LAPLACIAN-BASED BLUR DETECTION ALGORITHM FOR DIGITAL BREAST TOMOSYNTHESIS IMAGES IN IMPROVING BREAST CANCER DETECTION

Harron NA[1], Sulaiman SN[1,2], Osman MK[1], A. Karim NK[3,4], and Isa IS[1]

[1]Centre for Electrical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500 Penang, Malaysia

[2]Integrative Pharmacogenomics Institute (iPROMISE), UiTM Puncak Alam Campus, Bandar Puncak Alam, 42300 Puncak Alam, Selangor, Malaysia

[3]Department of Biomedical Imaging, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Bertam, 13200 Kepala Batas, Pulau Pinang, Malaysia

[4]Breast Cancer Translational Research Programme (BCTRP), Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, 13200 Kepala Batas, Pulau Pinang, Malaysia

Correspondence:
Noor Khairiah A. Karim
Department of Biomedical Imaging, Advanced Medical and Dental Institute,
Universiti Sains Malaysia, Bertam,
13200 Kepala Batas, Pulau Pinang, Malaysia
Email: drkhairiah@usm.my

Abstract

The most challenging aspect of working with digital images captured in an uncontrolled environment is to determine whether the image is of sufficient quality to be studied further. One of the most frequent reasons for a decrease in the quality of digital images is the presence of blur artefacts, especially when the images are taken from various angles of the x-ray source with limited angular range such as in digital breast tomosynthesis (DBT). The unwanted artefacts might substantially obscure the breast cancer location, especially in extremely dense fibroglandular breast tissue. It is almost impossible to differentiate breast cancer lesions in blurry and low-contrast DBT images, thereby reducing the accuracy of lesion diagnosis. Due to this blurry artefact issue, this study aims to assess the performance of Laplacian-based Blur Detection (LbBD) algorithm for the blurry detection of DBT images. The LbBD algorithm is designed and developed using MATLAB R2021a software. The algorithm identifies the amount of blurriness by calculating the image variance; the farther the variance is from the threshold value, the less blurry is the image. On the contrary, the lower the variance level is from the threshold value, the greater the blur level. An online survey was conducted with an expert to assess the quality of 20 DBT images using two subjective measurements of blur or non-blur to ascertain the algorithm's performance. Three threshold values were used to compare the outcome with the output of the algorithm's status. With a low error rate (0.05), an accuracy of 95% at the ideal threshold value of 200, and an image size reduction of 10%, the system successfully predicts blurry images.

Keywords: Blur Detection, Digital Breast Tomosynthesis, Laplacian-based Algorithm

Introduction

Automatic quality analysis of medical images has emerged as a significant research area. However, researchers continue to pay little attention to automated blur detection in digital breast tomosynthesis (DBT) images despite the quality of the obtained images having an impact on image interpretation and identification of breast lesions. DBT reassembles a sequence of slices with limited angular range, resulting in reconstructed images that are both blurry and crisp. DBT images are traditionally scanned slice by slice, requiring significant effort from radiologists. To identify cancer locations on DBT slices, well-experienced radiologists will review each slice from the DBT z-Stacks, which takes a long time given the number of slices per DBT scanning (1). Radiologists, however, will ignore the blurred slice and instead focus on the non-blurred slices. Consequently, blur detection in DBT, also known as pre-processing to a good Computer-Aided Detection (CAD) System, should be the first step before restoration.

Most of the scanned images are made up of two regions: blur and sharp. Blur is further divided into; defocus blur, also known as out-of-focus blur, and motion blur (2). The quality of an image can be assessed by using different blur detection algorithms, while image restoration can be applied to a blurry image. According to (3), public domain research on blur images is divided into three primary phases: detection, classification, and image restoration. Researchers have proposed numerous blur detection algorithms, which are used in the pre-processing stage of image restoration.

Apart from the previous study, (4) has also studied 36 different techniques or focus measure operators for computing the blurriness metric of an image. Some are straightforward and use only grayscale pixel intensity statistics, while others are more advanced with feature-based, so the image's local binary patterns can be evaluated. Furthermore, (5) discussed three distinct strategies for blur detection, including the Fast Fourier transform, the Laplacian operator, and the Haar Wavelet transform. In rapid Fourier transform, the algorithm analyses the frequencies in the image at various locations and determines whether the image is blurry based on frequency level. When there is a low quantity of high frequency, the image is declared blurry, and the programmer determines whether there will be a high or low amount of high frequency (4). In the Haar wavelet transform, the images are divided into NxN by iterating over each tile in the two-dimensional Haar wavelet transform and tiles are grouped diagonally, vertically, and horizontally into clusters of tiles with noticeable changes. Then, small clusters of images are designated to be blurred (6).

One of the primary distinctions between Laplacian and Sobel or Kirsch operators is that Laplacian operators remove inward and outward edges, whereas other operators remove edges in a defined direction. The fundamental difference between Sobel, Kirsch, and the Laplacian operator is that Sobel and Kirsch derivatives are first order, resulting in thicker edges and increased sensitivity to noise. In contrast, the Laplacian operator is a second-order derivative mask that results in an increasing sharpening effect on images. In this study, the Laplacian operator is examined and evaluated as a tool for DBT image edge detection.

Laplacian attempts to de-emphasize portions in an image by gradually varying grey levels and emphasizing grey-level discontinuities (7). This derivative operator produces images with grey edge lines and some discontinuities on a black background, creating an image's exterior and interior edges. The Laplacian technique's variance returns a single floating-point value representing the image's "blurriness" (5). To date, the radiologist uses no standard parameters to quantify the degree of blurriness in clinical diagnosis. It is a common practice to rely on the radiologist's subjective evaluation of the image quality, which is based on human judgment. Therefore, this study contends that a quantitative standard measurement of the level of blurriness is indeed required.

Methods

This study entails the creation of a dataset of blurry and non-blurry DBT images. The algorithm was programmed and tested in MATLAB R2021a, using a notebook with AMD Ryzen 5 3500U CPU @ 2.10GHz processor and 8.0 GB RAM with Windows 10, 64bit operating system.

Dataset

This study intends to assess the quality of DBT images obtained from DBT devices regarding their level of blurriness. Samples of DBT images were selected randomly from the database provided by the Advanced Medical and Dental Institute (AMDI), USM, picture archiving and communication (PAC) imaging unit system database. The Dicom samples consist of reconstructed series of slices with blur and non-blur images for 20 subjects, and each subject consists of 200 slices of images.

Data Processing

The source Dicom DBT slices were split and converted into single 1089 x 2457 pixel Grayscale images. Then, the source image was resized to increase the focus for the Laplacian filter because the default kernel in the algorithm is 3x3, which is too small for the size of the source image and is primarily affected by noise. Accurate image size is crucial because the variance value can be impacted, leading to inaccurate results. Thus, an experiment was conducted to find the best image size by reducing the image gradually from <25% to <10%. The MATLAB Imresize function, which is part of the algorithm line, was used to resize each DBT slice. The results are discussed in the next section.

Laplacian-based Algorithm

This method was used to find edges in an image. The zero crossings of the second derivative of the image intensity can be used to discover an image of Laplacian-based edge detection sites (7). The Laplacian operator was further separated into two classifications: positive and negative Laplacian operators (8). The method convolves the input image with the Laplacian operator and computes the variance. The image is considered blurry if the variance exceeds a predetermined value. A high variance in a standard representative in-focus image indicates a large number of responses for both non-edge-like and edge-like (9). A low detected variance indicates little response dispersion, implying that the image has few edges. Therefore, it is possible to conclude that an image is blurred if it has a few numbers of edges. Figure 1 represents a sample of DBT images used for testing the LbBD algorithm.

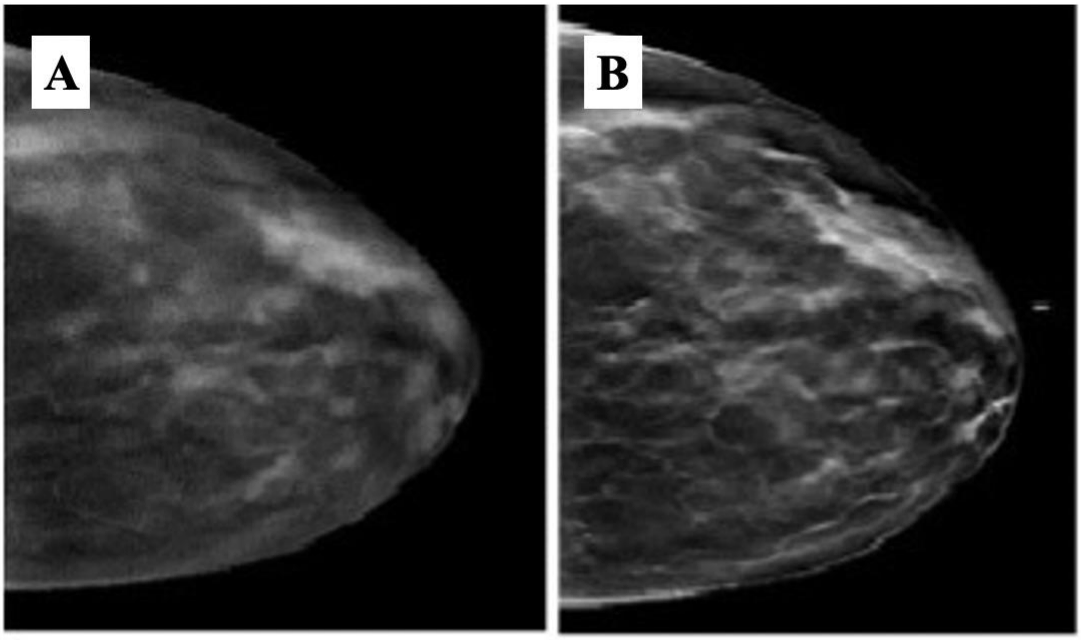


Figure 1: (A) Blurred Image (Out-focus) (B) Non-Blurred Image (In-focus)

A second-order Laplacian filter mask was convolved with the input frame to calculate the variance for edge detection. The Laplacian operator was used in this study to investigate the blurriness level that would be classified as blur or not-blur in the DBT image blur detection process. This level of blurriness is also referred to as a blurry vector. There are currently no standard parameters used by radiologists in clinical procedures to quantify the degree of blurriness.

It is a common practice to rely on the radiologist's qualitative assessment, which is subjective. Therefore, this study contends that a quantitative standard measurement of the level of blurriness is indeed required. Figure 2 depicts the proposed LbBD algorithm and Pseudo-code of the MATLAB programme (Table 1).

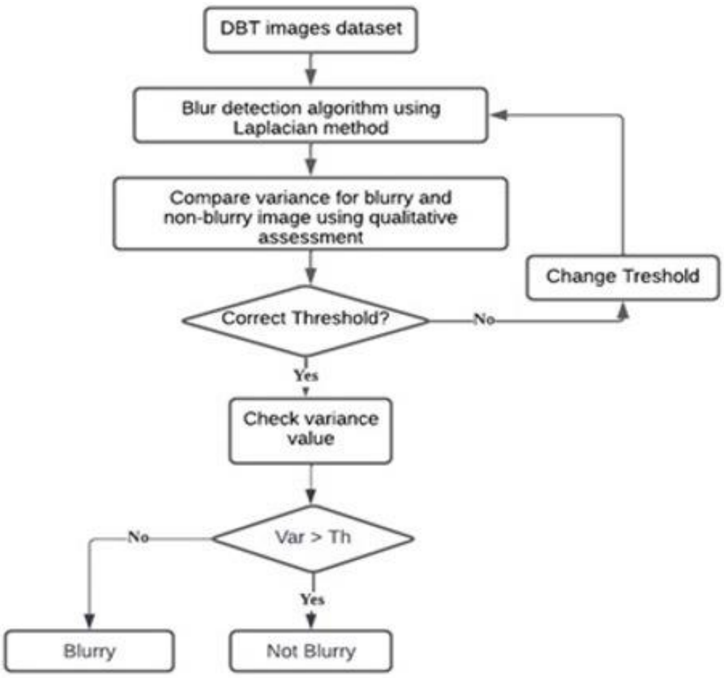


Figure 2: Proposed LbBD algorithm

Table 1: Pseudo-code of LbBD Algorithm and used MATLAB function.

Pseudo Code	**Function in MATLAB**
Initialize Inputs Images Dataset Register Get files length	 >> imageDataStore(sample directory); >> numberOfImages = length(Ori.Files);

Start Loop
Read image File name
Resize the image 189x246 pixels
Convert image type RGB to grayscale
Get double value
Set threshold value
Call variance subroutine
Get variance value
 	
IF variance < threshold
Set status as ''Blurry'
Write image to blurry folder
ELSE
Set status as ''NotBlurry'
Write image to Not blurry folder
End IF
End Loop
	
	>> inputFileName = Ori.Files{k};
	>> imread(inputFileName);
	>> imresize(Im,[189 246]);
	>> rgb2gray(rgbImage);
	>> double (Image);
	>> Th = 210;
	>> Variance_Inew = varian(Image);
	
	>> status = 'Blurry';
	>> imwrite(image,folder directory);
	
	>> status = 'NotBlurry';
	>> imwrite(image,folder directory);

Variance Subroutine Get Laplacian filter Filter the image Compute the filtered image variance End Subroutine	 >> J = fspecial('laplacian',0); >> imfilter(I,J); >> var(INew,0,'all');

The algorithm decisions entirely based on the threshold value set. Selecting a suitable threshold value is entirely reliant on the domain. If the threshold is set incorrectly, images will be inaccurately labelled; for example, a non-blurred image will be marked as blurry and vice versa. The variance value of 20 expert-evaluated DBT images was used to determine the threshold value in this study. Two weight centres, one for each class, are used in the division point calculation. The weighted mean of the classes is then chosen as the division point. Many threshold values that are close to the division point are examined to obtain the optimal outcome with a low error rate.

Results

First, to set some ground truth image data, an online survey was conducted with an expert to assess the quality of 20 DBT images using two subjective measurements of blurred or non-blurred images. Then, an experiment was conducted to determine an appropriate image size that produces a relevant variance (total number of edges) compared to the expert-evaluated status. The outcome is shown in Table 2. The findings indicate that the variance value is randomly changed on different image sizes once the image is reduced (resized) by factors ranging from 25% to 10%.

Table 2: Variance value for Image size reduction by factors ranging from 25% to 10%

Image No.	**Expert Evaluate Status**	**Not accurate**		**Possible Threshold range: 150<Th<250**	
		Size <25%	**Size <20%**	**Size <15%**	**Size <10%**
		P[473 614]	**P[378 491]**	**P[284 369]**	**P[189 246]**
1	***BLURRY***	44.89	57.01	**84.37**	**152.08**
2	SHARP	***113.77***	***159.8***	242.79	393.53
3	SHARP	87.77	125.94	200.43	332.43
4	***BLURRY***	32.07	35.43	45.3	71.63
5	***BLURRY***	***150.99***	***146.1***	105.01	119.74
6	SHARP	223.5	260.95	297.65	463.71
7	***BLURRY***	150.89	143.66	102.68	114.71
8	***BLURRY***	47.38	49.23	60.1	89.71
9	***BLURRY***	63.99	75.93	**103.28**	**167.63**
10	SHARP	***77.7***	***97.95***	**141.33**	**241.3**
11	SHARP	79.8	100.78	**145.25**	**236.62**
12	***BLURRY***	68.82	89.08	**127.47**	**205.52**
13	SHARP	124.61	174.39	259.74	413.65
14	SHARP	189.81	256.73	357.14	533.51
15	***BLURRY***	57.54	66.96	**91.23**	**153.21**
16	SHARP	84.84	108.82	158.48	266.79
17	SHARP	105.02	135.23	188.56	291.06
18	***BLURRY***	146.01	155.64	**138.52**	**183.22**
19	***BLURRY***	127.09	124.28	87.16	86.07
20	***BLURRY***	131.24	132.27	99.83	108.52

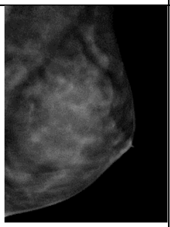
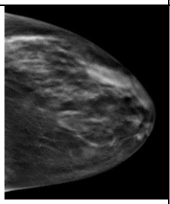
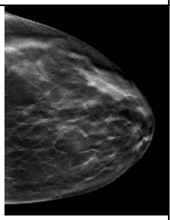
Note: Th = threshold, P = 'Pixel's dimension

The variance values were unstable for image pixel reduction of 25% to 20%, i.e., the variance of images 2, 5, and 10 have the same value for blurry and sharp images, respectively. In addition, the variance was too small for the image that had been evaluated as sharp images by the expert. The predicted variance of 25% to 20% image reduction was inaccurate, making it challenging to choose the threshold value. Then, for image reductions of 15% to 10%, the proposed LbBD algorithm effectively calculated the variance and classified the image status in accordance with expert assessment by the suitable range of threshold value setup. Based on the evaluation, the possible threshold range to be selected was 150<Threshold<250. Consequently, the 10% resized image was used for further testing.

The blur vector obtained was compared with a threshold value to make the blur decision on the image. In Table 3, two threshold values of 150 and 200 were used to detect whether the image is blurred. Different threshold values may influence the effect of blur detection. The third image in the table generates various statuses for each threshold. As shown in Table

4, with an appropriate image size and threshold value, the algorithm predicts blurred images efficiently and accurately, with a significantly low rate of error. A comparison of three threshold value findings revealed the impact of threshold value on the effectiveness of blur detection. The accuracy decreases, and the error rate rises if the threshold value is set either too high or too low. Therefore, 200 was the ideal threshold number for the DBT images dataset to achieve a 95% accuracy and 0.05 error rate, while the accuracy was reduced at the lower and upper threshold values of 90% and 75%, respectively.

Table 3: Sample Result of blur detection with LbBD using two Threshold value

DBT Image	Variance	Status (Th = 150)	Status (Th = 200)
	86.07	Blurry	Blurry
	119.74	Blurry	Blurry
	**152.08**	**Not Blurry**	**Blurry**
	332.43	Not Blurry	Not Blurry

Note: Th = threshold

Table 4: Qualitative Survey for Image Quality of DBT images

Image No.	Experts evaluate status	Algorithm status (Th=250)	Algorithm status (Th=200)	Algorithm status (Th=150)	Variance
1	**Blurry**	Blurry	Blurry	***Not blurry***	**152.08**
2	Not blurry	Not blurry	Not blurry	Not blurry	393.53
3	Not blurry	Not blurry	Not blurry	Not blurry	332.43
4	Blurry	Blurry	Blurry	Blurry	71.63
5	Blurry	Blurry	Blurry	Blurry	119.74
6	Not blurry	Not blurry	Not blurry	Not blurry	463.71
7	Blurry	Blurry	Blurry	Blurry	114.71
8	Blurry	Blurry	Blurry	Blurry	89.71
9	**Blurry**	Blurry	Blurry	***Not blurry***	**167.63**
10	**Not blurry**	***Blurry***	Not blurry	Not blurry	**241.3**
11	**Not blurry**	***Blurry***	Not blurry	Not blurry	**236.62**
12	**Blurry**	Blurry	***Not blurry***	***Not blurry***	**205.52**
13	Not blurry	Not blurry	Not blurry	Not blurry	413.65
14	Not blurry	Not blurry	Not blurry	Not blurry	533.51
15	**Blurry**	Blurry	Blurry	***Not blurry***	**153.21**
16	Not blurry	Not blurry	Not blurry	Not blurry	266.79
17	Not blurry	Not blurry	Not blurry	Not blurry	291.06
18	**Blurry**	Blurry	Blurry	***Not blurry***	**183.22**
19	Blurry	Blurry	Blurry	Blurry	86.07
20	Blurry	Blurry	Blurry	Blurry	108.52
Error rate		**0.1**	***0.05***	**0.25**	
Accuracy		**90%**	***95%***	**75%**	

Image data with a smooth trend generated false alarms, whereas images that were blurred but had an abrupt change from high detail to low detail resulted in missed detection. Some of the samples that were not detected as reported by LbBD are illustrated in Figure 3. Since the LbBD algorithm inspects the singularities in the image that are influenced by image blurriness, it resulted in 7 missed detections among three compared thresholds.

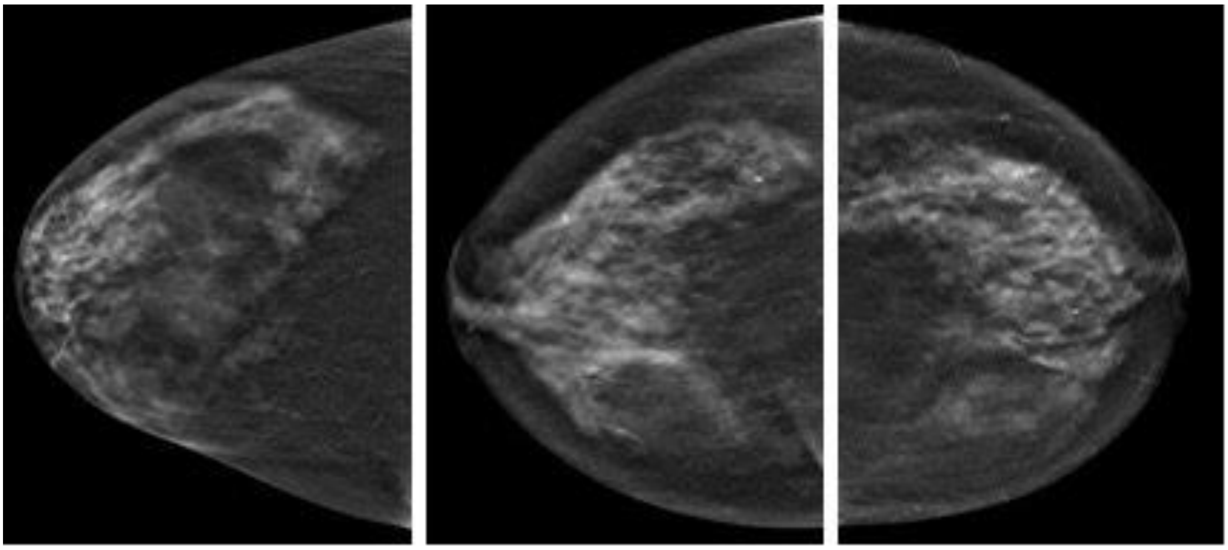


Figure 3: Sample missed detection by LbBD algorithm for images 9, 12 and 15

Discussion

Low image quality significantly impacts image interpretation and breast cancer diagnosis in DBT images (10). Therefore, it is critical to detect and comprehend the potential image quality concerns that may affect the visibility of breast lesions. An image blur is a mathematical convolution between the source image and the point spread function known as the blurring kernel. Blurring happens when each pixel in the image spreads over the surrounding pixels. This spreading process is more often referred to as smearing out. The LbBD algorithm computes variance by simply convolving the input picture with the Laplacian operator. The image is blurry if the variance is smaller than a certain amount. A significant variation in a typical representative in-focus picture implies the existence of a large number of non-edges-like and edge-like responses. In contrast, a low detected variance suggests little response dispersion, meaning that the images have only a few edges. Therefore, it is possible to conclude that an image is blurred if it contains few edges.

The original DBT image's in Dicom format size is 1890 x 2457 Pixel. The size has the drawback of being larger than the Laplacian kernel size. Because the image's borders are substantially larger than 3 x 3 pixels, the noise will significantly impact the original image with a kernel size of three. In other words, the 3 x 3 kernel emphasises a much higher special frequency than the details' low "special frequency". Scaling the image may therefore result in a more precise variance value. When the image is scaled from 25% to 20%, as shown in Table 1, the variances are erroneous since the focus is inadequate. Some variance values differ between blurry and non-blurry images compared to expert assessment. The image is then downsized from 15% to 10% of its original size. The value is accurate with the chosen threshold value.

The detection of blur is based on the precise identification of the edges. Edge detection includes applying a threshold value to evaluate whether or not a given pixel is blurred. Inaccurate threshold settings will result in poor edge recognition, yielding an incorrect decision. Additionally, photos taken as a result of inadequate convergence of light from an object on the image sensor plane produce what are known as out-of-focus images. A defocus blur on image results from a circle-shaped distribution of pixel intensities around its neighbours. There are several transitions from high grey level intensities to low grey level intensities in missed detection cases, which lead to false edge detection and a high value of variance reporting a missed detection. Furthermore, the difference between the smallest and highest variances was mathematically discovered to be quite large. Therefore, the minimum and maximum variance ratio resulted in a minimal number that determined the sharpness factor and classified the image as blurred. A thorough analysis of false alarms produced by the LbBD algorithm showed that they were brought on by images with very low contrast or linear characteristics, as illustrated in Figure 3.

Conclusion

In this study, a new image blur detection technique with a Laplacian-based algorithm for detecting blurriness in DBT images is proposed. This algorithm demonstrates how a simple step involving a Laplacian filter and a variance convolve can achieve excellent blur detection performance on a medical image such as a DBT dataset. The result will assist in removing the uncertain images from the diagnostic procedure. Based on the results, the correct threshold value is critical to achieve good accuracy and error rate performance. If the threshold value is chosen correctly, the algorithm can accurately identify blurred and non-blurred images with a low error rate. Besides, the algorithm is also sensitive to image size due to the fixed kernel in the Laplacian filter function. As discussed in the previous section, the accurate variance calculation achieved at the image size reduction range is between <15% to <10%. Significant performance can be observed by the algorithm effectively predicting blurry images with a low error rate (0.05), an accuracy of 95% at a selected threshold value of 200 and an image size reduction of <10%.

As part of the long-term project goal, a Convolutional Neural Network (CNN) blur detection method will be studied and applied to compare the performance of the DBT blur detection system. To the best of our knowledge, there is currently no publicly accessible image quality database that searches for DBT images to deblur the quality of the image or its blurry vector. As a result, this study entails the creation of a dataset of blurry and non-blurry DBT images. This dataset will be used in the next section, and the blurred vector produced by this method will be utilised to improve the blur detection model.

Acknowledgement

This research work is financially supported by the Ministry of Higher Education Grant Scheme (FRGS) "A New Cascaded Convolutional Neural Network Model for Deblurring and Contrast Enhancement of Extremely Dense Breast Tissue in Digital Breast Tomosynthesis Images" (Ref: FRGS/1/2021/TK0/UiTM/02/19). The authors would like to express their gratitude to members of the Advanced Control System and Computing Research Group (ACSCRG), Advanced Rehabilitation Engineering in Diagnostic and Monitoring Research Group (AREDiM), Integrative Pharmacogenomics Institute (iPROMISE) and Centre for Electrical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang for their assistance and guidance during the fieldwork. The authors are grateful to Universiti Teknologi MARA, Cawangan Pulau Pinang for their immense administrative and financial support. The authors are also grateful to the staff of Advanced Medical & Dental Institute, USM and Breast Cancer Translational Research Programme (BCTRP), Advanced Medical & Dental Institute, USM for their continuous support.

Competing Interests

All authors declare no competing interests.

Ethical Clearance

Ethics approval was obtained from Universiti Sains Malaysia (USM/JEPeM/21090622). Data from 20 subjects with underlying breast cancer were collected in a retrospective study from December 2019 until December 2021.

Informed Consent

Verbal informed consent was obtained from all subjects before the study. Written consent was not obtained because the data were collected retrospectively, and the subjects were not physically present at the healthcare facility. The images were the property of the institution and local approval has been obtained for both retrieval and publication.

References

1. Caumo F, Zorzi M, Brunelli S, Romanucci G, Rella R, Cugola L *et al*. Digital breast tomosynthesis with synthesized two-dimensional images versus full-field digital mammography for population screening: Outcomes from the verona screening program. Radiology. 2018;287(1):37-46
2. Ali U, Mahmood MT. Analysis of blur measure operators for single image blur segmentation. Appl Sci. 2018;8(5)
3. Koik BT, Ibrahim H. A literature survey on blur detection algorithms for digital imaging. In: Proceedings - 1st International Conference on Artificial Intelligence, Modelling and Simulation, AIMS 2013. ; 2013:272-277
4. Pertuz S, Puig D, Garcia MA. Analysis of focus measure operators for shape-from-focus. Pattern Recognit. 2013;46(5):1415-1432
5. Bansal R, Raj G, Choudhury T. Blur image detection using Laplacian operator and Open-CV. Proc 5th Int Conf Syst Model Adv Res Trends, SMART 2016. Published online 2017:63-67
6. Tran GS, Nghiem TP, Doan NQ, Drogoul A, Mai LC. Fast parallel blur detection of digital images. 2016 IEEE RIVF Int Conf Comput Commun Technol Res Innov Vis Futur RIVF 2016 - Proc. Published online 2016:147-152
7. Shrivakshan GT, Chandrasekar C. A Comparison of various Edge Detection Techniques used in Image Processing. Int J Comput Sci Issues. 2012;9(5):269-276. Accessed November 16, 2021.
8. Pagaduan RA, R. Aragon MC, Medina RP. iBlurDetect: Image Blur Detection Techniques Assessment and Evaluation Study. In: International Conference on Culture Heritage, Education, Sustainable Tourism, and Innovation Technologies (CESIT2020). ; 2020:286-291
9. Szandala T. Convolutional Neural Network for Blur Images Detection as an Alternative for Laplacian Method. 2020 IEEE Symp Ser Comput Intell SSCI 2020. Published online 2020:2901-2904
10. Sujlana PS, Mahesh M, Vedantham S, Harvey SC, Mullen LA, Woods RW. Digital breast tomosynthesis: Image acquisition principles and artifacts. Clin Imaging. 2019;55(June 2018):188-195