Buletin SK 8(2)(2024): 1 - 10 ISSN 2250-1852

Artikel Asli/*Original Article*

K-means clustering and thresholding for breast phantom imaging using ImageJ

Penggunaan Kluster K-Means dan Ambang untuk Pemprosesan Imej Fantom Payudara dengan ImageJ

NUR AQILA MAZLAN, SITI HAJAR ZUBER, NORLAILI AHMAD KABIR

ABSTRACT

This work aims to analyse diagnostic images and improve the quality via post-processing technique. Breast calcification phantom images were utilised, with ImageJ as a tool for post-processing and a tool in analysing the images. K-means clustering was performed on breast phantom images and contrast to noise ratio (CNR) was calculated. Data was statistically analysed using SPSS Statistics 20. For all the images, there were negative correlation between noise and tube voltage (kVp) adjustment while positive correlation was observed on CNR when current-exposure time (mAs) was adjusted, with K-means clustering method shows significant improvement for the contrast of breast phantom images. The plot profile of the images after post-processing with thresholding method showed that there were more discrete edges of the breast calcification compared to the plot-profile of original images. It is imperative to note that it is easier to distinguish the calcification from the normal tissue of the breast phantom after post-processing technique took place, suggesting overall improvement post-processing.

Keywords: breast calcification phantom, image quality, K-means clustering, post-processing

ABSTRAK

Kajian ini bertujuan untuk menganalisis imej diagnostik dan menambah baik kualiti melalui teknik pasca-pemprosesan. Imej fantom kalsifikasi payudara telah digunakan, dengan ImageJ sebagai alat untuk pasca-pemprosesan dan juga sebagai alat untuk menganalisis imej tersebut. Kluster K-means telah dijalankan ke atas imej fantom payudara dan nisbah kontras kepada hingar (CNR) telah dikira. Data dianalisis secara statistik menggunakan SPSS Statistik 20. Bagi semua imej, terdapat korelasi negatif antara hingar dan pelarasan voltan tiub (kVp) manakala korelasi positif diperhatikan pada CNR apabila arus-masa dedahan (mAs) dilaraskan. Kaedah kluster K-means menunjukkan peningkatan yang ketara dalam kontras imej fantom payudara. Profil plot imej selepas pasca-pemprosesan dengan kaedah ambang menunjukkan terdapat lebih banyak kalsifikasi payudara tepi yang jelas berbanding dengan profil plot imej asal. Adalah penting untuk dinyatakan bahawa kalsifikasi lebih mudah dibezakan daripada tisu normal fantom payudara selepas teknik pasca-pemprosesan dijalankan, sekaligus mencadangkan peningkatan keseluruhan selepas pemprosesan.

Kata kunci: fantom kalsifikasi payudara, kualiti imej, kluster K-means, pasca-pemprosesan

INTRODUCTION

In the diagnosis of any abnormality or disease in patient, medical imaging (Dawar et al., 2025; Lawson et al., 2025) such as general X-rays (Irede et al., 2024), mammogram, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) (van der Geest et al., 2024), or ultrasound (US) (Okawa, 2025) are among the procedures a patient must go through. Therefore, analysis of the image is a great significance for better diagnosis and treatment. Medical image processing can be utilised to further improve the interpretability of the depicted contents on medical images (Abhisheka et al., 2024). This may involve enhancement of the image itself to improve the visualisation of certain features as well as the automated or manual extraction of information.

In image analysis, qualitative analysis is based on observer perception, where image is usually evaluated by experienced radiographer or physicians. It is vital to have a very good quality of medical images with a well-defined edge, good contrast, (Yadav et al., 2024) and resolution with low noise to avoid misinterpretation that will lead to wrong treatment solution. Quantitative analysis can be done by measuring and calculating the contrast, noise, contrast-noise ratio, signal-noise ratio, and many other calculation and measurement related to pixel intensity. Diagnostic value of low-quality images can be improved with some process of image enhancement. For example, in breast biopsy procedure, it is challenging to determine the edge between normal tissue and calcification on mammogram images. This is because mammogram produce low contract images due to small difference in density between various breast tissues. Hence, image post processing may be used to enhance the edge of calcification or lesion. Accurate analysis of calcification can improve the diagnosis and prognosis (Tong et al., 2024).

In this work, thresholding and k-means clustering and thresholding is used to analyse the breast phantom images using ImageJ. By combining these methods, this study aimed to improve the visibility of critical features in the breast phantom images, aiding in better assessment and interpretation.

METHODS AND MATERIALS

In this work, breast phantom images were obtained from previous research conducted at Universiti Sains Malaysia. All acquired images were converted to 8-bit format. Four types of breast phantoms were used for image acquisition, each fabricated with two layers of agar of different densities to simulate adipose and fibro-glandular tissues. All phantoms contained calcifications embedded at varying depths and distribution patterns. Source-image-distance was fixed to 100 cm and the size of cassette used was 15 cm × 30 cm. There were 7 images acquired with different tube voltage (kVp) and current-exposure time (mAs). Potential difference at 40 and 50 kVp were used as breast mammogram requires low kVp. Low kVps were preferred because breast tissues which consist of adipose tissue and fibroglandular tissue have similar density value.

The mimicking properties for fibroglandular tissue are with the density of 1.02 g/cm3, mass attenuation coefficient of 0.279 cm2g-1 and effective atomic number of 7.2774, while the mimicking properties for adipose tissue are with the density of 0.95 g/cm3, mass attenuation coefficient of 0.277 cm2g-1 and effective atomic number of 7.2452. ImageJ version 1.51j8 was used for post-processing. The region of interest (ROI) was selected to analyse pixel intensity, from which the mean and standard deviation were measured. The histogram and pixel data for the entire image were then extracted using ImageJ software.

K-MEANS CLUSTERING

K-means clustering was performed for the segmentation of different types of tissues in breast. All pixels were assigned into clusters based on the closest its pixel intensity to the random mean values. Three new means were calculated based on the clustered pixels in the newly formed image. Then, all pixels were reassigned into three new clusters based on the closest its pixel intensity to the second mean values. The steps will be repeated until the mean values no longer change.

THRESHOLDING

Thresholding is an image segmentation technique where pixels with intensities above the upper threshold are converted to black, while those below the lower threshold are converted to white. The images were processed and displayed using ImageJ software, where thresholding was applied by setting specific upper and lower threshold values. The lower threshold was fixed at 10 for all breast phantom images. The upper threshold was set to 223 for images in rows 1, 2, 3, and 6 as shown in Table 1, while images in rows 4, 5, and 7 had upper thresholds of 213, 233, and 193, respectively. SPSS Statistics 20 was used to analyse the acquired data.

RESULTS AND DISCUSSION

The graphs were plotted in Figure 1 and Figure 2 to display the relation of exposure factor with the image quality, which was evaluated by the noise and CNR. Based on the graph in Figure 1, it is observed that at a constant mAs, an increase in kVp results in a decrease in noise. This trend is also reflected in the Pearson Correlation analysis, which examines the relationship between kVp and noise. In this work, at a constant kVp, an increase in mAs leads to a higher CNR. Pearson Correlation analysis using SPSS revealed a positive correlation. These findings align with previous research, which demonstrated that increasing technical factors enhances the CNR value in the image (Pauwels et al., 2014).

K-MEANS CLUSTERING

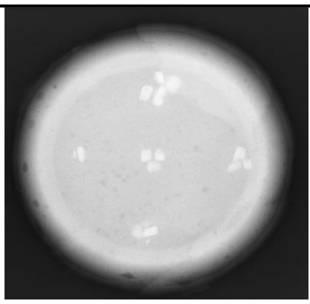
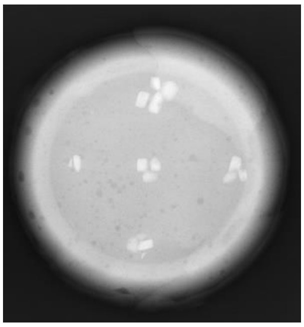
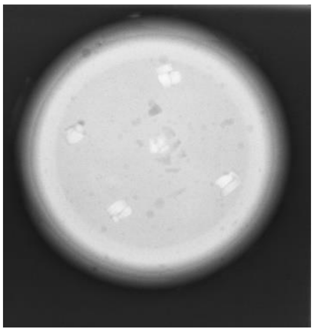
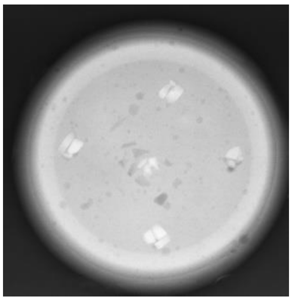
K-means clustering was used as a post-processing technique to enhance contrast in breast phantom images. The effectiveness of this method was evaluated by calculating the CNR of the images before and after applying the post-processing technique. Figure 3 illustrates the images before and after post-processing using k-means clustering. Based on Table 2 though Table 4, the highest contrast after k-means clustering can be seen in images acquired with 40 kVp and 30 mAs while the lowest contrast is found in image acquired with 40 kVp and 20 mAs.

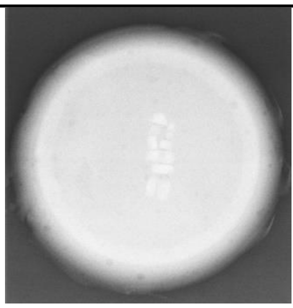
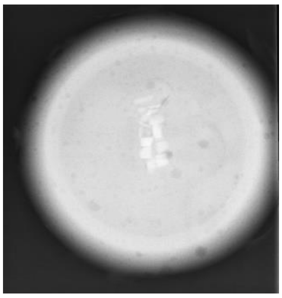
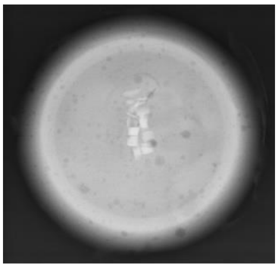
Pearson Correlation showed a negative correlation, indicating that CNR decreased as kVp increased. However, the decrease was not significant, as the p-value exceeded 0.05.

THRESHOLDING

Thresholding is the post-processing technique used on breast phantom images for edge enhancement. This was analysed qualitatively and quantitatively by histogram analysis. Figure 4 illustrates the breast image before and after thresholding.

TABLE 1 List of different acquisition parameters and images of breast phantom

No.	Images	kVp	mAs
1		40	20
2		40	30
3		40	20
4		40	30

5		40	30
6		40	20
7		50	20

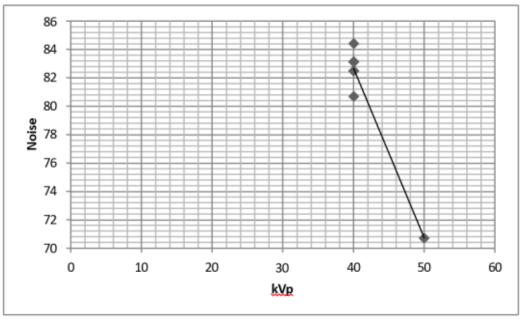


FIGURE 1 Graph showing the variation of image noise with increasing tube voltage (kVp) at a constant mAs for a breast phantom. The data demonstrate an inverse relationship, where image noise decreases as kVp increases

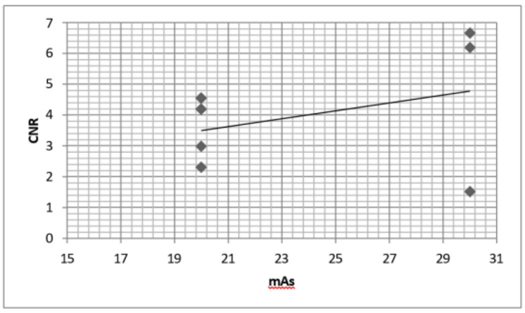
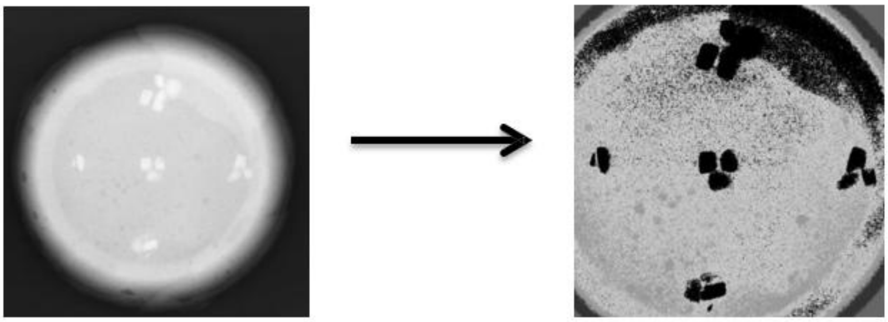


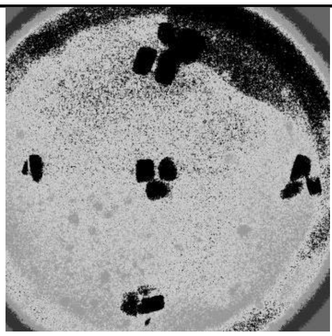
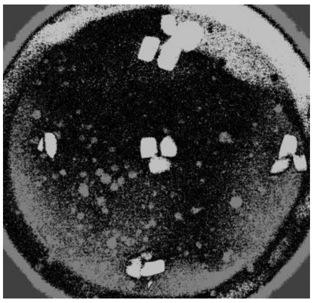
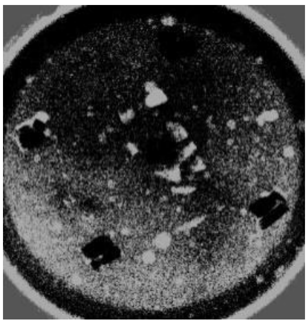
FIGURE 2 Graph showing the relationship between contrast-to-noise ratio (CNR) and varying mAs values at a constant kVp setting for a breast phantom. The data demonstrate a general trend of increasing CNR with increasing mAs



A B

FIGURE 3 Example of image before (A) and after post-processing (B) with k-means clustering

TABLE 2 Breast phantom images after k-means clustering

No	Images	kVp	mAs
1		40	20
2		40	30
3		40	20

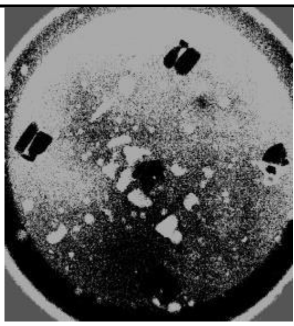
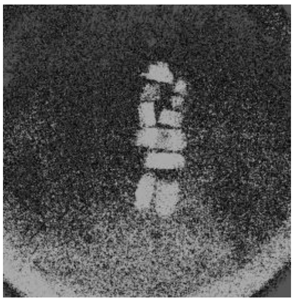
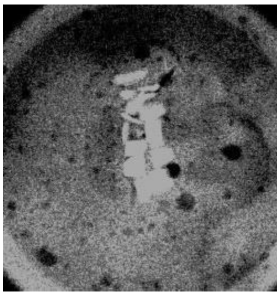
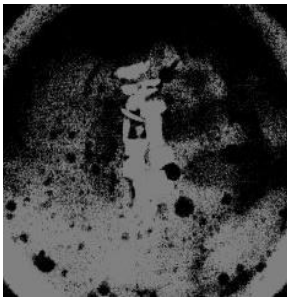
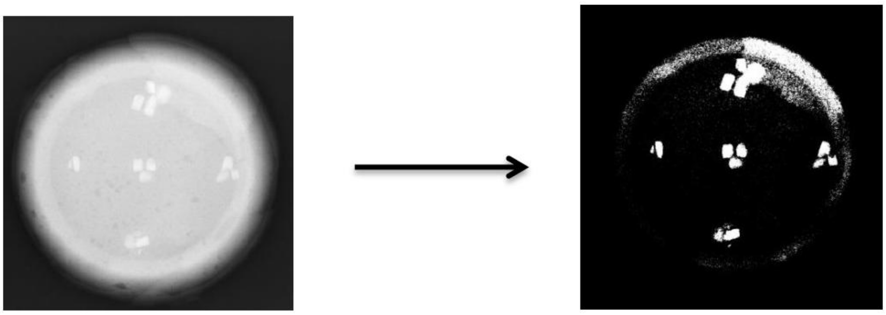
4		40	30
5		40	30
6		40	20
7		50	20

TABLE 3. Contrast and CNR measurement of the original images

No	kVp	mAs	Contrast	Contrast to noise ratio
1	40	20	0.344	4.547
2	40	30	0.340	6.177
3	40	20	0.335	4.185
4	40	30	0.406	6.649
5	40	30	0.163	1.506
6	40	20	0.226	2.983
7	50	20	0.178	4.969

TABLE 4. Contrast and CNR measurement of image after k-means clustering

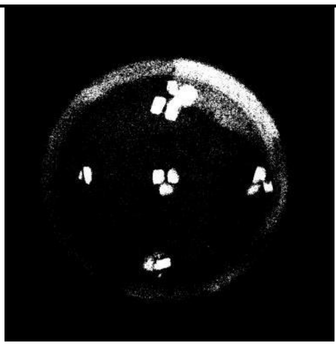
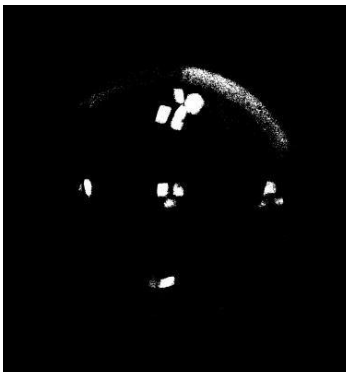
No	**kVp**	**mAs**	**Contrast**	**Contrast to noise ratio**
1	40	20	0.902	2.363
2	40	30	0.890	4.222
3	40	20	0.686	0.827
4	40	30	0.878	2.034
5	40	30	0.457	1.258
6	40	20	0.302	1.761
7	50	20	0.624	1.766

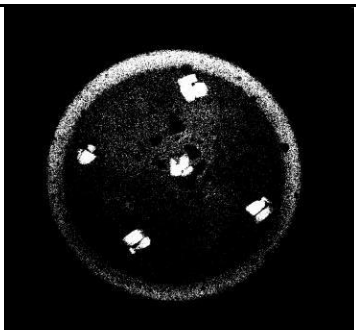
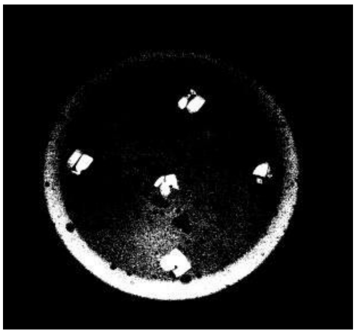
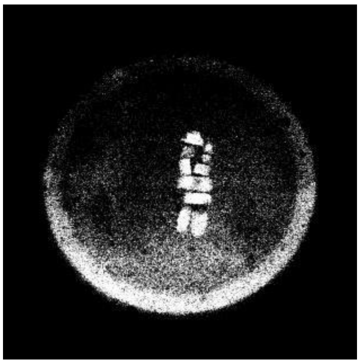
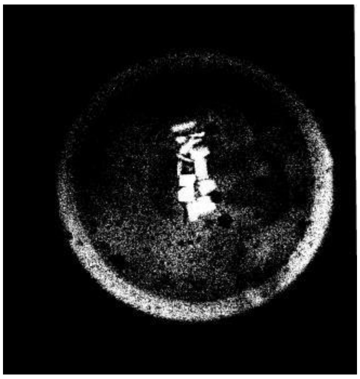


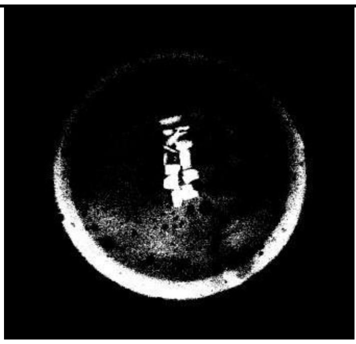
A B

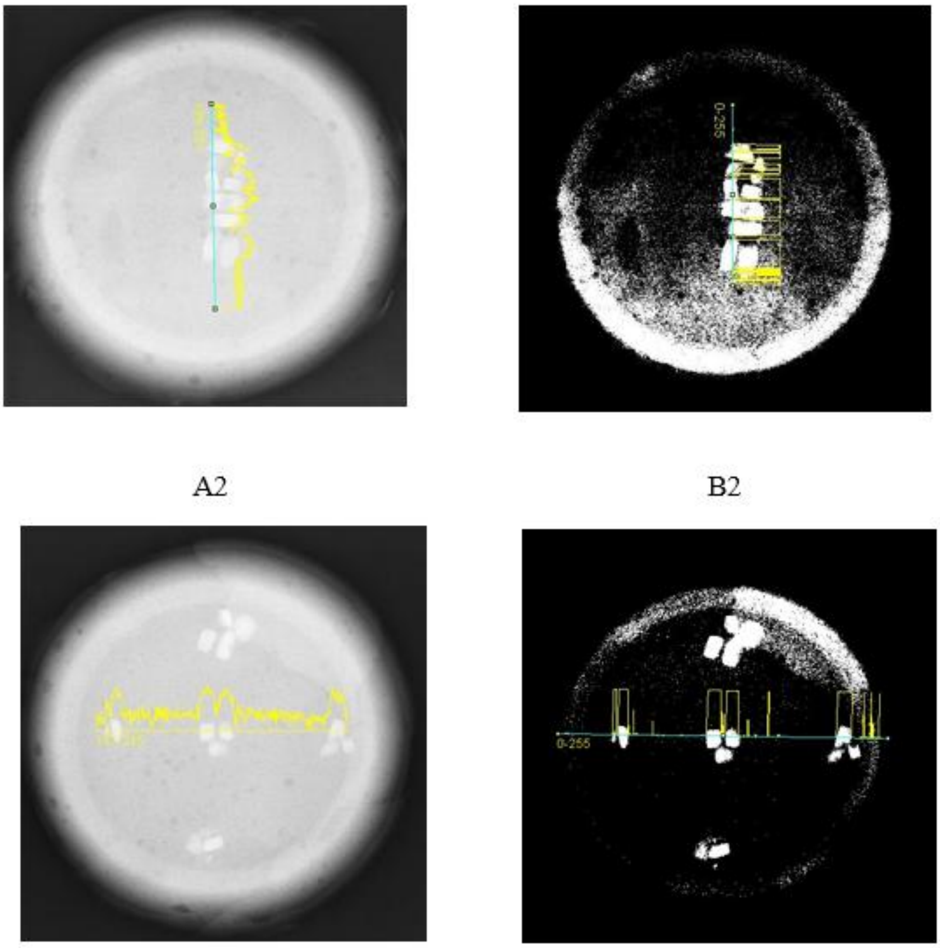
FIGURE 4 Example of breast phantom image before (A) and after (B) thresholding

TABLE 5 phantom images after thresholding

No	**Images**	**kVp**	**mAs**
1		40	20
2		40	30

3		40	20
4		40	30
5		40	30
6		40	20

7		50	20



A2 B2

A1 B1

FIGURE 5. Line profile across the image before (A1, A2) and after thresholding (B1, B2)

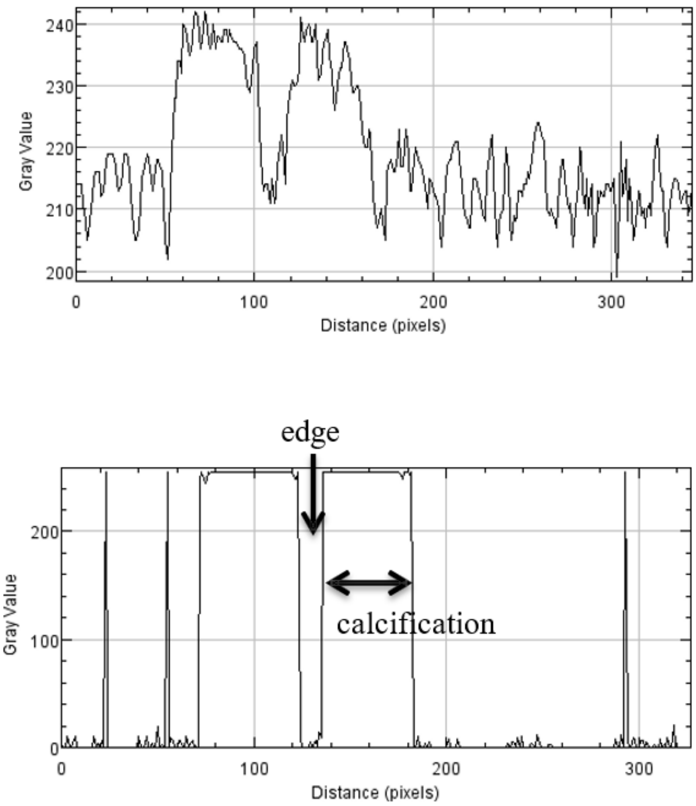


Figure 6. Plot-profile of the original breast phantom image and after thresholding

The images in Table 5 show improved differentiation between calcification and normal breast tissues, indicating that the image has undergone segmentation. The calcification edges are now more clearly defined, which may assist radiologists in marking areas for surgery or biopsy. The primary objective of edge detection is to extract important features from an image. To objectively analyse the image, a line profile was drawn across the region of interest, as illustrated in Figure 5. The line profile in Figure 6 reveals significant differences in the results obtained using the thresholding approach. The previously uniform pixel intensity distribution becomes discrete along the calcification border, indicating effective image segmentation. The calcification boundaries can be clearly identified in all images processed with the thresholding method. By analysing the line profile, the image was evaluated based on pixel distribution or gray value. A distinct difference in gray value between the foreground and background should be observed at the material's edge, where the material extent is marked by a sharp increase followed by a sudden decrease in gray value.

In research done by Cho, Yoon, and Yoon (2016), the edge response analysis is used in evaluating the sharpness of the boundary between the dark and bright areas in the acquired image. The image obtained by taking a picture of the actual object with a distinct boundary has blurred edge due to the blur phenomenon. The point spread function (PSF), was used to indicate the quality of the imaging system. In a PSF function, the sharpness was evaluated based on the FWHM size that corresponds to the width at half of the maximum amplitude, the better the sharpness, the smaller the full-width half maximum (FWHM) (Cho et al., 2016).

This evaluation of image sharpness is clinically significant, especially in the context of breast cancer imaging. Accurate delineation of tissue boundaries, enabled by a sharper edge response can improve lesion detectability and characterisation. Enhanced image sharpness facilitates the identification of subtle distortions or microcalcifications, which are critical for early detection of breast cancer. Furthermore, clearer imaging can guide more precise targeting during image-guided biopsies, thereby increasing diagnostic accuracy and reducing the risk of sampling errors.

CONCLUSION

In conclusion, the image quality of breast phantom images improved after post-processing. An increase in kVp reduced noise, while a higher mAs led to an increase in CNR. K-means clustering significantly enhanced the CNR of the images. Additionally, the edges of breast calcifications became more distinct following the thresholding approach. Qualitative observations indicate that post-processing made it easier to differentiate calcifications from normal breast tissue in the phantom images. These methods could also benefit other imaging interpretation, for example in lung disease detection from CT scan or brain tumor localisation from MRI.

ACKNOWLEDGEMENT

We thank School of Physics, Universiti Sains Malaysia for allowing this work to be conducted.

REFERENCES

Abhisheka, B., Biswas, S.K., Purkayastha, B., Das, D., Escargueil, A., 2024. Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. Multimed Tools Appl 83, 43035–43070.

Cho, H.-W., Yoon, H.-J., Yoon, J.-C., 2016. Analysis of Crack Image Recognition Characteristics in Concrete Structures Depending on the Illumination and Image Acquisition Distance through Outdoor Experiments. Sensors. https://doi.org/10.3390/s16101646

Dawar, Ria, Grimm, L.J., Sonnenblick, E.B., Dontchos, B.N., Coffey, K., Goudreau, S., Reig, B., Jacobs, S.A., Shah, Z., Mullen, L., Dialani, V., Dawar, Reema, Sayre, J., Dodelzon, K., Parikh, J.R., Milch, H.S., 2025. Mammography Home Workstations and Remote Diagnostic Breast Imaging: Current Practice Patterns and Planned Future Directions. J Breast Imaging wbae087. https://doi.org/10.1093/jbi/wbae087

Irede, E.L., Aworinde, O.R., Lekan, O.K., Amienghemhen, O.D., Okonkwo, T.P., Onivefu, A.P., Ifijen, I.H., 2024. Medical imaging: a critical review on X-ray imaging for the detection of infection. Biomedical Materials & Devices 1–45.

Lawson, M.B., Zhu, W., Miglioretti, D.L., Onega, T., Henderson, L.M., Rauscher, G.H., Kerlikowske, K., Sprague, B.L., Bowles, E.J.A., O'Meara, E.S., Tosteson, A.N.A., diFlorio-Alexander, R.M., Hubbard, R.A., Lee, J.M., Lee, C.I., 2025. Disparities in Standard-of-Care, Advanced, and Same-Day Diagnostic Services among Patients with Abnormal Screening Mammography. Radiology 314, e241673. https://doi.org/10.1148/radiol.241673

Okawa, Y., 2025. Ultrasound Imaging Equipment for Evaluating Chronic Constipation in Home Healthcare: A Review Article. Gastrointestinal Disorders. https://doi.org/10.3390/gidisord7010012

Pauwels, R., Silkosessak, O., Jacobs, R., Bogaerts, R., Bosmans, H., Panmekiate, S., 2014. A pragmatic approach to determine the optimal kVp in cone beam CT: balancing contrast-to-noise ratio and radiation dose. Dentomaxillofac Radiol 43, 20140059. https://doi.org/10.1259/dmfr.20140059

Tong, Y., Jiang, J., Chen, F., Guo, G., Zhang, C., Deng, T., 2024. A novel approach for segmentation and quantitative analysis of breast calcification in mammograms. Front Oncol 14. https://doi.org/10.3389/fonc.2024.1281885

van der Geest, K.S.M., Gheysens, O., Gormsen, L.C., Glaudemans, A.W.J.M., Tsoumpas, C., Brouwer, E., Nienhuis, P.H., van Praagh, G.D., Slart, R.H.J.A., 2024. Advances in PET Imaging of Large Vessel Vasculitis: An Update and Future Trends. Semin Nucl Med 54, 753–760. https://doi.org/https://doi.org/10.1053/j.semnucmed.2024.03.001

Yadav, P.S., Gupta, B., Lamba, S.S., 2024. A new approach of contrast enhancement for Medical Images based on entropy curve. Biomed Signal Process Control 88, 105625. https://doi.org/https://doi.org/10.1016/j.bspc.2023.105625

Nur Aqila Mazlan[1, 3]
Siti Hajar Zuber[2]
Norlaili Ahmad Kabir[3]

[1]Oncology and Radiotherapy Unit, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 13200,
Kepala Batas,
Malaysia.
[2]Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300,
Kuala Lumpur,
Malaysia.
[3]School of Physics, Universiti Sains Malaysia, 11800,
Penang,
Malaysia.

Corresponding Author: Norlaili Ahmad Kabir
E-mel: norlailikabir@usm.my
Tel: +604-653 5317
Fax: +604-657 9150

Received: 26 February 2025
Revised: 30 May 2025
Accepted for Publication: 6 June 2025