UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

Visión Electrónica
*Más que un estado sólido*

https://doi.org/10.14483/issn.2248-4728

VISIÓN ELECTRÓNICA

A RESEARCH VISION

# Measurement on MRI for prostate cancer diagnosis

## *Medición sobre MRI para diagnóstico de cáncer de próstata*

*Natalia Andrea Ramírez-Pérez* [1], *Lilia Edith Aparicio-Pico* [2], *Camilo Andrés Pérez-Triana* [3]

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Enviado: 08/10/2020
Recibido: 25/10/2020
Aceptado: 15/12/2020

Keywords:
Diagnosis
Machine learning
Measurement
Prostate cancer



Palabras clave:
Diagnóstico
Aprendizaje automático
Medición
Cáncer de próstata

ABSTRACT

The male reproductive system has a gland located below the bladder and in front of the rectum: the prostate. It surrounds the urethra and has the function of producing a fluid component in the seminal fluid. Over time, this gland tends to enlarge and block the urethra, making it difficult to urinate or sexual function. This alteration is known as harmless prostatic hyperplasia, which is corrected with surgery. Sometimes it is confused with prostate cancer due to the similarity of the symptoms, which is frequent in men.
Diagnosis of this disease is generally made using a manual technique called a digital rectal examination and a laboratory test that measures PSA levels in the blood. It is a substance found in the blood of someone who usually has prostate cancer. Additionally, the diagnosis is supported by a transrectal ultrasound through a catheter. This comprehensive process helps to determine the extension of prostate cancer and designate the correct treatment. The status of prostate injury is assessed by practicing a Magnetic Resonance Imaging (MRI). It is a procedure performed by radio waves and a computer that creates detailed prostate areas' images. It analyzes the prostate condition and determines the procedure or treatment according to the injury's status, for example, surgery, radiation therapy, or monitored observation. To define what kind of treatment, it is essential to analyze the different disease stages and the Gleason Score, a measurement of the histological grade, ranging from 2 to 10, that indicates the probability of spreading or extending the tumor.
This research focuses on the analysis and the extraction of measurements to classify forms of prostate lesions to support its diagnosis. It considers the PI-RADS categorization, which currently determines the probability of suffering from clinically significant prostate cancer. For this purpose, an analysis was made using a geometric interpretation from different categorizations of cancer (4-5). A digital processing of Python images on T2, ADC, and DWI was made applicating the concept of the curve, Zernike moments, fractal dimension, Caliper dimension, the total absolute curvature, the energy bending, direction, convexity, circularity, compactness, Hu moments, dimension, eccentricity, extent, solidity, orientation, largest axis length, smallest axis length, radius, center, centroid, length, area.

RESUMEN

El aparato reproductor masculino tiene una glándula ubicada debajo de la vejiga y frente al recto: la próstata. Rodea la uretra y tiene la función de producir un componente líquido

[1] Mathematician, Universidad Distrital Francisco José de Caldas, Colombia, MA in Information Sciences and Communications, Universidad Distrital Francisco José de Caldas, Colombia. Chief of Data Science at Buho media, Colombia. Current position: Assistant Professor at Institución Universitaria Pascual Bravo. E-mail: naaramirezp@correo.udistrital.edu.co

[2] BA in Special Education Science, PhD. in Technical Sciences, MSc. in Teleinformatics, Specialist in Education Project Management. Current position: Permanent Professor at Universidad Distrital Francisco José de Caldas. E-mail: medicina@udistrital.edu.co

[3] Mathematician, Universidad Distrital Francisco José de Caldas, Colombia. MA in Mathematics, Universidad de los Andes, Colombia. E-mail: ca.perezt@uniandes.edu.co

Cite this article as: N. A. Ramírez-Pérez, L. E. Aparicio-Pico, C. A. Pérez-Triana, "Measurement on MRI for prostate cancer diagnosis", *Visión Electrónica*, vol. 14, no. 2, pp. 196-206, July 2020. https://doi.org/10.14483/22484728.17965

en el líquido seminal. Con el tiempo, esta glándula tiende a agrandarse y bloquear la uretra, lo que dificulta la micción o la función sexual. Esta alteración se conoce como hiperplasia prostática, que se corrige con cirugía. En ocasiones se confunde con el cáncer de próstata por la similitud de los síntomas, que es frecuente en los hombres.
El diagnóstico de esta enfermedad generalmente se realiza mediante una técnica manual llamada tacto rectal y una prueba de laboratorio que mide los niveles de PSA en la sangre. Es una sustancia que se encuentra en la sangre de una persona que suele tener cáncer de próstata. Además, el diagnóstico se apoya en una ecografía transrectal a través de un catéter. Este proceso integral ayuda a determinar la extensión del cáncer de próstata y a designar el tratamiento correcto. El estado de la lesión de próstata se evalúa mediante la práctica de una resonancia magnética (MRI). Es un procedimiento realizado por ondas de radio y una computadora que crea imágenes detalladas de áreas de la próstata. Analiza la condición de la próstata y determina el procedimiento o tratamiento de acuerdo con el estado de la lesión, por ejemplo, cirugía, radioterapia u observación monitoreada. Para definir qué tipo de tratamiento es fundamental analizar los diferentes estadios de la enfermedad y el Gleason Score, una medida del grado histológico, que va de 2 a 10, que indica la probabilidad de diseminación o extensión del tumor.
Esta investigación se centra en el análisis y la extracción de medidas para clasificar formas de lesiones prostáticas que apoyen su diagnóstico. Considera la categorización PI-RADS, que actualmente determina la probabilidad de padecer cáncer de próstata clínicamente significativo. Para ello, se realizó un análisis utilizando una interpretación geométrica de diferentes categorizaciones de cáncer (4-5). Se realizó un procesamiento digital de imágenes de Python en T2, ADC y DWI aplicando el concepto de curva, momentos Zernike, dimensión fractal, dimensión Caliper, la curvatura absoluta total, la flexión de energía, dirección, convexidad, circularidad, compacidad, momentos Hu, dimensión, excentricidad, extensión, solidez, orientación, longitud del eje más grande, longitud del eje más pequeño, radio, centro, centroide, longitud y área.

## 1. Introduction

It is crucial to obtain a correct and prompt diagnosis because it would determine the future tumor stage [1]. Taking timely decisions for the patient's health, since there are no symptoms in the early stage of the disease, require a PSA test, a rectal examination (DRE), and a biopsy to get a precise diagnosis. As these processes require time and extension, it is urgently required to investigate and analyze objective prediction tools that can produce an early diagnosis of cancer, prevent the disease's advance, and affect the prostate to a high degree [2].

An analysis and an application were developed to aim at the early diagnosis of prostate cancer based on Magnetic Resonance images (MRI)in T2 and diffusion sequences. Likewise, images of patients with lesions detected in the prostate, linked to the Prostate Imaging Reporting System (PI-RADS).

The use of MRI for prostate cancer diagnosis began its implementation in the 80s. In 2007, it gained more space to improve the detection and characterization of prostate cancers. MRI has been the most used method for diagnosing prostate cancer, with high-resolution T2-weighted sequences and functional diffusion sequences (DWI), which improve the PSA test's sensitivity and specificity [3]. For this reason, the mathematical-computational support for the diagnosis of prostate cancer using the MRI is essential and a relevant research area to explore [4]. At this point, it is important to mention that the prostate produces the seminal fluid that transports and feeds the sperm, and when this seminal fluid accumulates, it produces alterations in the prostate gland, generating prostate cancer. Externally, cancer causes difficulties in urinating due to the obstruction of the urethra. Prostate cancer is the second most common cancer in men and the most common worldwide -reaching 1.6 million cases-. The lack of an initial diagnosis that determines a timely and adequate treatment can also cause the death of people affected. [5].

The implementation of modern tools and technological advances have improved the diagnosis of prostate cancer. It has caused a considerable decrease in deaths linked to this cancer worldwide. However, it has not happened in Latin America, where the mortality rate has risen and become a preeminent cause of mortality in the region. According to statistics of deaths per 100,000 inhabitants, the most affected countries by prostate cancer are The Bahamas with 78.86, followed by Barbados with 76.8. Mexico, Brazil, and Chile have the most prostate cancer cases registered. [6].

The first part of this article summarizes the computational models' implementation surged from innovation processes to support prostate cancer early diagnosis.

The first stage of prostate cancer is asymptomatic or manifests similar symptoms that can confuse another anomaly called benign hyperplasia. As the MRI is the current system most used for diagnosing prostate cancer, this study aims at contributing, with the help of an expert system, to identify a faster and more accurate cancer diagnosis lesions by locating them and preventing their evolution. [7][8].

The second part exposes the relation of measurements in action for the geometric study using MRI with sequence weighting (T2W), diffusion (DWI-ADC). It is complemented with an analysis of the structures affected through the extraction of geometric characteristics such as the lesion marking measurement in three images (T2, DWI, and ADC). It takes into account eccentricity, Caliper dimension, total absolute curvature, energy bending, direction, convexity, circularity, compactness, Hu moments, area, arc length, endpoints, centroid, convexity, radius, largest and smallest axis length, orientation, solidity, extension, diameter, and fractal dimension by box count. Data is attached showing coordinates of the lesion's position, the description matrix of the orientation, and the image scale, where the values of i, j, and k represent the column, the row, and the cut coordinate to find. Furthermore, the vector with spacing scalars x, y, z, with the position of the lesion TZ or Transitional Zone and PZ or Peripheral Zone with AS or Fibromuscular Stoma, and the PI-RADS evaluation categories determine the probability of clinically significant prostate cancer.

With the help of MRI, the aid system's use for prostate cancer diagnosis was implemented. The PI-RADS evaluation is different for each lesion's location since, in the transition zone, the PI-RADS evaluation is determined by T2W with sometimes modification of the DWI score, while the peripheral zone has a PI-RADS evaluation determined with the DWI and ADC score.

The contribution of this project combines machine learning of images through supervised classifiers. It is a modern advance in medicine that uses algorithms to extract characteristics from medical images. The project uses data sets with their respective validation to classify prostate cancer in patients, obtaining a diagnosis with a certain predictive value [9].

Using data of 35 patients and logistic regression models, neural network, decision tree, and random forests on categories four and five of PI-RADS, the benefits obtained with this classification model would be presented with the extraction of additional geometric characteristics to others studied. It would generate a system that will be managed by doctors in charge of a cancer diagnosis. This system's effectiveness would provide tranquility for patients and allows doctors to improve their decision-making to treat the disease. The certainty of applying the most appropriate timely treatment would eradicate the problem, stabilize it, and provide a better life expectancy for patients.

## 2. Materials and Methods

In diagnosing prostate cancer based on the levels of Prostate Specific Antigen (PSA), which identifies the protein produced by prostate cells found in a blood test, it was determined that PSA is not related to age but linked to prostate cancer. When the disease is detected, the PSA levels rise [10]. Even though the diagnosis is not precise, there are also elevated PSA levels when a benign infection occurs, for instance, prostatic hyperplasia or benign enlargement of the prostate. There are neuroendocrine prostate cancer and another small cell prostate cancer, which have lower PSA levels since it spreads outside the prostate, representing greater aggressiveness. [11].

Today, the magnetic resonance technique with a high-resolution anatomical assessment of the prostate and metabolic information is the most used method in studies of prostate cancer surgeries due to it grants the anatomical and metabolic combination. The most important purpose of this method is to locate prostate cancer, and in the future, there is a possibility of merging ultrasound, magnetic resonance, and PET images to establish new parameters in prostate cancer treatment [12].

New magnetic resonance techniques such as spectroscopy use conventional technology, adding metabolic information from the prostate gland. It gives more reliability to the examination when it is in a pre-surgery state, with a specificity of 90% when there are cancer rates. In distinction, the Dynamic MRI uses rapid sequences after administration of intravenous contrast (gadolinium). Preliminary studies have revealed an intense uptake of the contrast in the first minute after its administration in the neoplastic tissue compared to a normal gland. It is characteristic of the angiogenesis of a tumor. More studies and technological improvements are necessary to assess the intravenous contrast technique's potential utility in prostate cancer's MRI studies.

### *2.1. Assessment Categories of PI-RADS*

The assignment of a PI-RADS assessment category for each lesion is based on the T2, DWI / ADC, and DCE sequence score and according to zonal anatomy.

**Peripheral Zone (PZ)**

The peripheral zone is located on the prostate's posterior and lateral side, surrounding the transition

zone. For the peripheral zone, the DWI / ADC is the primary determining sequence (dominant technique) for assigning the PI-RADS evaluation category.

Since the dominant sequence for PI-RADS assessment in the peripheral zone is different from the transition zone, identifying the zonal location of a lesion is vital.

**Transition Zone (TZ)**

The transition zone surrounds the prostatic urethra and becomes enlarged in maturing men due to benign prostatic hyperplasia. For the transition zone, the T2 image is the primary determining sequence (dominant technique) for assigning the PI-RADS evaluation category.

**PI-RADS in Peripheral Area**

In the peripheral zone, the PI-RADS assessment category of a lesion is determined principally in DWI / ADC, and it is correlated with T2W images. Examples of PI-RADS 1-5 are provided in Figure 1.

**PI-RADS in Transition Zone**

In the transition zone, the PI-RADS assessment category of a lesion is determined primarily on T2 images, and these are correlated with DWI / ADC. Examples of PI-RADS 1-5 are given in Figure 2.

## 3. Topic development

Regarding automatic classification models, it is offered a model capable of discriminating in [13], automatically, the histological samples of prostates of healthy patients compared to sick patients who present cancer at an early stage. The final objective is to provide a diagnostic aid system that reduces the workload and the subjectivity level of specialized doctors in pathological anatomy when analyzing a specific sample resulting from prostate biopsy.

There is an application of clustering techniques in order to find elements of interest to identify gland units. A novel segmentation method known as "Locally Constrained Watershed Transform (LCWT)" [14] has been implemented. It generates three types of objects: artifacts, benign glands, or pathological glands. At this point, machine learning techniques and deep learning techniques are applied to compare the results provided by both approaches and determine the type of each segmented element.

Deep feature extraction is made using shape and texture descriptors, combined with contextual features and fractal analysis. 297 characteristics obtained were subjected to an exhaustive statistical analysis to select only relevant information regarding the correlation between variables and discriminatory capacity concerning their nature. A rigorous data partitioning strategy is implemented to guarantee the robustness of the models built from the available samples of 25 patients (3,200 artifacts, 3,195 benign glands, and 3,000 pathological glands). It ensures the reliability of the results obtained. Once the data partitioning is performed, the problem is approached using a multiclass perspective by applying, on the one hand, linear and non-linear classifiers of classical machine learning (Support Vector Machine and Multilayer Perceptron) and, on the other hand, Convolutional neurons networks, typical of deep learning and based on the VGG19 architecture [15]. The results obtained are compared with those published by other state-of-the-art studies to provide new information to fulfill the final objective, aiming to develop a system to help the early diagnosis of prostate cancer [16].

Supervised classification models are applied with the measurements extracted for each image. It includes logistic regression, decision tree, Random forests, Naive-Bayes (Gaussian), K-neighbors, multilayer perceptron, Naive-Bayes (Bernoulli), classifier of support vectors, stochastic descending gradient, support vector classifier (Linear), support vector classifier (Nu-support), vector support machines. It emphasizes the additional theoretical concepts to [17] and [18], the knowledge and research that support the medical diagnosis of this type of cancer, taking into account the experience and history of many studies that use data for experimentation classification in the diagnosis of prostate cancer.

The inputs will be the measurements resulted from the process applied to each of the images per patient. It consists of filtering the image to highlight and detect contours, which will be the constructed data set fields. These measures in the application of the curve concept, fractal dimension, Caliper dimension, total absolute curvature, energy bending, direction, convexity, circularity, compactness, Hu moments, eccentricity, area, arc length, extreme points, centroid, convexity, center and radius, largest axis length, smaller axis length, orientation, solidity, extent, equivalent diameter and fractal dimension by box-counting along with the data of the coordinates regarding the position of the lesion, the matrix that describes the orientation and scale of the image, the values of i, j and k being the

column, row, and coordinate of the cut to be found respectively. Additionally, the vector with x, y, z spacing scalars, and the lesion location: TZ transitional zone, PZ peripheral zone, and AS fibromuscular stroma.

**Figure 1.** PI-RADS IN Peripheral Zone (PZ, [13][14].

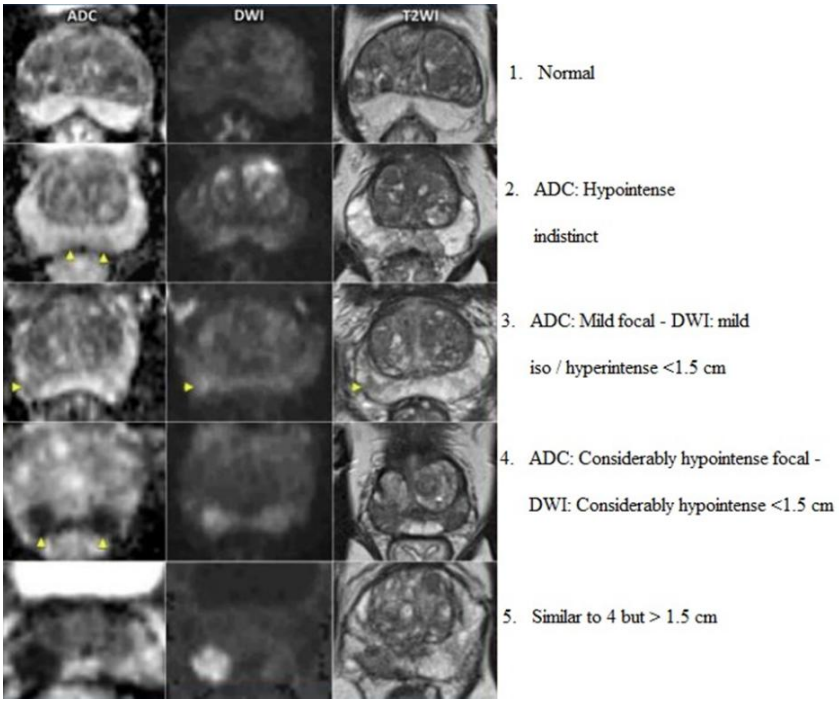
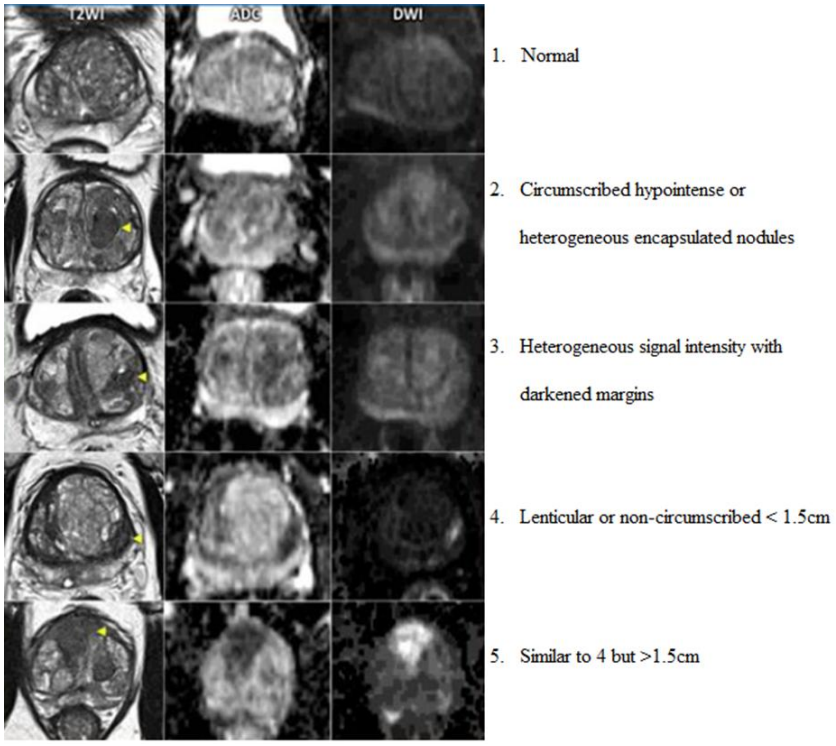



**Figure 2.** PI-RADS TRANSITION ZONE (TZ), [13][14].



All the mentioned measurements are in pixels and consider the prostate imaging data and report system (PI-RADS), which was created to help detect and localize prostate cancer with MRI. It intends to reduce the variability in the interpretation of prostate magnetic resonance studies.

The processing of MRI consists firstly of the identification and tracing of the lesion by a grayscale conversion for each image and the extraction of all the measurements on the binarized images.

The classification report determines the precision of the different models implemented in the prediction on a data set of 20 patients without any value in the PI-RADS categorization. It has better performance in terms of precision using the forest classifier randomized, with 94% precision. The classification model demonstrates precision for the calculated predictions, plus the sum of classified positive and negative predictions, divided over the total data. The random forests are adequately performed with data loading and reading, cleanup and anomalies, missing data, test and training set, predictions in the test set, performance metrics for predictions, model interpretation and results reporting, variable importance, and decision tree visualization [19]. Finally, MIRs for private use were studied with the Cancer Institute's collaboration, obtaining a system's development as a contribution to support medicine.

Random forests create a forest by randomizing, generating multiple decision trees, and combining them to obtain a more precise and stable prediction. It is an algorithm that adds randomness, giving extension to the trees instead of looking for a more critical characteristic by dividing a node. It searches for the best characteristic among the random subset of characteristics, giving a wide diversity and a better classification model [20].

## 4. Equations

### *4.1. Tumor Extension*

It is defined as the ratio of the area of the outline regarding the area of the bounding box:

$$Extension = \frac{lesion\ zone}{area\ of\ the\ bounding\ box}$$

### *4.2. Compactness measurement*

It is a numerical quantity representing the degree to which a shape is compact. This paper is the same isoperimetric quotient: the ratio of the area of the shape to the area of a circle (the most compact shape) having the same perimeter. More precisely, for a given closed curve. The compactness is calculated by:

$$C := \frac{4\pi A}{L^2}$$

Where the area of a circle is A, and L is the perimeter. Moreover, the isoperimetric inequality says (the measure takes a maximum value of 1 when the curve is a circumference) and equivalently, the isoperimetric ratio $\frac{L^2}{A}$ is at least 4π for every curve.

Objects which have an elliptical shape, or a boundary that is irregular rather than flat, will decrease the measure, and in a particular square, a curve has compactness equal to $\frac{\pi}{4}$

### *4.3. Roundness or Circularity*

Another prevalent shape factor is the circularity that belongs to the same family of the isoperimetric quotient. Can be obtained as the ratio:

$$R: = \frac{4\pi A}{P^2}$$

where A is the contour area, and P is the convex hull perimeter. This measure excludes local irregularities and again satisfies the isoperimetric inequality for an object that departs from circularity, except that it is relatively insensitive to irregular boundaries.

### *4.4. Convexity*

Convexity is the relative amount that an object differs from a convex object. A convexity measure should have the following basic properties:

- A convexity measure is a number from (0, 1].
- The convexity measure of a given shape is equal to 1 if and only if it is convex.
- There are shapes whose convexity measure is arbitrarily close to 0, implying no gap between 0 and the measurement's minimum possible value.
- The convexity measure is invariant under appropriate geometric transformations of the shape.

Sometimes less robustness is desirable, and it is crucial to have substantial effects on the measure's value, even if the object changes only slightly.

So, for a given planar shape K, where $K \subseteq R^2$ is a compact and connected polygon, we define convexity by:

$$C := \frac{Per(CH(K))}{Per(K)}$$

Where Per is the usual perimeter of an object, and CH is the surrounding convex of K.

### *4.5. Caliper dimension*

So called Caliper-Feret diameters is the distance between parallel tangents touching opposite sides of an object. Given the orientation $\boldsymbol{\theta}$, the Caliper diameter is:

$$max\ (x\ cos(\theta) + y\ sin\ (\theta))\ -\ min\ (xcos\ (\theta) + y\ sin(\theta))$$

$$(x,y) \in K \qquad\qquad (x,y) \in K$$

In special, the height and weight of the object can be calculated. Indeed if $\boldsymbol{\theta} = 0$:

$$width := max(y) - min(y)$$

$$(x,y \in K \qquad (x,y \in K$$

Or if: $\theta = \frac{\pi}{2}$: $\qquad width := max(x) - min(x)$

$$(x,y) \in K \qquad (x,y) \in K$$

### *4.6. Direction*

It is a property that makes sense in elongated regions only. If the region is elongated when the ratio between the object bounding box's length and width (is a rectangle which circumscribes the object) is close to 1. Thus, this measure is the direction of the longer side of a minimum bounding rectangle.

If the shape moments are known, the direction $\boldsymbol{\theta}$ can be computed as:

$$\theta := \frac{1}{2} tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right)$$

Here, $\boldsymbol{\mu_{pq}}$ are the $p+q$ order central moments. Thus, from this definition, the direction is independent of all linear transformations which do not include rotation. Additionally, we know that the second-order moments measure how dispersed the pixels in an object are from their centroid:

- $\boldsymbol{\mu_{20}}$ measures the object is spread over rows.
- $\boldsymbol{\mu_{02}}$ measures the object is spread over columns.
- $\boldsymbol{\mu_{11}}$ is a cross-product term representing spread in the direction in which both row and column indices increase

### *4.7. Boundary Descriptors: Bending energy and total absolute curvature*

For an irregularly shaped object, the contour direction is a better representation, although it is not directly used as a shape descriptor like centroid, orientation, area, among others.

Consecutive points on the contour of an object give the position or direction, and therefore the concept of chain codes is necessary. Chain codes are a notation for engraving on a list of points that make up an edge along a contour.

They specify the direction of a contour on each edge in the edge list and quantize one of the 8 directions, starting at the first value in the list and going clockwise around the contour. The direction is the values in the chain code for the 8-edge neighborhood (or 8 connected number schemes). It also represents a kind of slope of the contour [21]. The rate of change from one slope to another is called curvature. Since a digital contour is generally irregular, getting an exact curvature measurement is difficult. Its adjacent line segments can define the curvature at a contour point, and the difference between slopes of two adjacent segments is a good measure of curvature at the point of intersection. This idea then suggests that curvature is a local characteristic of an object.

The curvature of the boundary at $\boldsymbol{p} = (\boldsymbol{x_i}, \boldsymbol{y_i})$ can be estimated from the change in the slope is given by:

$$k(p) := \left\{tan^{-1}\left(\frac{y_{i+k} - y_i}{x_{i+k} - x_i}\right) - tan^{-1}\left(\frac{y_i - y_{i-k}}{x_i - x_{i-k}}\right)\right\}(mod 2\pi)$$

Note that a vertex point is in a convex segment when the slope change is positive; otherwise, that point is in a concave segment if there is a negative change in slope.

The descriptor called bending energy is obtained by integrating the squared curvature k(p) through the boundary length L and intuitively is the energy

necessary to bend a rod to the desired shape. It is a robust shape descriptor and can be used for matching shapes:

$$E_c = \frac{1}{L}\sum_{p=1}^{L} k(p)^2$$

Moreover, $\boldsymbol{E_c}$ s satisfies $\frac{2\pi}{R} \leq \boldsymbol{E_c} \leq \infty$ and the value $\frac{2\pi}{R}$ would be obtained for a perfect circle with radius R, and the value will be higher for an irregular object.

Finally, total absolute curvature is the curvatures added along with the boundary points divided by the boundary length.

$$K_{total} = \frac{1}{L}\sum_{p=1}^{L} |k(p)|$$

Here, $K_{total}$ satisfies $2\pi \leq K_{total} \leq \infty$. As the convex object will have the minimum value, a rough object will have a higher value.

### *4.8. Zernike moments*

Zernike moments are a kind of complex orthogonal moment. Its kernel is a set of Zernike complete orthogonal polynomials defined over the unit disc's interior in the polar coordinates space. Two-dimensional Zernike moment $Z_{pq}$ is defined by:

$$Z_{pq} := \frac{p+1}{\pi} \iint D\, f(x,y) V_{pq}^*(x,y) dA$$

where D is the unitary disk, *f(x,y)* is the image intensity function,* denotes the complex conjugate, $V_{pq}$ are Zernike polynomials, and *p, q* both are integers such that $(p - |q|)$ is even and $|q| \leq p$. The integer p is the order for Zernike moment. If the integral is expressed in polar coordinates:

$$Z_{pq} := \frac{p+1}{\pi} \int_0^{2\pi} \int_0^1 f(r,\theta) V_{pq}^*(r,\theta) r\, dr\, d\theta$$

and the Zernike polynomials can be expressed as:

$$V_{pq}(r,\theta) = R_{pq}(r) e^{-iq\theta}$$

Here, $R_{pq}(r)$ is the orthogonal radial polynomial defined by:

$$R_{pq}(r) := \sum_{s=0}^{\frac{p-|q|}{2}} (-1)^s \left[ \frac{(p-s)!}{s!\left(\frac{p+|q|}{2}-s\right)!\left(\frac{p-|q|}{2}\right)!} r^{p-2s} \right]$$

All of the above was in the continuous case, but it is crucial to care about the discrete case. So, all the integrals are replaced by sums:

$$Z_{pq} := \frac{p+1}{\pi} \sum r \sum \theta\, f(r,\theta) V_{pq}^*(r,\theta) r$$

To calculate the limits of sums, the image (or region of interest) is first mapped to the unit disc using polar coordinates, where the center of the image is the origin of the unit disc. Those pixels falling outside the unit disc are not used for the calculation. [22].

## 5. Analysis of results

Notice that in addition to the variables mentioned in [23] and [19] and convexity, other variables such as Zernike moments, Caliper dimension, total absolute curvature, energy bending, and direction were considered as well. The evaluation of the relevance of attributes constructed from the extraction of measurements on the image sequences excludes those with a lower value of importance since the classification models' precision decreased. The different moments are contained in the Hu moments as these are invariant under translations, rotations, and reflections. Therefore, convexity generated only one value and did not contribute to the distinction between the PI-RADS contemplated. However, because these variables were excluded, greater precision was obtained by the random forest classification model, increasing precision because of the relevance value for the Zernike moments 1,6,7 and 12, which provide more information.

## 6. Conclusions

For image processing, its invariants can be extracted from moments generally. For example, these can calculate a centroid's position, the area, the orientation, among others of an image or, more formally, of a mesh of pixel intensities. Furthermore, it is well known that Hu moments, unlike central moments, have the property that these do not change under translations, scaling, or rotations of the object to be analyzed. It makes Hu moments one of the most suitable tools for geometric analysis. However, other types of more complete analytical methods can be found, such as the Zernike moments. These appear from the application of a more formal concept of Zernike polynomials, which

provides enough information to reconstruct the object in an "orthogonal" way in the context of shape analysis. It means that the values calculated are not redundant, nor they overlap because of polynomials' perpendicularity. Even though these are not translation or scaling invariant, these are independent of rotation (for the fact that these are radial functions). For all this, it makes the two moments (Hu and Zernike) have greater significance than the other parameters chosen in this work for the analysis.

The variables of the data set are included in the classification models, and thus their performance in precision is made known, achieving the comparison of final results. Supervised classifiers such as random forests stand out for handling large amounts of data with more dimensionality. To handle thousands of input variables and identify the most significant ones and show the variable's importance surpassing the logistic regression method. Naive-Bayes, k neighbors, multilayer perceptron, support vector classifier (SVM). The evaluation categories the PI-RADS with the Magnetic Resonance images of T2 weighting and DWI-ADC diffusion denoted precision of 0.94%.

Machine learning as a subgroup of Artificial Intelligence (AI) refers to computers' learning through databases, generating algorithms for decision making. In medicine, it is a challenge to train algorithms to recognize categories and patterns for the medical diagnosis of prostate cancer. The use of MRI is linked to supervised classifiers and the Prostate Imaging Reporting Data System or system PI-RADS. Using this, a system aimed at supporting the diagnosis of prostate cancer is possible. The process is made conventionally and considers the result of the analysis of a Prostate Specific Antigen (PSA) levels. The substance is secreted from the prostate. A transrectal ultrasound or prostate biopsy, manual digital rectal examination, and MRI is made to analyze the results. These methods have improved due to science and technology advancement, but the diagnosis continues to depend more on the professional doctor's experience responsible for interpreting the images. The machine learning method and the PI-RADS to classify images support prostate cancer diagnosis, a valuable contribution to a medical diagnosis that speeds up decision-making.

The Prostate Imaging Reporting and Data System (PI-RADS) is based on the findings visualized in high-resolution T2 morphological sequences, functional diffusion sequences, and ADC maps. It is also contrasted with perfusion images, which allow identifying suspicious lesions of clinically significant cancers. The detailed description of the findings in magnetic resonance images using a common language with the other clinicians involved approaching the patient's most appropriate individualized management. PI-RADS recommends follow-up for patients according to the score that can be 1, 2, 3, 4,5 to assign the step to follow in the treatment. It achieves the extraction of geometric characteristics and the numerical characterization in pixels of the injury using the PI-RADS evaluation categories. The professional doctor in charge of issuing the final prostate diagnosis would have a useful tool available to analyze images obtained by magnetic resonance. The numerical value represents the categories in the evaluation of the patient of the extracted variables as a reference for the classifiers.

The modeling and numerical simulation of the body is an investigation that has increased in recent years. The objective is to gain knowledge about the human body's mechanical and physiological behavior and be able to design tools for numerical simulation on a computer to make predictions. It allows correct decisions in medicine, and mathematics' contribution offers precision by the quantification of methods. People can establish applications and calculations when diagnosing, performing surgeries, and other medical procedures. So, geometry is used to analyze and process images, applying angles, planes, and fundamental measures. When making a medical diagnosis, the basis of the scientific development of health based on mathematics is a tool that helps to understand numbers to combat diseases. Since doctors can process, interpret, communicate, and act based on their numerical models based on specific numerical aspects, geometric characteristics extraction can obtain the information from each image and the classification module, starting from the characteristics of Magnetic resonance imaging (T2, DWI, ADC) obtained in the previous module. It requires studying the different characteristics obtained from the images and a learning algorithm that allows classifying them after extracting their content.

## Acknowledgments

Institución Universitaria Pascual Bravo, Grupo de Investigación en Telemedicina (GITEM++), Universidad Distrital Francisco José de Caldas. Python 3.7.1  .

## References

[1] G. M. Carrillo, S. P. Carreño, y L. M. Sánchez, "Competencia para el cuidado en el hogar y carga

en cuidadores familiares de adultos y niños con cáncer", *rev. investig. andin.*, vol. 20, no. 36, pp. 87-101, 2018.

[2] C. Ramos, A. Mercado, "Detección precoz de cáncer de próstata: Controversias y recomendaciones actuales", *Revista Médica Clínica Las Condes*, vol. 29, no. 2, pp. 128-135, 2018. https://doi.org/10.1016/j.rmclc.2018.02.013

[3] T. Hambrock, et al., "Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort", *European Urology*, vol. 61, no. 1, pp. 177-184, 2012. https://doi.org/10.1016/j.eururo.2011.08.042

[4] K. Nguyen, B. Sabata, A. Jain, "Prostate cancer grading: Gland segmentation and structural features", *Pattern Recognition Letters*, vol. 33, no. 7, pp. 951-961, 2012. https://doi.org/10.1016/j.patrec.2011.10.001

[5] Medimaging, "Una herramienta nueva de aprendizaje automático predice con exactitud el cáncer de próstata - Industria - Medimaging.es". [Online]. Available at: https://www.medimaging.es/industria/articles/294777132/una-herramienta-nueva-de-aprendizaje-automatico-predice-con-exactitud-el-cancer-de-prostata.html

[6] J. C. Batlle et al., "Diagnóstico del cáncer de próstata mediante espectroscopia de resonancia magnetica endorectal", *Arch. Esp. Urol.*, vol. 59, no. 10, pp. 953-963, 2006. https://doi.org/10.4321/S0004-06142006001000003

[7] C. Vargas, "Diferenciación entre prostatitis y cáncer de próstata utilizando el sistema PI-RADS", 2016. [Online]. Available at: https://cbseram.com/2016/06/22/diferenciacion-entre-prostatitis-y-cancer-de-prostata-utilizando-el-sistema-pi-rads/

[8] T. Hambrock et al., "Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort", *Eur. Urol.*, vol. 61, no. 1, pp. 177-184, 2012. https://doi.org/10.1016/j.eururo.2011.08.042

[9] R. E. Valero Vargas, J. J. Palacios Rozo, R. González Silva, "Tecnologías de la Información y la Comunicación y los Objetos Virtuales de Aprendizaje: un apoyo a la presencialidad", *Revista vínculos*, vol. 16, no. 1, pp. 82–91, jun. 2019. https://doi.org/10.14483/2322939X.15537

[10] SEOM, "Cáncer de Próstata - SEOM: Sociedad Española de Oncología Médica", 2019. [Online]. Available at: https://seom.org/info-sobre-el-cancer/prostata?showall=1

[11] A. B. Rosenkrantz and S. S. Taneja, "Radiologist, be aware: Ten pitfalls that confound the interpretation of multiparametric prostate MRI", *American Journal of Roentgenology*, vol. 202, no. 1. pp. 109-120, 2014. https://doi.org/10.2214/AJR.13.10699

[12] Radiology, "The Radiology Assistant: Prostate Cancer - PI-RADS v2". [Online]. Available at: https://radiologyassistant.nl/abdomen/prostate/prostate-cancer-pi-rads-v2

[13] P. Guzmán, A. Messina, "Cáncer de próstata, el problema del diagnóstico ¿Es la resonancia multiparamétrica de próstata la solución?", *Rev. Chil. Radiol.*, vol. 25, no. 2, pp. 60-66, 2019. https://doi.org/10.4067/S0717-93082019000200060

[14] I. Robles, "Identificacion de Biomarcadores Predictivos, Pronósticos y de Respuesta al Cancer de Prostata", 2018. [Online]. Available at: http://digibug.ugr.es/bitstream/handle/10481/52748/29070302.pdf?sequence=4

[15] N. A. Ramirez, E. Gómez, O. Forero, "Clasificadores supervisados del cáncer de próstata a partir de imágenes de resonancia magnética en magnetic resonance images in T2 sequences", 14th Iberian Conference on Information Systems and Technologies (CISTI), 2019. https://doi.org/10.23919/CISTI.2019.8760647

[16] J. I. Díaz, "Matemáticas y Ciencias de la Salud", pp. 65-67, 2005. [Online]. Available at: http://www.mat.ucm.es/imi/documents/ActasMatematicasCienciasDeLaSalud.pdf

[17] S. L. Goldenberg, G. Nir, and S. E. Salcudean, "A new era: artificial intelligence and machine learning in prostate cancer", *Nat. Rev. Urol.*, vol. 16, no. 7, pp. 391-403, 2019. https://doi.org/10.1038/s41585-019-0193-3

[18] R. Cuocolo et al., "Machine learning applications in prostate cancer magnetic resonance imaging", *Eur. Radiol. Exp.*, vol. 3, no. 1, 2019. https://doi.org/10.1186/s41747-019-0109-2

[19] N. Ramírez, E. Aparicio, E. Gómez, "Supervised classifiers of prostate cancer. A geometric study on magnetic resonance images T2 weighted (T2W), by diffusion (DWI-ADC)," Congr. Int. electrónica, Control y telecomunicaciones, p. 51, 2018.

[20] E. L. Díaz Gutiérrez y C. F. Valderrama García, "Evaluación de la usabilidad de los EVA (entornos virtuales de aprendizaje) a partir de la experiencia de usuarios aplicando lógica difusa", *Revista vínculos*, vol. 15, no. 2, pp. 150–159, oct. 2018. https://doi.org/10.14483/2322939X.14006

[21] S. Yoo, I. Gujrathi, M. A. Haider, and F. Khalvati, "Prostate Cancer Detection using Deep Convolutional Neural Networks", *Sci. Rep.*, vol. 9, no. 1, pp. 1-10, 2019. https://doi.org/10.1038/s41598-019-55972-4

[22] I. Simon, C. R. Pound, A. W. Partin, J. Q. Clemens and W. A. Christens-Barry, "Automated image analysis system for detecting boundaries of live prostate cancer cells", *Cytometry*, vol. 31, no. 4, pp. 287-294, 1998. https://doi.org/10.1002/(SICI)1097-0320(19980401)31:4<287::AID-CYTO8>3.0.CO;2-G

[23] S. Sarkar and S. Das, "A Review of Imaging Methods for Prostate Cancer Detection", *Biomed. Eng. Comput. Biol.*, vol. 7, no. 1, 2016. https://doi.org/10.4137/BECB.S34255