

Detection of malignant breast tumors on the background of fibrocystic breast disease

Yu A Podgornova¹ and S S Sadykov¹

¹Murom Institute (branch) Federal State Budgetary Educational Institution of Higher Education Vladimir State University named after Alexander Grigoryevich and Nickolay Grigoryevich Stoletovs, Orlovskaya Street 23, Murom, Russia, 602264

Abstract. The incidence of breast cancer is increasing every year. The mammography screening is the only method of diagnosing the disease. The mammograms analysis is complicated by the existence of various types of mastopathy - the most common breast disease caused by changes in the woman's endocrine profile. More than 40 per cent of the female population is exposed to this disease according to the WHO statistics. An algorithm for the detection of malignant neoplasms was developed by the authors in the context of fibrosis-cystic mastopathy. The experimental research on mammography images from the MIAS database showed that malignant tumors were available in 80 per cent of cases (21 mammograms), in other cases (5 mammograms) the tumor area was vaguely separated from the surrounding tissue.

1. Introduction

In compliance with the statistical data provided by the Russian oncologists, every 5 woman is affected by breast cancer (hereinafter BC) which is, on average, the cause of 16,000 fatal cases per annum [1, 2]. At the early stages breast cancer is virtually impossible to reveal as it is not generally accompanied by any pain. In addition, it is a high risk of missing the diseases signs at mammography screening. One more reason of such missing is the presence of fibrosis-cystic mastopathy (hereinafter referred to as FCM) in a breast. FCM is a group of heterogeneous diseases characterized by a wide range of proliferative and retrograde changes in the breast tissue with a violated ratio of epithelium and connective tissue components.

The radiographic appearance of this mastopathy form is characterized by a mass calcification of a glandular triangle and an abnormal structural pattern. Instead of a usual radiary streaking from the breast base towards a nipple one can observe a chaotic and rough trabecular structure.

Diffuse FCM forms are of interest for research not only from the point of view of a significant incidence of this pathology but mainly because they provide a background for breast cancer development. In the opinion of many authors, malignant changes at non-proliferative FCM forms is 0.9%, at moderately intense proliferation - 2%, at strong proliferation - 31% [3].

Therefore, various mastopathy forms detected in the image complicate the neoplasm diagnostics, especially at the early stages of cancer when the size of a tumor is not more than 1 cm. Applying contemporary information technologies to process mammograms, both digital and digitized, helps to improve the image quality, highlight and magnify the section of interest, automatically highlight the areas of malignant and benign neoplasms. Therefore, as of today there is a problem consisting in urgent necessity to enhance the quality and plausibility of breast cancer detection on the background

of mastopathy at analog mammography images and the absence of contemporary technology to implement these functions in various medical institutions. In view of the foregoing, improving the quality and plausibility of breast cancer detection with the help of contemporary ITs is a relevant applied scientific and technical problem.

The purpose of the paper is to develop an algorithm for breast cancer revealing on the background of FCM allowing for the visualization of malignant tumours to find their exact location and necessary texture characteristics.

2. Algorithm for the detection of cancer tumors on mammography images in the FCM background

Figure 1 shows the proposed algorithm to detect malignant tumors affected by FCM.

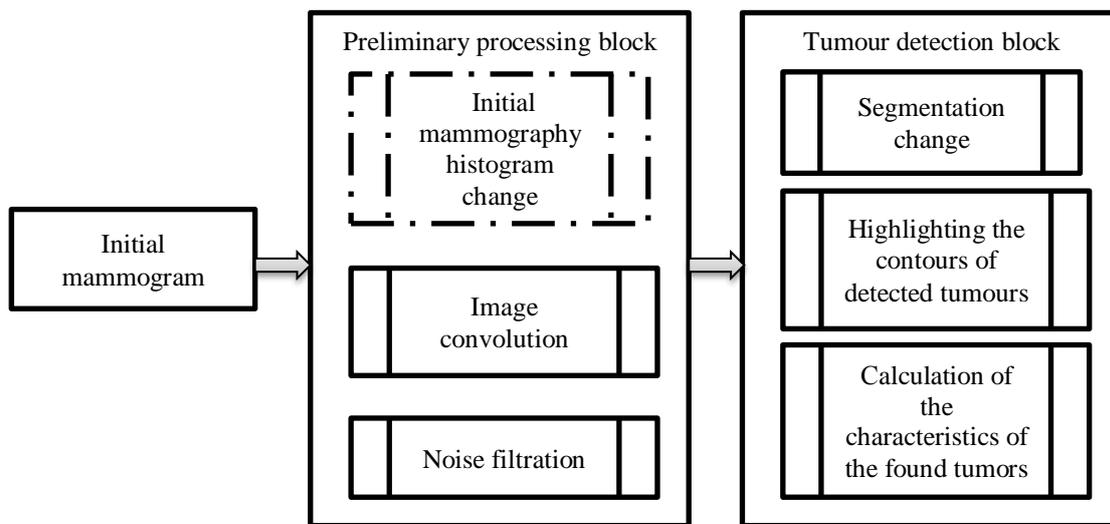


Figure 1. General block diagram for the algorithm of highlighting the breast cancer area on the FCM background.

The initial mammography image is a half-tone image $I(x,y)$ digitized by a high-precision scanner and having the dimension of $m \times n$ dots. The pixel brightness gradation is distributed within the range $[0..255]$.

Fibrous-cystic calcification areas present at mammography images significantly complicate diagnostics as it is impossible to visually detect the neoplasm lesions located under FCM. In this case a radiologist cannot state a plausible diagnosis to a patient. The example of the specified radiogram is a mammography image in Figure 3(б). The authors developed an algorithm for radiographic opacification permitting a specialist to remove a fibrous-cystic calcification layer from the image.

The method is based on histogram transformations [4, 5] of a radiographic image which can be represented as follows:

$$F(x, y) = R \cdot t[I(x, y)] \cdot (m \cdot n)^{-1} - t[0] + (F_{\max} - F_{\min}) \cdot (I(x, y) - I_{sr}) \cdot (I_{sr} - I_{\min})^{-1} \quad (1)$$

where $I(x,y)$ – current brightness value of the initial image, I_{\max} , I_{\min} , I_{sr} – maximum, minimum and average values of the initial mammography values, correspondingly, $F(x,y)$ – brightness value of the output image dot, F_{\max} , F_{\min} – required maximum and minimum brightness values of the output image dots, correspondingly, $R = 1.5$ – coefficient ($1 \leq R \leq 2$).

$$t[i] = \sum_{j=0}^i (v[j] \cdot (m \cdot n)^{-1}), \quad (2)$$

where i – brightness gradation value of the output image dot ($i=0..255$), j – brightness gradation value of the initial image dot ($j=0..255$, $j \leq i$), $v[j]$ – value of the initial image histogram element at the brightness j ($v[j] \in V$), T – area of the output image histogram element change $F(x,y)$, $t[i]$ – value of

the output image histogram element at the brightness i ($t[i] \in T$), $t[0] = t[I(x,y)=0]$ – value of the processed image histogram element at the point brightness equal to 0, $m \times n$ – image dimension.

If we consider a mammography image showing a breast in the form of a fat involution (Figure 3(a)), the image opacification is not necessary.

The next stage of a preliminary processing is the image mask filtration.

$$d(x, y) = \sum_{k,l} F(x-k, y-l) \cdot g(k, l), \quad (3)$$

where $d(x,y)$ – brightness value of the output image dot, k, l – mask dimensions ($k=0..2, l=0..2$), $F(x,y)$ – brightness value of the initial image dot, $g(k,l)$ – mask element.

Multiple research have shown the areas of malignant neoplasms can be highlighted with the help of a convolution of a mammography image using the mask given in Figure 2.

$$\begin{array}{ccc} 2 & -1 & 2 \\ 1 & -3 & 1 \\ 2 & 1 & 2 \end{array}$$

Figure 2. Mask for mammography convolution.

Such transformation of a mammography image causes some noise, that is why a median filtering is applied with the help of the mask 3×3 [6].

$$r(x, y) = M[d(x, y)] \quad (4)$$

where $r(x,y)$ – output image, M – median filtering operator.

The contour extraction block in the image locates a necessary area with clear contours.

Before extracting the contours in the image, the image texture map is subject to binarization. The binarization algorithm is presented as follows:

$$b(x, y) = \begin{cases} 1 & \text{если } r(x, y) \geq Q \\ 0 & \text{иначе} \end{cases}, \quad (5)$$

where $b(x,y)$ – output image, Q – threshold value equal to the maximum brightness value of the image dots $r(x,y)$,

$$Q = \max[r(x, y)] \quad (6)$$

Contour extraction is implemented with the help of the algorithm from [5].

$$k(x, y) = L[b(x, y)], \quad (7)$$

where $r(x,y)$ – output image, L – contour extraction operator.

The block of characteristics identification [7, 8] calculates the following:

- breast cancer area square S_k (number of dots in the area);
- area perimeter P – approximately equal to the number of the area contour dots;

- form coefficient: $K_f = \frac{P^2}{S_k}$

and other geometrical and texture characteristics of malignant neoplasms.

3. Experimental research of the proposed algorithm

The authors chose the mammography database MIAS for the experiment as all its images have a confirmed diagnosis as well as a clear description of each image with the specification of the tumour location. Figure 3 gives two initial mammography images.

In the radiographic image (Figure 3 (a)) in the breast upper quadrant projection one can identify a lesion with spicular contours, the skin over the them is thickened, the nipple is inverted. Conclusion: breast cancer.

Figure 3(b) shows that, on the background of breast tissue fat involution, one can identify a moderately expressed FCM with predominant fibrosis. In the upper quadrant one can suspect a lesion

presence. At the description of the screening mammography image there is a high possibility to miss the pathologic signs.

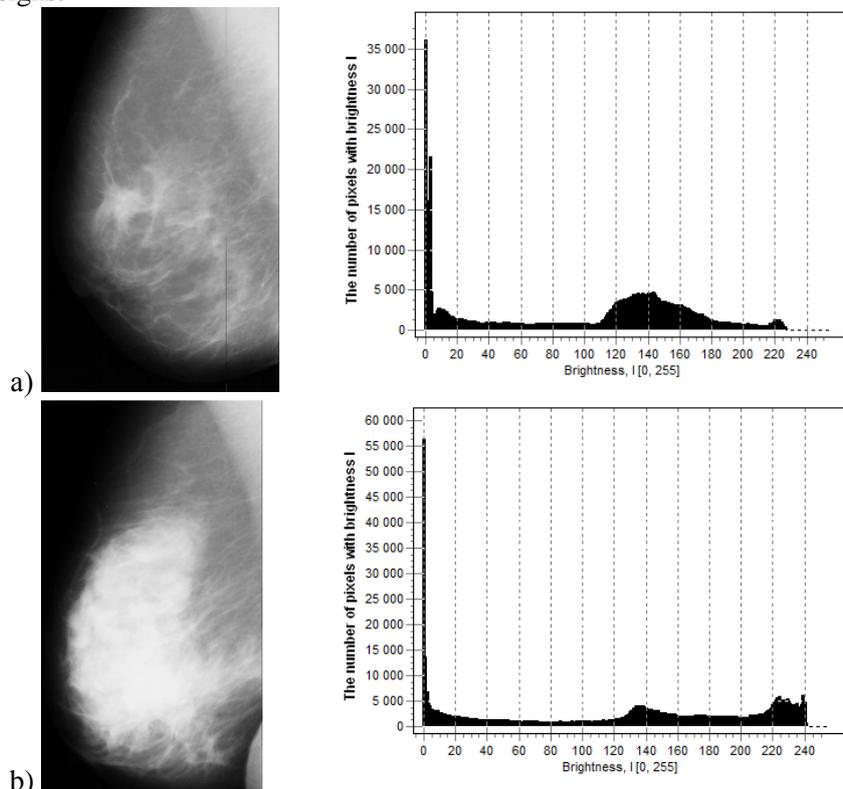


Figure 3. a, b – initial mammography images and their histograms, correspondingly.

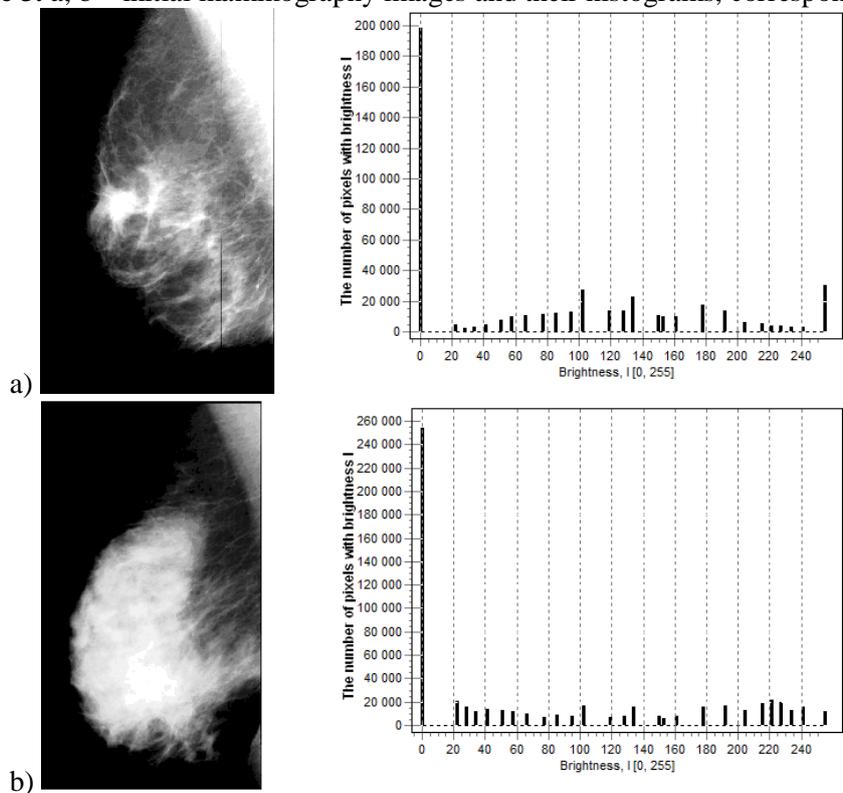


Figure 4. The results of the proposed histogram transformation: a) proposed algorithm of the histogram transformation for the Figure 3(a), b) proposed algorithm of the histogram transformation for the Figure 3(b).

See Figure 3: in the image processing using the proposed opacification method, the neoplasm described above is better visualized due to a larger difference of neoplasm and surrounding tissue densities.

At the opacification of Figure 3(b) one can see a high-density lesion with “blurred” contours, suspected malignant neoplasm. The next processing stage is the transformed image convolution with the help of the mask (Figure 3).

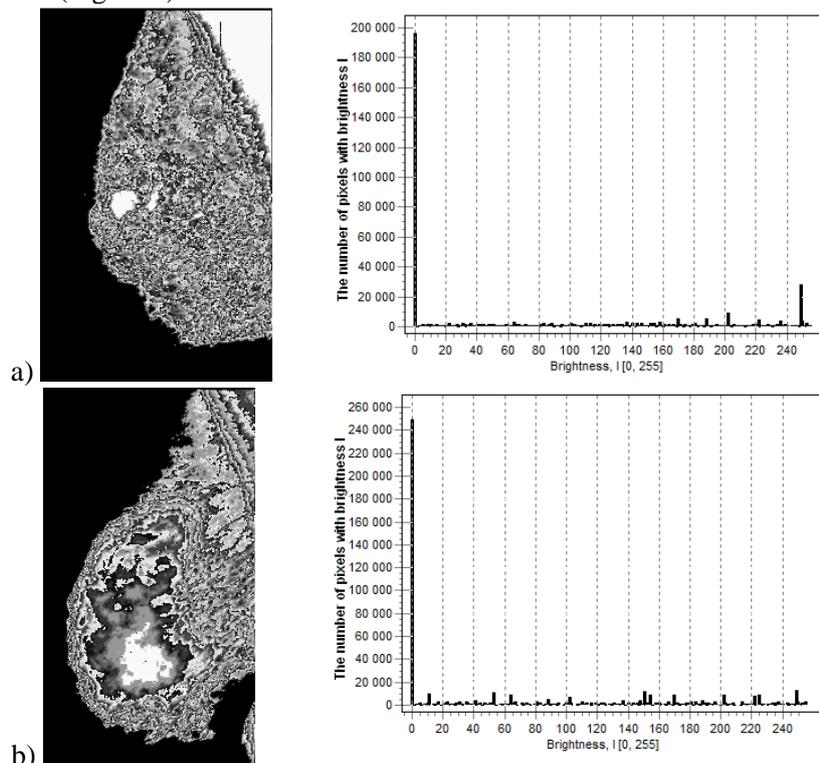


Figure 5. The image convolution result: a) image convolution in Figure 4(a), b) image convolution in Figure 4(b).

Applying the mask convolution method for the image in Figure 4(a), we obtain highlighted lesions on the breast tissues’ background, and small lesions are the tumor “seedings” poorly seen in the analog mammography image.

When processed with the help of the proposed method, this neoplasm (in Figure 5(b)) is highlighted on the background of surrounding tissues which points to a malignant neoplasm with high probability. Such characteristics as maximum and minimum brightness values, mean square deviation and entropy were calculated for all images and united into the Tables 1 and 2 for the purpose of comparison.

The values provided in the tables demonstrate the proposed algorithm of histogram transformations significantly improves the initial images as it increases the MSD value and decreases the entropy value of both initial images.

Table 1. Summary table of image characteristics for the image No. 1.

| Image | Minimum brightness | Maximum brightness | Average brightness value | Mean square deviation | Entropy |
|---------------|--------------------|--------------------|--------------------------|-----------------------|---------|
| Initial image | 0 | 240 | 97 | 68.60 | 7.11 |
| Hist. transf. | 0 | 255 | 80 | 84.5 | 3.51 |
| Convolution | 0 | 255 | 83 | 91.3 | 5.21 |
| Filter. | 0 | 249 | 83 | 82.7 | 4.91 |

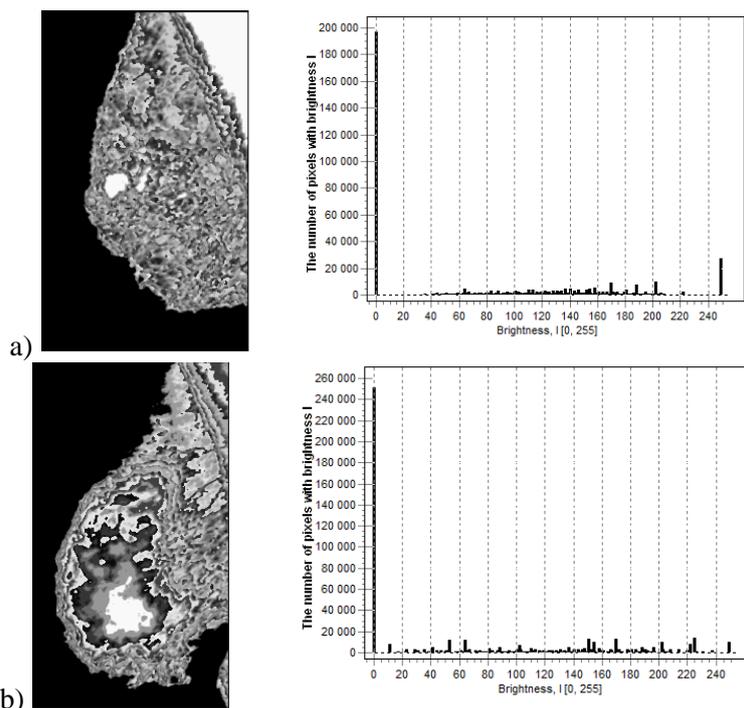


Figure 6. The image median filtration result: a) image median filtration in Figure 5(a), b) image median filtration in Figure 5(b).

Table 2. Summary table of image characteristics for the image No. 2.

| Image | Minimum brightness | Maximum brightness | Average brightness value | Mean square deviation | Entropy |
|---------------|--------------------|--------------------|--------------------------|-----------------------|---------|
| Initial image | 0 | 241 | 122 | 83.4 | 7.38 |
| Hist. transf. | 0 | 255 | 79 | 90.5 | 3.6 |
| Convolution | 0 | 255 | 77 | 87.3 | 5.05 |
| Filter | 0 | 249 | 77 | 81.4 | 4.82 |

Further, we apply binarization to the filtered image in Figure 6(a) and Figure 6(b) and extract the contours of neoplasm highlighted areas.

4. Conclusion

Summing up the aforesaid, one can draw the following conclusions on the research:

1. The proposed algorithm has been used to process 43 mammography images including: 26 images with previously detected and histologically confirmed cancer, 17 mammography images with detected changes assessed as benign neoplasms. At the mammogram processing the authors obtained:

- areas with malignant neoplasms are identified in 80% of cases (21 mammography images), in other cases (5 mammography images) the neoplasm area was vaguely separated from surrounding tissues;

- in 15 cases of 17 ones the lesion is not detected by the proposed algorithm. In 2 mammography images the lesion was highlighted, the interpretation of these results needs additional research conduct.

2. The proposed algorithm for the highlight of a breast cancer area in a mammography image allows detecting the breast cancer area even at the presence of a strong shade from FCM.

3. A set of algorithms for histogram filtering, median filtration, texture segmentation, binarization and contour highlight is presented in the form of the algorithm for highlighting the breast cancer area in mammography images with FCM and corresponds to all requirements of radiologists.

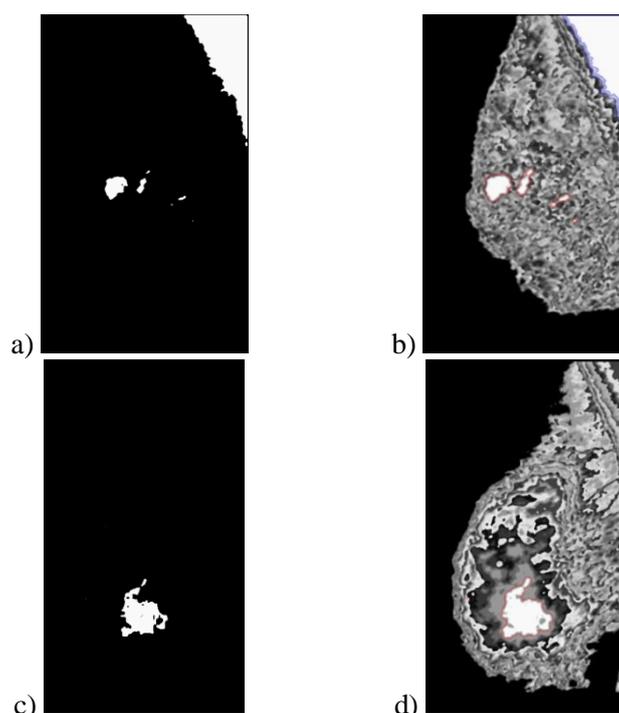


Figure 7. a, c – Image binarization in Figure 6(a) and (b) correspondingly, b, d - Contour application in the image in Figures 6(a), (b), correspondingly.

4. The further application of the proposed algorithm with breast cancer area highlighting on the FCM background provides for the opportunity to automatically detect breast cancer at early stages even at the presence of a severe FCM in a breast.

5. References

- [1] Sadykov S S, Bulanova Yu A and Zakharova Ye A 2014 Computer Diagnostics of Neoplasms in Mammography Images *Computer Optics* **38(1)** 131-138
- [2] Kaprin A D, Starinskiy V V and Petrovaya G V 2015 *Cancer Care Situation in Russia in 2014* (Moscow: Herzen Moscow Oncology Research Institute) p 236
- [3] Korzhenkova G P 2004 *Comprehensive Radiological and Sonographic Diagnostics of Breast Diseases* (Moscow: Firma “STROM”, LLC) p 128
- [4] Sankar P R and Srinivasa B K N 2007 Parallel Architecture for Implementation of Contrast Limited Adaptive Histogram Equalization *Advanced Engineering Sciences and Technologies* **10(1)** 047-051
- [5] Rajesh K R and Puran G 2012 Balvant Singh Underwater Image Segmentation using CLAHE Enhancement and Thresholding *Emerging Technology and Advanced Engineering* **2** 118-123
- [6] Pratt U 1982 *Digital Image Processing* (Moscow: Mir)
- [7] Gaidel A V and Krashennnikov V R 2016 Feature selection for diagnosing the osteoporosis by femoral neck x-ray images *Computer Optics* **40(6)** 939-946 DOI: 10.18287/2412-6179-2016-40-6-939-946
- [8] Gaidel A V 2016 Matched polynomial features for the analysis of grayscale biomedical images *Computer Optics* **40(2)** 232-239 DOI: 10.18287/2412-6179-2016-40-2-232-239

Acknowledgments

The research was conducted at the financial support of the RFBR in the framework of a scientific project No. 16-37-00227. Podgornova Yu.A. expresses her sincere gratitude to her academic advisor, Dr. of Tech. Sci., Professor Sadykov S.S. for his help at the research conduct, valuable recommendations in relation to their planning and article preparation as well as for his moral support.