2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Thermal Evaluation to Identify Nodules Using Semivariogram Curves*

C. G. Grassmann, J. C. P.Coninck, W. L. Ripka, and L. Ulbricht

Abstract—**Thermography can contribute to the early diagnosis of tumors by identifying nodules that need to be analyzed. The objective of this paper was to verify possible semivariogram curves to identify the possible spatial behavior centered in the region with the nodule and capture the thermal behavioral information surroundings of this point. For this, we used the resources of R-Studio and theoretical basis in semivariogram models. As results we verified that the spatial technique indicates Gaussian behavior in both healthy and tumor areas, but the thermal averages differ from each other. Based on the cases analyzed, all tumors evaluated were thermally significantly different in relation to the healthy point, even when presenting the same model for the semivariogram**

***Clinical Relevance*—Thermographic is a non-invasive technique and has been widely studied to aid in the early diagnosis of neoplasms by identifying nodules that need to be evaluated. Without this early identification, only larger (palpable) tumors would be evaluated, and thus many patients are at risk of developing metastasis before treatment begins.**

I. Introduction

Over a lifetime, one in eight women will be diagnosed with cancer, with more than half (52%) of cases and 62% of deaths occurring in developing countries [1]. Thyroid cancer most commonly affects women and is the most common type of neoplasm of the endocrine system. Early detection can reduce mortality. For this, it is necessary to adopt technologies to assist in this early diagnosis [2].

Thermoregulation is a natural phenomenon of body temperature maintenance mediated by the autonomic nervous system and has the potential to detect diseases that have thermal properties, such as cancer. Tumor cells absorb more nutrients, are metabolically more active, and reproduce in a disorganized way. Thus, the heat transferred in these cells leads to different temperatures than in the rest of the body. Furthermore, the temperature in these regions is higher, which enables identification through thermal cameras [3].

In addition, thermography is inexpensive when compared to other imaging exams and has shown promise as a technique. It is based on the principle of measuring the emitted radiation by an object or surface through an infrared camera to determine its temperature. All bodies above absolute zero temperature emit thermal radiation which can be computationally converted to "pixels" of temperature.

Infrared cameras, or thermographic cameras, capture the thermal radiation emitted by the body and convert it into an image that represents the distribution of surface temperatures of that body [4]. When there is an abnormality in the body tissue, such as a tumor (either malignant or benign), there is a change in the temperature in this tissue [5]. Thus, the goal of this study was to use mathematical modeling to do a tracking and identify possible tumor regions.

II. Methodology

A descriptive study was carried out with 32 patients (55.2±11.0 years old) with nodules in the thyroid region, confirmed by ultrasound (US) and fine needle aspiration puncture (FNAB), in a specialized cancer hospital. The experimental project involving human subjects was approved by the Ethics Committee of the Hospital before being initiated.

The data collection environment was 22°C, with a margin of error of 1°C, controlled by an air conditioner and monitored with a digital thermometer. The images were collected with a Model Ti32 Fluke Thermography camera. Thermal stress was performed with a gel pack for 30 seconds in the region. After three minutes of reheat, the thermal image was collected for data analysis (Fig.1). The analyses were performed in the SmartView 4.3 software that came with the camera, to perform temperature extraction and mean analysis.

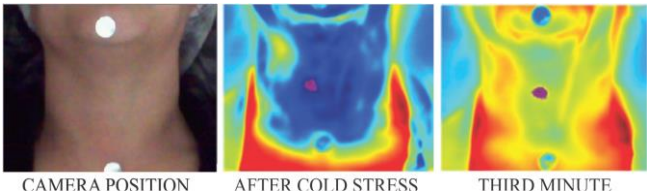

Figure 1. Representation of neck original position (left); thermal representation immediately after cold stress and in third minute.

Both the tumor areas (T) and the adjacent region (healthy tissue - H) were delimited and analyzed. R software was used to perform the analysis of the thermal variations and the analysis of the spatial behavior (variograms). To verify

*Research supported by the National Council for Scientific and Technological Development (CNPq).

C. Grassmann is an Undergraduate Student at UTFPR - Federal University of Technology - Paraná/Brazil. (Corresponding author phone: 55 41 95016697; email: camilagrassmann@alunos.utfpr.edu.br).

J. Coninck. Professor at DAEST – Academic Department of Statistics at UTFPR - Federal University of Technology - Paraná/Brazil (e-mail: coninck@utfpr.edu.br).

W. L. Ripka. Professor at Graduate Program in Biomedical Engineering at UTFPR - Federal University of Technology and School of Health at Centro Universitário Uniopet- Paraná/Brazil (e-mail: ripka.w@gmail.com).

L. Ulbricht. Professor at Graduate Program in Biomedical Engineering at UTFPR - Federal University of Technology - Paraná/Brazil (e-mail: leandraulbricht@utfpr.edu.br).

descriptive information between healthy and tumor tissue, exploratory analyses of data (mean, maximum and minimum temperatures) were performed [6].

In the second step the study of fit models of semivariance around the tumor and healthy regions was evaluated. To choose the theoretical semivariogram model with the best fit, the cross-validation technique was used for the values estimated by different theoretical models, given the ordinary kriging [7].

The semivariogram function [8] is one of the most widely used tools to represent the space expedition of a random function Z(u) in the direction of a spatial distance vector h between two points. This function is widely used in geostatistics to determine to explore spatial patterns or continuities. The variogram provides a precise meaning of the concept zone of influence of a sample. It is a function given by a curve. It is increasing with increasing distance "h" that separates pairs of samples, such that the more distant the samples are from each other, the greater the difference between their levels, and therefore less continuity, or spatial dependence, between them [9]. Statistical methods are not influenced by the spatial location of the samples, considering them spatially independent, taking into account only the variability of the data set. Geostatistical methods, on the other hand, consider the spatial correlation among the samples. This function is widely used in geostatistics to determine how to explore spatially distributed patterns or continuities [10]. Each semivariogram is associated with a mathematical theoretical model and can be: Circular (Cir), Spherical (Esf), Exponential (Exp), Gaussian (Gauss), Rational Quadratic (QR), Cardinal Sine (SC), K-Bessel (KB) and Stable (Est) to the isotropic semivariograms experiments [10]. Each theoretical model explains spatially distributive behavior. We can estimate by (1):

$$\gamma(h) = \frac{1}{2}E\{[Z(u) - Z(u+h)]^2\} \qquad (1)$$

The estimation is defined as the mathematical expectation (E) of the square of the difference between the values of the points in space, separated by the distance between vectors (h). The Gaussian variogram exhibits a parabolic behavior near the origin and represents extremely continuous phenomena. The amplitude in practice is approximately equal to A=a√3 and will reach a plateau C=c0+c asymptotically. The Gaussian model has the following (2):

$$(\mathrm{h}) = \mathrm{c}(0) + \mathrm{c}\{1 - \exp exp\,(-(\mathrm{ha}^2)\} \qquad (2)$$

Fig. 2 is a neck geographic visual example of the variograms. From the equation we obtain the best fit for the ratio of sampled versus estimated values [11]. After the verification of the best fit, we evaluated the test of equality between the thermal averages between the tumor and healthy regions via the Student's T-test to verify the null hypothesis (3):

$$H_0 = \mu_{healthy} \quad = \mu_{tumor} \qquad (3)$$

In effect, it is expected that the temperature around any point will not change significantly with respect to the centered value. However, when the comparison is made between healthy and tumor tissues, some difference is expected.

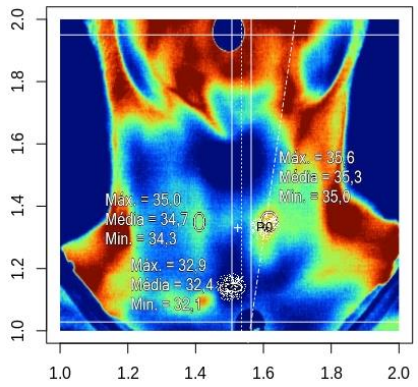

Figure 2. Example of areas of demarcation for quantification temperature. P0 area, in the right side of image, is a nodule area.

III. RESULTS AND DISCUSSION

Initially we show the results of mean temperature and standard deviation for the main tumor classifications found in the 32 patients (Table I). The mean healthy thermal values are lower than the tumor values, regardless of tumor type.

TABLE I. MEAN VALUES IN HEALTHY AND TUMOR REGIONS BY TUMOR TYPE PRESENTED AT CELSIUS DEGREES (ºC)

Tumor Type	n	Healthy		Tumor	
		Mean	*Standard deviation*	*Mean*	*Standard deviation*
Papillary carcinoma	7	28.92	0.22	31.68	0.09
Cyst in the isthmus	1	30.31	1.79	32.36	1.92
Atypia Sofun	4	32.79	0.74	36.98	0.88
Follicular Neoplasm	2	31.02	0.23	36.98	0.88
Multinodular goiter	2	34.89	0.10	39.14	0.19
Adenomatoid nodule	1	34.29	0.12	39.53	0.17
Follicular Adenoma	1	31.95	0.41	35.15	0.50
Benign Follicular Nodule	10	31.88	0.10	36.46	0.30
Colloid nodule	1	32.18	0.22	32.53	0.42
Medullary Carcinoma	1	36.06	0.54	37.82	0.40
Suspected malignancy	1	32.24	0.26	35.18	0.17
Suspected malignant neoplasm	1	35.21	0.92	36.63	0.378

a. Significative differences in temperature were verified for all tumor and healthy areas

To assess how much the averages differ between the two temperature groups, healthy and tumor region, we applied the T-Student test which indicated significant difference with p-value << 5% between the thermal averages in the healthy region (32.64ºC) and tumor region (35.87ºC). The residuals for the test are independent, show normality according to the Shapiro Wilk test, and proved to be homoscedastic, given the Bartlett's test in which the variances in each of the sample groups are equal.

That is, the null hypothesis of thermal equality around the samples does not hold, thus demonstrating a significant

thermal difference between the groups (Table I). Therefore, all the mean temperatures differ from each other.

Furthermore, Fig.3 represents the spatial distribution around each case, tumor and healthy. The tumor temperature is significantly higher than the healthy temperature, but the behavior around the sample center in both cases follows the Gaussian variogram model.

To better evaluate these thermal differences, it was performed confidence intervals by Turkey Test for the average of each of the points under analysis, which confirmed that the averages differ. In addition, it is verified the mean temperatures without interval intersection of the thermal results, in the internal, healthy point [31.29ºC - 33.99ºC], and in the analyzed point of the tumor tissue [34.66ºC- 37.67ºC].

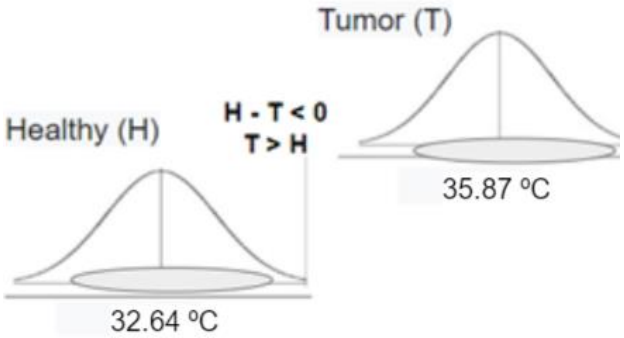

Figure 3. Spatial representation of temperature (degree Celsius) difference between healthy and tumor areas. Data represent the mean values for all volunteers.

Distributions are calculated using actual thermal data. Each thermal pixel is associated with a u (x, y) coordinate. This information is being used in the modeling equation (1). The difference between healthy and tumoral cases is in the mean temperature and not in the type of the adjusted distribution, in this case the Gaussian distribution. Therefore, it was possible to verify the thermal concentration around the chosen point (healthy and tumor), confirming the result obtained in the literature [12]. This result implies the possibility of using more robust discriminant statistical techniques, since the sample behavior around each point has Gaussian properties. This spatial characteristic is independent of the classification of the region, whether healthy or tumor, as well as between classifications regarding the type of tumor.

Additionally, three volunteer's thermal data were analyzed (one malignant tissue, one benign and one healthy tissue). These data were important to demonstrate that regardless of the tumor classification and theoretical Gaussian model.

In Fig.4 it is possible to visualize the compose result of a volunteer who was diagnosed with a Papillary Carcinoma, which is the most common neoplasm of the thyroid [7]. The results represent the spatial random sampling around the tumor (x and y coordinates) (A). Additionally, blue spots could be interpreted as coldest point as compared to the yellow and red sports in the center. On the other hand, Fig.4B represents the tumor thermal distribution

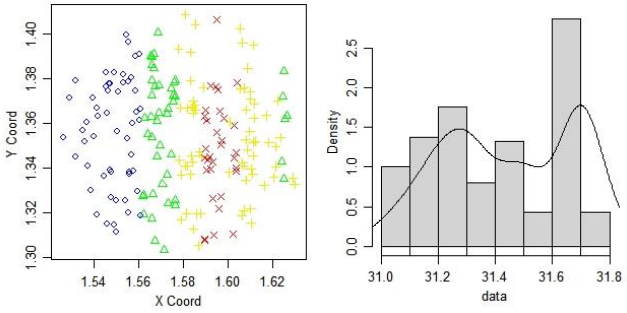

Figure 4. Individual analysis of a malignant tumor. A) Spatial random sampling around tumor (left); B) thermal distribution of the tumor (right).

In the same way, Fig. 5A and B showed the analysis for a volunteer with cyst in the isthmus and in Fig. 6A and B, for a healthy tissue.

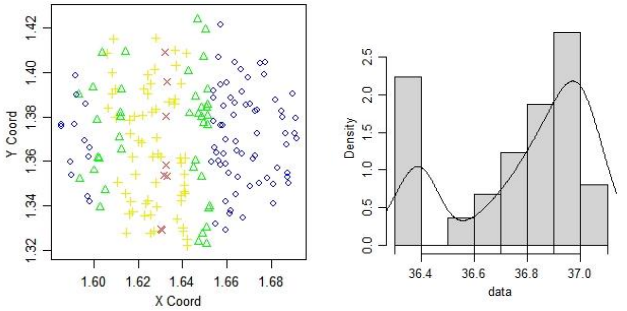

Figure 5. Individual analysis of benign nodule. A) Spatial random sampling around tumor (left); B) thermal distribution of the tumor (right).

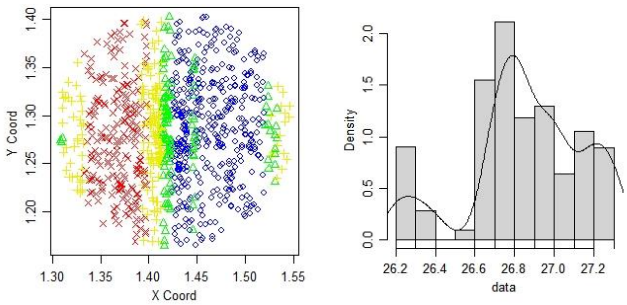

Figure 6. Individual analysis of healthy tissue. A) Spatial random sampling around tumor (left); B) thermal distribution of the tumor (right).

Finally, in Fig.7, we have the semivariogram of a volunteer diagnosed with Papillary Carcinoma (malignant), in which the theoretical model that had more correspondence, was the Gaussian Model, this process was repeated for each volunteer. The analyses performed about 87.5% of the cases (28 patients), were Gaussian.

A. Theoretical Gaussian Model

The analyses were performed for each patient, generating two new samples, one in the point of the tumor (T), and another in the point in a healthy area of the patient (H). After placing the mean thermal values (in ºC), we also indicated the result of the variogram model found in each tumor

classification. In both sample (tumor tissue and health tissue) sets we obtained the Gaussian model, the latter with the best fit given in the semivariogram curve.

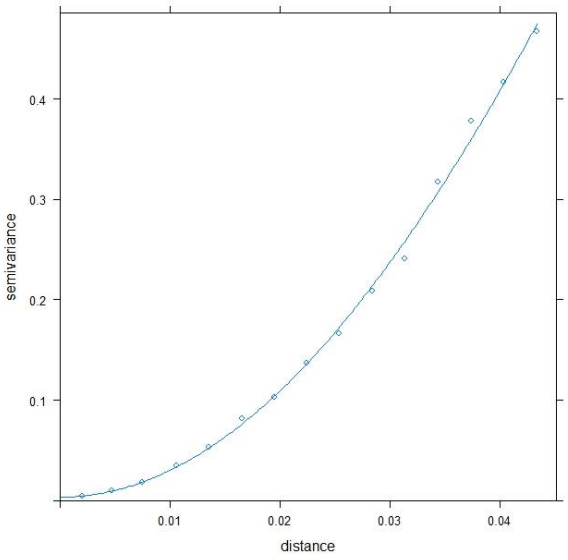

Figure 7. Example of Gaussian Semivariogram Model of a volunteer diagnosed with a malignant tumor.

Thus, for future work there are two possible techniques that can be applied in this work: a) Logistic regression, which applies the logistic sigmoidal function by evaluating the gradient of the function that minimizes the cost for a group of parameters [13]; b) Fisher's discriminant analysis which consists of forming discriminating functions from linear combinations of the original variables assuming p random variables with the goal of minimizing the variability between and within groups whose eigenvectors must satisfy matrix algebraic conditions. [14]

The discriminant analysis consists of a useful technique in the classification of sample elements in a population, that is, to discriminate or classify objects through a rule [15]. The main objective is to evaluate the differences between groups and, through the characteristic found, allocate the new observations to one of the classified groups, so with the discriminant analysis it would be possible to tie a new variable and correlate the information obtained [15].

IV. Conclusion

The healthy and the tumor tissues obtained semivariograms as a single theoretical Gaussian model. However, the mean temperatures at each point were distinct, in the healthy internal point ranging from 32.64ºC and in the tumor tissue point 35.87ºC (p-value << 5%). Therefore, with these data we show that the thermal distribution around the analyzed points concentrates spatially distributed independent of the chosen regions not presenting abrupt values, being smooth with similar spatial characteristics, which opens the possibility of using discriminant techniques in future studies to test the possibility of distinguishing benign and malignant tumors.

Furthermore, the difference between the thermal averages centered on the analysis region (healthy and tumor tissue) was confirmed by observing the comparative intervals in the difference of the group mean levels using Turkey's Test. Thus, sweeping using the variogram technique to distinguish between regions (healthy and tumor) can be used as a screening, or active search for tumors, aiding in their early identification.

Acknowledgment

C. G. thanks the National Council for scientific and Technological Development (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) -Finance Code 001.

References

[1] L. Sun, et al. " Uterine Fibroids Increase the Risk of Thyroid Cancer", Int. J. Environ.Res. Public Health, vol 17. nº 3281, 2020, pp 01-09.

[2] K. Segkos, et. al. "Neck Ultrasound in Patients with Follicular Thyroid Carcinoma". Hormones and Cancer. vol. 9 nº 6, 2018, pp 433-439.

[3] Y. Zhuang, et al. “A novel TIRADS of US classification”. BioMedical Engineering Online, v. 17, nº 1, 2018, pp. 1–17.

[4] C. H. Neal, et al. “Breast Imaging Outcomes following Abnormal Thermography”. Academic Radiology, v. 25, nº 3, 2018, pp. 273–278.

[5] T. Rago, et al. “Thyroid ultrasonography reporting: consensus of Italian Thyroid Association (AIT)”, Italian Society of Endocrinology (SIE), Italian Society of Ultrasonography in Medicine and Biology (SIUMB) and Ultrasound Chapter of Italian Society of Medical Radiology (SIRM). Journal of Endocrinological Investigation, v. 41, nº 12, 2018, pp. 1435–1443.

[6] R Core Team (2020). A: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https: //www.R-project.org/.

[7] A. R. S. Andrade, et al. “Geostatistics applied to spatial variability and patterns in the precipitation time series in the harsh Pernambuco”. Journal of Environmental Analysis and Progress, 2018, pp. 126-145.

[8] M. F. Triola, Introduction to Statistics: Technology Update. 11. ed. [s.l.] LTC Editora, 2013.

[9] W. R. J. Soares, Estimation of the variogram using maximum likelihood. 2017.

[10] S. Dionãzio. "An Introduction to Semivariogram: And its Importance Within Geostatistics." National Symposium on Geography and Territorial Management and Academic Week of Geography at the State University of Londrina, 2018, pp 539-549.

[11] C. S. Keretch, “Thermograph assisting the identification of thyroid nodules”. Dissertation (Master's) - Federal Technological University of Paraná. Graduation Program in Biomedical Engineering [S. l.], 2020

[12] A. Legey, M, Fernando; and J. Inoue, “Supervised classification methods.” [S. l.], 2016, pp. 1-3. Available at: http://www.leg.ufpr.br/~eferreira/CE064/pratica6.html. Accessed on: 20 abr. 2021

[13] L. Vicini and A. M. Souza, “Multivariate analysis from theory to practice”. Monography. Santa Maria, UFSM, CCNE, 2005. Available at: http://w3.ufsm.br/adriano/livro/Cadernodedaticomultivariada-LIVROFINAL1.pdf

[14] H. S. Souza, et al. “Multivariate analysis of chemical and physical soil attributes in agroforestry system with teak”. Scientia agraria, vol. 19, nº 1, 2018, pp. 87-93.

[15] C. Loesch, and M. Hoeltgebaum. Multivariate statistical methods. Saraiva Educação SA, 2017.