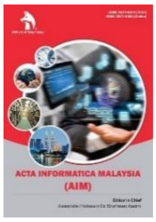
ZIBELINE INTERNATIONAL™ PUBLISHING

ISSN: 2521-0874 (Print)
ISSN: 2521-0505 (Online)
CODEN: AIMCCO

Acta Informatica Malaysia (AIM)

DOI: http://doi.org/10.26480/aim.01.2024.26.31



CrossMark

REVIEW ARTICLE

THERMAL IMAGES PRE-PROCESSING FOR EARLY DETECTION OF BREAST CANCER: A PROGRESSIVE REVIEW

Azman Brahim Llaguno Garcia[a]*, Arturo Orellana Garcia[b], Vivian Estrada Senti[c]

Informatics, Networks and Educative Technology Group, Medical Sciences Branch, Medical Sciences University, 21 Street number 32-B between 26 and 28 Streets, 26th July District, Sagua la Grande, Villa Clara, Cuba.
Medical Informatics Center, Process Mining Research Group, Informatics Science University, San Antonio de los Baños Road, Kilometers, Torrens District, Boyeros, Havana, Cuba.
Department of Scientific Degrees, Informatics Science University, San Antonio de los Baños Road, Kilometers 2 ½, Torrens District, Boyeros, Havana, Cuba.
Corresponding Author Email: azmanllg@gmail.com, azmanllg@infomed.sld.cu

ARTICLE DETAILS

Article History:

Received 18 January 2024
Revised 21 February 2024
Accepted 24 March 2024
Available online 28 March 2024

ABSTRACT

The aim of this study is to identify the pre-processing stage as a necessary step in the thermal image processing to support the decision-making in health institutions. Digital infrared thermal imaging allows the identification of unrecognized diseases in an apparently healthy population, through a thermal map generated by the temperature of the affected area. The methods used to identify the main aspects that affect the quality of the thermal image were inductive-deductive, historical-logical, analytical-synthetic and document analysis. As a result, the necessary initial conditions for the acquisition and the logical sequence to be develop in the pre-processing stage were presented.

KEYWORDS

computer diagnosis, computer aided diagnosis, image diagnosis, thermography, acquisition stage, pre-processing stage.

1. Introduction

Information and communication technologies have reached all spheres of society and revolutionized dissimilar areas of knowledge, within which was found from the outset, medicine, which has been defined as: "Conjunto de conocimientos y técnicas aplicados a la predicción, prevención, diagnóstico y tratamiento de las enfermedades humanas y, en su caso, a la rehabilitación de las secuelas que puedan producir" [A set of knowledge and techniques applied to the prediction, prevention, diagnosis and treatment of human diseases and, where appropriate, the rehabilitation of the aftermath they may produce] (Real Academia Española, 2023).

Among the main health problems affecting the population today is cancer, in which according to data presented by the International Agency for Research on Cancer a total of 19 292 789 new cases in the world were foreseen for both sexes in 2020, of which a total of 9 958 133 deaths were estimated, with a prevalence of 50 550 287 cases in the last five years; Table 1 elaborated from GLOBOCAN reflects the situation in different regions of the world (GLOBOCAN, 2023a; GLOBOCAN, 2023b; GLOBOCAN, 2023c.

Table 1: Global distribution of cancer cases by region.

Region	2020			2040	
	Incidence	Prevalence (last 5 years)	Mortality	Estimated Incidence	Estimated mortality
Asia	9 503 710	20 606 063	5 809 431	**28 887 940**	**16 180 202**
Europe	4 398 443	13 496 763	1 955 231		
North America	2 556 862	9 456 199	713 414		
Latin America and the Caribbean	1 470 274	3 837 718	711 429		
Africa	1 109 209	2 166 740	699 274		
Oceania	254 291	986 804	69 354		
Total	**19 292 789**	**50 550 287**	**9 958 133**		

Quick Response Code	Access this article online	
	Website: www.actainformaticamalaysia.com	**DOI:** 10.26480/aim.01.2024.26.31

Cite The Article: Azman Brahim Llaguno Garcia, Arturo Orellana Garcia, Vivian Estrada Senti (2024). Thermal Images Pre-Processing for Early Detection of Breast Cancer: A Progressive Review. *Acta Informatica Malaysia, 8(1): 26-31.*

Among the different types of cancer is breast cancer, which commonly develops in women, although there are also cases reported in men; a total of 2 261 419 new cases are expected worldwide and in the female sex by 2020, of which 684 996 deaths were estimated, with a prevalence of 7 790 717 cases in the last 5 years; Table 2 elaborated from GLOBOCAN reflects the situation in different regions of the world in relation to this pathological condition (GLOBOCAN, 2023e, GLOBOCAN, 2023f; GLOBOCAN, 2023d).

Table 2: Global distribution of female breast cancer cases by region.

Region	2020			2040	
	Incidence	**Prevalence (last 5 years)**	**Mortality**	**Estimated Incidence**	**Estimated mortality**
Asia	1 026 171	3 218 496	346 009	**3 025 471**	**1 037 723**
Europe	531 086	2 138 117	141 765		
North America	281 591	1 189 111	85 787		
Latin America and the Caribbean	210 100	710 039	57 984		
Africa	186 598	429 220	48 407		
Oceania	25 873	105 734	5 044		
Total	**2 261 419**	**7 790 717**	**684 996**		

Source: Own elaboration.

Breast cancer was the leading cause of female death worldwide by 2020, an approach derived from the analysis of the data in Figure 1.

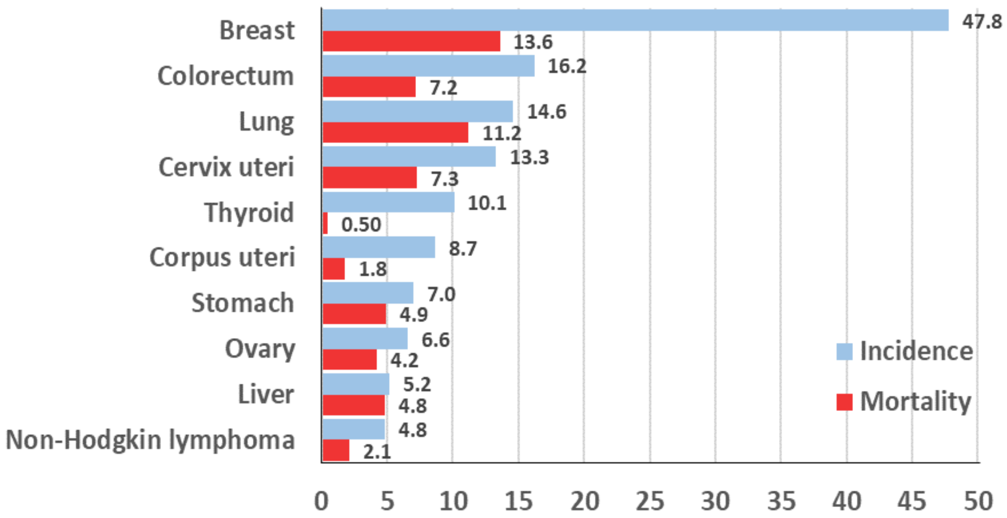


Figure 1: Incidence and mortality rates per 100 000 inhabitants in women with cancer by 2020. Source: GLOBOCAN (GLOBOCAN, 2023g).

The incidence and mortality rates of breast cancer presented above reflect the need to continue researching techniques based on early detection to locate this pathology in its initial stages, which regardless of the progress made, continues to be a necessity, due to the fact that in countries that diagnose tumors using these techniques, such as Spain, 35 001 new cases of breast cancer were estimated in both sexes in 2023 and 6 614 cases died with this pathology in both sexes in 2021, while in the United States a total of 300 590 new cases were predicted for 2023, of which 43 170 deaths were estimated (Sociedad Española de Oncología Médica, 2023; Sociedad Española de Oncología Médica, 2023; Seigel et al., 2023). Cuba, a country that advances in early detection through CAD, in 2021, tumors constituted the second cause of death with a total of 26 791 deaths for a gross rate of 239.5 per 100 000 inhabitants, which, with respect to 2020, represented an increase, since it had a total of 26 289 deaths for a gross rate of 234.7 per 100 000 inhabitants (Dirección de Registros Médicos y Estadísticas de Salud, 2022).

2. Methodology

The documentary review process in the present research consists of seven steps:

- Collection of research articles related to and based on DITI for early detection of breast cancer using CAD.
- Explicit investigation of the methods and approaches of the collected research articles.
- Description of the steps involved in DITI processing.
- Discussion of advantages and disadvantages of the methods used in the stages of DITI processing.
- Analysis of the acquisition stage, description and identification of the necessary capture protocols to be taken into account.
- In-depth analysis of the pre-processing stage, shortcomings and existing needs.
- Logical sequence of steps in future methods and algorithms to be developed.

An exhaustive search was performed to identify the main related research articles. The main scientific journals, databases, doctoral thesis repositories and conference paper presentation consulted were GLOBOCAN of the International Agency for Research on Cancer, CA: Cancer Journal for Clinicians, Spanish Society of Medical Oncology; World Health Organization; Journal of Healthcare Engineering; Journal of Engineering Science; Thermal Science and Engineering Progress; PLoS ONE; Sensors; Quantitative Thermography Journal; BioMed Research International; Computers in Biology and Medicine; Research in Computing Science; Research on Biomedical Engineering; BioMedical Engineering Online; Springer; Journal of Image and Graphics; Journal of Medical Imaging and Health Informatics; IEEE Transactions on Systems, Man and Cybernetics; Teaching and Research Digital Archive from the University of the Basque Country; Doctorals Thesis on Xarta; International Conference on Innovation in Information, Embeddeds and Communication Systems is being explored with some keyword like ['Digital Infrared Thermal Image' | 'DITI'], ['Pattern Recognition Technique' | 'Infrared Thermal Image Processing'], ['Artificial Intelligence Technique' | 'Infrared Thermal Image Processing'], ['Early diagnosis' | 'Breast Cancer'], ['Computer aided diagnosis' | 'Breast Cancer'], etc.

Through the use of the keywords described above in the exploration of scientific journals, databases, doctoral thesis repositories and conference paper, a total of 68 articles are collected, from which 42 are selected for the initial study selected for the criteria of:

- The research focuses on identifying the main problems affecting the identification of breast cancer in early stages through the use of digital infrared thermal imaging.
- The researchers identify the main methods, techniques and algorithms used in the different stages of infrared thermal image processing.
- The investigators identify the potential benefits of developing methods and algorithms at the pre-processing stage of digital infrared thermal imaging based on pattern recognition techniques.
- The research shows a logical sequence of steps to keep in mind in the development of methods and algorithms for the pre-processing stage of infrared thermal imaging.

3. Document Analysis And Literature Review

3.1 Digital infrared thermal imaging in early detection of breast cancer

Early detection is the presumptive identification of an unrecognized disease in an apparently healthy and asymptomatic population. Its application increases the likelihood of successful treatment by focusing on the detection of symptomatic patients as soon as possible; and it is composed of three stages as shown in Figure 2 (World Health Organization, 2023; World Health Organization, 2017).

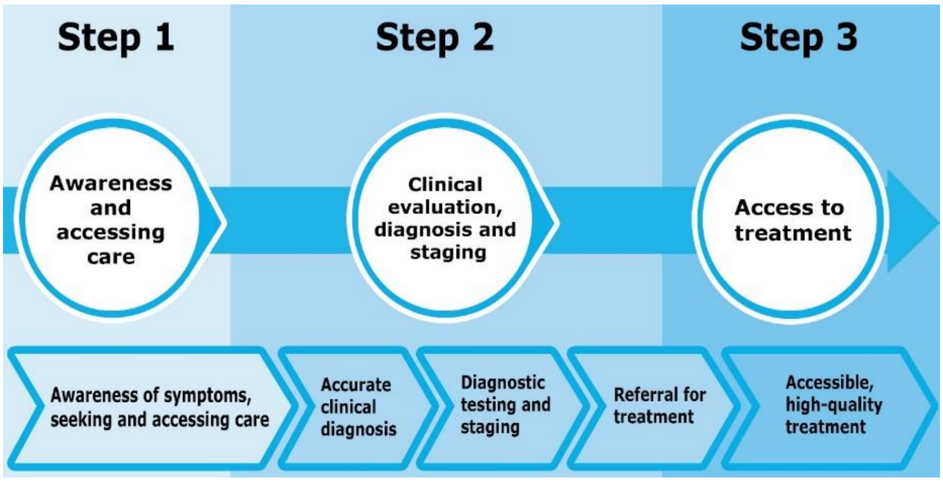


Figure 2: Stages of early detection. Source: World Health Organization (Word Health Organization, 2017).

Within the methods for early detection of breast cancer, are: self-exploration, mammography, ultrasound, magnetic resonance imaging and thermal imaging; being mammography the reference method, which by means of X-rays generates a high contrast and resolution image; however, it presents as limiting factors, discomfort in the patient due to breast compression, affection in the sensitivity of the images obtained by breast density, less effectiveness in women under 50 years of age and continuous exposure to radiation that can damage the cancerous tissue and cause its mutation (Pavithra et al., 2018; Santana et al., 2018; Dey et al., 2017; Wakankar et al., 2016; Mammoottil et al., 2022; Mashekova et al., 2022; Mohamed et al., 2022; Tsietso et al., 2022; Teja Kakileti et al., 2019; Hossam et al., 2018).

The development and research in early detection techniques such as DITI directly linked to the detection of breast cancer, is essential to avoid the unnecessary diagnostic indication of a mammography, in addition to requiring less time of cancer evolution for detection, not irradiate the patient, and do not require a minimum volume of breast to fix the area under study and acquire the image, because its use is based on identification of specific features of breast heat patterns, like, highly asymmetric temperature distributions between the left and right breasts, localized hot spots indicative of abnormalities, changes in hypothermic vascular patterns due to tumor growth and variation in the areolar and periareolar regions (Mammoottil et al., 2022; Mashekova et al., 2022; Mohamed et al., 2022; Tsietso et al., 2022; Rodríguez Guerrero et al., 2014).

3.2 Computer aided diagnosis

DITI, is the most efficient technique for the study of temperature distribution in the skin, since it provides information on the normal or abnormal functioning of the sensory and sympathetic nervous system, vascular dysfunction, myofascial trauma and local inflammatory processes product to which all body or organ emits infrared radiation due to its thermal condition (Mammoottil et al., 2022; Mashekova et al., 2022; Mohamed et al., 2022; Tsietso et al., 2022; Dey et al., 2017; Kamath et al., 2015).

This technique presents as output a digital medical image that constitutes a representative source of information, because it is captured and used by specialists to obtain information on the health status of patients, constituting a function of two dimensions f(x,y) where x and y are spatial coordinates and the amplitude of f in each pair of coordinates (x,y) represents the level of intensity, whose values depend on constituting an image in black and white, grayscale or color; for black and white images f(x,y) has a binary value where 0 is black and 1 is white; for grayscale images f(x,y) can take values from 0 (black) to 255 (white) with all intermediate values; while for color images f(x,y) is a vector of 3 individual components for red, green, and blue, with values from 0 to 255 for each component (Tsietso et al., 2022; Pavithra *et al.*, 2018; Al-shamasneh *et al.*, 2017).

Beside its advantages, it has shortcomings in the analysis carried out by specialists, since quality depends on experience, visual acuity and knowledge of the characteristics of the object to be analyzed; and the identification of anomalies is limited by the different elements present in the surface under study; which place medical personnel before difficulties in the identification of pathologies in complex backgrounds associated with image quality, where relevant features, visualization environments and experience can cause erroneous results. Approach according to the criteria issued by (Mammoottil et al., 2022) "[...] Earlier thermal images were analyzed by humans making the process strenuous and inaccurate. With the emergence of artificial intelligence (AI) and machine learning (ML) algorithms, thermal images can be intercepted and used in a whole new different way. [...]" (p.2),, the criteria issued by (Roa Martínez *et al.*, 2016): "El diagnóstico médico a partir de imágenes médicas digitales en numerosas ocasiones tiene una gran influencia subjetiva, porque generalmente se basa en la extracción de una determinada información o elementos simples sobre fondos complejos" [The medical diagnosis from digital medical images often has a great subjective influence, because it is generally based on the extraction of a particular information or simple elements on complex backgrounds] (p.652) and to the criteria issued by (Resmini *et al.*, 2021): "[...], human interpretation of those images is complex and, for this reason, artificial intelligence algorithms must be used to help the medical team to identify the differences between sick and healthy breast patterns." (p.1). For this reason, DITI needs to be analyzed and processed more efficiently using digital image processing techniques that allow the extraction of the main features for an accurate diagnosis of breast cancer patients (Mammottil et al., 2022; Mohamed et al., 2022; Padmappriya et al., 2018; Dey et al., 2017).

CAD is the procedure performed using computational algorithms like digital image processing, artificial intelligence, personalized data and physical modeling and machine detection of temperature distribution patterns based on medical knowledge of specific diseases created to identify suspicious regions of malignancy on DITI and achieve an accurate clinical diagnosis, for this reason the are many studies on its improvement as a prognostic adjunctive tool for early detection of breast cancer, and the trend for future research seems to be in the performance of the CAD system for identify a correct highest number of cases, for its advantages according to the criteria emitted "[...] advantages of CAD systems are: 1) it is less dependent on human subjective opinion; 2) it is a quantitative based system that can reduce false positives or false negatives; 3) it reduces the cost of additional medical procedures [...]" (Mashekova et al., 2022; Huang et al., 2018; Jean Mambou et al., 2018; Liu et al., 2018).

Interpretation of DITI using CAD systems requires three steps (Mashekova et al., 2022; Huang et al., 2018; Jean Mambou et al., 2018; Liu et al., 2018; Qin et al., 2018):

- Manual interpretation of the digital image by the specialist.
- Digital processing of the image by means of CAD, for the identification of the possible affected areas and to emit a final result classifying the image as pathological or not.
- Revaluation of the first digital medical image analysis result compared to the second. Aspect that constitutes a fundamental step for the correct performance of the CAD system, and of necessary presence, as an intermediate point that allows before the final result to re-evaluate a previous result, with the possibility that the system can correct the decision in case it is erroneous in correlation with the opinion of the specialist.

3.3 Digital infrared thermal image processing

The treatment of the medical image differs from the way in which images of natural scenes are acquired, a product that can present problems in the differentiation of internal organs and tissues represented by aspects such as noise, not available or fuzzy image, low contrast, incorrect intensity values, and ill-defined and blurred edges (Mammoottil et al., 2022; Tsietso et al., 2022; Bereciartua Pérez et al., 2016). This process is composed of five stages, acquisition, pre-processing, segmentation, feature extraction and classification (Mashekova et al., 2022; Tsietso et al., 2022; Hossam et al., 2018; Pavithra et al., 2018; Santana et al., 2018; Dey et al., 2017).

3.4 Acquisition

Several parameters need to be taken into account to detect the thermal contrast between cancerous and non-cancerous areas, this principle is based on the fact that the temperature patterns of the cancerous tissue in the breast remain the same, while the healthy regions are affected by cold stress. It is given by the use of a physical environment sensitive to changes in ambient temperature, atmospheric temperature, relative humidity, patient position, air currents and previous rest time, so it is necessary to acquire it under certain protocols, which have been proposed by researchers such as (Mashekova et al., 2022):

- A group research obtained a DITI from the free access database developed which proposed for the acquisition rest by patients two hours before the examination without alcohol, caffeine, nicotine, physical exercise, no application of cream, oil or any type of deodorant in areas of the breast or armpit; a controlled temperature in the room in a range of 20°C to 22°C, no windows, no wind, no direct air flow to the patient, and the use of only fluorescent bulbs in the lighting, need for the non-presence on the patient of earrings, necklaces or other accessory that may interfere with the correct taking of the thermal image, the check of body temperature with a clinical thermometer and hair gathered towards the upper area of the head; for image capture pre-cooling of the chest and axilla area with a fan until the thermal stress stop condition is met, distance between the camera and the patient of one meter, recording of relative air humidity and temperature for input as parameters in the camera used (Mammoottil et al., 2022 and Tello-Mijares et al., 2019; Silva et al., 2014). The clinical data of each patient that was recorded using web scraping.
- A group researchers used the online database for research in mastology with infrared imaging DMR-IR which was elaborated from images acquired under the protocols of prior rest of patients for 10 minutes in a room with controlled temperature between 20°C and 22°C and hands over the head, and the private database of the Federal University of Pernambuco UFPE which was elaborated from images acquired under the protocols of prior rest of patients for 10 minutes and presence of hands on the top of the head (Resmini et al., 2021).
- A group researchers proposed a controlled temperature in a range of 20 to 22 ± 0.1°C, a relative humidity of 60 ± 5%; rest in patients 15 minutes prior to the test, use of a wide gown to not restrict air flow, all within the recommended period on days 5 to 12 and after 21 of the menstrual cycle (Santana et al., 2018).
- In other study, authors proposed a temperature of 18 to 22°C, relative humidity of 60%, dark room to minimize interference from infrared sources, rest of 10 to 20 minutes in patients, no alcohol consumption, no physical exercise, no application of cosmetics on the surface of the breast on the day of acquisition of images; emissivity of 0.97 to 0.98 for an ambient atmospheric temperature of 17°C (Pérez et al., 2016).

3.5 Pre-processing

It consists of tackling one or more problems, including noise elimination, increasing contrast, emphasizing relevant edges and structures, detecting points of interest and uniform lighting. Some operations include converting RGB scale to Grayscale, since less information is needed for each pixel and allows RGB colors to contain equal intensity values. With respect to the techniques used at this stage there are different criteria issued (Tsietso et al., 2022; Al-shamasneh et al., 2017; Bereciartua Pérez, 2016):

- Some researchers remove from the dataset the patients that had not available or fuzzy image in the static protocol to replace them with the corresponding patients image from the dynamic protocol, remove image if there are blurry and barely visible, if have presence of injury, if proper protocols were not followed during the process of data collection or did not have all five views, remove data with any anomalous entries, and transform all images to a 640 x 640 size using the resize functionality provided by PyTorch (Mammoottil et al., 2022).
- in the input the thermal images of size 680 × 480 pixels and its computation time will be high due to the limitation of the PC capabilities used in this study (Mohamed et al., 2022). So, the thermal images are resized to a smaller size of 228 × 228 pixels for faster computation.
- attenuates distortions by removing noise through the use of a Gaussian filter (3x3) and contrast enhancement through the change of the image from RGB scale to grayscale which in turn is subsequently converted into binary format by using a predefined threshold of 0.25; defines the left and right margins using Canny's edge detection algorithm to constitute the initial elliptical points of the GVFS, the left and right margins or lines, and the left and right breast curvature function boundaries (Tello-Mijares et al., 2019).
- remove the noise present and then perform the conversion from RGB scale to grayscale (Hossam et al., 2018).
- convert the image to grayscale and extract the ROI belonging to the area of both breasts of the image with its subsequent conversion into a matrix of characteristics to be processed by moving the area most likely to present cancer to the entry of the next component (Jean Mambou et al., 2018).
- make adjustments from RGB scale to gray scale (Santana et al., 2018).
- perform gray level histogram filtering on the entire image to improve future steps in image processing and remove unnecessary backgrounds and regions with the extraction of the ROI (Dey et al., 2017).
- take steps to remove the bottom, change size and in turn remove the lower body (Wakankar et al., 2016).

3.6 Segmentation

It is given by the identification of multiple segments due to its greater significance and easy analysis, so it is necessary to examine pixel by pixel and label each of them, depending on whether the value of the gray level is greater or less than the threshold. At this stage, the images are divided into several groups of pixels to facilitate their representation in another image easier to analyze (Padmappriya et al., 2018; Mahmoud, 2016). In the breast cancer study, through the DITI, one of the most useful indicators for medical personnel is the comparison of temperatures in symmetrical anatomical regions, so the diagnosis can be based on automatically locating certain characteristics that indicate possible temperature asymmetries between the right and left regions of both breasts, identifying borders, or segmenting regions, lines or curves (Meena Prakash et al., 2017).In the investigation of procedures used at this stage, we have identified the use of Watershed techniques that allow us to extract the borders of the regions present in the image and to distinguish highly complex objects which cannot be easily processed by conventional algorithms; grouping using the K-means algorithm based on colour characteristics with extraction of the hot region in two steps, first grouping the pixels on the basis of their colour and spatial characteristics from which the grouping process is carried out, and a second step where the segmentation of the hot region is carried out separating the background from the image; Fuzzy C-means and Level Set techniques; Gaussian mix models; and Hough transform algorithm (Tello-Mijares et al., 2019; Meena Prakash et al., 2017; Wakankar et al., 2016).

3.7 Feature extraction

It is a key stage for the correct detection of pathologies, and depending on the type of characteristics obtained, its space can be divided into three categories (Tello-Mijares et al., 2019; Jean Mambou et al., 2018; Mahmoud, 2016; Pérez et al., 2016).

- Intensity Characteristics: Consists of using the gray level values of each ROI to characterize it.
- Geometric characteristics: It is based mainly on the shape of the ROI, thus calculating area, perimeter, convexity, moments of first, second and third order, obliquity, among others.
- Texture Characteristics: Obtained from co-occurrence matrices or gray level dependency matrices, and Haralick texture descriptors.

Among the characteristics to have present in this stage are asymmetry, kurtosis, entropy and area of the lesion; extracted as characteristics of textures by means of descriptors of geometric moments, Legendre, Hu, Zernike or Haralick texture from the co-occurrence matrix of the image, which constitute the entrance to the next stage. Characteristics extracted from low or high level can help discrimination of various regions; in turn the decrease in computational time depends on the selection of the most significant characteristics (Santana et al., 2018; Wakankar et al., 2016).

Of the techniques previously referred to, one of the most promising are the Haralick texture descriptors, developed in the 1970s by Robert M. Haralick, which analyze textures in an image by using the co-occurrence matrix of gray tones to calculate fourteen texture descriptors of a statistical nature; where for their calculation, it is necessary to assume that the totality of the textural information of an image is contained in the spatial relationships that occur between the different gray levels of an object (Haralick et al., 1973).

3.8 Classification

The techniques used in this stage allow, from a set of characteristics obtained from the region of segmented study, the assignment to one group or another, which is called class; this assignment, in the case of medical imaging can be the identification of a type of tissue or lesion; in addition, measurement and quantification operations can be added to the image to obtain more accurate data that are of value in the diagnosis (Bereciartua Pérez, 2016).

Classification techniques such as CNN, deep learning networks, fuzzy models, artificial neural networks, FNN, PNN, support vector machine, extreme learning machines, BPA, bayesian networks, GMM and decision trees; have been used for their high precision starting from the input parameters coming from previous stages (Mammoottil et al., 2022; Mashekova et al., 2022; Teja Kakileti et al., 2019; Hossam et al., 2018; Jean Mambou et al., 2018; Santana et al., 2018; Wakankar et al., 2016).

4. Analysis And Discussion

In the studies led by a method based on CNN to detection of breast cancer using five different views, front, left 45°, right 45°, left 90°, and right 90° were developed, it is found that the addition of clinical data increased the ability of the model to classify a patient as healthy or sick correctly (Mammoottil et al., 2022). However, the study and implementation of techniques for better and more robust pre-processing like data augmentation and image segmentation would be interesting, since the model would improve its interaction with the availability of better and less limited data.

In research presented by methods were developed in the stages of segmentation, feature extraction and DITI classification which reach breakpoints due to the lack of deep treatment in the pre-processing stage; there is a lack in the literature of procedures and techniques that provide better quality (Resmini et al., 2021; Dey et al., 2017; Mahmoud, 2016; Wakankar et al., 2016).

In studies conducted is not taken into account within the performance of the CAD systems proposed; the re-evaluation of the result achieved with the one issued by the specialist, which allow the feedback of the system in the cases classified in an erroneous way, in order to increase their precision and effectiveness in obtaining accurate results in future analyses of DITI (Resmini et al., 2021; Tello-Mijares et al., 2019; Teja Kakileti et al., 2019; Hossam et al., 2018; Jean Mambou et al., 2018; Sanchez-Ruiz et al., 2018; Santana et al., 2018; Bastida et al., 2017; Meena Prakash et al., 2017; Mahmoud, 2016; Pérez et al., 2016; Wakankar et al., 2016; Kamath et al., 2015 and Rodríguez Guerrero et al., 2014).

So, the authors of this article have identified the need for:

- Development of algorithms using advanced methods of region growth using contour techniques to improve the segmentation stage of regions in thermal imaging.
- Development of characterization algorithms using statistical correlation methods to analyze the texture of the ROI and improve comparison at the classification stage.
- Development of methods for image enhancement adjustments, noise reduction, distortion attenuation, image smoothing and ROI extraction in the pre-processing stage that provide a higher quality DITI for later stages of segmentation, feature extraction and classification.
- Proportion of volumetric resolution in the ROI to demarcate the specific location of the possible injury.

And based on the results obtained in the analyzed investigations, have identified two fundamental stages in the processing of digital infrared thermal images for the diagnosis of breast cancer; in which the correct performance, treatment and compliance with acquisition protocols are the key to the future performance of subsequent stages in order to arrive at an early and accurate diagnosis.

The first one, acquisition stage; as a fundamental conditions are proposes, in order that the captured image has the least possible influence from external sources that introduce undesirable factors, such as protocols to have in mind, a convergent point between the criteria emitted (Mashekova et al., 2022; Santana et al., 2018; Tello-Mijares et al., 2019; Pérez et al., 2016; Rodríguez Guerrero et al., 2014); these are:

- Controlled ambient temperature in a range of: 18 to 22°C.
- Relative humidity in the room: 60 ± 5%.
- Room for the capture: Hermetized and closed for the non-influence of external sources of radiation, with constant air flow.
- Prior rest in patients without contact with the study area before testing: 10 to 20 minutes.
- Patient position: Front of the camera, with the waist area facing upwards uncovered, arms raised with the palm of the hands resting on the back of the head.
- Distance between the camera and the patient: 1 meter.
- No consumption of alcohol, nor application of cosmetics in the area of the surface of the breasts, nor accomplishment of physical exercises the day of accomplishment of the test.
- Acquisition of images on the following days: From 5 to 12 and after 21 of the menstrual cycle.
- Emissivity of human skin for the camera used: 0.98.
- Distance between the camera and the patient: 1 meter.

The second one, pre-processing stage; an arduous task that requires attention and efforts in the methods used in it, the operations to be performed must go beyond resize the input thermal images and perform gray level histogram, sharpness enhancement, color correction, attenuation and correction actions since it is a crucial step for the subsequent performance of the CAD systems in the early detection of breast cancer, in the stages of segmentation, extraction of characteristics and finally the classification of the image analyzed according to the study being developed.

For this reason, operations must be carried out for the attenuation and correction of product distortion to require prior treatment to reduce effects such as erroneous pixels, random noise from the electronics used in the measurements; or optical problems of framing and focusing of the image or non-uniform heating, by sources of excitation that independent of having taken into account protocols at the time of acquisition, may continue to appear, followed by the extraction of ROI since the presence of areas outside the study area that show non-relevant information in these images may be considered as noise accompanied by the fact that they bring with them a higher computational cost and time in later stages and its subdivision into a certain number of quadrants that allow comparing symmetrical anatomical areas in both DITI to identify different temperature patterns which based on medical criteria may constitute an alarm, to only continue to later stages with those DITI of patients that are more likely to constitute a pathological images and move on to analyze other characteristics, and to reach a faster and more accurate diagnosis since the DITI of patients that at this point do not constitute an alarm can go directly to the classification stage like a not pathological images.

Based on the above, the authors of this article propose a logical sequence of six stages to be taken into account in the development of future methods and algorithms for the pre-processing stage of DITI for the early detection of breast cancer.

The first stage would be related to an algorithm to discriminate images without potential risk of a pathological condition, which by subdividing them into a specific number of quadrants referring to both symmetrical anatomical areas of the breasts and comparing them can detect temperature differences that may be the starting point for the presence of a pathological abnormality and thus not having to perform the following stages and be able to move directly to the classification stage in the processing of DITI, the second stage would be composed of actions for the reconstruction of defective pixels, the third stage would present actions for the attenuation of noise introduced by external sources at the time of acquisition, the fourth stage would allow the enhancement of contrast in the DITI taking into account the non-alteration of the initial image, the fifth stage would allow the enhancement of features associated with the area of possible pathological condition and the sixth stage would lead to the extraction of the ROI and the feedback with the criteria issued by the specialist; as shown on Figure 3.

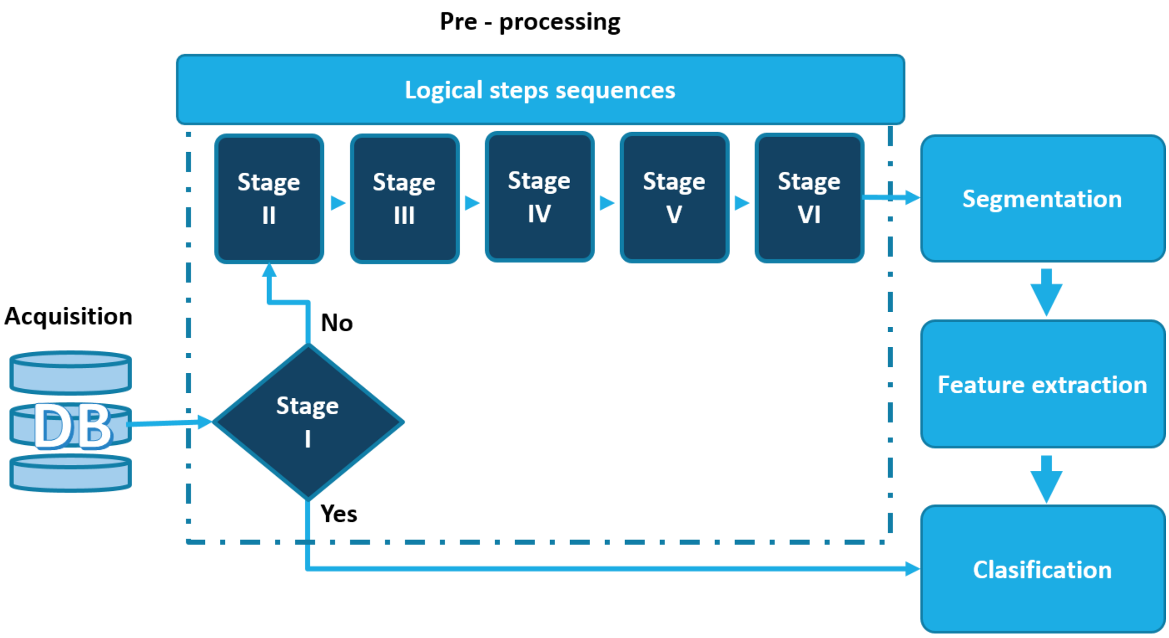


Figure 3: Proposed logical steps sequences diagram for the pre-processing stage of DITI.

5. Conclusion

The methods and techniques used in the processing of DITI with CAD support for the early detection of breast cancer were discussed, showing the need for continued research in this field by the high numbers of incidence and mortality associated with this disease in different regions of the world, and the existence of needs identified in the pre-processing stage of the DITI by the erroneous results achieved in research that enhance segmentation, feature extraction and classification.

In order to solve these needs, several needed steps for acquisition stage and a logical steps sequences of operations and actions to be taken into account in the pre-processing stage was presented in order to improve the quality of the digital infrared thermal image and increase the performance of the existing algorithms in the following stages of segmentation, feature extraction and classification, in order to support the decision making process of the diagnostic imaging staff of health care institutions.

References

Al-shamasneh, A.R.M., Obaidellah, U.H., 2017. Artificial Intelligence Techniques for Cancer Detection and Classification: Review Study. European Scientific Journal, ESJ, 13 (3), Pp. 342-370. https://doi.org/g8w2

Bastida, M.B., Cabello, M.E., Rodríguez, A.L., and García, J., 2017. Sistema de apoyo para la toma de decisiones en termografía de glándulas mamarias. Revista mexicana de ingeniería biomédica, 38 (1), Pp. 166-187. https://bit.ly/33soQTF

Bereciartua, P.M.A., 2016. Desarrollo de algoritmos de procesamiento de imagen avanzado para interpretación de imágenes médicas. Aplicación a segmentación de hígado sobre imágenes de Resonancia Magnérica multisecuencia [Doctoral Thesis, Universidad del País Vasco]. Archivo Digital Docencia e Investigación. https://bit.ly/3s9GWnw

Dey, N., Ashour, A.S., and Althoupety, A.S., 2017. Thermal Imaging in Medical Science. In V. Santhi (Ed.), Recent Advances in Applied Thermal Imaging for Industrial Applications, pp. 87-117. IGI Global. https://doi.org/hcpm

Dirección de Registros Médicos y Estadísticas de Salud. 2022. Anuario estadístico de Salud. La Habana, Cuba: Ministerio de Salud Pública. https://bit.ly/3Navwcn

GLOBOCAN. 2023a. Cancer Today. Estimated number of new cases in 2020, worldwide, both sexes all ages. International Agency for Research on Cancer. https://bit.ly/3xYuBUh

GLOBOCAN. 2023b. Cancer Today. Estimated number of new cases in 2020, all cancer, both sexes, all ages. International Agency for Research on Cancer. https://bit.ly/31DVhxl

GLOBOCAN. 2023c. Cancer Tomorrow. Estimated number of new cases from 2020 to 2040, both sexes, ages [0-85+]. International Agency for Research on Cancer. https://bit.ly/3Gh9HTg

GLOBOCAN. 2023d. Cancer Today. Estimated number of new cases in 2020, worldwide, female, all ages. International Agency for Research on Cancer. https://bit.ly/3otDPF2

GLOBOCAN. 2023e. Cancer Today. Estimated number of new cases in 2020, breast, females, all ages. International Agency for Research on Cancer. https://bit.ly/3DyJ7TH

GLOBOCAN. 2023f. Cancer Tomorrow. Estimated number of new cases from 2020 to 2040, females, age [0-85+]. International Agency for Research on Cancer. https://bit.ly/31uT9Il

GLOBOCAN. 2023g. Cancer Today. Estimated age-standardized incidence and mortality rates in 2020, worldwide, females, all ages. International Agency for Research on Cancer. https://bit.ly/3GzBAX1

Haralick, R.M., Shanmugam, K., and Dinstein, I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 (16), Pp. 610-621. https://doi.org/bdqvtn

Hossam, A., Harb, H., and Abd El Kader, H., 2018. Automatic image segmentation method for breast cancer analysisi using thermography. JES. Journal of Engineering Sciences, 46 (1), Pp. 12-32. https://doi.org/hcpp

Huang, Q., Zhang, F., and Li, X., 2018. Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey. BioMed Research International, 2018 (1), Pp. 1-10. https://doi.org/gdfndb

Jean, M.S., Maresova, P., Krejcar, O., Selamat, A., and Kuca, K., 2018. Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model. Sensors, 18 (9), Pp. 1-19. https://doi.org/hcpv

Kamath, D., Kamath, S., Prasad, K., and Rajagopal, K.V., 2015. Segmentation of Breast Thermogram Images for the Detection of Breast Cancer - A Projection Profile Approach. Journal of Image and Graphics, 3 (1), Pp. 47-51. https://doi.org/g8ws

Liu, S., Zeng, J., Gong, H., Yang, H., Zhai, J., Cao, Y., Liu, J., Luo, Y., Li, Y., Maguire, L., and Ding, X., 2018. Quantitative analysis of breast cancer diagnosis using a probabilistic modelling approach. Computers in Biology and Medicine, 92, Pp. 168-175. https://doi.org/gcxhtx

Mahmoud, M.A.M., 2016. Development of advanced computer methods for breast cancer image interpretation through texture and temporal evolution analysis [Doctoral Thesis, Universitat Rovira I Virgili]. Tesis Doctorals en Xarxa. https://bit.ly/3dKcesS

Mammoottil, M.J., Kulangara, Ll.J., Cherian, A.S., Mohandas, P., Hasikin, K., Mahmud, M., 2022. Detection of Breast Cancer from Five-View Thermal Images Using Convolutional Neuronal Networks. Journal of Healthcare Engineering, (1), Pp. 1-15. https://bit.ly/3Cx4Qxy

Mashekova, A., Zhao, Y., Ng, E.Y.K., Zarikas, V., Fok, S.C., and Mukhmetov, O., 2022. Early detection of the breast cancer using infrared technology - A comprehensive review. Thermal Science and Engineering Progress, 27 (1), Pp. 101-142. https://bit.ly/469BBP0

Meena, P.R., Bhuvaneshwari, K.S., Divya, M.N., Sri, J., and Sulaiha, B.A., 2017. Segmentation of thermal infrared breast images using K-means, FCM and EM algorithms for breast detection [Paper presentation]. International Conference on Innovation in Information, Embeddes and Communication Systems (ICIIECS), Coimbatore, India. https://doi.org/g85f

Mohamed, E.A., Rashed, E.A., Gaber, T., and Kram, O., 2022. Deep learning model for fully automated breast cancer detection system from thermograms. PLoS ONE, 17 (1), Pp. e0262349. https://bit.ly/43ID1ht

Padmappriya, S., and Sumalatha, K., 2018. Digital Image Processing Real Time Applications. International Journal of Engineering and Science Invention, 2018 (1), Pp. 46-51. https://bit.ly/3dKa9NB

Pavithra, P., Ravichandran, K., Sekar, K., and Manikandan, R., 2018. The Effect of Thermography on Breast Cancer Detection. Systematic Reviews in Pharmacy, 9 (1), Pp. 10-16. https://bit.ly/31C9nzZ

Pérez, M.G., Conci, A., Aguilar, A., Sánchez, Á., and Andaluz, V.H., 2016. Detección temprana del cáncer de mama mediante la termografía en Ecuador. Maskana, 5, Pp. 111–123. https://bit.ly/3ylEknK

Qin, C., Yao, D., Shi, Y., and Song, Z., 2018. Computer-aided detection in chest radiography based on artificial intelligence: a survey. BioMedical Engineering OnLine, 17, 113(2018). https://doi.org/gh2h7s

Real Academia Española. 2023. Medicina. In Diccionario de la lengua española. Retrieved June 15, 2023, from https://dle.rae.es/?id=OkukRhl

Resmini, R., Silva, L., Araujo, A.S., Medeiros, P., Muchaluat-Saade, D., and Conci, A., 2021. Combining Genetic Algorithms and SVM for Breast Cancer Diagnosis Using Infrared Thermography. Sensors, 21 (1), Pp. 1-24. https://doi.org/hcpt

Roa, M.S.M., Borsetti, G.V.S.A. and Vicentini, J.M.J., 2016. Representación conceptual de imágenes médicas digitales: Integración de Contexto y Contenido Visual. Revista General de Información y Documentación, 26 (2), Pp. 651-672. https://doi.org/g8w4

Rodríguez, G.S., Loaiza, H., and Restrepo, G.A.D., 2014. Automatic segmentation of thermal images to support breast cancer diagnosis. Revista R&T, 12 (30), Pp. 25-34. https://doi.org/hcps

Sánchez-Ruiz, D., Olmos-Pineda, I., and Olvera-López, J.A., 2018. Automatic Segmentation in Breast Thermographic Images Based on Local Pattern Variations. Research in Computing Science, 147 (11), Pp. 55-66. https://bit.ly/3oQBe85

Santana, M.A., Pereira, J.M., Silva, F.L., Lima, N.M., Sousa, F.L., Arruda, G.M., Lima, R.C., Silva, W.W., and Santos, W.P., 2018. Breast cancer diagnosis based on mammary thermography and extreme learning machines. Research on Biomedical Engineering, 34 (1), Pp. 45-53. https://doi.org/hcpn

Seigel, R.L., Miller, K.D., Wagle, N.S. and Jemal, A., 2023. Cancer Statistics, 2023. CA: A Cancer Journal for Clinicians, 73 (3), Pp. 1-112. https://bit.ly/43Fp3gv

Silva, L.F., Saade, D.C.M., Sequeiros, G.O., Silva, A.C., Paiva, A.C., Bravo, R.S., and Conci, A., 2014. A new database for breast research with infrared image. Journal of Medical Imaging and Health Informatics, 4 (1), Pp. 92-100. https://doi.org/hcpw

Sociedad Española de Oncología Médica. 2023. Las Cifras del Cáncer en España. Madrid, España: Sociedad Española de Oncología Médica. http://bit.ly/3Nu68zA

Teja Kakileti, S., Dalmia, A., and Manjunath, G., 2019. Exploring deep learning networks for tumour segmentation in infrared images. Quantitative InfraRed Thermography Journal, 17 (3), Pp. 153-168. https://doi.org/hcpq

Tello-Mijares, S., Woo, F., and Flores, F., 2019. Breast Cancer Identification via Thermography Image Segmentation with a Gradient Vector Flow and a Convolutional Neural Network. Journal of Healthcare Engineering, 2019 (1), Pp. 1-13. https://doi.org/gh569q

Tsietso, D., Yahya, A., and Samikannu, R., 2022. A Review on Thermal Imaging-Based Breast Cancer Detection Using Deep Learning. Mobile Information Systems, (1), Pp. 1-19. https://bit.ly/42Hwxy3

Wakankar, A.T., and Suresh, G.R., 2016. Automatic Diagnosis of Breast Cancer using Thermographic Color Anaolysis and SVM Classifier. In J.M. Corchado Rodríguez, S. Mitra, S.M. Thampi, E. El-Alfy (Ed.), Intelligent Systems Technologies and Applications 2016 (pp. 21-32). Springer. https://bit.ly/3pFhL9L

World Health Organization. 2017. Guide to Early Cancer Diagnosis. Geneva, Switzerland: WHO Document Production Services. https://bit.ly/3y1Tf6z

World Health Organization. 2023. Cancer. https://bit.ly/3X6NdxX