JOURNAL LA MULTIAPP

VOL. 03, ISSUE 02 (037-043), 2022
DOI: 10.37899/journallamultiapp.v3i2.567

Comparative Analysis of Mammography Image Segmentation Strategies

Areej Rebat Abed[1], Karim Q. Hussein[2]

[1]*Computer Science, Iraqi commission for computer and informatics, Informatics Institute for Postgraduate, Studies, Baghdad, Iraq*
[2]*Assist. Prof, Computer Science Dept./Faculty of Science, Mustansiryha University, Baghdad, Iraq*

**Corresponding Author: Areej Rebat Abed*
Email: areejaa372@gmail.com

Article Info

Article history:
Received 22 February 2022
Received in revised form 25 March 2022
Accepted 30 March 2022

Keywords:
Breast Cancer
Segmentation
Mammography
Computer-Assisted Diagnosis (CAD)

Abstract

Breast cancer is a serious medical problem that affects women all over the world, and it is one of the most well-known tumors that kill women. The specialists of Breast cancer Prefer to use imaging methods such as a mammography to speed up recovery and reduce the risk of breast cancer. An ROI describe the tumor will be retrieved from the image that is entered to detect a malignant tumor. One of the basic techniques used to classify breast cancer is segmentation. Segmentation may be difficult in the presence of noise, blurring or low contrast. Pre-processing aids in the removal of extraneous data from a picture or the enhancement of image contrast in the early stages. Classification is greatly influenced by segmentation. Recent research have presented automatic and semi-automated segmentation algorithms for extracting the region of interest (ROI), lesions, and masses to check for breast cancer. In this study provides high-level overview of approaches of segmentation, with a focus on mammography images from current research. The datasets that were available were discussed as well as the problems encountered during the segmentation operation for the identification of breast cancer.

Introduction

The most frequent malignancy in women over the age of 40 is breast cancer, and early detection is critical to lowering fatality rates. Mammography, a low-dose X-ray imaging technology, is one of the imaging techniques used for breast cancer screening and early diagnosis. It exposes the existence of breast cancer from the surrounding infiltrating breast tissues (Divyashree, & Kumar, 2021). Mammography is one of the most often used diagnostic methods for detecting and classifying breast abnormalities. Ultrasound, magnetic resonance imaging (MRI), X-ray imaging, and novel technologies such as molecular breast imaging and digital breast tomosynthesis are all alternatives for evaluating the breast (DBT). Mammography is a type of imaging that uses low-dose X-ray technology to inspect the breast. It's the most accurate approach to spot breast cancer before it becomes clinically palpable (Heck & Herzen, 2020). Mammography is the current imaging method for detecting and diagnosing breast cancer in its early stages. False positives occur when a mammography detects abnormal regions in the breast that appear to be cancer but are actually normal. Because mammogram pictures are difficult to interpret, a computer-assisted diagnosis (CAD) is becoming a more important tool to aid radiologist in mammographic lesion interpretation (Kshema, 2017). Death rates have been reduced by 30–70% as a result of screening programs (Nguyen et al., 2019). In mammography image analysis, preprocessing is considered to be a crucial step. The success of the remaining processes, such as segmentation and classification, is determined by the accuracy of the preprocessing (Kshema, 2017). These methodologies are classified as region-based segmentation paradigms (gathering pixels into similar areas in enormous computation time due to higher resolution) (Jan et al., 2020), contours-based segmentation (Because of the Shading asymmetry, can identification the edges of the image which divides the image into distinct areas and can also identification the level of opacity).Grouping-based segmentation (Grouping of pixels that clustered which have similar characteristics) (Shih, 2017). strategies of thresholding (The foreground is separated from the background using data that global as a graph) and energy function related techniques (advancement of curves to recognize objects).

Each of these treatments have hold set of advantages and disadvantages. Because of higher resolution, the region-based technique is widely commonly used; however, it takes longer. When it comes to grouping, edge-based and thresholding techniques are easily affected by noise since they rely on initial clusters and an energy function. Table1 provides overview of their benefits and drawbacks. This work primarily take a look of several segmentation techniques propose in the previous works, primarily for mammography image. In Section II, there is overview about the datasets that used in segmentation, and breast cancer statistics. Section III contains new researches on mammography image segmenting by a number of authors, and Section IV conclusions drawn.

Table 1. Main Classes for Methods of Segmentation for Mammography Technique

Mechanism	Benefits	Drawbacks
Region-based (Karthick et al., 2014; Bandyopadhyay, 2010).	It takes more time; seed point selection is required.	Noise-resistant: For the homogeneity component, increment is favored over splitting and merging; otherwise, splitting and merging are required. Closed borders are provided by a watershed.
Edge-based (Dharani et al., 2016; Bandyopadhyay, 2010).	Under large or improperly Defined boundaries, The situation becomes complicated. Segmentation is poor, and the edges are frayed.	Fast: copes well with images with high contrast
Clustering-based (Mazur et al., 2016; Zafar & Ilyas, 2015).	Sensitive to early clusters and outliers, cosy	It works well with overlapping data.
Thresholding- based (Guruprasad, 2020)	Doesn't perform well with noise and low contrast; thresholding is complicated.	It's as simple as that.

Breast Cancer Statistics and Datasets

From of chronic diseases a mong women is Breast cancer which affected hug numbers women each year and resulting in a large number of deaths. Breast cancer claimed lives of 627,000 women in 2018, accounting for roughly 15% of all illness fatalities among women. Breast cancer averages are higher through women, spatially in greatly develop places, and they increase in practically in each country all over the world. According to WHO statistics de Martel et al. (2020) the top five diseases have the highest incidence and mortality rates (breast, lung, stomach, liver, and prostate) as shown in figure 1.

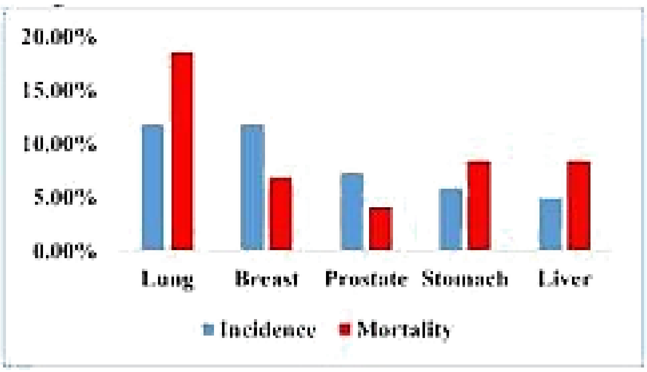

Figure 1. The Most Prevalent Types of Cancer

Several datasets and databases have been developed, and transferred it to various locations for the purpose of storing throughout the last few years. Some datasets are open to the public and can be used. Several of the datasets can be come in CSV format or in image format, most likely. Through the study of images of breast cancer, these databases are utilized. Various researchers, for example, rely heavily on the Mammography Image Analysis Society (MIAS) database. Figure 1 shows a database of 322 image samples, with 208 conventional images and 114 abnormal images (51 malignant and 63 benign instances). Also there is another dataset that is extensively used and well-known for recording DDSM which is the Digital Database, which has 2500 images. Table 2 shows a summary of the most commonly cited and updated breast cancer databases.

Table 2. Datasets Used to Segment Mammography

Datasets	Size of dataset	Format
DDSM	55,890	jpeg
Breast histopathology	277,524	png
Breast cancer proteomes	64	CSV
CBIS-DDSM	10239	DICOM
Breast US image	250	-
MIAS	322	pgm
Breast cancer Wisconsin 569 CSV (Diagnostic)	569	CSV
NKI Breast cancer Data	272	CSV
UMD-BRCAI/BRCA2	11793	-
InBreast	410	XML
Araujo	260	-

Literature Review

Different segmentation strategies are examined in Table3 depending on methodologies used, datasets used, and other factors to determine the best procedure for mammography image segmentation.

Table 3. Different approaches for segmentation along with its advantages

Ref	Technique	Illustrate	Datasets	Advantages
(de Martel et al., 2020)	Cropping augmentation, and balancing pre processing processes should be used.	Cropping We crop each image to remove black areas in the mammography image, then apply label-preserving changes to the original images to generate additional examples, resulting in 8 new labels for each subsequent mammogram. After that, evenly distribute each category.	mini-MIAS and BCDR	Reduce the amount of time it takes to train the classifier and improve its performance
(Hepsa Et al., 2017)	ROI images that have been extracted	Using the binary segmentation masks given in the CBIS-DDSM dataset, the ROI images are recovered from the mammography image.	CBIS-DDSM dataset	improve a Fine Tuned model's performance
(Falconí et al., 2020)	Novel multi-Level nested Pyramid network (MNPNet)	The MNPNet includes an encoder and a decoder. The former encodes contextual information, low-level detail information, and high-level semantic information in a multi-level multi-scale manner by multi- level nesting at rous spatial pyramid pooling (ASPP) module on the feature pyramid Generated by modified ResNet34.	INbreast and DDSM-BCRP	Refine the Segmentation results along mass boundaries
(Wang et al., 2019)	Global thresholding and Otsu segmentation	To distinguish between the breast area and the background in CBIS-DDSM, Otsu segmentation is used. In the case of IN breast, global thresholding is used to separate the breast region from the background, and all right breasts are horizontally mirrored to maintain the same orientation.	CBIS-DDSM IN breast	To make the breast area stand out from the background
(Agarwal et al., 2019)	A semantic segmentation method for breast	The contrast limited adaptive histogram equalization method was used to	The first dataset contains of 264 pictures, 100 of which are malignant and	The results of The experiments

ISSN: 2716-3865 (Print), 2721-1290 (Online)

	ultrasound (BUS) images	enhance the BUS images after they were resized. The pre processed image was then encoded using the variant enhanced block. Finally, The segmentation mask was created by concatenating the convolutions	164 of which are benign BUS. The second dataset contains 830 photos, 487 of which are malignant, 210 benign, and 133 of which are normal	reveal that the network can segment breast lesions more precisely than any of the competitors
(Xue, et al., 2021)	Applying efficient morphological based techniques	The objects are divided into two sorts of areas, one for the foreground object and the other for the background object. The Otsu global thresholding approach is used to convert the input cytology image into a binary image	A actual data set of 400 photos was Obtained from different patients at the Lady Reading Hospital's pathology department in Peshawar, Pakistan	Identification of cancer cell
(Bandyopadhyay, 2010)	Applying efficient morphological based techniques	The objects are divided into two sorts of areas, one for the foreground object and the other for the background object. The Otsu global thresholding approach is used to convert the input cytology image	A actual data set of 400 photos was obtained from different patients at the Lady Reading Hospital's pathology department in Peshawar, Pakistan	Requires no prior knowledge of the image's contents or training data

The more substantial part of the techniques are primarily applied on publicly accessible datasets to extract the region of interest(ROI), as evidenced by the correlation . Every strategy has its own set of benefits. Low contrast, noise, foggy edges, and fluctuations in intensities make a few operations difficult to handle. However, a few processes provide excellent precision while working with a single dataset. One can take use of their advantageous conditions to improve outcomes while also addressing their shortcomings by offering various solutions.

Conclusion

Breast cancer is by far the most common cause of death in women. It is critical to have early detection approaches for breast cancer employing modern algorithms. For further analysis, these models can use the image produced by other tests such as mammography. Preprocessing, segmentation, feature extraction, and other steps are used to classification breast cancer kind, for example. A few mammography image segmentation algorithms are examined in this work, including how the methods are used, their roles, and the databases used, in addition to suitable conditions. Bad contrast, noise, blurring, and some of the complex structures, all have an impact on the segmentation process. In classifying breast cancer into separate categories, segmentation accuracy is very important. Hybrid approaches can be used to show evidence of increase in segmentation accuracy after examining recent research publications on the subject.

References

Agarwal, R., Diaz, O., Lladó, X., Yap, M. H., & Martí, R. (2019). Automatic mass detection in mammograms using deep convolutional neural networks. *Journal of Medical Imaging*, *6*(3), 031409.

Bandyopadhyay, S. K. (2010). Pre-processing of mammogram images. *international journal of engineering science and technology*, *2*(11), 6753-6758.

de Martel, C., Georges, D., Bray, F., Ferlay, J., & Clifford, G. M. (2020). Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *The Lancet Global Health*, *8*(2), e180-e190.

de Martel, C., Georges, D., Bray, F., Ferlay, J., & Clifford, G. M. (2020). Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *The Lancet Global Health*, *8*(2), e180-e190.

Dharani, T., Aroquiaraj, I. L., & Mageshwari, V. (2016). Comparative Analysis of Edge Detection Algorithms Based on Content Based Image Retrieval With Heterogeneous Images. *International Journal of Computational Intelligence and Informatics, ISSN*, 2349-6363.

Divyashree, B. V., & Kumar, G. H. (2021). Breast cancer mass detection in mammograms using gray difference weight and mser detector. *SN Computer Science*, *2*(2), 1-13.

Falconí, L., Pérez, M., Aguilar, W., & Conci, A. (2020). Transfer learning and fine tuning in mammogram bi-rads classification. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 475-480). IEEE.

Guruprasad, P. (2020, June). Overview of different thresholding methods in image processing. In *TEQIP Sponsored 3rd National Conference on ETACC*

Heck, L., & Herzen, J. (2020). Recent advances in X-ray imaging of breast tissue: From two- to three-dimensional imaging. *Physica Medica*, *79*, 69-79.

Hepsağ, P. U., Özel, S. A., & Yazıcı, A. (2017, October). Using deep learning for mammography classification. In *2017 International Conference on Computer Science and Engineering (UBMK)* (pp. 418-423). IEEE.

Jan, M. M., Zainal, N., & Jamaludin, S. (2020). Region of interest-based image retrieval techniques: a review. *IAES International Journal of Artificial Intelligence*, *9*(3), 520.

Karthick, S., Sathiyasekar, K., & Puraneeswari, A. (2014). A survey based on region based segmentation. *International Journal of Engineering Trends and Technology*, *7*(3), 143-147.

Kshema, M. J. (2017). George, and DAS Dhas,“. In *Preprocessing filters for mammogram images: A review,” in 2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*(pp. 1-7).

Mazur, V. V., Barmuta, K. A., Demin, S. S., Tikhomirov, E. A., & Bykovskiy, M. A. (2016). Innovation clusters: Advantages and disadvantages. *International Journal of Economics and Financial Issues*, *6*(1S), 270-274.

Nguyen, P. T., Nguyen, T. T., Nguyen, N. C., & Le, T. T. (2019, October). Multiclass breast cancer classification using convolutional neural network. In *2019 International symposium on electrical and electronics engineering (ISEE)* (pp. 130-134). IEEE.

Shih, F. Y. (2017). *Image processing and mathematical morphology: fundamentals and applications*. CRC press.

Wang, R., Ma, Y., Sun, W., Guo, Y., Wang, W., Qi, Y., & Gong, X. (2019). Multi-level nested pyramid network for mass segmentation in mammograms. *Neurocomputing*, *363*, 313-320.

Xue, C., Zhu, L., Fu, H., Hu, X., Li, X., Zhang, H., & Heng, P. A. (2021). Global guidance network for breast lesion segmentation in ultrasound images. *Medical image analysis*, *70*, 101989

Zafar, M. H., & Ilyas, M. (2015). A clustering based study of classification algorithms. *International journal of database theory and application*, *8*(1), 11-22.

ISSN: 2716-3865 (Print), 2721-1290 (Online)