Contribution to the Pre-processing Method for Image Quality Improving: Application to Mammographic Images

Y. Nadji *nadji_yeltongar@yahoo.fr*
Faculty of Exact and Applied Sciences
University of N'Djamena
N'Djamena, Chad

J. Mbainaibeye *mbai_jerome@yahoo.com*
Faculty of Sciences and Technics
University of Doba
Doba, Chad

G. Toussaint *tgnouwa@yahoo.fr*
Mother and Child Hospital
N'Djamena, Chad

Abstract

Breast cancer is the most common type of cancer of women worldwide, but it can be cured if diagnosed at an early stage. Mammography is the main means of cancer screening and provides useful information on the signs of cancer, such as microcalcifications, masses, architectural distortion etc., which are not easy to distinguish due to certain defects in mammographic images, including low contrast, high noise, blurring and confusion. These challenges could be overcome by proposing a new preprocessing model. This work proposes a pre-processing method using different techniques and their combination in order to minimize the above-mentioned defects in mammographic images and make them usable for further processing. The different techniques range from filtering, thresholding, histogram, blur masking, morphological operations, thresholding and cropping. The aim is to put the mammographic images into a representation that will facilitate the detection of microcalcifications and the classification of healthy and cancerous images. Algorithms were developed and tested using the publicly available international database of the Mammographic Image Analysis Society (MIAS), which contains 322 samples. The results obtained on the regions of interest using four samples clearly show the background of the image and the objects (the part of the pectoral muscle and the suspicious area). These results show that much of the adipose tissue, fat mass and some of the features observed in the zoomed-in part of the image are significantly reduced. Furthermore, the results obtained in terms of visual quality compared to the literature show that they are better**.**

Keywords: Breast Cancer, Digital Mammography, Region of Interest.

1. INTRODUCTION

Breast cancer is one of the leading causes of death of women worldwide. It represents 15% of all cancer deaths in women and affects 2.1 million women worldwide each year (Sahar et al, 2021). International Agency for Research on Cancer (IARC) estimates that the incidence of breast cancer will increase by more than a third, to over three million new cases annually by 2040. This increase will be mainly due to demographic factors and lifestyle changes (Katarzyna, 2022).According to world health organization (WHO) in 2018, it could claim 10.30 million lives per year if left unchecked. Screening and early diagnosis are two key strategies to reduce mortality and treatment costs by detecting the disease at an early stage (Sahar et al, 2021). The most commonly used screening tool is mammography (analog or digital). Mammography can reveal whether or not cancer is present in a breast. However, the mammographic image obtained from the breast is difficult to use because of fatty masses (Iman et al, 2022). Therefore, several researchers have worked to improve the contrast of mammographic images by pre-processing. Preprocessing is an important step in a computer-aided diagnostic system (Khan et al, 2020), (Mudassar et al, 2018). Thus, Saleem Z. Ramadan proposed a method based on bi-histogram equalization. This method divides the image into two sub-images of medium-high and low brightness based on the average intensity of all pixels and then histogram equalization is applied to each sub-region independently (Saleem et al 2020). ArdymulyaIswardani and WahyuHidayat made use of negative image transformation and histogram equalization to improve the image (Ardymulya et al, 2018). Btissam ZERHARI used a hybrid approach combining two approaches, a sequential approach and a parallel approach in order to take advantage of both (Btissam, 2020). The authors in (Saraswathi et al, 2017) in combining the ACWE model with the Fuzzy C means clustering method to handle the intensity in homogeneity of images and reduce the presence of noise (Noor et al, 2023). Although various techniques have been proposed in the literature to improve the visibility of the obtained images, no single method gives complete satisfaction (Yosra, 2017). In this work, we propose a pre-processing method using different techniques and their combination in order to reduce as much as possible the defects of the mammographic images and make them usable for further processing.

The rest of this paper is organized as follows: Section 2 presents the material used and the method developed. The implementation of the proposed method is presented in section 3. Section 4 presents the results obtained. The discussions of the obtained results are presented in section 5. Finally, the conclusion and the perspectives of the work are presented in section 6.

2. MATERIAL AND METHODS

2.1 Material

We used the Mammographic Image Analysis Society (MIAS) digital mammography database. MIAS is an organization of research groups interested in mammography. It is based in the United Kingdom (UK) and has produced a database of digital mammograms (Asma, 2021). The X-ray films in the database have been carefully selected from the UK National Breast Screening Programmed and digitized using a Joyce-Lobel scanning microdensitometer to a resolution of 50 μm x 50 μm; each pixel being 8-bit coded. The database contains images of the left and right breast of 161 patients. In total, there are 322 images, belonging to three types, namely normal, benign and malign. There are 208 normal, 63 benign and 51 malign (abnormal) images.

We used this database because it presents several pieces of information about the abnormality, namely the class of the lesion, its location and its size.
For our work, we used the Intel(R) Core (TM) i5 CPU, 4 GB RAM and Matlab R2017b software.

2.2 Methods

First, we performed an exhaustive bibliography search in order to be up to date with the state of the art in mammographic image preprocessing. Then, we tested all the preprocessing methods found in this bibliography. This allowed us to identify the limitations of these methods, and we propose three methods with different combinations. Finally, we have built a database of digital mammography from known digital sites for medical applications, with public access. We will first present a brief state of the art on pre-processing techniques. In a second step, we will present our contribution focusing on a number of combinations of the existing techniques. In a third step, we will test our developed method on a few images from the MIAS database.

2.2.1 “Blur mask” Filter

This filter enhances the edges and increases the sharpness of the image. To obtain this filter, a fraction of the high-pass filtered image is added to the original image as shown give the reference of the figure.

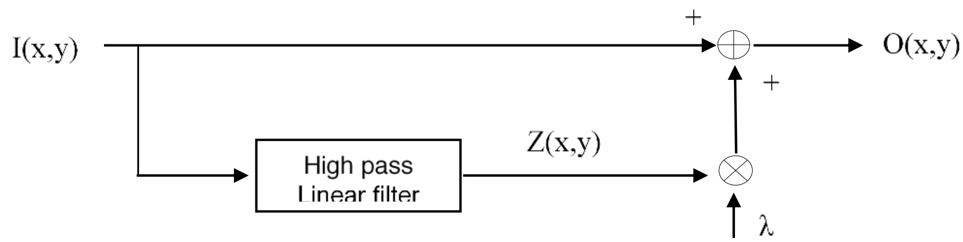


FIGURE 1: Block diagram of unsharp masking filter.

Thus, the image O(x,y) is obtained from the input image I(x,y) according to the equation (1):

$$O(x, y) = I(x, y) + \lambda \times Z(x, y) \quad (1)$$

Where: $Z(x, y)$ is the correction signal calculated by the linear high-pass filter (Sobel filter), λ is the enhancement factor that controls the level of contrast enhancement obtained at the output. The processed images are sharper because the low-frequency information in the mammogram is reduced in intensity while the high-frequency details are amplified (Xiang et al, 2018). This makes microcalcifications more visible on mammograms.

2.2.2 Morphological Operations

Filters are obtained by combining two simple operations, erosion and dilation. These operators are called opening and closing. Opening is defined as erosion followed by dilation while closing is dilation followed by erosion. The opening and closing can be used to obtain two other techniques namely the top-hat shape transform, which is defined as the difference between the original image and its opening, while the bottom-hat shape transform is the difference between the closing image and the original image (Sunil et al, 2022). Image addition and subtraction operations are then performed using the top-hat and bottom-hat transforms to obtain a mammographic image containing much more visible microcalcifications (Rafsanjany et al, 2017).

2.2.3 Thresholding

Thresholding consists of cancelling the coefficients that amplitude is lower than a given threshold. There are two types of thresholding in the literature, namely hard and soft thresholding.

Hard thresholding consists in eliminating the coefficients that values are lower than the threshold; the other coefficients remain unchanged. On the other hand, for soft thresholding, the coefficients that amplitude is lower than the threshold are cancelled; the other coefficients have their amplitude reduced by the value of the threshold.

Equations (2) and (3) define hard and soft thresholding respectively (Jingsong et al, 2019):

$$D_{dur}(w, \lambda) = \begin{cases} 0 & si\ |w| \le \lambda \\ w & si\ non \end{cases} \quad (2)$$

Where: w represents the image coefficients and λ is the threshold.

$$D_{doux}(w, \lambda) = \begin{cases} 0 & if\ |w| \le \lambda \\ sgn(w)(|w| - \lambda) & else \end{cases} \quad (3)$$

Where $sgn(.)$ is the sign function.

The coefficients that amplitude is less than or equal to five, are cancelled; the others have their amplitude reduced by five. According to these two types of thresholding, how the choice of the threshold λ is done?

This choice is done by Donoho. It depends only on the image size and the square root of the variance of the image size, according to equation (4).

$$\lambda = \sigma\sqrt{2}\log(T) \quad (4)$$

Where λ is the threshold, σ is the square root of the variance and *T* is the image size.

2.2.4 Histogram Equalization

One of the most commonly used techniques in image pre-processing is histogram equalization (Yuanmin et al, 2019). Histogram equalization is a contrast adjustment method that aims at transforming the initial histogram into a "standardized" histogram. It consists in distributing the intensities of the pixels of an image over the whole range of possible values in order to increase the contrast. Among the most commonly used techniques in the literature, there are standard histogram equalization, adaptive histogram equalization and bi-histogram equalization with brightness preservation.

3. IMPLEMENTATION OF THE PROPOSED METHOD

We have proposed a method that consists in a combination of the different methods presented above. We describe each of the combinations.

3.1 Combination 1

Combination 1is performed as follows: mathematical morphology operations (opening, closing) followed by fuzzy mask filtering and then soft thresholding. We used this combination in order to eliminate the white and black lines from the image captures by the morphological operation. Then, the blur mask allows amplifying the high frequencies of the image, which correspond to the microcalcifications. Finally, the cropping operation allows the grey levels of the image to be better distributed when they are not well distributed in order to improve the contrast of the pre-processed image. Figure 4 shows the steps of combining 1.

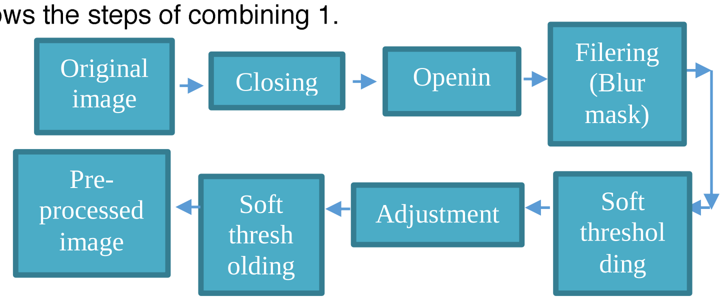


FIGURE 4: Steps of the combination 1.

3.2 Combination 2

Combination 2 starts with the crazy mask filtering followed by the mathematical morphology operations and the fitting. We pursue the same goal as in combination 1. Figure 5 presents the steps of combination 2.

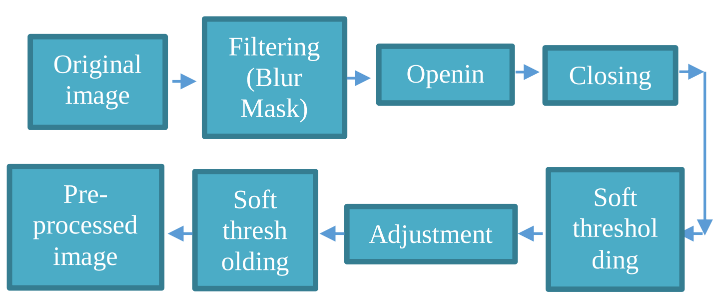


FIGURE 5: Steps of the combination 2.

3.3 Combination 3
Combination 3 starts with mathematical morphology operations (opening, closing) followed by fuzzy mask filtering and then adaptive histogram equalization. The aim is the same as in combination 1 and 2. We represent the steps of combination 3 in figure 6.

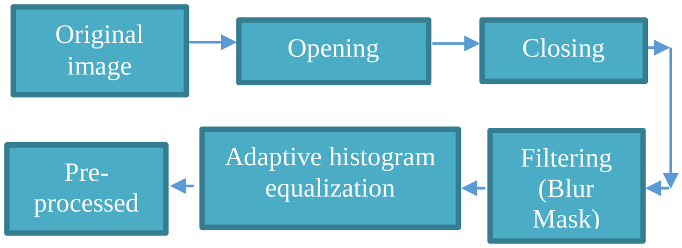


FIGURE 6: Steps of the combination 3.

Algorithms are developed and implemented for the above combinations and are tested using some samples of the MIAS database.

4. RESULTS OBTAINED
Four samples from the MIAS database are used to evaluate our proposed method. We identified and zoomed in on the suspect area for each image. This area is used to test the three combinations described above. Figures 7 to 10 show the results obtained by each of the three combinations**.**

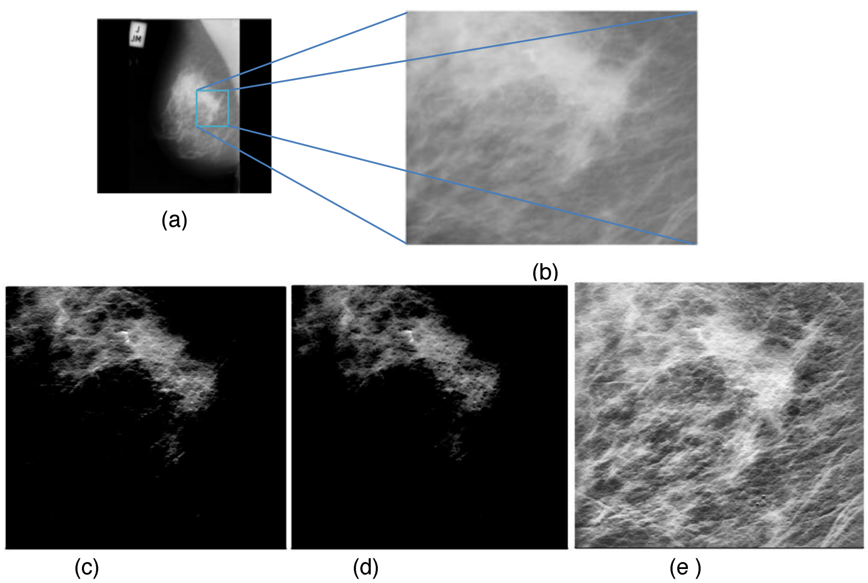


FIGURE 7: (a) original image, (b) zoomed region of interest, (c) image of combination 1, (d) image of combination 2 and (e) image of combination 3, (mdb209).

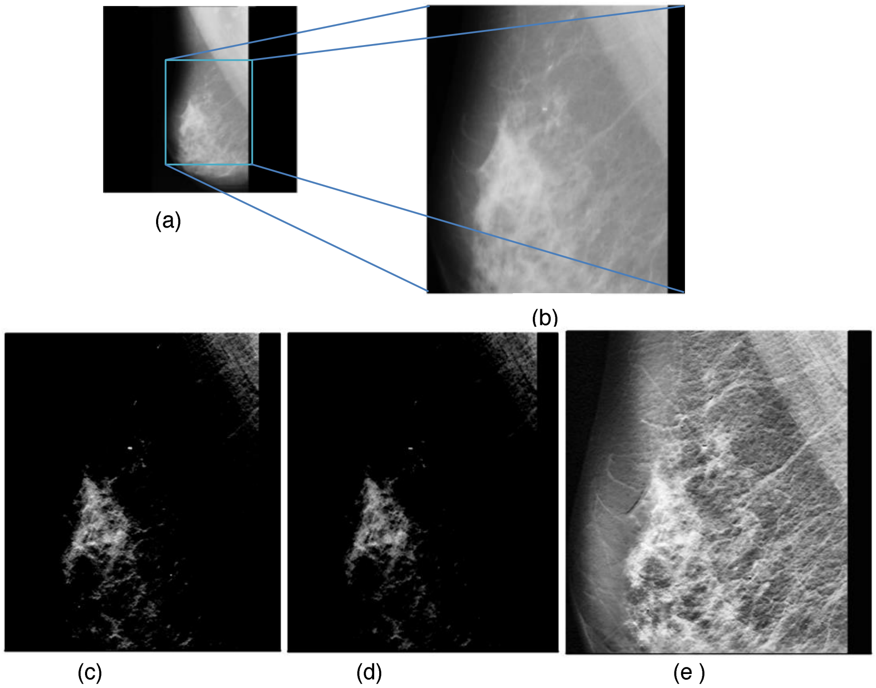


FIGURE 8: (a) original image, (b) zoomed region of interest, (c) image of combination 1, (d) image of combination 2 and (e) image of combination 3 (mdb249).

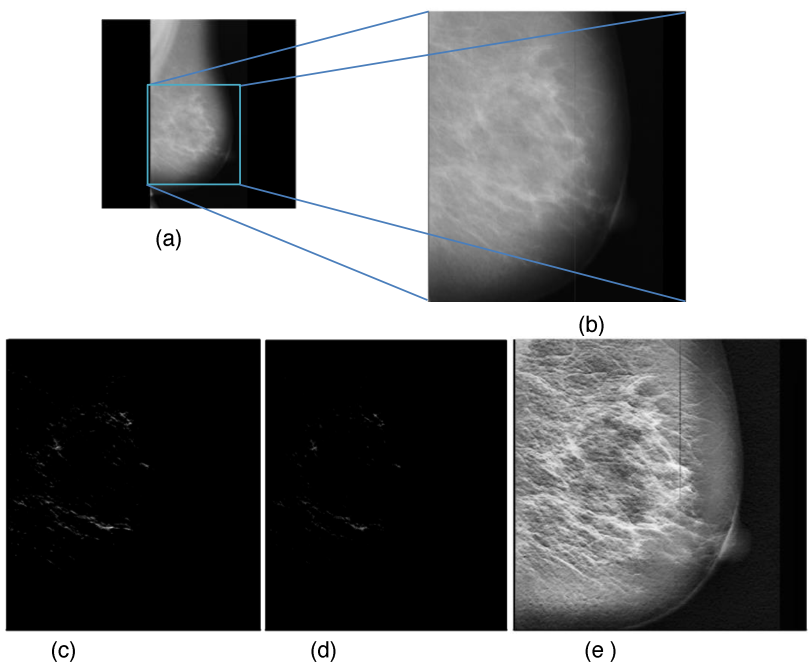
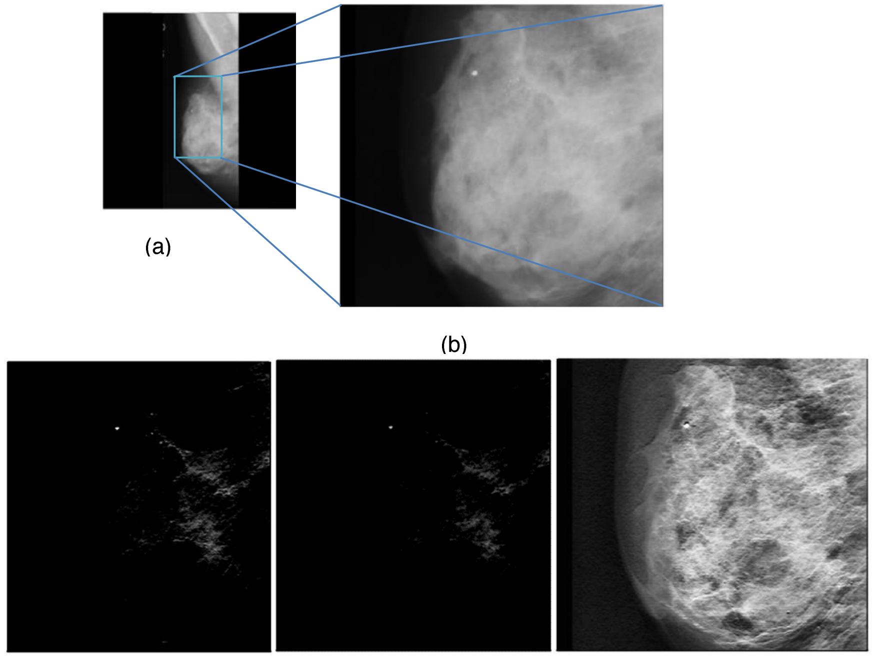


FIGURE 9: (a) original image, (b) zoomed region of interest, (c) image of combination 1, (d) image of combination 2 and (e) image of combination 3, (mdb223).

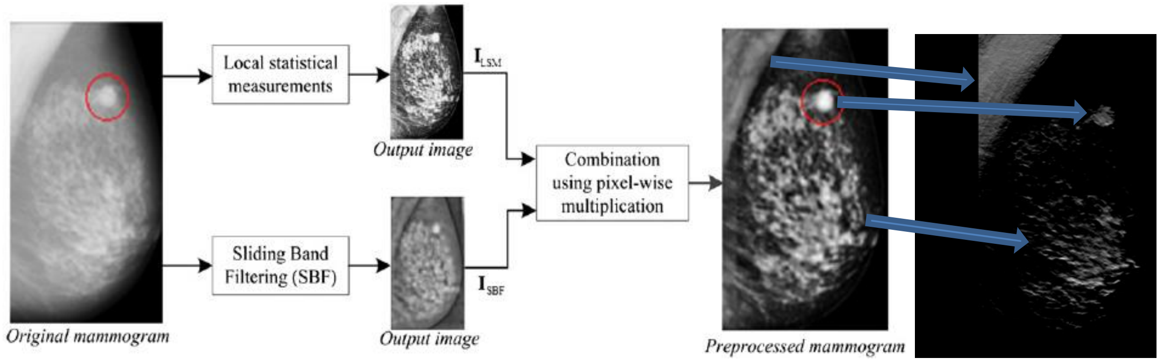


(a) (b) (c) (d) (e)

FIGURE 10: (a) original image, (b) zoomed region of interest, (c) image of combination 1, (d) image of combination 2 and (e) image of combination 3 (mdb238).

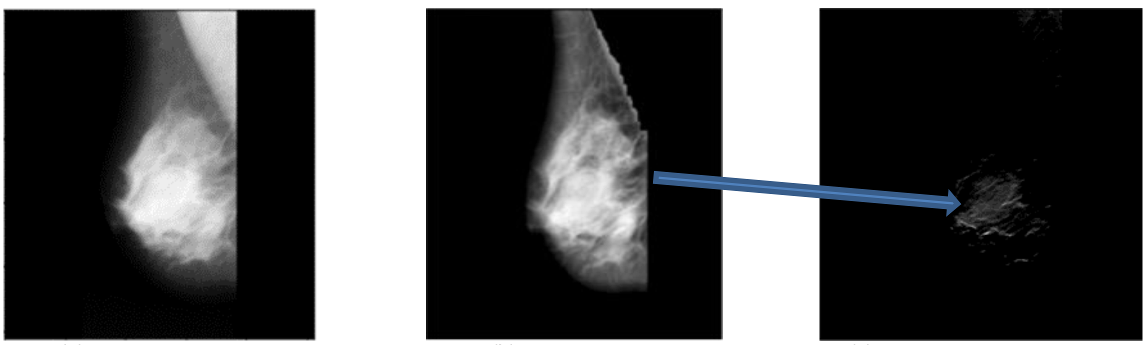
5. DISCUSSIONS

It results for the four shown images that the proposed combinations significantly improve the mammographic images. These results obtained make the background of the image and the objects (the part of the pectoral muscle, the suspicious area) stand out distinctly. Much of the adipose tissue, fat mass and some of the features seen in the zoomed-in part of the image are significantly reduced. For example, in figure 9 b where a vertical line, the breast with the nipple, fatty tissue etc. are seen, after the treatment has been performed, many of these observed parts disappear. Only the points that initially appear to be suspicious spots are visible (cf. figure 9 c and d). In figure 8 b, we have seen a part that represents the breast on which there are stains, bright spots, adipose tissue and fat mass and a part of the pectoral muscle. After application of our method, the adipose tissue and fat mass have disappeared while the bright spots and stain are enhanced and appear distinctly, as well as the pectoral muscle (cf. Figure 8 c and d). Similarly, looking at figure 10 b, we see the fat mass and a very bright spot. We applied the methods performed (especially combinations 1 and 2); so, the fat mass is clearly reduced and the points hidden in this mass appear (cf. figure 10 c and d). The same observations are made in figure 7. These results obtained, compared to those presented in (Yuliana et al, (2021), (Ilhame, (2017), (Jae, 2014)) and(Cai et al, 2019), are better. For example, Figures 11 and 12 below clearly illustrate the difference between our proposed method and those encountered in the literature. We can see that in Figure 11 (a) and (b) and Figure 12 (a) and (b), which are results from ((Jae, 2014), (Ilhame, (2017)) respectively, the fat mass and adipose tissue are not sufficiently reduced while the results obtained by our method (Figure 11 (c) and Figure 12 (c)) have sufficiently reduced the part representing the fat mass and adipose tissue. Thus, the results presented in Figure 11, (c) and Figure 12 (c) clearly show an area of interest that is easy to exploit while the other results (Figures 11 (a), (b) and 12 (a), (b)) show an area of interest with a lot of fat mass and the risk of having false positives seems to be high.



(a) (b) (c)

FIGURE 11: (a) original image, (b) preprocessed image (Jae, 2014),
(c) image preprocessed by our method.



(a) (b) (c)

FIGURE 12: (a) original image, (b) preprocessed image (Ilhame, (2017),
(c) image preprocessed by our method.

Although our proposed method significantly reduces the fat mass and enhances the suspicious points, some objects of no interest are still present in the obtained results.

6. CONLUSION AND PERSPECTIVES

We proposed a mammography image preprocessing method based on combinations of the classical image preprocessing techniques existing in the literature. Before realizing the combination, we tested more than five preprocessing techniques encountered in the literature in order to realize three different image preprocessing methods. The realized methods are used to preprocess four images taken from the MIAS mammography image base. The results obtained showed their efficiency and allowed to improve the quality of the images in order to use them for the detection of microcalcifications which are the warning signs of breast cancer. Compared to some literature methods, the results obtained by our methods are better. Even though some unnecessary objects have not been completely eliminated, their presence does not compromise the further exploitation of the results to design and implement a decision support system. This system will help health care personnel in the fight against breast cancer. In the future, we will develop the algorithm using wavelet transform and multi-scale products to eliminate the residual unwanted objects and obtain the microcalcifications, which will be used as decision features. Classification, determination of false positive acceptance and false negative rejection ratios for malignancy or benignity decision will also be considered.

7. REFERENCES

American Cancer Society, Breast Cancer Facts & Figures 2019–2020. Available online: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-andstatistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020. pdf (accessed on 30 March 2023).

ArdymulyaIswardani, WahyuHidayat, (2018)"Mammographic Image Enhancement using Digital Image Processing Technique", *International Journal of Computer Science and Information Security (IJCSIS)*, Vol. 16, No. 5.

Asma TOUIL, (2021)"Fusion collaborative de classifieurs : Application à la détection de pathologies en mammographie", Thèse, Institut Supérieur de l'informatique et des technologies de la communication de Hammam Sousse*, Unité de recherche : Latim (IMT Atlantique – France) et LATiS (Eniso, Sousse – Tunisie),* Thèse N° : 2021IMTA0263.

Breast Cancer Awareness Month 2021, WHO. Available online: https://www.iarc.who.int/featured-news/breast-cancerawareness-month-2021/ (accessed on 30 March 2023).

Cai H., Q. Huang, W. Rong, Y. Song, J. Li, J. Wang, J. Chen, and L. Li, (2019) "Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms", *Computational and mathematical methods in medicine*, vol. 2019.

IlhameAitLbachir, RachidaEs-salhi, ImaneDaoudi, SaadiaTallal, (2017) A New Mammogram Preprocessing Method for Computer-Aided Diagnosis Systems, *Conference Paper*,DOI: 10.1109/AICCSA.2017.40.

Iman Mohammed Burhan, Rahman NahiAbid, Mustafa AbdalkhudhurJasim, Refed Adnan Jaleel, (2022)"Improved Methods for Mammogram Breast Cancer Using by Denoising Filtering", article, Babylon University, *Iraq, Webology*, Volume 19, DOI: 10.14704/WEB/V19I1/WEB19099.

Jae Young Choi, Dae Hoe Kim, Konstantinos N Plataniotis and Yong Man Ro, "Computer-aided detection (CAD) of breast masses in mammography: combined detection and ensemble classification", *Phys. Med. Biol.* 59 (2014) 3697–3719, doi:10.1088/0031-9155/59/14/3697.

Jingsong Yang, Jie Lia , Han Wang, (2019), "A Seismic Signal Denoising Method Based on Wavelet Comprehensive Threshold", *issues ITM Web Conf.*, Volume 25 (2019),ITM Web Conf., 25 (2019) 01014.

KatarzynaKwiatkowska, Piotr Rhone, KatarzynaWrzeszcz, Barbara Ruszkowska-Ciastek, (2022) "High Post-Treatment Leptin Concentration as a Prognostic Biomarker of the High Risk of Luminal Breast Cancer Relapse: A Six-Year Comprehensive Study", *Life 2022*, 12, 2063. https://doi.org/ 10.3390/life12122063.

Khan, Muhammad Attique; Sharif, Muhammad; Akram, Tallha; Raza, Mudassar; Saba, Tanzila; Rehman, Amjad (2020), "Hand-crafted and deep convolutional neural network features fusion and selection strategy: An application to intelligent human action recognition".

Mudassar Raza, Muhammad Sharif, Mussarat Yasmin, Muhammad Attique Khan, Tanzila Saba, Steven Lawrence Fernandes, (2018) "Appearance based pedestrians' gender recognition by employing stacked auto encoders in deep learning. *Future Generation Computer Systems*", 88, 28-39.

Noor Ain SyazwaniMohd Ghani, Abdul KadirJumaat,Rozi Mahmud, MohdAzdiMaasar, FarizuwanaAkmaZulkifle, Aisyah Mat Jasin, (2023)"Breast Abnormality Boundary Extraction in Mammography Image Using Variational Level Set and Self-Organizing Map (SOM)", MDPI, *Mathematics* 2023, *11*(4), 976; https://doi.org/10.3390/math11040976.

RafsanjanyKushol, Md. HasanulKabir, MdSirajus Salekin & A. B. M. Ashikur Rahman, (2017) "Contrast Enhancement by Top-Hat and Bottom-Hat Transform with Optimal Structuring Element: Application to Retinal Vessel Segmentation", *International Conference Image Analysis and Recognition (ICIAR)*, pp 533–540.

SaharBayoumi, Sanaa Ghouzali, Souad Larabi-Marie-Sainte, Hanaa Kamel, (2021) "Diagnostic de stadification du cancer du sein à deux niveaux",*Artificial Intelligence and Internet of Things,* 1st edition, pages 20.

Saleem Z. Ramadan, (2020)"Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms":*A Review, Journal of Healthcare Engineering*, Volume 2020, Article ID 9162464, 21 pages.

Saraswathi, D.; Srinivasan, E.; Ranjitha, P. (2017) "An Efficient Level Set Mammographic Image Segmentation using Fuzzy C Means Clustering*", Asian J. Appl. Sci. Technol.* 2017, *1*, 7–11.

Sunil B., Nakerakanti Y., Puppala D. and Kethavath S., (2022) "Opening and closing in morphological image processing", *World Journal of Advanced Research and Reviews*, 14(03), 687–695.

Xiang Lin, Jian-Hua Li, Shi-Lin Wang, Alan-Wee-Chung Liew, Feng Cheng, Xiao-Sa Huang, (2018) "*Recent Advances in Passive Digital Image Security Forensics: A Brief Review"*, Volume 4, Issue 1, Pages 29-39.

Yosra Ben Salem, (2017) "Fusion d'images multimodales pour l'aide au diagnostic du cancer du sein", thèse, IMT Atlantique, *Université Bretagne Loire, École Doctorale Mathématiques et STIC et en cotutelle avec l'Ecole Nationale d'Ingénieurs de Tunis*, 148 pages.

YuanminXie, Lichuan Ning, Mingqi Wang, Chengcheng Li, (2019) "Image Enhancement Based on Histogram Equalization", *IOP Conf. Series: Journal of Physics: Conf. Series* 1314 (2019) 012161, doi:10.1088/1742-6596/1314/1/012161.

Yuliana Jiménez Gaona, M. J. Rodríguez-Álvarez, Jimmy Freire, Darwin Castillo, VasudevanLakshminarayanan, (2021) "Preprocessing fast filters and mass segmentation for mammography images," *Proc. SPIE 11842, Applications of Digital Image Processing XLIV*, Vol. 11842, 1184213, doi: 10.1117/12.2593939.