Effective Strain Image Sequence Selection by Using Semi-Automated Image Processing Technique

Gulam Mahfuz Chowdhury, Md. Mahedi Hasan, Asif Ahmed, Md. Wahid Tousif Rahman
and Md. Taslim Reza

Abstract—One fourth of the cancer detected in women worldwide is breast cancer which leads this as a major threat for women. There are many methods of detecting cancer among which ultra-sound strain imaging is one of the promising techniques. However, in strain sequence, not all the frames show clear tumor visibility. Consequently, in this paper we tested some well-defined algorithms to find only those frames where the tumor is comparatively clearly visible. We have used Mean Pixel Difference (MPD) and Gray-Level Co-occurrence Matrix (GLCM) to find the frames with better tumor visibility. We have tested our methods in several real-life cases and the results have been examined by a professional doctor. The MPD has an accuracy of 96.2% and the GLCM. Contrast has that of 55.55%.

Index Terms—B Mode image, Computer Aided Diagnosis (CAD), Contrast, Elastography, GLCM, Strain Image.

I. INTRODUCTION

BREAST cancer is the most frequent form of cancer in women worldwide, with about 25% of all cancer diagnosed in women being breast cancer [1]. Around one in every eight women in the United States will suffer from breast cancer over the course of her lifetime. It is also the second major cause of death among women worldwide [2] [3] [4], claiming 500,000 lives every year [5]. Early detection and treatment are pre-eminent in reducing the death toll due to the disease [6-8]. Due to its cost- effectiveness and preclusion of needles and radiation,

G.M. Chowdhury was with Islamic University of Technology (IUT), Gazipur, Bangladesh. He is now with the Department of Electrical and Electronic Engineering, Leading University, Sylhet, Bangladesh. E-mail: *gulammahfuz@iut-dhaka.edu*

Md. M. Hasan is with IUT, Gazipur, Bangladesh. E-mail: *mahedihasan@iut-dhaka.edu*

A. Ahmed is with IUT, Gazipur, Bangladesh. E-mail: *asifahmed@iut-dhaka.edu*

Md. W. T. Rahman is with IUT, Gazipur, Bangladesh. E-mail: *wahidtousif@iut-dhaka.edu*

Md. T. Reza is with IUT, Gazipur, Bangladesh. E-mail: *taslim@iut-dhaka.edu*

ultrasonography presents itself as an auspicious cancer diagnosis tool [9-11]. Ultrasonography carries with its no known risk of exacerbating cancer and the process can be carried out on the patient with little to no prior preparation. Fully automated Computer Aided Diagnosis (CAD) is difficult to achieve because of the challenges posed by the breast tissue density, the location of the tumor, and speckle/noise intrinsic to the ultrasound image [12]. As such, a semi-automatic diagnosis methodology is proposed in this paper which would compare the frames found from the sequences and will give the most clearly visible frames from it. For simplicity, we have shown the best 3 of those in the tables of the result section.

US transducers initially convert electrical signals into pressure waves. Reflection and scattering are caused by the differences in tissue density and speed of sound of the transmit pulses, and it reflects a portion of the sound waves towards the transducer. They are then converted into a receive electrical signal which is composed of multiple reflections. This received signal is commonly known as RF data. Passing the RF signal through an envelope detector yields 'A mode' images. Multiple 'A mode' images are then converted into a 2D grayscale image known as a 'B mode' image. As malignant tumors have no specific shape and cannot be easily differentiated from the surrounding breast tissue, 'B mode' images are sometimes insufficient to facilitate their detection. Applying pressure over the area of interest results in the formation of strain images, which are likely to present a better view of the tumor in cases where 'B mode' images prove ineffective [9]. When tissue is pressed, Ultrasonic strain imaging mimics manual palpation to feel the alterations in hardness. US strain imaging demonstrates great sensitivity for breast tumor differentiation unlike said manual palpation [27].

The nature of tissues subjected to mechanical forces is the basis of elastography [13-14]. While taking pre- and post-compression echo signals, the tissue displacements are measured from the time delays between them [15-16]. Strain is then calculated from the estimated displacement maps. They are developed from the tissue displacement information form pre- and post-compression RF data [17]. Applied pressure does not result in the shape of the tumor changing much owing to its higher density relative to

surrounding tissue. The strain images show the surrounding tissue changing on a higher scale than the tumor, thereby making the tumor more visible. The strain images are obtained in the form of an image sequence where not every frame offers well tumor visibility. Poor visibility in certain frames may result from noise introduced during the strain estimation. It follows that only frames with good tumor visibility be considered during diagnosis.

In this paper, two image processing methods (MPD & GLCM) have been used to detect the clearly visible frames among those image sequences produced from strain sequences. We have used these methods because these have not been used in this context before and we have got a very promising result. In the first step, we have extracted the frames from the strain image sequence and then the frames have been preprocessed and enhanced to reduce speckles and noises in the second step. The frames have been sorted based on tumor visibility as described by Mean Pixel Difference (MPD) and Gray-Level Co-occurrence Matrix also known as GLCM (contrast) [18] in the third step. The GLCM features homogeneity, energy and contrast of healthy and malignant cells that are noticeably distinct from one another [19].

The choice of these methods is for their promising results, algorithmic simplicity, and pioneer use in this context. Manual sorting of a lot of frames consumes time and energy. A doctor/radiologist can easily sort them out using an efficient algorithm. Insufficient research in this area motivated us to work on this.

II. METHODOLOGY

A. Frame Extraction

The strain image sequence is formed from pre- and post-compression ultrasound RF data. The frames are extracted from the sequence and stored in the hard drive for processing.

B. Pre-processing Frames

Pre-processing reduces speckles and lowers the noise level. Running each frame through a Hybrid Filter, consisting of a median filter and Gaussian filter for seven times considerably reduces the noise level and makes it easier to process [20]. It also reduces the speckles. The demonstration of employing and restraint of such filters in biomedical image processing has been done in [21-22]. For convenience the frame shown in Fig. 1 has an SNR of 10.12 dB in the pre-processed case and 12.67 dB in the post-processed case.

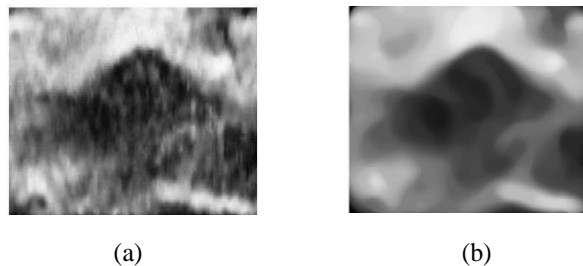(a)   (b)

Fig. 1. Pre-processing steps: (a) before processing, (b) After processing Ultrasound taken at 10.0 MHz frequency and 5.0 cm depth.

C. Using MPD for Clearly Visible Frame Selection

The tumor lesions have much lower pixel values in the ultrasound images than the surrounding tissue, owing to their higher density. This means that tumors appear dark while the surrounding tissue appears whitish. Thus, it follows that the average pixel value of the whole frame be lower than the region to which the tumor is confined, i.e., the region with the tumor in it exhibits high mean pixel density. If the tumor is poorly visualized then a large mean pixel difference will not be observed for that particular frame.

This is where the semi-automated CAD comes in.

The whole frame and roughly the region with the tumor in it are selected manually. Based on a single viewing of the strain image sequence, an area of interest where the tumor may be located can be assumed. The mathematical mean of the pixel values within the whole frame and the area of interest are found using the following equation:

$$\text{avg} = \frac{\sum_{i=1}^{N} i^{th} pixel\ value}{N} \qquad (1)$$

Where N = Number of pixels within the area

The tumor may not retain its position during the period of the ultrasonography and may move around as a result of the application of pressure. To account for this change, the area of interest is moved downward, upward, to the left, to the right, upward and to the right, upward and to the left, downward and to the right and downward and to the left. The minimum pixel average from each of these nine positions is taken as the tumor will have lower pixel values. Thus, if the tumor moves from its initial position while collecting the RF data, it can be compensated.

After calculating the mean pixel values of both the entire frame and the area of interest, their difference is found and stored for each frame. This difference is the Mean Pixel Difference (MPD). Frames are then sorted by MPD, with higher values of MPD corresponding to better tumor visibility. This process is applied to various strain image sequences which yielded auspicious results.
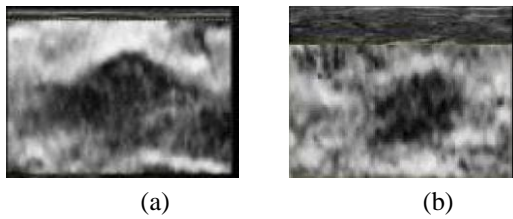
(a)   (b)

Fig. 2. Some output frames from strain videos using MPD method
  (a) - Ultrasound taken at 10.0 MHz frequency and 5.0 cm depth.
  (b) - Ultrasound taken at 6.6 MHz frequency and 5.0 cm depth.

D. Using GLCM.Contrast for Clearly Visible Frame Selection

The other method of selecting the clearly visible frames uses their GLCM properties as image descriptors. GLCM characterizes the texture profile of a given image. The Gray level co-occurrence matrix (GLCM) is a function that can compute the probability density function (pdf) of an image f(a,b) for all pair of pixels (x) and (y) in distance (d) where the angular displacement is $(\theta) = 0, 45, 90$ and 135 degree [20][23]. The contrast statistic of the GLCM defers the calculation of the intensity contrast linking pixel and its neighbor over the whole image. For a constant image, the value of contrast is 0. While measuring contrast, exponential increase of weight takes place (0,1,4,9) as persists from the diagonal [18][24].

$$contrast = \sum_{x,y=0}^{n-1} p(x,y)\,|x-y|^2 \qquad (2)$$

Equation (2) is used to calculate contrast from the Gray Level Co-Occurrence Matrix of a grayscale image. Contrast is known as the difference in luminance (color) from which we can distinguish an object. Contrast is regulated by the difference in the color and brightness between the main object and other objects within the same field of view in the real world. We can define a contrast as the difference between the light and dark parts of an image [25]. An idea of the intensity contrast between a pixel and its neighbor (over the entire image) can be obtained from GLCM.Contrast. A uniform and visible tumor will result in a low value. If the tumor is obscured by noise, there is no uniformly distributed tumor region and the GLCM.Contrast value is relatively high. It follows that frames with high GLCM.Contrast values exhibit higher tumor visibility, making the tumor more distinguishable from the surrounding tissue in such frames.

Following the process, GLCM.Contrast is calculated for each of the frames which are then sorted in ascending order of GLCM.Contrast. The required frames are found in this manner.
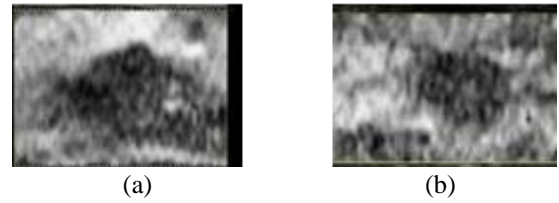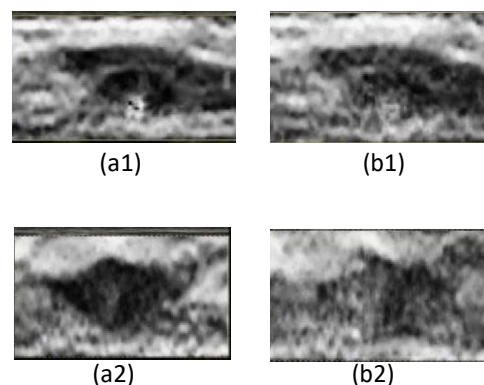

(a)   (b)

Fig. 3. Some output frames from strain videos using MPD method
  (a) - Ultrasound taken at 10.0 MHz frequency and 5.0 cm depth.
  (b) - Ultrasound taken at 6.6 MHz frequency and 5.0 cm depth.

III. RESULTS

Videos are compiled using the proposed algorithm via MATLAB R2017a running on a personal computer. The in vivo breast data that we are using in these tests are chosen from a database of 7 cases (age group: 20 to 75 years). These data were obtained by using a sonix sP500 scanner integrated with a 114–5/38 probe operating at 10 MHz (nominal) at the University of Vermont Medical center.

Now, these images are mostly unclear as they are comprised of a lot of noises, and hence these methods are implemented to select which of the frames give better tumor visibility. Then these frames are marked (which is obviously by computer programming) based on their clear visibility which happens automatically. After that, these frames are sorted based on the markings. Thus, the top 10 best frames are selected. Now some of the best images that our systems detected from 7 different sequences are given below:
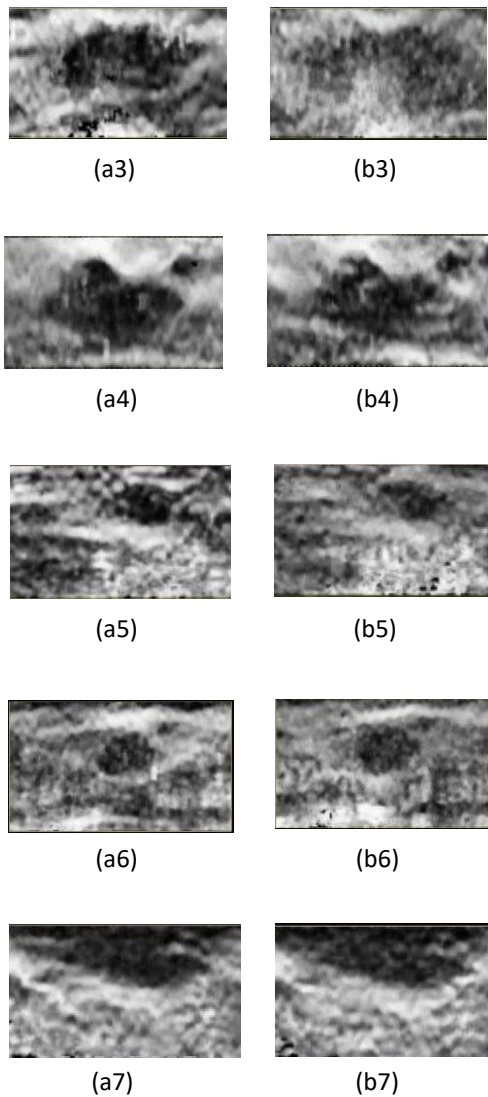

(a1)   (b1)

(a2)   (b2)

(a3) (b3)

(a4) (b4)

(a5) (b5)

(a6) (b6)

(a7) (b7)

Fig. 4. Some best output frames obtained using MPD method (a1-a7) and GLCM. Contrast method (b1-b7) for 7 different strain image sequences.

In case of the strain image sequence of frames (a2) and (b2) in Fig. 4, the range of value of MPD is from 4.8766 to 55.79 and the value of the most clearly visible three frame obtained through this method are 55.79, 52.846 and 51.5195. Again, the range of the values of GLCM.Contrast is from 5.9047 to 15.47079 and the value of the most clearly visible three frames obtained through the method are 5.9047, 5.9806 and 5.9825.

To understand how effective the system is, a standard is chosen where manually frames with best tumor visibility are selected by the doctor and radiologist (Dr. Rosy Parveen, Associate Professor (c.c.) & Head of the Dept. [Radiology and Imaging], Medical College for women & Hospital, Uttara, Dhaka). Then, the frames that are given by the system are compared with the manually selected frames. If the system selected frame falls within 3 close frames

of the manually selected frame then the frame is taken as selected automatically otherwise not selected. Here each frame number indicates the location of the frame in the particular frame sequence.

TABLE I
ERROR COMPARISON BETWEEN MPD AND GLCM METHODS.

Video no.	Manually selected (Frame Number)	Selected by MPD (Frame Number)	Selected by GLCM. Contrast (Frame Number)	No. of missing frames	
				MPD (out of 3)	GLCM (out of 3)
1	43	43	44	0	2
	213	214	220		
	248	249	243		
2	146	146	146	0	1
	124	124	124		
	115	113	126		
3	228	227	225	0	0
	174	176	174		
	171	170	173		
4	26	26	26	0	1
	269	269	265		
	309	309	309		
5	79	79	95	0	1
	211	209	211		
	102	101	99		
6	41	41	40	1	1
	95	96	94		
	106	121	106		
7	26	25	31	0	1
	105	104	107		
	219	220	218		

TABLE II
PERCENTAGE (%) OF ERROR COMPARISON BETWEEN MPD AND GLCM METHODS.

Video no.	Percentage (%) of error by MPD	Percentage (%) of error by GLCM
1	0	66.67
2	0	33.33
3	0	0
4	0	33.33
5	0	33.33
6	33.33	33.33
7	0	33.33

TABLE III

Confusion Matrix of MPD.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	26	0
	Negative (0)	1	0

TABLE IV

Confusion Matrix of GLCM.Contrast

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	15	0
	Negative (0)	12	0

While using MPD, the average pixel value is calculated to find out the clearly visible frames. But in GLCM.Contrast method, we calculate the pixel-to-pixel variation which is greatly affected by the noise level of the frame. This makes the GLCM.Contrast method more prone to error. We form confusion matrices for both cases and find out that the MPD has an accuracy of 96.2% and the GLCM.Contrast has an accuracy of 55.55%.

IV. Conclusion

Ultrasonography has gained traction in recent years and consequently elastography has seen widespread use among clinicians [26]. Two methods of using ultrasound imaging have been proposed in this paper that results in a better detection of breast frames that have very good tumor visibility. Generally, doctors or radiologists go through the whole strain image sequence and find out the clearly visible frames manually. This is time consuming and prone to human error. Our tested methods can significantly reduce the required time. Significant improvements can be made to the proposed method by using better pre-processing techniques and developing the means to adapt to the movement of the tumor. Using a more suitable filter can reduce the noise level which can improve the accuracy of GLCM.Contrast method. Also, the selection of ROI (Region of Interest) can be made automated. Moreover, Machine Learning can be implemented on these output frames from where a decision on the tumor being benign or malignant or the probability of being cancerous can be drawn.

References

[1] American Medical Association, "Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability Adjusted Life-years for 32 Cancer Groups, 1990 to 2015," JAMA Oncology, vol. 3, pp. 524-548, 2017.

[2] H. Cheng, J. Shan, W. Ju, Y. Guo and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," Pattern Recognition, vol. 43, no. 1, pp. 299-317, 2010.

[3] J. Ferlay, I. Soerjomatara, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman and F. Bray, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," International Journal of Cancer, vol. 136, no. 5, pp. E359-E386, 2014.

[4] Breast cancer, Understanding Breast Cancer Risk http://www.breastcancer.org/risk/understanding.

[5] M. M. Saleck, A. El Moutaouakkil and M. Mouçouf, "Tumor Detection in Mammography Images Using Fuzzy C-means and GLCM Texture Features," in *2017 14th International Conference on Computer Graphics, Imaging and Visualization*, Marrakesh, 2017.

[6] J. Chamberlain, D. Coleman, R. Ellman, S. Moss, M. Alderson, C. Faulder, J. Hayward, M. Jack, C. Joslin, C. Ray, C. Spicer, J. Swanson Beck and A. Tucker, "First results on mortality reduction in the UK Trial of Early Detection of Breast Cancer. UK Trial of Early Detection of Breast Cancer Group," The Lancet, vol. 8608, no. 332, pp. 411-416, 1988.

[7] L. Tabár, A. Gad, L. Holmberg, U. Ljungquist, K. C. P. Group, C. Fagerberg, L. Baldetorp, O. Gröntoft, B. Lundström, J. Månson, Ö. C. P. Croup, G. Eklund, N. Day and F. Pettersson, "Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare.," *The Lancet,* vol. 325, no. 8433, pp. 829-832, 1985.

[8] S. Preston-Martin, M. C. Pike, R. K. Ross, P. A. Jones and B. E. Henderson, "Increased Cell Division as a Cause of Human Cancer," *Cancer Research,* vol. 50, no. 23, pp. 7415-7421, 1990.

[9] J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi and X. Li, "Elastography: A quantitative method for imaging the elasticity of biological tissues," *Ultrasonic Imaging,* vol. 13, no. 2, pp. 111-134, 1991.

[10] S. Franchi-Abella, C. Elie and J.-M. Correas, "Ultrasound elastography: Advantages, limitations and artefacts of the different techniques from a study on a phantom," *Diagnostic and Interventional Imaging,* vol. 94, no. 5, pp. 497-501, 2013.

[11] F. Tranquart, N. Grenier, V. Eder and L. Pourcelot, "Clinical use of ultrasound tissue harmonic imaging," *Ultrasound in Medicine & Biology,* vol. 25, no. 6, pp. 889-894, 1999.

[12] B. A. Abrahim and Y. Kadah, "Speckle noise reduction method combining total variation and wavelet shrinkage for

clinical ultrasound imaging," in *1st Middle East Conference on Biomedical Engineering*, Sharjah, 2011.

[13] J. Ophir *et al.*, "Elastography: Ultrasonic estimation and imaging of the elastic properties of tissues," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine,* voL 213, no, 3, pp,203-233, aL 1999.

[14] L. Gao, K, 1. Parker, R. ~, Lerner, and S, F. Levinson, "Imaging of the elastic properties of tissue-A review," Ultrasound in Medicine & Biology, voL 22, no, 8, pp, 959-977, Jan, 1996.

[15] L. S. Wilson and D. E. Robinson, "Ultrasonic measurement of small Displacements and Deformations of tissue," *Ultrasonic Imaging,* voL 4, no, l, pp, 71-82, Jan, 1982.

[16] Irteza Enan Kabir, Rasheed Abid, Abdullah Salmon Ashik, Kazi Khairul Islam, S. Kaisar Alam, "Improved strain estimation using a novel 1.5D approach: Preliminary results", 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec).

[17] M. O'Donnell, A. Skovoroda, B. Shapo and S. Emelianov, "Internal displacement and strain imaging using ultrasonic speckle tracking," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control,* vol. 41, no. 3, pp. 314-325, 1994.

[18] E. K. Sharma, E. Priyanka, E. A. Kalsh and E. K. Saini, "GLCM and its Features," *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE),* vol. 4, no. 8, pp. 2180-2182, 2015.

[19] M. B. Al Rasyid, Yunidar, F. Arnia and K. Munadi, "Histogram statistics and GLCM features of breast thermograms for early cancer detection," in *2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON)*, Chiang Rai, 2018.

[20] A. Abdelrahman and O. Hamid, "Breast Ultrasound Images Enhancement Using Gray Level Cooccurrence Matrices Quantizing Technique," *International Journal of Information Science,* vol. 2, no. 5, pp. 60-64, 2012.

[21] Christos P. Loizou and Constantinos S. Pattichis, Despeckle Filtering Algorithms and Software for Ultrasound Imaging, Morgan & Claypool.USA.2008.

[22] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Second Edition, Prentice Hall, New Jersey, 2002.

[23] Robert M. Haralick, "Statistical and Structural Approaches to Texture" IEEE Proceedings of Vol. 67, no. 5, May 1979.

[24] D. Narain Ponraj, M. K. Evangelin, Jennfier P Poongodi, J Samuel Manoharan, "A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer" Journal of Emerging Trends in Computing and Information Sciences, vol. 2, NO. 12, December 2011.

[25] Marius Pedersen, Alessandro Rizzi, Jon Yngve Hardeber and Gabriele Simone, "Evaluation of contrast measures in relation to observers perceived contrast" Society for Imaging Science and Technology, 2008.

[26] K. Parker, M. Doyley and D. Rubens, "Imaging the elastic properties of tissue: the 20-year perspective," Physics in Medicine & Biology, vol. 56, no. 1, pp. 513514, 2010.

[27] Jingfeng Jiang, Timothy J. Hall, and Amy M. Sommer, "A Novel Image Formation Method for Ultrasonic Strain Imaging", Ultrasound Med Biol. 2007 Apr; 33(4): 643–652.

University of Technology (IUT), Gazipur in 2018. He has been a lecturer with the Electrical and Electronic Engineering Department, Leading University, Sylhet. His Research interests include Biomedical Image Processing, Renewable Energy, Power Electronics.

Md. Mahedi Hasan was born in Dinajpur, Bangladesh. He received his B.Sc. Engineering degree in Electrical and Electronic Engineering from IUT, Gazipur in 2018. He is serving as a data scientist at W4 solution since 2019.

Asif Ahmed was born in Sirajganj, Bangladesh. He received his B.Sc. Engineering degree in Electrical and Electronic Engineering from IUT, Gazipur in 2018. He is currently doing his M.Sc. degree from the same university. His research interests include Biomedical Image Processing and Plasmonics.

Md. Wahid Tousif Rahman was born in Rajshahi, Bangladesh. He received his B.Sc. Engineering degree in Electrical and Electronic Engineering from IUT, Gazipur in 2018. He has been working as an adjunct lecturer at the Sonargaon University, Dhaka.

Md. Taslim Reza, working as an Associate Professor in the department of Electrical and Electric Engineering of IUT since 2012. He has done his PhD in the field of Ultrasound Imaging from department of EEE of IUT. He has done his M.Sc. from Tampere University of Technology, Finland in the field of Biomedical Engineering. Currently he is working on Medical Imagining, Medical Instrumentation, Signal Processing

Gulam Mahfuz Chowdhury was born in Sylhet, Bangladesh. He received his B.Sc. Engineering degree in Electrical and Electronic Engineering from the Islamic