International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:3, No:8, 2009

Breast Cancer Treatment Evaluation based on Mammographic and Echographic Distance Computing

M. Caramihai, Irina Severin, H. Balan, A. Blidaru, and V. Balanica

***Abstract*—**Accurate assessment of the primary tumor response to treatment is important in the management of breast cancer. This paper introduces a new set of treatment evaluation indicators for breast cancer cases based on the computational process of three known metrics, the Euclidian, Hamming and Levenshtein distances. The distance principals are applied to pairs of mammograms and/or echograms, recorded before and after treatment, determining a reference point in judging the evolution amount of the studied carcinoma. The obtained numerical results are indeed very transparent and indicate not only the evolution or the involution of the tumor under treatment, but also a quantitative measurement of the benefit in using the selected method of treatment.

***Keywords*—**Breast cancer, Distance metrics, Cancer treatment evaluation.

I. Introduction

BREAST CANCER (BC) is the major health problem because of its high incidence and unpredictable evolution. Early detection by screening and improving treatment could be the solution. The research and development of several examination methods and techniques, as well as the imaging instruments, underwent a high rate of progress in the last century in order for the routine screening process to go as fast and accurate as possible in diagnosing the tumour. A large range of technologies and instruments have evolved based on X-rays analysis, ultrasound evaluations and magnetic resonance techniques, among which mammography, echography and the magnetic resonance imaging offer the best qualitative results in performance and input-output ratio.

Current international developments in the area of imaging technologies are directly influenced by the continuous improvement of the high-frequency transducer, the utilization of high harmonics and the possibility of joining real time extracted space data. In fact, due to such state-of-the-art breakthroughs, the escalated imaging resolution gave way to an accurate recognition and examination of ducts, even of the smallest ones, as well as visualization of the breast functional unit – the terminal duct-lobular unit (TDLU), the area of appearance of majority of mammary formations. All these accumulations allow cancer diagnosis at early stages, on one hand due to a more accurate determination of the extensive malign conditions and on the other hand due to an improved negative predictive capacity (emphasizing of benign characters).

Based on imaging results, all identified lesions - masses, architectural distortions, asymmetric densities and calcifications - are assessed on a scale between benign and malignant, for example in terms of the BI-RADS Assessment Category (Breast Imaging Reporting and Data System), allowing the estimation of the cancer risk. The characteristics of the discovered lesions are analyzed for the cases evaluated with a high risk of cancer and a treatment procedure is recommended after the histopathologic diagnostic of cancer is established. Using the aforementioned imaging techniques, the post-treatment cases are again examined using the same equipment after-treatment and the mammographic and echographic modifications and structural alterations are assessed, [1].

The aim of this paper is to develop an evaluation method of the breast cancer treatment, based on before- and post-treatment mammographic and echographic images through Hamming, Euclidian and Levenshtein distances evaluation. The resulted distances, computed using the parametric descriptors of the mammograms and echograms at two (or more) different ill steadies, are analyzed and then a supplementary help indicator is given about the status of the tumour, whether the breast cancer evolved or involved as result of the applied treatment.

II. Breast Cancer Diagnosis

A. The Mammographic Examination for BC Diagnosis

The breast radiological exploration has a special importance in the early diagnostic of cancer and has taken a great ampleness lately due to the improvement of radiological technique, [2]. Although it is the oldest para-clinical investigation, the mammography remains by far the most useful and utilized imagery method in the diagnostic of mammary diseases.

M. Caramihai, Irina Severin and V. Balanica are with University POLITEHNICA Bucharest, Bucharest, CO 060042 Romania (phone: +4021 402 9167; e-mail: m.caramihai@ieee.org, irina_severin2003@yahoo.com, vicord20011@gmail.com),

H. Balan is with University of Medicine "C. Davila", Str. Dionisie Lupu 37, Bucharest, Romania (e-mail: drhoriabalan@yahoo.com)

A. Blidaru is with Oncology Institute "Alex. O. Trestioreanu", Sos. Fundeni, 252, Bucharest, Romania (e-mail: alexandrublidaru@yahoo.com).

B. BC Risk Classification

Even though the cancer lesions can vary considerably, many screening cases have been studied and analyzed and a set of characteristics has been chosen as best defining the tumour. Based on these characteristics oncologists evaluate the state of the breast cancer and classify it on the BI-RADS scale. BI-RADS is an unitary system designed for helping medical professionals assess, interpret and classify mammographies, echographies and magnetic resonance imaging in a concise and unambigous and standardized way, [8], by assigning numbers or numerical codes to different risk categories. The Assessment Categories numbered from 0 to 5 are described below:

- 0: Incomplete – Additonal imaging data and evaluation is necessary
- 1: Negative – Breasts are symmetrical and there is no suspecious lesions
- 2: Benign finding(s) – There is no sign of malignancy, although benign lesions are found (like involuting fibroadenomas, fat containing oil cysts, lipomas, etc.)
- 3: Probably benign – Benign lesions are found and short interval follow-up procedures are needed for more evolution data
- 4: Suspicious abnormality – Detected lesions are not cancer, but have a definite probability of being malignant
- 5: Highly suggestive of malignancy – Lessions have a high probability of being cancer.

C. Mammary Echography for BC Diagnosis

The nosologic frame of mammary echographic pathology needs also an innovating approach in the sense of unification of lesions classes with those that define mammar-graphical modifications and that were utilized for a long time, namely BI-RADS. The screening purpose is to decelate cancer for a large population level or for non-symptomatic subjects of the mammary, and the utilization of echography for diagnostic purpose refers to the examination of some mammary abnormalities decelated either mammo-graphically or by palpation.

Unlike the BI-RADS categories for mammography, in the case of echographic classification, the 4th category was subdivided into two subcategories 4a and 4b, because of the importance, from the medical-legal point of view, of the adjective “probable” which is associated to category 4b (risk over 50%).

III. Mammography and Echography Treatment Evaluation

A. Parametrization Process of Lesion Characteristics

For a complete analysis of the trend of cancer evolution – basically the tumour evolution or the involution – the hypothesis shown below has been verified in every aspect based on the experience of medical staff.

The essential sign of tumor response, (based on image analysis), is the reduction of its size and dimensions, [4]. Moreover, many other characteristics can be taken into account, e.g.: contrast of the lesion, spiculations, angulations, posterior shadow, margins, calcifications, fragmentation, structure, etc.

Based on the aforementioned aspects, the tabel showed below (Table I) presents the lesions characteristics that have been selected to further parametrise the tumour from the digital analysis point of view. Hence, every characteristic has a weight attached in order to describe its relevance vs. the final diagnostic:

TABLE I
Relevance of tumour characteristics

Characteristics of lesion	Mammography - M Echography - E
Dimension E	30%
Dimension M	30%
Contrast M	20%
Spiculations M	10%
Fragmentation M	10%

The dimension is a characteristic expressed in terms of surface or volume that is precisely measured by the machine: the mammograph or the echograph. The ranges of this characteristic can vary dramatically, but in this study we consider a scalable value from “0” to “100”. The value “0” reflects the absence of any cancer lesion or visible carcinoma, while “100” represents the evidence of the highest level of cancer risk possible for a tumour. As dimension is the most important indicator of cancer evolution, this parameter was analysed and evaluated from both the mammographic and echographic perspectives.

The contrast of the lesion indicates the level of active tumoural cells in a carcinoma and is measured in grey levels between “0” and “255”. A lighter grey color (near the value 0) indicates the absence of the carcinoma (i.e.the malign cells becoming less in number than in the beginning). This treatment result shows that the healing method prevailed as many invasive cells were distroyed. On the other hand, a darker grey shows that the carcinoma evolved, i.e. the number of active cells increased and more blood is being supplied to the lesion. This is the discouraging result of a treatment, when the healing method wasn’t powerful enough to disrupt the cancer cells.

The presence or the absence of spiculations arising from a detected mass, on a mammography or echography, gives a very powerful indicator of the malignancy. This fact was largely verified in many studies in the field […], confirming that the existance of spiculations implies a high level of malignancy. In this study, this characteristic is considered to be a boolean typed parameter, where the value “0” expresses the absence of any spiculations and “1” the presence of spiculations.

Fragmentation of tumour is also a very important characteristic, its existance suggesting the amelioration of the tumoural lesion. This characteristic was also boolean parametrised, the value “0” representing the absence of any

fragmentation, while "1" represens the presence of fragmentation.

In short, the next table presents a short review of the considered value range variation for the chosen characteristics, given in the incremental order of the cancer degree / intensity / evolution:

The elements described above materialize the hypotesis of the present study. Hence, the mammography and echography characteristics are further assessed in accordance to the stated parametrisation.

TABLE II
RANGE OF VALUES FOR THE SELECTED CHARACTERISTICS

CHARACTERISTICS OF LESION	MINIMAL VALUE (ABSENCE OR LOW INTENSITY / DEGREE OF CANCER)	MAXIMAL VALUE (HIGH INTENSITY / DEGREE OF CANCER)
Dimension E	0	100
Dimension M	0	100
Contrast M	0	255
Spiculations M	0	1
Fragmentation M	1	0

B. Treatment Evaluation Based on Distance Computation

This paper presents the computational process of the Euclidian, Hamming and Levenshtein metrics for one pair of mammograms and one pair of echograms, regarding the same ill person. The distance results give in all cases a hint about the evolution direction of the tumour and more information about the level of progress regarding the treatment effect: the tumour lesions have evolved or involved under a specific treatment and how significant the change was.

The Euclidean distance, or the Euclidian metric, represents the root of square differences between coordinates of two points/strings of equal length [1], i.e. the Euclidian distance for n-dimensional $X = (x_1, x_2, ...,x_n)$ and $Y = (y_1, y_2, ...,y_n)$ is computed as:

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2} \quad (1)$$

The Hamming distance represents the amount of difference for two strings of equal length, $X = (x_1, x_2, ..., x_n)$ and $Y = (y_1, y_2, ...,y_n)$, computed by counting the minimum number of substitutions needed to change one string into the other [2].

The Levenshtein distance is applicable also for strings of different length and is computed by counting the minimum number of operations (insertion, deletion or substitution of single characters) [3] needed to turn one string into the other.

In order to assess the modifications and structural alterations of the inspected lesions, each pair of images - consisting of one image taken before treatment and one (or more) during the treatment – is recorded using the same clinical investigation method, same equipment and the same environmental conditions. Without this supposition, the image analysis results may be questionable.

The first and most important determination is the evolution direction of the tumour. For a pair of images, either mammograms or echogramms, the dimension alteration is the most relevant sign of evolution.

The following section presents a detailed simulation of the considered evaluation steps, highlighting the advantages and benefits of using the distance measurement methodology in the breast cancer treatment evaluation.

IV. SIMULATION OF DISTANCE EVALUATION FOR BC IMAGES

A. Simulation Analysis

Fig. 1 presents a set of two mammogramms taken for the same pacient before treatment and post-treatment. The recommended treatment was the chemotherapy, which was regularly followed, [13].

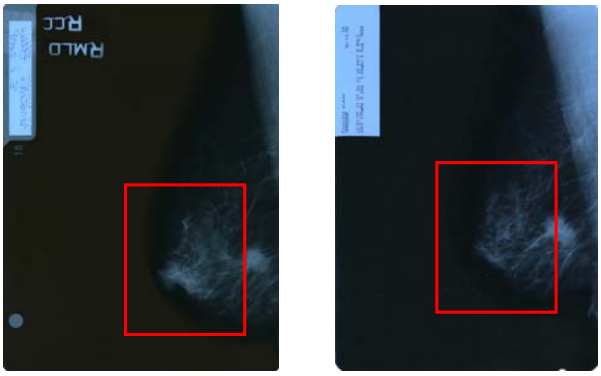

Fig. 1 Pair of mammograms

First of all, the lesion dimension for each mammography was recorded by the mammograph at each of the two points of time when the mammography was taken (same for echography). The dimension values, presented in Table III, suggest that the treatment was successful and the dimension has diminished. In this way, the evolution direction of the cancer is determined.

TABLE III
CHARACTERISTICS OF THE MAMMOGRAMS

CHARACTERISTICS OF LESION	BEFORE TREATMENT RESULTS	POST- TREAT MENT RESULTS
Dimension E	21	17
Dimension M	20	15
Contrast M	180	130
Spiculations M	0	0
Fragmentation M	1	0

The identification of the lesion, seen also in Fig. 1, is followed by the evaluation of lesion characteristcs: contrast, spiculations and fragmentation. This is done with a help of a developed software application, which takes an image as input and returns the assessed values for each of the enumerated characteristics as output. Of course, the medical staff can write / modify the values based on its own experience (e.g. Table III). The progress of each characteristic is an individual indicator of the cancer evolution, [3]. The presented values indicate an improvement of the patient's condition after

treatment by a diminished dimension and contrast of the lesion and also the appearance of fragmentation process.

TABLE IV
NORMALIZED VALUES OF THE CHARACTERISTICS

CHARACTERISTICS OF LESION	BEFORE TREATMENT RESULTS	POST- TREATMENT RESULTS
Dimension E	0.790	0.830
Dimension M	0.800	0.850
Contrast M	0.294	0.490
Spiculations M	0	0
Fragmentation M	1	0

During the linear interpolation normalization, the values for all the characteristics are computed based on the formula $c_i = \frac{b_i^{\max} - b_i}{b_i^{\max} - b_i^{\min}}$, where c_i is the normalized value, b_i is the assessed value of the characteristic and $[b_j^{min}, b_j^{max}]$ represents the interval of possible values for the investigated characteristic, assigned in Table II. The normalization process applies only for the first two selected characteristics, dimension and contrast, as the second two, spiculations and fragmentation, are already normalized (being boolean parameters). For this case, the obtained results are listed in Table V.

TABLE V
NORMALIZED AND WEIGHTED VALUES OF THE CHARACTERISTICS

CHARACTERISTICS OF LESION	BEFORE TREATMENT RESULTS	POST- TREATMENT RESULTS
Dimension E	0.237	0.249
Dimension M	0.240	0.255
Contrast M	0.058	0.098
Spiculations M	0.000	0.000
Fragmentation M	0.100	0.000

For the presented case, the amount of change in the cancer evolution can be predicted based on the computational process of the above discussed metrics. For both the mammograms recorded before and after treatment, the assessed values of the specific selected characteristics are first normalized, using the linear interpolation normalization based on the value ranges assigned in Table II, and then recalculated, based on the weights assigned in Table I. These computational steps make sure that the procedure is consistent and uniform applicable for any range of values set for the chosen characteristics, and the computed results are universally valid and reflects the significance of tumour characteristics evaluation.

Next computational step implies the calculation of the weighted value of normalized results. The obtained data, presented in Table V, is further used in determining the Euclidian, Hamming and Levenshtein distances.

The vectors lengths for the selected characteristics are the same in both cases, (i.e. Hamming and Levenshtein distances); hence, the distances values will show the same result, namely 3. The Euclidian distance is computed based on equation (1). The distance results of the Euclidian, Hamming and Levenshtein metrics are displayed in Table VI.

TABLE VI
METRICS COMPUTATION FOR MAMMOGRAMS

METRICS	DISTANCE
Euclidian	0.l094
Hamming	3
Levenshtein	3

B. Discussions regarding Distance Results

In order to judiciously appreciate the obtained results, Table VII presents the range of possible values for the discussed distances, ranges directly computed based on a hypothetical case which presents the full evolution of cancer, starting with the absence of the tumour and ending with the full negative manifestation of all selected characteristics. The pair wise comparison of the obtained results with the respective interval or set of possible values indicates the amount of progress or change in the tumour evolution.

TABLE VII
DISTANCE RANGES FOR METRICS

METRICS	DISTANCE RANGE
Euclidian	[0, 2.24]
Hamming	{0, 1, 2, 3, 4}
Levenshtein	{0, 1, 2, 3, 4}

The Euclidian metric is the only one that precisely indicates the amount of progress in the evolution / involution of cancer, the direction of cancer development being established by comparing the dimensions and, if applicable, the contrast of the mammograms or echograms before and after treatment. The Hamming and Levenshtein metrics give only a measurement of the change / alteration level as consequence of the treatment procedures, representing only a rough indicator of the lesion evolution trajectory.

For the presented case, the obtained Hamming and Levenshtein distances, with the computed value of ”3”, suggests a significant change level in the development of cancer as consequence of the treatment. The Euclidian distance, with the computed value of “0.11”, indicates an amelioration of the tumour with almost 5%. This final result is relevant and extremely transparent for all interested people involved in BC treatment. Moreover, in this manner, the treatment evaluation process can be assessed in an computerized way, and can improve the computer use in oncology care.

V. CONCLUSION

The main objective of this paper was to present a new computational methodology for BC treatment evaluation. The concepts of Euclidian, Hamming and Levenshtein distances applied to pairs of digitally assessed characteristics of mammographies and echographies allow the evaluation of a set of quantitative indicators capable to evaluate the treatment efficiency of BC. The obtained results offer another way of

computer use in primary tumour response evaluation of to the treatment.

Although this paper is the result of qualified research in breast cancer cases, the idea can be further investigated and extended to all the other types of cancer. Further development in this field includes also the calibration of the obtained results to reflect also the appropriate category of the BI-RADS scale. Last but not least, the software application that assesses the characteristics of BC images and evaluates the oncologic treatment can be further extended and new medical tools can be implemented.

REFERENCES

[1] B.K. Verma, *Heredity & Cancer: Breast cancer as a model*, IASTED RM'96, Hawaii, USA, 1999, pp. 84-88.
[2] Carlos Andres Pena et all, *Proceedings Information Processing in Medical Imaging*, IPMI"03, 2003.
[3] K. S. Woods, C. C. Doss, *Comparative Evaluation of Pattern Recognition Techniques M 16L-171PI for Detection of Microcalcifications in Mammography*, International Journal of Pattern Recognition and Artificial Intelligence, 1993.
[4] K.S. Woods, C.C Doss, *Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography*, International Journal of Pattern Recognition and Artificial Intelligence, 2003, pp. 80-85.
[5] Fitzgibbons PL, Page DL, Weaver D, et all, *Prognostic factors in breast cancer*, pp. 124:966–978, College of American Pathologists Consensus Statement. Arch Pathol Lab Med, 1999, 2000.
[6] Mirza AN, Mirza NQ, Vlastos G, et all, *Prognostic factors in node-negative breast cancer: a review of studies with sample size more than 200 and follow-up more than 5 years*, Ann Surg, 2002, 235:10–26.
[7] Greene FL, Page DL, Fleming ID, et all, *AJCC Cancer Staging Handbook. TNM Classification of Malignant Tumors*, 6th ed. New York: Springer Verlag, 2002.
[8] Sobin LH, Wittekind CH, eds. TNM, *Classification of Malignant Tumours*, 6th ed. New York: John Wiley & Sons, 2002.
[9] Adair F, Berg J, Joubert L, et all, *Long-term follow-up of breast cancer patients: the 30-year report*, Cancer, 1974, pp. 33:1145–1150.
[10] Abraham. Kandel, *Fuzzy Expert Systems*, CRC Press, 2002.
[11] Lance Chambers, Practical Handbook of Genetic Algorithms, CRC Press, 2004.
[12] Stephen I. Gallant, Neural Network Learning and Expert Systems, MIT Press, 2006.
[13] *** Case studies and clinical documentation, Institute of Oncology Bucharest.