANTs); breast mask

I. Introduction

Breast cancer is a condition in which the breast's cells proliferate out of control. According to the Indian Council of Medical Research (ICMR) [1], more women than males are receiving cancer diagnoses in India, where 1.46 million people are expected to be diagnosed with the disease this year and 1.57 million in 2025. The World Health Organisation (WHO) [2] says that according to recent global cancer estimates from the International Agency for Research on Cancer (IARC), 2.3 million female breast cancer cases have been discovered. Female breast cancer has overtaken all other cancer types as the most frequently detected malignancy. Breast cancer is four to six times more likely to occur in women with thick breast tissue. Breast density is significant in the clinical assessment of the risk of breast cancer since it is nearly usually the breast tissue, not the fatty tissue, that develops cancer.

In the paper [3], the current scenario and challenges of breast cancer in India are presented. Early identification of breast cancer is crucial because it frequently increases the likelihood that the patient will be able to choose from a wider range of extremely effective, less invasive treatment choices [4]. Different methods related to detection of breast cancer are explained in paper [5], [6]. These techniques include using x-rays, ultrasound, radio waves, gamma rays, microwave imaging, and thermal methods. The standard practice is to have the radiologist examine the mammograms and segment out the suspicious regions, so that the patients can be investigated further. However, the number of radiologists is limited, which makes it hard for radiologists to evaluate huge numbers of images, and further, radiologists do not all classify in the same manner [inter-rater variability]. In fact, the same rater segments suspicious regions differently at different times [intra-rater variability]. Consequently, there is a need for automatic techniques that can detect breast densities and suspicious areas consistently, accurately, and swiftly.

Given this need, it is unsurprising that many different diagnosis methods including machine learning based techniques [7], deep learning-based techniques, based on fractal properties [8], U-Net related [9], logistic regression based [10] etc, have been used. However, before the breast tissue can be segmented out, it is important to pre-process the mammograms for further processing [11].

In this paper, mammograms are considered. Mammograms are x-ray images that are taken using x-ray machines. Mammography is the process of taking x-ray images. In front of a specialized X-ray machine, the person will stand, and the breast will be put on a horizontal plate by a technician. The breast will be tightly pressed from above by another plate. The breast will be flattened by the plates, and x-ray beams will pass through them. To create a side view of the breast, the stages are tilted and steps are repeated. There are mainly two standard views for mammography. Mediolateral oblique (MLO) view and craniocaudal (CC) view. MLO view is taken from the side, usually at an angle of 45 degrees, and the angle changes between 45 and 60 degrees, whereas CC view is taken from top to bottom.

Breast density is now the most reliable indicator of breast cancer, and accurate segmentation of breast masses is crucial to lowering mortality rates. The first step towards this is the extraction of the breast area from the mammogram. Segmentation of mammograms includes segmentation of the breast area and breast tissue so as to determine the density since it plays a pivotal role in breast cancer. Breast cancer is 4–6 times more likely to occur in women with thick breast tissue. Segmenting the breast from its surroundings is the first stage. There are various segmentation techniques [12], which include those based on edges, thresholds, regions, machine learning, and deep learning. Edge-based segmentation includes techniques like Canny edge detection, energy minimization, active contour, Sober, Prewitt, etc. Threshold-based segmentation includes overall and distinct thresholds, histogram detection, adaptive thresholds, etc. Region-based segmentation is connected to the growing of regions, their splitting, merging, etc. Support vector machines and the k-means algorithm are examples of supervised learning. Deep learning-based segmentation is represented by U-Net, SegNet, and others.

The first step to segmenting the breast tissue is to segment the breast area. There are not many techniques that explicitly segment the breast region. Methods like boundary detection and the edge-based method are commonly used. But here, an atlas-based segmentation method is used to extract the breast area and tissue. A number of imaging modalities and domains, including lung imaging, cardiac imaging, and neuroimaging, have successfully adopted the atlas-based approach. This is a topic that hasn't yet garnered significant attention for mammography breast region segmentation.

This paper focuses on the procedures for Atlas-based algorithms and how to use them in MLO and CC perspectives to get segmented images. Papers like [13], [14], [15], [16], and [17] deal with segmentation using atlases. Paper [13] is a survey on atlas-based segmentation. The purpose of the paper is to provide a categorization scheme and a list of distinct techniques for atlas-based segmentation. In paper [14], it deals with methods to enhance Atlas Based Automatic Segmentation (ABAS) in clinical settings. Paper [15] deals with every aspect of atlas-based segmentation. In [16], a multi atlas approach for estimating the position of the nipple and the breast muscle region in digital mammograms is provided. In paper [17], multi-atlas segmentation algorithms are reviewed, along with studies that have used these algorithms to solve diverse biomedical issues.

Section I provides an introduction to breast cancer and various breast segmentation methods. Section II mentions atlas-based segmentation. Section III discusses the proposed solution, Section IV discusses the performance metrics, Section V is the discussion, and the last section wraps up the paper.

II. Atlas Based Segmentation

Atlas-based segmentation is an algorithm used to segment the target image. The atlas-based segmentation will separate the image even when there is no clear relationship between the brightness of different regions and individual pixels. Atlas is a template or a model image that is segmented by a skilled operator. For later comparison and integration, different breast maps might be superimposed on atlases as standardized templates. There are mainly two phases involved in segmenting images using the Atlas: [15], [16]. They are atlas creation and segmentation using an atlas. The two stages of creating an atlas are image selection and template construction. Image registration and segmentation are the two components of segmentation. Image registration includes rigid, affine, and deformable registrations. Fig. 1 shows the flow chart of atlas-based segmentation.

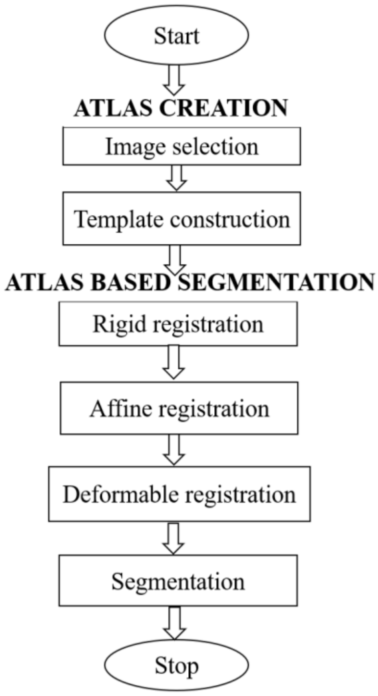

Fig.1: Flow chart for Atlas-based segmentation

A. *Image Selection*

There are four distinct methods for choosing images for creating an atlas. They are: division using a single and unique atlas, using the individual atlases of the most similar segments, using an averaged atlas, and using various atlases. In division using a single, unique atlas, one image is randomly selected as an atlas. But the selection depends on some conditions, like the quality of the image, images with fewer artifacts, etc. In the individual atlases of most similar segments, the input is compared with the pool of images or datasets, and the closeness between the images after the registrations or mappings to one another is considered. In utilizing average form models, the average shape atlas is chosen. In using various images, multiple individual atlases are overlapped with the input image and combined together by decision fusion to produce the output.

B. *Template Construction*

Template construction is done by averaging the set of selected images. Based on the image selection method, images are selected and a template is made. This template will serve as the atlas or reference upon which other images will be overlapped.

C. *Image Registration*

The process of overlaying images to achieve maximum correspondence is known as image registration. Both the reference image, also known as the moving image or the template image, and the static image, sometimes known as the fixed image, or the input image, are synchronized. The energy function (E) is utilized to define correspondence between them, and output is then employed to control the transformation function (T). If the energy function (E) is zero, there will be maximum correspondence between the template and input images.

1) Rigid registration: A model of linear transformation is called rigid registration. It serves as the initialization for both rigid and affine transforms, and the orientation of a figure won't change in any way. Sliding motion, changes in position with respect to the x and y axis, mirror imaging, etc. are included in rigid transformation or registration.

2) Affine registration: It is a linear mapping technique that will maintain and conserve points and lines. It will allow the sliding motion, enlarging or shrinking, rotation, and shearing.

3) Deformable registration: Stretching one of the picture volumes to match the other will result in deformable registration. It has the power to alter the image's size. Shear and dilation are two deformable transformations.

Fig. 2 represents the deformation of the input image with respect to the moving image. Here the moving image or the template is randomly selected from the input images. These images show how deformation or registration happens. In fig. 2 the fixed image or the input image to be segmented is denoted as (a). Image (b) is the moving image or the template. The template will deform with respect to the input images and overlap with each other for segmentation. When closely observed, the template becomes aligned with the fixed image. When overlapped and segmented, the unwanted artifacts and muscles marked in red will be removed.

D. Segmentation

The template, which is the reference image, maps to the input image and will undergo rigid, affine, and deformable transformation and registration. Similarly, an expert from the medical field will mark the breast area out of the template called the breast mask, which will undergo the same deformation as the template. The input image is overlapped with the breast mask to segment the breast area.

III. Proposed Solution

We register the breast mammogram against the constructed breast atlas and then segment the breast from the surrounding regions. Advanced normalization tool (ANTs) [20] is used to register breast mammograms against the atlas. In our experiment, mammograms imaged in the craniocaudal (CC) view and mediolateral oblique (MLO) view were utilized to construct the atlas and have exactly 50 CC images and 15 MLO images as datasets. The atlas is created from 6 CC breast mammogram images and 3 MLO images. Once the atlas is created, the breast area and breast tissues of the template is marked out by an expert [breast mask image]. Then, we register the template to the breast mammogram under question using both affine and deformable registration. Using those transformations for the breast mask of the template, the breast region in the mammogram under question is obtained. The breast tissue in this breast region is used for further training of deep learning-based networks to estimate the breast density. Fig. 4 clearly indicates the template built using an atlas, the deformation of images using the Advanced Normalization Tool, the deformed breast mask, and the overlapped image.

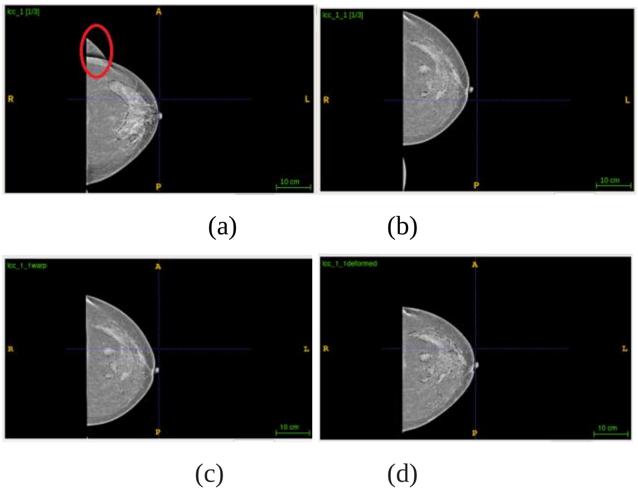

Fig. 2: (a) is the input image or the fixed CC image, (b) is the moving image (randomly selected) that undergoes affine transformation (c) and deformable transformation (d).

A. Template and breast mask constructions

Fig. 3 represents the left and right craniocaudal (CC) template and corresponding breast mask. Template is built using the Advanced normalization tool (ANTS). It will extract data from intricate datasets, such as image data. The main building blocks of ANTs are image registration using variable transformation models (rigid, affine, and deformable) and similarity metrics (landmarks, cross-correlation, and mutual information). antsMultivariateTemplateConstruction2 which is part of ANTS registration is used to build the template. antsMultivariateTemplateConstruction.sh, antsMultivariateTemplateConstruction2.sh, buildtemplateparallel.sh etc. are some of the commands to build the template. Once the template is built, an expert, normally a radiologist will mark the breast area. Here, the breast area is marked using ITK-SNP software. ITK-SNAP is a multi-platform, free, open-source software program used to separate out structures in 3D and 4D biological pictures. This template will serve as the reference for any input images.

B. Atlas based segmentation

1) Breast area segmentation: Once the template is constructed by the above method, it will undergo deformations and warps to the input image. The breast mask (black and white) will undergo the same transformation to fit the input image. Once the deformation is done, the input image and the deformed masks are overlapped and segmented out. Image (f) in Fig. 4 and Fig. 5, shows how the artifacts are removed from the input and required breast area is segmented out.

2) Breast tissue segmentation: Once the breast area is extracted from the mammogram, the same transformation as the mammogram underwent to conform to the input image is applied to the breast tissue mask of the template [atlas], to extract the breast tissues from the input mammogram. The breast tissue mask of the input image, and the corresponding

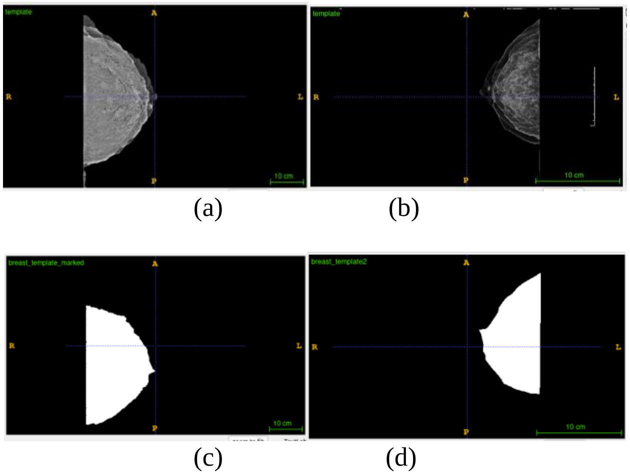

Fig. 3: (a) and (b) are the left and right CC templates made by combining 6 CC images, (c) and (d) are the marked left and right CC templates

breast mask is overlapped and a pixel-wise AND operation is performed to extract the exact breast tissue of the input image. Fig 6 represents the overlapped breast tissue on the original image and overlapped breast tissue on the deformed breast mask image, removing the false breast tissue segmentation outside the breast regions.

IV. RESULTS

A. Qualitative analysis

Fig. 4 and Fig. 5 represent the final outcomes of the proposed method. Both MLO and CC images are considered for segmentation. In Fig. 4, which is a CC image, the image (a) is the input with some artifacts in the form of muscles. These artifacts have to be removed to extract the breast area alone. Image (b) is the template formed by the above proposed method. This template will undergo deformations. The breast mask, which is segmented by an expert, will undergo the same transformations and overlap with the input image. Image (d) is the overlapped image, and the blue shaded region gives the segmented output. Similarly, in Fig. 5, MLO images are considered. From the input image, the pectoral muscles and artifacts have to be removed to extract the breast area for further calculation. Pectoral muscles appear hyperintense in mammogram images, which cause false positives when breast tissue is considered. So, it is very important to remove the muscles. The template will deform with respect to the input image. Image (d) is the breast mask which indicates the breast area alone. Image (d) indicates the overlapped image, and blue shaded shows the segmented part. Due to the limited number of inputs used to construct the template, the deformed template in Fig. 5 is slightly warped. As a result, the breast tissues are not easily visible. Thus, the template cannot be employed for the examination of breast tissue in the MLO image.

B. Performance Metrics

Several metrics—also referred to as performance metrics or evaluation metrics—are used to assess the effectiveness or quality of the model. Some of the performance matrices are accuracy, precision, recall, F1 score, dice similarity coefficient,

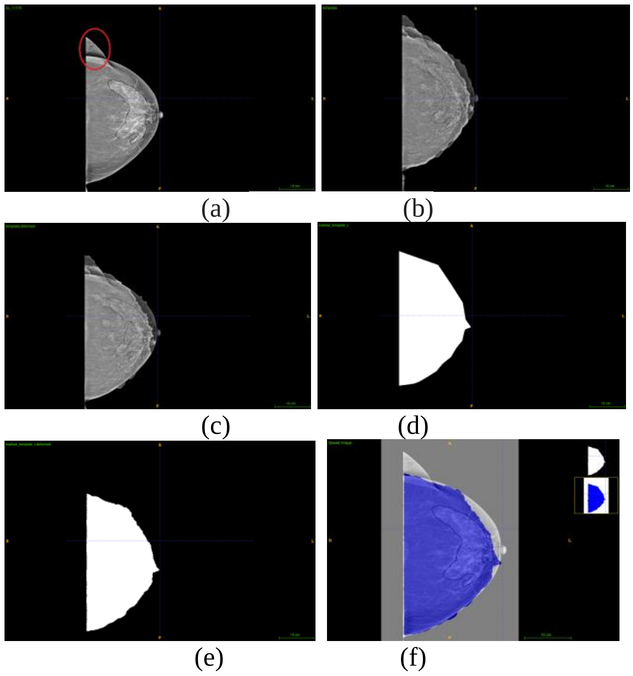

Fig. 4: (a) is the input image or the fixed image, (b) is the template (c) is the deformed template with respect to the input image, (d) is breast mask (e) is the deformed mask with respect to the input image and (f) is the overlapped image

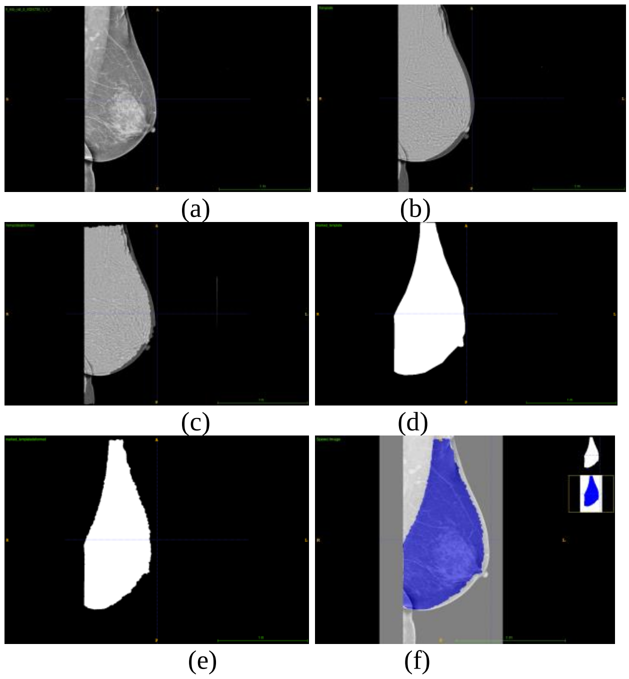

Fig. 5: (a) is the input image or the fixed image, (b) is the template (c) is the deformed template with respect to the input image, (d) is breast mask (e) is the deformed mask with respect to the input image and (f) is the overlapped image

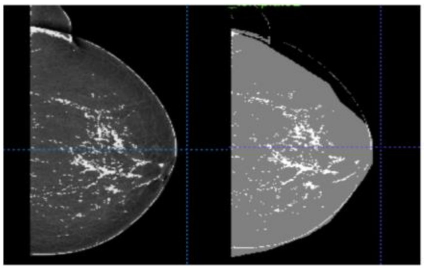

Fig. 6: Breast tissue segmentation

confusion matrix etc. In this paper dice similarity coefficient is considered for measuring the similarity. It is a statistical matrix that is used to assess how similar two samples are to one another. It has a scale from 0 to 1, with 0 denoting no overlap and 1 denoting complete overlap. A higher Dice coefficient value suggests that the two sets are more comparable.

Here, A and B are two sets of data. $A \cap B$ represents the intersection of A and B. $|A| + |B|$ represent the cardinality of A and B. From the above method, it was able to segment the input image with a dice similarity coefficient of 0.87 which is 87% similarity between the input image and deformed mask for CC images and 76% for MLO.

V. Results

In this paper, we present an atlas-based registration and segmentation method that aims to segment the breast area and tissue from mammograms. The proposed algorithm was tested in 50 CC images, 15 MLO and obtained a similarity of 87% for CC images and 76% for MLO images. From the dataset, 6 CC breast mammogram images and 3 MLO images are used for template construction. The segmentation accuracy of MLO images can be enhanced by increasing the number of images used for template construction. Even though the time requirement is more in atlas construction, it can be resolved by registering the image after down sampling, implementing a parallel computing environment like the Graphics Processing Unit (GPU) or by reducing the number of pixels processed. However, the atlas-based method has to be tested using numerous images to get better accuracy and speed, which lies in future research.

Conclusion

The atlas-based method has been effectively used in a variety of imaging modalities and fields, including pulmonary imaging, cardiac imaging, and neuroimaging. For mammography breast region segmentation, this hasn't yet received much attention. But in these years, it has been an important topic for research. This paper tried to explain about atlas-based registration, template constructions, atlas-based segmentation, breast tissue segmentation, performance metrics and its limitations. We believe that this wide topic from the field of image analysis has provided one concrete starting point for future investigations into atlas-based segmentation.

References

[1] P. Mathur, K. Sathishkumar, M. Chaturvedi, P. Das, K. L. Sudarshan, S. Santhappan, V. Nallasamy, A. John, S. Narasimhan, F. S. Roselind et al.,"Cancer statistics, 2020: report from national cancer registry programme, india," JCO global oncology, vol. 6, pp. 1063–1075, 2020

[2] J. FERLAY, "Lam breast cancer overtakes lung cancer in terms of number of new cancer cases worldwide," 2021

[3] R. Mehrotra and K. Yadav, "Breast cancer in india: Present scenario and the challenges ahead," World Journal of Clinical Oncology, vol. 13, no. 3, p. 209, 2022

[4] D. Khadakban, T. Gorasia-Khadakban, D. Vijaykumar, K. Pavithran, and R. Anupama, "Factors associated with better survival after surgery in metastatic breast cancer patients," Indian journal of surgical oncology, vol. 4, pp. 52–58, 2013

[5] T. Swathi, S. Krishna, and M. V. Ramesh, "A survey on breast cancer diagnosis methods and modalities," in 2019 international conference on wireless communications signal processing and networking (WiSPNET). IEEE, 2019, pp. 287–292

[6] S. Sharma, K. Hiran, K. Pavithran, and D. Vijaykumar, "A pilot study to assess the feasibility of evaluation of markers of response to chemotherapy at one day & 21 days after first cycle of chemotherapy in carcinoma of breast: a prospective non-randomized observational study," World journal of surgical oncology, vol. 7, no. 1, pp. 1–11, 2009

[7] S. S. Shastri, P. C. Nair, D. Gupta, R. C. Nayar, R. Rao, and A. Ram, "Breast cancer diagnosis and prognosis using machine learning techniques," in Intelligent systems technologies and applications. Springer, 2018, pp. 327–344

[8] S. Don, D. Chung, K. Revathy, E. Choi, and D. Min, "A new approach for mammogram image classification using fractal properties," Cybernetics and information technologies, vol. 12, no. 2, pp. 69–83, 2012

[9] S. Yellari, J. P. Dharmarajan, S. S. Karve, D. Vijaykumar, and N. Subbanna, "Estimation of breast density from human mammograms using u-nets," in 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA. IEEE, 2022, pp. 1–6

[10] T. Singh, R. Jha, and R. Nayar, "Mammogram classification using multinominal logistic regression," in International Conference on Communication and Signal Processing-ICCSP'17 at Adhiparasakthi Engineering College Melmaruvathur, 2017

[11] R. Shwetha and B. Rajathilagam, "Super resolution of mammograms for breast cancer detection," International Journal of Applied Engineering Research, vol. 10, no. 1, pp. 21 453–21 465, 2015

[12] E. Michael, H. Ma, H. Li, F. Kulwa, and J. Li, "Breast cancer segmentation methods: current status and future potentials," BioMed Research International, vol. 2021, pp. 1–29, 2021

[13] H. Kalinic, "Atlas-based image segmentation: A survey," Croatian Scientific Bibliography, pp. 1–7, 2009

[14] D. Ciardo, M. A. Gerardi, S. Vigorito, A. Morra, V. Dell'Acqua, F. J. Diaz, F. Cattani, P. Zaffino, R. Ricotti, M. F. Spadea et al., "Atlas-based segmentation in breast cancer radiotherapy: evaluation of specific and generic-purpose atlases," The Breast, vol. 32, pp. 44–52, 2017

[15] T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and C. R. Maurer Jr, "Quo vadis, atlas-based segmentation?" in Handbook of Biomedical Image Analysis: Volume III: Registration Models. Springer, 2005, pp. 435–486

[16] J. E. Iglesias and N. Karssemeijer, "Robust initial detection of landmarks in film-screen mammograms using multiple ffdm atlases," IEEE transactions on medical imaging, vol. 28, no. 11, pp. 1815–1824, 2009

[17] J. E. Iglesias and M. R. Sabuncu, "Multi-atlas segmentation of biomedical images: a survey," Medical image analysis, vol. 24, no. 1, pp. 205–219, 2015