



Exterior Means for Premature Recognition of Breast Cancer

*AJIBOLA, OOE; BABAFEMI OM; ADEDEJI, SA; FOLORUNSO, OP

Department of Systems Engineering, University of Lagos, Lagos, Nigeria
**Corresponding Author Email: waleisiro@gmail.com; Tel: +2348034488877*

ABSTRACT: This research is aimed at the early detection of breast cancer. Breast cancer is a form of cancer that is found in the breast tissue. Though mostly common in women, research has shown that it can also be diagnosed in men. The increasing trend in the mortality rate attributable to breast cancer can be traced to inefficient methods of detecting the cancer in its early stage. Current methods involve invasion, exposure of the patient to radiation, and/or compressing the breast. These may cause pain or even invest patient with other forms of ailments. In this work, infrared (IR) thermography has been deployed to detect breast cancer in its early stage. IR thermography uses IR radiation to measure heat patterns of human skin. It is passive in nature and it neither emit harmful radiation nor subject the patient to further risks. The thermograms captured by infrared camera are analysed by a software in stages, viz: extraction of region of interest, detection and masking of warm region among others. The resulting image in grayscale is sampled by comparing the white pixels (warm region) to the dark pixels (cool region). The software subsequently compares the outcome with a predefined threshold to predict the chances of occurrence of cancer and displays result for further diagnosis by medical expert. Our proposed modality is cost effective, safe and user friendly.

DOI: <https://dx.doi.org/10.4314/jasem.v22i5.24>

Copyright: Copyright © 2018 Ajibola *et al.* This is an open access article distributed under the Creative Commons Attribution License (CCL), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dates: Received: 01 April 2018; Revised: 22 April 2018; Accepted: 25 April 2018

Keywords: Breast cancer, thermography, temperature variations, infrared thermal imaging

Breast cancer is a type of cancer that develops from the breast tissues comprising of lobules and the ducts that connect lobules to the nipple. According to National Cancer Institute (2017), estimated new cases of cancer in 2016 was about 246,660, and 14.6% of the population of cancer patients are cases of breast cancer. Estimated cancer induced death cases were 40,450 and 6.8% of those deaths were due to breast cancer. The disease starts when cells in the breast begin to grow out of control forming a tumor that can often be seen on an x-ray or felt as a lump. The tumor is malignant if the cells invade surrounding tissues or spread (metastasize) to distant areas of the body (Fernandez *et al.*, 1998; Bhojal and Kennedy, 2005; Bartella *et al.*, 2006). Breast cancer occurs almost entirely in women, it can also manifest in men. Breast cancers may start from different parts of the breast. Most breast cancers begin in the ducts that carry milk to the nipple (ductal cancers). Some start in the glands that make breast milk (lobular cancers). Other types of breast cancer are less common (Amy, 2010). For instance, a small number of cancers start in other tissues in the breast. These cancers are called sarcomas and lymphomas and are not really thought of as breast cancers. Although many types of breast cancer manifest as a lump in the breast, not all do. However, a lump in the breast may not necessarily be attributed to cancer, they made be benign. Benign breast tumors

are abnormal growths, but they do not spread outside of the breast and they are not life-threatening but some benign breast lumps can increase a woman's risk of developing breast cancer. Any breast lump or change needs to be checked by a healthcare provider to determine whether it is benign or malignant, and whether it might impact patient's future cancer risk (Oesterreich and Fuqua, 1999).

The traditional way of checking breast cancer is by self-examination of the breast. This only gives the patient an idea of when the breast has developed a lump or glaring signs. Some women do not even know how to properly carry out this examination while some do not notice the changes occurring. Although several treatment methods exist, it is said that early detection would require less invasive treatment techniques and engender a higher survival rate but patients with the disease hardly detect it early enough (Fisher *et al.*, 1975; Oesterreich and Fuqua, 1999). Breast cancer is diagnosed in various ways. Some popular machines used for its diagnosis include mammogram, ultrasound, MRI, biopsy and laboratory tests (Fisher *et al.*, 2002; Hankare, 2006). Handy devices used for detecting BC include: chonobra by Hugh Simpson, the high-tech bra, the Braster's by Braster science team, the smart bra by First Warning Systems and molecular breast imaging (MBI). The above listed devices use

*Corresponding Author Email: waleisiro@gmail.com; Tel: +2348034488877

temperature difference as their indicator of breast cancer, this has been helpful to some extent but it has been observed that there are other factors such as the weather and individual's state of health that can cause changes in temperature of the body. Thus, some of the aforementioned devices may report false positive results (Sarkar and Mandal, 2011; Selvarasu *et al.*, 2012). Some of these devices have been criticized for their invasive means of diagnosis while others make use of X-ray and Ultraviolet rays in detecting the breast tumour. All ultraviolet frequencies have been classified as Group 1 carcinogens by the World Health Organization (WHO) as ultraviolet radiation from sun exposure is the primary cause of skin cancer (WHO, 2017). Ultraviolet rays, X-rays and gamma rays, are referred to as ionizing radiation due to the ability of photons of these radiations to produce ions and free radicals in materials, living tissues inclusive (WHO, 2017).

Infrared (IR) thermography uses thermal imager to capture the IR radiation and measures the heat pattern of the object surface, human skin inclusive. It neither emit harmful radiation nor subject the patient to any risk (Egorov and Sarvazyan, 2008). To this end, thermography is gaining prominence as a physiological test for normal and abnormal physiologic functioning of the nervous, vascular and muscular systems for imaging local inflammatory processes compared to anatomical tests such as mammography. While IR radiation is invisible to human eye, it can be detected and displayed by special IR cameras. Thermal imager is operated with computer software for the interpretation of the IR images. The test is ideal for detecting hot and cold spots or areas of different emissivity on the skin surface since humans radiate IR energy very efficiently. Normal tissue that is non-cancerous has a blood supply automatically controlled by the central nervous system. For a cell to become cancerous, its surrounding tissues start to create new blood vessels (Mohamed, *et al.*, 2014). To sustain the rapid growth of these precancerous and cancerous cells a constant supply of nutrients are needed. In order to maintain this supply, the cancerous cells release chemicals (Chemokines) into the surrounding area, which keep existing blood vessels open, awakening dormant ones and creating new ones in a process known as

angiogenesis (Kapoor *et al.*, 2012). However, infrared thermography has the ability to detect the temperature or more importantly, actually map out the hot spots associated with chemical and angiogenic vessels both in precancerous as well as the cancerous breast tissue (Hankare, 2006; Lodish *et al.*, 2000). Consequently, thermography can be a first indicator of the pathogenesis of cancer since in many cases it takes between 4 and 10 years before it can be detected by any other method, mammography inclusive (Shah *et al.*, 2011).

This study proposes a solution to the aforementioned problem using the anomaly detection algorithm. Considering emissions from the breast as black-body radiation, the algorithm is designed to read a thermogram and compare its features to those of standard breast thermograms based on a classifier. Black-body radiation refers to the thermal electromagnetic radiation within or surrounding a body in thermodynamic equilibrium with its environment, or those emitted by an opaque and non-reflective body called the black body (Brar *et al.*, 2015). It has a specific spectrum and intensity that depends only on the body's temperature. In other words, all matters with positive absolute temperature emit thermal radiation. In infrared thermography, images are captured using an infrared camera. Irregular heat patterns in the breast may result from environmental factors other than angiogenesis therefore, some standard conditions are imposed before thermograms are taken to avoid errors. The thermograms used in this work were acquired from Visual Lab Database for Mastology Research (DMR): an open source database. And the thermograms underwent image preprocessing to improve the quality of the images before the computational processing. The purpose was to filter noise from images: normalize the intensity of individual particles of the image, enhance edges and remove blurriness to mention but a few. Image preprocessing improves the quality of the features extracted in an image and the results of image analysis. Breast IR thermograms usually capture larger areas than the required. Therefore, it is necessary to dress the thermogram images before further processing. Thereafter, the feature images obtained from normal thermograms are used to train the classifier. These processes are as illustrated in Figure 2.

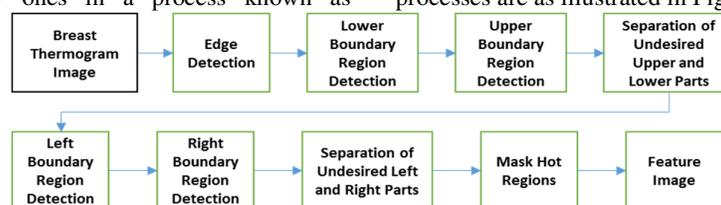


Fig 1: Workflow for the pre-processing images of the breast

Edge Detection: Edges are sudden changes of discontinuities in images. Edge detection helps remove less relevant information in the image while preserving the important content of the image. Edges contain meaningful features and relevant information of images. In this study, Canny edge detection algorithm developed by John F. Canny in 1986 was used. It is a multistage algorithm that accepts gray scale images as input but produces an output image that shows the position of tracked intensity discontinuities. Edges detection is important because both the horizontal projection profile (HPP) and the vertical projection profile (VPP) require the upper and lower, and the left and right boundaries respectively for computational analyses. The projection profile of an image in a particular direction refers to the running sum of pixels in that direction. The algorithm was selected because the gradient operator works on both x and y axis while some algorithms work in just one direction

Lower Boundary Region Detection (Inframammary Fold): The inframammary fold is the angle of deflection where the breast tissue meets the chest wall below the breast (Pease, Jr., 1996). It is the region with the most edges. For inframammary fold, the lower region was detected by increasing the count of white pixels. Row-wise scanning is done continuously until a value greater than the HPP is obtained. The row number with value corresponding to the first high HPP was used as the lower limit for the segmentation of the breast thermogram image.

Upper Boundary Region Detection: The breast height was calculated here by standardizing image heights as the height of the image varies depending on the structure and size of the breast. From study and observation, the height of the breast was calculated as follows:

1. If the distance between the bottom of the image and lower limit of the breast is less than 100 pixels, then $h = 2/3m$ where h is the height of the image and m is the total number of rows present in the image.
2. If the distance between the bottom of the image and lower limit of the breast is greater than 100 pixels then $h = 1/2m$

Therefore, the upper region of the breast is the breast height subtracted from the position of the breast.

Crop the Image: The image is cropped leaving the region bounded between the detected lower and upper bounds.

Left and Right Boundary Detection: After the upper and lower regions were detected and image cropped, the left and right boundaries were detected using

Vertical Projection Profile (VPP). VPP counts the number of white pixels in a column of the edge image and detects each boundary by the sudden increase in number of white pixels.

Crop the Image from VVP: The image is cropped leaving the region bounded between the detected left and right bounds.

Determine the spectrum for regions of higher temperature: The breasts do not generate much heat on their own. Healthy breasts appear purple or blue during a thermographic examination. Red, orange or yellow structures that appear in the breast during a thermographic exam may indicate the presence of malignant cells. Image thresholding was performed on the image. The simplest property that pixels in a region can share is intensity. So, a natural way to segment such regions is through thresholding, the separation of light and dark regions. Thresholding creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one. A threshold is set using the HSV (Hue Saturation Value) color space because producing filters based on RGB (Red Green Blue) values would be a bit more difficult. The color boundary for the hue are defined and used to mask the image.

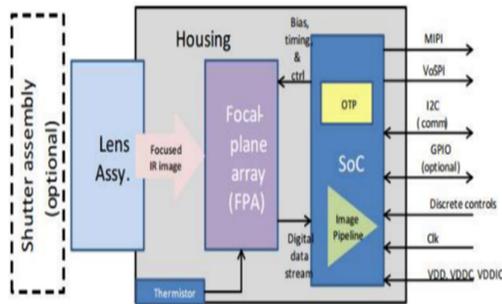
Mask hot regions of extracted region of interest: Masking here means identifying only the hot or cold regions; regions between within the threshold set. It involves coloring pixels that lie within the specified spectrum white and any pixel not in the spectrum black. Hence, any hot region displayed is shown as a white pixel.

The resulting feature thermograms then form the database for the training algorithm of the device's software which outcome is the probability of future occurrence of breast cancer. According to Rafferty et al (2017), it has been proven that more than 90% of women diagnosed with breast cancer at their earliest stage survive the disease for at least 5 years as against 15% for women diagnosed with the most advanced stage of the disease. Of concern in this study therefore, is the design of a safe, easy to use device that detects breast cancer in its early stages even before a lump is formed based on infrared thermography.

MATERIALS AND METHODS

Materials: The design of the hardware is based on the open source DIY – Thermocam V2 from 2016 by Max Ritter. The DIY – Thermocam is a do-it-yourself infrared camera based on the FLIR Lepton long-wave infrared array sensor. To assemble the device the following tools are needed: a soldering iron, some

solder, a cutting pliers, a nipper, a screwdriver and a multimeter. The subcomponents are Microcontroller, Visual Camera, FLIR Lepton, Spot Sensor (MLX90614), display (connected over 40-pin header), 5v booster, lithium battery charger, coin cell battery, SD card and several resistors used as pull – ups for 12C bus and voltage dividers for the lithium battery gauge. All the aforementioned components are arranged as demonstrated in Figure 2.



Source: zdnnet.com
 Fig 2: Schematics of the device

Methods: This study revolves around eight processes executed in three stages, namely: extraction of region of interest, masking of hot regions, and, training of location detectors and filtering warm region noise.

Region of Interest Extraction: This step helps to extract regions to focus on the breast region needed for analysis in the thermogram while discarding irrelevant parts. The processes involved are: edge detection, Boundary Region Detection, image cropping and

Mask Hot Regions: This step involved converting the region of interest already extracted to a mask that highlights the regions with higher temperatures. This is necessitated by the intrinsic thermal white noise (low frequency noise) arising from heat exchange fluctuations. Thus, filtering was used as a preprocessing step to reduce the noise. It involve two processes namely: determining the spectrum for regions of higher temperature and the masking of the hot region.

Training Location Detectors and Filtering Warm Region Noise: This is because thermograms usually have hot regions in the inframammary fold, neck and armpit regions which do not necessarily indicate cancer. Warm region noise needs to be filtered out from these areas. Hence, region detectors are trained for these regions and normal data is needed to do this. After which the processed data are fed into the training algorithm.

Training Algorithm: The features extracted from the previous stage form the criteria for the training the algorithm for the classification of breasts as normal or cancerous. It involves feature extraction from pre-processed images. For efficiency, there must be sufficient number of images for the training of the algorithm in two sets: the training and the test data. Traditionally, the ratio of training to test thermograms is 7:3 of the population of selected thermograms with the number of normal breast thermograms overwhelmingly outweighing that of cancerous breasts (Ajibola *et al*, 2011; Ibiwoye *et al*, 2012). The anomaly or outlier detection algorithm is trained in the same manner an artificial neural network is trained using the Gaussian distribution function to identify thermograms that do not conform to an expected dataset. The same algorithm can be adapted to other forms of medical diagnosis, fraud detection, manufacturing, structural defect and many other such events.

Supervised anomaly detection methods require labelling dataset usually as “normal” and “abnormal” when training the classifier. In anomaly detection, we have a dataset which contains normal data used as reference point to detect anomalous events. Using the training dataset, a model is built and can be accessed using $p(x)$ which defines the probability that the data x is normal. After building the model, a threshold value is set depending on the outcome of the training of our algorithm with thermograms from normal breasts.

Given training set of m events $\{x^1, x^2, \dots, x^m\}$, where each event is a feature vector. The probability $p(x)$ was modeled from the dataset such that

$$p(x) = p(x_1; \mu_1, \sigma_1^2) * p(x_2; \mu_2, \sigma_2^2) * \dots * p(x_n; \mu_n, \sigma_n^2)$$

The probabilities of features were multiplied since features are known to conjoin. Given that each of the features is normally distributed, the probability of each feature x_i is given by:

$$p(x_i; \mu_i, \sigma_i^2) = \frac{1}{\sigma\sqrt{(2\pi)}} \int^{x_i} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$$

The following algorithm is then deployed to simulate biometric indices of the breast:

1. Choose features x_i that is indicative of anomalous events
2. Calculate fit parameters $\mu_1, \dots, \mu_n, \sigma_1^2, \dots, \sigma_n^2$, the standard deviation, σ , and the mean μ , for each feature x
3. For a new events x , compute $p(x)$ such
4. Set a threshold, ξ , for all events such that event is anomaly if $p(x) > \xi$

RESULTS AND DISCUSSION

This study investigated early detection of breast cancer as panacea for cancer induced deaths especially in women diagnosed with the disease. It identified IR thermography as a safe method of diagnosis of the disease and designed a modality for early detection of breast cancer with the ultimate goal of annihilating cancer related deaths. The essence of the device’s user interface is to enable the user interacts with the device. It contains descriptive icons designed to assist intended users to navigate through the application. The default interface of the application consists of a browse image button, process button, result and recommendation screen and an output button to display a graphical outcome. Figure 3 is the workflow of the graphical user interface (GUI).



Fig 3: Workflow of GUI

The software evaluates and collates the probability of manifestation of breast cancer in each user and plots the graph: usually a Gaussian normal curve. The process of assigning probability to a new thermogram to predict an occurrence of breast cancer involves computational analysis of the stochastic variate involved based on Gaussian density function that culminate in the display of the outcomes in GUI of Figure 4 or 5 for normal and cancerous breast thermal images respectively.

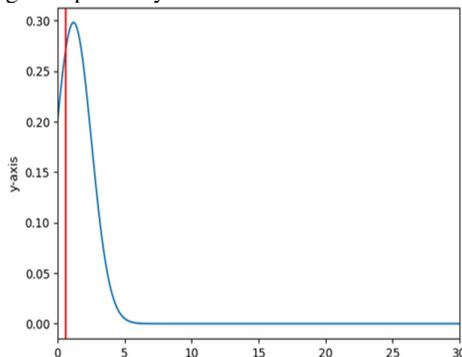


Fig 4: Gaussian graph representing normal breast thermogram

It should be recalled that the number of normal breast thermograms used to derive our Gaussian normal curve overwhelmingly outweighed those of cancerous breasts; good reason why the curve is skewed. Therefore, the threshold, representing the smallest detectable sensation as depicted by the vertical red line is biased towards normal breast thermograms. It was observed that the thresholds obtained from a set of labelled normal thermograms have very small standard deviations from the mean (Figure 4) while the standard

deviations for those from a set of abnormal thermograms are very far off from the mean (Figure 5) which interprets to a cancer prone breast.

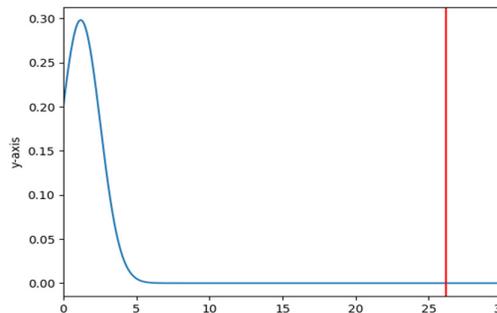


Fig 5: Gaussian graph representing cancerous breast thermogram

The outcomes are very vital for early diagnosis of cancer. For treatment of cancer, especially breast cancer to be realistic more importantly in Africans, it must be detected at its formative stage (Kakushadze *et al.*, 2017; Weinberg, 2016). And that is the focus of this study. Earlier works dwelt on diagnosis of breast cancer after lump(s) have been formed (Ali, *et al.*, 2014; Amy, 2010; Egorov and Sarvazyan2008; Fisherm *et al.*, 1975, 2002)

This paper proposes a novel diagnostic paradigm as a springboard for the management of breast cancer. According to Kakushadze *et al.*, (2017), early diagnosis which translates to cost effectiveness is a key factor in facilitating the treatment of the disease especially in a poverty stricken continent as Africa. This work seeks to provide some succour for the victims of cancer worldwide irrespective of their social status and financial capabilities.

Conclusion: In this study, we have deployed infrared thermography as an evolving trend for detection of breast cancer in its formative stage. We have also developed a software with a classifier with the capacity to identify and analyse temperature difference as a parameter for evaluating the probability of occurrence of breast cancer. The software developed evaluates the probability of occurrence of as a precursor for prediction of likelihood of occurrence of breast cancer in the near future. It also recommends action to be taken based on the result of the stochastic process. A medical expert will always be required so as to determine whether further examination(s) should be conducted to facilitate early treatment of the condition. We hope our algorithm and consequently the device will provide a succor to the teeming population of Nigerians who desire to know their status before the manifestation of breast tumor.

REFERENCES

- Ajibola, OOE; Olunloyo, VOS; Ibidapo-Obe, O (2011): Artificial Neural Network Simulation of the Arm Gait of Huntington's Disease Patient. *Int. J. Biomech. Biomed. Robotics*, 1(3): 133–140.
- Ali, S; Sethi S; Ali, AS; Philip, PA; Sarkar, FH (2014). Molecular Pathogenesis of Breast Cancer and the Role of MicroRNAs. In *miRNAs and Target Genes in Breast Cancer Metastasis*. Springer. pp. 7–22.
- American cancer society (2015). Non-cancerous breast condition 2015. Retrieved from <https://www.cancer.org/cancer/breast-cancer.html>.
- American cancer society (2017). Cancer fact and figures 2017. Retrieved from <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2017.html>.
- Amy, D (2010). Lobar ultrasound of the breast. In *Breast cancer*. Springer. pp. 153–162
- Bartella, L; Morris, EA; Dershaw, DD; Liberman, L; Thakur, SB; Moskowitz, C; Huang, W; Jennifer G (2006). Proton MR spectroscopy with choline peak as malignancy marker improves positive predictive value for breast cancer diagnosis: preliminary study. *Radiol*, 239(3): 686–692.
- Bhojal, J; Kennedy, JF (2005). *Essential Cell Biology Volume 2: Cell Function*. John Davey, Mike Lord (Eds.). Oxford University Press, Oxford, UK.
- Brar, VW; Sherrott, MC; Jang, M.S; Kim, S; Kim, L; Choi, M; Sweatlock, LA; Atwate, HA (2015). Electronic modulation of infrared radiation in graphene plasmonic resonators. *Nature Communications*. Pp. 1 – 7
- Egorov, V; Sarvazyan, AP (2008). Mechanical imaging of the breast. *IEEE Trans. Med. Imag.* 27(9): 1275–1287.
- Fernandez, PL; Jares, P; Rey, MJ; Campo, E; Cardesa, A (1998). Cell cycle regulators and their abnormalities in breast cancer. *Molecular Pathology*, 51(6), 305.
- Fisher, B; Anderson, S; Bryant, J; Margolese, RG; Deutsch, M; Fisher, ER; Jeong JH; Wolmark, N (2002). Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *New England Journal of Medicine*, 347(16): 1233–1241.
- Fisher, ER; Gregorio, RM; Fisher, B; Vellios, F; Sommers, SC (1975). The pathology of invasive breast cancer a syllabus derived from findings of the National Surgical Adjuvant Breast Project (Protocol no. 4). *Cancer*, 36(1), 1–85.
- Hankare, P; Shah, K; Nair, D; Nair, D (2016). Breast cancer detection using thermography. *Int. Res. J. Eng. Technol*, 4(3): 2356–2395.
- Ibiwoye, A; Ajibola, OOE; Sogunro, AB (2012). Artificial Neural Network Model for predicting Insolvency in Insurance Industry. *Int. J. Mgt. Bus. Res.* 2(1): 33-36.
- Kakushadze, Z; Raghubanshi, R; Yu, W (2017). Estimating Cost Savings from Early Cancer Diagnosis. *Data*. 30(2): 1 - 16.
- Kapoor, P; Prasad, S; Patni, S (2012). Automatic analysis of breast thermograms for tumor detection based on biostatistical feature extraction and ANN. *International Journal of Emerging Trends in Engineering and Development*, 2(7): 245–255.
- Kapoor, P; Bhayana E; Prasad SVAV (2010). Real time intelligent thermal analysis approach for early diagnosis of breast cancer.
- Kamath, D; Kamath, S; Prasad, K; Rajagopal, KV (2015). Segmentation of Breast Thermogram Images for the Detection of Breast Cancer--A Projection Profile Approach. *Journal of Image and Graphics*, 3(1): 47 - 51.
- Lodish, H; Berk, A; Zipursky, SL; Matsudaira, P; Baltimore, D; Darnell, J (2000). *Molecular cell biology 4th edition*. *National Center for Biotechnology Information's Bookshelf*.
- Mohamed, MM; Al-Raawi, D; Sabet, SF; El-Shinawi, M (2014). Inflammatory breast cancer: New factors contribute to disease etiology: A review. *J. Adv. Res.* 5: 525–536
- Oesterreich, S. and Fuqua, S. A. (1999). Tumor suppressor genes in breast cancer. *Endocrine-Related Cancer*, 6(3): 405–419.
- National Cancer Institute (2017), *Cancer of the Breast (Female): Cancer Stat Facts*. NCI Surveillance, epidemiology and results program. NCI's

- Division of Cancer Control and Population Sciences. <https://cancercontrol.cancer.gov/>.
- Pease, Jr. RW (1996). Merriam Webster's Medical Dictionary. Merriam Webster, Incorporated, USA.
- Rafferty, EA; Rose, EL; Miller, DP; Durand, MA; Conant, EF; Copit, DS; Friedewald, SM; Plecha, IL; Hayes, OMK; Carlson, KL; Cink, TM; Barke, LD; Greer, LN; Niklason, LT (2017). Effect of age on breast cancer screening using tomosynthesis with digital mammography. *Breast Cancer Research and Treatment*. 164,(3): 659 - 666.
- Sarkar, S; Mandal, M (2011). Breast Cancer: Classification Based on Molecular Etiology Influencing Prognosis and Prediction. In *Breast Cancer-Focusing Tumor Microenvironment, Stem cells and Metastasis*. InTech. USA.
- Selvarasu, N; Nachaippan, A; Nandhitha, NM (2012). Image processing techniques and neural networks for automated cancer analysis from breast thermographs-A review. *Indian J. Comp. Sci. Engr.* 3(1): 133–137.
- Shah, TH; Siores, E; Daskalakis, C (2011). Non-Invasive Devices for Early Detection of Breast Tissue Oncological Abnormalities Using Microwave Radio Thermometry. In *Adv. Cancer Therapy*. InTech. USA.
- Weinberg, R (2016). *The biology of cancer*. Garland science. Taylor and Francis Group, UK.