K-Means and Active Contour Approach to Detecting and Segmenting Masses in Breast MR Images

Justin Kennedy[1], William Hsu[2]

[1]Applied Mathematics, [2]Radiological Sciences and Bioengineering, University of California, Los Angeles

INTRODUCTION

While the five-year relative survival rate for women with stage 0 or stage I breast cancer is 100%, the survival rate of cancers caught at stage III and stage IV drops to 72% and 22%, respectively[1]. Effective methods for early detection, particularly in high risk women with extremely dense breasts, can significantly improve patient outcomes. As a systematic, non-invasive approach to examining the malignancy potential of masses in magnetic resonance (MR) images, image quantification has proven to be an effective method for evaluating MR images and efficiently scaling results to multiple series and patients.

In this work, we implore the use of machine learning approaches such as k-means clustering and active contours to automatically detect and segment masses in breast MR images. To evaluate the efficacy of the algorithm, we compare the algorithm's performance to human annotations of the same studies. Our goal is to create an effective supplemental tool for clinical decision making in regards to the evaluation of breast MR images.

DATASET

The images used in this study were MR dynamic contrast enhanced (MR-DCE) images from the Cancer Imaging Archive (TCIA) dataset (n=87)[2]. The cases contained high-risk normals, ductal carcinoma in situ (DCIS), fibroids, and lobular carcinomas acquired using 3 or more distinct MR pulse sequences on a Phillips 1.5T scanner. Specifically, the sequences labelled BLISS were processed and used for training and testing the algorithm. BLISS is an MR technique that improves fat suppression and achieves high spatial resolution of the breast. The volume of Bayer gadolinium contrast injected into the brachial vein is amount based on 10% of the patient's weight in pounds, and the injection is 6 or 7 seconds at a rate of 3cc per second. The first dynamic sequence is started 1 minute after the injection is started.

METHODS

Data

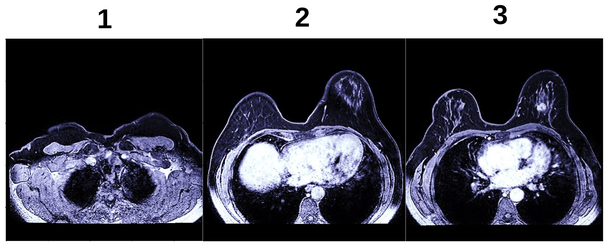
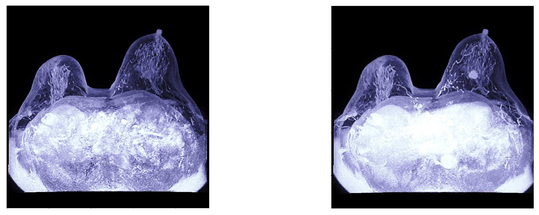


Figure 1. The input DICOM image is represented as an array of different sequences, separated by Trigger Time (e.g., time after contrast is injected).

Maximum Intensity Projection (MIP)

Figure 2. The MIP projects the highest pixel values from each vertical ray of the 3-D image set to one 2-D image.

Peak Trigger Time



(a) **(b)**

Figure 3. The MIP is then taken for each trigger time array to find the 'peak trigger time', or when the resolution with contrast shows the greatest difference with the previous trigger time. In this case (b) is the Peak Trigger Time (PTT) MIP and (a) is the MIP of the trigger time before it.

K-Means Clustering

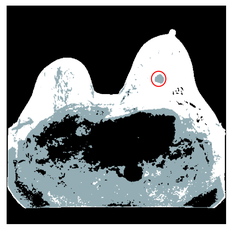


Figure 4. The PTT MIP is then clustered using k-means clustering (k=4) with three feature vectors to better isolate potential masses from the rest of the breast and chest cavity. This MIP is then compared to the segmented MIPs of the previous trigger times to derive a coordinate-based location and radius for the mass(es).

Active Contouring Segmentation

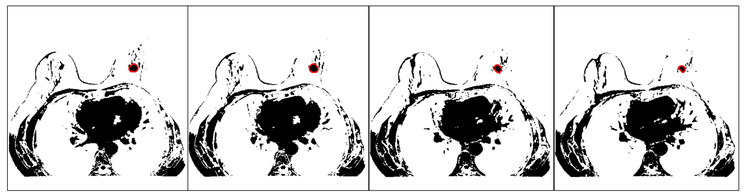


Figure 5. The original image array is looped through to find images containing the detected mass. A Morphological Snake Active Contouring Algorithm[4] initializes at the center of the mass in each image and expands outward until detecting the mass' boundary.
The snake algorithm works by seeking a contour boundary for separating an image into a foreground and background through solving a system of partial differential equations on a binary circle level set. The PDE system is based on the general form:

$$\frac{\partial u}{\partial t} = |\nabla u| \left(\mu \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) - \nu - \lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2 \right).$$

Binary Segmentation Output

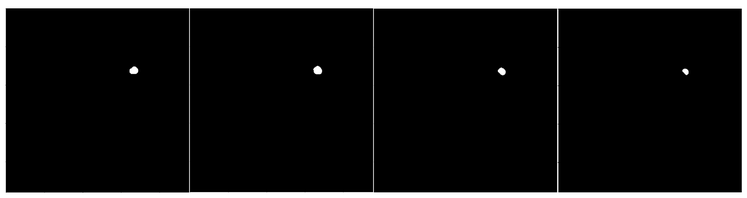


Figure 6. The resulting mass segmentations are converted to binary output images.

Evaluation

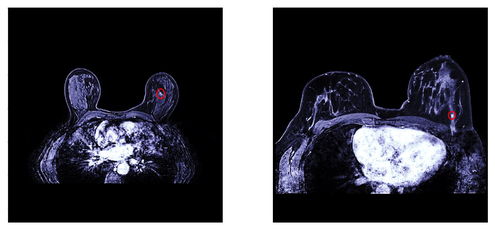
A subset of 10 BLISS MR sequences was used for training the algorithm. 10 patient cases were randomly selected from the TCIA dataset for testing. The algorithm's accuracy was determined through comparison with manual segmentations of 30 slices using the Sorensen-Dice Coefficient, a statistic commonly used to measure the similarity between binary data sets.

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$,where X and Y are the elements of each set.

RESULTS

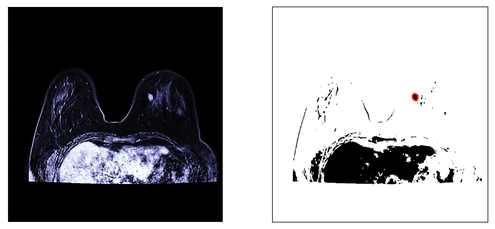
	TCIA Dataset	
	DICE AVG	DICE STDEV
Proposed	**0.91**	**0.05**

Table 1. Results of mass segmentations for proposed method (n=30).



(a) **(b)**

Figure 7. Examples of DICOM image slices where the contained masses were not detected by the algorithm.



(a) **(b)**

Figure 8. Example of a DICOM image slice where the contained mass was detected and segmented by the algorithm.

DISCUSSION

The results of the segmentation comparisons showed the algorithm was effective at segmenting the breast MR masses it detected. A current limitation of the algorithm is a difficulty in detecting smaller masses not as visible to the naked eye in the original DICOM images.

For future work, with the help of radiologists to identify more common features of tumors, the algorithm can be trained to better detect masses of smaller size, specifically in adjusting the clustering component. Furthermore, in being able to better differentiate mass tissue versus extraneous tissue such as fat through the addition of insightful features, the algorithm's performance can be improved, particularly on masses more heavily shrouded by excess tissue.

Additionally, quantitative features such as shape and texture can be extracted from the resulting 3-D segmentations of the algorithm to create models predicting probabilities of malignancy for the detected masses.

Acknowledgements:
We would like to thank Dr. Stephanie Lee-Felker for providing helpful feedback on identifying characteristics of tumors and trigger time properties in MR Images.

References:

[1] American Cancer Society. "Breast Cancer Survival Rates". https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html (accessed May 2018).
[2] Bloch, B. Nicolas, Jain, Ashali, & Jaffe, C. Carl. (2015). Data From BREAST-DIAGNOSIS. The Cancer Imaging
[3] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057
[4] *Morphological Snakes*. Luis Álvarez, Luis Baumela, Pablo Márquez-Neila. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2010 (CVPR10).